Proceedings of SAIP2011,

the 56th annual conference of the South African Institute of Physics

Edited by
Ilsa Basson and André E. Botha
Department of Physics, UNISA

SOUTH AFRICAN INSTITUTE OF PHYSICS
UNISA
college of science, engineering and technology

PROCEEDINGS EDITORS: Ilsa Basson and André E. Botha

PROCEEDINGS EDITORIAL COMMITTEE AT UNISA:
Ilsa Basson
André Botha
Moritz Braun
Simon Dhlamini
Leslie Lekala
Sabata Moloi
Moses Mothudi
Joel Rampho
Derck Smits
Malebo Tibane
Srinivasu Vallabhapurapu

PUBLISHER: University of South Africa (UNISA)

COVER DESIGNER: Dawid Kahts

DUPLICATION of CD: UNISA Unit for Sound, Video & Photography, Pretoria, South Africa

The Proceedings of SAIP2011, the 56th annual conference of the South African Institute of Physics (SAIP) will only be available electronically on compact disk (CD) and on the SAIP website www.saip.org.za.

ISBN: 978-1-86888-688-3

SAIP2011

Proceedings of SAIP2011,
the 56th annual conference of the
South African Institute of Physics
(SAIP)

Hosted by the Department of Physics,
University of South Africa (UNISA)

12 - 15 July 2011
St George Hotel
Gauteng
South Africa

Edited by
Ilsa Basson and André E. Botha
Department of Physics, UNISA

TABLE OF CONTENTS

Message from the Conference Chair … xv

Message from the Review Committee at UNISA ... xvi

Conference Chairs and Committees ... xvii

List of Reviewers ... xviii

Full Research Papers

Division A– Condensed Matter and Material Science

Dielectric barrier discharge CO_2 TEA laser operated at frequencies up to 400 Hz
Peter Baricholo, Timo Stehmann, Erich G Rohwer, Mike Collier, Dumisani J Hlatywayo and Hubertus M von Bergmann... 2

Determining the bulk concentration of S in Fe-S: An Auger electron spectroscopy study
P E Barnard, J J Terblans, H C Swart, B G Anderson and M J H Hoffman... 8

Investigation of Sulphur diffusion in bulk Iron – A DFT study
P E Barnard, J J Terblans, H C Swart, B G Anderson and M J H Hoffman... 14

Energetic stability and charge states of 3d transition metals in diamond: Towards a diamond based diluted magnetic semiconductor
E M Benecha and E B Lombardi... 20

Tuning carbon nanotubes for photovoltaic applications
G Bepete, D Wamwangi, N J Coville, and Z Chiguvare... 26

Electrical and surface morphological studies of palladium and ruthenium Schottky diodes on *n*-Ge (100)
A Chawanda, F D Auret, W Mtangi, J M Nel, M Diale and C Nyamhere... 32

Thermal properties of micron size Ag-added PDMS composites: A fuel cell perspective
Hao Chen, Ionel Botef, V Vasudeva Rao and V V Srinivasu... 39

Non-linear Fowler-Nordheim plots in thin film polymer-fullerene composite devices (transition from hole-only to electron-only conduction)
Zivayi Chiguvare... 44

Photoluminescence surface mapping as a probe for interface disorder in quantum well structures
K G Chinyama, P N Njingana and E L Meyer... 50

Magnetic vector charges in the mystery of a circular current's *pair* of distinct Cartesian elemental magnetic dipoles
M Chirwa..... 56

Computational modelling of Zr-Nb alloys
M H Chuma, P E Ngoepe and H R Chauke..... 62

Measuring the effect of surface evaporation on the segregation process in a Sb/Cu system
S Cronjé, W D Roos, R E Kroon and J K O Asante..... 68

Synthesis and photoluminescence properties of $CaSi_xO_y:Tb^{3+}$
F *B Dejene and M A Kebede*..... 74

X-ray Photoelectron Spectroscopy analysis and characterization of $Gd_2O_2S:Tb^{3+}$ phosphor powder
J J *Dolo, F B Dejene, J J Terblans, O.M. Ntwaeaborwa and H C Swart*..... 79

Blended thin films of tris-(8-hydroxyquinoline) aluminium (Alq_3) embedded in polymethyl methacrylate (PMMA).
M M Duvenhage, O M Ntwaeaborwa and H C Swart..... 85

Synthesis and characterization of a green $CaAl_xO_y:Tb^{3+}$ phosphor using solution combustion method
K E Foka, F B Dejene and H C Swart..... 90

MOVPE grown self-assembled InSb quantum dots: Structural characterization
M Godbole, E J Olivier, C Heiligers, J H Neethling and J R Botha..... 95

Spectroscopic studies of nanofluorides doped with Ln^{3+} synthesized *via* thermal decomposition of organic precursors
M A Gusowski and H C Swart..... 100

Spin-density-wave properties of $(Cr_{84}Re_{16})100_{-x}V_x$ alloys
B S Jacobs, A R E Prinsloo1, C J Sheppard and A M Strydom..... 105

The surface orientation dependence of the pre-exponential factors extracted from the segregation profiles of a Cu(111/110) bi-crystal
C J Jafta, W D Roos, J J Terblans, J K O Asante and K I Ozoemena..... 111

A positron annihilation study of F and Ba sub-lattices in superionic BaF_2 at elevated temperatures
T P Jili, E Sideras-Haddad, D Wamwangi, F Tuomisto and B Kibirige..... 116

Synthesis and characterization of $Ba_xMg_yAl_2O_4$: Eu,Dy nanophosphors prepared using solution - combustion method
M A Kebede and F B Dejene..... 121

Characterization of nitrogen-doped carbon nanospheres using electron magnetic resonance
Jonathan M Keartland, Makhosonke B Dubazane, Vincent Marsicano, Nikiwe Kunjuzwa and Neil J Coville..... 127

Temperature variation of the electric field gradient in mercuric chloride
Jonathan M Keartland and Eric Newby..... 134

Cr/ α-Cr_2O_3 monodispersed spherical core-shell particles based solar absorbers
S Khamlich, Z Y Nuru, R Nemutudi, O Nemraoui, K T Roro, N Mongwaketsi, R McCrindle, N Cingo and M Maaza..... 140

Vibrational properties of mass produced graphene monolayer by chemical method
M Khenfouch, M Maaza and M Baïtoul..... 146

Synthesis and characterization of PbS nanostructures using the chemical bath deposition method
L F Koao, F B Dejene and H C Swart..... 151

Ion irradiation effects on the formation of metal nanoparticles in crystals
Anna Kozakiewicz, Trevor Derry, Paul Franklyn, S R Naidoo and Chris Theron..... 156

The impurity levels of lanthanide ions in silica
R E Kroon, H A A Seed Ahmed and M A Gusowski..... 162

Magnetic properties of carbon nanospheres at low temperatures
Makhosonke B Dubazane, Jonathan M Keartland and Yosief Mishgina..... 168

Low temperature synthesis of ZnO nanoparticles and nanorods via wet chemistry route
P S Mbule, G H Mhlongo, H C Swart and O M Ntwaeaborwa..... 174

Effects of ZnO and Ce^{3+} incorporation on the photoluminescence and cathodoluminescence intensity of Pr^{3+} doped SiO_2
G H Mhlongo, M S Dhlamini, O M Ntwaeaborwa, H C Swart, R E Kroon and K T Hillie..... 179

The influence of Pr^{3+} co-doping on the photoluminescence and cathodoluminescence properties of SiO_2:Eu^{3+}/Tb^{3+}
G H Mhlongo, M S Dhlamini, O M Ntwaeaborwa, H C Swart and K T Hillie..... 186

Nanoscale manipulation of lamellar copolymers using electric fields
Simiso K Mkhonta..... 193

Optical properties of $SrGa_2S_4$:Ce^{3+} thin films prepared by pulsed reactive cross laser ablation (PRCLA) technique
P A Moleme, J J Dolo, H C Swart and O M Ntwaeaborwa..... 198

XRD, DRS and SEM studies of the effects of metal dopants (Pt and Au) on the structural and optical properties of TiO_2
J R Moloantoa, K E Rammutla, T E Mosuang, R M Erasmus and K T Hillie..... 203

Ab *initio* structural and electronic study of metals on diamond (111)-(1 x 1) surface
I Motochi1, N W Makau, G O Amolo, B A Mathe and S R Naidoo..... 207

Equivalent parameters for the empirical pseudopotential and *k·p* models
Kagiso Mpshe, Wilnelia E Adams, André E Botha and Enrico B Lombardi..... 214

Spin-density-wave behaviour of the $(Cr_{100-x}Al_x)_{95}Mo_5$ alloy system
B Muchono, A R E Prinsloo, C J Sheppard, H L Alberts and A M Strydom..... 220

Statics and dynamics of single polymers confined between two corrugated walls
Debashish Mukherji and Martin H Müser..... 225

Low voltage electron beam induced degradation and surface chemical changes of $Zn_3(PO_4)_2$:Tb phosphor
I M Nagpure, Shreyas S Pitale, Liza Coetzee, O M Ntwaeaborwa, J J Terblans and H C Swart..... 232

Effect of pH on the morphology and orientation of Fe_2O_3 nanostructures grown using aqueous chemical growth
C L Ndlangamandla, C B Mtshali, M J Sithole, B D Ngom, O M Ndwandwe and M Maaza..... 238

Investigation of broken symmetry of Sb/Cu(111) surface alloys by VT-STM
G F Ndlovu, J K O Asante, W D Roos and K T Hillie..... 243

The decoration of vicinal copper polycrystalline surfaces by Antimony
G F Ndlovu, J K O Asante, W D Roos and K T Hillie..... 249

Enhancement of Pr^{3+} red emission by adding In^{3+} as co-dopant in $CaTiO_3$:Pr^{3+} phosphor
L L Noto, S S Pitale, M M Gusowki, J J Terblans, O M Ntwaeaborwa and H C Swart..... 254

Computational study of rutile tin-oxide (SnO_2)
J N Ntimane, T E Mosuang and K E Rammutla..... 258

Microstructure and spectral selectivity of Pt-Al_2O_3 nanocoatings for high temperature applications
Z Y Nuru, S Khamlich, K Roro, T F G Muller, C J Arendse and M Maaza..... *262*

Reactive DC magnetron sputter deposition and characterization of ZrN thin films
T G Nyawo and M O Ndwandwe..... 267

Structure-property relationship of sol-gel synthesised zinc-oxide nanoparticles
N V Peterson, C J Arendse, K T Hillie and M Dhlamini..... 274

Effect of annealing on the Ce^{3+}/Ce^{4+} ratio measured by XPS in luminescent SiO_2:Ce
H A A Seed Ahmed, W D Roos, O M Ntwaeaborwa, H C Swart and R E Kroon..... 280

Synthesis and photoluminescence properties of Tb^{3+}-doped $Sr_{0.5}$ $Zn_{0.5}$ Al_2O_4 phosphor prepared via combustion process
S K K Shaat, F Roelofse, H C Swart and O M Ntwaeaborwa..... 285

Magnetic behaviour of the Cr-Al alloy system around the triple point
C J Sheppard, A R E Prinsloo, H L Alberts, A M Strydom and B Muchono..... 291

Synthesis and characterization of reproducible stoichiometry of cobalt sulfide nanoparticles using sulphur containing single-source precursors
S B Sibokoza, M J Moloto and N Moloto..... 296

The magnetocaloric effect in ferromagnetic PrSi: evidence of a novel magnetic ground state and higher order exchange interactions
J L Snyman and A M Strydom..... 302

Crystalline electric field effects in $PrNiGe_2$
J L Snyman and A M Strydom..... 308

Magnetocaloric effect of ferromagnetic CeAuGe
J L Snyman, B M Sondezi-Mhlungu and A M Strydom..... 314

Thermodynamic properties and inelastic neutron scattering of ferromagnetic CeCuSi
B M Sondezi-Mhlungu, A M Strydom and D T Adroja..... 319

Critical behavior near the ferromagnetic Curie phase transition in CeCuGe
B M Sondezi-Mhlungu and A M Strydom..... 324

Low-lying magnetism in strongly correlated $CeRh_2Sn_2$
A M Strydom..... 329

Thermal and electronic transport studies of the Kondo energy scale in the heavy-fermion system $CeCu_{5-x}Al_x$
A M Strydom and D Britz..... 334

Synthesis and characterization of the ZnO nanoparticles and the polyvinyl pyrrolidone (PVP) encapsulated ZnO nanoparticles
M A Tshabalala, B F Dejene and H C Swart..... 338

Enhanced green emission from UV down-converting Ce^{3+} and Tb^{3+} co-activated $ZnAl_2O_4$ phosphor
K G Tshabalala, S-H Cho, J-K Park, I M Nagpure, J H Neethling, J R Botha, H C Swart and O M Ntwaeaborwa..... 343

On the orbital rehybridization in tetrahedral amorphous carbon
Aniekan Magnus Ukpong..... 348

Synthesis and characterization of structural and luminescent properties of long afterglow $CaAl_2O_4$: Eu^{2+}, Nd^{3+}, Dy^{3+} phosphors by solution – combustion technique.
A H Wako, F B Dejene and H C Swart..... 354

Division B– Nuclear, Particle and Radiation Physics

A century of nuclear theory: severely condensed
Ken Amos..... 361

Quasifree alpha cluster knockout studies
A A Cowley..... 369

Low-lying states in ^{17}C from a multi-channel algebraic scattering theory
S Karataglidis, K Amos, L Canton, J P Svenne, D van der Knij and P R Fraser..... 375

Characterization of cluster states in ^{16}O with the (*p*; *t*) reaction
J Mabiala, E Z Buthelezi, A A Cowley, S V Förtsch, M Freer, T T Ibrahim, J P Mira, R Neveling, P Papka, F D Smit, G F Steyn, J A Swartz, I Usman, J J Van Zyl and S M Wyngaardt..... 381

Study of rare modes of "Collinear Cluster Tri-Partition" of ^{252}Cf(sf)
V D Malaza, N M Jacobs, D V Kamanin, Yu V Pyatkov, A A Aleksandrov, I A Aleksandrova, N A Kondratyev, E A Kuznetsova, O V Strekalovsky, A O Strekalovsky and VE Zhuchko..... 385

Radon escape from mine tailings dams
J N Ongori, R Lindsay, R T Newman and P P Maleka..... 391

Division C– Lasers, Optics and Spectroscopy

Cross phase modulation induced depolarization of a probe signal and its impact on polarization mode dispersion compensators
R R G Gamatham, T B Gibbon and A W R Leitch..... 398

Quantum cryptography for satellite communication
M Mariola, A Mirza and F Petruccione..... 403

Spectroscopy with a mode-locked femtosecond laser frequency comb
Vela Mbele and Elias Sideras-Haddad..... 409

Efficiency of Tm^{3+}-doped silica triple clad fiber laser
W I Ndebeka, A M Heidt, H Schwoerer and E G Rohwer..... 417

Evaluation of WC-9Co-4Cr laser surface alloyed coatings on stainless steel
B A Obadele, P A Olubambi, S L Pityana and A P I Popoola..... 421

Polarisation encoded quantum key distribution in fibre
S Pillay, A R Mirza and F Petruccione..... 426

Magneto-optical trapping: first step towards BEC
M Semonyo, M Rehn and F Petruccione..... 432

Phase transition in hydrothermal synthesized VO2(M) nanocrystals:an X-ray diffraction study
A Simo, R Madjoe, L Edomwonyi-Otu, A Gibaud and M Maaza..... 437

Division D– Astrophysics and Space Science

Searching for signatures of nearby sources of cosmic rays in their local chemical composition
D *Bisschoff, I Büsching and M S Potgieter*..... 444

Evaluation of satellites LAGEOS I and II; general relativistic accelerations in the Schwarzschild field of Earth
L Combrinck..... 450

Optical study of PMS stars in the southern hemisphere high mass star forming region RCW 34
R J Czanik, D J van der Walt and S I Loubser..... 456

Pulsating B stars in the LMC
C A Engelbrecht and FAM Frescura..... 462

Lie derivatives in stellar perturbation theory
F A M Frescura and C A Engelbrecht..... 467

An African VLBI network of radio telescopes
M J Gaylard, M F Bietenholz, L Combrinck, R S Booth, S J Buchner, B L Fanaroff, G C MacLeod, G D Nicolson, J F H Quick, P Stronkhorst and T L Venkatasubramani..... 473

The neutron superfluid and a possible application to neutron star interiors
G J Kemp, C A Engelbrecht and F A M Frescura..... 479

Daytime modelling of VLF radio waves over land and sea, comparison with data from DEMETER Satellite
S G Meyer, A B Collier, C J Rodger..... 484

The effect of rotation on the shape of stars
P R Nicol, F A M Frescura and C A Engelbrecht..... 489

Statistical analysis of outer electron radiation belt dropouts: geosynchronous and low earth orbit responses during stream interfaces
Olakunle Ogunjobi, Andrew B Collier and Craig J Rodger..... 495

Pulsar-like hard X-ray emission in AE Aquarii
B Oruru and P J Meintjes..... 501

X-ray timing and spectral analyses of the unusual magnetic cataclysmic variable AE Aquarii
B Oruru and P J Meintjes..... 507

Atomic processes in gaseous nebulae
A Prozesky..... 513

Genetic algorithms in astronomy and astrophysics
Vinesh Rajpaul..... 519

Atmospheric turbidity over Soweto
T Sethabane and H Winkler..... 525

Constraining viewing geometries of pulsars with single-peaked γ-ray profiles using a multiwavelength approach
A S Seyffert, C Venter, T J Johnson and A K Harding..... 531

Constraining the phase space for chameleon dark energy
Muzikayise E Sikhonde and Amanda Weltman..... 537

Eclipsing contact binary stars - period analysis using SuperWASP data
P L Skelton..... 543

What can we learn from phase alignment of γ-ray and radio pulsar light curves?
C Venter, T J Johnson and A K Harding..... 549

Campaign for vicarious calibration of SumbandilaSat in Argentina
L M Vhengani, D Griffith and M Lysko..... 555

Brightest cluster galaxies – single or composite stellar populations?
D N Viljoen and S I Loubser..... 561

The unusual Seyfert Markarian 926 – a link to the LINERs?
H Winkler..... 567

Division E - Physics Education

Contrasts between student and examiner perceptions of the nature of assessment tasks
D Clerk, D Naidoo and S Shrivastava..... 574

Visualizing physics: good vibrations
D Fish, T R Anderson..... 579

Active learning in the third year statistical physics module at the University of the Witwatersrand
Jonathan M Keartland..... 585

Simple pendulum: a first year student's dilemma
Frank Komati and Itumeleng Phage..... 591

"Seeing is believing": visual perceptions and the learning of kinematics
Miriam Lemmer..... 596

Learners' understanding of ammeter and voltmeter in direct current schematic circuits
R T Mautjana and T T Netshisaulu..... 602

The pre-knowledge of the physics I students on vectors
M R Mhlongo, M E Sithole and A N Moloi..... 608

The transition from high-school physics to first-year physics: how well prepared are our students?
P Molefe and B M Sondezi-Mhlungu..... 614

Instructional design principles applied to physics laboratory and tutorial courses
D Naidoo and D Clerk.....619

National curriculum statement achievement levels – can they serve as a measure of science students' preparedness for university study?
Sam Ramaila, Padmanabhan Nair and Leelakrishna Reddy..... 624

Lightning – scientific knowledge versus mythological beliefs
Sam Ramaila, Padmanabhan Nair, Leelakrishna Reddy..... 630

Foundation provision: any difference in student's performance?
B M Sondezi-Mhlungu, P Molefe..... 637

Visualising and quantifying students' academic disappointment in science courses
Anitta Thomas and André E Botha..... 643

Division F - Applied and Industrial Physics

Raman characterization of the phase transition of ion implanted hexagonal boron nitride to cubic boron nitride Nanoparticles
E Aradi, S R Naidoo, R M Erasmus and T E Derry..... 652

Application of electroluminescence and thermal imaging in defect identification in photovoltaic modules
J L Crozier, E E van Dyk and F J Vorster..... 658

Comparative efficiency study of a solar trough receiver: hot mirror and selective coating
M C Cyulinyana and P Ferrer..... 664

Protein localization and folding mechanisms revealed by molecular dynamics simulations
James Gumbart..... 671

Laser irradiation: a complementary treatment for wounds
N N Houreld and H Abrahamse..... 677

Resin phantoms as skin simulating layers
Aletta Elizabeth Karsten, Ann Singh and Ivy Mankekolo Ndhundhuma..... 683

Enhancement of hydrogen production using biomass gasification process
A I Mabuda, N S Mamphweli and E L Meyer..... 689

Investigating the thermal performance of a hybrid photovoltaiv solar system
B Mtunzi, E L Meyer and M Simon..... 694

Batch anaerobic co-digestion of cow dung and donkey manure
P Mukumba, G Makaka and S Mamphweli..... 701

Adipose derived stem cells (ADSCs) and low intensity laser irradiation (LILI): potential use in regenerative medicine
B Mvula and H Abrahamse..... 707

Transmission electron microscopy of radiation damage by keV multi-implantation in single-crystal diamond
E K Nshingabigwi, J H Neethling, C M Levitt, S R Naidoo and T E Derry..... 711

Determining optical performance and current generation of a CPV as a function of intensity distribution
R D Schultz, F J Vorster and E E van Dyk..... 715

Inflammatory response of injured diabetic fibroblasts after low intensity laser irradiation at a wavelength of 830 nm
P R Sekhejane, N N Houreld and H Abrahamse..... 721

Optical coherence tomography as a diagnostic tool
Ann Singh and Aletta Elizabeth Karsten..... 728

Indoor temperature predictions in an energy efficient solar house
S Ziuku and E L Meyer..... 733

Division G - Theoretical and Computational Physics

Wave – packet scattering off a soliton
A M H H Abdelhady and H Weigel..... 740

Density-matrix renormalization group study of the electro-absorption in conjugated polymers
A O Abdelwhab, E Jeckelmann and A Artoli..... 746

Representation of the quantum and classical states of light carrying orbital angular momentum
Humairah Bassa and Thomas Konrad..... 752

Optimised periodic and hyperchaotic modes of a triple pendulum
André E Botha and Guoyuan Qi..... 758

Finite element calculations for molecules with multiple Coulomb centers
Moritz Braun..... 766

Non-locality without inequality and generalized non-local theory
Sujit K Choudhary..... 774

Alternative derivation of the master equation for a particle in an external field subject to continuous measurement
K Garapo and T Konrad..... 780

Entanglement dynamics in an oscillating bipartite Gaussian state coupled to reservoirs at different temperatures
A Ghesquière, I Sinayskiy and F Petruccione..... 786

Monte Carlo simulation using GEANT 4 of the EMU μMSR facility at ISIS
J M Hartman and S H Connell..... 792

The physics of exceptional points
W D Heiss..... 799

Analytic density functionals
S Karataglidis and B G Giraud..... 804

Towards the unconditional security proof for the coherent-one-way protocol
Mhlambululi Mafu, Adriana Marais and Francesco Petruccione..... 811

Many-boson quantum walks on graphs with shared coins
Dibwe Pierrot Musumbu, Ilya Sinayskiy and Francesco Petruccione..... 817

Few-nucleon systems studies with extended antisymmetrized molecular dynamics
G J Rampho and S A Sofianos..... 823

Electromagnetic form factors of ^{3}H and ^{3}He systems
G J Rampho, S A Sofianos, S Oryu and T Watanabe..... 827

Unravelling of open quantum random walks
Ilya Sinayskiy and Francesco Petruccione..... 833

A simple model of decoherence-assisted transport
Ilya Sinayskiy, Adriana Marais and Francesco Petruccione..... 837

Message from the Conference Chair

It is with pride that the Department of Physics at the University of South Africa (UNISA) presents to you the Proceedings of SAIP2011 the 56th Annual Conference of the South African Institute of Physics (SAIP). This is the first time in the 56 year history of SAIP conferences that proceedings with peer reviewed research papers are formally published. We are grateful to authors for the unexpected number of papers submitted, as well as all the reviewers that ably supported us to get the reviews done. The creation of the SAIP2011 Proceedings is a milestone for the Physics community in South Africa, our international collaborators and Physics at UNISA. The purpose of the proceedings is twofold; firstly to disseminate original research and new developments in Physics and related fields and secondly to provide an opportunity for Masters and Doctoral students to publish their research work. The proceedings will be available via the SAIP website www.saip.org.za.

The Department of Physics at UNISA was delighted to host the annual conference of the South African Institute of Physics in the winter of 2011. We welcomed delegates from within the country, some parts of Africa as well as abroad. The contributions of the plenary speakers need special mention. We learned from the lifelong experiences of the 2010 SAIP Gold medal winner Professor George Ellis and the following experts in their respective fields: Professors Wolfgang Christian (USA), Claude Carignan (UCT), Gyözö Garab (Hungary), Debashish Mukherji (Germany) and Ken Amos (Australia). A large number of postgraduate students participated. This is particularly pleasing, since one of the SAIP and UNISA's objectives is to build research capacity.

This year we celebrated three centenary milestones on the Physics calendar; the discoveries of the nucleus and superconductivity and together with the Chemists in the International Year of Chemistry, the remarkable work in radioactivity and the discovery of two elements of one of the ladies of science, Marie Curie.

Our Department at UNISA presented this year's Winter School on Computational Physics and High Performance Computing in collaboration with the Centre for High Performance Computing (CHPC) in Cape Town. Our guest speakers were Professors Wolfgang Christian (USA) and Fransisco Esquembre (Spain) on computational approaches in Physics tuition, Debashish Mukherji (Germany) and Moritz Braun (Physics UNISA) on computational approaches in theoretical Physics and Enrico Lombardi (Physics UNISA) and Dr Daniel Moeketsi (CHPC) on high performance computing in Physics. The CHPC offered for the first time Masters and Doctoral prizes for students that utilised High Performance Computing in their research.

We thank all participants of SAIP2011. The success of this conference is attributed to the combined efforts of all. The support of the SAIP Council and Office and the University of South Africa needs mention as well as the contributions of our sponsors.

I also sincerely thank all the staff of the Department of Physics at UNISA for their effort and work before, during and after the conference. It was a great accomplishment for a Department consisting of only eleven people to host more than 450 conference delegates, 405 oral and poster presentations and to publish 145 peer-reviewed papers!

Ilsa Basson

SAIP2011 Conference Chair
Chairperson Department of Physics, UNISA

Message from the Review Committee at UNISA

The review committee considered 185 contributions towards the proceedings, and subjected 167 of these to a full peer-review process. A total number of 141 subject specialists assisted with this review process, 129 of whom hold at least a Ph.D. in Physics or in a related field. Each paper received two or more independent reviews. Conflicts of interest were handled by the review committee. For each contribution, the review committee selected two independent reviewers whom they felt could provide impartial expert opinions. In addition, declarations were made by the reviewers themselves, to indicate possible conflicts of interests that may have been overlooked by the review committee. Several reviewers declined to review because of such unforeseen conflicts of interest.

The final selection of papers was made by the Editors, based on the outcome of the peer-review process. In cases where two reviewers' reports were in conflict, a third independent review was solicited, before an editorial decision was made. A total of 145 contributions were finally accepted (78%).

The style of these proceedings is that of the British Institute of Physics Conference Series. In the initial call for contributions the conference organisers indicated that the final responsibility for editing and producing the PDF version of each contribution would lie with the author(s). Nevertheless, the editorial committee made every effort to assist authors with the production of correctly formatted contributions. In some cases, however, it was impossible to complete this process within the allocated time and thus, in addition to small differences that may be observed between contributions typeset in LaTex compared to MSWord, there are still a few contributions which do not strictly conform to the prescribed style.

The review committee is very grateful for the diligence of the reviewers. The quality of feedback which was received undoubtedly contributed to the quality and overall high standard of the research reported in these proceedings. All the research papers published in these proceedings have been through a thorough peer-review process. Consequently each paper reports on original research which has not been published previously.

The committee would like to thank all the authors who submitted papers to SAIP2011, as well as the reviewers who generously gave of their time and expertise to ensure that each contribution received a critical review.

André E. Botha

On behalf of SAIP2011 Review Committee
Department of Physics, UNISA

Conference Chairs and Committees

SAIP2011 Conference Chair
Ilsa Basson, Department of Physics, University of South Africa

SAIP2011 Programme Chair
André Botha, Department of Physics, University of South Africa

SAIP Division Chairs
Division A: Japie Engelbrecht (Nelson Mandela Metropolitan University)
Division B: Simon Mullins (iThemba LABS)
Division C: Erich Rohwer (University of Stellenbosch)
Division D: Patrick Woudt (UCT), Andrew Collier (University of KwaZulu Natal)
Division E: Sam Ramaila (University of Johannesburg)
Division F: Freddie Vorster (Nelson Mandela Metropolitan University)
Division G: Frederik Scholtz (National Institute for Theoretical Physics)

Finance Chair
Brian Masara, SAIP Office, CSIR, Pretoria, South Africa

Logistics
Carla de Jager, Carlamani Conferences and Events

Proceedings Editorial Committee at Unisa
Ilsa Basson
André Botha
Moritz Braun
Simon Dhlamini
Leslie Lekala
Sabata Moloi
Moses Mothudi
Joel Rampho
Derck Smits
Malebo Tibane
Srinivasu Vallabhapurapu

List of Reviewers

Dr. ABRAMS, Cameron	Drexel University
Dr. ACKERMANN, Lutz	University of Limpopo
Prof. ARENDSE, Chris	University of the Western Cape
Dr. ASANTE, Joseph	Tshwane University of Technology
Dr. BARK, Robert	iThemba Labs
Prof. BERGMANN, Hubertus	University of Stellenbosch
Prof. BHATTACHARYYA, Somnath	University of the Witwatersrand
Prof. BOTHA, Andre	University of South Africa
Prof. BOTHA, Reinhardt	Nelson Mandela Metropolitan University
Prof. BRAUN, Moritz	University of South Africa
Prof. BUCHLEITNER, Andreas	University of Freiburg
Ms. BUCHNER, Sarah	HartRAO
Prof. CARIGNAN, Claude	University of Cape Town
Dr. CHAUKE, Hasani	Materials Modelling Centre
Prof. CHAWANDA, Albert	University of Pretoria
Prof. CHETTY, Nithaya	University of Pretoria
Dr. CHIGUVARE, Zivayi	University of the Witwatersrand
Dr. CHINYAMA, Kaumba	University of Fort Hare
Dr. CHIRWA, Max	Walter Sisulu University
Mr. CLERK, Douglas	University of the Witwatersrand
Prof. CLEYMANS, Jean	University of Cape Town
Dr. COETZEE, Liza	University of the Free State
Prof. COLAFRANCESCO, Sergio	University of the Witwatersrand
Prof. COMBRINCK, Ludwig	HartRAO
Prof. COWLEY, Anthony	Stellenbosch University
Dr. DANGBEGNON, Julien	Nelson Mandela Metropolitan University
Dr. DAVYDOV, Evgeny	Joint Insitute for Nuclear Research
Prof. DE MELLO KOCH, Robert	University of the Witwatersrand
Prof. DEJENE, Francis	University of the Free State
Prof. DERRY, Trevor	University of the Witwatersrand
Dr. DHLAMINI, Simon	University of South Africa
Dr. DIALE, Mmantsae	University of Pretoria
Mr. DOLO, Jappie	University of the Free State
Dr. ENGELBRECHT, Christian	University of Johannesburg
Mr. ERASMUS, Rudolph	University of the Witwatersrand
Dr. FERRER, Phil	University of the Witwatersrand
Mr. FISH, Derek	Unizul Science Centre
Prof. FORBES, Andrew	Council for Scientific and Industrial Research
Prof. FOUCHE, Willem	University of South Africa
Prof. FRESCURA, Fabio	University of the Witwatersrand
Dr. GAYLARD, Michael	HartRAO
Dr. GHESQUIERE, Anne	National Institute for Theoretical Physics
Dr. GIRALDI, Filippo	School of Physics and NITheP, UKZN
Dr. GOLDSTEIN, Kevin	University of the Witwatersrand
Dr. GREBEN, Jan	Council for Scientific and Industrial Research
Dr. HARRIS, Richard	Mintek
Dr. HATO, Joseph	Council for Scientific and Industrial Research
Dr. HEIDT, Alexander	University of Stellenbosch
Dr. HOURELD, Nicolette	University of Johannesburg

Prof. JONES, B	University of Lancaster
Dr. KEARTLAND, Jonathan	University of the Witwatersrand
Dr. KROON, Ted	University of the Free State
Dr. LEEUW, Lerothodi	University of Johannesburg
Dr. LEMMER, Miriam	North-West University (Potchefstroom)
Prof. LINDSAY, Robert	University of Cape Town
Dr. LITWINSKI, Christian	Rhodes University
Prof. LOMBARDI, Enrico	University of South Africa
Dr. LOUBSER, Ilani	North-West University
Dr. MACHATINE, Augusto	University of Pretoria
Dr. MACLEOD, Gordon	Department of Science and Technology
Dr. MAMPHWELI, Sampson	University of Fort Hare
Dr. MANYALA, Ncholu	University of Pretoria
Dr. MARTINEZ, Peter	SAAO
Dr. MATHE, Bhekumusa	University of the Witwatersrand
Prof. MEINTJIES, Petrus	University of the Free State
Dr. MEYER, Walter	University of Pretoria
Ms. MHLONGO, Gugu	Council for Scientific and Industrial Research/UFS
Dr. MISHRA, Dilip K.	IMMT, Bhubaneswar, India
Dr. MKHONTA, Simiso	University of Swaziland
Dr. MOEKETSI, Daniel Mojalefa	Council for Scientific and Industrial Research
Dr. MOJI, Cable	University of Pretoria
Dr. MOLA, Genene	University of Kwazulu-Natal
Dr. MOLEFI, J A	Sasol
Dr. MOLOI, Sabata	University of South Africa
Dr. MOLOTO, Nosipho	University of the Witwatersrand
Dr. MOODLEY, Mathew	Council for Scientific and Industrial Research
Dr. MOSUANG, Thuto	University of Limpopo
Dr. MOTAUNG, David	Council for Scientific and Industrial Research
Dr. MOYO, Thomas	University of Kwazulu-Natal
Dr. MUKHERJI, Debashish	Max-Planck Institute for Polymer Research
Dr. MULLER, T F G	University of Western Cape
Prof. MULLER-NEDEBOCK, Kristian	University of Stellenbosch
Dr. MWAKIKUNGA, Bonex	Council for Scientific and Industrial Research
Dr. MYER, Martin	Council for Scientific and Industrial Research
Dr. NAGABHUSHANA, K	University of Mysore, India
Dr. NAGPURE, Indrajit	University of the Free State
Dr. NAIDOO, Deena	University of the Witwatersrand
Dr. NAIR, Padmanabhan	University of Johannesburg
Dr. NEETHLING, Pieter	University of Stellenbosch
Dr. NEL, Jackie	University of Pretoria
Prof. NETSHISAULU, Thomas	University of Limpopo
Dr. NISTOR, Razvan IBM T.J.	Watson Research Center
Dr. NTOAHAE, Peter	University of Limpopo
Prof. NTWAEABORWA, Odireleng	University of the Free State
Dr. NYAMHERE, Cloud	Nelson Mandela Metropolitian University
Dr. Ocaya, R	University of the Free State
Mr. OKENG'O, G	University of Nairobi
Mr. OLIPHANT, Clive	University of the Western Cape
Dr. OLUWAFEMI, O	Walter Sisulu University
Mr. ORURU, Bosco	University of the Free State

Mr. PHAGE, Itumeleng	Central University of Technology
Dr. PILLAI, Sreejarani	Council for Scientific and Industrial Research
Ms. PROZESKY, Andri	University of South Africa
Mr. RAMAILA, Sam	University of Johannesburg
Prof. RAMMUTLA, Koena	University of Limpopo
Dr. RAMPHO, Gaotsiwe	University of South Africa
Prof. RAKITIANSKI, Sergey	University of Pretoria
Dr. REDDY, Leelakrishna	University of Johannesburg
Prof. ROOS, Wiets	University of the Free State
Dr. RORO, Kittessa	Council for Scientific and Industrial Research
Dr. SCRIBA, M	University of Cape Town
Dr. SHEPPARD, Charles	University of Johannesburg
Dr. SIKAKANA, Ike	Vaal University of Technology
Dr. SINAYSKIY, Ilya	University of KwaZulu-Natal
Ms. SKELTON, Patricia	University of South Africa
Dr. SMIT, Jacoba	Council for Scientific and Industrial Research
Prof. SMITS, Derck	University of South Africa
Dr. SPARROW, Raymond	Council for Scientific and Industrial Research
Dr. STEENKAMP, Christine	University of Stellenbosch
Dr. STRAUSS, Hencharl	Council for Scientific and Industrial Research
Prof. STRYDOM, Andre	University of Johannesburg
Prof. SWART, Hendrik	University of the Free State
Prof. TERBLANS, Koos	University of the Free State
Prof. THERON, Chris	University of Pretoria
Dr. TIBANE, Malebo	University of South Africa
Dr. TINARWO, David	University of Venda
Dr. TYAGI, A. K.	BARC, Mumbai, India
Dr. UYS, Hermann	Council for Scientific and Industrial Research
Prof. VALLABHAPURAPU, Srinivasu	University of South Africa
Prof. VALLURI, Ram	University of Western Ontario
Prof. VAN DER WALT, Johan	North-West University
Prof. VAN DYK, Ernest	Nelson Mandela Metropolitan University
Dr. VENTER, Christo	North-West University
Prof. VENTER, Andre	Nelson Mandela Metropolitan University
Prof. VON BERGMANN, Hubertus	University of Stellenbosch
Dr. VORSTER, Frederik	Nelson Mandela Metropolitan University
Prof. WAGENER, Pieter	Nelson Mandela Metropolitan University
Prof. WALTERS, Piet	University of Stellenbosch
Dr. WATERMEYER, Jean	University of Cape Town
Prof. WINKLER, Hartmut	University of Johannesburg
Prof. WOUDT, Patrick	University of Cape Town

Division A – Condensed Matter and Material Science

Dielectric barrier discharge CO_2 TEA laser operated at frequencies up to 400 Hz

Peter Baricholo[1], Timo Stehmann[2], Erich G. Rohwer[2], Mike Collier[1], Dumisani J. Hlatywayo[1], Hubertus M. von Bergmann[2]

1. Department of Applied Physics, National University of Science and Technology, P. O. Box AC939, Ascot, Bulawayo, Zimbabwe
2. Laser Research Institute, Department of Physics, University of Stellenbosch, South Africa

Corresponding author: pbaricholo@nust.ac.zw

Abstract. A dielectric barrier discharge CO_2 TEA laser excited by a thyratron driven power supply has been developed and characterized. Laser output was observed at frequencies up to 25 Hz for an electrode separation of 10 mm with 1.8 mm suprasil glass covering the electrodes. At this gap separation, pulsed power output of about 9 W was detected for gas pressures between 100 and 400 mbar. The laser has a pulse duration was about 0.3 μs with a risetime of ~100 ns. Changing the electrode separation to 5 mm and using 1.4 mm suprasil glass dielectric increased the output power to 23 W and enabled laser output to be observed at gas pressures up to 700 mbar and maximum pulse excitation frequencies of up to 400 Hz. The developed laser does not require water cooling since the system operates in burst pulse mode.

1. Introduction

Despite current advances in high power fibre laser development, the CO_2 laser remains the preferred laser when high beam quality and focusability are required. CO_2 lasers are excited by d.c. [1, 2], high frequency (10 kHz–3 MHz) [3-6], radio frequency (13–1500 MHz) [7] and microwave (2.45 GHz) [8, 9] power supplies. Dielectric barrier discharges (DBDs) have also been used in the excitation of CO_2 lasers but to date, only high frequency excitation and radio frequency excitation has resulted in cw lasing [6, 10-12] in these systems. Such lasers make use of special gas mixtures which usually has 30 % helium and more than 50 % nitrogen content. In most instances, DBD excited lasers use a trigger unit [11, 13] as a form of preionization to enable the generation of a uniform discharge and in other cases, a damped oscillation discharge is used to excited the DBD excited system [14].

DBD excitation of the transversely excited atmospheric (TEA) CO_2 laser has the following advantages. Critical alignment is avoided. Compact laser structures that do not require preionization can be employed and the metal electrode is not inside the laser discharge cavity. This can increase the lifetime of sealed-off CO_2 lasers by reducing oxide formation [15]. The self ballasting effect of the dielectric makes this possible.

We have developed a small CO_2 TEA laser excited by a DBD which can operate at frequencies up to 400 Hz without the use of preionization. This laser uses either the 1:1:3 or 1:1:8 CO_2:N_2:He gas mixtures. We believe this is the first time that such a laser is being demonstrated at these frequencies. This work aims at prolonging the electrode life of mini CO_2 TEA lasers which have arcing and pitting problems from continuous use as well as demonstrate that DBD excited lasers do not require pre-ionization when operated at atmospheric pressures (≤500 mbar).

2. Description of a dielectric barrier discharge excited CO_2 TEA laser

Self adhesive aluminium foil electrodes, 40 cm long and 18 mm wide, used with this laser are covered with suprasil glass and are neither in contact with the plasma nor the vacuum as shown in Figure 1. Suprasil glasses 1.8 mm and 1.4 mm in thickness have been used with either a gap separation of 5 mm or 10 mm measured between the dielectrics. At the center of the discharge cavity, the discharge is estimated to have a width of 3 mm, resulting in estimated mode volumes of 6 cm^3 and 12 cm^3 respectively. The developed low cost DBD excited CO_2 TEA laser has ZnSe windows fitted at both ends of the discharge cavity. A 99.8 % half inch ZnSe output coupler and a copper full back reflector of 1.99 m radius of curvature demarcates the 71 cm resonant cavity. A PEM detector (Vigo System, PEMI) has been employed for measurement of the optical output of the DBD excited TEA CO_2 laser.

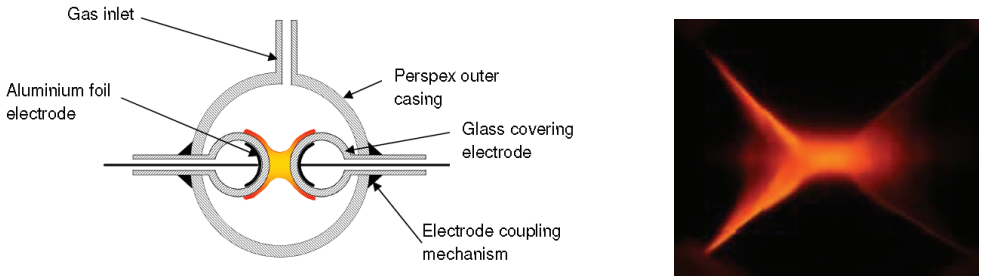
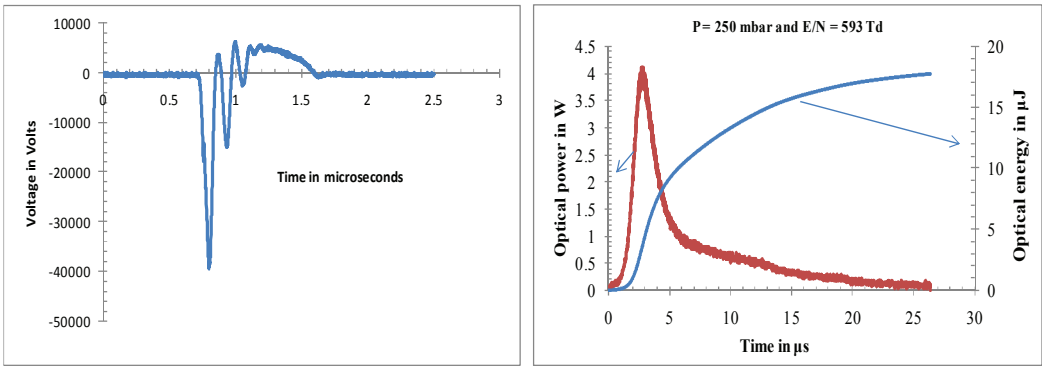


Figure 1: Sketch showing discharge cavity with electrodes inside the glass cylindrical dielectrics. The picture shows the end view of the discharge observed through a ZnSe window.

A thyratron (Triton electron technology, model F-189) driven high voltage power supply (Lambda EMI, Model 402L) with variable output voltage and repetition frequency was used to provide the excitation pulses. A capacitor bank comprising of 5 capacitors of 0.92 nF capacitance each connected in parallel was selected resulting in a energy storage of the capacitor bank of 2 J at a charging voltage of 30 kV. The 5 mm electrode gap in our DBD system gives a gas load capacitance of 12.74 pF.

3. Results, analysis and discussion

Figure 2(a) show the voltage pulse used during excitation and Figure 2(b) show the measured optical power and the integrated energy obtained with this laser. Two different CO_2 laser gas mixtures, the 1:1:3 and 1:1:8 CO_2:N_2:He were used during the characterization of the developed laser system.



(a) b)

Figure 2: a) Voltage signal used in the excitation of the laser b) optical power and energy of the DBD excited laser. The 1:1:3 laser gas mixture at 500 mbar constituted the laser active medium.

Breakdown occurred at about 18 kV and the pulse duration was about 0.3 µs with a risetime of ~100 ns. In Figure 2(b) peak optical power of 9 W and pulse energy of 8 µJ was detected of the laser when the electrode gap was 10 mm.

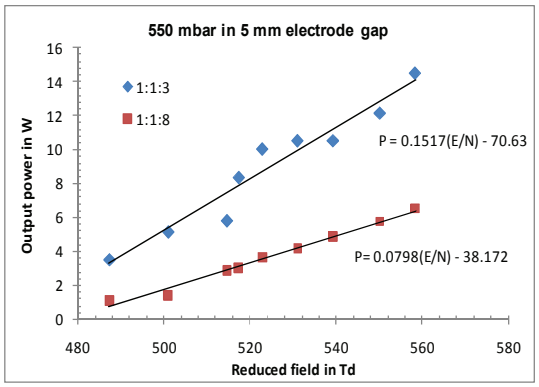
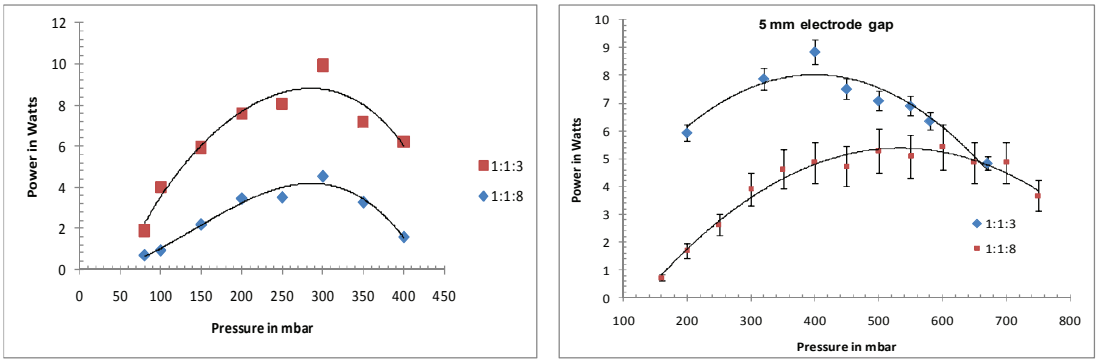


Figure 3: Variation of output power with reduced field for gas pressure of 550 mbar and electrode gap of 5 mm using the 1:1:3 and 1:1:8 CO_2:N_2:He gas mixtures.

A threshold reduced field (*E/N*) exists for this laser as reflected from Figure 3. The graph was obtained with a 1.4 mm dielectrics and a gap separation of 5 mm. Power output values plotted against *E/N* in Figure 3 show a linear relationship between the two quantities. Optical power is a function of the reduced field. The threshold reduced field required before any optical output could be detected for

the two gas mixtures used is 470 ± 8 Td (1 Td = 10^{-17} Vcm^2). High values of optical pulse power e.g. 50 W would require a reduced field of 800 Td.

The constant of proportionality for the pressure and *E/N* relationship is 2 times smaller at 300 mbar as compared to 500 mbar. This means by raising the pressure by 200 mbar from 300 mbar, the population of the upper laser level is almost doubled.

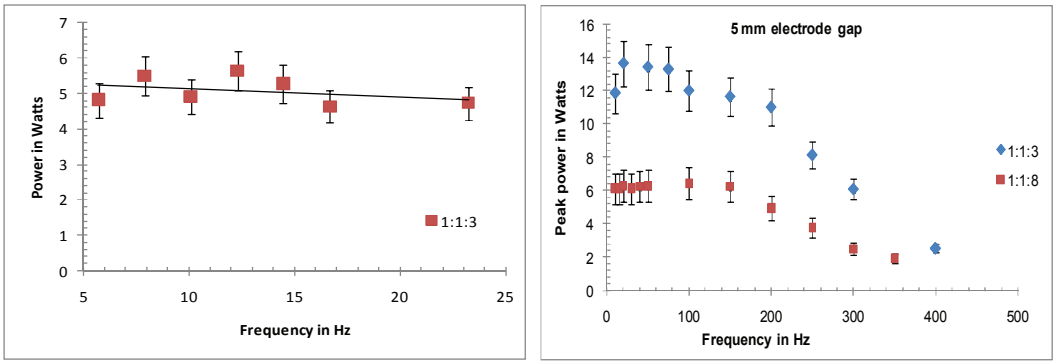


(a) (b)

Figure 4: Variation of power output with pressure a) with a 10 mm electrode gap and dielectric thickness of 1.8 mm and b) 5 mm electrode gap with a dielectric thickness of 1.4 mm.

The optimum pressure observed for maximum optical output energy in Figure 4(b) was 300 mbar for a gap separation of 10 mm and 400 mbar for a 5 mm electrode gap with a 1:1:3 gas mixture. Maximum pulse power of 14.0 ±1.0 W was obtained with the 1:1:3 CO_2 laser gas at a pressure of 300 mbar. Optimum pressure for the 10 mm electrode gap with a 1.8 mm thick glass dielectric is 300 mbar for the two gas mixtures used in this research as shown in Figure 4(a). As pressure was increased, optical output power increased up to a maximum value before decreasing again. In Figure 4(b) the variation of power output with pressure for an electrode gap of 5 mm and with 1.4 mm glass dielectrics shows that higher optical output could be obtained at higher pressures. The optimum operating pressure was 550 mbar and at this pressure 7 μJ of optical pulse energy with a gas mixture of 1:1:8 was measured. It is evident that, after the modifications, the laser could give an optical output at pressures above 500 mbar and an optimum pressure of 400 mbar for the 1:1:3 gas mixture. The maximum optical power did not vary significantly in magnitude as compared with prior to the modifications. Pressure above the optimum values resulted in lower population levels of the upper laser levels and therefore in gain reduction. The steep decrease of the output at pressures above 300 mbar for the 10 mm gap and above 500 mbar for the 5 mm gap is attributed to the fact that the voltage across the discharge was not high enough resulting in a highly filamentary discharge.

The developed laser gives high values of output optical power for frequencies up to 100 Hz for the 5 mm electrode separation as shown in Figure 5. Results in Figure 5(a) show that laser output is obtained at frequencies less than 25 Hz. With half the initial electrode separation and double the initial pressure the optical output was increased two fold as observed in Figure 5(b). Exciting the laser using higher frequencies resulted in reduced optical power output which could be a result of electron trapping in the discharge gap preventing the bulk of the electrons from crossing the gap. Gas heating resulting in increased CO_2 dissociation is also one of the reasons for the decrease in output.



(a) (b)

Figure 5: Variation of optical power with excitation frequency for a) a 1 cm electrode gap and 1.8 mm thickness glass dielectric and gas pressure of 250 mbar and b) 0.5 cm electrode gap and 1.4 mm thickness glass dielectric and gas pressure of 500 mbar.

When the frequency was varied, the output power changed as reflected in Figure 5. Exciting the DBD CO_2 (TEA) laser with an electrode gap of 5 mm increased the frequency window for which the laser can be used as exposed by Figure 5(b). High optical powers where obtained at low frequencies of up to 150 Hz. The power output decreased as excitation frequency was further increased and no output was observed at frequencies above 400 Hz. By increasing the pulse repetition rate, we increased the energy deposition into the gas. Without water cooling, the gas temperature increased for pulse repetition rates greater than 150 Hz. Once dielectric electrodes charge up, the discharge current terminates and the large part of the stored energy remains in the charging capacitor and the dielectric electrodes without being transferred to the laser medium. The optical output is also sensitive to gas mixture. Using the 1:1:3 CO_2:N_2:He gas mixture, the optical power increased two fold when compared with results obtained using 1:1:8 CO_2:N_2:He gas mixture.

4. Conclusion

A DBD excited CO_2 TEA laser operating at frequencies normally associated with DBD excited excimer lamps has been demonstrated. Laser output was observed in pulsed mode with maximum

energy per pulse being 10.0 ± 1.0 μJ and a maximum of 400 Hz pulse frequency for a 5 mm electrode gap separation. The laser optical pulse was observed about 5 μs after the excitation pulse and its duration was 3 to 5 μs. Our results show that it is possible to prolong electrode life of mini CO_2 TEA lasers by covering the metal electrodes with a dielectric and continue operating such lasers in a real pulsed mode. To improve the optical power output, besides increasing the excitation voltage, the electrode gap could be further reduced, a barium titanate ($BaTiO_3$), a ceramic with high dielectric constant of more than 3000 could be used instead of glass or a thinner dielectric would be a viable option. These modifications increase the dielectric capacitance and should result in increased pulse energy. Also different gas mixtures with varying proportions of CO_2 and N_2 could be investigated in an attempt to increase the optical output.

Acknowledgements
The authors would like to thank the African Laser Center for the financial support received towards this project and the Laser Research Institute at University of Stellenbosch for the facilities used in carrying out this research.

References

[1] Patel C N K 1964 *Phys. Rev.* A **136** 617- 9
[2] Witteman W J 1987 *The CO_2 laser* ed Enoch J M, MacAdam D L *et al* Berlin Springer 81 -192
[3] Tanaka M, Yagi S and Tabata N 1985 High frequency silent discharge and its application to CW CO_2 laser excitation *Proc. 8th Int. Conf. on Gas Discharges and their Applications (Oxford 16 -20 September 1985)* 551 - 554
[4] Pace P W and Cruinkshank J M 1980 *IEEE J. Quantum Electron.* **QE-16** 937 – 44.
[5] Brown R T 1973 *IEEE J.Quantum electron.* **QE-9** 1120 - 122
[6] Wieneke S and Viöl W 2000 *J. Phys D: Appl. Phys.* **33**:1282 - 286
[7] Abe T, Yasuda T, Tanizaki M, Yamaguchi S, Nanri K, Miyakawa N and Fujioka T 2008 *Rev. Laser Eng.* **Supp. Vol.** 261-64
[8] Viöl W and Uhlenbuschz J. 1996 *J. Phys. D: Appl. Phys.* **29** 57–67
[9] Wester R and Seiwert S 1991 *J.Phys. D: Appl. Phys.* **24** 1102-107
[10] Wieneke S, Uhrlandt C and Viöl W 2004 *Laser Phys. Letts.* **1** 241 - 47
[11] Yasui K M K and S Ogawa, M. Tanaka and S. Yagi 1989 *IEEE J. Quantum Electron.* **25** 836 - 40
[12] Yagi S, Kuzumoto M, Tanaka M and Ogawa S 1990 *IEEE J. Quantum Electron.* **26** 1130 - 34
[13] Wester R, Seiwert S and Wagner R. 1991 *J. Phys. D: Appl. Phys.* **24** 1796-802
[14] Hokazono H, Midorikawa K, Obara M and Fujioka T 1982 Rev. Sci. Instrum. **54** 719-21
[15] Cohn D B 1980 Appl. Phys. Letts. **37** 771-73

Determining the bulk concentration of S in Fe-S: An Auger electron spectroscopy study

P E Barnard[1], J J Terblans[1], H C Swart[1], B G Anderson[2], M J H Hoffman[1]

[1]Department of Physics, University of the Free State, P. O. Box 339, Bloemfontein, ZA-9300, South Africa
[2] Sasol Technology R&D, P. O. Box 1, 1 Klasie Havenga Road, Sasolburg, ZA-1947, South Africa

E-mail: terblansjj@ufs.ac.za

Abstract. Auger electron spectroscopy constant temperature segregation runs have been performed on Fe-S-C for the temperatures 823 – 1073 K in 50 K increments. For the temperatures 823 – 973 K, the segregation of Carbon (C) is initially observed, but soon replaced by Sulphur (S) which dominates the Iron (Fe) surface once it start segregating. At the high temperatures 1023 K and 1073 K only the segregation of S was observed on the surface. Fitting of the adapted $t^{\frac{1}{2}}$ equation to the segregation profile of S for each of the temperatures delivered an average bulk concentration of 14.07 ± 0.1946 ppm S in the Fe sample. The results indicate the use of AES to determine the bulk concentration of S in polycrystalline Fe.

1. Introduction

Fe having such a vast amount of applications, has captured the attention of both scientist and engineers alike. The presence of S in the Fe, however causes a weakening of these properties. At high temperatures , the S segregates to the grain boundaries causing grain boundary embrittlement which ultimately leads to failure of the machinery [1, 2]. The Fischer-Tropsch process makes use of an Fe catalyst to convert syngas into hydrocarbons which is used in the petrochemical industry. The presence of S in the Fe catalyst segregates to the surface and occupies surface positions, causing a decreased conversion rate and ultimately leads to the deactivation of the catalyst [3, 4].

The application of Fe in the above mentioned cases, where Fe is operated at high temperatures, calls for a study of Fe under high temperatures in order to investigate the diffusion of impurities such as S. Such studies have been performed on polycrystalline Fe by Reichl *et al*. [5] and also on the Fe(100), Fe(110) and Fe(111) surfaces [6]. Using the models of Fick, Reichl *et al*. was able to determine the diffusion parameters, D_0 and Q, as well as the segregation energies, ΔG for the impurities C and S.

As mentioned above the fitting of Fick`s model to experimental data delivers the diffusion parameters for a sample of known concentration. The results presented here made use of Fick`s adapted $t^{\frac{1}{2}}$ equation [7] to extract values for the bulk concentration of S in the polycrystalline sample of Fe by making use of diffusion parameters from literature. A similar study has been performed by

Terblans *et al*. [8] on a Cu(Bi,S) system in order to determine the bulk concentration of S in the sample.

2. Theory

The surface segregation of elements, including the changes that occur to the bulk concentration during sputtering, has been shown by du Plessis *et al*. [7] to be described by equation 1

$$C^{\phi} = C_0^{\phi} + C^B \left[1 + \frac{2}{d}\left(\frac{D(t+t_0)}{\pi} \right)^{\frac{1}{2}} \right] - C^B \left[\frac{2}{d}\left(\frac{Dt_0}{\pi} \right)^{\frac{1}{2}} \right] \quad (1)$$

where C^{ϕ} is the surface concentration of the segregated species, C_0^{ϕ} is the concentration of the segregated species that was not sputter in time t_0 and remained on the surface, C^B is the bulk concentration of the sample, d is the interlattice spacing of the sample, D is the diffusion coefficient, t_0 is the time the sample was sputter cleaned and, t is the time in which the segregation measurements were performed.

Equation 1 can be fitted to experimental data in order to extract important parameters such as the diffusion parameters or as was done in this study the bulk concentration. In order to extract the bulk concentration values, equation 1 was written in the linear form given by equation 2, where the concentration, C_0^{ϕ} was considered negligible

$$C^{\phi} = \frac{2C^B}{d}\left(\frac{D}{\pi} \right)^{\frac{1}{2}} (t+t_0)^{\frac{1}{2}} - C^B \left[\frac{2}{d}\left(\frac{Dt_0}{\pi} \right)^{\frac{1}{2}} - 1 \right] \quad (2)$$

plotting of C^{ϕ} against $\left(t + t_0^{'}\right)^{\frac{1}{2}}$ gives a straight line with

$$\text{Gradient} = \frac{2C^B}{d}\left(\frac{D}{\pi} \right)^{\frac{1}{2}};$$

$$\text{y-intercept} = C^B \left[\frac{2}{d}\left(\frac{Dt_0}{\pi} \right)^{\frac{1}{2}} - 1 \right]$$

using the gradient and substituting in the parameters *d* and *D,* the bulk concentration of the sample was determined for each of the temperature.

3. Experimental

All measurements were performed on the PHI 590 Scanning Auger Microscope, consisting of an electron gun, model 18-085 and a single pass cylindrical mirror analyser, model: 25-110. The ion gun used for sputter cleaning of the sample is a model 11-065 from Perkin-Elmer. The Auger spectrum was recorded in the differentiated mode, by using a lock-in amp with a time constant of 0.1 s and a sensitivity value of 100 ×. In conjunction with the lock-in amp an electron multiplier control unit with modulation energy set to 2 eV was used. Data was recorded using an electron beam voltage of 5 keV and a ion current of 0.95 μA. Sputter cleaning of the surface was performed on an raster area of 2×2 mm with an ion beam voltage of 2 kV and a current of 0.073 μA. The system was kept in the UHV region with a pressure value of 10^{-9} Torr, by making use of an ion pump. Only during sputter cleaning did the pressure rise to a maximum value of 2×10^{-7} Torr.

Polycrystalline Fe samples of dimension 10 mm diameter and 0.5 mm in thickness were mechanically polished on diamond suspension down to 1 μm and subsequently cleaned in an ultrasonic bath of ethanol. The sample was then mounted onto the heating stage located on the sample holder of the AES system. During the heating experiments, the temperature was measured using a K-type thermocouple located on the heating stage beneath the sample.

Constant temperature heating was performed at temperatures in the range 823 − 1073 K in increments of 50 K. Samples were heated to the desired temperature and once there, kept at that temperature for 1 hour to ensure an equilibrium state within the bulk of the crystal. Afterwards the surface was sputter cleaned for 10 min and the data acquisition started. After each segregation run the sample was heated at 1073 K for 3 hours in order to eliminate any depletion of subsurface layers that might have originated from the previous segregation run. The following peaks were monitored during the segregation run: S (153 eV), C (275 eV), N (389 eV), O (510 eV) and Fe (50 eV and 705 eV). Peaks were analysed afterwards in order to calculate the true Auger peak-to-peak height (APPH) and ensure that no peak overlapping occurred. Data was then quantified using the method for thin overlayers, described by Briggs and Seah [9] and the equations of Shimizu [10] for the calculation of the backscattering factors using elemental standards of known composition to calculate sensitivity factors. For the calculation of the inelastic mean free path (IMFP) of Auger electrons, the TPP-2 method was used [11].

4. Results and discussion

Figures 1 and 2 show the segregation profiles for temperatures 823 K and 1073 K respectively representing the segregation at lower temperatures and at the higher temperatures. Three elements were observed for the temperatures 873 K, 923 K and 973 K, namely C, S and Fe. Looking at figure 1, initially the segregation of C is observed, after some time the S starts segregating and is dominating the surface. This segregation profile can be explained by looking at the diffusion mechanisms of C and S. C is known to segregate via an interstitial diffusion mechanism [12] which is a faster mechanism than the substitutional mechanism of S [13]. The dominance of S on the surface can be explained at the hand of the segregation energies obtained from literature, for C this value is -85kJ/mol and for S it is -190kJ/mol [5, 6]. Thus the segregation of S is energetically more favourable that the segregation of C.

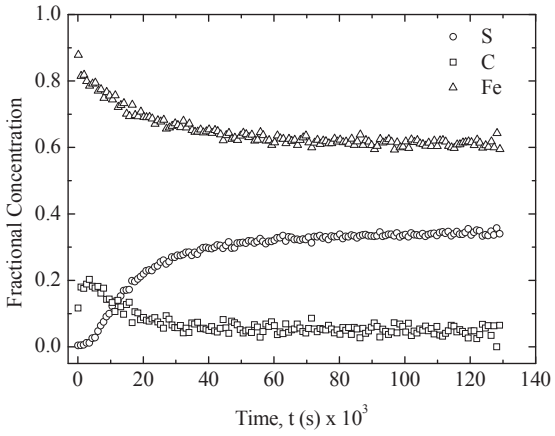


Figure 1: Segregation profile of the Fe sample for constant temperature heating at 823 K.

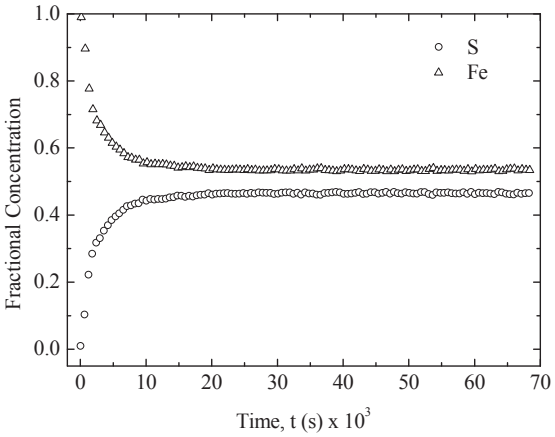


Figure 2: Segregation profile of the Fe sample for constant temperature heating at 1073 K.

Figures 3-5 shows the best fit of equation 2 to the kinetic part of the segregation profiles. The diffusion parameters used for the fit was obtained from literature [5].

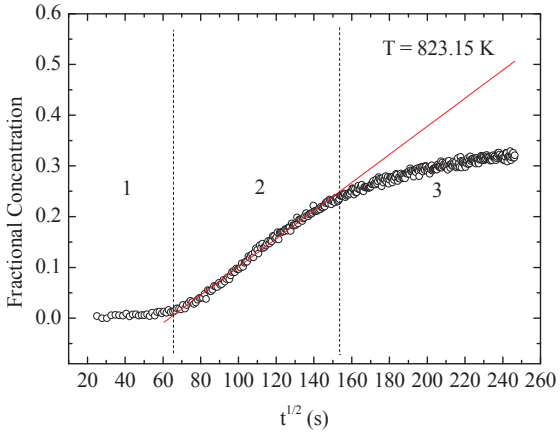


Figure 3: Linear fit to the segregation profile of S obtained at a temperature of 823 K. Three regions in the segregation profile can be distinguished: the first region as a result of sputtering, the second is the kinetic part of the segregation run and the third is the region in which segregation slows down.

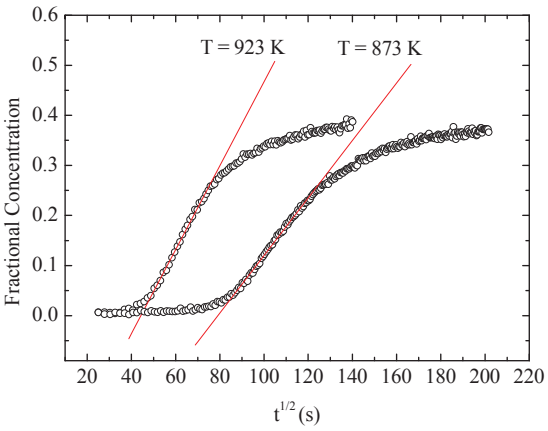


Figure 4: Linear fits to the segregation profiles of S obtained at temperatures 873 K and 923 K.

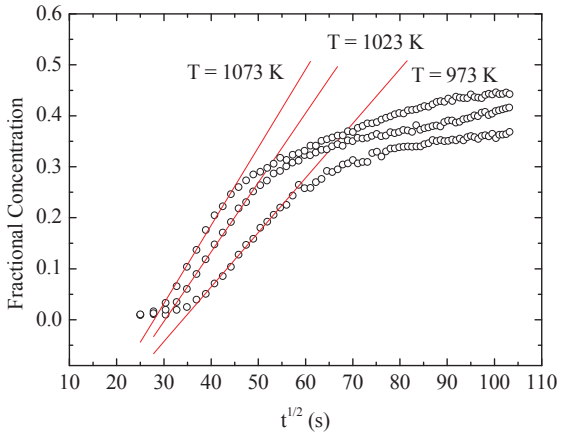


Figure 5: Linear fits to the segregation profiles of S obtained at temperatures 973 K, 1023 K and 1073 K respectively. An increased gradient of the linear fit is observed as the temperature is increased from one segregation run to the next, indicative of a larger diffusion coefficient for the higher temperatures.

Looking at the data, there are three distinctive regions, illustrated in figure 3: The first region is as a result of sputtering, during the sputtering process a depletion layer is created beneath the surface. As the S starts to segregate out from the bulk, this depletion region is filled, once filled the surface is populated with the segregating S. The second region is marked by an increasing S concentration, immediately after sputtering the Fe surface is poor in S, but as more and more S segregates to the surface an increasing rate is observed. This region is the rate determining region in the segregation run, and also the region in which the fitting of equation 2 was performed. The third region is seen as a plato area, this is where the surface has reached a maximum S coverage and the segregation process has reached an equilibrium state. From figures 4 and 5, an increased gradient of the fits can be seen as the temperature is increased, indicating an increase in the diffusion coefficient as expected. Bulk concentration values as extracted numerically for each of the temperatures are given in table 1 below.

Table 1: Values for the bulk concentration of S in the Fe-S sample, as numerically extracted from experimental data using the adapted $t^{\frac{1}{2}}$ equation describing segregation.

Temperature, T (K)	Bulk concentration, C_B (ppm)
823	17.10 ± 0.1500
873	19.40 ± 0.1528
923	16.40 ± 0.2278
973	12.90 ± 0.2200
1023	10.60 ± 0.1990
1073	7.99 ± 0.2181
Average bulk concentration	**14.07 ± 0.1946**

5. Conclusion

The use of the $t^{\frac{1}{2}}$ equation to describe the kinetic part of a segregation profile has been fitted in the linear form for temperatures in the range 823−1073 K increased in 50 K increments. Using values for the diffusion parameters obtained from literature, the concentration of S as an impurity in the Fe was determined as 14.07 ± 0.1946 ppm. As expected, an increase in the gradient of the fit was observed going from one temperature to the next higher temperature, indicating an increased diffusion coefficient. At 873 K, the segregation of both C and S was observed. At first the C segregates due to its faster diffusion mechanism which follows an interstitial path. After sufficient time the substitutional diffusing S segregates to the surface and dominates the surface due to its large segregation energy of -190 kJ/mol over the segregation energy of -85 kJ/mol for C as obtained from literature. At 1073K only the segregation of S was observed on the surface, this illustrates the dominance of S on the surface at high temperatures.

Acknowledgments

The authors would like to thank Sasol, the National Research Foundation (NRF) and the University of the Free State for financial support of this work.

References

[1] Rez P and Alvarez J R 1999 *Acta Mater.* **47** 4069-75
[2] Braithwaite J S and Rez P 2005 *Acta Mater.* **53** 2715-26
[3] Arabczyk W, Moszyński D, Narkiewicz U, Pelka R and Podsiadły M 2007 *Catal. Today* **124** 43-8
[4] Kritzinger J A 2002 *Catal. Today* **71** 307-18
[5] Reichl B M, Eisl M M, Weiss T and Störi H 1994 *Surf. Sci.* **331-333** 243-8.
[6] Dowben P A and Miller A 1990 *Surface segregation phenomena* (Florida: CRC press) pp 233-238
[7] du Plessis J, Viljoen P E and van Wyk G N 1991 *Surf. Sci.* **244** 277-84
[8] Terblans J J and Swart H C 2009 *e-J. Surf. Sci. Nanotech.* **7** 480-5.
[9] ed Briggs D and Seah M P vol 1 *Practical Surface Analysis* 1983 (New York: John Wiley & Sons) pp 211-212
[10] Shimizu R 1983 *Japan. J. Appl. Phys.* **22** 1631-42
[11] Tanuma S, Powell C J and Penn D R 1993 (IMFPs) *Surf. Interface Anal.* **20** 77-89
[12] Shewmon P G 1963 *Diffusion in Solids* (New York: McGraw-Hill) pp 43-44
[13] Hong S Y and Anderson A F 1988 *Phys. Rev.* B **38** 9417-24

Investigation of Sulphur diffusion in bulk Iron – A DFT study

P E Barnard[1], J J Terblans[1], H C Swart[1] , B G Anderson[2], M J H Hoffman[1]

[1]Department of Physics, University of the Free State, P. O. Box 339, Bloemfontein, ZA-9300, South Africa
[2] Sasol Technology R&D, P. O. Box 1, 1 Klasie Havenga Road, Sasolburg, ZA-1947, South Africa

E-mail: terblansjj@ufs.ac.za

Abstract. The diffusion of Sulphur (S) in bulk Iron (Fe) was investigated by means of Density Functional Theory (DFT) using the QUANTUM Espresso code. The ground state properties of Fe was accurately described by the PW91 pseudopotential, delivering a lattice parameter of 2.852 Å, a bulk modulus of 154.4 GPa and a magnetic moment of 2.47 μB. The vacancy formation energy and binding energy of Fe was calculated as -2.658 eV and -4.878 eV respectively. Binding energies for the interstitial and substitutional positions of S in Fe were calculated as -1.660 eV and -3.605 eV respectively. The minimum energy path and migration energy of S was calculated using the Climbing Image Nudged Elastic Band (CI-NEB) method. The larger substitutional binding energy of S indicates a substitutional or interstitial-substitutional dissociative mechanism of diffusion. It was found that S prefers the highly coordinated sites (positions were S can from a large number of bonds with Fe) in the Fe lattice and will diffuse via a interstitial-substitutional dissociative diffusion mechanism. The migration energy and activation energy for the diffusion of S in bcc Fe was calculated as -1.003 eV and -3.661 eV respectively.

1. Introduction

The valuable properties offered by Fe have made it a popular metal in industry for use in the manufacturing of alloys and also as an catalyst. The biggest problem faced when working with Fe is the impurities sulphur (S) and phosphorus (P) in the metal, which causes a weakening of both mechanical [1, 2] and catalytic properties [3, 4]. Fe exposed to high temperatures results in the segregation of these impurities to the grain boundaries, leading to grain boundary embrittlement [1, 2]. The presence of S in the Fe catalyst causes the catalyst to deactivate, as the S occupies positions on the surface [3, 4]. The Fe-S system has received a vast amount of attention, in order to determine the diffusion properties of the system. Theoretical studies have been performed by Hong *et al* [5] and Tsuru *et al.* [6] to determine the segregation properties of S in bulk Fe. The question of the diffusion mechanism remains unclear, with some authors claiming a interstitial and others a substitutional diffusion mechanism [5, 6]. The study presented here will investigate the possibility of a interstitial, substitutional and a interstitial-substitutional dissociative diffusion mechanism. Properties of interest such as the binding energies, vacancy formation energies and the migration energies are calculated.

2. Theory

The Nudged Elastic Band Method (NEB) [7] is used to determine the minimum energy path (MEP), for a transition from one stable state to the next. A number of images are chosen along a path, from the initial to the final state, usually a linear initial path is chosen. The atom then moves along the path according to the coordinates given by the images and is optimized with respect to the NEB force given by equation 1

$$\overline{F}_i^{NEB} = \overline{F}_i^{\perp} + \overline{F}_i^{S||} \tag{1}$$

where $F_i^{\perp}$ is the force projected perpendicular to the elastic band between two images, and is called the true force. This is the force experienced by the atom as a result of the potential energy surface in which it is located and is given by

$$\overline{F}_i^{\perp} = -\nabla \overline{R}_i + \nabla \overline{R}_i \bullet \hat{\tau}_i \hat{\tau}_i \tag{2}$$

The force $\overline{F}_i^{S||}$ is the spring force parallel to the elastic band, experienced by the images as a result of the elastic band. Equation 3 gives the expression for the parallel spring force

$$\overline{F}_i^{S||} = k\left(\left|\overline{R}_{i+1} - \overline{R}_i\right| - \left|\overline{R}_i - \overline{R}_{i-1}\right|\right)\hat{\tau}_i \tag{3}$$

where $\hat{\tau}_i$ is the tangent along the direction, $\overline{R}$, of the path from one image to a neighbouring image of higher energy and *k* is the elastic constant.

The elastic band is optimized in order to obtain the minimum energy path for the transition, this is done by minimizing the total force given by equation 1. The CI-NEB method [8] used in the calculations done here, is further modified by allowing the image of highest energy, *l*, to climb up the saddle point via a reflection in the force along the tangent $\hat{\tau}_l$. The force experienced by atom *l* is given by

$$\overline{F}_i^{CI} = \overline{F}_l - 2\overline{F}_l \bullet \hat{\tau}_l \hat{\tau}_l \tag{4}$$

This image does not experience any spring forces and can climb freely along the saddle point. The CI-NEB method as implemented in the QUANTUM Espresso code was used in all calculations to calculate the migration energy of diffusion.

3. Computational Details

All calculations were performed using the QUANTUM Espresso code [9], which performs fully self-consistent DFT calculations to solve the Kohn–Sham equations [10] within the generalized-gradient spin approximation (GGSA), using the functional of Perdew and Wang (PW91) [11]. The electronic wavefunctions are expanded as linear combinations of plane waves, truncated to include only plane waves with kinetic energies below the energy cut-off, Ecut, of 28 Ry. Core electrons are replaced by

ultrasoft pseudopotentials [12] to render the computations tractable as well as to enhance efficiency. k-space sampling is performed using a Monkhorst-Pack mesh [13] of 8x8x8 for all single unit cell calculations and a mesh of 3x3x3 for all calculations on the 3x3x3 supercell and for surfaces in the z direction a 3x3x1 mesh was used. For metallic smearing, the scheme of Methfessel and Paxton [14] with a degauss value of 0.04 Ry was used. A starting value for the magnetization resulting in the lowest energy, 0.4 was used to include the effects of magnetism. To ensure that the calculations for the supercells are reliable, the bulk properties have been converged to $< 2\times10^{-4}$ eV/atom with respect to the computational parameter: Ecut, k-point sampling, smearing and degauss on a single unit cell.

For the CI-NEB calculations a 6 image initial path was chosen and atomic positions were relaxed according to the Broyden scheme along this path. The maximum and minimum values for the elastic constant, *k*, was chosen as 0.1542 a.u and 0.0366 a.u respectively. This choice of a variable elastic constant increases the resolution at the transitions state with sufficient elasticity to avoid constraining the elastic band. The convergence criteria for the CI-NEB calculations was chosen as 0.05 eV/Å.

4. Results and discussion

4.1. Fe bulk properties

Convergence with respect to the computational parameters stated above, resulted in the ground state properties of bcc Fe shown in column 1of table 1, column 2 is the calculated values obtained from literature and column 3 contains the experimental values.

Table 1: Bulk properties of bcc Fe, values in column 1 is that of the current work and column 2 and 3 is values calculated by other authors and experimental results respectively.

	Calculated in this work	Literature [15]	Experimental [15]
Lattice parameter (Å)	2.852	2.869	2.866
Bulk modulus (GPa)	154.4	140	168
Magnetic moment (μB)	2.47	2.37	2.22

4.2. Binding and Vacancy formation energy

For the 3x3x3 supercell of Fe, a binding energy of -4.878 eV was obtained, which is in good agreement with the literature value of -4.893 [17] and the experimental value of -4.28 [16]. The vacancy formation in Fe was simulated as the formation of a Schottky defect [18, 19] on a Fe(100) surface. The first 5 layers of the 3x3x3 supercell was allowed to relax while keeping the 2 bottom layers fixed. To eliminate interactions between one surface and the bulk of the next periodic cell, a vacuum spacing of 28 Bohr was used. The calculation resulted in a vacancy formation energy of -2.658 eV which is in good agreement with the values obtained from literature -1.7 – -2.33 eV [20, 21, 22, 23] and the experimental value of -2eV [20]. The binding energies for interstitial and substitutional S in bcc Fe was calculated as -1.660 eV and -3.605 eV respectively, which is in good agreement with the literature values. Tsuru *et al*. calculated a interstitial binding energy of -1.53 eV [6], while a substitutional binding energy of -3.10 eV was calculated by Hong *et al.*[5].

4.3. Migration and Activation energy

The large substitutional binding energy compared to the interstitial binding energy of S indicates a substitutional or interstitial-substitutional diffusion mechanism. The substitutional diffusion of S in Fe, whereby a S atom will diffuse from a substitutional position to a nearest neighbour vacancy, was calculated as an unfavourable diffusion path. Instead of diffusing towards the vacancy, the S atom would diffuse in a direction were it can form a large number of bonds with the Fe lattice, thus S prefers a high coordination number (large number of bonds) with the Fe lattice.

The interstitial-substitutional dissociative diffusion mechanism, whereby the substitutional S atom diffuses along a interstitial path to a next nearest neighbour vacancy, has been shown to be the diffusion mechanism whereby S will diffuse. This reaction path ensures that a large number of coordination sites is available to the diffusing S atom. Also the S atom is small enough so that it can easily diffuse via an interstitial path, with only a small distortion in the surrounding lattice atoms. Using the CI-NEB method, a migration energy of -1.003 eV was calculated and taking the vacancy formation energy of -2.658 eV into account, a activation energy of -3.661 eV was calculated for the diffusion of S in bcc Fe. The migration energy of the S atom as a function of the reaction coordinates along with the crystal structures of each image is given in figure 1.

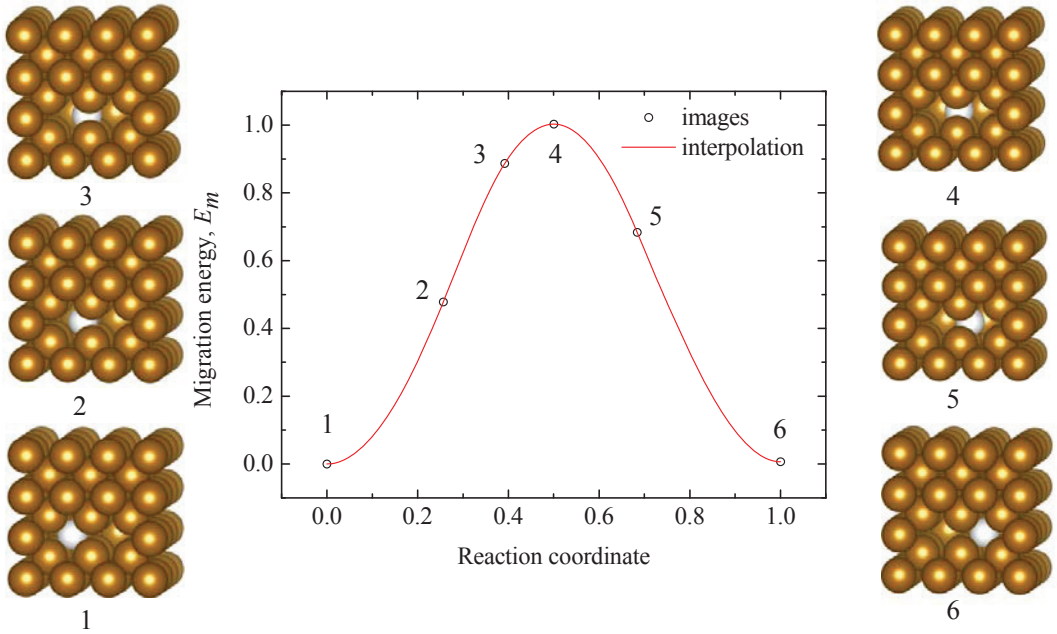


Figure 1: The migration energy as a function of the reaction coordinates for the interstitial-substitutional dissociative diffusion mechanism of S in bcc Fe. Crystal structure of each image is included for illustration [24].

5. Conclusion

The diffusion of S in bcc Bulk Fe has been investigated using the computational technique DFT. The large binding energy of S (-3.605 eV) in a substitutional lattice site over the interstitial binding energy (-1.660 eV) indicated a higher probability of S diffusing via a substitutional or interstitial-substitutional mechanism. The diffusion of S via the substitutional mechanism was found to be energetically unfavourable, due to the lack of coordination sites for the S atom along a substitutional path. The interstitial-substitutional mechanism provides such a path, with a migration energy of -1.003 eV and taking the vacancy formation energy of -2.658 eV into account, a activation energy of -3.661 eV for the diffusion of S in Fe was obtained.

Acknowledgements

The authors would like to thank Sasol, the National Research Foundation (NRF) and the University of the Free State for financial support of this work. A special thanks to Jan van der Berg and his colleagues at Sasol, for useful discussions on DFT modelling.

References

[1] Rez P and Alvarez J R 1999 *Acta Mater.* **47** 4069-75
[2] Braithwaite J S and Rez P 2005 *Acta Mater.* **53** 2715-26
[3] Arabczyk W, Moszyński D, Narkiewicz U, Pelka R and Podsiadły M 2007 *Catal. today* **124** 43-8
[4] Kritzinger J A 2002 *Catal. Today* **71** 307-18
[5] Hong S Y and Anderson A B 1988 *Phys. Rev. B* **38** 9417-24
[6] Tsuru T, Suzuki C, Kaji Y and Tsukada T 2010 *J. Appl. Phys*. **107** 061805
[7] Sheppard D, Terrell R and Henkelman G 2008 *J. Chem. Phys.* **128** 134106
[8] Henkelman G, Uberuaga B P and Jónsson H 2000 *J. Chem. Phys.* **113** 9901-4
[9] Giannozzi P, Baroni S, Bonini N, Calandra M, Car R, Cavazzoni C, Ceresoli D, Le Chiarotti G, Cococcioni M, Dabo I, Del Corso A, de Gironcolli S, Fabris S, Fratesi G, Gebauer R, Gertsmann U, Gougoussis C, Kakolj A, Lazzeri M, Martin-Samos L, Marzari N, Mauri F, Mazzarello R, Paolini S, Pasquarello A, Paulatto L, Sbraccia C, Scandalo S, Sclauzero G, Seitsonen A P, Smogunov A, Umari P and Wentzcovitch R M 2009 *J. Phys. Cond. Mat.* **21** 1-19
[10] Kohn W and Sham L J 1965 *Phys. Rev. B* **140** 1133-8
[11] Perdew J P and Wang Y 1992 *Phys. Rev. B* **45** 13244-9
[12] Vandebilt D, 1990 *Phys. Rev.B* **41** 7892-5
[13] Monkhorst H J and Pack J D 1976 *Phys. Rev. B* **13** 5188-92
[14] Methfessel M and Paxton A T 1989 *Phys. Rev. B* **40** 3616-21
[15] Hung A, Yarovsky I, Muscat J, Russo S, Snook I and Watts R O 2002 *Surf. Sci.* **501** 261-9
[16] Kittel C 2005 *Introduction to Solid State Physics* (New York: John Wiley & sons) p 50
[17] Coa W 2009 *Adsorption of surface active elements on the iron (100) surface* (Stockholm: Department of Material Science and Engineering) p 16
[18] Terblans J J 2002 *Surf. Interface Anal*. **33** 767-70

[19] Terblans J J 2003 *Surf. Interface Anal.* **35** 548-51
[20] ed Yip S 2005 *Handbook of Materials Modelling* (Dordrecht: Springer) p 281
[21] Sato K, Ihara T, Sakurai H, Yoshiie T and Xu Q 2009 *Comp. Mater. Sci*. **47** 521-5
[22] Olsson P A T 2009 *Comp. Mater. Sci*. **47** 135-45
[23] Mendelev M I, Han S and Srolovitz D J 2003 *Philos. Mag.* **83** 3977-94
[24] Momma K and Izumi F 2008 *J. Appl. Crystallogr.* **41** 653-8

Energetic stability and charge states of 3*d* transition metals in diamond: Towards a diamond based diluted magnetic semiconductor

E.M. Benecha[1] and E.B. Lombardi[1]

[1]*Department of Physics, University of South Africa (UNISA), P.O Box 392, UNISA 0003, Pretoria, South Africa*

E-mail: ebenecha@gmail.com

Abstract. Energetic stability of magnetically ordering dopants in diamond holds the prospect of achieving a diamond based diluted magnetic semiconductor which, in addition to diamond's extreme properties may successfully be considered for spintronic device applications. However, one of the problems to be addressed is to predict the energetically most stable lattice configurations or charge states in which such dopants will remain electrically or optically active, and at the same time induce magnetic moments when incorporated into diamond. We report *ab initio* DFT calculations on the formation energies of isolated 3*d* transition metal impurities at various lattice sites and charge states in diamond, and show that these impurities introduce deep donor and acceptor levels in the band gap of diamond. We further show that their formation energies and magnetic ordering properties in diamond are critically dependent on the position of the Fermi level in the diamond band gap. The formation energies across the 3*d* series are lower in *n*-type or *p*-type diamond compared to intrinsic diamond, thus showing that co-doping 3*d* transition metals with other impurities such as boron, nitrogen or phosphorus will considerably enhance their energetic stability in diamond.

1. Introduction

Understanding energetic properties of defects in semiconductors is fundamental for the development of semiconductor based technology. In particular, the magnetic properties of Diluted Magnetic Semiconductors (DMS) depend on the energetic positions of the induced defect levels which arise from interactions between the impurity and the host semiconductor. The energy of the induced defect levels in the host semiconductor's band gap, together with the on-site correlation energy, determine the electrical activity of an impurity defect and the behaviour of its charge or spin states in the presence of other impurities within the semiconductor matrix [1].

Diamond based DMSs have been predicted to possess high Curie temperatures and are expected to offer superior performance in high temperature, high power and high frequency applications due to diamond's excellent physical and electrical properties [2]. In particular, diamond's high intrinsic hole and electron mobilities (3800 $cm^2V^{-1}s^{-1}$ and 4500 $cm^2V^{-1}s^{-1}$, respectively [3]), thermal conductivity (22 $Wcm^{-1}K^{-1}$ [4]) and breakdown field (10000 $KVcm^{-1}$ [5]), compared to other semiconductors [4,6], make it an ideal material for fabrication of semiconducting devices with a wide range of applications.

For successful application of diamond in spintronic devices, it is essential to understand the properties of candidate magnetically ordering defects, such as transition metals, in diamond, as well as

their formation energies and spin properties in different charge states and in different types of doped diamond.

Impurity defects in semiconductors occur in multiple charge states, with different charge states corresponding to distinct induced magnetic states. Transitions between different charge states of the defect centre often form the basis for experimental detection of the defects, and therefore prediction of the induced defect levels in the host semiconductor's band gap is important for the understanding and development of a diamond based DMS [7]. Further, the influence of other common dopants in diamond, such as B, P or N, needs to be considered, since these dopants will influence the position of the Fermi level in the diamond host, impacting the charge state and other properties of these transition metal defects.

Therefore, an important priority in the development of diamond based spintronic applications is to determine which charge states of the impurity defects are stable under specific doping conditions, while at the same time exhibiting useful magnetic ordering, since each charge state may correspond to a different spin configuration, some of which may order ferromagnetically in their ground state, while others not.

Previous studies on the energetic stability of transition metal defects in diamond have focused mostly on those transition metals (Mn, Fe, Ni, Co, Cr) which are commonly used as catalysts during High Pressure-High Temperature (HP-HT) growth of synthetic diamond. However, no explicit explanation exists as to why some of these transition metals are experimentally found to be preferentially incorporated into diamond containing other impurities, such as nitrogen or boron [2]. Although complexes of transition metals with these co-dopants may form in diamond during growth or after annealing, it has previously been noted that [8,9] their role is to shift the Fermi level of the system, therefore influencing the charge states of these defects in the diamond crystal. In this study, we determine the formation energies of isolated 3*d* transition metal defects in diamond at various lattice sites and charge states, and demonstrate that their energetic stability depends on the charge states, and the type of diamond doping (i.e. intrinsic, *p*-type or *n*-type). Their spin configurations and magnetic ordering are also determined, allowing us to predict candidate transition metals which may be used as magnetically ordering dopants in the development of diamond based spintronic applications.

2. Method

Ab initio Density Functional Theory (DFT) calculations have been performed for 3*d* transition metal-doped diamond using a pseudopotential plane wave approach, as implemented in the CASTEP computational code [10]. The generalized gradient approximation (GGA) of Perdew-Burke-Ernzerhof (PBE) [11] was used, with Ultrasoft Vanderbilt pseudopotentials [12] in reciprocal space representation, and a plane wave cut-off energy of 310 eV. Optimized $4 \times 4 \times 4$ Monkhorst-Pack grid sampling [13] for integration over the Brillouin zone (32 k-points in the irreducible wedge of the Brillouin zone) was employed for geometry optimization, together with a $48 \times 48 \times 48$ FFT grid for electron density representation [14]. Increasing the number of k-points or plane wave cut-off energy resulted in insignificant changes in total energy and structural relaxation (less than 10^{-4} eV and 10^{-3} Å, respectively).

All calculations were carried out using a 64-atom diamond supercell constructed from $2 \times 2 \times 2$ conventional fcc unit cells with an optimized lattice constant of 3.569 Å, which is in close agreement with the experimental value of 3.567 Å [15]. Full geometry optimization was performed without any symmetry or spin restrictions, with the transition metal (TM) atom in different charge states: (q = +2, +1, 0,-1, -2) placed at the divacancy (TM_{2V}), substitutional (TM_S), tetrahedral interstitial ($TM_{Td\text{-}i}$) or hexagonal interstitial ($TM_{Hx\text{-}i}$) sites in the diamond supercell. For each lattice site and charge state, various initial positions and spins of the TM atom were considered in order to systematically determine the most energetically stable configurations with respect to geometry and spin state.

To determine the energetically most favourable lattice site for the TM atoms in diamond, their formation energies (E_f) at each charge state and lattice site in the diamond supercell were calculated using [7]:

$$E_f[TM]^q = E_T[n_C, n_{TM}]^q - n_C\mu_C - n_{TM}\mu_{TM} + q(\varepsilon_V + \varepsilon_F)$$

where $E_T[n_C, n_{TM}]^q$ is the total energy of the supercell in a charge state q containing n_C carbon and n_{TM} transition metal atoms, with chemical potentials μ_C and μ_{TM}, respectively. The Fermi energy ε_F is measured relative to the energy of valence band maximum ε_V. The chemical potential of each atomic species was determined from their respective pure crystalline phases [16,17].

3. Results and discussion

3.1 Formation energy and chemical trends

Figure 1 presents the calculated formation energies of isolated 3*d* transition metals in the neutral charge state at substitutional and divacancy lattice sites in intrinsic diamond, as well as the predicted formation energies of the most stable charge states in *p*-type boron doped (E_V + 0.37 eV), *n*-type nitrogen doped (E_C − 1.6 eV) and *n*-type phosphorus doped (E_C − 0.6 eV) diamond. We find the majority of transition metals to be most stable at the divacancy site compared to the substitutional or interstitial sites. Transition metals at the interstitial sites are highly unstable for any position of the Fermi level, indicating that the interstitial species are unlikely to be observed in significant concentrations. A distinctive trend of the calculated formation energies is seen across the 3*d* series, with transition metals in the middle of the series predicted to be energetically most stable compared to those early or late in the series – a result which can be explained by the electron filling of the *d* orbital [1] and the relatively large free atomic radius of transition metal atoms compared to diamond's lattice constant.

Importantly, the formation energy at any lattice site is strongly dependent on the charge state of the transition metal impurity or the type of diamond doping (intrinsic, *p*-type or *n*-type) with reference to the pinning of the Fermi level by co-doping with B, N or P, as illustrated in Figure 2 for the case of Fe. We find that all the 3*d* transition metals have lower formation energies in *n*-type or *p*-type diamond compared to intrinsic diamond. The formation energies of transition metals in the double negative charge state in *n*-type diamond are predicted to be considerably lower than the neutral charge state, consistent with experimental results where transition metal ions are often observed in diamond containing nitrogen (which is a deep donor in diamond) [2].

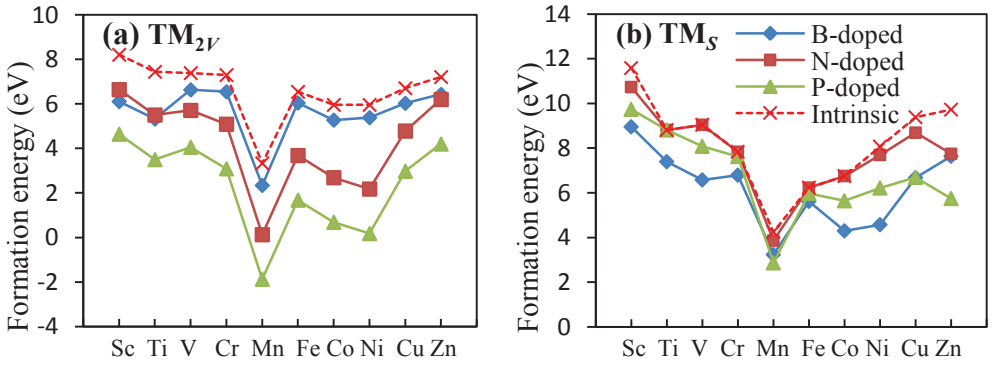


Figure 1. Predicted formation energies of 3*d* transition metals in intrinsic diamond as well as in *p*-type B doped (E_V + 0.37 eV), *n*-type N doped (E_C – 1.6 eV) and P doped (E_C – 0.6 eV) diamond at (a) divacancy, TM_{2V} and (b) Substitutional, TM_S lattice sites. Data points correspond to the calculated formation energies, while lines show the trend.

In intrinsic and *p*-type B doped diamond, the formation energy difference between substitutional and divacancy V, Cr, Mn, Fe and Co is relatively small (~0.5 eV), implying that these transition metals may be observed at both of these sites. This result is consistent with recent experimental data for Fe implanted diamond [18] in which 65 % of Fe ions were found to be located at the substitutional site, in agreement with our prediction of $(Fe_S)^{+1}$ being more stable by ~0.4 eV compared to $(Fe_{2V})^{+1}$ in *p*-type B-doped diamond (Figure 2).

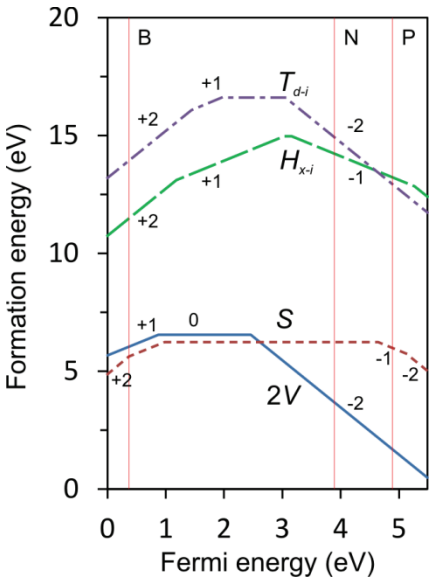


Figure 2. Formation energies of Fe in diamond as a function of the Fermi level, at divacancy (*2V*), substitutional (*s*), tetrahedral interstitial (T_{d-i}) and hexagonal interstitial (H_{x-i}) lattice sites. Only line-segments corresponding to respective lowest-energy charge states as a function of Fermi level are shown; the slope of these segments indicates the charge state (q = +2, +1, 0, -1, -2); changes in slope correspond to charge transition levels (double donor (++/+), donor (+/0), acceptor (0/-), ..., levels). The vertical lines indicate the position at which the Fermi level would be pinned if diamond samples were co-doped with Boron (*Ev* + 0.37 eV), Nitrogen (*Ec* -1.6 eV) or Phosphorus (*Ec* - 0.6 eV). Note the corresponding charge states Fe would assume in diamond co-doped with these dopants.

3.2 Defect levels

Figure 3 illustrates the thermodynamic charge transition levels of 3*d* transition metals in diamond with respect to the valence band maximum and conduction band minimum at divacancy and substitutional lattice sites, respectively, for the charge states q = +2, +1, 0,-1, -2. It is seen that the majority of transition metals induce deep donor and acceptor levels in the diamond band gap, thus highlighting the fact that electronic and magnetic properties of transition metals in diamond will be strongly dependent on their charge state [19]. The deep nature of these levels is particularly important in achieving highly localized and controllable quantum states well isolated from sources of decoherence, similar to the well-known $(N\text{-}V)^{-1}$ centre in diamond which is among the leading candidates currently being considered for spintronic and quantum computing applications [19,20].

In substitutional Sc, Ti, V Cr, and Mn, we find that the double acceptor (-1/-2) impurity levels occur as resonances inside the conduction band; the bands associated with these levels may hybridize with shallow donors, thus potentially influencing magnetic ordering properties, similar to theoretical predictions in transition metal-doped ZnO [21].

To determine which 3*d* transition metals and charge states are likely to lead to collective magnetic ordering in diamond, Table 1 summarizes the induced magnetic moments and point symmetries of the most stable charge states in different types of diamond doping as given in Figure 1. Importantly, the magnitude of the induced magnetic moment generally depends on the doping site and charge state. The majority of the 3*d* transition metals at any lattice site induce non-vanishing magnetic moments, thus indicating that incorporation of these impurities into diamond may lead to collective magnetic ordering [22]. However, to determine the nature of this magnetic ordering, impurity-impurity interaction also need to be taken into account in order to obtain the magnetic ground states and magnetic stabilization energies for the most stable charge states. Previous theoretical studies on the magnetic ordering of transition metals in diamond have predicted that ferromagnetic ordering in Mn-

doped diamond is unlikely [16], while Co has been predicted to order ferromagnetically with a small magnetic moment of $0.4\mu_B$ per Co atom [23].

Based on the results of this study we predict that V, Cr, and Fe may successfully be considered as candidate dopants for creation of a diamond based DMS, due to their relatively low formation energies, implying that they may be incorporated into diamond in reasonable concentrations, together with non-zero magnetic moments in most of their stable charge states (Table 1). We predict that $(V_{2V})^{+2}$ and $(Fe_S)^{+2}$ will be of particular interest since, in addition to their favourable formation energies, they also possess integer magnetic moments per impurity ion – which is a necessary condition for half-metallic ferromagnetic ordering [24]. Hence these defects may attain up to 100% spin polarisation when doped into diamond.

Table 1. Induced magnetic moments together with the point symmetries for the most stable charge states (q) of 3d transition metals in diamond at Fermi levels corresponding to p-type B doped (E_V + 0.37 eV), n-type N doped ($E_C - 1.6$ eV) and P doped ($E_C - 0.6$ eV) diamond at (a) divacancy and (b) substitutional lattice sites.

(a) DIVACANCY SITE

	Charge state, Magnetic Moment (μ_B) and Symmetry		
	B-doped	N-doped	P-doped
Sc	(+2) 0.1 D_{3d}	(-2) 0.6 D_{3d}	(-2) 0.6 D_{3d}
Ti	(+2) 0.0 D_{3d}	(-2) 0.0 D_{3d}	(-2) 0.0 D_{3d}
V	(+2) 1.0 C_{2h}	(-1) 0.0 C_2	(-2) 1.2 D_{3d}
Cr	(+2) 2.5 D_{3d}	(-2) 2.5 D_{3d}	(-2) 2.5 D_{3d}
Mn	(+2) 3.0 D_{3d}	(-2) 3.0 D_{3d}	(-2) 3.0 D_{3d}
Fe	(+1) 1.0 D_{3d}	(-2) 2.0 D_{3d}	(-2) 2.0 D_{3d}
Co	(+1) 0.0 C_s	(-2) 1.0 C_2	(-2) 1.0 C_2
Ni	(+2) 0.0 D_{3d}	(-2) 0.0 D_{3d}	(-2) 0.0 D_{3d}
Cu	(+2) 0.3 D_{3d}	(-1) 0.0 D_{3d}	(-2) 1.0 D_{3d}
Zn	(+2) 0.1 D_{3d}	(-2) 1.8 D_{3d}	(-2) 1.8 D_{3d}

(b) SUBSTITUTIONAL SITE

	Charge state, Magnetic Moment and Symmetry		
	B-doped	N-doped	P-doped
Sc	(+2) 1.0 C_{2v}	(-1) 0.0 T_d	(-1) 0.0 T_d
Ti	(+2) 0.0 C_s	(0) 0.0 C_{3v}	(-1) 1.0 T_d
V	(+2) 0.1 C_{3v}	(0) 1.0 T_d	(-1) 2.0 T_d
Cr	(+2) 0.5 C_{3v}	(0) 2.0 T_d	(-1) 1.0 T_d
Mn	(+1) 0.0 T_d	(-1) 0.0 T_d	(-1) 0.0 T_d
Fe	(+2) 1.0 T_d	(0) 0.0 T_d	(-1) 1.0 C_{3v}
Co	(+1) 0.0 T_d	(0) 1.0 C_{3v}	(-1) 0.0 T_d
Ni	(+2) 0.0 T_d	(-1) 3.0 T_d	(-2) 2.0 T_d
Cu	(+2) 0.5 T_d	(-2) 1.0 T_d	(-2) 1.0 T_d
Zn	(+2) 0.0 T_d	(-2) 1.8 T_d	(-2) 1.8 T_d

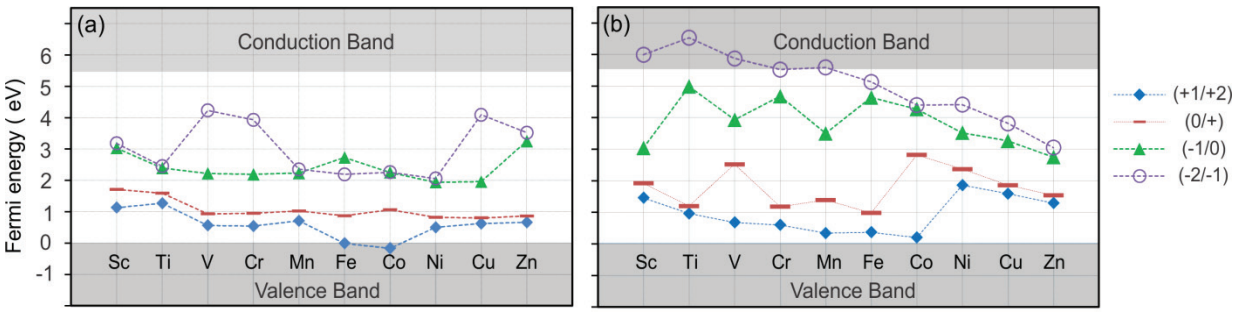


Figure 3. Thermodynamic charge transition levels of 3d transition metals at (a) divacancy and (b) substitutional lattice sites in diamond, relative to the valence band maximum and conduction band minimum.

4. Conclusion

The energetic stability of 3*d* transition metals and their defect levels in diamond at different lattice sites and charge states have been investigated using *ab initio* pseudopotential DFT methods. Significantly, the formation energies and magnetic moments of the transition metals are shown to be critically dependent on the charge state, and the type of doped diamond they are incorportated into (intrinsic, *n*-type or *p*-type). At each lattice site, the transition metal formation energies are predicted to be lower in *p*-type (B-doped) or *n*-type (N- or P-doped) diamond compared to intrinsic diamond. The majority of transition metals are found to be energetically stable at the divacancy site in any charge state, with the formation energy of transition metals in the middle of the 3*d* series being lowest at any lattice site.

Most of the transition metals are found to induce deep donor and acceptor levels in the diamond band gap, thus indicating that the electronic and magnetic properties of transition metals in diamond will be strongly dependent on their charge state. The magnitude of the induced magnetic moments is also found to depend on the doping site and charge state, with the majority of the 3*d* transition metals at any lattice site inducing non-vanishing magnetic moments, thus indicating that incorporation of these impurities into diamond may lead to collective magnetic ordering.

References

[1] Dietl T 2002 Semicond. Sci. Technol. **17** 377
[2] Sussmann R S 2009 *CVD Diamond for Electronic Devices and Sensors* (John Wiley and Sons Ltd, United Kingdom)
[3] Isberg J, Hammersberg J, Johansson E, Wikström T, Twitchen D J, Whitehead A J, Coe S E and Scarsbrook G A 2002 Science **297** 1670
[4] Wort J H and R S Balmer 2008 Mater. Today **11** 22
[5] Geis M 1990 Mat. Res. Soc. Symp. Proc. **152** 15
[6] Kalish R, 2007 J. Phys. D: Appl. Phys **40** 6467
[7] Van de Walle C G and Neugebauer J 2004 J. Appl. Phys. **95** 3851
[8] Larico R, Justo J F, Machado W V M and Assali L V C 2004 Brazilian J. of Phys. **34** 669
[9] Davies G 1999 Physica B **273** 15
[10] Clark S J, Segall M D, Pickard C J, Hasnip P J, Probert M J, Refson K and Payne M C 2005 Zeitschrift fuer Kristallographie **220** 567
[11] Perdew J P, Burke K and Ernzerhof M 1996 1990 Phys. Rev. Lett. **77** 3865
[12] Vanderbilt D 1990 Phys. Rev. B **41.**7892
[13] Monkhorst H J and Pack J D 1976 Phys. Rev. B **13** 5188
[14] Fabris S, de Girancoli S, Baroni S, Vicario G and Balducci G 2005 Phys. Rev. B **72** 237102
[15] Donnay J D H and Ondik M 1973 *Crystal Data: Determinative Tables vol 2, 3rd ed.* (US Department of Commerce - National Bureau of Standards - JCPDS)
[16] Erwin S C and Hellberg C S 2003 Phys. Rev. B **68** 245206
[17] Jeong J W and Oshiyama A 2001 Phys. Rev. B 64 235204
[18] Bharuth-Ram K, Wahl U and Correia J G 2003 Nucl. Instrum. Meth. B **206** 941
[19] Weber J R, Koehl W F, Varley J B, Janotti A, Buckley B B, Van de Walle C G and. Awschalom D D 2010 Proc. Nat. Acad. Sci. **107** 8513
[20] Awschalom D D, Epstein R and Hanson R 2007 Scientific American **297** 84
[21] Raebiger H, Lany S and Zunger A 2009 Phys. Rev. B **79** 165202
[22] Zunger A and Lany S 2010 Physics **3** 53
[23] Lombardi E B 2008 Diamond Rel. Mater. **17** 1345
[24] de Groot R A, Muller F M, Engen P G and Bushchow K H J 1983 Phys. Rev. Lett. **50** 2024

Tuning carbon nanotubes for photovoltaic applications

G. Bepete[1,2], D. Wamwangi[2], N. J. Coville[1], and Z. Chiguvare[2]

[1]DST/NRF Centre of Excellence in Strong Materials and Molecular Sciences Institute, School of Chemistry, University of the Witwatersrand, WITS 2050, Johannesburg, South Africa

[2]DST/NRF Centre of Excellence in Strong Materials and Material Physics Research Institute, School of Physics, University of the Witwatersrand, WITS 2050, Johannesburg, South Africa

Zivayi.Chiguvare@wits.ac.za

Abstract. We sought to control the diameter and nitrogen doping levels in the synthesis of nitrogen doped carbon nanotubes (N-CNTs) with the aim of incorporating these in organic solar cells. This was achieved by careful selection of catalyst, precursor mixture and chemical vapour deposition (CVD) synthesis temperature. A solution of methanol and acetonitrile (20% acetonitrile) was pyrolysed over a Fe/Al/MgO catalyst in a CVD reactor at 850 ^{o}C. The product was characterised by TEM, TGA and Raman spectroscopy. N-CNTs with well graphitised morphology, small diameter, and narrow diameter distribution were obtained. Catalyst support and metal particles were removed completely during purification without any change in morphology of the material. Preliminary organic solar devices made from P3HT and the synthesized N-CNTs gave $V_{oc} \sim 0.3$ V and $J_{sc} \sim 7$ μA/cm^2 but poor fill factors of around 30 % under AM1.5 white light intensity.

1. Introduction

Multiwalled carbon nanotubes structurally comprise of several coaxially arranged graphene cylinders of different radii with an inter-tube separation of approximately 0.34 nm. Generally multiwalled carbon nanotubes are synthesized with a range of diameters up to as big as 100 nm [1]. The thickness of the bulk heterojunction organic layer in organic photovoltaic devices is 200 nm. Thus for efficient exciton splitting in the devices the composite film between donor polymer and acceptor material should be nanostructured with less than 20 nm between the acceptor material and the donor material. Because of this, carbon nanotubes of diameters less than 20 nm are most suitable for application in organic photovoltaic devices. During the synthesis of carbon nanotubes for application in organic photovoltaic devices control of diameters then becomes a critical issue. Semiconducting carbon nanotubes are particularly attractive for use in organic photovoltaic devices [2]. It is very difficult to control whether resulting carbon nanotubes are purely metallic or semiconducting as they are both produced as a mixture during synthesis. Separation of semiconducting from metallic carbon nanotubes is a very difficult process. For single walled carbon nanotubes, whether they are metallic or semiconducting depends on their chirallity [3]. MWNTs are more complex than SWNTs because it is

not only the outer wall which determines the conducting properties of carbon nanotubes. No matter the chirallity of the outer wall MWNTs are always metallic in nature [4].

Incorporating heteroatoms in MWNT matrix results in the tuning of the electronic properties of the resulting MWNTs. Doping with boron and nitrogen is the preferred method for changing the electronic properties of carbon nanotubes. Boron doping yields purely p-type semiconducting carbon nanotubes whilst doping with nitrogen yields purely n-type semiconducting carbon nanotubes [5]. During nitrogen doping incorporating pyridinic nitrogen type in carbon nanotubes is desired. However, in a CVD reaction atmosphere it is not possible to control the type of nitrogen incorporated and hence graphitic, quaternary, oxidized nitrogen and molecular nitrogen are also incorporated within the carbon nanotubes structure.

The diameter of grown MWNTs is influenced by catalyst and temperature used during synthesis [6]. The use of ethanol has been investigated as a co-solvent in the N-CNT synthesis method and it was found that it influences the crystalinity of the produced MWNTs [7]. OH radicals are produced during the reaction and they etch away amorphous carbon material which might be deposited during growth.

We sought to control the diameter and nitrogen doping levels in the synthesis of nitrogen doped carbon nanotubes (N-CNTs) with the aim of incorporating these in organic solar cells.

2. Experimental

To make nitrogen doped carbon nanotubes N-CNTs 300 mg catalyst was placed in a quartz boat inserted in the centre of a quartz tube reactor. The catalyst used in this study was Fe/MgO and Fe/Al/MgO. The reactor was heated to 850^{o} C in 5% H_2 in Ar flow at 240 ml/min. At 850 ^{O}C the carbon and nitrogen source (20% acetonitrile in ethanol) was introduced by means of 20 ml syringe driven by a pump at 1.2 ml/min for 20 minutes. Finally the reactor was cooled to ambient temperature under a flow of 5% H_2 in Ar. The synthesized carbon nanotubes were purified by washing for several hours in dilute NaOH in an ultrasonic bath followed by washing several times with water, refluxing for several hours in hydrochloric acid at 100 ^{O}C and lastly by washing in water and drying in an oven at 120 ^{O}C for 12 hours. The morphology and quality of the as synthesized N-CNTs was characterized by Transmission Electron Microscopy (Tecnai G^2 Spirit TEM at 120 kV). TGA was also used to characterize the extent of purification of material and the stability of N-CNTs before purification and after purification, using a Perkin Elmer Pyris Thermogravimetric Analyser 1 instrument under air flow.

To fabricate devices, the ITO electrode on glass substrate (8-12 Ω/sq) was patterned by etching in acid and cleaned in an ultrasonic bath using detergent, followed by several organic solvents and lastly with distilled water. Solutions were prepared by dissolving poly 3-hexythiophene (P3HT), purchased from Sigma-Aldrich Co. and the N-CNTs (synthesized by the authors) in chlorobenzene and sonicated for ten minutes followed by stirring for twelve hours in a dark inert atmosphere. The device was fabricated by spin coating Poly (3, 4-ethylenedioxythiophene) poly (styrenesulfonate) PEDOT: PSS a hole collecting layer at 2000 rpm followed by the P3HT/N-CNT active material at 2000 rpm and lastly an aluminium top contact was deposited by evaporating in a vacuum evaporator at 10^{-5} mbar. Absorption spectra of the P3HT and the P3HT/N-CNT blend were taken using a Varian Cary UV-VIS-NIR spectrometer and I-V characteristics of the devices made were measured in the dark and under illumination (solar simulated radiation with a AM1.5 filter) using an HP 4141 Source Measure Unit.

3. Results and discussion

3.1. Characterization of N-CNTs

Chemical vapor deposition CVD method for synthesizing carbon nanotubes has the advantage of controlling the properties of as synthesized multi-walled carbon nanotubes. We were able to control the properties of the as synthesized carbon nanotubes used in this study using good catalyst selection, suitable precursor material and growth conditions. The optimum temperature for synthesizing good

quality carbon nanotubes with small diameter and low yields (<2%) of amorphous material was obtained at 850 ^{o}C. TEM studies revealed that the N-CNTs obtained contained bamboo-like morphologies. When using an iron catalyst the presence of the bamboo morphology confirms the presence of nitrogen incorporation in the carbon nanotubes matrix [8]. The pyridinic like nitrogen is the one responsible for inducing bamboo morphology in nitrogen doped carbon nanotubes [9]. The presence of the pyridinic like nitrogen in nitrogen doped carbon nanotubes introduces electrons in the form of electron lone pair from the pyridinic like nitrogen. This makes nitrogen doped carbon nanotubes purely n-type semiconducting material [5]. Figures 1 (a) and (b) show a comparison of the multi-walled N-CNTs produced using Fe/MgO and Fe/Al/MgO catalysts respectively. Fig. 1(b) revealed that Fe/Al/MgO produces N-CNTs with smaller and more uniform diameters. The role of Al is to modify the Fe/MgO surface and prevent sintering of the Fe. We suspect that the inclusion of aluminium in the catalyst material helps in the formation of smaller catalystic iron particles on the MgO catalyst support material. Smaller iron catalyst material form N-CNTs of smaller diameters compared to N-CNTs synthesized from catalyst material without aluminium.

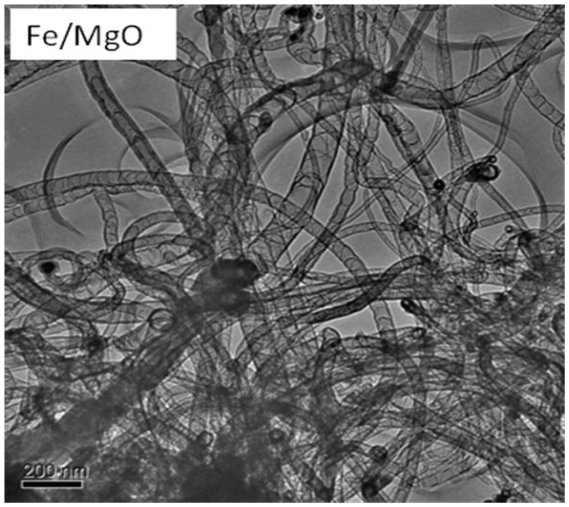


Figure 1(a). Purified N-CNTs synthesized from Fe/MgO catalyst

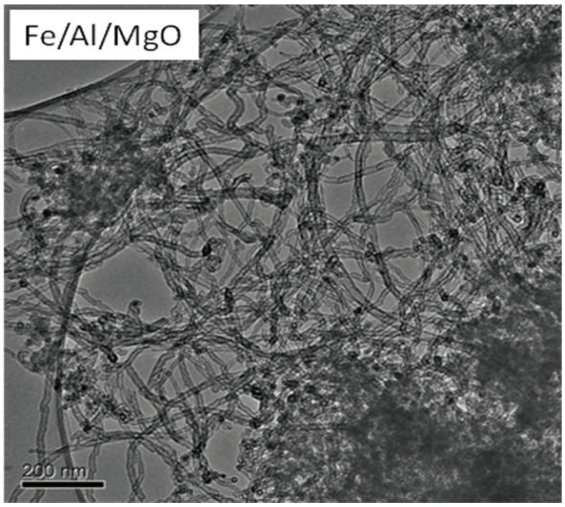


Figure 1(b). Purified N-CNTs synthesized from Fe/Al/MgO catalyst

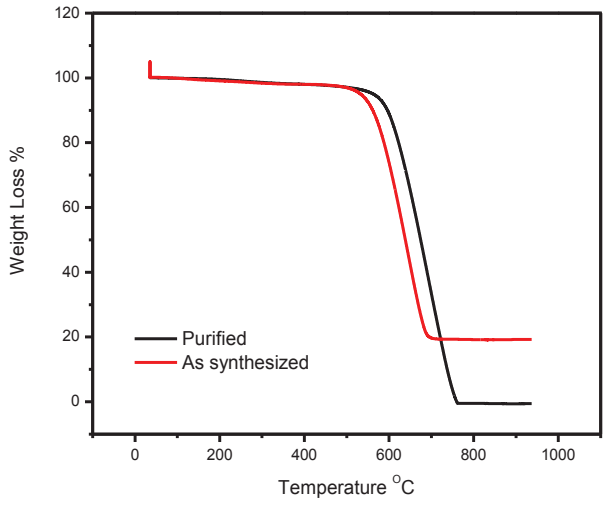


Figure 2(a). TGA of as synthesized N-CNTs from Fe/Al/MgO catalyst and N-CNTs after purification

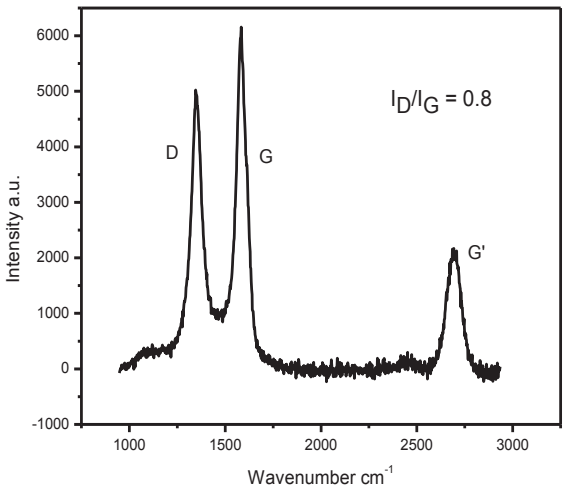


Figure 2(b). Raman spectra of N-CNTs synthesized from ethanol and acetonitrile mixture at 850 ^{o}C over a Fe/Al/MgO catalyst

From the TEM images in figure 1, very little amorphous material is seen, and thermogravimetric analysis (Fig. 2(a)) confirmed this result. The synthesized N-CNTs were thermally stable with a decomposition temperature of 650 ^{O}C. After purification using NaOH solution and HCl no additional defects were added on the N-CNT structure as shown by the similarities of the TGA profile of the as synthesized N-CNTs and the purified N-CNTs. Raman spectroscopy (figure 2(b)) showed an I_D/I_G of 0.8 for N-CNTs produced from Fe/Al/MgO showing that the N-CNTs have graphitic character.

3.2 Solar cells

The thin film photovoltaic devices fabricated were made with the architecture shown in figure 3 (a), to evaluate the influence of incorporating N-CNTs in P3HT on the device performance.

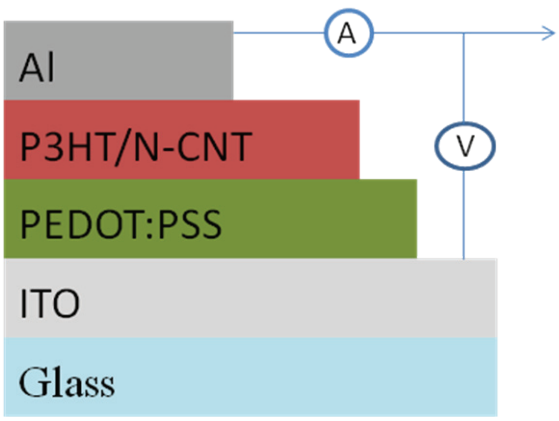


Figure 3(a). Schematic representation of fabricated P3HT/N-CNT solar cells

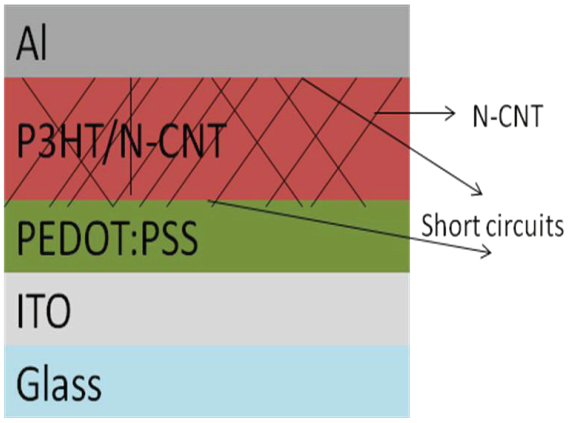


Figure 3(b). Schematic representation of possible source of poor performance in fabricated solar cells

The 1% N-CNT in P3HT and 5% N-CNT in P3HT showed a decrease in I_{SC} when N-CNTs were introduced in the active layer but both gave an increase in V_{OC}. A higher concentration of N-CNTs in the devices (more than 5 %) resulted in decrease of both the I_{SC} and the V_{OC}. The devices made from P3HT alone showed an open circuit voltage V_{OC} of 0.27 V and a short circuit current density J_{SC} of 9.1 μA/cm^2, while the devices made from P3HT/ 1% N-CNT showed V_{OC} of 0.3 V and J_{SC} of 7μA/cm (see figure 4). Incorporating N-CNTs in the P3HT active layer improved charge separation resulting in the small increase in V_{OC}. An increase in V_{OC} results from efficient exciton splitting and charge separation. The change in the open circuit voltage could be attributed to the energy offset between the HOMO (of the donor) and the LUMO (of the acceptor) levels which dissociates the photo generated excitons into free carriers at the interface of the acceptor and donor materials. The control of the energy levels is thus critical to induce efficient charge separation.

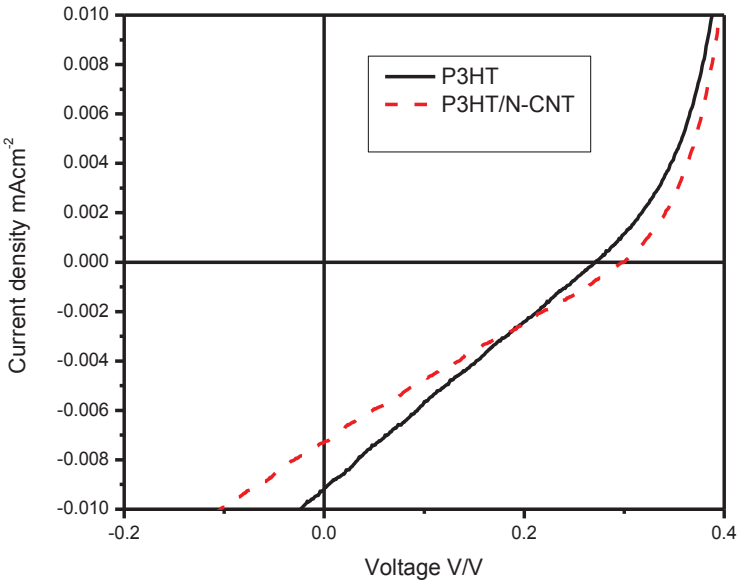


Figure 4. *J-V* characteristics of solar cells made from P3HT and P3HT/ 1% N-CNT active materials

In this work we synthesized N-CNTs which are said to be purely n-type semiconducting [5]. Interaction between the n-type N-CNTs and the p-type P3HT resulted in good electron hole pair dissociation. The I_{SC} does not behave the same way as the V_{OC}. A decrease in I_{SC} after incorporation of N-CNTs was noted. The N-CNTs used in this study were very long, generally more than 1µm long (figure1 (b)). We suspect that this resulted in them spanning both electrodes and shorting the devices as shown in figure 3(b). The photogenerated charge carriers in the N-CNTs travel to both the anode and cathode thus diminishing the photogenerated current as it is possible that the current will be flowing in opposing directions.

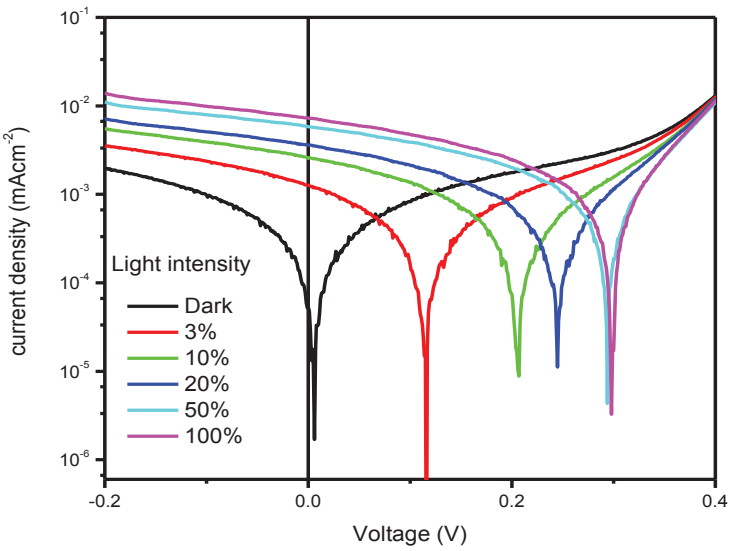


Figure 5. Semi log plot of *J-V* characteristics of solar cells made from P3HT/1% N-CNT at different light intensities

Under AM 1.5 spectral illumination with light of different intensities the P3HT/N-CNT showed a photo-response as shown in figure 5. The efficiencies so far obtained were in the range of 0.09 %. The results obtained for solar cells are comparable to Nogueira *et al.* [10], who used poly(3-octylthiophene) as the donor polymer in their photovoltaic devices. Further trials and improvement in device fabrication procedures should improve these values significantly.

4. Conclusions

The ability to control the diameter of N-CNTs using catalyst selection and carbon and nitrogen source has been demonstrated. The use of the synthesized materials in the fabrication of solar cells yielded P3HT/N-CNT solar cells with enhanced V_{OC} but diminished I_{SC} as compared to the devices made out of P3HT alone. Optimization of concentration of N-CNTs in P3HT matrix is needed to obtain solar cells with improved performance. Improved device fabrication methods and application of short N-CNTs in organic photovoltaic devices is expected to enhance performance. Shortening of N-CNTs will be done during synthesis by keeping short reaction times or after synthesis by ultrasonication for long hours.

Acknowledgements

The authors wish to thank the DST/NRF Centre of Excellence in Strong Materials and the University of the Witwatersrand for financial support.

References

[1] Shang S, Yang X and Tao X 2009 *Polymer* **50** 2815
[2] Bindl D, Safron N, and Arnold M 2010 *ACS Nano* **4** 5657
[3] Hatton R A, Miller A J and Silva S R P 2008 *J. Mater. Chem.* **18** 1183
[4] Ago H, Kugler T, Cacialli F, Salaneck W R, Shaffer M S P, Windle A P and Friend R H 1999 *J. Phys. Chem. B* **103** 8116
[5] Panchakarla L S, Govindaraj A and Rao C N R 2010 *Inorganica Chimica Acta* **363** 4163
[6] Usoltseva A, Kuznetsov V, Rudina N, Moroz E, Haluska M and Roth S 2007 *Phys. stat. sol. (b*) **11** 244
[7] Bystrzejewski M, Huczko A, Byszewski P, Domańska M, Rümmeli M H, Gemming T and Lange H 2009 *Fullerenes, Nanotubes and Carbon Nanostructures* **17** 298
[8] Sumpter B, Meunier V, Romo-Herrera J, Cruz-Silva E, Cullen D and Terrones H 2007 *ACS Nano* **1** 369
[9] Liu H, Zhang Y, Li R, Sun X, De´silets S, Abou-Rachid H, Jaidann M and Lussier L S 2010 *Carbon* **48** 1498
[10] Nogueira A F, Lomba B S, Soto-Oviedo M A, Duarte Correia C R, Corio P, Furtado C A and, Huemmelgen I A 2007 *J. Phys. Chem. C* **111** 18431

Electrical and surface morphological studies of palladium and ruthenium Schottky diodes on *n*-Ge (100)

A. Chawanda[1,2], F.D. Auret[2], W. Mtangi[2], J.M. Nel[2], M. Diale[2], and C. Nyamhere[3]

[1]Department of Physics, Midlands State University, Gweru, Zimbabwe

[2]Department of Physics, University of Pretoria, 0002, Pretoria, South Africa

[3]Department of Physics, Nelson Mandela Metropolitan University, Box 7700 , Port Elizabeth 6031, South Africa

E-mail: albert.chawanda@up.ac.za

Abstract. Palladium (Pd) and ruthenium (Ru) Schottky barrier diodes were fabricated on (100) Sb-doped *n*-type germanium using resistive evaporation and electron beam deposition systems, respectively. Electrical characterization of these diodes was performed using current-voltage (*I-V*) measurements under various annealing conditions. The morphological evolution of the surface was analysed using the scanning electron microscopy. The variation of the electrical and structural properties of these Schottky diodes are attributed to combined effects of interfacial reaction and phase transformation during the annealing process. Thermal stability of both the Pd/*n*-Ge (100) and Ru/*n*-Ge (100) Schottky diodes is maintained up to annealing temperature of 550°C. Results also indicate that the onset temperature for agglomeration in Pd/*n*-Ge (100) system occurs between 500-600°C and in Ru/*n*-Ge (100) system occurs between 600-700°C.

1. Introduction

As device dimensions are scaled to sub-micrometer dimensions in silicon-based microelectronics, new processes and materials are becoming necessary to overcome the limitations of the conventional methods [1]. Of interest are silicon compatible materials that provide better device performance. Germanium (Ge) is a promising material for high mobility devices due to its higher and more symmetric carrier mobility compared with silicon [2], and its excellent compatibility with high-*k* materials [3]. The lack of a stable native Ge oxide has been the obstacle for the use of Ge in CMOS devices [4]. However, recent developments of next generation deposited high-k dielectrics, germanium oxynitride (GeON), Zirconia oxide (ZrO_2), Aluminum oxide (Al_2O_3) and Hafnium oxide (HfO_2) allow for the fabrication of high performance Ge-based metal-oxide semiconductor field effect transistors (MOSFETs) [5]. Metal germanides may be used as contact materials in future germanium technology. Compared with silicides that have been extensively investigated in the past [6], formation of germanides on single crystal germanium surface attracted less attention. Therefore optimal implementation of germanium technology will require an optimal understanding of metal-germanium interaction from both metallurgical and electronic standpoints. Most of the studies on metal-Ge reaction up to date have been carried out using in-situ annealing by slowly-ramping annealing

temperature or rapid thermal annealing processing, rather than using furnace annealing, and also with less emphasis on morphological evolution [7].

Metal-semiconductor (MS) interfaces are an essential part of virtually all semiconductor electronic devices [8]. The MS structures are important research tools in the characterization of new semiconductor materials [9]. Their interface properties have a dominant influence on the performance, reliability and stability of device [10]. These applications include microwave field effect transistors, radio-frequency detectors, quantum confinement devices and space solar cells [11, 12].

In this study we investigated the change in electrical properties, morphological evolution and microstructure stability of Pd/- and Ru/*n*-Ge (100) Schottky diodes at different furnace annealing temperatures in the temperature range 25-700°C.

2. Experimental procedures

To study the thermal annealing behavior of the Schottky diodes, we have used bulk-grown (100) oriented, *n*-type Ge, doped with antimony (Sb) to a density of 2.5×10^{15} cm^{-3} supplied by Umicore. Before metallization the samples were first degreased and then etched in a mixture of H_20_2: H_2O (1:5) for 1 minute. Immediately after cleaning, the samples were inserted into a vacuum chamber where AuSb (0.6% Sb), 100 nm thick, was deposited by resistive evaporation as an ohmic contact. A ten-minute anneal at 350°C in argon (Ar) to lower the barrier height and increase the ohmic behavior of the contact was performed. Before Schottky contact deposition, the samples were again chemically cleaned as described above. Pd Schottky diodes were resistively deposited under vacuum below 10^{-6} Torr, while Ru Schottky diodes were deposited by using an electron beam deposition system. The diodes were (0.60±0.05) mm in diameter and 30 nm thick. The thickness of the diodes and deposition rate were monitored by using an INFICON XTC 751-001-G1 quart crystal thickness monitor. AuSb, Pd and Ru films were deposited with an unintentional heating of the substrate. After the diodes fabrication, the samples were characterized by current-voltage (*I-V*) measurements. The electrical characterization was repeated after every annealing cycle in Ar ambient for 30 minutes between 25°C and 575°C. Characterization of the films for as-deposited and after different annealing temperatures was accomplished using a JEOLJSM-5800LV and ZEISS ULTRA PLUS scanning electron microscopy (SEM) systems operating at 5 kV and 1 kV, respectively.

3. Results

The Schottky barrier heights (SBHs) of the diodes were calculated from *I-V* characteristics, which were analyzed by the thermionic emission (TE) model given by [10,13]:

$$I(V) = I_0 \exp\left(\frac{qV}{nkT}\right)\left[1 - \exp\left(-\frac{qV}{kT}\right)\right] \tag{1}$$

where I_0 is the reverse saturation current given as [14]:

$$I_0 = AA^*T^2 \exp\left(-\frac{q\Phi_B}{kT}\right), \tag{2}$$

obtained from the intercept of ln *I* at $V = 0$, A^* is the effective Richardson constant, *A* is the effective diode area, *T* is the measurement temperature in Kelvin, *k* the Boltzmann constant, Φ_B is the zero bias effective Schottky barrier height (SBH), *q* is the electronic charge and *n* the ideality factor. The ideality factor is determined accurately from the slope of the linear part of an ln *I* versus *V* plot. Assuming pure thermionic emission, *n* can be obtained from equation (1) as

$$n = \frac{q}{kT}\frac{dV}{d(\ln(I))} \tag{3}$$

The value of *n* is equal to 1.0 for an ideal diode and usually has a value greater than unit.

Figure 1 shows the variation of the SBH and reverse current at -1 V with annealing temperature for Pd Schottky diodes on *n*-Ge (100).

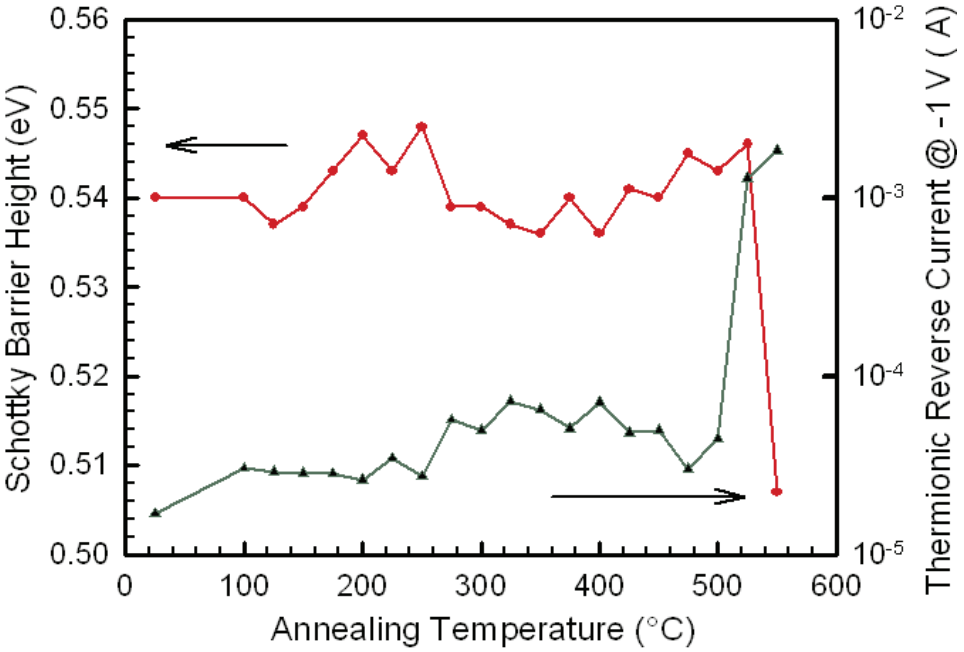


Figure 1. Plot of the Schottky barrier height and reverse current at -1 V as a function of annealing temperature for Pd/n-Ge (100) Schottky diodes.

The SBH and reverse current at a bias voltage of -1 V for as-deposited Pd/*n*-Ge (100) Schottky diodes were found to be (0.540 ± 0.005) eV and (17.00 ± 0.02) μA, respectively. The variation of Pd/n-Ge (100) Schottky diodes barrier height between 100-500°C is approximately constant within experimental error. According to Gaudet et al [15], only one germanide phase, PdGe exist for Pd on *n*-Ge (100) in this temperature range. This germanide is stable over a wide range of temperature.
The value of ideality factor for as-deposited Pd/n-Ge (100) diodes was determined to be 1.14. The ideality factors were between 1.20 and 1.50 at annealing temperatures between 25°C and 525°C.

The variation of SBH and reverse current at a bias of voltage -1 V with annealing temperature for Ru/*n*-Ge (100) Schottky diodes is shown in figure 2. The values of SBH and reverse current at -1 V for as deposited Ru/*n*-Ge (100) Schottky diodes were determined to be (0.557 ± 0.05) eV and (5.79 ± 0.02) μA, respectively. Figure 2 indicates nearly a constant SBH in the temperature range (50-150°C) At annealing temperatures higher than 150°C the SBH decreases with annealing temperature reaching a low SBH of (0.467 ± 0.005) eV after a 275°C anneal, depicting a significant reaction between Ru and the Ge substrate forming a germinide Ru+Ru_2Ge_3 [15]. We propose that, after subjecting the Ru Schottky diodes on *n*-Ge (100) to isochronal annealing, the first phase of Ru germanide is formed in the temperature range 150-275°C, as the SBH decreases significantly above 150°C annealing. The increase in SBH after annealing at temperatures in the range of 325-525°C depicts the formation of the Ru germinide Ru_2Ge_3 [15]. After a 550°C the diodes became near-ohmic and further evaluation was impossible. The as-deposited ideality factor for the Ru/*n*-Ge (100) Schottky diodes was found to be 1.08. This ideality factor is almost a constant within experimental error up to annealing temperature of 175°C. The ideality factor was greater than 1.1 for annealing temperatures higher than 175°C.

The interface states and chemical reactions between metals and semiconductors at the interface play an important role in the electrical properties of devices [16]. Since during the annealing process, metals may react with semiconductors and new compounds would form. Therefore, the change in SBHs may be attributed to the combined effects of interfacial reaction and phase transformation [16].

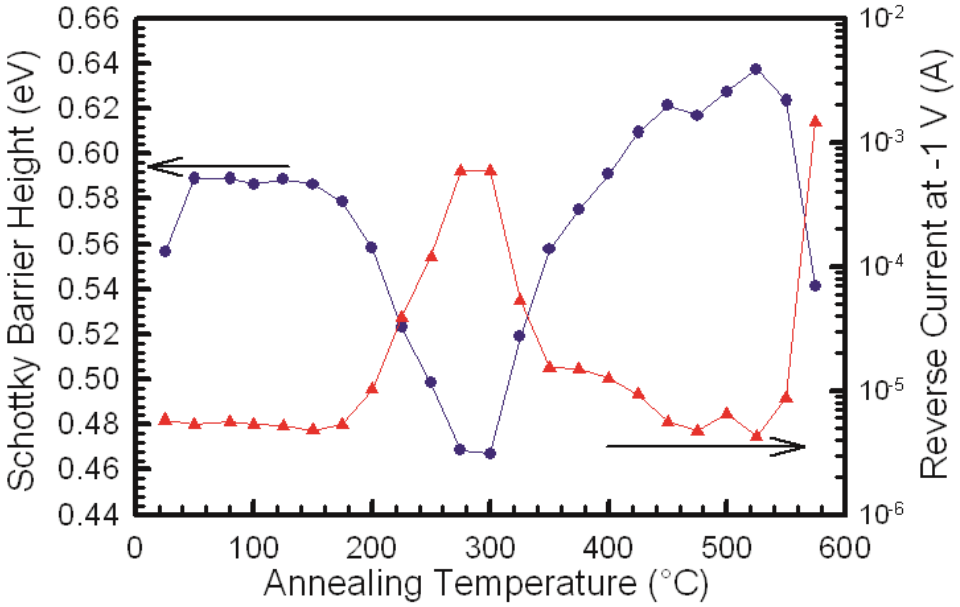
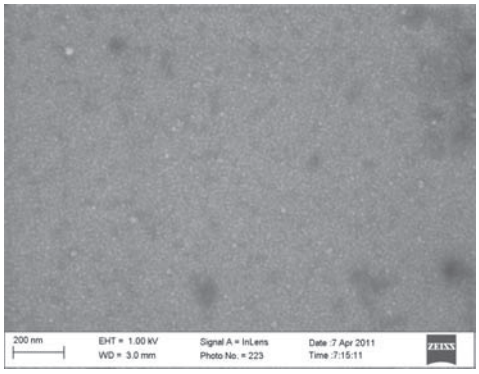


Figure 2. Plot of the Schottky barrier height and reverse current at -1 V as a function of annealing temperature for Ru Schottky contacts on *n*-Ge (100).

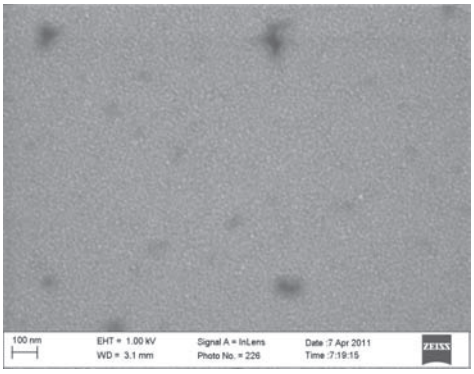
For ideal Schottky diode the ideality is 1.0. The deviation from ideality may be due mostly to the states associated with the defects near the surface. Interface states, inter-diffusion, compound formation, etc can all be derived from thermodynamics due to thermal annealing [17]. These may lead to recombination centres [18], and SBH inhomogeneities [19], which may cause a flow of excess current leading to a deviation from the ideal TE behavior.

Scanning electron microscopy (SEM) observations were conducted for Pd/- and Ru/*n*-Ge (100) samples, as-deposited and annealed at different temperatures. Results are shown in figures 3 and 4. Figure 3 shows the morphological evolution of Pd films on *n*-Ge (100). In figure 3 (a) and (b) the metal surface shows little change when annealed below 400°C. Grain growth at the surface was evident after 400°C anneal and severe grain grooving was observed after a 600°C anneal (Figure 3 (e)). From these results we suggest a good morphological stability for Pd films on *n*-Ge (100). From these observations we conclude that the onset temperature for agglomeration in 30 nm Pd/n-Ge (100) system occurs between 500-600°C.

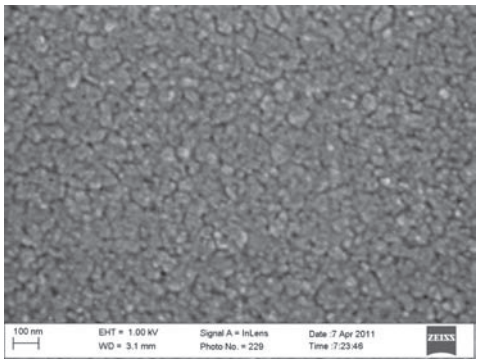
SEM observations of the Ru films on *n*-Ge (100) for as-deposited and morphological evolution of the samples after annealing at different temperatures are shown in figure 4. In figure 4 (a), (b) and (c) the metal surface shows little change when annealed below 500°C. Grain growth was evident after 600°C anneal (Figure 4. (d)). After annealing at 700°C (see figure 4 (e)), film continuity was interrupted as indicated by dark spots caused by exposed Ge regions. It can be concluded from these SEM micrographs that the onset temperature for agglomeration in 30 nm Ru/*n*-Ge (100) system is between 600-700°C.



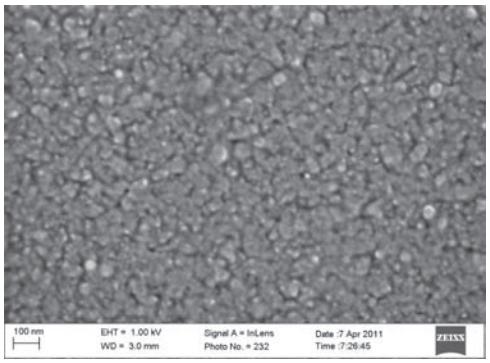
(a)



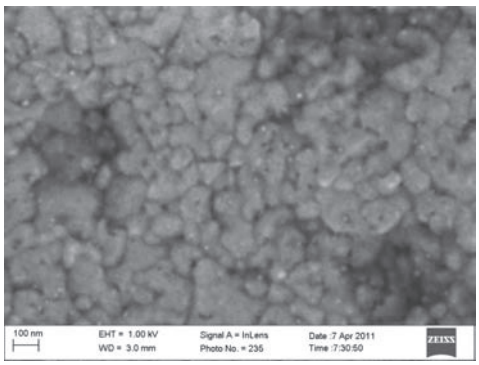
(b)



(c)

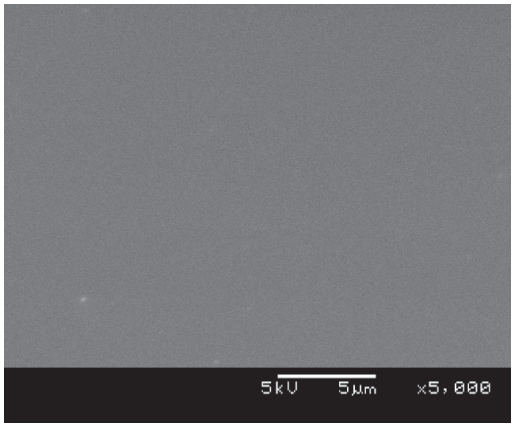


(d)

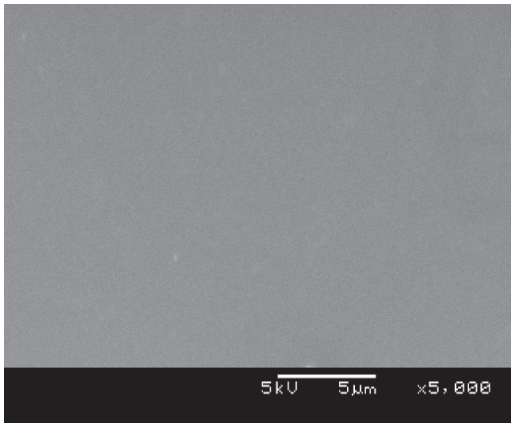


(e)

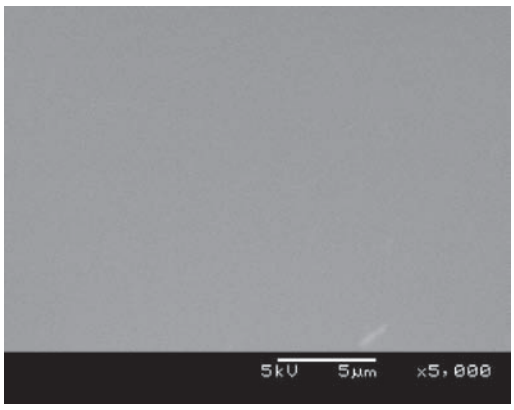
Figure 3. SEM observation for Pd films deposited on *n*-Ge (100) after isochronal thermal treat at different annealing temperatures: (a) as-deposited, (b) 200 (c) 400 (d) 500 and (e) 600°C



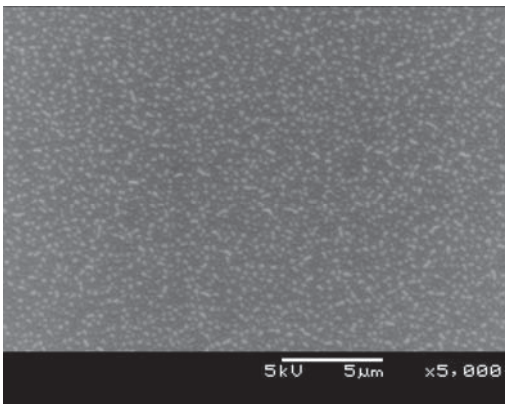
(a)



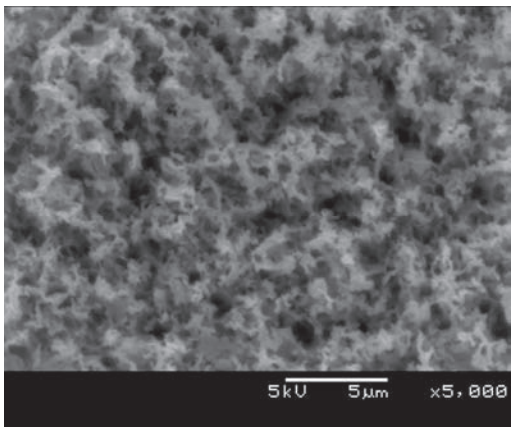
(b)



(c)



(d)



(e)

Figure 4. SEM observation for Ru films deposited on germanium after isochronal thermal treat for 30 min at different annealing temperatures: (a) as-deposited, (b) 400, (c) 500, (d) 600 and (e) 700°C.

4. Summary and Conclusions

Pd/*n*-Ge (100) Schottky diodes were fabricated by resistive deposition. Ru/*n*-Ge (100) diodes were fabricated by electron beam deposition. The Schottky behaviour was investigated under various annealing conditions. SEM observations were carried out for samples annealed at different temperatures. The variation of SBHs and ideality factors with annealing temperature may be due to interfacial reactions of metals (Pd, Ru) with germanium and the phase transformation of the metal-germanides during annealing process. The results show that Pd/- and Ru/*n*-Ge (100) Schottky diodes are thermally stable over a wide range of temperatures. From SEM observations, it can be concluded that the onset temperature for agglomeration in Pd/*n*-Ge(100) system occurs between 500-600°C, and in Ru/*n*-Ge (100) system occurs between 600-700°C.

Acknowledgements

This work has been made possible by financial assistance from South African National Research Foundation.

References

[1] Ashburn S P, Öztürk M, Harris G and Maher M 1993 *J. Appl. Phys.* **74** 4455
[2] An X, Fan C, Huang R and Zhang X 2009 *Chin. Phys. Lett.* **26** 087304
[3] Sun H P, Chen Y B, Pan X Q, Chi D Z, Nath R and Floo Y L 2005 *Appl. Phys. Lett.* **86** 071904
[4] Chi D Z, Yao H B, Liew S L, Tan C C, Chua C T, Chua K C, Li R and Lee S J 2007 *International Workshop on Junction Technology* 81
[5] Chui C O, Ramanathan, Triplett B B, McIntyre P C and Saraawat K C 2002 *IEEE Electron Dev. Lett.* **23** 473
[6] Zhang S L and Ostling M 2003 *Crit. Rev. Solid State Mater. Sci.* **28** 1
[7] K.Y. Lee, S.L. Liew, S.J. Chua, D.Z. Chi, H.P. Sun and X.Q. Pan 2004 *Mat. Res. Symp.Proc.* **810** C.2.4.1
[8] Tung R T 2001 *Mater. Sci. Eng. R* **35** 1
[9] Yüksel Ö F 2009 *Physica B* **404** 1993
[10] Sze S M 1981 *Physics of Semiconductor Devices* (New York: Wiley- Interscience)
[11] Wermer J H and Güttler H H 1991 *J. Appl. Phys.* **69** 1522
[12] Chand S and Kumar J 1996 *J. Appl. Phys.* **80** 288
[13] Yahia I S, Fadel M, Sark G B, Yakuphanoglu, Shenouda S S and Farooq W A 2011 *J. Alloys Compd.* **509** 4414
[14] Uslu H, Altindal S, Aydemir U, Dokme I and Afandiyeva 2010 *J. Alloys Compd.* **503** 96
[15] Gaudet S, Kellock, Desjardins P and Lavoie C 2006 J. Vac. Sci. Technol. A **24** 474
[16] Sun Y, Shen X M, Wang J, Zhao D G, Feng G, Fu Y, Zhang S M, Zhang Z H, Feng Z H, Bai Y X and Yang H 2002 *J. Phys. D: Appl. Phys.* **35** 2648
[17] Sands T 1988 *Appl. Phys. Lett.* **52** 197
[18] Dogan H, Yildirim N, Turut A 2008 *Microelectron. Eng.* **85** 197
[19] Tung R T, Sullivan J P, Schrey F 1992 *Mater. Sci. Eng.* **14** 266

Thermal properties of micron size Ag-added PDMS composites: A fuel cell perspective

Hao Chen[1,2], Ionel Botef[2], V. Vasudeva Rao[3] and V.V. Srinivasu[4*]

[1]Mechantronics & Micro Manufacturing, Materials Science and Manufacturing, CSIR, P.O. Box 395, Pretoria, South Africa

[2]School of Mechanical, Industrial and Aeronautical Engineering, University of Witwatersrand, P.O. WITS, Johannesburg 2050, South Africa

[3]Department of Mechanical Engineering, Sreenidhi Institute of Science and Technology, Yamnampet, Ghatkesar, Hyderabad 601 301, AP, India.

[4]Department of Physics, University of South Africa, P O Box 392, UNISA, Pretoria 0003, South Africa

[*]**E-mail: vallavs@unisa.ac.za**

Abstract. We consider polydimethylsiloxane (PDMS) as a prospective structural material candidate for a Fuel Cell fabrication. In this perspective, our work focuses on the tailorability and optimization of thermal properties of PDMS with micro/nano fillers such as silver particles. Thermal conductivity and thermal stability of micron size silver particle added PDMS composites (μAg-PDMS) were studied. The Ag filler fraction was varied in the range of 50 to 77 wt%. The thermal conductivity was measured using 'Cut-bar' technique in the temperature range 50 to 150 °C. We found that in general silver addition in PDMS improved the thermal conductivity and in particular the μAg-PDMS composite with 77 wt% silver filler showed a remarkable enhancement of almost 10 times as compared to the pure PDMS. This is very good for heat management and quick uniform distribution of heat generated in a fuel cell. Further we studied the thermal stability using TGA. We found that our μAg-PDMS composites are highly stable with only less than 1% weight loss. This feature is very good for a fuel cell fabrication using these μAg-PDMS composites.

1. Introduction

Electrical energy for powering small devices is widely expanding, and more focus is put on developing advanced micro fuel cells. Developing fuel cells for potable communication system faces a new set of difficult engineering challenges, which include miniaturizing reactant delivery system and fabricating conductive polymers to be used in bipolar plates for fuel cell stacks. Traditionally, metals were used to fabricate bipolar plates. Many problems emerged when applying metals, as a result electrical conductive polymers were more preferred. Electrically conducting polymers are mostly used in fuel cells as flow field plates, i.e. bipolar plates and flow field plates; the characteristics of the polymer makes them more preferred than metals. Thus miniaturization of the microscopic flow fields

in fuel cells can help to achieve higher energy density which is crucial for the long lasting of micro fuel cells.

Most research on conducting polymers is focused on the electrical conductivity [1-4]. However, one has to consider the thermal properties of the polymer composite as well, especially in a fuel cell application as heat is a byproduct of the fuel cell operation. When conductive filler like carbon black, carbon nanotubes or metal particles are incorporated in the polymer matrix, not only the electrical but also the thermal conductivity is improved. The polymer used in the fabrication process is called polydimethylsiloxane (PDMS), which has advantageous properties such as high flexibility, chemical resistance, and biocompatible making it popular in micro manufacturing [5]. Improved thermal conductivity will allow removal of excess heat and therefore is very important for heat management as well as uniform distribution of heat generated in a fuel cell. Unmodified PDMS has a poor thermal conductivity (0.17 $Wm^{-1}K^{-1}$). In this perspective, our work focuses on the tailorability and optimization of thermal properties of PDMS with micro fillers. Previous work has been reported on the thermal properties of PDMS-Carbon Black composite [6,7]. In this paper, thermal stability and conductivity of PDMS- Silver composite has been studied. PDMS composite samples with 50wt% to 77wt% silver loadings were prepared.

2. Materials and experimental methods

2.1. Materials

PDMS (Sylgard 184 Silicone Elastomer) is supplied by Dow Corning as two-part liquid component kits comprised of a base and a curing agent. The two parts mix in the ratio of 10: 1, when the two liquids are mixed thoroughly, the mixture cures at 60°C to form a flexible elastomer. Silver powder (327085, silver powder, 2-3.5 μm, ≥99.9% trace metals basis) was supplied by Sigma Aldrich, it has a particle diameter of 2-3.5 μm

2.2. Test specimen fabrication

The silver particles and PDMS polymer blend were first ultrasonicated for 30min, then an in-house designed mechanical mixer was used to mix the blends for 24 h to obtain homogeneous dispersion, the crosslinker was added last. A vacuum degassicator was used to remove the bubbles caused by the mixing process. The blend was casted in the mould and oven-baked for 8 h at 60°C to cure properly. Samples with silver loading of 50wt% to 77wt% were prepared. A disk-shaped mould with diameter of 25 mm and thickness of 1 mm was designed for the specific requirement of the thermal conductivity measurement.

2.3. Characterisation

A widely used Cut-Bar technique (ASTM E 1225-87 methodology) was used for thermal conductivity measurement for samples with varied silver loading, as well as thermal conductivity dependence on temperature with the temperature range of 50–150°C. Thermal Gravimetric Analysis (TGA) for thermal stability/degradation was performed using a TA Instrument TGA Q500 under air atmosphere. The heating rate was 10°C/min and the flow rate for oxygen and nitrogen are 5 and 30 ml/min, respectively. All experiments were conducted in the temperature range 30–800°C.

3. Results and discussion

3.1. Thermal conductivity

Thermal conductivity (k) of the test specimens is measured using a cut-bar test facility under steady-state conditions. The system is assumed to have attained steady-state conditions when all the temperature sensors (thermocouples) indicate a constant value with a variation of less than 0.5°C. Variation of thermal conductivity as a function of mean bulk temperature is determined. Complete details of experimental procedure are described elsewhere [8] . Figure 1 shows the variation of thermal conductivity k as a function temperature for PDMS composites filled with silver with different filler

fraction ranging from 50 wt% to 77 wt%. The measurements of the thermal conductivity are made in the temperature range of 50–150°C.

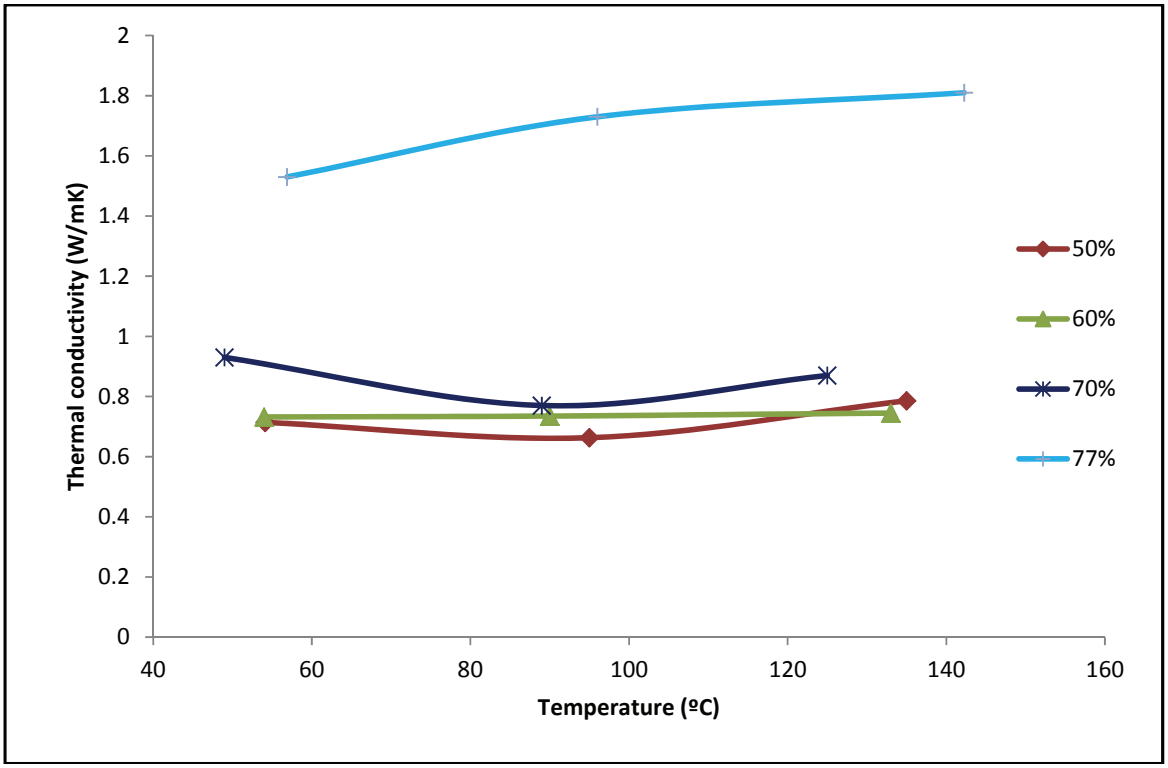


Figure 1. Thermal conductivity VS temperrature for silver filled PDMS composites.

Results have shown a general trend for all samples that thermal conductivities remain relatively stable with minor variations upon increasing temperature. PDMS is a polymer with low thermal conductivity of 0.17 $Wm^{-1}K^{-1}$[9]. However, for PDMS composite with silver fillers incorporated (77wt% sample), a remarkable enhancement of almost 10 times was noted. At lower silver loadings (50, 60 and 70wt %) samples showed similar thermal conductivity, no significant difference was found. When silver loading of 77wt% is achieved, thermal conductivity is twice of 50wt% sample (0.71 $Wm^{-1}K^{-1}$) and one and half times of 70wt% sample (0.93 $Wm^{-1}K^{-1}$). This could be attributed to the increase in the conductive pathway and network density (obtaining the percolation threshold) at 77wt% of filled silver particles [10]. Similarly, Cong et al reported improvement of thermal conductivity with silver addition in PDMS [2]. However, we cannot compare their results with ours as their measurements were taken as a function of volume concentration.

3.2. Thermal stability (TGA)

To maintain thermal stability is important for materials to be used in a fuel cell application at various operating temperatures. Therefore, the thermal degradation of silver filled PDMS composite is investigated by TGA. Figure 2 illustrates the results of TGA analysis.

From figure 2, results have shown that all four samples maintained a high weight percentage until 220°C with only less than 1% weight loss. Above this temperature, samples start to degrade at a higher rate. Samples with higher silver filler loading showed better thermal stability, the weight remaining of samples matches the increasing order of filler concentration. Apparent weight loss occurred between 430 and 500°C and then all samples appeared to be stable up to 800°C. 77wt% silver loading sample

had 81% weight remaining while 50wt% silver loading had 57% weight remaining at 800°C, this also indicates that silver is very stable at this temperature range and weight loss for these composite are mainly due to the PDMS polymer matrix. This kind of thermal stability of the material is very good for fuel cell fabrication as normally the fuel cell has optimum performance at operational temperatures around 100 - 200 °C depending on the types of fuel cell.

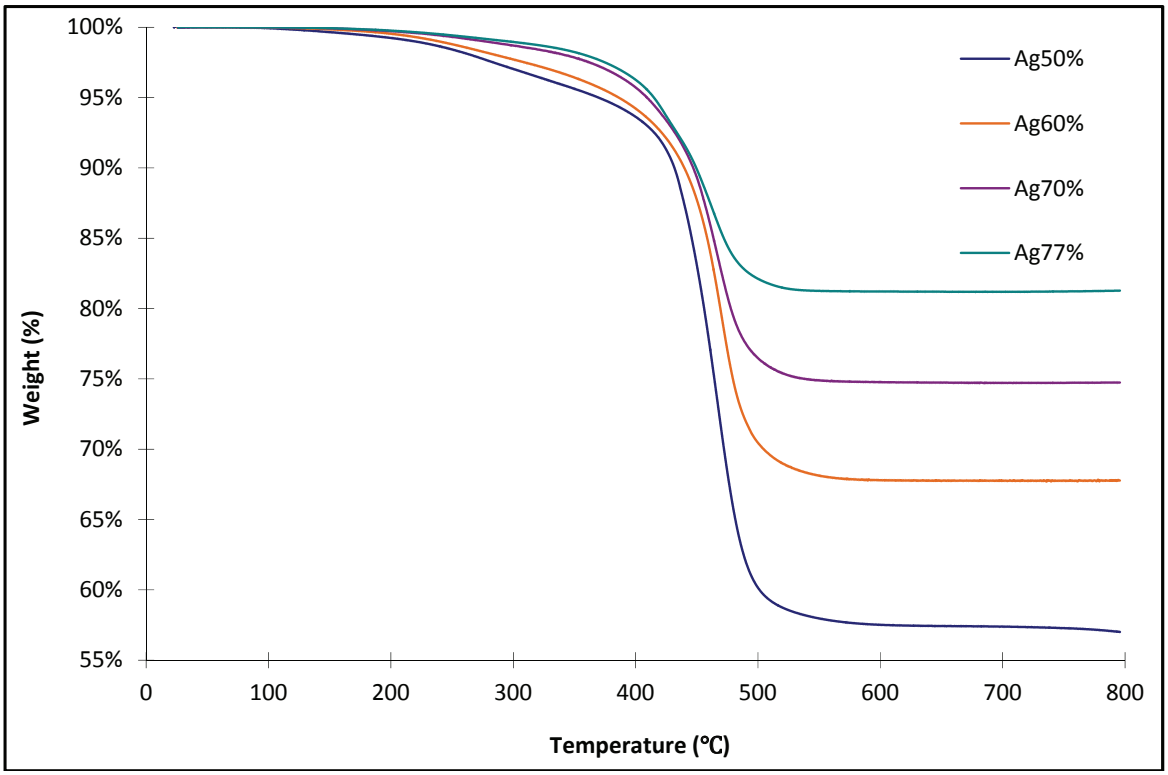


Figure 2. Thermal conductivity vs temperature for silver filled PDMS composites.

4. Conclusion

Thermal conductivity measurements were carried out on silver filled PDMS composites as functions of temperature as well as filler fractions. When compared with pure PDMS polymer, thermal conductivity of sample with 77wt% silver loading was noticed to have an order of magnitude improvement as compared to that of pure PDMS. Furthermore, the composite blends have also shown synergistic improvement in thermal stability. The presence of silver filler in the PDMS polymer has resulted in retardation of the thermal decomposition of the composite. A study of these composites with utilisation in the fuel cell bipolar plate is still at the research and development stage. Further work will include the study of the electrical and mechanical properties of the composites.

References

[1] Niu X, Peng S, Liu L, Wen W and Sheng P 2007 *Adv. Mater.* **19** 2682-6
[2] Cong H and Pan T 2008 *Adv. Funct. Mater.* **18** 1912-21
[3] Gojny F H, Wichmann M H G, Fiedler B, Kinloch I A, Bauhofer W, Windle A H and Schulte K 2006 *Polymer.* **47** 2036-45
[4] Huang J, Baird D G and McGrath J E 2005 *J. Power. Sources.* **150** 110-9

[5] Land K, Mbanjwa M, Hugo S, Chen J and Govindasamy K 2010
[6] Chen H, Botef I, Zheng H, Maaza M, Rao V and Srinivasu V V 2011 *Int. J. of Nanotechnology*. **8** 437-45
[7] Chen H, Botef I, Guduri B and Srinivasu VV 2010 *AIP Conference Proceedings*. **1276** 243-8
[8] Rao V V, Bapurao K, Nagaraju J and Murthy M V 2004 *Meas. Sci. Technol.* **15** 275-8
[9] Yamamoto T, Fujii T and Nojima T 2002 *Lab Chip.* **2** 197-202
[10] Burden A P, Guo W and Hutchison J L 1998 *Polymer.* **39** 4211-7

Non-linear Fowler-Nordheim plots in thin film polymer-fullerene composite devices (transition from hole-only to electron-only conduction)

Zivayi Chiguvare

University of the Witwatersrand, DST/Centre of Excellence in Strong Materials and Material Physics Research Institute, School of Physics, Private Bag 3, Wits 2050, Johannesburg, South Africa

Zivayi.Chiguvare@wits.ac.za

Abstract. We studied charge injection and transport mechanisms in blends of poly(3-hexylthiophene) (P3HT) and [6,6]-phenyl C61-butyric acid methylester (PCBM), by analyzing dark, temperature dependent current-voltage characteristics of the P3HT:PCBM blend thin films sandwiched between aluminium electrodes in a MIM configuration. We present a general method of interpreting Fowler - Nordheim plots of metal/semiconductor/metal devices with pronounced non-linear characteristics by dividing them into several regions based on physical origins. We show that by applying appropriate electric fields it is possible to switch from electron-only conduction to hole-only conduction in a single Al/P3HT:PCBM/Al device. We affirm that electrons can be selectively transported through the lowest unoccupied molecular orbital of PCBM at low applied voltages and low temperatures; and alternatively holes can be transported through the highest occupied molecular orbital of P3HT at higher applied voltages and high temperature, within a single device.

1. Introduction

Thin films of polymer- fullerene blends are interesting because electronic devices such as light emitting diodes, field effect transistors, and solar cells made using them can be flexible, light-weight, and far cheaper as compared to those based on inorganic materials such as silicon, or germanium. For instance the efficiency of solar cells based on P3HT:PCBM blends has risen from less than 1% in the early 1990's to above 5 % in 2009, prompting companies like Konarka[§] to start commercial manufacture of organic photovoltaic cells for consumer products with up to 1GW capacity per annum envisaged. The Physics underlying charge carrier injection and transport in such devices is however still not yet fully understood.

Charge injection in MIM (e.g. polymer) devices takes place by thermionic emission from metal electrode to some transport level in the polymer, and by quantum mechanical tunneling of charge carriers at high enough applied electrical fields. Charge transport is however determined by the characteristics of the bulk material itself, such as mobility and density of charge carriers, and the existence of impurities. The essential assumption of the Richardson-Schottky (RS) model of thermionic emission is that an electron from the metal can be injected once it has acquired sufficient thermal energy to cross the potential maximum that results from the superposition of the external and the image charge potential. For low bias the number of injected carriers is smaller or comparable to the

thermally generated intrinsic charge carriers, and the current – voltage characteristics can be described by Ohm's law:

$$J_{Ohm} = q(n_p\mu_p + n_e\mu_e)\frac{V}{d} \tag{1}$$

In the absence of traps, or when all traps are filled, J has a quadratic dependence on V, and follows the Mott-Gurney law:

$$J_{TFSCLC} = \frac{8}{9}\mu_p N_v \varepsilon\varepsilon_0 \frac{V^2}{d^3} \tag{2}$$

Fowler–Nordheim tunneling is the wave- mechanical tunneling of an electron through an exact or rounded triangular barrier. The F-N model for tunneling injection ignores image charge effects and invokes tunneling of electrons from the metal through a triangular barrier into unbound continuum states. It predicts:

$$J_{FN}(F) = \frac{q^3}{16\pi^2\hbar\phi}F^2 \exp\left[-\frac{4(2m_{eff})^{1/2}\cdot\phi^{3/2}}{3\hbar eF}\right] \tag{3}$$

independent of temperature. Here m_{eff} is the effective mass of the carrier inside the dielectric.

Observation of straight line fits of *J(V)* data in F-N plots is usually taken as confirmation that quantum mechanical tunneling of charge carriers dominates the conduction mechanisms of a given single carrier device. The physical origin of nonlinearity in the Fowler–Nordheim plot is still a debatable issue. Non-linear F-N curves were observed by several authors [1, 2], where the F-N plot is curved downwards for low applied fields, with positive slope, and upwards at high fields with negative slope, and approaching the sought after straight line approximation [3]. Various reasons have been suggested for the occurrence of the first inflexion point: F-N plot curvature can be relatively small [4], so where marked curvature occurs in experimental plots, usually some other effect must be operating, such as the presence of vacuum space charge, electron supply limitation inside the emitter, or statistical effects associated with a many emission- sites electron source [5]. We have shown that the first inflexion point leading to the appearance of the first minimum is due to the transition from conduction by thermionically injected charge carriers to conduction by field emission injected charge carriers, the sum of which gives a single minimum in a F-N plot [6]. Both F-N electron and hole tunneling [7] have been observed in experiments.

We note that a second inflexion point at higher fields is evident in the experimental results of several studies [8, 9], but the authors have not referred to it, limiting their discussion to the straight line approximation. A flattening of the F-N plots, linked to the second inflexion point, was suggested to be due to several factors such as a finite reflection coefficient, a finite energy width of emitted electrons, and space charge effects [10]. We have observed even a third inflexion point as applied electric field is increased and will elucidate the physical origins of the observed non-linearities in this paper. Such behaviour was referred to as oscillatory behaviour of the F-N plots, by Khlifi *et al.* [11] who analysed more than 3 inflexion points and attributed the oscillations to some 'excess' current.

2. Materials and Methods

Glass substrates were cleaned in deionized water, acetone, toluene and isopropanol, respectively, in a hot ultrasonic bath. The Al electrode in contact with the glass substrate was deposited by thermal evaporation in better than $5x10^{-7}$ mbar vacuum. A chloroform-toluene based 1:2 polymer – fullerene blend by mass solution (10 mg/ml) was then spin-coated in the nitrogen atmosphere of a glove box, O_2 – 2 ppm and H_2O – 0.01 ppm giving films of thicknesses around 80 nm. The aluminium top electrodes were deposited by thermal evaporation in high vacuum, better than $5x10^{-7}$ mbar at ~ 0.2 nm/s rate.

Note that the Al electrodes are each in contact with both P3HT and PCBM. All devices were stored in nitrogen atmosphere prior to measurement. The formulae of P3HT and PCBM, and energy level diagrams of the studied Al/P3HT:PCBM/Al devices (under non – equilibrium conditions) are shown in Fig. 1. All devices were heated slowly up to380 K before measurement was initiated.

3. Results

Typical current-voltage characteristics of Al/P3HT:PCBM/Al thin film devices are shown in Fig. 2. The Al/P3HT:PCBM/Al devices are electron only devices at low applied electric fields. At temperatures above absolute zero, electron-hole pairs are generated by thermal excitation resulting in electrons in the LUMO of P3HT, and holes in its HOMO. The electrons are immediately transferred to LUMO PCBM through an ultrafast charge transfer (~10^{-15} s). The resulting quasi stable state is that of positive P3HT radicals in the vicinity of negative PCBM radicals, where electron-hole recombination is inhibited by potential barriers at heterojunctions distributed within the bulk. Such a configuration favours the transport of electrons through the LUMO of the PCBM, while holes are transported through the HOMO of the P3HT.

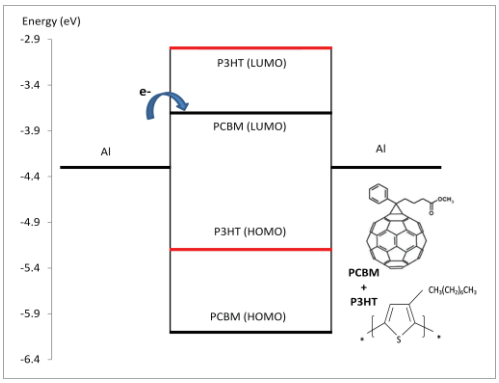


Fig. 1 Energy band diagram of an Al/P3HT:PCBM/Al device. The inset shows the chemical formulae of the materials making up the composite layer.

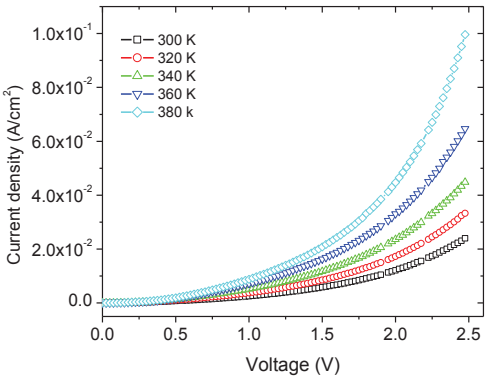


Fig. 2 Typical current voltage characteristics of an Al/P3HT:PCBM/Al device in linear scale.

Double log plots of *J(V)* characteristics show that ohmic and trap free SCLC (TFSCLC)analysis alone cannot describe completely the behaviour of these devices. The curves of Fig. 3 do not exhibit the slope ~ 1 region even at the lowest applied voltages. Such behaviour indicates that even at low voltages, there is some significant injection of charge carriers from the electrodes; in this case, electrons are injected from the Al electrode into the LUMO of PCBM. The slopes are constant between 1 and 2, suggesting a mixture of ohmic and space limited currents at these low voltages. Slopes increase beyond 2, but the lines have an oscillatory nature and are not straight as would be expected for SCLC conduction in the trap filling regime. Clearly TFSCLC cannot adequately describe the conduction mechanism in these Al/P3HT:PCBM/Al devices.

We investigate the charge injection mechanisms and their contribution to the total measured current, and observe non-linear F-N curves as indicated in Fig. 4. Each F-N curve has variable positive slope at low applied electric fields, which changes sense and becomes negative close to the highest applied fields, for lower temperatures. At intermediate applied voltages, the F-N curves have negative slopes, which again change to positive and finally

straight line regions of negative slopes are observed as voltage is increased further. Thus three inflexion points resulting in two minima are observed in the F-N curves.

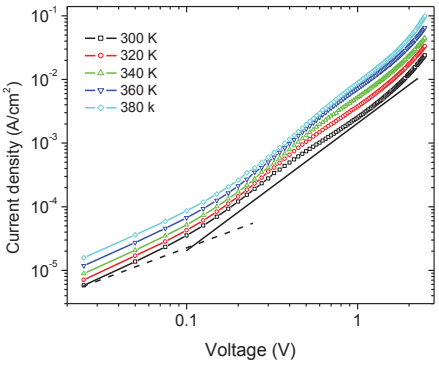


Fig. 3 Double log plot of *JV* curves for Al/P3HT: PCBM/Al devices. Dashed lines indicate slope = 1, and dotted lines indicate slope = 2

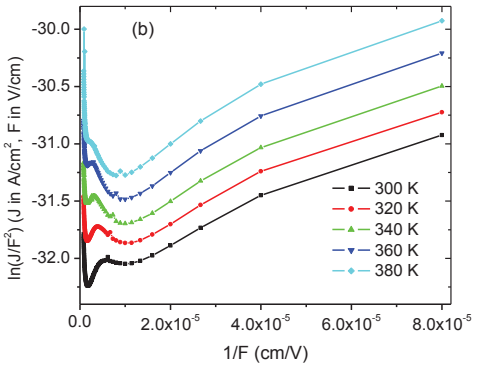


Fig. 4 Fowler Nordheim tunneling curves for Al/P3HT:PCBM/Al devices

Two negative slope regions with straight line tendencies are clearly observed; suggesting two tunnelling regimes in these devices (see Fig. 5). We argue that the first minimum as we increase voltage is due to significant tunnelling of electrons from Al electrode into the LUMO of the PCBM. In a single carrier device, straight lines should then be observed and characteristics of the tunnelling barrier studied. However, a second point of inflexion again changes the slope to positive, suggesting a reduction in the rate of increase of charge carriers available for current conduction. This can only happen if some of the injected charge carriers are not reaching the opposite electrode. We attribute this to recombination of charge carriers within the bulk of the film

The opposite electrode also injects charge carriers (holes) into the HOMO of P3HT, thereby shifting the recombination plane from the positive electrode towards the negative electrode. This is illustrated in Fig. 6 where a typical energy diagram of a MIM device with symmetric electrodes, under non-equilibrium conditions is shown; hole injection barrier φ_2 is greater than electron injection barrier φ_1; (b) Small negative bias on *metal 1* modifies the energy band structure, and both electrons and holes can be injected by thermionic emission over the trapezoidal barriers. The effective thickness traversed by charge carriers is the same as the film thickness and tunnelling is negligible; (c) At intermediate applied voltages, the effective thickness for electron conduction is smaller than the device thickness, and the electrode of smaller injection barrier φ_1, injects holes significantly by F-N tunnelling (electron-only device); (d) At high V_a both barriers are now smaller than qV_a, and tunneling injection takes place from both electrodes, leading to ambipolar conduction. As the voltage is further increased the injection of holes becomes so significant that the recombination plane is shifted until it reaches the negative electrode, in which case current conduction through the film is now due to holes through HOMO of P3HTonly. The straight line region can then be used to study the hole injection barrier at the Al/P3HT interface.

4. Model

We note that the first point of inflexion originally suggested to being the transition from direct to F-N tunnelling [12] is rather due to changeover, from thermionic emission dominated

injection, to FN tunneling dominated injection, of electrons across the Al/PCBM LUMO barrier.

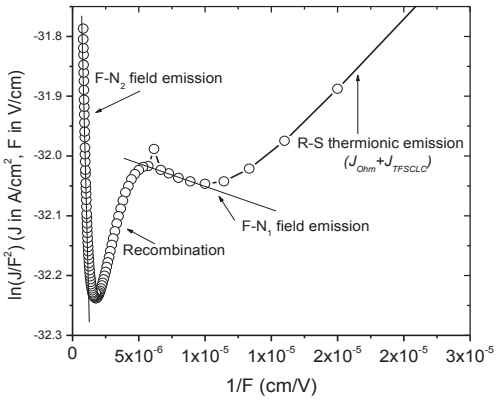


Fig. 5: 300 K F-N plot for Al/P3HT:PCBM/Al device, showing the four different regimes: R-S thermionic emission, electron tunneling through the lower barrier electrode, recombination regime, and hole tunneling through higher barrier electrode.

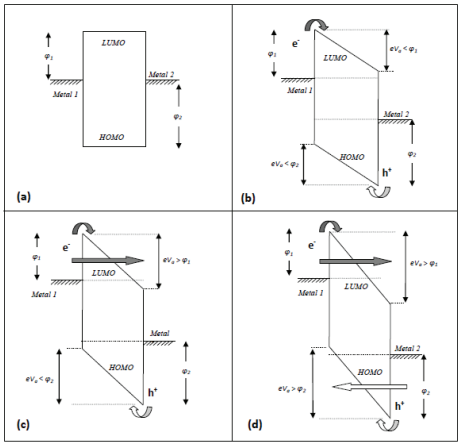


Fig. 6: Typical energy diagram of a MIM device with symmetric electrodes, **(a)** under non-equilibrium conditions; (b) under low bias; (c) under intermediate bias; and (d) under high bias.

We propose that the total current in the studied devices can be modelled as:

$$J_{tot} = J_{TE} + J_{FE} - J_{rec}\,, \tag{4}$$

Where J_{TE} is the current due to thermionically emitted charges, J_{FE} is the current due to tunneling injection, and J_{rec} is the current due to charge carrier recombination.

$$J_{tot} = J_{Ohm} + J_{SCLC} + J_{FN1} + J_{FN2} - J_{rec} \tag{5}$$

We propose a recombination current of the form:

$$J_{rec} = \alpha \frac{J_{FN1} \cdot J_{FN2}}{(J_{FN1} + J_{FN2})}\,, \tag{6}$$

where α is a temperature dependent constant associated with the position of the recombination plane. The value of α is minimal at the electrodes, $\alpha = 0$, but can take values greater than1 becoming maximum at a plane where $\mu_p n_p = \mu_e n_e$. The recombination current cannot be measured by an ammeter because the injected charges never reach the electrodes. Instead they annihilate each other resulting in the emission of a photon, or of a phonon, which may be absorbed by the lattice, increasing the temperature of the device. We have employed Eqs. (5) and (6) in an in-house developed MATLAB programme to fit our data and obtained reasonable fits as indicated in Fig. 7.

In Fig. 8 we illustrate the effect of varying the recombination constant α while keeping all other factors constant. For high α we obtain a recombination regime for a larger span of applied fields, while this becomes smaller when α is reduced. At $\alpha = 0$, the recombination regime is not observed at all, and an apparent smooth transition from electron only to hole only conduction should be observed. In this case it can be difficult to determine the two slopes due to F-N tunneling.

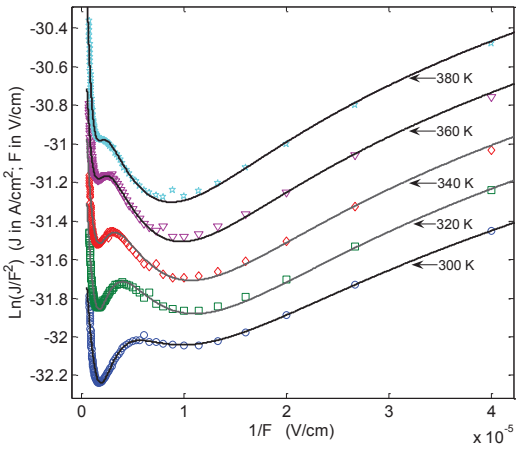


Fig. 7: Model fits to data.

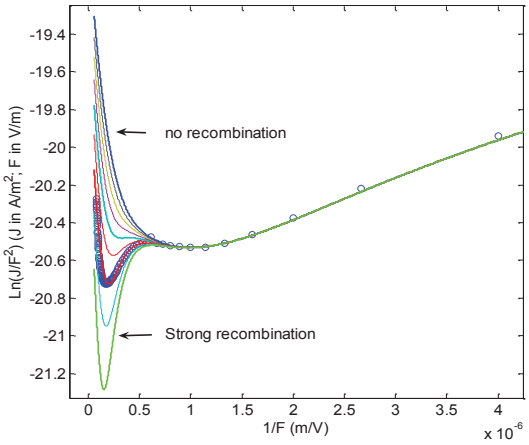


Fig.8: The slope of the FN curves for tunneling injection through lower barrier electrode is influenced by the recombination strength.

This slope cannot be used to estimate the electron tunneling barrier height, but on the contrary the tunneling regime at the higher barrier electrode is constant for all temperatures. If sufficiently high fields were applied, then the straight line observed would be attributed to hole tunneling in this case. It is therefore reasonable to expect that both n and μ increase with temperature for our Al/P3HT:PCBM/Al thin film devices (see Fig. 7).

5. Conclusions

We conclude that the nonlinearity in the FN plots originates from the transition from TE at low field to FE at moderately high electric fields(first minimum); from the transition from single to ambipolar conductivity (also associated with significant recombination) at sufficiently high fields (first maximum); and finally from the transition from ambipolar to single carrier conductivity of the opposite charge carrier of opposite sign (second minimum) after which the recombination plane shifts completely to the opposite electrode.

References

[1] C. Lee, J.Y. Park, Y.W. Park, Y.H. Ahn D. S. Kim, D.H. Hwang, T. H. Zyung, J. Kor. Phys. Soc. **35**, S291 (1999).
[2] A.J. Heeger, I.D. Parker, Y. Yang, Synth. Met. **67**, 23-29 (1994).
[3] S. H. Choi, B. Kim, C. D. Frisbie, *Science* **320**, 1482 (2008).
[4] C. J. Edgcombe, and N. de Jonge, J. Vac. Sci. Technol. B **24**, 869 (2006).
[5] R. G. Forbes, and J. H. B. Deane, Proc. R. Soc. A, **463**, 2907 (2007).
[6] Z. Chiguvare, J. Parisi, V. Dyakonov, J. Appl. Phys. **94**, 2440 (2003).
[7] R. K. Chanana, K. McDonald, M. Di Ventra, S. T. Pantelides, and L. C. Feldman, G. Y. Chung, C. C. Tin, and J. R. Williams, R. A. Weller, Appl. Phys. Lett., **77**, 16 (2000).
[8] C. Lee, J.Y. Park, Y.W. Park, Y.H. Ahn D. S. Kim, D.H. Hwang, T. H. Zyung, J. Kor. Phys. Soc. **35**, S291 (1999).
[9] G. Lai, Z. Li, L.Cheng, and J. Peng, J. Mater. Sci. Technol., **22**, 677 (2006).
[10] Y. Gohda, Y. Nakamura, K. Watanabe, and S. Watanabe, Phys. Rev. Lett. **85**, 1750 (2000).
[11] Y. Khlifi, K. Kassmi, L. Roubi, and R. Maimouni, Phys. Stat. Sol. (A) **182**, 737 (2000).
[12] J. M. Bebee, B. S. Kim, J. W. Gadzuk, C. D. Frisbie, and J. G. Kushmerick, Phys. Rev. Lett. **97**, 026801 (2006).

[§] http://www.konarka.com

Photoluminescence surface mapping as a probe for interface disorder in quantum well structures

K G Chinyama, P N Njingana and E L Meyer

Fort Hare Institute of Technology, University of Fort Hare, Private Bag X1314, Alice, 5700, South Africa.

KChinyama@ufh.ac.za

Abstract. We show here that a simple photoluminescence surface mapping across the sample indicates whether the nano-quantum well layer is interface disordered or alloy-disordered. Two sets of CdSe-ZnSe nano-quantum well structures, one grown by metalorganic vapour phase epitaxy and atomic layer epitaxy are analysed both structurally and optically. The former structures are found to be predominantly interface disordered whereas the latter samples are found to be alloy-disordered. Photoluminiscence mapping across the sample surface reveals that there is a strong correlation between the spatial variations in the exciton band peak energy position and the linewidth (~ 60 – 100 meV and 40 – 70 meV, respectively) in the well-width disordered systems compared to the alloy-disordered or quantum dot systems. A contour plot of the photoluminiscence shows for the rough structures orderly patterned contour lines for both the peak energy and linewidth. On the contrary, contour plots for the quantum dot samples show uncorrelated maps with the linewidths scattered in relation to those of the peak energy. We show that even without structural characterisation, one can tell the quality of the sample under study (whether rough or quantum dot) by carrying out a surface mapping of the emission and studying the contour maps and scatter plots of peak energies and their corresponding linewidths. These results can assist in shedding more light on similarly lattice-mismatched material combinations in the III-Vs and III-Nitrides nanostructures. This knowledge is crucial for understanding not only of what is behind the optical properties, but also of the growth dynamics of these systems and can be utilised in calibrations of the growth process to optimise sample quality.

1. Introduction

Problems of structural disorder associated with growth in ultra-thin structures include interface roughness, interfacial alloy formation, and well width fluctuations. Nanostructure heterointerfaces between quantum wells and barriers are quite rough within certain lateral regions that change in size at small terraces or islands with heights generally being in monolayers so that excitonic photoluminescence (PL) linewidths are closely related to the heavy-hole excitons bound in the potential well/barrier caused by the fluctuations formed by submonolayer deposition [1-13].

Here, we draw attention to photon emission characteristics of CdSe/ZnSe ultra-thin nanostructures surveyed on single quantum well structures using surface mapping measurements of the PL spectra. While a great deal has already been learned about the luminescence from these II-VI heterostructures, insufficient attention has been given of the correlation between the luminescence and the structural details of the CdSe insertion layer responsible for the optical effects reported. This knowledge is

crucial for understanding not only of what is behind the optical properties, but also of the growth dynamics of this lattice-mismatched system. This presentation tries to demonstrate that before any further investigation, surface mapping of the sample is important in determining the overall behaviour of the sample in terms of uniformity and nano-quantum well width fluctuations. In some areas where the structural variations are large over a sample, PL mapping can yield information which would otherwise require several samples to be grown.

2. Samples

The structures studied in this work have been grown by two methods: atomic layer epitaxy (ALE) for samples KC1 and KC2 and metal-organic vapour phase epitaxy (MOVPE) for samples KC3 and KC4. All QWs in this case are ZnSe/CdSe/ZnSe systems. All the samples are single QWs (SQWs) consisting of the GaAs substrate, followed by a ZnSe buffer layer, the CdSe QW, and finally a ZnSe capping layer. Details of the sample structures are given in Table 1 below.

Table 1: Summary of the sample structure

Sample Name	ZnSe buffer thickness (nm)	CdSe QW thickness (ML*)	ZnSe cladding thickness (nm)
KC1	150	0.5	50
KC2	150	1.0	50
KC3	150	1.8	50
KC4	150	1.0	50

*ML = monolayer

3. PL measurements and surface mapping

The emission was excited using a 357 nm argon-ion continuous-wave (cw) laser. The laser beam first passed through a mechanical chopper, which modulates the beam at 330Hz beam then through a 395 nm interference filter to remove the laser plasma lines, after which it was focused with a converging lens onto the sample held at a fixed temperature in a continuous flow cryostat. The sample was placed at an angle such that the direct reflection of the laser is not collected by the collimating lens between the sample and the spectrometer. The luminescence from the sample was then focused with the help of two lenses onto the slits of a grating spectrometer, of which the slitwidth could be varied (giving an appropriate spectral resolution) depending on the intensity of the luminescence from the sample in the wavelength range of interest. Generally the slit width was between 200 – 500 μm giving resolutions of ~1 meV. The output from the spectrometer was detected using a photomultiplier tube which was linked via an amplifier to the photon counter also fed with the 330 Hz reference signal. The data was then collected by a computer which, in addition to data acquisition, also regulated the temperature controller and drove the spectrometer motor.

The sample was divided up into a grid of equal small point areas. Then the PL spectra of each of the point areas across the entire surface were measured. From these spectra, the peak photon energy and full width at half maximum (FWHM) of the exciton band for each point was extracted. An analysis of the results yielded a relation between the peak energy and FWHM. Also 2-D contour maps of these parameters across the *xy*-plane of the surface were plotted. PL peak intensity mapping was not considered because of fluctuations in the intensity of the measuring incident laser beam.

4. PL Results

Figure 1 shows the PL spectra of all the samples studied in this work. All spectra are obtained at ~ 12K and the emissions are quite bright compared with standard ZnSe. Figure 1A shows spectra for KC3 and KC4 samples grown by MOVPE and Figure 1B shows that for ALE-grown KC1 and KC2. The exciton bands of KC3 and KC4 with nominal well-widths 1.8 and 1.0 ML sit on peak energies ~ 2.60 and 2.75 eV, respectively, with linewidths of ~ 150 and 60 meV, respectively. The emission bands in the ALE-grown single QWs KC1 and KC2 with nominal well-widths of 1.0 ML and ½ ML, sit on ~ 2.72 and 2.76 eV, respectively. Unlike the MOVPE-grown samples, these have much narrower PL lines of linewidths ~ 6 and 10 meV, respectively. A comparison of all the samples reveals that in the ALE systems, the PL lines feature relatively weaker exponential tails on the low energy side extending between ~ 20-50 meV, compared to the much stronger tails in the MOVPE samples.

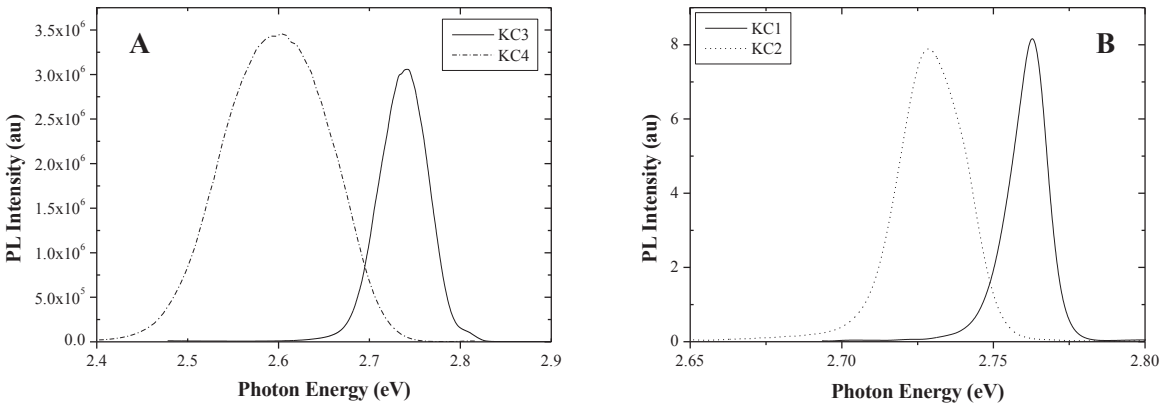


Figure 1. PL spectra of the studied samples taken at temperature of 12 K.

5. PL surface maps

In this part, spatial variations in the exciton band PL spectra were measured by surface mapping of the emission by scanning the sample laterally in the plane of the QW, i.e., across the sample area. The luminescence intensity was then mapped as a function of the emission peak energy and emission linewidth and the spatial coordinates *x* and *y*, resulting in contour plots shown in Figure 2. Even at the scale of our plots, it can be seen that there is some good correlation in the contour patterns of the peak energy and linewidth in the MOVPE structures with large variations in these values across the sample compared to the ALE structures with little variatians. In the former samples, the variations in peak energy and linewidth across the sample area depict well matched patterns showing a strong spatial correlation between the variations in the peak shift and linewidth. These large variations of the order of magnitude ~50 meV and ~30 meV for peak energy and linewidth, respectively, are associated with structures that predominantly exhibit large well width fluctuations across the plane of the well region, in good agreement with the structural results previously reported [14]. On the other hand the latter structures exhibit peak energy and linewidth variations across the sample of the order of ~1.5 meV, respectively, a far smaller change compared to that in the former. Emission in the ALE samples is expected predominantly from self-assembled quantum dots (QDs). The rather narrow Pl lines and very weak fluctuations in the peak energy and linewidth with position suggests two possibilities: (i) either there are too few of these QDs for their size-effects to be noticed or/and (ii) the dots possess size homogeinity. From structural results, we suggest that an interplay between low dot densities and less spread in size of the QD recombination centres influence the quality of the luminescence.

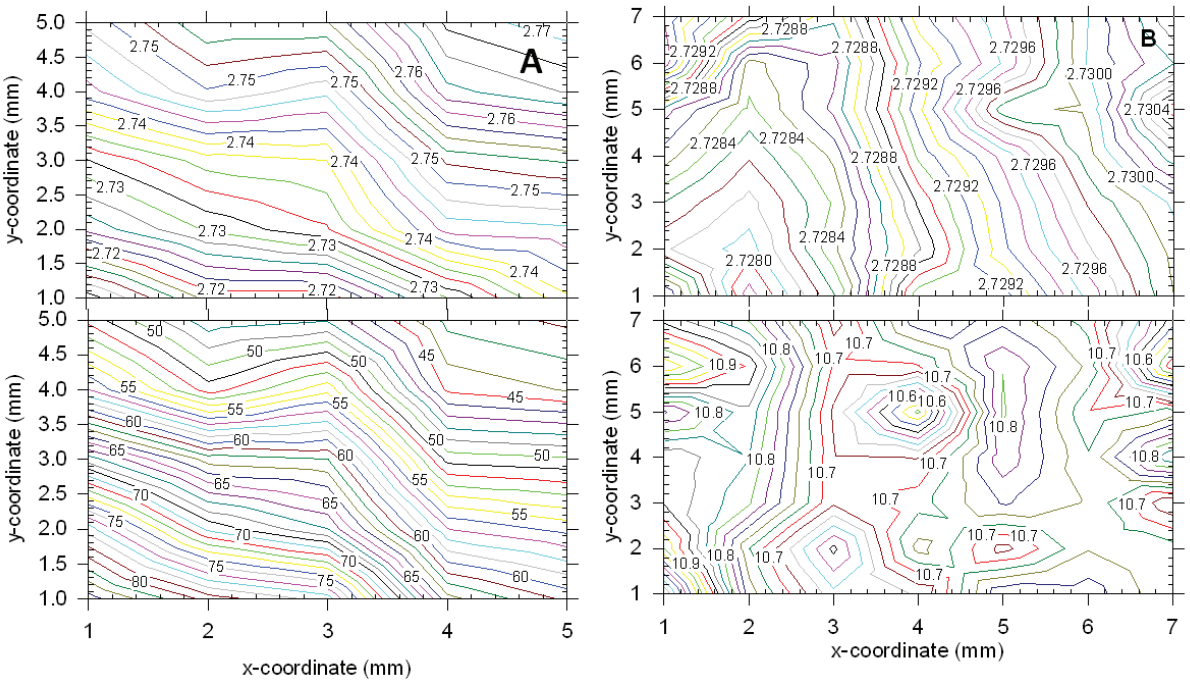


Figure 2. Contour maps showing the variation of the exciton peak energy and linewidth across the (A) MOVPE and (B) ALE grown samples.

Figure 3 gives a plot of linewidth versus peak energy. This shows more vividly the difference in quality between the MOVPE and ALE grown structures with the former systems possessing large PL linewidths than the latter. The strong correlation between the well thickness and linewidth in the former case shows a general decrease in linewidth with increasing exciton band peak energy in these systems. This means that regions with smaller well-widths have narrower PL bands, and this pattern is generally reflected in the overall behaviour of both the MOVPE and ALE systems. This is contrary to the III-V and other II-VI, systems reported in which the opposite is observed to be the case [15-20]. Furthermore, within the range of well-width sizes studied here, ½ - 3.5 ML, there appears to be a linear relationship between the well-width size and its exciton band linewidth. From the mapping, we reveal a well-defined linear variation of well-width and PL linewidth across the samples, with linear fits, $\Gamma(E) \sim \beta_0 + \beta_1 E$ (where the β's coefficients or empirical parameters are related to the sample quality) to the data points yielding individual slope values (giving a measure of the roughness) of -0.64, and –0.73, respectively. As already observed earlier, the data points for the more homogeneous ALE series are clouded over the same point, emphasising the uniformity of the active layer. For clarity, we apply the curve fits shown. A linear relationship would entail the line width continuously increases with peak energy. On the other hand, a more detailed analysis reveals, $\Gamma(E) \sim \beta_0 + \beta_1 E - \beta_2 E^2$, i.e. a quadratic dependence offers a more befitting model considering that for bulk materials, the PL linewidth is generally less than for QW structures. If we were to extrapolate the $\Gamma(E)$ curve down towards the band-edge energy 1.84 eV for bulk strained CdSe, an ideal peak energy dependence of linewidth should approach bulk values so that a quadratic function, with accurately defined β–coefficients, provides a more sensible approximation to the overall behaviour. A precise account of the β-coefficients would require a more detailed and rigorous study which is not within the scope of this work. Nonetheless, a closer analysis of this quadratic form leads us to conclude the parameters are strong functions of the growth procedure and depend a great deal on the interface and alloy disorder in the systems.

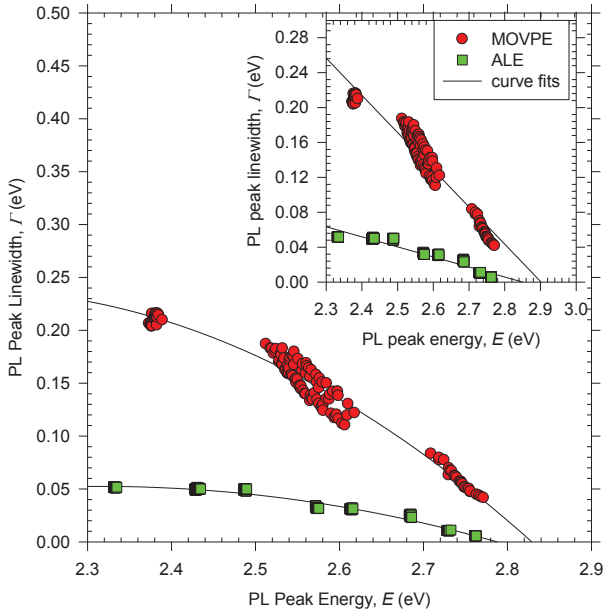


Figure 3. Overall plot of all the mapping data points comparing the MOVPE and ALE grown systems. The insert shows the linear fits to the points.

6. Conclusion

The luminescence of the QWs, the alloy-disordered and rough structures, has also revealed some new information that further distinguishes these two categories. Our PL characterisation shows that knowledge about whether the structures are alloy-disordered or rough can be derived by a simple PL surface mapping across the sample area. In rough structures, there is strong correlation between variations in peak energy position and the linewidth of the main exciton band. There is a generally a linear decrease in linewidth with increasing peak energy across the sample, displaying the typical behaviour of the CdSe-ZnSe QW systems, that thinner well widths give narrower PL lines than thicker ones. On the other hand, in alloy-disordered structures, there is virtually no correlation between the peak energy and linewidth. The variations in peak energy and linewidth across the sample area is weak and very random. In fact this is because they do not show any significant spatial fluctuations either in peak energies or linewidths across the sample, at least relative to the rough structures. A contour plot of the luminescence will show, for the rough structures, orderly patterned contour lines for both peak energy and linewidth, with their overall contour maps well correlated. On the contrary, contour plots for the QD samples show uncorrelated maps, with the linewidth contour lines rather very randomly distributed in relation to those of the peak energy. So for the first time we have shown that, even without structural characterisation, one can tell the quality of the sample (whether alloy-disordered or rough) by carrying out a prior surface mapping of the emission and producing a contour map of peak energies and linewidths. Surface mapping has also revealed variations in peak energy across the rough samples do not only correspond to monolayer fluctuations in well-width, but fluctuations in steps of submonolayers are very common. This explains why thinner QWs give narrower lines compared to thicker one in these systems. Further, this mapping analysis has shown alloy-disordered structures give good quality structures in terms of optical properties in comparison to the rough ones. We see more intense and less broadened exciton emission bands in the former relative to the latter.

Clearly the PL peak energy is a strong function of the growth sequence so that it is still impractical to accurately tie calculated and measured QW exciton-band peak energies. In other words, a ‘universal’ PL peak energy for QW insertions that can be precisely tied with a particular well-width

does not exist for these II-VI structures, unlike the well established III-V systems where it may be possible because of the high quality of the produced structures.

References

[1] Parbrook P J, Wright P J, Cockayne B, Cullis A G, Henderson B and O'Donnell K P 1990 *J. Cryst. Growth* **106** 503
[2] Zhu Z, Yoshihara H, Takebayashi K and Yao T 1994 *J. Cryst. Growth* **138** 619
[3] Fujita S, Wu Y, Kawakami Y and Fujita S 1992 *J. Appl. Phys.* **72**(11), 5233 (1992).
[4] Rosenauer A, Reisinger T, Steinkirchner E, Zweck J and Gebhardt W 1995 *J. Cryst. Growth* **152** 42
[5] Xin S H, Wang P D, Yin A, Kim C, Dobrowolska M, Merz J L and Furdyna J K 1996 *Appl. Phys. Lett.* **69** (25) 3884
[6] Zhang B P, Yasuda T, Segawa Y, Yaguchi H, Onabe K, Edamatsu E and Itoh T 1997 *Appl. Phys. Lett.* **70** (18) 2413
[7] Ledentsov N N, Krestnikov I L, Maximov M V, Ivanov S V, Sorokin S L, Kop'ev P S, Alferov Zh I, Bimberg D and Sotomayor Torres C M 1997 *Appl. Phys. Lett.* **70** (20) 2766
[8] Leonardi K, Heinke H, Ohkawa K, Hommel D, Selke H, Gindele F and Woggon U 1997 *Appl. Phys. Lett.* **71** (11) 1510
[9] Rose M, Lowisch M, Kreller F and Henneberger F 1997 *Phys. Stat. Sol. (b)* **202** 1817
[10] Zhu Z, Kurtz E, Arai K, Chen Y F, Bagnall D M, Tomasini P, Lu F, Sekiguchi T, Yao T, Yasuda T and Segawa Y 1997 *Phys. Stat. Sol. (b)* **202** 827
[11] Hommel D, Leonardi K, Heinke H, Selke H, Ohkawa K, Gindele F and Woggon U 1997 *Phys. Stat. Sol.(b)* **202** 835
[12] Suemune I, Uesugi K, Suzuki H, Nashiki H and Arita M 1997 *Phys. Stat. Sol (b)* **202** 845
[13] Shen M Y, Goto T, Kurtz E, Zhu Z and Yao T 1997 *J. Phys. Condens. Matter* **10** L171
[14] Chinyama K G, O'Donnell K P, Rosenauer A and Gerthsen D 1999 *J. Cryst. Growth* **203** 362
[15] Ivanov S V, Toropov A A, Shubina T V, Sorokin S V, Lebedev A V, Sedova I V and Kop'ev P S 1998 *J. Appl. Phys.* **83** 3168
[16] Flack F, Samarth N, Nikitin V, Crowell P A, Shi J, Levy J and Awschalom D D 1996 *Phys. Rev.* **B 54** R17 312
[17] Matsumoto T, Iwashita T, Sasamoto K and Kato T 1994 *J. Cryst. Growth* **138** 63
[18] O'Donnell K P and Middleton P G 1997 in *Wide Bandgap II-VI Semiconductors*, ed. R. Bhargava, EMIS Datareviews Series No.**17** 33
[19] Parbrook P J 1990 *PhD. Thesis* University of Strathclyde Glasgow Scotland
[20] Lozykowki H J and Shastri V K 1991 *J. Appl. Phys.* **69** 3235

Magnetic vector charges in the mystery of a circular current's *pair* of distinct Cartesian elemental magnetic dipoles

M Chirwa

Physics Department, Faculty of Science, Engineering and Technology, Walter Sisulu University, P/Bag X1, Nelson Mandela Drive, Mthatha 5117, South Africa.

E-mail: mmchirwa@wsu.ac.za

Abstract. A circular electric current, perpendicularly bisected by a *field plane*, was modelled as a continuum of pairs of distinct Cartesian component elemental magnetic *vector* charges normal and parallel to the field plane. The Cartesian component elemental magnetic *vector* charges normal to the field plane pair up into Cartesian elemental magnetic dipoles with intra-dipolar displacements parallel to the field plane. These dipoles generate the overall magnetic vector potential. The Cartesian elemental magnetic *vector* charges parallel to the field plane form Cartesian elemental magnetic dipoles (with intra-dipolar displacements perpendicular to the field plane) which individually and collectively contribute nothing to the magnetic vector potential. Each continuum of these two sets of Cartesian elemental magnetic dipoles independently yields the traditionally renowned "magnetic dipolar moment of a circular current". However, together their distinct magnetic fields and their distinct magnetic torques respectively constitute the circular current's overall magnetic field and magnetic torque. These results reconcile only if the magnetic dipolar moments of both sets are endorsed, that is a circular current of any spatial size is a continuum of pairs of distinct Cartesian elemental magnetic dipoles. In addition the customary *ad hoc* definition of magnetic dipole moment is deceptively erroneous, thus prompting a review of many relations involving it. These include the magnetic torque and magnetic field generated by it, and the classical magneto-mechanical ratio.

1. Introduction

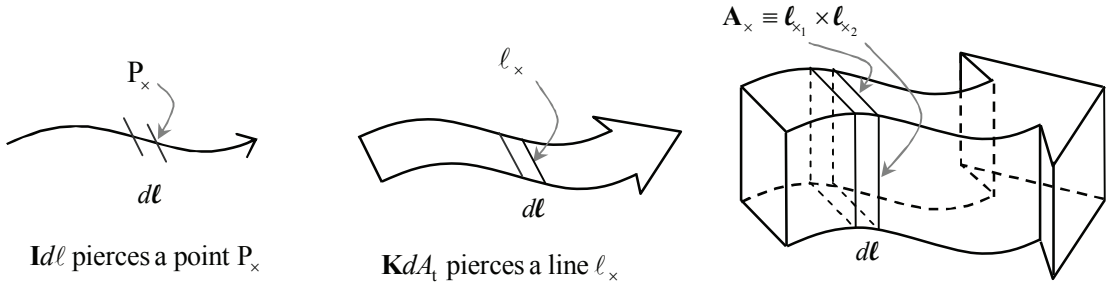
This article arose from the humble aim of verifying from first principles, that the new theory of representing any source of magnetic fields as a distribution of elemental magnetic vector charges [1] could reproduce the traditionally acclaimed magnetic properties of or due to a circular electric charge current [2–7]. Further, one is amazed at the traditional definition of the magnetic moment either in terms of a hitherto non-existent scalar quantity (Gilbert model of 1600 or Dirac's magnetic charge of 1931) or in terms of a *complete* current loop only (Ampère model) [2–7]. Both models are not exactly in line with the moments of other physical quantities.

After identifying the various continuous distributions of magnetic vector charges and modeling a circular line current as one of these, we demonstrate by deriving from first principles that a circular line current has two distinct Cartesian elemental magnetic dipoles which must be fully considered as essential elements in deriving its various magnetic properties, and that the above traditional assumptions are faulty. Here easier evaluation of vector cross-products is due to expressing vectors in

terms of Cartesian unit vectors, with the yz-plane as the field plane and the source positions in the xy-plane.

2. Continuously distributed electric currents and magnetic vector charges

Figure 1 shows our models of continuously distributed elements of electric current vectors in elemental spaces of line length $d\ell$, tangential surface area dA_{t}, and space volume dv with respective cross-sections of point $\mathrm{P}_{\times}$, line length $\ell_{\times}$, surface area $A_{\times}$, and spatial current density vectors $\mathbf{I}$, $\mathbf{K}$, $\mathbf{J}$. By the harmonized definition, the line, surface and volume elemental magnetic vector charges are



$\mathbf{I}d\ell$ pierces a point $\mathrm{P}_{\times}$ $\quad$ $\mathbf{K}dA_{\mathrm{t}}$ pierces a line $\ell_{\times}$ $\quad$ $\mathbf{J}dv$ pierces a plane $A_{\times}$

Figure 1. Elemental spaces $d\ell$, dA_{t}, dv, cross sections $\mathrm{P}_{\times}$, $\ell_{\times}$, $A_{\times}$ and *current densities* $\mathbf{I}$, $\mathbf{K}$, $\mathbf{J}$.

$$d\mathbf{Q} = \begin{cases} \mu_0 \mathbf{I} d\ell \equiv \mu_0 I d\boldsymbol{\ell}, & \text{along a line} \\ \mu_0 \mathbf{K} dA_{\mathrm{t}} = \mu_0 \mathbf{K} \left| \boldsymbol{\ell}_{\times} \times d\boldsymbol{\ell} \right| \equiv \left| \mu_0 \mathbf{K} \times \boldsymbol{\ell}_{\times} \right| d\boldsymbol{\ell} = \mu_0 I_{\ell_{\times}} d\boldsymbol{\ell}, & \text{over a tangential surface} \\ \mu_0 \mathbf{J} dv = \mu_0 \mathbf{J} \left(\mathbf{A}_{\times} \cdot d\boldsymbol{\ell} \right) \equiv \left(\mu_0 \mathbf{J} \cdot \mathbf{A}_{\times} \right) d\boldsymbol{\ell} = \mu_0 I_{\mathbf{A}_{\times}} d\boldsymbol{\ell}, & \text{in a volume} \end{cases} \qquad (1)$$

Note the alternative representations on the extreme right hand sides where $\mu_0 I$, $\mu_0 I_{\ell_{\times}}$ and $\mu_0 I_{\mathbf{A}_{\times}}$ are magnitudes of the magnetic vector charges *per unit length* of the line, surface and volume spatial distributions respectively, and in each case the direction is given by the elemental displacement $d\boldsymbol{\ell}$.

Traditionally [2–7], I the magnitude of $\mathbf{I}$ is inaptly called current, while *simultaneously* accepting $I\mathrm{d}\boldsymbol{\ell}$ as a *current element* (that is, an elemental line current). Then $\mathbf{K}$ and $\mathbf{J}$ are confusingly viewed as "surface" and "volume" densities of I over perpendicular components of the cross sections $\ell_{\times}$ and $A_{\times}$. This is an apparent mix-up between the space occupied and the cross section of that space, due to not recognizing the two alternative representations of current distributions. Rather I should be referred to as the current-flux, so that $\mathbf{I}$, $\mathbf{K}$ and $\mathbf{J}$ can also be called *point*, *line* and *surface* current-flux densities respectively.

3. Magnetic moments and magnetic torques of the distinct Cartesian magnetic dipoles

In figure 2 each of the line elemental magnetic vector charges $d\mathbf{Q}_j = \hat{\boldsymbol{\varphi}}_j Q_0 d\phi, j = 1, 2, 3, 4$, with $Q_0 = \mu_0 \rho I$, represents an azimuthal line current element $\mathbf{I}_j d\ell = \hat{\boldsymbol{\varphi}}_j I \rho d\phi$ at point P_j, which has the radial position $\hat{\boldsymbol{\rho}}_j \rho$ on a circle of radius ρ and centred at origin O. As each of the four position vectors is inclined at an angle ϕ to the x-axis, the angular position ϕ_j of P_j takes on the respective values of $2\pi - \phi$, ϕ, $\pi - \phi$ and $\pi + \phi$. Hence each $d\mathbf{Q}_j$ has Cartesian components selectable from the following Cartesian elemental magnetic vector charges that are normal and parallel to the yz-plane (the field plane):

$$\pm d\mathbf{Q}_a = \pm \hat{\mathbf{x}} dQ_a = \pm \hat{\mathbf{x}} Q_0 \sin\phi \, d\phi \qquad \text{and} \qquad \pm d\mathbf{Q}_b = \pm \hat{\mathbf{y}} dQ_b = \pm \hat{\mathbf{y}} Q_0 \cos\phi \, d\phi \qquad (2)$$

When the $+d\mathbf{Q}_a$ at P_1 pairs up with the $-d\mathbf{Q}_a$ at P_2, while the $+d\mathbf{Q}_b$ at P_2 pairs up with the $-d\mathbf{Q}_b$ at P_3, they constitute the *a* and *b* Cartesian elemental magnetic dipoles with respective *di*polar moments [8]

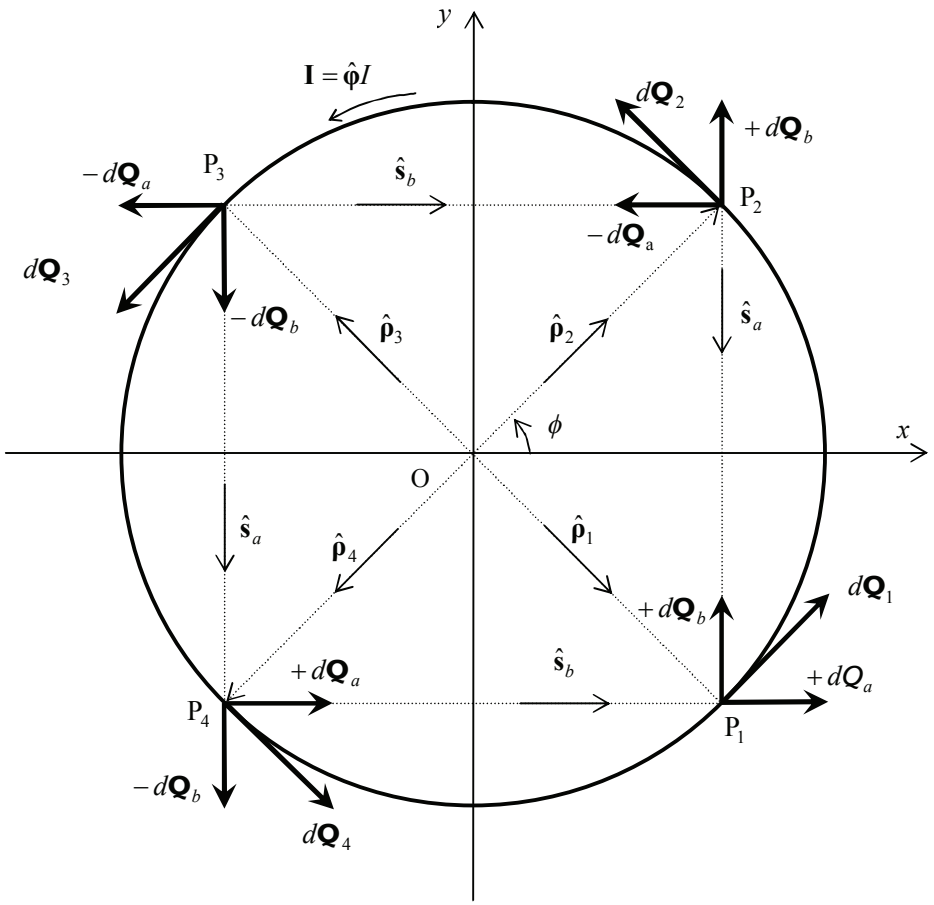


Figure 2. Separated Cartesian elemental magnetic vector charges pair up into magnetic dipoles.

$$dm_a = dm_{a_+} + dm_{a_-} = (\hat{\boldsymbol{\rho}}_1 - \hat{\boldsymbol{\rho}}_2)\rho \times d\mathbf{Q}_a = \mathbf{s}_a \times \hat{\mathbf{x}} dQ_a = -\hat{\mathbf{y}} s_a \times \hat{\mathbf{x}} dQ_a = \hat{\mathbf{z}} dm_a \quad (3a)$$

$$dm_b = dm_{b_+} + dm_{b_-} = (\hat{\boldsymbol{\rho}}_2 - \hat{\boldsymbol{\rho}}_3)\rho \times d\mathbf{Q}_b = \mathbf{s}_b \times \hat{\mathbf{y}} dQ_b \equiv +\hat{\mathbf{x}} s_b \times \hat{\mathbf{y}} dQ_b = \hat{\mathbf{z}} dm_b \quad (3b)$$

The second level subscripts + and − in (3a) and (3b) signify the *mono*polar contributions (to the magnetic moment) of the related individual Cartesian elemental magnetic vector charges specified in (2). This notation is similarly used in subsequent discussions below.

The two Cartesian elemental magnetic dipolar moments in (3a) and (3b) are typically distinct as

$$s_a = 2\rho \sin\phi \neq 2\rho \cos\phi = s_b \quad \text{and} \quad dm_a = s_a dQ_a \neq s_b dQ_b = dm_b \quad (4)$$

Integrating (3a) and (3b) from $\phi = 0$ to $\phi = \pi$ accounts for entire loop to yield

$$m_a \equiv \hat{\mathbf{z}} m_a = \hat{\mathbf{z}} \pi \rho Q_0 \quad (5a)$$

$$m_b \equiv \hat{\mathbf{z}} m_b = \hat{\mathbf{z}} \pi \rho Q_0 \quad (5b)$$

Thus the overall magnetic dipolar moment $m = m_a + m_b = \hat{\mathbf{z}} 2\pi \rho Q_0$ is twice the traditional value for a circular current [2–7].

In a magnetic field $\mathbf{H}$, the *a* and *b* Cartesian elemental magnetic dipoles are characterized by the magnetic forces $d\mathbf{F}_a = d\mathbf{Q}_a \times \mathbf{H}$ and $d\mathbf{F}_b = d\mathbf{Q}_b \times \mathbf{H}$ and their opposites. Then the coupled elemental magnetic torques [8] on the dipoles are the moments of the magnetic forces:

$$d\boldsymbol{\tau}_a = d\boldsymbol{\tau}_{a_+} + d\boldsymbol{\tau}_{a_-} = (\hat{\boldsymbol{\rho}}_1 - \hat{\boldsymbol{\rho}}_2)\rho \times d\mathbf{F}_a = \mathbf{s}_a \times d\mathbf{F}_a \equiv -\hat{\mathbf{y}} s_a \times (\hat{\mathbf{x}} dQ_a \times \mathbf{H}) \quad (6a)$$

$$d\boldsymbol{\tau}_b = d\boldsymbol{\tau}_{b_+} + d\boldsymbol{\tau}_{b_-} = (\hat{\boldsymbol{\rho}}_2 - \hat{\boldsymbol{\rho}}_3)\rho \times d\mathbf{F}_b = \mathbf{s}_b \times d\mathbf{F}_b \equiv +\hat{\mathbf{x}} s_b \times (\hat{\mathbf{y}} dQ_b \times \mathbf{H}) \quad (6b)$$

After expressing the magnetic field as $\mathbf{H} = \hat{\mathbf{x}} H_x + \hat{\mathbf{y}} H_y + \hat{\mathbf{z}} H_z$, the above torques become

$$d\boldsymbol{\tau}_a = -\hat{\mathbf{y}} \times (\hat{\mathbf{z}} H_y - \hat{\mathbf{y}} H_z) s_a dQ_a = -\hat{\mathbf{x}} H_y d m_a \neq d m_a \times \mathbf{H} \quad (7a)$$

$$d\boldsymbol{\tau}_b = +\hat{\mathbf{x}} \times (-\hat{\mathbf{z}} H_x + \hat{\mathbf{x}} H_z) s_b dQ_b = +\hat{\mathbf{y}} H_x d m_b \neq d m_b \times \mathbf{H} \quad (7b)$$

Note that $\mathrm{d}\boldsymbol{\tau}_a \neq \mathrm{d}\boldsymbol{\tau}_b$ and $\boldsymbol{\tau}_a \neq \boldsymbol{\tau}_b$. Then the overall coupled torque is

$$\boldsymbol{\tau} = \boldsymbol{\tau}_a + \boldsymbol{\tau}_b = -\hat{\mathbf{x}} m_a H_y + \hat{\mathbf{y}} m_b H_x \equiv m_a \times \mathbf{H} \equiv m_b \times \mathbf{H} \neq m \times \mathbf{H} \equiv 2\boldsymbol{\tau} \quad (8)$$

The equivalences are *accidental* and non-basic relations (cannot be used to define torque) because, although $m_a = m_b$, the differentials dm_a and dm_b differ in size and originating elemental dipoles. Thus the traditional choice [2–7] of m_a or m_b as the total magnetic dipolar moment is unjustified.

4. Magnetic vector potentials and magnetic fields of the distinct Cartesian magnetic dipoles

When the magnetic vector charge distributions in figure 2 become the sources of the magnetic vector potential and magnetic field at a field point P, the magnetic vector charges and their positions are signified by primed symbols. Thus, as shown in figure 3, the displacements of any field point P in the yz-plane from the origin O and the source points P'_j, $j = 1, 2, 3$ on the circle are respectively

$$\mathbf{r} = \hat{\mathbf{r}} r \qquad \text{and} \qquad \mathbf{R}_j = \hat{\mathbf{R}}_j R_j = \hat{\mathbf{r}} r - \hat{\boldsymbol{\rho}}'_j \rho' \quad (9)$$

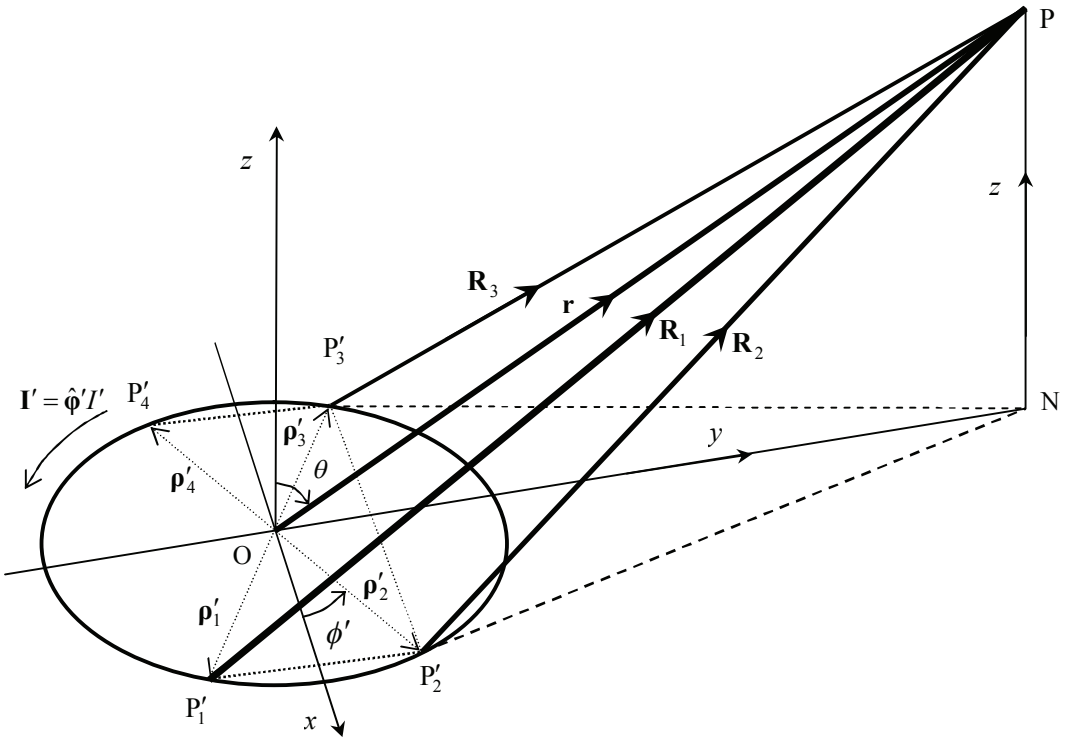


Figure 3. Positions of source points P_1, P_2, P_3 in the xy-plane, and field point P in the yz-plane.

These are expressible in terms of r, θ, ρ', ϕ', $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$. Below, all ensuing first order approximations for $\rho' << r$ are after binomial expansions of the indicated geometrical factors

$$\frac{r^n}{R_j^n}$$

where $n = 1$ or 3, and application of equations (3a), (3b) and (4) for the elemental magnetic dipole moments.

Then the two distinct Cartesian elemental magnetic dipoles on the circle will generate at P the harmonized magnetic vector potentials

$$d\mathsf{A}_a = d\mathsf{A}_{a_+} + d\mathsf{A}_{a_-} = \frac{\hat{\mathbf{x}} dQ'_a}{4\pi\mu_0 r}\left(\frac{r}{R_1} - \frac{r}{R_2}\right) \approx \frac{-\hat{\mathbf{x}} dm'_a \sin\theta}{4\pi\mu_0 r^2} \equiv \frac{\hat{\mathbf{z}} dm'_a \times \hat{\mathbf{r}}}{4\pi\mu_0 r^2} \tag{10a}$$

$$d\mathsf{A}_b = d\mathsf{A}_{b_+} + d\mathsf{A}_{b_-} = \frac{\hat{\mathbf{y}} dQ'_b}{4\pi\mu_0 r}\left(\frac{r}{R_2} - \frac{r}{R_3}\right) = \frac{\hat{\mathbf{y}} dQ'_b}{4\pi\mu_0 r}\left(\frac{r}{R_2} - \frac{r}{R_2}\right) = \mathbf{0} \tag{10b}$$

This vanishing of $d\mathsf{A}_b$ is due to the fact that the elemental magnetic vector charges $+d\mathbf{Q}'_b$ and $-d\mathbf{Q}'_b$ are equidistant to any point in the field plane, so that their individual magnetic vector potentials $d\mathsf{A}_{b_+}$ and $d\mathsf{A}_{b_-}$ are identical in magnitude but opposite in direction; it does not mean an absence of dm'_b [8]. Then the total magnetic vector potential is

$$\mathsf{A} = \mathsf{A}_a + \mathsf{A}_b = \mathsf{A}_a = \frac{\hat{\mathbf{z}} m'_a \times \hat{\mathbf{r}}}{4\pi\mu_0 r^2} \equiv \frac{\hat{\mathbf{z}} m'_b \times \hat{\mathbf{r}}}{4\pi\mu_0 r^2} = \frac{\hat{\boldsymbol{\varphi}} m'_a \sin\theta}{4\pi\mu_0 r^2} \tag{11}$$

Traditional wisdom [7] uses (11) to assign m'_a the status of "magnetic dipole moment of the loop" [2–7], thus inadvertently ignoring the existence of m'_b. Such a conclusion can be likened to stating that, since $\mu_0 \boldsymbol{\mu}_s \cdot \mathbf{H} = \mu_0 \mu_{s_z} H$ is the electron's spin magnetic energy, the electron's total spin magnetic dipole moment μ_s reduces to its z component μ_{s_z}!

Similarly, the magnetic fields at point P due to the two distinct Cartesian elemental magnetic dipoles are

$$d\mathbf{H}_a = d\mathbf{H}_{a_+} + d\mathbf{H}_{a_-} = \frac{\hat{\mathbf{x}} dQ'_a}{4\pi\mu_0 r^3} \times \left(\mathbf{R}_1 \frac{r^3}{R_1^3} - \mathbf{R}_2 \frac{r^3}{R_2^3}\right) \approx \frac{dm'_a}{4\pi\mu_0 r^3}\left[3\sin\theta\left(\hat{\mathbf{y}}\cos\theta - \hat{\mathbf{z}}\sin\theta\right) + \hat{\mathbf{z}}\right] \tag{12a}$$

$$d\mathbf{H}_b = d\mathbf{H}_{b_+} + d\mathbf{H}_{b_-} = \frac{\hat{\mathbf{y}} dQ'_b}{4\pi\mu_0 r^3} \times \left(\mathbf{R}_2 \frac{r^3}{R_2^3} - \mathbf{R}_3 \frac{r^3}{R_3^3}\right) = \frac{\hat{\mathbf{y}} dQ'_b}{4\pi\mu_0 r^3} \frac{r^3}{R_2^3} \times \left(\mathbf{R}_2 - \mathbf{R}_3\right) \approx \frac{dm'_b}{4\pi\mu_0 r^3} \hat{\mathbf{z}} \tag{12b}$$

Unlike for the magnetic vector potentials $d\mathsf{A}_{b_+}$ and $d\mathsf{A}_{b_-}$ in (10b), here in (12b) the individual magnetic fields $d\mathbf{H}_{b_+}$ and $d\mathbf{H}_{b_-}$, which are due to the same $+d\mathbf{Q}'_b$ and $-d\mathbf{Q}'_b$, do not cancel out.

Integrating (12a) and (12b), and then transforming unit vectors from Cartesian to spherical systems yields

$$\mathbf{H}_a = \frac{m'_a}{4\pi\mu_0 r^3}\left(\hat{\mathbf{r}}\cos\theta + \hat{\boldsymbol{\theta}} 2\sin\theta\right) \tag{13a}$$

$$\mathbf{H}_b = \frac{m'_b}{4\pi\mu_0 r^3}\left(\hat{\mathbf{r}}\cos\theta - \hat{\boldsymbol{\theta}}\sin\theta\right) \tag{13b}$$

Then since $m'_a = m'_b$ (equations 5a and 5b), the overall magnetic field becomes

$$\mathbf{H} = \mathbf{H}_a + \mathbf{H}_b = \frac{m'_a}{4\pi\mu_0 r^3}\left(2\hat{\mathbf{r}}\cos\theta + \hat{\boldsymbol{\theta}}\sin\theta\right) \equiv \frac{m'_b}{4\pi\mu_0 r^3}\left(2\hat{\mathbf{r}}\cos\theta + \hat{\boldsymbol{\theta}}\sin\theta\right) \tag{14}$$

This cannot justify taking m'_a or m'_b as the circular current's *only* magnetic moment. Also applying the relation $\mathbf{H} = \nabla \times \mathsf{A} \equiv \nabla \times \mathsf{A}_a$ on (11) to get (14), tradition [7] barely "evades" this scrutiny.

5. Conclusions

Depicting an electric charge current in a circle as a continuum of paired Cartesian elemental magnetic vector charges, parallel and normal to the field plane perpendicularly bisecting the circle, provides realistic models of the structure and attributes of an elemental magnetic dipole. It shows that, contrary to tradition, a circular current, whatever its spatial size, consists of two continuous distributions of distinct Cartesian elemental magnetic dipoles. Each has a distinct nonzero contribution to the overall magnetic dipole moment (which is twice the renowned traditional value), the magnetic field and magnetic torque. Characteristically, the total magnetic vector potential is entirely due one of the continuous distributions of Cartesian elemental magnetic dipoles. Other fundamental outcomes of the above theory include:

(1) Doubling of the classical magneto-mechanical ratio which agrees with Dirac's relativistic electron theory [3].

(2) The traditional analogies between the structures and torques of electric and magnetic dipoles are deceptively erroneous.

(3) Quite generally relations involving the traditional magnetic moment need to be looked at.

References

[1] Chirwa M 2009 A New theory of elemental electromagnetic scalar and vector charges *Brochure of Programme and Abstracts*, 54th Annual Conf. of the South African Institute of Physics (Durban, South Africa, 6–10 July 2009) abstract 73

[2] Jackson J D 1975 *Classical Electrodynamics* 2nd ed. (New York: John Wiley & Sons) pp178–86

[3] Griffiths D J 1999 *Introduction to Electrodynamics* 3rd ed. (London: Prentice–Hall) pp 242–58

[4] Pollack G L and Stump D R 2002 *Electromagnetism* (Cape Town: Addison-Wesley) pp 284–8

[5] Wangsness R K 1986 *Electromagnetic Fields* 2nd ed. (New York: John Wiley & Sons) pp 297–308

[6] Panosfky W K H and Phillips M 1962 *Classical Electricity and Magnetism* (London: Addison-Wesley) pp 128–33

[7] Lorrain P, Corson D P and Lorrain F 1987 *Electromagnetic Fields and Waves* 3rd ed. (New York: W H Freeman & Co) pp 331–41

[8] Chirwa M 2010 Status of electric and magnetic dipole moments in a complete characterization of electromagnetic first order moments *Advancing Knowledge for Empowerment Through Innovations in Research: Proc. 2nd WSU Int. Research Conf.* (East London, South Africa, 25–27 August 2009) (Walter Sisulu University) pp 88–96

Computational modelling of Zr-Nb alloys

MH Chuma, PE Ngoepe and HR Chauke

Materials Modeling Centre, University of Limpopo, Turfloop campus, Private Bag x1106, Sovenga, 0727, SA

Corresponding author: hr.chauke@ul.ac.za

Abstract. We use density functional theory to investigate the structural, elastic properties and energetic stabilities of Zr, Nb and its alloys in the alpha and beta phases, employing pseudopotential plane wave methods within the Pardew-Burke-Ernzerhof Generalized Gradient Approximation (PBE-GGA). The structures were fully optimized in a non-magnetic state, allowing atomic positions, cell volume and shape to change. Lattice parameters for the pure phases gave excellent agreement with the available experimental data. We also found that the phonon dispersions curves display soft-modes for the metastable beta-Zr phase which is lacking on the alpha-Zr and beta-Nb, condition of mechanical stability. The virtual crystal approximation calculations show that an increase in the Nb contents stabilizes the Zr structure. Interestingly, the calculated elastic moduli for Zr-2.3%Nb are in good agreement to the experimental results of Zr-2.5%Nb.

1. Introduction

Recently, zirconium has attracted lot of attention due to its high resistance to corrosion and high melting point 1855 °C. It is soft, ductile and a malleable metal, solid at room temperature, though it becomes hard and brittle at lower purities. However, its mixture with niobium (high melting point of 2,468 °C) has desirable physical and thermodynamic applications; it exhibits superelasticity properties, superconductivity properties and forms dielectric oxide layers [1].

Zirconium alloys are extensively used in various types of fission reactors both light and heavy water types for different applications, since 1960 zirconium alloys were the principal cladding material due to the good irradiation stability, high mechanical strength, the superior neutron economy and corrosion resistance etc. These alloys are available to designers of high level nuclear waste disposal containers as internal components or external cladding [2].

There are different compositions of zirconium alloys, the most common alloys that has been used in the nuclear technology are zrcaloy-2 which is composed of Zr-1.5%Sn-0.15%Fe-0.1%Cr-0.05%Ni and has been predominantly used as fuel cladding in Boiling Water Reactors (BWR), the other is zrcaloy-4 which is composed of Zr-1.5%Sn-0.2%Fe-0.1%Cr and is typically used as fuel cladding in Pressurized Water Reactors (PWR) and Canadian Deuterium Uranium (CANDU) reactors and Zr-2.5%Nb which has been of interest as the pressure tube material in CANDU and RBMK reactors [3,4,5]. Zr-2.5%Nb alloy plays a significant role in structural material for pressure tubes of channel nuclear power reactors, and have been widely investigated for various mechanical properties such as tensile and creep resistance, resistance to plastic flow and rupture to reactor pressure types [6,7,8].

However more advanced Zr-based alloys are required for the more severe operating conditions such as higher burn-up, increased operation temperature, and high pH operation. Most new Zr alloys that are

being developed contain the Nb element, for example Zirlo (Zr-1.0Nb-1.0Sn-0.1Fe) [3], M5 (Zr-1Nb-O) [4] and NDA (Zr-0.1Nb-1.0Sn-0.27Fe-0.16Cr) [5]. The tendency to select Nb as the major alloying element in such alloys is a common characteristics for the newly developed fuel cladding material [9]. Other research findings on Nb-Zr alloys, revealed that the microstructural and corrosion test with variation of cooling rate and Nb content in alpha matrix plays important role to enhance the corrosion resistance than the supersaturated Nb, the beta phase and the precipitates [3], on the other hand Kim et.al found that the Nb-Zr alloys showed a good corrosion resistance which was not affected by the annealing temperature [10].

In this paper, we report DFT results on the structural, electronic, energetic and mechanical properties of Zr, Nb and Zr-Nb alloys using virtual crystal approximation (VCA). Firstly, their energies and elastic constants will be calculated for pure Zr and Nb phases. Secondly, we will find that the α-Zr is vibrationally stable while β-Zr is mechanically unstable due to presence of imaginary frequency in the phonon dispersion spectrum. Lastly, the elastic constants and bulk properties of Zr-Nb alloys will be determined.

2. Methodology

In the present study, we use density functional theory [11,12] employing two plane wave pseudopotential codes CASTEP [13] and VASP [14], to study the structural, electronic and mechanical properties of pure Nb, Zr structures and their alloys. All calculations have been performed using the same gradient-corrected functional, namely, that of Perdew, Burke, and Ernzerhof (PBE) [15]. The phonon spectra were calculated using PHONON code [16] as implemented by Materials Design within their MedeA software platform [17] in which the DFT force constants are obtained via the VASP program [14,18]. The structures were optimized allowing atom positions to relax, cell volume and cell shape to change and calculations were performed using non-magnetic magnetism, a cutoff energy of 400eV and Monkhorst-Pack special k-point mesh of 18x18x10 for alpha Zr and 20x20x20 for beta Nb was used [19]. Zr-Nb alloys calculations were performed using the virtual crystal approximation [20] as implemented in the CASTEP code [13], to study the effect of Nb doping on the Zr site. VCA calculations are based on a weighting of the contribution of the pseudopotentials according to the site occupancies [21], and the current approach uses the mixture atom representation of the Zr and Nb compositions in a solid solution. We have investigated three possible alloys: Zr-1.6%Nb, Zr-2.3%Nb and Zr-2.6%Nb, where Nb is incorporated on the P6/mmm Zr sites constituting three atoms per unit cell. The elastic constants were calculated using six different values of the strain ±0.002, ±0.003 and ±0.004, using CASTEP on Materials Studio interface [13]. Note that the application of strain on the lattice implies a lowering of symmetry from that of the crystal, therefore very accurate total-energy calculations are required, since the energy differences involved are of the order 10 to 1000 μeV/atom. Thus an equivalent number of k-points were used for Zr-Nb alloys.

3. Summary of results and conclusion

We show the results of equilibrium lattice parameters and volume, the elastic constants and bulk moduli in table 1. Our predicted parameters for the pure zirconium structure agree very well with the experimental data to within 0.4%, while the Zr-Nb alloys show reasonable values. It should be noted that the Zr (P6/mmm) phase was found to be the most favorable structure since it had the lowest energy, in agreement with the experimental phase diagram [22]. Furthermore, the Zr-Nb alloys results show that an increase in the Nb contents stabilizes the Zr structure since the energy is lowered. It is evident from the table that addition of 1 to 1.26 atomic percent niobium content on the P6/mmm zirconium structure gives desirable results. The current observations suggest that the Zr-2.6%Nb is more energetically favorable.

The mechanical stability of these systems was evaluated using the elastic constants, Cij, and the phonon dispersion curves. Firstly, all the independent elastic constants (Cij) in the Zr and α-Zr pure

structures as well as the solid solutions (Zr-Nb alloys) are positive satisfying the stability conditions as described elsewhere [23,24,25]. However, the β-Zr phase is elastically unstable since it gives a negative tetragonal shear modulus, **C'**. Secondly, the predicted Cij for the Zr-Nb alloy give good agreement with the experimental results of Zr-2.5%Nb [26]. Lastly, the phonon dispersion spectrum of α-Zr show real phonon modes (positive frequency), that is no soft modes were detected along high symmetry directions on the Brillouin zone indicating mechanically stability, while on the other hand the β-Zr display negative frequency phonon modes, condition of mechanically instability, in agreement with the predicted tetragonal elastic modulus. In contrast to β-Zr phase, the β-Nb phonon spectra display real frequency indicating that the structure is mechanically stable at absolute zero.

Table 1. Lattice parameters, volume, total energy, elastic properties and bulk moduli of pure Zr and Z-Nb alloys, at 0K. The experimental values for Zr-2.5%Nb are also given for comparison [25].

	Zr	α-Zr	β-Zr	Zr-1.6%Nb	Zr-2.3%Nb	Zr-2.6%Nb	Zr-2.5%Nb Exp. [25]
Symmetry	P6/mmm	$P6_3/mmc$	Im-3m	P6/mmm	P6/mmm	P6/mmm	
Lattice parameters(Å)							
a	5.061	3.231	3.095	5.057	5.055	5.056	-
b	5.061	3.231	3.095	5.057	5.055	5.056	-
c	3.156	5.171	3.095	3.148	3.143	3.148	-
Angles							
α	90	90	109.47	90	90	90	-
β	90	90	109.47	90	90	90	-
γ	120	120	109.47	120	120	120	-
Volume ($Å^3$/atom)							
	70.020	46.733	22.822	63.756	69.549	69.697	
Energy (eV/atom)							
	-1281.09	-1281.08	-640.497	-1286.47	-1289.11	-1289.13	
Elastic constants (GPa)							
C11	149.56	151.99	72.168	153.87	153.48	152.79	147.08
C12	62.693	56.10	90.423	70.746	65.12	63.206	74.46
C13	45.867	-	-	45.975	45.364	46.257	75.38
C33	177.04	-	-	171.51	182.71	184.64	145.33
C44	30.740	26.09	28.248	30.539	32.091	31.436	33.81
C'	43.433	47.95	-9.127	41.562	44.188	44.792	-
Bulk modulus, B (GPa)							
	87.16	93.06	84.34	89.37	88.99	88.95	98.7

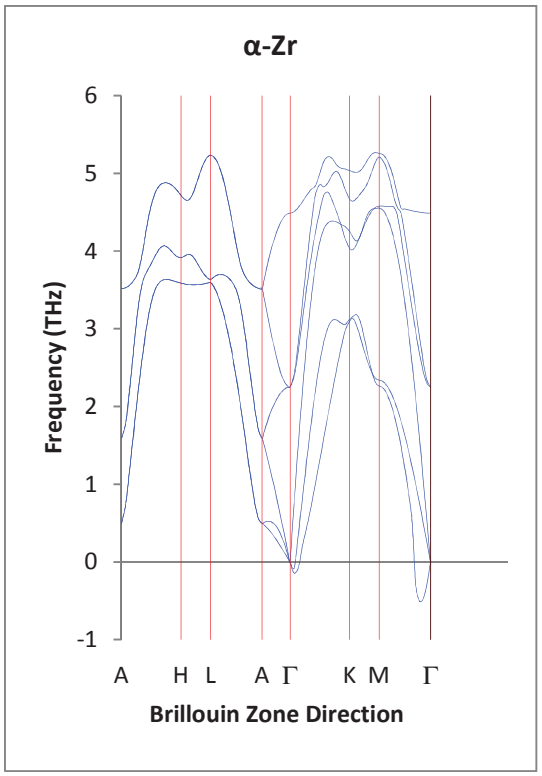


Figure 1(a). Phonon dispersion spectra of $P6_3/mmc$ α-Zr structure.

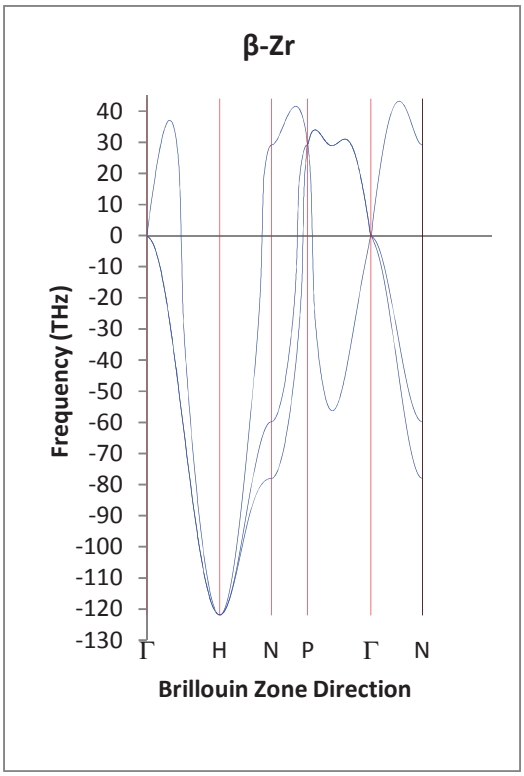


Figure 1(b). Phonon dispersion spectra of Im-3m β-Zr structure.

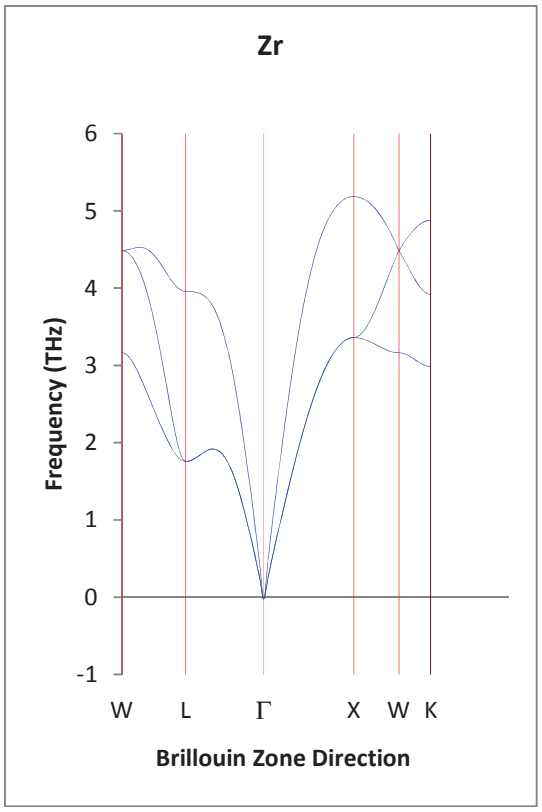


Figure 1(c). Phonon dispersion spectra of P6/mmm Zr structure.

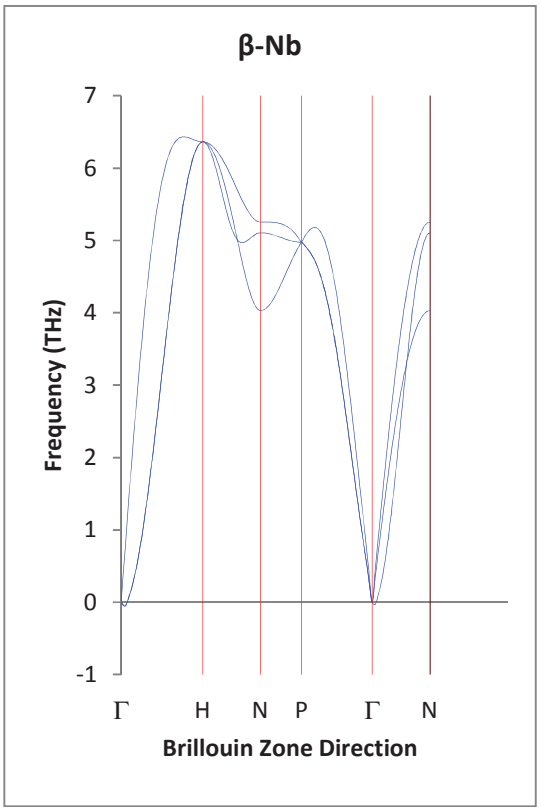


Figure 1(d). Phonon dispersion spectra of Im-3m β-Nb structure.

Acknowledgements
The computations were performed at the Materials Modelling Centre (MMC), University of Limpopo. We also acknowledge the support of the South African Research Chair Initiative of the Department of Science and Technology and the National Research Foundation.

References

[1] Emsley J, Nature's Building Blocks 2001 Oxford University Press. 506
[2] Joeng Y H, Lee K O, Kim H G 2002 *J. Nucl. Mater.* **302** 9
[3] Sabol G P, Kilp G R, Balfour M G and Roberts E 1989 *ASTM spec. Tech. Publ*. **1023** 227
[4] Yamate K, Oe A, Hayashi M, Okamoto T, Anada H and Hagi S 1997 in: Proceedings of the 1997 International Topical Meeting on LWR Fuel Performance, Portland, OR, 2-6 March. 318
[5] Ardon J P, Garner G, Beslu P, Charquer D and Senevat J 1997 in: OR, Proceedings of the 1997 International Topical Meeting on LWR Fuel Performance, Portland, OR, 2-6 March. 405
[6] Rodchenkov B S and Semenov A N 2005 Nuclear Engineering and Design. **235** 2009
[7] Cheong Y M, Kim S S and Kim Y S 2002 *J. Nuc. Mater*. **3030** 83
[8] Cai S, Daymond M R, Khan A K, Holt R A and Oliver E C 2009 *J. Nucl. Mater*. **393** 67
[9] Froideval A, Degueldre C, Segre C U, Pouchon M A and Grolimund D 2008 *Corros. Sci*. **50** 1313
[10] Kim H G, Park S Y and Lee M H 2008 *J. Nucl. Mater*. **373** 429
[11] Kohn W and Sham L J 1965 *Phys. Rev*. **140** 1133A
[12] Hohenberg P and Kohn W 1965 *Phys. Rev*. **136** 871

[13] Milman V, Winkler B, White J A, Pickard C J, Payne M C, Akhmatskaya E V and Nobes R H 2000 *Int. J. Quantum Chem*. **77** 895
[14] Kresse G and Furthmuller J, 1996 *Comput. Mater. Sci*. **6** 15
[15] Perdew J P, Burke K and Ernzerhof M 1996 Phys. Rev. Lett. **77** 3865
[16] Parlinski K, Li Z Q and Kawazoe Y 1997 *Phys. Rev. Lett*. **78** 4063
[17] http://www.materialsdesign.com
[18] Kresse G and Furthmuller J 1996 *Phys. Rev*. **B54** 11169
[19] Monkhorst H J and Pack J D 1976 *Phys. Rev*. **B13** 5188
[20] Bellaiche L, Vanderbilt D 2000 *Phys. Rev.* **B61** 7877
[21] Wilson DJ, Winkler B, Juarez-Arellano EA, Friedrich A, Knorr K, Pickard CJ, Milman V 2008 *J. Phys. Chem. Solids* **69** 1861
[22] Abriata J P and Bolcich J C 1982 Binary Alloy Phase Diag. 2789
[23] Zhou Z and Joos B 1996 *Phys. Rev*. **B54** 3841
[24] Qiao Y, Zhang H, Hong C and Zhang X 2009 *J. Phys*. **D42** 1
[25] Wu Z J, Hao X F, Liu X J and Meng J 2007 *Phys. Rev.* **B75** 054115
[26] Cheong Y M, Kwon S C, Ahn S B, Choo K N and Kim Y S 1999 Transactions of the 15th International Conference on Structural Mechanics in Reactor Technology (SMiRT-15), Seoul, Korea, 15-20 August. 269

Measuring the effect of surface evaporation on the segregation process in a Sb/Cu system

S Cronjé[1], W D Roos[1], R E Kroon[1] and J K O Asante[2]
[1]Department of Physics, University of the Free State, Bloemfontein, South Africa
[2]Department of Physics, Tshwane University of Technology, Pretoria, South Africa

E-mail: CronjeS@ufs.ac.za

Abstract. Metallurgical products play an important role in everyday life. The search for alloys with better material properties such as strength, wear and corrosion resistance continues to this day. In addition to these desirable properties, the search for ways to reduce production costs and time has led to a large amount of research being conducted on the processes which determines the material properties of metals and alloys. One of these processes is known as segregation. To improve segregation studies the influence of surface evaporation should be considered. As experimental segregation studies are performed under high vacuum conditions, certain elements are prone to this neglected phenomenon. Although some attempts have been made to develop segregations models that take surface evaporation into account, these models can only predict segregation in either the kinetic region or when equilibrium has been reached. In addition the effect of evaporation has received little attention in previous experimental studies. In this study surface evaporation during segregation of a Sb/Cu system was measured. These results together with the modified Darken model will be used to simultaneously predict kinetic and equilibrium segregation including the effects of surface evaporation. A first approximation will be discussed as well as modifications made to an Auger system in order to measure the actual surface evaporation that takes place during segregation.

1. Introduction

Metallurgical products play an important role in everyday life. The search for alloys with better material properties such as strength, wear and corrosion resistance continues to this day. In addition to these desirable properties, the search for ways to reduce production costs and time has led to a large amount of research being conducted on the processes which determines the material properties of metals and alloys. One of these processes is known as segregation.

Segregation of one or more components of an alloy to interfaces and surfaces can influence both the chemical and physical properties of the alloy. Surface segregation is commonly regarded as the redistribution of solute atoms between the surface and the bulk of the material, resulting in a solute surface concentration that is generally higher that the solute bulk concentration. The redistribution comes about so that the total energy of the crystal is minimized.

In previous investigations a typical surface segregation study might have had the following procedure:

- Prepare binary alloys of single crystals doped with a foreign element
- Measure the concentration of the segregant on the surface using Auger Electron Spectroscopy (AES) with the method of linear temperature ramp (LTR).
- Extracting the segregation parameters by doing simulations of the experimental data.

The segregation models commonly used in the above mentioned type of study are based on the macroscopic transport equations. These are the Fick, Bragg-Williams, Guttman and the Darken models [1]. The models in one way or another describe the change in atom flux between the surface layer and the first atom beneath the surface layer as:

$$J_{10} = N_s X_1 \vartheta_1 \exp[\Delta G_s / RT] \quad (1)$$

Where N_s is the total number of available surface sites, $X_1$is the surface concentration of the segregant,ϑ_1 is the effective jump frequency, ΔG_s is the segregation energy as shown in figure 1. This equation is not taking any possible evaporation into account. A modification by Stinespring *et.al* [2] suggests a third flux of atoms from to surface layer into the vacuum as:

$$J_{10} = N_s X_1 \vartheta_1 \exp[-(\Delta G_v - \Delta G_s / RT] \quad (2)$$

where ΔG_v is the evaporation energy.

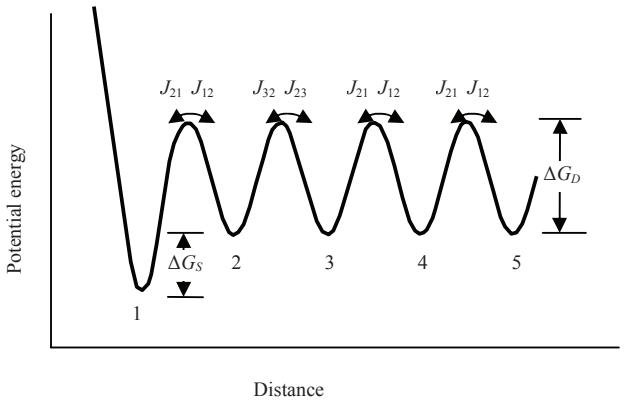


Figure 1. The potential energy diagram as a function of depth into the crystal indicating a very steep potential barrier for atoms escaping into the vacuum.

The evaporation process is incorporated into the models by ΔG_v which is illustrated in figure 2. The potential barrier between the surface layer and vacuum is thus finite, making evaporation possible. For evaporation to take place the energy barrier ΔG_v - ΔG_s has to be exceeded. Remember that the conventional thermodynamic definition of zero energy is the minimum of the internal free energy. Therefore ΔG_v will be positive and ΔG_s be negative.The Gibbs free energy necessary for this additional flux of atoms is illustrated in figure 2. Expanding the segregation model the experimental challenge is to measure this third flux of atoms leaving the surface into the vacuum under various conditions of temperature and pressure. As AES is an ultra high vacuum technique (~10^{-9} Torr) it is possible that at high temperatures (>400K) certain elements might evaporate from the surface. From literature [3] it is known that a common element used for segregation studies namely antimony (Sb) is expected to evaporate under these conditions.

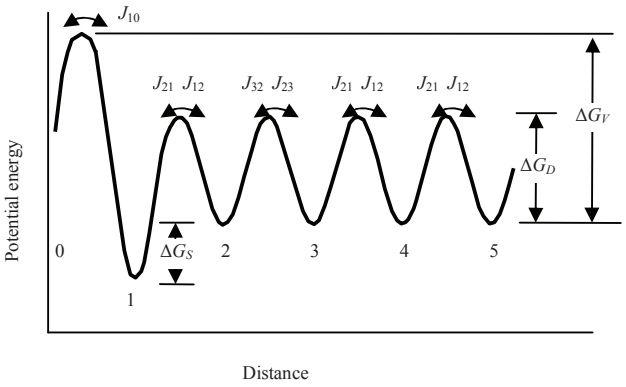


Figure 2. The potential energy diagram as a function of depth into the crystal indicating a finite potential energy barrier between the surface layer (1) and vacuum (0) allowing evaporation to take place.

By qualifying and quantifying the segregant that evaporates off a crystal surface during a segregation run, the accuracy of the segregation parameters obtained from such a measurement for the specific system can be improved upon.

2. Experimental Setup

An AES system was modified to allow the user to perform standard segregation measurements, but also to measure the evaporation rate of any material that might evaporate from the surface of a sample during a temperature ramp. This was done by employing an Inficon XTC/3s Deposition Controller which utilizes a gold coated crystal to detect any material that is deposited onto the crystal surface. This thickness monitor was mounted in the AES system in such a way that the specimen mounted on a resistance heater could be repositioned from in front of the electron gun/analyzer to a position underneath the gold coated crystal. This was achieved by making use of a manipulator. See figures 3 and 4

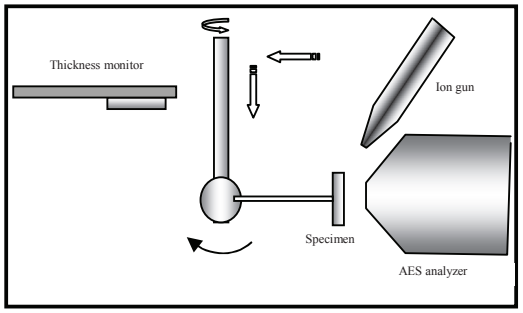


Figure 3. A schematic of the thickness monitor and AES analyzer. The arrows indicate the movements necessary to align the specimen for evaporation measurements.

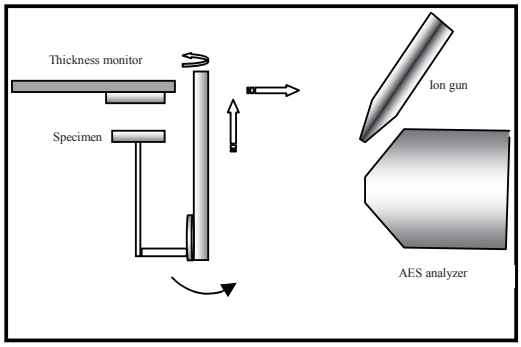


Figure 4. A schematic of the evaporation monitor and AES analyzer. The arrows indicate the movements necessary to align the specimen for AES measurements.

A Cu(110)crystal doped with Sb was used in this experiment. After performing a LTR to confirm that segregation of Sb did indeed take place, the sample was sputtered at room temperature for a few minutes using 2 keV Ar^+ ions and rastered over an area of 3 mm × 3 mm. This was done to rid the sample surface of any contaminants, as well as to remove any residual Sb. The sample was then positioned below the gold coated crystal, and a temperature ramp was performed. This ramp was performed by increasing the temperature in steps of 20K every 5 minutes from room temperature up to 830K, while simultaneously monitoring the evaporation rate of any material deposited onto the gold coated crystal of the thickness monitor. At this stage it was assumed that any material deposited onto this surface would be Sb which evaporated from the sample surface.

3. Results and discussion

The evaporation rate measured is shown in figure 5.

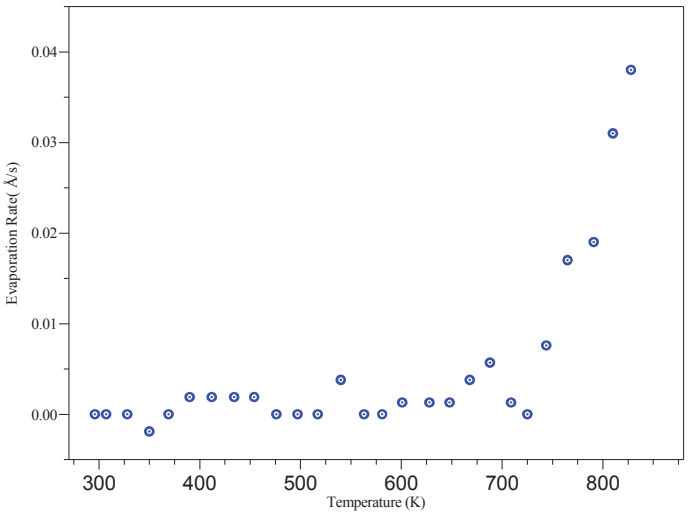


Figure 5. Change in evaporation rate as a function of temperature measured by the thickness monitor during a segregation run.

A typical exponential increase in the flux rate of atoms escaping from the surface can be seen. According to evaporation graphs [3] for pure Sb at a pressure of 10^{-8} Torr Sb sublimates at approximately 520K. The evaporation run was repeated several times indicating a start of evaporation from the Cu substrate at approximately 700K.

Because evaporation at lower temperatures is controlled by the segregation term (see equation 2) the high sublimation temperature is a result of very low surface concentrations of Sb at lower temperatures. AES segregation runs show a sufficiently high concentration of Sb only at a temperature of 700K. Adding to this, the Sb escaped from a Cu substrate suggesting stronger interaction between Sb-Cu than between Sb-Sb atoms on the surface. This positive Sb-Cu interaction was measured by Asante [4].

The gold coated crystal was then removed from the thickness monitor and transferred to a PHI Versaprobe X-Ray Photoelectron Spectroscopy (XPS) spectrometer. To identify the deposited material, XPS analyses were performed on the used and unused gold coated crystals in the flux monitor. It is clear from the XPS results shown in figure 6 and 7 that the deposited material was Sb, and that no significant evaporation of Cu took place. It can thus be concluded that the change in flux measured by the thickness monitor was due to Sb evaporating off the sample surface.

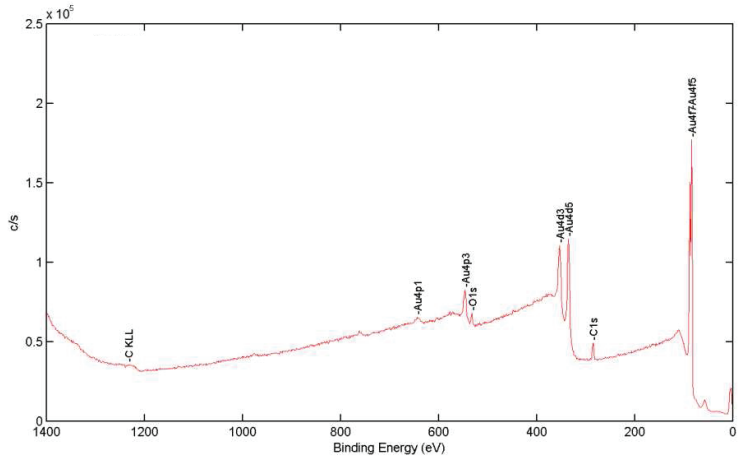


Figure 6. XPS spectrum of the gold coated crystal used in the monitor before a segregation run

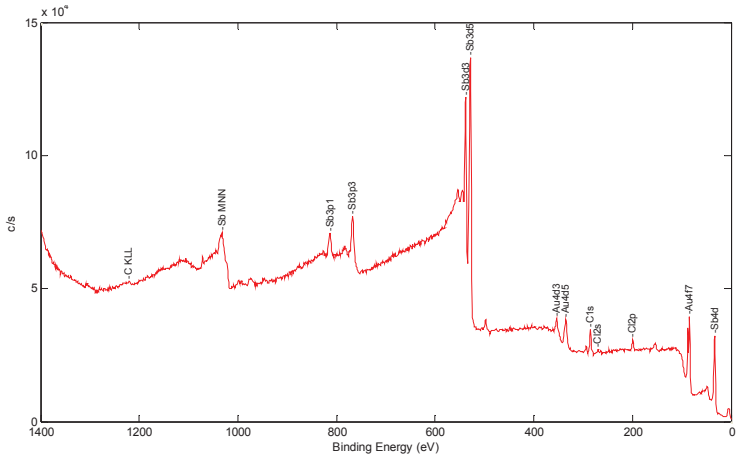


Figure 7. XPS spectrum of the gold coated crystal used in the monitor after a segregation run

4. Conclusion

An AES system was successfully modified to not only perform standard segregation experiments, but also to monitor any segregant which evaporates from the sample surface during such an experiment. This additional evaporation data may change the exiting segregation parameters obtained so far and give a more realistic model for segregation in metal substrates.

References

[1] Terblans J J 2001Modellering en Eksperimentele ondersoek van Sb-oppervlak segregasie in Cu-enkelkristalle *PhD Thesis, University of the Free State, South Africa*
[2] Stinespring C D and Lawson WF 1985 *Surf. Sci.* **150** 209
[3] Honig R E and Kramer D A 1969 *RCA Rev*. **30** 285
[4] Asante J K O 2005 Surface segregation of Sn and Sb in the low index planes of Cu *PhD Thesis, University of the Free State, South Africa*

Synthesis and photoluminescence properties of $CaSi_xO_y:Tb^{3+}$ phosphors prepared using solution-combustion method.

F. B. Dejene* and M. A. Kebede
Department of Physics, University of the Free State(Qwaqwa Campus), Private Bag X13,Phuthaditjhaba, 9866, South Africa
*Corresponding author: dejenebf@qwa.ufs.ac.za

Abstract. Effect of Tb^{3+} ion concentrations on the structural and persistence luminescence properties of $CaSi_xO_y:Tb^{3+}$ crystals were evaluated using powders grown by the solution combustion technique. The XRD study indicates the change of phase from $CaSiO_3$ to $Ca_3Si_2O_7$ as the terbium concentration increase. Broad band excitations peaking between 280 - 360 nm derived from excited states of Tb^{3+}ions were observed for all powders grown from various Tb compositions. The green emission peak at 545 nm due to an electronic transition of $^5D_4$-$^7F_5$ was found to be more prominent and intense. The persistent luminescent property was greatly influenced by the composition of the activator ions. The maximum concentration of Tb^{3+} ions that is necessary to obtain the optimal PL intensity was 1%.

1. Introduction

Phosphors are widely used in emissive displays. However, all currently used phosphors still need considerable improvement in material properties such as its lower current saturation, high efficiency, and better chromaticity [1]. Oxide phosphors (including silicate phosphor) are more chemically stable than conventional sulfide phosphors under high Coulomb loading. Metal silicates have been widely reported as promising host materials for rare earth and transition metal ions with excellent luminescence properties in the blue, green and red spectral regions [2]. Calcium silicate phosphors are ideal from the manufacturing point of view because both calcium and silica are abundant and are relatively inexpensive. The Tb^{3+} ion shows good luminescence performance in different bulk host materials and through the sol–combustion high quality powders with good crystallinity and homogeneous composition can be obtained at a relatively low temperature. These phosphor materials are usually manufactured using solid state reaction process, which include a mixing process of raw material powders, sintering process under high temperature (1800°C) and high pressure and pulverization of the sintered ceramics compacts to make micron sized phosphor powders. In recent years, the experimental methods, such as sol-gel-microwave process [3], co-precipitation [4,5], combustion reaction or microwave combustion methods [6], have been investigated to synthesize phosphors. Among these methods, solution combustion reaction is attractive because of low synthesis temperature and short reaction time. It has also been reported that other liquid processes results into amorphous calcium silicate, hence the use of solution combustion to prepare crystalline $CaSi_xO_y:Tb^{3+}$ materials. Therefore, in the present paper we report on the synthesis of $CaSi_xO_y:Tb^{3+}$ green luminescent powders via the solution–combustion process, and investigate their morphological, structural and luminescence properties as a function of the Ca:Si molar ratios and the Tb^{3+} doping concentration.

The endothermic processes during the solution-combustion process are determined by decomposition and removal of nitric oxides and significantly vary depending on the precursor composition and ratio of nitrate to urea. These physical-chemical processes results in various crystalline phases from the single host with corresponding effects on the luminescent properties.

2. Experimental

Tb^{3+}-activated calcium silicate ($CaSi_xO_y:Tb^{3+}$) nanophosphors have been synthesized by means of simple and low temperatures (500°C) using the solution-combustion of corresponding metal nitrate, TEOS and urea solution mixtures. To determine the optimum amount of silica a series of phosphors was made in which the mole ratio of cation to silica was varied. The amount of activator that could be added to these phosphors was also varied from 0 to about 4%. XRD patterns of as-synthesized samples were recorded on an x-ray diffractometer with Cu K_α = 1.5406 Å, which was operated at 40 kV voltage and 40 mA anode current. Data were collected in 2θ values from 20° to 80°. Photoluminescence (PL) measurements were performed at room temperature on a Cary Eclipse fluorescence spectrophotometer (Model: LS 55) with a built-in 150 W xenon lamp as the excitation source and a grating to select a suitable wavelength for excitation.

Results and Discussion

3. Structure

An XRD pattern of $CaSiO_3$ phosphors, shown in Fig. 1a, reveals that the whole spectrum can be indexed to a crystalline hexagonal $CaSiO_3$ phase (JCPDS 84-0655). From the XRD pattern of as-prepared $CaSi_xO_y$:Tb shown in Fig. 1c, it was found that the structure of the sample becomes $Ca_3Si_2O_7$ (JCPDS 76-0623). The XRD study indicates that new peaks appear as the terbium concentration increase results in phase changes from $CaSiO_3$ to $Ca_3Si_2O_7$. The later phase may be favored as a result of some Si sites substituted by the Tb ions, as the concentration of Tb was changed from 1 to 4 mole %. Due to the differences in size of the two atoms this substitution results in the shrinkage of $CaSiO_3$ lattice therefore causing phase transitions. On the other hand, the substitution of Tb^{3+} ions were supposed to enhance the hole trapping ability of oxygen vacancies that increases the persistent luminescent property.

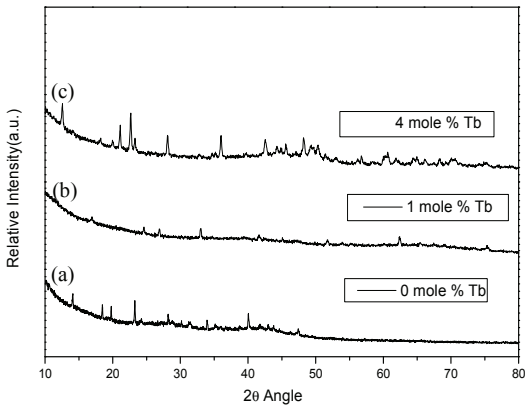
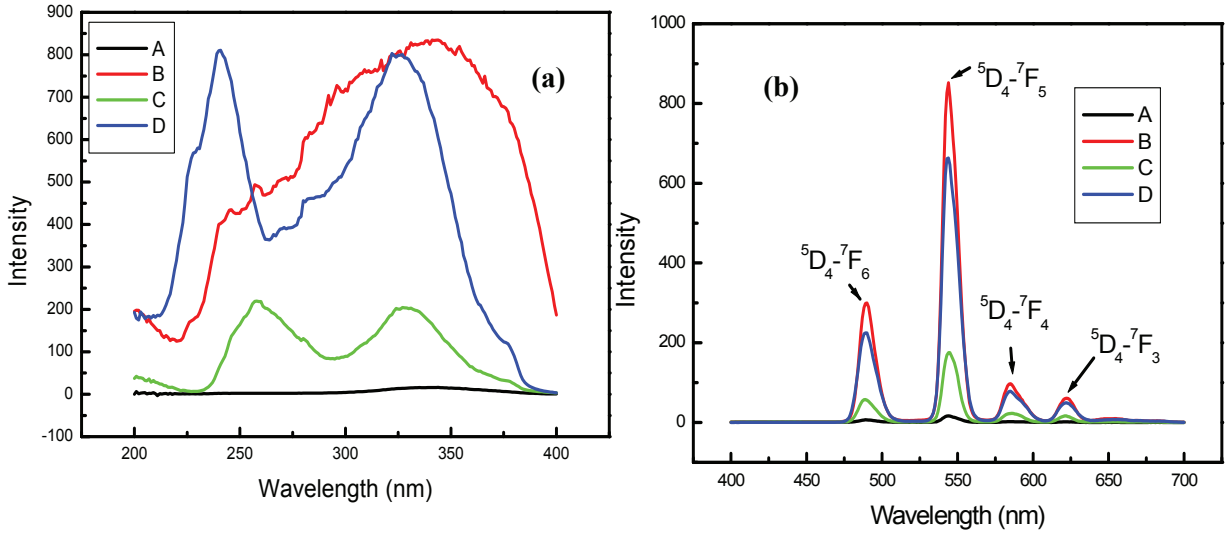


Figure 1: The XRD pattern for $CaSi_xO_y:Tb^{3+}$ 1(a) Tb=0, 1(b) Tb=1%, 1(c) Tb=4%

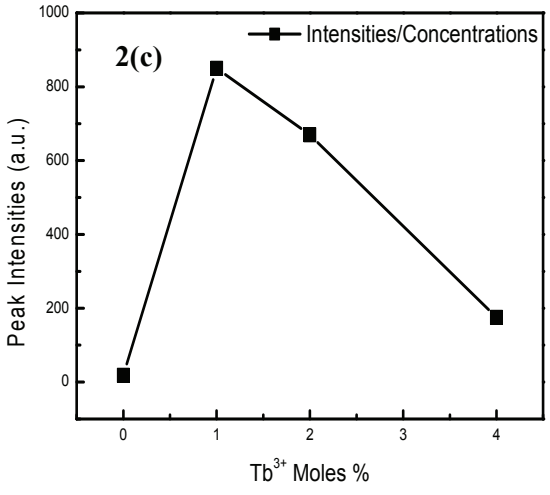
3.1. Photoluminescence

In Fig. 2(a), the excitation spectra of the samples monitored at 545 nm show two bands at about 240 nm and 280–360 nm. It is well known that there should be a strong absorption peak at 244 nm from spin-allowed transition of $^7F \rightarrow {}^5D$ of Tb^{3+}. The intensity and slight shift in the 240 nm peak presumably resulted from the randomicity of intense combustion process, and the strong influence of

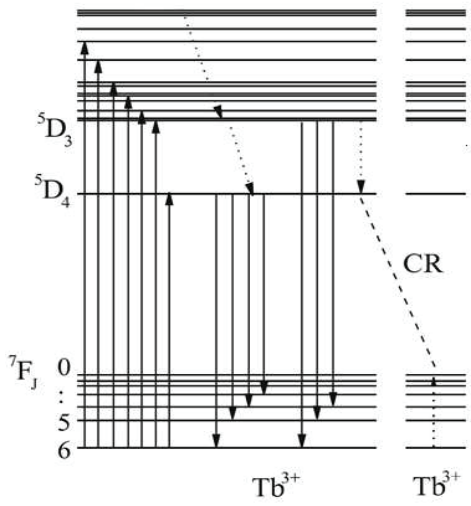
f→d transition of Tb^{3+} when subjected to two kind of host lattice and incompletely transform from Tb^{4+} to Tb^{3+} owing to lack of reductive atmosphere. The group of weak peaks covering 280–360 nm belongs to *f→f* transition of Tb^{3+} in the two host matrix of $CaSiO_3$ and $Ca_3Si_2O_7$. The persistent luminescence emission spectra of the $CaSi_xO_y$:Tb phosphors grown with various Tb^{3+} ions doping are shown in Fig. 2(b) that were obtained by exciting the samples at the wavelength 257 nm, and exhibit emission transitions of ($^5D_4 \rightarrow {}^7F_{3,4,5,6}$). Particularly, the green emission peak at 545nm has been found to be more prominent and intense which is due to an electronic transition of $^5D_4$-$^7F_5$ [7-9]. Such strong green emission displays that the phosphor powder can find applications in the development of coated screens in certain electronic systems.



Fig**ure** 2(a) Excitation and (b) emission spectra of $CaSi_xO_y$:Tb^{3+} A Tb=0, B Tb=1%, C Tb=4% and D Tb=2%.



Fig**ure** 2(c) Graph of Tb^{3+} ions concentration versus maximum peak intensity



Fig**ure** 2(d) Energy level scheme of Tb^{3+} ions in a crystal showing cross relaxation process [10]

The position of the emission peak in the emission spectra almost shows no change, regardless of the concentrations of the Tb^{3+} ions doping, implying that no obvious change occurs in the crystal field effect on 5d electron states of Tb^{3+} ions, even though the molar ratio of Tb^{3+} ions doping varies. The PL spectrum of $CaSiO_3$ phosphors prior to Tb doping is shown in Fig. 2b, which has a band around 540nm in a green region. The Figure depicts the PL spectrum of $CaSiO_3$:Tb^{3+} (1%Tb) which is the highest intensity than the other Terbium concentrations. The effects of Tb^{3+} ions concentrations on the maximum peak intensities are clear in Figure 2(c). The graph of maximum PL intensity of $CaSi_xO_y$ as a function of the Tb^{3+} concentrations is shown in Fig. 2(c). An increase in Tb^{3+} concentrations up to 1% Tb^{3+} ions increases the luminescence intensity possibly due to an increase in activator centres. Further increase in Tb^{3+} ions above 1% result in luminescence quenching caused by cross relaxation (CR) processes see Fig. 2(d). The CR involves a process in which an electron in an excited state undergoes a nonradiative transition to an intermediate excited state accompanied by an energy-conserving transition of a neighbouring ion from its ground state to some excited state. The CR process depends on ion-ion separations and can occur by multipole –multipole or by exchange interactions. At more dilute concentrations typical of phosphors, the short range exchange-mediated transfer rate will have decreased to the point that the dipole-dipole (or other multipole-multipole) interactions, which fall off more slowly as R^6, will dominate. This idea has been known for a long time, and serves as one of the main reason for the main mechanisms for concentration quenching. Persistent luminescence curves of the phosphor powders were shown in Fig. 3. It can be seen from the curves that the powders show differences in initial intensity and medium persistence when the powders were efficiently activated by fluorescent lamp. The results indicate that the initial persistent luminescence intensity and the decay time of phosphors are enhanced with increase in Tb^{3+} concentrations. The sample with Tb =1% significantly enhanced the initial persistent luminescence property of the green luminescence.

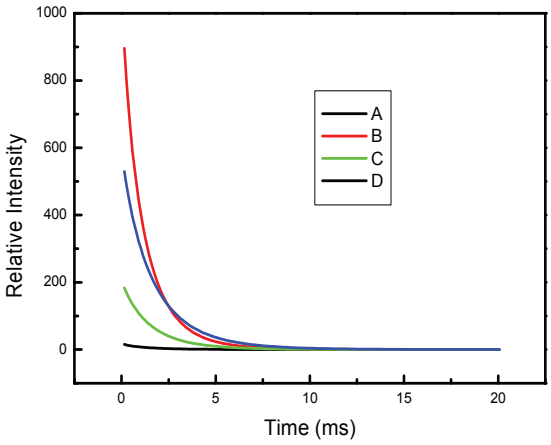


Figure 3 Afterglow characteristics of $CaSi_xO_y$:Tb^{3+} A Tb = 0, B Tb = 1%, C Tb = 2 % and D Tb = 4%

Tb moles %	0	1	2	4
Component	Decay constants (ms)			
Fast (τ_1)	0.89	0.31	0.27	0.99
Medium (τ_2)	2.21	1.03	1.30	0.99
Slow (τ_3)	1.38	1.93	2.36	2.18

Table 1 Results for fitted decay curves of the phosphor powders with different Tb^{3+} moles

The decay behaviour can be analysed by curve fitting [11,12], relying on the following triple exponential equation:

$$I = I_1 \exp(\frac{t}{\tau_1}) + I_2 \exp(\frac{t}{\tau_2}) + I_3 \exp(\frac{t}{\tau_3}) \qquad \textbf{(1)}$$

where I represents the phosphorescent intensity; I_1, I_2 and I_3 are constants; t is the time; τ_1, τ_2 and τ_3 are the decay constants, deciding the decay rate for the rapid, medium and slow exponentially decay components, respectively. The results of the parameters τ_1, τ_2 and τ_3 for the fitted decay curves are listed in Table 1. Three components namely slow, medium, and fast component are responsible for the persistent luminescence from the synthesized phosphor. A trend can be observed that the decay constants of the phosphors decrease gradually with the increasing of Tb^{3+} ions doping to $CaSi_xO_y$:Tb phosphors due to reduced hole trapping ability of oxygen vacancies.

4. Conclusion

The $CaSi_xO_y:Tb^{3+}$ phosphor was synthesized by solution combined process. Different crystalline structures were obtained depending on Ca:Si mole ratios without further calcinations after solution combustion process. The luminescent properties of the phosphors can be greatly modified by the concentration of Tb^{3+} ions although the peak position observed doesn't change with amount of Tb^{3+} ions doping. The phosphors doped with 1% of Tb^{3+} result in a very intense and bright initial phosphorescent emission. The persistent luminescence decay time is found to be influenced greatly by the terbium composition of the phosphor.

Acknowledgement
The authors send gratitude to the National Research Foundation (NRF) for funding the project and the University of Free State Physics (Physics department, Center for microscopy and Geology department) for the research techniques used in this study.

References

[1] Yocom P N, 1996 *J. of SID* **4**, 149
[2] Harrison D E and Hoffman V J 1959 *Electrochem. Soc.*, **106**, 800
[3] Zhang W S and Jinzong S Y 2007 *Mater. Chem. Phys.*, **102**(1): 80
[4] Chang C K, Yuan Z X and Mao D L 2006 *J. Alloys Compd.*, **415**(1-2): 220
[5] Gabelkov S V, Tarasov R V, Poltavtsev N S, Kurilo Y P, Starolat M P, Andrievskaya N F, Mironova A G, Ledovskaya E G, Litvinenko L M and Belkin F V 2007 *Inorg. Mater.*, **43**(3): 398
[6] Hanyan Du, Gengshen Li and Sun Jiayu 2007 *Journal of Rare Earths*, **25**(1): 19
[7] Yan B and Xiao X Z 2006 *Opt. Mater.* **28** 498
[8] Amekura H, Eckau A, Carius R and Buchal C 1998 *J. Appl. Phys.* **84** 3867
[9] Schaik W V, Lizzo S, Smit W and Blasse G 1993 *J. Electrochem. Soc.* **140** 216
[10] Guo H 2006 *Appl. Phys. B* **84**, 365–369
[11] Sakai R, Katsumata T, Komuro S and Morikawa T, 1999 *J. Lumin.* **85** (1–3) 149
[12] Yuan Z X, Chang C K, Mao D L and Ying W J, 2004 *J. Alloys Compd.* **377** 268

X-ray Photoelectron Spectroscopy analysis and characterization of $Gd_2O_2S:Tb^{3+}$ phosphor powder

J.J. Dolo*, F.B. Dejene, J.J Terblans, O.M. Ntwaeaborwa and H.C. Swart*

Physics Department, University of the Free State, P.O. Box 339, Bloemfontein, 9300, South Africa

*Corresponding author: dolojj@qwa.ufs.ac.za or SwartHC@ufs.ac.za

Abstract. This paper presents the X-ray Photoelectron Spectroscopy (XPS) analyses of electron beam degraded and undegraded $Gd_2O_2S:Tb^{3+}$ phosphor powder. The XPS data was collected from $Gd_2O_2S:Tb^{3+}$ phosphor powders before and after electron beam degradation. The data confirms the presence of Gd_2O_3 and Gd_2S_3 from both the degraded and undegraded powders. In addition, S-O bonding was also detected in degraded powders. This clearly indicates that a surface reaction occurred during prolonged electron bombardment in an oxygen atmosphere.

1. Introduction

Terbium doped gadolinium oxysulfide ($Gd_2O_2S:Tb^{3+}$), one of the rare earth oxysulfide group of phosphors, is known to be an efficient phosphor and has been put to practical application in low voltage cathodoluminescent and X-ray devices because of its high conversion efficiency (12–25%) of the exciting radiation [1-3]. $Gd_2O_2S:Tb^{3+}$ is a well-known green-emitting photoluminescence and cathodoluminescence phosphor used in high resolution and projection television screens [4-7]. We report the XPS characterization of commercial terbium doped gadolinium oxysulfide ($Gd_2O_2S:Tb^{3+}$) green phosphor, which was evaluated for possible application in cathode ray tube CRT and field emission display (FED) screens.

2. Characterization

Auger electron spectroscopy (AES) and Cathodoluminescence (CL) spectroscopy were used to monitor changes on the surface and the CL properties. The AES measurements were taken in an UHV chamber using a PHI model 549 Auger spectrometer. The chamber was first evacuated to $2.8x10^{-9}$ Torr before backfilling with oxygen to $1x10^{-6}$ Torr. Scanning Electron Microscopy (SEM) images were taken with a Gemini LEO 1525 Model to determine the particle morphology. The crystalline structure of the phosphor powders were investigated using a Burker D8 (Burker Co, German) X-ray diffractometer with Cu Kα = 1.5406 Å. The 100 μm, 25 W, 15 kV energy X-ray beam was used to analyze the S 2p, O 1s, Gd 3d and Gd 4d binding energy peaks (pass energy 11.8 eV, analyser resolution ≤ 0.5 eV). The possible chemical states were identified with the Multipak version of 8.2c computer software [8] using Gaussian-Lorentz fits.

3. Results and Discussion

Figure 1 shows the XRD pattern of the $Gd_2O_2S:Tb^{3+}$ phosphor powder. The position and relative intensity of the XRD lines are in good agreement with the data of JSPDS file No.26-1422, which shows the pure Gd_2O_2S hexagonal structure.

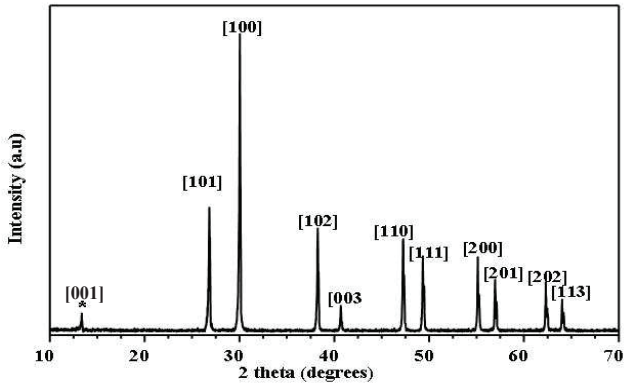


Figure 1: XRD spectrum for the $Gd_2O_2S:Tb^{3+}$ phosphor powder.

Figure 2 (a) shows the SEM image of the $Gd_2O_2S:Tb^{3+}$ phosphor powder. The particles are polyhedron in shape and agglomerated, showing relatively good close packing which is one requirement for the CRT or X-ray intensifying screens. The particles also differ in size and shape. EDS data in figure 2 (b) confirms the presence of all the elements (Gd, O and S) together with the adventitious carbon. Tb^{3+} ions were not detected, probably due to their relatively low concentration in the $Gd_2O_2S:Tb^{3+}$ matrix.

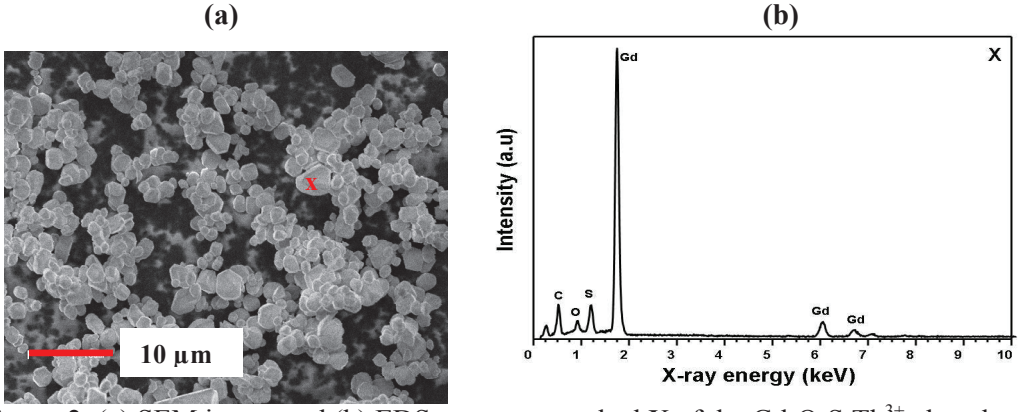


Figure 2: (a) SEM image and (b) EDS spectrum marked X of the $Gd_2O_2S:Tb^{3+}$ phosphor powder.

Figure 3 (a) shows the CL spectra before degradation and after degradation as function of wavelength. The CL decreased due to the formation of a new non-luminescent surface oxide layer. The main emission peak due to the $^5D_4 \rightarrow {}^7F_5$ transition is at a wavelength of 545 nm. Less intense emission peaks at 490 nm, 585 nm and 620 nm due to the $^5D_4 \rightarrow {}^7F_J$ (J = 0, 1, 2, 3,…) transitions are also shown. The main emission peak at 545 nm was only about 45% of the initial intensity after degradation. Figure 3 (b) shows the PL spectra of $Gd_2O_2S:Tb^{3+}$ powder phosphors excited at 254 nm. The luminescence peaks in the figure arise from the transitions of the $^5D_4$ excited state levels to ^7F_J (J = 0, 1, 2, 3, 4, …) ground state levels, and belong to the characteristic emission of Tb^{3+}. The emission

line at 490 nm corresponds to the $^5D_3 \rightarrow {}^7F_6$ transitions, and the emission lines between the 585 and 620 nm corresponds to the $^5D_4 \rightarrow {}^7F_4$ and $^5D_4 \rightarrow {}^7F_3$ transitions respectively. The peak at 545 nm arising from the $^5D_4 \rightarrow {}^7F_5$ transition has the highest intensity.

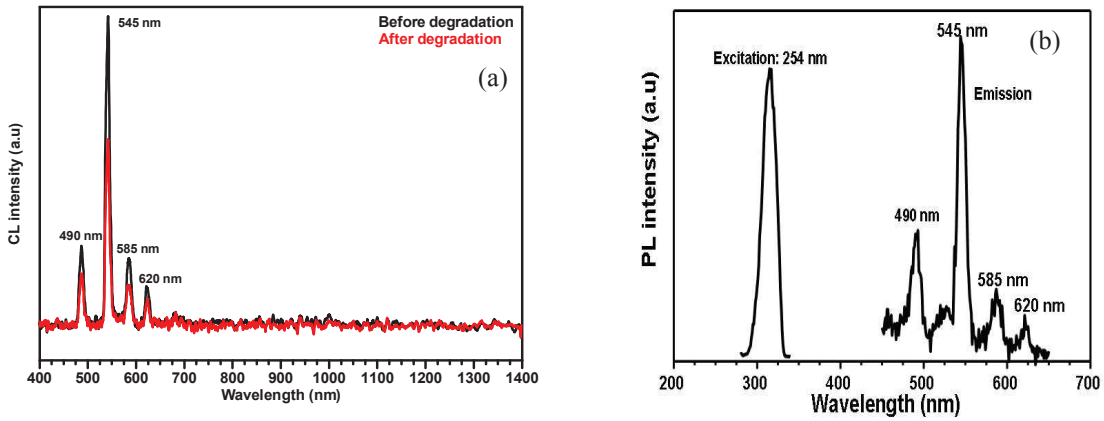


Figure 3: (a) CL before and after degradation 1x10^{-6} Torr and (b) PL spectra for $Gd_2O_2S:Tb^{3+}$ phosphor powder.

Figure 4 shows fitted high resolution S 2p XPS spectra (a) before and (b) after degradation at 1x10^{-6} Torr O_2. Note that it is well known that the S 2p consists of $2p_{3/2}$ (BE = ~ 163 eV) and $2p_{1/2}$ (BE = ~ 165 eV) peaks. XPS spectrum in figure 4 (a) indicates that sulfur was primarily present as Gd_2O_2S (BE = ~ 165 eV) plus a small amount of sulfide species (BE = ~ 158 eV). The peak at BE = ~ 158 eV can be assigned to Gd_2S_3 [9]. In addition, the fitted data shows an evidence of oxides species (SO_2) at BE = ~ 168 eV. The degraded spectrum in figure 4 (b) shows an increase in the Gd_2S_3 and SO_2 peak intensity suggesting that an electron beam induced surface chemical reaction occurred between S and Gd, and S and O.

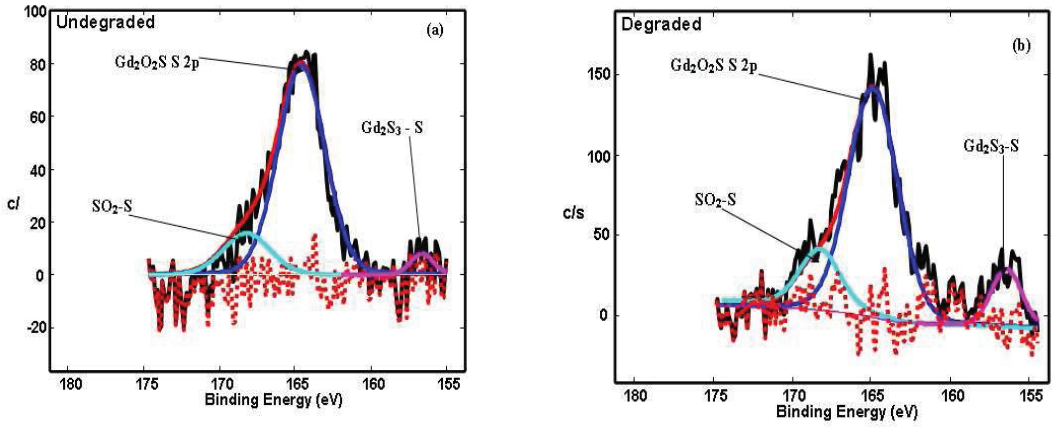


Figure 4: Fitted XPS for S 2p of $Gd_2O_2S:Tb^{3+}$ phosphor powder peaks (a) before and (b) after degradation for 1x10^{-6} Torr.

Figure 5 shows fitted XPS for O1s peaks (a) before and (b) after degradation for 1x10^{-6} Torr obtained at O 1s at 531.3 eV. It can be noted that both the degraded and undegraded powder spots only two binding energy peaks were identified. The Gd_2O_3 peak, however, shows an increase in intensity for the degraded sample. The growth in Gd_2O_3 after degradation is due to the oxide formation on the

surface as a result of the (Electron Stimulated Surface Chemical Reaction) ESSCR process as observed in the (Auger Peak-to-Peak Height) APPH results [9]. It is therefore clear that a chemical reaction occurred during the degradation process. The binding energy assignment of the mixed oxides is based on a large extent to the core level screening that occurs in O^{2-} anion compared to the other oxygen species that are present on the surface and the relative magnitude of this peak compared to other O 1s peaks present [10].

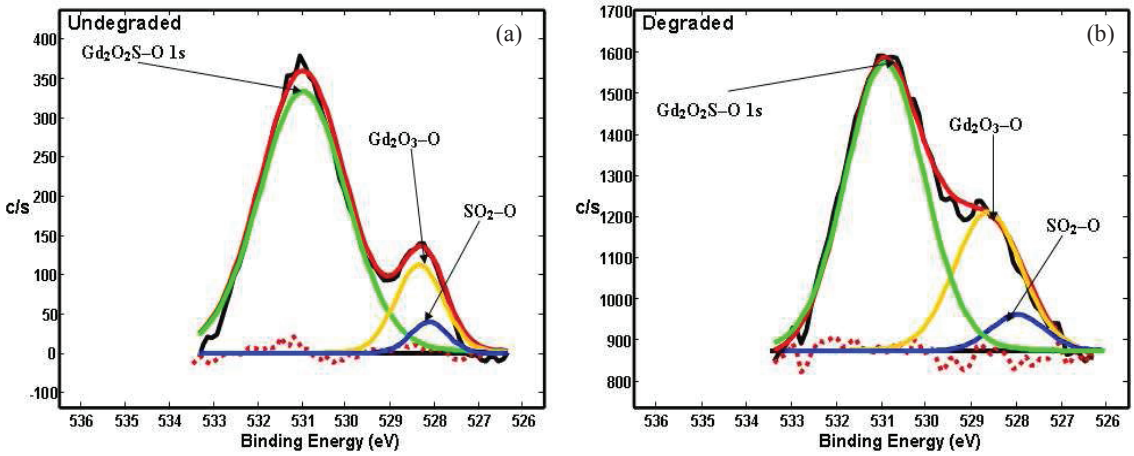


Figure 5: Fitted XPS for O 1s of $Gd_2O_2S:Tb^{3+}$ phosphor powder peaks (a) before and (b) after degradation for $1x10^{-6}$ Torr.

Figure 6 shows the fitted results from an XPS spectrum of the Gd 3d peaks (a) before and (b) after degradation. The peak shape changed due to an extra peaks of Gd_2O_3 (1189.0 eV) and Gd_2S_3 (1192.2 eV) that developed at higher binding energies. Gd_2O_2S-Gd 3d peaks were measured at 1185.20 eV. Both the degraded powder and undegraded spots recorded one peak of Gd_2O_2S-Gd 3d. The peaks for Gd_2O_3 and Gd_2S_3 after degradation have grown and this clearly shows that the surface reaction did occur after degradation.

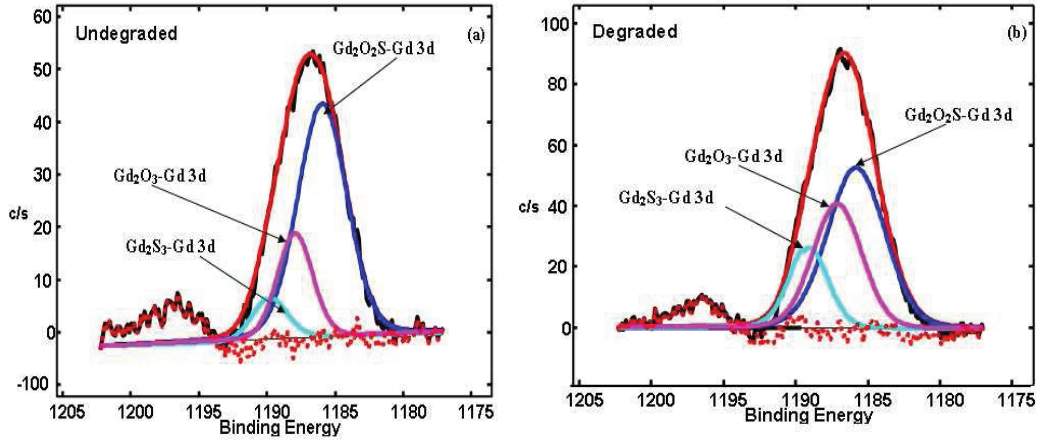


Figure 6: Fitted XPS for Gd 3d of $Gd_2O_2S:Tb^{3+}$ phosphor powder peaks (a) before and (b) after degradation for $1x10^{-6}$ Torr.

Figure 7 shows the Gd 4d peaks position for the Gd_2O_2S (Gd_2O_2S-4d peaks) (a) before and (b) after degradations. Six peaks can be identified for the Gd 4d core level spectrum after deconvolution of the experimentally measured curve. Gd $4d_{3/2}$ and Gd $4d_{5/2}$ peaks of the Gd_2O_2S are located at 146.7 eV and 141.6 eV. In addition, there are two peaks measured at 142.2 eV and 147.3 eV that can be associated with Gd_2O_3 and the small peaks measured at 147.9 eV and 144.8 eV can be associated with Gd_2S_3. There is also an increase in relative ratio of the peaks, this suggests that a surface chemical reaction occurred and another possibility for the presence of the peaks (Gd_2O_3 and Gd_2S_3) could be chemical decomposition of the material [11].

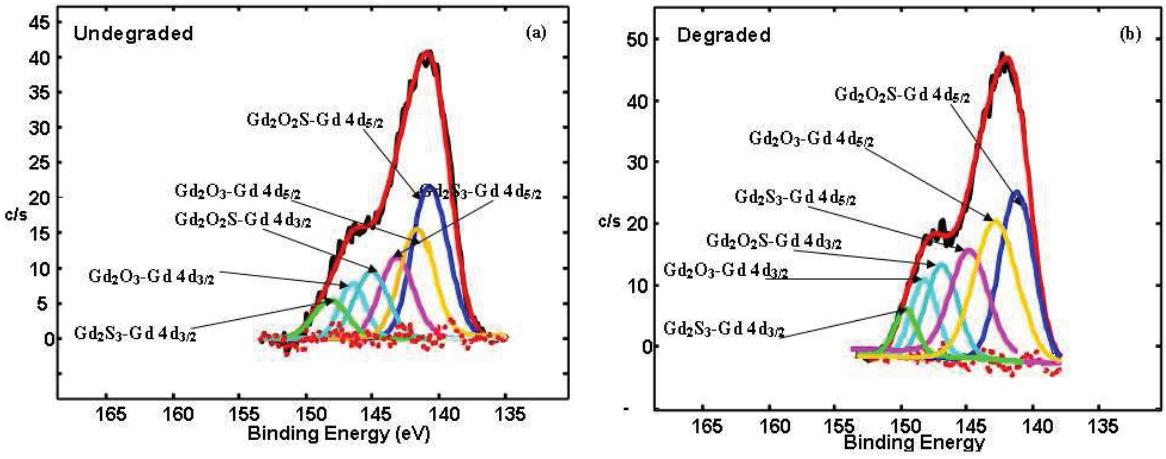


Figure 7: Fitted XPS for Gd 4d of $Gd_2O_2S:Tb^{3+}$ phosphor powder peaks (a) before and (b) after degradation for $1x10^{-6}$ Torr.

4. Conclusion

The XPS results confirmed the presence of Gd_2O_3 and Gd_2S_3 on the degraded $Gd_2O_2S:Tb^{3+}$ powder spots. The XRD pattern of $Gd_2O_2S:Tb^{3+}$ powder shows hexagonal phase structure. The EDS confirms the presence of all elements of the host matrix (Gd_2O_2S) as well as the adventitious carbon from the surface. The PL properties were also investigated. The SEM confirms that the particles are polyhedron in shape and agglomerated.

Acknowledgements

The authors send gratitude to the National Research Foundation (NRF) for funding the project and the University of Free State (Physics department, Center for Microscopy and Geology department) for the research techniques used in this study.

References

[1] Royce MR, 1968 US Patent 3 **418** 246
[2] Royce MR, Smith AL, 1968 Electrochem. Soc., Spring Meeting Extended Abstracts, **34** 94
[3] Giakoumakis GE, Pallis A, 1989 J., Solid State Comm. **70** 419
[4] Blasse G, Grabmaier BC, 1994 Luminescence Material, Springer-Verlag, New York
[5] Popovici EJP, Muresan L, Hristea-Simoc A, Indrea E, Vasilescu M, Nazarov M, Jeon DY, 2002 Opt. Mat., **29** 632
[6] Jiang Y, Wu Y, 2000 J. Am. Ceram. Soc., **83** 2628
[7] Yamamoto H, Kano T, 1979 J. Electrochem. Soc., **126** 305
[8] Moulder F, Stickle WF, Sobol PE, Bombe KD, 1995 Handbook of X-ray Photoelectron Spectroscopy, ULVAC-PHI, Inc., 370 Enzo, Chigasaki, **143** Japan

[9] Dolo JJ, Swart HC, Coetsee E, Terblans JJ,
Ntwaeaborwa OM and Dejene BF, 2010 Hyperfine Inter., **197** 129
[10] Hashimoto AK, X-ray photoelectron spectra of several oxides, **1977** Corros. Sci., 17 559
[11] Datta P Majewski, F Aldinger, 2009 Mat. Charac., **60** 138

Blended thin films of tris-(8-hydroxyquinoline) aluminium (Alq_3) embedded in polymethyl methacrylate (PMMA).

M M Duvenhage, O M Ntwaeaborwa and H C Swart[1]

Physics Department, University of the Free State, P.O. Box 339, Bloemfontein, 9300, South Africa

Email: swarthc@ufs.ac.za

Abstract. Alq_3 is widely used in organic light emitting diodes (OLEDs) as emission and electron transport layer. During the fabrication of OLEDs the current trend is to use solution-processing; however, Alq_3 must be vacuum deposited. One possible way to overcome this problem is to use different polymers containing the Alq_3. In this study, photoluminescent (PL) properties of blended Alq_3:PMMA thin films with different molar concentration of Alq_3 ranging from 0.5% to 5% were investigated. The PL data were record when the films were exposed to 355 nm UV light. PL emission was characterized by broad spectra with maximum at ~ 515 nm corresponding to the well known green photoluminescence from Alq_3 in the solid state. As observed from the scanning electron microscopy images, the films were smooth with numerous voids caused by evaporating gases while the atomic force microscopy images showed that there were outcrops suggesting that the films were not absolutely smooth..

1. Introduction

Since its first report in 1987 by Tang and Van Slyke [1], Alq_3 has been used as light-emitting and electron transporting layer in organic light emitting diodes (OLEDs). The current trend during the fabrication of OLEDs is to use solution-processing. However, with this technique Alq_3 must be vacuum deposited [2]. A possible solution to this problem is to dope different polymers with Alq_3. These polymers should contain the optical properties of Alq_3 while the processability of a polymer is maintained. This will allow for low cost manufacturing techniques such as ink-jet printing and solution processing [3]. In this study polymethyl methacrylate (PMMA) is doped with different moles of Alq_3 ranging from 0.5 - 5 mol %. The effects of Alq_3 concentration on the photoluminescence (PL) intensity and emission wavelength are investigated. The surface morphology and structure of the blended films were studied to evaluate their effects on the PL intensity.

2. Experimental

2.1 Synthesis

Alq_3 phosphor powder was synthesized using a co-precipitation method explained in ref. [4] by adding 0.625g of 8-hydroxyquinoline (8-Hq) to a mixture of 6.5 ml H_2O and 6.5 ml of glacial acetic

[1] To whom any correspondence should be addressed.

acid. The mixture was stirred for 15 minutes. 0.5 g $Al(NO_3)_3$ was added to 20 ml H_2O and this mixture was also stirred for 15 minutes. The $Al(NO_3)_3$ solution was added drop-wise to the 8-Hq solution with vigorous stirring. The resulting brown mixture was stirred for further 15 minutes. 5 ml of NH_4OH was added drop-wise to the mixture while stirring. A yellow green precipitate was formed. The precipitate was filtered and washed 8 times with distilled water was left to dry overnight in an oven maintained at 80°C. After drying the precipitate was ground into a fine powder. The powder was dissolved in 10 ml acetone and left to recrystalize in air at room temperature.

A blended thin film of Alq_3:PMMA at a concentration ratio of 1% Alq_3 was obtained in the following way: 2g of PMMA was dissolved in 25 ml of chloroform. A solution of 0.026 g Alq_3 in 15 ml of chloroform was added to the PMMA solution. The blend was stirred and poured into a stainless steel dish and was left to dry in air for 24 hours. The films were then cut into 1x1 cm pieces for further characterization. The blended thin films of Alq3:PMMA at concentration ratios of 0.5% - 5% were obtained in a similar way.

2.2 Characterization

The structure and morphology of the samples were determined by scanning electron spectroscopy (SEM) using a Shimadzu SSX550 and atomic force microscopy (AFM) by using the Shimadzu SPM – 9600 system. The PL emission data was collected with a Cary Eclipse fluorescence spectrophotometer equipped with a Xenon flash lamp.

3. Results and discussion

Fig. 1 shows the SEM images of the Alq_3:PMMA (Alq_3 = 1 mol%) blended thin film at x1000 and x5400 magnification. The surface was smooth with small holes scattered all over the film. These holes were formed when gasses escaped during evaporation of chloroform (the solvent). At higher magnifications it can be seen that the holes are not all the same size. Bigger holes were probably formed early in the solidification process due to rapid rate of evaporation of the solvent while small holes were probably formed toward the end when the was less solvent and the rate of evaporation was slow.

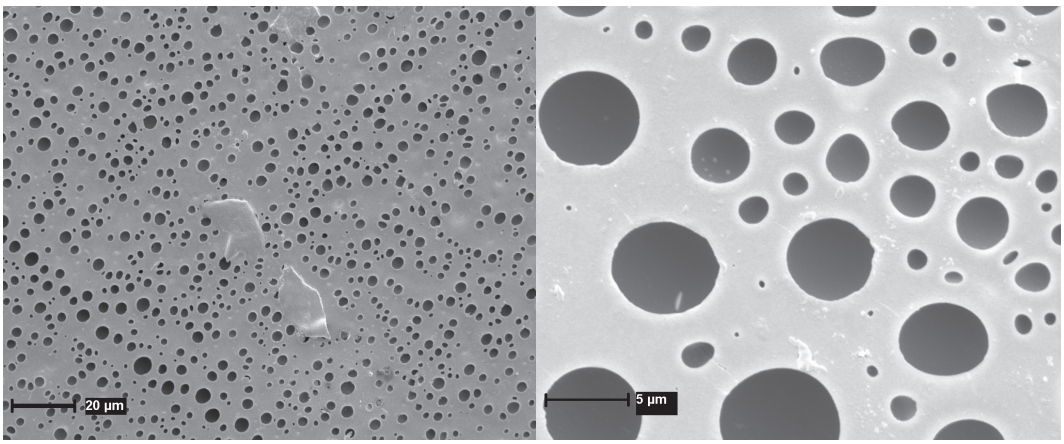


Fig. 1: SEM images of 1% Alq_3:PMMA at x1000 and x5400 magnifications. Holes that had formed during the evaporation of the solvent can clearly be seen.

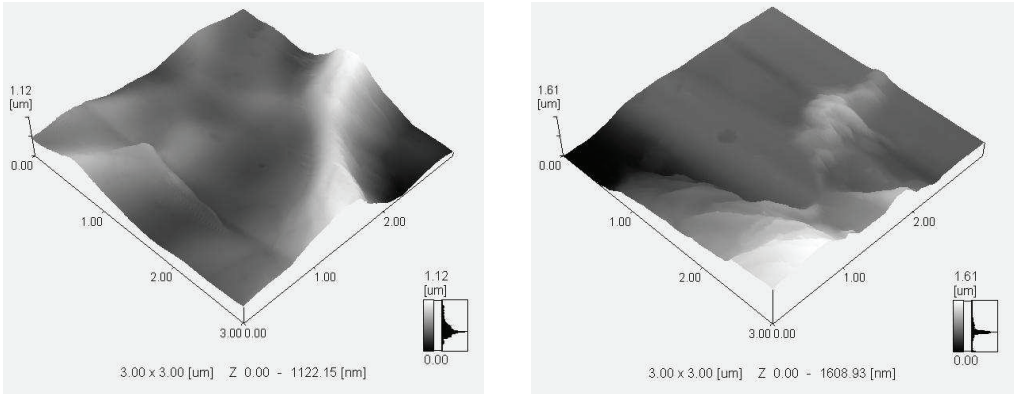


Fig. 2: AFM images of two different spots on the Alq_3:PMMA (Alq_3 = 1 mol%) film.

Fig. 2 shows the AFM images of two different spots on the Alq_3:PMMA film. From the images, it can be seen that the surface of the film is not as smooth as the SEM image suggests. Outcrops can be seen ranging between 1-2 µm in height. No distinct particles can be seen confirming that the Alq_3 powder was completely mixed with the PMMA

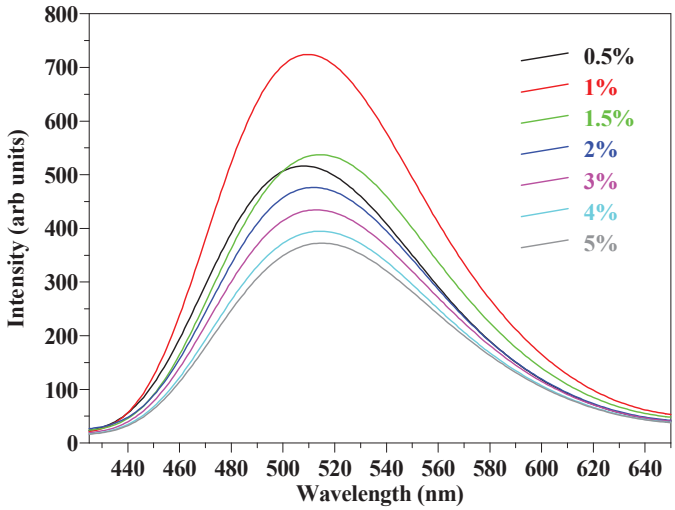


Fig. 3: Photonluminescent spectra of different concentrations of Alq_3 in a PMMA matrix, just after synthesis.

Fig. 3 shows the PL spectra of different concentrations of Alq_3 in a PMMA matrix just after preparation of the films. The films were excited at 355 nm. All the peaks show a broad spectrum with a maximum intensity at around 515 nm (figure 4(a)). This corresponds to the emission of Alq_3 in the solid state [5]. This emission is associated with the $S_1 \rightarrow S_0$ transition localized on the a-quinolate ligand [11]

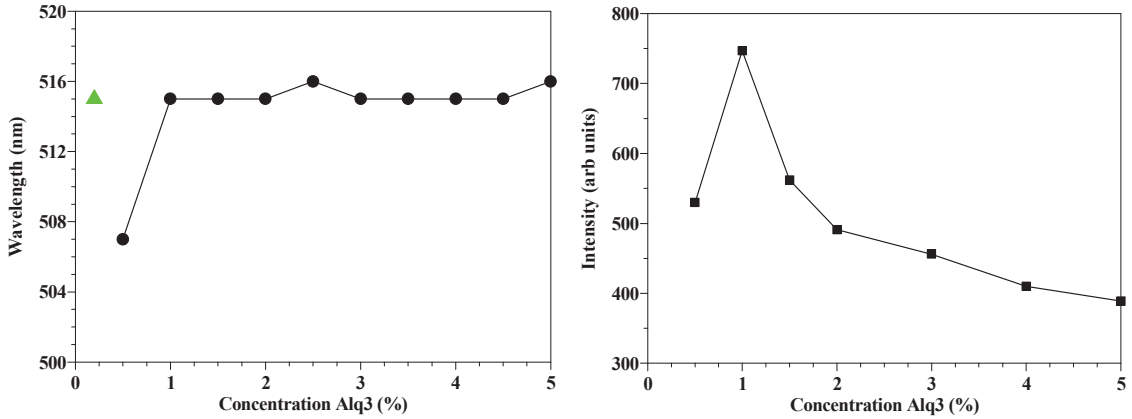


Fig. 4: (a) A plot of the emission wavelength at maximum intensity as function of the Alq_3 concentration. The triangle represents the emission of Alq_3 in the solid state. (b) Plot of PL maximum intensity as a function of concentration of Alq_3.

The peak of the 0.5 mol% sample is blue-shift by 8 nm (Fig. 4(a)). Previous reports by Cölle et. al. [6] and Levichkova et.al. [7] attributed the blue-shift to either the formation of the δ-phase of Alq_3 or the transformation of the meridional form of Alq_3 to the facial form while Shukla and Kumar [12] attributed the blue shift in Alq_3 thin films to the change from the 3D to 2D excitonic states with decreasing film thickness. Although the film thicknesses were not evaluated, it is also possible that the blue-shift of the Alq_3 film in this study was due the change in excitonic states. The emission of solid state Alq_3 is due to the relaxation of an excited electron from the S_1-S_0 level. A lot of S_n levels exist above the S_1 level. At low concentrations the emission is most probable due to relaxation from the S_n-S_0 level causing a blue shift Fig. 4(b) shows that the sample doped with 1 mol% of Alq_3 has the highest intensity. The intensity then decreases with an increase in concentration. This same decrease in intensity was reported by Meyers and Weck [2]. The decrease (or quenching) of luminescence at higher concentrations is due to concentration quenching effects. This is the result of clustering or cross relaxation [13].

4. Conclusion

Blended thin films of Alq_3:PMMA were successfully synthesized. The morphology and structure measurements of the films showed that the surface had a lot of holes. The AFM images showed that the surface of the films was not smooth, but rather rough with a lot of "hills and valleys". This rough surface contributed to the high luminescence intensity that was observed due to less internal reflection that occurred within the film. A broad emission band was observed, with its maximum at 515 nm when excited with 355 nm photons. This is consistent with emissions of Alq_3 in the solid state. A blue shift in the emission was observed for the 0.5 mol % sample and it was attributed to change in excitonic states. The 1 mol % sample showed the highest intensity and the intensity then decreased with an increase in doping concentration due to concentration quenching effects.

Acknowledgments

The authors wish to thank the National Research Foundation (NRF) and the University of the Free State for financial support of this work. Prof P.W. van Wyk from the CCEM for help with the SEM images.

References

[1] Tang C W and Vanslyke S A 1987 *Appl. Phys. Lett.* **51** 913
[2] Meyers A and Weck M 2003 *Macromolecules* **36** 1766
[3] Friend R H, Gymer R W, Holmes A B, Burroughes, J H, Marks R N, Taliani C, Bradley D D C, Dos Santos D A, Brédas J L, Lögdlund M and Salaneck W R 1999 *Nature* **397** 121
[4] Mahakhode J G, Bahirwar B M, Dhoble S J and S.V. Moharil 2006 *Proc. Of ASID.* 237
[5] Shukla V K, Kumar S and Deva D 2006 *Synth. Met.* **156** 387
[6] Colle M, Gmeiner J, Milius W, Hillebrecht H and Brutting W 2003 *Adv. Func. Mat.* **13** 108
[7] Levichkova M M, Assa J J, Fröb H and Leo K 2006 *Appl. Phys. Lett.* **88** 201912
[8] Lee T and Lin M S 2007 *Cryst. Growth Des.* **7** 1803
[9] Blasse G 1988 *Prog. Sol. State Chem.* **18** 79
[10] Walser A D, Priestley R and Dorsinville R 1999 *Synth. Met* **102** 1552
[11] Halls M D and Schlegel HB 2001 *Chem. Mater*. **13** 2632
[12] Shukla V K and Kumar S 2010 *Synthetic Metals* **160** 450
[13] Blasse G and Grabmaier B C 1994 *Luminescent Materials* Springer Verlag ISBN 3-540-58019-0 New York

Synthesis and characterization of a green $CaAl_xO_y:Tb^{3+}$ phosphor using solution combustion method

K.E Foka[a]*, F.B Dejene[a]* and H.C Swart[b]

[a]Physics Department, University of the Free State (Qwqwa Campus), Private Bag x13, Phuthaditjhaba, 9866, South Africa

[b]Physics Department, University of the Free State, P.O. Box 339, Bloemfontein, 9300, South Africa

* Corresponding author: E-mail: dejenebf@ufs.ac.za

Abstract. $CaAl_xO_y:Tb^{3+}$ green phosphors were obtained at low temperature (500 ^{0}C) by a solution- combustion method. The morphology and structural properties were characterized by scanning electron microscopy (SEM) and X-ray diffraction (XRD), and the structural analysis revealed the presence of both $CaAl_4O_7$ and $CaAl_2O_4$ monoclinic. The characteristic luminescence properties were investigated using emission spectra. The optimal intensity was obtained when the concentration of Tb^{3+} was increased to 2 mol%.

1. Introduction

Long lasting phosphor can be used as a type of energy storing material, which can absorb the visible light, store the energy, and then release energy as visible light which lead to a long persistence afterglow in darkness [1, 2]. Afterglow phosphors have attracted considerable attention due to their potential application in various fields, including emergency signs, light sources, luminous paints, Optical data storage and dial plates of glow watches, etc. [3, 4]. Among various host matrix materials, aluminates are considered to possess high energy efficiency in a wide excitation wavelength range and a high quenching temperature range [5]. The $CaAl_2O_4$ host belongs to the stuffed tridymite structure and the framework consists of AlO_4 tetrahedras, with Ca^{2+} ions in the cavities to balance the charge [6]. Terbium is one of the most widely use rare earth ion which emits strongly in the green region. The Tb^{3+} ions luminescence under UV excitation usually consists of four main emission bands around 490, 545, 580 and 620 nm [1]. Compared with other synthesizing methods, like Solid state reaction, Sol-gel technique, Microwave heating techniques, etc. the combustion method is very simple, energy saving and takes only a few minutes [7,8] to prepare. It must be kept in mind that the combustion process involves an exothermic reaction with a large amount of gas liberation which might involve a safety risk. In this study, we report on the synthesis and characterization of the phosphor calcium aluminate activated with different concentration of Tb^{3+} ions that was prepared with the solution combustion method.

2. Experimental procedure

Tb^{3+} doped $CaAl_xO_y$ phosphor was prepared by the solution combustion method. The $Ca(NO_3)_2.4H_2O$, $Al(NO_3)_3.9H_2O$, $Tb(NO_3)_3.6H_2O$, and $CO(NH_2)_2$ were taken as starting materials and were dissolved in 5ml of deionized water and stirred for 30 minutes. The doping concentrations of the Tb^{3+} ions were

0.25, 0.4, 0.5, 1, 1.5, 2 mol%, respectively. After stirring, the transparent solution was taken into a muffle furnace maintained at a temperature of 500 ^{0}C. The solution boiled and underwent dehydration, followed by decomposition with the escape of large amount of gasses (oxides of carbon, nitrogen and ammonia) and then a white foamy and voluminous ash was produced. The whole process was over within less than 5 min, after the product was cooled to room temperature. The foamy powder was crushed into a powder using a pestle mortar and the obtained white powder was used for characterization. The techniques used for characterization were X-ray diffraction (XRD), Scanning electron microscope (SEM), Energy dispersive spectroscopy (EDS) and Photoluminescence (PL).

3. Results and Discussion

3.1 X-ray diffraction and morphology of the sample

The XRD of the $CaAl_xO_y$ host phosphor doped with different concentration of Tb^{3+} are shown in Fig. 1. The XRD peaks revealed both the $CaAl_2O_4$ monoclinic (JCPDS) card number 70-0134 and $CaAl_4O_7$ monoclinic (JCPDS) card number 74-1467 structure. The XRD patterns show diffraction peaks of $CaAl_4O_7$ for all concentrations. With an increasing Tb^{3+} concentration the monoclinic $CaAl_2O_4$ phase began to appear and its peaks became sharp. This indicates that the proportion of the monoclinic form was increase and lead to $CaAl_2O_4$ monoclinic as single phase for the 2% doping. The doping of the different concentration of Tb^{3+} ions on the host did not cause much significant change in the structure. Fig.2 shows the SEM images of $CaAl_xO_y$:Tb^{3+} with a 2 and 1.5 mol% concentration of Tb^{3+} ions. The surface aspect of the particles is found to be nonspherical and irregular. It was observed that the different concentration of the Tb ions have no effect on the morphology of the samples.

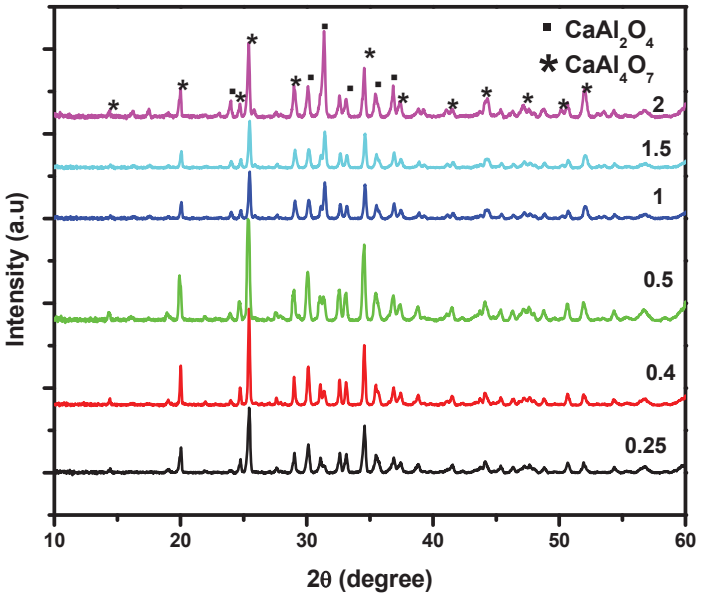


Figure 1: X-ray diffraction patterns of $CaAl_xO_y$ with different concentrations of Tb^{3+} ions.

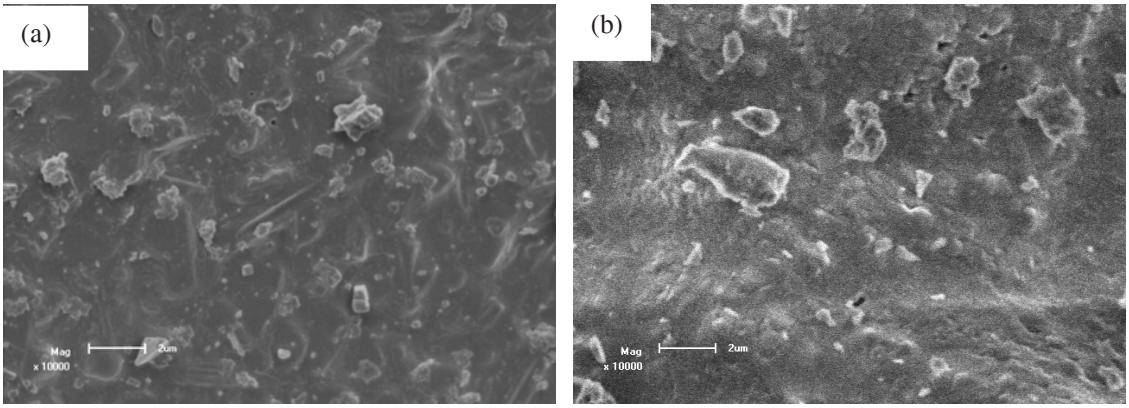


Figure 2: SEM image of $CaAl_xO_y:Tb^{3+}$ phosphor (a) 2 mol% and (b) 1.5 mol% at high magnification

Fig. 3 shows the EDS spectrum of the $CaAl_xO_y:Tb^{3+}$ phosphor. The presence of $CaAl_xO_y$ in the sample is confirmed with the Ca, Al, and O peaks. The C peak is coming from the carbon tape on which the sample was mounted. The Tb small peak is visible in the spectra due to the high concentration of Tb^{3+} in the 2 and 1.5 mol% samples. For 0.5 and 0.25mol% of Tb^{3+} in the sample the amount of Tb was too small to be detected by EDS.

3.2 Photoluminescence spectra

Fig. 4 illustrates the emission spectra of $CaAl_xO_y:xTb^{3+}$ (where x= 0.25, 0.4, 0.5, 1.0, 1.5 and 2.0 mol%). The emission spectra show the peaks mainly at 493, 545, 588 and 622 nm respectively when

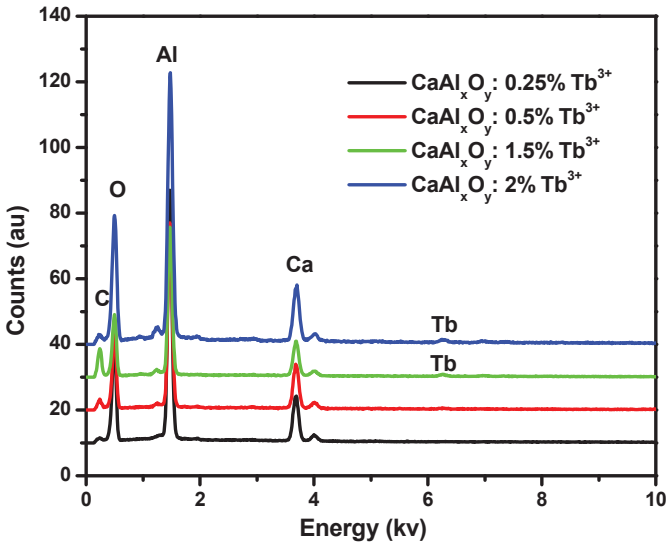


Figure 3: EDS spectra of $CAl_xO_y:Tb^{3+}$ with different concentration of Tb^{3+} ions.

excited with a 239 nm excitation wavelength. The emission peaks are from transition of the $^5D_4$ state to the ^7F_J (J = 6, 5, 4, 3) state. It was found that with an increase in the concentration of Tb in the host lattice there was a significant increase in the intensity of all the peaks. The excitation spectra of the $CaAl_xO_y:Tb^{3+}$ were recorded at an emission of 545 nm. It consists of 4f—5d transitions [2]. The broad band was centered at 239 nm. The intensity increased with an increase in the concentration of Tb^{3+}. The PL intensity of $CaAl_xO_y$: xTb^{3+} where x is the different mol percentages of Tb^{3+} ($0.25 \leq x \leq 2.0$) as a function of the Tb^{3+} concentrations is shown in Fig. 4(c). The emission peaks intensity increased when the amount of Tb increased, and a maximum value for the concentrations tested in our case was found for x = 2 mol%. By increasing Tb^{3+} concentration further, the PL intensity normally will be quenched. In this study, however, the Tb^{3+} concentration was only varied up to 2%. Future work need to be done to determine the real maximum.

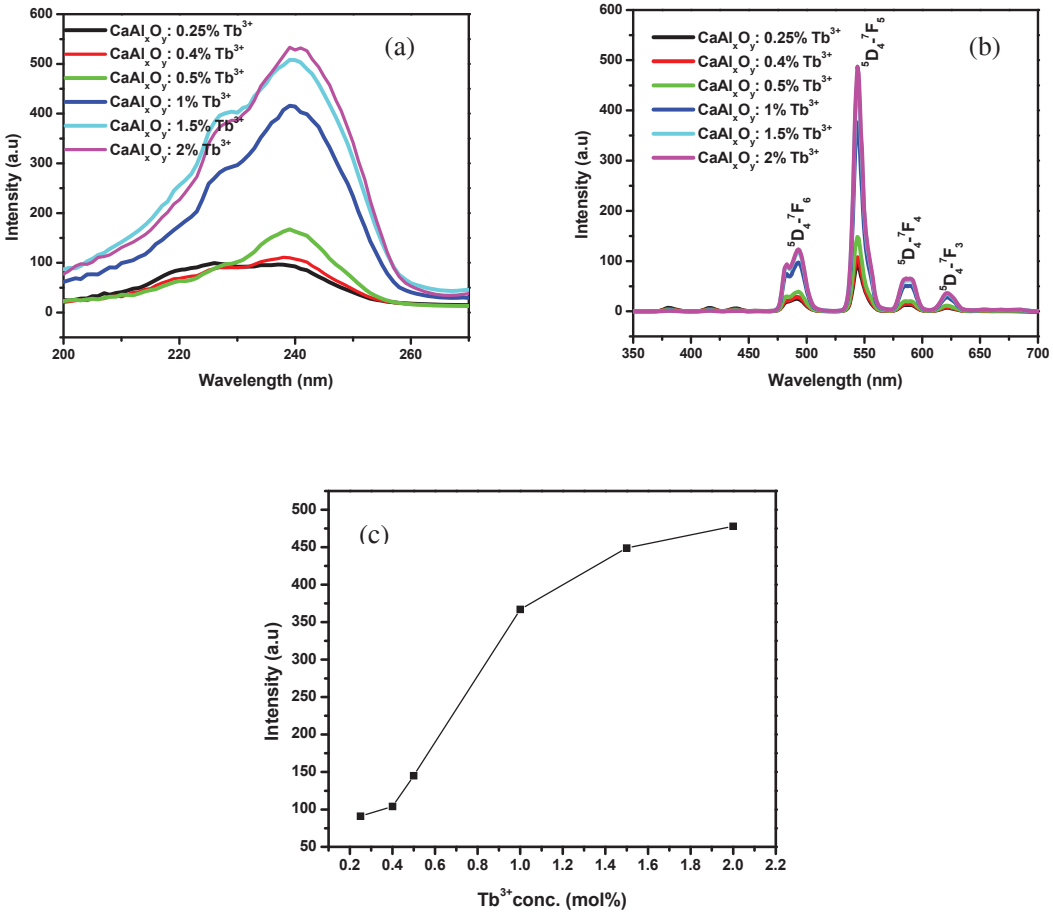


Figure 4: (a) Photoluminescence excitation (b) emission spectra of $Cal_xO_y:Tb^{3+}$ and (c) PL intensity as a function of Tb^{3+} concentration.

Fig. 5 shows the decay curves of the $CaAl_xO_y:Tb^{3+}$ phosphor with a Tb concentration of 0.25 and 2 mol%. The results indicate that the intensity of the phosphorescence decreased quickly at first and then at a slower rate. The curve of the sample doped with the higher concentration displays a higher intensity and long afterglow time than the curve for the lower Tb concentration($CaAl_xO_y:0.25\%Tb^{3+}$). The initial luminescence brightness was improved from 48 a.u to 147 a.u and the long afterglow time was prolonged from 10 ms to 20 ms.

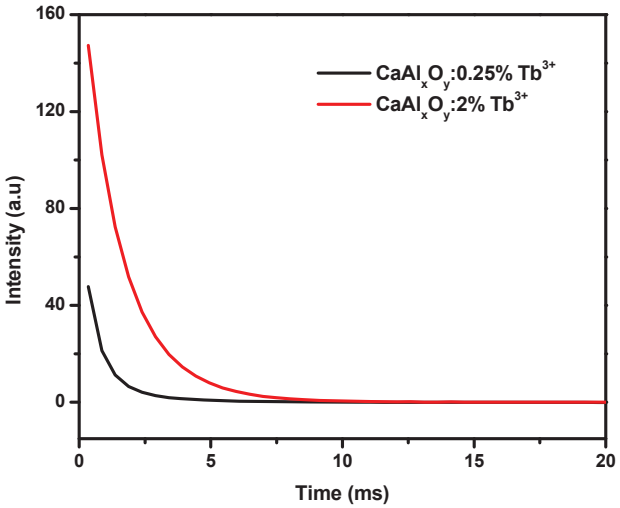


Figure 5: Decay curves of $CaAl_xO_y:Tb^{3+}$ phosphor for different Tb concentration as indicated.

4. Conclusion

$CaAl_2O_4:Tb^{3+}$ phosphor was successfully synthesized by the combustion method. The XRD results implied that the phosphor has a monoclinic structure. The increase of the Tb^{3+} concentration in the sample showed a significant enhancement in the emission intensity of the phosphor. The brightness and the afterglow properties were enhanced with a higher doping concentration of Tb^{3+}.

Acknowledgement

The authors send gratitude to the National Research Foundation (NRF) for funding the project and the University of Free State Physics (Physics department) for the research techniques used in this study.

References

[1] Huajie S, Donghua C, Wangjun T, Yuhua P 2008 *Displays* **29** 41
[2] Zhengwei L, Yingliang L 2005 *Materials Chemistry and Physics* **93** 129
[3] SHideo H, Takeru K, Hiroshi K, Masaaki Y, Yoshinori Y, Naruhito S 1998 *J. Phys.: Condens. Matter* **10** 9541
[4] Xuhui X, Yuhua W, Xue Y, Yanqin L, Yu G *2011 J. Am. Ceram. Soc.* **94** 160
[5] Zhang Y, Li L, Zhang X, Wang D, Zhang S 2008 *Journal of Rare Earths* **26(5)** 656
[6] S Shin-Hei C, Nam-Hoon Y, Young-Hoon Y and Sung-Churl C 2006 *Journal of Ceramic Processing Research* **7(1)** 62
[7] Huajie S, Donghua C 2007 *Luminescence* **22** 554
[8] Barrows B S, Melo P S, Kiminami R H G A, Costa A C F M, de Sá G F, Alves Jr S 2006 *J Mater Sci* **41** 4744

MOVPE grown self-assembled InSb quantum dots: Structural characterization

M Godbole, E.J. Olivier, C Heiligers, J.H. Neethling, J.R. Botha

Department of Physics, Nelson Mandela Metropolitan University, PO Box 77000, Port Elizabeth, 6031, South Africa

E-mail: mohit.godbole@nmmu.ac.za

Abstract. InSb quantum dots were grown on GaSb and GaInSb surfaces using metal-organic vapour phase epitaxy (MOVPE) at atmospheric pressure.The effects of surface morphology and indium mole fraction in the buffer layer on the height of the islands were investigated. Quantum dots grown on GaSb (100) substrates etched for 120 s using a tartaric acid solution are flatter and have a lower surface density than those etched for 120 s using a hydrochloric acid solution. The indium mole fraction in the buffer layer was varied from 0 to ~0.08 in order to study the change in height of the dots. As the percentage of indium increases it leads to a reduction in lattice mis-match and thus lateral strain at the interface, resulting in a higher density of dots with smaller average height.

1. Introduction

In recent years there has been a tremendous development in the field of self assembled quantum dots (QDs) due to their unique physical and optical properties [1]. Self-assembled QDs are normally formed using the Stranski-Krastanov growth mode [2]. An example of a QD system, suitable for mid-infrared optoelectronic devices, is InSb in a GaSb matrix [3]. Most of the work performed to date on this system has been by molecular beam epitaxy (MBE) [4, 5]. Comparatively little work has been done on MOVPE grown structures, especially when considering InSb QDs grown on GaInSb surfaces. It has been shown theoretically that a change in the aspect ratio of InAs QDs grown on GaAs surface results from a change in the indium mole fraction in the InGaAs capping layer [6]. Since the lateral alignment, surface distribution, size and height of QDs play an important role in the design of mid-infrared detectors [7], it is important to test the influence of strain on these properties for the InSb/GaSb QD system.

In this paper the effects of the surface morphology of GaSb and GaInSb surfaces on the size and surface distribution of InSb QDs are reported. The effect of indium mole fraction in the buffer layer on the height of the QDs is also investigated.

2. Experimental procedure

The QDs were grown by MOVPE at atmospheric pressure, using a GaSb (100) substrate. Trimethylindium, triethylgallium and trimethylantimony were used as precursors. A 200 nm thick GaSb buffer layer was first grown at a constant susceptor temperature of 513 ^{0}C (typical growth rate of ~1 nm/s). The temperature was then decreased to the QD growth temperature of 450 ^{0}C and InSb was deposited at a V/III ratio of 7.9 for times of 5 s to 40 s. In order to study the effects of substrate morphology on the surface distribution and dot size, the samples were grown using the same procedures and parameters mentioned above, but the substrate was etched using different etchants as mentioned in table 1.

Table 1. Types of etchants and etch times used for etching GaSb (100) substrate

Etchant	Etch mixture	Etch time	Etch rate
Hydrochloric acid solution	HCl : H_2O_2 (80 : 1)	120 s	~ 33 nm/s
Citric acid solution	aq. $C_6H_8O_7$: H_2O_2 (2 : 1)	120 s	~ 0.01 nm/s
Tartaric acid solution	$H_2C_4H_4O_6$: H_2O_2 : HF (20 : 10 : 1)	120 s	~ 166 nm/s

In order to study the effects of indium mole fraction on the surface distribution and dot size, the samples were grown using the same procedures and parameters mentioned above, but with two different indium mole fractions of 0 and 0.3 in the vapour phase during growth of the 200 nm thick GaInSb buffer layer.

A JEOL 7001F field emission scanning electron microscope (SEM) was used to investigate the morphologies of the QDs. The elemental composition of the dots and buffer layer was analysed using a Perkin-Elmer ELAN 6100 inductively coupled plasma/mass spectrometer (ICP-MS). The surface features of the structures were studied using atomic force microscopy (AFM) using a SIS contact mode AFM housed on a CSM Instruments nano-indentor assembly. Strain and defects at the interface between the buffer and wetting layer were studied using a Philips CM20 transmission electron microscope (TEM).

3. Results and discussion

The surface morphology of the substrate plays an important role in surface migration and coalescence of quantum dots [8].The influence of surface morphology on quantum dot height and surface density was studied by using different etchants to etch the GaSb (100) surface prior to growth. Three samples of GaSb substrate were etched for 120 s using the three solutions mentioned in Table 1. Figure 1 shows the surface morphology of the GaSb substrate surface etched with hydrochloric acid solution and tartaric acid solution for 120 s each. The tartaric acid solution gives the best surface morphology. The fine etch (~4 μm) obtained using the hydrochloric acid solution and superfine etch (~ 1.2 nm) obtained using the citric acid solution is insufficient in smoothing out the roughness of the as-received substrate. It was found earlier that an etch depth of > 10 μm is required to achieve a clear improvement in the surface morphology [9]. It should be noted, however, that prolonged wet chemical etching has been found to produce a convex surface, hence the total etch depth needs to be kept as low as possible.

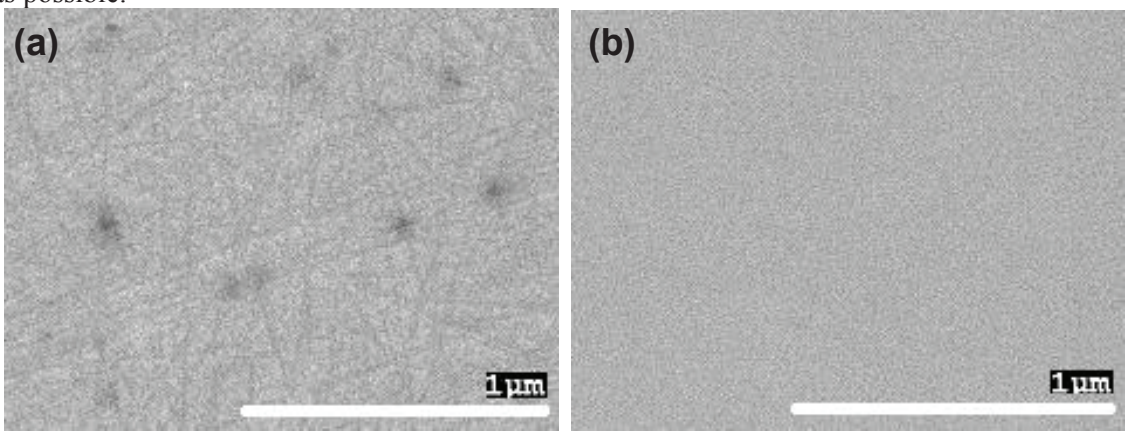


Figure 1. SEM images of GaSb (100) substrate surface etched in (a) hydrochloric acid solution and (b) tartaric acid solution for 120 s each.

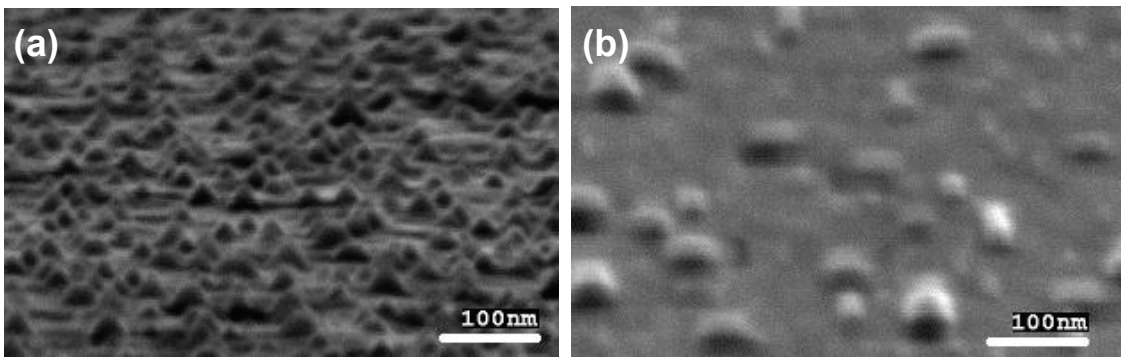


Figure 2. SEM images of InSb islands grown on GaSb (100) substrate surface etched in (a) hydrochloric acid solution and (b) tartaric acid solution for 120 s each.

In Figure 2 SEM micrographs are shown of samples in which InSb islands were grown for 6 s directly on an etched GaSb (100) surface using (a) the hydrochloric acid solution and (b) the tartaric acid solution. The density of dots observed for the substrate etched in the hydrochloric acid solution is considerably higher than for substrate etched in the tartaric acid solution. It is suggested that in the former case, the surface migration of indium atoms will be reduced compared to the latter case (due to the high density of nucleation sites), leading to a high density of small sized islands. The improved morphology of the surface after etching in the tartaric acid solution will promote surface diffusion; reduce the nucleation density, leading to a lower density of dots, as seen in Figure 2(b). The shape of the islands is also different: the pyramids shown in Figure 2(b) are truncated ("flatter").

The defects and strains at the interface were also investigated. In Figure 3(a) and 3(b) cross-sectional TEM micrographs of the interface between the InSb layer and GaSb (100) substrate are shown. The lattice mismatch of 6.3 % between the InSb wetting layer and the GaSb buffer layer can be clearly seen by the strain contrast images near the interface. Planar defects in the form of stacking faults can also be seen in Figure 3(a). The stacking faults are due to atomic stacking errors occurring during growth as a result of the lattice misfit [4].

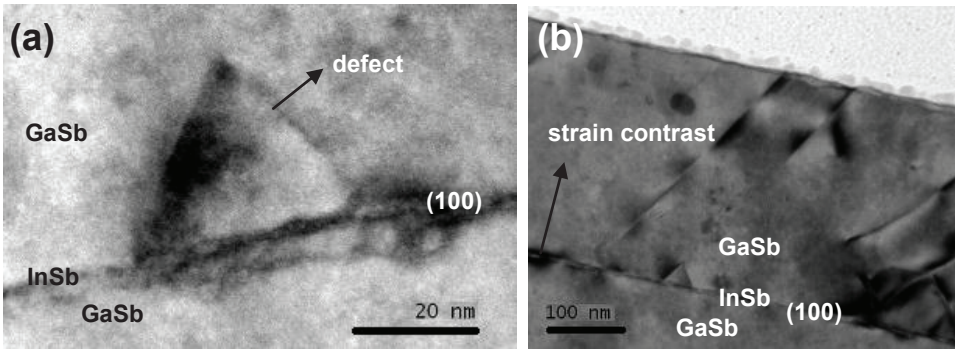


Figure 3. TEM images of InSb in a GaSb matrix showing (a) planar defects (b) strain contrasts at interface.

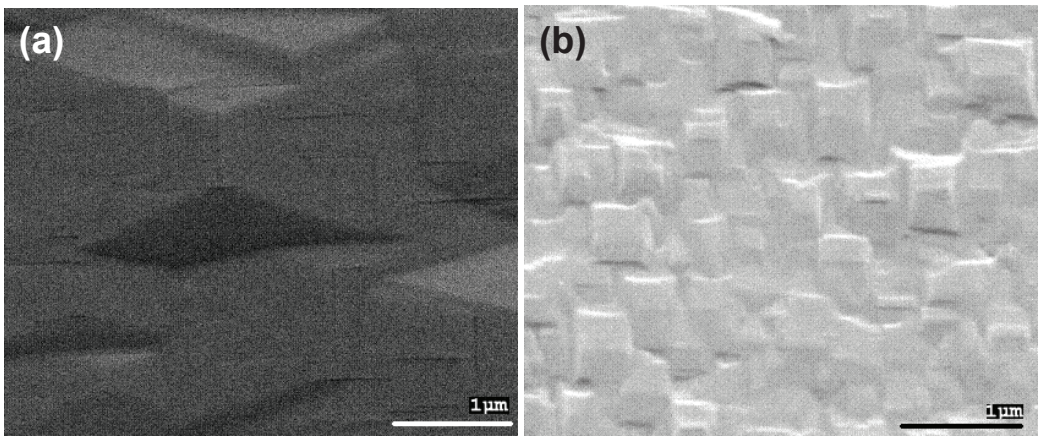


Figure 4. SEM images of surface of (a) ~ 200 nm GaSb and (b) ~ 200 nm GaInSb grown on GaSb (100).

In Figure 4(a) and (b) SEM micrographs of the surfaces of ~200 nm thick GaSb and GaInSb buffers can be seen. The buffer layers were grown under nominally identical growth conditions on substrate etched in the tartaric acid solution, but the surface morphology of the GaSb buffer is better than that of the GaInSb buffer due to an exact lattice match in the former case. In both cases the surface is uneven; the hills present on the GaSb surface, however, are larger in size as compared to those on the GaInSb surface. The facets of these broader hills act as flat surfaces for the nucleation and growth of InSb islands of size < 50 nm, which will lead to a lower density of dots. This is evident from the AFM images and accompanying fast Fourier transforms (FFT) of the AFM images shown in Figure 5 and 6. The intensity shown on the FFTs is a representation of the density and spatial separation between

dots. It is seen from the FFTs that the density of islands grown on GaSb is lower than those grown on GaInSb. Due to less strain between InSb/GaInSb as compared to between InSb/GaSb, the grown islands also have a lower height. Statistical analysis performed on the AFM images revealed an average QD height of 29 nm and 54 nm for samples grown on GaInSb and GaSb, respectively. An approximation of the indium content in the GaInSb buffer layer was done using ICP-MS chemical analysis, yielding a value of ~ 8 % (in the solid phase).

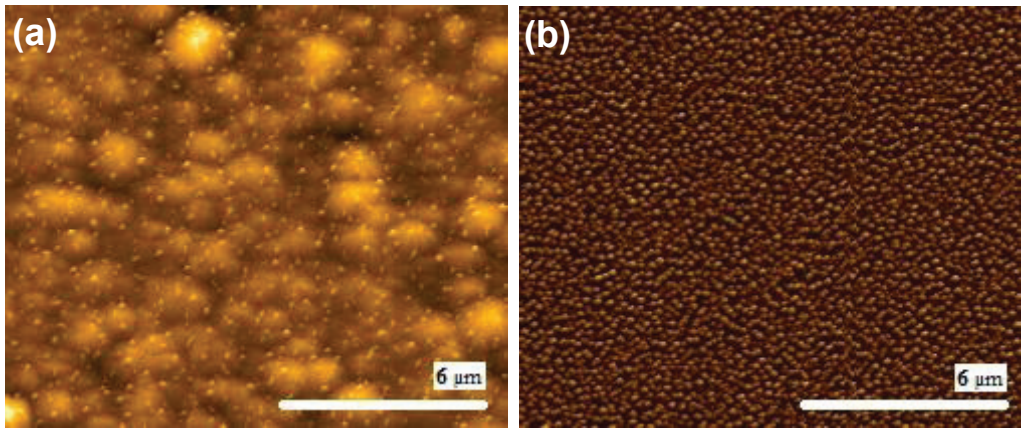


Figure 5. AFM images of InSb islands grown on (a) ~ 200 nm GaSb and (b) ~ 200 nm GaInSb buffer, both grown on GaSb (100).

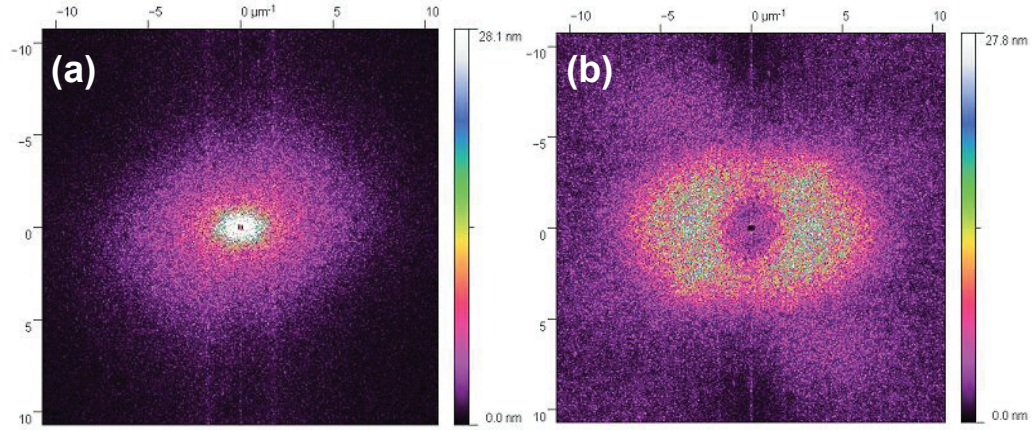


Figure 6. FFT AFM images depicting surface density of InSb islands grown on (a) ~ 200 nm GaSb and (b) ~ 200 nm GaInSb buffer, both grown on GaSb (100).

4. Conclusions

Single layer InSb self-assembled QDs were grown at 450 ^{0}C using atmospheric pressure MOVPE on GaSb and GaInSb surfaces. The effects of substrate etching and indium mole fraction in the underlying buffer layer on the areal density of the QDs were investigated. QDs have a higher density on a substrate etched with hydrochloric acid solution, due to the high density of nucleation sites on such a relatively rough surface. As the indium concentration in the buffer layer increased, the lateral strain between the buffer and the dots was reduced, leading to a reduced height of the islands. This is a successful step towards the growth of multilayer QDs with strain engineering using spacer layers for detector applications.

Acknowledgements

This work is based upon research supported by the SA Research Chairs Initiative of the Department of Science and Technology and the National Research Foundation, South Africa. The financial support of Armscor and the Nelson Mandela Metropolitan University is also gratefully acknowledged. We would like to thank Dr. Richard Betz and Dr. Eric Hosten from the Chemistry Department, NMMU for many useful discussions and for ICP-MS measurements. We would also like to thank

Prof. M.E. Lee and Mr. William Goosen for SEM measurements; Miss Tasneem Limbada, and Miss Tanya Liebenberg for sample preparation and Mr. Johan Wessels for technical assistance.

References

[1] V.A. Shchukin, N.N. Ledentsov and D. Bimberg 2004 *Epitaxy of Nanostructures*, Springer, Berlin.
[2] D. Bimberg, M. Grundmann and N. Ledentsov 1999 *Quantum Dot Heterostructures,* John Wiley and Sons, England.
[3] A. Krier 2006 *Mid-infrared Semiconductor Optoelectronics*, Springer, Berlin.
[4] N. Deguffroy , V. Tasco, A.N. Baranov, E. Tournie, B. Satpati, A. Trampert, M.S. Dunaevskii, A. Titkov and M. Ramonda 2007 *J. Appl. Phys.* **101** 124309.
[5] V. Tasco, N. Deguffroy, A.N. Baranov, E. Tournie , B. Satpati, A. Trampert, M. Dunaevski and A. Tiktov 2006 *Appl. Phys. Lett.* **89** 263118.
[6] M. Usman, D. Vasileska and G. Klimeck 2010 *Physics of Semiconductors* **1119** 527.
[7] R. Paiella 2006 *Intersubband Transitions in Quantum Structures,* McGraw Hill, USA.
[8] P. Sutter, E. Mateeva-Sutter and L. Vescan 2001 *Appl. Phys. Lett.* **78** 1736.
[9] M. Godbole, E.J. Olivier, E. Coetsee, H.C. Swart, J.H. Neethling and J.R. Botha 2011 submitted to *Physica B: Condensed Matter* (under review).

Spectroscopic studies of nanofluorides doped with Ln^{3+} synthesized *via* thermal decomposition of organic precursors

M A Gusowski and H C Swart

Department of Physics, University of the Free State, P.O. Box 339 (IB51), Bloemfontein 9300, South Africa

E-mail: gusowskima@ufs.as.za

Abstract. The results of recent investigation of wide band-gap luminescent cubic $NaYF_4$ nanocrystals doped with the optically active rare earth ions, RE=Pr, Eu, Er, as the converters of VUV radiation into visible emission, are presented. The materials have been synthesised by the thermal decomposition of the alkaline metals trifluoroacetate and lanthanide trifluoroacetate precursors. The high-boiling octadecene was selected as the growth medium for the nanoparticles and oleic acid was employed as the surfactant. The crystal structures of the compounds have been confirmed by X-ray diffraction powder diffraction. The morphology study of the materials was performed by transmission electron microscope. The spectroscopic properties of the crystals have been determined based on the excitation and emission spectra recorded in the UV–VIS–NIR spectral region. The optical and morphological properties of materials have been presented and discussed.

1. Introduction

There has been a growing demand for quantum cutting and upconverting efficient nanoluminescent materials during the last decades. The wide band-gap luminescent materials, especially fluorides, are one of the major role players. Due to a number of advantages, such as the excellent transparency in the infrared (IR) to far ultraviolet (UV) range and low-energy phonons as well as low toxicity they have recently gathered considerable attention for their potential applications in bioimaging and diagnostics. The same group of materials having down-conversion characteristic can be applied to the improvement of solar cell efficiency in the quantum cutting processes. In our work we present the facile way of their nanocrystal synthesis. The aim of our work is to show the possible ways of the energy relaxation in the nanocrystals doped with luminescence active ions after the excitation with the high energy photons. The influence of the nanocrystal's size on the luminescence and excitation spectra of $NaYF_4 : RE^{3+}$ have been also presented.

2. Experimental

2.1. Synthesis of nanocrystals $NaYF_4 : RE^{3+}$

The nanocrystals of $NaYF_4$:RE^{3+} (NC) were synthesized via a thermal decomposition method. Firstly, the trifluoroacetates of selected rare earths were prepared. The lanthanide trifluoroacetate precursors were prepared from the corresponding lanthanide oxides and trifluoroacetic acid in butanol. The corresponding amount of sodium trifluoroacetate and RE (III) trifluoroacetates was added to the

reaction vessel with octadecene and oleic acid. The mixture of sodium trifluoroacetate and oleic acid (90%, technical grade) in molar ratio 1:50 and corresponding amount of 1-octadecene (90%) were mixed in the three-necked flask. The resulting solution was then stirred and heated for 30 min at 120 °C. During this time the flask was purged with dry argon to remove water and oxygen. The solution was heated to 300 °C at a rate of about 10 °C/min. under argon gas protection and kept at this temperature under vigorous stirring for 2 h. A capillary was used to let the argon gas in and out during the synthesis. The mixture was cooled to room temperature. The nanoparticles were isolated via centrifuge at 3000 rpm prior to the addition of absolute ethanol. The nanoparticles were washed several times with ethanol and isolated via centrifuge at 3000 rpm each time. They were stored in ethanol and centrifuged for experiments if needed.

2.2. *Structure and morphology*.
The size and morphology of the nanocrystals were characterized by transmission electron microscopy (TEM) using a Philips CM100 microscope. The powder X-ray diffraction (XRD) patterns were recorded with a Buker D8 ADVANCE diffractometer using Cu Kα radiation (0.15418 nm). The 2θ angle of the XRD spectra was recorded at a scanning rate of 1.2°/min.

2.3. *Spectroscopic techniques*.
The measurements of the luminescence excitation and emission spectra were performed at the Deutsches Elektronen Synchotron (DESY, Hamburg) using the facility of the SUPERLUMI station at HASYLAB22. The luminescence spectra at room temperature and at 10 K were measured within the 300-1000 nm spectral range using 0.3 m (Acton Research Corporation) Spectra Pro308 monochromator-spectrograph in Czerny-Turner mounting equipped with a liquid nitrogen cooled (Princeton Instruments, Inc.) CCD detector. The spectral resolution of the analyzing monochromator was ~0.5 nm. The emission spectra were not corrected for the detector sensitivity and monochromator transmission. High resolution time-resolved luminescence excitation spectra were scanned within 60-300 nm with the primary 2 m monochromator in 15° McPherson mounting (3.2 A resolution) using (Hamamatsu R6358P) a PMT detector with a secondary ARC monochromator. The excitation spectra have been corrected for the incident photon flux compared to that of a reference sample of NaSal whose quantum efficiency is assumed to be about 58% and is constant over the excitation wavelength range from 4 to 25 eV[1]. The temperature could be varied between 8 and 300 K by means of a liquid-helium cryostat (Cryovac GmbH).

3. Results

3.1. Morphology and structure results.
The morphology of the materials was evaluated by the TEM. **Figure 1** shows the TEM images of $NaYF_4$:RE nanocrystals synthesized by thermal decomposition of trifluoroacetates precursors.

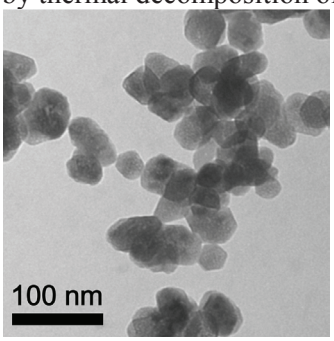


Figure 1. TEM images of $NaYF_4$:RE powders displaying uniformity of the particles.

The synthesized nanoparticles are equiaxed and clearly facetted which in turn is a good indication of their crystallinity. The crystallinity of the synthesized nanoparticles has been subsequently confirmed by powder XRD. The obtained nanocrystals belong to the cubic system, space group Fm3m. The diffraction patterns presented in **Figure 2** disclose reflexes which can be assigned to the (111), (002), (022), (113) (004) (133) and (224) of the (hkl) planes of the standard ICSD-60257 cubic $NaYF_4$.

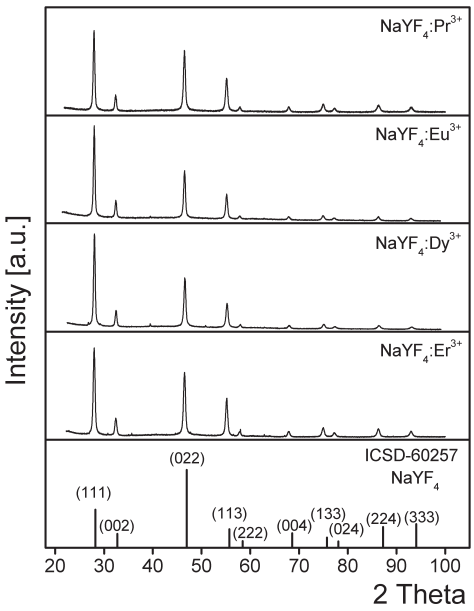


Figure 2. X-ray powder diffraction pattern of $NaYF_4$ nanocrystals doped with RE^{3+} ions compared with the pattern generated based on the monocrystal $NaYF_4$ cubic structure ICSD-60257 standard.

The pattern generated based on cubic $NaYF_4$ monocrystal data from ICSD-60257 is shown for comparison (the bottom diagram of **Figure 2**). It is evident that the prepared nanomaterial is highly crystalline. The positions of the XRD peaks correspond to the standard ICSD-60257 cubic pattern of $NaYF_4$ and no peaks from other phases or impurities have been observed. The XRD peaks have very broad widths originating from the nanosize effect. The average crystallite size of the nanocrystals was calculated according to Scherrer's equation.

$$D = K\lambda / \beta \cos\theta \quad \textbf{(eq 1)}$$

where D is average crystal size (in nanometers), K=0.89, λ is the wavelength of the Cu Kα radiation equal to 0.154 nm, ß is the full width at half maximum of the diffraction peak (FWHM 2θ = 0.392 = FWHM Θ = 0.196 = 0.0034 radians) and θ is Bragg's angle of the diffraction peak. For our calculations the most intense peak at 2θ equal to 28°, giving Bragg θ equal to 14°, has been taken. According to the equation **(eq 1)** the average nanocrystal size is calculated to be about 42 nm, which is consistent with the TEM results.

3.2. Spectroscopic results and discussion

Figure 3A presents the room temperature excitation spectrum monitoring the 396 nm $^1S_0 \rightarrow {}^1I_6$ transition of Pr^{3+} ions in $NaYF_4$ nanocrystals. The intensity of the $^1S_0 \rightarrow {}^1I_6$ transition increases for incident photons with energy corresponding to around 210 nm wavelength where the onset of interconfiguration $4f^2 \rightarrow 4f^5d$ transitions begins. The $^1S_0 \rightarrow {}^1I_6$ luminescence increases with the change of the excitation wavelengths up to a maximum at 180 nm. There is no evidence of any transfer from the matrix to the 4f5d configuration which can increase the luminescence intensity of $^1S_0 \rightarrow {}^1I_6$ transitions.

In the luminescence spectrum presented in Figure 3B eight bands are clearly resolved, having maxima at 249, 269, 334, 396 nm assigned to $^1S_0 \rightarrow {}^3H_6$, $^3F_4$, $^1D_2$, $^1I_6$ transitions and at 482, 530, 600, 636 nm which correspond to $^3P_0 \rightarrow {}^3H_4$, $^3H_5$, $^3H_6$, $^3F_2$ transitions of the Pr^{3+} ions, respectively.[2] When Pr^{3+} ions are excited into the 4*f*5*d* configuration, firstly relaxation from the lowest 4*f*5*d* level to the $^1S_0$ level takes place. It is followed by a photon cascade emission, i.e. the emissions of $^1S_0 \rightarrow {}^1I_6$ and followed by the emission of $^3P_0 \rightarrow {}^3H_J$,$^3F_J$.

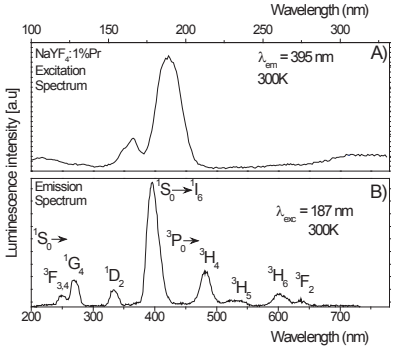


Figure 3. Excitation and emission spectrum of $NaYF_4$:1%Pr NC.

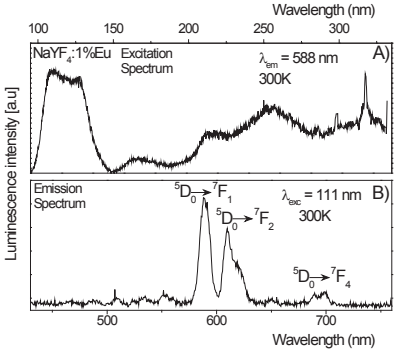


Figure 4**.** Excitation and emission spectrum of $NaYF_4$:1%Eu NC.

The excitation spectrum monitoring the $^5D_0 \rightarrow {}^7F_1$ emission is presented in Figure 4A. In the region above 150 nm the spectrum is dominated by the broad band which origin can be ascribed to the absorption of NaYF4 nanocrystals matrix overlapping with the *fd* band. This can be close to the host lattice absorption edge which has been already observed in $LiYF_4$.[3] The broad band observed in the region 150 to 200 nm with a maximum at 166 nm is assigned to the fluoride-europium charge-transfer band. The origin of the broad bands in the region 200 -300 nm were assigned to the defect and the contaminations of the nanocrystals by the surfactant used for the synthesis. Interesting is the observation that they also transfer the energy to the optically active centers and feed the emission from the $^5D_0$ level. In Figure 4B the luminescence of $NaYF_4:Eu^{3+}$ nanocrystals is presented. The main emissions channels are $^5D_0$ - $^7F_1$ and $^5D_0$ - $^7F_2$ but in the region 450 to 550 nm very week transitions originating from $^5D_1$ and $^5D_2$ are also observed.

Figure 5A presents the VUV excitation spectrum of $^4F_{9/2}$ to $^6H_{13/2}$ emission of Dy^{3+} in $NaYF_4$ nanocrystals. The edge of the matrix absorption is evident at the 125 nm. The band responsible for matrix absorption similar to this one presented in the Er excitation picture has been removed for clarity. In this spectrum the lowest low-spin fd transition of Dy^{3+} is observed at the 175 nm. Its structure is very similar to the transitions of Dy^{3+} observed in the $LiYF_4$ crystals.[4] The expected fd high-spin transitions of Dy^{3+} around 190 nm are fully covered by the intense broad band transitions originating from the traces of surfactant-covered nanoparticles. Figure 5B shows the emission spectrum of $NaYF_4$: 1%Dy NCs measured after the excitation of the band gap at 91 nm which has been registered at room temperature. In this spectrum four Dy^{3+} bands are clearly resolved, having maxima at 477, 573, 663, 754 nm, assigned to $^4F_{9/2} \rightarrow {}^6H_{15/2}$, $^6H_{13/2}$, $^6H_{11/2}$ and $^6H_{9/2}$+$^6F_{11/2}$ transitions.

In Figure 6A the excitation spectrum of $^4S_{3/2} \rightarrow {}^4I_{15/2}$ in $NaYF_4$ doped with Er^{3+}, registered at 10 K is presented. Here the host lattice absorption at wavelengths shorter than 120 nm is clearly seen. The energy from the matrices is efficiently transferred to the luminescence levels of Er3+ and has determined the main channel of the excitation of $^4S_{3/2} \rightarrow {}^4I_{15/2}$ emission. In the excitation spectrum the bands in the region 120 - 300 nm are not well resolved and do not show the fine structure. In the emission spectrum in Fig.6B the main emissions are due to the Er^{3+} $^2P_{3/2} \rightarrow$ $^4I_{13/2}$, $^4I_{11/2}$, $^4I_{9/2}$ transitions (399, 467 and 521 nm) and $^4S_{3/2} \rightarrow {}^4I_{15/2}$ transition (around 550 nm). The lines around 525 nm can be partly due to the emission from the $^2H(2)_{11/2}$ state, which lies just above the $^4S_{3/2}$ level and has been thermally populated at room temperature, as well as from the $^2P_{3/2}$ level. In the spectrum we can also

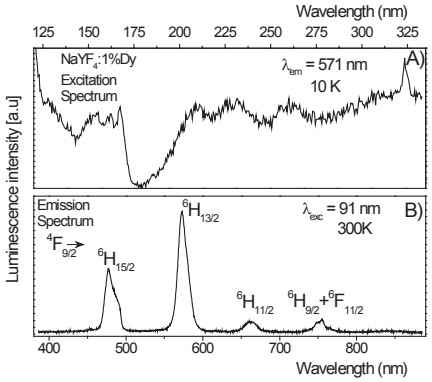


Figure 5**.** Excitation and emission spectrum of $NaYF_4$:1%Dy NC.

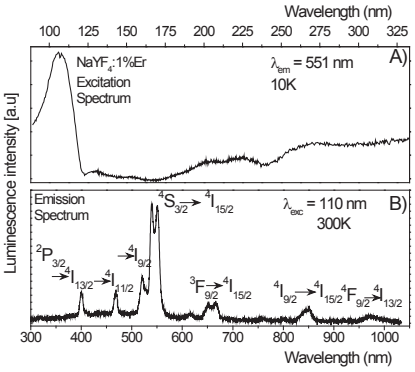


Figure 6**.** Excitation and emission spectrum of $NaYF_4$:1%Er NC.

observe the transitions $^4F_{9/2} \rightarrow {}^4I_{15/2}$ and the transitions within $^4I_{9/2} \rightarrow {}^4I_{15/2}$ and $^4F_{9/2} \rightarrow {}^4I_{13/2}$. The band assignments are done based on the previously studied YAlO3 and LiGdF4 systems.[5-6]

4. Conclusion

The crystalline nanocrystals of the $NaYF_4$ doped with RE were successfully synthesized by the thermal decomposition of trifluoroacetates precursors. The size and the crystalinity of the particles has been confirmed by the TEM and XRD techniques. The luminescence characteristics of prepared nanoparticles have been studied. In the case of the $NaYF_4$:1% Pr the quantum-cutting (QC) effect on the photoluminescent (PL) spectra of colloidal $NaYF_4$ nanocrystals doped with 1% Pr^{3+} ions has been clearly evident under interconfiguration $4f^2 \rightarrow 4f5d$ excitation. The main channel of luminescence is around 396 nm which originates from the $^1S_0 \rightarrow {}^1I_6$ transition being responsible for the feeding of the photon cascade.

Acknowledgements

The authors send gratitude to the National Research Foundation (NRF) for funding the project and the University of Free State Material and Nano Science cluster for financial support.

References

1. Berkowitz JK and Olsen JA 1991 *Journal of Luminescence*, **50** (2) 111.
2. Kück S, Sokólska I, Henke M, Scheffler T and Osiac E 2005 *Physical Review B*, **71** (16) 165112.
3. van Pieterson L, Reid MF, Wegh RT, Soverna S and Meijerink A 2002 *Physical Review B*, **65** (4) 045113.
4. van Pieterson L, Wegh RT, Meijerink A and Reid MF 2001 *The Journal of Chemical Physics,* **115** (20) 9382.
5. Duan C-k, Tanner PA, Makhov VN and Kirm M 2007 *Physical Review B*, **75** (19) 195130.
6. Wegh RT, van Loef EVD and Meijerink A 2000 *Journal of Luminescence*, **90** (3-4) 111.

Spin-density-wave properties of $(Cr_{84}Re_{16})_{100-x}V_x$ alloys

BS Jacobs, ARE Prinsloo[1], CJ Sheppard and AM Strydom

Department of Physics, University of Johannesburg, PO Box 524, Auckland Park, 2006

E-mail address: alettap@uj.ac.za

Abstract. This paper reports on the spin-density-wave properties of the $(Cr_{84}Re_{16})_{100-x}V_x$ alloy system, with $0 \leq x \leq 10.4$. The $Cr_{84}Re_{16}$ alloy represents an important and interesting position on the magnetic phase diagram of CrRe, just below the critical concentration where the Re concentration suppresses the antiferromagnetism in this alloy system. The possible coexistence of antiferromagnetism and superconductivity in CrRe was previously studied in samples with concentrations in the range of 14 to 25 at.% Re. In this study the $Cr_{84}Re_{16}$ alloy is tuned to a possible critical point through the addition of V. The Néel transition temperatures (T_N) for samples with x = 0.0, 5.7, 8.5 and 10.4 are obtained using electrical resistivity (ρ) and magnetic susceptibility (χ) measurements. A plot of T_N versus V concentration indicates the existence of a possible quantum critical point at $x \approx 10.5$ at.% V.

1. Introduction

The concentration versus temperature magnetic phase diagram of the CrRe alloy system exhibits three spin-density-wave (SDW) phases, the longitudinal (L) incommensurate (I) SDW, the transverse (T) ISDW and the commensurate (C) SDW [1]. A triple point is situated at $c_t \approx 0.30$ at.% Re, where the ISDW, CSDW and paramagnetic (P) phases coexist. Only a nearly linear ISDW–P phase line exists for $c < c_t$, but for $c > c_t$ both ISDW–CSDW and CSDW–P phase lines, which are strongly non-linear, appear on the phase diagram [1].

The CrRe system shows rather normal behaviour, fairly well understood within the framework of the canonical model, where Re behaves as an electron-donor in the Cr matrix [1]. As the Re concentration is increased above the triple point concentration, the electron and hole octahedral Fermi surfaces become closer in dimension. As both surfaces are almost of equal size, an improvement of the nesting of these two surfaces occurs with the concomitant increase in the Néel temperature (T_N) reaching saturation at approximately 7 at.% Re. At about 7 at.% Re, perfect nesting between the electron and hole surfaces occurs [2, 3]. Higher concentrations of Re result in a reduction of T_N. This behaviour is presently not well understood, but corresponds to that observed for CrRu [1].

Previous research on CrRe has focussed on alloys and single crystals close to the triple point concentration [1, 4], as well as on concentrations in the range of 14 to 25 at.% Re [1]. Interest was drawn to CrRe alloys with concentrations above 14 at.% Re after the work of Muheim and Müller [5] on the nature of the phase boundary and the possible coexistence of the SDW and the superconducting phases in $Cr_{100-y}Re_y$ alloys, indicating superconductivity above 20 at.% Re. The coexistence of antiferromagnetism and superconductivity was found in a sample with 17 at.% Re [6], with a Néel temperature of approximately 160 K and superconducting transition temperature typically around 3 K. Superconductivity was also later found [7] in samples containing concentrations as low as 14 at.% Re,

with a superconducting transition of approximately 1 K. This research was extended by Nishihara *et al.* [8] investigating the magnetic and superconducting properties of samples containing between 14 to 25 at.% Re.

Metallurgical problems associated with preparation of CrRe samples of uniform composition have been reported [5,9]. This led Damaschke and Felsh [7] to caution against the conclusion that the observation of a Néel transition in samples with concentrations higher than 14 at.% Re is used as firm evidence for the coexistence of the SDW and superconducting phases.

The behaviour around critical points of the Cr alloys can be probed by fine tuning of the parameters that influence SDW formation in these systems. The first such parameter is the effect of electron concentration on the area of the electron and hole Fermi surface sheets of Cr that nests [1] when the antiferromagnetic phase is entered. This nesting decreases the energy of the system through electron-hole pair condensation and results in the appearance of SDW energy gaps at the Fermi surface in certain directions of k-space on cooling through T_N. The nesting area, and concomitantly the stability of the SDW state, depends on the electron concentration per atom (*e/a*) which can easily be tuned by alloying Cr (*e/a* = 6) with elements like V (*e/a* = 5) or Mn (*e/a* = 7) to respectively decrease or increase the electron concentration. This effect plays a dominant role in affecting the SDW state in Cr alloys when alloyed with 3-d solutes that have an *e/a* number different from that of 3-d Cr. The second important parameter is the electron-hole pair breaking effect due to electron scattering by solute atoms. This parameter plays a dominant role when Cr is alloyed with the isoelectronic 4-d metals Mo (*e/a* = 6) or W (*e/a* = 6) [1] that leaves the electron concentration intact. Taking this into account there thus seems to be several options to tune $Cr_{100-y}Re_y$ through a critical point, using chemical doping with a third element as tuning parameter. This creates the opportunity to investigate the possibility of a quantum critical point in the CrRe alloy system, as well as the phase boundary between the antiferromagnetic and superconducting phase.

Utilizing the above mentioned characteristics Alberts *et al.* [10] investigated the magnetic and superconducting properties of $(Cr_{100-y}Mo_y)_{75}Ru_{25}$ with y = 0, 3, 6 and 10 at.% Mo. The results of this investigation was however inconclusive since these samples were superconducting, but none of them showed any evidence for the Néel transitions.

Recent work on CrRu [11, 12] reignited the interest in the CrRe alloy system, as much correspondence is seen between these two alloy systems [1]. Firstly, the *T-c* magnetic phase diagrams of these two systems are very similar and in the second place the possibility of the coexistence of the SDW phase together with the superconducting phase has also been considered in both these alloys previously [1]. The in-depth studies on $Cr_{86}Ru_{14}$ reported by Reddy *et al.* [11] resulted in investigations into the magnetic and quantum critical behaviour in a CSDW antiferromagnetic $(Cr_{86}Ru_{14})_{100-x}V_x$ system [12]. The concentration–temperature (y-T) magnetic phase diagram obtained for this system depicts a critical point at x = 10.4, classified as a CSDW type of quantum critical point. This study indicated a peculiar difference in the mechanism responsible for driving this system and the $Cr_{100-z}Ru_z$ system to a quantum critical point using *e/a* as a tuning parameter. Our present work seeks to further illuminate this question.

The work by Reddy *et al.* [12] is strongly aligned with the current interest in quantum criticality in Cr alloy systems, as is reflected in recent literature [13, 14, 15, 16]. For a more comprehensive view, it is of interest to broaden studies on quantum criticality to a wider range of the different possible magnetic phase diagrams, and regions amenable to the quantum critical ground state of Cr alloy systems.

The $Cr_{84}Re_{16}$ alloy represents an important and interesting position on the magnetic phase diagram of CrRe, just below the concentration where Re addition suppresses antiferromagnetism in this system and superconductivity is also expected. By adding V to this alloy, the $(Cr_{84}Re_{16})_{100-x}V_x$ system can slowly be tuned to a critical point. In this way it is possible to investigate not only the phase boundary between the antiferromagnetic and superconducting phases, but also the possibility of a quantum critical point in this alloy system.

The present study is a preliminary investigation aiming to extend these studies to the CrRe system in order to better understand the quantum critical behaviour in Cr alloys in general.

2. Experimental

Pseudo-binary $(Cr_{84}Re_{16})_{100-x}V_x$, with $0 \leq x < 11$ alloys were prepared by arc melting in a purified argon atmosphere from Cr of mass fractional purity of 99.99 %, Re of mass fractional purity of 99.99 % and V of mass fractional purity of 99.8 %. Special emphasis was placed on synthesizing specimens of high homogeneity and metallurgical quality in an effort to resolve magnetic and anticipated superconductivity instabilities and phase changes in the phase diagram of the system. This is of great importance as previous studies on CrRe pointed out metallurgical problems encountered in preparing samples of uniform composition [5, 9], attributed to the high vapor pressure of Cr at the liquidus temperature [1, 5]. The unusually high melting temperatures of refractory elements such as Re inevitably demands that adequate heat be supplied at the time of arc-melting in order to coerce full solution of the dopant elements into the Cr matrix. The alloys were annealed in an ultra-high purity argon atmosphere at 1343 K for seven days and quenched into iced water. Powder X-ray diffraction (XRD) analyses were used to confirm that the samples were single phase. The approximate chemical composition of the individual alloys was determined using a scanning electron microprobe (SEM) and energy-dispersive X-ray spectrometry (EDS). The actual elemental composition and homogeneity were determined using electron microprobe analyses. Electrical resistivity (ρ) and magnetic susceptibility (χ) were measured for 2 K $\leq T \leq$ 390 K, using standard Physical Properties Measurement System (PPMS) incorporating appropriate measuring options and a SQUID-type magnetometer based on the Magnetic Properties Measurement System (MPMS) platform of Quantum Design [17]. For magnetic susceptibility measurements the samples were cooled to 2 K in zero field and the measurements were done upon subsequent heating of the samples in a field of 100 Oe. Resistivity measurements for temperatures above 390 K were performed using resistive heating in an inert environment using the standard dc-four probe method and current reversal with Keithley instrumentation.

3. Results and discussion

Figure 1 shows the XRD patterns for the $Cr_{84}Re_{16}$ sample. The entire profile in the $Cr_{84}Re_{16}$ spectra is well fitted to the XRD pattern of pure bcc Cr by adjusting the lattice parameter of pure Cr (0.28839 nm) to 0.29350 nm for this alloy, which corresponds well with that obtained previously in CrRe studies [18]. No additional peaks are detected to within instrumental resolution in the CrRe alloy spectra, indicating that the alloy formed in the bcc phase of pure Cr.

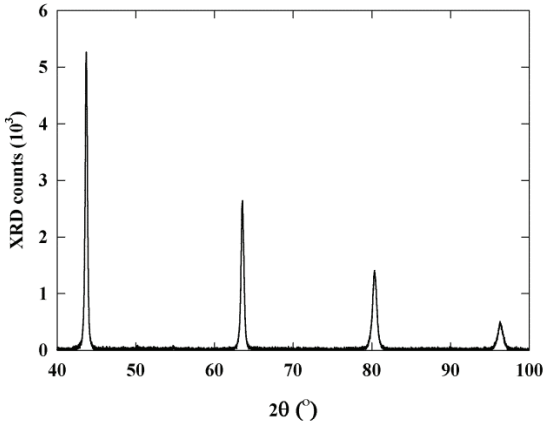


Figure 1. The xrd pattern for the $Cr_{84}Re_{16}$ sample with (hkl) Miller indices of the various reflections expected for the profile of bcc Cr indicated.

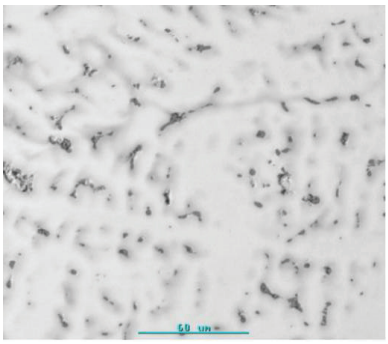


Figure 2. The backscattered-electron image of the $Cr_{84}Re_{16}$ alloy.

The backscattered-electron image of the $Cr_{84}Re_{16}$ annealed mother alloy is shown in figure 2. The image indicates that the mother alloy consists of a CrRe matrix with Cr concentration of (84 ± 1) at.% and Re concentration of (16 ± 1) at.%. Also clear in this image are dark spots. Wavelength dispersive X-ray spectra analyses showed the presence of oxygen within these dark inclusions. These inclusions cover approximately 3 % to 4 % of the image having a surface area of about 180 μm^2. For the individual $(Cr_{84}Re_{16})_{100-x}V_x$ samples used in this study, the V concentrations were found to be 5.7, 8.5 and 10.4 at.%.

The $\rho(T)$ curves for $Cr_{84}Re_{16}$ and $(Cr_{84}Re_{16})_{100-x}V_x$, with x = 5.7, 8.5 and 10.4, are shown in figure 3(a) and figure 3(b). A well defined anomaly in the form of a minimum is seen in the $\rho(T)$ curve of the $Cr_{84}Re_{16}$ sample, attributed to an induced SDW energy band gap at the Fermi energy on cooling through T_N [1]. T_N is often defined for Cr and its dilute alloys as the temperature of the minimum in $d\rho(T)/dT$ accompanying the magnetic phase transition [1]. This definition is also used for the present $(Cr_{84}Re_{16})_{100-x}V_x$ system. The inset in figure 3(a) depicts the temperature dependence of $d\rho(T)/dT$, obtained from the ρ-T curve of the mother alloy shown in figure 3(a), with the position of T_N marked by an arrow. The SDW anomaly is better defined in $d\rho(T)/dT$ than in $\rho(T)$ itself, especially for the alloys with higher V concentrations as an increase in the V concentration not only decreases the Néel temperature, but also results in a reduction in the size of the anomaly seen in these curves. T_N values obtained in this manner are plotted on the magnetic phase diagram of figure 5. The inset in figure 3(b) shows a broad low-temperature minimum for the $(Cr_{84}Re_{16})_{89.6}V_{10.4}$ sample, which is absent in the other three alloys. T_N for this sample was taken at the temperature associated with this minimum. The experimental error in the absolute value of ρ amounts to 5 % and originates mainly from errors in determining the sample dimensions, while our instrumentation permitted a resistivity resolution of 0.5 %. No evidence of superconductivity was seen in these samples down to 2 K but measurements below 2 K are needed to clarify this issue.

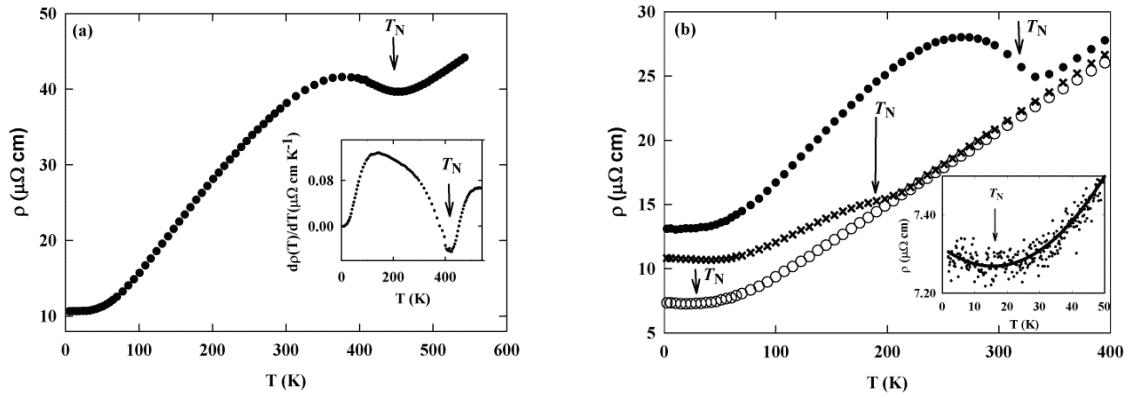


Figure 3. The temperature dependence of the electrical resistivity, ρ, of (a) the annealed $Cr_{84}Re_{16}$ mother alloy and (b) the $(Cr_{84}Re_{16})_{100-x}V_x$ alloys, with x = 5.7 (●) , x = 8.5 (✕) and x = 10.4 (○). The inset in (a) shows the temperature derivative ($d\rho(T)/dT$) of the $Cr_{84}Re_{16}$ alloy. The Néel temperatures (T_N) for the various samples, shown by arrows, are obtained from the minima of $d\rho(T)/dT$. The inset in (b) shows the broad minimum associated with T_N observed for the x = 10.4 sample and the line is a guide to the eye through the data.

The magnetic susceptibility (χ) as function of temperature for the samples containing x = 8.5 and x = 10.4 is shown in figures 4(a) and (b), respectively. In correspondence with results from previous studies [19] broad anomalies are seen in the χ-T curves, becoming more pronounced as the V concentration is increased. Although only a broad minimum was observed in the ρ-T curve of the $(Cr_{84}Re_{16})_{89.6}V_{10.4}$ sample, a clear peak can be seen in the χ-T curve of this alloy giving $T_N \approx 28$ K. The

Néel temperatures (T_N), shown by arrows in these figures, were obtained from the resistivity measurements. The transition temperatures from the χ-T curves were taken at the point where a decrease in χ occurs, as the antiferromagnetic phase is entered. The broken line in figure 4(a) is a fit to the data obtained in the paramagnetic phase, indicating the trend of the curve if the sample was to remain paramagnetic. T_N was taken at the point where the experimental values deviate from the broken line curve and the transition temperatures obtained in this manner correspond well with that determined from the $(d\rho(T)/dT)$ curves. T_N values obtained from the magnetic susceptibility measurements of the $(Cr_{84}Re_{16})_{100-x}V_x$ alloys with x = 5.7, 8.5 and 10.4, were plotted on the magnetic phase diagram of figure 5. The transition temperature of the $Cr_{84}Re_{16}$ sample could however not be determined in the present MPMS setup, as it is limited to a maximum temperature of 400 K.

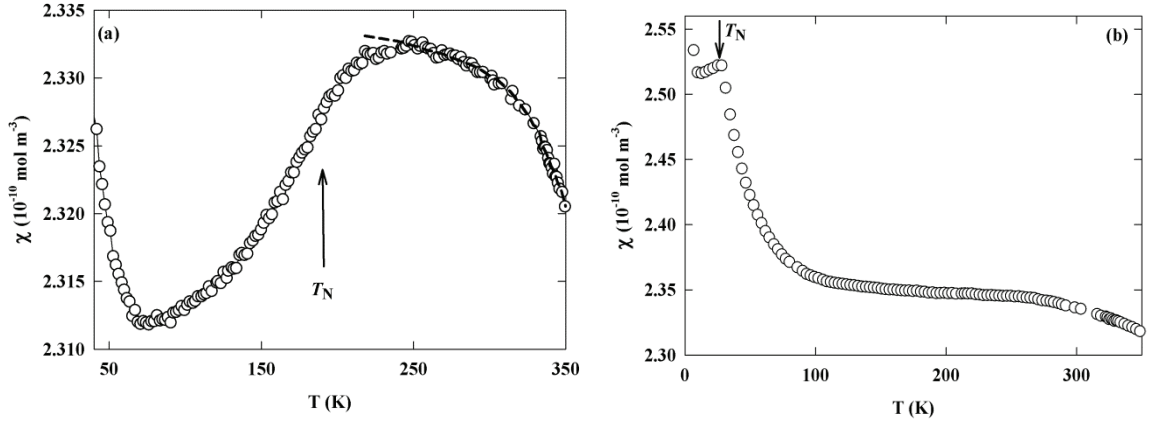


Figure 4. The temperature dependence of the magnetic susceptibily, χ, for the $(Cr_{84}Re_{16})_{100-x}V_x$ alloys, with (a) x = 8.5 and (b) 10.4 at.% V. The Néel temperatures (T_N), shown by arrows in these figures, are obtained from the resistivity measurements. The broken line in (a) is a fit to the experimental data in the paramagnetic region.

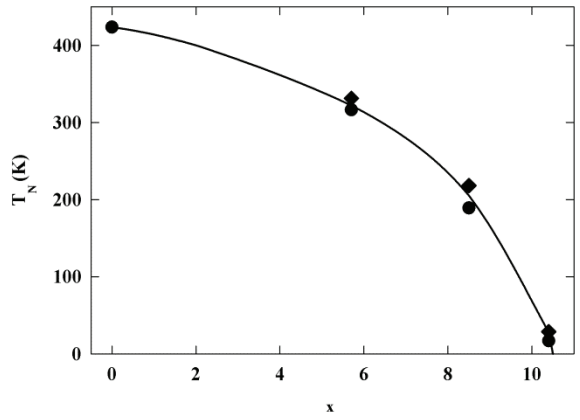


Figure 5. The magnetic phase diagram for the $(Cr_{84}Re_{16})_{100-x}V_x$ alloy system as a function of the V concentration, x, showing T_N values obtained through electrical resistivity (●) and magnetic susceptibility (◆) measurements. The error in the T_N values falls within the size of the data points. The solid line is a guide to the eye.

Figure 5 shows the T_N values, obtained from the electrical resistivity and magnetic susceptibility measurements, as a function of V concentration for the samples investigated. The error in the T_N values falls within the size of the data points. The solid line is a guide to the eye through the data points. Extending the present results (see the figure for $x \geq 10.4$ at.% V) it seems that the SDW is suppressed down to 2 K at $x \approx 10.5$ at.% V. This suggests a possible quantum critical point at $x \approx 10.5$ at.% V. Measurements of additional physical properties are necessary in order to clarify this projection.

4. Conclusion.

The present study was devised to investigate the way in which antiferromagnetic ordering could be suppressed in the $(Cr_{84}Re_{16})_{100-x}V_x$ system. Our results are conducive to an interpretation of concentration-dependent magnetic ordering in this system, specifically through the addition of V as an electron dopant into the CrRe pseudo-binary alloy, and hence we conclude that a first and important ingredient of putative quantum criticality in this itinerant electron Cr system has been satisfied. Investigations are in progress to include more alloy concentrations, as well as, for example, Hall coefficient and magnetic susceptibility studies in order to explore the notion of criticality in this alloy system. These are frequently used parameters [15, 20, 21] used in assessing quantum critical behaviour in Cr alloy systems. Furthermore, measurements to temperatures below 2K are highly desirable for investigating the possibility of a superconducting dome at very low temperatures in the phase diagram.

Acknowledgements

This work was supported by the NRF of South Africa under Grant No. 2072956 and 61388.

References

[1] Fawcett E, Alberts HL, Galkin VY, Noakes DR and Yakhmi JV 1994 *Rev. Mod. Phys.* **66** 25
[2] Trego AL and Mackintosh AR 1968 *Phys. Rev.* **166** 495
[3] Arajs S, Moyer CA and Abukay D 1980 *Phys. Stat. Sol. (b)* **101** 63
[4] Boshoff AH, Alberts HL, Du Plessis P and Venter AM 1993 *J. Phys.: Condens. Matter* **5** 5353
[5] Muheim S and Müller J 1964 *Phys. Kondens. Matter* **2** 377
[6] Nishihara Y, Yamaguchi Y, Kohara T and Tokumoto M 1985 *Phys. Rev.* B **31** 5775
[7] Damaschke B and Felsch W 1986 *Z. Phys.* B **63** 179
[8] Nishihara Y, Yamaguchi Y, Tokumoto M, Takeda K and Fukamichi K 1986 *Phys. Rev.* B **34** 3446
[9] Kohara T, Asayama K, Nishihara Y and Yamaguchi Y, 1984 *Solid State Commun.* **49** 31
[10] Alberts HL, McLachlan DS, Germishuyse T and Naidoo M 1991 *J. Phys.: Condens. Matter* **3** 1793
[11] Reddy L, Alberts HL, Prinsloo ARE and Venter AM 2006 *J. Alloys Comp.* **426** 83
[12] Reddy L, Alberts HL, Strydom AM, Prinsloo ARE and Venter AM 2008 *J. Appl. Phys.* **103** 07C903-1
[13] Yeh A, Soh Y-A, J Brooke J, Aeppli G, Rosenbaum TF and Hayden SM 2002 *Nature* **419** 459
[14] Lee M, Husmann A, Rosenbaum TF and Aeppli G 2004 *Phys. Rev. Lett.* **92** 187201
[15] Jaramillo R, Feng Y, Wang J and Rosenbaum TF 2010 *Proc. Nat. Acad. of Sci. USA (PNAS)* **107** 13631
[16] Sheppard CJ, Prinsloo ARE, Alberts HL and Strydom AM 2011 *J. Appl. Phys*. **109**(7) 07E104
[17] Quantum Design Inc, 6325 Lusk Boulevard, San Diego, USA
[18] Nishihara Y, Yamaguchi Y, Waki S and Kohara T 1983 *J. Phys. Soc. Japan* **52**(7) 2301
[19] Arajs A, Kote G, Moyer CA, Kelly JR, Rao KV and Anderson EE 1976 *Phys. Stat. Sol. (b)* **74** K23
[20] Yeh A, Soh Y, Brooke J, Aeppli G and Rosenbaum TF 2002 *Nature* **419** 459
[21] Paschen S 2006 *Physica* B **378-380** 28

The surface orientation dependence of the pre-exponential factors extracted from the segregation profiles of a Cu(111/110) bi-crystal

C J Jafta[1,2], W D Roos[1], J J Terblans[1], J K O Asante[3] and K I Ozoemena[2]

[1]Department of Physics, University of the Free State, Bloemfontein 9301, South Africa
[2]Energy and Processes Unit, Materials Science and Manufacturing, Council for Scientific and Industrial Research (CSIR), Pretoria 0001, South Africa
[3]Department of Physics, Tshwane University of Technology, Pretoria 0001, South Africa

E-mail: cjafta@csir.co.za

Abstract. Previous experimental investigations have only shown, without explanation, that the pre-exponential factor (D_0), in the diffusion coefficient of Sb segregating in Cu, is dependent on the surface orientation of a crystal. In this study, the surface concentration of Sb in a Sb doped Cu(111/110) bi-crystal was measured using a method combining Auger Electron Spectroscopy (AES) and linear temperature heating. Segregation parameters, including the D_0's are extracted from the experimental data of surface concentration versus temperature using the modified Darken model. The difference in the two pre-exponential factors D_0 (Sb in Cu(111)) and D_0 (Sb in Cu(110)) is explained thermodynamically in terms of entropy change ΔS that is calculated, for the first time, for a Cu(111/110) bi-crystal.

1. Introduction

Bulk-to-surface segregation studies are often carried out to determine the bulk diffusion coefficients (D) of the segregating species using the techniques Auger Electron Spectroscopy (AES), Low Energy Electron Diffraction (LEED), Ion Scattering Spectroscopy (ISS), etc. [1 – 12]. These diffusion and segregation studies of elements in alloys play a major role in the industrial treatment of alloys.

The need to develop improved catalysts for use in connection with environmental protection and the creation of viable alternative energy systems have led to an increasing use of metal alloys as heterogeneous catalysts, in which surface concentration plays a key role in controlling such important factors as activity and selectivity. Segregation in alloys can be used to engineer the properties of a catalytic surface by allowing the active elements to segregate to the surface [13]. Therefore an understanding of these surface phenomena requires knowledge of not only the structure of the surface, but also of the surface composition. The surface composition, in turn, depends on the surface segregation thermodynamics.

One of the parameters that the surface composition depends on is the diffusion coefficient. The pre-exponential factor (D_0), in the diffusion coefficient, depends on the formation and migration entropies. In this study a Cu bi-crystal was doped with Sb and the segregation behaviour used to measure the differences in D_0 for different surface orientations.

2. Theory

It is shown [14] that the diffusion coefficient can be written as:

$$D = f_c \frac{\rho \gamma_A d^2}{6} \tag{1}$$

where f_c is a factor correlating subsequent atomic jumps ($f_c \leq 1$), ρ a probability factor, γ_A the atomic jump frequency and d the jump distance.

For vacancy diffusion ρ is the probability that there is a neighbouring vacancy available to mediate the jump.

The temperature dependence arises from two distinct thermodynamic factors:

- The vacancy formation energy term (it cost thermal energy to create the vacancy).
- The atomic migration energy term (it cost thermal energy to transfer an atom from one site to another).

The Gibbs free energy of vacancy formation is:

$$\Delta g_v = \Delta h_v - T\Delta s_v \tag{2}$$

where Δh_v is the formation enthalpy, T the temperature and Δs_v the formation entropy.

The probability (ρ) of finding a vacancy adjacent to an atom is:

$$\rho = \exp(-\Delta g_v / k_b T) \tag{3}$$

The migration free energy (Δg_m) can be decomposed into enthalpy and entropy terms as:

$$\Delta g_m = \Delta h_m - T\Delta s_m \tag{4}$$

The probability of having sufficient energy to surmount the energy barrier occurs with a frequency proportional to the Boltzmann factor, $\exp(-\Delta g_m / k_b T)$.

The jump frequency of the atom (γ_A) is now equal to the product of the atomic vibrational frequency (v_A) and the probability of executing a jump:

$$\gamma_A = v_A \exp(-\Delta g_m / k_b T) \tag{5}$$

and therefore the diffusion coefficient D can be expanded and expressed as:

$$D = v_A f_c \frac{d^2}{6} \exp[(\Delta s_m + \Delta s_v)/k_B] \times \exp[-(\Delta h_v + \Delta h_m)/k_B T] \tag{6}$$

$$= D_0 \exp[-Q/k_B T]$$

with

$$D_0 = v_A f_c \frac{d^2}{6} \exp[(\Delta s_m + \Delta s_v)/k_B] \tag{7}$$

which is temperature independent and also known as the pre-exponential factor, and

$$Q = \Delta h_v + \Delta h_m \tag{8}$$

also known as the activation energy.

It is clear from equation (7) that, for the same substrate, the D_0's will be determined by the jump distance and vacancy formations entropy terms.

The D_0 for each surface orientation (xxx) can be written as:

$$D_0^{(xxx)} = v_A f_c \frac{d^2_{(xxx)}}{6} \exp\left[\left(\Delta s_m + \Delta s_v^{(xxx)}\right)/k_B\right] \quad (9)$$

Using equation (9) for both orientations and a few mathematical manipulations the following equation is derived:

$$\Delta S_{\Delta s} = \Delta s_v^{(111)} - \Delta s_v^{(110)} = k_B\left[\ln\left(\frac{D_0^{(111)}}{D_0^{(110)}}\right) - \ln\left(\frac{d^2_{(111)}}{d^2_{(110)}}\right)\right] \quad (10)$$

where $\Delta s_v^{(111)} - \Delta s_v^{(110)}$ is the difference in the formation entropies ($\Delta S_{\Delta s}$), which can be calculated if the D_0's are known.

3. Experimental

Electron beam physical vapour deposition was used to deposit a thin layer (≈ 20 kÅ) of Sb on the back face of a Cu(111/110) bi-crystal[15], which was purchased from MaTeck GmbH in Germany. The crystal was annealed in a custom build annealing system as described in reference [16]. The homogeneously distributed Sb in the crystal was 0.1 at%. A unique experimental setup and measuring procedure [17] was used to determine the concentration of the segregant as a function of temperature.

A Perkin Elmer Auger spectrometer with a PHI 15-110B single pass cylindrical mirror analyzer was used to measure the Sb surface composition on the bi-crystal. The primary electron beam was alternatively deflected from one orientation to the other as the temperature was ramped and the surface compositions of both orientations were recorded almost simultaneously. Two different linear temperature rates of 0.1 K.s^{-1} and 0.075 K.s^{-1} were used.

The setup ensures the same experimental conditions for both orientations, allowing direct comparison of the segregation data on two different surface orientations. A modified Darken segregation model was used to fit the measured segregation profiles where the segregation parameters were extracted.

4. Results and Discussion

Figure 1 shows the experimental results and segregation model fits for Sb segregation in the Cu(111/110) bi-crystal at a linear temperature ramp of 0.1 K. s^{-1}.

The pre-exponential values used in the segregation model, for the best fits, are summarized in table 1.

From equation (7) it is clear that D_0 is a function of the change in entropy ($\Delta s_m + \Delta s_v$), which includes the terms for, (i) the change in entropy for the migration of an atom in the bulk (Δs_m) and (ii) the change in entropy for vacancy formation (Δs_v). The change in entropy for the migration of an atom in the bulk is assumed to be the same for the (111) and (110) surfaces. The change in entropy for vacancy formation, on the contrary, is different for the two surfaces due to the number of vacancies in the surfaces which results in different pre-exponential factors. It is now possible to calculate the difference in the formation entropies ($\Delta S_{\Delta s}$) between the two surfaces.

Using the parameters in table 1 the difference in formation entropies between (111) and (110) orientations in Cu is calculated as 4.32 k_B ($= 3.7 \times 10^{-4}$ eV.K^{-1}.atom^{-1}).

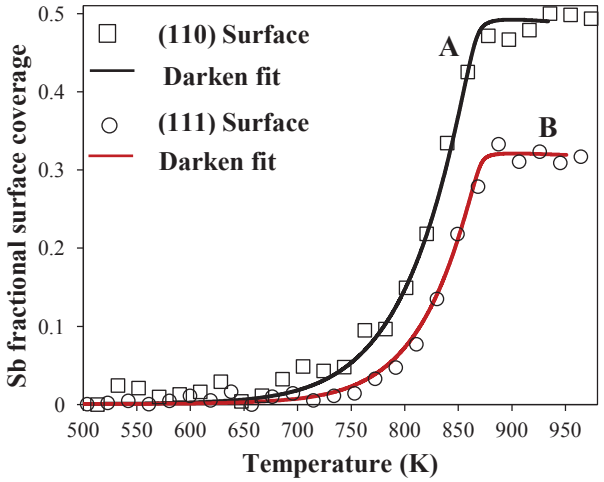


Figure 1. The measured Sb segregation profiles in (A) Cu(110) and (B) Cu(111) at a heating rate of 0.1 K. s^{-1} and their corresponding modified Darken fits.

Table 1. A summary of the pre-exponential factors experimentally extracted.

Orientation	Rate ($K.s^{-1}$)	D_0 ($m^2.s^{-1}$) ($\pm 5\%$)	Average D_0 ($m^2.s^{-1}$)
Cu(110)	0.075	2.3×10^{-5}	2.2×10^{-5}
	0.1	2.0×10^{-5}	
Cu(111)	0.075	1.1×10^{-3}	1.1×10^{-3}
	0.1	1.1×10^{-3}	

5. Conclusion

Pre-exponential factors for a Sb doped Cu(111/110) bi-crystal was extracted for the first time. The experimental D_0's measured in this study are in the same range as were calculated and reported by Wynblatt [18].

Comparison of the theoretical calculations of vacancy formation entropy of copper, evaluating the effect of the surface, shows clearly an influence on the formation entropy and that the presence of the surface decreases the entropy approximately by 0.4 k_B [18].

From the pre-exponential factors $D_0^{(110)} = 2.2 \times 10^{-5}$ $m^2.s^{-1}$ and $D_0^{(111)} = 1.1 \times 10^{-3}$ $m^2.s^{-1}$ the difference in the formation entropies was found to be 4.32 k_B. Although the atom jump distance, $d_{(111)}$ and $d_{(110)}$, also contribute to the change in pre-exponential factors, it was demonstrated thermodynamically that the difference is mainly due to the difference in vacancy formation entropy.

Acknowledgements

The authors would like to thank the NRF for financial assistance.

References

[1] Du Plessis J and Viljoen E C 1992, *Appl. Surf. Sci.* **59** 171
[2] Hoffman S and Erlewein J 1978 *Surf. Sci.* **77** 591
[3] Du Plessis J and Viljoen E C 1996 *Appl. Surf. Sci.* **100-101** 222
[4] Asante J K O, Terblans J J and Roos W D 2005 *Surf. Int. Anal.* **37** 517
[5] Giordano H and Aufray B 1996 *Surf. Sci*. **352-354** 280
[6] Jablonski A, Overbury S H and Somorjai G A 1977 *Surf. Sci.* **65** 578
[7] Anton R, Eggers H and Veletas J 1993 *Thin Solid Films* **226** 39
[8] Swartzfager D G, Ziemecki S B and Kelley M J 1981 *J. Vac. Sci. Technol.* **19** 185
[9] Kuntze J, Speller S, Heiland W, Deurinck P, Creemers C, Atrei A and Bardi U 1999 *Phys. Rev.* B **60** 9010
[10] Kuntze J, Speller S, Heiland W, Atrei A, Rovida G and Bardi U 1999 *Phys. Rev.* B 60 1535
[11] Aschoff M, Speller S, Kuntze J, Heiland W, Platzgummer E, Schmid M, Varga P and Baretzky B 1998 *Surf. Sci.* **415** L1051
[12] Koel B E, Sellidj A and Paffett M T 1992 *Phys. Rev.* B **46** 7846
[13] Loukrakpam R, Luo J, He T, Chen Y, Xu Z, Njoki P N, Wanjala B N, Fang B, Mott D, Yin J, Klar J, Powell B, and Zhong C 2011 *J. Phys. Chem.* C **115** 1682
[14] Elliot S and Elliot S R 1998 *The Physics and Chemistry of Solids* (John Wiley and Sons)
[15] Jafta C J, Roos W D and Terblans J J May 2011 *Measuring segregation on a Cu bi-crystal using AES* submitted for publication to *Appl. Surf. Sci. (ACSIN 11 conference series)*
[16] Jafta C J, Roos W D, Terblans J J and Hugo A B 2010 *Instr. Sci. Technol.* **38** 261
[17] Jafta C J, Roos W D and Terblans J J May 2011 Sb *Segregation in a Cu(111/110) bi-crystal* submitted for publication to *Appl. Surf. Sci.* (*ACSIN 11 conference series*)
[18] Wynblatt P 1969 *J. Phys. Chem. Solids* **30** 2201

A positron annihilation study of F and Ba sub-lattices in superionic BaF_2 at elevated temperatures

T P Jili[1,2], E Sideras-Haddad[2], D Wamwangi[2], F Tuomisto[3] and B Kibirige[1]

[1] Department of Physics, University of Zululand, P/B X1001, Kwa-Dlangezwa, 3886, South Africa

[2] Department of Physics, University of the Witwatersrand, Johannesburg, PO Wits 2025, South Africa

[3]Department of Applied Physics, Aalto University, P.O. Box 11100, Otakaari 1 M, FI-02150, Espoo, Finland

E-mail : tjili@pan.uzulu.ac.za

Abstract. There is a general misunderstanding regarding the creation of Frenkel pairs in the vicinity of the observed critical temperature T_c in superionic BaF_2. The measured conductivity increases sharply at a temperature coinciding with a deviation of the temperature-dependent lattice constant from 0.62056 nm Frenkel pairs responsible for superionic conduction are gradually generated well below the critical temperature and this is informed by the calculation of the S-parameter through the measurements of Doppler broadening at various temperatures. It is interesting to note that the lattice constant plays a pivotal role in the superionic conductivity threshold. Positron annihilation spectroscopy, through the determination of positron lifetime and Doppler broadening, reveals that the generation rate of Frenkel pairs becomes prominent 113 K well below the critical temperature of 693 K. This is also a clear indication of a continuous disordering of F sub-lattice noticeable at a temperature of 580 K. The fact that the defect positron lifetime is constant in the temperature range 300K to 900K, confirms a non-distortion of Ba sub-lattice.

1. Introduction

Flourites, including BaF_2, are class-II ionics in which the transition to a highly ionic conducting material is not accompanied by any structural change in the crystal lattice but associated with an anomaly in the heat capacity. The diffuse transition temperature, T_d, is associated with the temperature well below the melting point at which a peak in the heat capacity occurs [1, 2] and at which negative deviation from the Arrhenius behaviour occurs [3, 4].

Upon entering a sample, positrons have several states associated with different kinds of defects, each of which gives a characteristic lifetime. The positron lifetime spectrum is thus a sum of exponential decay components given by

$$-\frac{\mathrm{d}n(t)}{\mathrm{d}t} = \sum_i I_i \lambda_i \mathrm{e}^{-\lambda_i t} \quad (1)$$

where I_i are relative intensities, $n(t)$ is the probability that a positron is still alive at time t after its emission from the source and λ_i are the annihilation rates at various positron states. Doppler broadening of the annihilation line shape, characterized by the S-parameter, which is defined as the area under the central part of the annihilation photopeak divided by the total area [5], is a measure of the electron momentum at the annihilation site. If the vacancy charge is positive, and is the only defect-type, we would be observing Frenkel pairs, rather than Schottky defects, since only one type of interstitials are formed.

The objective of the present work is to determine the temperature value at which interstitials responsible for superionic conductivity, specifically for BaF_2 crystal, are created, and to investigate the distortion of both F and Ba sub-lattices in the temperature range of 300 K to 900 K..

2. Experiment

The positron source, with activity of 10.5 μCi, was achieved by evaporating a few drops of sodium chloride solution on a thin (7 μm) nickel foil. The salt was then covered by another identical nickel foil. The foiled source was then sandwiched between two identical silicon BaF_2 crystals (each of size 10 mm×8 mm ×2 mm). Two high purity germanium detectors with energy resolution of 1.2 keV full width at half maximum (FWHM) at 511 keV were used to obtain Doppler broadening profiles of the annihilation radiation. The samples were annealed at different selected temperatures up to 400 K.

The lifetime measurement was carried out using a standard coincidence setup by employing two fast scintillator detectors (XP2020) for the start (1275 keV) and stop (511 keV) signals. Samples were kept under the pressure of 10^{-5} Torr. The time resolution of the positron lifetime coincidence setup used for the measurements was of the order of 220 ps at FWHM. The 1.27 MeV γ-ray should not exhibit any broadening, as shown in figure 1, since it is not associated with the annihilation but with the birth of a positron during a β^+ -decay of ^{22}Na into ^{22}Ne. The peak position of 1.27 MeV γ-ray should be used as a measure of how stable the electronics are, as shown in figure 2.

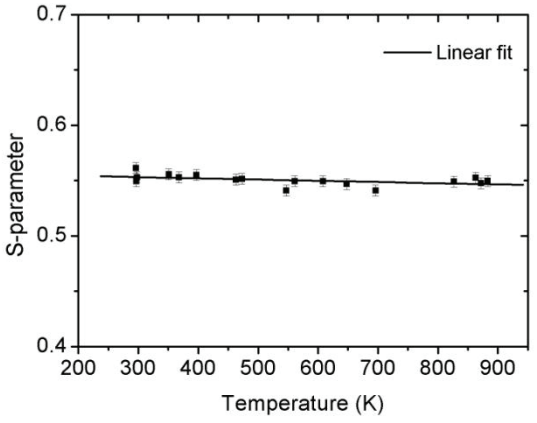


Figure 1. The temperature dependence of the 1.27 MeV gamma ray S-parameter does not exhibit any Doppler broadening variation

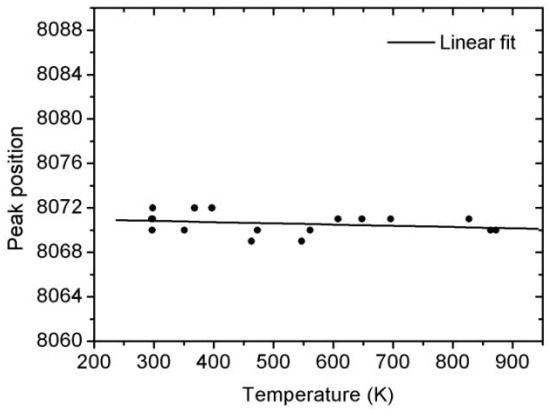


Figure 2. The temperature dependence of the measured peak positions indicate the stability of the electronics used

The positron annihilation spectra (5×10^5) were accumulated over a temperature range from 300 K to 900 K. The vacancy-type defects act as trapping sites for positrons and annihilation with low energy

valence electrons at these defects results in a narrowing of the photopeak corresponding to an increase in the S-parameter. Since a fraction of positrons, 10% to 15 %, annihilate in the source material and source-sample interfaces (foils), it is crucial to subtract these additional lifetime components from the lifetime spectrum. Nickel foil has a lifetime component of 180 ps and the salt NaCl crystal introduces a lifetime of 450 ps. For the foil intensities, the Bertolucci-Zappa empirical formula [6]

$$I_{foil}(\%) = k = 0.324 Z^{0.93} \times \Omega^{\alpha} \qquad (2)$$

was utilized to obtain reliable values of lifetimes and intensities in a sample, where Z is the average atomic number of the sample, Ω the surface density of the foil (thickness of the foil multiplied by volume density of the foil) in mg cm^{-2} and $\alpha = 3.45 \times Z^{-0.4}$.

3. Results and discussion

The central parameter S is defined as the ratio of the counts in the central region of the annihilation line to the total number of counts under the annihilation curve. From figure 3, it is clear that although there is no appreciable increase in S-parameter from room temperature to about 598 K (~ΔS = 0.001), there is a sharp increase of thermal vacancies between 598 K and 877 K (~ΔS = 0.024). However, defect lifetime component from room temperature to 700 K, shown in figure 4, remains constant, indicating that only positive defects, which are not visible to probing positrons, are created. If negative Ba vacancies were created, non constant defect lifetime components would have been observed, since positrons would have been trapped in negative vacancies. Therefore no negative Ba vacancies were observed in the temperature range 300K to 900 K.

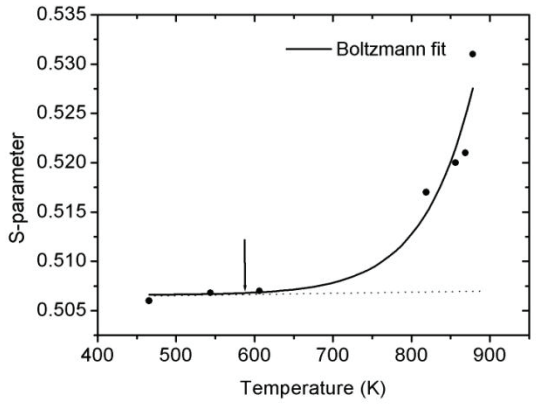


Figure 3. S-parameter versus temperature indicates that the generation of Frenkel defects responsible for superionic region commence earlier at about 580 K

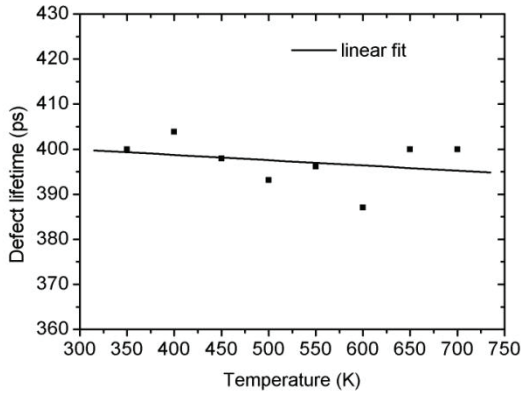


Figure 4. Second positron lifetime component confirms the non-distortion of Ba sub-lattice over a temperature range.

The S-parameter and the behaviour of superionic current at different temperatures confirms that although F sub-lattice distorts, superionic current, mainly due to high mobility of F ions, cannot be established until an appreciable increase of lattice constant reaches at least 0.62056 nm. The linear expansion of the lattice constant a reported in [7] is given by

$$a = a_m[1 + b(T - T_m)] \text{ Å} \qquad (3)$$

where $a_m = 0.620$ nm, $b = 2.5 \times 10^{-5}$ K^{-1} and $T_m = 293.2$ K

Accurate neutron diffraction studies [8] indicate more reliable temperature dependence of the lattice constant at high temperatures as shown in figure 7.

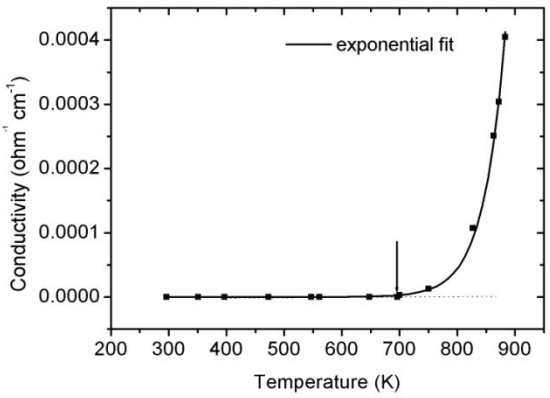


Figure 6. Measured conductivity as a function of temperature. Conductivity enters superionic region at about 693 K.

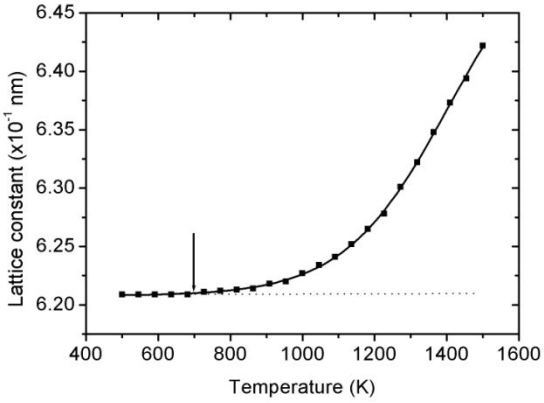


Figure 7. The dependence of BaF_2 lattice constant on temperature curve [8]

The S-parameter threshold at threshold temperature of 580 K, shown in figure 3, indicates that Frenkel defects are generated 113 K below the critical temperature. The threshold temperature (693 K) at which the ionic current enters superionic region (see figure 6) is observed to agree very well with the temperature at which the lattice constant begins to increase beyond 0.62056 nm as shown in figure 7.

4. Conclusion

Positron annihilation lifetime measurements reveal the existence of the distortion of F sub-lattice in the temperature range 300 K to 900 K and that Ba sub-lattice stays intact for this temperature range. ΔS in the range 300 K to 600 K, represents only 4 % of the total change in the S-parameter which suggests a non drastic variation in defect creation. In the temperature range from 600 K to 900 K, the 96 % change in the S-parameter is a clear indication of the enhanced interstitial space as indicated by both lattice constant curve and conductivity curve. The creation of defects in BaF_2 starts at 580 K, a temperature 113 K below the transition temperature, T_c (693 K). At the transition temperature, T_c, the generation of defects responsible for superionic region becomes prominent mainly due to the variation of lattice constant with respect to temperature. Therefore a conclusion can be made that defects responsible for superionic conduction are created well below a transitional temperature, T_c.

Acknowledgement

The authors wish to thank iThemba LABS (Gauteng) for the facilities, the University of Zululand and NRF (National Research Foundation) for their financial support. We would also like to thank Positron Laboratory staff at Aalto University in Finland for sharing with us their knowledge and positron lifetime equipment.

References

[1] Schroter W 1979 *PhD Dissertation* University of Gottingen
[2] Belosludov V R, Efremova R I and Matizen E V 1974 *Fiz. Tela* Leningrad **16**, 1311
[3] Boyce J B and Huberman B A 1979 *Phys. Rep*. **51**, 189
[4] Oberschmidt J and Lazarus D 1980 *Phys. Rev*. B **21**, 5823

[5] Krause-Rehberg R and Leipner H S 1999 *Positron Annihilation in Semiconductor* Springer-Verlag
[6] Bertolaccini M and Zappa L 1967 *Nuovo Cimento* B **52** 487
[7] Chadwick A V 1983 *Solid State Ionics* **8** 209
[8] Aityan S K, Ivanov-Shitz A K, Buchstab A S and Loshmanov A A 1990 *Superionic State of BaF2:Molecular dynamic simulation and neutron diffraction study* N1536-B90 Moscow VINIT 1-37

Synthesis and Characterization of $Ba_xMg_yAl_2O_4$: Eu,Dy nanophosphors Prepared Using Solution - Combustion Method

M. A. Kebede[1,2] and F.B. Dejene[1*]

[1]Department of Physics, University of the Free State (Qwaqwa Campus), Private Bag X13, Phuthaditjhaba, 9866, South Africa
[2]Energy and Processes, Materials Science and Manufacturing, Council for Scientific and Industrial Research, P.O.Box 395, Pretoria, 0001, South Africa

E-mail: dejenebf@ufs.ac.za

Abstract. Europium-doped barium magnesium aluminate ($Ba_xMg_yAl_2O_4$:Eu) phosphors were obtained at low temperature using the solution - combustion of corresponding metal nitrate–urea solution mixtures. The particle sizes, morphology, structural and luminescent properties of the as-synthesized phosphors were examined by means of scanning electron microscopy (SEM), X-ray diffraction (XRD) and photoluminescence (PL). It was found that the change in Ba:Mg molar ratios showed great influence not only on the particle size and morphology, but also on their PL spectra and crystalline structure. The structure of $Ba_xMg_yAl_2O_4$ nanophosphors changes from a monoclinic structure for Ba 26 mole% to orthorhombic for Mg 26 mole% and a small change in peak position at high angles due to differences in size between Ba and Mg ions. The peak of the emission band occurs at longer wavelength (around 615nm) with increase in Mg concentration but display a broad band emission at 498 nm for lower Mg concentration. The blue-green emission is probably due to the influence of 5d electron states of Eu^{2+} in the crystal field because of atomic size variation causing crystal defects while the red emission is due to f-f transitions. This finding clearly demonstrate the possibility of fine tuning the colour emission and solid solubility limit in $Ba_xMg_yAl_20_4$:Eu phosphors through the simple and a cheap process.

1. Introduction

Luminescent phosphors have gained abundant interest in their production for their potential applications in the development of different luminescent display systems [1–6]. Phosphors based on silicates, aluminates, germinate and other related oxides are of more interest nowadays. Aluminates generally generate more defect-related traps when they are doped with rare earth ions. Several aluminate compositions are investigated and used as photoluminescence (PL), catholuminescence and plasma display panel phosphors for their high quantum efficiency in the visible region. Eu^{2+}-doped phosphors usually show intense broad band PL with a short decay time of the order of tens of nanoseconds [7]. The emission of Eu^{2+} is strongly dependent on the host lattice and can occur from the ultraviolet to the red region of the electro-magnetic spectrum [8-9]. This is because the 5d↔4f transition is associated with the change in electric dipole and the 5d excited state is affected by crystal

* To whom any correspondence should be addressed.

field effects. Notably barium and strontium aluminates have been reported to be good host material. Accordingly, the property of phosphor depends on their crystal structure of the mother phases.

Compared with sulfide phosphorescent phosphors, $BaAl_2O_4$:Eu,Dy phosphor possesses safer, chemically stable, very bright and long-lasting photoluminescence with no radiation [10, 11]. The tuning of colors is controllable by the subtle structure modifications [12, 13]. For instance, the emissions of Eu^{2+}-doped MAl_2O_4 (M = Ba, Ca, Sr or Mg) are all in range of blue and green [14]. For long afterglow phosphors, it is well known that defect-related traps and trapping dynamics, which can be effectively adjusted by doping ions with various valences, play an important role in their luminescence and afterglow properties [14]. Also, controlling trap kinds and amount by co-doping ions is an efficient technique to study the long afterglow mechanism. In this paper, Mg^{2+} ion was used to study the co-doping effect on the luminescence and afterglow behavior of $BaAl_2O_4$:Eu,Dy phosphor. The grain size of phosphor powders prepared through solid-state reaction method is in several tens of micrometers. Phosphors of small particles must be obtained by grinding the larger phosphor particles. Those processes easily introduce additional defects and greatly reduce luminescence efficiency [15]. With the development of scientific technologies on materials, several chemical synthesis techniques, such as co-precipitation [16], sol–gel [17] and combustion synthesis methods [18], have been applied to prepare nano-sized $BaAl_2O_4$ and/or its phosphors. Among the various wet chemical routes, the solution combustion technique has been regarded as one of the effective and economic methods due to its convenient processing, simple experimental setup and significant time-saving and high purity products [19-21]. In the present work, luminescence properties of Eu^{2+} doped $BaAl_2O_4$–$MgAl_2O_4$ ternary system prepared by a solution–combustion process is reported through emission, excitation, lifetime and structural from X-ray diffraction (XRD), SEM.

2. Experiments

$Ba_xMg_yAl_2O_4:Eu^{2+}$ phosphor powders were prepared by a solutions combustion method without the use of flux material. $Ba(NO_3)_2$, $Mg(NO_3)_3{\cdot}9H_2O$, $Al(NO_3)_3{\cdot}9H_2O$, $Eu_2(NO_3)_3$ (4N) and CH_4N_2O were used as starting materials. The procedure used to prepare $Ba_xMg_yAl_2O_4$:Eu;Dy had the following stages. Firstly, $Al(NO_3)_3{\cdot}9H_2O$, $Mg(NO_3)_3{\cdot}9H_2O$, $Ba(NO_3)_2$, $Eu_2(NO_3)_3$ and $CO(NH_2)_2$ were dissolved into deionised water and stirred at ambient temperature for ½ hr to obtain transparent solution. After that, the precursor solution was introduced into a furnace preheated at 500°C and then white powders were obtained. To investigate the effect of the Ba/Mg molar ratios on the structural and PL properties of the $Ba_xMg_yAl_2O_4:Eu^{2+}$ phosphor, samples with Ba/Mg molar ratios of 0 to 26 mol % were prepared under atmospheric pressure without post heat treatment. The doping concentration of Eu and Dy were fixed at 1% and 2% of the Ba + Mg component, respectively. The solutions were prepared by dissolving various amounts of the nitrate precursors of each component into a minimum amount of distilled water. The resulting solution was transferred into a china crucible, which was then introduced into a muffle furnace maintained at 500 ^oC for 5-6 minutes. The voluminous and foamy combustion ash was easily milled to obtain the final $Ba_xMg_yAl_2O_4:Eu^{2+}$ phosphor powders. XRD patterns of as-synthesized samples were recorded on an x-ray diffractometer with Cu K_α = 1.5406 Å, which was operated at 40 kV voltage and 40 mA anode current. Data were collected in 2θ values from 20° to 80°. The morphologies and sizes of the particles were examined using a PHI 700 Nano Scanning Auger Microprobe (Nano SAM) and a Shimadzu model ZU SSX – 550 Superscan scanning electron microscope (SEM), coupled with an energy dispersive x-ray spectrometer (EDS). Photoluminescence (PL) measurements were performed at room temperature on a Cary Eclipse fluorescence spectrophotometer (Model: LS 55) with a built-in 150 W xenon lamp as the excitation source and a grating to select a suitable wavelength for excitation.

3. Results and discussion

The present investigation aims at examining the relationship between the structural, morphological, photoluminescence properties and the compositions in the Eu^{2+} doped $BaAl_2O_4$-$MgAl_2O_4$ ternary and

quaternary systems. In our experiment, the Eu^{2+} doping concentration was fixed 1.5 % (molar ratio of Ba +Mg ions) for all the compositions, so that any changes in properties considered to be originating from the change in Ba/Mg molar ratios. Fig. 1(a) and (b) show representative pictures of the SEM micrographs of the $Ba_xMg_yAl_2O_4:Eu^{2+}$ sample. It can be seen that the morphology of the synthesized powders is unconsolidated, reflecting the inherent nature of the combustion process (Fig. 1).

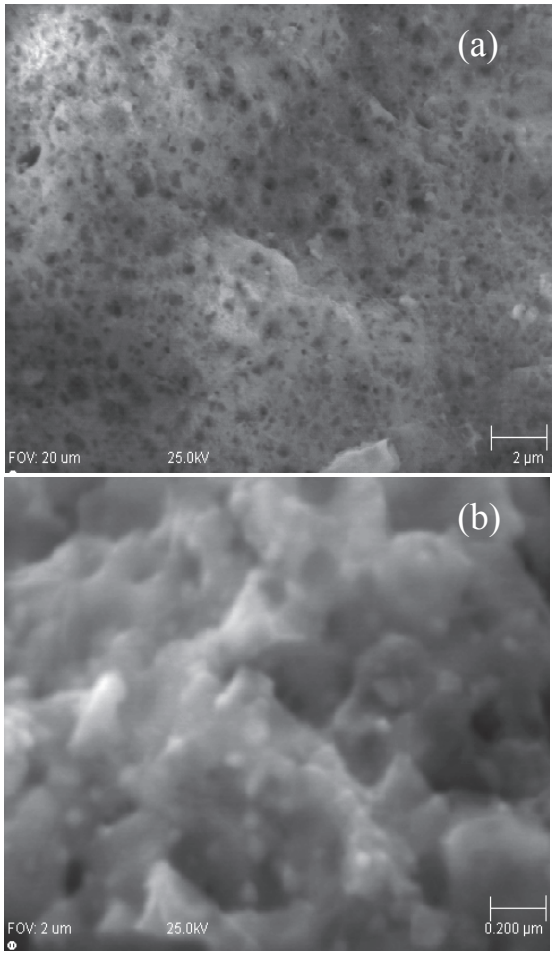


Figure 1 (a) and (b) are SEM image of $Ba_xMg_yAl_2O_4:Eu^{2+}$.

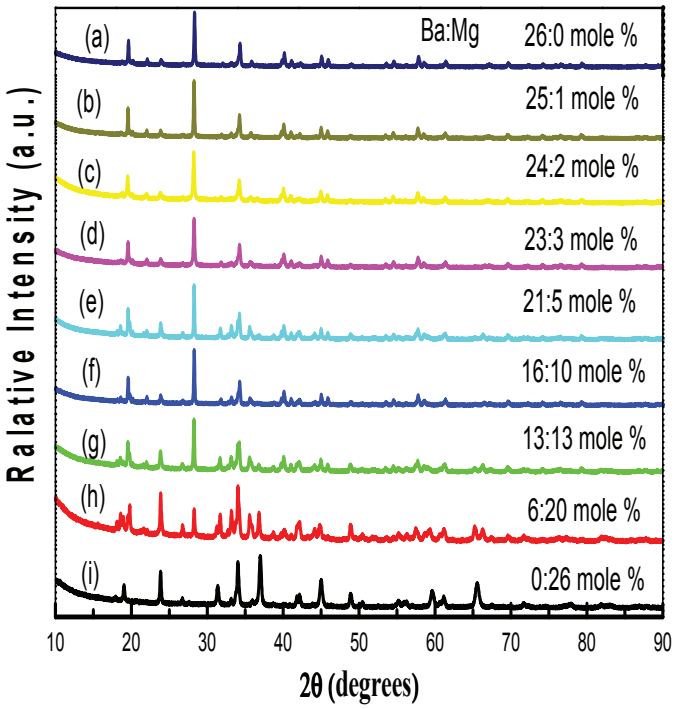


Figure 2 XRD patterns of $Ba_xMg_yAl_2O_4:Eu^{2+}$ Ba:Mg mole% (a)26:0 , (b)25:1, (c)24:2 , (d) 23:3, (e) 21:4, (f) 16:10, (g) 13:13, (h) 6:20, and (i) 0:26.

The surfaces of the foams show a lot of cracks, voids and pores formed by the escaping gases during combustion reaction. In order to achieve accurate data of the grain size of $Ba_xMg_yAl_2O_4:Eu^{2+}$ as-synthesized powders, its SEM image at high magnification recorded in Fig. 1(b). Most of the particles of $Ba_xMg_yAl_2O_4:Eu^{2+}$ phosphor powders appear to be irregularly elliptical with aggregation and an average particle size of 30–50 nm (Fig. 1(b)). It is observed that the presence of Mg ions has minimum significance on the morphology of phosphors. A typical XRD patterns recorded for the $Ba_xMg_yAl_2O_4:Eu^{2+}$ compositions with varying Ba/Mg concentrations were shown in Fig. 2. The examination of the diffraction patterns confirms that pure monoclinic phase $BaAl_2O_4$ (JCPDS: 82-2001) and pure orthorhombic $MgAl_2O_4$ (JCPDS:47-0254) are formed at Ba 26 mol% (Mg 0mol%) shown in Fig. 2(a) and Ba 0 mole%(Mg 26 mole%) displayed in Fig. 2(i), respectively. No other product or starting material was observed, implying that the phase composition of the synthesized powders are all low-temperature monoclinic phase for Ba 26mole % and orthorhombic for Mg 26mole% , and the little amount of doped rare earth ions have no noticeable change on the $BaAl_2O_4$ and $MgAl_2O_4$ phase composition [16]. Although no flux is added, $Ba_xMg_yAl_2O_4:Eu^{2+}$ phase with high purity can be obtained at 500°C through the solution-combustion process to the starting materials, whereas it is

impossible to happen for solid-state reaction method due to impurities are formed at lower temperatures.

The excitation spectra of $Ba_xMg_yAl_2O_4:Eu^{2+}$ powder phosphors are shown in Fig. 3(a) with strong excitation bands which could be attributed to a charge transfer band (CTB) of Eu^{3+}– O band in the short ultraviolet region (240 nm) [22,23] for high Mg concentration and it can be seen clearly that the excitation spectra consist of a broad band with a maximum at about 340 nm, which can be attributed to the $Ba_xMg_yAl_2O_4$ host excitation band for low Mg concentrations. The f–f transitions within the Eu^{3+}, $4f^6$ configuration in longer wavelength region with $^7F_0 \rightarrow {}^5D_4$ (361 nm), $^7F_0 \rightarrow {}^5L_7$, (379 nm) and $^7F_0 \rightarrow {}^5L_6$ (393 nm) as the most prominent group. Fig. 3(b) shows the PL emission spectra of BaxMgyAl2O4:Eu2+ powder phosphors excited at 240nm for low Mg concentration and 325 nm for high Mg concentration. The spectra tunes the emission colour from red to green as Ba:Mg concentration changes from 0:26 mole% to 26:0 mole%. The emission spectrum for Ba:Mg 0:26 mole% consists of strong and sharp peaks which are attributed to $^5D_0 \rightarrow {}^7F_J$ emission transitions, indicating the existence of Eu^{3+} ions in the $MgAl_2O_4:Eu^{2+}$ matrix. The spectra has emission peaks at 578, 592, 615, 656 and 690 nm which can be assigned to $^5D_0 \rightarrow {}^7F_J$ (J=0, 1, 2, 3 and 4) [24,25] transitions of Eu^{3+} ions namely the $^5D_0 \rightarrow {}^7F_0$ (578 nm), $^5D_0 \rightarrow {}^7F_1$ (592 nm), $^5D_0 \rightarrow {}^7F_2$ (615 nm), $^5D_0 \rightarrow {}^7F_3$ (656 nm), and $^5D_0 \rightarrow {}^7F_4$ (694 nm), respectively. Since 4f electrons are shielded by 5s and 5p electrons in the outermost shells of the europium ion, as a result narrow emission peaks are expected, consistent with the sharp and intense peak around 615 nm and is due to the $^5D_0 \rightarrow {}^7F_2$ transition, based on selection rules [26,27]. The intensity for $^5D_0 \rightarrow {}^7F_2$ red (615 nm) emission is much stronger than that of the orange 592nm emission.

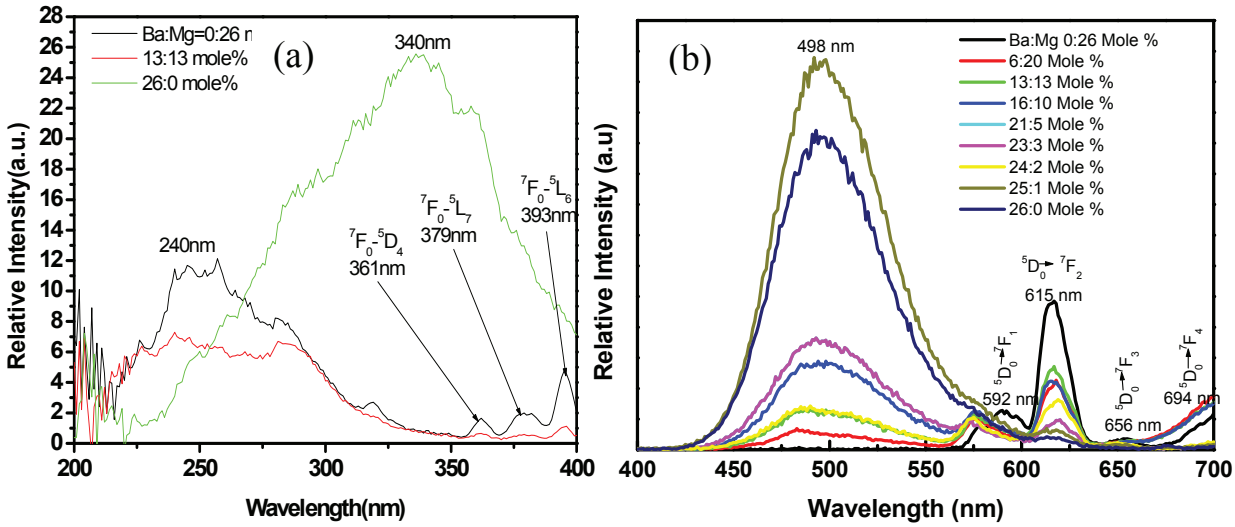


Figure 3 (a) Excitation spectra, and (b) Emission spectra of $Ba_xMg_yAl_2O_4:Eu^{2+}$ powder phosphors.

On the other hand, the symmetrical band blue-green emission at 498 nm for Ba:Mg 26:0 mole% is attributed to the typical transition between the ground state ($4f^7$) and the excited state ($4f^6\ 5d^1$) of Eu^{2+} ions [28]. They compare very well with the results of Lu et al. [29].

The Luminescent decay curve of $Ba_xMg_yAl_2O_4$:Eu,Dy phosphor monitored at 615 nm for different Ba:Mg mole ratio was shown in Fig. 4. The decay behaviour analysed using curve fitting [30], relying on the following triple exponential equation:

$$I = I_1 \exp(\frac{t}{\tau_1}) + I_2 \exp(\frac{t}{\tau_2}) + I_3 \exp(\frac{t}{\tau_3}) \qquad (1)$$

where I represents the phosphorescent intensity; I_1, I_2 and I_3 are constants; t is the time; τ_1, τ_2 and τ_3 are the decay constants, deciding the decay rate for the fast, medium and slow exponentially decay components, respectively.

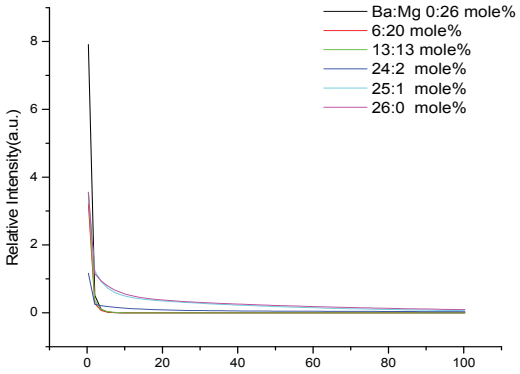


Figure 4 Afterglow characteristics of $Ba_xMg_yAl_2O_4$:Eu,Dy powder

Ba:Mg mole ratio	0:26	6:20	13:13	24:2	25:1	26:0
Component	Decay Constants (s)					
Fast(τ_1)	1.20	0.57	1.73	0.39	0.21	0.31
Medium(τ_1)	0.35	0.22	0.63	9.14	3.33	59.58
Slow (τ_1)	0.55	1.48	0.21	127.86	47.41	4.97

Table 1 Results for fitted decay curves of the phosphor powders with different Mg:Ba molar ratios

The fitting results of parameters of τ_1, τ_2 and τ_3 are listed in Table 1. Three components namely slow, medium, and fast component are responsible for the persistent luminescence from the synthesized phosphor.

4. Conclusion

In conclusion, excellent tuning green-red $Ba_xMg_yAl_2O_4$:Eu2+ nanophosphors were successfully synthesized at low temperature 500˚C using the solution - combustion with optimized condition Eu^{2+} doping concentration fixed as 1.5 % (molar ratio of Ba +Mg ions). XRD data was analyzed for the phase stability and crystallinity and observed that pure monoclinic phase diffraction peaks of $BaAl_2O_4$ are predominant for low Mg concentration and orthorhombic $MgAl_2O_4$ for low Ba mole concentration. The synthesized $Ba_xMg_yAl_2O_4$ phosphor doped with Eu^{2+} ions can be efficiently excited by 325nm UV radiation for low Ba concentration and by 240nm shorter UV for low Mg concentration and gives dominant emission band peaking at 615nm(red) and 498nm(green), respectively. The nature and structural details of this phosphor have been investigated from the measurement of XRD and SEM. Based on the emission analysis results; we suggest that it is a novel tuning green-red colors emitting optical material with potential luminescent applications.

References

[1] L. Zhou, W.C.H. Choy, J. Shi, M. Gong and H. Liang 2006 Mater. Chem. Phys. **100** 372
[2] B. Yan and X.Q. Su 2007 Opt. Mater. **29** 547
[3] W.B. Im, Y.I. Kim, J.H. Kang and D.Y. Jeon 2005 Solid State Commun. **134** 717
[4] Y.Q. Li, C.M. Fang, G. de With and H.T. Hintzen 2004 J. Solid State Chem. **177** 4681
[5] M. Yu, J. Lin, J. Fu, H.J. Zhang and Y.C. Han 2003 J. Mater. Chem. **13** 1413.
[6] M. Leskel and L. Niinist 1992 Mater. Chem. Phys. **31** 7
[7] G. Blasse, B.C. Grabmaier, Luminescent Materials, Springer, Berlin, 1994.
[8] T. Matsuzawa, Y. Aoki, M. Takeuchi and Y. Murayama ‘‘A New Long- Phosphorescent Phosphor with High Brightness, SrAl2O4;Eu2+,Dy3+’’; p. 160 in *Proceedings of 188th Meeting of the Electrochemical Society* (Chicago, IL, Oct. 8–13, 1995).
[9] Y. Murayama 1996 Sci. Am. Jpn., **26** 20
[10] Y. Lin, Z. Zhang, F. Zhang, Z. Tang and Q. Chen 2000 Mater. Chem. Phys. **65** 103

[11] W. Jia, H. Yuan, L. Lu, H. Liu and W.M. Yen 1998 J. Lumin. **76–77** 424
[12] X.Q. Piao, K.-I. Machida, T. Horikawa, H. Hanzawa, Y. Shimomura and N. Kijima 2007 Chem. Mater. **19** 4592
[13] Y.Q. Li, A.C.A. Delsing, G.D. With and H.T. Hintzen 2005 Chem. Mater. **17** 3242
[14] F.C. Palilla, A.K. Levine and M.R. Tomkus 1968 J. Electrochem. Soc. **115** 642
[15] R.P. Rao 1996 J. Electrochem. Soc. **143** 189
[16] Y. Lin, Z. Zhang, F. Zhang, Z. Tang and Q. Chen 2000 Mater. Chem. Phys. **65** 103
[17] L.K. Kurihara and S.L. Suib 1993 Chem. Mater. **5** 609
[18] J.J. Kingsley, K. Suresh and K.C. Patil 1990 J. Mater. Sci. **25** 1305
[19] S.T. Aruna and A. S. Mukasyan 2008 Curr. Opin. Solid ST. M. **12** 44
[20] K. C. Patil, S. T. Aruna and T. Mimani 2002 Curr. Opin. Solid ST. M. **6** 507
[21] R. Ianos and I. Lazau 2009 Mater. Chem. Phys. **115** 645
[22] M. Yu, J. Lin, Y.H. Zhou, S.B.Wang, 2002 Mater. Lett. **56** 1007
[23] G. Blasse 1966 J. Chem. Phys. **45** 2356
[24] H. You, M. Nogami 2005 J. Phys. Chem. B **109** 13980
[25] B. Vengala Rao, U. Rambabu and S. Buddhudu 2006 Physica B **382** 86
[26] B.R. Judd 1962 Phys. Rev. **127** 750
[27] G.S. Ofelt 1962 J. Chem. Phys. **37** 511
[28] G. Blasse and B.C. Grabmaier, ‘‘Luminescent Materials’’ (Springer-Verlag, Berlin, 1994) p.33.
[29] Y.Q. Lu, Y.X. Li, Y.H. Xiong, D. Wang and Q.R. Yin 2004 Microelectron. J. **35** 379
[30] Sakai R, Katsumata T, Komuro S and Morikawa T 1999 J. Lumin. **85** 149

Characterization of nitrogen-doped carbon nanospheres using electron magnetic resonance

Jonathan M. Keartland[1,2], Makhosonke B. Dubazane[1,2], Vincent Marsicano[2], Nikiwe Kunjuzwa[1,3], Neil J. Coville[1,3]

[1]DST/NRF Centre of Excellence in Strong Materials, WITS 2050, South Africa
[2]School of Physics and Materials Physics Research Institute, WITS 2050, South Africa
[3]Molecular Sciences Institute and School of Chemistry, WITS 2050, South Africa

E-mail: Jonathan.Keartland@wits.ac.za

Abstract. Carbon nanospheres (CNS) were produced using two different sets of chemical vapour deposition (CVD) apparatus - a vertical and a horizontal reactor. Nitrogen was introduced into the samples using several sources of nitrogen. Electron magnetic resonance (EMR) was used to characterize a range of samples of varying concentrations of nitrogen at room temperature. The spheres doped with nitrogen show a strong narrow paramagnetic peak at $g \simeq 2$, indicating that the nitrogen takes up substitutional sites in the carbon matrix. Careful analysis enables us to determine the nitrogen content in each of the samples by integration of the resonance peak, and normalising to the mass of the sample. Comparison with a reference sample allows us to extract g for each sample. Power saturation experiments show the the relaxation rates of the nitrogen ions are high in all the samples studied. This paper reports on the synthesis of, and the EMR characterization results for, the nitrogen-doped nanospheres.

1. Introduction

Since the studies on carbon nanotubes by Iijima [1] nanomaterials have received considerable attention from scientists and engineers. Carbon nanomaterials exhibit many remarkable physicochemical properties [2]. In addition, carbon nanomaterials are produced in a large range of morphologies, including tubes, spheres and horns. Carbon spheres come in a range of sizes, from the fulleride family and graphitic carbon onions (2 - 20 nm), to the less graphitic carbon spheres (50 nm - 1 μm), to carbon beads (1 μm upwards) [3]. In this paper we focus on the second category - carbon nanospheres (CNS) that consist of graphene flakes. In particular we are interested in the properties of CNS that contain nitrogen.

Synthesis of CNS may be accomplished using a number of techniques, including chemical vapour deposition (CVD). A recent review [2] provides a current summary of these techniques, as well as a comprehensive summary of the broad current and potential applications of these materials. The study of undoped and doped CNS is consequently of considerable scientific and technological significance. A recent paper [4] has reported on the synthesis and characterization of CNS that contain boron (B-doped CNS). Carbon materials containing substitutional boron or nitrogen occur naturally (in diamond, for example) as these two elements are adjacent to carbon in the periodic table.

It is well-known that electron magnetic resonance (EMR), which is also known as electron paramagnetic resonance (EPR) or electron spin resonance (ESR), may be used to characterize

materials that have unpaired electrons. In the past these investigations have included a studies of the properties of nitrogen centres in diamond [5]. EMR has recently been used to successfully characterize boron-doped multi-walled carbon nanotubes [6], and has also been used in the characterization of other nanomaterials [7]. Nitrogen containing CNS should be sensitive to EMR, particularly if the nitrogen is incorporated into the carbon matrix providing a paramagnetic centre. Incorporation of substitutional nitrogen into the carbon matrix should result in the appearance of a narrow paramagnetic peak, as is seen in diamond for single substitutional nitrogen defects [5]. In this paper we report on the EMR characterization of nitrogen-doped CNS, and provide evidence that the nitrogen is present in substitutional sites. EMR is also used to determine the properties of the nitrogen impurity sites, and to estimate the concentration of the paramagnetic centres.

2. Sample synthesis

Two CVD reactors were used to produce doped and undoped CNS samples using a variety of carbon and nitrogen sources, No catalyst was required to produce CNS samples using either reactor, in contrast to carbon nanosphere CVD synthesis [6].

2.1. Vertical CVD Reactor (v-CVDR)

This apparatus was designed and fabricated in the Nanotechnology Laboratory of the School of Chemical and Metallurgical Engineering, University of the Witwatersrand [8]. It consists of a vertical silica plug flow reactor, immersed in a furnace with a sensitive temperature regulator. A system of rotameters, pressure controllers and valves control the flow of gases into the reactor. The upper end of the reactor is connected to a condenser, which leads to the two delivery cyclones, where the carbon structures produced are collected. Full details of the synthesis procedure are given elsewhere [9]. Doping was achieved using a range of nitrogen sources - the details are given in Table 1.

2.2. Horizontal CVD reactor (h-CVDR)

Carbon spheres were synthesized in a horizontal reactor. The reactor was placed horizontally in an electronically controlled furnace, allowing production of the carbon material. A syringe is used to inject the carbon and nitrogen sources into the quartz reactor. The heating rate, the reaction temperature and the gas flow rates may all be controlled electronically. Full details of the synthesis procedure and a schematic diagram of the reactor are given elsewhere [9]. Doping was achieved by injecting a mixture of pyridine and toluene into the reactor (see Table 1) with all the other reactor parameters kept constant.

3. Experimental considerations

The CNS obtained were characterized using a suite of analytical tools including elemental analysis (see Table 1), Transmission Electron Microscopy (TEM), Thermogravimetric Analysis (TGA), Powder X-ray diffraction (XRD) and Raman Spectroscopy. No catalyst was required for the synthesis of the CNS (unlike for other carbon nanostructures) and hence no purification was required. This also means that no background EMR signal from the catalyst residue was observed [10].

EMR measurements were done at room temperature using a Bruker ESP 300E X-band spectrometer operating in the frequency range 9.4 - 9.9 GHz. All of the EMR experiments were performed using standard continuous wave (CW) techniques. Polycrystalline 2,2-diphenyl-1-picrylhydrazyl (DPPH) reference samples were employed to determine the g-factors and the spin concentrations of nitrogen in the samples. The size of the DPPH sample used varied depending

Table 1. Details of the samples presented for EMR analysis. Acetylene was used as a carbon source in the v-CVDR. Sample NK5 was synthesized at 850 °C. Samples NK6, NK7, NK8 and NK9 were synthesized at 1000 °C. The nitrogen content was obtained from elemental analysis measurements conducted by the Institute for Soil, Climate and Water, Pretoria, South Africa.

Sample	Nitrogen Source	Reactor	Nitrogen %
NK1	Undoped	Vertical	< 0.02
NK2	NH_3	Vertical	0.19
NK3	NH_4^+ ions	Vertical	3.08
NK4	Pyridine	Vertical	0.15
NK5	100:0 Pyridine:Toluene	Horizontal	-
NK6	100:0 Pyridine:Toluene	Horizontal	5.00
NK7	10:90 Pyridine:Toluene	Horizontal	1.48
NK8	0:100 Pyridine:Toluene	Horizontal	0.13
NK9	90:10 Pyridine:Toluene	Horizontal	3.52

on the size of the EMR signal produced by the CNS sample. All of the data were obtained using phase-sensitive detection, and the EMR spectra shown later are the first derivative of the absorption spectrum. Standard power saturation measurements [11] were performed on all samples, in order to obtain estimates of the spin-spin and spin-lattice relaxation rates at room temperature. Small samples are necessary for EMR measurements, as it becomes impossible to tune the microwave cavity for large samples. In addition, changing the sample results in de-tuning of the microwave cavity. In order to re-tune the cavity the microwave frequency must be adjusted.

4. Results and Discussion

Full details of the results of other characterization techniques employed in the analysis of these samples are provided elsewhere [9]. This paper will concentrate on the results obtained using CW EMR.

4.1. CW EMR characterization

A summary of the EMR results is provided in Table 2. All of the samples showed evidence of unpaired electrons, with $g \simeq 2$ (corresponding to a resonance field of approximate 3470 G) in all cases. No additional EMR features were seen over a wide field scan (400 G to 6400 G) in any of the samples. The linewidth and estimates of the g-shift are given for all samples. More detailed discussion of the results for samples produced by the two reactors is provided below. As will be seen the CNS products produced from the two reactors are significantly different.

A comparison between an undoped (NK1) and a nitrogen-doped (NK2) sample is provided in Fig. 1. The appearance of a narrow paramagnetic line with a linewidth of 5.0 G in the spectrum for NK2 indicates that the nitrogen has been incorporated into the carbon matrix. This feature is a characteristic of localized spins [10]. The relatively broad features in both samples may be ascribed to delocalized spins and are probably due to conduction electrons [10]. The delocalized spins in NK1 have a smaller linewidth than the broad background in NK2. This may be evidence that the delocalized spins have been affected by the paramagnetic impurities. Conductivity measurements in B-doped CNS [4] indicate that the conductivity decreases with boron doping, and it is logical to speculate that the broadening of the line for the delocalized spins in NK2 may be related to a decrease in the conductivity. Conductivity measurements on the CNS samples in the present study would therefore be of interest. Fig. 2 compares the

spectra for NK2, NK3 and NK4 (all containing nitrogen according to the elemental analysis measurements). The spectra have been normalized to the background signal provided by the delocalized spins so that the relative magnitude of the paramagnetic peaks may be visually compared. NK3 exhibits a much larger paramagnetic signal, indicating that these spheres have a larger paramagnetic nitrogen component, as is suggested by the elemental analysis results for the v-CVDR samples.

Table 2. Summary of the EMR characterization results of the linewidth (ΔB) and the relative g-shift (Δg) for all samples.

	NK1	NK2	NK3	NK4	NK5
ΔB (G)	9.5 ± 0.5	5.0 ± 0.2	3.2 ± 0.1	1.8 ± 0.1	2.4 ± 0.2
Δg (%)	1.5 ± 0.5	- 3.4 ± 0.1	- 1.4 ± 0.1	- 1.4 ± 0.1	- 1.1 ± 0.1
	NK6	NK7	NK8	NK9	
ΔB (G)	1.36 ± 0.01	0.57 ± 0.01	12.5 ± 0.5	1.14 ± 0.01	
Δg (%)	- 5.09 ± 0.02	- 1.5 ± 0.02	- 8 ± 1	- 5.24 ± 0.02	

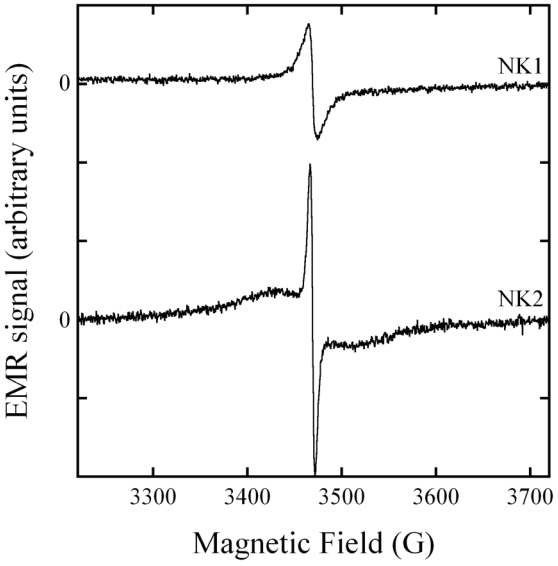


Figure 1. A comparison of the EMR spectra for undoped (NK1) and nitrogen-doped (NK2) CNS produced in the v-CVDR. Note the presence of a strong paramagnetic peak (indicating localized spins - substitutional nitrogen)) in the spectrum for NK2, as well as the broadening of the signal due to delocalized spins [10].

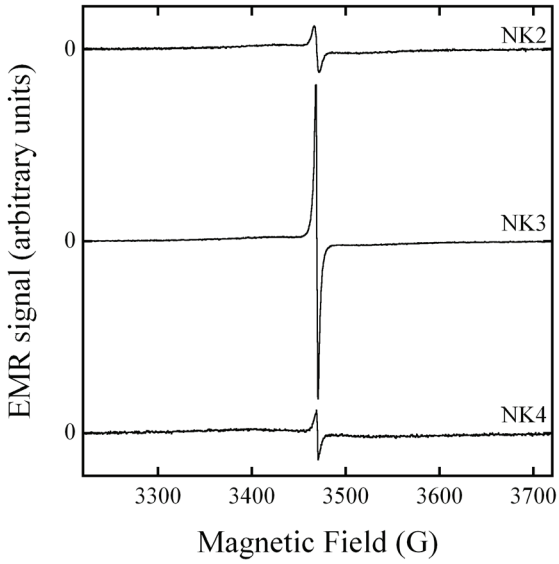


Figure 2. A comparison of the EMR spectra for the three nitrogen-doped samples produced in the v-CVDR. As expected from the results of the elemental analysis (see Table 1) the sample doped with NH_4^+ ions (NK3) has the largest paramagnetic component in the spectrum.

Three of the samples produced using the h-CVDR are qualitatively and quantitatively very different from those produced using the v-CVDR (NK6, NK7 and NK9). The other two samples

(NK5 and NK8) have some similarities. Sample NK8 (0% pyridine) has features that are comparable with NK1 (the undoped sample produced using the v-CVDR). The EMR signal for NK8 may therefore be assumed to originate from delocalized spins (conduction electrons) [10], as is suggested for NK1. The spectrum for NK5 (produced using the h-CVDR at 850 °C) has features similar in character to NK3, but the broad background signal in this sample is relatively much smaller than the paramagnetic peak. From the evidence it would appear that the concentration of nitrogen in this sample is of the same order of magnitude as the nitrogen-doped CNS produced using the v-CVDR.

The other nitrogen-doped CNS samples prepared using the h-CVDR (at 1000 °C) showed a much larger (approximately two orders of magnitude) paramagnetic peak. Fig. 3 compares the spectra obtained for NK6, NK7 and NK9. Several features are of note. As the concentration of pyridine in the source increases the EMR spectrum becomes progressively broader, and the spectrum becomes more asymmetric (or more Dysonian in character). In common with NK5, the broad background observed in samples produced using the v-CVDR is less significant, in these samples it is completely swamped by the signal from the paramagnetic nitrogen, indicating that the concentration of delocalized spins is smaller. This may well have an effect on the electrical properties, leading to a decrease in the conductivity in common with the effect seen previously in boron-doped CNS [4], and discussed previously in this work in relation to the samples synthesized using the v-CVDR. In Fig. 4 spectra of composite samples of the CNS and the reference sample - DPPH - are shown. These spectra show clearly the g-shift of the paramagnetic ions. A computer routine was written to deconvolute the composite spectra, using the experimentally obtained spectrum for the CNS sample and the spectrum for the appropriate DPPH sample, so no assumptions were made regarding the theoretical lineshape in this process. The results of deconvoluting these composite spectra for all samples studied are provided in Table 2.

No evidence was found for power saturation in any of the samples. This indicates that relaxation rates are high, and that the spin-spin and spin-lattice relaxation times are, at most, of the order of 10^{-9} s. Relaxation times may be accurately determined using pulsed EMR techniques, but these techniques are not available at present.

Finally, an attempt was made to determine the relative nitrogen concentration, and the spin concentration relative to a reference sample (DPPH) for NK6, NK7 and NK9. Careful measurements of the mass of each of NK6, NK7 and NK9 were made. Due to the small masses involved (< 10 mg) in the EMR measurements, the mass determination has large inherent errors (~ 20 %). The double integral of the spectrum obtained for each sample was determined, normalized to the mass and compared. The results are tabulated in Table 3, together with the results of elemental analysis for these samples. The results show that the concentration may be controlled by the mixture of different amounts of pyridine and toluene in the source, but these do not correspond to the initial concentrations. Comparison with DPPH shows that the paramagnetic spin concentration is of the order of 3% of DPPH in the highest concentration sample (NK6). It is suggested that the concentration of substitutional nitrogen is of the order of a few percent, at most, in the h-CVDR samples and at least an order of magnitude lower in the samples produced using the v-CVDR.

The CW EMR results are in qualitative agreement with the elemental analysis for all samples. However, the elemental analysis result for NK3 is anomalously high when compared with the results for NK6, NK7 and NK9 (which show clear evidence of a much higher paramagnetic nitrogen component). It is important to note that elemental analysis cannot distinguish between the different forms of nitrogen present in a sample, and the presence of other forms of nitrogen cannot be ruled out. By its nature, EMR provides evidence of only paramagnetic nitrogen.

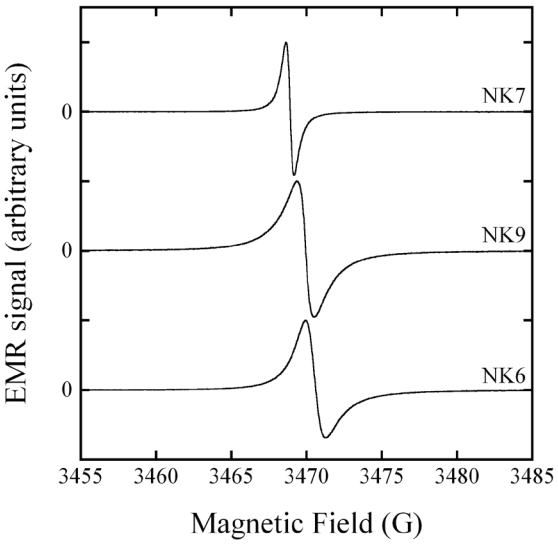


Figure 3. A comparison of the EMR spectra for nitrogen-doped CNS produced in the h-CVDR. As the nitrogen content of the spheres increases the spectrum becomes broader and progressively more Dysonian in character. Shifts in the resonance field from sample to sample are due to re-tuning of the microwave cavity, rather than true g-shifts.

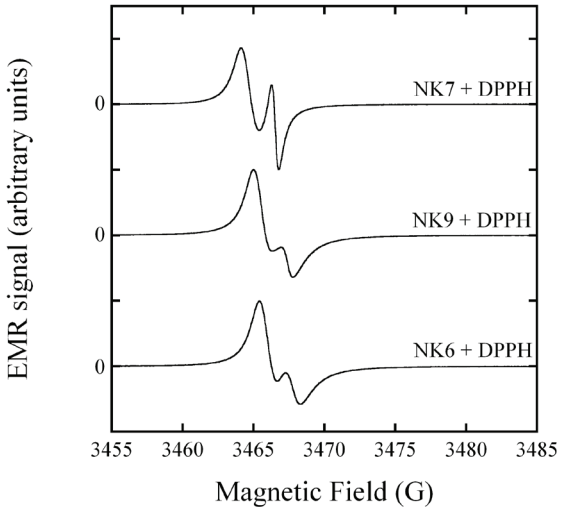


Figure 4. A comparison of the EMR spectra for the composite samples of the DPPH reference and the three nitrogen-doped samples produced in the v-CVDR. Shifts from the resonant fields in comparison with Fig. 3 are due to re-tuning of the microwave cavity following changing the sample. The g-shifts given in Table 2 were obtained from composite spectra such as these.

Table 3. Concentration of paramagnetic nitrogen centres. The relative concentration is calculated with respect to NK6, while the absolute concentration is found by comparing the sample to a polycrystalline DPPH sample. Error estimates are not provided, but these are expected to be large due to the error in the determination of the mass of the samples ($\sim$ 20 %). The results obtained from elemental analysis are shown for comparison.

Sample	Elemental analysis (%)	Relative concentration (%)	Absolute concentration(%)
NK6	5.00	100	3.2
NK9	3.52	16.2	0.55
NK7	1.48	8.4	0.3

5. Conclusions

CNS have been prepared using two CVD methods, and characterized using a suite of techniques, including EMR. The presence of a strong narrow paramagnetic signal in the EMR results for all but two of the samples studied are strong evidence that nitrogen is present in substitutional sites in the N-doped samples to concentrations not exceeding a few percent. The concentration of substitutional nitrogen in the CNS prepared using the h-CVDR is much higher, in general, than those produced using the v-CVDR method. Different sources change the nitrogen concentration in the product, but it would appear that this cannot be done in a controlled fashion at this stage. The reasons for this are not clear, and further investigations are in progress. Increasing the nitrogen concentration changes the character of the EMR spectrum, the spectrum becoming

progressively more Dysonian in character as the concentration increases. It is suggested that the addition of nitrogen in relatively large quantities may affect the conductivity of the pristine CNS, as has been observed in B-doped CNS samples. Electrical transport measurements on this suite of samples would therefore be of interest. Power saturation measurements show that the spin relaxation rates of all samples are high at room temperature. It is also of interest to note that the paramagnetic signals have remained stable since the start of this investigation over two years ago, indicating that the nitrogen remains incorporated in the carbon matrix for at least this length of time.

Acknowledgments
The authors gratefully acknowledge the financial support of the DST/NRF Centre of Excellence in Strong Materials, the National Research Foundation, and the School of Physics and the Materials Physics Research Institute of the University of the Witwatersrand.

References
[1] Iijima S 1991 *Nature* **354** 56.
[2] Nieto-Márquez A, Romero R, Romero A and Valverde J L 2011 *J. Mater. Chem* **21** 1664.
[3] Serp P, Feurer R, Kalck P, Kihn Y, Faria J L and Figueiredo J L 2001 *Carbon* **39** 621.
[4] Mondal K C, Strydom A M, Tetana Z, Mhlanga S D, Witcomb M J, Havel J, Erasmus R M and Coville N J 2009 *Mat. Chem. Phys.* **114** 973.
[5] Smith V V, Sorokin P P, Gelles I L and Lasher G J 1959 *Phys. Rev* **115** 1546.
[6] Mondal K C, Strydom A M, Erasmus R M, Keartland J M and Coville N J 2008 *Mat. Chem. Phys.* **111** 386.
[7] Corzilius B, Dinse K-P and Hatab K 2007 *Phys. Chem. Chem. Phys.* **9** 6063.
[8] Iyuke S E and Danna A B M 2005 *Micropo. Mesopo. Mater.* **84** 338.
[9] Kunjuzwa N 2009 M.Sc dissertation University of the Witwatersrand Johannesburg.
[10] Chipara M, Iacomi F, Zaleski J M and Bai J B 2006 *J. Optoelec. Adv. Mat.* **8** 820.
[11] Castner T G 1959 *Phys. Rev.* **115** 1506.

Temperature variation of the electric field gradient in mercuric chloride

Jonathan M. Keartland and Eric Newby

School of Physics and Materials Physics Research Institute, WITS 2050, South Africa

E-mail: Jonathan.Keartland@wits.ac.za

Abstract. The temperature dependence of the ^{35}Cl quadrupolar resonance frequency has been measured in a high purity sample of mercuric chloride ($HgCl_2$) over the temperature range 4 K - 460 K. The results allow the temperature dependence of the electric field gradient (efg) at the two inequivalent sites of the chlorine atoms to be determined. Several models were considered to describe the experimental observations, including both librational and vibrational modes. The Bayer model for librational modes provides a satisfactory description of the data over the entire temperature range, when the temperature variation of the lattice modes is taken into account, while the other mechanisms considered provide less convincing descriptions. A single librational mode with a wavenumber of approximately 26 cm^{-1} can account for the variation of the electric field gradient over the temperature range investigated. Our conclusions are consistent with the results of optical spectroscopy investigations.

1. Introduction

$HgCl_2$ crystallises in a distorted rhombohedron structure. It is a molecular solid with essentially straight Cl-Hg-Cl molecules that form herring bone layers in the mirror plane of the *Pnma* space group at ambient pressures [1]. This results in two distinct non-cubic sites for the Cl atoms in the crystal, and hence a unique electric field gradient (efg) at each site. In this paper the two sites will be labeled A and B, where A is the site with the larger electric field gradient. The efg at both sites is known to have small deviations from axial symmetry. The interaction between the efg and the quadrupole moment of the Cl nuclei (both isotopes have $I = \frac{3}{2}$) lifts the degeneracy of the nuclear angular momentum energy levels, giving rise to resonance frequencies in the radio frequency region of the electromagnetic spectrum. Pure nuclear quadrupole resonance (NQR) may therefore be used to probe temperature-induced changes in the efg. For nuclei with $I = \frac{3}{2}$, in an efg with lower than axial symmetry, a single resonance line is observed with a frequency given by:

$$\nu_Q = e^2 q Q \sqrt{1 + \frac{\eta^2}{3}}, \tag{1}$$

where ν_Q is the NQR frequency. The quantities eq and η describe the magnitude and asymmetry of the efg respectively and eQ is the quadrupole moment of the nucleus. The results of a pulsed NQR investigation of the quadrupole resonance frequencies over a wide temperature range (4 K - 460 K) are presented here. From the results the temperature variation of the efg

Table 1. Wavenumbers (in cm^{-1}) for vibrational modes of $HgCl_2$ at ambient pressures as measured by Adams and Appleby [10].

T (K)	Translatory modes				Rotatory modes			ν_1	ν_3
295	18	26	43	74	124		167	315	383
150	18.5	26, 29	48	77	126	134		317	388

at the sites A and B are extracted, and the results compared with theoretical models for the variation of the efg with temperature.

NQR was first used to study $HgCl_2$ by Dehmelt *et al* [2], who established the resonance frequencies for both sites A and B, and both Cl isotopes (^{35}Cl and ^{37}Cl) at 303 K. A subsequent study by Dinesh and Narasimhan [3] (DN) investigated the temperature variation over a limited temperature range at several fixed points, and based their data analysis on the theory of Bayer [4] as modified by Kushida *et al* [5] and Brown [6]. They predicted a low wave number lattice frequency that had not been observed previously using optical spectroscopy. Dinesh and Smith (DS) [7] extended the temperature range of the measurements of the NQR frequencies in $HgCl_2$ to the temperature range 77 K - 400 K. They made no further analysis of the temperature dependence of the efg at the two sites despite the results differing markedly from the results published previously by DN. No further measurements or analysis of the temperature dependence of the NQR frequency have been reported. In particular, the temperature region below 77 K has not been explored at all. The work described here addresses the lack of published data for ν_Q in $HgCl_2$ in this temperature region. A more recent NQR investigation [8] has revisited the asymmetry parameter of the efg at room temperature and provided values consistent with theoretical calculations [9]. It is generally accepted that there are small distortions from axial symmetry in $HgCl_2$, but there is no definitive determination of the temperature dependence of this parameter.

The temperature variation of the efg in molecular solids arises from modulation of the static efg by lattice vibrations. The lattice dynamics of $HgCl_2$ have been studied as a function of temperature and pressure using Raman and infrared spectroscopy by Adams and Appleby [10], who have made a complete assignment of the vibrational spectrum from measurements made at 295 K and 150 K. In a more recent study by Voyiatzis and Papatheodorou [11] Raman spectra of mercuric chloride were taken at 77 K and 545 K. Neither of these data sets were available to DN in their data analysis, but the results confirm their prediction that low wave-number modes play a dominant role in the temperature variation of the efg. The results obtained by Adams and Appleby [10] are summarised in Table 1. The values obtained for the mode at approximately 26 cm^{-1} by Adams and Appleby [10] and Voyiatzis and Papatheodorou [11] are given in Table 2 and are plotted in Fig. 1. The significance of this particular mode will become clear in the discussion of our data that follows.

A more recent X-ray structure study by Subramanian and Seff [12] at room temperature has confirmed the general observations of the crystal structure. In particular, they have shown that the Cl-Hg bond-lengths are marginally larger than the parameters used by Brill *et al* [9] in their calculations.

2. Experimental details

A commercially obtained high purity powder sample of $HgCl_2$ was sealed in a quartz ampoule under vacuum. No further purification of the sample was undertaken. NQR frequencies for the ^{35}Cl and ^{37}Cl nuclei were measured using a standard coherent pulsed radio frequency spectrometer operating in the range 17 - 23 MHz. Quadrupolar spin echoes were obtained using the standard $\pi/2 - \pi$ sequence with a pulse separation of the order of 500 μs. Quadrature

Table 2. The temperature variation of the wavenumber of the mode at approximately 26 cm^{-1} as determined by Adams and Appleby [10] and Voyiatzis and Papatheodorou [11].

T (K)	k_1 (cm^{-1})
77	26.5
150	26.0
295	26.0
545	25.0

detection and phase cycling allowed measurement of the resonance frequency to an accuracy of 1 kHz for the ^{35}Cl nucleus.

Measurements were made using three experimental arrangements. A copper cryostat, fitted with a heater and a copper-constantan thermocouple in thermal contact with the sample, was used to obtain data in the range 77 K - 300 K. The cryostat was placed in an partially sealed nitrogen dewar above the level of the cryogen. A heater was immersed in the liquid to provide dense cold vapour, and hence exchange gas cooling of the cryostat. Temperature control was achieved using the cryostat heater and a copper-constantan thermocouple, together with an Oxford Instruments temperature controller calibrated with respect to the thermocouple. Above room temperature (300 K - 460 K) an oil bath was employed in conjunction with the copper-constantan thermocouple and the temperature controller. An Oxford Instruments continuous flow cryostat (CFC) was used in the low temperature region (4 K - 80 K). In these experiments the temperature was monitored using the Fe+7%Au-chromel thermocouple and a calibrated carbon-glass resistance thermometer. Temperature stability of the CFC was found to be well within 0.1 K in the range 4 K - 100 K, and the sample temperature remained constant during measurements within the precision of ν_Q.

3. Theoretical Models

The variation of the NQR frequency with temperature has been studied extensively in many molecular solids. In most cases a modified form of the Bayer model for molecular librations is sufficient to explain the observed temperature dependence, in the absence of a structural phase transition. In some cases the acoustic phonon contribution to the modulation has been significant. A brief review of the underlying theoretical models is presented.

3.1. Molecular Librations

A model for the temperature variation of the efg due to rotations about the axis of symmetry of the efg for a single mode was first developed by Bayer [4]. For a single mode the expression is of the form

$$\nu_Q(T) = \nu_0 \left[1 - \frac{3\hbar^2}{4Ik_B\Theta_T} \coth\left(\frac{\Theta_T}{2T}\right) \right], \tag{2}$$

where ν_0 is the limiting or static value of the resonance frequency, I is the moment of inertia of the rotating molecule and Θ_T is the characteristic temperature of the oscillation frequency. Kasprzak and Nogaj [13] have provided an illuminating discussion of the various models based on the Bayer model, and these models are presented briefly below. In each of these models the quantity a represents the the resonance frequency at 0 K. The other co-efficients may be determined from fitting the models to the data. These models do not take into account explicitly the asymmetry of the efg. This point will be addressed in more detail in the discussion that follows.

3.1.1. Kushida-Benedek-Bloembergen (KBB) model This model takes into account the effect of changes in the sample volume with temperature. Kushida *et al* [5] have correctly pointed out that NQR measurements are usually made at a constant pressure, not at a constant volume. The model has the following form:

$$\nu_Q(T) = \nu_0\left(1 + bT + \frac{c}{T}\right). \tag{3}$$

3.1.2. Brown Model Brown [6] accounted for the anharmonicity of the librations by assuming a linear temperature dependence of the mode frequencies. The model has the form:

$$\nu_Q(T) = \nu_0\left(1 + bT + cT^2\right). \tag{4}$$

3.1.3. Combined Brown and KBB (B-KBB) model This model takes into account both the changes in sample volume and the anharmonicity of the librations as follows:

$$\nu_Q(T) = \nu_0\left(1 + bT + cT^2 + \frac{d}{T}\right). \tag{5}$$

3.1.4. Extended Brown Model In many measurements of the lattice dynamics a quadratic dependence of the libration frequencies has been observed. In an attempt to take this non-linearity into account a fourth term was added to the Brown model as follows:

$$\nu_Q(T) = \nu_0\left(1 + bT + cT^2 + dT^3\right). \tag{6}$$

3.2. Acoustic Vibrations

A model for the contribution of acoustic phonons to the temperature variation of the efg has been developed by Stahl [14] and Stahl and Tipsword [15]. More recently this model has been applied to $KBrO_3$ [16]. For an axially symmetric efg the variation of the quadrupolar resonance frequency with temperature is of the form:

$$\nu_Q(T) = \nu_0\left[1 - S\left(\frac{T}{\Theta_D}\frac{c_v}{9R} + \frac{1}{\exp\frac{\Theta_D}{T} - 1}\right)\right], \tag{7}$$

where the dimensionless quantity S represents the ratio of the Boltzmann energy at the Debye temperature to the rotational energy, Θ_D is the Debye temperature of the solid, and

$$\frac{c_v}{9R} = \left(\frac{T}{\Theta_D}\right)^3 \int_0^{\frac{\Theta_D}{T}} \frac{x^4 e^x}{\left(e^x - 1\right)^2}\, dx. \tag{8}$$

In the low temperature limit this model predicts a T^4 dependence, which changes to a linear dependence at high temperatures (for all practical purposes the linear dependence is apparent for $T \geq \Theta_D$).

4. Results and Discussion

The results for the entire temperature range studied for the ^{35}Cl nucleus at both sites A and B are given in Fig. 2. There is a smooth change in the efg with temperature over the entire temperature range, confirming that no phase transitions occur. The results are in excellent agreement with those of DS, but there are large discrepancies with those of DN at high temperatures, and in the region between 77 K and 273 K. All of the models discussed in the previous section were fitted to

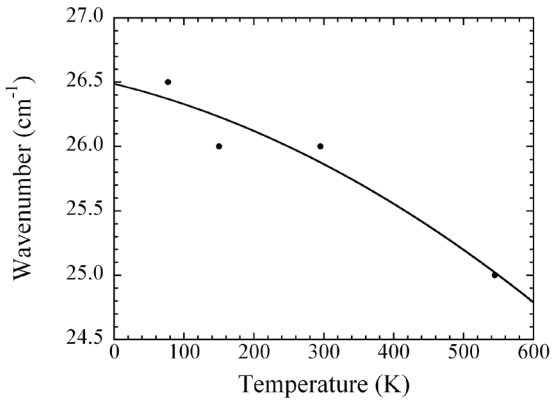


Figure 1. Values for the lattice mode at approximately 26 cm^{-1} obtained by Adams and Appleby [10] and Voyiatzis and Papatheodorou [11]. A quadratic is fitted to the data.

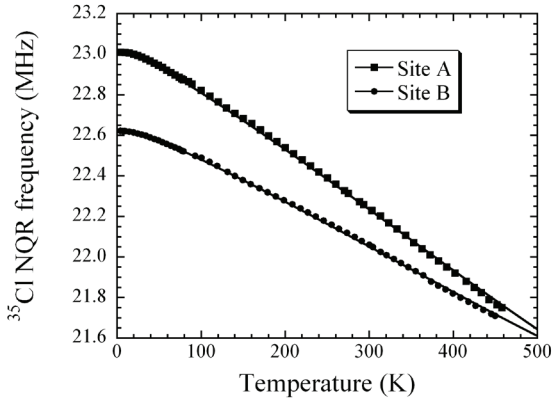


Figure 2. NQR resonance frequency data over the entire temperature range. The Bayer model for a single vibrational mode is fitted to the data. The fit parameters are given in Table 3.

our data. The KBB and B-KBB models both produce unphysical results at low temperatures, as the $1/T$ term becomes dominant. The Brown model shows significant departures from the data at low temperatures, while the acoustic phonon model provides a markedly better fit in this temperature range. The Debye temperature extracted from the acoustic phonon fits is 120 $\pm$ 5 K. A value for the Debye temperature for $HgCl_2$ could not be sourced from the literature for comparison. Significant discrepancies between the data and the acoustic phonon model at high temperatures were noted, as there is a clear non-linear behaviour within the precision of the data. The Brown model provides a better description of the data at high temperatures. The extended Brown model produces a better fit to the data at low temperatures, as might be expected.

A consideration of the assignment of the symmetry of the Raman modes by Adams and Appleby [10] will allow determination of the modes most likely to contribute to thermal modulation of the efg. The rotatory modes are both of high wave-number and their contribution is insignificant. In addition, there appear to be large changes in the character of these modes with temperature (see Table 1), and this is not reflected in the high temperature measurements of this work and those of DS. Of the low energy translatory modes, the two at approximately 43 cm^{-1} and 74 cm^{-1} have been assigned to modes parallel to the molecule axis, and are therefore unlikely to contribute to rotations about the z-axis of the efg. The two lower modes (approximately 18 cm^{-1} and 26 cm^{-1}) have been assigned to modes perpendicular to the z-axis, and so these may contribute to rotations of the molecules. The presently accepted experimental values of the asymmetry parameter at room temperature are $\eta_A = 0.037$ and $\eta_B = 0.012$ [8]. Setting $\eta = 0$ for further analysis is unlikely to result in significant errors. Assuming that the temperature variation of η is negligible we apply a modified Bayer model for the single mode (2) at 26 cm^{-1} to the data for sites A and B as described below. The geometrical value for the rotational inertia was obtained from the structure data of Subramanian and Seff [12] and is $I = 6.187 \times 10^{-45}$ kg.m^2. The values of k_1 of the mode at approximately 26 cm^{-1} obtained experimentally [10, 11] are plotted against T in Fig. 1 and a quadratic of the form $k_1(T) = a_0 + a_1 T + a_2 T^2$ is fitted to the data to obtain a_0, a_1 and a_2. Θ_T is proportional to $k_1(T)$ so Θ_T is of the form

$$\Theta_T = \alpha \left(a_0 + a_1 T + a_2 T^2 \right). \tag{9}$$

We substituted (9) into the Bayer model (2), and the resulting expression was fitted to data

Table 3. Fit parameters for the Bayer model for a single mode over the entire temperature range.

Site	ν_0 (MHz)	α
A	23.070	1.59
B	22.676	1.83

for both lattice sites. The values for α and ν_0 extracted from the fit are given in Table 3. The fit provides an excellent description of ν_Q over the entire temperature range, except for small deviations at the highest and lowest temperatures (above 400 K and below 20 K respectively). We conclude that the temperature-induced changes in the efg at the two lattice sites are dominated by rotations of the molecules about the axis of symmetry of the efg for the mode at approximately 26 cm^{-1}.

5. Conclusions

The temperature dependence of the efg in $HgCl_2$ at the two inequivalent lattice sites has been deduced from measurements of the NQR frequencies over a wide temperature range. The data for the resonance frequency is best described by the model for molecular librations developed by Bayer [4] for a single mode with a wave number of approximately 26 cm^{-1}, although deviations above 400 K and below 20 K are noted. It is probable that these discrepancies are related to the necessary approximations employed in determining the temperature dependence of this mode obtained from the Raman data. None of the other models considered provide a convincing fit to the data over the entire temperature range. Our results suggest that further investigations of the lattice dynamics of $HgCl_2$ at lower temperatures would be of interest.

Acknowledgments

The authors gratefully acknowledge the financial support of the the National Research Foundation, and the School of Physics and the Materials Physics Research Institute of the University of the Witwatersrand. Technical support from the staff of the Physics Workshop has been invaluable. Fruitful discussions with C. Kasl on the experimental arrangement are also acknowledged.

References

[1] Hostettler M and Schwarzenbach D 2005 *C. R. Chimie* **8** 147–156
[2] Dehmelt H G, Robinson H G and Gordy W 1954 *Phys. Rev.* **93** 480–482
[3] Dinesh and Narasimhan P T 1966 *J. Chem. Phys.* **45** 2170–2174
[4] Bayer H 1951 *Z. Phys.* **130** 227–238
[5] Kushida T, Benedek G B and Bloembergen N 1956 *Phys. Rev.* **104** 1364–1377
[6] Brown R J C 1960 *J. Chem. Phys.* **32** 116–118
[7] Dinesh and Smith J A S 1974 *Advances in Nuclear Quadrupole Resonance* **1** 31–36
[8] Sengupta S, Giezendanner D and Lucken E A C 1980 *J. Magn. Reson.* **38** 553–554
[9] Brill T B, Hugus Z Z and Schreiner A F 1970 *J. Phys. Chem.* **74** 469–475
[10] Adams D M and Appleby R 1977 *J. Chem. Soc., Dalton Trans.* **11** 1530–1535
[11] Voyiatzis G A and Papatheodorou G N 1994 *Ber. Bunsenges. Phys. Chem.* **98** 683–689
[12] Subramanian V and Seff K 1980 *Acta Crystallogr. B* **36** 2132–2135
[13] Kasprzak J and Nogaj B 1993 *Phys. Stat. Sol. (a)* **136** 195–206
[14] E A Stahl J 1975 *J. Chem. Phys.* **63** 5055–5056
[15] E A Stahl J and Tipsword R F 1990 *J. Magn. Reson.* **28** 95–98
[16] E A Stahl J and Herzog S 2003 *Magn. Reson. Chem.* **41** 996–999

Cr/ α-Cr_2O_3 monodispersed spherical core-shell particles based solar absorbers

S. Khamlich[1,2*], Z. Y. Nuru[1], R. Nemutudi[1] , O. Nemraoui[1], K.T Roro[3], N. Mongwaketsi[1], R. McCrindle[2], N. Cingo[1,3] and M. Maaza[1,2,]

[1] Nanosciences African Network, Materials Research Dept.- iThemba LABS, National Research Foundation,1 Old Faure road, Somerset West, South Africa

[2] Faculty of Sciences, Pretoria-Tshwane University of Technology, Private Bag X 680, Pretoria, South Africa.

[3]African Laser Centre, CSIR campus, P.O. Box 395, Pretoria, South Africa

E-mail address: skhamlich@tlabs.ac.za, skhamlich@gmail.com

Abstract. Monodispersed spherical core-shell particles of Cr/α-Cr_2O_3 cermet ACG coatings investigated within this contribution could be successfully employed in thermal converters. Their selectivity depends on their chemical, physical and structural characteristics and related optical properties like reflectance, emittance, solar light absorption, or absorptance and transmittance. Parameters such as particles size, aging-time and substrate roughening will greatly influence their relevant optical properties as reported. The coated Cr/α-Cr_2O_3 spherical particles on rough copper substrates by a simple self-assembly-like method were characterized by scanning electron microscopy, energy dispersive spectrometry, Raman spectroscopy, and diffuse reflectance UV-VIS-NIR spectroscopy. The samples aged for more than 40h at 75°C exhibit the targeted high absorbing optical characteristic "Black chrome" while those aged for ≤ 40h show a significant high diffuse reflectance.

1. Introduction

The development of new and low-cost selective solar photothermal absorbers is gaining momentum and efforts in a number of research laboratories all over the world. Spectrally selective coatings used in solar collector are known to enhance the efficiency of solar thermal conversion. In photo-thermal systems, the solar energy is generally obtained in the form of heat, usually absorbed by a black body like surface such as surfaces coated with a thin film of black paint. To enable effective solar energy utilization, the absorption of the incident solar radiation should be maximized while the thermal losses from the collector minimized. Hence, the solar absorbing surface is one of the key components of a solar collector, and its optical properties and quality influence both the heat losses and gains. Therefore, it is important to develop a spectral-selective solar energy absorbing surface which is characterized by a maximum absorptance (α) over the solar spectrum (0.3–2.5 μm) and low thermal emittance (ε) at the operating temperature. This is translated by a low reflectance of absorbing surface in the solar main irradiation spectrum (nearly zero) and high reflectance (close to one) in the IR region. For practical reasons, a good selective coating features optical properties of α ~ 0.9 and ε ~ 0.2 [1]. More recently, many efforts have focused on developing new types of ceramic-metal nano-composite materials for improving their performance and thermal stability as solar selective absorbers [2]. Metal-dielectric composites, known as cermets, exhibit a strong absorption peak in the main solar radiation

region while remaining substantially transparent in the thermal infrared region. A large number of different metal-dielectric composites [3,4], of Cu, Au, Ni, Mo, Cr, Co, Pt, or W in SiO, SiO_2, Al_2O_3, or MgO have been studied as selective absorbers with certain detail, and some of them have reached commercial application. Electroplated black chromium, which employs a Cr-Cr_2O_3 cermet material, is the most widely used solar absorber of this kind [5,6], but preparation of a black chromium solar coating using an electroplating technique is not suitable due to environmental pollution.

In this paper, we describe the preparation of large surface coatings of black chromium-chromium (III) oxide cermet by an environmentally friendly and novel cost effective physical-chemical method: the Aqueous Chemical Growth "ACG". This method presents several advantages such as being simple, easy to scale up, cost effective and versatile in operation [7,8]. More accurately, we report on the influence of aging time on the formation of Cr on the surface of the uniform fine "in the range nano to micron" spherically shaped core-shell particles of Cr/α-Cr_2O_3 with high adhesion onto metallic substrates using the ACG low temperature-soft physical-chemistry preparative technique and as well as on the effect of aging on the optical properties.

2. Experimental techniques

The growth of the monodisperse spherical particles of hydrated Cr_2O_3 was conducted according to Ref. 7. The synthesis was performed with reagent grade chemicals. An aqueous solution (MilliQ, 18.2 MΩ cm) of 1mM of chromium potassium sulfate dodecahydrate ($KCr(SO_4)_2.12H_2O$) at pH 6.5 mixed in a glass bottle with autoclavable screw cap (e.g., Duran laboratory) containing a rough copper substrate was heated in a laboratory oven at a constant temperature of 75 ºC with different aging times. Subsequently, the durable and scratch resistant thin films are thoroughly washed with water to remove any residual salts. Afterward, the particulate coatings are heat treated in a flowing H_2 gas at 500 °C, based on thermal analysis data, for 1h to obtain Cr/α-Cr_2O_3 quasi-monodisperse and spherically shaped core-shell particles. The structural characterization of Cr/α-Cr_2O_3 cermet nano-coatings are investigated using Raman spectroscopy "model Horiba Jobin Yvon, LabRAM HR UV/Vis/NIR" and Perkin Elmer Spectrum one FTIR with diamond ATR accessory. The particles and the coating morphologies are investigated using a Leo-StereoScan 440 scanning electron microscope "SEM". The chemical composition is determined using electron energy dispersive spectroscopy "EDS". Diffuse reflectance spectra were recorded in the wavelength range 300–2500 nm using Varian Associated Cary 500 double beam spectrophotometer. The infrared emissivity (ε) was measured in the wavelength range of 3 to 30 μm by the D&S Emissometer, Model AE1 within ±0.01 emissivity units, which is close to the required prediction accuracy.

3. Results and descutions

3.1. Raman spectroscopic study

The Raman spectra of Cr/α-Cr_2O_3 monodispersed spherical core-shell particles are depicted in Fig. 1. However, all the detected peaks exhibit only three E_g and one A_{1g} modes of chromium (III) oxide, and no Cr peaks due to its metallic behavior. Moreover, chromium (III) oxide has a corundum structure that belongs to the D^6_{3d} space group and consists of a hexagonal close packed array of oxide anions with 2/3 of the octahedral holes occupied by chromium. The site symmetry for the Cr atoms is C_3, whereas the O atoms are on sites having C_2 symmetry. The irreducible representations for the optical modes in the crystal are $2A_{1g}$, $2A_{1u}$, $3A_{2g}$, $2A_{2u}$, $5E_g$ and $4E_u$ vibrations with only two A_{1g} and five E_g vibrations are Raman active [9]. Hart *et al*. [10], in their Raman investigations of Cr_2O_3 single crystals, observed seven Raman bands. In this present work, we found that all peaks belonging to samples aged at 20, 30, 40 and 50h showed only three E_g and one A_{1g} modes of chromium (III) oxide (see Fig. 1). The observation of these only four modes in the samples could be caused by the polycrystallinity of Cr/α-Cr_2O_3 (micron to nano)-spherical particles. On other hand, it is evident from figure 1 that Raman intensity increases with increase of ageing time and it become nearly constant after 50h in line with a

high crystallinity. These results reveal that the size and amount of the particles is nearly stable after 50h, this is also confirmed by the scanning electron microscopy results in our previous study on the growth kinetics of α-Cr_2O_3 spherical particles [8].

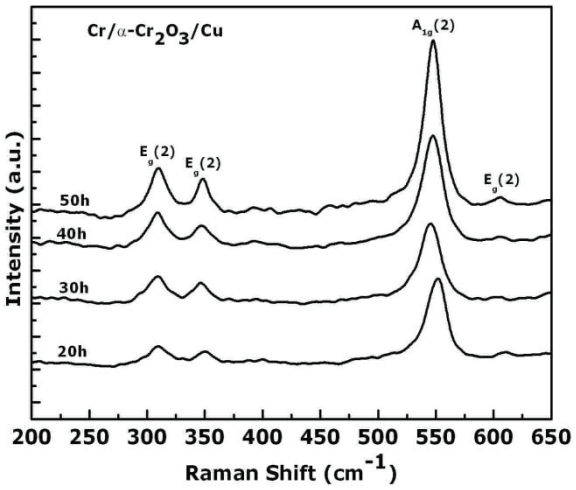


Figure 1. The Raman spectra of Cr/α-Cr_2O_3 monodispersed spherical core-shell particles

3.2. Attenuated Total Reflection (ATR) study

Figure 2 depicts the Attenuated Total Reflection (ATR) spectra for Cr/α-Cr_2O_3 monodisperse spherical core-shell particles deposited on copper at different aging time, and heated at 500 °C in a flowing H_2 gas. As one could notice that a gradual disappearance of Cr-O and O-O bands with increasing aging time is observed. This may be due to the black thin layer of chromium metal coated on the surface of the particles. Furthermore, samples aged for 20 and 30h reveal two strongest bands centered at 539 and 611cm^{-1}. These values fall well within the range reported in the literature for samples having the exact α-Cr_2O_3 stoichiometry [11]. Both bands are associated with Cr–O stretching modes (symmetrical stretching and antisymmetrical stretching) in α-Cr_2O_3 and are due to various combinations of O^{2-} and Cr^{3+} displacements in the lattice. The pair of sharp bands at 442 and 411 cm^{-1} corresponds to two specific O^{2-} displacements in the lattice. From the ATR and Raman results we concluded the existence of pure α-Cr_2O_3. Moreover chromium metal doesn't show any peaks in the Raman/ATR spectra in the wave-number range studied which suggested the use of electron energy dispersive spectroscopy technique to investigate the existence of Cr as a shell in Cr/α-Cr_2O_3 monodisperse spherical core-shell particles

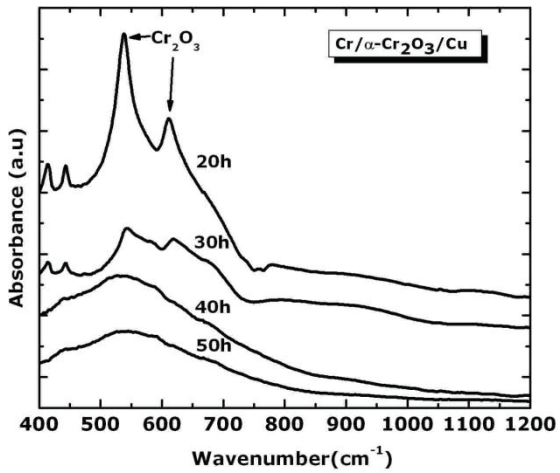


Figure 2. The Attenuated Total Reflection spectra for Cr/α-Cr_2O_3 monodisperse spherical core-shell particles

3.3. Scanning Electron Microscopy (SEM) and EDS analysis

SEM micrographs in the inset of Figure 3 reveal that the Cr/α-Cr_2O_3 coatings consists of spherical particles with an average size of about 498, 990, 1010 and 1185 nm respectively. This result demonstrated that Cr/α-Cr_2O_3 can be produced with well-defined surface morphology, narrow size distribution and excellent shape control not only onto standard amorphous substrates but also on metallic ones. The chemical composition analysis of the Cr/α-Cr_2O_3 monodisperse spherical core-shell particles by EDS shows that the coatings contain black chromium "see Figure 3", but since no chromium was found by Raman/ATR measurements, it is confirmed that the film contains chromium. A closer EDS analysis confirmed that the percentage of chromium increases with increasing aging time. In view of discriminating the shell of Cr from the core of α-Cr_2O_3, a fine electron probe beam of 3 nm in the EDS system was used. Thereafter, EDS spectra were acquired by positioning at different parts of the large-size particles. Figure 4 reports a comparison of EDS spectra acquired by positioning the electron probe through only the shell and through the core and shell "see Figures 4a and 4b", respectively. However, the atomic percent composition of the shell is 76.2 at.% chromium and 23.8 at.% oxygen. On the other hand, the atomic percent composition of the core is 29.6 at.% chromium and 70.4 at.% oxygen. It is apparent that the shell is dominated by Cr, while the core is rich in α-Cr_2O_3, suggesting that the shell of the large-size particle is Cr.

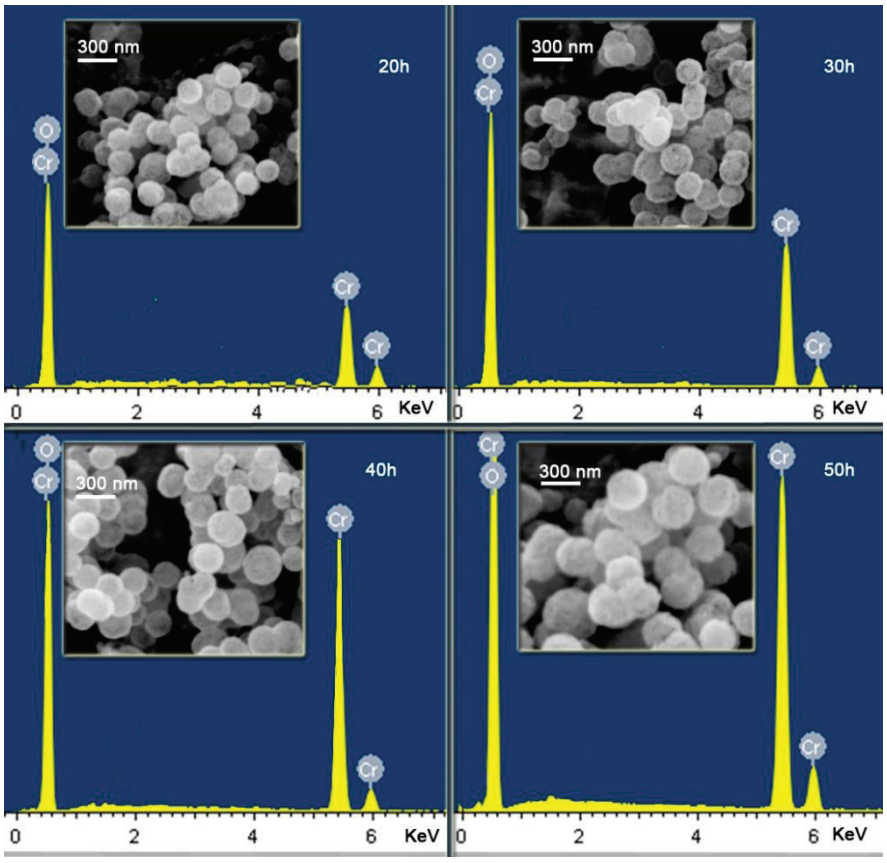


Figure 3. SEM/EDS of Cr/α-Cr_2O_3 monodispersed spherical core-shell particles

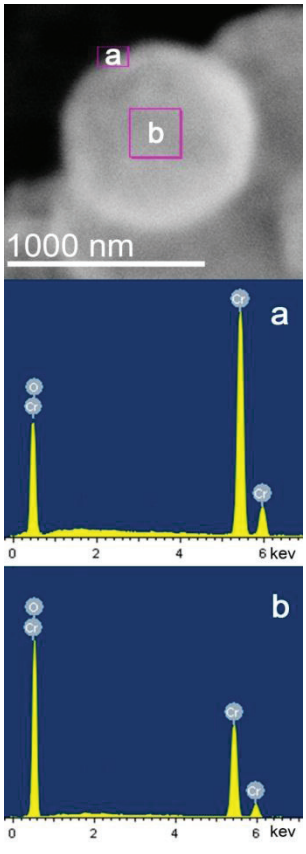


Figure 4. Investigation of shell and core of one Cr/α-Cr_2O_3 monodispersed particle

3.4. Diffuse reflectance and the infrared emissivity (ε): preliminary study

Figure 5 represents the total diffuse reflectance "specular and non-specular reflectivities" of different samples over the UV-VIS-NIR spectral range. In general, very low reflectance spectra are obtained with samples aged for more than 40h at 75°C. Furthermore, it is observed that the spectral reflectance of the materials decreases with increasing aging time. A quantitative comparison of the total solar absorptance [calculated by numerical integration using measured total diffuse reflectance spectrum of figure 5 in the solar radiation region of 300–2500nm and the solar irradiance at Air Mass 1.5 (AM1.5)] and the measured infrared emissivity of all the samples is presented in Table 1. The absorptivity and emissivity of sample aged for high aging time t ~50h is very high, but its emissivity can be reduced by a suitable heat mirror such as SnO_2:F. For samples aged at lower aging time t~20h the absorptivity is relatively high and the emissivity low, hence such nano-coatings are potentially competitive for solar absorbers application. However, this difference in the optical properties is strongly dependent on the aging time of the samples and can be attributed to many factors such as the film microstructure, surface composition and morphology.

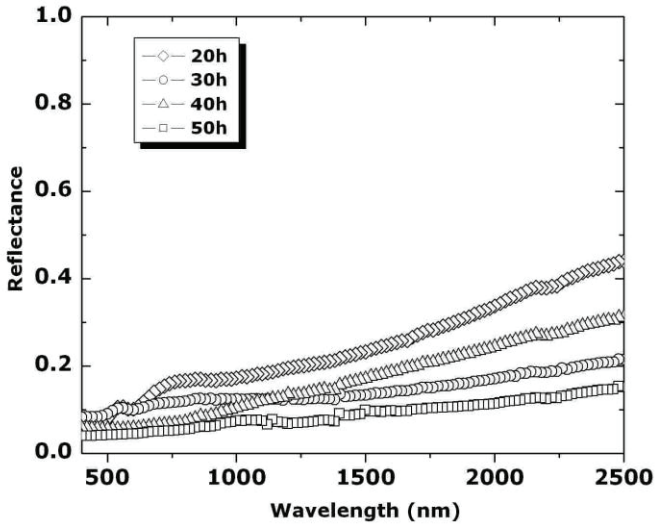


Figure 5. The total diffuse reflectance of different samples over the UV-VIS-NIR range.

	Aging time (h)	Absorptance (α)	Emittance (ε)
Sample 1	20	0.84	0.49
Sample 2	30	0.88	0.77
Sample 3	40	0.90	0.60
Sample 4	50	0.94	0.80

Table 1. The solar selectivity of Cr/α-Cr_2O_3 monodispersed spherical core-shell particles coated on Copper substrates

4. Conclusion

Chromium-Chromium (III) oxide Cr/α-Cr_2O_3 monodispersed spherically shaped core-shell particles were deposited by simple self-assembly method with high adhesion to the substrate: the Aqueous Chemical Growth "ACG". This ACG technique has demonstrated that it is a potential method to produce reproducible high quality-low cost coatings. Samples deposited for longer aging time show high chromium percentage on the surface of the particles, high absorptivity and emissivity but this can be reduced by a suitable heating mirror. Samples deposited for lower aging time show high absorptivity and low emissivity. As conclusion, these preliminary results indicate that this type of nano-coatings onto Cu or stainless steel rough substrates, are good candidates for solar absorbers applications at temperature lower than 500^{o}C.

Acknowledgements

This work was sponsored by the African Laser Centre (ALC) in collaboration with iThemba LABS-MRD/ National Research Foundation of South Africa, Tshwane University of Technology and the Council for Scientific and Industrial Research in South Africa, as well as the Abdus Salam ICTP-Trieste., Italy.

References

[1] Mastai Y, Polarz S , Antonietti M 2002 *Adv. Funct. Mater*. **12** 197–202
[2] Zhao S, Avendaňo E, Gelin K, Lu J and Wäckelgãrd E 2006 *Sol. Enegy Mater. & Sol Cells* **90** 308
[3] Fan J C and Zavracky P M 1977 *Appl. Phys. Lett.* **29** 478
[4] Blain J, Bel C L, Saint-Jacques R G and Rheault F 1985 *J. Appl. Phys.* **58** 490
[5] Agnihotri O P and Gupta B K 1981 Solar Selective Surfaces, *Wiley-Interscience Publication*, New York
[6] Pekruhn W, Thomas L K, Broser I, Schroder A and Wenning U 1986 *Sol. Energy Mater.* **13** 199
[7] Vayssieres L, and Manthiram A 2003 *J. Phys. Chem, B.* **107** 2623-2625
[8] Khamlich S, Manikandan E, Ngom B D, Sithole J, Nemraoui O, Zorkani I, McCrindle R, Cingo N, Maaza M. 2011 *J. Phys. Chem. Solids.* **72** 714-718
[9] Mougin J, Bihan T. Le, Lucazeau G. 2001 *J. Phys. Chem. Solids* **62** 553-563
[10] Hart T.R, Aggarwal R.L, Lax B, 1971 *Proc. Int. Conf. on Light Scattering Solids, Paris*, 174-179
[11] Tabbal M, Kahwaji S, Christidis T C, Nsouli B, Zahraman K, 2006 *Thin Solid Films* **515** 1976-1984

Vibrational properties of mass produced graphene monolayer by chemical method

M Khenfouch[1], M Maaza[2] and M Baïtoul[1]

[1] University Sidi Mohamed Ben Abdellah, Faculty of Sciences Dhar el Mahraz, Laboratory of Solid state Physics, Group of Polymers and Nanomaterials, BP 1796 Atlas Fez 30 000, Morocco

[2] Nanosciences Laboratories, Materials Research Department, iThemba LABS, Old Faure Road, P.O. Box 722, Somerset West 7129, South Africa.

khenfouch@yahoo.fr

Abstract. Graphene is a two-dimensional crystal of carbon atoms arranged in a honeycomb lattice. It is a zero band gap semimetal with very unique electronic optical and mechanical properties which make it useful for many applications such as ultra-high-speed field-effect transistors, p-n junction diodes, terahertz oscillators, low-noise electronic, NEMS and optical sensors. When the high quality mass production of this nanomaterial is a big challenge, we have used chemical method which helped to achieve this goal. Hummers method was followed by purification with alcohols and different heat treatment steps to quantitatively produce good quality graphene powder. Uv-Vis, Raman and FTIR vibrational spectroscopies, Scanning and transmission electron microscopies, Energy Dispersive X-ray were investigated to the examination of the structure and the evolution of the prepared graphene sheets quality.

1. Introduction

Graphene has received much attention recently in the scientific community because of its distinct properties and potentials [1]. The recent emergence of this stable two-dimensional (2-D) carbon, graphene, has initiated a rapid exploration of its potential in a broad range of fields such as nanoelectronics, sensors, catalysis and nanocomposites [2].

A central challenge of rapid graphene implementation is its mass production and researchers are investigating a number of alternative means for producing the nanomaterial. These methods have shown great promise for industrial-scale production of graphene and the chemical exfoliation of graphite through oxidation and then dispersion in water, down to single graphene sheets, is one of the potential methods of achieving this goal [3].

Raman and FTIR spectroscopies have been utilized as a powerful tool for the characterization of graphene, as it can identify the number of layers, the electronic structure, the edge structure, the type of doping and any defects in the graphene [4].

In this work we developed a simple chemical method for producing large quantities of the promising nanomaterial graphene. This method needs little processing, and paves the way for cost-effective mass production of graphene. Raman and FTIR spectroscopies were investigated to study the vibrational properties in the goal to determine structural properties of our prepared graphene powder.

2. Experimental

The graphite used in this work was purchased from Sigma-Aldrich. We synthesized graphene powder using a modified Hummers method which involved the exfoliation of graphite in the presence of strong acids and oxidants [5]. The obtained solution was reduced with hydrazine at 40 ^{o}C with stirring for 6h. Poly(vinyl-pyrrolidone) (PVP) dissolved in ethanol-water solution was added to the reduced graphene oxide. This mixture was stirred and sonicated for 30 minutes and heated at 60°C for a slow evaporation. The obtained powder, washed with alcohol, was dispersed in PVP by sonication and calcinated at 900°C under Argon atmosphere (Fig.1).

To analyse our synthesized graphene powder we have used Bruker Sentarra RAMAN spectrometer with an excitation wave length of 785 nm, Bruker Vertex 70 Fourier transform infrared (FTIR) Spectrometer and SEM images were obtained by a Nova NanoSEM 230. Fei High resolution TEM. The absorption spectrum was recorded at room temperature using a Cary UV-visible-near-infrared spectrometer. Lastly, the Element composition was investigated with an Energy Dispersive X-ray (EDX).

Figure 1. Graphene powder

3. Results and discussion

3.1. Morphology and structural properties

To examinate the structural properties of the graphene powder UV-VIS spectroscopy, SEM, TEM and EDX were investigated.

3.1.1. *UV-VIS spectroscopy.*

As shown in (Fig.2), the absorbance peak at 272nm indicates that the electronic conjugation within the reduced graphene sheets was revived upon reduction of graphene oxide (GO) which its absorbance band normally situated at 230nm was red_shifted after hydrazine attack. This result is in agreement with the one done in the paper [6] confirming the formation of reduced graphene oxide (RGO).

3.1.2. *SEM imaging.*

The SEM images obtained using Nova NanoSEM 230 Electron Microscope show typical large flakes of graphene powder (Fig.3). The transparency of these flakes confirms the successful exfoliation of graphene sheets and indicates that each flake is composed of a few individual graphene sheets as well as the large scale production.

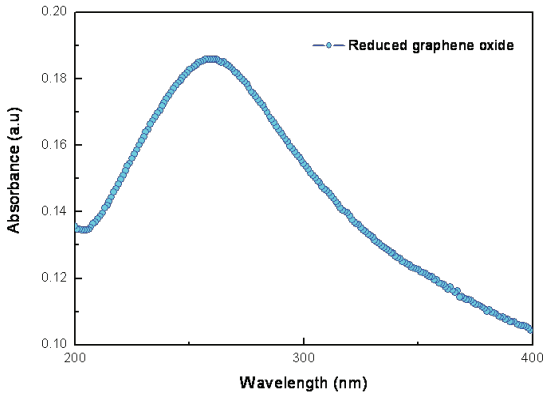


Figure 2. Uv-Vis spectrum shows the reduction of the exfoliated graphene oxide.

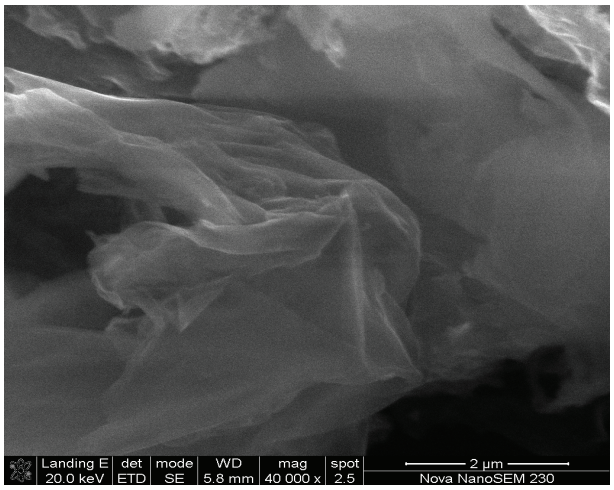


Figure 3. These two figures have been placed side-by-side to save space. Justify the caption.

3.1.3. *HRTEM imaging.*

High-resolution TEM of our graphene powder dispersed in ethanol revealed a wide and flat individual graphene sheets without any fold along its length, with straight and nearly atomically smooth edges without any discernible roughness or parallel lines related to a double layer, or more layers. HRTEM images confirm that graphene sheets are fully exfoliated and the final graphene product is composed of single layers.

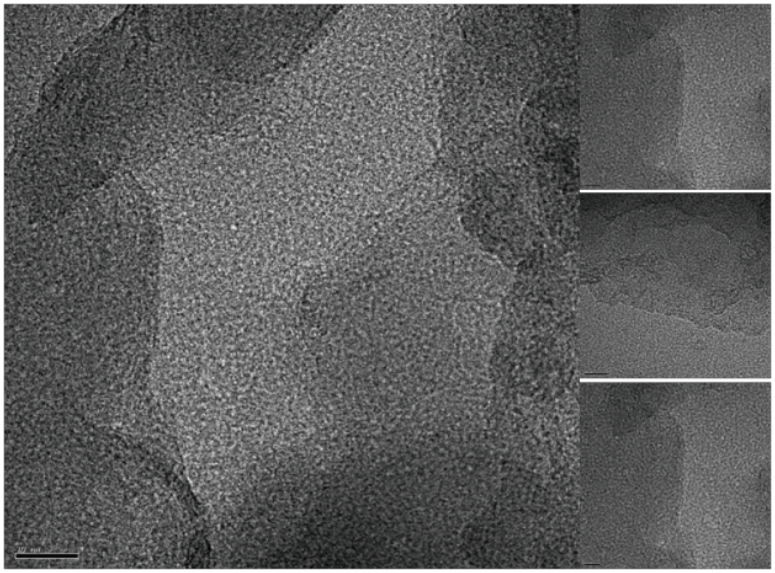


Figure 4. HRTEM of our prepared monolayered graphene powder

3.1.4. EDX analysis.

EDX analysis was done for different area of the graphene powder. From the EDX spectrum results in (Fig.5) chemical residues and Oxygen atoms are totally removed and the percentage of carbon atoms in our samples was higher than 99% confirming the purity of this product.

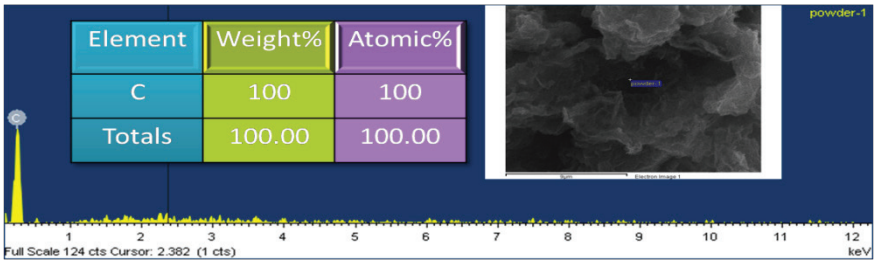


Figure 5. Graphene powder

3.2. Vibrational properties

3.2.1. Raman spectroscopy

During the oxidation of graphite powder process in the presence of concentrated acids using Hummers method, the original extended conjugated π-orbital system of the graphite was destroyed and oxygen-containing functional groups were inserted into the carbon skeleton. After reduction with hydrazine, oxygen groups were then replaced with nitrogen ones. However, after a purification treatment, these functional groups derived from the intensive oxidation and hydrazine hydrate were eliminated, which can be proved by the existence of D, G and G' bands only, reflecting the skeletal vibration of graphene sheets[7] and the remarkable decrease of D band intensity due to purification. The G band that originates from in-plane vibration of sp^2 carbon atoms is doubly degenerate (TO and LO) phonon mode (E_2g symmetry) at the Brillouin zone center. The 2D band originates from a two phonon double resonance Raman process (figure 6). The D band is attributed to defects in graphene nanosheets. These results indicate that the graphite oxide has been successfully exfoliated and purified to good quality graphene sheets.

3.2.2. FTIR spectroscopy.

FTIR spectra of GO, RGO and purified RGO are shown in figure 7. The absence of many peaks related to the remaining of oxygen groups generated by the chemical oxidation and the dramatic decrease of other peaks intensities is a strong indication of the exfoliation and the good quality of our purified RGO, at 3400 cm^{-1} (O-H stretching vibrations), at 1620 cm^{-1} (skeletal vibrations from unoxidized graphitic domains), at 1220 cm^{-1} (C-OH stretching vibrations), at 1060 cm^{-1} (C-O stretching vibrations) [8], and stretching vibrations from C=O at 1720 cm^{-1} due to the remaining of carboxyl groups even after hydrazine reduction and the vibration modes of epoxide (C-O-C) at (1230-1320 cm^{-1}) [6].

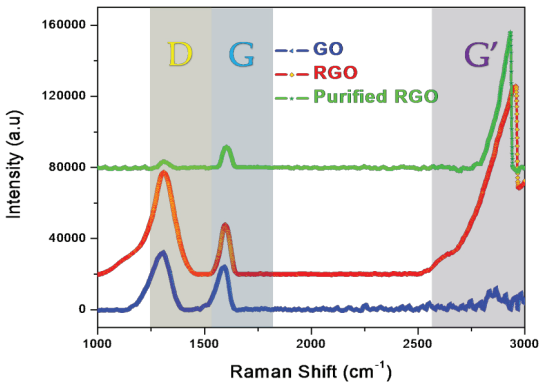


Figure 6. Raman spctra of graphene oxide and graphene.

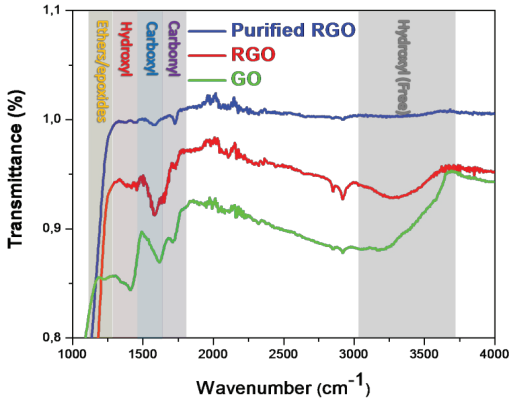


Figure 7. FTIR spectra of graphene oxide and graphene.

4. Conclusion

The obtained results in this work confirm that graphene sheets are fully exfoliated and the final graphene product is composed of a few good quality layers. This work opens the way to combine graphene properties with other nanostructured materials to generate promising nanocomposites for optoelectronic applications as well as its integration in large fields of applications.

References

[1] Geim A K and Novoselov K S 2007 Nat. Mater. 6 183
[2] Robinson J T, Perkins F K, Snow E S, Wei Z Q and Sheehan P E 2008 Nano Lett. 8 3137
[3] Jung J, Dikin D, Piner R D and Ruoff R S 2008 Nano Lett.12 4283
[4] Ferrari A C, Meyer J C, Scardaci V, Casiraghi C, Lazzeri M, Mauri F, Piscanec S, Jiang D, Novoselov K S, Roth S and Geim A K 2006 Phys. Rev. Lett. 97 187
[5] Hummers W S Jr and Offeman R E 1958 J. Am. Chem. Soc 80 1339
[6] Li D, Muller M B, Gilje S, Kaner R B and Wallace G G 2008 Nature Nanotech.3 101
[7] Stankovich S, Dikin D A, Piner R D,. Kohlhaas K A, Kleinhammes A, JiaY, Wu Y, Nguyen S T, Ruoff R S 2007 Carbon 45 1558
[8] Xu Y X, Bai H, Lu G W, Li C and Shi G Q 2008 J. Am. Chem. Soc 130 5856

Synthesis and characterization of PbS nanostructures using the chemical bath deposition method

L.F. Koao[1], F. B. Dejene[1*] and H.C. Swart[2].

[1]Department of Physics, University of the Free State (Qwaqwa Campus), Private Bag X13, Phuthaditjhaba, 9866, South Africa
[2]Department of Physics, University of the Free State, P.O. Box 339, Bloemfontein, 9300,South Africa.

*Corresponding author: Tel: +27 58 718 5307; Fax: +27 58 718 5444;
E-mail: dejenebf@ufs.ac.za

Abstract. Crystalline lead sulfide (PbS) nanorods were synthesized by chemical bath deposition (CBD) method. The lead acetate concentrations, the pH and deposition time were varied in the deposition process to determine the effect thereof. The final yields were characterized for structural, morphology and optical properties. The X-ray diffraction (XRD) spectra of the PbS nanorods correspond to the various planes of a single phase cubic PbS. The average particle sizes calculated using the XRD spectra were found to be 14±0.5 and 16±0.5 nm for samples stirred for 10 and 5 minutes respectively. The particle size was found to be dependent on the pH and the concentration of lead acetate used. Scanning electron microscopy (SEM) micrographs depict nanorod structures for a high molar concentration of lead acetate and a spherical shape for a low molar percentage. Energy dispersive X-ray (EDS) analyses confirm the presence of all the expected elements. The solid powder nanorods show good optical properties with high absorptions in the UV and visible regions. The band gap energies were estimated to be 2.07 to 3.02 eV which are higher than the bulk PbS band gap of 0.37 eV.

1. Introduction

It is well known that the synthesis of semiconductor one-dimensional nanostructures has attracted much attention in nanotechnology for several years. Designing extremely narrow nanorods with well defined shapes and sizes in the strong confinement regime (rod radius, r<< Bohr radius, a_B) remains a big challenge in nanotechnology. The need for such nanorods of different materials is increasing due to their possible application as nanoscale building blocks for different applications.The interest is particularly true for the IV-VI class of materials which can cover the technologically important wide spectral range from mid-infrared to the visible region, yet have been studied to a lesser extend compared to II-VI semiconductors [1]. Lead sulfide has a narrow and direct band gap (0.37 eV) at 300 K. PbS nanorods are potentially useful in electroluminescent devices such as light emitting diodes, high speed switching and optical sensors because of large band gap of above 2 eV [2 and 3]. The band gap energy of PbS can be widened from UV to visible region by forming nanocrystal because of their quantum size effects [4 and 5]. The chemical bath deposition (CBD) is a simple, cheap and convenient process to prepare semiconducting materials. The more recent interest in all things 'nano' has provided a boost for CBD, since it is a low temperature, solution (almost always aqueous) technique and crystal size is often very small. The aim of this paper is to investigate the effect of different growth parameters on the material properties of PbS.

2. Experimental

The PbS nanorods were prepared by changing different molar concentrations of lead acetate and holding constant thiourea and ammonia. The preparation of PbS nanorods was carried out using the following procedure: 50 ml of X Lead acetate solutions (X is the different molar concentrations of lead acetate) were taken in a 200ml beaker. Then, 50 ml of thiourea solution was added in the reaction bath

and the mixture was stirred for few minutes, Following that 50 ml of ammonia solution was added drop wise into the mixture, while continuing stirring for different times (for few minutes or immediate after the precipitates has formed, 5 and 10 minutes). The temperature of the bath was then allowed to increase up to 80°C. After the precipitates were formed the pH was measured and the precipitates were filtered and washed with 50 ml of ethanol. The resulting powders were dried at ambient conditions for 3 days and ready to be characterized using various characterization technique.

3. Results and Discussion

3.1 Structural and Composition analysis

Fig. 1 (a) shows XRD pattern of the PbS nanorods synthezed at constant molar concentration of lead acetate of 0.06M and different stirring times, respectively. The four main peaks observed in the diffractogram around 25.99°, 30.11°, 43.09° and 51.03° reveal a cubic lattice structure of PbS [6].

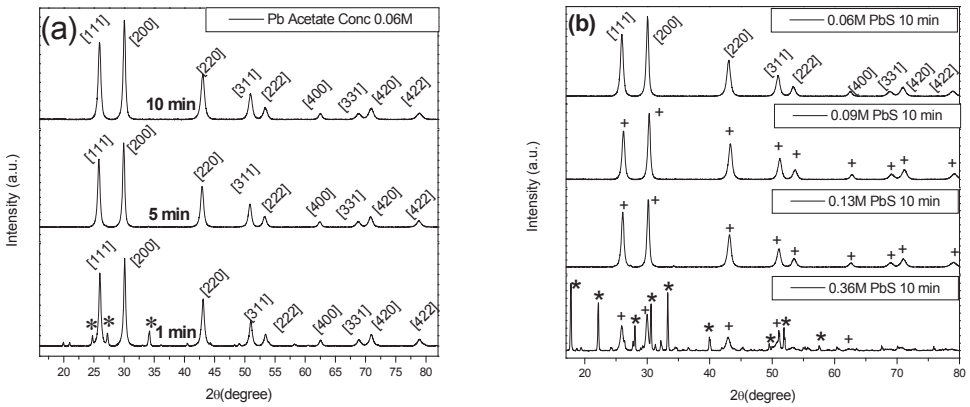


Figure 1: X-ray powder diffraction patterns for undoped PbS prepared for (a) different stirring times at constant molar concentrations and (b) different molar concentrations of lead acetate at constant time (10 minutes).

There were also additional peaks (*) which may be attributed to unreacted Pb $(CH3COO)_2.2H_2O$, $(NH_2)_2CS$ and NH_3 precursors or other impurities phases such as PbO and PbOH. The calculated value of the cell constant of 5.9597 Å matched perfectly with standard data available in JCPDS card no. (05-0592, a=5.936 Å). The average size of the as-prepared nanocrystals can be calculated from the full width at half maximum (FWHM) of the diffraction peaks using the Debye-Scherrer formula [7]:

$$t = \frac{0.9\lambda}{\beta \cos\theta_B} \quad (1)$$

where t is the particle diameter of nanocrystallites, λ is the X-ray wavelength (0.154 nm), β is the diffracted full-width at half maximum (in radian) caused by the crystallites and θ is the Bragg angle. All major diffraction peaks for all samples are chosen to estimate the average size of the nanocrystals by least square method. The average particle size of the PbS synthesized using a 0.06M lead acetate for 1, 5 and 10 minutes are estimated to be 16±0.2, 15±0.4 and 14±0.5 nm, respectively. The values increased to 18±0.7, 17±0.5 and 16±0.1 nm, after the lead acetate concentration increased to 0.13M. The sizes of the products are affected by the precursor concentration in the solution and stirring time. In Fig. 1 (b) for the sample stirred for 10 minutes the only difference is that at the high molar concentration of 0.36M of lead acetate there are many peaks that may be due to the clustered lead

acetate (*) but at low molar concentrations of lead acetate many peaks diminishes and the other peaks are clearly distinguishable which shows the PbS (+) and the line width broadens.

It is clear that the average crystallite size increased with an increase in the molar concentration of lead acetate, Fig. 2 (a) (for molar concentrations of 0.04M, 0.06M, 0.09M and 0.13M respectively). It was also observed that the estimated average particle size decreases with an increase in the pH (See Fig 2(b)). The reason for this is not clear at this stage.

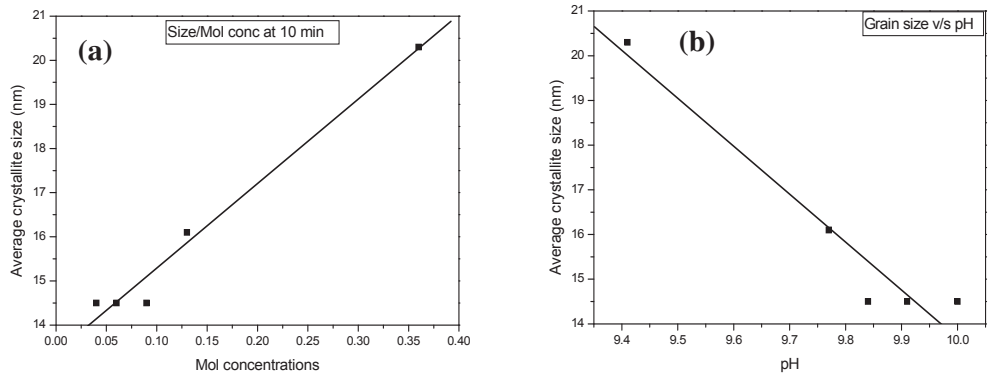


Figure 2: The graph of average crystallite sizes dependence on the (c) molar concentrations and (d) pH all samples stirred for 10 minutes and dried at room temperature.

3.2 Surface morphological analysis

Fig. 3 (a) and (b) show SEM images of the as prepared PbS nanocrystals, composed of nanorods and spherical particles respectively.

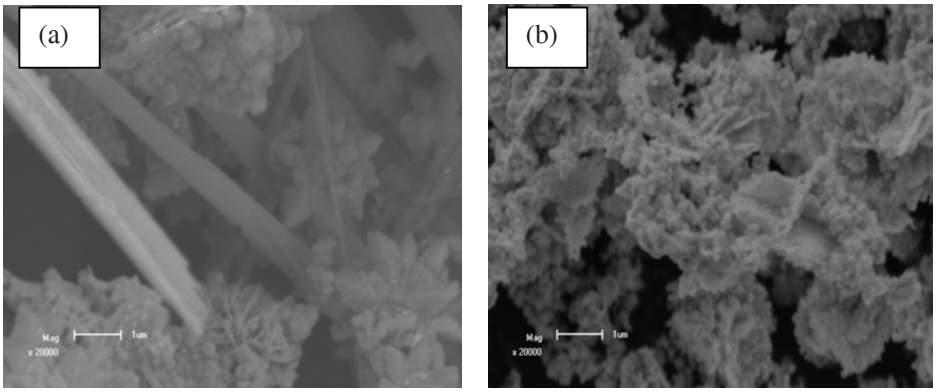


Figure 3: SEM images of PbS nanostructures prepared using (a) 0.46M and (b) 0.04M lead acetate both stirred for 5 minutes.

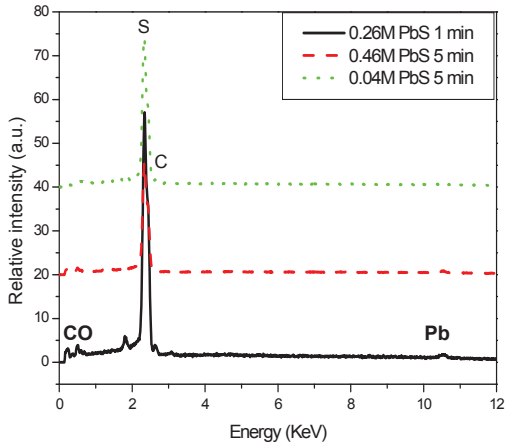


Figure 4: A representative EDX spectrums of the undoped PbS nanocrystals.

From the micrographs, it is clear that the use of low concentration of lead acetate resulted into clustered composed of nearly spherical morphology, while high concentration resulted into nanorods with different diameter and length. The EDS spectrum in Fig. 4 shows the presence of Pb and S, with slight excess of sulphur for all the samples. The existence of O and C impurities is believed to have originated from the surface contamination in the atmosphere and also the double sided carbon tape.

3.3 Optical analysis

Fig. 5(a) shows the typical variation of absorbance with wavelength for the PbS prepared at different molar concentration of lead acetate but synthesized at constant time, respectively. At low molar concentrations of 0.04M and 0.13M of lead acetate the UV-visible spectrum shows the absorption bands with λ_{max} at 247 and 565 nm, 247 and 573 nm, respectively. This implies the presence of anisotropic nanoparticles like PbS spherical particles. λ_{max} at 247 nm may be due to the impurities such as PbOH and PbO. At high molar concentrations of 0.36M of lead acetate the UV-visible spectrum shows one broad absorption band with λ_{max} at 412 nm indicates the presence of PbS particles with the wide size distribution or anisotropic PbS nanorods. The energy band gap of these materials was estimated using the following equation [8],

$$E_{gN} = \frac{hc}{\lambda_{max}} \qquad (2)$$

where E_{gN} (energy gap of nanoparticles and nanorods) is the separation between the valence and the conduction bands, *h* is the Plank constant, *c* is the speed of light in vacuum and λ_{max} is the wavelength at the maximum absorption. The value of the optical band gap energy of the PbS nanorods and nanoparticles synthesized with 0.04M, 0.13M and 0.36M lead acetate have been estimated to be 2.20 eV, 2.17 eV and 3.02 eV, respectively, as calculated using Eq.(2). It is found that the band gap energy of the resulting nanoparticles and nanorods shows a noticeable increment as compared with that of PbS bulk materials (0.37 eV). The band gap energy of the PbS nanorods is bigger than the bandgap of the nearly spherical PbS nanoparticles. Fig. 5(b) shows the typical variation of absorbance with wavelength for the PbS prepared at different synthesis times but at constant molar concentration of lead acetate of 0.06M, respectively. At low molar concentrations of 0.06M of lead acetate for (i) 1 minute, (ii) 5 minutes and (iii) 10 minutes the UV-visible spectrum shows the absorption bands with λ_{max} at 244 and 559 nm, 244 and 573 nm, 244 and 584 nm, respectively. This implies the presence of anisotropic particles like PbS spherical nanoparticles and nanorods. Their estimated band gap energy values are 2.22 eV, 2.14 eV and 2.07 eV.

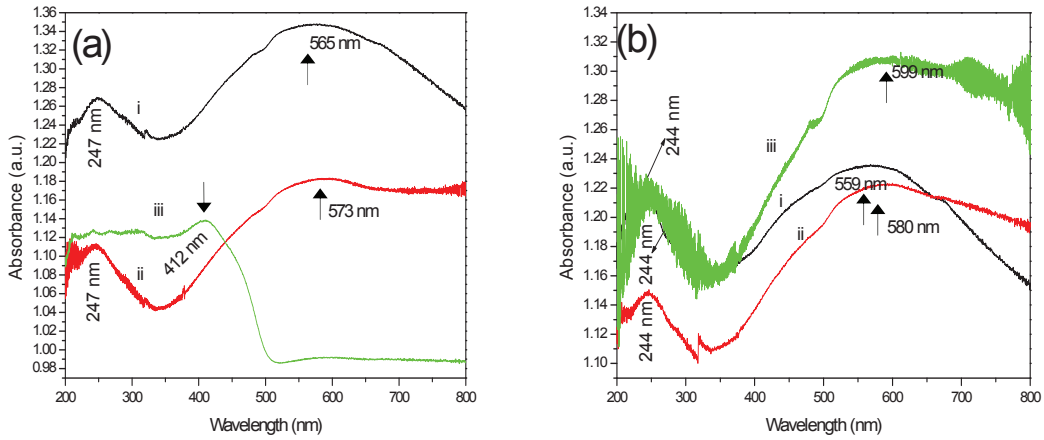


Figure 5: The absorbance spectra of PbS nanorods, prepared at (a) different molar concentration of lead acetate but synthesized at constant time and (b) constant molar concentration of lead acetate but synthesized at different times.

4. Conclusion

The PbS nanostructures have been successfully synthesized by the chemical bath deposition technique at 80°C. XRD showed that the structure of material is a single cubic phase. UV showed that the band gap energy of the nanostructures depend on the molar concentration of lead acetate. It was observed that the nanorods structure have bigger than the band gap energy than nearly spherical nanoparticles.

Acknowledgements

The author would like to acknowledge the National Research Foundation and Department of science and technology and the University of the Free State for financial support. We are also grateful to people currently of Centre for Microscopy at UFS for EDS, SEM and X-ray diffraction measurements.

References

[1] Acharya S, Gautam U K, Sasaki T, Bando Y, Golan Y, Ariga K 2008 *J. Am. Chem. Soc.* **130** 4595
[2] Jing S, Xing S,Wang Y, Hu H, Zhao B, Zhao C 2008 *Mater. Lett.* **62** 977
[3] Wu M, Zhong H, Jiao Z, Li Z, Sun Y 2008 *Colloids Surf. A: Physicochem. Eng. Aspects* **35** 313
[4] Chestnoy N, Hull R, Brus LE 1986 *J. Chem. Phys.* **85** 2237
[5] Wang Y, Suna A, Mahler W, Kawoski R 1987 *J. Chem. Phys.* **8** 7315
[6] Wang Z, Zhao B, Zhang F, Mao W, Qian G, Fan X 2007 *Mater. Lett.* **61** 3733
[7] Cullity BD 1978, 1956 *Elements of X-ray Diffraction (2nd Ed)*, (Addison Wesley) p 285-284.
[8] Khiewa PS, Radimana S, Huanga NM, Soot Ahmada Md, Nadarajahb K *2005 Mater. Lett.* **59** 989

Ion irradiation effects on the formation of metal nanoparticles in crystals

Anna Kozakiewicz[1], Trevor Derry[1], Paul Franklyn[2], S R Naidoo[1] and Chris Theron[3]

[1]School of Physics and DST/NRF Centre of Excellence in Strong Materials, University of the Witwatersrand, Johannesburg, South Africa

[2]School of Chemistry, University of the Witwatersrand, Johannesburg, South Africa

[3]Department of Physics, University of Pretoria, Pretoria, South Africa

E-mail: anna.kozakiewicz@students.wits.ac.za

Abstract. Metal nanoparticles (MNP) were synthesized by room temperature ion implantation of high fluences (from $6x10^{16}$ to $5x10^{17}$ cm^{-2}) of 150 keV Ag and 15 - 22 MeV Au ions into Al_2O_3 and MgO single crystals. Optical absorption (OA) spectra show surface plasmon resonance (SPR) bands characteristic of the implanted metal ions. Upon annealing (300°-1200°C) the optical response of the metal nanostructures changes, related directly to their morphology, shape and size. High resolution transmission electron micrographs indicate that the particles are crystalline. The implanted ions profiles were obtained from 1.6 MeV He^+ Rutherford Backscattering Spectroscopy (RBS) for the silver implanted substrates, and High Resolution Transmission Electron Microscopy (HRTEM) revealed buried layers of implanted ions in the form of metal nanoparticles (MNP).

1. Introduction

A study of properties of colloidal small particles has a long history and there is a large body of experimental [1] – [6] and theoretical work [7] - [10] in this field of study. Metal nanoparticles can be produced by various methods and among them is ion implantation. Ion implantation is a well established technique, which by the injection of energetic ions into the near surface region of a substrate modifies optical properties of the implanted material. The formation of nanoclusters requires implantation with high fluences, above a threshold value. The depth and distribution of the implanted ions were calculated from SRIM [12] and verified by Rutherford Backscattering Spectrometry.

The ion implantation technique, which overcomes the substrate solubility limit of high dopants concentration, is well suited for inducing formation of metal – host medium nanocomposites.

The mechanism of defect creation for two ion species used in this work is different. The nuclear stopping is responsible for damage in 150 keV Ag ions while electronic energy loss is dominant in the MeV range for Au implants. Defects produced by nuclear collisions cause displacement of atoms from their lattice positions. The electronic, nonelastic mechanism

involves electronic excitations and the defects are formed mainly in the anion sublattice during lattice relaxation processes.

Irradiation of the oxides by energetic particles leads to the displacement of an atom into an interstitial position leaving a vacancy behind, forming a Frenkel defect pair. In the oxygen sublattice this could result in the formation of three kinds of vacancy-interstitial defect pairs, depending on the charge state with which interstitial oxygen leaves the vacancy: either a neutral interstitial O atom and vacancy or a F-like centres which consists of a vacancy with one or two trapped electrons (the so-called F^+ and F centres) and their complementary O^- and O^{2-} defects [7]. The aggregation of F-type defects is responsible for formation of the nanoparticles.

The sapphire lattice has different diffusion rates for defects for different crystal orientations. The highest mobility of point defects and introduced ions is along the c axis. Hence the defect migration and their separation is direction dependant.

The diffusion rates are dependent on the ion species and are much lower for Au ions.
The as-implanted and annealed samples were studied by RBS, HRTEM and optical absorption to get a detailed picture of the structures formed during the implantation. This knowledge can be used to synthesize new composite materials of desired optical properties. Plasmon resonances play a very important role in the creation of new materials.

2. Experimental details

We have implanted Al_2O_3 and MgO single crystals from MTI and Crystec of different orientations using 150 keV Ag^+ and C^+ ions using a 200-20A2F Ion Implanter.

Implants of 15 MeV Au^{3+} were performed on a Van de Graaff tandem accelerator. All implantations were done without temperature control at iThemba LABS (Gauteng). A fluence of $3x10^{17}$ ions/cm^2 was used for carbon implantation. Gold implants were done with ~$1x10^{17}$ cm^{-2} fluences and silver with a range of doses from $3x10^{16}$ to $5x10^{17}$ ions/cm^2.

The annealing process was carried out in a horizontal tube furnace for 1h with intervals of 100° from 300° up to 1200°C in reducing atmosphere (Ar + 4% H) and allowed to cool in air to ambient temperature. After each annealing step, optical absorption (OA) was measured on a Varian Cary 500 UV – VIS double beam spectrometer with a non-irradiated sample in the reference beam.

RBS spectra were obtained on the Van De Graaff accelerator at the University of Pretoria using 1.6 - 2 MeV He^+ beams with Si surface barrier detector at 165° in IBM geometry.

HRTEM specimens were prepared from sapphire platelets of the average thickness of 200 nm and diameter ranging from 3 – 10 µm, deposited and implanted on TEM carbon film coated grids and analyzed on 200 kV Tecnai F20 at the University of the Western Cape.

3. Results and discussion

Annealing from 300° to 1100°C removed F-type centre absorption bands from the OA spectra in both materials. Colour centres aggregate upon heat treatment producing optical absorption bands in the visible range of the spectrum with peak positions which correspond to surface plasmon resonance (SPR) bands of the implanted ions.

In 15 MeV Au^{3+} implanted sapphire crystals a band appears at 530 nm (2.3 eV) after annealing at 1000°C. The peak reaches its maximum after 1100°C annealing and the transparent, colourless sample acquires a pink colour. The peak position is associated with the SPR of the Au MNP absorption band (Figure 1).

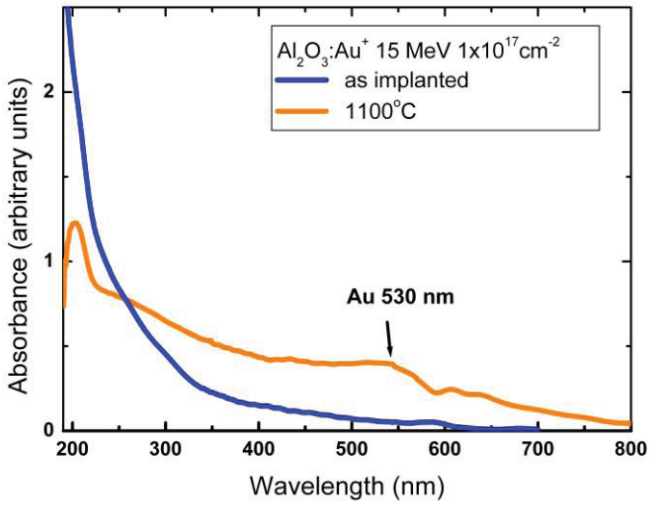


Figure 1. Optical absorption spectra of 15 MeV gold implanted sapphire (fluence of $1x10^{17}$ ions/cm^2) as implanted and after annealing to 1100°C in reducing atmosphere.

Al_2O_3 implanted with a high fluence of silver ions shows a peak at 426 nm (2.9 eV) (Figure 2) in as-implanted samples. This peak is attributed to SPR absorption by silver nanoparticles. Similar behaviour is observed in MgO crystals (Figure 3) implanted with Ag^+ ions.

However for lower Ag fluences of $3x10^{16}$ Ag ions/cm^2 in MgO, the band is formed only after 1100°C annealing (Figure 3).

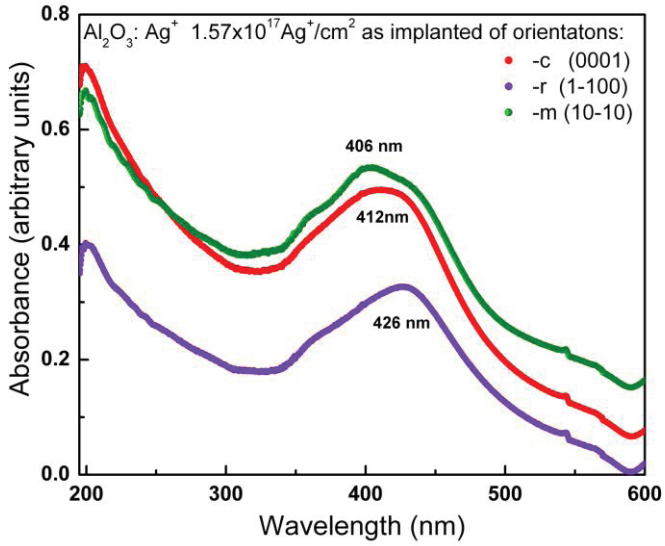


Figure 2. Al_2O_3 with different crystal orientations implanted with high fluence ($1.6x10^{17}$ ions/cm^2) of silver ions. Peaks at 412 nm - 426 nm are SPR OA bands due to the presence of colloidal silver. Their spectral position depends on the

crystal orientation. Spectra were measured at room temperature (RT) from as-implanted material.

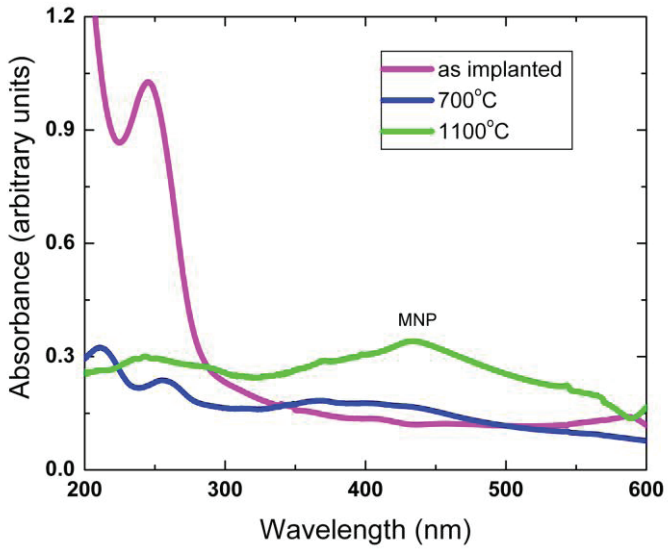


Figure 3. MgO implanted with $3x10^{16}$ Ag ions/cm^2 and annealed. The SPR band appears after 900°C annealing and reaches its maximum intensity of 0.34 (arb.units) after 1100°C anneal.

In the carbon implanted sapphire samples (with a fluence of $3x10^{17}$ C/cm^{-2}) (not illustrated here) the 245 - 260 nm band develops after annealing in reducing atmosphere at 700° - 900° C depending on the crystal orientation. After the band intensity reaches maximum its position shifts gradually to shorter wavelengths with decreasing intensity. The band origin is possibly due to Mie scattering from carbon nanoparticles [11].

High Resolution TEM micrographs (Figure 4) of silver implanted samples indicate the presence of Ag crystalline nanoparticles in both the alumina and, possibly, also in the holey carbon film. The sizes varied from 3 to 7 nm for as-implanted specimens. Image analysis by comparison of simulated diffraction patterns and the power spectra of the images, indicated that the alumina retained its crystallinity. The depth profile obtained from RBS measurement has a projected range of 41 nm, in agreement with SRIM calculations [12].

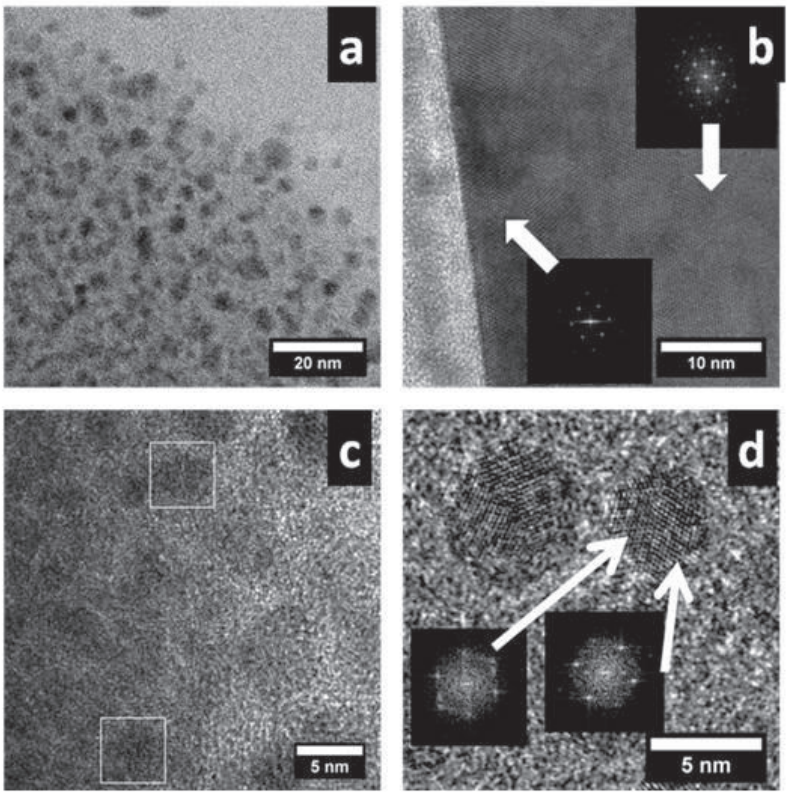


Figure 4. (a) Sapphire substrate with silver nanoparticles after RT ion implantation at high fluence $1x10^{17}$ ions/cm^2. (b) HRTEM image of the edge of a sapphire crystal showing the distortion (lower power spectrum) near the edge where the Ag is implanted compared to the bulk (upper power spectrum). (c) HRTEM image showing Ag nanoparticles in the sapphire (left) and carbon support film (right) with the boxed areas highlighting crystalline Ag nanoparticles in the sapphire. (d) Power spectra showing the twinning of the Ag nanoparticles in the carbon support film

4. Conclusions

Low diffusivity of gold ions results that the formation of MNP in 15 MeV gold ion implants in sapphire with fluences near $1x10^{17}$ ions/cm^2, takes place after thermal annealing to temperatures above 1000 °C. A similar observation is made for low fluence ($3x10^{16}$ ions/cm^2) 150 keV Ag implants in MgO.

In contrast, MNP were formed in "as implanted" specimens for high fluence (near $1.5x10^{17}$ ions/cm^2) of Ag ions. This result can be expected if the ion fluence is above a threshold value for spontaneous colloid precipitation in an implanted substrate. By implanting sapphire platelets under similar conditions, it is observed using HRTEM that the implanted sapphire substrate retains its crystallinity and the Ag MNP that are formed are crystalline.

References

[1] Alves E, Marques C, da Silva R C, Monteiro T and McHargue C 2009 *Surface & Coatings Technology* **203** 2357

[2] McHargue C J, Farlow G C and White C W, Appleton B R, Angelini P and Naramoto H 1985 *Nuclear Instruments and Methods in Physics Research* B **10/11** 569

[3] Schnohr C S, Wendler E, Gärtner K, Wesch W and Ellmer E 2006 *Journal of Applied Physics* **99** 123511

[4] Marques C, Alves E, da Silva R C, Silva M R and Stepanov A L 2004 *Nuclear Instruments and Methods in Physics Research B* **218** 139
[5] Marques C, da Silva R C, Wemans A, Maneira M J P, Kozanecki A and Alves E 2006 *Nuclear Instruments and Methods in Physics Research* B **242** 104
[6] Stepanov A L and Khaibullin I B 2005 *Review of Advanced Materials Science* **9** 109
[7] Kotomin E A, Popov A I 1998 *Nuclear Instruments and Methods in Physics Research* B **141** 1
[8] Kreibig U and Vollmer M 1995 *Optical Properties of Metal Clusters* (Berlin: Springer)
[9] Hughs A E and Jain S C 1979 *Advances in Phys*ics **28** 717
[10] Doyle W T 1958 *Physical Review* **111** 1067
[11] Alves E, Marques C, Sáfrán G and McHargue Carl J 2009 *Nuclear Instruments and Methods in Physics Research* B **267** 1464
[12] Ziegler J F and Biersack J P SRIM 2008 Stopping and range of ions in matter. Monte Carlo code http://www.srim.org/.

The impurity levels of lanthanide ions in silica

R E Kroon, H A A Seed Ahmed and M A Gusowski

Department of Physics, University of the Free State, Bloemfontein 9300, South Africa

E-mail: kroonre@ufs.ac.za

Abstract. Silica is a basic material of technological importance for optics, microelectronics, photonics and fibre optics and it has been used as a host material for a variety of luminescent lanthanide ions due to its chemical stability and non-hygroscopic nature. Although for lanthanide ions the 4f electron energy levels are shielded from the host environment by the filled outer 5s and 5p orbitals, making the f-f transitions relatively insensitive to the host, f-d transitions involving the unshielded 5d state are host dependent. The absolute positions of the 4f and 5d states relative to the valence and conduction bands of the host also affect quenching and charge trapping phenomena and so they are required for proper modelling of phosphor performance. Recently a model has been formulated to determine the energy levels of any of the thirteen divalent lanthanides relative to the band edges using only three host dependent parameters. Another two host dependent parameters are required to also establish the absolute levels of the trivalent ions. In this paper a scheme for the energy levels of both the divalent and trivalent lanthanide ions in silica is proposed making use of experimental data. It is suggested that although the location of the divalent europium ion f-level above the valence band can be located by using the charge transfer energy of trivalent europium, this process cannot be generalized to find the location of the trivalent cerium ion f-level above the valence band using the charge transfer energy of tetravalent cerium as has been suggested.

1. Introduction

Silica is a basic material of technological importance for optics, microelectronics, photonics and fibre optics. Its high absorption edge energy makes it particularly useful for UV applications and it has been used as a host material for a variety of luminescent lanthanide ions due to its chemical stability and non-hygroscopic nature. Specific applications include optical amplifiers doped with erbium for the telecommunications industry [1] and x-ray scintillation material doped with cerium for radiation sensing [2].

For lanthanide ions the 4f electron energy levels are shielded from the host environment by the filled outer 5s and 5p orbitals, so that the transitions between the 4f states (and therefore the luminescence wavelengths) are relatively insensitive to the host. For this reason little attention has been paid until recently to the location of the impurity levels of the lanthanide ions within the energy gaps of their hosts. However, luminescence from some lanthanide ions, e.g. cerium, occurs due to f-d transitions from the unshielded 5d state of which the energy relative to the f-states is therefore host dependent. The absolute positions of the 4f and 5d states relative to the energy gap of the host also affect quenching and charge trapping phenomena and so they are required for proper modelling of phosphor performance [3]. In addition, the interaction of lanthanide ions with defect levels of a host can only be studied if the energy levels of the lanthanide ions are known.

Although the use of lanthanides in luminescence phosphors became widespread about half a century ago [4], it is only recently that a phenomenological model capable of predicting the lanthanide energy levels has become available. This is primarily due to the work of Dorenbos and was applied initially to CaF_2 and YPO_4 [5] and more recently to other materials e.g. GaN [6] and $Al_xGa_{1-x}N$ [7] where it was possible to explain the lack of blue luminescence from terbium ions unless $x > 0.38$ due to overlap of the 4f $^5D_3$ level with the conduction band. Only five host-specific parameters are required: the bandgap of the host and four others, two each for divalent and trivalent ions to locate the ground states of the f and the d levels respectively. In principle the data can be obtained from any of the lanthanide ions, since the data compiled by Dorenbos [8-14] relates to the values between the different lanthanide ions. In practice Eu^{+2} is usually used for the divalent ions and Ce^{3+} for the trivalent ions, as this data is most readily obtained experimentally. Determining the parameters to locate the levels for a particular host is not always straightforward. In this paper a scheme for the energy levels of both the divalent and trivalent lanthanide ions in amorphous silica is proposed.

2. Model for the lanthanide impurity levels

As a first step to the model, Dorenbos [8] used experimental data to show that the differences between the f-d transition energies of trivalent lanthanide ions were independent of the host. This allowed new estimates of the free ion f-d transition energies and, using at least one experimental f-d transition energy for any trivalent lanthanide ion in a host, to give the crystal field depression $D(A,3+)$ for that host A [9]. The process was analogous for the divalent ions [10], yielding the crystal field depressions $D(A,2+)$ for divalent ions in different hosts [11].

Given the f-d transition energies E_{fd}, if one can determine either the energy from the valence band to the f-level E_{Vf} or from the d-level to the conduction band E_{dC} one can place the levels absolutely. Note that

$$E_{Vf} + E_{fd} + E_{dC} = E_{VC}$$

where E_{VC} is the bandgap. To first approximation, an electron promoted to the 5d level of any lanthanide ion interacts with the crystal in the same manner, since the differences between the lanthanides occurs only for the inner shielded 4f electrons. Therefore the 5d energy levels of all the lanthanides are approximately equal [5]. In addition it was found that the charge transfer (CT) energy (E^{CT}) of a trivalent lanthanide ion provides a fair measure of E_{Vf} for the corresponding divalent ion [12]. By comparing small calculated changes in E_{dC} across the divalent lanthanides [13] with changes in E_{Vf} based on CT data, and using the known relative f-d transition energies together with

$$\Delta E_{Vf} + \Delta E_{fd} + \Delta E_{dC} = 0,$$

Dorenbos [12] was able to find the variation ΔE_{Vf} across the divalent lanthanides. Then if either the f- or the d-energy level of any divalent lanthanide ion can be found relative to either the valence or the conduction band, all the f- and d-levels for all the divalent ions can be computed. The value of E^{CT} for Eu^{3+}, giving $E_{Vf}(Eu^{2+})$, is often most easily obtained experimentally [14] and so ΔE_{Vf} is generally expressed relative to this ion. Dorenbos [12] also predicted the variation ΔE_{Vf} across the trivalent lanthanides. Although E^{CT} for tetravalent lanthanide ions is expected to be a fair measure of E_{Vf} for the corresponding trivalent lanthanide ions, this data is not so readily available and so instead $E_{dC}(Ce^{3+})$ is usually estimated and for the trivalent ions ΔE_{Vf} is generally expressed relative to cerium.

Therefore in addition to the bandgap of a particular host, one needs the crystal field depression of the f-d transitions and any one known absolute energy level to predict all the f- and d-levels of all the lanthanide ions. The last two parameters are needed for both divalent and trivalent ions, giving five host specific parameters in total. The free ion f-d transition energies and their relative values, together with the values of ΔE_{Vf} and ΔE_{fd} are given in Table 1. Note that recently Dorenbos [15] has published new estimates of ΔE_{Vf} for the divalent and trivalent ion which are used here. Older values have been

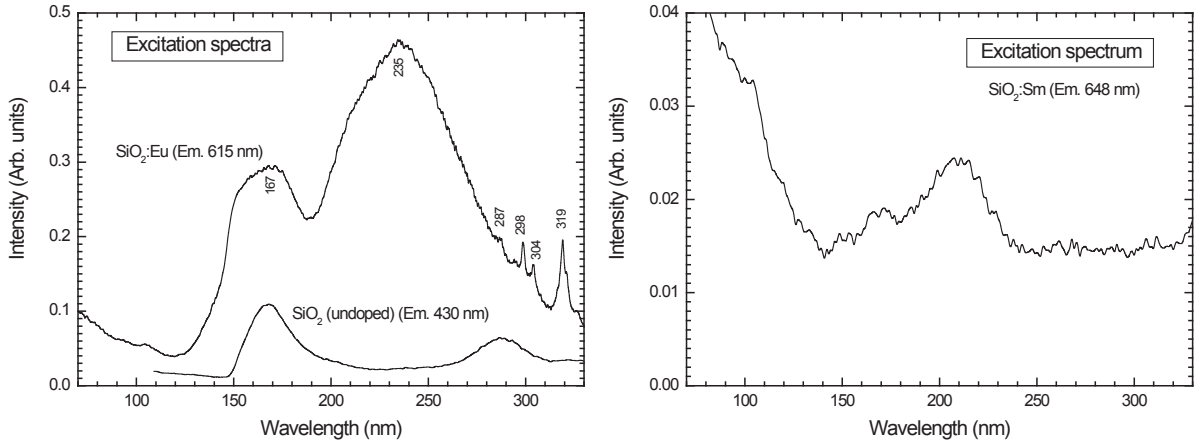
applied successfully to oxide and fluoride hosts [14], but for a sulphide host it the model did not work well and for that work the values were adapted to give a constant value for E_{dC} [16].

Table 1. Host independent data for energy level diagrams of lanthanide ions. The columns labelled Free give the f-d transition energy of the free ions. The next columns gives the relative values of the f-d energy, the energy from the valence band to the f-level ground state and the energy from the d-level ground state to the conduction band [8-15]. All values are in electron-volts.

Ion	Divalent ions, Ln^{2+} (changes relative to Eu^{2+})				Trivalent ions, Ln^{3+} (changes relative to Ce^{3+})			
	Free	ΔE_{fd}	ΔE_{Vf}	ΔE_{dC}	Free	ΔE_{fd}	ΔE_{Vf}	ΔE_{dC}
La	-0.94	-5.16	5.69	-0.53	-	-	-	-
Ce	0.35	-3.87	4.07	-0.20	6.12	**0.00**	**0.00**	**0.00**
Pr	1.56	-2.66	2.83	-0.17	7.64	1.52	-1.94	0.42
Nd	1.93	-2.29	2.50	-0.21	8.94	2.82	-3.29	0.47
Pm	1.96	-2.26	2.32	-0.06	9.31	3.19	-3.63	0.44
Sm	3.00	-1.22	1.24	-0.02	9.40	3.28	-3.74	0.46
Eu	4.22	**0.00**	**0.00**	**0.00**	10.58	4.46	-4.92	0.46
Gd	-0.10	-4.32	4.67	-0.35	11.81	5.69	-6.17	0.48
Tb	1.19	-3.03	3.23	-0.20	6.97	0.85	-0.88	0.03
Dy	2.12	-2.10	2.30	-0.20	8.58	2.46	-2.53	0.07
Ho	2.25	-1.97	2.44	-0.47	9.72	3.60	-3.55	-0.05
Er	2.12	-2.10	2.62	-0.52	9.46	3.34	-3.41	0.07
Tm	2.95	-1.27	1.77	-0.50	9.46	3.34	-3.18	-0.16
Yb	4.22	0.00	0.49	-0.49	10.83	4.71	-4.15	-0.56
Lu	-	-	-	-	12.22	6.10	-5.64	-0.46

3. Lanthanide impurity levels in amorphous silica

Amorphous silica samples doped with lanthanide ions were prepared with the sol-gel process using TEOS as a precursor and annealed at 1000°C. Emission and excitation measurement were performed on a Cary Eclipse fluorescence spectrophotometer having a xenon lamp as well as with synchrotron radiation using the SUPERLUMI station of HASYLAB at DESY. All measurements were made at room temperature, and excitation spectra from SUPERLUMI were corrected for the incident flux by comparison to the excitation spectrum of sodium salicylate.



(a) (b)

Figure 1. Excitation spectra for sol-gel silica measured at SUPERLUMI. (a) undoped and 1% Eu^{3+} doped silica (b) 1% Sm^{3+} doped silica. Samples were annealed at 1000°C in air.

Figure 1(a) shows the excitation spectrum of a sample doped with 1% Eu while monitoring the emission due to Eu^{3+} ions at 615 nm. The small sharp features around 300 nm are due to f-f transitions of Eu^{3+} [17], while the broad band at 235 nm (5.28 eV) is the CT band. A similar excitation spectrum was obtained by Yi *et al.* [18], but they only measured down to 250 nm. The band near 170 nm was also observed in undoped silica as well as when doping with different lanthanides and is attributed to defects in the silica [19]. The value of E^{CT} (Eu^{3+}) is used for the energy E_{Vf} (Eu^{2+}) and from the differences in Table 1, the absolute f-levels of the divalent lanthanide ions are established as in figure 2(a). The value found here agrees well with that for silicates listed by Dorenbos [14]. From figure 2(a) one may predict that E^{CT} (Eu^{3+}) will lie 1.24 eV above that of Eu^{3+} at 6.51 eV (190 nm). Recently Gutsov *et al.* [20] attributed a broad absorption peak in Sm doped sol-gel silica at 35714 cm^{-1} (4.42 eV, 280 nm) to E^{CT} (Sm^{3+}). However, this energy lies lower than that for Eu^{3+} and is not consistent with our results. Figure 1(b) shows the excitation spectrum of SiO_2:1% Sm^{3+} and the CT band is found at 210 nm, fairly close to the 190 nm predicted from the model. An excitation band near 280 nm, similar to that found by Gutsov *et al.* [20], is also found for undoped silica (figure 1(a)) and may be associated with defects [19].

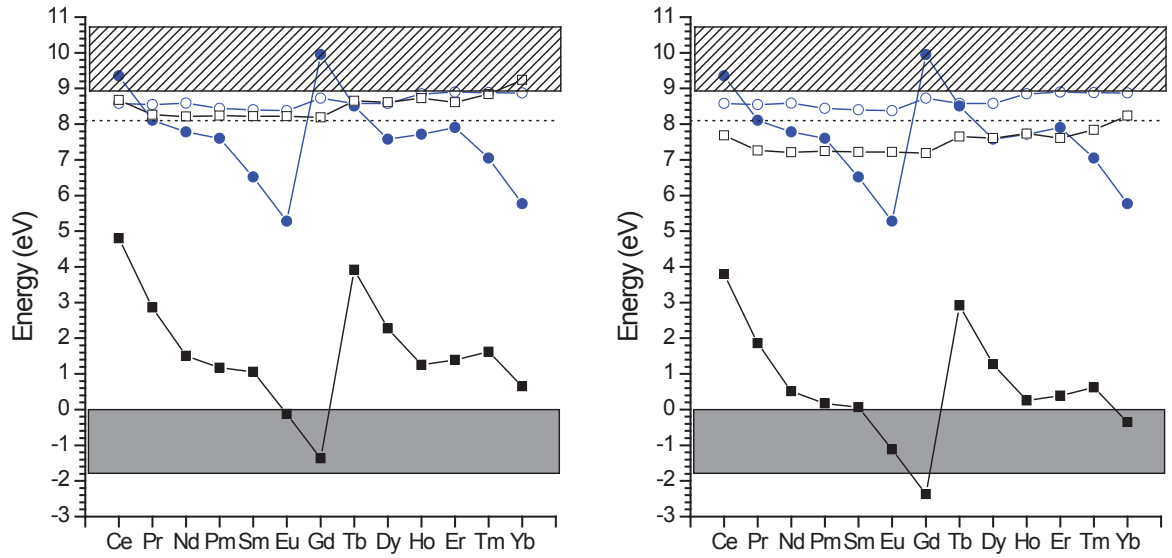
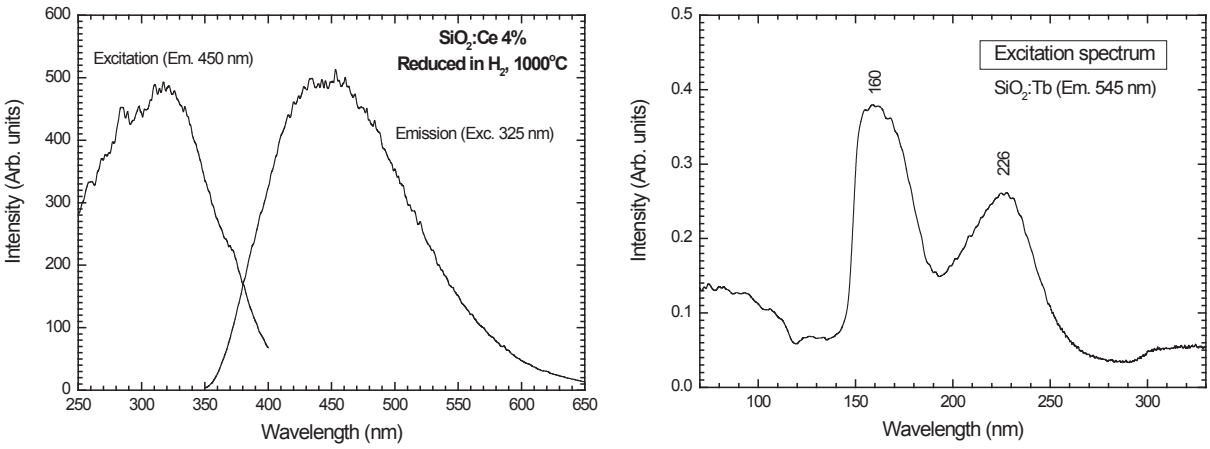


Figure 2. Location of the energy levels of divalent and trivalent lanthanide ions in silica. (a) Scheme based on associating E^{CT} (Ce^{4+}) with the value E_{Vf} (Ce^{3+}). (b) More realistic scheme obtained by reducing E_{Vf} (Ce^{3+}) in (a) by 1 eV.

To obtain the divalent lanthanide d-levels, the crystal field depression $D(A,2+)$ is required, but this value is not given for silica in the extensive tabulations of Dorenbos [11]. Hu *et al.* [21] give an excitation spectrum for Eu^{2+} f-d luminescence near 450 nm from sol-gel silica. Since the lowest d-level overlaps with higher levels the excitation peak value should not be used, but rather the 15-20% onset value on the long wavelength side [11], which is about 400 nm (3.1 eV). Comparing this to the free ion value of 4.22 eV gives $D(A,2+)$ = 1.12 eV, from which all the f-d transition energies and hence the d-levels can be calculated, as shown in figure 2(a).

Although Ce^{4+} is non-luminescent, absorption experiments can be used to determine its CT band, and a value of about 260 nm (4.8 eV) has be reported [22-23]. If E^{CT} (Ce^{4+}) is interpreted as E_{Vf} (Ce^{3+}) then the absolute f-levels of the trivalent lanthanide ions can be determined. However, there are some indications that this is not the case. Firstly, let us consider the d-levels of the trivalent ions based on this value. To do so we require $D(A,3+)$, which can be obtained by comparing an f-d transition energy

in silica with that of the free ion listed in Table 1. Figure 3(a) shows the luminescence spectra of trivalent Ce^{3+} in silica.



(a) (b)

Figure 3. (a) Excitation and emission spectra of 4% Ce doped silica after annealing in reducing atmosphere at 1000°C for 2 h, measured with Cary Eclipse. (b) Excitation spectrum of 0.1% Tb doped silica (annealed at 1000°C in air) measured using synchrotron radiation at SUPERLUMI.

The excitation maximum occurs near 320 nm (3.88 eV), giving $D(A,3+) = 2.24$ eV, from which the d-levels can be derived as shown in figure 2(a). Here the d-levels of the trivalent ions lie about equal in energy to the divalent ones, whereas they should lie slightly lower. Dorenbos [9] does tabulate a value of $D(A,3+)$ for silica, but we believe our value is more accurate since in the reference used by him the sample had not been annealed at high temperature to incorporate the cerium in the silica. To further check the value, consider the excitation spectrum of a sample doped with Tb while monitoring the emission due to Tb^{3+} ions at 545 nm (figure 3(b)). The peak at 226 nm (5.49 eV) is the f-d absorption band, while the peak near 160 nm is due to defects as mentioned before. This f-d energy cannot, however, be compared directly to that of the free ion because the lowest f-d transition for Tb^{3+} is spin-forbidden. The experimental value corresponds to the first spin-allowed f-d transition which occurs 0.78 eV higher [8,24]; therefore the crystal field depression can be estimated as (6.97 + 0.78) – 5.49 = 2.26 eV, in excellent agreement with the value from cerium. Additionally, Dorenbos [25] has shown that the crystal field depressions for divalent and trivalent ions are related by $D(A,2+) = 0.64D(A,3+) - 0.233$, which holds reasonably well for the values obtained here. It therefore appears that the crystal field depression is accurate, so the higher than expected trivalent d-levels seem to be as a result of overestimation of $E_{Vf}(Ce^{3+})$. A second factor that seems to indicate this is the difference in the f-levels of trivalent and divalent Eu: generally it varies from above 7 eV in poorly polarisable compounds like the fluorides to below 6 eV in strongly polarisable compounds like bromides and sulphides, and for the metal it is 5.4 eV [15]. In the model of figure 2(a) it is only 5.4 eV, and a realistic value is at least 1 eV higher. Decreasing $E_{Vf}(Ce^{3+})$ to 3.8 eV achieves this and also brings the trivalent 5d levels 1 eV below the divalent ones, creating the more realistic energy structure in figure 2(b). Since we have no reason to doubt the accuracy of the Ce^{4+} CT absorption energy (or any of the other measurements) to the extent of the correction applied, it is concluded that the value of $E_{Vf}(Ce^{3+})$ cannot be associated with the Ce^{4+} CT absorption energy. Although using the CT absorption energies of tetravalent ions was suggested initially [12], to our knowledge it has not been applied before and generally a value of $E_{dC}(Ce^{3+})$ is estimated instead. It has been acknowledged that identifying the trivalent CT energies with the absolute energies of the divalent ions above the valence band is an

assumption that can be disputed [15], and although it appears to work well to find values of $E_{Vf}(Eu^{2+})$, in this work we have shown that a similar method is not as successful for determining $E_{Vf}(Ce^{3+})$.

To complete the energy level diagrams in figure 2 only the bandgap of amorphous silica is required. This has been determined as 9.0 eV by photoconductivity measurements [26], with the fundamental absorption edge at 8.1 eV [27] shown by a dotted line in figure 2.

References

[1] Tanabe S 2005 *IPAP Books* **2** 101
[2] Chiodini N *et al.* 2009 *Proc. SPIE* **7136** 731614
[3] Dorenbos P 2007 Lanthanide level locations and its impact on phosphor performance. *Phosphor Handbook* 2nd ed. ed W M Yen, S Shionoya and H Yamamoto (Boca Raton: CRC Press) p 139
[4] Dorenbos P and Bos AJJ 2008 *Rad. Meas.* **43** 139
[5] Dorenbos P 2004 *J. Lumin.* **108** 301
[6] Dorenbos P and van der Kolk E 2006 *Appl. Phys. Lett.* **89** 061122
[7] Dorenbos P and van der Kolk E 2008 *Opt. Mat.* **30** 1052
[8] Dorenbos P 2000 *J. Lumin.* **91** 91
[9] Dorenbos P 2000 *J. Lumin.* **91** 155
[10] Dorenbos P 2003 *J. Phys: Condens. Matter* **15** 575
[11] Dorenbos P 2003 *J. Lumin.* **104** 239
[12] Dorenbos P 2003 *J. Phys: Condens. Matter* **15** 8417
[13] Dorenbos P 2003 *J. Phys: Condens. Matter* **15** 2645
[14] Dorenbos P 2005 *J. Lumin.* **111** 89
[15] Dorenbos P 2009 *J. alloys and compounds* **488** 568
[16] Bressiere A, Dorenbos P, van Eijk CWE, Yamagishi E, Hidaka C and Takizawa T 2004 *J. Electrochem. Soc.* **151(12)** H254
[17] Carnall WT 1979 The absorption and fluoresence spectra of rare earth ions in solution. *Handbook on the Physics and Chemistry of Rare Earths* vol 3 ed K A Gschneider Jr and L Eyring (Amsterdam: North Holland) p 171
[18] Yi W, Lansheng L, Huiqun Z and Ruiqin D 2006 *J. Rare Earths* **24** 199
[19] Sakurai Y and Nagasawa K 2001 *J. Non-Cryst. Solids* **290** 189
[20] Gutzov S, Ahmed G, Petkova N, Fügleinand E and Petkov I 2008 *J. Non-Cryst. Solids* **354** 3438
[21] Hu X, Fan J, Li T, Zhang D, Chen W, Bai J and Hou X 2007 *Opt. Mat.* **29** 1327
[22] Patra A, Kundu D and Ganguli D 1997 *J. Sol-Gel Sci. Techn.* **9** 65
[23] Fasoli M, Vedda A, Lauria A, Moretti A, Rizzelli E, Chiodini N, Meinardi F and Nikl M 2009 *J. Non-Cryst. Solids* **355** 1140
[24] Dorenbos P 2003 *J. Phys: Condens. Matter* **15** 6249
[25] Dorenbos P 2003 *J. Phys: Condens. Matter* **15** 4797
[26] DiStefano TD and Eastman DE 1971 *Solid State Comm.* **9** 2259
[27] Guzzi M, Martini M, Paleari A, Pio F, Vedda A and Azzoni CB 1993 *J. Phys: Condens. Matter* **5** 8105

Magnetic Properties of Carbon Nanospheres at Low Temperatures

Makhosonke B. Dubazane[1,2], Jonathan M. Keartland[1,2], Yosief Mishgina[2]

[1]DST/NRF Centre of Excellence in Strong Materials, WITS 2050, South Africa
[2]School of Physics and Materials Physics Research Institute, WITS 2050, South Africa

E-mail: dubazane@psi.phys.wits.ac.za

Abstract. The magnetic properties of two nitrogen-doped carbon nanosphere samples are probed using Electron Magnetic Resonance (EMR) techniques as a function of temperature in the range 4 K - 320 K. EMR measurements on these samples showed a large paramagnetic peak at room temperature, implying that nitrogen forms paramagnetic impurities within the carbon matrix. Continuous wave EMR measurements of the peak amplitude and the linewidth have allowed us to to study the resonance line and the spin dynamics of the nitrogen impurities. The line width increases steadily in both samples as the temperature increases, and reaches a steady plateau at approximately 150 K. The variation of the amplitude of the EMR signal with temperature may be understood in terms of interactions of the paramagnetic ions with lattice vibrations. At low temperatures the direct process dominates, and the signal amplitude decreases as the temperature increases as expected from the Boltzmann factor. The signal amplitude reaches a broad minimum in the range 150 K - 200 K, and then increases markedly to the highest temperatures reached. At high temperatures our model suggests that an Orbach process with $\Delta \simeq 3500$ K accounts for the increase in the signal amplitude.

1. Introduction

Carbon nanotubes were discovered by Iijima [1] in the 1990s, and since then a considerable amount of attention has been devoted to the exposure of their underlying physical properties [2]. Carbon nanomaterials are produced in a large range of morphologies [3]. In this paper we focus on the fundamental physical properties of nitrogen-doped carbon nanospheres (CNS) that consist of graphene sheets. Synthesis of CNS may be accomplished using a number of techniques, including chemical vapour deposition (CVD). A recent review [2] provides a current summary of these techniques and a comprehensive summary of the broad current and potential applications of these materials. The study of undoped and doped CNS is consequently of considerable scientific and technological significance. A recent paper has reported on the synthesis and characterization of CNS containing boron [4], but no reports of their fundamental physical properties at low temperatures are in the literature as yet.

Electron magnetic resonance (EMR) - also known as electron paramagnetic resonance (EPR) or electron spin resonance (ESR) - is likely to provide information on the spin dynamics of the nitrogen centres in CNS. In the past similar investigations have probed the properties of shallow donor centres in silicon [5]. EMR has recently been used to successfully characterize boron-doped multi-walled carbon nanotubes [6], and has also been used in the investigation of

the fundamental physical properties of other nanomaterials [7]. In this paper we report on EMR investigations of two nitrogen-doped CNS samples in the temperature range 4 K - 320 K. Our results suggest that the spin-lattice relaxation rate of the paramagnetic impurities is dominated by relaxation of the paramagnetic spins via interactions with the lattice vibrations.

2. Sample preparation

The two CNS samples (labelled NK7 and NK9) were prepared in a chemical vapour deposition (CVD) reactor in the School of Chemistry at Wits, from mixtures of toluene and pyridine. Initial concentrations of 10 % pyridine and 90 % toluene for NK7, and 90 % pyridine and 10 % of toluene for NK9 were used. Full details of the the CNS synthesis are given elsewhere [8].

3. Experimental details

Continuous wave (cw) EMR measurements were performed using Bruker ESP300E X-band spectrometer operating in the range 9.53 - 9.57 GHz. This corresponds to $\Delta_Z \simeq 0.46$ K. The high temperature approximation to the Boltzmann factor in (4) applies to all temperatures in our experiment. Measurements were done using a high Q microwave cavity placed within an Oxford continuous flow cryostat, operating in conjunction with an Oxford Instruments ITC503 temperature controller. The CNS samples were placed in tubes made of borosilicate glass. Liquid nitrogen (80 K - 320 K) and liquid helium (4 K - 80 K) were used as cryogens. Temperatures were maintained to a precision of better than 0.3 K above 80 K, and to a precision of better than 0.1 K below 80 K. All measurements were taken with the microwave power set in the range 7.90 mW to 7.94 mW as this power range allowed for the most satisfactory tuning of the cavity for the entire temperature range. Phase sensitive detection results in the spectrometer displaying the first derivative of the EMR absorption peak as shown in Fig. 1. The spectra were used to obtain the peak-to-peak linewidth and the peak-to-peak amplitude as indicated in Fig. 1.

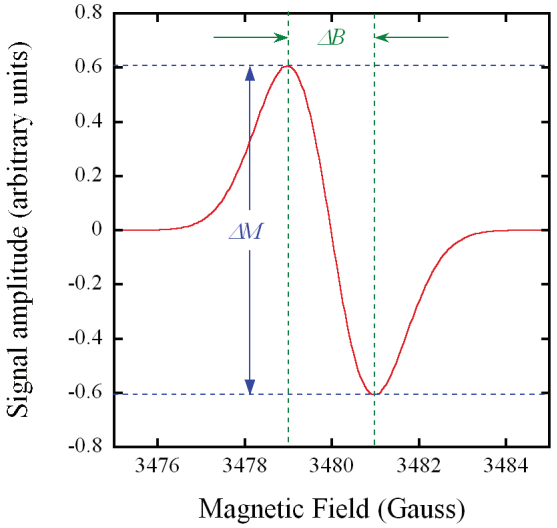


Figure 1. A simulated derivative spectrum for a single EMR line with $g \simeq 2$. The parameters ΔB (linewidth) and ΔM (signal amplitude) obtained from such a spectrum are defined.

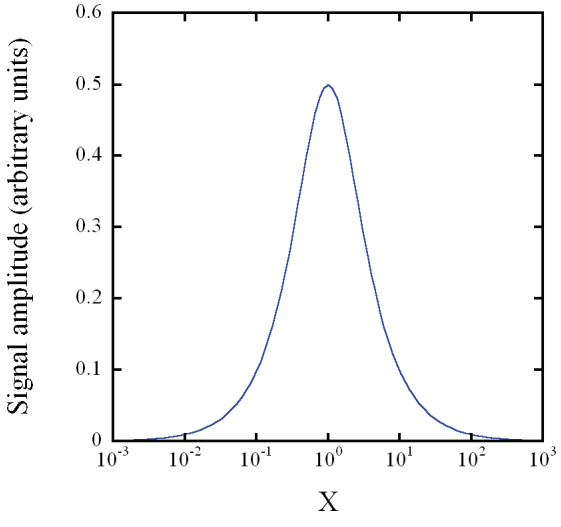


Figure 2. The power saturation curve for a homogeneously broadened EMR line [9]. The parameter X depends on the microwave power and the relaxation times of the spins as defined in (1).

The implications of performing the experiment at a constant power should be considered. The amplitude of the EMR spectrum depends on the gyromagnetic ratio γ, the amplitude of the

perturbing microwave magnetic field B_1, the spin-lattice relaxation time T_1 and the spin-spin relaxation time T_2 [9]. For homogeneously broadened lines this may be written in the form

$$\Delta M(T) \propto \frac{X}{1+X^2}, \tag{1}$$

where $X = \gamma B_1 \sqrt{T_1 T_2}$. A plot of the signal amplitude as a function of X is shown in Fig. 2. For constant microwave power B_1 is constant, and if we further assume that changes in T_2 are negligible then the signal amplitude, in the absence of other considerations, may be written in terms of T_1 as follows:

$$\Delta M(T) = M_0 \frac{A\sqrt{T_1}}{1+A^2 T_1}, \tag{2}$$

where M_0 and A are parameters that depend on the experimental arrangement. A discussion of the expected temperature dependence of T_1 for lattice vibrations will be provided in the next section.

4. Theoretical Considerations

4.1. The Boltzmann factor

The EMR signal for an ensemble of spin-$\frac{1}{2}$ particles (in the absence of any spin-lattice relaxation effects) in a sample at a temperature T placed in a uniform magnetic field will be determined by the Boltzmann factor:

$$\Delta M(T) \propto p_0 \left[1 - \exp\left(-\frac{\Delta_Z}{T}\right)\right], \tag{3}$$

where Δ_Z is the characteristic temperature of the Zeeman interaction of the spin with the applied magnetic field, and p_0 is the population of the lower spin energy level. In the high temperature limit ($\Delta_Z \ll T$) this may be re-written as:

$$\Delta M(T) \propto p_0 \frac{\Delta_Z}{T} \Rightarrow \Delta M(T) = \frac{B_Z}{T}, \tag{4}$$

where $B_Z \propto p_0$ is determined experimentally. Empirical determination of the Boltzmann parameter B_Z is used in magnetic thermometry [10], where care is taken to ensure that it is determined under conditions where the time between sampling $t \gg T_1$ where T_1 is the spin-lattice relaxation time. In the experiments described here the microwave power was kept constant, and so the signal will depend on the spin-lattice relaxation rate, as pointed out in the previous section.

4.2. Spin-lattice relaxation due to lattice vibrations

The relaxation of paramagnetic ions due to lattice vibarations has received considerable attention, with much of the seminal work described by Orbach [11]. For temperatures well below the Debye temperature there are generally three contributions - the direct process, the Raman process and the Orbach process respectively. The relaxation of a paramagnetic ion may be written in the form:

$$\frac{1}{T_1} = \frac{T}{A_D} + \frac{T^7}{A_R} + \frac{1}{A_O(e^{\frac{\Delta}{T}} - 1)}, \tag{5}$$

where A_D, A_R, A_O and Δ are determined from fits to the experimental data. The amplitude of the EMR signal cam then be determined by including the Boltzmann factor given in (4), and taking T_1 into account by inserting it into (2).

5. Results and Discussion

5.1. EMR Linewidth

The linewidth was obtained directly from the derivative of the EMR line obtained for each measurement. The results are summarized in Fig, 3 and Fig, 4. The linewidth for the two samples showed similar characteristics. Above approximately 150 K the linewidth is essentially constant. Below 150 K the lines for both samples narrow. For NK7 the linewidth changes from approximately

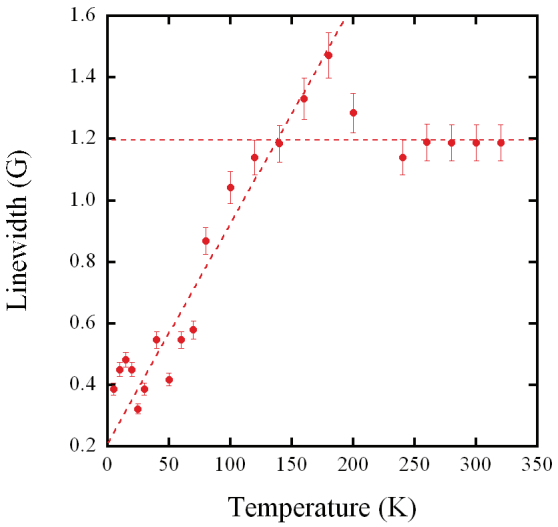


Figure 3. Dependence of the linewidth of the EMR spectrum for sample NK7. The dotted lines are to guide the eye of the reader. Two regions are identified - a plateau above approximately 150 K, followed by a steady decrease in the linewidth down to the lowest temperature (5 K).

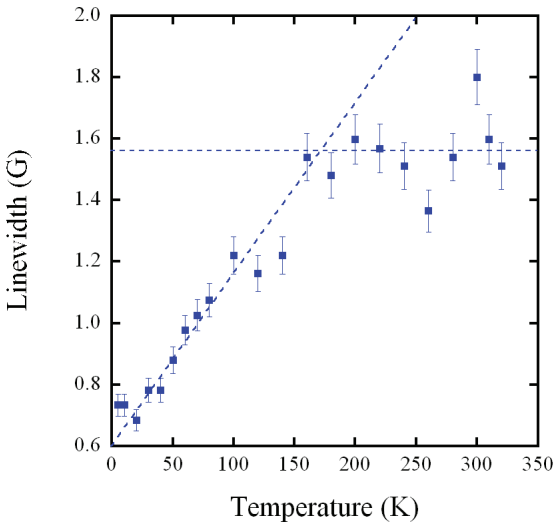


Figure 4. Dependence of the linewidth of the EMR spectrum for sample NK9. The dotted lines are to guide the eye of the reader. Two regions are identified - a plateau above approximately 150 K, followed by a steady decrease in the linewidth down to the lowest temperature (5 K).

5.2. EMR Signal Amplitude

The peak-to-peak amplitude for the derivative curves of the EMR lines was obtained for each spectrum. The results for the entire temperature range studied for NK9 and NK7 are given in Fig. 5 and Fig. 6 respectively. The samples show similar behaviour as the temperature decreases. Between 320 K and approximately 200 K the size of the signal decreases markedly (by a factor of approximately 4 in both samples). Between 200 K and a lower temperature (approximately 120 K for both samples) the signal amplitude is essentially constant, with the signal amplitude then increasing markedly as the lowest temperatures are reached (a factor of approximately 6 for NK7, and a factor of approximately 12 for NK9).

As we have pointed out, all measurements were taken at the same microwave power. As result the signal amplitude depends on both the Boltzmann factor and the dominant spin-lattice relaxation mechanisms. For a sample in which the relaxation rate is approximately constant we would expect the signal amplitude to follow the Boltzmann factor closely for measurements taken at the same applied microwave power, which is the case at low temperatures where we believe the direct process dominates. In Fig. 2 this corresponds to taking measurements in the tail of the saturation curve (large X). As T_1 decreases markedly we move along the saturation curve toward smaller values of X. This results in an increase in the observed EMR signal amplitude.

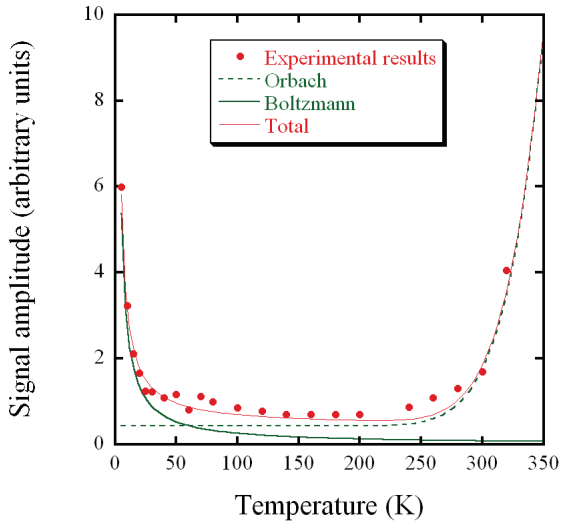


Figure 5. Temperature dependence of the signal amplitude for sample NK7. The observed variation of the signal amplitude with temperature is described by a model involving the Boltzmann factor (4) and the power saturation curve as determined by spin-lattice relaxation processes due to lattice vibrations (2).

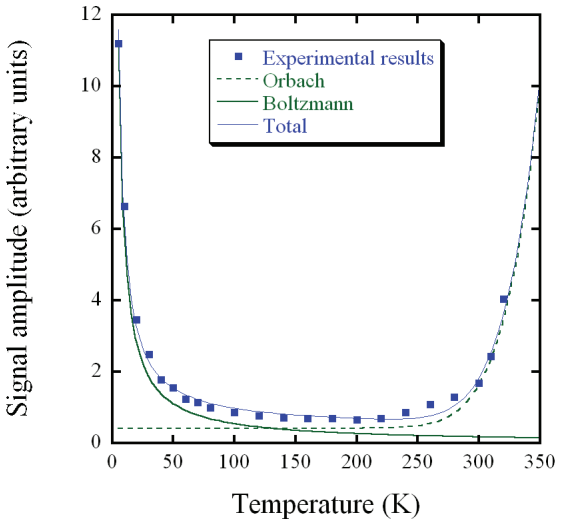


Figure 6. Temperature dependence of the signal amplitude for sample NK9. The observed variation of the signal amplitude with temperature is described by a model involving the Boltzmann factor (4) and the power saturation curve as determined by spin-lattice relaxation processes due to lattice vibrations (2).

Table 1. Fit parameters for the low temperature region for both NK7 and NK9 in which the Boltzmann factor dominates the signal amplitude. These correspond to large values of X in Fig. 2.

Sample	A	B_Z
NK7	0.34 ± 0.09	28 ± 1
NK9	0.67 ± 0.21	54 ± 1

At low temperatures (between 5 K and approximately 60 K) the signal amplitude follows a 1/T dependence, which is identical to that the predicted Boltzmann factor dependence. This suggests that X is largely temperature independent in this temperature region, and that decrease in the signal amplitude is entirely due to the increasing population difference between the spin energy levels as indicated in equation (4). We have fitted an equation of the form:

$$\Delta M = A + \frac{B_Z}{T}, \tag{6}$$

the data in this temperature regime, and the results of the fit are given in Table 1. The ratio of B_Z for the two samples should provide reasonable estimates of the concentration of the paramagnetic ions in each of these samples. We find that the concentration of nitrogen in NK7 is approximately half that of the concentration in NK9. This is in excellent agreement with careful EMR measurements made at room temperature for both of these samples. The results are provided elsewhere in these proceedings [12].

As the temperature increases above 200 K the signal amplitude increases markedly. We have simulated the effect of spin-lattice relaxation processes on the signal amplitude, and the

Table 2. Parameters for the lattice vibration contributions to the spin-lattice relaxation time of contributions obtained from simulations of the experiment. The model is highly sensitive to changes in Δ, but less sensitive to changes in the other parameters.

A_D (s.K^{-1})	A_R (s.K^{-7})	A_O (s)	Δ (K)
3	10^{10}	10^{-7}	3500

parameters that best describe the observed changes in the EMR signal amplitude for both samples are given in Table 2. The simulations are particularly sensitive to Δ, and the inclusion of the Orbach process is necessary to provide a satisfactory description of the data in the high temperature regime. It would therefore be of considerable interest to extend the EMR measurements to higher temperatures, and to measure the relaxation times in these materials directly using pulsed EMR methods.

6. Conclusions

Continuous wave EMR measurements at a constant power over the temperature range 5 K - 320 K have been made for two nitrogen-doped CNS samples. Changes in the linewidth and the amplitude of the signal have been observed. The changes in the amplitude may be ascribed to spin-lattice relaxation effects, whereby the paramagnetic nitrogen ions interact with the lattice vibrations. It would be of interest to perform pulsed EMR measurements in this temperature range, and thus measure the relaxation rates directly.

Acknowledgments

The authors gratefully acknowledge the financial support of the DST/NRF Centre of Excellence in Strong Materials, the National Research Foundation, and the School of Physics and the Materials Physics Research Institute of the University of the Witwatersrand.

References

[1] Iijima S 1991 *Nature* **354** 56
[2] Nieto-Márquez A, Romero R, Romero A and Valverde J L 2011 *J. Mater. Chem* **21** 1664
[3] Serp P, Feurer R, Kalck P, Kihn Y, Faria J L and Figueiredo J L 2001 *Carbon* **39** 621
[4] Mondal K C, Strydom A M, Tetana Z, Mhlanga S D, Witcomb M J, Havel J, Erasmus R M and Coville N J 2009 *J. Mat. Chem. Phys.* **114** 973
[5] Castner T G 1963 *Phys. Rev* **130** 58
[6] Mondal K C, Strydom A M, Erasmus R M, Keartland J M and Coville N J 2009 *J. Mat. Chem. Phys.* **111** 386
[7] Likodimos V, Glenis S and Lin C L 2005 *Phys.Rev. B* **72** 045436
[8] Kunjuzwa N 2009 *M.Sc dissertation University of the Witwatersrand Johannesburg*
[9] Castner T G 1959 *Phys. Rev* **115** 1506
[10] Blum K I 1988 *Experimental Techniques in Condensed Matter Physics at Low Temperatures* (*Frontiers in Physics* vol 67) ed Richardson R C and Smith E N (Redwood City: Addison-Wesley)
[11] Orbach R 1961 *Proc. Roy. Soc. London Ser. A* **264** 485
[12] Keartland J M, Dubazane M B, Marsicano V, Kunjuzwa N and Coville N J 2011 *These proceedings*

Low Temperature Synthesis of ZnO nanoparticles and Nanorods via Wet Chemistry Route

P.S. Mbule[1], G.H. Mhlongo[1,2], H.C. Swart[1] and O.M. Ntwaeaborwa[1,3]

[1] Department of Physics, University of the Free State, Bloemfontein, ZA9300, South Africa

[2]National Centre for Nano-structured Materials, CSIR, PO Box 395, Pretoria, ZA0001

[3]Corresponding author e-mail: ntwaeab@ufs.ac.za

Abstract: In this study, we report the synthesis and characterization of colloidal ZnO nanoparticles (NPs) and nanorods. The colloids were prepared by precipitation method at room temperature. The X-ray diffraction (XRD) patterns of nanoparticles and nanorods, revealed hexagonal wurtzite structure and the diffraction peaks matched well with the JCPDS (card no 80-0075) standard data. The Scanning electron microscopy (SEM) revealed the spherical nanoparticles and clustered nanorods. The optical absorption of the nanoparticles at 340 nm was more intense than that of the nanorods. The green emission from defects was more intense than the excitonic UV emission in the case of the nanoparticles, while the excitonic UV emission was more intense in the case of the nanorods. The ZnO nanoparticles and nanorods were investigated for possible applications in transparent electrodes in solar cells and gas sensors.

1. **Introduction**

ZnO optical and structural properties have drawn so much attention in the past decades. ZnO nanostructures such as quantum dots (QDs), nanoparticles (NPs), nanobelts, nanowires and nanorods have been widely investigated [1]. Moreover, it has been reported that nano-sized ZnO exhibit a unique luminescence properties different from that of bulk ZnO because nanoparticles posses an enormous surface-to-volume ratio [2]. At room temperature ZnO typically exhibit a UV emission due to recombination of free excitons and emission in the visible spectral range attributed to defects [3]. Different physical and chemical processes such as chemical bath deposition, solid state and hydrothermal technique have been used to prepare ZnO particles with different morphologies, However wet chemical synthesis technique offers many advantages in comparison with other techniques because of their realiability, reproducibility at lower temperatures, higher yields and well defined nanostructures [1,4]. With its well-known wide bandgap (~3.37 eV at room temperature), large exciton binding energy (60 meV) and excellent chemical stability, ZnO has attracted extensive attention due its potential applications in laser diodes, solar cells, field emission display and gas sensors [5]. In this study a simple precipitation method was used to synthesize ZnO NPs and nanorods and we report the structure, morphology and photoluminescence of the ZnO nanostructures.

[3] Corresponding author: email address: ntwaeab@ufs.ac.za

2. Experimental

2.1. ZnO nanoparticles synthesis

Zinc acetate and tetramethylammonium hydroxide (TMAH) were used as precursors. TMAH dissolved in 30ml ethanol was added dropwise to 2g of zinc acetate dissolved in dimethylsulfoxide (DMSO) solution followed by an hour stirring at room temperature.The resulting precipitation was washed at least three times in a mixture of heptane:ethanol (1:3) and ZnO NPs were dispersed in ethanol. Colloids of ZnO nanoparticles with different concentrations of zinc acetate ranging from 0.1 – 0.8 M were prepared and the UV-vis absorption and fluorescence spectra of the colloids were recorded. The precipitate was later separated by centrifugation and was dried in oven at 100^{o}C for 1hour for XRD characterization.

2.2. ZnO nanorods synthesis

3g of zinc nitrate tetrahydrate was dissolved in 100 ml of distilled water and stirred for 30 minutes. Simultaneously, a 10 ml of sodium hydroxide (NaOH) was added dropwise into aqueous zinc nitrate solution and stirred for an hour. The resulting precipitate was washed at least three times in a mixture of heptane:ethanol same ratio as in 2.1. The colloids of ZnO nanorods with different concentrations of NaOH ranging from 0.1 – 1.0 M were prepared and the UV-vis absorption and fluorescence spectra of the colloids were recorded. The precipitate was later separated by centrifugation and was dried in oven at 100^{o}C for 1hour for XRD characterization.

The crystalline structure was analyzed by XRD using Bruker AXS D8 Advance diffractometer. The optical absorption was recorded by Perkin Elmer Lambda 950 Spectrophotometer. SEM images were obtained by JEOL-JSM 7500F, Field Emission SEM and PL measurements were carried out at room temperature using Varian Cary Eclipse fluorescence spectrophotometer.

3. Results and discussion

Diffraction peaks corresponding to the hexagonal wurtzite structure of ZnO were observed from both dried ZnO nanoparticles and nanorods. The spectra in figure 1, show the XRD pattern with the nanorods having a narrow peaks indicating better crystallinity as compared to that of nanoparticles. Average particle size estimated by Scherer equation was ~10-20±1 nm for nanoparticles and 30-50±2 nm for nanorods in diameter. The peaks matched well with the JPCDS standard data (card no 80-0075).

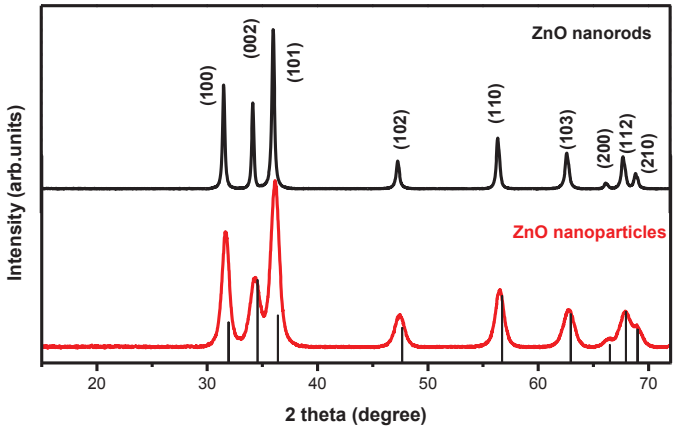


Figure 1: XRD patterns of ZnO nanorods, nanoparticles and standard JPCDS match (card no 80-007)

SEM images in figure 2(a) and (b) showed spherical nanoparticles and clustered nanorods. The clustering may be due to the sample preparation or the process of rods formation might be incomplete resulting to particles and chunks of clustered nanorods. The absorption spectra in figure 3, shows a well defined excitonic absorption for nanoparticles and the weak absorption for nanorods colloids both at ~ 340 nm (3.65 eV). This absorption peak is blue-shifted from the bulk absorption at 3.37 eV. This blue-shift can be ascribed to the confinement of charge carriers in restricted volume of smaller particles [6,7]. The PL spectra measured at room temperature using Xe lamp as an excitation source with an excitation wavelengths of 260 nm and 200 nm is shown in figures 4 (a) and (b) respectively. The colloidal nanoparticles exhibited dual emission in the UV and visible regions. The UV emission consists of two bands at ~364 nm (3.41 eV) and ~380 nm (3.26 eV) which can be ascribed to donor-acceptor transition and recombination of free excitons respectively [8,9,10]. The intense green emission ~527 nm (2.35 eV) is ascribed to oxygen or zinc vacancies. The inset of figure 4 (a) shows how the maximum PL intensity varies with zinc acetate concentration. In Figure 4 (b), the colloidal nanorods on the other hand, showed an intense near band-edge emission and a weak red-shifted defects emission at ~723 nm. Similarly, the near band-edge emission consists of two bands at at ~364 nm (3.41 eV) and ~380 nm (3.26 eV). The red shifted emission at ~723 nm is attributed to excess oxygen [3] and OH groups attached the nanostructures [9,11]. Note that the origin of different defect emissions in ZnO is not fully established yet and it is still a subject of intensive research.

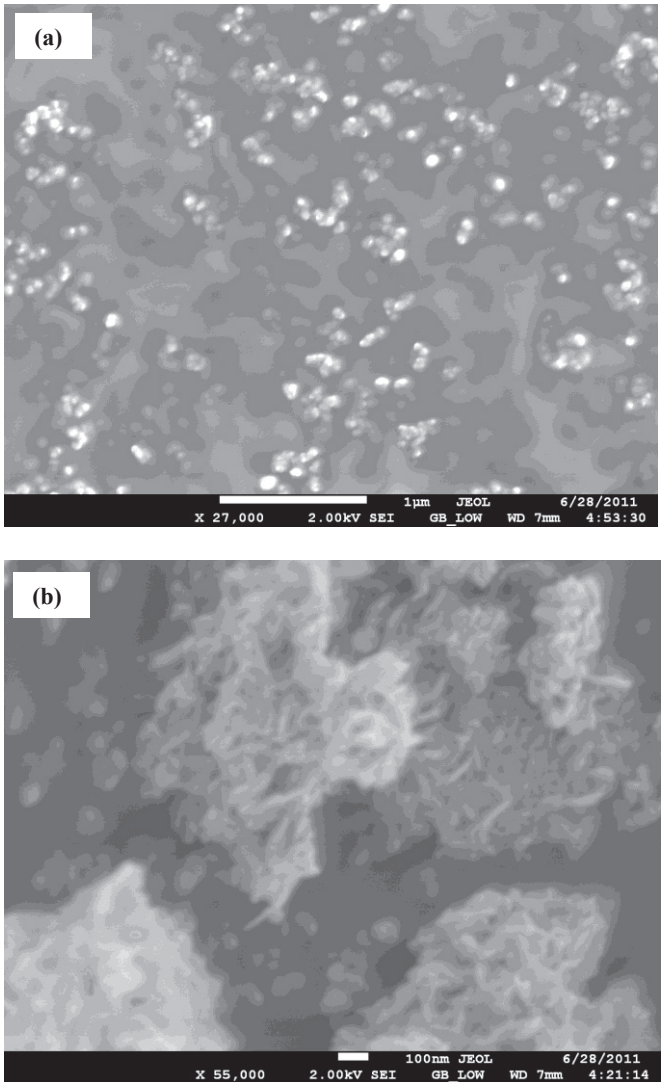


Figure 2: SEM images of ZnO (a) Nanoparticles and (b) clustered Nanorods

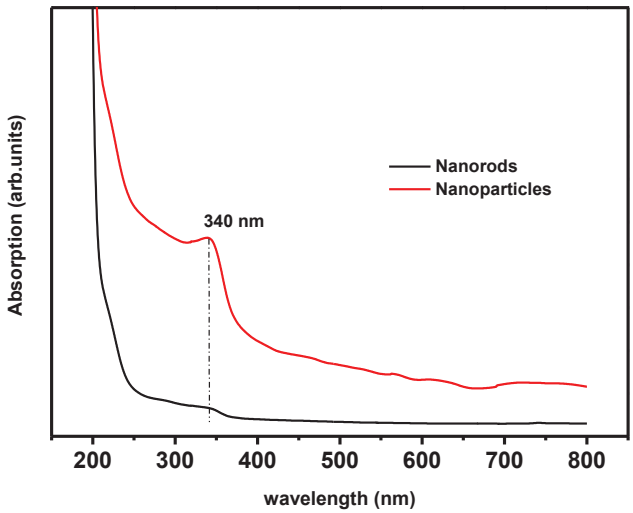


Figure 2: Absorption spectra of both colloidal ZnO nanorods and nanoparticles.

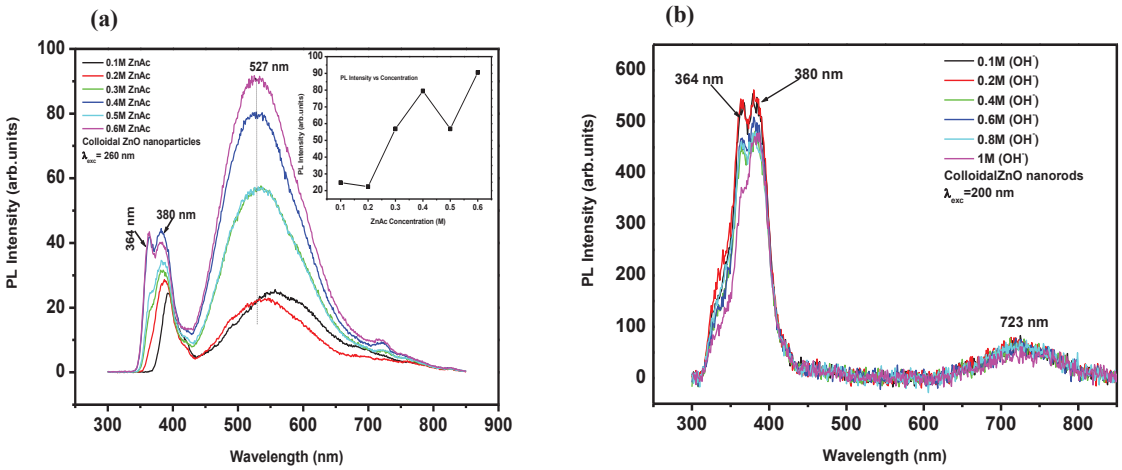


Figure 4: The PL spectra of colloidal (a) ZnO nanoparticles, the insert show the PL intensity vs concentration of zinc acetate, (b) ZnO nanorods.

4. Conclusion

Colloidal ZnO nanoparticles and nanorods were synthesized by a simple wet chemistry route at low temperatures. Both nanoparticles and nanrods showed dual emission in the UV and visible region of the electromagnetic spectrum associated with excitonic recombinations and defects. In the case of ZnO nanoparticles, the green defects emission was more intense than the UV emission while the UV emission from the nanorods was more intense than the red-shifted defects emission. The red-shifting of the defects emission was attributed to excess oxygen and surface OH group.

5. Acknowledgement

The authors would like to thank the University of The Free Sate and South African National Research Foundation (NRF) for financial support.

References

[1] Kumar D, Kumar S, Bhatti H.S, Gupta A and Sharma J.K, 2008, *Journal of Ovonic Research*, **4** 101-105
[2] Pen Y-Y, Hsieh T-E and Hsu C-H, 2006, *Nanotechnology*, **17** 174-180
[3] Djurisic A.B, Leung Y.H, Tam K.H, Hsu Y.F, Ding L, Ge W.K, Zhang Y.C, Wong K.S, Chan W.K, Tam H.L, Cheah K.W, Kwok W.M and Phillips D.L,2007, *Nanotechnology*, **18** 095702(8pp)
[4] Kajbafvala A, Shayegh M.R, Mazloumi M, Zanganeh S, Lak A, Mohajerani M.S and Sadrnezhaad S.K, 2009, *Journal of Alloys and Compounds* **469** 293-297
[5] Ntwaeaborwa O.M, Kroon R.E, Kumar V, Dubroca T, Ahn J-P, Park J.K and Swart H.C,2009, *Journal of Physics and Chemistry of Solids* **70** 1438-1442
[6] Hung C-H and Whang W-T, 2003, *Materials Chemistry and Physics*, **82** 705-710
[7] O.M Ntwaeaborwa and P.H Holloway, 2004, *Nanotechnology* **16** 865
[8] Kumari L and Li W.Z, 2010, *Crystal Research and Technology*, **45** 311-315
[9] Hu Q.R, Wang S.L, Jiang P, Xu H, Zhang Y and Tang W.H, 2010, *Journal of alloys and Compounds*, **496** 494-499
[10] Zhang Y, Du G, Liu D, Wang X, Ma Y, Wang J, Yin J, Yang X, Hou X and Yang S 2002, *Journal of Crystal Growth* **243** 439
[11] Tam K.H, Cheung C.K, Leung Y.H, Djurisic A.B, Ling C.C, Beling C.D, Fung S, Kwok W.M, Chan W.K, Phillips D.L, Ding L, Ge W.K, 2006, *Journal of Physics and Chemistry* **110** 20865–20871

Effects of ZnO and Ce^{3+} incorporation on the photoluminescence and cathodoluminescence intensity of Pr^{3+} doped SiO_2

G H Mhlongo[1,2], M S Dhlamini[3], O M Ntwaeaborwa[2*], H C Swart[2], R E Kroon[2], K T Hillie[1, 2]**

[1]National Centre for Nano-structured Materials, CSIR, PO Box 395, Pretoria, ZA0001

[2]Department of Physics, University of Free State, Bloemfontein, ZA9300.

[3]Department of Physics, University of South Africa, P.O. Box 392, Pretoria, ZA0003

E-mail: ntwaeab@ufs.ac.za, THillie@csir.co.za

Abstract. The successful incorporation of ZnO nanoparticles and Ce^{3+} ions in Pr^{3+} doped SiO_2 using a sol-gel process is reported. SiO_2:Pr^{3+} gels, with or without ZnO and Ce^{3+}, were dried at room temperature and annealed at 600 ^oC. The Field Emission Scanning Electron Microscopy indicated clustered SiO_2 nanoparticles with particle size diameter ranging from 20 to 120 nm while the energy dispersive x-ray spectroscopy (EDS) revealed the presence of Zn, Ce, and Pr clustures enveloped in the SiO_2 matrix. The emission intensity from SiO_2:Pr^{3+}-Ce^{3+} at 494 nm was slightly enhanced compared to Ce^{3+} singly doped SiO_2. The red emission of Pr^{3+} was enhanced when ZnO nanoparticles were incorporated in SiO_2 suggesting that energy was transferred from ZnO to Pr^{3+}. Possible energy transfer mechanisms between $Pr^{3+}$and Ce^{3+} and between ZnO and Pr^{3+} are discussed in detail.

1. Introduction

Among rare earths, trivalent praseodymium (Pr^{3+}) has been used as an activator in various host matrices to prepare phosphors [1-4] that can be used in different types of light emitting devices. The possibility of enhancing luminescence efficiency via energy transfer between activator ions (i.e. the sensitizer and the acceptor) was predicted by Dexter [5] in the 1960s. Studies of energy transfer process between Pr^{3+} and other ions co-doped in the same host matrix have been reported [6,7]. Interest has also been focused on nanocrystalline semiconducting quantum dots as they are considered as another good choice of sensitizing centres for radiative relaxation of activator ions since their excitation cross sections are very high due to the efficient band-to-band absorptions. Rare earth ion doped glass matrices containing nanocrystal semiconducting quantum dots prepared by the sol-gel method have been reported [8,9]. In such cases, enhanced luminescence from rare earth ions can be achieved by efficient energy transfer from the quantum dots to rare earth ions. In the present work, we report the enhanced emission of Pr^{3+} induced by energy transfer from ZnO nanoparticles incorporated

1 O.M Ntwaeaborwa, ntwaeab@ufs.ac.za
2 T. K. Hillie, thillie@csir.co.za

in situ in sol-gel silica under vacuum ultra violet (VUV) and electron beam excitation. Energy transfer between Pr^{3+}-Ce^{3+} co-doped in SiO_2 was also studied under electron beam excitation.

2. Experimental

SiO_2:Pr^{3+} co-doped with different concentrations of Ce^{3+} phosphor samples were prepared by mixing 0.05 mol of tetraethylorthosilicate (TEOS), 0.1 mol of H_2O, 0.1 mol of ethanol, and 0.145 mol of dilute nitric acid. The mixture was stirred at room temperature for 1 hour. The resulting transparent sol was mixed with a desired amount of $Pr(NO)_3.6H_2O$ dissolved in 5 ml of ethanol and was stirred for 30 minutes. A desired amount of $Ce(NO)_3.6H_2O$ dissolved in 5 ml of ethanol was also added to the solution and stirred further for 30 minutes until a gel formed. The gel was then dried at room temperature for 3-8 days. The dried gel was ground and annealed in air at 600 ^oC for 2 hrs. The $ZnO.SiO_2$:Pr^{3+} nanocomposite phosphor was prepared by mixing the SiO_2:Pr^{3+} sol with ZnO nanoparticles suspended in ethanol. The detailed preparation of ZnO nanoparticles has been reported [8].

The particle morphology was analyzed by JEOL 2100 High Resolution Transmission Electron Microscopy (HRTEM) and JEOL JSM-7500F, Field Emission Scanning Electron Microscope (FESEM). The cathodoluminescence (CL) data of the phosphor samples were recorded using an Ocean Optics S2000 spectrometer coupled with an ultra high vacuum chamber of the Physical Electronics PHI 549 Auger electron spectrometer. The photoluminescence (PL) data was recorded at the Deutsche Elektronen Synchrotron (DESY) using the setup at SUPERLUMI experimental station of HASYLAB in Hamburg. The emission spectrum of ZnO was obtained using a LS 55, Fluorescence spectrometer.

3. Results and discussion

3.1. Particle morphology, size, and phase structure

The FE-SEM image in fig. 1 illustrates the morphologies of the SiO_2 nanoparticles calcined at 600 ^oC for 2 hrs. The image exhibits the agglomeration of mostly spherical SiO_2 particles with an average particle size in the range of ~20 to 120 nm in diameter. Similar morphology was observed for $ZnO.SiO_2$:Pr^{3+} and SiO_2:Pr^{3+}-Ce^{3+} samples. Fig. 2 (a) and (b) show the XRD patterns of SiO_2 nanoparticles and ZnO nanoparticles. The XRD patterns of the ZnO nanoparticles presented in fig. 2 (b) are consistent with a hexagonal phase of ZnO referenced in JCPDS file No. 80-0075. The broadening of the ZnO diffraction peaks is attributed to the smaller particle sizes. The average crystallite size of the ZnO nanoparticles estimated using Scherrer's equation was ~ 4 nm. Fig. 2 (a) shows only one broad diffraction peak of amorphous SiO_2 at 22^o. The diffraction peaks of SiO_2:Ce^{3+}-Pr^{3+} and $ZnOSiO_2$:Pr^{3+} samples resembled that of pure SiO_2 probably due to the relatively low concentration of $Pr^{3+}$and the ZnO nanoparticles and/or high amorphous scattering background from the SiO_2 matrix [8]. The presence of the Pr^{3+},Ce^{3+} ions and ZnO nanoparticles in the phosphor powders (not shown) was confirmed by Electron Dispersive Spectroscopy (EDS).

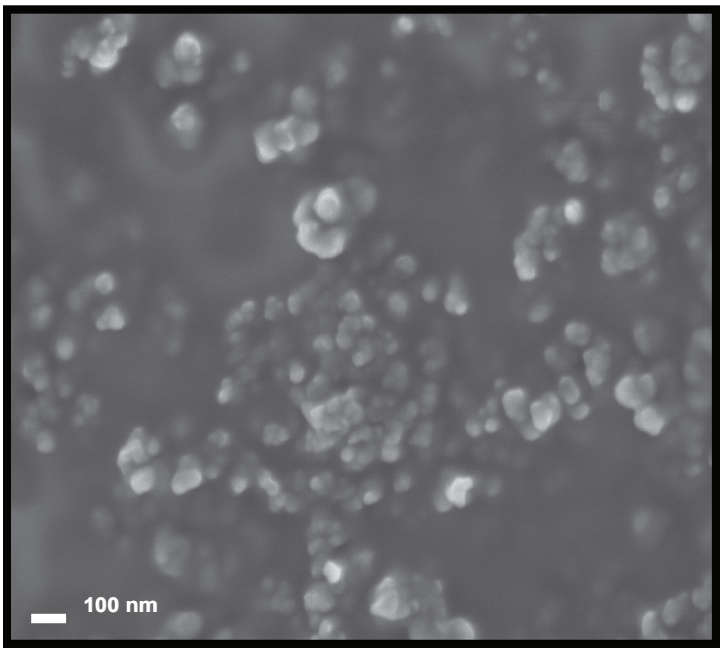


Fig. 1. FE-SEM image of SiO_2.

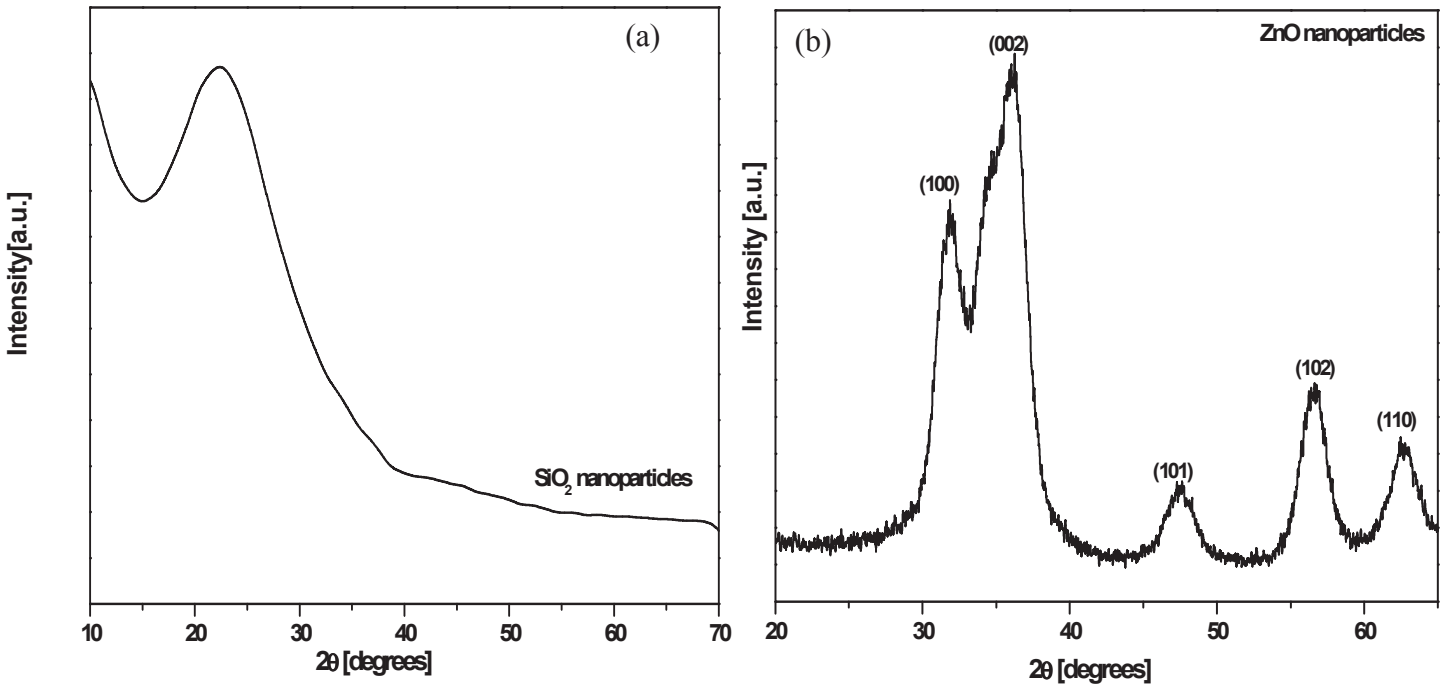


Fig. 2. XRD patterns of the (a) SiO_2 nanoparticles after calcination at 600 ^oC and (b) ZnO nanoparticles dried at 90 ^oC.

3.2. PL and CL properties of ZnO.SiO$_2$:Pr^{3+}.

CL spectra of the SiO_2, SiO_2:0.2 mol% Pr^{3+} and ZnO.SiO_2:0.2 mol% Pr^{3+} nanophosphor powders excited by 2 keV electrons, at 20 µA beam current in a high vacuum chamber at a base pressure of ~ 1.6 x 10^{-8} Torr are shown in fig. 3. The CL spectrum of SiO_2 showed the emission peak at 445 nm which can be assigned to structural defects in the SiO_2 network or charge transfer between O and Si atoms [8]. The characteristic emission peaks with maximum emission at 614 nm that can be associated with transitions in Pr^{3+} were observed from both SiO_2:Pr^{3+} and ZnO.SiO_2:Pr^{3+}samples. This emission can be assigned to the transitions originating from the $^3P_0$ and $^1D_2$ energy levels to the $^3H_{(J = 6, 5, 4)}$ and $^3F_{(J = 2, 3, 4)}$ energy levels all localized in the $4f^2$ intra-configuration of the Pr^{3+} ions with a dominant emission at 614 nm corresponding to the $^3P_0 \rightarrow {}^3H_6$ transition. It was also noticed that there was no emission from ZnO nanoparticles (usually at ~370 and 560 nm), instead the red emission of Pr^{3+} was enhanced when ZnO nanoparticles were incorporated. Lack of emission from ZnO nanoparticles and the subsequent enhancement of the red emission of Pr^{3+} demonstrate sensitization of Pr^{3+} emission centres by ZnO nanoparticles. That is, energy was transferred non-radiatively from ZnO to Pr^{3+}. The transfer was most probably by phonon mediated processes as previously reported [8-10].

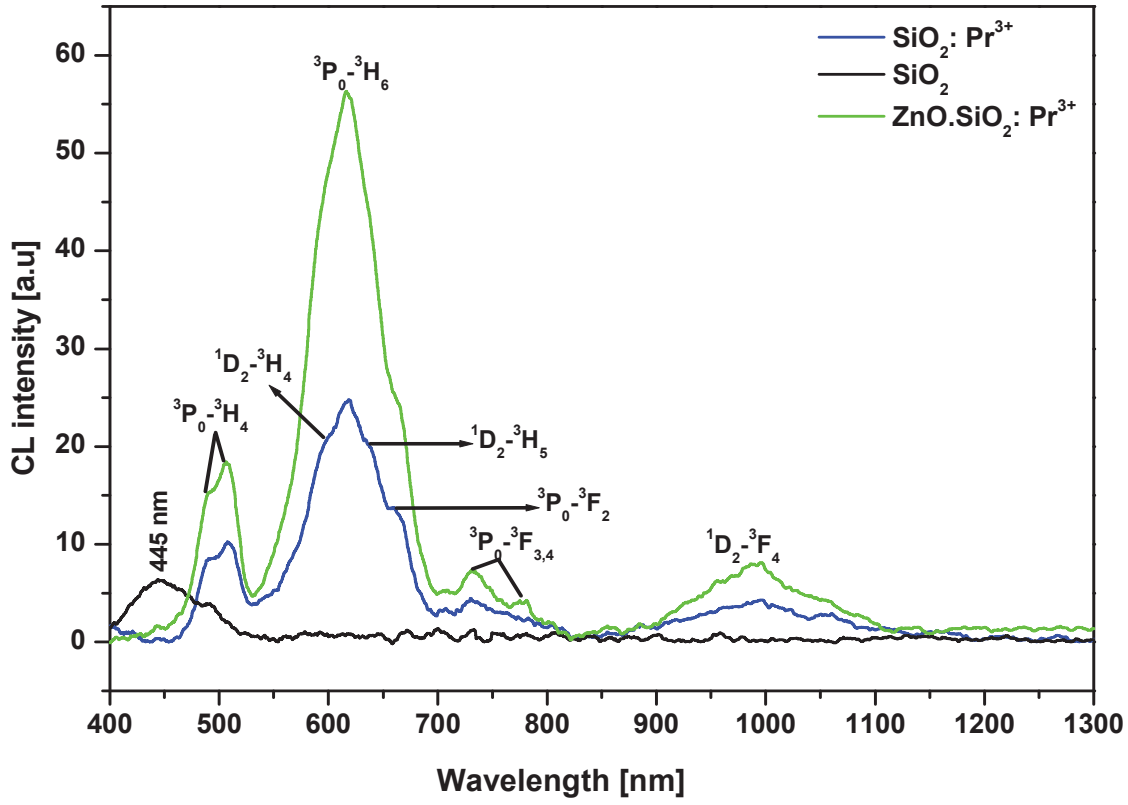


Fig. 3. CL spectra SiO_2, SiO_2:0.2 mol% Pr^{3+} and ZnO.SiO_2:0.2 mol% Pr^{3+} irradiated with 2 keV, 20 µA beam of electrons in a high vacuum chamber containing 1.6 x 10^{-8} Torr.

Fig. 4 presents the PL emission spectra of SiO_2:0.2 mol% Pr^{3+} and ZnO.SiO_2:0.2 mol% Pr^{3+} phosphors under excitation at 90 nm using synchrotron radiation. Also, only characteristic emissions from Pr^{3+} ions with the main red emission centred at 614 nm were detected from both SiO_2:Pr^{3+} and ZnO.SiO_2:Pr^{3+} phosphors. The PL intensity of ZnO.SiO_2:Pr^{3+} was enhanced with ZnO incorporation indicating energy transfer from ZnO to Pr^{3+} ions.

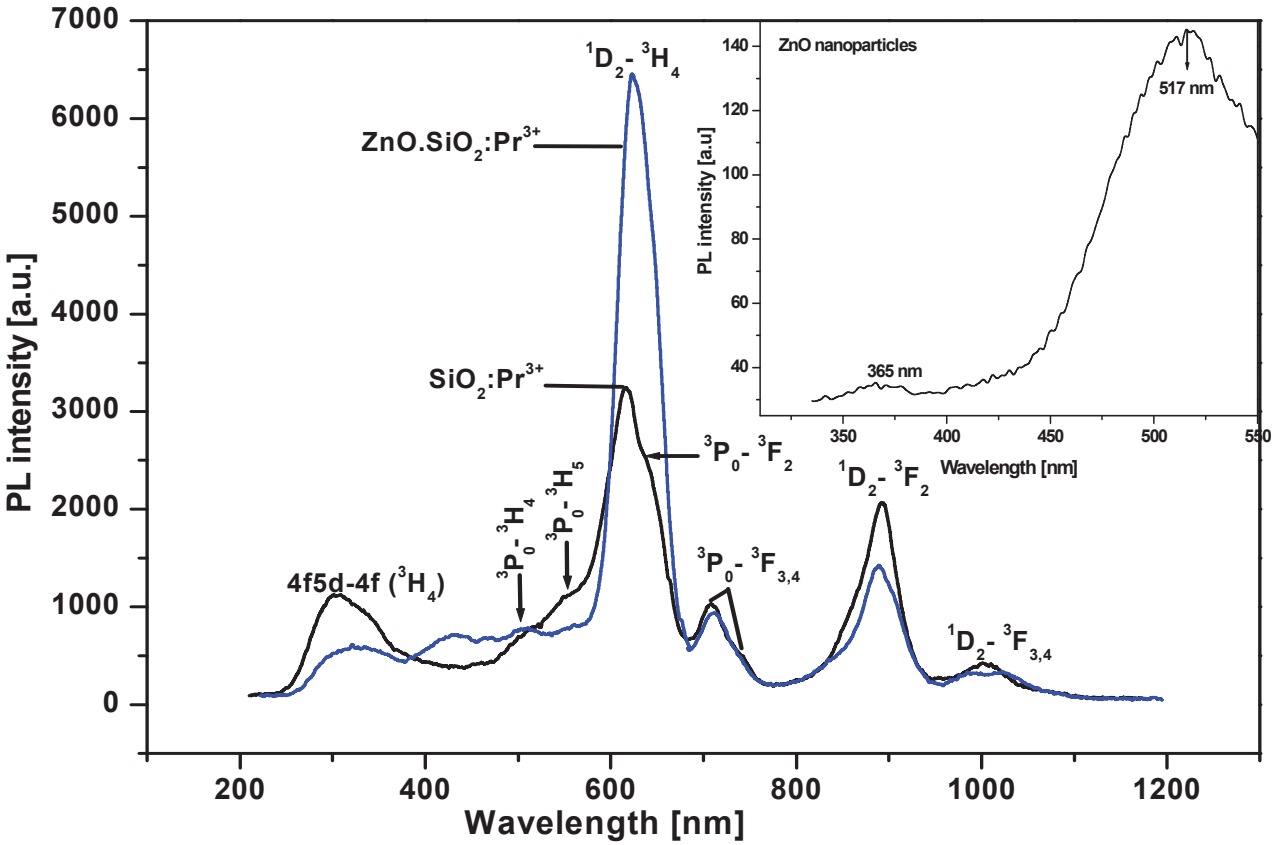


Fig. 4. PL emission spectra of SiO_2:0.2 mol% Pr^{3+} and $ZnO.SiO_2$:0.2 mol% Pr^{3+} after VUV excitation at 90 nm using synchrotron radiation. The inset shows the PL emission spectrum of ZnO nanoparticles after excitation at 325 nm.

Note that the emission spectrum from ZnO could not be measured due to charging when exciting using synchrotron radiation. However the spectrum was recorded when a monochromatized xenon lamp (λ_{exc} = 325 nm) was used as shown in the inset in fig. 4. The direct band gap emission and green emission from the ZnO nanoparticles was measured at 365 nm and 517 nm, respectively. Based on these results it was then concluded that the energy transfer from ZnO could be due to bandgap absorption and relaxation to the defects states of ZnO [9].

3.3. CL properties of SiO_2: Ce^{3+}-Pr^{3+}.

Fig. 5 shows the CL emission spectrum of SiO_2:1 mol% Ce^{3+} phosphor powder under irradiation with 2 keV electrons, 8.5 μA beam current in a high vacuum chamber at a base pressure of 1.2 x 10^{-8} Torr. The CL emission spectrum from SiO_2:Ce^{3+} consists of two blue bands located at 452 nm (shoulder) and 494 nm (strong). The two bands at 452 nm and 494 nm can be associated with $^2D_{3/2}$ -$^2F_{7/2}$ and $^2D_{3/2}$ -$^2F_{5/2}$ of Ce^{3+}, respectively [7].

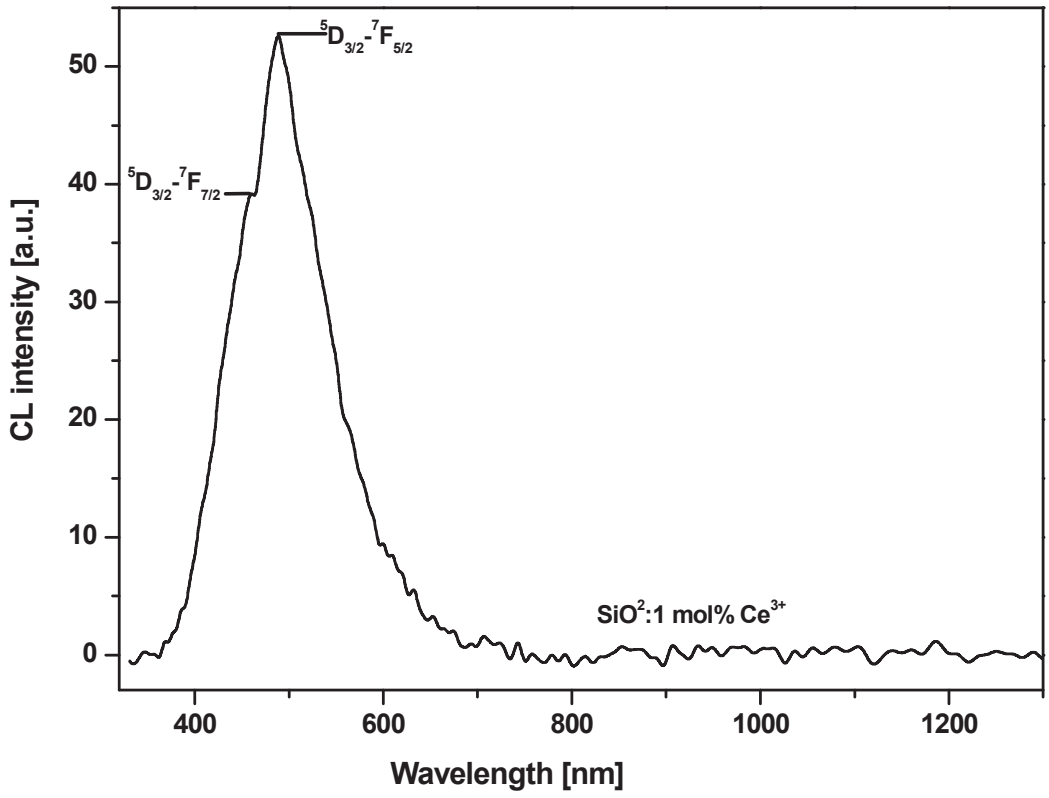


Fig. 5. CL emission spectrum of SiO_2:Ce^{3+} (1mol %) irradiated with 2 keV, 8.5 μA beam of electrons in a high vacuum chamber containing a base pressure of 1.2 x10^{-8} Torr.

The CL emission spectra of the SiO_2:Pr^{3+}, SiO_2:Ce^{3+}, and SiO_2:Pr^{3+}.Ce^{3+} (with different Ce^{3+} concentrations) phosphor powders under irradiation with 2 keV electrons, 8.5 μA beam current in a high vacuum chamber at a base pressure of 1.2 x 10^{-8} Torr are shown in fig. 6. The CL emission spectra of SiO_2:Ce^{3+}-Pr^{3+} showed both bands from Ce^{3+} in the blue spectral region and a small shoulder from Pr^{3+} in the red spectral region when the amount of Pr^{3+} and Ce^{3+} were both 0.2 mol%. The small shoulder from Pr^{3+} was slowly quenched with increasing Ce^{3+} concentration and only blue emission from Ce^{3+} could be observed. Blue emission from Ce^{3+} was slightly enhanced with addition of 1 mol% Ce^{3+} into 0.2 mol% Pr^{3+} compared to singly doped SiO_2:Ce^{3+}. These results suggest the energy transfer from Pr^{3+} ions to its nearest neighbouring Ce^{3+} ions. In addition, it appears that at 1 mol% of Ce^{3+}, the Ce^{3+} neighbouring ions are close enough to each other for energy transfer from Pr^{3+} to Ce^{3+} to take place. An increase in Ce^{3+} concentration led to the shortening of the distance between its neighbouring ions as a result of non-radiative loss of excitation energy between Ce^{3+} ions favoring quenching effect thus decreasing luminescence intensity. The quenching of the CL intensity of SiO_2:Pr^{3+}.Ce^{3+} with increasing Ce^{3+} concentration is demonstrated in the inset of fig. 6.

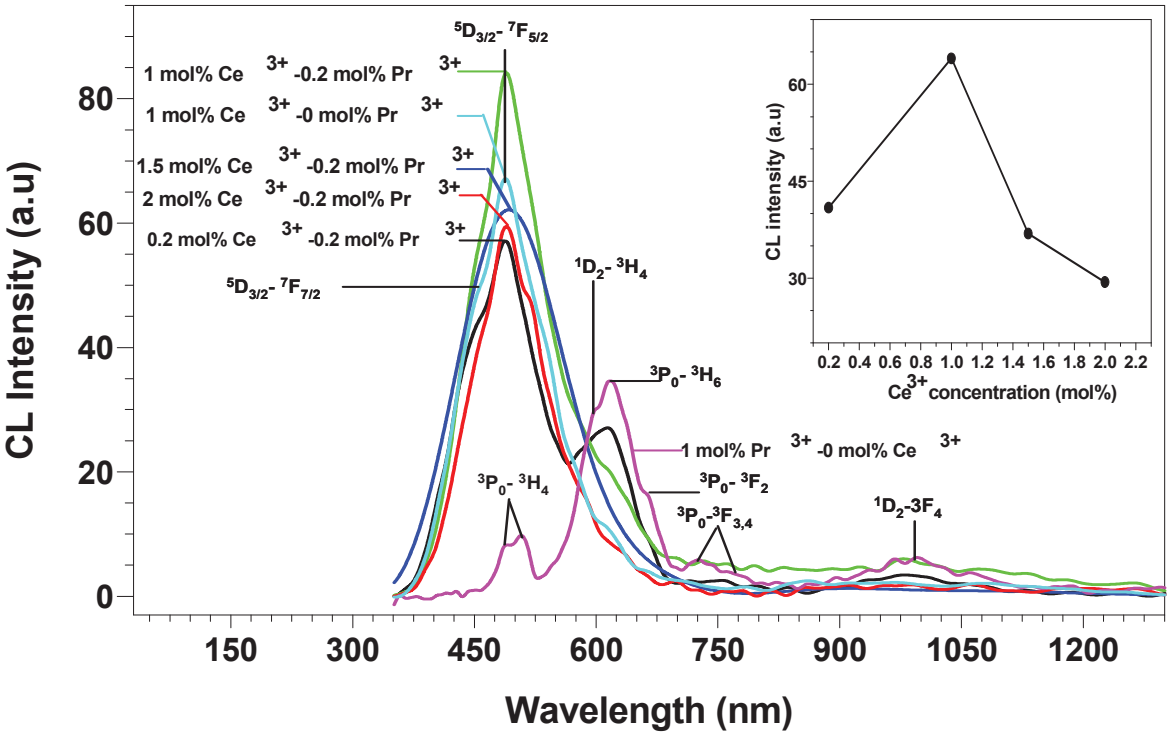


Fig. 6. CL emission spectra of the SiO_2:Pr^{3+} (1mol %), SiO_2:Ce^{3+} (1mol %), and SiO_2:Pr^{3+} (0.2mol %) - Ce^{3+} (0.2, 1, 1.5, 2mol %) irradiated with 2 keV, 8.5 µA beam of electrons in a high vacuum chamber containing a base pressure of 1.2 x10^{-8} Torr. The insert shows the CL intensity as a function of different Ce^{3+} concentrations.

4. Conclusion

Enhanced red and blue emissions from Pr^{3+} and Ce^{3+} due to energy transfer from ZnO and Pr^{3+} respectively were demonstrated. The luminescence enhancement of Pr^{3+} due to energy transfer from ZnO to Pr^{3+} in ZnO.SiO_2:Pr^{3+} was confirmed by both PL and CL analysis. Addition of 1 mol% Ce^{3+} to 0.2 mol% Pr^{3+} led to quenching of Pr^{3+} emission while the blue emission from Ce^{3+} was slightly enhanced as compared to SiO_2:Ce^{3+} indicating an energy transfer from Pr^{3+} to Ce^{3+}. The excitation energy was most probably transferred by phonon mediated processes.

5. Acknowledgements

This work was financially supported by the Department of Science and Technology of South Africa and the Council for Scientific and Industrial Research of South Africa.

6. References

[1] Rai V K , Kumar K, Rai S B, 2007 *Opt. Mater.* **29** 873-878
[2] Biswas A, Chakrabarti S, Acharya H N, 1997 *Mater. Sci. Eng.* ***B*** **49** 191
[3] Okamoto S, Kobayashi H, 1999 *J. Appl. Phys.* **86** 10.
[4] Nikl M, Ren G, Ding D, Mihokova E, Jary V, Feng H, 2010 *Chem. Phys. Lett.* **493** 72-75
[5] Blasse G, Grabmaier B C, 1994 Luminescent Materials, *Springer-Verlag*, Berlin
[6] Meijerink A, Wegh R, Vergeer P, Vlugt T, 2006 *Opt. Mat.* **28** 575-581
[7] Jang H S, Im W B, Lee D C, Jeon D Y, Kim S S, 2007 *J. Lumin.* **126** 371-377
[8] Mhlongo G H, Ntwaeaborwa O M, Dhlamini M S, Swart H C, Hillie K T, 2010 *J. Mater. Sci.* **45** 5228-5236
[9] Bang J, Yang H, and Holloway P H, 2005 *J. Chem. Phys.* **123** 084709
[10] Ntwaeaborwa O M, Holloway P H, 2005 *Nanotechnol.* **16** 865-868

The influence of Pr^{3+} co-doping on the photoluminescence and cathodoluminescence properties of SiO_2:Eu^{3+}/Tb^{3+}.

G H Mhlongo[1,2], M S Dhlamini[3], O M Ntwaeaborwa[2*], H C Swart, K T Hillie[1, 2]**

[1]National Centre for Nano-structured Materials, CSIR, PO Box 395, Pretoria, ZA0001

[2]Department of Physics, University of Free State, Bloemfontein, ZA9300.

[3]Department of Physics, University of South Africa, P.O. Box 392, Pretoria, ZA0003

E-mail: ntwaeab@ufs.ac.za, THillie@csir.co.za

Abstract. Tb^{3+}- Pr^{3+}, and Eu^{3+}- Pr^{3+} ion pairs co-doped in a SiO_2 matrix were prepared by a sol-gel method. Red and green photoluminescence associated with $^5D_0 \rightarrow {}^7F_2$ $^5D_4 \rightarrow {}^7F_5$ transitions of Eu^{3+} and Tb^{3+} singly doped in SiO_2 were observed at 614 and 543 nm respectively. Co-doping of Eu^{3+} and Tb^{3+} ions with Pr^{3+} in SiO_2 resulted in the quenching of Eu^{3+} and Tb^{3+} emissions with increasing Pr^{3+}concentrations. The quenching was also confirmed by cathodoluminescence measurements recorded from the same powders. The quenching effects are discussed.

1. Introduction

Enhanced luminescence induced by energy transfer between trivalent rare earth co-activators in host matrices has been demonstrated in nanophosphors. The enhanced luminescence is investigated in a variety of studies for potential application in different types of light emitting devices. Such studies have provided information about the interaction between the activators and the host matrix, as well as an insight in regard to the long-range interactions and energy transfer between the activator ions. For example, Mahato et al [1] studied energy transfer between Pr^{3+} and Tb^{3+} and observed that increasing the Pr^{3+} concentration reduced the emission intensity of Tb^{3+} while that of Pr^{3+} was enhanced indicating the energy transfer from the $^5D_4$ level of Tb^{3+} to the $^3P_0$ level of Pr^{3+}. Wang et al [2] observed the energy transfer from both $^1S_0$ and $^3P_0$ states of Pr^{3+} ions to Er^{3+} ions co-doped in $CaAl_{12}O_{19}$. In some cases, co-activators quench the luminescence intensity instead of enhancing it. For example, Meijerink et al [3] observed quenched luminescence of Pr^{3+} when Eu^{3+} and Pr^{3+} were co-doped in YF_3.

In the present paper, we have synthesized Pr^{3+} co-activated SiO_2:Eu^{3+}/Tb^{3+} using the sol-gel method. The effect of Pr^{3+} co-doping in the luminescence properties under excitation of ultraviolet (UV) light and an electron beam was demonstrated. Fluorescence quenching effects due to Pr^{3+} co-doping are discussed.

[1] O.M Ntwaeaborwa, ntwaeab@ufs.ac.za
[2] T. K. Hillie, thillie@csir.co.za

2. Experimental

SiO_2 phosphors co-doped with different concentrations of Pr^{3+}-Tb^{3+} ion pairs were prepared by mixing 0.05 mol of tetraethylorthosilicate (TEOS), 0.1 mol of H_2O, 0.1 mol of ethanol, and 0.145 mol of dilute nitric acid. The mixture was stirred at room temperature for 1 hour. The resulting transparent sol was mixed with a desired amount of $Pr(NO)_3.6H_2O$ dissolved in 5 ml of ethanol and was stirred for 30 minutes. A desired amount of $Tb(NO)_3.6H_2O$ dissolved in 5 ml of ethanol was also added to the solution and stirred further for 30 minutes until a gel formed. The gel was then dried at room temperature for 3-8 days. The dried gel was ground and annealed in air at 600 ^oC for 2 hrs. A similar procedure was used to prepare SiO_2:Pr^{3+}-Eu^{3+}phosphors with different concentrations of $Eu^{3+}$and Pr^{3+}.

The cathodoluminescence (CL) data of the samples were recorded using an Ocean Optics S2000 Spectrometer. The photoluminescence (PL) data were recorded using a DongWoo Optron Co. setup equipped with a Xe Arc lamp DL180-Xe. The particle morphology was analyzed using a JEOL 2100 High Resolution Transmission Electron Microscopy (HRTEM) while the chemical composition was analyzed by Energy Dispersive Spectrometer (EDS).

3. Results and discussion

3.1. HRTEM, particle morphology and size

Fig. 1 (a), (b), and (c) show the HRTEM images of SiO_2:Tb^{3+}-Pr^{3+}, SiO_2:Eu^{3+}-Pr^{3+}, and pure SiO_2, respectively. As indicated in Fig. 1 (c), pure SiO_2 was amorphous while fig. 1 (a) and (b) revealed the presence of smaller spherical particles enveloped in amorphous SiO_2. As confirmed by EDS data (not shown) these particles were composed of Pr, Eu, and Tb.

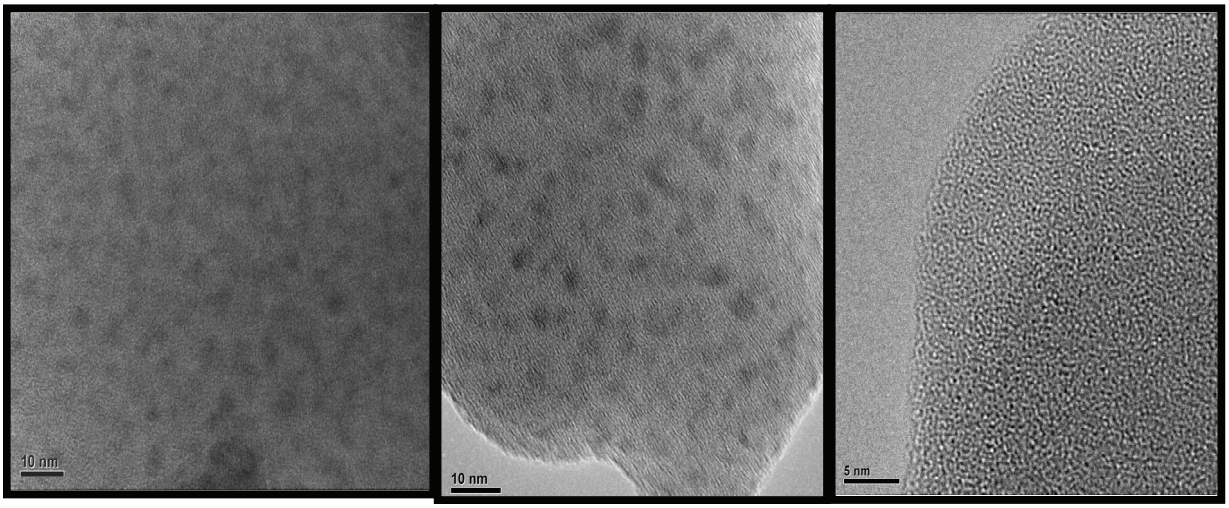


Fig. 1. HRTEM images of (a) SiO_2:Tb^{3+}-Pr^{3+}, (b) SiO_2:Eu^{3+}-Pr^{3+}, and (c) pure SiO_2.

3.2. Photoluminescence (PL) properties of SiO_2:Eu^{3+}-Pr^{3+} and SiO_2:Tb^{3+}-Pr^{3+}

In order to evaluate the effect of Pr^{3+} co-doping in SiO_2:Eu^{3+}, the PL properties were investigated. In fig. 2, the PL emission spectra of SiO_2:1 mol% Eu^{3+} and SiO_2:1 mol% Eu^{3+}-0.2 mol% Pr^{3+} obtained after excitation at 395 nm using a Xenon lamp are compared. It can be seen from the spectra that the emission associated with the $^5D_0 \rightarrow {}^7F_{J\ (J=0,1,2,3,4)}$ transitions of Eu^{3+} with the intense red emission at 614 nm corresponding to the$^5D_0 \rightarrow {}^7F_2$ transition was observed. The assignment of these emission peaks is in accordance with the previous work reported by Changqing et al [5] and Wang et al [6] for SiO_2:Eu^{3+} phosphors. It can be noticed from the spectra that the Eu^{3+} emission is well pronounced even after Pr^{3+}

co-doping and no emission from Pr^{3+} was observed, which indicates no contribution from Pr^{3+}. Also, the relative intensities of SiO_2:Eu^{3+} and SiO_2:Eu^{3+}-Pr^{3+} show that the luminescence from Eu^{3+} ion is strongly quenched due to addition of 0.2 mol% of Pr^{3+}. These results confirm the observation of Kandpal et al [7] for Pr^{3+}-Eu^{3+} ion pairs co-doped in a dimethylsulphoxide (DMSO) system.

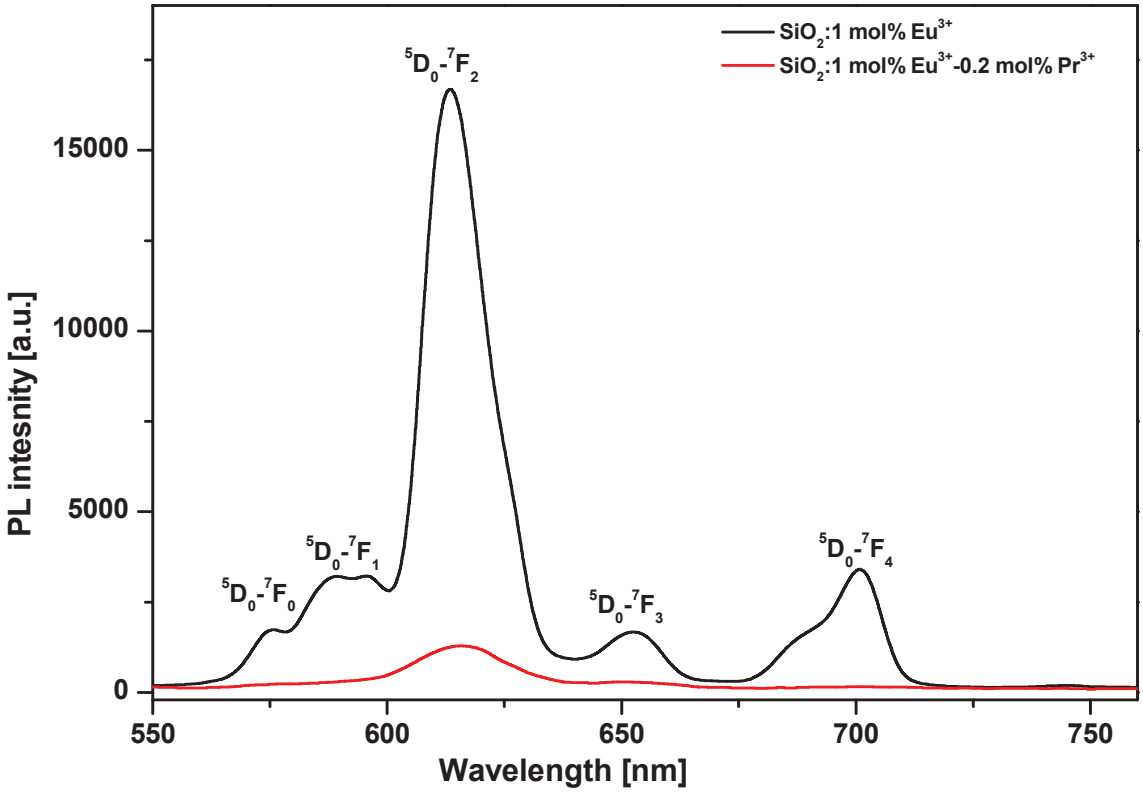


Fig. 2. Emission spectra of SiO_2:1 mol% Eu^{3+} and SiO_2:1 mol% Eu^{3+}-0.2 mol% Pr^{3+} obtained during excitation at 395 nm at room temperature.

The emission spectra of SiO_2:1 mol% Tb^{3+} and SiO_2:1 mol% Tb^{3+}-0.2 mol% Pr^{3+} are compared and shown in figure 3. These spectra were collected during excitation at 371 nm at room temperature. Both spectra exhibit green emission at 543 nm which can be attributed to the $^5D_4 \rightarrow {}^7F_5$ transition and the minor emission peaks associated with $^5D_4 \rightarrow {}^7F_{J\ (J = 0,1,2,3,4)}$ transitions of Tb^{3+} were also observed at 489, 543, 549, 584, 622 nm, respectively. These results agrees well with results reported elsewhere [1,8-10]. Again, when adding Pr^{3+} the luminescence intensity from Tb^{3+} was significantly quenched but to a lesser extent when compared to the data in Figure 2.

The strong quenching of the PL intensity is apparent from all Pr^{3+} co-doped SiO_2:$Eu^{3+/}Tb^{3+}$ phosphor powders. There are two possible reasons which may cause this. First, the aggregation of Pr^{3+} ions or pairing of Pr^{3+} ions with Eu^{3+}/Tb^{3+} ions. Such a process leads to migration of excitation energy from one activator ion to its nearest ion by nonradiative transitions and via a large number of transfer steps, before the energy is released. During the transfer process, the energy can be transferred to defects which act as energy sinks within a transfer chain thus leading to a significant decrease in the PL intensity. In this case, it is most likely to happen that the energy may be exhausted by killer or/and defects during the energy transfer process among one kind of Pr^{3+} ions prior to the other kind of Tb^{3+}/Eu^{3+} ion. Secondly, it is also possible that the interaction between Pr^{3+}-Eu^{3+} and Pr^{3+}-Tb^{3+} ion pairs is through a Pr^{4+}-Eu^{2+} and Pr^{4+}-Tb^{2+} metal-to-metal charge transfer state and this energetically overlaps the Tb^{3+} ($^5D_4$) level or Eu^{3+} ($^5D_0$) level hence the strong quenching [3].

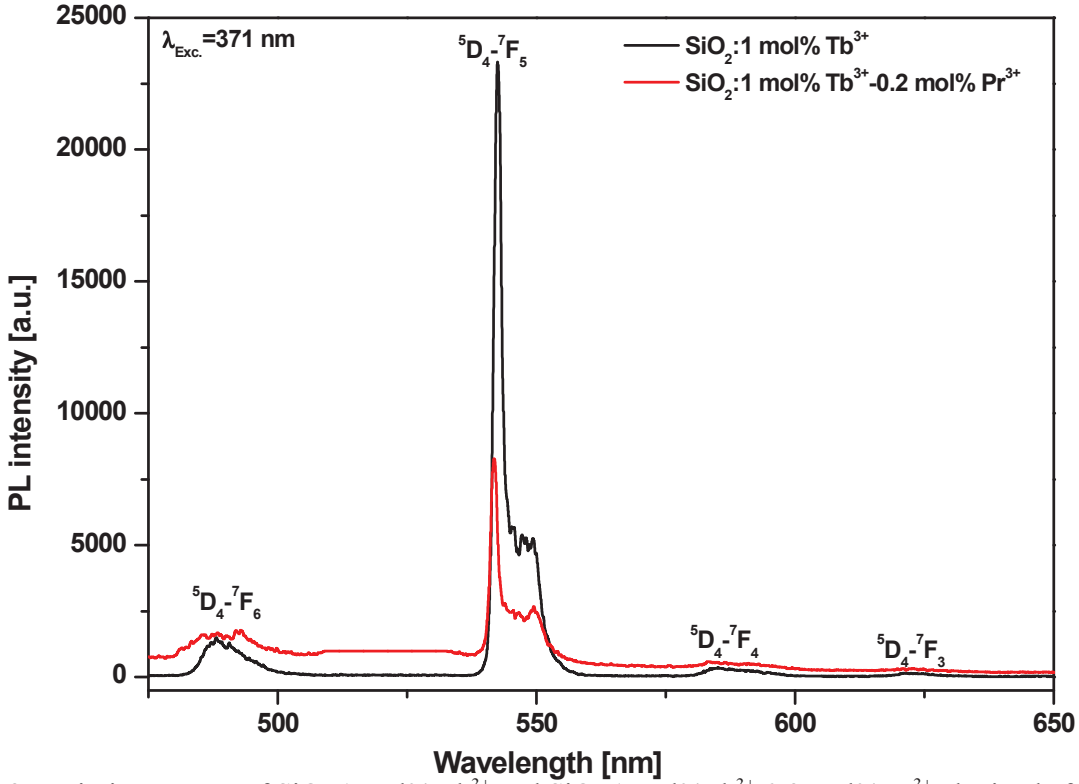


Fig. 3. Emission spectra of SiO_2:1 mol% Tb^{3+} and SiO_2:1 mol%Tb^{3+}-0.2 mol% Pr^{3+} obtained after excitation at 371 nm at room temperature.

3.3. CL properties of SiO_2: Eu^{3+}-Pr^{3+} and SiO_2: Tb^{3+}-Pr^{3+}

The CL properties of Pr^{3+} co-doped in SiO_2:Eu^{3+}/Tb^{3+} phosphor powders were also investigated under electron beam bombardment. In Fig. 4, the CL spectra of SiO_2:1 mol% Tb^{3+}, SiO_2:1 mol% Tb^{3+}-0.2 mol% Pr^{3+} phosphor powders obtained under irradiation with a 2 keV beam of electrons and 20 µA beam current in a high vacuum chamber at a base pressure of ~ 1.6 x 10^{-8} Torr are compared. It was realized that under electron beam excitation, both SiO_2:Tb^{3+} and SiO_2:Tb^{3+}-Pr^{3+} phosphor powders exhibited typical emission lines of Tb^{3+} and a significant decrease in Tb^{3+} emission with addition Pr^{3+} concentration was also observed. However, in comparison with the SiO_2:Tb^{3+}, SiO_2:Tb^{3+}-Pr^{3+} phosphor powders excited by UV light, extra emission peaks associated with the $^5D_3 \rightarrow {}^7F_{3,4}$ transitions of Tb^{3+} in the blue region can be obviously observed for identical samples in the CL spectrum in fig. 4. The insert in fig. 4 shows the CL intensity behavior of the phosphor powders when the concentration of Tb^{3+} is kept constant at 1 mol% while that of Pr^{3+} is varied from 0.2, 0.5, and 1 mol%. The Tb^{3+} emission was shown to significantly decrease with increasing Pr^{3+} concentration.

Fig. 5 presents the CL spectra of SiO_2:1 mol% Eu^{3+}, SiO_2:1 mol% Eu^{3+}-0.2 mol% Pr^{3+} phosphor powders under irradiation with a 2 keV, 20 µA electron beam in a high vacuum chamber containing a base pressure of ~ 1.6 x 10^{-8} Torr. Under electron beam irradiation, the addition of 0.2 mol% of Pr^{3+} into SiO_2:1 mol% of Eu^{3+} completely suppressed the Eu^{3+} emission and only emission from Pr^{3+} associated with *f-f* transitions from $^3P_0 \rightarrow {}^3H_4$, $^3H_6$, $^3F_2$, $^3F_{3,4}$ and $^1D_2 \rightarrow {}^3H_4$, $^3H_5$, $^3F_4$ with the main red emission at 614 nm assigned to $^3P_0 \rightarrow {}^3H_6$ was observed. It was realized that the CL intensity of SiO_2: Eu^{3+}-Pr^{3+} kept on dropping with increasing concentrations of Pr^{3+} (0.5 and 1 mol%) as shown in an insert in fig. 5. Such results confirm the observations of Zachau et al [11].

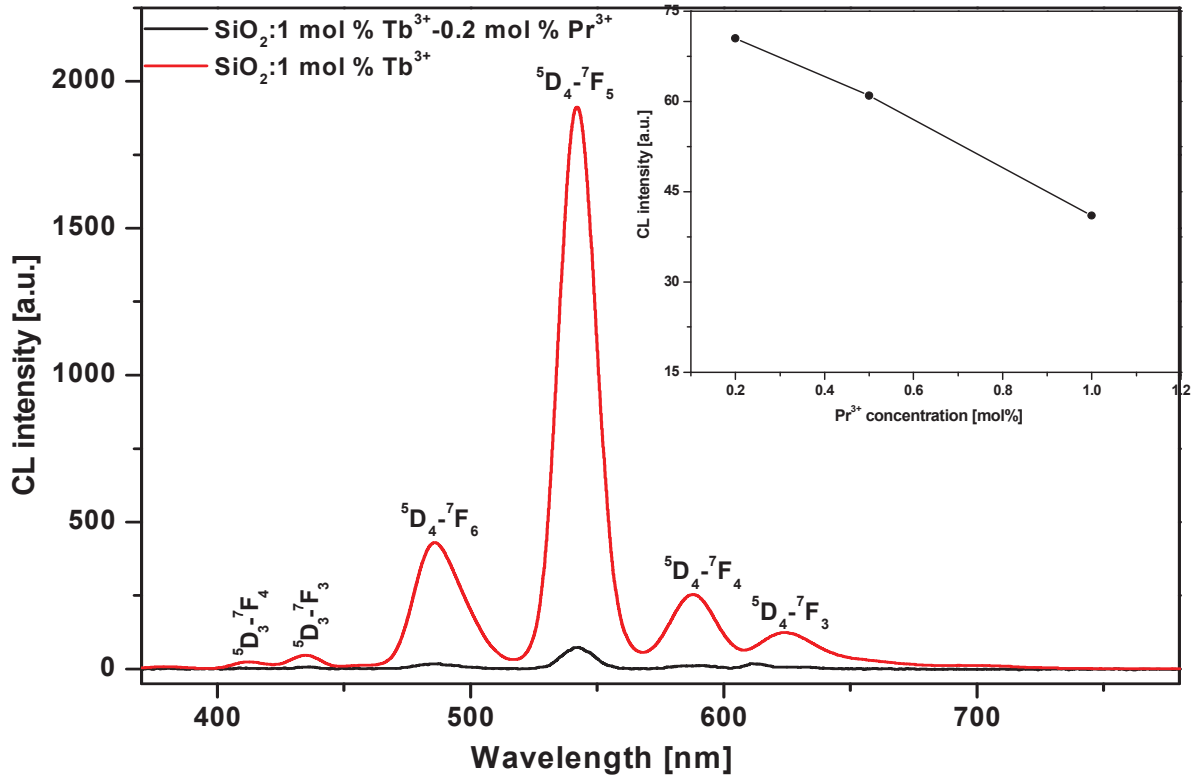


Fig. 4. CL intensities of $SiO_2$1 mol% Tb^{3+} and SiO_2:1 mol% Tb^{3+}-0.2 mol% Pr^{3+} at 2 keV, 20 μA in a high vacuum pressure of 1.6 x10^{-8} Torr. The inset shows the CL intensity from SiO_2:Pr^{3+}-Tb^{3+} as a function of 0.2, 0.5, and 1 mol% Pr^{3+} concentrations.

The CL analysis also revealed the strong quenching with Pr^{3+} addition of the luminescence from both SiO_2: Eu^{3+}-Pr^{3+} and SiO_2: Tb^{3+}-Pr^{3+} phosphor powders. However, we observed an obvious difference between the CL and PL spectra. Such difference between the CL and PL spectra obtained from phosphor samples may be attributed to two possible factors. First, the differences of the excitation mechanisms involved during the PL and CL. During the PL process, the photons used for excitation have energy less than the bandgap of the host matrix so that no electron-hole pairs are generated. While in the case of CL, the generation of charge carriers (electrons and holes) under irradiation by energetic electrons occurs through impact ionization. As a result, the accelerated electrons create a multiplicity of charge carriers and the excitation energy is transferred by excitons to luminescent centres. Secondly, in comparison with the UV excitation, the excitation intensity and excitation density of luminescence centres are very high under electron beam bombardment. For example, in the case of Tb^{3+} doped samples, the cross relaxation $^5D_3+^7F_6 \rightarrow ^5D_4+^7F_0$ may be depressed to some extent, hence, the appearance of the $^5D_3$ emission with electron beam irradiation.

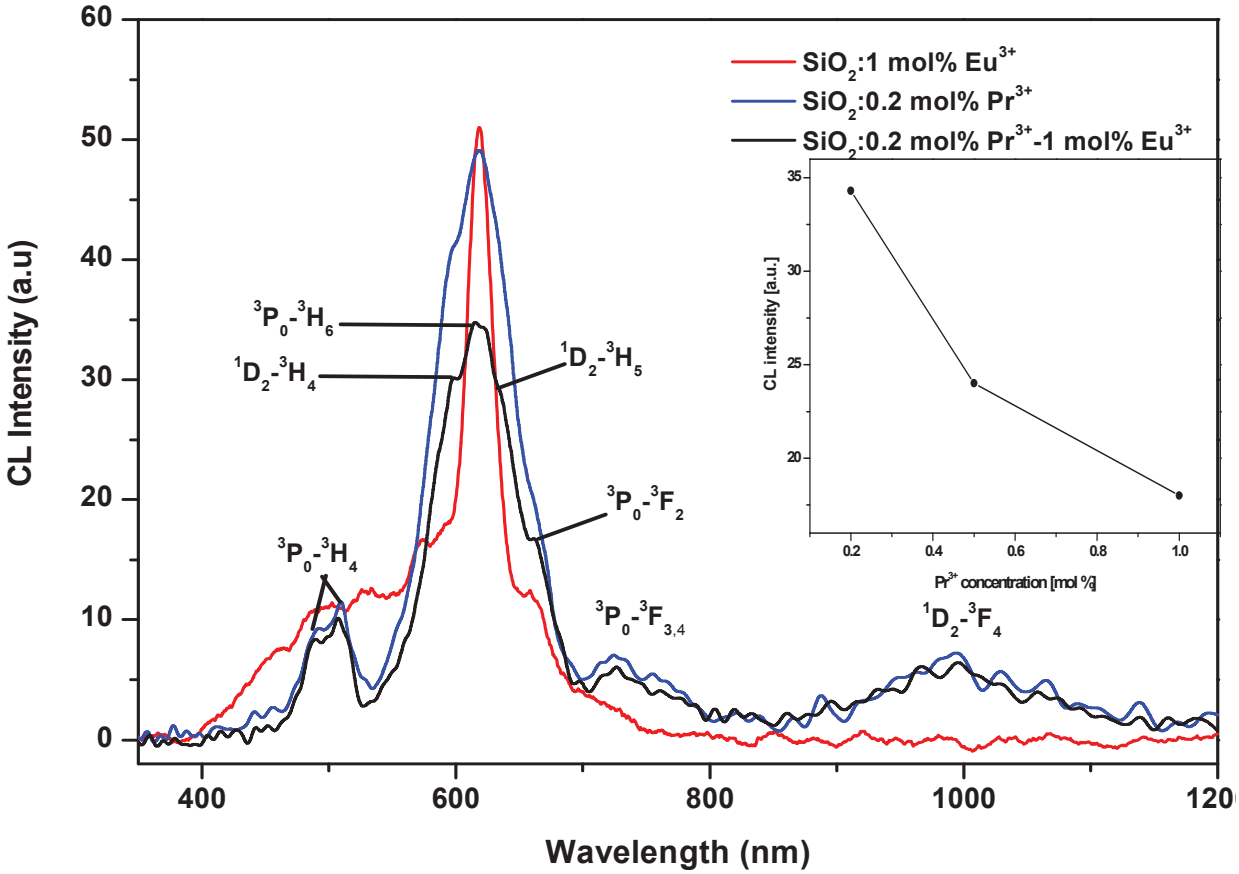


Fig. 5. CL intensities of SiO_2:1 mol% Eu^{3+}, SiO_2:0.2 mol% Pr^{3+}, and SiO_2:1 mol% Eu^{3+}-0.2 mol% Pr^{3+} at 2 keV, 20 μA in a high vacuum pressure of 1.6 x10^{-8} Torr. The inset shows the CL intensity from SiO_2:Pr^{3+}-Eu^{3+} as a function of 0.2, 0.5, and 1 mol% Pr^{3+} concentrations.

4. Conclusion

The CL and PL properties of SiO_2:Tb^{3+}/Eu^{3+} co-doped with Pr^{3+} derived from sol gel method were studied by varying the concentration of Pr^{3+}. The strong fluorescence quenching with increasing Pr^{3+} concentration was observed from both SiO_2:Tb^{3+}/Eu^{3+} co-doped with Pr^{3+} phosphor powders and no sign of Pr^{3+} emission was observed. The possible reasons which might have been the main cause of this quenching are discussed in detail.

5. Acknowledgements

This work was financially supported by the Department of Science and Technology of South Africa and the Council for Scientific and Industrial Research of South Africa. Authors would like to thank Dr P Solarz and Prof. W Ryba-Romanowski from the Institute of Low Temperature and Structure Research in Polish Academy of Sciences, Poland for assisting with PL measurements.

6. References

[1] Mahato K K, Rai S B, 2000 *Spectrochim. Acta Part A* **56** 2333-2340
[2] Wang X–J, Huang S, Lu L, Yen W M, Srivastava A M, Setlur A A, 2001 *Optics Comm.* **195** 405-410
[3] Meijerink A, Wegh R, Vergeer P, Vlugt T, 2006 *Opt. Mater.* **28** 575-581
[4] Mhlongo G H, Ntwaeaborwa O M, Dhlamini M S, Swart H C, Hillie K T, 2011 *J. Alloys Compd.* **509** 2986-2992
[5] Changqing H, Ting S, Weijian T, Baoping Z, 2006 *J. Rare Earths* **24** 134-137
[6] Wang R Y, 2004 *J. Lumin.* **106** 211-217
[7] Kandpal H C, Tripathi H B, 1979 *Solid State Comm.* **29** 139-142
[8] Liu X, Pang R., Li Q, Lin J, 2007 *J. Solid State Chem.* **180** 1421-1430
[9] Silversmith A J, Boye D M, Brewer K S, Gillespie C E, Lu Y, Campbell D L, 2006 *J. Lumin.* **121** 14-20
[10] Wakefield G, Keron H A, Dobson P J, Hutchison J L, 1999 *J. Phys. Chem. Sol.* **60** 503-508
[11] Zachau M, Zwaschka F, Kummer F, 1998 Proccedings- *Electrochem. Soc.* **97(**29) 314–324

Nanoscale manipulation of lamellar copolymers using electric fields

Simiso K Mkhonta

Department of Physics, University of Swaziland, Private Bag 4, Kwaluseni, Swaziland

E-mail: smkhonta@uniswa.sz

Abstract. Block copolymers can self-assemble into lamellar and cylindrical phases with a fundamental period of 10–100nm. These nanoscopic phases are a basis for a great range of applications in nanotechnology. We describe how electric fields can be utilized to tune the lamellar period of copolymers that consist of liquid-crystal sub-units. We evaluate the critical electric field required to tilt the liquid-crystal director relative to the lamellar normal in terms of the repulsive interaction between the chemical dissimilar copolymer blocks. The tilting of the liquid-crystal director leads to the tuning of the lamellar spacing. This electrically induced changes of the block copolymer microstructure can lead to interesting applications.

1. Introduction

Block copolymers consists of two or more incompatible polymers that are covalently bonded together. To minimize the interfacial energy between dissimilar blocks, a melt of block copolymers often self assemble into nano-periodic structures. Typical ordered phases of block copolymers include lamellar, cylindrical, and cubic phases [1]. The lattice parameter of these block copolymer crystals depends on the size of the polymer chains and lies within the range of 10–100nm.

Block copolymers are important in condensed matter physics research due to their rich phase behavior. They also offer promising applications in nanotechnology due to their regular nanometer-scale patterns [2]. Research efforts in block copolymers have focused on how to direct the self-assemble process in order to obtain defect-free patterns and as well as on achieving particular orientations of anisotropic copolymer structures. Recently, experimental research has been carried on block copolymers with tunable micro-domain sizes [3, 4]. A nanostructured material with tunable density modulations is desirable in the fabrication of photonic crystals with a full color display.

In this work we investigate the behavior of electrically tunable block copolymers that incorporate liquid-crystal molecules. Flexible polymers adopt a uniaxial ellipsoidal conformation in the presence of orientationally ordered liquid-crystals as illustrated in Fig. 1. The uniaxial direction of the ellipsoidal chains can be controlled by external fields. For example, in the lamellar phase the uniaxial direction of the ellipsoidal chains is expected to be parallel to the lamellar direction [5]. However, an electric field can induce relative rotations leading to thinner lamellae as shown in Fig. 2. This mechanism is expected to occur at modest electric fields due to the large dielectric anisotropy of the liquid-crystals. Our analytical calculations show that

the amplitude of the layer contractions depends on the aspect ratio of the ellipsoidal chains and can be unusually large for highly anisotropic chains.

2. Model

We consider a system of A-B diblock copolymers with side-chain liquid crystals as illustrated in Fig. 1. We can describe this system using two fields: $\psi(\mathbf{r})$ and $\hat{\mathbf{n}}(\mathbf{r})$, the deviation of the local density of one block species from its spatial average and the local director of the liquid-crystal molecules, respectively. The liquid-crystal director is a unit vector that describes the preferred orientation of the liquid crystals. We assume that it is confined to the xy plane: $\mathbf{n} = [\cos\theta(\mathbf{r}), \sin\theta(\mathbf{r}), 0]$. In the weak segregation limit, this system can be described by the following free energy: [5]

$$\begin{aligned}\frac{Na^3}{k_BT}F_{ab} &= \int d^3r \,\{\tau\psi^2 + \psi^4/12 + \\ &+ \xi^2\left[q_0^2\psi + \mathcal{A}(\theta)\partial_{xx}\psi + \mathcal{B}(\theta)\partial_{yy}\psi + \mathcal{C}(\theta)\partial_{xy}\psi + \partial_{zz}\psi\right]^2 + \\ &+ K_1(\nabla\cdot\hat{\mathbf{n}})^2 + K_3(\hat{\mathbf{n}}\times\nabla\times\hat{\mathbf{n}})^2 + \alpha\left(\hat{\mathbf{n}}\cdot\vec{\nabla}\psi\right)^2\}/2, \qquad (1)\end{aligned}$$

where $\mathcal{A}(\theta) = \cos^2\theta + \kappa^2\sin^2\theta$, $\mathcal{B}(\theta) = \sin^2\theta + \kappa^2\cos^2\theta$, and $\mathcal{C}(\theta) = (\kappa^2 - 1)\sin 2\theta$. In the orientationally ordered liquid-crystalline state, $\theta(\mathbf{r}) \approx$ constant, the structure factor that describes the density fluctuations of the polymeric chains is anisotropic. The characteristic wavevectors of the density fluctuations lie within an ellipsoidal shell. This anisotropy reflects the two characteristic lengths of individual ellipsoidal chains: their diameter $d = 2\pi/q_0$ and length $l = 2\pi\kappa/q_0$. The parameter $\kappa > 1$ corresponds to the aspect ratio of the polymer chains when the liquid-crystal sub-units are uniformly orientated. The nematic elasticity of the liquid-crystal is defined by two moduli: K_1 and K_3, the splay and bend moduli, respectively. The parameter $\tau = 2N(\chi_c - \chi)$, measures the distance to the critical point ($\tau = 0$) in terms of the Flory-Huggins interaction parameter and $N = N_A + N_B$ is the total chain polymerization index. At the critical point $N\chi_c \approx 10.49$. In addition, a is the Kuhn statistical segment and $\xi^2 = 6c^2/q_0^4$ where c is a constant of order 1. The coefficients λ and u depend on the block ratio $f_A = N_A/N$. Below we would restrict ourselves to the case where the block ratio is symmetric $f = 1/2$. In this limit $\lambda = 0$ and the lamellar phase is the only stable ordered phase above the critical value χ_c. For convenience we set $c = u = 1$.

Finally, the last term in Eq. (1) couples of the liquid-crystal director to the orientation of the block copolymer microstructure. It favors an orientation of the liquid-crystals that minimizes the interfacial energy between the A-B domains.

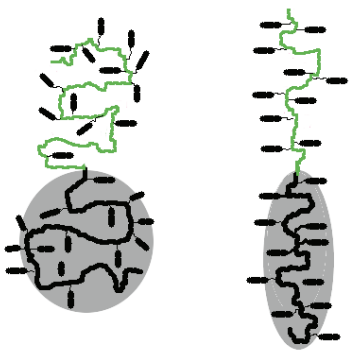


Figure 1. Schematic diagram of a liquid-crystal diblock copolymer chain. The shape of the copolymer chains depends on the orientational order of the side-chain liquid-crystals. Nematic ordering of the liquid-crystals leads to an ellipsoidal chain conformation.

2.1. Electric field contribution

An applied uniform electric field will have an alignment effect both on the liquid-crystals and copolymer microstructure. The lamellar phase has an inhomogeneous dielectric constant due to the stacking of chemically distinct polymers and is also dielectric anisotropic due to the liquid crystals. The electric field contribution to the free energy is then [6, 7],

$$\frac{Na^3}{k_BT}F_{el} = \frac{\beta}{2}\int d^3r \left(\vec{\nabla}\psi/|\vec{\nabla}\psi|\cdot\vec{E}\right)^2\psi^2 - \frac{\Omega}{2}\int(\mathbf{E}\cdot\mathbf{n})^2 d^3r, \tag{2}$$

where $\beta = Na^3(\varepsilon_A - \varepsilon_B)^2/[k_BT(\varepsilon_A + \varepsilon_B)]$ for symmetric AB diblock copolymer, ε_A and ε_B are the dielectric constants for the A/B blocks, k_B is the Boltzmann constant, and T is the temperature of the system. This means that there is a energy cost when the dielectric interfaces are perpendicular to the electric field. We also have $\Omega = Na^3\Delta\varepsilon\varepsilon_0/(k_BT)$, where ε_0 is the permittivity of free space and $\Delta\varepsilon = \varepsilon_{\|} - \varepsilon_{\perp}$ is the dielectric anisotropy, the difference between the liquid-crystal dielectric constant along their long and short axes.

3. Results

To understand the effects of a uniform field on the microstructure of liquid-crystalline block copolymers, we consider a lamellar phase with the liquid crystals oriented parallel to the layers as illustrated in Fig. 2. We assume that the applied uniform fields $\mathbf{E} = E\hat{\mathbf{x}}$ and that the liquid-crystals have a negative dielectric anisotropy and thus the electric field will favor a 90° reorientation of the liquid-crystal director.

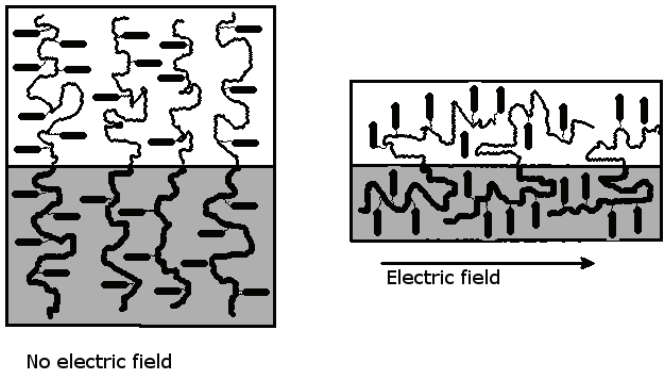


Figure 2. Schematic diagram of an electrically tunable block copolymer. The side-chain liquid-crystal molecules have a negative dielectric anisotropy, they prefer to orient perpendicular to the field. As the nematic director tilts so does the coupled anisotropic chains leading to the lamellar contraction.

3.1. Mean field Calculations

The energetics of the system are then described by the total free energy

$$\mathcal{F} = F_{ab} + F_{el}. \tag{3}$$

For a given electric field strength, E, we can find the mean field solutions of $\psi(\mathbf{r})$ and $\theta(\mathbf{r})$ that minimize the total free energy. Very close to the orientationally ordered state, i.e., $\theta(\mathbf{r}) \approx$ constant, the terms involving the gradients of $\hat{\mathbf{n}}$ have a vanishingly small contribution to the total free energy of the system.

Within the weak segregation limit the local density profile of lamellar diblock copolymers can be approximated by a sinusoidal function [7]

$$\psi(y) = D\cos(qy), \tag{4}$$

where $2\pi/q$ is the fundamental period of the lamellae and D is the amplitude of the density modulations. The lowest energy state is obtained under the condition that the amplitude D and the characteristic wavenumber q minimize the free energy of the system. Substituting Eq. (4) into Eq. (3) we obtain the free energy per unit volume

$$f = \frac{Na^3}{k_B T}\frac{\mathcal{F}}{V} = \frac{1}{4}\left[-|\tau| + \alpha q^2 \sin^2\theta + \xi^2\left(\mathcal{B}q^2 - q_o^2\right)^2\right] D^2 + \frac{1}{64}D^4 + \frac{|\Omega|}{2}E^2\cos^2\theta. \quad (5)$$

Minimizing the above free energy density with respect to D and q, we obtain the equilibrium amplitude D_m and characteristic wavenumber q_m. In the off state ($E = 0$) we have $D_m^b = 2\sqrt{-\tau}$ and $q_m^b = q_0/\kappa$. In the on-state ($E \neq 0$)

$$D_m^a = \sqrt{8[|\tau| - \alpha q^2 \sin^2\theta - \xi^2(q_0^2 - \mathcal{B}q^2)^2]} \quad (6)$$

and

$$q_m^a = \sqrt{\frac{q_0^2}{\mathcal{B}} - \frac{\alpha \sin^2\theta}{2\xi^2\mathcal{B}^2}}, \quad (7)$$

where $\mathcal{B} = \kappa^2 + (1-\kappa^2)\sin^2\theta$. The value of $\theta(\mathbf{r})$ that minimizes the free energy is naturally a strong function of the applied electric field. Substitute Eq. (6) and Eq (7) into Eq. (5) we obtain the free energy density as the functional of the liquid-crystal director angle, $f(\theta)$. In the small field limit θ (and $\sin(\theta)$) is small and the free energy can be expanded as follows,

$$\begin{aligned} f(\theta) &= -|\tau^2| + \frac{|\Omega|E^2}{2} + \left(2|\tau|\alpha\frac{q_0^2}{\kappa^2} - \frac{|\Omega|E^2}{2}\right)\sin^2\theta \\ &\quad + \frac{\alpha q_0^2}{\kappa^2}\left[2|\tau|(1-\kappa^{-2}) - \frac{\alpha q_0^2}{\kappa^2}\left(\frac{|\tau|}{2\xi^2 q_0^4} + 1\right)\right]\sin^4\theta + \mathcal{O}(\sin^6\theta). \end{aligned} \quad (8)$$

It can be seen in Eq. (8) that the electric fields alters the coefficient of the $\sin^2\theta$ term. The coefficient changes sign at the critical electric field

$$E_c = 2\sqrt{\frac{|\tau|\alpha}{|\Omega|}}\frac{q_0}{\kappa}. \quad (9)$$

above which it becomes energetically favorable for electric field to tilt the liquid-crystals molecules relative to block copolymer layers. When E is slightly above E_c the free energy is minimized when,

$$\sin^2\theta = \frac{(E/E_c)^2 - 1}{2(1-\kappa^{-2}) - \alpha q_0^2[1/|\tau| + 1/(2\xi^2 q_0^4)/\kappa^2]}. \quad (10)$$

A consequence of the liquid-crystal reorientation induced by the electric field is the change of the lamellar spacing. The strain associated with layer contraction is

$$\delta = (t_b - t_a)/t_b \quad (11)$$

where $t_b = 2\pi/q_m^b$ and $t_a = 2\pi/q_m^a$ are the layer width at $E < E_c$ and $E > E_c$ respectively. Utilizing Eq. (7) and Eq. (10) we can obtain the strain that is achieved at a full 90° liquid-crystal director rotation

$$\delta_{max} = 1 - \kappa^{-1}\left(1 - \frac{\alpha}{2\xi^2 q_0^2}\right)^{-1/2}. \quad (12)$$

In the limit of a soft anchoring of the liquid crystals to the lamellar surfaces, $\delta_{max} = 1 - \kappa^{-1}$. This means the lamellar spacing of a system with large aspect-ratio ellipsoidal chains, $\kappa >> 1$, can be tuned approximately by 100% in a reversible way.

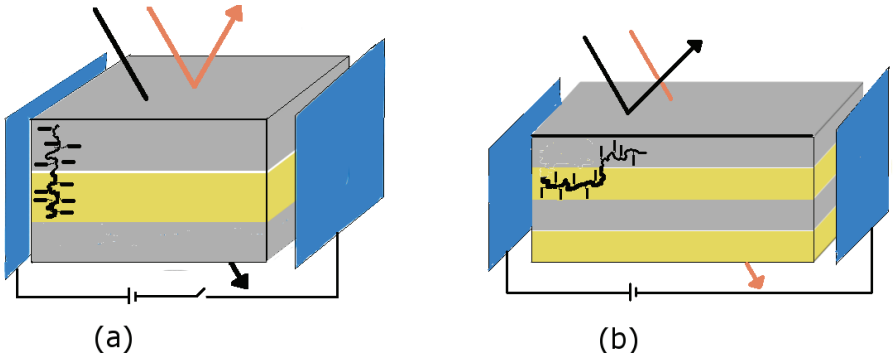


Figure 3. Schematic diagram of an electrically tunable block copolymer photonic crystal exhibiting colored stop bands.
(a)Off-state: the lamellar sample completely reflects long-wavelength in the visible spectrum (infrared). (b) On state: contraction of the lamellar spacing leads to complete reflection of shorter wavelengths in the visible spectrum (ultraviolet). The inset diagram shows the orientation of the anisotropic copolymer chains within the lamellar structure.

3.2. Applications
The tunability of the block copolymer microstructure can lead to useful applications. For example in photonic applications. The wavelength of light, λ, that is coherently reflected by the polymer photonic crystal depends on the lamellar periodicity d_l:

$$\lambda \approx 2d_l\sqrt{n_e^2 - \sin^2\phi} \tag{13}$$

where n_e is the effective refractive index of the material and ϕ is the angle of incidence of light with respect to the lamellar normal. When the electrically induced liquid-crystal rotation changes $d_l n_e$ from $350 - 700$ nm, we can obtain a polymer photonic crystal that operates in the visible light spectrum, coherently reflecting different colors as the electric field intensity is varied as shown in Fig. 3.

4. Summary
In summary, we have presented a model that describes microphase separation of block copolymers that incorporated liquid-crystals. We have shown that the electrically induced relative rotations of the liquid-crystal sub-units can lead to unusually large changes of the characteristic spacing of the block copolymer microstructure. This electric field effect may enhance the use of block copolymers in photonic applications.

Acknowledgments
The author is grateful to Ken Elder and Martin Grant for valuable discussions.

References
[1] Hamley I W *Development in block copolymer Science and Technology* 2004 (London: John Wiley and Sons).
[2] Park C, Yoon J, and Thomas E L 2003 *Polymer* **44** 6725-60
[3] Lu Y, Xia H, Zhang G, and Wu C 2009 *J. Mater. Chem.* **19** 5952-55
[4] Kang Y, Walish J, Gorishnyy T and Thomas E L 2007, *Nature Materials* **6** 957-960
[5] Mkhonta S K, Elder K R, and Grant M, *Eur. Phys. J. E* 2010 **32**, 349-355
[6] Kleman M and Lavretovich O 2003 *Soft Matter Physics: An Introduction* (New York: Springer-Verlag) p 87
[7] Tsori Y, Tournilhac F, Andelman D, and Leibler L 2003 *Phys. Rev. Lett.* **90** 145504

Optical properties of $SrGa_2S_4:Ce^{3+}$ thin films prepared by pulsed reactive cross laser ablation (PRCLA) technique

P A Moleme, J J Dolo, H C Swart and O M Ntwaeaborwa[1]

Department of Physics, University of the Free State, P O Box 339, Bloemfontein ZA 9300, South Africa.

e-mail: ntwaeab@ufs.ac.za

Abstract. $SrGa_2S_4:Ce^{3+}$ thin films were prepared by the pulsed reactive cross laser ablation (PRCLA) technique. Characterization of the films was carried out with scanning electron microscopy (SEM), atomic force microscopy (AFM) and X-ray diffraction (XRD). Cathodoluminescence (CL) and photoluminescence (PL) properties of the films were investigated. The films were prepared using different substrate temperatures, number of pulses and working pressures. A highly crystalline $SrGa_2S_4$ layer was observed when the substrate temperature of 400°C was used. The XRD patterns showed that the growth of the films was sensitive to substrate temperature. PL and CL showed one broad band that can be fitted with two Gaussian peaks according to the two Ce^{3+} radiative transitions at lower substrate temperatures, low and high pulses and in O_2 atmosphere. At high substrate temperature and in Ar atmosphere, the emission peaks were red-shifted to resemble Ce^{3+} emission in a SrS host for both UV and high energy electrons excitation. The AFM images before annealing exhibited a smooth surface at a low substrate temperature, which became rougher at high substrate temperatures and after annealing in vacuum at a temperature of 700°C. Non-uniformity in particles of the films and smooth surfaces were observed from the SEM images.

1. Introduction

Thin films of commercial $SrGa_2S_4:Ce^{3+}$ phosphor are promising candidates for full colour electroluminescence (EL) and field emission displays (FED) because of the phosphor's good optical properties [1]. These films were previously prepared using different techniques such as Radio Frequency (RF) sputtering, molecular beam epitaxy (MBE), reactive multi-source deposition (MSD), deposition from binary depositions (DVB), reactive multisource deposition, flash evaporation method, metal-organic chemical vapour deposition (MOCVD), two target pulse pulsed-electron-beam evaporation (EBE) and Pulsed Laser Deposition (PLD) techniques [2], [3]. In this study the films were prepared using the pulsed reactive cross laser ablation (PRCLA) technique in which the gas pulse crosses the laser plume. The PRCLA technique has advantage over the conventional PLD because of a much brighter plume due to an increased collision between the ionized gas pulse and reactive ablated species from the target material. The structure, morphology, topography and luminescent (photoluminescence and cathdoluminescent) properties of the films were investigated.

2. Experimental details

Silicon (Si) (100) substrates were first cleaned with KIMTECH SCIENCE KIMWIPES obtained from Kimberly Clark Professionals. A pellet with a 2.4 cm diameter and 6 mm thickness was prepared by pressing the $SrGa_2S_4:Ce^{3+}$ powder for 1 hour at a pressure of 1.96×10^7 mbar. The pellet was then annealed for 6 hours at 600°C in vacuum to improve its hardness. It was then mounted on a rotating holder lying diagonally to a heater on which Si substrates were mounted. The distance between the

[1] Corresponding author: ntwaeab@ufs.ac.za

target and the substrates was maintained at 4 cm during the deposition of each film. The Lambda Physik EMG 203 MSC 308 nm XeCl excimer laser was used to ablate the target. The films growth was carried out in a chamber which was first evacuated to a base pressure of 8 x 10^{-5} mbar before backfilling to pressures of 1.0 x 10^{-2} mbar Ar and 1.0 x 10^{-2} mbar O_2, where Ar and O_2 were used as pulse gases. The films were deposited at different substrate temperatures ranging from 400°C to 600°C using 28 800 and 57 600 pulses. The laser beam was operated at an 8 Hz repetitive rate. The substrate temperature, number of pulses and the working pressure were varied during the deposition of the thin phosphor films. Characterization of the films was carried out with scanning electron microscope (SEM), atomic force microscopy (AFM) and X-ray diffraction (XRD). Cathodoluminescent (CL) and photoluminescent (PL) spectra were recorded with a S2000 Ocean Optics spectrometer and Varian Cary Eclipse fluorescence spectrophotometer respectively. Auger electron spectroscopy (AES) was used to analyze the elemental (chemical) composition of the films.

3. Results and discussions

Shown in figure 1 are the XRD patterns and the Miller indices of the $SrGa_2S_4:Ce^{3+}$ films prepared at different substrate temperatures ranging from 400 to 600°C. A highly crystalline $SrGa_2S_4$ layer was observed at the growth temperature of 400°C with 28 800 pulses in an Ar environment. From the comparison with the standard powder pattern of $SrGa_2S_4$ (JCPDS file no. 77-1189) all the peaks, except an identified impurity peak (marked x) at $2\theta = \sim 47^{o}$, in figure 1 (a) were found to belong to the orthorhombic $SrGa_2S_4$ crystal structure. The (10 4 4) diffraction peak is more intense than that observed from $SrGa_2S_4:Ce^{3+}$ powder reported elsewhere [4].

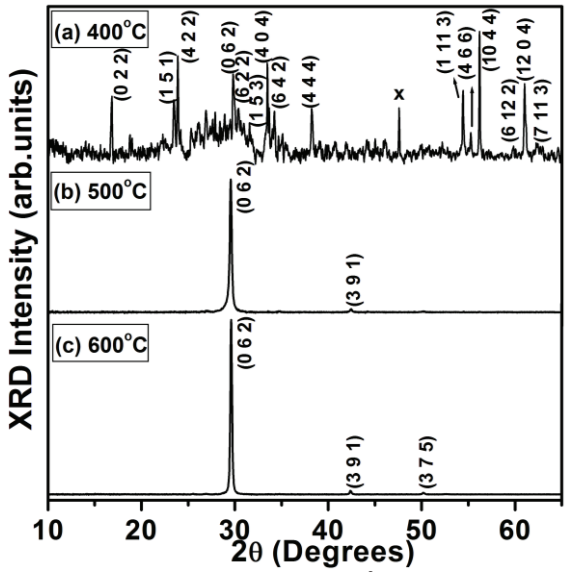


Figure 1. The diffraction patterns of the $SrGa_2S_4:Ce^{3+}$ films deposited at different substrate temperatures.

A preferential growth along the orientation (0 6 2) was observed when the substrate temperature was increased to 500°C and 600°C. In addition, small signatures of the (3 9 1) and (3 7 5) peaks which also belong to the orthorhombic $SrGa_2S_4$ crystal structure were detected. Similar results as the once observed at 400°C and 500°C substrate temperatures were reported by Heikkinen et al [5]. The XRD patterns in figure 1 show that the film growth is sensitive to the substrate temperature. The intensity of the (0 6 2) peak was observed to increase with an increase in the substrate temperature from $T_{substrate}$ = 400°C to 600°C. These results also indicate that crystallinity can be achieved even without post-deposition annealing. Tanaka et al [6] investigated the crystallinity of the $SrGa_2S_4:Ce^{3+}$ thin films grown on quarts glass substrates with the MBE technique. The films were grown at substrate

temperatures of 400°C to 600°C. They reported a gradual decrease in the XRD peak intensity at $T_{substrate} > 600°C$ and $T_{substrate} < 500°C$ and found the best substrate temperature for $SrGa_2S_4$ film growth to be around 560°C. Yang et al [7] reported the same substrate temperature as the best for $SrGa_2S_4$ films growth. The XRD patterns for the films deposited at different number of pulses and in O_2 environment exhibited poor crystallinity.

For the range of all deposition conditions investigated, the SEM images of all films exhibited small and big spherical particles distributed unevenly on the smooth surfaces of the films. Auger survey was performed on the films and the major elements, namely, Sr, Ga, and S were detected from all the films. In addition, atmospheric O and C were also detected. Shown in figure 2 (a) is the SEM image of the surface of the $SrGa_2S_4:Ce^{3+}$ film deposited at 400°C and in figure 2 (b) is the Auger survey spectrum for the film.

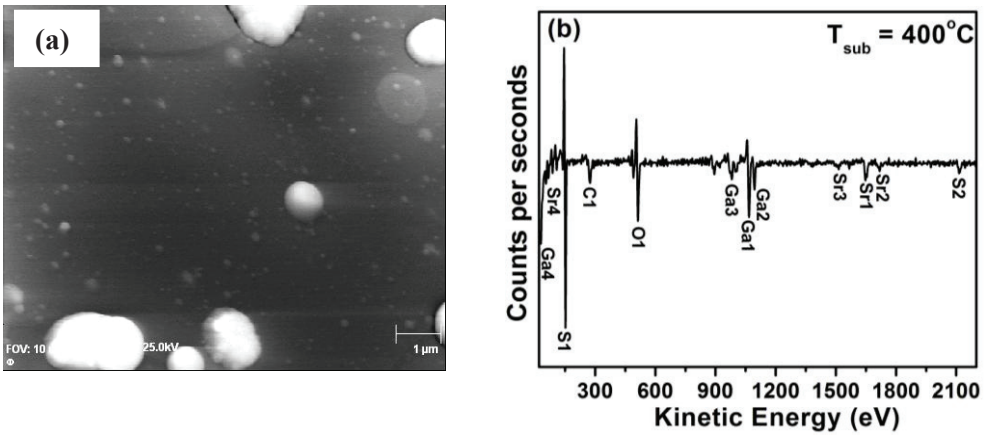


Figure 2 (a) The SEM micrograph and **(b)** Auger survey spectrum of the $SrGa_2S_4:Ce^{3+}$ films deposited at 400°C.

Figure 3 (a) – (c) and (d) – (e) show respectively the AFM images of the films deposited at the substrate temperatures of 400, 500, and 600°C before and after annealing at 700°C in vacuum.

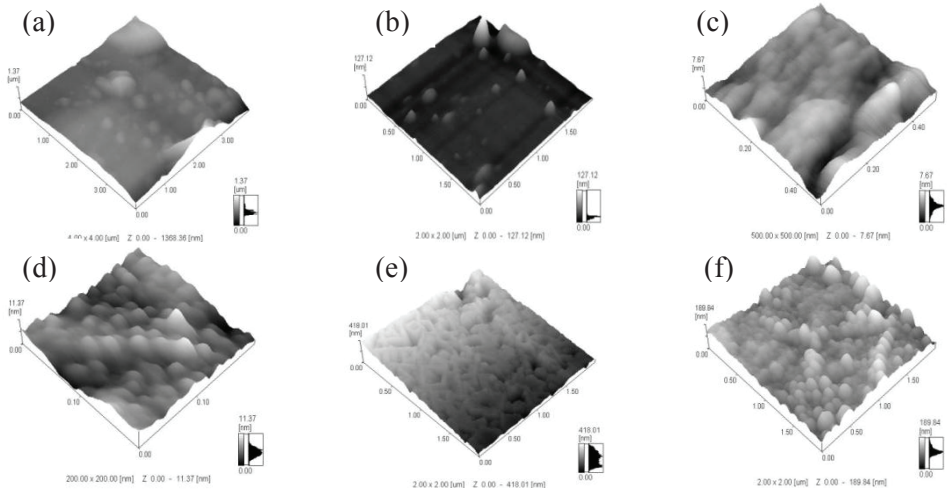


Figure 3. AFM images for the un-annealed $SrGa_2S_4:Ce^{3+}$ films prepared at **(a)** 400°C, **(b)** 500°C, **(c)** 600°C. Images **(d), (e)** and **(f)** show, respectively, the films annealed in vacuum at 700°C.

At 400°C (figure 3 (a)), the film was smooth with uneven distribution of few spherical particles of different sizes on the surface. After annealing (figure 3 (d)) the surface became rough with an increased number of spherical particles on the surface. The smooth surface with steps was observed at 500°C as depicted in figure 3 (b), with few non-uniform cylindrical particles of different sizes. Generally, all the annealing films were fairly rougher than the unannealed ones.

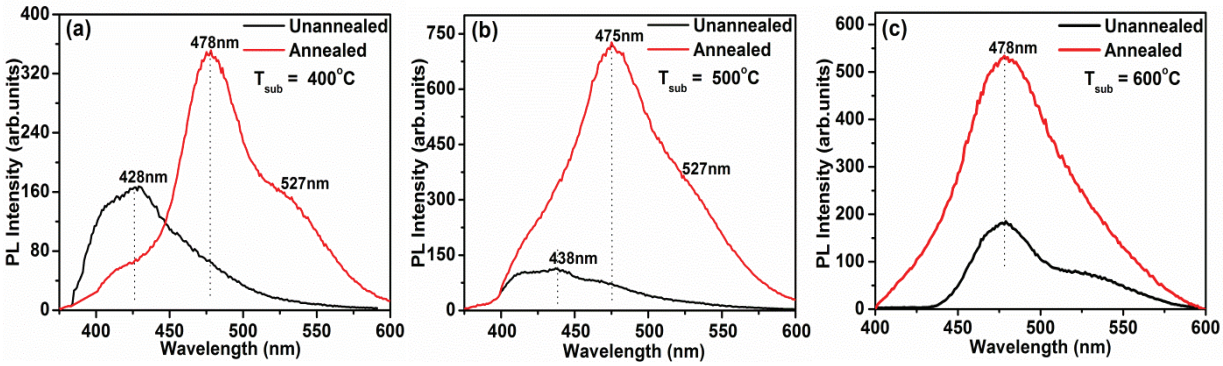


Figure 4. PL spectra of the $SrGa_2S_4:Ce^{3+}$ films prepared at **(a)** different substrate temperatures, **(b)** different number of pulses, and **(c)** in Ar and O_2 atmospheres.

Shown in figure 4 (a-c) are the PL emission (λ_{exc} = 230 nm) spectra for the films deposited at different substrate temperatures and after annealing at 700°C. At 400°C, one broad band that can be fitted with two Gaussian peaks according to the two Ce^{3+} emission peaks which are known to originate from 5d (T_{2g}) → 4f ($^2F_{5/2}$) and from 5d (T_{2g}) → 4f ($^2F_{7/2}$) radiative transitions of Ce^{3+} were observed at 428 nm and were similar to the data published elsewhere [8]. After annealing, the emission peak was red-shifted to 478 nm and a subsequent development of the satellite peak at 527 nm. These two emission peaks are associated with the crystal field splitting of the Ce^{3+}. Similar results were reported elsewhere [9], [10]. The red-shifted spectrum resembles the Ce^{3+} emission in a SrS host [11]. While the possible phases relating to SrS were not detected from the XRD data, the X-ray photoelectron spectroscopy data in ref [4] confirmed that SrS was formed. Similar observation was made from the sample prepared at 500°C before and after annealing while the main PL emission peak from the film prepared at 600°C substrate temperature was stable at 478 nm, suggesting that the SrS phase was already formed even before post-deposition annealing. Note that in addition to red-shifted emission, the PL intensity of all the films was increased considerably after annealing. Note that all the post-deposition annealed films were rougher with bigger particles than the unannealed films as shown in the AFM images of figure 3. It is therefore reasonable to attribute the increase in the PL intensity to rougher surface and improved particle sizes at high substrate temperature and post-fabrication annealing.

Depicted in figure 5 are the CL emission spectra exhibiting the well known two emission peaks of the Ce^{3+} for the $SrGa_2S_4$: Ce^{3+} films deposited at different parameters. At 400°C (fig. 5(a)), the film showed the same pattern as the CL emission of the powder phosphor reported in ref [2]. That is, the two well known emission peaks associated with 5d (T_{2g}) → 4f ($^2F_{5/2}$) and from 5d (T_{2g}) → 4f ($^2F_{7/2}$) radiative transitions of Ce^{3+} were observed at 443 and 585 nm. Again there was a red-shift of the emission peaks with increasing substrate temperature. The highest CL intensity was observed from the film deposited with 28 800 pulses compared to that deposited at 57 600 pulses as shown in figure 5 (b). The AFM data not shown confirmed that the film deposited at 28 800 pulses was rougher that that deposited at 57 600, resulting in relatively higher PL intensity. The film deposited in Ar atmosphere at 600°C was more intense than that deposited in O_2 atmosphere as shown in figure 5 (c). The inset of

figure 3(c) is the stand alone spectrum of the film deposited in O_2 showing that the emission spectrum resemble those of the film deposited at a 400°C (figure 1(a)) and different number of pulses (figure 2).

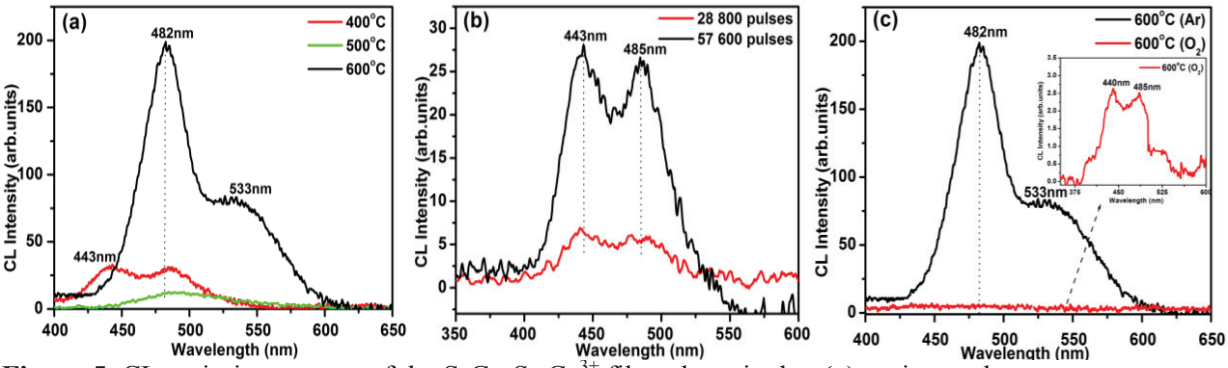


Figure 5. CL emission spectra of the $SrGa_2S_4:Ce^{3+}$ films deposited at **(a)** various substrate temperature, **(b)** different number of pulses and **(c)** in Ar and O_2 atmosphere.

4. Conclusion

$SrGa_2S_4:Ce^{3+}$ films were prepared successfully for the first time by the PRCLA technique. The growth temperature of 400°C showed a highly crystalline $SrGa_2S_4$ layer and also CL emission pattern similar to the one of the $SrGa_2S_4:Ce^{3+}$ powder was observed at this substrate temperature. At higher substrate temperatures, a shift to SrS emission occurred for both PL and CL emission peaks. Non-uniformity in particles (big and small) of the films and smooth surface were observed from SEM images. The AFM images showed smooth surface before annealing at low substrate temperature and at high substrate temperature the films surface became rougher as well as after annealing in vacuum at 700°C temperature resulting in an enhanced PL intensity.

References

[1] Nakanishi Y, Uekura N, Nakano F and Hatanaka Y 1997 *J. of Lumin.* **72-74** 373-374
[2] Gambarov E F and Bayramov A, Academy of Sci. of Azerbaijan, Institute of Physics.
[3] Bènalloul P, Barthou C and Benoit J 1998 *J. Alloy Compd.* **275-277** 709-715
[4] Moleme PA, Swart HC and Ntwaeaborwa O M 2011 MSc Dissertation, University of the Free State, South Africa
[5] Heikkinen H, Johansson L-S, Nykänen E and Niinistö L 1998 *Appl. Surface Sci.* **133** 205-212
[6] Tanaka K, Inoue Y, Okamoto S and Kobayashi K 1995 *J. of Crystal Growth* **150** 1211-1214
[7] Yang T, Wagner B, Chaichimansour M, Park W, Wang Z, and Summers C 1996 *J. Vac. Sci. Technol.* B **14** 3
[8] [9] Mhlongo G H, Ntwaeaborwa O M, Dhlamini M S and Hillie K T 2011 *J. Alloy Compd.* **509** 2986 - 2992
[9] [11] Chunxiang X, Zhidong L, Zheng X and Xurong X 1998 *Chinese Sci. Bull.* **43** 7 603-605
[10] Chunxiang X, Zheng X, Xurong X and Zhidong L 1999 *Solid State Comm.* **109** 183-187
[11] Kumar V, Pitale S S, Biggs M-M, Nagpure I M, Ntwaeaborwa O M and Swart H C 2010 *Mater. Lett.* **64** 752 - 754

XRD, DRS and SEM studies of the effects of metal dopants (Pt and Au) on the structural and optical properties of TiO_2

JR Moloantoa[1], KE Rammutla[1], TE Mosuang[1], RM Erasmus[2] and KT Hillie[3]
[1]Department of Physics, University of Limpopo, SOVENGA, 0712, RSA
[2]School of Physics, University of the Witwatersrand, JOHANNESBURG, 2050, RSA
[3]Centre for Nanostructured Materials, CSIR, PRETORIA, 0001, RSA

E-mail: Erasmus.rammutla@ul.ac.za

Abstract. Au and Pt doped TiO_2 nanocrystals were prepared using the sol gel method. Diffuse Reflectance Spectroscopy (DRS) was used to study the characteristics of these precious metals and X-ray diffraction (XRD) with calculated lattice parameters was used to study the location of these precious metal dopants in nanocrystalline TiO_2. The effects of these dopants on the phase transformation and grain growth were observed. Electron Microscopy was used to study the morphology. The results of the lattice parameter calculations from the XRD patterns of metal loaded TiO_2 are almost the same as for the undoped.

1. Introduction

Titania is a cheap and nontoxic polymorphic material of current interest for a variety of technological applications like gas sensing [1] and photovoltaic cells [2]. Generally, TiO_2, with a band gap of 3.2 eV, can only be excited by a small UV fraction of solar light, which accounts for only 3 - 5 % of the solar energy. Various strategies have been pursued including doping with metallic elements (e.g. Fe) [3] or nonmetallic elements (e.g. N) [4] with the aim of shifting the absorption into the visible range. Their potential applications depend strongly on their crystalline structure, morphology and particle size [5]. The main polymorphs of titania are rutile and anatase. It is now well established that anatase phase transforms into rutile at high temperatures [6]. Although many attempts have been made to control the process of phase transformation in titania, the transformation occurs over a range of temperatures due to the complexity of the crystal growth during the sample calcining process. In the present contribution, Pt and Au were incorporated into TiO_2 to study their effects on the structural and optical properties.

2. Experiments

Appropriate amounts of $HAuCl_4.3H_2O$ and $PtCl_4$ solutions were added to titanium iso-propoxide solution and stirred for 60 minutes on the magnetic stirrer. Then, appropriate amount of water was added, the precipitates filtered and dried overnight at room temperature. The solid was then dried at 80 °C, ground and portions calcined at 300 °C, 600 °C and 900 °C. Samples were prepared to give 5 wt% dopants.

X-ray powder diffraction patterns of the samples were recorded using a conventional laboratory diffractometer, a Phillips PW1720, using a Cu K_α tube operating at 35 KV and 20 mA. The average

grain size of the samples, L, was determined by the broadening of peaks on the XRPD pattern, using the Scherrer equation.

The Scanning Electron Microscopy (SEM) measurements were captured from the JEOL-JSM 7500F Field Emission Scanning Electron Microscope.

DRS measurements in the spectral range of 200-800 nm were recorded at room temperature using a Varian Cary 300 spectrometer equipped with a standard diffuse reflector unit.

3. Results and Discussions

3.1. XRD

The structure of all the metal incorporated TiO_2 samples that were synthesized by standard sol-gel methods appeared to be of anatase phase. Post-synthesis calcination at various temperatures resulted in the sharpness of peaks which clearly indicates the improvement in the degree of crystallinity corresponding to the formation of larger particles. The anatase to rutile (A –R) transformation of pure TiO_2 normally occur around 700 ^{0}C [6]. In the case of samples doped with some metals, the phase transformation temperature is altered. In figure 1 we show XRD patterns for Pt-TiO_2. It is clearly seen that the A-R phase transformation reduces significantly since at 600 ^{0}C, the phase transformation had already completely occurred. Au-TiO_2 behaves in a similar manner as the undoped sample.

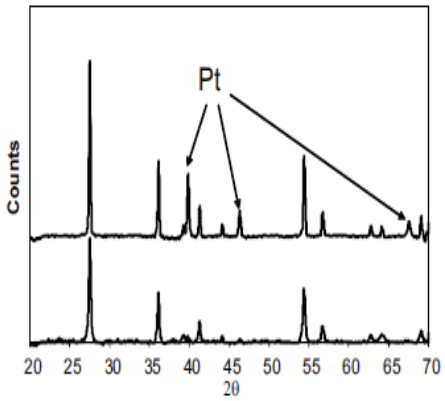


Figure 1. XRPD spectra of 5% Pt-doped TiO_2 annealed at 600 °C (bottom) and 900 °C (top).

Obvious diffraction peaks associated with metal ion dopants were observed indicating that they have not been fully integrated into the basic structure of TiO_2. The lattice parameters of metal loaded TiO_2 tabulated in table 1 were determined using the X-ray diffraction peaks of crystal planes [101] at 2θ=25° and [200] at 2θ=48° for anatase and also [101] at 2θ=36° and [200] at 2θ=39° for rutile.

Table1. Lattice parameters of pure and metal loaded TiO_2

Sample	**Calcination temperature**	**Crystal Phase**	**Lattice parameters**			
			a (Å)	**b (Å)**	**C (Å)**	**Unit cell Volume (Å^3)**
Undoped TiO_2	Untreated	Anatase [7]	3.788	3.788	9.53	136.76
		Rutile [8]	4.594	4.594	2.959	62.45
TiO_2 (5% Au)	600 °C	Anatase	3.76	3.76	9.535	134.80
TiO_2 (5% Pt)		Rutile	4.62	4.62	2.936	62.66
TiO_2 (5% Au)	900 °C	Rutile	4.60	4.60	2.962	62.67
TiO_2 (5% Pt)		Rutile	4.62	4.62	2.936	62.66

The lattice parameters of pure and doped TiO_2 are almost the same, indicating that the metal dopants are merely placed on the surface of the TiO_2.

3.2. DRS

The spectrum of undoped TiO_2 (not shown) consists of a single intense absorption below approximately 400 nm and this is usually ascribed to a charge-transfer process from the valence band to the conduction band. Beyond 400 nm it is almost totally reflective indicating that it cannot absorb the visible region of the spectrum. To shift the absorption capability, it is normally doped with some metallic dopants [3,4]. In the present work Pt and Au were used as dopants. In figure 2, we show DRS spectra of Au and Pt doped TiO_2 both at 600 and 900 ^{0}C. The spectrum of Au doped sample at 600 ^{0}C is typical of an anatase structure whereas the one at 900 ^{0}C is associated with the rutile phase [9]. For Pt doped sample, both spectra (i.e. at 600 and 900 ^{0}C) follow a similar trend and are associated with the rutile structure. The results are consistent with the XRD.

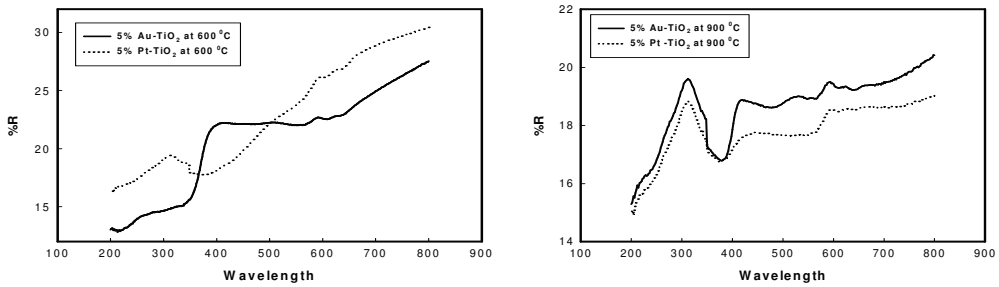


Figure 2. DRS spectra of Au and Pt-TiO_2 both at 600 ^{0}C and 900 ^{0}C.

The DRS spectra of the metal loaded TiO_2 reveal substantial reduction in reflection (i.e. they show enhanced absorption) from 400 nm to the entire visible region. The observed enhanced absorption in metal doped systems could be due to the generation of new energy levels resulting from the injection of impurities (i.e. metal clusters) within the band gap and/or the generation of oxygen vacancies.

3.3. *SEM*

In figs. 3(a) and (b) we show the typical SEM images of Pt-TiO_2 samples at 300 ^{0}C and 900 ^{0}C. The images clearly show that the morphologies of two samples differ significantly.

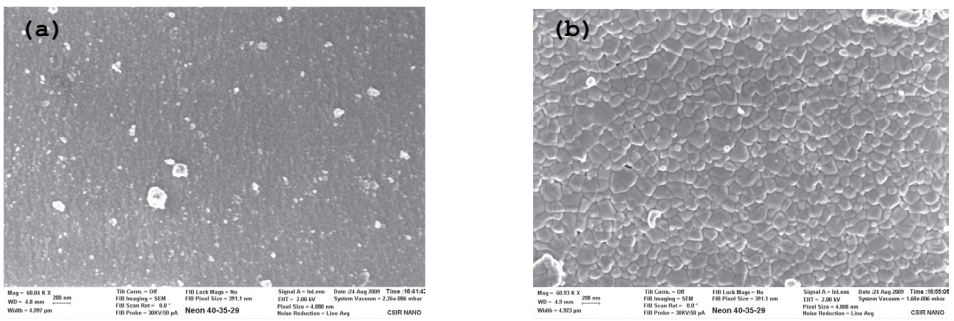


Figure 3. SEM images of Pt-TiO_2 samples at (a) 300 ^{0}C and (b) 900 ^{0}C.

The lower temperature Pt–TiO_2 composites revealed a relatively compact surface morphology with low porosity. Annealing of the TiO_2 sample at 900 ^{0}C produced more homogeneous spherical particles, which aggregated to form bigger particles. Similar results were obtained for Au loaded TiO_2. In general, the characteristic particle sizes become larger at higher calcination temperatures with a corresponding decrease in surface area. The elemental composition of the metal loaded samples was estimated by EDS. The EDS spectra (not shown) showed signals associated with Au and Pt metals.

4. Conclusions

From the XRD results, we calculated the lattice parameters of the metal loaded samples and found that they are almost similar to the undoped sample. The EDS spectra also showed signals associated with Au and Pt metals. We can therefore conclude that the metals have not entered into the lattice but are merely placed on the surface of the crystal. Platinum significantly affects the anatase to rutile transformation temperature whereas the presence of gold does not affect the phase transformation temperature and the results are corroborated by DRS.

5. Acknowledgements

RJM would like to acknowledge the financial support of NRF and University of Limpopo and also Tom Malwela of CSIR Centre for Nanostructured Materials for helping with the SEM measurements.

6. References

[1] Kim D, Rothschild A, Yang DJ and Tuller HL 2008 *Sensor Actuat. B - Chem*. **130**, 9
[2] Reyes-Coronado D, Rodrıguez-Gattorno G, Espinosa-Pesqueira M.E, Cab C, de Coss R and Oskam 2008 *Nanotechnology* **19**, 145605
[3] Zhang X W and Lei L C 2008 *Mater. Lett.* ***62***, 895
[4] Asahi R, Morikawa T, Ohwaki T, Aoki K and Taga Y 2001 *Science* ***293***, 269
[5] Wang W, Gu B, Liang L, Hamilton W A, Wesolowski DJ 2004 *J. Phys. Chem.* B **108**, 39
[6] Ma W, Lu Z and Zhang M 1998 *Appl. Phys.* **A66**, 621
[7] Guillard C, Disdier J, Monnet C, Dussaud J, Malato S, Blanco J, Maldonado MI and Herrmann JM 2003 *Appl. Catal. B* **46**, 319
[8] Reufer M, Feldmann J, Rudati P, Ruhl A, Muller D, Meerholz K, Karnutsch C, Gerken M, and Lemmer U 2005 *Appl. Phys. Lett.* **86**, 221102
[9] Nolan N, Seery M, Hinder S, Healy L and Pillai S 2010 *J. Phys. Chem.* **C114**, 13026

Ab initio structural and electronic study of metals on diamond (111)-(1×1) surface

I. Motochi[1,2], N. W. Makau[1], G. O. Amolo[1], B. A. Mathe[2], S. R. Naidoo[2]

[1]Computational MSG, Physics department, Moi University, P.O Box 1125-30100 Eldoret, Kenya.
[2]DST/ NRF CoE-SM and School of physics, University of the Witwatersrand, Wits 2050, Johannesburg, South Africa

E-mail: Isaac.Motochi@students.wits.ac.za

Abstract. Metal contacts on clean diamond surfaces have been studied using *ab initio* Density Functional Theory. Metal contacts modeled from five metals, i.e., gold, titanium, tantalum, vanadium and palladium on these surface were studied. Gold and palladium were found to form weak bonds on clean diamond (111) surface compared to the other three metals. The clean diamond surface was found to have surface states which were modified by oxygen but removed by hydrogen. Density of states revealed that on adsorption, these metals removed the band gap on clean diamond surface. A peak that was characteristic of diamond was observed at ≈ -11.8 eV. For the clean diamond surface terminated with Au and Pd, states due to Au-2p orbitals were observed at -2.5 eV and ≈ -1.5 eV, while Pd-5s orbitals were located at ≈ -1.7 eV and -0.5 eV. Ti, Ta and V showed unique states at high binding energies; ≈ -38 eV for V, ≈ -34 eV for Ta and ≈ -32.5 eV for Ti, which may explain their strong bonding.

1. Introduction

Semiconducting diamond has found many applications in the electronics industry. The fabricated diamond electronic devices often connects to metallic conductors in an electric circuit. Hewett *et al* [1] argue that contacts formed on diamond, must as a primary requirement, have a very low contact resistance, they must also be strongly adhesive, be able to withstand the harsh environments for which diamond devices are intended and they should also be compatible with conventional device processing techniques.

Techniques have already been developed to measure ohmic contact characteristics on naturally occurring p-type diamond (type IIb) when adsorbed with some carbide forming metals [2]. Empirically, circular transmission line method (c-TLM) of Reeves [3] together with an end resistance measurement has been used to study the properties of Au, Au/Ti and Al/Ti on hydrogenated CVD diamond [4].

This work provides a theoretical treatment to what is observed experimentally [2] with regard to the formation of ohmic contacts. In doing this, the study examines the atomic monolayer of carbide forming metals; Ti, Ta and V on diamond, and two non-carbide forming metal, Au and Pd.

2. Computational procedure

Calculations in this work were performed using the *Quantum-ESPRESSO* computer code [5, 6]. It is based on Density Functional Theory (DFT) employing plane waves and pseudopotentials. Conjugate gradient electronic minimization was used in energy minimization, done by moving atoms in order to reduce the net forces (the gradients of potential energy) on the atoms to negligible values. Within DFT the electronic ground state is given by the minimum of the energy functional and is usually obtained by solving the Kohn-Sham (KS) equations. Electron-ion interactions were described using ultrasoft pseudopotentials. Norm-conserving pseudopotentials are derived from an atomic reference state, requiring that the pseudo- and all-electron valence eigenstates have the same energies and amplitude outside a chosen core cutoff radius r_c. Pseudopotentials with larger cutoff radius are said to be soft hence converge more rapidly. Ultrasoft pseudopotentials attain much smoother (softer) pseudo-wavefunctions, so use considerably fewer plane-waves for calculations of the same accuracy. The exchange and correlation energies were calculated with the Perdew-Burke-Ernzerhof (PBE) form of the generalized gradient approximation (GGA) [7, 8]. Geometry optimizations were performed by the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. The structures were relaxed to an accuracy better than 10^{-5} Ry in their total energies. The plane wave cutoff energy was thoroughly tested and an optimized value of 37 Ry was used for all calculations done in this work. Optimized **k**-points of 4×4×4 for bulk and 4×4×1 for surfaces were used. The cell parameters were fixed at the experimental value of the bulk crystal during geometry optimizations while the total energies were minimized with symmetries.

3. Results

Results of calculations for bulk diamond, which were in good agreement with experiment are summarized in Table 1.

Table 1. Calculated DFT-GGA parameters of bulk diamond together with their experimental values and their % deviation with respect to the corresponding experimental values.

Parameter	Calculated Value	Expt. value [Ref]	% deviation
Lattice constant,a_o, (Å)	3.568	3.567 [9]	0.03
C-C atom distance (Å)	1.540	1.540 [9]	0.00
Bulk Modulus (Mbar)	4.220	4.420 [9]	-4.54
Cohesive energy (eV/atom)	-9.10	-7.37 [10]	23.47

The study considered the energetics of relaxed geometries of the C(111)-(1×1) surface with and without the adsorbed atoms. The ground state energies of free H, O, Au, Ti, Ta, Pd and V atoms were calculated since they are required in determination of cohesive energies of these elements on diamond.

3.1. Structural Changes of Relaxed Diamond (111)-(1×1) Surface

A five bi-layer C-atom slab was taken to represent the diamond surface because it makes the calculations less expensive computationally and was found to be representative of the diamond's bulk and surface properties. Following system relaxation, bond length changes and bond angles for the surface, near surface and bulk regions were obtained using the Xcrysden program [11].

Alternate contraction and expansion of bonds in the surface region in the clean surface as seen in both Table 2 and Figure 1 suggests a bulk symmetry breakdown. This however seems to be stabilized by hydrogen and partly by oxygen termination (see Figure 2).

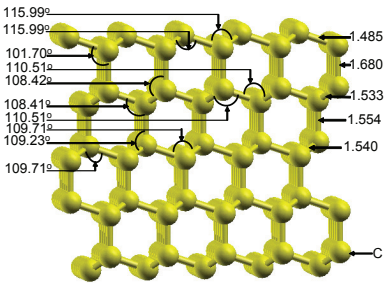


Figure 1. Relaxed clean diamond (111)-(1×1) surface. The bond lengths are in Angstroms.

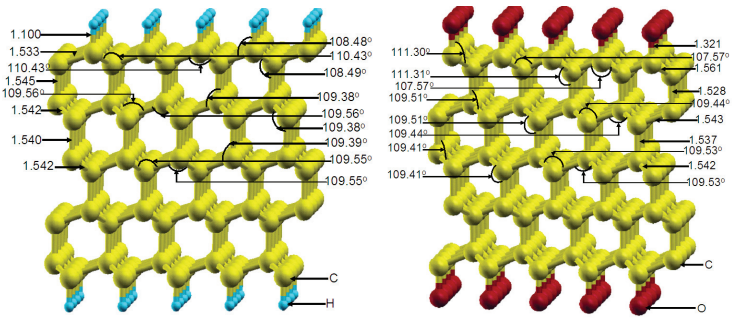


Figure 2. Relaxed hydrogenated and oxygenated diamond (111) surfaces. Hydrogen and oxygen atoms occupy a full ML coverage at the on-top site.

Table 2. Bond lengths and their % changes in relaxed; clean (1×1), hydrogenated and oxygenated diamond surface. $r_{(C-H)}$ $r_{(C-O)}$ are the C-H and C-O bond lengths while $r_{1(C-C)}$, $r_{2(C-C)}$ ... are the C-C bond lengths between the first and second, second and third atomic layers, etc. The % bond length changes (in brackets) are relative to the experimental C-C bond length of 1.540 Å for C-C, 1.100 Å for C-H and 1.360 Å for C-O.

Bond	Relaxed bond length (Å) Clean surface	Relaxed bond length (Å) Hydrogenated surface	Relaxed bond length (Å) Oxygenated surface
$r_{(C-H)}$		1.100(0.00)	
$r_{(C-O)}$			1.321 (-2.87)
$r_{1(C-C)}$	1.485 (-3.57)	1.533 (-0.45)	1.561 (+1.36)
$r_{2(C-C)}$	1.680 (+9.09)	1.545 (+0.32)	1.528 (-0.78)
$r_{3(C-C)}$	1.533 (-0.45)	1.542 (+0.13)	1.543 (+0.19)
$r_{4(C-C)}$	1.554 (+0.91)	1.540 (0.00)	1.537 (-0.19)
$r_{5(C-C)}$	1.540 (0.00)	1.545 (+0.13)	1.542 (+0.13)

3.2. Structural Properties of Diamond-Metal Interfaces.

3.2.1. The C-metal bond lengths. The surfaces were modeled ensuring that each had full monolayer termination of adsorbates (see Figure 3), to avoid any influence of dangling bonds which may otherwise affect the electrical conductivity of the surfaces as well as maintain the bulk-like structure. This keeps the surface from reconstructing as is the case with a clean C (111)-(2x1) [12]. Unavailable bond length values were approximated by the sum of the covalent radii of the atoms forming the bond, using equation (1).

$$r_{cov} = r_a + r_b, \tag{1}$$

where r_{cov} is the interatomic distance of a molecule made up of elements a and b whose covalent radii are r_a and r_b, respectively.

The systems were then allowed to relax, sample structures for the (1×1) surface terminated with a carbide/non-carbide forming metal, (Ta)/Au) are shown in Figure 3 and results recorded in Table 3.

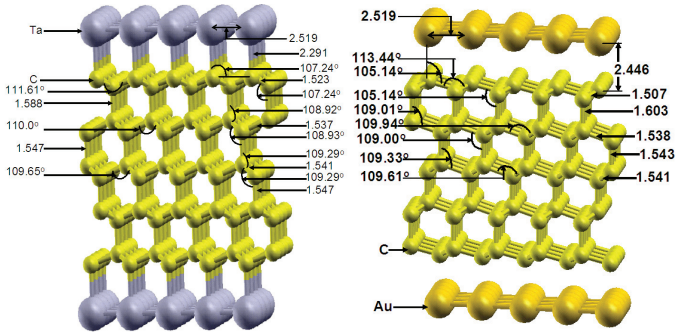


Figure 3. Tantalum and gold on clean diamond (111) surface showing metallization of diamond surface. The bond lengths are in Angstroms.

Table 3. Calculated bond lengths (between bolded atoms) when ohmic contacts form on clean, hydrogenated and oxygenated diamond (111)(1×1) surfaces. The % change in bond lengths are given with respect to the approximated bond lengths, since no experimental or theoretical data was available.

Adsorbing species	Approximated bond length (Å)	Relaxed Bond length (Å)	% change in bond length
C-Au	2.21	2.45	10.86
C-Pd	2.08	2.24	7.69
C-Ta	2.15	2.29	6.51
C-Ti	2.13	2.31	8.45
C-V	2.02	2.21	9.41

C-Metal bonds for all metal atoms showed significant bond elongation with respect to modeled values, the largest increase of ≈ +11 % occurring in the C-Au and the shortest of 6.5 % in C-Ta. This was, however, not unexpected since gold is basically inert, and hence less reaction was expected at the interface.

In surfaces, cohesive energy is the energy required to remove an adsorbed atom. Cohesive energy therefore gives an indication of whether the adsorbate is chemisorbed or physisorbed. The cohesive energies for adsorbates on clean diamond surface were obtained using equation (2) and results shown in Figure 4.

$$E_{coh}(x) = \frac{1}{N}\Big\{E_{slab}(x) - E_{clean-surf} - N[E_{atom}(x)]\Big\} \times 13.6, \tag{2}$$

where $E_{slab}(x)$ is the total energy of the clean diamond slab terminated with an atom of element x, $E_{clean-surf}$, is the total energy of the clean surface. $E_{atom}(x)$ is the total energy of a free

atom of element x and N is the number of element x atoms per surface unit cell; $N = 2$, since the calculations were for full monolayer coverages, x is the terminating species.

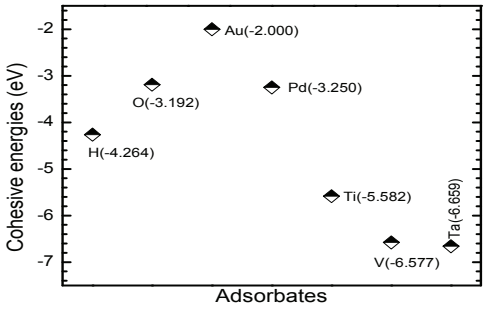


Figure 4. Calculated cohesive energies of the different ohmic contacts (metal adsorbates) on clean diamond surfaces. Cohesive energies of oxygen and hydrogen are also shown though they do not form ohmic contacts on diamond surface.

3.3. Electronic Properties

3.3.1. Density of States (DOS) for Clean Diamond Surface. It was necessary to determine the density of states of the clean surface first, so that the effect of the adsorbates could then be established. Figure 5 shows the DOS and PDOS for a clean diamond surface which are constituted of broad valence bands and a slightly narrow conduction band with some states at the Fermi level i.e., at 0 eV attributed to the presence of dangling bonds on the clean diamond surface. Olguín [13] argues that a surface introduced on an infinite periodic system, makes the system semi-infinite, with the boundary conditions dictating that the wave function has to be zero at the surface. The new boundary conditions manifest themselves in changes in the energy spectrum and in the occurrence of new states. The valence bands had a band width of $\approx$ 21.5 eV, a slight deviation from that observed by Zhen *et al* [14] of 21.71 eV. This value also compares well with 21.0$\pm$ 1.0 eV obtained by Himpsel *et al* [15]. The empty conduction bands which were narrower than the valence bands had a band width of $\approx$ 11.5 eV. A band gap of 4.56 eV was obtained, though it was much smaller than the bulk value of 5.5 eV [9]. Hafner *et al* [16], obtained a band gap of 4.25 eV using first principles approach. The deviation observed between our computed value and the experimental one was attributed to then fact that both local density approximation (LDA) and generalized gradient approximations (GGA) tend to underestimate the fundamental band gap [17].

A state that is characteristic of diamond was observed at slightly above -11.8 eV, a value not very far from -12.5 eV obtained by Derry *et al* [18] from DFT calculations, and very close to -12.0 eV obtained by Zhen *et al* [14] from first principles, and -13.2 eV obtained by Reinke *et al* [19] from their experiments. There was no evidence of any surface states effect extending into the bulk. This explains the preservation of diamond bulk structure as seen in Figure 1 where there were minimal changes in bulk bond length and bond angles, for deep lying layers.

3.3.2. Carbide Forming Metals on Clean Diamond Surface. The DOS for these metals (Ti, Ta and V) showed similar characteristics to each other (see Figure 6 for Ti on clean surface). These metals removed the energy band gap completely, hence rendering the diamond surface to exhibit a metallic character. It was also established that apart from the usual broad valence and narrow conduction bands on clean diamond surface, there were new states at high binding energies occurring at $\approx$ -37.5 eV for vanadium, $\approx$ -34 eV for tantalum (not shown here) and $\approx$

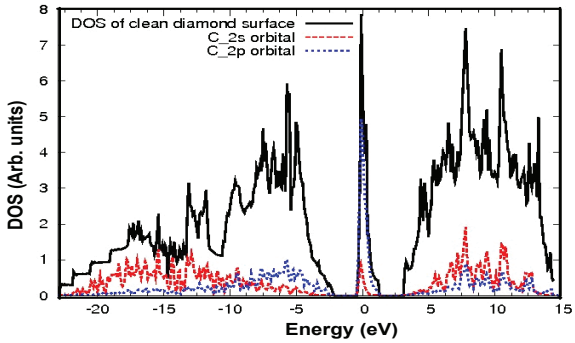


Figure 5. DOS for clean five bilayer C(111)-(1×1) surface.

-32 eV for titanium, respectively. In the case of titanium on clean diamond surface, these states were due to Ti-2p orbitals. The d orbitals were observed to have their effect mostly around the Fermi level, resulting in peaks at ≈ -0.4 eV and other peaks which manifest themselves as surface states just above the Fermi level. The states at high binding energies may explain the high cohesive energies observed for these metals. The conduction band was slightly wider as compared to that of the clean diamond surface thereby enhancing conductivity. The surface states observed on clean diamond surface were diminished, implying that the metals facilitate easy flow of charge carriers from the semi-conductor to the metal and vice versa.

3.3.3. Non-Carbide Forming Metals on Clean Diamond Surface. Unlike the carbide forming metals which had states at high binding energies, gold and palladium did not have these states. The absence of these states could explain the low cohesive energies exhibited by the two metals as shown in Figure 7 for Pd.

The valence bandwidth was ≈ 22 eV and a wider conduction band of ≈ 14 eV was observed. The fairly enlarged conduction bands explain why gold electronically forms good contacts on a diamond surface, though with poor adherence hence not suitable for application in harsh environments. Though the band gap is not present, there is a clear distinction between the valence and conduction bands which are joined by some diminished states where a band gap existed in the the unterminated surface. Just above the Fermi level in both cases of gold and palladium terminated clean surface there are low lying states dominated by Au 2p and Pd 5s orbitals (see Figure 7).

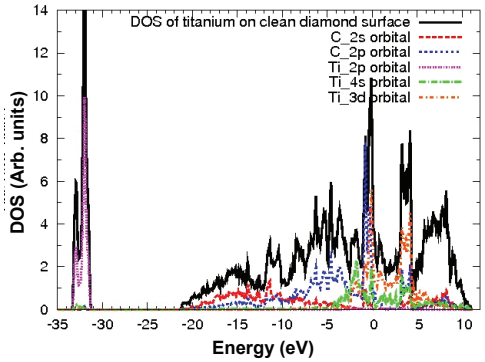


Figure 6. DOS for titanium on clean diamond surface.

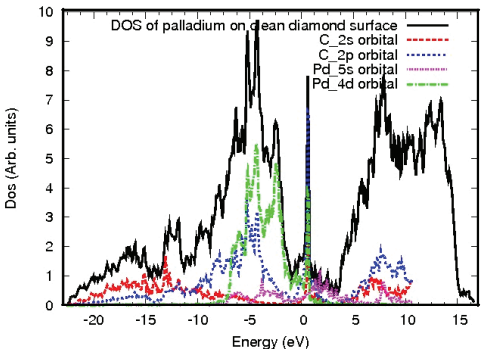


Figure 7. DOS for Palladium on clean diamond surface.

4. Conclusions

The study has revealed a clear correlation between C-metal bond lengths and their respective cohesive energies. A shorter bond length between atoms resulted in stronger bonds, hence a correspondingly high cohesive energy e.g. the C-Ta case. The states responsible for the strong bonding in the case of Ti, Ta and V (at high binding energies) were absent in the case of Au and Pd. In spite of this, they still had expanded conduction bands just like the case of Ti, Ta and V. These characteristics fit in well with diamond electronic devices that are often used or are envisaged to be used under harsh environments.

Like gold, palladium also showed a weak bond on diamond surface possibly due to its chemistry of having filled d orbitals, hence can not be good contacts under harsh conditions. The known carbide forming metals on the other hand had strong bonds formed between them as confirmed by their high cohesive energies.

According to this study, Tantalum, vanadium and titanium are therefore proposed for use as ohmic contacts in diamond-based electronics due to their strong adherence properties.

References

[1] Hewett C A, Taylor M J, Zeidler J R and Geis M W 1995 *J. Appl. Phys.* **77** 755-760
[2] Tachibana T, Williams B E and Glass J T 1992 *Phys. Rev. B.* **45** 11968
[3] Revees G K 1980 *Solid State Electron.* **23** 487
[4] Jackman R B, Looi H J, Pang L Y S,Whitefield M D and Foord J S 2000 *Diamond Relat. Mater.* **9** 975-981
[5] Giannozzi P, Baronim S, Bonini N, Calandra M, Car R, Cavazzoni C, Ceresoli D, Chiarottii G, Cococcion M and Dabo I 2009 *J. Phys.: Condens. Matter.* **21** 395502
[6] Scandolo S, Giannozzi P, Cavazzoni C, de Gironcoli S, Pasquarello A and Baroni S 2005 *Zeit. krist.* **220** 574-579
[7] Perdew J P,Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865-3868
[8] Perdew J and Wang Y 1992 *Phys. Rev. B.* **45** 13244
[9] Ristein J 2006 *Surf. Sci.* **600** 3677
[10] Kim E and Chen C F 2004 *Phys. Letts.* **A 326** 442-448
[11] Kolkaji A 2003 *Comput. Mater. Sci.* **28** 155
[12] Yan Y, Zhang S B and Al-Jassim M M 2002 *Phys. Rev. B.* **66** 201401
[13] Olguín D and Baquero R 2003 *Revista Mexicana de Física.* **49** 1-5
[14] Zhen C M, Wang Y Y, Guo Q F, Zhao M, He Z W and Guo Y. P 2002 *Diamond Relat. Mater.* **11** 1709-1712
[15] Himpsel F J, van der Veen J and Easterman D E 1980 *Phys. Rev. B.* **22** 1967
[16] Hafner J, Furthmüller J and Kresse G 1994 *Phys. Rev. B.* **50** 15606
[17] Tran F, Laskowski R, Blaha P and Schwarz K 2007 *Phys. Rev. B.* **75** 115131
[18] Derry T E, Makau N W and Stampfl C 2010 *J. Phys.: Matter.* **22** 265007
[19] Reinke P, Francz G and Oelhafen P 1996 *Phys. Rev. B.* **54** 7067-7073

Equivalent parameters for the empirical pseudopotential and *k·p* models

Kagiso Mpshe, Wilnelia E Adams, André E Botha and Enrico B Lombardi

Department of Physics, University of South Africa, P.O. Box 392, Pretoria 0003, South Africa.

E-mail: mpshek@unisa.ac.za

Abstract. After thirty-five years the empirical pseudopotential electronic band structures that were fitted by James R. Chelikowsky and Marvin L. Cohen (Phys. Rev. B 14, 556, 1976) are still widely used and regarded by some workers as being among the best. Although Chelikowsky and Cohen's paper has received more than 600 citations to date, not all subsequent workers were able to reproduce their band structures accurately. In a few of the citing papers, significant errors were introduced by, for example, incorrectly adding the effects of non-locality in the core potentials and spin-orbit coupling. In the original work both of these effects were added as perturbations in order to reduce the computational time. With present computers however, computational time is no longer an issue, and we have shown that the original calculations of Chelikowsky and Cohen can be reproduced accurately, without the use of perturbation theory. This reproduction requires refitting some of the parameters, but subsequently allows the original band structures to be used with great ease on any desktop computer. As an example, the original band structures are used as input to an optimization calculation which produces equivalent fourteen-band ***k·p*** parameters for GaAs. In future work we plan to compare the original band structures to those that can be obtained through state-of-the-art *ab initio* calculations.

1. Introduction

1.1. Empirical ***k·p*** model

Developed during the early days of semiconductor physics by E.O. Kane [1, 2], the ***k·p*** model has frequently been used to calculate the electronic band structure and electronic properties of bulk semiconductors and their associated layered semiconductor structures [3]. Because of its theoretical simplicity (using a relatively small number of bands) the ***k·p*** model has been used extensively in the past and has provided much of the theoretical insight responsible for the development of technology based on layered semiconductor devices during the 1980's and 1990's [3, 4, 5]. For example, in reference [3], Winkler and Rössler use the 8-band ***k·p*** model for a self-consistent calculation of spin-split electron sub-bands in a metal-oxide semiconductor (MOS) inversion layer on InSb. Botha and Singh [6] have also employed the 14-band ***k·p*** model to study the interaction between bulk and structural inversion asymmetry in type-II quantum wells.

1.2. Empirical pseudopotential model

Since its introduction by J. C. Phillips et al. [7–9], between 1958 and 1960, the empirical pseudopotential model (EPM) has become another important tool for both electronic band structure studies of solids and for understanding the behaviour of electrons in the crystal. In 1963 Cohen and Bergstresser [10] extended the model to include zincblende semiconductors. Bloom and Bergstresser [11] went on to include the effects of spin-orbit coupling in their full-zone band structure calculations for α-Sn, InSb and CdTe. Their work was based on and modified the model proposed by Weisz [12].

Based on the work by Bloom and Bergstresser, Walter et al. [13] calculated the band structures for ZnTe and ZnSe, including for the first time the spin-orbit interactions. In this work the imaginary part of the frequency-dependent dielectric function, the reflectivity and the logarithmic derivative of the reflectivity were calculated and fitted to the available experimental data.

The aforementioned studies were all based on a simplified "local" approximation of the core potentials. However, later it was demonstrated through XPS and UPS data that local empirical pseudopotential calculations yielded incorrect values for valence band widths. Chelikowsky and Cohen [14] conducted a study which included nonlocal corrections. They recalculated the band structure calculations of 11 diamond and zincblende semiconductors, which included all of the materials studied previously. The resulting band structures were in such excellent agreement with the experimental results that they are still widely used and published in reference books, such as the well-known Landolt-Börnstein tables.

1.3. Levenberg-Marquardt minimization algorithm

The Levenberg-Marquardt (LM) algorithm [15, 16] is an iterative technique that locates the minimum of a multivariate real-valued function, where the function is expressed as a sum of the squares of nonlinear real-valued functions. It can be thought of as a combination of the Steepest Decent and Gauss-Newton methods, and has become a standard technique for dealing with nonlinear least-squares problems.

2. Method for numerical calculations

2.1. Reproduction of the EPM band structures

Our calculations were based on the fitted empirical pseudopotential parameters obtained by Chelikowsky and Cohen [14]. In order to effect modifications to the original codes, the radial wave functions, $R_{n,l}(r)$, were first recalculated using the updated, freely-available Herman-Skillman code [17]. Once these wave functions were obtained for each material, the momentum dependence of the spin-orbit parameters and the nonlocality of the core potentials were included. Since processing time did not pose any problems, as it must have for Chelikowsky and Cohen in 1976, the use of perturbation theory could entirely be avoided. Thus it was found that the accuracy of our more modern implementation of the Chelikowsky and Cohen calculation matched or exceeded that of the original work by our use of a larger plane-wave basis, i.e., 274 reciprocal lattice vectors. Once the codes had been modified, the band structure of each material could be accurately reproduced (to better than 0.1 %) by using optimized parameters.

2.2. Extracting $k{\cdot}p$ parameters via Levenberg-Marquardt minimization

In order to compute the 14-band $\boldsymbol{k{\cdot}p}$ parameters, the Python programming language was used. The Scientific Python (SciPy) package, includes a minimisation routine (scipy.optimize.leastsq) which implements the Levenberg-Marquardt algorithm (see section 1.3). A Python code was developed, using leastsq, to fit the $\boldsymbol{k{\cdot}p}$ parameters. It produced the best agreement between the 14-band model and our reproduction of Chelikowsky and Cohen's [14] band structures.

It should be mentioned that the 14-band model is of course not capable of an accurate description of the bands over the full zone and, as such the fitting procedure described above only makes use of a small portion of the band structure near the zone centre.

3. Results and discussion

For brevity the results presented here will be in the form of a comparison between the band structure parameters obtained for only one of the materials (GaAs) which is common to this project and the work by Winkler [18]. The recalculated energy band structures from the empirical pseudopotential method [14] are shown in figure 1.

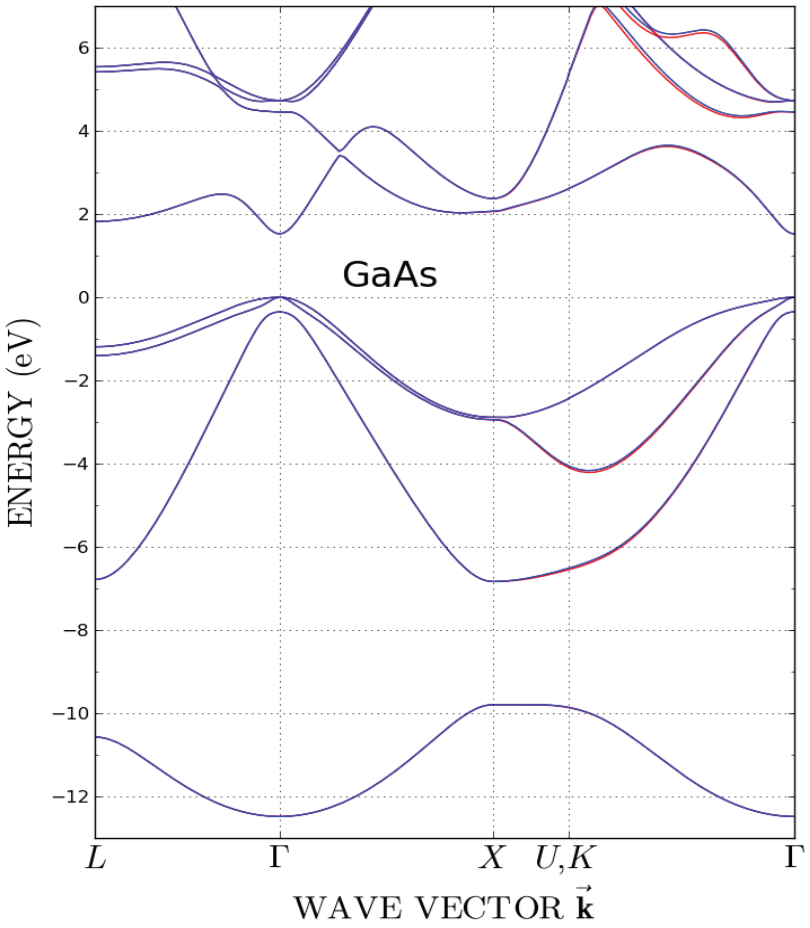


Figure 1. Recalculated energy band structures of GaAs using the model by Chelikowsky and Cohen [14]. In this calculation a 2×137 plane wave basis was used to obtain agreement to within 0.1%, without the use of perturbation theory and with only small adjustments to the non-local and spin-orbit parameters. As noted in the main text, the reproduction of the original band structures by Chelikowsky and Cohen [14] are made difficult by the fact that the original calculations used perturbation theory, most likely in an attempt to save computational time.

It is interesting to note that in order to save computational time, the original band structures were plotted by using fewer points along the path followed in ***k***-space. For this reason many of the sharper features (kinks) that can bee see in figure 1 are absent in the corresponding original figures of [14].

The pseudopotential parameters corresponding to the recalculated band structure of GaAs are listed in table 1. The adjusted parameters were also obtained via Levenberg-Marquardt minimization, using the original data points supplied by Chelikowsky and Cohen [14].

Table.1 Pseudopotential parameters for GaAs in Rydberg atomic units. The cation and anion d-radii (R_2^C and R_2^A) are listed in Angstrom. To facilitate the comparison we use the same units as Chelikowsky and Cohen [14]. For this material a Gaussian nonlocal correction was used. The lattice constant for GaAs is 5.65Å.

Form factors					
$V^S(\sqrt{3})$	$V^S(\sqrt{8})$	$V^S(\sqrt{11})$	$V^A(\sqrt{3})$	$V^A(\sqrt{4})$	$V^A(\sqrt{11})$
-0.2114	0.0139	0.0667	0.0554	0.0388	0.0000
Non-local parameters					
μ	α	A_2^C	A_2^A	$R_2{}^C$	$R_2{}^A$
0.0001	1.377	0.1859	0.6518	1.1626	1.1850

Figure 2 shows the optimised band structure for GaAs near the zone centre.

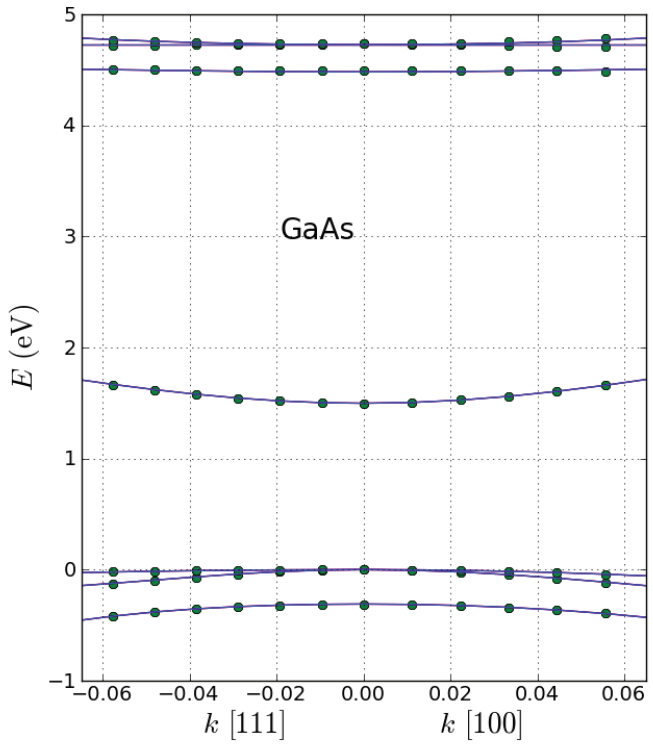


Figure 2. Energy band structure of GaAs near the zone centre. The green dots show the recalculated data points from the work by Chelikowsky and Cohen [14], while the solid blue lines is the bands obtained from the extended 14-band ***k·p*** model of Winkler [18], using the optimised (fitted) parameters. This fit was done by making use of a cubic mesh of 343 ***k***-points near the zone centre. The distance between the points was successively halved until the optimised parameters all converged to within 0.1% of their previous values.

The corresponding ***k·p*** parameters that were extracted via Levenberg-Marquardt minimisation are listed in table 2. The optimised parameters are also compared to those listed in Winkler [18].

Table 2. Comparison of the extended (14-band) ***k·p*** model parameters for GaAs.

Parameter	*Results from the literature [18]*	*Present work*
$E_0/[eV]$	*1.519*	*1.498*
$E_0'/[eV]$	*4.488*	*4.491*
$\Delta_0/\,[eV]$	*0.341*	*0.312*
$\Delta_0'/\,[eV]$	*0.171*	*0.241*
$P/\,[eV\,Å]$	*10.493*	*10.762*
$P'/\,[eV\,Å]$	*4.780i*	*6.210i*
$Q/\,[eV\,Å]$	*8.165*	*0.096*
$C_k/\,[eVÅ]$	*-0.0034*	*0.0000*
$m^*/[m^0]$	*0.0665*	*0.0661*
γ_1	*6.85*	*6.33*
γ_2	*2.10*	*1.68*
γ_3	*2.90*	*2.57*

Table 2 shows that the parameters P and P' are in relatively good agreement, while Q is not. The fact that $Q \approx 0$ shows the lack of non-sphericity in the pseudopotential of Chelikowsky and Cohen [14]. The other discrepancies seen here are for the Luttinger parameters. These discrepancies can be accounted for by the fact that the non-local corrections were not explicitly included in the extended ***k·p*** model. The effect of non-locality thus has to be 'absorbed' into the parameters γ_1, γ_2 and γ_3.

4. Summary and outlook

The aim of this work was to compare two important methods for computing electronic energy-band structure in semiconductor materials, namely the multiband $\boldsymbol{k\cdot p}$ method and the empirical pseudopotential method. The results obtained here were based on a modernised recalculation of the work by Chelikowsky and Cohen [14].

An outline of the underlying theory and the numerical basis of our calculations were provided, together with a brief overview of the existing literature on the application of the empirical pseudopotential and $\boldsymbol{k\cdot p}$ methods to important semiconductor materials.

As an example of the method the fourteen-band $\boldsymbol{k\cdot p}$ parameters were obtained for GaAs via a novel application of Levenberg-Marquardt minimisation for the following material. The results obtained from the minimisation were satisfactory since the calculated 14-band parameters provided an accurate description of the input energy-band structure near the zone centre. For the case of GaAs, for example, the conduction band effective mass ($m^* = 0.0661$) is found to be in excellent agreement with the experimentally determined value ($m_e^* = 0.0665$). In future work we plan to use the re-calculated band structures and the minimization method developed here to compare and possibly improve on the energy band structures that have been calculated through *ab initio* (density functional) methods [19].

Acknowledgements

This work is based upon research supported by the National Research Foundation (NRF) of South Africa. Any opinion, findings and conclusions or recommendations expressed in this material are those of the authors and therefore the NRF do not accept any liability in regard thereto.

References

[1] Kane E O 1957 *J. Phys. Chem. Solids* **1** 249
[2] Kane E O 1956 *J. Phys. Chem. Solids* **1/2** 82
[3] Winkler R and Rössler U 1993 *Phys. Rev. B* **48** 12
[4] Cardona M Christensen N E and Fasol G 1988 *Phys. Rev. B* **38** 3
[5] Wood D M and Zunger A 1996 *Phys. Rev. B* **53** 12
[6] Botha A E and Singh M R 2003 *Phys. Rev. B* **67** 195334
[7] Phillips J C 1958 *Phys. Rev.* **112** 685
[8] Phillips J C and Kleinman L 1959 *Phys. Rev.* **116** 287
[9] Kleinman L and Phillips J C 1960 *Phys. Rev.* **118** 1153
[10] Cohen M L and Bergstresser T K 1966 *Phys. Rev.* **141** 2
[11] Bloom S and Bergstresser T K 1968 *Solid State Commun.* **6** 465
[12] Weisz G 1966 *Phys. Rev.* **149** 504
[13] Walter J P, Cohen M L Petroff Y and Balkanski M 1970 *Phys. Rev. B* **1** 6
[14] Chelikowsky J R and Cohen M L 1976 *Phys. Rev. B* **14** 2
[15] Marquardt D 1963 SIAM *Journal of Applied Mathematics* **11** 431
[16] Levenberg K 1944 *Quarterly of Applied Mathematics* **2** 164
[17] An overview of the Herman-Skillman program is available at: http://hermes.phys.uwm.edu/projects/elecstruct/hermsk/HS.Overview.html
[18] Winkler R 2003 Spin-Orbit Coupling Effects in Two–Dimensional Electron and Hole System Springer-Verlag Berlin
[19] Tran F and Blaha P 2009 *Phys. Rev. Lett.* **102** 226401-1

Spin-density-wave behaviour of the $(Cr_{100-x}Al_x)_{95}Mo_5$ alloy system

B Muchono, A R E Prinsloo, C J Sheppard, H L Alberts and A M Strydom
Physics Department, University of Johannesburg, P.O. Box 524, Auckland Park, Johannesburg 2006, South Africa

E-mail address: alettap@uj.ac.za

Abstract. Electrical resistivity, Seebeck coefficient and specific heat measurements on a $(Cr_{100-x}Al_x)_{95}Mo_5$, $0 \leq x \leq 8.1$ at.% Al alloy system are reported. The results indicate two possible quantum critical points in the magnetic phase diagram of this system. One is an incommensurate spin-density-wave – paramagnetic quantum critical point situated at $x \approx 1.5$ at.% Al and the other a paramagnetic – commensurate spin-density-wave critical point at $x \approx 5$ at.% Al. We forward experimental evidence that this system harbours two spin-density-wave related quantum critical points which presents an unusually rich case study for magnetic quantum criticality of the itinerant kind.

1. Introduction

Chromium and its dilute alloys are exceptional examples of spin-density-wave type itinerant electron antiferromagnetic systems. The $Cr_{100-x}Al_x$ alloy system is of particular interest. It presents both incommensurate and commensurate spin-density-wave antiferromagnetic phases, as well as a paramagnetic phase in its magnetic phase diagram. There exists a triple point at $x_c \approx 2$ at.% Al, where these three phases coexist [1]. Previous studies [2] comprising electrical resistivity and magneto-elastic measurements on alloys of $Cr_{100-x}Al_x$ with Mo in ternary $(Cr_{100-x}\ Al_x)_{95}\ Mo_5$ suggest that spin-density-wave antiferromagnetism is suppressed down to at least 4 K in the concentration range $2.0 \leq x \leq 5.0$ at.% Al. The current interest in quantum criticality of Cr and its dilute alloys [3, 4] therefore warrants further detailed investigations on the $(Cr_{100-x}Al_x)_{95}Mo_5$ system, particularly on critical effects and the possibility of having simultaneously two types of spin-density-wave quantum critical points in the same Cr alloy system. Here we report results of the temperature dependence of the electrical resistivity as well as of the Seebeck and Sommerfeld coefficients of thermoelectricity and specific heat, respectively, for a more comprehensive range of alloys in the concentration range $0 \leq x \leq 8.1$ at.% Al.

2. Experimental

Ternary $(Cr_{100-x}Al_x)_{95}Mo_5$ alloys were prepared by arc melting in a purified argon atmosphere from 99.999 at.% pure Cr, 99.999 at.% pure Al and 99.99 at.% pure Mo. The alloys were annealed in an ultra-high pure argon atmosphere at 1300 K for three days and quenched in iced water. The actual elemental composition and homogeneity were checked by electron microprobe analyses. Electrical resistivity (ρ), Seebeck coefficient (S) and specific heat (C_p) were measured in the range $2 \leq T \leq 350$ K, using standard Quantum Design PPMS equipment incorporating appropriate measuring options [5].

3. Results and discussion

Figure 1 shows three typical examples of $\rho(T)$ curves, one each for the commensurate spin-density-wave, paramagnetic and incommensurate spin-density-wave alloys in the $(Cr_{100-x}Al_x)_{95}Mo_5$ system. For incommensurate spin-density-wave alloys ($x \lesssim 1.5$ at.% Al), $\rho(T)$ depicts a well defined anomaly in the form of a minimum near the Néel transition temperature (T_N), figure 1(a), which is typically also observed in other incommensurate spin-density-wave Cr alloys [1]. This is attributed to an induced spin-density-wave energy band gap at the Fermi energy upon cooling through T_N [1]. The $\rho(T)$ anomaly is absent for alloys in the concentration range between 1.5 and 5.3 at.% Al, indicating that alloys in this concentration range remain paramagnetic down to 2 K, as indicated in figure 1(b). The reappearance of antiferromagnetism, accompanied by a $\rho(T)$ magnetic anomaly of commensurate spin-density-wave origin is conspicuous in figure 1(c) for $x \gtrsim 5$ at.% Al [1, 2]. The $\rho(T)$ anomaly in the commensurate spin-density-wave phase is however much weaker than that in the incommensurate spin-density-wave phase, contrary to expectations for commensurate spin-density-wave Cr alloys [1].

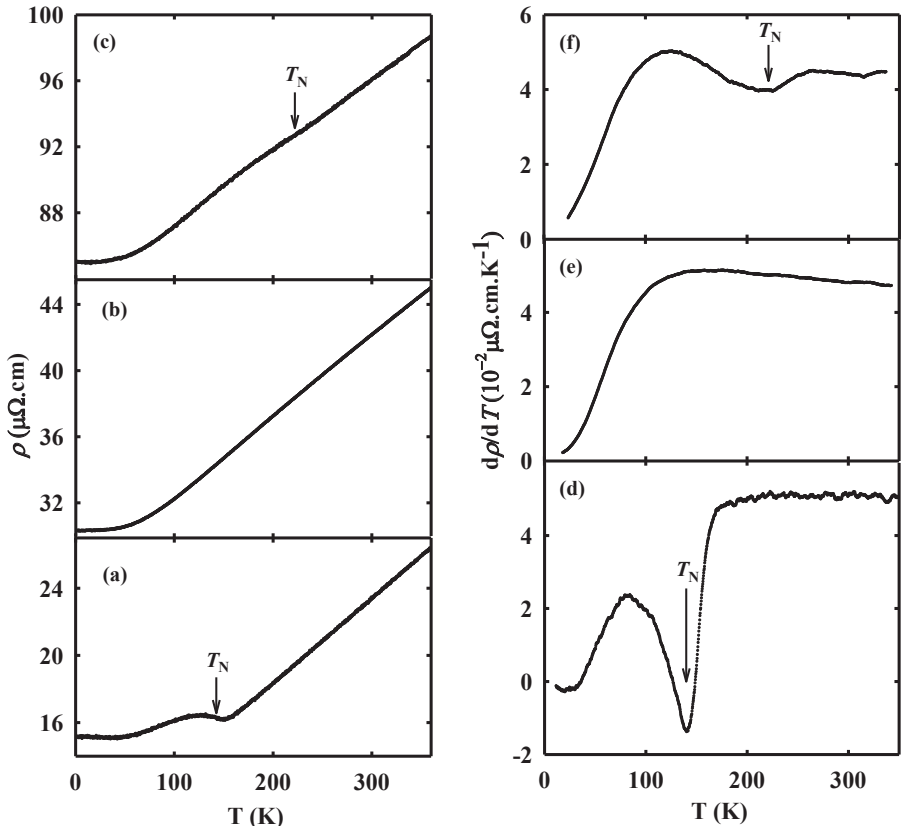


Figure 1: Typical examples of the temperature dependence of the electrical resistivity, $\rho(T)$, and its temperature derivative $(d\rho(T)/dT)$ of the $(Cr_{100-x}Al_x)_{95}Mo_5$ alloy system. Shown are an incommensurate spin-density-wave alloy with x = 1.0 [(a) and (d)], a paramagnetic alloy with x = 2.8 [(b) and (e)] and a commensurate spin-density-wave alloy with x = 6.1 [(c) and (f)] at.% Al. The Néel temperature (T_N), shown by arrows, is obtained from the $d\rho(T)/dT$ minimum. The experimental error in the absolute value of ρ amounts to ≈ 5% and originates mainly from errors in determining the sample dimensions, while changes in ρ of 0.5% or better could be detected as a function of temperature.

T_N is often defined for Cr and its dilute alloys as the temperature of the minimum in $d\rho(T)/dT$ accompanying the $\rho(T)$ magnetic anomaly [1] and this definition is appropriately also used for the present $(Cr_{100-x}Al_x)_{95}Mo_5$ system. Figures 1(d), (e) and (f) depict the temperature dependence of $d\rho(T)/dT$, obtained from figures 1(a), (b) and (c), respectively, with the position of T_N marked by an arrow. The spin-density-wave anomaly is clearly better defined in $d\rho(T)/dT$ than in $\rho(T)$ itself. T_N values thus obtained are plotted on the magnetic phase diagram displayed in figure 4(a).

It may be mentioned that T_N is in some instances obtained for Cr and its alloys by back extrapolation of the $\rho(T)$ curves, from temperatures high up in the paramagnetic phase down to 0 K [6]. This generates a paramagnetic base-line curve, should the alloy remains paramagnetic down to 0 K, from which the magnetic component, $\Delta\rho(T)$, of $\rho(T)$ can be extracted. T_N is then taken at the temperature where $\Delta\rho(T)$ tends to zero. This was tested on the incommensurate spin-density-wave x = 1.0 at.% Al alloy, giving results that compare to within 10 K with that obtained from $d\rho(T)/dT$. Applying this method to the commensurate spin-density-wave alloys is however problematic, as the present measurements do not extend to high enough temperatures above T_N for a reliable back extrapolation.

Seebeck coefficient, $S(T)$, measurements are useful and complementary [1] to $\rho(T)$ measurements for obtaining T_N of Cr alloys, particularly for those alloys showing weak spin-density-wave $\rho(T)$ anomalies near T_N. The reason is the fact that the carrier diffusion component of $S(T)$ depends on the energy derivative of the electrical conductivity at the Fermi energy, resulting in a much stronger $S(T)$ anomaly on spin-density-wave formation than that for $\rho(T)$ [1]. $S(T)$ measurements on the $(Cr_{100-x}Al_x)_{95}Mo_5$ alloys are therefore also reported here. Figure 2(a), (b) and (c) show typical examples for incommensurate spin-density-wave, paramagnetic and commensurate spin-density-wave alloys, indicating the higher sensitivity of $S(T)$ for the Néel transition.

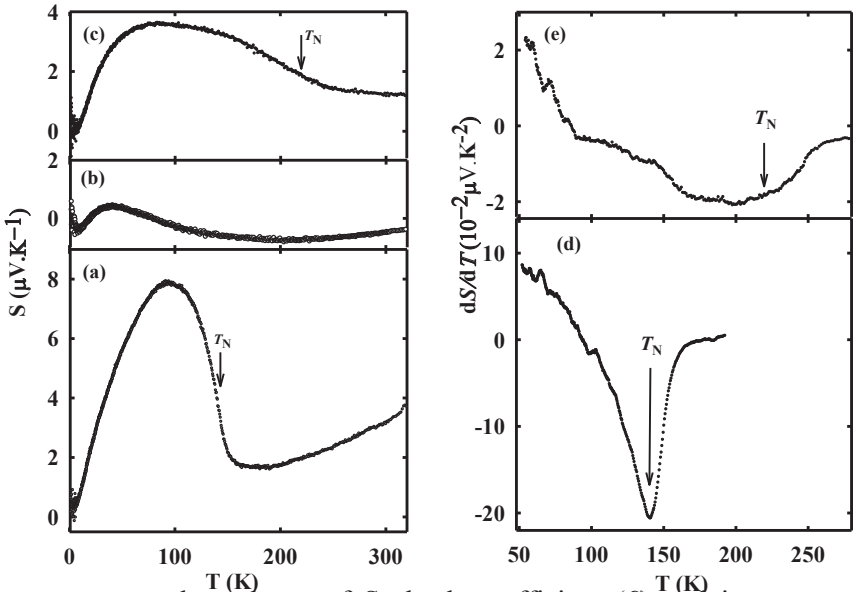


Figure 2: The temperature dependence of Seebeck coefficient (S) and its temperature derivative ($dS(T)/dT$) of the $(Cr_{100-x}Al_x)_{95}Mo_5$ alloy system. Shown are an incommensurate spin-density-wave alloy with x = 1.0 [(a) and (d)], a paramagnetic alloy with x = 2.8 (b) and a commensurate spin-density-wave alloy with x = 6.1 [(c) and (e)] at.% Al. The Néel temperature (T_N), shown by the arrows, is taken at the minimum point of $d\rho(T)/dT$. The standard deviation in the measurement of the thermal voltage in the calculation of S ($= dV/dT$) is typically less than 0.5%.

T_N values obtained from $d\rho(T)/dT$ are indicated by arrows in these figures. $dS(T)/dT$ curves for the two antiferromagnetic alloys of figures 2(a) and (c) are shown in figure 2(d) and (e), respectively, together with T_N obtained from $d\rho(T)/dT$. There is reasonably good correspondence between values of T_N obtained from the temperature derivatives of $\rho(T)$ and $S(T)$ respectively, which provides confidence in our method for obtaining T_N.

Figure 3 shows three typical examples, one each for incommensurate spin-density-wave, commensurate spin-density wave and paramagnetic alloys, of low temperature C_P/T vs. T^2 plots obtained from $C_P(T)$ measurements on the $(Cr_{100-x}Al_x)_{95}Mo_5$ alloys in the temperature range $2 \leq T \leq 60$ K. The curves are fitted rather well by the low-temperature approximation of the Debye formulation of specific heat, $C_p(T) = \gamma T + \beta T^3$, where γ represents the Sommerfeld electronic specific heat coefficient and the last term is representative of the lattice specific heat contribution. $\gamma(x)$, obtained from linear C_P/T vs. T^2 plots for the various alloys are shown in figure 4(b). The curve shows interesting behaviour that is related to influences of effects of spin-dinsity-wave formation on the electronic density of states at the Fermi energy and spin-fluctuation effects.

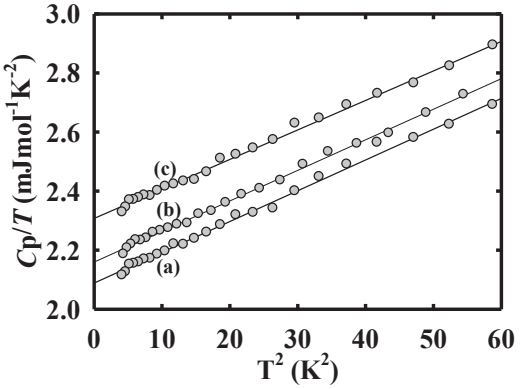


Figure 3: Graphs of $C_P(T)/T$ vs. T^2 plotted for three typical $(Cr_{100-x}Al_x)_{95}Mo_5$ alloy examples at low temperatures: (a) an incommensurate spin-density-wave alloy with x = 1.0, (b) a commensurate spin-density wave alloy with x = 6.1 and (c) a paramagnetic alloy with x = 2.8 at.% Al. The solid lines through the data points represent least-square linear fits. The experimental error in $C_P(T)/T$ is about 1%.

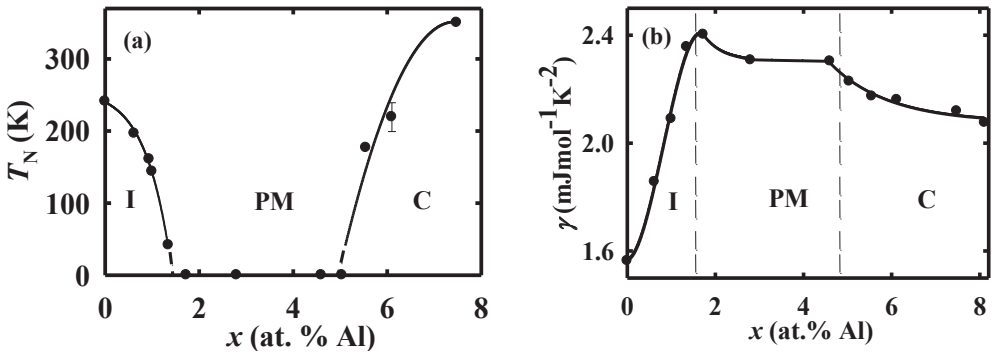


Figure 4: (a) The magnetic phase diagram and (b) the Sommerfeld coefficient, γ, as a function of Al concentration, x, for the $(Cr_{100-x}Al_x)_{95}Mo_5$ alloy system. I, PM and C denotes incommensurate spin-density-wave, paramagnetic and commensurate spin-density-wave phases, respectively. The experimental error in T_N is shown by an error bar while the experimental error in γ is within the size of the experimental points. The solid lines are guides to the eye.

Of interest is the sharp rise of $\gamma(x)$ in the incommensurate spin-density-wave phase that is followed by a slow decrease, becoming nearly flat, in the paramagnetic phase and eventually decreases relatively sharply on entering the commensurate phase. This behaviour is understood by considerations of Takeuchi's [7] application of Hasegawa's [8] spin fluctuation theory for γ of itinerant electron antiferromagnetic systems, to the case of $Cr_{100-y}V_y$ alloys. This alloy system depicts [4] an incommensurate spin-density-wave – paramagnetic quantum critical point on the magnetic phase diagram at y_c = 3.4 at.% V, a point up to which $\gamma(y)$ increases sharply, followed by a rather slow decrease in the paramagnetic phase [7]. It was shown that the slow decrease of $\gamma(y)$ in the paramagnetic phase, $y > y_c$, of this system is well explained by spin-fluctuation effects alone, while the sharp decrease at $y < y_c$ resulted from a combination of spin fluctuation and incommensurate spin-density-wave energy gap influences, the latter outweighing that of spin-fluctuations [7]. Following this reasoning one then expects $\gamma(x)$ of the $(Cr_{100-x}Al_x)_{95}Mo_5$ system also to follow this trend on increasing x through the two quantum critical points: a sharp rise up to a peak at the incommensurate spin-density-wave – paramagnetic quantum critical point, followed by a rather slower decrease in the paramagnetic phase, and the possibility of a second small peak, before $\gamma(x)$ decreases again on entering the commensurate spin-density-wave phase through the paramagnetic – commensurate spin-density-wave quantum critical point, as shown in figure 4(b). $\gamma(x)$, which is related to the electronic density of states at the Fermi energy, thus appears to be an excellent parameter for indications of quantum critical behaviour in this alloy system.

4. Conclusion

The present study substantially expands and corroborates earlier studies on the magnetic phase diagram of $(Cr_{100-x}Al_x)_{95}Mo_5$. The results, particular the behaviour of the concentration dependence of the Sommerfeld electronic specific heat coefficient, are in evidence of the presence of two quantum critical points on the magnetic phase diagram of the $(Cr_{100-x}Al_x)_{95}Mo_5$ alloy system. It is then rather exceptional to observe both incommensurate spin-density-wave – paramagnetic and paramagnetic – commensurate spin-density-wave quantum critical points in the same Cr alloy system. Hall coefficient and magnetic susceptibility are critical parameters [3, 4, 9] for exploring quantum criticality in Cr and its alloys and further investigations in this regard as well as exploring the possibility of a second peak in $\gamma(x)$ (figure 4 (b)) are underway for the present $(Cr_{100-x}Al_x)_{95}Mo_5$ alloys.

Acknowledgments

Financial support from the South African National Research Foundation (Grant numbers 61388, IFR2009021800008 and 2072956) is acknowledged.

References

[1] Fawcett E, Alberts H L, Galkin V Yu, Noakes D R and Yakhmi J V 1994 *Rev Mod. Phys*. **66** 25
[2] Smit P and Alberts H L 1993 *J. Phys.: Condens. Matter* **5** 6433
[3] Jaramillo A, Feng Y, Wang J, and Rosenbaum T F 2010 *Proc. Nat. Acad. of Sci*. USA **107** 16361
[4] Yeh A, Soh Y, Brooke J, Aeppli G and Rosenbaum T F 2002 *Nature* **419** 459
[5] Quantum Design, 6325 Lusk Boulevard, San Diego, CA 92121-3733, USA
[6] Alberts H L, Prinsloo A R E and Strydom A M 2010 J. *Magn. Magn. Mater*. **322** 1092
[7] Takeuchi J, Sasakura H and Masuda Y 1980 *J. Phys. Soc. Jpn*. **49** 508
[8] Hasegawa H 1975 *J. Phys. Soc. Jpn*. **38** 107
[9] Paschen S 2006 *Physica* B **378-380** 28

Statics and dynamics of single polymers confined between two corrugated walls

Debashish Mukherji

Max-Planck Institut für Polymerforschung, Ackermannweg 10, 55128 Mainz, Germany

E-mail: mukherji@mpip-mainz.mpg.de

Martin H. Müser

Jülich Supercomputer Center, Institute for Advanced Simulation, FZ Jülich, 52425 Jülich, Germany

Abstract. Using molecular dynamics simulations, we study static and dynamic properties of isolated linear polymers in good solvent conditions, which are confined between two parallel, corrugated walls. If the distance between the confining walls is so small that the accessible space for a monomer, normal to the surface, is comparable to the monomer size itself, the diffusion constant D is found to follow Rouse scaling, i.e., $D \approx N^{-1}$, with the degree of polymerization N. The proportionality constant is sensitive to wall-wall and wall-polymer commensurability. Upon further increasing the confinement (or decreasing the inter-wall distance) a stronger scaling is observed, where $D \approx N^{-3/2}$. Static properties, such as out-of-plane monomer density profile $n(z)$ and radius of gyration $R_{\rm g}$, obey the scaling laws predicted by Flory's mean field theory.

1. Introduction

Understanding the diffusion of polymers that are confined or constrained by solid walls is not only theoretically challenging [1, 2] but also relevant in various technological applications such as in tribology [3, 4], narrow-channel macromolecular devices [5, 6], and for the possibility to separate macromolecules of different size via confinement [7]. In this context two cases are widely studied: (a) adsorbed linear polymers [8, 9, 10, 11, 12, 13, 14, 15] and (b) polymers confined between two surfaces [16, 17, 18, 19, 20] Experimental advances in this field are made possible due to the application of fluorescence spectroscopy, which allows one to measure, for instance, the mobility of DNA [9, 18, 19, 20] and that of phospholipids [12] on lipid bilayers with the prospect of their use in nano-technology.

Static and dynamic properties of polymers are commonly rationalized in terms of scaling laws [21]. In this paper, we will present molecular dynamics (MD) simulations of polymers confined between two parallel, corrugated walls, and examine the effect of wall friction on the center of mass diffusion of the polymers in terms of scaling laws. To produce the correct polymer configurations, we will include the solvent implicitly in the (thermodynamic) interactions between different monomers, but neglect any (implicit) hydrodynamic interaction mediated by the solvent. The latter has to be done in order for us to single out the wall-induced damping. As our simulations will be carried out in equilibrium, and as we are interested in a linear response quantity, namely diffusion, the two damping mechanism can be considered to be additive.

Table 1. Different cases studied in this work. A relative orientation of 90° renders two hexagonal surfaces incommensurate. d is nearest-neighbor distance between surface atoms. $\bar{b}$ represents the intrinsic bond length for a bead spring polymer.

cases	wall orientation	d	b
case 1	incommensurate	1.201	0.97
case 2	commensurate	1.201	0.97
case 3	incommensurate	0.970	0.97
case 4	commensurate	0.970	0.97

For our study, we will employ the same generic model as in our previous works [13, 14, 15] on polymer diffusion on a single surface. These studies reproduced successfully a variety of experimental observations, among others: (a) The $D \propto N^{-1.5}$ relationship for polymers on a solid surface [10], (b) the $D \propto N^{-1}$ relationship for polymers on an on-overage flat liquid surface [9] and/or lattice models, and (c) the non-monotonic change of mobility of polymers with concentration [11]. An important aspect in our previous work, in particular that on the diffusion of single polymers, turned out to be the commensurability between polymer and surface. This is why bond lengths commensurate and incommensurate with the lattice constant will be considered here too. In addition, the commensurability between the walls may play a role, as even an ideal gas can pin two commensurate walls that do not interact directly with one another [4]. Therefore, we will investigate both, commensurate and incommensurate confining walls.

2. Model and method

As usual, we assume that chemical detail does not affect the values of critical exponents. This is why we will base our study on a simple, albeit widely used particle-based continuous-space representation of a polymer molecule developed by Kremer and Grest [22] also known as bead-spring model. A more detailed model related discussion is provided in references [13, 14, 15]. In this model, good solvent conditions can be mimicked by having each two monomers interact via a truncated and shifted (purely repulsive) Lennard-Jones (LJ) potential. Adjacent monomers also interact with an additional finitely extensible nonlinear elastic (FENE) potential. The parameters of the potential (i.e., LJ + FENE) is chosen such that no unphysical bond crossing is allowed [22]. In the following, we will express all physical quantities in units of Lennard-Jones (LJ) energy ε, LJ length σ, and mass m of individual monomers. This model gives rise an effective bond length of 0.97 [22].

Temperature is controlled using a Langevin thermostat, which is only applied to the z-component of the monomer's motion, i.e., normal to the interface. Unless stated otherwise, the thermal energy is set to 0.5. The damping coefficient is set to 0.1 in reduced units.

The bead-spring chain is confined between two surfaces, each of which consists of a discrete hexagonal plane, e.g., a (111) plane of a face-centered-cubic solid. Wall atoms are pinned to their hexagonal lattice sites. The geometry of the surface is square with fixed linear dimensions 75.0. In the case of the nearest neighbor distance 0.97, we increased the number of unit cells in the lateral dimension to attain same linear dimensions 75.0. Periodic boundary conditions are applied only along the lateral directions, i.e., in the (x, y)-plane. In each MD run, the separation δ between the two walls is kept fixed. The orientations of the two surfaces are usually parallel, however, upon extreme confinement, i.e., for $\delta \leq 7.5$, we also investigate effectively incommensurate interfaces. These are produced by rotating one of the two hexagonal walls by 90°. We studied four cases as listed in table 1.

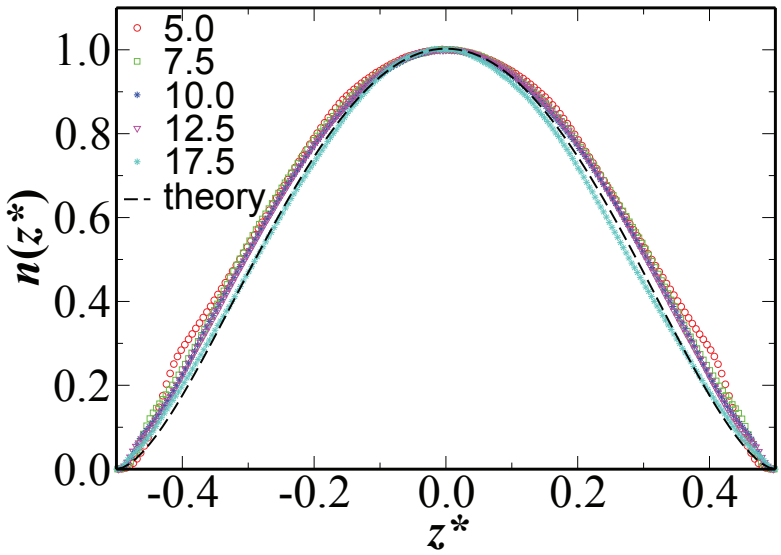


Figure 1. Out-of-plane monomer density profile, $n(z^*)$, for different inter-wall distances, δ. Here $z^* = \tilde{z}/\tilde{\delta}$ with $\tilde{\delta} = \delta - 1.5$ and $\tilde{z} = z - 1.5/2$. The dashed curve is a fit based on equation 1 using the exact value $\nu_3 = 0.587$ [21].

The interaction between a wall atom and a monomer is the same LJ potential as the one acting between two monomers. When investigating the dependence of an observable as a function of δ, we restrict the degree of polymerization to the two values $N = 160$ and $N = 220$. The distance between the walls is varied from $\delta = 2.5$ to $\delta = 26.5$. When keeping δ fixed to study chain length effects, we vary N within $40 \leq N \leq 310$.

The equations of motion are integrated using a fifth order predictor-corrector method with time step $\Delta t = 0.005$. Simulations are usually equilibrated for a few hundred million MD time steps and then observation is carried out over another $50 \cdot 10^6$ MD time steps, i.e., over a time span $2.5 \cdot 10^5$ (in LJ units). During this observation period, quantities of interest such as the the radius of gyration R_g, diffusion coefficient D and out-of-plane monomer density profile $n(z)$ are measured.

3. Model validation

We would like to start by demonstrating that the used model produces the proper polymer configurations. For this purpose, we will investigate the density profile as well as the in-plane and out-of-plane radius of gyration, which depend on δ via well-established power laws [21].

3.1. Monomer density profile

The monomer density $n(\Delta z)$ of a single polymer near a flat, impenetrable wall has been predicted by Flory in a mean-field theory to follow a $n(\Delta z) \propto \Delta z^{1/\nu_3}$ power law, where ν_d is given as, $\nu_d = 3/(d+2)$, so that $\nu_3 = 0.6$. Here Δz denotes the distance of a monomer from the wall. The value for the exact exponent, namely $\nu_3 = 0.587$ [21], is fairly close to the mean-field value.

When two walls are present the following dependence of the density on the z-coordinate is observed [21]:

$$n(z^*) \propto \left(\frac{1}{4} - z^{*2}\right)^{1/\nu_3}, \qquad (1)$$

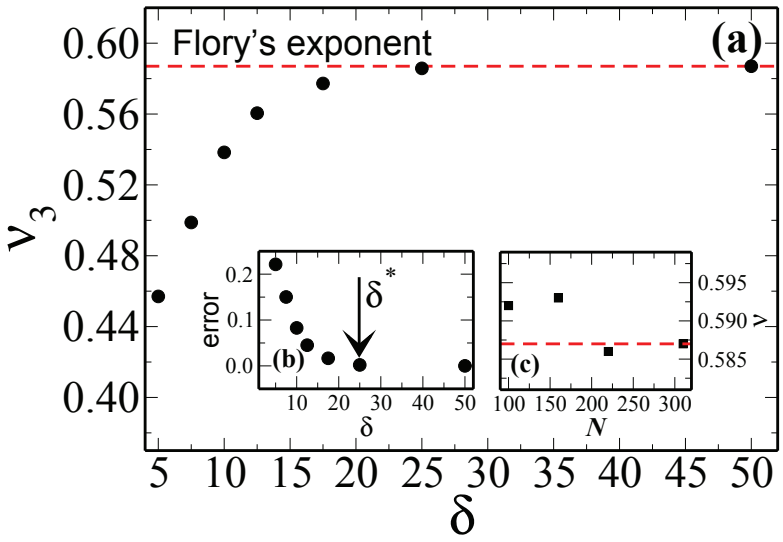


Figure 2. (a) The (apparent) exponent ν as obtained for different inter-wall distances δ at $N = 220$. ν_3 is calculated by fitting $n(z)$ to the power-law z^{1/ν_3} at small values of z/δ. (b) Absolute, relative difference between apparent and exact exponent ν_3, $(0.587 - \nu_3)/0.587$, as a function of δ. The arrow indicates the value for δ that satisfies equation (2). (c) (Apparent) exponent ν_3 for different chain lengths calculated for $\delta = \delta^*$. The exact exponent is represented by the horizontal dashed line.

where we have normalized the value for z with δ and also subtracted a correction due to the finite size of wall atoms, specifically $z^* = \tilde{z}/\tilde{\delta}$ with $\tilde{\delta} = \delta - 1.5$ and $\tilde{z} = z - 1.5/2$. The rationale for this correction is the following one: The atoms of the bottom solid sit at $z = 0$ while those at the top solid sit at $z = \delta$. Monomers can approach these positions only up to a distance that is defined by the wall-atom monomer Lennard-Jones interaction distance $\sigma_{\rm sp}$, which is chosen as unity in our work to mimic a reasonably rough surfaces [15]. The correction factor of 1.5 is obtained from the density profile, which shows that monomers can reach upto a distance 0.75 from a "corrugated" surface (for $\sigma_{\rm sp} = 1$) and hence for two surfaces we have a value of 1.5. Precise values cannot be stated, because walls are slightly corrugated and also because the walls are not ideally hard. In the figure 1, we show the normalized data for $n(z^*)$. It can be appreciated that the data for different cases collapse onto one master plot which is consistent with equation (1).

It is well known, that the exponent ν_3 can only be obtained at sufficiently large inter-wall separation [17, 21]. In our simulations, we observe that the exact exponent can be reproduced within 1% if the following inequality is obeyed:

$$\delta \geq \delta^* = 2R_{\rm gb} + 1.5\sigma_{\rm sp}, \tag{2}$$

where $R_{\rm gb}$ is the bulk radius of gyration. To validate the claim that the choice $\delta \geq \delta^*$ will lead to a very good approximation to the exact value of ν_3, we included part (a) in figure 2, whose main part shows how the apparent exponent ν_3 changes with δ at fixed $N = 220$. The more detailed analysis in inset (b) shows that the deviation between our data and the predicted exponent decreases with δ, while inset (c) demonstrates that Flory's exact exponent can be observed within 1% accuracy for different chain lengths N as long as our rule, equation (2), is obeyed.

Table 2. Apparent exponent ν_3 for different wall separations and chain lengths. Included is information on the radius of gyration in the bulk solution, $R_{\mathbf{gb}}$.

N	R_{gb}	δ^*	ν_3
100	7.5 ± 0.3	16.5	0.591
160	9.6 ± 0.4	21.5	0.593
220	11.6 ± 0.7	25.0	0.586
310	14.2 ± 0.8	30.0	0.587
Theory			0.587

The data leading to part (c) of figure 2 is shown in table 2. Included are also the values we obtained for the bulk radius of gyration R_{gb}, which was calculated from single polymers in the absence of any confining walls.

3.2. Chain conformation

We now direct our focus on the chain conformation. The quantity that best describes the polymer conformation in is the tensor of gyration $T_{\alpha\beta}(\delta)$,

$$T_{\alpha\beta}(\delta) = \frac{1}{N}\left\langle \sum_{i=1}^{N} (R_{i\alpha} - \bar{R}_\alpha)(R_{i\beta} - \bar{R}_\beta) \right\rangle, \tag{3}$$

where α and β indicate the Cartesian indices. $R_{i\alpha}$ is the α^{th} component of the i^{th} monomer position, and $\bar{R}_\alpha$ is the α^{th} component of the center-of-mass of the chain. In the following, we will distinguish between the in-plane radius of gyration, $R^2_{\mathrm{g}||} = T_{xx} + T_{yy}$, and the out-of-plane radius of gyration, $R^2_{\mathrm{g}\perp} = T_{zz}$.

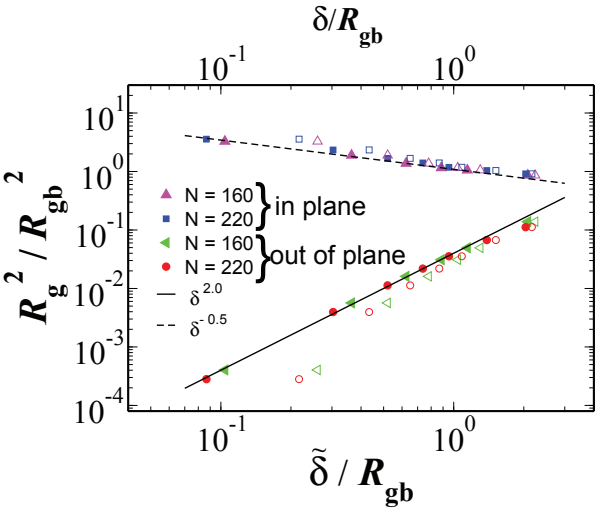


Figure 3. Normalized radius of gyration squared as a function of normalized inter-wall distance, $\tilde{\delta}/R_{\mathrm{gb}}$ for corrected data (solid symbol) and δ/R_{gb} for uncorrected data (open symbol), for two different chain lengths. Lines are power law fits according to equations. (4-5) and mean-field values for Flory's exponents.

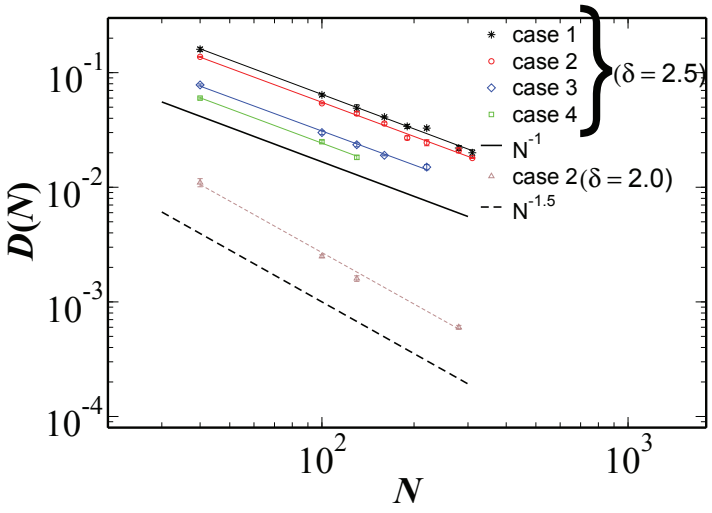


Figure 4. In-plane chain lateral diffusion D as a function of chain length N. Results are shown for both commensurate and incommensurate surfaces as mentioned in table 1 for $\delta = 2.5$. For cases 3 and 4 we increase the temperature from $T = 0.5$ to $T = 0.6$. Line are fits with the exponent shown in caption.

The following scaling laws are established [17, 21] for in-plane and out-of-plane radius of gyration in the limit of small δ:

$$R^2_{g||}(\delta) \propto \delta^{-2(\nu_2-\nu_3)/\nu_3}, \tag{4}$$

and

$$R^2_{g\perp}(\delta) \propto \delta^2, \tag{5}$$

where $2(\nu_2 - \nu_3)/\nu_3 = 1/2$.

In the analysis of our simulation data, we will compare the numerical data to the mean field value of ν_d rather than to the exact exponent ν [21], because our error bars are not sufficiently small to discriminate between the two. Results are depicted in figure 3 for two different chain lengths. By normalizing the value of δ, $R_{g||}$, and $R_{g\perp}$ with R_{gb}, results for the two different chain length could be collapsed. In addition, by subtracting the excluded-volume "offset" of $1.5\sigma_{sp}$ from δ, the power-law dependence $R^2_{g||} \propto \tilde{\delta}^{-1/2}$ and $R^2_{g\perp} \propto \tilde{\delta}^2$ could be borne out on a larger domain at small wall separations. Of course, at large values of $\tilde{\delta}$, the scaling laws valid for small $\tilde{\delta}$ must break down, because for $\tilde{\delta} \geq R_{gb}$, the configurations become bulk-like, and thus $R^2_{g||}$ approaches $2R^2_{gb}/3$ and $R^2_{g\perp}$ approaches $R^2_{gb}/3$.

4. Results for polymer dynamics

Having reproduced the known static behavior of a confined polymer chain, we now direct our focus on dynamic properties. We investigate how D changes with N at fixed separation between two walls. Results for $\delta = 2.5$ are summarized in figure 4, in one case we also reduce $\delta = 2.0$. For polymer bond lengths commensurate with the walls, we had to increase the temperature from $T = 0.5$ to $T = 0.6$, in order to speed up polymer diffusion for accurate measurements. The data can be fit via Rouse dynamics, namely $D \propto N^{-1}$, in all cases for $\delta = 2.5$. Looking into the simulation snapshot we find that polymers essentially form a flat monolayer configuration. Therefore, this flat polymer is "rubbing" against the repulsive walls so that the damping (i.e., the inverse diffusion constant) of the polymers centroid is simply proportional to the number of

monomers in contact with the surfaces. This argument explains why we observe a scaling linear in N.

It may be worth comparing these results to those found when a single polymer is strongly adsorbed onto only one wall, which however is so strongly adhesive that the polymer collapses into a single layer. For the strongly adhesive single wall, commensurability plays a role, while in the new study it does not. Specifically, for a single adhesive wall, we had found $D \propto N^{-1}$ for a polymer whose bond lengths are commensurate with the substrate and $D \propto N^{-1.5}$ for the incommensurate case. Thus, our new simulations show similarity with the commensurate, single-wall case and so there appears to be a contradiction: If two walls are commensurate with each other but incommensurate with the polymer, then in the limit of a single-layer confinement, one should produce the $N^{-1.5}$ and not the $1/N$ scaling of D. This issue can be resolved as follows: In our previous work we had shown that the $N^{-1.5}$ is only born out if N is sufficiently large and/or the walls sufficiently rough. Small roughness first produced N^{-1} scaling at small N before crossing over the $N^{-3/2}$ dependence. The small-roughness regime is what we see in the present study at $\delta = 2.5$. If we reduce the separation between the two walls to $\delta = 2$, we recuperate the $N^{-3/2}$ scaling law for an incommensurate polymer between incommensurate surface, because at this small separation, corrugation becomes large. We have also compared the out-of-plane radius of gyration in this case to that of the adsorbed polymers [15] and they are found to be quantitatively the same.

5. Conclusions

We use molecular dynamics simulations to study the static and dynamics of confined polymer chain. Static properties, such as radius of gyration and out-of-plane monomer density profile, follow the predicted scaling behavior. Chain length dependent lateral dynamics follow Rouse behavior, unless the confinement is extreme, in which case $D \approx N^{-3/2}$ for incommensurate polymer-wall geometries.

References

[1] S. S. Patel and M. Tirrel 1989 *Annu. Rev. Phys. Chem.* **40** 597.
[2] E. Eisenriegler *Polymers near surfaces* (World Scientific, Singapore, 1993).
[3] S. Granick et. al. 2003 *J. Pol. Sci. B* **41** 2755.
[4] M. H. Müser and M. O. Robbins 2000 Phys. Rev. B **61** 2335.
[5] O. B. Bakajin, T. A. J. Duke, C. F. Chou, S. S. Chan, R. H. Austin, and E. C. Cox 1998 *Phys. Rev. Lett.* **80** 2737.
[6] W. Reisner, K. J. Morton, R. Riehn, Y. M. Wang, Z. Yu, M. Rosen, J. C. Sturm, S. Y. Chou, E. Frey, and R. H. Austin 2005 *Phys. Rev. Lett.* **94** 196101.
[7] J. Han and H. Craighead 2000 *Science* **288** 1026.
[8] A. Milchev and K. Binder 1996 *Macromolecules* **29** 343.
[9] B. Maier and J. O. Rädler 1999 *Phys. Rev. Lett.* **82** 1911.
[10] S. A. Sukhishvili, Y. Chen, J. D. Müller, K. S. Schweizer, E. Gratton, and S. Granick 2000 *Nature* **406** 146.
[11] J. Zhao and S. Granick 2004 *J. Am. Chem. Soc.* **126** 6242.
[12] L. Zhang and S. Granick 2005 *Proc. Nat. Acad. Sci. U.S.A.* **102** 9118.
[13] D. Mukherji and M. H. Müser 2006 *Phys. Rev. E.* **74**, 010601(R).
[14] D. Mukherji and M. H. Müser 2007 *Macromolecules* **40** 1754.
[15] D. Mukherji, G. Bartels and M. H. Müser 2008 *Phy. Rev. Lett.* **100** 068301.
[16] F. Brochard and P.-G. de Gennes 1977 *J. Chem. Phys.* **67** 52.
[17] A. Milchev and K. Binder 1998 *Eur. Phys. Jour. B* **3** 477 (1998).
[18] Y. L. Chen, M. D. Graham, J. J. de Pablo, G. C. Randall, M. Gupta and P. S. Doyle 2004 *Phys. Rev. E.* **70** 60901.
[19] A. Balducci, P. Mao and P. S. Doyle 2006 *Macromolecules* **39** 6273.
[20] A. Balducci, J. Tang, and P. S. Doyle 2008 *Macromolecules* **41** 9914.
[21] P.-G. de Gennes *Scaling Concepts in Polymer Physics* (Cornell University Press, London, 1979).
[22] K. Kremer and G. S. Grest 1990 *J. Chem. Phys.* **92** 5057.

Low voltage electron beam induced degradation and surface chemical changes of $Zn_3(PO_4)_2$:Tb phosphor

I. M. Nagpure*, Shreyas S. Pitale, Liza Coetzee, O. M. Ntwaeaborwa, J. J. Terblans and H. C. Swart**
Department of Physics, University of The Free State, P. O. Box 339, Bloemfontein, 9300, Republic of South Africa

Corresponding authors: *indrajitnagpure@yahoo.co.in (I M Nagpure) / **swarthc@ufs.co.za (HC Swart)

Abstract: Tb activated α-$Zn_3(PO_4)_2$ phosphor powders have been subjected to low voltage electron bombardment of 2 keV-10 μA at an oxygen pressure of 1 x 10^{-6} Torr. The phosphor consists out of the monoclinic structure. Zn is present at two different sites Zn_1 and Zn_2 in the $Zn_3(PO_4)_2$ host, as inferred from the crystallographic study. Cathodoluminescence (CL) and Auger electron spectroscopy of the phosphor excited by the same electron beam were used to monitor changes in the surface state during electron bombardment. A direct correlation between the surface reactions and the degradation of the CL brightness was observed. The surface chemical composition measurements of the powder before and after electron degradation indicated the formation of new compounds on the surface. A metal rich P and/or metal oxide layer of phosphorous were formed on the degraded surface, which was responsible for the degradation. The layer formed according to the electron stimulated surface chemical reactions mechanism.

1. Introduction

The development of flat panel displays device such as Field emission displays (FEDs) requires highly efficient cathodoluminescent (CL) materials [1-2]. Rare earth ion activated phosphors such as Ce^{3+} co-doped with Mn^{2+} and Tb^{3+} in $Zn_3(PO_4)_2$ [3] exhibited favorable spectroscopic properties for use in various optical applications. The phosphate based $Zn_3(PO_4)_2$:Mn,Tb red emitting phosphor has been prominently used as a phosphor for color television screens [2]. The phosphate chemistry has attracted a lot of attention during the past decades and has undergone a lot of developments. Interest in this material is also generated due to the crystallographic arrangement of the crystal. Zinc orthophosphate $Zn_3(PO_4)_2$ has three crystalline modifications, labeled as α, β and γ [4]. A lliterature survey confirms that the β and γ phases of zinc orthophosphate are more chemically and thermally stable compared to the α phase [4]. The ideal apatite α-$Zn_3(PO_4)_2$ crystallizes into the monoclinic system with a uni-axial lattice that belongs to the crystallographic group $C_{12}/c_1(15)$ [5]. The structure consists of two non-equivalent crystallographic sites for the Zn, namely Zn_1 (Zn in a column) with a C_s symmetry and Zn_2 (Zn screw axis) with a C_3 symmetry [5].

232 SA Institute of Physics 2011 ISBN: 978-1-86888-688-3

Mn and Tb doped α-$Zn_3(PO_4)_2$ phosphor was tested for good thermal stability [2] and showed good CL efficiency for cathode ray tube colour television. There is, however, no report to confirm the CL stability of the phosphor under continuous electron beam bombardment. This study deals with the CL emission and chemical stability of the phosphor under continuous electron bombardment for low voltage FED applications. The surface chemical reactions and influence on the CL intensity induced by prolonged electron beam bombardment were also monitored using *in situ* Auger electron spectroscopy (AES) combined with CL spectroscopy.

2. Synthesis and characterization methods

2.1 Synthesis method

The gel-combustion method was used for the synthesis of the α-$Zn_3(PO_4)_2$:Tb crystalline phosphors. Zinc nitrate ($Zn(NO_3)_24H_2O$), ammonium dihydrogen phosphate ($NH_4H_2(PO_4)$), citric acid ($C_6H_8O_7H_2O$) and $Tb(NO_3)_3.5H_2O$, all of analytical purity were used as starting materials. Hydrated citric acid was the source of the citrate anions that were used as both chelating agent to metal cations and fuel for the combustion. The zinc nitrate, ammonium dihydrogen phosphate and terbium nitrate were combined to yield a composition with the general formula $(Zn_{1-x}Tb_x)_3(PO_4)_2$ with x = 1m%. The ratio of nitrate to citrate used in the present work is 1:1. The mixed solution was heated at 120^{o}C and continuously stirred using a magnetic stirrer for 3-4 hours. The auto-combustion took place accompanied by the evolution of a brown fume and formation of a citrate complex. The solution turned to a bluish brown sticky gel. The bluish brown sticky gel was then heat treated at 900^{o}C for 2 h in a muffle furnace in air. The evolution of a gaseous fume was observed during annealing and the final product appeared to be a white crystalline powder.

2.2 Characterization methods

The as-prepared sample was characterized with a Bruker X-ray diffractometer (XRD) with CuKα radiation (λ = 1.541 Å). PL measurements were carried out on a Varian Cary Eclipse fluorescence spectrophotometer at room temperature. The Auger measurements were made in an UHV chamber with a PHI Model 549 system. The phosphors powder was excited for CL measurements by the same primary electron beam (beam voltage 2 kV, current of 10 μA) which was use for Auger excitation. CL measurements were done on the emitted light at an angle of 60^o to the incident electron beam. CL intensity data for the peak at 541 nm were collected up to an electron dose of590 C/cm^2 with a PC2000 Spectrometer using OOI Base32 computer software. The CL experiment was done at an oxygen pressure of 1 x 10^{-6} Torr.

3. Results and discussion

3.1 Structural analysis

XRD data of the $Zn_3(PO_4)_2$ matrix is shown in figure 1. The crystalline lines are in agreement with the pure crystalline α-$Zn_3(PO_4)_2$ phase referenced in ICCDs data file No. 029-1390. It ishaving a monoclinic structure with the space group C_{12}/c_1 (15) and Z =4 [5]. The unit cell of the α-$Zn_3(PO_4)_2$ structure is shown in figure 2. In the pyrophosphate family the α-$Za_3(PO_4)_2$ belongs to oxyphosphate, has the highest symmetry with the space group corresponding to C_{12}/c_1 (15) which can be attributed to the occurrence of oxygen ions in the planes of triangles constituted out of alkaline earth atoms [5]. The Zn_1 atoms (C_s symmetry) occupies the 4e Wyckoff position along the plane that forms the tetrahedral with the oxygen atoms and is present in between the Zn_2 (C_3 symmetry) and the oxygen tetrahedral at the 8f Wyckoff position [5]. Polarization of the Zn^{2+} divalent metal ions in the triangles by the small O^{2-} ions might play the role of reducing the effective radius in the direction towards the O^- ion showing more effective PL and CL properties. This structure can accommodate a variety of substitutions such as rare earths, transition metals and alkaline earths.

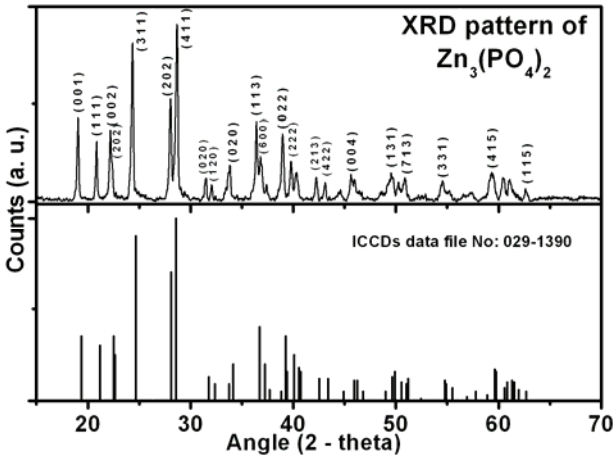


Figure-1: XRD crystal peaks of the α-$Zn_3(PO_4)_2$ powder

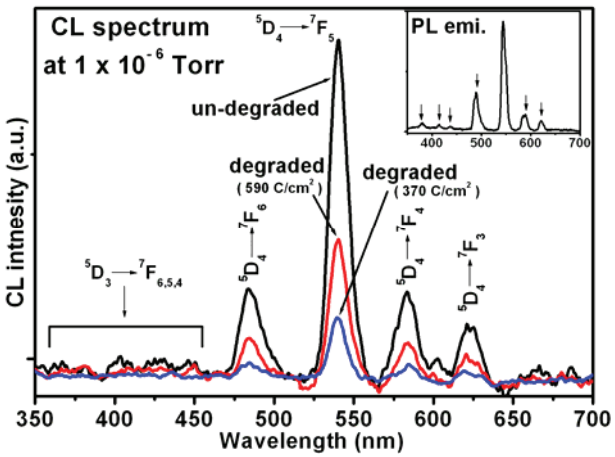


Figure-3: The CL spectra of the $Zn_3(PO_4)_2$:$Tb_{1mol.\%}$ phosphor, (PL emission spectrum is shown in the inset).

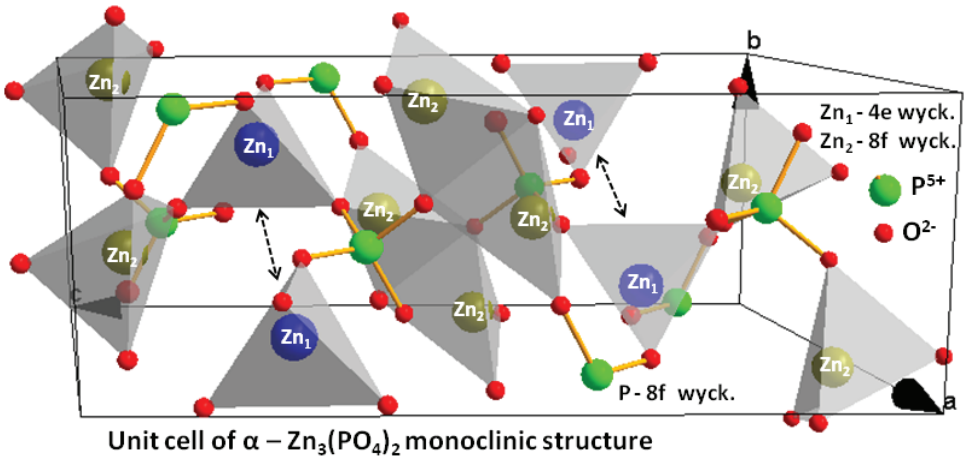


Figure-2: Crystal structure of the unit cell of the α-$Zn_3(PO_4)_2$ host.

3.2, CL and PL analysis

Figure 3 illustrate the CL emission spectrum of α-$Zn_3(PO_4)_2$:Tb excited with a 2 keV-10 μA electron beam. The CL spectra were measured at 1 x 10^{-6} Torr when the chamber was backfilled with oxygen. The CL spectrum consists of several emission peaks in the 470-700 nm range. The CL emission peaks due to the $^5D_4 \rightarrow {}^7F_6$, $^7F_5$, $^7F_4$ and $^7F_3$ transitions are observed at 485, 541, 583 and 622 nm, respectively [6]. The green CL emission at 541 nm due to the $^5D_4 \rightarrow {}^7F_5$ transition of Tb^{3+} appeared to be more prominent compared to the blue and red emissions. The PL emission spectrum of α-$Zn_3(PO_4)_2$:Tb measured at a 230 nm excitation wavelength is shown in the inset of figure 3. The PL emission spectrum also has weak blue emission peaks in the region between 350-480 nm. The weak PL emission peaks at 380, 420, 440 nm are due to the transitions from the excited level $^5D_3$ to the $^7F_6$, $^7F_5$, $^7F_4$ levels of the Tb^{3+} ions, respectively. These peaks are almost completely absent in the CL emission spectra. The prominent PL emission peaks observed at 490, 543, 584 and 620 nm are also due to the $^5D_4 \rightarrow {}^7F_6$, $^7F_5$, $^7F_4$ and $^7F_3$ transitions of the Tb^{3+} ions, respectively [6]. Also shown in figure 3 are the CLspectra after different electron doses. The intensity clearly decreased during electron bombardment.

3.3, CL degradation and APPHs analysis

CL degradation measurements were carried out by monitoring the 541 nm Tb ion CL emissions peak simultaneously with the Auger peak to peak heights (APPH's). Figure 4 shows the normalized CL intensity and APPHs spectra as a function of electron dose at an oxygen pressure of 1 x 10^{-6} Torr. Figure 4 shows an initial decrease of the CL intensity up to an electron dose of 70 C/cm^2 followed by a rapid decrease from 70 to 370 C/cm^2, which later increased and stabilized after 470 C/cm^2. This behaviour is due to the prolonged electron bombardment of the sample, which caused changes in the surface chemistry of the sample as indicated in the APPH changes as shown in figure 4. The adventitious carbon (C) desorbed from the surface as CO_2 and CH_4 as a result of the electron beam exposure (decrease in C APPH's) [7]. The concentration of the oxygen (O) and zinc (Zn) rapidly decreases with an increase electron dose. Simultaneously with the decrease in C, Zn and O a dramatic change in the phosphorous (P) elemental concentration was observed. Elemental APPH analysis indicates that P initially increases (i.e from 0 to 30 C/cm^2) and then stabilized (i.e from 30 to 220 C/cm^2) and which latter start to increases from 380 C/cm^2. It is therefore clear that electron stimulated surface chemical reactions (ESSCRs) [8] occurred on the surface of the α-$Zn_3(PO_4)_2$:Tb phosphor during prolonged electron bombardment.

3.4, AES analysis

The differentiated Auger spectra (AES) of $Zn_3(PO_4)_2$:$Tb_{1m\%}$ phosphor before and after 490 C/cm^2 electron exposure at 1 x 10^{-6} Torr oxygen pressure is shown in figure 5. After prolonged electron exposure the Auger spectra confirm the removal of adventitious carbon (C) and desorption of O from the surface as observed in the APPHs of figure 5. After an electron exposure of 490 C/cm^2, the P peak at 98 and 118 eV is altered in shape and shift in a peak position to a lower energy as shown in figure 6. This indicates that a phosphorous (P) rich surface was formed on the degraded surface [9]. The shift in the energy position clearly shows that during the surface modification the work function of the material changed and charging occurred. The charging has a detrimental effect in the AES peak intensity and shape and also on the CL intensity. Band bending will occur if charging takes place. Electron and holes are swept apart before recombination and less light will be excited, with a consequence loss in CL intensity. The change in concentration on the surface of the sample can be attributed to the formation of new chemical co-ordinates related to the pure P metal or as oxide based P on the surface as explained by the ESSCR mechanism [8].

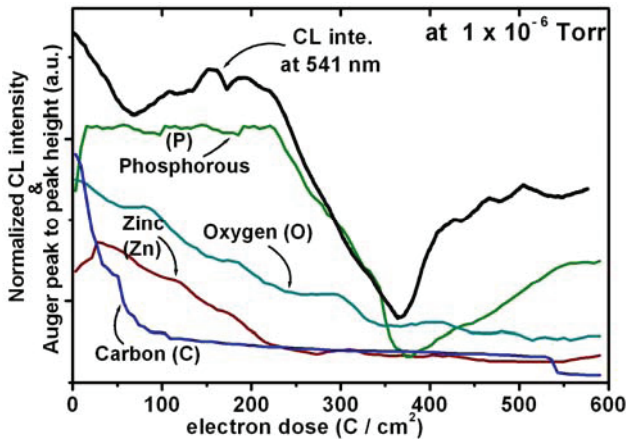


Figure-4: The APPHs variation and normalized CL intensity of the $Zn_3(PO_4)_2$:$Tb_{1\ mol.\%}$ phosphor.

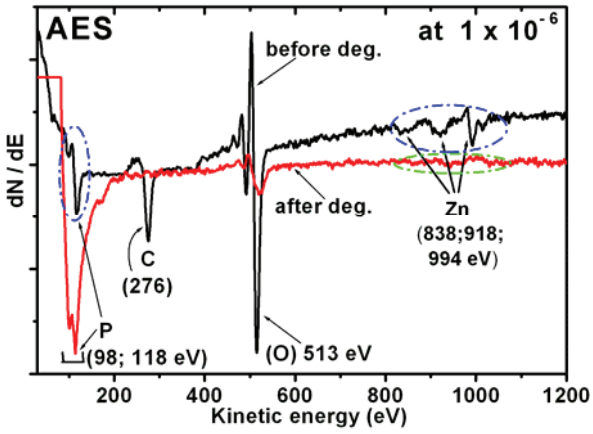


Figure-5: The AES measurements before and after degradation

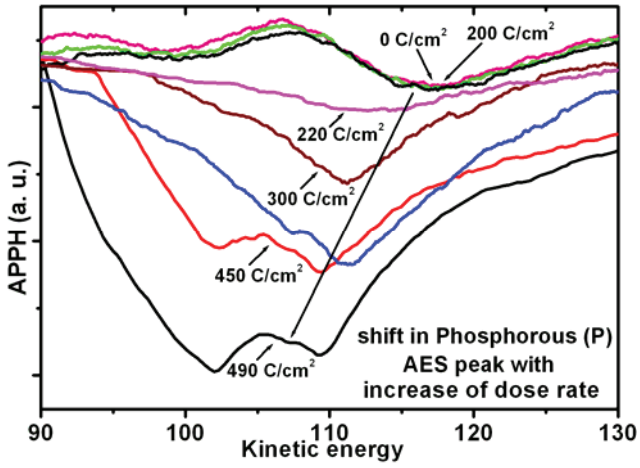


Figure-6: The P AES peak after different coulomb doses as indicated.

According to the ESSCR model the Zn–O and P–O bonds are likely to be broken into free oxygen, zinc and phosphorous when irradiated with a beam of electron. This will be followed by a chemical reaction resulting in new chemical compounds forming on the surface. In most cases the new oxide layers are non-luminescent and therefore will reduce the CL intensity. Simultaneous with the O_2 desorption during CL degradation, it is most likely that P (metallic), and P_2O_7 [9] mixed layers were formed on the surface according to the ESSCR mechanism. In this case these layers may form according to the following chemical reaction as mentioned in equation-1:

$$7.Zn_3(PO_4)_2 \leftrightarrow 21Zn + 2P + 7P_2O_7 + CH_4\uparrow + CO_2\uparrow (\Delta H_{f298K}) \text{ --- Eq-1}$$

$$(-270.31) \leftrightarrow (-137.35) + (-6.19) + (-104.36) + (0)\uparrow + (0)\uparrow \ (\Delta H_{f298K}) = 22.39 \text{ kcal/mol } O_2)$$

4. Conclusion

Sol-gel combustion synthesized α-$Zn_3(PO_4)_2$:Tb phosphor was used for the CL degradation study. The optical properties such as PL and CL were evaluated and compared for the α-$Zn_3(PO_4)_2$:Tb phosphor. The green CL emission at 541 nm due to the $^5D_4 \rightarrow {}^7F_5$ transition of Tb^{3+} appeared to be the more prominent emission. The APPHs and AES results of the degraded sample at 10^{-6} Torr oxygen pressure confirmed the formation of new coordinated related metal rich P species on the surface during degradation with a direct consequence of CL intensity degradation. The CL degradation nature is directly related to the change in P elemental concentration on the surface. It is clear that ESSCR's were responsible for the degradation in CL intensity and the effect of band bending due to charging is not excluded.

Acknowledgement: Authors are thankful to the University of the Free State (UFS) and South African National Research Foundation (NRF) for the financial support.

References

[1] Swart HC, Greeff AP, Holloway PH and Berning GLP 1999 *Appl. Surf. Sci.* **140** 63
[2] Franz KA, Hehr WG, Siggel A 1985 *Luminescent materials*, Ullmann's Encyclopedia of Industrial Chemistry A15, VCH, Veinheim, Germany
[3] Wang Jing, Wang Shubin and Su Qiang 2004 *J. Mater.* Chem. **14** 2569
[4] Song Yan-Hua, Zou Hai-Feng, Gan Shu-Cai, Deng Yue-Feng, Hong Guang-Yan, Meng Jian 2007 *J. Mater Sci.* **42** 4899
[5] Calvo Crispin 1965 *Canadian J. Chem.***43** 436

[6] Chung-Hsin, Godbole Shrikant Vasant and Qureshi Mohammad 2006 *Jpn. J. Appl. Phys.* **45** 2606
[7] Holloway PH, Sebastian JS, Trottier TA, Swart HC and Peterson RO 1995 *Solid State Technology*, Augustus, 47
[8] Swart HC, Sebastian JS, Trottier TA, Jones SL, and Holloway PH, 1996 *J. Vac. Sci. Technol. A* **14(3)** 1697
[9] Shah KV, Goswami M, Manikandan S, Shrikhande VK and Kothiyal GP 2009 *Bull. Mater. Sci.* **32 (3)** 329

Effect of pH on the morphology and orientation of Fe_2O_3 nanostructures grown using aqueous chemical growth

C L Ndlangamandla [1,2], C B Mtshali[1, 2], M J Sithole[2], B D Ngom[3], O M Ndwandwe[1] and M Maaza[2]

[1] Department of Physics and Engineering, University of Zululand, Kwa-Dlangezwa, South Africa

[2] Nanosciences Laboratories, Material Research Department, iThemba LABS, South Africa

[3]CSIR National Laser Centre, PO Box 395, Pretoria 0001, South Africa

E-mail : cebo@tlabs.ac.za

Abstract. Fe_2O_3 nanostructures were synthesized for various pH values on corning glass substrates by aqueous chemical growth (ACG) using a solution of ferric chloride ($FeCl_3$). It was found that increasing the pH from 1.15 to 3.0 of the solution significantly leads to a modification of the Fe_2O_3 morphology or orientation from randomly parallel needle shaped rod-like to randomly perpendicular rectangular structure on to a substrate

1. Introduction

Hematite (alpha-Fe_2O_3) as an n-type semiconductor has a band gap of 1.9 - 2.2 eV which can absorb approximately 40% of solar incident radiation and stable in most electrolytes with wide range of pH > 3 [1], has a high content of iron and is natural abundance, non-toxic, environmental friendly, photo-corrosive resistance and photo-catalytic [2], which makes it an ideal candidate for water splitting using sunlight compared to alternatives which are abundant and inexpensive such as TiO_2, ZnO, and WO_3 with wide band gaps and are therefore limited in general to terrestrial solar energy efficiency to less than ~4%[3]. The disadvantages of this material is that it has a high resistivity for a single crystal, low electron mobility and very short hole diffusion length of 2 to 4 nm which exhibit high rate of charge carrier recombination and therefore poor charge transfer .

The efficient photo-voltaic properties have been demonstrated through designing a thin film of hematite consisting of crystalline arrays of oriented nanorods [4]. These designs exhibited substantial photo-current efficiency due to a better transport and collection of photo-generated electrons through a designed path. Hematite has a conduction band gap edge which is too low than the chemical potential for hydrogen gas (H_2/H^+) evolution [5,6] without an external bias, although the valence band edge is located at a lower energy level, which is in correspondence with the chemical potential for oxygen gas

(H_2O/O_2) evolution. Therefore this energy position of the band edge level can be modified using the electronegativity of the dopants and solution pH [7].

ACG [7,8,9] is a quite simple technique for the deposition of nanostructures at relatively low temperatures, presenting several advantages such as the use of non-expensive equipment, the requirement of cheap and non-toxic reagents and the presence of non-hazardous by-products. Moreover, the morphological and structural characteristics of the grown samples can be controlled by adjusting parameters such as the concentration of the solution, the reagents stoichiometry, the temperature and the pH [10,11]. In this work, we investigate the possibility of controlling the nanostructures morphology grown using ACG by adjusting the pH of the solution. Towards this scope, results are presented related to the deposition of Fe_2O_3 nanostructures on corning glass, at a deposition temperature of 95 °C, for a range of growth periods and various pH values.

2. Experiment

Hematite (alpa-Fe_2O_3) nanostructures was grown by employing aqueous chemical growth technique, using 0.1 M aqueous solution of ferric chloride ($FeCl_3$) as precursors at various values of pH adjusted using a concentrated hydrochloric acid (HCl) solution. The solution and the glass substrate was placed in Pyrex glass bottles with polypropylene autocleavable screw caps and heated at a constant temperature of 95°C for 24 h in a regular laboratory oven. After this generation period, the samples were taken out and washed thoroughly with MilliQ de-ionized water to reduce residual salts. The homogeneous films of akaganiete (FeOOH) was obtained first on the substrate and annealing it on air at 500°C for 1hr to transform FeOOH to alpa-Fe_2O_3 as confirmed using x-ray diffractometer (model Bruker AXS D8 Advance). The morphology investigation was studied using a Leo-Stereo Scan 440 Scanning Electron Microscope (SEM).

3. Results and discussion

Figure 1 presents the XRD patterns of Fe_2O_3 powder deposited for 24 h collected from the bottom of the autoclivable bottle and the presented graph are for pH 1.15 to pH 5.0. The patterns exhibit all the characteristic peaks of hematite alpha-Fe_2O_3 rhombohedral R-3C space-group structure, according to JCP2. CAT: file No. 73-0603 or 72-0469. This suggested that only hematite phase was formed, no impurities present on the sample.

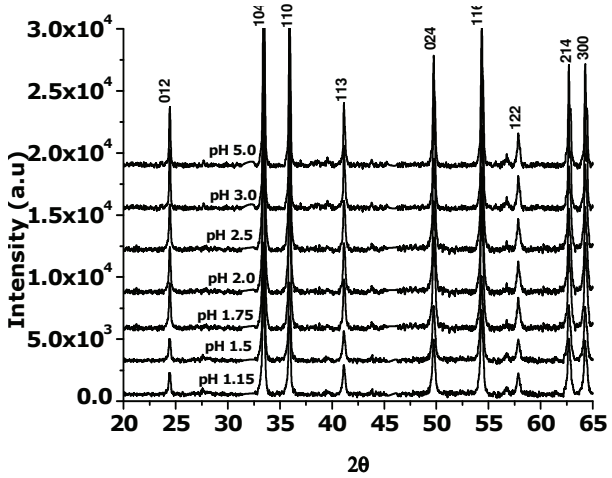


Figure 1. XRD analysis of Fe_2O_3 powder from a 0.1 M solution of $FeCl_3$ at 95 °C for a deposition of pH 1.15 to 5.0

SEM images of Fe_2O_3 hematite nanostructures grown for 24 h on corning glass using an aqueous solution at 95 °C and for various pH values ranging from pH 1.15 to pH 5.0 are presented in figure.2.

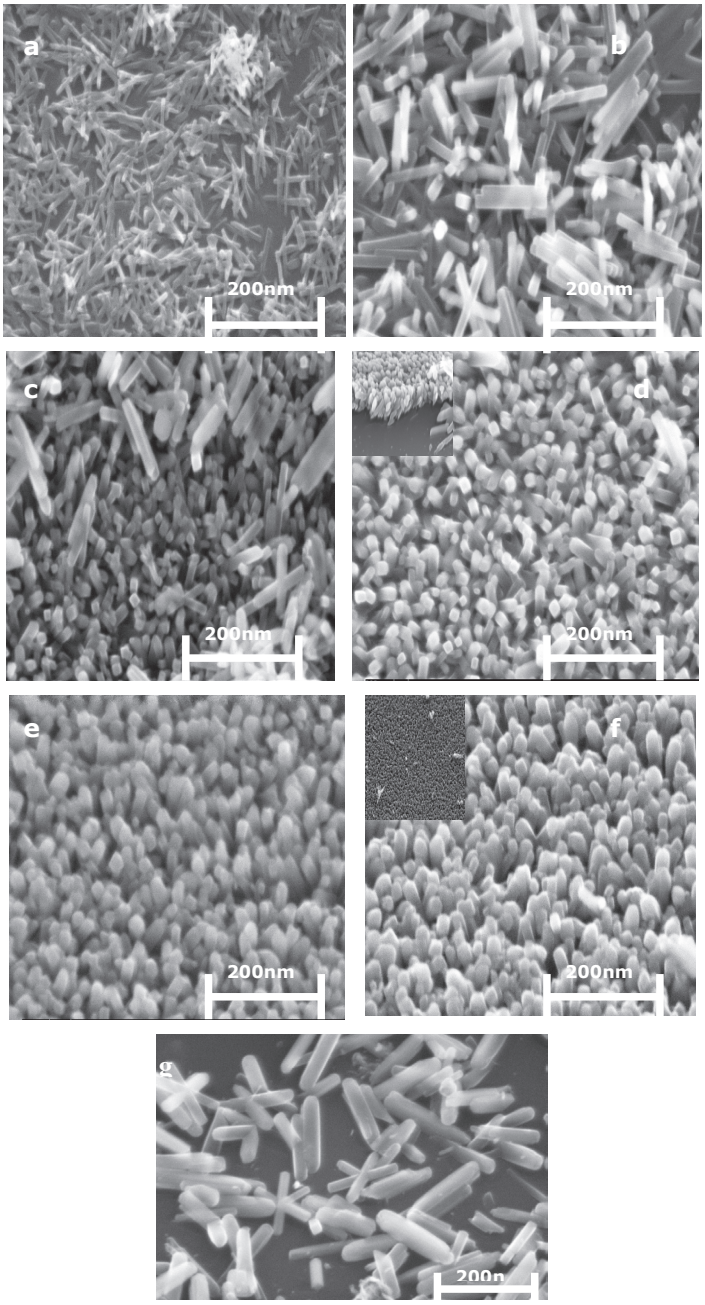


Figure 2. SEM images of Fe_2O_3 deposited on corning glass using ACG for a deposition time of 24 h at various pH values, (a) pH 1.15, (b) ph1.5, (c) pH 1.75, (d) pH 2.0, (e) pH 2.5, (f) pH 3.0 and (g) pH 5.0, which presenting the morphology and the orientation of the nanostructures all the images were taken at a scale of 200nm.

As can be seen, the morphology and orientation of the obtained Fe_2O_3 nanostructures can be affected by the pH of the solution. At a pH of 1.15 needles like nanorods random oriented parallel onto the substrate are observed. At a pH of 1.5 the nanorods are cubic at the tip and the orientation is mixed random and perpendicular oriented on to a substrate. Therefore this shows that the perpendicular orientation of cubic nanorods starts at pH 1.5. Increasing the pH to a value of 1.75, a most observed orientation of the nanorods is perpendicular on the substrate, while only cubic structured perpendicular nanorods are observed at the pH of 2.0. At the pH of 2.5 to the pH of 5.0, nanorods with a spherical shape at the tip are observed. Finally, at pH 5.0, the random oriented nanorods structures are observed. Following the above observations, it is clear that the morphological characteristics of as-prepared Fe_2O_3 nanorods arrays are controlled by the concentrated hydrochloric acid. This is due to the fact that increasing the pH, the increase in the precipitate is observed after every generation time leading to the random orientation of these nanostructures.

Furthermore, as the morphology and the orientation of the nanostructures changes with increasing pH, the diameter of the nanorods increases with an increase in the pH values while the lengths decreases. It shows from figure 3 (a) that the best engineered rods are found at pH 1.5 to pH 2 where the length is longest. All these changes observed are due to the pH of the solution. To summarize the changes in the morphology of the Fe_2O_3 nanostructures, the graphs of length and diameter as a function of the pH value are shown in figure.3 (a, b). In any case, these are only preliminary results and further investigations, are required in order to clarify both the pH influence on the Fe_2O_3 nanostructure morphology and the actual shape of the individual structures.

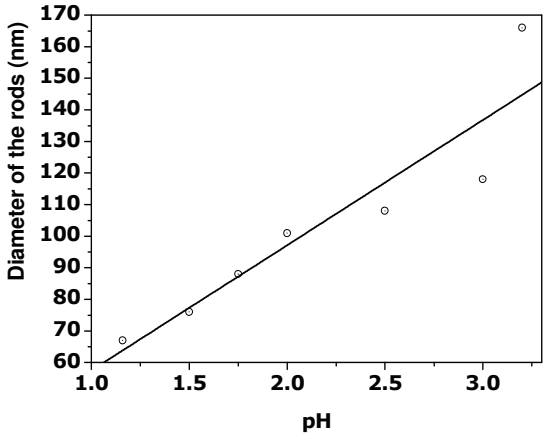


Figure 3 (a). The diameter as a function of the pH values

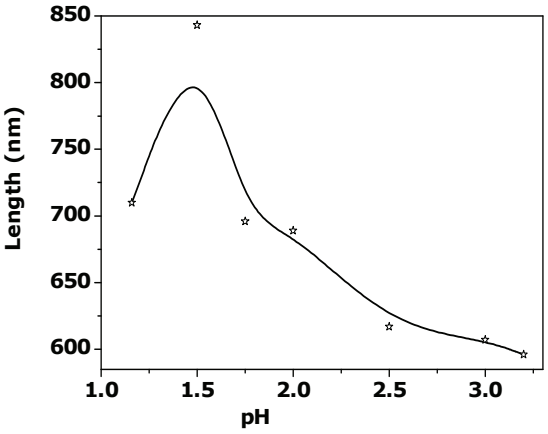


Figure 3 (b). The length of the rods as a function of pH values

4. Conclusion

The pH effect on the morphology and orientation of hematite nanostructures grown by ACG using aqueous solution of $FeCl_3$ as precursors at 95 °C was observed. It was found that pH significantly influences the shape of the Fe_2O_3 nanostructures, leading to a modification of the morphology orientation from randomly parallel oriented rod-like to perpendicular oriented nanorods structures on the substrate. Therefore, one can control the shape of the Fe_2O_3 nanostructures by adjusting the pH of the solution. However, these are preliminary results and further investigation is required to verify the existing pH effect on the properties of Fe_2O_3 samples.

Acknowledgement
The authors wish to thank the financial support From South African Nuclear Human Asset Research Program (SANHARP) and also to iThemba LABS a facility of the National Research Foundation (NRF) of South Africa for the use of their facilities.

References

[1] Kleiman S A, Huda M N, Walsh A, Yan Y, Stucky G D, Hu Y S, Al-Jassim M M and McFarland E W 2010 *Experiment and Theory Chem. Mater*. **22** 510–17
[2] Vayssieres L 2006 *C. R. Chimie* **9** 691–701
[3] Kleiman S A, Hu Y-S, Forman A J, Stucky G D and McFarland E W 2005 *J*. Com. *Chem*. **7** 264
[4] Matsumoto Y 1996 *J. Solid State Chem.***126** 227-34
[5] Vayssieres L 2006 *Solar Hydrogen and Nanotechnology, SPIE Proceedings* **6340**, 1-336
[6] Glasscock J A., Barnes P R F, Plumb I C, Bendavid A and Martin P J 2006 *Solar Hydrogen and Nanotechnology, SPIE Proceedings* **6340**, 1-336
[7] Beermann N, Vayssieres L, Lindquist S E, Hagfeldt A, 2000 *J. Electrochem. Soc.* **147** 2456-61
[8] Vayssieres L 2003 *Adv. Mater.* **15** 464
[9] Li Q 2005 *Chem. Mater* **17** 1001.
[10] Liu X, Jin Z, Bu S, Zhao J and Yu K 2006 *Mater. Sci. Eng., B, Solid-State Mater.Adv. Technol.* **129** 139
[11] Zhang H 2005. *Mater. Lett*. **59** 1696

Investigation of broken symmetry of Sb/Cu(111) surface alloys by VT-STM

G F Ndlovu[1,2,4], J K O Asante[3], W D Roos[2], K T Hillie[1,2]

[1]National Centre for Nano-structured Materials, CSIR, PO Box 395, Pretoria, ZA0001
[2]Department of Physics, University of Free State, Bloemfontein, ZA9300.
[3]Department of Physics, Tshwane University of Technology, Pretoria, ZA0001

E-mail: gndlovu@csir.co.za, rooswd@ufs.ac.za

Abstract. This work present an in situ Variable Temperature Scanning Tunneling Microscopy (VT-STM) study of the Sb/Cu(111) system studied at various temperatures. The experimental data support a structural model for the Cu(111) $(\sqrt{3}$ x $\sqrt{3})R30^\circ$ - Sb in which Sb atoms displace up to 1/3 of the first layer of Cu atoms and incorporate them into the first Cu layer. In contrast to the clean Cu(111) substrate the surface of the surface alloy is corrugated due to the outward relaxation of the alloy atoms thus breaking the surface structural symmetry. The study report for the first time to our knowledge the observation of the Cu(111)(2 x 2) -Sb superstructure obtained at elevated temperatures.

1. Introduction

Recent developments in nanoscience make it possible to engineer artificial structures at surfaces and to gain control over matter at the atomic scale. The unique ability of scanning tunneling microscopy (STM) to reveal details at the atomic level has been particularly valuable. A great variety of structures are formed after deposition of one metal on the surface of another. In many metal-on-metal systems, it is of interest to consider the possibility of bulk or surface alloy formation [1-4]. Elements which are immiscible in the bulk have been found to form stable two dimensional mixtures at the surfaces [1-6]. Some surface adsorbate species such as antimony (Sb) are known to act as surfactants in both homo- and hetero-epitaxy; these adsorbates appear to induce layer-by-layer growth in systems which otherwise tend to island growth [7-9]. The adsorption of Sb on the (111) plane of noble metals surfaces such as copper (Cu) and silver (Ag) has long been of practical interest in the fields of surface science and technology. Theoretical calculations and experimental observations have suggested that the energetics of the Sb/Ag(111) system are such that in the ordered 0.33 ML Ag(111) $(\sqrt{3}$ x $\sqrt{3})R30^\circ$ - Sb phase, the Sb atoms substitute one-third of the outermost Ag atoms to produce an ordered surface alloy [1,3-5]. The structures of these monolayer surface alloys results from a complex interplay of electronic, stress and geometric effects, all related to each other. This study reports on a similar system, Sb/Cu(111) studied by STM, LEED and Auger with a particular emphasis on the structure of the Cu(111) $(\sqrt{3}$ x $\sqrt{3})R30^\circ$ - Sb surface phase.

2. Experimental

The experiments were performed in a UHV chamber with a base pressure below 5×10^{-10} Torr, equipped with a STM, SPECTA-LEED, EFM3 evaporator system from Omicron NanoTechnology (Germany, GmbH). The Cu(111) crystal were acquired from MaTeck with orientation accuracy of less than 1°. The Cu single crystal polished to 0.03 μm, was mounted on a Ta baseplate and cleaned by repeated cycles of 2 keV Ar+ ion sputtering at 45° incidence followed by annealing to 530°C, until the surface was confirmed clean and well ordered as confirmed by AES, LEED and STM. Sb was evaporated from an alumina crucible heated resistively in a Knudsen cell at a crucible temperature of ~700 °C. The Sb was evaporated for 3 s at the deposition rates of 0.1 ML s^{-1}. The VT-STM was operated in the constant-current mode utilizing electrochemically etched tungsten (W) tips. The STM data is displayed in a top-view with the gray-scale representing surface features, the darker levels corresponding to lower lying areas and brighter areas corresponding to higher lying areas. Measurements were taken at room temperature after two different annealing temperatures at 300 ^{o}C and 700^{o}C for 12 hours at each annealing temperature.

3. Results and discussion

3.1. Sb growth on Cu(111) at room temperature

Figure 1(a) shows a high resolution STM image of atomically resolved well ordered hexagonal structured surface after sputtering. The distance between the protrusions is measured to ~ 0.257 ± 0.001 nm which is within the experimental uncertainty identical to the Cu spacing of 0.256 nm along the [110] direction [10]. The high resolution image was taken with a negative bias of (-1 mV). The "grey" scale next to the image illustrates the depth or height of the features on the image ranging from 0.0 to 39.6 pm. The corresponding line profile (inset) taken along A to B in (a) depicts the corrugation of the Cu surface at atomic scale. The STM data for all recorded images was acquired at different scanning directions to exclude any tip-induced artifacts. Figure 1(c) depicts the low energy electron diffraction (LEED) pattern of the clean Cu(111) surface in reciprocal space. The (1×1) pattern is represented by white spots illustrating the periodicity of a well ordered Cu(111) surface.

The STM image in figure 1(b) was acquired after deposition of 0.3 ML Sb on the atomically clean Cu(111) surface, at a temperature of 300°C. The STM image depicts atomically resolved Sb atoms as bright spots on the Cu surface (figure 1(b)). The Sb atoms are more or less randomly distributed and embedded on the Cu surface. The high resolution image was taken with a negative bias of (-1 mV). The on figure 1(b) illustrates the corrugations of the Sb atoms on the Cu surface along the line C to D. The corresponding LEED pattern (figure 1(d)) is similar to that of a clean Cu(111) surface with no extra spots visible on the pattern. The spots remain sharp and with low background after growth. LEED is an averaging technique, thus the Sb atoms are not arranged in a regular pattern within the Cu surface therefore Sb atoms make no contribute to the electron diffraction process.

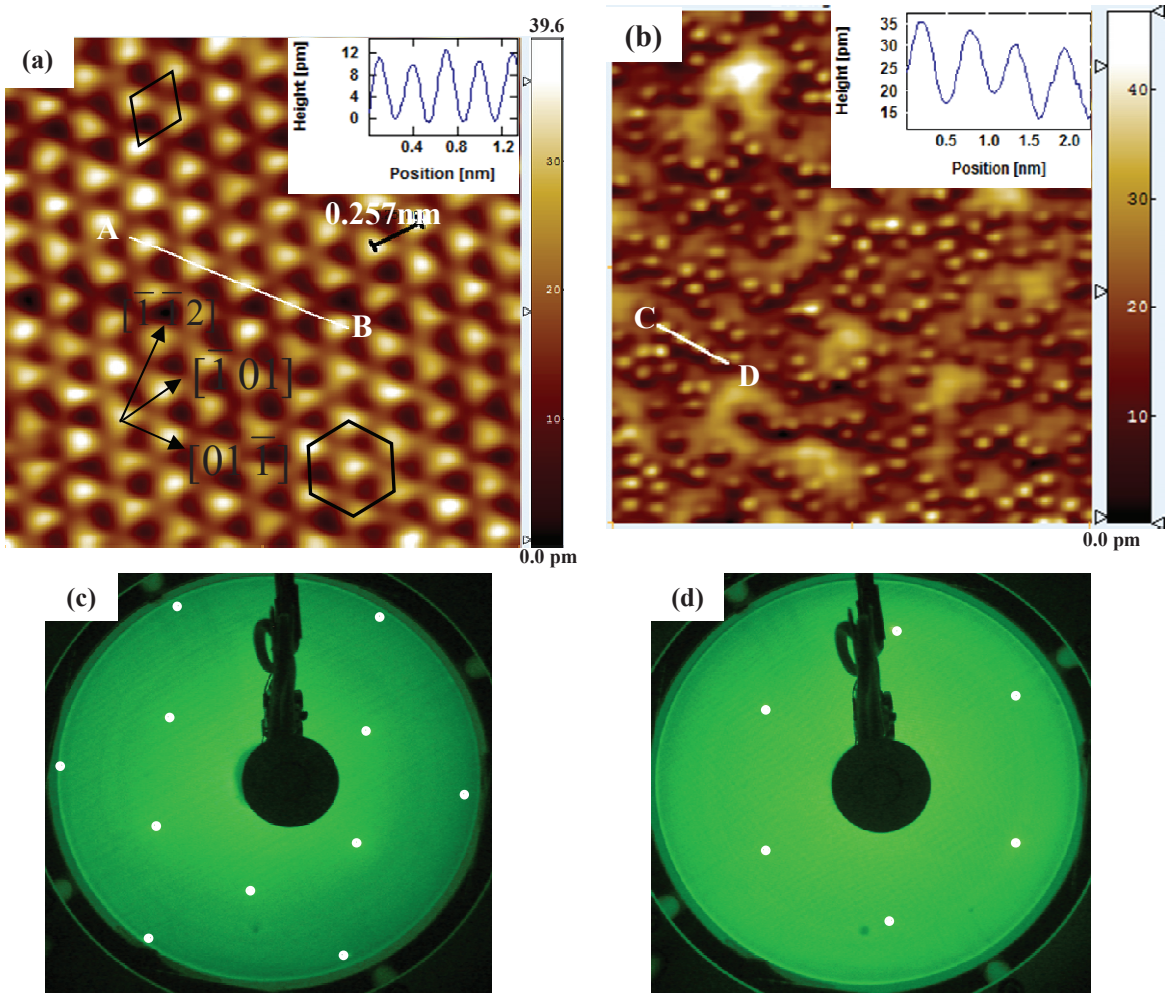


Figure. 1. (a) STM image of atomically clean Cu(111) surface (2.50 nm × 2.50 nm) after cycles of sputtering and annealing at 300°C (V_{bias} = -1 mV, I_{tun} = 2 nA). The black rhombus indicates the primitive unit cell and the hexagon is superimposed to illustrate the arrangement of Cu atoms at the surface. (b) Atomic resolution STM image (13.8 nm × 13.8 nm) upon deposition of ~0.3 ML of Sb (V_{bias} = 5 mV, I_{tun} = 2 nA). The s on the top right of each image show the line scan along A to B and C to D respectively. (c-d) Corresponding LEED patterns of the bare Cu(111) surface before (beam energy 114 eV) and after growth (beam energy 116 eV), respectively. the white dots correspond to spots resulting from Cu atoms in reciprocal space.

3.2. Sb dissolution on atomically clean Cu(111) at 400 °C

After the sample was annealed at 400 °C for 12 hours, successive STM images (9 nm × 9 nm, V_{bias} = 1 mV, I_{tun} = 2 nA) captured the Sb dissolution and showed a perfect $(\sqrt{3} \times \sqrt{3})R30^\circ$ superstructure (figure 2(a)). The corresponding LEED pattern (figure 2(e)) show the $(\sqrt{3} \times \sqrt{3})R30^\circ$ reconstruction when taken at the same beam energy (114 eV) as in the clean Cu(111). Extra spots (white dots) are visible on the diffraction pattern compared to the clean Cu surface of figure 1(c). In both the STM and LEED data, the surface showed long range behaviour as the $(\sqrt{3} \times \sqrt{3})R30^\circ$ structure and form locally hexagonal arrangement of the Sb atoms. The vertical corrugation of figure 2(a) represent the variation in tunneling current ranging from 0.0 to 800 pA as illustrated by the grey scale. A closer look of figure 2(b) it is possible to discern that Sb atoms occupy substitutional sites surrounded by six Cu atoms, three of them being displaced from the lattice site positions of the (111) plane.

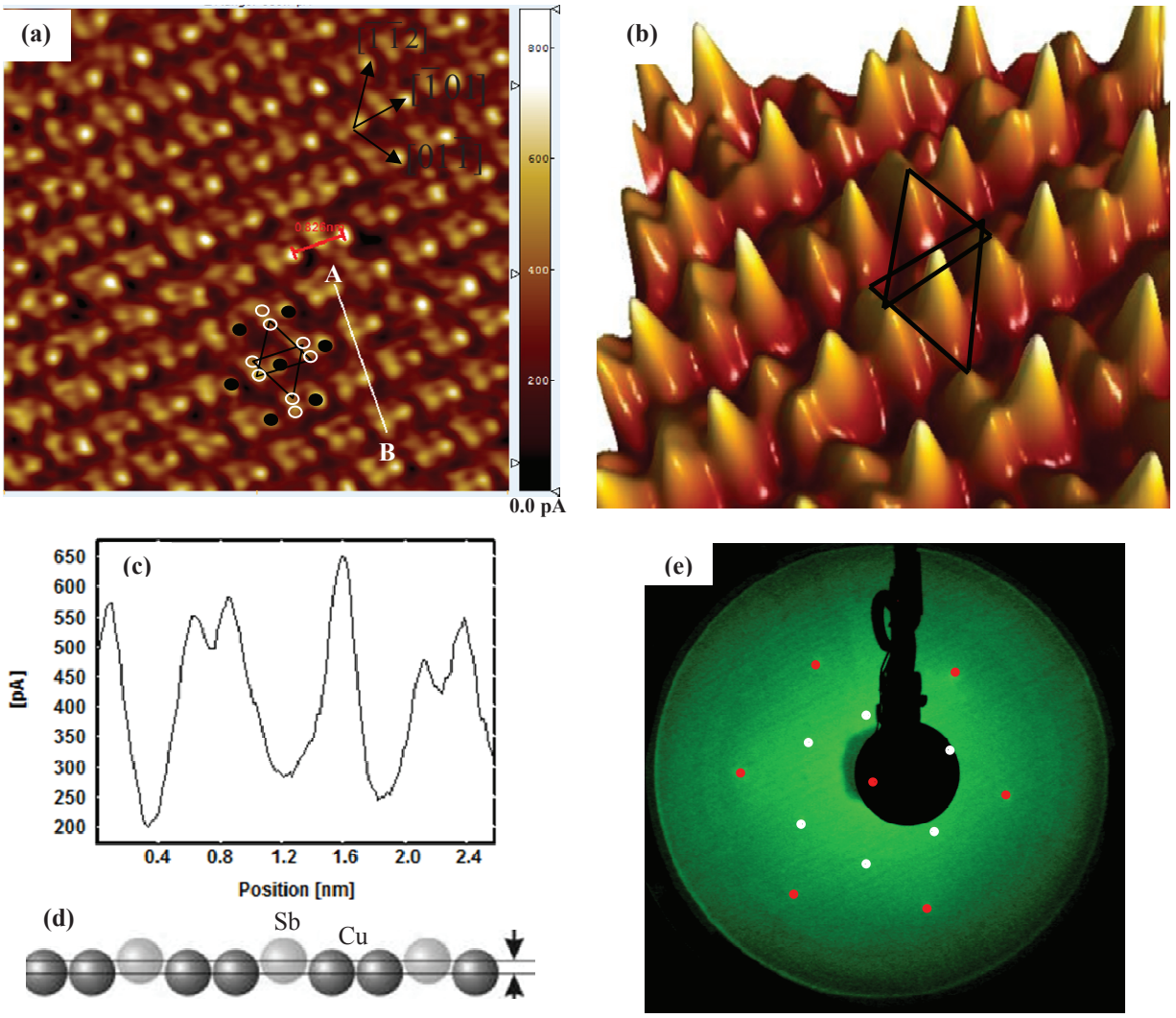


Figure 2.(a) An STM image (8.01 nm × 7.65 nm) clearly showing the six Cu atoms that surrounds each Sb atom in a Sb/Cu(111) $(\sqrt{3} \times \sqrt{3})R30^{\circ}$-Sb reconstruction ($V_{bias}$ = 1 mV, I_{tun} = 2 nA). (b) Three-dimensional (3D) view of the high magnification STM image of the image in (a). The white circles and the black dots are superimposed to represent the position of the Cu and Sb atoms respectively. The superimposed triangles are associated to two in-equivalent Cu atoms and a Sb atom at the centre resulting in surface asymmetry. (c) Line trace along the white line along A to B in (a) showing the amplitude of the corrugation in picoamps vs distance in nm. (d) Schematic representation of the CuSb atomic arrangement at the surface showing rippling. (e) Experimental LEED pattern of the sample surface at 114 eV. Red dots illustrate position of the Sb atom while the white dots correspond to spots resulting from Cu atoms in reciprocal space.

From the STM data, Sb atoms buckle outwards (figure 2(d)) by ~0.23 nm (rippling) compared to the previous reported theoretical and calculated values of 0.25 nm [4]. The hexagonal surface symmetry is then broken by this displacement as illustrated in the three-dimensional picture in figure 2(b). From the STM images (figure 2(a-b)), the Sb atoms remain centred in the <112> -type directions and followed by two lower Cu atoms evident in the line scan of figure 2(c). The out-of-plane displacement (rippling) of Sb incorporated into the copper surface is consistent with the picture of the larger antimony atoms substituting for copper as illustrated by the schematic in figure 2(d).). The inclusion of the Sb in the Cu matrix induces strain in the Cu matrix which is reduced by the outward relaxation.

The outward relaxation is linked to the strongly enhanced spin-splitting in these surface alloys [2,5,12,13]. This is due to the additional planner component of the potential difference provided by this configuration of a high atomic number (Sb) element surrounded by low atomic number element (Cu).

3.3. The Cu(111) -Sb superstructure

The sample was further annealed at ~700 °C for 12 hours and an STM study of the sample surface was conducted at sample temperature of ~ 100 °C. Figure 3(a) illustrates an STM image (6.94 nm x 6.76 nm) of the CuSb surface obtained as the sample was cooling down to room temperature showing Cu(111) $(\sqrt{3} \times \sqrt{3})R30^{\circ}$ - Sb and Cu(111) (2 x 2) – Sb superstructure.

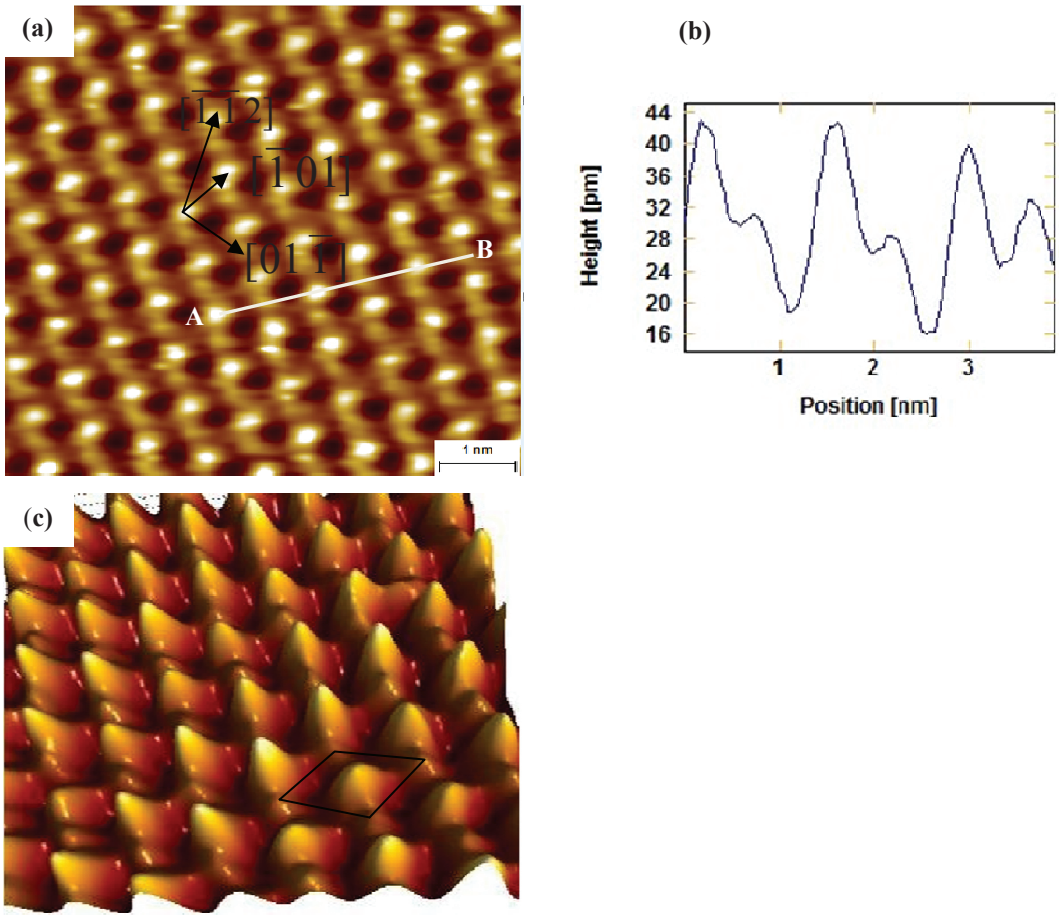


Figure 3. (a) STM image (V_{bias} = 0.7 mV, I_{tun} = 4 nA) CuSb surface after annealing at ~700 °C (6.94 nm × 6.76 nm) forming hexagonal arrangement of a Cu(111)(2 x 2) -Sb. (b) line profile taken along the line in (a), (c) Three-dimensional (3D) view of the STM image.

From a careful examination of the STM image in figure 3(a) and the corresponding line profile in figure 3(b), it possible to distinguish between the Sb and Cu atoms. The interatomic spacing between the Sb atoms is 0.787 ± 0.02 nm. The images reveal only four Cu atoms instead of the six surrounding each Sb atom. The Cu atoms appears to be centered in the <110> - type directions. The line scan in figure 3(c) taken along the line AB in figure 3(a) depicts the relative corrugations between the Sb and Cu atoms.

The $(\sqrt{3}\ \mathrm{x}\ \sqrt{3})R30^\circ$ structure is considered to be the more thermodynamically stable configuration [1-3], thus the Cu(111) (2 x 2) Sb structure in figure 3(a) observed after annealing at high temperatures might be due to kinetic limitations at the surface. At this elevated temperature, the LEED pattern reverted back to that observed in figure 2(e) showing the $(\sqrt{3}\ \mathrm{x}\ \sqrt{3})R30^\circ$ structure. It is highly probable that at this annealing temperature, the Cu atoms desorbed from the surface and desegregate to the bulk to leave a lower copper concentration on the surface. The STM measurements are not enough to distinguish between these possibilities, thus complementary techniques are required for the understanding of the interesting overlayer structure.

4. Conclusion

VT STM was used to characterize the behaviour of Cu(111) following deposition of 0.3 ML of Sb. Annealing of the CuSb film was found to give rise to two distinct ordered surface structures, Cu(111) $(\sqrt{3}\ \mathrm{x}\ \sqrt{3})R30^\circ$-Sb and Cu(111) (2 x 2) -Sb. These observations are in agreement with previous studies on CuSb systems. Annealing at ~700°C produces a structure containing less Sb atoms. The Sb that is lost from the near-surface is believed to diffuse into the bulk of the crystal or alternatively desorbs from the surface. The data clearly favours a structural model based on Sb atoms occupying substitutional rather than overlayer sites within the top Cu layer.

Acknowledgement

This work was supported by the Department of Science and Technology and the National Centre for NanoStructured Materials, CSIR under the HGERA4S project.

References

[1] Giordano H and Aufray B 1996 *Surf. Sci* **352-354** 280-284
[2] Woodruff D and P Robinson J 2000 *J. Phys. Condens. Matter* ***12*** 7699–7704
[3] Noakes T C Q, Hutt D A, McConville C F and Woodruff D P 1997 *Surf. Sci*. **372** 117
[4] Bailey P, Noakes T C Q D and Woodruff D P 1999 *Surf. Sci.* **426** 358
[5] Oppo S Fiorentini V and Scheffler M 1993 *Phys. Rev. Lett*. **71** 2437
[6] Terblans J J and vanWyk G N 2004 *Surf. Int. Anal*. **36** 935–937
[7] Aufray B, Giordano H and Seidman D N 2008 *Surf. Sci*. **447** 180-186
[8] Soares E A, Paniago R, Carvalho V E de, Lopes E L, Abreu G J P and Pfannes H D 2005 *Phys. Rev*. B **73** 035419
[9] Eaglesham D J, Unterwald F C and Jacobson D C 1993 *Phys. Rev. Lett.* **70** 966
[10] Börnstein L 1971 *Springer* Berlin **6**
[11] Meunier I, Gay J M, Lapena L, Aufray B, Oughaddou H, Landemark E, Falkenberg G, Lottermoser L and Johnson R L 1999 *Surf. Sci.* **422** 42
[12] Pacilé D, Ast C R, Papagno M, da Silva C, Moreschini L, Falub M, Seitsonen, Ari P and Grioni M 2006 *Phys. Rev.* B **73** 245429
[13] Ast C R , Henk J, Ernst A, Moreschini L, Falub M, Pacilé D, Bruno P, Kern K and Grioni M 2007 *Phys. Rev. Lett.* **98** 186807

The decoration of vicinal copper polycrystalline surfaces by Antimony

G F Ndlovu[1,2,4], J K O Asante[3], W D Roos[2], K T Hillie[1,2]

[1]NationalCentre for Nano-structured Materials,CSIR, Pretoria, ZA0001
[2]Department of Physics, University of Free State, Bloemfontein, ZA9300.
[3]Department of Physics, Tshwane University of Technology, Pretoria, ZA0001

E-mail: gndlovu@csir.co.za, rooswd@ufs.ac.za

Abstract. An Ultra-high Vacuum Variable Temperature Scanning Tunnelling Microscope was used to study the growth mechanism of Antimony on vicinal Cu polycrystalline samples. The STM data after deposition of 0.3 ML Sb at 300°C showed localization of Sb atoms in the vicinity of the step edges illustrative of the Frank van der Merwe type of growth. Kinks provide for the adsorbates a site with a coordination that is higher than for sites at the straight step edge. Thus, kinks act as efficient nucleation sites for the growth of rows during and after growth.

1. Introduction

Polycrystalline films received a rapidly growing interest in the past decades due to their increasing area of application in advanced technologies [1-6]. The properties of polycrystalline materials are controlled by the structure and interfaces they possess – e.g., grain boundaries, atomic steps and kinks. Growth on steps is actively explored as a promising alternative approach for planar nanostructures on surfaces in addition to lithography-based methods [7]. In this respect, STM with its capability to image conducting surfaces with unprecedented resolution provides a unique possibility to study processes for which surface imperfection plays a key role during and after growth processes. Previous studies have shown that introduction of a surfactant -a suitable adsorbate that remains at the free surface during growth- alters the surface free energy and can thus change the growth mode of a film to achieve layer-by-layer growth in both metal and semiconductor systems [8-13]. Copper alloys are important commercial materials which are often used at temperatures where diffusion processes has a huge influence in their properties. An important characteristic of these alloys is that segregation of one component to the alloy surface causes the surface composition to differ significantly from the bulk [13]. The present paper presents a VT-STM study of ~0.3 ML Sb on Cu polycrystalline surfaces studied at various temperatures.

[4] gndlovu@csir.co.za

2. Experimental

All experiments were conducted in a UHV chamber with a base pressure below 5×10^{-10} Torr, equipped with a STM, SPECTA-LEED, EFM3 evaporator system from Omicron Nanotechnology (Germany, GmbH). The Cu polycrystal polished to 0.25 μm was purchased from MaTeck. The Cu sample was mounted on a Ta baseplate and cleaned by repeated cycles of 2 keV Ar+ ion sputtering at 45° incidence followed by annealing to 530°C until the surface was confirmed clean. Antimony (Sb)was evaporated from an alumina crucible heated resistively in a Knudsen cell. The Sb was evaporated for 3 s at the deposition rates of 0.1 ML s^{-1}. The VT-STM was operated in the constant-current mode utilizing electrochemically etched tungsten (W) tips. The STM data is displayed in a top-view gray-scale representation with darker levels corresponding to lower lying areas. Measurements were acquired at room temperature after two separate annealing temperatures of 500 °C and 700 °C.

3. Results and discussion

3.1. Sb growth on Cu polycrystalline

After deposition of 0.3 ML Sb onto a cleaned Cu polycrystalline surface, a typical STM image acquired over an area of (9.30 x 9.13 nm) exhibits Sb atoms as bright spots on the atomically resolved stepped Cu polycrystalline surface (figure1(a)). Single embedded Sb atoms decorate the monoatomic step edge. The nearest neighbour distance of Sb atoms along the step edge is 0.515 nm ± 0.003 nm. At the top of the terrace in figure 1(a), Sb atoms occupy substitutional sites, where one Sb atom replaces one Cu atom. Given that the STM images the local density of states, Sb (Z = 51) atoms which appear brighter than Cu (Z = 29) atoms can be easily distinguished. The oval shapes (figure 1(a)) indicate the hollow core of screw dislocations.

A careful examination of the screw dislocation on the far left of figure 1(a) shows that the channel is isotropic. The isotropy is the manifestation of the line tension of the steps with respect to their crystallographic orientation [17]. Since steps with different orientations have different line tensions, this observation leads to the conclusion that the scanned area of the crystal was of the same orientation, possibly (111) plane. Based on the STM data of figure 1 it can be seen that the steps are one atom high and atoms join the crystal at the edges of the steps. Thus the crystal grows by spreading existing layers and generating new ones at the screw dislocations. The measured kink length (L) is 0.515 nm and the kink separation (S) is 1.03 nm (figure 1(a)).

The line scan (figure 1(b)) acquired along line A to B in figure 1(a) illustrates the step spacing *d* which varies from ~0.12 nm for the smallest step spacing to 2.5 nm for the larger step width. At a growth temperature of 300 ˚C, the diffusivity of Sb adatoms on the Cu surface increases. As a result, Sb adatoms are likely to reach a step edge within a given time (*t*). If the time between deposition of Sb is larger than *t*, Sb adatoms will typically reach a step edge before another Sb atom is deposited, and thus no Sb nucleation will occur on the surface. Therefore, increasing the temperature move a system from the nucleation-and-growth regime to the step-flow regime as seen in figure 1.

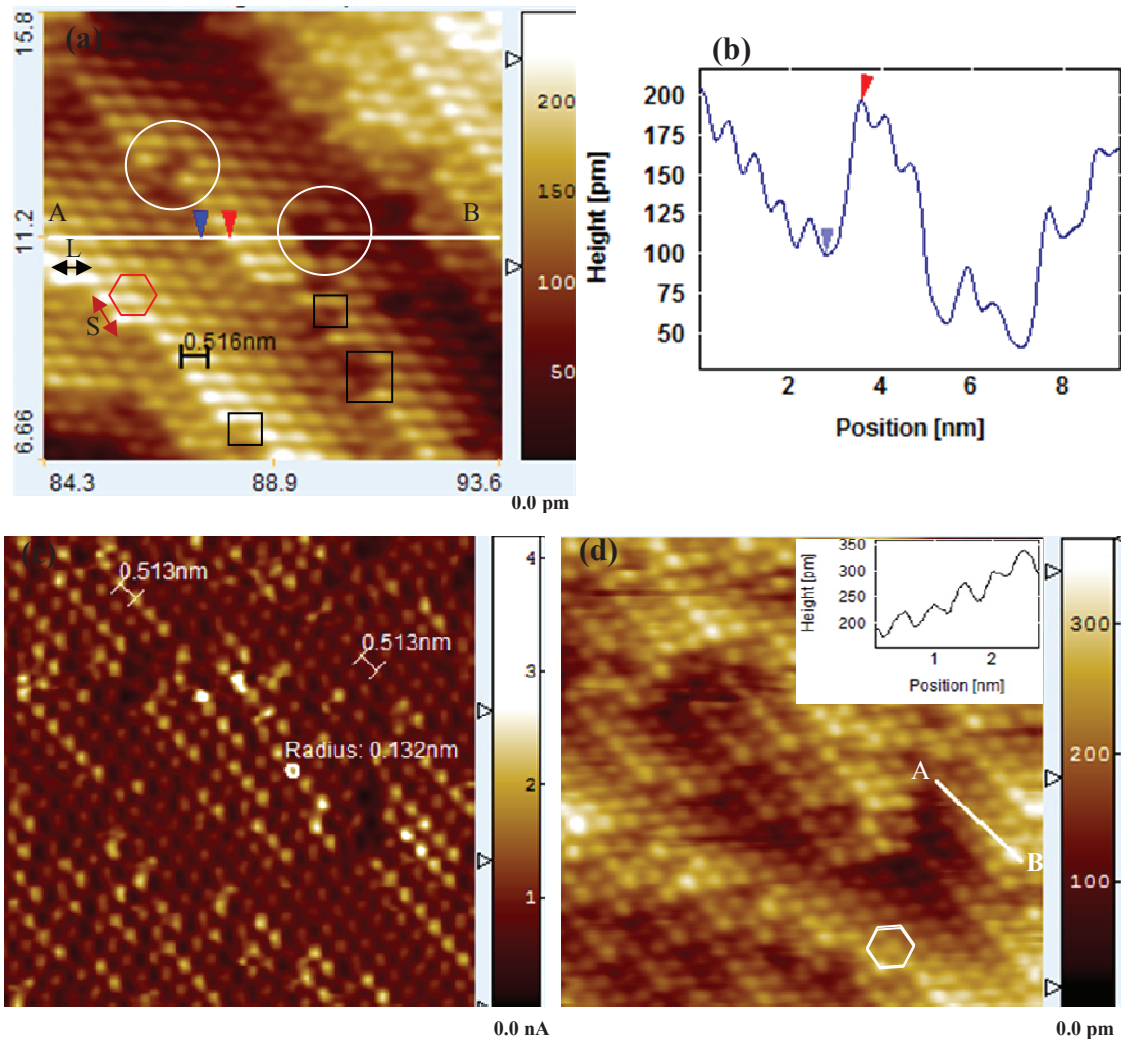


Figure 1. (a) High resolution STM image showing monatomic steps of Cu polycrystalline sample decorated by Sb (-1mV, 2 nA) acquired at an area of 9.30 nm x 9.13 nm after 0.3 ML growth of Sb. (b) Line profile acquired along the line A to B in (a) showing variations of terrace length (step spacing) and step height. (c-d) Constant current STM images (11.3 nm × 11.3 nm) acquired with a positive bias (1mV) The insert in (d) depicts a line scan of the corrugations of a single chain of Sb atoms (from A to B) on a Cu step.

STM images of figure 1(c and d) were acquired simultaneously to illustrate the positions and arrangement of the Sb atoms with respect to the Cu surface (11.3 nm x 11.3 nm). Figure 1(c) represents the variations in tunneling current resulting from both Sb (bright protrusions) and the Cu (darker protrusions) atoms. The Sb atomic radius was measured to be 0.132 nm (compared to the reported 0.133 nm[18]) utilizing the scanning probe image processor software. The measured Sb-Sb distances are comparable to the Cu-Cu atomic distances (0.513 nm). Figure 1(d) is the corresponding height image for figure 1(c). The STM images (figure 1 (c-d)) depict the decoration of atomic steps by a chain of Sb atoms and a large number of kinks are visible. The insert on the top right corner of figure 1(d) shows the corrugation of the about 5 Sb atoms decorating a Cu polycrystalline step.

3.2. Annealing of SbCu poly sample

By increasing the temperature it is well known that the rate of mass transport is substantially increased in regions of low atomic coordination, e.g. at surfaces and in grain boundaries [14-16]. The diffusivity at the surface can be increased by several orders of magnitude as compared with the bulk. The CuSb sample was annealed at 500 ˚C in UHV to study the diffusion characteristics of the Sb atoms on the sample surface. After annealing, successive STM images acquired over an area of (19.3 nm x 19.1 nm) at a positive bias of 0.1 mV show a smooth evolution from very wide terraces with a less smaller number of steps and long step spacing *d*, to short terrace length and increase in step density (figure 2(a)). A closer look of the terraces in figure 2(a) reveals migration of Sb atoms from the step edge to the terrace.

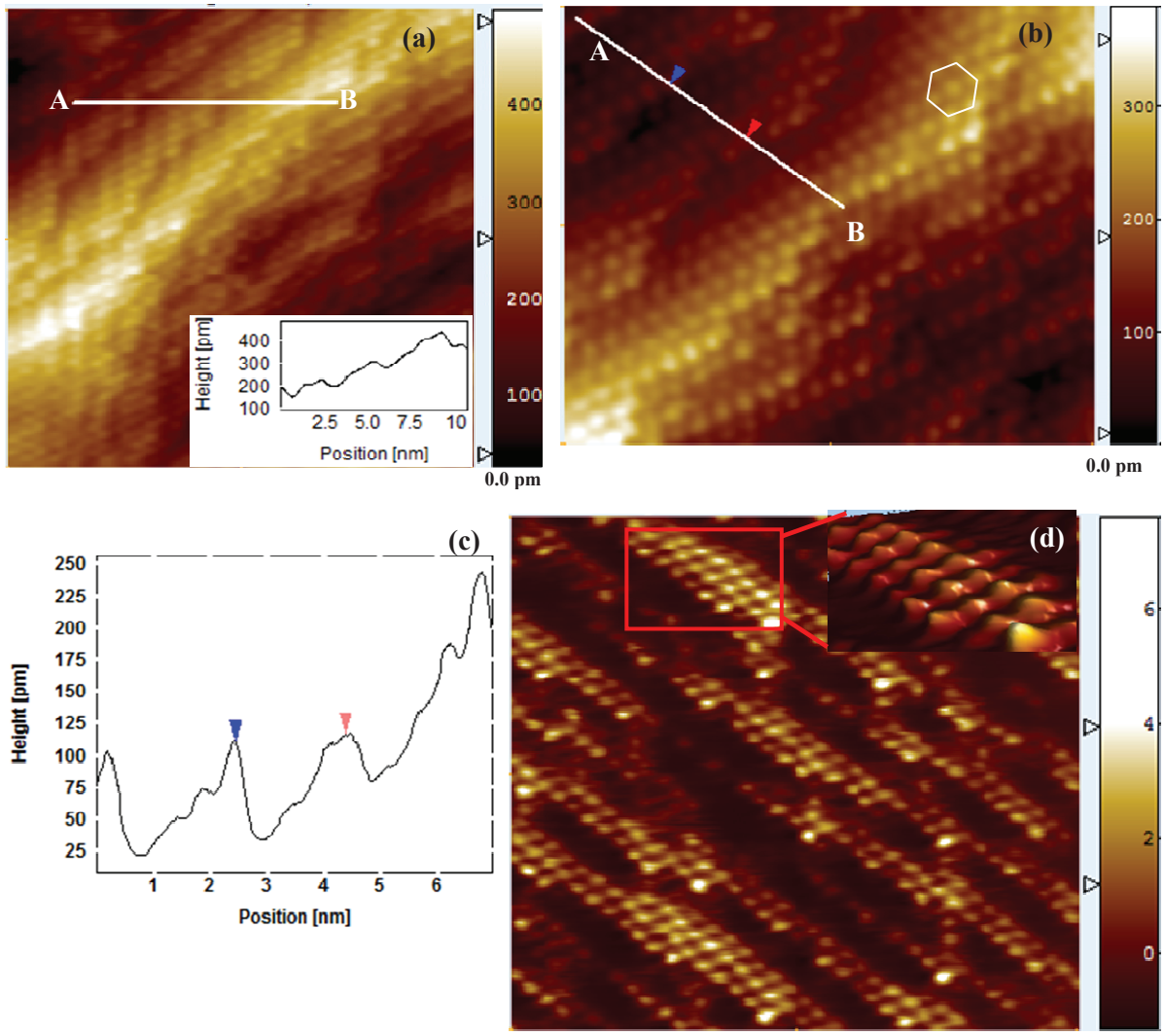


Figure 2. (a) An STM image (19.3 nm x 19.1 nm) of CuSb surface after annealing at 500 °C showing reduced terrace width and increase in step density (V_{bias}: 0.1 mV, I_{tun}: 2.2 nA). The insert on the lower right corner depicts the step height acquired along the line from A to B. (b) The zoomed-in image of (a). (c) shows the line scan acquired in (b) along line AB. The hexagon is superimposed to show atomic arrangement at the terrace. (d) High resolution STM image of the surface after annealing at 700°C for 12 hours acquired over an area of (15.0 nm x 13.6 nm) at a positive bias of 1 mV and tunnelling current of 3nA. The insert show a 3D view of a small portion of the Sb atoms on the surface.

The formation of step bunches in figure 2(a) is a manifestation of instabilities that occur during step generation or growth of Sb. Step density is seen to vary considerably within the same sample after annealing (figure 2(a)). Thus the evolution of the morphology with the increasing annealing temperature exhibits roughening of the sample surface. It can be deduced from figure 2(b) that even the presence of a neighboring step at the shortest possible distance has no effect on the kink creation energy as shown by the increased number of kinks at the surface. The kinks provide for the adsorbate a site with a coordination that is higher than for sites at the straight step edge. This effect can be seen in figure 2(d) after annealing at 700°C. Sb atoms diffused to the terrace covering ~ 4 atomic rows. Even at this high temperature the Sb atoms retain the arrangement of the underlying Cu surface and there is no change in the Sb – Sb atomic distances. The site with the next highest binding energy is a site close to an atom in a kink position, which may in this case efficiently immobilize the Sb adsorbates. In the initial stages of growth, the kink site belong to the substrate, but it soon become part of a growing one-dimensional adsorbate chain of Sb atoms at the Cu step edge.

4. Conclusion

A VT STM Sb was used to study the growth mechanism of Sb on vicinal Cu polycrystalline surfaces. The morphology of the deposited Sb depends strongly on the defect structure of the Cu substrate before, during and after growth. The sample surface showed roughening with increase in temperature. This temperature dependence of the surface morphology reflects the interplay between diffusion, and step incorporation, which are thermally activated processes. Perfect step decoration can be obtained if the potential energy of Sb adatom at the step edge is significantly lower than on the terrace and the temperature should be high enough such that the essential diffusion processes (terrace diffusion, step diffusion, corner rounding) are active. Step decoration of vicinal surfaces can be tailored and exploited to grow for example quantum wires.

Acknowledgement

The authors would like to acknowledge the support from the Department of Science and Technology and the National Centre for NanoStructured Materials, CSIR under the HGERA4S project.

References

[1] Howard J K 1968 *J. Vac.Sci. Technol*. **4** 1
[2] Mhi M Zu, Chen T Q and Judy J H 1992 *IEEE Trans. Magn*. **28** 3255
[3] Jones G A, Srinivasan D K, Grundy P J and Keeler G J 1992 *J.Mat.Sci*. **27**, 6244
[4] Ye X G and Zhu J G 1994 *J.Appl.Phys*. **75**, 6135
[5] Cheng-Zhang L, Lodder J C and Szpunar J A 1994 *J.Magn.Magn.Mater*. **131** 427
[6] L. Bai, J. Tersoff and F Liu 2004 *Phys. Rev. Lett*. **92** 22550
[7] Eaglesham D J, Unterwald F C and Jacobson D C 1993 *Phys. Rev. Lett.* **70** 966
[8] Aufray B, Giordano H and Seidman D N 2000 *Surf. Sci* **447** 180-186
[9] Snyder C W and Orr B G 1993 *Phys. Rev. Lett*. **70**1030
[10] van Gastel R et al 2001 *Phys. Rev. Lett*. **86** 1562
[11] Darby T P and Balluffi R W 1977 *Phil. Mag*, **36** 53
[12] Sorensen M R, Mishin Y and Voter A F 2000 *Phys. Rev. B* **62** 3658
[13] Woodruff D 2002 *Surface Alloys and Alloy Surfaces,* Elsevier
[14] Copel M, Reuter M C, Horn von Hoegen M and Tromp R M 1990 *Phys. Rev. B* **42** 11682
[15] Horn von Hoegen M. et al. 1993 *Surf. Sci*. **24** 53
[16] Lapujoulade J 1994 *Surf. Sci. Rep.* **20** 191
[17] Adaulian I, Kirilka S,Harald I and Margret G 2007 *Phys. Rev. B* **75**, 245411
[18] JohnsonW R 2007 *Atomic Structure Theory* Springer Berlin

Enhancement of Pr^{3+} red emission by adding In^{3+} as co-dopant in $CaTiO_3:Pr^{3+}$ phosphor

L.L. Noto*, S.S Pitale, M.M. Gusowki, J.J Terblans, O.M. Ntwaeaborwa and H.C. Swart*

Physics Department, University of the Free State, P. O. Box 339, Bloemfontein, 9300, South Africa

*Corresponding author: NotoLL@ufs.ac.za or SwartHC@ufs.ac.za

Abstract. A single red emission peak of calcium titanate doped with praseodymium ions is observed around 613 nm. Photoluminescence (PL) enhancement of $CaTiO_3:Pr^{3+}$ was obtained by the addition of In^{3+} as a co-dopant. The orthorhombic structure of the phosphor was confirmed by X-ray diffraction spectroscopy, and the PL properties were investigated using a PL spectrometer. The maximum concentration of Pr^{3+} ions that is necessary to obtain the optimal PL intensity was 0.2 mol% and 0.1 mol% for In^{3+} ions.

1. Introduction

The red emitting orthorhombic perovskite (figure 1a) calcium titanate ($CaTiO_3$) doped with trivalent praseodymium (Pr^{3+}) ions, is known to have a single narrow red emission peak at 613 nm coming from the $^1D_2 \rightarrow {}^3H_4$ transition of Pr^{3+}[1,2]. It is promoted by the complete quenching of the $^3P_0$ level (figure 1b) through intervalence charge transfer state (IVCT) induced by orbital overlapping of Pr^{3+} and Ti^{4+} ions inside the $CaTiO_3$ matrix [1,2].

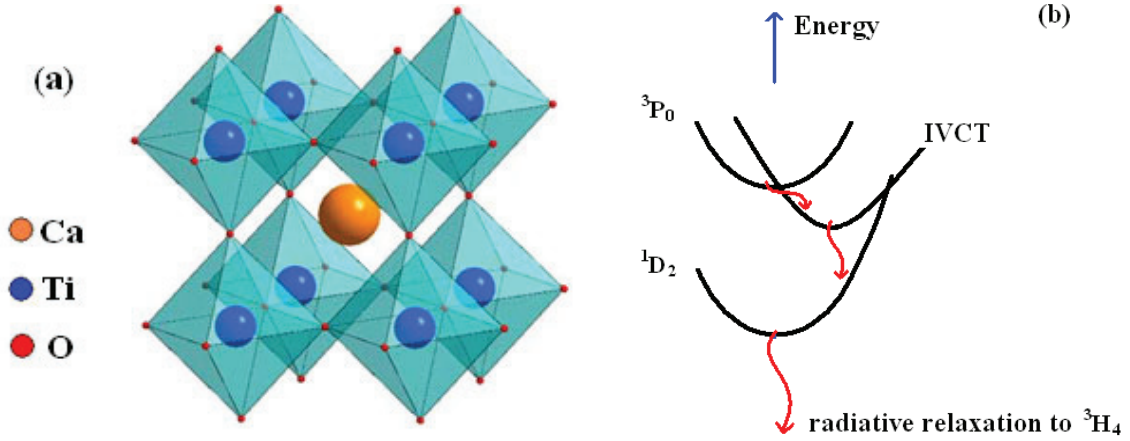


Figure 1: Orthorhombic $CaTiO_3$ crystal structure (a) and the IVCT model (b) [7]

This phosphor material has been around for some time, and the challenging goal for researchers is to enhance its emission intensity to make it suitable for application purposes. This problem has been tackled in various ways, such as preparing $CaTiO_3:Pr^{3+}$ via different synthetic routes, and through charge compensation using different metal cations [3,4,7]. In the present work, we present the enhancement of $CaTiO_3:Pr^{3+}$ through charge compensation, using In^{3+} cations.

2. Experiments

$CaTiO_3$: Pr^{3+} phosphors co-doped with different In^{3+} concentrations were prepared by solid state reaction at 1200 ^oC for 4h, to fabricate $CaTiO_3$. Activation was achieved by doping the matrix with Pr^{3+} ions from $PrCl_3$ and charge compensation by adding In^{3+} ions from $In(NO_3)_3$. The final composition can be related by the following stochiometric equation:

The phase was identified using a Bruker, AXS D8 Advance X-ray diffractometer (XRD). The luminescence properties were investigated using a Varian Carry-Eclipse fluorescent spectrometer for Photoluminescence (PL) measurements, PerkinElmer Lambda 950 UV/VIS spectrometer for diffuse reflectance.

3. Results and Discussion

3.1. Structure

The phase formation of the synthesized $CaTiO_3:Pr^{3+}$ phosphors co-doped with different In^{3+} concentrations are shown in XRD patterns (figure 2), which match the well known orthorhombic perovskite structure of $CaTiO_3$ phase described by the JCPDS card no. 22-0153. According to these patterns, full phases were formed for each material and the effects of Pr^{3+} doping and In^{3+} co-doping into the matrixes had a negligible effect.

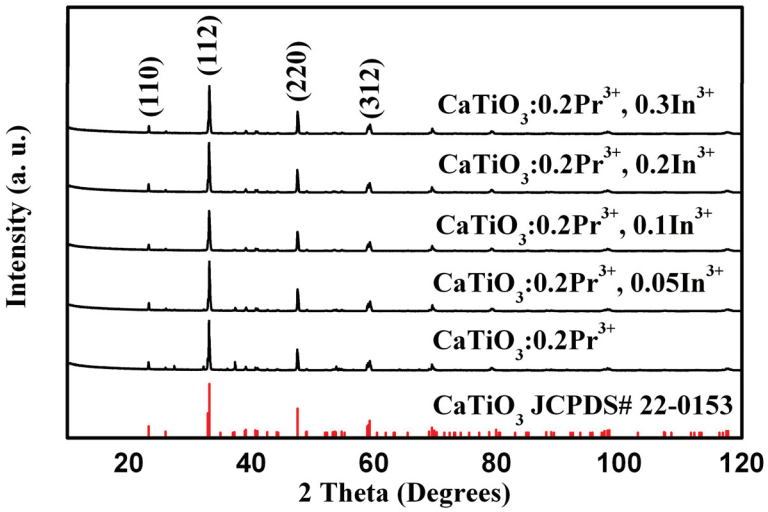


Figure 2: XRD pattern of $CaTiO_3:Pr^{3+}$ with different In^{3+} concentrations as indicated.

3.2. Pr^{3+} emission enhancement

Substitution of the trivalent Pr^{3+} ions in the site of the divalent Ca^{2+} ions brings about charge imbalance, by generating more positive charge in the system. This forces the matrix to compensate for such by generating positive O vacancies, and negative Ti and Ca vacancies that act as luminescence

quenching centers. At higher Pr^{3+} concentrations there are more negative vacancies being formed and the distance between neighboring Pr^{3+} ions becomes smaller. This leads to more quenching centers than at lower concentration, and the shortened distance between neighboring Pr^{3+} increases the chances of the two Pr^{3+} ions to trap an electron between them. The trapped carrier between Pr^{3+} ions will then non-radiatively de-excite to ground state because of phonon interaction. Figure 3a&b, show spectra of $CaTiO_3$ doped with different Pr^{3+} concentrations, and were measured using a medium photon multiplier tube (PMT) voltage. The optimal concentration is obtained at 0.2 mol%, and for concentrations higher than 0.2 mol% the luminescence lowers. This is attributed to concentration quenching [3,4].

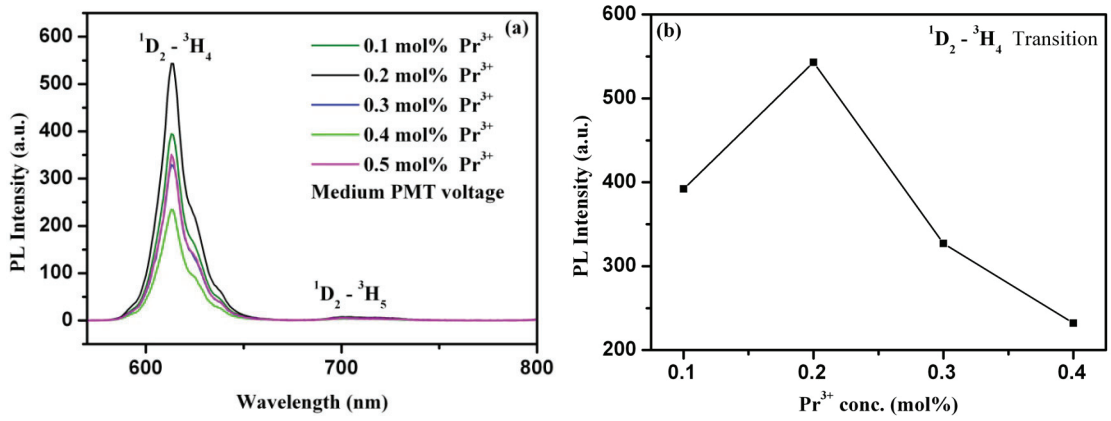


Figure 3: (a) PL spectra of $CaTiO_3$ doped with different Pr^{3+} concentration and (b) PL intensity as a function of Pr^{3+} concentration.

The 613 nm single red emission of Pr^{3+} ion doped in $CaTiO_3$ (figure 3a & 4a) is related to the intervalence charge transfer state (IVCT) that provides an efficient quenching channel for the non – radiative depopulation of ^3P_o level by a cross over to the $^1D_2$ level at room temperatures, and this results the observed single red emission from $^1D_2 \rightarrow {}^3H_4$ transition [5].

3.3. Charge compensation

$CaTiO_3$ self compensates the charge imbalance by generating negatively charged calcium vacancies, by reducing Ti^{4+} to Ti^{3+}, and also by generating positively charged oxygen vacancies. The Ti^{3+} state induces the d → d transition that acts by re-absorbing the population of electrons from the excited states of Pr^{3+}, which later de-excite non-radiatively to the valence band, and thereby acting as a luminescence quenching center. Addition of charge compensators to $CaTiO_3$:Pr^{3+} acts by promoting energy transfer to the excited states of Pr^{3+} from the host material. When In^{3+} ions are co-doped in $CaTiO_3$:Pr^{3+}, they substitute in the sites of Ti^{4+}, and assume a role of reducing the number of local sites where Pr^{3+} incorporation in matrix generates an additional positive charge. This has an effect of reducing the density of calcium defect related luminescence quenching centers, and thus reinforce the number of electrons that fall straight to Pr^{3+} luminescence centers.

Hence addition of In^{3+} cations enhances the PL intensity (figure 4a&b) of $CaTiO_3$:Pr^{3+} [1,4,5,6,8,9]. The spectra of In^{3+} co-doped $CaTiO_3$:Pr^{3+} (figure 4a) were measured using low PMT voltage because the emission intensity resulting from In^{3+} saturates the detector at high PMT voltage.

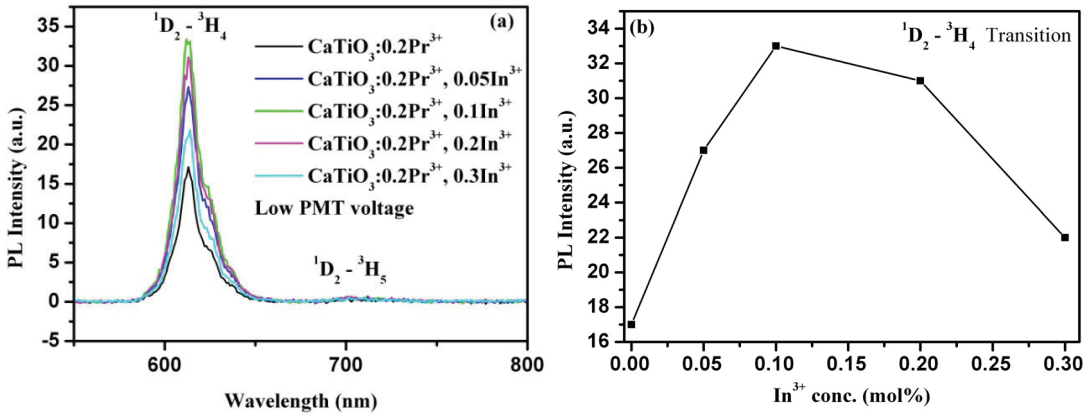


Figure 4: (a) PL spectra of $CaTiO_3$:0.2Pr^{3+} doped with different In^{3+} concentration and (b) PL intensity as a function of In^{3+} concentration.

4. Conclusion

$CaTiO_3$:Pr^{3+}, In^{3+} with a single red narrow emission peak at 613 nm was synthesized using solid state reaction at 1200 ^oC for 4h. Addition of In^{3+} in $CaTiO_3$:0.2Pr^{3+} showed significant enhancement in the emission of the phosphor. The optimal doping mole percentage of In^{3+} was found to be 0.1 mol%.

Acknowledgements

The authors send gratitude to the National Research Foundation (NRF) for funding the project and the University of Free State Physics (Physics department, Center for microscopy and Geology department) for the research techniques used in this study.

References

[1] Diallo PT, Jeanlouis K, Boutinaud P, Mahiou , Cousseins JC, 2001 *J.Alloys and Compd.*, **218** 323
[2] Boutinaud P, Sarakha L, Cavalli E, Bettinelli M, Dorenbos P, Mahiou R, 2009 *J.Phys. D: Appl. Phys.*, **42** 045106
[3] Shengyu Yin, Donghua Chen, Wanjun Tang, 2007 *J.Alloy and Compd.*, **441** 327
[4] Tang Wanun, Chen Donghua, 2009 *Materials Research Bulletin*, **44** 836
[5] Phillipe Boutinaud, Rachid Mahiou, Enrico Cavalli, Marco Bettinelli, 2006 *Chem. Phys. Lett.*, **418** 185
[6] Shun Li, Xixi Liang, 2008 *J Mater Sci: Mater Electron*, **19** 1147
[7] Bautinaud P, Pinel E, Oubah M. Mahiou R, Cavalli E dn Betinelli M, 2004 *J. Opt. Mater.*, **210** 1016
[8] Xianmin Zhang, Jiahua Zhang, Xia Chen, Shaozhe Lu, Xiao-Jun Wong, 2007 *J. Lumin.* **122** 958
[9] Wanjun Tang and Douglas Chen, 2009 *Phys. Status Solidi A* **206 No. 2** 229

Computational study of rutile tin-oxide (SnO_2)

JN Ntimane, TE Mosuang[1], KE Rammutla[1]

1. Department of Physics and Geology, University of Limpopo, Private Bag X 1106, Sovenga, 0727, Polokwane, South Africa

thuto.mosuang@ul.ac.za

Abstract. Rutile structured tin-oxide ceramics have been intensively studied in recent years because of their potential in sensing and fuel cells. The present work uses classical molecular dynamics simulations focused on the structure and possible transformation of rutile tin-oxide to other phases. The empirical Buckingham potential has been used to describe the interatomic interactions in tin-oxide. The total energy of the NPT hoover ensemble at various temperatures has been calculated in order to determine the transition temperature and pressure. The results obtained showed an energy increase with temperature which was constantly compared with experiments. The radial distribution functions for the two structures suggest the transformations at temperature above 900 K in agreement with the experiments.

1. Introduction

Metal-oxides like tin-oxide (SnO_2) are increasingly being used in the chemical, pharmaceutical, ceramic, and electronic industries. These materials can be synthesized using chemical and physical methods. The chemical methods include the sol-gel, whereas the physical methods involve vapour and pulsed laser deposition. In both methods, the production of qualitative materials plays a great role and is influenced by the interatomic interactions and the thermodynamics environment of the material. Molecular dynamics (MD) is often used together with these experimental methods to obtain the reliable method of synthesis.

SnO_2 is an attractive *n*-type semiconducting material with a wide band gap. It has outstanding applications due to its structural, electronic, optical, and chemical properties [1]. SnO_2 can be applied in catalytic reactions, in optoelectronic devices, in gas sensing, as an electrode for flat panel displays, and in solar cells. Under normal conditions SnO_2 crystallizes in tetragonal rutile structure, which belongs to the $P4_{2/mnm}$ space group. Experiments and calculations have shown that under high different pressures SnO_2 transforms from rutile to pyrite, to fluorite and then cotunnite structures [1,2].

This works intends to explore rutile SnO_2 and its kinetics through high pressure, high temperature environments. This can be attained by exploring the energy-temperature relations and the radial distribution functions effects in this material. It is apparent from Alvarez and Valladares [3] that the atomic topology also determines the electronic properties of the structured materials, and therefore any understanding of the RDF's and the atomic distribution is relevant in the characterization of the structural and optical properties of tin-oxide. In the process, information about the nanoparticle grain size and morphology necessary for the sensor application could be obtained. Sverjensky [4] has theorized that the bulk crystal and interfacial dielectric constant play an important role in controlling whether hydrated ions adsorb directly onto the oxide surface or form electrostatically adsorbed outer-sphere complexes, which is crucial in gas sensing.

In this work, structural and thermodynamic properties of rutile SnO_2, using the MD simulation based on the Buckingham empirical interatomic potential is being investigated. The RDF's of rutile SnO_2 are also being investigated, specifically different temperature Sn-Sn, Sn-O, and O-O pair distributions.

2. Computational details

The DL_POLY package [5] has been used to perform all the empirical bond-order molecular dynamics calculations of SnO_2. A supercell with a 4.55 Å cutoff, 480 atoms, and a sufficiently large number grid points for the fast Fourier transformations (kmax1 = 11, kmax2 = 11, and kmax3 = 6) has been used throughout the calculations. The ewald convergence parameter of 0.6537, on a Noose-Hoover NPT ensemble allowing the simulation supercell to change has been applied. The thermostat relaxation was set at 0.1 whilst the barostat was at 0.5. The simulation was allowed to run for more than 100 000 steps, with a simulation step of 0.001 ps. The controlled experimental crystal structure for bulk SnO_2 is according to Bolzan *et al*. [6]. The material as is used for the MD modeling is described by its lattice parameters as listed in Table 1, and a set of parameters required for the Buckingham potential are taken from AV Bandura *et al*. [7] and P Amstrong *et al*. [8].

Table 1 Lattice parameters and relative sites for anions and cations in a rutile tin-oxide crystal lattice [8].

$a = b = 4.7372$ (Å)		$c = 3.1863$ (Å)
$u = 0.306$		
anions:	$(u, u, 0)$,	$(-u, -u, 0)$,
	$(0.5 + u, 0.5 - u, 0.5)$	$(0.5 - u, 0.5 + u, 0.5)$
cations:	$(0, 0, 0)$	$(0.5, 0.5, 0.5)$

3. Results and discussion

A bulk SnO_2 in the rutile form was studied with the MD code as described in the previous section. In order to check the stability of rutile SnO_2 structure at various temperatures within the Buckingham potential, the pair distribution functions of rutile SnO_2 at 400 K and 900 K temperatures are presented in Figure 1 and 2. From the peak positions, the most probable distances between the various atoms can be determined. In the 400 K temperature the peaks appear around 1.85, 2.15, and 3.65 Å, respectively for the Sn-O, O-O, and Sn-Sn bond-lengths, whereas at 900 K the peaks appear around 1.85, 2.65, and a double peak around 3.85, 3.90 Å for the very same bond-lengths. There are obvious peak shift at 900 K for O-O and Sn-Sn bonds and a step within a Sn-O peak at g(r) = 5 suggest that it might also shift to the right. It can also be noted that the averaged rdf's for Sn-Sn, Sn-O, and O-O at r > 4.5 Å, is normalizing towards unity. This suggests a possible phase transformation in agreement with the first principle calculations of Yanlu Li *et al* [1] together with Fan and Reid [2] measurements.

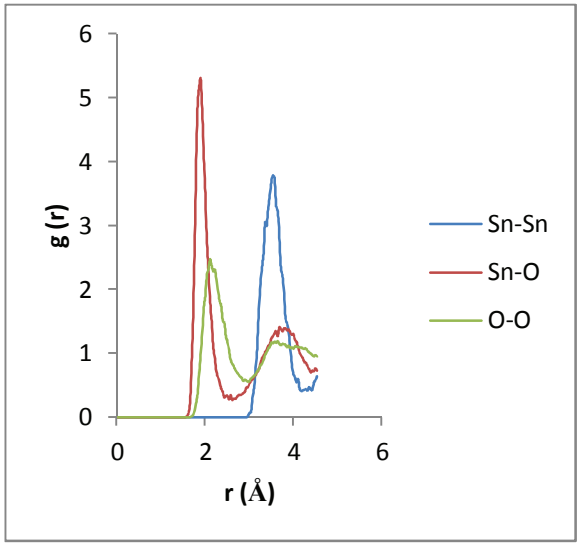


Figure 1 Pair correlation function for rutile tin-oxide at 400 K.

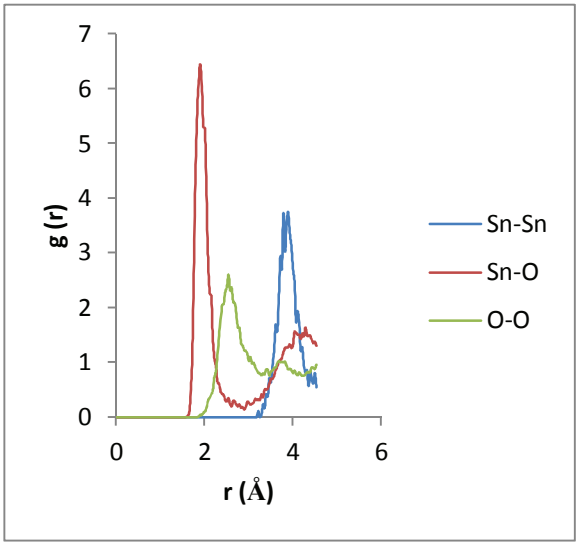


Figure 2 Pair correlation function for rutile tin-oxide at 900 K.

Figure 3 shows the volume of rutile SnO_2 as a function of temperature. The simulation data was fitted to the polynomial functions of temperature as

$$V(T) = 1.0 \times 10^{-4}T^2 + 0.0554T + 5963.4.$$

The thermal expansion coefficient is then calculated from the temperature derivative of the volume,

$$\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right).$$

The volume thermal expansion coefficient is determined from the above equation to be 19.4 x 10^{-6} K^{-1}. Peercy and Morosin [9], have measured the volume thermal expansion of SnO_2 from 93 to 700 K using Raman spectroscopy to be 11.7 x 10^{-6} K^{-1}. It has also been noted that the volume-temperature data for the rutile SnO_2 is quite disordered. A huge difference could be associated with the anisotropy of the rutile structure along the c-axis and the proposed charge distribution among the Sn and O atoms. Although in Figure 3 the behaviour of the graph is rather turbulent, it can be seen that the steady slump followed by a gradual rise around1000 K suggest a possible phase shift in accordance with the pair correlation plots. The two suggestions are in agreement with the Fan and Reid [2] experiments using the Alvarez and Valladares [3] approach.

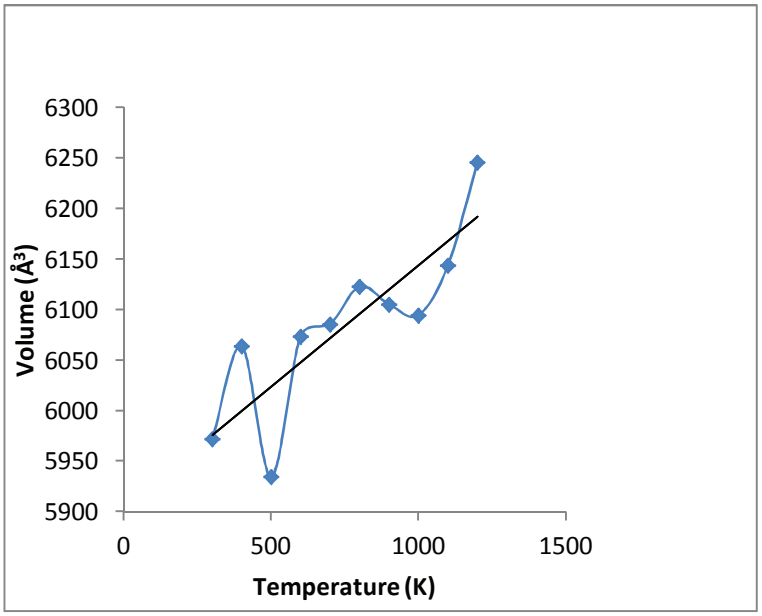


Figure 3 The rutile tin-oxide volume as a function of temperature.

4. Conclusions

The validity of the Buckingham empirical potential for rutile SnO_2 has been investigated using the structural and pair distribution analysis. Both the volume-temperature and the pair distribution functions suggest a possible phase transformation at high temperatures. According to the volume-temperature graph (Figure 3) the phase transformation is most probable nearby 1000 K. The ionic nature of SnO_2 should be considered in future. The Sn-O bond peak experiences a positive shift from 1.85 Å in the 900 K temperature region, despite the evident Sn-Sn and O-O bond peak shift. This is in agreement with the calculations if atomic charges are not considered [7].

Acknowledgements

NRF is thanked for their financial support. CHPC is thanked for their computational facilities.

References

[1] Yanlu Li *et al.* 2010 *Journal of Physical Chemistry* A **114** 1052
[2] H Fan and SA Reid 2003 *Chemistry of Materials* **15** 564
[3] F Alvarez and AA Valladares 17 Oct 2001 *arXiv:cond-mat*/0110332v**1** [*cond-mat.mtrl-sci*]
[4] DA Sverjensky 1993 *Nature* **364** 776
[5] W Smith, TR Forester, and IT Todorov 2009 *The DL_POLY 2 User Manual*, STFC Daresbury Laboratory, Daresbury, Warrington WA4 4AD Cheshire, UK
[6] AA Bolzan, C Fong, BJ Kennedy, and CJ Howard 1997 *Acta Crystallographica* **53** 373
[7] AV Bandura, JO Sofo, and JD Kubicki 2006 *Journal of Physical Chemistry* B **110** 8386
[8] P Amstrong, C Knieke, M Mackovic, G Frank, A Hartmaier, M Göken, and W Peukert 2003 *Acta Materialia* **57** 3060
[9] PS Peercy and D Morosin 1973 *Physical Review* B **7** 2779

Microstructure and Spectral selectivity of Pt-Al_2O_3 nanocoatings for high temperature applications

Z Y Nuru[1-2], S Khamlich[1,] K. Roro[1], T F G Muller[2], C J Arendse[2] & M Maaza[1-3]
[1]Nanosciences African Network, MRD-iThemba Labs, National Research Foundation, Old Faure road, Somerset West, South Africa
[2]Dept. of Physics, University of Western Cape, Private Bag x 17, Belleville, 17, Belleville, South Africa
[3]Faculty of Sciences, Pretoria-Tshwane University of Technology, Private Bag X 680, Pretoria, South Africa.
E-mail: zebib@tlabs.ac.za

Abstract. In terms of high photo-thermal conversion efficiency and thermal stability, multilayered structures based on metal-dielectric composites i.e. ceramic-metal (cermets) are the most attractive candidates for receivers operating at high temperature (above 400°C). Pt-Al_2O_3 cermet nano-composites are a representative family as per their high temperature chemical inertness and stability. This contribution reports on the microstructure and optical property of Pt-Al_2O_3 cermet nano-composites deposited in a multilayered tandem structure. The radio-frequency sputtered Pt-Al_2O_3 cermet consisting of stainless steel substrate/ Mo coating layer/ Pt-Al_2O_3 absorbing layer /Al_2O_3 protective layer and stainless steel substrate/ Mo coating layer /Pt-Al_2O_3 absorbing layer for different composition and thickness of Pt-Al_2O_3 coatings. This optimized coating exhibits high solar absorptance (α ~0.95/0.93) and low thermal emittance (ϵ ~0.11/0.17 at 89 °C) with and without Al_20_3 protecting layer, respectively, which are stable up to 650°C in air. X-ray diffraction, atomic force microscopy, effective medium theory and UV-VIS-NIR total reflectance were used to characterize the microstructure, morphology, theoretical modeling and optical property of these coatings.

1. Introduction

Solar collectors are environmentally friendly and non-polluting technological way of converting solar radiations into thermal energy to fulfil the ever the current increasing demand of energy supply [1]. Basically, a solar collector's photo-thermal conversion efficiency depends on the solar selective absorbing coating. A variety of selective absorbers have been proposed and fabricated for photo-thermal applications [1-3]. Selective absorbing surfaces can be classified into three main types from a structural point of view: (i) intrinsic, (ii) optical trapping surface and (iii) tandem. Tandem selective surfaces can further be classified into three: (i) semiconductor-metal tandems (ii) multilayered absorbers, and (iii) metal – dielectric composite tandem. An intrinsic solar absorbing material is wavelength selective. However, there is no naturally occurring absorbing material. Optical trapping surfaces can produce high solar absorptance by multiple reflections among them one could quote needle-like, dendritic, or porous microstructures. Tandem absorbers consist of at least two layers with different optical properties: Semiconductor-metal tandems absorb short wavelength radiation because of the semiconductor band gap and have low thermal emittance as a result of the metal layer. Multilayered absorbers use multiple reflections between layers to absorb light and can be tailored to be efficient selective absorbers. Metal-dielectric composites called cermets type of solar absorbers exhibit high absorption as well as high reflection in the UV-VIS and NIR respectively over a large solar

spectrum range. In addition, when such a composite cermet-coating is formed on a highly reflecting metal surface, the resulting tandem coating has a good spectral selectivity [4-7]. Concerning the fabrication of such solar absorbers, several techniques such as electroplating, physical vapour deposition, chemical vapour deposition and so on, have been developed to produce spectrally selective absorbing surfaces. However, among these sputtering approach is one of physical vapour deposition which is the most commonly and widely used [4-12]. The optical properties of such composites can be intermediated between those of metals and of the dielectric. Effective dielectric permeability of the composite, which can be related to the constituents in the effective medium theories, can be used to estimate the optical performance of the cermet. This contribution reports on the optimization and further selectivity improvement of radio-frequency sputtered graded Pt-Al_2O_3 deposited onto a Mo base layer exhibiting a high solar absorptance and low thermal emittance with a significant thermal stability up to 650^{o}C in air with and without antireflection layers.

2. Experiments and synthesis

Al_2O_3 disc "~13 cm in Ø" with circular small Pt pellets "~5 mm in Ø" placed on it was used as a target. The Pt pellets were placed in a hexagonal array on the Al_2O_3 disc target to ensure an isotropic deposition of Pt and Al_2O_3. The composition of the films was varied according to the number of Pt pellets used. The optimized working pressure was fixed to ~10^{-2} Torr without heating the samples' stainless steel substrates as was substantiated by early studies on pure Pt-Al_2O_3 samples [7-11]. The optimized synthesized cermet samples were characterized from morphological, crystallographic and optical view points by using scanning electron microscopy "SEM", atomic force microscopy "AFM", X-rays diffraction "XRD". Total reflectance was collected by an integrated sphere system, that includes both specular and non-specular components, was measured with a double beam Vary Cary 500 spectrophotometer to calculate how Pt-Al_2O_3 cermet films absorb light using an AM1.5 and a direct measurement of emittance *using* an emissometer model AE1 which has an accuracy of about ± 0.01 emittance units was used to measure the emissivity of the devices.

3. Results and discussions

The optimized Pt concentration in the Pt-Al_2O_3 cermet layer i.e. in terms of the Pt filling factor "*f* value", was deduced by modeling using the Brugeman effective medium approximation. The Pt-Al_2O_3 cermet layer was treated as an isotropic inhomogeneous medium with a random mixture of metallic Pt nano-particles in the host dielectric matrix with a filling factor *f* and dielectric constants ϵ_{pt} and ϵ_{Al2O3} as a homogeneous medium with an effective dielectric constant ϵ_{eff} given by:

$$\epsilon = f\left(\frac{\epsilon_{pt} - \epsilon_{eff}}{\epsilon_{pt} + 2\epsilon_{eff}}\right) + (1-f)\left(\frac{\epsilon_{Al2O3} - \epsilon_{eff}}{\epsilon_{Al2O3} + 2\,\epsilon_{eff}}\right) \quad (1)$$

Using the tabulated values of ϵ_{Pt} and ϵ_{Al2O3} from the standard Palik's database, the preliminary modeling calculations allowed the determination of the value of the Pt filling factor "*f*" at about ~0.34 and the effective optical constants (refractive index, n and extinction coefficient, k) given by:

$$n^2 = \left(\frac{\sqrt{\epsilon_{pt}{}^2 + \epsilon_{Al2O3}{}^2}}{2} + \frac{\epsilon_{pt}}{2}\right) \quad (2)$$

and

$$k^2 = \left(\frac{\sqrt{\epsilon_{pt}{}^2 + \epsilon_{Al2O3}{}^2}}{2} - \frac{\epsilon_{pt}}{2}\right) \quad (3)$$

As reported in Figure 1, the simulation demonstrates a prominent feature of the optical constants calculated from the above equation that illustrated how the Pt-Al_2O_3 composite changed the optical constants from the constituents: Pt and Al_2O_3. Moreover, it indicates that the effective refractive index was greater than the extinction coefficient while both of them were positive. This indicates that, in the UV-VIS-NIR spectral region, light was not rejected but absorbed. In case of equivalent coatings of pure Pt and Al_2O_3, the light would be reflected and transmitted respectively. The optimized Pt-Al_2O_3 thin film thickness was found to be about ~70 nm with a Mo buffer layer of ~150 nm onto the considered stainless steel substrates.

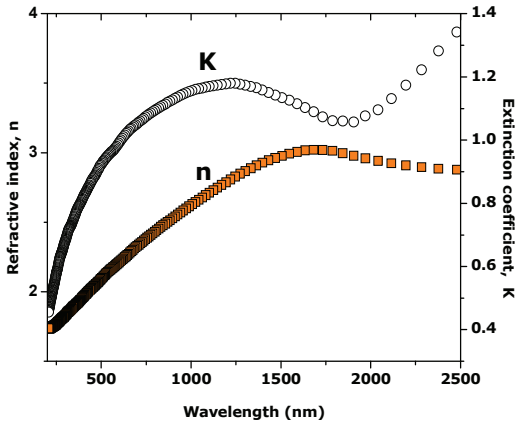


Figure 1. Optical constants of Pt-Al_2O_3 composite calculated using Bruggeman model. The Pt filling factor in the Pt-Al_2O_3 cermet coatings was ~0.34.

Figure 2 reports a typical SEM surface morphology of Pt-Al_2O_3 thin film/Mo buffer layer/stainless steel substrate with f~0.34. The corresponding cermet film exhibits a tortuous surface morphology with Pt nano-particles distributed isotropically in the basal plane indicated by the arrows. Their average diameter and inter-particles distance were statistically about 4-6 nm and 7-10 nm respectively.

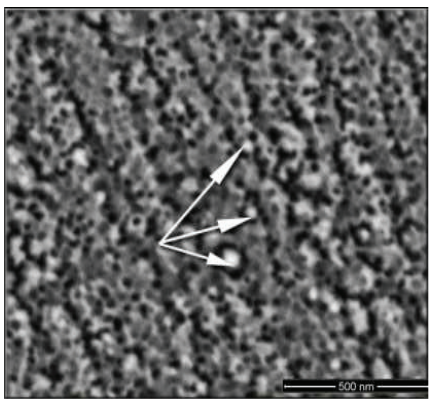


Figure 2. Scanning electron microscopy of the optimized cermet sample: anti-reflecting layer of ~ 110 nm Al_2O_3/ ~70 nm Pt-Al_2O_3/~150 nm Mo/~0.5 mm Stainless steel substrate.

Figure 3 depicts a representative AFM surface scanning of the cermet sample's surface exhibits two type of surface topographies; highly disordered and semi-ordered stripes type zones. These latter regions of semi-disordered "stripes" have an average length of ~0.41 nm consisting of 1-D chains-like of length of about ~0.41 nm. This 1-D chains-like were spatially ordered and consisted of crystallites with an average diameter of about 250 nm. The crystallites, in the disordered regions have approximately an identical average size. The average roughness value was of about 8.81 nm. Relatively speaking, this roughness value was comparable to the average diameter and inter-particles distance which were found to be about 4-6 nm and 7-10 nm respectively. Hence, one could deduce that the surface topography was controlled by the Pt nano-particles.

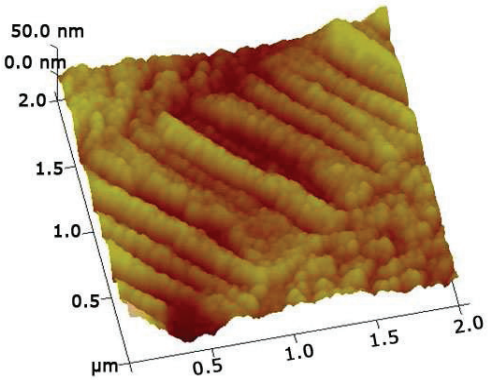


Figure 3 . Atomic force microscopy topography of the optimized cermet sample: anti-reflecting layer of ~110 nm Al_20_3/~70 nm Pt-Al_2O_3/~150 nm Mo/~0.5 mm Stainless steel substrate.

Figure 4 reports the crystallographic orientations of the optimized Al_2O_3/Pt-Al_2O_3/Mo/Stainless steel. Taking into account the anti-reflecting layer of Al_20_3, cermet layer of Pt-Al_2O_3 as well as the buffer IR reflective metallic layer of Mo is thin; the probing X-rays impinging the samples do penetrate and reach the stainless steel substrate. Indeed as shown by Figure 4, while the stainless steel substrate was highly crystalline with a net (111), (200) and (220) texturing, the buffer IR reflecting layer of Mo as well as the host Al_2O_3 matrix seemed to be fully amorphous as there was no corresponding Bragg peaks in the recorded 2Θ angular range of 30-80 deg. In contrast, and within the same angular range, one observes the (111), (200) and (220) Pt Bragg peaks. These peaks were quite wide corresponding to Pt grains with an average size "using the Debye-Scherer approximation" of about 3.2-4.7 nm in diameter hence corroborating with the values deduced from the SEM and AFM investigations within the error bars.

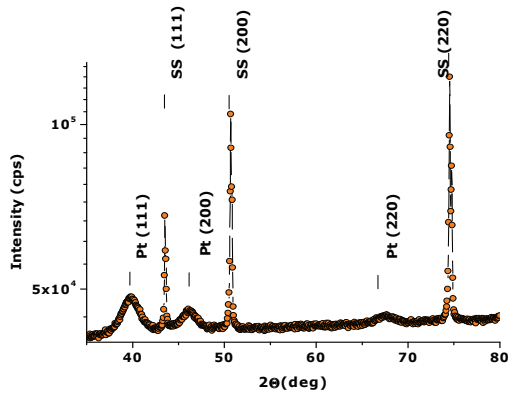
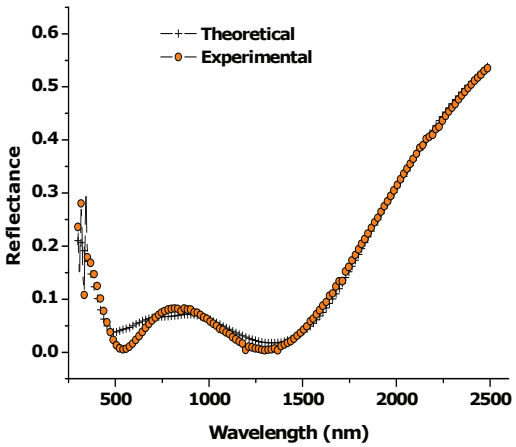


Figure 4. X rays diffraction patterns of the optimized cermet samples with and without the anti-reflecting layer of ~110 nm Al_2O_3/~70 nm Pt-Al_2O_3/~150 nm Mo/~0.5 mm Stainless steel substrate.

The optical reflectance spectra of simulated and experimental Pt-Al_2O_3 composites in the UV-VIS-NIR range of 300-2500 nm in particular for the Pt- Al_2O_3/Mo/Stainless steel substrate tandems without and with Al_2O_3 antireflection layer are shown in figure 5 (a) and (b) respectively. Hence the tandem without the Al_2O_3 anti-reflecting layer exhibited an absorptance of 0.93 and thermal emittance of 0.17, the corresponding one of the tandem with the Al_2O_3 anti-reflecting layer obtained an absorptance of 0.95 and thermal emittance of 0.11. The average total reflectance in the first spectral region was less than 10% it was decreased to 5% in the second one. These very low total reflectance values, which demonstrate the feasibility of the current considered optimized cermet based tandems, should be attributed not only to the optimized nature of the tandems with or without the Al_2O_3 reflecting layer but to the specific surface morphology and texture of these coatings as well.



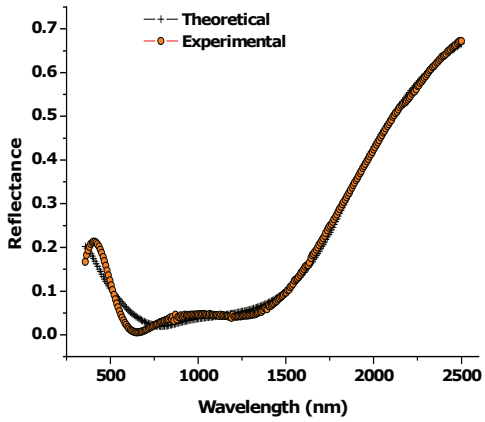


Figure 5. a) Experimental and calculated total Reflectivities "specular + non sepcular" of the optimized cermet samples without anti-reflection layer of ~70 nm Pt- Al_2O_3/~150 nm Mo/~0.5 mm Stainless steel substrate.

b) Experimental and calculated total Reflectivities "specular+ non sepcular" of the optimized cermet samples with anti-reflecting layer of ~110 nm Al_2O_3/~70 nm Pt- Al_2O_3/~150 nm Mo/~0.5 mm Stainless steel substrate.

4. Conclusion

In this work, optimized multilayered Pt-Al_2O_3 cermet nano-coatings onto Mo IR reflecting buffer layer coated stainless steel substrate with and without anti-reflecting layers of Al_2O_3 exhibited an excellent spectral selectivity of 0.95/0.11 which was consistent with a computer simulation. These cermet films are found to be thermally stable up to 650 ^oC in air. Combining such superior optical properties and the refractory nature of alumina as well as the chemical inertness of Pt, this type of coatings are competitive candidates for high temperature solar power plant applications.

Acknowledgment

We would like to thank for the finance support from (i) the Organization of Women Scientists for Developing World-Trieste "OWSDW", the Abdus Salam ICTP-Trieste, the Nanosciences African Network "NANOAFNET"-Cape Town, as well as iThemba LABS-National Research Foundation of South Africa.

References

[1] Granqvist C G 1991 J. Appl. Phys. A **52** 83
[2] Granqvist C G and Niklasson G A 1978 Phys. Rev. B **18** 554
[3] Niklasson G A and C G Granqvist 1984 J. Appl. Phys. **55** 3382
[4] Sella C and Lafait J 1988 J. Appl. Surf. Sci. **33/34** 942
[5]. Kenndy C E 2002 NREL/TP-520-31267
[6]. Thronton J A, Penfold A S and Lamb J L 1980 Thin Solid Films **72** 101
[7]. Berthier S, Lafait J, Sella C and Vien T K 1985 Thin Solid Films **125** 171
[8]. Maaza M, Nemraoui O, Sella C , Lafait J, Gibaud A, Pischedda V 2005 J. Phys. Letters A **344**
[9]. Sella C, Maaza M, Pardo B, Dunsteter F and Sainte Catherine M C 1997 J. Phys. A **241** 192
[10]. Gibaud A, Sella C, Maaza M, Sung L, Dura J A, Satija S.K 1999 Thin Solid Films **340** 153
[11]. Sella C, Maaza M, Pardo B, Dunsteter F, Martin J C and Kaba A 1997 J. Surface and Coatings Technology **97** 603
[12]. Maaza M, Nemraoui O, Sella C, Lafait J, Gibaud A, Baruch-Barak A and Beye A C 2006 J. Solid State Comms. **137** 166

Reactive DC magnetron sputter deposition and characterization of ZrN thin films

T G Nyawo[1] and M O Ndwandwe[1]

[1] Department of Physics and Engineering, University of Zululand, Private Bag X1001, KwaDlangezwa, 3886, South Africa.

E-mail: tnyawo@pan.uzulu.ac.za

Abstract. ZrN films were deposited on Si<100> substrates, Al, brass strips and glass by DC magnetron sputtering under varying conditions of power, pressure, argon and nitrogen gas flow rates as well as temperature, and were characterized by SEM, RBS and resonant RBS. The films were transparent and semiconducting. They tend to absorb oxygen. The films were found to adhere well to the substrates. The colours of the films varied depending on deposition conditions and have been expressed in the colour space. Potential uses of such films are as protective hard coatings as well as decorative layers.

1. Introduction

The interest in the nitrides and carbides of the transition metal elements such as titanium, hafnium and zirconium is due to their extreme physical and chemical properties. Some of these desirable properties are a high melting point, hardness and good corrosion resistance [1]. ZrN has been named as a possible inert nitride matrix material to burn plutonium or transmute long-lived actinides in accelerator driven submicron nuclear systems or fast nuclear reactors [2]. Because of the properties of some transition metal nitrides, they also have been used as field emitters [3]. Required properties for an emitter material are chemical stability, a high melting point, low vapour pressure, low work function, no deformation by ion bombardment and relative ease for making the material into needle-like arrays [3]. It has been suggested that ZrN films may be used as barrier layers for oxygen [4]. ZrN has been found by industry to make good decorative films [6]. Due to their extreme wear resistance they have been used as hard-coatings to expand the lifetimes of mechanical compounds in cutting tools and dies. Their gold-like colour makes them excellent candidates to compete with gold in coatings [6].

The most common method of depositing ZrN films is by DC magnetron sputtering. However other methods have also been used. Some of these are: plasma nitridation, vacuum arc deposition, ion-beam assisted deposition, ion plating, pulsed laser deposition, cathodic arc evaporation and high vacuum sputtering [7]. It has been reported that ZrN has higher contamination susceptibility when compared for example to other nitrides such as TiN. There have therefore been few studies of ZrN when compared to other transition metal nitrides. In this paper we report the growth and characterization of ZrN films on silicon, aluminum and brass strips as well as on Cu on Si, Al and brass. We also study the uptake of oxygen by ZrN films using resonant Rutherford Backscattering Spectrometry (RBS), a technique that has not been applied to ZrN before.

2. Experimental procedure

Silicon wafers, brass or aluminum strips were chemically cleaned using methanol, followed by acetone, then trichloroethylene, then acetone, and finally methanol. These were then rinsed in de-ionized water and dried in air. The samples were loaded in AJA's Orion 5 Sputtering System where a thin layer of ZrN was deposited on the substrates. The chamber was evacuated to a base pressure of $2x10^{-6}$ Torr. The sputtering target used was a 2" Zr metal. Argon was used as a process (sputtering) gas. Nitrogen was introduced to act as a reactive gas so as to form ZrN. It is clear that for relatively

high magnetron power (a lot of Zr atoms will be produced from the target) and a low N_2 flow rate, a metallic film will form. For relatively low magnetron power and higher flow rates of N_2, a reactive mode will be reached and a compound ZrN_x (x>0) will form. It must be noted however that very low power means fewer Zr atoms sputtered and lower growth rates for the film. The design of the sputtering system is such that there is a valve between the chamber and the evacuation pumps (turbo pump and rotary pump). Running the sputtering system in a throttling mode (partially closing the valve between chamber and pumps) allows for lower inflow rates of both process and reactive gases. It also saves the pumps from overworking (leads to longer pump lives).

3. Results and discussion

First we used Rutherford Backscattering Spectrometry (RBS) in order to find layer thicknesses as well as to check whether the correct stoichiometry between Zr and N had been achieved. The films were deposited either at room temperature or higher temperatures (~ 200 ^oC) on a rotating Si wafer to ensure uniformity of the film. DC magnetron power applied to the Zr target was varied between 100 W and 200 W and the argon flow rate used was varied between 2 sccm and 8 sccm. Nitrogen was introduced next to the sample at values varying from 2 sccm to 20 sccm. The duration of deposition varied from 30 minutes to 2 hrs. After deposition samples were kept in air under dry conditions.

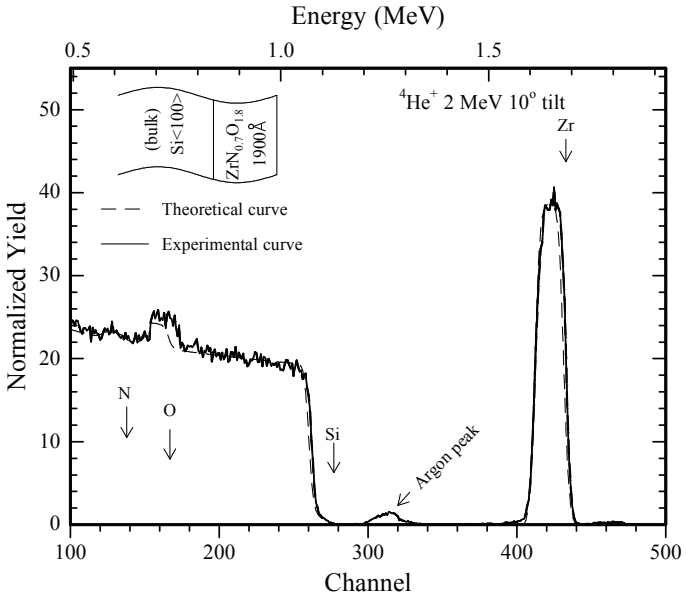


Figure 1. BS spectrum of a sample having configuration Si<100>/ZrN. According to the results of RBS analysis the atomic percentages are Zr = 28.6%, N = 20% and O = 51.4%. A simulation curve is also shown.

Figure 1 shows the RBS spectrum of a sample Oxygen made of Si<100> (bulk)/ZrN. The deposition conditions for sample O were a DC magnetron power of 200 W, argon flow rate of 4 sccm and nitrogen flow of 8 sccm. Deposition duration was 45 minutes. The ZrN layer was found by means of RBS characterization to be 1 900 Å. Surface positions of Zr, Si, O and N are shown on the figure. The peak shown at channel 315 corresponds to that of Ar (Argon was used as a process gas to sputter a Zr target in a nitrogen rich environment). The sample was found to have 28.6% zirconium, 20% nitrogen and 51.4% oxygen. Checking the literature we found that it is normal for sputtered ZrN films to have that much oxygen. The film was found to be well adherent to the silicon wafer. This was deduced from strip test measurements. We also looked at the film under a scanning electron microscope (SEM). The surface of the film appeared smooth and featureless. SEM results are not

shown here. The colour of the film appeared uniform to the eye –also implying a uniform thickness. We also used Energy Dispersive X-rays (EDX) to analyze the sample. The elements found agreed with that obtained from RBS measurements. Rutherford Universal Manipulation Program (RUMP) simulation showed that there was a significant amount of oxygen on all the samples. We decided to measure the amount of oxygen incorporated in ZrN much more accurately. The presence of oxygen had been mentioned in the literature but it had not been measured accurately.

The two samples A and B, shown in figure 2, were prepared under almost similar conditions. They were both deposited by magnetron sputtering from a Zr target using power of 200 W for a period of 45 minutes. They differ only in gas flow rates during deposition. The flow rate of argon during the deposition of sample A was 2 sccm while that of nitrogen was 8 sccm. The flow rate of argon was 2 sccm during the deposition of sample B, while that of nitrogen was 20 sccm. The chamber pressure during deposition was maintained at $3x10^{-3}$ Torr in both cases. One expects the amount of Zr deposited in the two cases to be the same since the same magnetron power of 200 W was used.

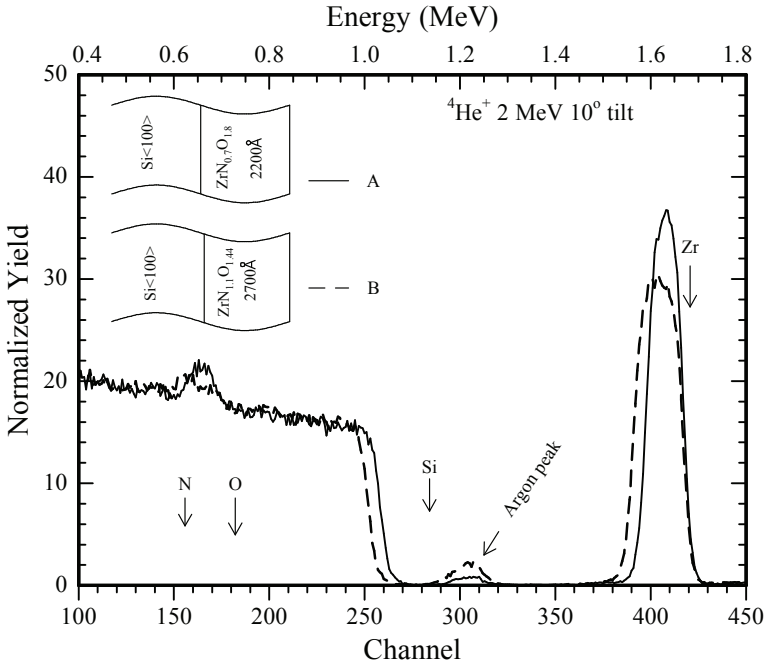


Figure 2. RBS spectra of sample A and B that were prepared under almost similar conditions. The only difference is the nitrogen flow rates (8 sccm for sample A and 20 sccm for sample B). The samples differed markedly in colour.

Figure 2 shows the spectra of both samples (A and B). Surface positions for elements Zr, Si, O and N are marked clearly by means of arrows on the figure. These are channels (or energy values) at which one expects the alpha particle energy when it has been backscattered from those elements, if the elements are on the surface of the sample. We integrated the area under the curve of the Zr signal in both cases (A & B) and found that the two areas are equal. This confirms that the same amount of Zr was deposited in each case. The thickness of the film in sample A was found from RUMP simulation to be 2 200 Å while that in sample B was 2 700 Å. The difference in the samples therefore is in their composition. RBS analysis shows the composition of sample A to be Zr = 28.6%, N =20% and O = 51.4%. Sample B has be found to be made up of Zr = 28.2%, N = 31.1% and O = 40.7 %. Sample B had a goldish colour while A looked greenish. It has been decided to measure the oxygen content of the samples much more accurately. To solve problem of oxygen resonant RBS at 3.05 MeV was done.

Resonant $^{16}O(^{4}He, ^{4}He)^{16}O$ was performed in a van De Graaf accelerator using He^{2+} ions with a kinetic energy of 3.05 MeV at a scattering angle of 165^{o} and a sample tilt of 10^{o} (to avoid channeling effects). The spectra collected from samples A and B are shown in figure 3.

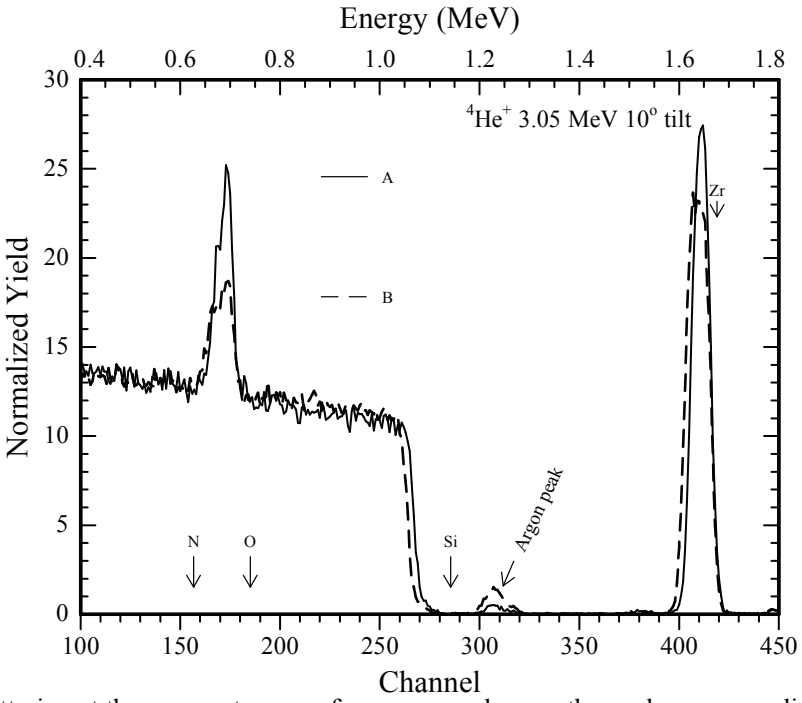


Figure 3. Scattering at the resonant energy for oxygen enhances the peak corresponding to oxygen. The amount of oxygen in sample B is 38% and the amount of oxygen in sample A is 62%. Sample B appears golden in colour whereas sample A appears bluish in colour.

It will be noticed that the spectra has shifted to higher energies (i.e. the energies of the backscattered α particles are higher since higher beam energy has now been used). The intensity of the peaks is generally lower than in figure 2. The peaks have shifted to high energies due to the fact that the surface barrier detectors used are less efficient at higher energies. Note however that the peak intensity of oxygen has increased. At energy of 3.05 MeV the scattering cross section for alpha particles scattering on oxygen is larger resulting in a higher yield, thus giving rise to a larger peak. Unfortunately one cannot use RUMP for the analysis of the spectra as it has no inbuilt ability for analysing a resonance at this energy. We can however compare the amounts of oxygen in the two samples by integrating the area under the curve of oxygen. The area under the oxygen curve for sample B is 62% that of sample A (i.e. sample B has 38% less oxygen when compared to sample A).

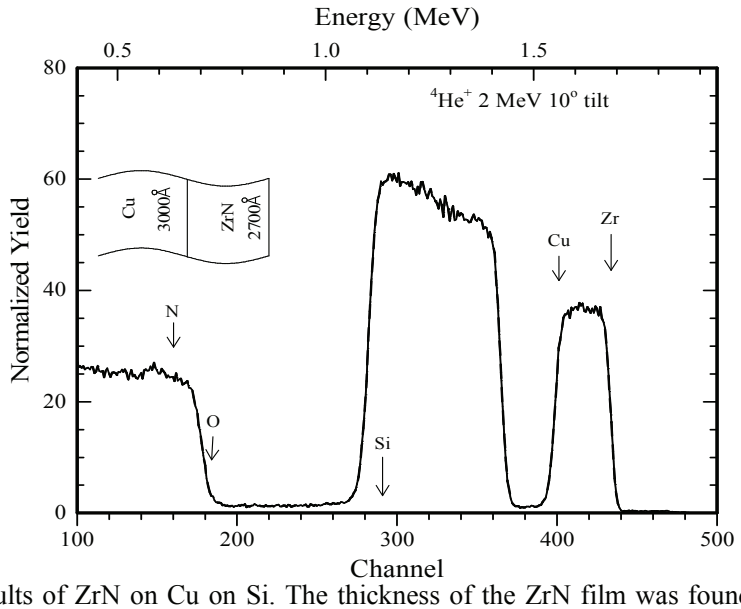


Figure 4. RBS results of ZrN on Cu on Si. The thickness of the ZrN film was found to be 2700 Å, while the thickness of Cu was 3000 Å. ZrN can therefore be used as a protective coating against corrosion. Even though ZrN films absorb oxygen, they however do not pass that oxygen to the under layer film**.**

4. ZrN as a protective and decorative coating

The ZrN film looked flat and featureless under the Scanning Electron Microscope (SEM results not shown here). There were no bumps or voids on the film that could have indicated failure of the film to adhere or bond to the silicon. Adhesion was further tested by means of a tape stripping. This also showed that the film had bonded well to the substrate. We also deposited Cu on two Al metal strips. We then deposited ZrN on top of the Cu film on Al. We wanted to find out whether the ZrN film will protect the Cu film from corrosion. Within two hours corrosion spots appeared on the Cu film not covered with ZrN. The film covered with ZrN showed no corrosion even after exposing it to air for three months.

The colour of the ZrN films changed depending on the deposition conditions. The colours of ZrN/Si<100> films were calibrated with a Konica Minolta colour spectrophotometer. The relationship in colour differences are given by colour space in terms of the following notation L*a*b*, where L* indicate lightness of the sample while a* and b* are chromaticity coordinates which are determining colour directions. Both figures 5a and 5b present the spectral reflectance vs. wavelength graphs of ZrN/Si<100> films together with measurements of their colour space. The sample B in figure 5a displayed a golden colour with the following values of colour difference L* = 2.82, a* = -0.06 and b* = -1.31. Figure 5a shows that at 350 nm wavelengths it reflected most of the light. As soon as the wavelength increased the reflectance drop until reflecting nothing because all incoming light were absorbed at wavelength around 420 nm to 440 nm. At around 450 nm the reflectance started to rise again and reached maximum peak at 600 nm. The sample A in figure 5b displayed a sky blue colour with the following corresponding values of colour difference values L* = 1.81, a* = -0.66 and b* = -1.08. Figure 5b shows reflectance that is proportional to the wavelength up until the maximum peak of reflectance was reached at wavelength of around 450 nm. After passing this peak immediately drops down if the wavelength is increasing further.

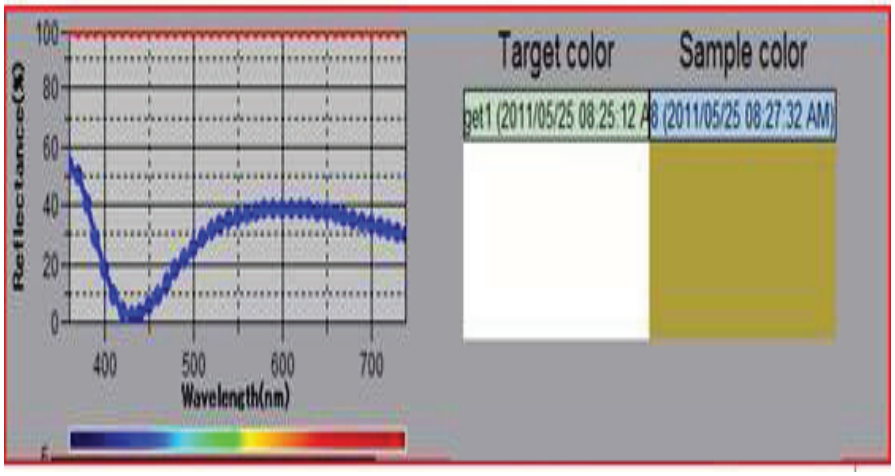


Figure 5a. Reflectance colour spectrum on sample B. It is noted that the maximum of the reflectance is at about 600 nm and the greatest absorption occurs at 450 nm wavelength. Sample B appears golden in colour.

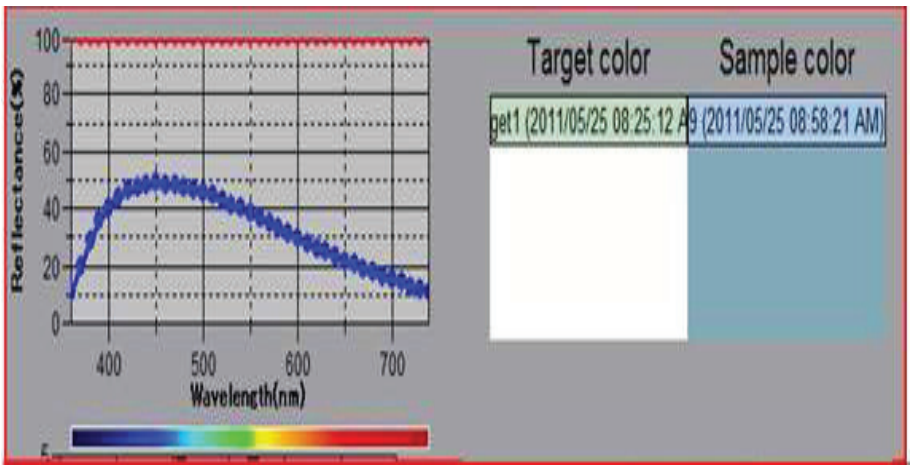


Figure 5b. The colour of sample A appeared bluish. We have compared this to a white colour standard. The graph shows a reflectance spectrum in the visible region. The maximum reflectance for this film is at about 450 nm wavelength.

5. Conclusion.

Films of ZrN we deposited on various substrates (S_i, Al, brass, glass) and characterized from a Zr target in an argon nitrogen atmosphere at a pressure of $3x10^{-3}$ torr. The samples were characterized by means of SEM, RBS, and resonant RBS colour spectrometer as well as by strip test (for adhesion). The surfaces of the films appeared smooth under SEM investigation and no bubble or peeling was observed. It was found therefore that the ZrN films adhere well on all substrates tested. We also found that the ZrN films take up oxygen. To better characterize the oxygen uptake under different process conditions resonant RBS was carried out at 3.05 MeV. This helped to enhance the oxygen signal and allowed us to compare amounts of oxygen in different samples. Samples with less oxygen and ratio of Zr to N of close to 50% appeared golden in colour. We analyzed the colour differences using a colour spectrometer. The values of colour difference of ZrN/Si<100> films were found in terms of L*a*b*.

We also found that ZrN forms a very good barrier to oxygen and may therefore be used to prevent corrosion by oxidation.

Acknowledgements:
We thank the (SA) National Research Foundation for financial assistance through project GUN of 2010. We also thank iThemba LABS for their hospitality and the use of their Ion Beam Analysis equipment.

References:

[1] Pichon L, Girardeau T, Straboni A, Lignon F, Guerin P, Perriere J 1999 *Applied Surface Science* **150** 115-124.
[2] Streit M and Ingold F 2005 *Journal of the European Ceramic Society*, **25** 2687-92.
[3] Endo M Nakane H and Adach H 1996 *Applied Surface Science*, **94/95** 113-16.
[4] Al-Shareef H N, Chen X, Lichtenwalner D J and Kingon A I 1996 *Thin Solid Films* **280** 265-70.
[5] Subramanian B, Ashok K, Sanjeeviraja C, Kuppsami P and Jayachandran M 2008 *Journal of Physics: Conference Series* **114** 012039.
[6] Niyomsoan S, Grant W, Olson D L and Mishra B 2002 *Thin Solid Films*, **415** 187-94.
[7] Huang J H, Ho C H and Yu G P 2007 *Materials Chemistry and Physics* **102** 31-8.

Structure-property relationship of sol-gel synthesised zinc-oxide nanoparticles

N V Peterson[1], C J Arendse[1,*], K T Hillie[2] and M Dhlamini[3]

[1]Department of Physics, University of the Western Cape, Private Bag X17, Bellville 7535, South Africa,

[2]DST/CSIR National Centre for Nano-Structured Materials, P. O. Box 395, Pretoria 0001, South Africa,

[3]Department of Physics, University of South Africa, P. O. Box 392, Muckleneuk Ridge UNISA 0003, South Africa

*cjarendse@uwc.ac.za

Abstract. Zinc-oxide nanoparticles are well known for their novel optical and electronic properties for applications in various fields such as solar cells, ultra violet shielding, gas sensors, paint and heat mirrors. We report on the relation between the structure and optical properties of ZnO nanoparticles synthesized via the sol-gel technique, with specific emphasis on the effect of growth and reaction temperatures. High-resolution microscopy techniques, complemented by x-ray diffraction, confirm that the crystallinity and particle size of ZnO nanoparticles is directly related to the synthesis conditions. Optical absorption and emission spectroscopy show that optical band gap and photoluminescence of the ZnO nanoparticles are intimately related to its structural properties, ascribed to the Ostwald ripening process. Photoluminescence spectroscopy confirm the emission peaks in the ultraviolet (380 nm) and visible (500 nm) region; the latter attributed to the presence of the singly ionized oxygen vacancies in the nanoparticle.

1. Introduction

ZnO is an important II-VI compound semiconductor with a wide direct band (3.37 eV) and a large excitation binding energy (60 meV) at room temperature [1 – 5]. Room temperature photoluminescence of ZnO typically exhibits dual emissions in the near UV and visible range. The UV near band edge emission is at approximately 370 nm, which is a broad UV emission that is sharp and highly intense. The other visible emission is in the range from 450 – 730 nm [6, 7] and it is relatively weak and broad. The broad emission is due to the existence of surface defects and the band edge emissions are found to be more stable for most of the II-IV semiconductors [7].

ZnO is widely used in various applications such as UV- photodetectors, bulk acoustic wave resonators, gas sensors, solar cells, laser diodes and optical waveguides [8 – 10]. There are several methods reported for the synthesis of ZnO nanoparticles, which include chemical vapor deposition [11], gas-phase method [12], spray pyrolysis, hydrothermal synthesis, micro emulsion, electrochemical method [13], pulsed laser deposition, microwave synthesis and the sol gel method [14, 15]. Recently, a variety of morphologies,

including thin films, nanospheres, nanorods, nanowires, nanobelts, nanotubes, nanoflowers and nanowhiskers [16 – 20], have been prepared by some of the methods above. ZnO nanoparticles are receiving worldwide interest because of the novel optical and electrical properties for various applications [21]. ZnO nanoparticles, produced by the sol gel method, show excellent optical properties in which good control of the size and morphology of the particles is achieved. This method costs is very low and it is reliable, repeatable and simple [4].

In this paper, the sol gel method is used to synthesis ZnO nanoparticles. Investigations of the effect of growth temperature and reaction temperature on the particles size, the morphology, crystallinity, and optical emission and absorption properties were carried out by using x-ray diffraction, transmission electron microscopy, photoluminescence spectroscopy and ultraviolet-visible spectroscopy.

2. Experiment

All chemicals were purchased from the supplier and were used without further purification. In the typical synthesis, 0.46g of zinc acetate (Zn $(CH_3CO_2)_2$) was dissolved in 30ml of ethanol (CH_3CH_2OH) under vigorous stirring at a temperature of 75 – 80°C for 1 hour. 0.22g of sodium hydroxide (NaOH) was dissolved in 10 ml of CH_3CH_2OH in the ultrasonic bath. Both clear solutions were placed in a water bath with respective temperature in the range 0 – 60^0C for 15 minutes. This temperature is hereafter referred to as the *reaction temperature*. The NaOH solution was added dropwise to the Zn $(CH_3CO_2)_2$ solution under vigorous stirring at temperatures from 0 – 60^0C. This temperature is hereafter referred to as the *growth temperature*. The solution was stirred for an extra 30 minutes. The solution was stored for 24 hours. Purification for the as-prepared ZnO nanoparticles was firstly washed in ethanol. Heptane ($CH_3C_5H_{10}CH_3$) was further added to wash the ZnO nanoparticles. White ZnO nanoparticles precipitated after centrifugation at 4000 rpm for 15 minutes and the unwanted ions were removed. The ZnO precipitate was redispersed in heptane using the ultrasonic bath. The above procedure of washing the ZnO nanoparticles in heptane was repeated three times. The ZnO precipitate was suspended in ethanol and it was dried in the oven at a temperature of 90^0C for 2 hours [22]. Thereafter the required analytical techniques were performed.

The crystallinity and the particles sizes of the ZnO nanoparticles were measured in reflection geometry of the PANalytical X'pert pro PW3040/60 X-Ray diffractometer. The system was operated with a CuKα (λ = 1.54060 Å) monochromatic radiation source at voltage 45.0 kV and the current at 40.0 mA. The XRD pattern was collected in the step scan mode with 2θ-values ranging from 10.0 to 120.0° with a step size of 0.026°. The internal structure, particle size and crystallinity of the ZnO nanoparticles were recorded on a JEOL 2100 high-resolution transmission electron microscopy (HR-TEM) operated at 200 kV accelerating voltage. The optical transmission spectra of ZnO nanoparticles were recorded using the Perkin-Elmer LAMDA 750S UV-VIS spectrophotometer in the range from 200 – 800 nm with a spectral resolution of 1 nm. The emission properties were performed using a Perkin-Elmer LS2 Fluorescence Spectrometer in the range from 340 – 800 with an excitation source of 320 nm.

3. Results and Discussions

3.1. Structural Properties

The XRD patterns of the ZnO nanoparticles synthesized at different reaction and growth temperatures are shown in figure 1. Several preferential orientations in the <100>, <002>, <101>, <102>, <110>, <103>, <112> and <202> directions are observed, which match well with the hexagonal wurtzite structure of ZnO, corresponding with Joint Committee for Powder Diffraction Standards (JCPDS) File no. 36- 1451 [23]. No other diffraction peaks for the reagents are observed, illustrating the enhanced purity of the ZnO nanoparticles. The (101) diffraction peak appears to be dominant in all spectra, implying that the majority of crystallites have a preferred growth along the (101) plane. After quantitative analysis of the (101)

diffraction peak, the interplanar spacing (*d*), lattice constants (*a* and *c*) and particle size were calculated for all spectra (not shown). From figure 1, it is clear that with increasing temperature the diffraction peaks become sharper and stronger, which imply that the crystallinity of the nanoparticles improves. The results show a systematic increase in the particle size with an increase in both the reaction and growth temperature from 3.6 – 4.9 nm and 3.5 – 4.0 nm, respectively. The *d*-spacing, and hence the lattice constants, remains relatively constant as both the reaction and growth temperature increases. However, *a* and *c* deviates from the standard reported values of 3.2475 - 3.2501 Å and 5.2042 - 5.2075 Å, respectively, and the ratio *c/a* varies in the range 1.593 - 1.6035 [24 - 26]. The ratio *c/a* relates the electronegativities of the zinc and oxygen constituents and the components undergoing a massive difference show large departure from the ideal ratio [24].

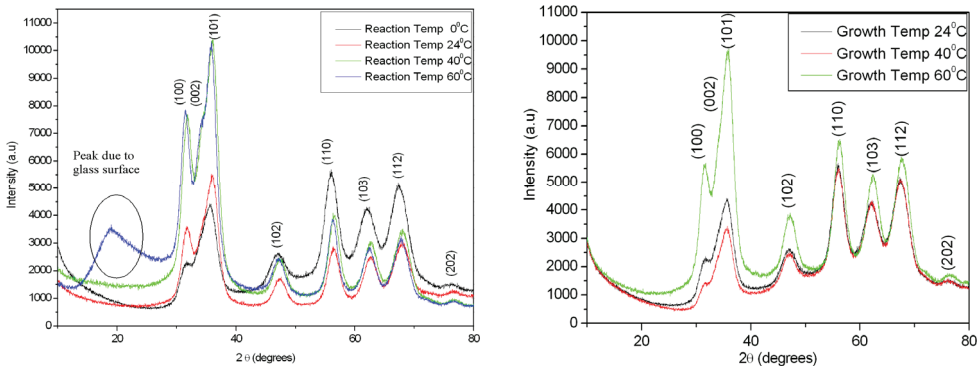
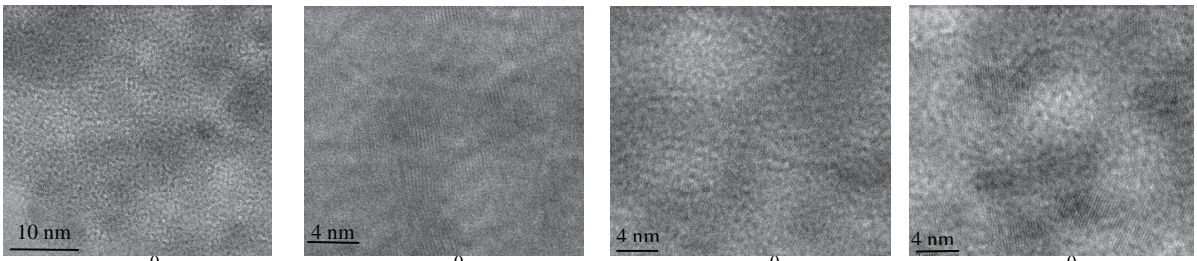


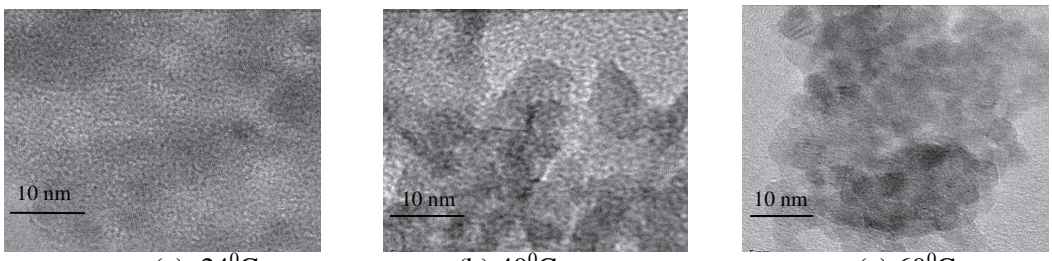
Figure 1. XRD spectra of ZnO nanoparticles synthesized at different reaction and growth temperatures

The crystalline nature and the internal structure of the nanoparticles have been further investigated by HR-TEM. Figures 2 and 3 depict the respective HR-TEM images of the ZnO nanoparticles synthesized at different reaction and growth temperatures, respectively. Clear lattice fringes are present for all samples, indicative of the superior crystallinity. The particle size was statistically measured from the low magnification TEM micrographs (not shown). As in the case for XRD, a growth in the particle size from 4.1 – 6.3 nm and 4.1 – 5.7 nm is observed with an increase in growth and reaction temperature, respectively. It should be noted that the lattice strain effects were neglected from the calculation of the particle sizes from the XRD spectra, which explains the discrepancy between the particle sizes determined from XRD and HR-TEM.



(a) 0^0C (b) 24^0C (c) 40^0C (d) 60^0C

Figure 2. TEM images of ZnO nanoparticles at their respective reaction temperatures at high magnification



(a) 24^0C (b) 40^0C (c) 60^0C

Figure 3. TEM images of ZnO nanoparticles at their respective growth temperatures at high magnification

3.2. Optical Properties

Figure 4 shows the optical absorbance spectra of ZnO nanoparticles synthesized at the respective reaction and growth temperatures. For qualitative analysis, the optical absorption edge can be estimated from the absorption peak centered around 320 nm. A systematic red shift in the absorption edge is observed with an increase in the reaction and growth temperature, indicative of a reduction in the optical band gap. The reduction in the optical band gap is attributed to a growth in the particle size and is explained by the Ostwald ripening process [27].

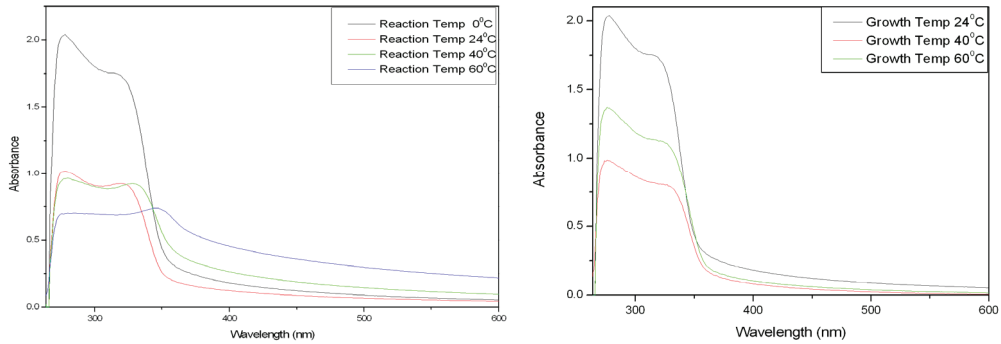


Figure 4. Optical absorption spectra of ZnO nanoparticles synthesized at their respective reaction and growth temperatures.

Figure 5 shows the optical emission (photoluminescence) spectra of ZnO nanoparticles synthesized at the respective reaction and growth temperatures. It should be noted that the concentration of the ZnO solutions used for the PL measurements were fixed for all samples. Two emission peaks are present in the spectrum: one consisting of ultraviolet peak located in the range from 375 – 396 nm, related to the near band gap transition of ZnO which comes from the recombination of free excitons [28]. The other peak is located in the visible emission in the range from 492 – 507 nm, representing the green emission due to the presence of the singly ionized oxygen vacancies. This is caused by the radiative recombination of a photo-generated hole with an electron occupying the oxygen vacancy.

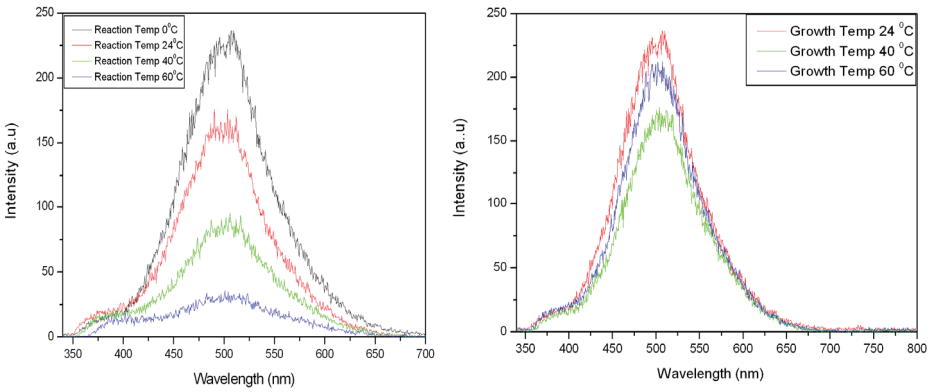


Figure 5. Optical emission (PL) spectra of ZnO nanoparticles synthesized at their respective reaction and growth temperatures.

The intensity of the visible emission band is higher than the ultraviolet emission band for all respective temperatures. A reduction in the intensity of the green emission peak is observed with an increase in the reaction and growth temperature, with the exception of the particles synthesized at a growth temperature of 60 °C. Dijken *et al* [28] reported that the key features of the visible emission process are a very fast trapping of a photogenerated hole at the surface site. It further continues to tunnel the surface trapped hole back into the bulk of the particle to the V^{*}_{o} oxygen vacancy and creates a V^{**}_{o} center where it recombines with a shallowly trapped electron to give rise to a trap emission. As the particle sizes increases the tunneling rate of the surface trapped holes to the V^{**}_{o} center decreases, resulting in the observed reduction in the intensity of the visible emission [28].

4. Conclusion

XRD and HR-TEM show that an increase in reaction and growth temperature results in a growth in the ZnO nanoparticle size with minimum effect on the crystalline structure, i.e. interplanar spacing and lattice constants. Optical absorption spectroscopy show an expected red shift in the absorption edge with an increase in particle size, which is attributed to the Ostwald ripening process. Emission spectroscopy show a reduction in the visible emission intensity, which has been related to a growth in the particle size.

References

[1] Peng W Q, Qu S C, Cong G W and Wang Z G 2006 *Mater. Sci. Semic. Proc.* **9** 156 - 159
[2] Yu Q, Fu W, Yu C, Yang H, Wei R, Sui Y, Lui Y, Lui Z, Li M, Wang G, Shao C, Lui Y and Zou G 2007 *J. Phys. D: Appl. Phy.* **40** 5592 - 5597
[3] Pal U, Serrano J G, Santiago P, Xiong G, Ucer K B and Williams R T 2006 *Opt. Mater.* **29** 65-69
[4] Vafaee M and Sasani Ghamsari M 2007 *Mater. Lett.* **61** 3265 - 3268
[5] Wu L, Wu Y and Lü W 2005 *Physica.* **28** 76 - 82
[6] Bera D, Qian L, Sabui S, Santra S and Holloway P H 2008 *Opt. Mat.* **30** 1233– 1239
[7] Wang Z G, Zu X T, Zhu S and Wang L M 2006 *Phy. E.* **35** 199–202
[8] Ni Y, Cao X, Wu G, Hu G, Yang Z and Wei X 2007 *Nanotechn.* **18** 155603 – 155609
[9] Baruah S and Dutta J 2009 *Sci. Technol. Adv. Mater.* **10** 45 - 48
[10] Wang L and Muhammed M 1999 *J. Mater. Chem.* **9** 2871 – 2878
[11] Jitianu M and Goia D V 2007 *J. Coll. Interf, Sci.* **309** 27 - 85
[12] Zhou J, Zhao F, Wang Y, Zhang Y and Yang L 2007 *J. Lumin.* **122 – 123** 195
[13] Zhang Y L, Yang Y, Zhao J H, Tan R Q, Cui P and Song W J 2009 *J. Sol-Gel Sci Technol.* **51** 198
[14] Sridevi D and Rajendran K V 2009 *Int. J. Nanotechn. Appl.* **3(2)** 43
[15] Bouvy C, Marine W, Sporken R and Su B L 2006 *Chem. Phys. Lett.* **428** 312

[16] Gumus C, Ozkendir O M, Kavak H and Ufuktepe Y 2006 *J. Optoel. Adv. Mater.* **8** 299 – 303
[17] Seow Z L S, Wong A S W, Thavasi V, Jose R, Ramakrishna S and Ho G W 2009 *Nanotech.* **20** 045604 - 045610
[18] Xia Y , Yang P, Sun Y, Wu Y , Mayers B, Gates B, Yin Y, Kim F and Yan H 2003 *Adv. Mater.* **15** 353 – 389
[19] Kharisov B I 2008 *Rec.Pat. Nanotech.* **2** 190 - 200
[20] Hu X, Masuda Y, Ohji T and Kato K 2008 *J. Cer. Soc. Japan.* **116** 384 - 388
[21] Yang J, Liu X, Yang L, Wang Y, Zhang Y, Lang J, Gao M and Wei M 2009 *J. Alloys & Comp.* ***485*** 743 - 746
[22] Dhlamini M S 2008 PhD Dissertation (University of Free State)
[23] Shen L, Bao N, Yanagisawa K, Domen K, Gupta A and Grimes C A 2006 *Nanotech.* **17** 5117
[24] Özgür Ü, Alivov Ya I, Liu C, Teke A, Reshchikov M A, Doğan S, Avrutin V, Cho S J and Morkoç H 2005 *J. Appl. Phy.* **98** 041301
[25] Morkoç H and Özgür Ü 2009 *Zinc Oxide: Fundamentals, Materials and Device Technology* (WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim)
[26] Jagadish C and Pearton S J 2006 *Zinc Oxide Bulk, Thin Films and Nanostructures: Processing, Properties and Application* (Elsevier Ltd)
[27] Yadav R S, Mishra P and Panday A C 2008 *J. Ult. Son. Chem.* **15** 863
[28] van Dijken A, Meulenkamp E A, Vanmaekelbergh D and Meijerink A 2007 *J. Lumin.* **87 – 89** 454

Effect of annealing on the Ce^{3+}/Ce^{4+} ratio measured by XPS in luminescent SiO_2:Ce

HAA Seed Ahmed, WD Roos, OM Ntwaeaborwa, HC Swart and RE Kroon

Physics Department, University of the Free State, PO Box 339, Bloemfontein 9300, South Africa

E-mail: KroonRE@ufs.ac.za

Abstract. Ce doped silica was prepared by the sol gel method and annealed in air or reduced with hydrogen, both at 1000°C. The reduced sample exhibited strong luminescence while the sample annealed in air did not, which has been attributed to Ce taking the tetravalent form which is non-luminescent in the latter case. In order to test this, x-ray photoelectron spectroscopy (XPS) was used to investigate the oxidation state of the Ce ion. The obtained spectra were fitted with 10 Gaussian peaks and fitted parameters showed the existence of the Ce^{3+} ion is about 20% and 73% in annealed and reduced samples respectively, confirming the increased luminescence is associated with reduction of the Ce. This is in contrast to other published results where samples subjected to rapid thermal annealing using an oxidising oxygen-hydrogen flame exhibited an increase in luminescence despite indications that the Ce^{4+} concentration had increased. Photoluminescence (PL) and UV–vis spectroscopy were performed on the reduced and annealed samples and the results are consistent with the XPS measurements.

1. Introduction

Ce doped silica has potential applications as a luminescent material as phosphors for cathodoluminescence, scintillators and detectors [1]. Ce ions can occur in a trivalent or a tetravalent state: only the trivalent Ce^{3+} state with a single 4f electron is optically active, while the tetravalent Ce^{4+} ion is non-luminescent. X-ray photoelectron spectroscopy (XPS) is a suitable technique to investigate the oxidation states of Ce in cerium oxides and such studies have been carried out because of the importance of CeO_2/Ce_2O_3 conversion in automotive exhaust catalysts [2]. However, the XPS Ce3d spectrum of cerium oxide is rather complex as it contains ten closely spaced and overlapping peaks on a strong background. The main challenge is to obtain accurate fits to experimental data while still maintaining a good physical basis for the fitting parameters [3]. The analysis of Ce in SiO_2:Ce is even more challenging since the Ce concentration for luminescent samples is only in the region of 1 mol%. Although it has been experimentally shown that to improve the luminescence efficiency of Ce doped silica it can be useful to anneal the glass in a reducing atmosphere [4], with the implication that this increases the concentration of Ce^{3+} luminescent ions, there is no evidence of XPS measurements that correlate the relative concentrations of the Ce^{3+} and Ce^{4+} ions to the luminescent properties. In this work, cerium doped silica was prepared by the sol-gel method. The effect of annealing atmosphere on the luminescent properties is correlated to XPS measurements of the oxidation state of Ce in the samples.

2. Experimental

Ce-doped SiO_2 glass was prepared by the sol–gel method using tetraethylorthosilicate (TEOS), water, ethanol (C_2H_5OH) and $Ce(NO_3)_3$ as starting materials and nitric acid (HNO_3) as a catalyst. The molar ratio of TEOS:H_2O:C_2H_5OH was 1:5:10, and the HNO_3 concentration was 0.015 M in water. TEOS was mixed with ethanol and stirred for 30 minutes after which the acidified water was added and stirring continued for another 30 minutes. After that $Ce(NO_3)_3$ was dissolved in a little ethanol and added to the mixture, which was stirred for a further 4 h. Then the mixture was stored in a closed container and transferred to a water bath at 50°C until a gel was formed. The gel was dried, crushed and divided into two parts, one part annealed in air at 1000°C for 2 h, and the other part annealed at flowing 4% hydrogen in argon atmosphere for the same time and temperature. XPS measurement were made using two different spectrometers, PHI 5400 ESCA and PHI 5000 Versaprobe both using Al x-rays. The powder sample was deposited onto a conductive carbon tape stuck to the sample holder. Charging shifts of the spectra were corrected using the known values of the C1s and/or Si2p peaks. Photoluminescence was measured at room temperature with a Cary Eclipse fluorescence spectrophotometer equipped with a xenon lamp. Diffuse reflectance spectra were recorded using a Lambda 950 UV–vis spectrophotometer with an integrating sphere.

3. Results and Discussions

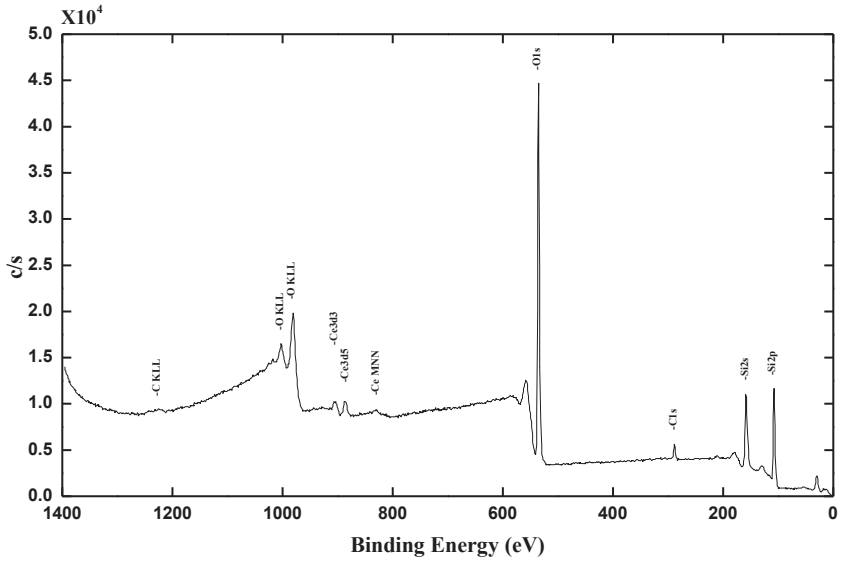
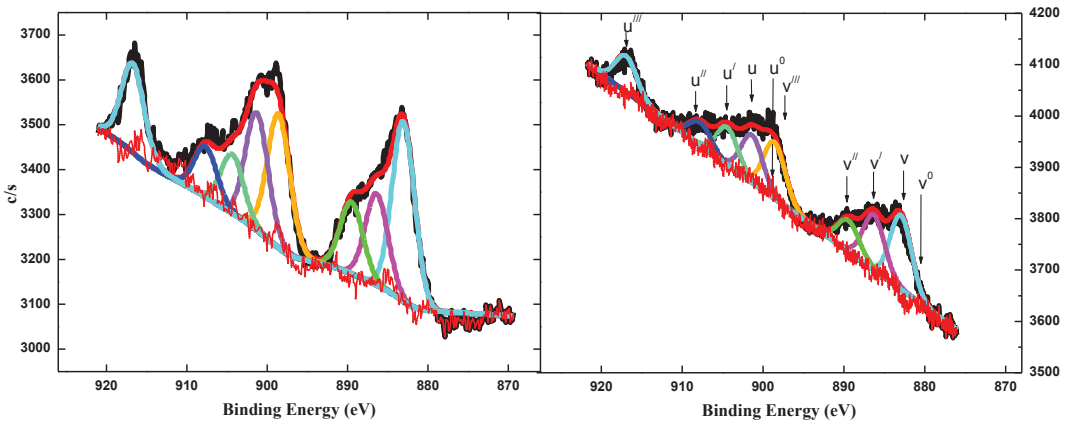


Figure 1. XPS survey scan spectrum of the reduced sample

3.1. XPS

For XPS analysis a higher dopant concentration (4 mol%) was used than what is required for good luminescence (0.5 mol%) to obtain a reasonable Ce3d signal. Figure 1 shows the full range survey scan spectrum of the reduced sample. The spectrum is dominated by four photoelectron peaks, corresponding to electrons originating in the 2p orbital of the Si, 1s orbitals of the C and O and 3d orbital of the Ce atoms in the sample surface. The existence of the C atoms on the sample surface can be ascribed to three possible reasons: from adventitious hydrocarbon nearly always present, or from the precursors, or from the carbon tape used to stick the powder to the sample holder. Multipak version 8.2c software was used to fit the Ce3d pattern. The spectrum was fitted with 10 Gaussian peaks, two doublet peaks of the spin–orbit split 5/2;3/2 components attributed to the final states of Ce^{3+} and three doublet peaks of the spin–orbit split 5/2;3/2 components attributed to the final states of



(a) (b)

Figure 2. Ce3d XPS spectra of the annealed sample obtained by (a) PHI 5000 (b) PHI 5400.

Table 1. Fit parameter of the annealed sample.

ion	$3d_{5/2}/3d_{3/2}$	pos	final state	%area	%total
Ce^{4+}	v/u	882.3/901.0	Ce $3d^9 4f^2$ O $2p^4$	36	80
	v''/u''	888.9/907.5	Ce $3d^9 4f^1$ O $2p^5$	16	
	v'''/u'''	898.2/916.8	Ce $3d^9 4f^0$ O $2p^6$	28	
Ce^{3+}	v^0/u^0	880.5/898.8	Ce $3d^9 4f^2$ O $2p^5$	0	20
	v'/u'	885.7/904.1	Ce $3d^9 4f^1$ O $2p^6$	20	

Ce^{4+} (see Tables 1 and 2). Peak position, spin–orbit splitting, area ratio and Full Width at Half Maximum (FWHM) are fitting parameters that need to be controlled in order to perform good fits with physically meaningful parameters. Figure 2(a) and (b) show the detailed Ce3d XPS spectra of the sample annealed at 1000°C obtained by PHI 5000 and PHI 5400 respectively. The spectra were fitted with 10 peaks and the FWHM of all the peaks was kept the same for each fit. The peaks positions and the spin–orbit splitting agree well with the literature [5] and the intensity ratio between $3d_{5/2}$ and $3d_{3/2}$ was taken as 3:2 [6]. Six peaks labelled as v/u , v''/u'' and v'''/u''' referring to three doublets of the spin–orbit split components are attributed to Ce^{4+} final states. The high binding energy doublet v'''/u''' at 898.2 and 916.8 eV are attributed to the final state of Ce $3d^9 4f^0$ O $2p^6$, doublet v''/u'' at 888.9 and 907.5 eV are attributed to the state of Ce $3d^9 4f^1$ O $2p^5$, and doublet v/u at 882.3 and 901.0 eV correspond to the state of Ce $3d^9 4f^2$ O $2p^4$. The other four peaks labelled as v^0/u^0 and v'/u' referring to two doublets of the spin–orbit split components are attributed to Ce^{3+} final states. The doublet v'/u' at 885.7 and 904.1 eV are corresponding to the final state of Ce $3d^9 4f^1$ O $2p^6$ and the doublet v^0/u^0 at 880.5 and 898.8 eV are corresponding to the final state of Ce $3d^9 4f^2$ O$2p^5$. Table 1 lists the fit parameters obtained from the spectra of the PHI 5000. The peak area percentage can be used to determine the relative concentrations [6]. From the table the relative total concentration of the Ce^{4+} is 80% while that of the Ce^{3+} is 20%. From the PHI 5400 spectrometer, the calculated relative concentrations of Ce^{4+} and Ce^{3+} are 75% and 25% respectively. The different values given by the two spectrometers can be taken as an indication of the accuracy of the technique. Figure 3(a) and (b) shows the Ce3d XPS spectra of the reduced sample obtained by the PHI 5000 and PHI 5400 respectively. The Ce^{4+} u''' characteristic peak has almost disappeared indicating that the sample contains less Ce^{4+}. The spectra can be fitted with the four peaks v^0/u^0 and v'/u' of the Ce^{3+} components, in addition of the

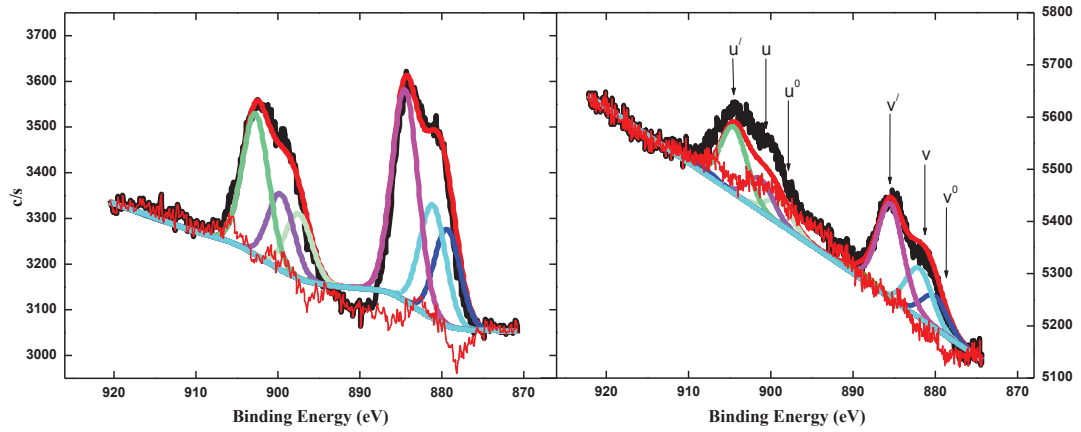


Figure 3. Ce3d XPS spectra of the reduced sample obtained with the (a) PHI 5000 and (b) PHI 5400 XPS systems.

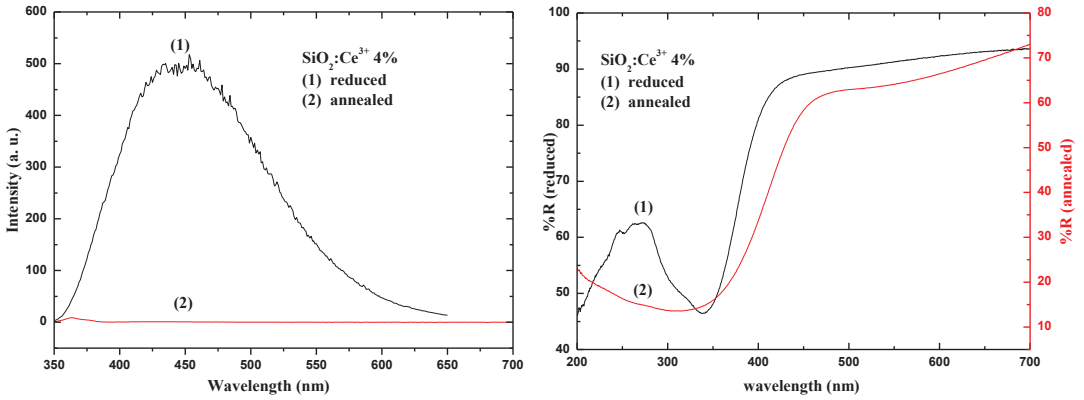
Table 2. Fit parameter of the reduced sample.

ion	$3d_{5/2}/3d_{3/2}$	pos	final state	%area	%total
Ce^{4+}	v/u	882.3/901.0	Ce $3d^9 4f^2$ O $2p^4$	27	27
	v''/u''	888.9/907.5	Ce $3d^9 4f^1$ O $2p^5$	0	
	v'''/u'''	898.2 /916.8	Ce $3d^9 4f^0$ O $2p^6$	0	
Ce^{3+}	v^0/u^0	880.5/898.8	Ce $3d^9 4f^2$ O $2p^5$	23	73
	v'/u'	885.7/904.1	Ce $3d^9 4f^1$ O $2p^6$	50	

doublet v/u peaks of Ce^{4+} in the valley between (v^0 and v') and (u^0 and u'). Table 2 lists the fit parameters for the spectra obtained by the PHI 5000 from the reduced sample. The total relative concentration of the Ce^{3+} is 73% which compares well with the 71% measured by the PHI 5400. There were some difficulties to fit these spectra. The software can not subtract the background well in the spectra obtained by PHI 5000, using either the Shirley or iterated Shirley algorithms. While the background was subtracted succesfully in the spectra obtained by PHI 5400, the two seperate bands couldnot fit well simultaneously without relaxing the physically important constraints. Ce3d XPS spectra and their fitting obviously show that reducing the sample with hydrogen can improve the Ce^{3+} concentration.

3.2. Photoluminescence and UV–vis spectroscopy

Figure 4(a) shows the PL of the reduced and annealed samples. The sample which was annealed in air does not show significant emission, while the reduced one exhibits a high intensity emission. The sample annealed in air may contain more Ce^{4+} which is optically in-active [7]. A significant PL emission is observed in the reduced sample which may containing more Ce^{3+} rather than Ce^{4+} and significant PL emission observed. These results are consistent with the XPS measurements done on the two samples. Figure 4(b) shows the reflectance spectra done on the annealed and the reduced samples. The optical reflectance profile from the annealed sample exhibits a wide absorption peak which corresponds to Ce^{4+} ion absorption due to charge transfer from O^{2-} to Ce^{4+} [8]. In the reduced sample an absorption peak appeared at around 338 nm, which corresponds to the 4f to 5d transition of Ce^{3+} [9,10].



(a) (b)

Figure 4. (a) Photoluminescence and (b) diffuse reflectance of reduced and annealed samples

4. Conclusion

Ce ions were successfully incorporated into SiO_2 by the sol-gel process. Ce 3d XPS spectra analysis indicates that the concentration of the Ce^{3+} was increased in the reduced sample. This is in contrast to other published results [11] where samples subjected to rapid thermal annealing using an oxidising oxygen-hydrogen flame exhibited an increase in luminescence despite indications that the Ce^{4+} concentration had increased. The reduced sample shows a high intensity PL emission compared to the annealed sample. UV–vis spectroscopy done on the reduced sample shows an absorption peak associated with the Ce^{3+} ion while the annealed samples shows an absorption peak associated with the Ce^{4+} ion. The photoluminescence and UV–vis spectroscopy results are consistent with the XPS measurements.

Acknowledgments

The authors thank Dr. E. Coetsee for her assistance to perform XPS measurements. The financial support from the South African National Research Foundation (NRF) is also gratefully acknowledged.

References

[1] Reisfeld R, Patra A, Panczer G and Gaft M 1999 *Opt. Mat.* **13** 81
[2] Larachi F, Pierre J, Adnot A, and Bernis A 2002 *Appl. Surf. Sci.* **195** 236
[3] Paparazzo E 2011 *Mat. Res. Bulletin* **46** 323
[4] Fasoli M, Vedda A, Lauria A, Moretti F, Rizzelli E, Chiodini N, Meinardi F and Nikl M 2009 *J. Non-Cryst. Solids* **355** 1140
[5] Limei Q, Fen L, Liangzhong Z, Ying M and Jiannian Y 2006 *Appl. Sur. Sci.* **252** 4931
[6] Moulder F, Stickle W F, Sobol P E and Bomben K D 1995 *Handbook of X-ray Photoelectron Spectroscopy* (Japan: Physical Electronics) p 10
[7] Xu G Q, Zheng Z X, Tang W M and Wu Y C 2007 *J. Lumin.* **124** 151
[8] Ligang Z, Anxian L, Chenggang Z and Weiqun S 2011 *J. Alloys Compounds* **509** 7789
[9] Paul A, Mulholland M and Zaman M S 1976 *J. Mater. Sci.* **11** 2082
[10] Weipingyz C, Yeyz Z and Lideyz Z 1998 *J. Phys Condens. Matter* **10** L473
[11] Chiodini N, Fasoli M, Martini M, Rosetta E, Spinolo G and Vedda A 2002 *Appl. Phys. Lett.* **81** 4374

Synthesis and photoluminescence properties of Tb^{3+}-doped $Sr_{0.5}$ $Zn_{0.5}$ Al_2O_4 phosphor prepared via combustion process

S.K K. Shaat[1], F. Roelofse[2], H.C. Swart[1] and O.M. Ntwaeaborwa[1,3]

Department of Physics, University of the Free State, P. O. Box 339, Bloemfontein, ZA9300, South Africa

Department of Geology, University of the Free State, P.O. Box 339, Bloemfontein, ZA9300, South Africa

[3]Corresponding author: ntwaeab@ufs.ac.za

Abstract. Tb^{3+} doped $Zn_{0.5}Sr_{0.5}Al_2O_4$ (ZSAO) phosphor was successfully synthesized by a combustion method at a relatively low temperature (500°C), using metal nitrates as precursors and urea as a fuel. The samples were annealed in air at 700°C for 4 hrs. The X-ray diffraction patterns showed phases associated with both $ZnAl_2O_4$ (ZAO) and $SrAl_2O_4$ (SAO) as referenced to standard JCPDS data files No. 05-0669 and 74-794, respectively. As confirmed from the scanning electron microscopy images, the characteristic platelet-like particles of the combustion method were formed. The photoluminescence emission and excitation spectra of $Zn_{0.5}Sr_{0.5}Al_2O_4$:Tb^{3+} before and after annealing were recorded and compared to those of $ZnAl_2O_4$:Tb^{3+} and $SrAl_2O_4$:Tb^{3+}. In all cases, PL emissions were due to $^5D_4 \rightarrow {}^7F_J$ (J = 0,1,2,3,4,5,6) transitions of Tb^{3+}. The PL emissions of $ZnAl_2O_4$:Tb^{3+} and $SrAl_2O_4$:Tb^{3+} were more intense than that of $Zn_{0.5}Sr_{0.5}Al_2O_4$:Tb^{3+}. The objective of the study was to prepare a new aluminate host for Tb^{3+} consisting of Zn and Sr and evaluate its luminescent properties and chemical study for application in lighting.

1. Introduction

In recent years, interests have been focused on the development of new luminescent materials (phosphors) such as lanthanide-doped nanomaterials. Among the known luminescent materials, lanthanide (Ln) ions doped inorganic materials offer a range of compounds with unique versatility. In particular, oxide matrices are attractive host materials for the study of development of advanced phosphors due to their ease of synthesis and chemical stability [1]. Among various host oxide matrix materials, aluminates are considered to possess high energy efficiency, wide excitation wavelength range, and high quenching temperature. Aluminium-based spinels are an intriguing class of oxide ceramics with important technological applications. Rare earth ions (RE) doped aluminates serve as an important class of phosphor for fluorescent lamp and plasma display application and phosphorescence. As one of the most promising phosphors, strontium aluminate spinel, $SrAl_2O_4$ (SAO), has been found to be an efficient host material with a wide band gap, which offers the possibility of generating broad band emission [2] and much attention has been paid to it owing to its higher radiation-resistance and

[3] Corresponding author: email address: ntwaeab@ufs.ac.za

thermal stability [3]. SAO is considered as an ideal host material for long-lasting phosphorescence. The synthesis and optical properties of SAO, in bulk form, or as films or nanoparticles, have been extensively explored during the past decades [1]. Zinc aluminate, $ZnAl_2O_4$ (ZAO), is a well-known wide-band gap semiconductor and due to its transparent and electro-conductive properties, it can be used for ultraviolet (UV) photoelectronic devices. Recently, ZAO has been extensively studied as a phosphor host material for application in thin film electroluminescence displays, mechano-optical stress sensors, and stress imaging devices [4]. It has a unique combination of attractive properties such as high thermal stability, better diffusion and ductility. Due to these unique properties, it is used as high temperature material and optical coating. In addition, ZAO is also widely used in many catalytic reactions, cracking, dehydration, hydrogenation and dehydrogenation, ceramic and electro-conductive materials because of its high thermal stability, high mechanical resistance, low surface acidity, and excellent optical properties [5]. The understanding of optical properties of Tb^{3+} in inorganic compounds is of great importance because these compounds have potential technological applications as functional photonic materials such as optical fiber amplifiers, lasers, and wavelength converting devices [6]. Tb^{3+} is a well known activator for green luminescence emitting materials and has been used widely in tricolor energy saving fluorescent lamp. The green emission of Tb^{3+} mainly originates from $^5D_4 \rightarrow {}^7F_J$ (J=0−6) transitions [7], and the blue emission and UV-light correspond to the $^5D_3 \rightarrow {}^7F_J$ [3]. The luminescent properties of phosphors depend strongly on the particle size, crystal structure, and uniform distribution of activators in the host lattice [8]. In recent years, a variety of synthesis technique routes have been used successfully to prepare pure spinel powders [4], such as precipitation [9], solid-state reaction [10], hydrothermal synthesis [11], sol-gel [12], sol-gel-microwave process [13], co-precipitation [14], and combustion reaction [15-17] or microwave combustion methods [18]. Among the aforementioned different synthesis techniques, combustion reaction can be used to prepare complex oxide ceramics, such as aluminates, ferrites and chromites [19]. The combustion reaction has been widely used to fabricate aluminates, due to its success in producing homogeneous and nanocrystalline aluminate phosphors with fine particles. In addition, the synthesis can be carried out at relatively low temperatures in a short processing time [20]. In comparison with other methods, combustion synthesis is the most feasible choice because it can guarantee high purity, compositionally uniform, single phase and small and uniform particle sizes [21]. Previously, Kingsley et al.[22,23] obtained aluminates with uniform small size and good dispersity at 500 or 350 ^oC by the combustion reaction of the relevant metal nitrates and urea. In this study, the combustion method was used to prepare a new aluminate matrix consisting of 1:1 ratio of Zn to Sr and was used as host for Tb^{3+} resulting in a green emitting $Zn_{0.5}Sr_{0.5}Al_2O_4:Tb^{3+}$ phosphors, which was evaluated for application in different types of light emitting devices. The samples prepared were characterized by X-ray diffraction, scanning electron microscopy (SEM), and a photoluminescence (PL) property of doped sample was investigated.

2. Experimental

Tb^{3+} 1mol% doped $Zn_{0.5}Sr_{0.5}Al_2O_4$ (ZSAO) phosphor was successfully prepared by combustion method at an initiating temperature of 500^oC. $Al(NO_3)_3.9H_2O$, $Sr(NO_3)_2$, $Zn(NO_3)_2.4H_2O$, $Tb(NO_3)_3.5H_2O$ and $CO(NH_2)_2$ precursors of AR grade purchased from Merck, South Africa were used as precursors and were used as obtained without further purification. The precursors were dissolved in distilled water and stirred at 65°C for 5hrs and a transparent solution was obtained. The resulting solution was introduced into a muffle furnace maintained at 500°C. Initially, the solution boiled and underwent dehydration, followed by decomposition with escaping large amounts of gases (oxides of carbon, nitrogen and ammonia) and then spontaneous ignition occurred and underwent smouldering combustion with enormous swelling, producing white foamy and voluminous ash, the whole process was over within 5 minutes. The product was cooled to room temperature and the ashes were ground gently. The resulting powders were annealed at 700°C for 4 hrs. The amount of Zn and Sr in the powder was 0.5 mol% each and Tb^{3+} was 1 mol%. Using the same procedure, $ZnAl_2O_4:Tb^{3+}$ and $SrAl_2O_4:Tb^{3+}$ were also prepared. The crystalline structure and phase composition of the

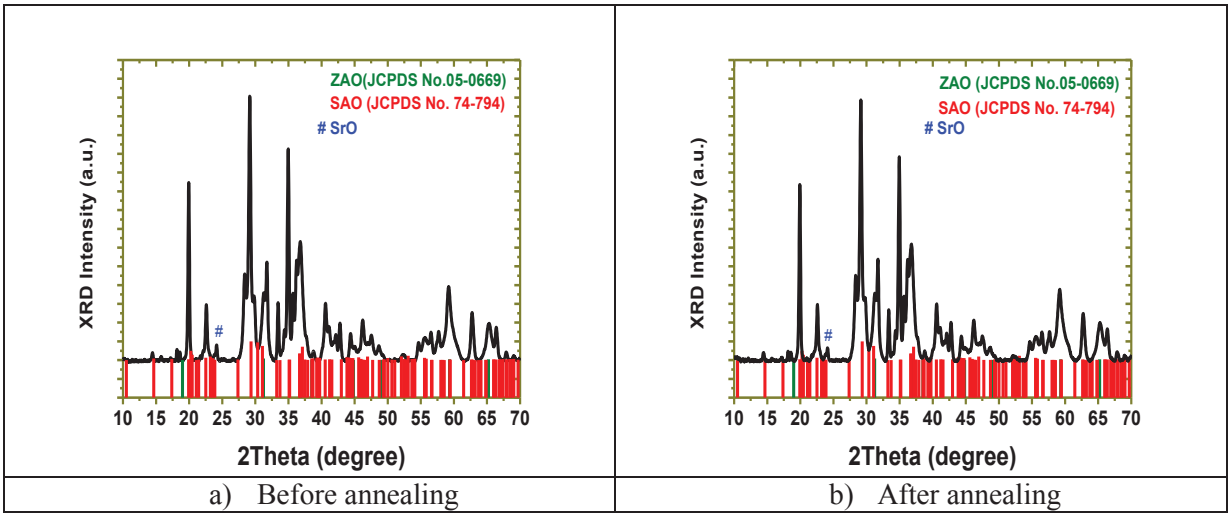
$Zn_{0.5}Sr_{0.5}Al_2O_4:Tb^{3+}$ were analysed by X-ray diffraction (XRD) using a Siemens D5005 diffractometer with Cu Kα radiation, λ = 1.5406 Å. Scanning electron microscopy pictures were collected using JEOL JSM-6610 Scanning electron microscopy, Energy dispersive X-ray (EDS) measurements were performed using energy dispersive X-ray spectrometer attached to JEOL JSM-6610 (SEM). Photoluminescence (PL) excitation and emission measurements were carried out at room temperature using the Varian Cary Eclipse fluorescence spectrophotometer.

3. Results and Discussion

4. Structure

4.1. XRD analysis

Fig. 1 displays the X-ray diffraction patterns of ZSAO: Tb^{3+} before and after annealing. The patterns are consistent with ZAO and SAO phases referenced in the Joint Committee of Powder Diffraction (JCPDS) file No. 05-0669 and 74-794, respectively. It is noticed that there were no differences on the X-ray diffraction before and after annealing. The broadening of the diffraction peaks points to the presence of nanocrystalline domains. The result indicates that the obtained products are of high crystallinity and the small amount of Tb^{3+}doped virtually has no effect on the phase structures. Secondary phase of SrO was detected.

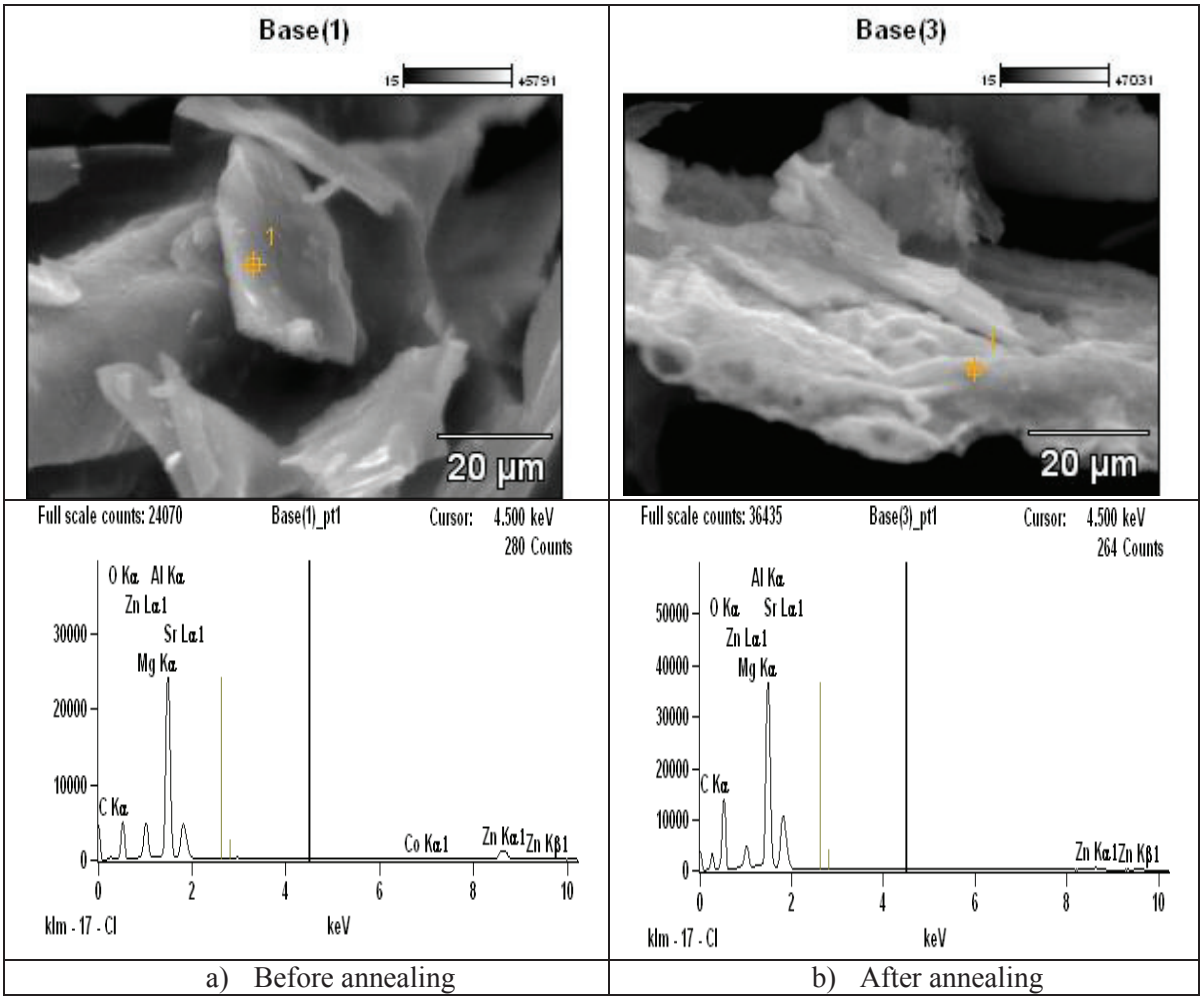


a) Before annealing | b) After annealing

Figure 1: XRD pattern along with the standard JCPDS file of combustion synthesized ZSAO:Tb^{3+} lmol.% phosphor.

4.2. Phosphor morphology

The morphological aspect of the resulting powders was examined by scanning electronic microscopy (SEM). The SEM images of ZSAO:Tb^{3+} before and after annealing are shown in Fig. 2a–b. The well-known semicontinuous folded dense platelet like morphology was observed and there were also voids resulting evolution and escape of large amount of gases during the combustion process. The particle shape was irregular; the surfaces of the foams show a lot of cracks, voids and pores formed by the escaping gases during combustion reaction. Compositional analysis on the nanostructures probed by EDS shows that the powders were composed of the C, Zn, Sr, Al and O elements (Fig. 2), and also, composed of a small amount of Mg and Co, which probably were incidental impurities in the precursors. Tb^{3+} was not detected due to its relatively low concentration. The results show that the doping process does not make significant changes to the morphology and size of the nanostructures.



a) Before annealing b) After annealing

Figure 2: SEM micrographs of ZSAO:Tb^{3+} lmol% , and the EDS spectrum of the combustion.

5. Photoluminescence properties

Photoluminescence (excitation and emission) spectra of Tb^{3+} in nano sized ZSAO, SAO and ZAO powder phosphors are shown in Fig 3a- b, before and after annealing. The excitation peak at 230 nm is attributed to the $4f^8 \rightarrow 4f^85d^1$ transitions of Tb^{3+} [15]. Fig.3a depicts the emission spectra ($\lambda_{excitation}$ = 230 nm) before annealing which consists of eight line emission bands. The 378, 410, 434 and 447 nm bands originate from the $^5D_3$–^7F_J (J= 6,5,4,3,2) transitions, whereas 487, 544, 584 and 619 nm bands come from $^5D_4$–^7F_J (J=6,5,4,3,2) transitions of Tb^{3+}. The emission spectra after annealing also consists mainly of eight broad band (Fig.3b). The PL intensity of the annealed sample in Fig 3b was considerably more intense than that of the as prepared sample. The highest intensity before and after annealing was observed from $ZnAl_2O_4$:Tb^{3+} and the least intensity was observed from $Zn_{0.5}Sr_{0.5}Al_2O_4$:Tb^{3+}. Although the PL emission from $Zn_{0.5}Sr_{0.5}Al_2O_4$:Tb^{3+} was less intense than those from $SrAl_2O_4$:Tb^{3+} and $ZnAl_2O_4$:Tb^{3+}, this preliminary data suggest that $Zn_{0.5}Sr_{0.5}Al_2O_4$, as a possible future host for rare-earths to prepare phosphors. This study is in progress and the effects of Zn and Sr concentration of the PL intensity will be evaluated in future experiments.

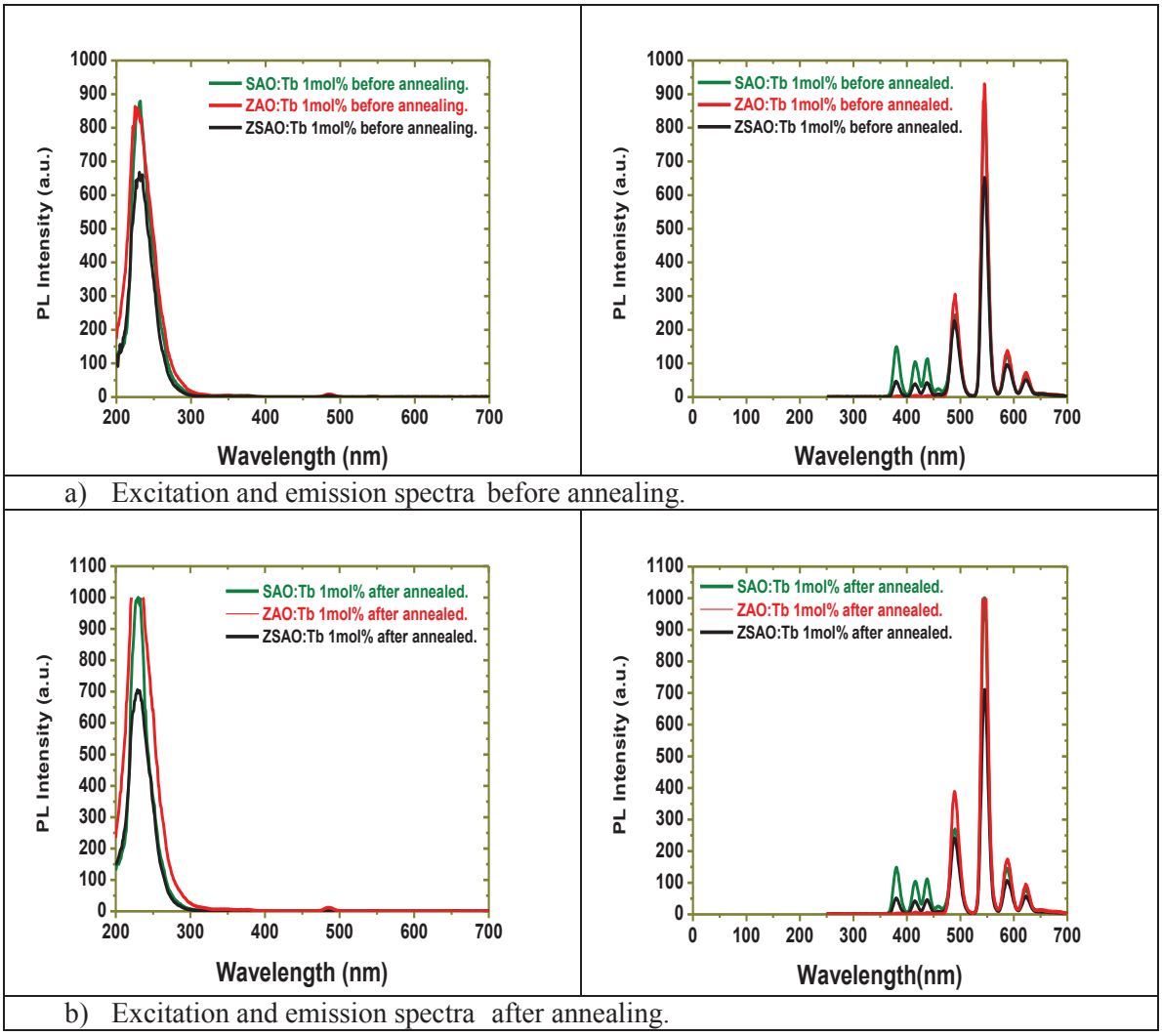


Figure 3: PL (excitation and emission) spectra of ZSAO:Tb^{3+} lmol% before and after annealing.

6. Conclusion

ZSAO:Tb^{3+} 1mol % was prepared using combustion method and the initiating temperature of 500 ^{o}C. The formation of the monoclinic and cubic spinel phase was confirmed by X-ray diffraction data. The well known characteristic platelet-like particles of the combustion method was confirmed from the SEM data. Characteristic blue and green luminescence from Tb^{3+} ions was observed in ZSAO:Tb^{3+} nanophosphors. The green line emission at 543 nm was more intense in annealed versus as prepared samples.

Acknowledgement

The authors are grateful to the financial support from the South African National Research Foundation (NRF).

References:

[1] Geng B Y, Ma J Z and Zhan F M 2009 *J.of Alloys and Compounds* **473** 530.
[2] Palilla F C, Levine A K and Tomkus M R 1968 *J. Electrochem. Soc.* **115** 642.
[3] Zhang Y, Lan L I, Zhang X, Wang D and Zhang S 2008 *J. of Rare Earths* **26** 656.
[4] Lou Z and Hao J 2005 *Appl. Phys. A* **80** 151.
[5] Xiulan D, Duorong Y, Xinqiang W and Hongyan X U 2005 *J. of Sol-Gel Science and Technology* **35** 221.
[6] Zuoling F U, Shihong Z and Siyuan Z 2005 *J. Phys. Chem. B* **109** 14396.
[7] Su Q, Liang H, Tao Y and Wang S *2001 J. of the Chinese Rare Earth Society (in Chin.)* **19**(6) 482.
[8] Mu-Tsun T, Yu-Xiang C, Pei-Jane T and Yen-Kai W 2010 *Thin Solid Films* **518** e9–e11.
[9] Adak A K, Pathak A and Pramanik P 1998 *J. Mater. Sci. Lett*. **17** 559.
[10] Valenzuela M A, Bosch P, Aguilar-rios G, Montoya A and Schifter I 1997 *J. Sol–Gel Sci. Technol*. **8** 107.
[11] Minami T et al In Extend. Abstr. 4th Int. Conf. 1998 *Science and Technology of Display Phosphors* Bend OR *370*.
[12] Matsui H, Xu C N and Tateyama H 2001 *Appl. Phys. Lett*. **78** 1068.
[13] Wu S, Zhang S and Yang J 2007 *Mater. Chem. Phys*. **102**(1**)** 80.
[14] Chang C K, Yuan Z X and Mao D L 2006 *J. Alloys Compd.* **415**(1-2) 220.
[15] Barros B S, Melo P S, Kiminami R H G A, Costa A C F M, Sa´ G F D and Alves S J 2006 *J. Mater Sci* **41** 4744.
[16] Esparza A, Garcia M and Falcony C 1998 *Thin Solid Films* **325** 14.
[17] Golego N and Cocivera M 1998 *Thin Solid Films* **322** 14.
[18] Hanyan D, Gengshen L and Jiayue S 2007 *J. of Rare Earths* **25**(1) 19.
[19] Huajie S and Donghua C 2007 *Luminescence* **22** 554.
[20] Pitale S S, Vinay K, Nagpure I M, Ntwaeaborwa O M and Swart H C 2011 *Applied Surface Science* **257** 3298.
[21] Shu F W, Feng G, Meng K L, Xiu F C, Wen G Z , Guang J Z, Shu M W and Yuan Y Z 2005 *J. of Alloys and Compounds* **394** 255.
[22] Kingsley J J and Patil K C 1988 *Mater. Lett.*, **6**(11-12), 427.
[23] Kingsley J J, Suresh K and Patial K C 1990 *J. Mater. Sci.* **25**(2) 1305.

Magnetic behaviour of the Cr-Al alloy system around the triple point

C J Sheppard[1], A R E Prinsloo, H L Alberts, A M Strydom and B Muchono
Department of Physics, University of Johannesburg, P O Box 524, Auckland Park, Johannesburg, Republic of South Africa, 2006

E-mail: cjsheppard@uj.ac.za

Abstract. Electrical resistivity and specific heat measurements on a series of $Cr_{1-x}Al_x$ ($0 \leq x \leq 0.034$) alloy single crystals are reported. The results indicate that the triple point, where the incommensurate and commensurate spin-density-wave phases coexist with the paramagnetic phase, is situated close to T = 0 K on the magnetic phase diagram.

1. Introduction

The magnetic phase diagram of the itinerant electron alloy system $Cr_{1-x}Al_x$ presents a triple point at a concentration $x_c \approx 2$ at. % Al, where the incommensurate (I) and commensurate (C) spin-density-wave (SDW) antiferromagnetic (AF) phases coexist with the paramagnetic (P) phase [1]. Recent studies [2] on a $Cr_{0.978}Al_{0.022}$ alloy single crystal suggest that the triple point of the $Cr_{1-x}Al_x$ alloy series might be a special type of critical point, for which ISDW, CSDW and P phases coexist at 0 K. This suggestion is further investigated here through electrical resistivity (ρ) and specific heat (C_p) measurements on a series ($0 \leq x \leq 0.034$) of $Cr_{1-x}Al_x$ single crystals that include x_c.

2. Theory and Experimental techniques

The single crystals (x = 0, 0.019, 0.022, 0.026, and 0.034) of this study were grown in our laboratories by a floating-zone technique using RF heating in an ultra-high purity argon atmosphere. Alloying Cr with Al forms dilute solid solutions in which Al atoms occupy substitutional cites in the body centred Cr lattice [3]. The as grown crystals where not submitted to applied magnetic fields and should be in the multi-$\bar{Q}$ domain state, where $\bar{Q}$ is the SDW wave vector. These crystals were previously used [4, 5, 6] for magnetoelastic (ME), neutron diffraction (ND) and in the case of x = 0.022 [2], also for ρ(T) and C_p(T) measurements. Measurements were done in the temperature range $2 \leq T \leq 350$ K, using standard PPMS [7] equipment. The electrical current for ρ measurements was directed along the cubic [100] crystallographic direction. The Sommerfeld specific heat coefficient (γ) was obtained for each x from a least squares fit to plots of C_p/T vs. T^2 at low temperatures. Such plots could be well fitted to the low-temperature approximation of the Debye formulation of specific heat $C_p(T) = \gamma T + \beta T^3$, where the first term represents the electronic specific heat component and the last term that of the acoustic modes of lattice contributions. Al impurities are not expected to produce local magnetic moments in Cr. Magnon contribution to C_p is therefore neglected in the above equation and γ is considered a measure of the density of states at the Fermi energy [8]. Theoretical models [9] for a simple two-lattice localized antiferromagnetic system predict a T^3 spin wave contribution to C_p, similarly to that of the

acoustic phonon contribution. In such an approximation an added magnon term to C_p will not influence γ.

Both ρ and γ are influenced by changes in the density of states at the Fermi energy. Such changes are introduced in Cr and its alloys by nesting of sections of the electron and hole Fermi surface sheets with a concomitant appearance of an energy band gap at the Fermi surface on SDW formation. These two physical parameters are thus ideally suited for the present investigation

3. Results

Figure 1 shows $\rho - T$ curves for the $Cr_{1-x}Al_x$ crystals. Although not too prominent for $x = 0.026$, these curves depict well defined ρ(T) – anomalies of SDW origin below the Néel temperature (T_N), which is shown by an arrow at values previously [4, 5, 6] obtained from neutron diffraction and magnetoelastic measurements (or ρ(T) - measurements in the case of pure Cr) on the same crystals. The crystals $x = 0$ and 0.019 display ISDW – P type Néel transitions [1, 5] and those with $x = 0.022$, 0.026 and 0.035 CSDW – P type [1, 4, 5], placing the $x = 0.022$ crystal probably just above the actual x_c.

The two ISDW alloys ($x = 0$ and 0.019) depict typical Cr-like ρ(T) anomalies in the form of a hump, with or without a minimum point, on cooling through T_N. The three CSDW alloys, however, behave anomalously, depicting weaker anomalies at T_N, in contrast to observations on most other CSDW Cr alloys [1]. This is notwithstanding the presence of very strong magnetoelastic anomalies, pointing to the complexity of the magnetic behaviour of the Cr-Al system.

Figures 1 (a) and (b) also show base-line paramagnetic curves, ρ_P(T), for the two $x = 0.019$ and 0.022 crystals, obtained from a theoretical back extrapolation of the ρ(T) data at $T > T_N$ down to 0 K. The $\rho - T$ curves for figure 1 (b) and (c) were fitted rather well in the paramagnetic region $T > 100$ K using the equation [2]:

$$\rho_P(\mathrm{T}) = \rho_0 + ATG\left(\frac{\theta_\mathrm{D}}{T}\right) + BT^3, \qquad (1)$$

where ρ_0 is the temperature independent impurity resistivity of an ideal non-magnetic $Cr_{1-x}Al_x$ crystal, the second term represents the lattice resistivity given by the Bloch-Grüneisen function $G(\theta_\mathrm{D}/\mathrm{T})$, and the third term represents an electron s-d scattering term that plays a role in dilute Cr alloys [2]. The constants ρ_0, A, B, and θ_D were treated as fitting parameters for a best fit at $T > T_N$. These fitting parameters were then used to obtain ρ_P(T) down to 2 K from equation (1). The back extrapolated ρ_P(T) curves represent ρ(T) should the two crystals remain paramagnetic down to 0 K in the absence of a P – SDW phase transition. Back extrapolation could however not be generated successfully for the other crystals, because the temperature range of the ρ(T) measurements at $T > T_N$ were too small. For these the open symbol linear back extrapolations in figure 1 only serve as indications of the P-behaviour.

The results of figure 1 nevertheless indicate fairly good correspondence of T_N, obtained from neutron diffraction and magnetoelastic measurements, with the temperature where the magnetic contribution, $\Delta\rho(\mathrm{T}) = \rho_{\mathrm{measured}}(\mathrm{T}) - \rho_\mathrm{P}(\mathrm{T})$, tends to zero. This confirms a previous definition [2] of T_N for the $x = 0.022$ crystal and a much deeper and sharper minimum at the triple point on the SDW-P phase line of the magnetic phase diagram, than that generally accepted [1].

Good quality single crystalline Cr alloys, like Cr-Ru [10], usually present a small first-order like step in ρ(T) on crossing the I-C phase line of the magnetic phase diagram. Such a step is absent in figure 1 for the CSDW alloys down to 2 K, confirming observations of neutron diffraction studies on the same crystals [4, 5] that the CSDW alloys remain CSDW without signs of an C-I transition down to 14 K. It is therefore concluded that, if present, the I-C phase line, starting at the triple point, (x_c, T_c), should have a very sharp slope, reaching a point (x, 0 K) in the range $0.019 < x < 0.022$, compared to the generally accepted [1] range $0.020 \lesssim x \lesssim 0.034$.

Figure 2(a) shows a plot of $\gamma(x)$ for the $Cr_{1-x}Al_x$ crystals, obtained from the C_p(T) measurements at $T \leq 6$ K. Typical examples of low temperature linear $C_p(\mathrm{T})/T - T^2$ plots are shown in figures 2(a)

and (b). Also shown in figure 2 are $\gamma(x)$ values [11] on polycrystalline $Cr_{1-x}Al_x$ alloys in the range $0 \leq x \leq 20$ at.% Al. The present study gives decisive detail around x_c. Of importance is the sharp $\gamma(x)$ peak around $x \approx 0.025$, confirming a previous [2] conjecture in this regard.

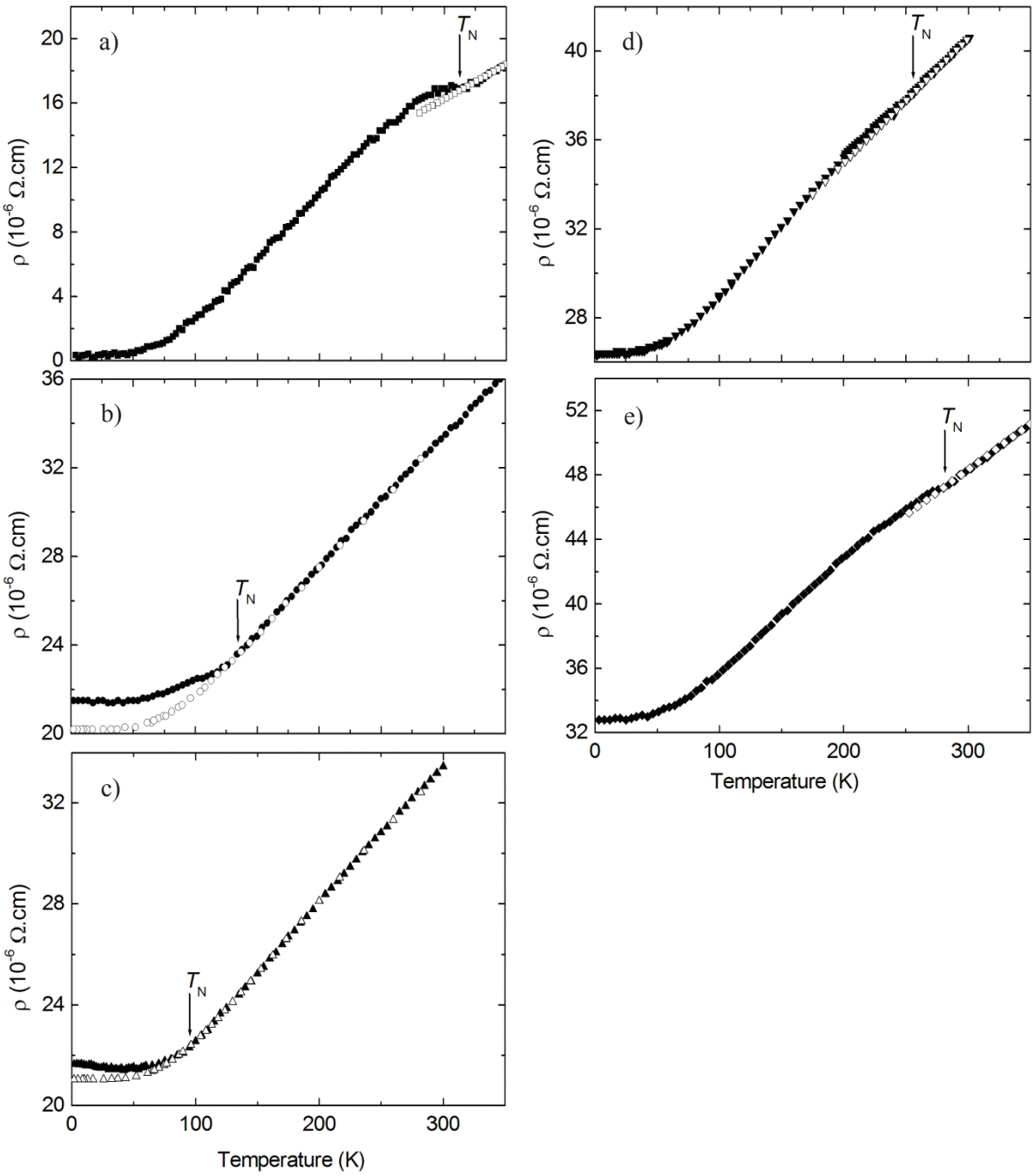


Figure 1: The electrical resistivity (ρ) as a function of temperature for $Cr_{1-x}Al_x$: (a) $x = 0$, (b) $x = 0.019$, (c) $x = 0.022$, (d) $x = 0.026$ and (e) $x = 0.034$. The arrows show the position of the Néel temperatures (T_N) obtained from magnetoelastic and neutron diffraction measurements. The open symbols for $x = 0.019$ and $x = 0.022$, depict the theoretical back extrapolated behaviour of each, should they remain paramagnetic down to 2 K [2]. The open symbols in (a), (d) and (e) only serve as "guides to the eye" back extrapolations, from the limited temperature range above the Néel transitions temperature.

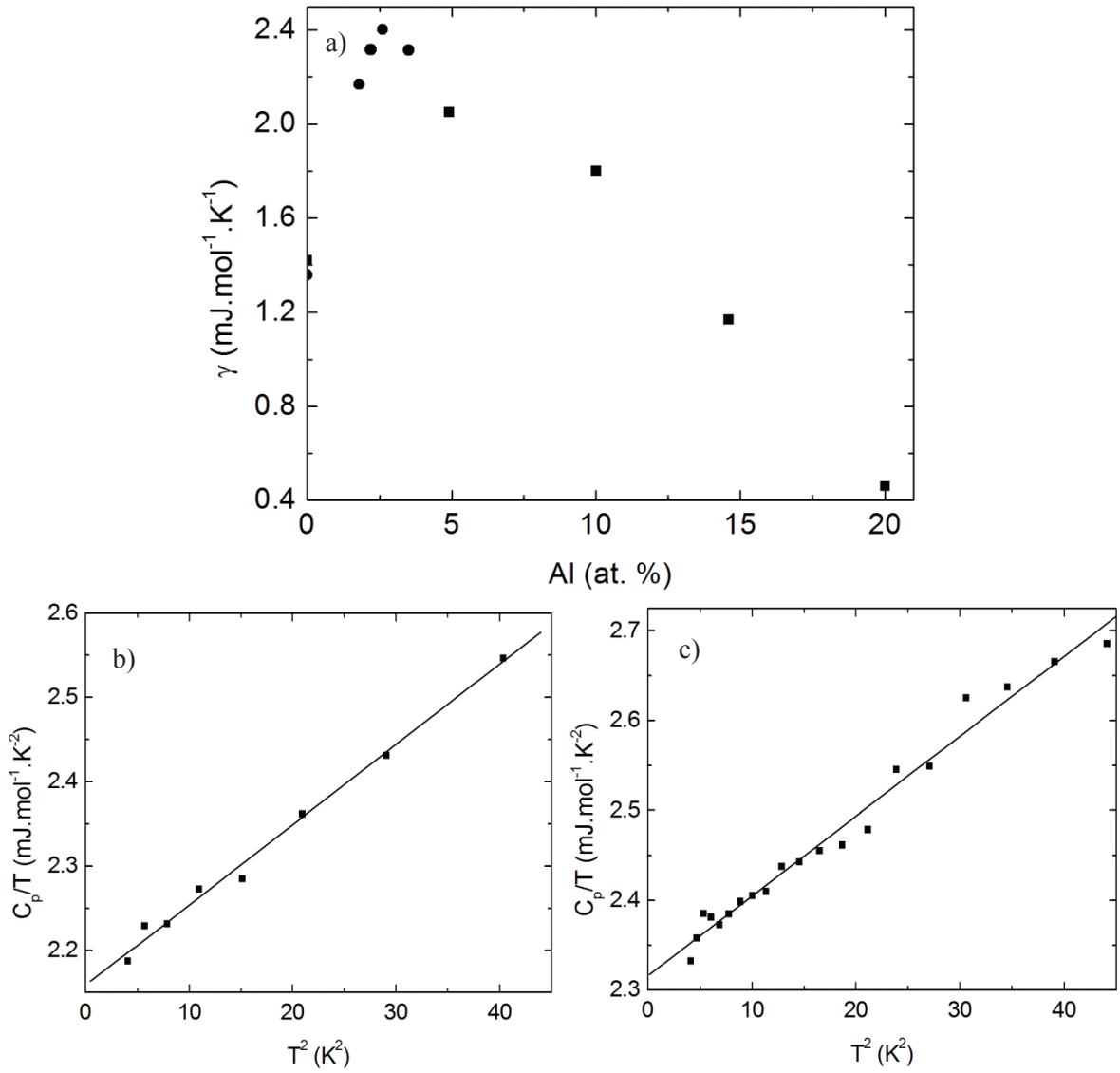


Figure 2: a) The Sommerfeld electronic specific heat coefficient (γ) as a function of Al-concentration for the $Cr_{1-x}Al_x$ alloy system: (●) present results, (■) Ref. [11]. Typical C_p/T vs. T^2 plot for b) x = 0.019 and c) x = 0.034 at low temperatures with the least-squares linear fit to the data points. The error for γ is within 0.2%.

It was previously [12] shown that the asymmetric $\gamma(y)$ peak observed in the $Cr_{1-y}V_y$ alloy system at the quantum critical point concentration, y_{QCP}, separating ISDW and P phases on the magnetic phase diagram at 0 K, is explained by combined influences of SDW energy gap and spin-fluctuation (SF) effects on $\gamma(y)$. In this system the slow decrease observed for $\gamma(y)$ in the P phase ($y \geq y_{QCP}$) is fully attributed to spin-fluctuation effects, while the sharp and large increase on increasing y up to y_{QCP} in the ISDW phase, mainly comes from effects of the SDW energy gap. We used this approach to interpret the sharp and large $\gamma(x)$ peak observed in figure 2.

In a scenario of a $Cr_{1-z}Al_z$ alloy system presenting a sequence ISDW-P-CSDW of phase transitions at 0 K on the magnetic phase diagram, one would, following the above discussion, expect $\gamma(z)$ to show two asymmetric peaks: one crossing the ISDW-P phase point and the other crossing the P-CSDW phase point on increasing z. $\gamma(z)$ should then increase sharply in the ISDW phase, followed

by a rather small variation in the P phase, and then by a down-turn as the CSDW phase is entered. Recent specific heat studies [13] in our laboratory show that this scenario is feasible in a $(Cr_{1-x}Al_x)_{95}Mo_5$ alloy system. If the width of the P phase region in this system could then be tuned to zero by tuning the Mo content, one would end up with a triple point at 0 K and a sharp peak in $\gamma(z)$, like that observed in the present $Cr_{1-x}Al_x$ system.

4. Conclusion

We presented a two-fold scenario for the behaviour of the $Cr_{1-x}Al_x$ alloy system around the triple point on the $T(x)$ magnetic phase diagram: either the phase boundary line separating ISDW and CSDW phases start at the triple point and reaches $T = 0$ K nearly vertically below this point, or this phase line is absent and the triple point is situated at or close to 0 K. The first scenario may be ruled out by the observation that a typical Cr alloy system with a triple point temperature $T > 300$K and an I-C phase line that stretches towards 0 K over a relatively large concentration range, like the $Cr_{1-v}Re_v$ system, does not indicate a significant influence of the I-C transition on $\gamma(v)$ [14]. This is in contrast with the observation in figure 2(a) for $Cr_{1-x}Al_x$. The second scenario is, however, supported by the present $\gamma(x)$ and previous neutron diffraction measurements, making $Cr_{1-x}Al_x$ a likely candidate material exhibiting the enigmatic critical point where ISDW, CSDW and P phases coexist at 0 K.

Acknowledgements

Financial support from the SA National Research Foundation (Grant No's IFR2009021800008, 61388 and 2072956) is acknowledged. A. M. Strydom acknowledges financial support from UJ-URC.

References

[1] Fawcett E, Alberts HL, Galkin V Yu, Noakes DR and Yakhmi JV 1994 Rev Mod. Phys. **66** 25
[2] Alberts HL, Prinsloo ARE and Strydom AM 2010 J. Magn. Magn. Mater. **322** 1092
[3] Takahashi J, Yamaguchi S, Fujino Y, Ozawa K, Mizuki J and Endoh Y 1980 J. Phys. Soc. Jpn. **49** 1480
[4] Smit P, Alberts HL, Venter AM 2000 J. Appl. Phys. **87** 4885
[5] Baran A, Alberts HL, Strydom AM, du Plessis P de V 1992 Phys. Rev. B**45** 10473
[6] Alberts HL and Smit P 1995 Phys. Rev. B**51** 15146
[7] Quantum Design, 6325 Lusk Boulevard CA 92121-3733, San Diego, USA]
[8] Galkin V-Y, Ortiz WA, Black R, Diederichs J, Spanga S, Fawcett E 2003 Physica B **339** 137
[9] Heiniger F, Bucher E, Muller J 1966 Phys. Kondens. Materie, **5**, 243
[10] Smit P and Alberts HL 1993 J. Phys.: Condens. Matter **5** 6433
[11] Pessal N, Gupta KP, Cheng CH and Beck P 1964 J. Phys. Chem. Solids **25** 993
[12] Takeuchi J, Sasakura H and Masuda Y 1980 J. Phys. Soc. Jpn **49** 508
[13] Muchono B 2011 private communication and presentation at this conference
[14] Heiniger F 1966 Phys kondens. Materie **5** 286

Synthesis and characterization of reproducible stoichiometry of cobalt sulfide nanoparticles using sulphur containing single-source precursors

SB Sibokoza[1], MJ Moloto[1]*, N Moloto[2]

[1]Department of Chemistry, Vaal University of Technology, Private Bag X021, Vanderbijlpark, 1900,
South Africa
[2]School of Chemistry, University of Witwatersrand, P O Wits, Braamfontein, 2050, South Africa

*Email: makwena m@vut.ac.za

Abstract. Cobalt sulphide nanoparticles are one of the most complex and challenging particles to synthesize due to strongly reducibility of the cobalt ion and the oxidizable nature of the sulphide ion. Cobalt sulphide therefore exist in various phases with different compositions including Co_4S_3, Co_9S_8, CoS, $Co_{1-x}S$, Co_3S_4, Co_2S_3 and CoS_2. Herein, we describe the synthesis of Co_xS_y nanoparticles; furthermore we interrogate the influence of the reaction conditions on the structure and optical properties of the nanoparticles. Co_xS_y nanoparticles were synthesized by thermolysis of N, N, N´, N´-tetramethylthiuram disulphide cobalt (II) complex in the presence of stabilizing agent, hexadecylamine (HDA). The advantage of this method is that the bond between the cobalt and sulphur atoms already exist, therefore decomposition of the complex at conditions will result in the formation of Co_xS_y nanoparticles. The cobalt sulfide nanoparticles were prepared by varying concentration (1.0, 0.5, 0.1 and 0.05 g) of the $CoCl_2[(CH_3)_2NCS_2S_2CN(CH_3)_2]_2$ complex at a temperature of 130 °C. The effect of temperature was also investigated by varying the temperature (80, 130, 200 and 250 °C). The nanoparticles were characterised by a combination of absorption spectroscopy, photoluminescence (PL), X-ray diffractometry (XRD) and transmission electron microscopy (TEM).

1. Introduction

Transition-metal semiconductor has shown to be one of the most active research areas [1]. Because of their particular optical, electronic, magnetic and chemical properties, these materials have been widely applied in many fields such as solid-state secondary lithium battery cathodes, industrial catalysts, solar energy conversion, and fluorescence devices [1-3]. In this study cobalt sulfide was synthesized using a single source precursor route. Single source precursors are molecules that contain all the elements required in the final material. These precursors were designed such that there is bond between metal and sulphur atoms, and including stoichiometry. The advantages of using a single source precursor route are that they reduce the chances of impurities in the final product, they minimize the synthetic time and they prevent the competition of functional groups from the surfactant [4]. In addition to that, these organometallic precursors are moisture-insensitive, air-stable, less toxic, and easy to synthesize and handle [5-7]. Semiconductor nanocrystals are synthesized from the precursor by controlled

thermal decomposition in the presence of a stabilizing surfactant [8]. Cobalt sulfides form a group of II-IV semiconductor materials with considerable potential application in electronic devices [6]. They are one of the most complex metal sulfides systems, with a number of phases and different composition. Cobalt sulfide include phases such as Co_4S_3, Co_9S_8, CoS, $Co_{1-x}S$, Co_3S_4, Co_2S_3 and CoS_2. The diversity in the stoichiometry poses some challenge in controlling the morphologies of small-(particle) [9].

2. Experimental section

2.1. Synthesis of complex

Cobalt(II) chloride tetramethyl thiuram disulfide complex was synthesized from $CoCl_2.6H_2O$ salts (4.25 mmol) which were dissolved in hot methanol (50 ml). This solution was added into the solution of tetramethylthiuram disulfide (prepared by dissolving tetramethylthiuramdisulfide (8.41 mmol) in hot methanol (50 ml)). The mixture of these solutions was heated to reflux with constant stirring for 60 minutes. The formed precipitates were filtered, washed with excess methanol and dried in open air environment. $C1_2H_{24}N_4Cl_2S_8Co$ (610.72): Anal Calc. C, 23.42; H, 3.96; N, 9.17; S, 42.00. Found: C, 22.97; H, 3.71; N, 8.12; S, 44.21. IR $\nu(cm^{-1})$: 2922.74(m), 2853.61(m), 1927.80(m), 1698.80(w), 1500.35(s),1375.87(s), 1237.79(s), 1142,53(s), 1044.90(s), 971.54(s), 955.85(w), 896.21(w), 848.38(m), 792.64(w). ^1H NMR ($CDCl_3$) δ: 3.101(s, 3H).

2.2. Synthesis of nanoparticles

Cobalt sulfide nanoparticles were synthesized at various concentrations of complex in about 5.0 g HDA. Different masses of the complex, (1.0, 0.5, 0.1 and 0.05 g) were dissolved into trioctyphosphine (TOP) (5 ml) separately. Each solution was injected into hot hexadecylamine (HDA) (5.0 g) which was heated and maintained at temperature of 130 °C under nitrogen atmosphere. This reaction was maintained at this temperature with constant starring for 30 minutes. Then, the solution was cooled to 70 °C and an excess of methanol was added. The precipitate formed were isolated by centrifugation, washed with methanol three times to remove some of the capping agent and dried in open air environment.

2.3. Characterization techniques

Several analytical techniques were used to characterize the as prepared materials. Elemental analysis (CARLO ERBA elemental analyzer for C, H, N and S), Powder X-ray diffractometry (XRD) (Phillips X'Pert materials research diffractometer using secondary graphite monochromated Cu Kα radiation (λ=1.54060 Å) at 40kV/50 mA), transmission electron microscopy (TEM) and energy-dispersive X-ray spectroscopy (EDS) (JEM-2100 JEOL electron microscope), Fourier-Transform Infrared spectra (FT-IR) (FT-IR Perkin Elmer 100 spectrometer), UV-visible spectroscopy (UV) (Analytikjena SPECORD 50 UV-Visible spectrophotometer) and Photoluminescence (PL) (LS 45 Perkin-Elmer Fluorimeter).

3. Results and discussion

The complex was prepared in a mole ratio of 1:2 between the cobalt salt and the ligand, tetramethylthiuram disulfide to produce a six coordinate with the binding to the metal through sulphide. A combination of elemental analysis and spectroscopic technique confirms the formation of the complex. Thermogravimetric analysis curve of complex (figure 1) shows that the precursor decompose into two steps between 655 °C – 826 °C and 1027 °C – 1225 °C. The observed residue was 92.85 % resulting from the loss of organic moiety from the ligand. The observed final residue was 20.47 %, which is in good agreement with the theoretical 20.15 % of CoS.

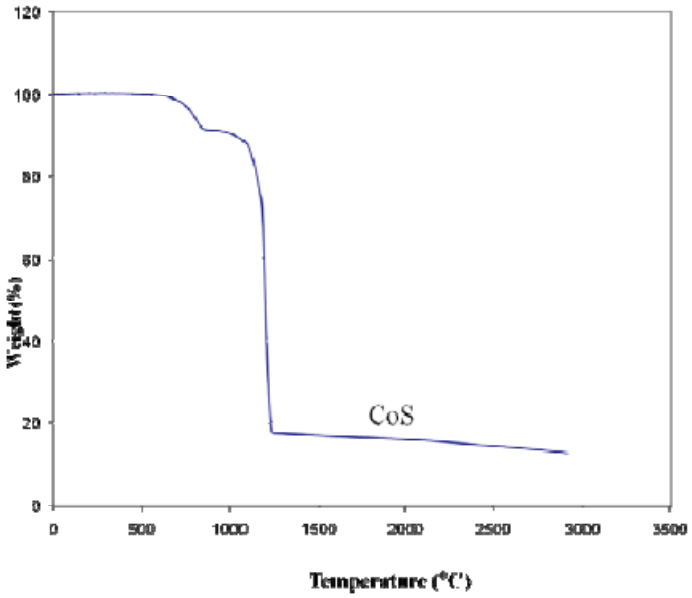


Figure 1. Thermogravimetric analysis curve for the Cobalt(II) chloride tetramethylthiuramdisulfide complex.

3.1. Optical properties of cobalt sulphide nanoparticles

Cobalt sulfide is known to have an optical band gap of 1078.26 nm with a cubic structure. The absorption spectra of cobalt sulfide nanoparticles prepared at different concentrations of complex are shown in figure 2(I). Cobalt sulfide nanoparticles prepared at high concentration (figure 2a and 2b) show two major absorption peaks, a hump at 418 nm and a maximum peak at 329 nm. The spectra results for the particles prepared at low concentration (figure 2c and d), reveals two peaks at 329 and 415 nm. The band edges of these particles are blue shifted from the bulk. Table 1 clearly shows that as concentration of the precursor decreases, the band edge also decreases. These results indicates that the size of the nanoparticles decrease with the decrease in concentration of the precursor.

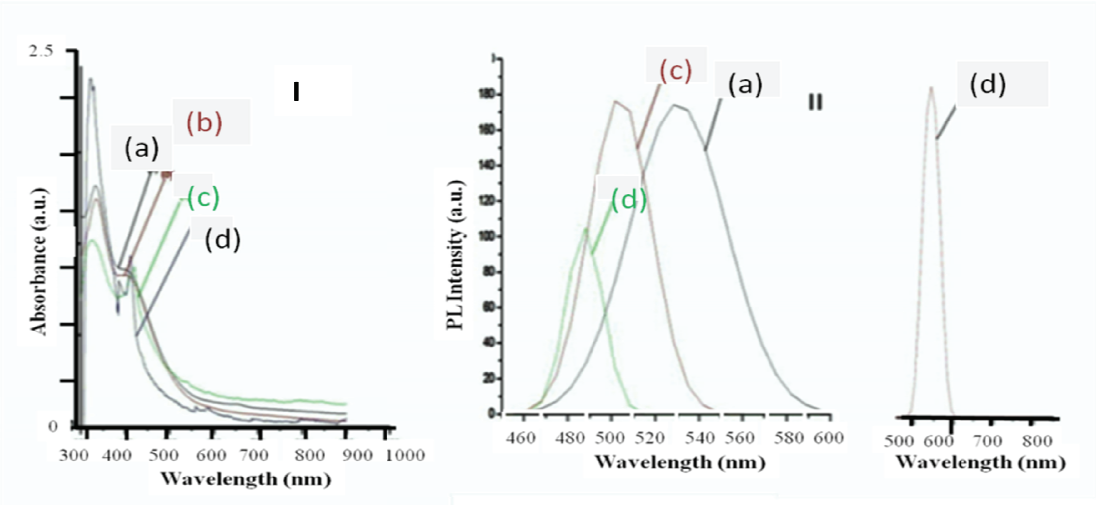


Figure 2. The absorption (I) and emission (II) spectra of Co_xS_y nanoparticles prepared at 130 °C, 30 minutes and various amount of the complex in 5.0 g HDA (a) 1.0 g, (b) 0.5 g, (c) 0.1 g, (d) 0.05 g.

Table 1: Absorption band edges and emission maxima of the cobalt sulphide nanoparticles prepared at 130 °C and 30 minutes using various amount of precursor in HDA.

Amount of precursor (g)	Absorption band edge (nm)	Emission maximum (nm)
1.0	543	565
0.5	524	547
0.1	486	506
0.05	459	483

The emission spectra of the samples prepared at various concentrations are given in figure 2(II). The emission maximum of the samples is red shifted from their respective band edges of the absorption spectra. The emission peaks have a narrow shape which indicates the monodispersity and good passivation of particles. The emission maximum shift to lower wavelength as the concentration of the precursor is decreased.

3.2. Structural characterization

Figure 3 shows the XRD patterns of the nanoparticles prepared at various concentration of precursor. Figure 3(a-d) shows four predominant peaks due to Co_3S_4 face-cantered cubic with 2θ values 15.19, 27.82, 38.45 and 65. 02. The two predominant peaks, 15.19 and 27.82 are indexed to 100 and 311 planes for cubic Co_3S_4.

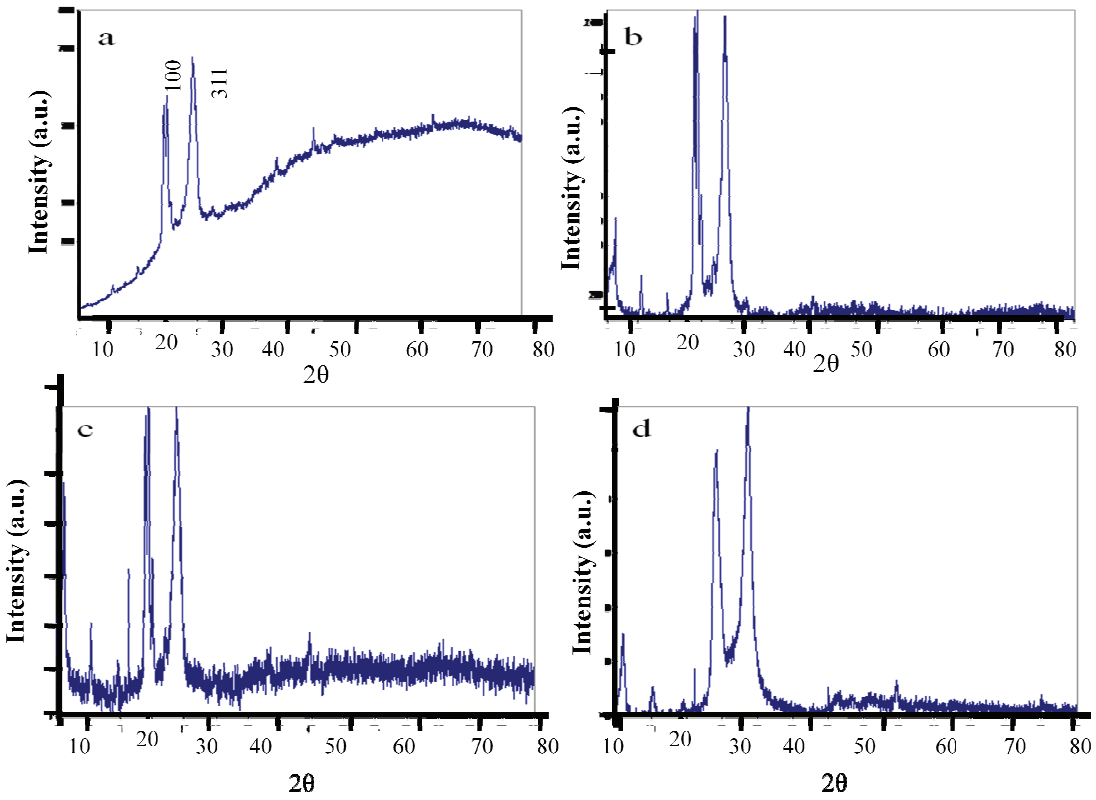


Figure 3. XRD pattern of cobalt sulfide prepared at 130 °C for 30 minutes and various amount of the complex in 5.0 g HDA (a) 1.0 g, (b) 0.5 g, (c) 0.1 g, (d) 0.05 g.

The crystallites sizes of selected cobalt sulfide nanoparticles were estimated using Scherrer equation.

$$d = 0.9 \lambda / B\text{Cos } \theta_B \quad (1)$$

In this equation, d represents the crystalline size (nm); λ is the X-ray wavelength; B is the width at half maximum of the diffraction peak at θ_B; θ_B is the diffraction angle. The average crystalline size for samples were determined to give an average values (0.05 nm) of 30.34 nm (1.0 g), 28.65 nm (0.5 g), 26.78 nm (0.1 g) and 23.94 nm (0.05 g). The TEM images of the compound prepared at various concentrations are given in figure 4(a-d). The TEM of these nanoparticles show an increased in particle sizes as the concentration of complex was increased. At concentration of 1.0 g (figure 4(a)), spherical particles were observed and the decrease to 0.5, 0.1 and 0.05 g gave mixture of spheres and rod shaped particles with the lowest concentration with irregular growth. The particle size trend conforms with the decrease from the highest to the lowest concentration promoted by the amount of precursor provided in the nucleation and growth of particles.

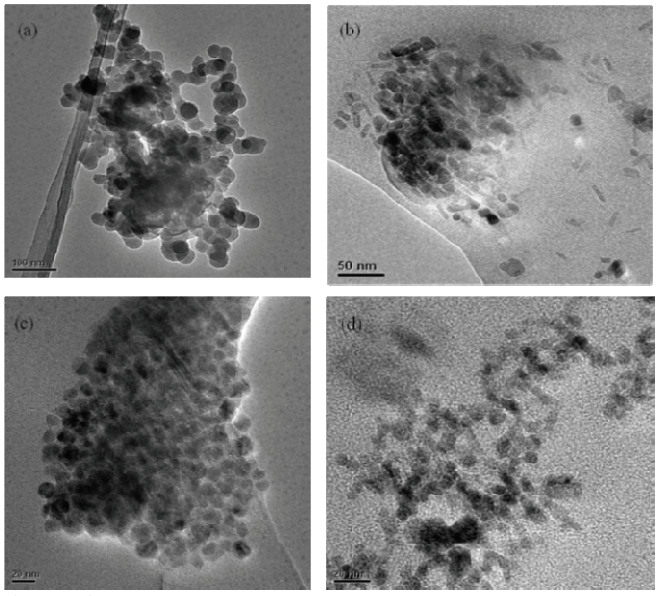


Figure 4. TEM images of cobalt sulfide prepared at 130 °C for 30 minutes and various amount of the complex in 5.0 g HDA (a) 1.0 g, (b) 0.5 g, (c) 0.1 g, (d) 0.05 g.

4. Conclusions

The cobalt sulfide nanoparticles were successfully prepared by single source precursor method. The particle size was found to decrease with the decrease of concentration of the complex. The shapes of the particles were found to be dominated by spheres at high concentration, rods and spheres at lower concentration. XRD patterns also proved that no change in formation of the stoichiometry of cobalt sulphide which indicate that concentration have little or no effect on the stoichiometry.

Acknowledgements

We would like to acknowledge the Vaal University Technology and NRF for their financial support. University of Johannesburg, University of Witwatersrand and CSIR NCNSM are thanked for the characterisation of the samples.

References

[1] Deng H, Cheng C, Peng Q and Li Y 2006 *Mat. Chem and Phys.* **100** 224

[2] Bao S-J, Li M C, Guo C-X and Qiao Y 2008 *Journal of Power Source.* **180** 676

[3] Liu B, Wei S, Xing Y, Liu D, Shi Z, Liu X, Sun X, Hou S and Su Z 2010 *J. Chem. Eur.* **16** 6625

[4] Moloto N, Revaprasadu N, Musetha P L and Moloto M J 2009 *Journal of Nanoscience and Nanotechnology*. **9** 4760

[5] Ramasamy K, Malik A M, Raftery J, Tun F and O'Brien P 2010 Chem. Mater. **22** 4919

[6] Ramasamy K, Maneerprakorn W, Malik A M and O'Brien P 2010 A. Phil. Trans. R Soc. **368** 4249

[7] Srouji F, Afzaal M, Waters J and O`Brien P 2005 *Chem. Vap. Deposition.* **11** 91

[8] Jung Y K, Kim J and Lee* J-K 2010 *J. Am. Soc.* **132** 132

[9] Bao S-J, Li Y, Li C M, Bao Q, Lu Q and Guo J 2008 *Crystal Growth & Design.* **8** 3745

The magnetocaloric effect in ferromagnetic PrSi: evidence of a novel magnetic ground state and higher order exchange interactions

J L Snyman and A M Strydom

Physics Department, University of Johannesburg, P.O. Box 524 Auckland Park, 2006, South Africa

E-mail: jlsnyman@uj.ac.za

Abstract. The discovery by Pecharsky et al. of the giant magnetocaloric effect (MCE) in $Gd_5Si_2Ge_2$, a member of the substitution series $Gd_5Si_{4-x}Ge_x$ has generated significant interest into MCE phenomena in rare earth intermetalic compounds. Much recent effort has gone into determining how the MCE in such systems is influenced by the many salient features associated with rare earth magnetism. Here the MCE in polycrystalline ferromagnetic PrSi is determined from specific heat measurements. While the magnitude of the MCE in PrSi is found to be modest when compared to other binary systems, the power of MCE analyses in probing electronic and ground-state properties in magnetic systems is emphasized in this work. We forward a description in terms of a novel magnetic ground state for the $4f$-electrons associated with the Pr^{3+}-ions in this particular compound. It is found that the MCE in this system can be accurately reproduced by modelling the system as a modified Ising-type ferromagnet with the addition of a significant higher order exchange term and assuming that the ground state of the $4f$-electrons is the unperturbed 9-fold degenerate free-ion spin-orbit coupled ground state multiplet. Such a ground state would imply that the usual splitting of the degenerate energy levels of the free ion multiplet by the crystalline electric field is absent in this system, which in turn implies that significant multipolar interactions may be responsible for the higher order exchange terms present in the magnetic Hamiltonian.

1. Introduction

The magnetocaloric effect (MCE) was first defined as a reversible temperature change induced in a material through the action of a varying magnetic field [1]. However it has now become standard practice to define the MCE in terms of two parameters describing specific trajectories within the system phase space [2]. The first is an isothermal change in entropy $\Delta S(T)_{\Delta H}$ induced in the system by magnetising the sample while maintaining good contact with a heat bath at temperature T. The second parameter is an adiabatic change in temperature $\Delta T(T_i)_{\Delta H}$ obtained by isolating the sample from its surroundings at temperature T_i and demagnetising the sample isentropically. Here H denotes the magnitude of the applied magnetic field.

Recent research regarding the MCE in rare earth intermetallic compounds may be classified in terms of two distinct objectives. The first is finding magnetocaloric materials which may be utilized for commercial applications (see for example the review by Tishin, Ref. [3]) while the second objective has been to gain a better understanding of how the MCE is influenced by some

of the features unique to rare earth magnetic systems [4]. It is within the latter context that this work is presented.

PrSi is known to order ferromagnetically at $T_C \simeq 54\,\mathrm{K}$ [5, 6]. However, the magnetic ground state of this compound has not yet been characterised in detail. The compound crystallizes in the orthorhombic FeB-structure (space group *Pnma*), as do the other members of the RSi (R = La-Sm, Gd-Tb) family of compounds. Within this family PrSi and NdSi are the only two members to exhibit ferromagnetic order [5]. Similar behaviour in the RB_4 series of compounds has been attributed to significant orbital angular momentum contributions to the magnetic moments localised at rare earth crystallographic sites [7]. However, from symmetry considerations alone, it is expected that the 9-fold degenerate $J = 4$ spin-orbit coupled ground state multiplet associated with the $4f$-electrons of the free Pr^{3+}-ion should be split into nine non-degenerate (and typically non-magnetic) singlet states through the action of the crystalline electric field (CEF) [8]. In such systems the emergence of cooperative magnetic behaviour requires a secondary interaction mechanism. One such example is the establishment of ferromagnetic order through the formation of a pseudo doublet, produced through the admixture of the two lowest lying singlet states. Such a mechanism has been shown to produce magnetic order in TmNi (see Ref. [8]), which also crystallises in the orthorhombic FeB-structure.

In this work we show that the specific heat of PrSi is suggestive of the full nine fold degenerate $J = 4$ multiplet being the correct $4f$-electron ground state in this system, contrary to the point charge model calculations reported in Ref. [8]. While the magnetic configurational entropy is amenable to our proposed ground state, the strongest evidence supporting our claim is found by analysing the MCE in the system in conjunction with a simple numerical model.

2. Experimental procedure

Two PrSi samples and one sample of the isostructural non-magnetic reference compound LaSi were prepared by melting together stochiometric amounts of the constituent elements in an arc-furnace under argon atmosphere. The first PrSi sample (hereafter referred to as sample A) was prepared from 99.99 wt.% purity Pr and 99.9999 wt.% Si. The second sample (hereafter sample B) was prepared from 99.9 wt.% purity Pr and 99.9999 wt.% Si. The LaSi sample was prepared from of 99.9 wt.% La and 99.9999 wt.% purity Si. During the melt samples were flipped and remelted several times to ensure homogeneity.

Powder X-ray diffraction was used to determine the phases present in all samples. X-ray diffraction was carried out on a Philips Panalytical X'pert Pro Instrument at room temperature using Cu radiation ($\lambda = 1.542\,\text{Å}$) for the incident beam. The initial characterisation revealed that sample A had formed in the desired structure, whilst sample B and the LaSi sample both had small amounts of an impurity phase present in their respective diffractograms. Subsequently both sample B and the LaSi sample were annealed at 1273 K for 10 days. In preparation for annealing, samples were wrapped in tantalum foil and sealed in evacuated quartz tubes. X-ray diffraction showed the annealed samples to be predominantly single phase, with trace amounts of an unidentified impurity phase present in the diffractograms of both samples.

The specific heat results are derived from heat capacity measurements executed using the heat capacity measurement option of a Physical Properties Measurement System [9].

The numerical results are calculated from a model magnetic Hamiltonian using a simple Metropolis algorithm [10] on a $10 \times 10 \times 10$-element cubic lattice.

3. Results and discussion

Figure 1 shows the X-ray diffractogram obtained for the as-cast sample of PrSi (sample A) together with a simulated diffractogram obtained for the orthorhombic *Pnma* FeB-type structure. The simulated diffractogram was obtained by using the lattice parameters and unit cell data reported for the FeB-structure type in [5, 11] as input to a full profile Rietfeld refinement

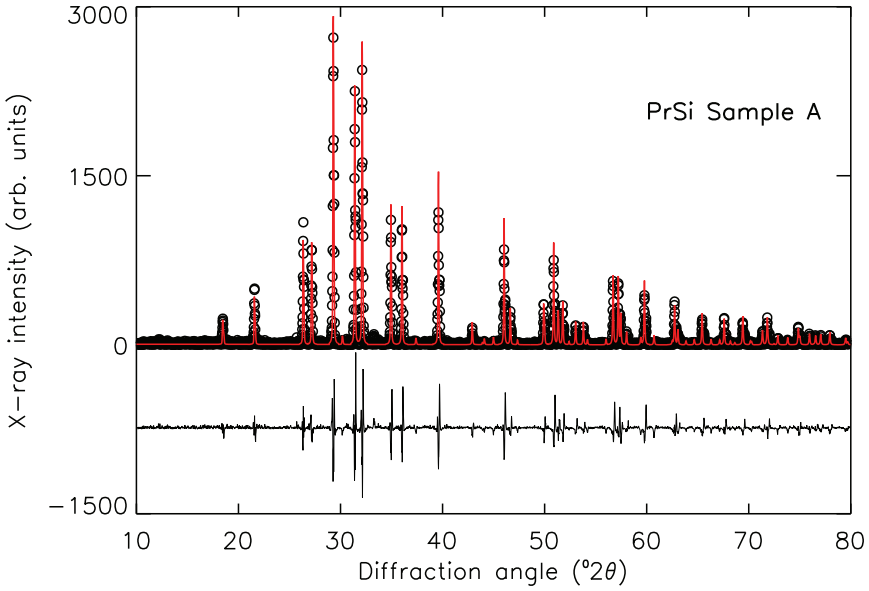


Figure 1. Powder X-ray diffraction spectrum (open circles) for PrSi sample A together with the spectrum calculated from a full profile least squares Rietveld refinement fit (solid line). The difference between the observed and simulated data is shown underneath the diffractogram.

program (GSAS, [12, 13]). The lattice parameters used to calculate the refined diffractogram are $a = 8.24362\,\text{Å}$, $b = 3.94167\,\text{Å}$ and $c = 5.92000\,\text{Å}$. A comparison between the simulated and measured diffractograms shown in figure 1 allows us to conclude that the sample formed in the desired structure. With the exception of a minor unknown impurity phase, similar results were obtained for sample B as well as the LaSi sample. As is shown below, the pressence of the latter impurity phase has a negligible effect on the specific heat of sample B.

The specific heat ($C_{p,\text{PrSi}}$ and $C_{p,4\,\text{kOe}}$) of PrSi in applied fields of zero and 4 kOe together with the specific heat of LaSi ($C_{p,\text{LaSi}}$) are shown in figure 2. The magnetic phase transition reported earlier [5, 6] is visible in the PrSi zero field specific heat as the prominent peak at $T_C = 51.74\,\text{K}$. The inherent ferromagnetic nature of this transition is evidenced by the behaviour of the specific heat in an applied field of 4 kOe. The magnetic field acts to suppress and broaden the peak while simultaneously shifting the apex towards higher temperatures.

An attempt was made to calculate the contribution to the specific heat from the $4f$-electron magnetic configuration and thermal population of CEF-split levels ($C_{p,4f}$). The LaSi specific heat ($C_{p,\text{LaSi}}$) was subtracted from the zero field specific heat of PrSi so that $C_{p,4f}(T) \simeq C_{p,\text{PrSi}}(T) - C_{p,\text{LaSi}}(T)$. A small Sommerfeld contribution ($C_s = \gamma T$ with $\gamma = 4 \pm 3\,\text{mJ.mol}^{-1}.\text{K}^{-2}$) could be identified and subtracted from $C_{p,4f}(T)$. As shown in figure 3 $C_{p,4f}(T)$ does not show the typical Schottky contribution [14] associated with the thermal population of higher lying CEF-split energy levels, implying that such a splitting is completely absent from the system or that the experimental temperature range ($T \leq 300\,\text{K}$) is less than a putative first thermal CEF excitation in PrSi. In order to test the first possibility it was assumed that $C_{p,4f}(T)$ is a purely magnetic configurational contribution to $C_{p,\text{PrSi}}$, allowing the calculation of the zero field magnetic configurational entropy $S_{4f}(T)$ from the general equation

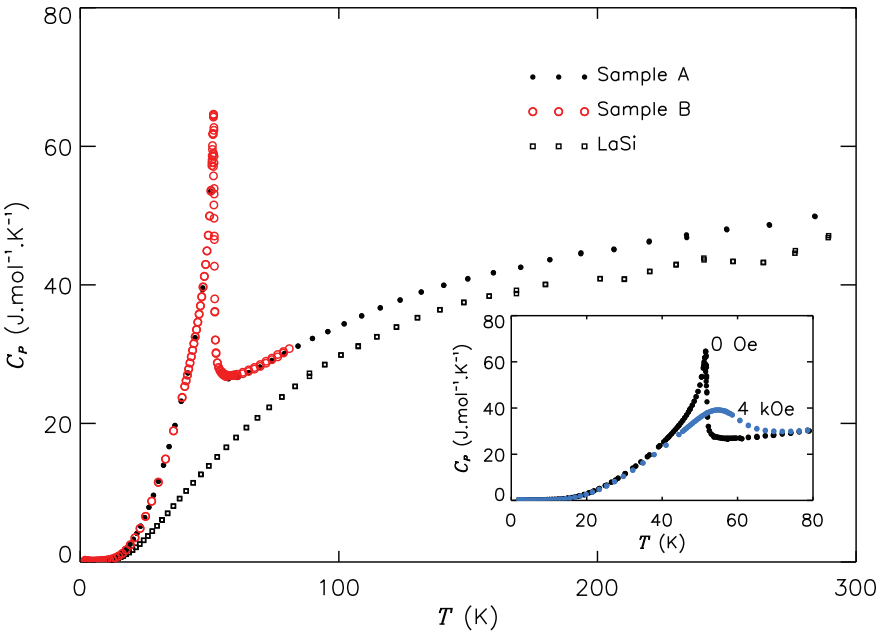


Figure 2. Specific heat C_p of PrSi and LaSi in zero applied field. Inset: PrSi specific heat in 0 and 4 kOe.

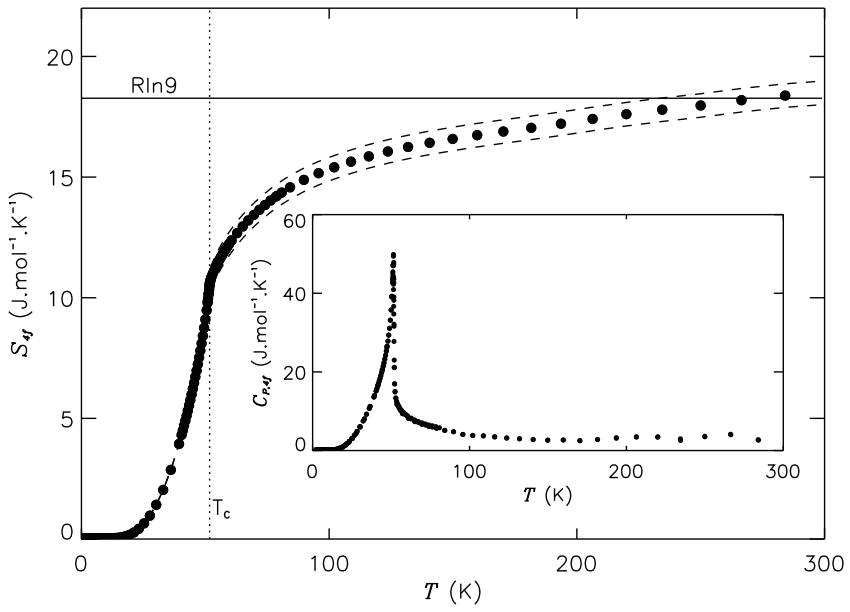


Figure 3. The $4f$-electron magnetic configurational entropy in PrSi. Inset: The $4f$-electron magnetic configurational specific heat contribution.

$$S_H(T') = \int_0^{T'} \frac{C_{p,H}(T)}{T} dT \tag{1}$$

The results are shown in figure 3. For a system containing magnetic moments where each moment has N possible orientations the saturation magnetic configurational entropy is equal to $R \ln N$. Saturation occurs at a temperature where the population of all possible orientations become statistically equally probable. In PrSi the magnetic configuration entropy at T_C is already well in excess of $R \ln 2 = 5.76\,\mathrm{J.mol^{-1}.K^{-1}}$, ruling out the possibility of a pseudo doublet ground state. While the saturation magnetic configurational entropy associated with the $J = 4$ spin-orbit coupled free-ion ground state multiplet is approached at high temperatures, figure 3 shows that the entropy does not saturate within the experimental temperature region. This implies that the description of the non-magnetic specific heat contribution adopted earlier may not be adequate, or that there are non-negligible magnetic interactions present above $T = T_C$ that inhibit saturation. A third reason might be optical lattice excitations present in LaSi. This would imply that $C_{p,4f}(T)$ is underestimated in our calculation.

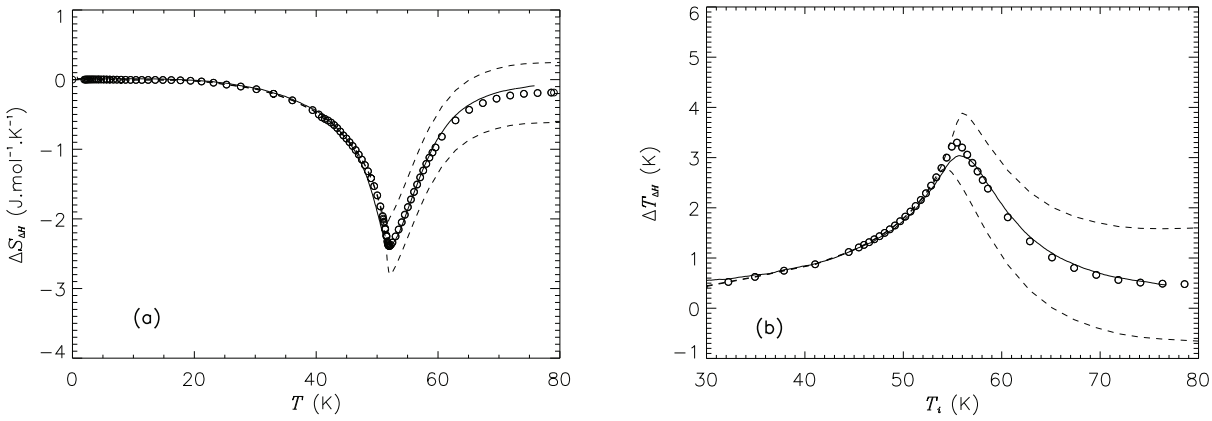


Figure 4. Isothermal entropy change upon magnetisation from 0 to 4 kOe (a) and adiabatic temperature change upon demagnetisation from 4 to 0 kOe (b). Uncertainties due to experimental errors are shown as dashed lines. Modelled results are shown by the solid lines.

In order to confirm the ground state of the $4f$-electrons in the system the MCE is calculated from the specific heat data (according to the method given in Ref.[15]) and analysed as is discussed below. Eq. 1 is used to calculate the isofield entropy from the specific heat data shown in figure 2 from which $\Delta S_{\Delta H}(T)$ and $\Delta T_{\Delta H}(T_i)$ are then calculated. The MCE thus calculated is shown in figure 4.

The shape of the MCE in the vicinity of a magnetic phase transition depends on the order of the transition [2], the behaviour of the correlation length associated with emerging magnetic order [16] as well as other site specific magnetic interactions such as those between localised moments and the CEF[4]. Through the Zeeman interaction the magnitude of the MCE depends strongly on the quantity $g \sum_i J_{z,i}$, where g is the Landé g-factor and $J_{z,i}$ is the component of the total angular momentum parallel to the quantization axis at site i.

Taking the above into account, an attempt was made to model the MCE in PrSi. Initial attempts to model the MCE in PrSi within the mean field [4] and pure Ising-model frame work assuming a pseudo doublet ground state failed to reproduce the experimental results accurately. The absence of CEF-effects in the $4f$-electron specific heat ruled out the addition of CEF-terms to the magnetic Hamiltonian. Instead, with reference to the $R\mathrm{B}_4$ series of compounds noted earlier, we added higher order exchange terms to the magnetic Hamiltonian, in an attempt to

better reproduce the MCE in PrSi. The best results were obtained by calculating the MCE in PrSi from a modified Ising model Hamiltonian containing higher order exchange terms, so that

$$\mathcal{H} = g\mu_B B \sum_i J_{z,i} + \mathcal{J} \sum_i \sum_j J_{z,i} J_{z,j} + \zeta \sum_i \sum_j J_{z_i}^2 J_{z,j}^2 \tag{2}$$

The first term in the Hamiltonian denotes the Zeeman interaction, where the assumption is made that the magnetic moments are test particles in a magnetic field $B = \mu_0 H$. The second term is the traditional Ising-model exchange interaction between nearest neighbour sites. The third term is a higher order exchange term, also limited to nearest neighbour sites, and may originate from magnetoelastic exchange [4] or multipolar [17] interactions. It is found that for $g = 0.45$, $\mathcal{J} = 0.076\,\text{meV}$ and $\zeta = 0.011\,\text{meV}$ the magnetocaloric effect in PrSi can be reproduced accurately by assuming the ground state of the system incorporates the full $J = 4$ multiplet, as shown in figure 4.

4. Conclusion

The $4f$-electron contribution to the specific heat of PrSi suggests the likelihood of a spin-orbit coupled free-ion $J = 4$ multiplet ground state in PrSi. A pseudo-doublet ground state can be ruled out on account of the value of the magnetic configurational entropy at the phase transition temperature. The MCE in this compound can be accurately reproduced if a $J = 4$ multiplet ground state is assumed, CEF-effects are neglected and higher order exchange terms are added to the magnetic Hamiltonian. These latter terms are necessitated by the large orbital contribution to the magnetic moment implied by a $J = 4$ ground state, and consistent with the ferromagnetic order exhibited by PrSi in a family of otherwise predominantly antiferromagnetic compounds, analogous to the $R\text{B}_4$ series. Detailed analyses on a single-crystal specimen that would allow the incorporation of magnetocrystalline anisotropy effects into our analyses are in progress.

Acknowledgements

JLS would like to acknowledge the financial support received through both the University of Johannesburg New Generation Scholarship program, as well as the NRF DoL Scarce Skills bursary program. AMS thanks the NRF (2072956) and the UJ-URC for financial support.

References

[1] Bitter F 1931 *Physical Review* **38** 528
[2] Pecharsky V K and Gschneidner, Jr K A 1999 *Journal of Applied Physics* **86** 565
[3] Tishin A M and Spichkin Y I 2003 *The Magnetocaloric Effect and its Applications* (IoP)
[4] de Oliveira I G and von Ranke P J 2010 *Physics Reports* **489** 89
[5] Nguyen V N, Tchéou F and Rossat-Mignod J 1977 *Solid State Communications* **23** 821
[6] Pinguet N, Weitzer F, Schuster J C and Noël H 2003 *Journal of Alloys and Compounds* **348** 1
[7] Buschow K H J 1980 *in: E.P. Wohlfarth (Ed.),Ferromagnetic Materials Volume 1* (North-Holland Publishing Company)
[8] Gignoux S, Rossat-Mignod J and Tcheou F 1972 *Physica Status Solidi* a **14** 483
[9] Hwang J S, Lin K J and Tien C 1997 *Review of Scientific Intruments* **68** 94
[10] Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H and Teller E 1953 *The Journal of Chemical Physics* **21** 1087
[11] Shobinger-Papamantellos P and Buschow K H J 1985 *Journal of the Less Common Metals* **111** 125
[12] Larson A C and Von Dreele R B 1994 *Los Alamos National Laboratory Report* **LAUR** 86–748
[13] Toby B H 2001 *Journal of Applied Crystallography* **34** 210
[14] Pashchenko V A, Borisenko V G, Khatsko E N, Cherny A S, Paulsen C, Feher A, Orendac M, Orendacova A, Jansen A G M and Wyder P 2002 *Journal of Physics: Condensed Matter* **14** 9693
[15] Pecharsky V K and Gschneidner, Jr K A 1997 *Physical Review Letters* **78** 4494
[16] Franco V, Blazquez J S and Conde A 2006 *Applied Physics Letters* **89** 222512
[17] Fedders P A and Myles C W 1979 *Physical Review* B **19** 1331

Crystalline electric field effects in $PrNiGe_2$

J L Snyman and A M Strydom

Physics Department, University of Johannesburg, P.O. Box 524 Auckland Park, 2006, South Africa

E-mail: jlsnyman@uj.ac.za

Abstract. Within the confines of the standard model of rare earth magnetism, the high temperature magnetic properties of rare earth ions in a crystalline environment are identical to the magnetic behaviour of free tri-positive rare earth ions. At low temperatures deviations from free-ion behaviour occur due to the action of an anisotropic crystalline electric field (CEF). For Pr^{3+}-ions occupying sites with orthorhombic point symmetry it is expected that the 9-fold degeneracy of the spin orbit coupled ground state multiplet associated with the free tri-positive ion should be completely uplifted by the CEF yielding a non-magnetic singlet ground state for the Pr^{3+} 4f-electrons. $PrNiGe_2$ is known to order ferromagnetically despite the fact that the Pr^{3+} ions occupy the low-symmetry $m2m$ sites in the $CeNiSi_2$-type structure. This has prompted the current study into the ground state properties of $PrNiGe_2$. We determined the structure of the CEF-split energy levels in this system from specific heat measurements. The analyses points to the formation of a pseudo-doublet ground state and we forward a conceptual explanation of the observed magnetic order in $PrNiGe_2$ in terms of a fortuitous merging of electronic singlet levels into a local level dispensation of higher degeneracy.

1. Introduction

The $4f$-electron ground state of free rare earth ions are determined by the spin-orbit interaction, which yields for the $4f$-electrons degenerate multiplets, well separated in energy, and characterised by the total angular momentum J. The ground state multiplet of free Pr^{3+} is nine fold degenerate, characterised by the total angular momentum number $J = 4$ and is separated from the first excited multiplet by an energy (in temperature units) of $\Delta/k = 3100\,\mathrm{K}$ [1]. However, the ground state of the rare earth $4f$-electrons is subject to additional interactions, one of which is the crystalline electric field (CEF) which adds non-spherically symmetric contributions to the $4f$-electron Hamiltonian [2], and which acts to uplift (either fully or partially) the degeneracy of the free-ion multiplet. The nature of this splitting depends crucially on the point symmetry of the rare-earth ion crystallographic site and on the orbital contribution to the free-ion $4f$-electron ground state.

$PrNiGe_2$ crystallises in a $CeNiSi_2$-type structure (space group $Cmcm$) [3, 4]. Within this structure the Pr^{3+}-ions occupy the $4c$ site (with orthorhombic $m2m$ site symmetry). For such a low symmetry site it is expected that the CEF should completely uplift the degeneracy of the free-ion ground state multiplet [5, 6]. However, a previous study has shown that $PrNiGe_2$ orders ferromagnetically at a Curie point of $T_C = 15\,\mathrm{K}$ [4]. Thus far very little is known about the ground state of this compound, the nature of the CEF-splitting, or the type of ferromagnetic order exhibited. Here we present a first investigation into these properties. We attempt to determine the CEF-levels and their degeneracies from heat capacity measurements

on a polycrystalline sample of $PrNiGe_2$ and show that the $4f$-electron ground state in this compound is a doublet.

2. Experimental procedure

A polycrystalline sample of $PrNiGe_2$ was prepared by melting together stoichiometric amounts of the constituent elements in an arc furnace under a purified argon atmosphere. The sample was prepared from 99.9 wt.% purity Pr, 99.99 wt.% Ni and 99.999 wt.% Ge. During the melt the button was flipped over and remelted several times to ensure homogeneity.

Powder X-ray diffraction was carried out on a Philips Panalytical X'Pert Pro instrument at room temperature. The sample stage is equipped with a sample spinner rotating the sample around the scattering vector to negate possible effects of preferred orientation in the powder. Cu radiation (Kα, $\lambda = 1.542\,$Å) was used for the incident beam.

The specific heat of $PrNiGe_2$ was measured by utilizing the Heat Capacity option on a Physical Properties Measurement System [7]. For details regarding the method used, see Refs. [8, 9].

3. Results and discussion

Figure 1 shows the X-ray diffractogram obtained for $PrNiGe_2$ together with a simulated diffractogram obtained from the orthorhombic *Cmcm* $CeNiSi_2$-type structure. The simulated diffractogram was obtained by using the lattice parameters and unit cell data reported in [4] as input to a full profile Rietveld refinement program, and refining the lattice parameters in order to obtain the best fit between the observed and calculated diffractograms. The Rietveld refinement was done by utilising the General Structure Analyses System (GSAS) [10, 11]. The lattice parameters used as input for the simulated data are $a = 4.23119\,$Å, $b = 16.7813\,$Å and $c = 4.18974\,$Å. A comparison between the calculated and observed diffractograms in figure 1 shows that the sample formed in the desired structure.

The specific heat of $PrNiGe_2$ is shown in figure 2. The low temperature anomaly (reaching a maximum of $17.90\,\mathrm{J.mol^{-1}.K^{-1}}$ at 12.33 K) is associated with the onset of magnetic order. Magnetisation measurements classified the ordering as ferromagnetic [4]. At temperatures well in excess of 12 K the specific heat of $PrNiGe_2$ resembles that of the well known Debye-model, implying that the specific heat is dominated by a phononic contribution at these temperatures. In order to isolate this contribution the Debye model specific heat (for acoustic modes only) was fitted to the high temperature data in figure 2. Subtraction of the phonon contribution yielded the $4f$-electron contribution to the specific heat similar to the data shown in the inset of figure 2.

After subtracting the phonon contribution the specific heat showed a clearly defined Schottky-anomaly (with a peak at $\sim 50\,$K), indicating the thermal population of CEF-split energy levels. The Schottky anomaly could not be accurately described using the well known two- or three level descriptions often found in literature (see for example Ref. [12]). In order to adequately parameterise the anomaly a first principles calculation involving up to the full set of nine CEF-levels for Pr^{3+} had to be done. As is discussed below this procedure not only accurately parameterised the Schottky-anomaly but also constrained the CEF-parameters to some extent.

Theoretically the Schottky contribution arises from the thermal population of a number of discrete energy levels. This contribution can readily be calculated from a partition function of the form

$$\mathcal{Z} = \sum_{i=0}^{n} g_i e^{-E_i/kT} \tag{1}$$

where E_i is the energy of the i'th level with degeneracy g_i, k is Boltzman's constant and T is the temperature of the system. The Schottky contribution to the specific heat was found

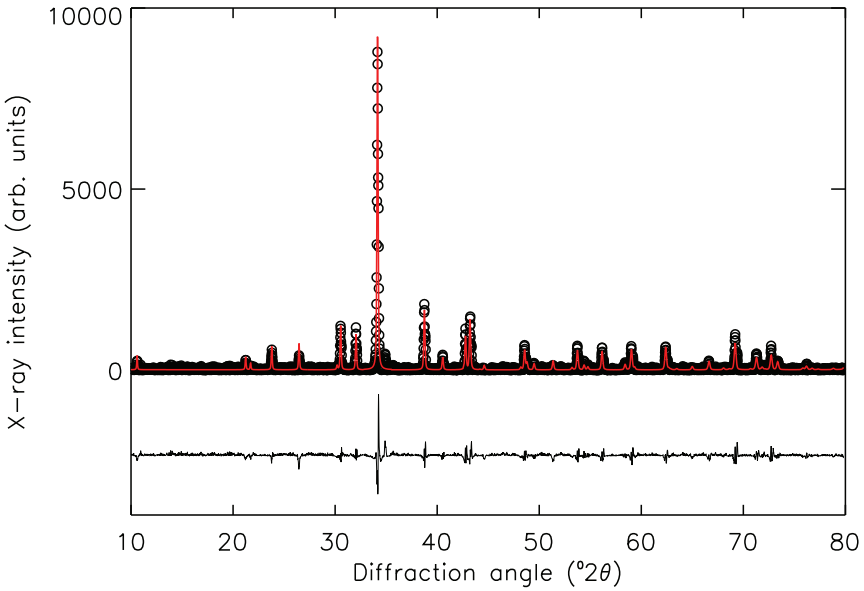


Figure 1. The powder X-ray diffraction spectrum for $PrNiGe_2$ (open circles) together with the spectrum calculated from a full profile least squares Rietveld refinement fit (solid red line). The difference between the simulated and observed data is shown below the diffractogram.

to extend up to the highest measured temperatures (250 K) . Therefore it was not deemed sufficient to fit a Debye and Sommerfeld contribution to the specific heat independently from the Schottky contribution (as was indeed done in the first analyses, mentioned above), and hence in order to parameterise the phononic (Debye), electronic (Sommerfeld) and Schottky contributions to the specific heat a Monte-Carlo algorithm was used to optimize all parameters involved simultaneously.

Initially the algorithm generated random instances of the Debye temperature (between 250 K and 350 K), Sommerfeld coefficients (between 0 and 0.1 $J.mol^{-1}.K^{-2}$) and nine values of E_i ($E_i/k \in [0\,K, 1000\,K]$), subject to the constraint that $E_0/k = 0\,K$. It was assumed that $g_i = 1$ for all i in Eq.1. This was done in order to find the most general description of the CEF-split levels in the compound. For example, if the CEF-split levels are one doublet and seven singlets, then the algorithm should yield two values of E_i which overlap within the error margins, while the other E_i values should show no overlap. Lastly, the large range for the E_i, $i = 0, 1, 2, ..., 8$ was used to compensate for the possibility that the total CEF splitting is much larger than the experimental temperature range considered. For example, if the Schottky contribution to the specific heat can be described by six singlets alone, then it is expected that the parameter set that yields the best fit should contain three levels at energies which are sufficiently high to negate any contribution at 250 K.

The algorithm was executed by generating 200000 random instances of each parameter (Debye temperature θ_D, Sommerfeld coefficient γ and nine energy level values E_i). After each instance the average difference $\langle\sqrt{(C_e - C_m)^2}\rangle$ between the experimental specific heat C_e and the modelled specific heat C_m was calculated over the range $17 \leq T \geq 250\,K$. The lower limit of 17 K was chosen to ensure that the magnetic contribution to the specific heat does not affect the calculation. For every 200000 instances the parameter set that gave the best fit was chosen. The whole procedure was repeated 30 times, yielding the 30 best parameter sets given the constraints

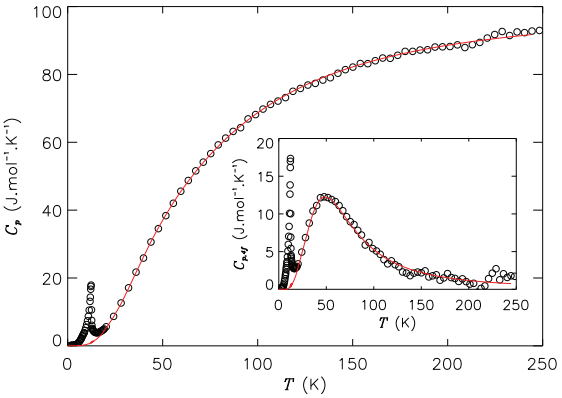


Figure 2. Total measured specific heat of $PrNiGe_2$ (open symbols). The solid line is a fit of the combined effect of the Schottky, Debye and Sommerfeld contributions to the specific heat of the compound calculated for $n = 7$ as is described in the text. Inset: $4f$-electron contribution to the specific heat of $PrNiGe_2$.

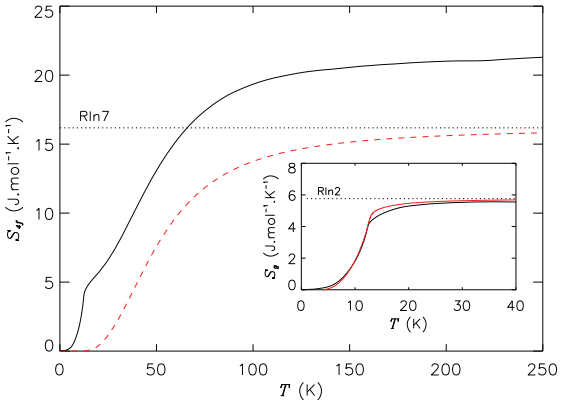


Figure 3. $4f$-electron contribution (solid line) to the entropy of $PrNiGe_2$ as calculated from the inset of figure 2. The dashed line corresponds to the entropy of the system due to thermal population of the energy levels listed for $n = 7$ in table 1. Inset: Magnetic configurational entropy, calculated by subtracting the Schottky-contribution from the total $4f$-entropy. Also shown is the magnetic configurational entropy of an isotropic three dimensional Ising-ferromagnet (red line).

mentioned above.

The data obtained from this simultaneous optimization immediately showed that the Debye temperature lies between 290 K and 320 K, irrespective of how many energy levels were considered in the calculation. Therefore, the procedure noted above was repeated by constraining the instances of θ_D to lie between 290 K and 320 K. The data thus obtained is summarised in table 1, where n denotes the number of discrete energy levels assumed to contribute to the partition function in Eq.1.

The data in table 1 requires careful analyses. For $n = 6, 7, 8$ and 9 the Schottky contribution to the specific heat could be accurately parameterised (within $\sim 0.45\,\mathrm{J.mol^{-1}.K^{-1}}$). However using fewer than six energy levels to calculate the Schottky contribution yields significantly poorer fits. Therefore it can be concluded that the Schottky anomaly seen in the measured specific heat data arises from the thermal population of at least six CEF-split singlets. While the data in table 1 does constrain the number of levels to some extent it does not adequately constrain the degeneracies of these levels. For example, for $n = 8$ E_5 and E_6 lie close enough together (given the uncertainties implied by the average deviation) to be either two close lying singlets, or a single doublet. However the data in table 1 allows the accurate *parameterisation* of the Schottky, Debye and Sommerfeld contributions to the $4f$-electron entropy as is shown next.

In general the entropy of a system can be calculated from its specific heat through

$$S(T')_H = \int_0^{T'} \frac{C_{P,H}(T)}{T} dT + S_0(H) \qquad (2)$$

where $C_{P,H}(T)$ is the specific heat in constant pressure P and applied field H at temperature

Table 1. The data obtained by simultaneously fitting the Debye temperature Θ_D, Sommerfeld coefficient γ and n energy levels E_i, $i = 0, 1, 2, ..., n-1$ to the total measured specific heat of $PrNiGe_2$. For every value of n the values in the left column denote the average taken over all 30 parameter sets noted above (with average deviations from the mean shown in brackets). $\langle \delta C \rangle = \langle \sqrt{(C_e - C_m)^2} \rangle$ is calculated as mentioned in the text and has units of $J{\cdot}mol^{-1}{\cdot}K^{-1}$. Θ_D and the E_i/k values are all quoted in Kelvin. The γ values shown in the table have units of $mJ{\cdot}mol^{-1}{\cdot}K^{-2}$.

Param.	$n = 9$	$n = 8$	$n = 7$	$n = 6$	$n = 5$
Θ_D	315 (2)	313 (2)	309 (1)	304 (2)	298 (2)
γ	5 (1)	5 (1)	5 (1)	4 (1)	3 (1)
E_0/k	0	0	0	0	0
E_1/k	99 (10)	105 (13)	111 (10)	112 (10)	115 (6)
E_2/k	130 (9)	134 (8)	130 (7)	134 (8)	125 (6)
E_3/k	156 (12)	148 (9)	147 (9)	146 (7)	141 (6)
E_4/k	172 (8)	170 (7)	166 (12)	162 (9)	159 (9)
E_5/k	187 (8)	187 (11)	187 (12)	181 (10)	-
E_6/k	208 (11)	208 (12)	225 (24)	-	-
E_7/k	227 (11)	249 (27)	-	-	-
E_8/k	296 (42)	-	-	-	-
$\langle \delta C \rangle$	0.431 (0.015)	0.441 (0.017)	0.456 (0.013)	0.471 (0.014)	0.531 (0.014)

T and $S_0(H)$ is the specific entropy at absolute zero [13]. In calculating the entropy it has been assumed that $S_0(H) = 0$ for all H and that

$$S(T')_H = \int_0^{T'} \frac{C_{P,H}(T)}{T} dT = 0 \text{ for } T' < T_0 \tag{3}$$

where T_0 is the lowest temperature at which experimental data is available. Shown in figure 3 is the total $4f$-contribution to the entropy for the case where $n = 7$. The inset shows the entropy after the Schottky, Debye and Sommerfeld contributions have been subtracted. Given the magnetic phase transition at $\sim 12\,K$, the entropy in the inset of figure 3 is taken to represent the magnetic configurational entropy associated with the onset of ferromagnetic order as is indeed suggested by the similarity with the magnetic configurational entropy of a three dimensional isotropic Ising-ferromagnet which is also shown. Clearly the magnetic configurational entropy saturates above T_C at a value close to $R \ln 2\,J.mol^{-1}.K^{-1}$, which suggest a magnetic doublet ($J = 1/2$) dispensation in the ground state.

It is noted that a similar ground state is found for all parameter sets $n = 6, 7, 8$ and 9 in table 1 as is summarised in table 2. Therefore it can be safely concluded that the magnetic ground state of $PrNiGe_2$ is a doublet. The usual mechanism by which magnetic order is established in low symmetry Pr-based compounds is through the admixture of two low-lying singlets, given that the exchange energy exceeds a critical value [14]. For example, such a mechanism has been shown to produce an Ising-like doublet ground state in TmNi [15].

4. Conclusion

A Monte Carlo algorithm permitted the parameterization of the Schottky, Debye and Sommerfeld contributions respectively, weighed against experimental data of the specific heat of $PrNiGe_2$. A subsequent calculation of the magnetic configurational entropy revealed that

Table 2. The saturation entropy S_{sat} calculated for each parameter set in table 1 and the percentage of the theoretical saturation value (equal to $R\ln 2$) represented by S_{sat}.

n	S_{sat} (J.mol^{-1}.K^{-1})	% of $R\ln 2$
9	5.43	94.2
8	5.54	96.2
7	5.56	96.4
6	5.52	95.8
5	5.44	94.5

the magnetic ground state is a doublet. The parameterisation of the Schottky contribution to the specific heat enabled us to constrain the CEF-level scheme as is shown in table 1. The parameterisation suggests that at least six energy levels are responsible for the Schottky contribution observed in the specific heat up to 250 K.

Acknowledgements
JLS acknowledges financial support received through the University of Johannesburg New Generation Scholarship program, as well as the NRF DoL Scarce Skills bursary program. AMS gratefully acknowledges financial support from the NRF (2072956) and from UJ-URC.

References

[1] Jensen J and Mackintosh A R 1991 *Rare Earth Magnetism* (Clarendon Press Oxford)
[2] Hutchings M T 1964 *Solid State Physics* **16** 227
[3] Francios M, Venturini G, Malaman B and Roques B 1990 *Journal of the Less-Common Metals* **160** 197
[4] Gil A, Szytula A, Tomkowicz Z, Wojciechowski K and Zygmunt A 1994 *Journal of Magnetism and Magnetic Materials* **129** 271
[5] Javorsky P, Nakotte H, Robinson R A and Kelley T M 1999 *Journal of Magnetism and Magnetic Materials* **186** 373
[6] Sarkar R, Ghoshray A and Ghoshray K 2007 *Journal of Physics* Condensed Matter **19** 086202
[7] *www.qdusa.com*
[8] Hwang J S, Lin K J and Tien C 1997 *Review of Scientific Intruments* **68** 94
[9] Dachs E and Bertoldi C 2005 *Eur. J. Miniral* **17** 251
[10] Larson A C and Von Dreele R B 1994 *Los Alamos National Laboratory Report* **LAUR** 86–748
[11] Toby B H 2001 *Journal of Applied Crystallography* **34** 210
[12] Collocott S J, Driver R and Andrikidis C 1992 *Physical Review* B **45** 945
[13] Pecharsky V K and Gschneidner, Jr K A 1999 *Journal of Applied Physics* **86** 565
[14] Andres K 1973 *Physical Review* B **7** 54295
[15] Gignoux S, Rossat-Mignod J and Tcheou F 1972 *Physica Status Solidi* a **14** 483

Magnetocaloric effect of ferromagnetic CeAuGe

J L Snyman, B M Sondezi-Mhlungu and A M Strydom

Physics Department, University of Johannesburg, P.O. Box 524 Auckland Park, 2006, South Africa

E-mail: jlsnyman@uj.ac.za

Abstract. The magnetocaloric effect (MCE) may become strongly enhanced in systems exhibiting a second order magnetic phase transition due to critical behaviour of the order parameter in close proximity of the transition region. Recently it has been shown that, analogous to the order parameter, the MCE shows scaling behaviour with the applied magnetic field along the critical isotherm. It has been argued that this behaviour should be indicative of the universality class of the system. Using ferromagnetic CeAuGe as a model system with high crystallographic and atomic order, we have determined the MCE from specific heat measurements in various applied fields. In the low field limit we observe scaling behaviour reminiscent of a mean field ferromagnet. However, in progressively larger applied fields the MCE in CeAuGe significantly exceeds that of the mean field reference system, indicating that the refrigeration capacity of CeAuGe is significantly larger than that of a model mean field ferromagnet. Our results are contextualized in terms of a more general amenability of local-moment magnetic compositions as model systems for magnetic refrigeration.

1. Introduction

The magnetocaloric effect describes the reversible cooling or heating of a material under the action of a changing applied magnetic field. It is customary to express the magnetocaloric effect in terms of two parameters, each denoting a specific magnetic field induced trajectory in phase space. These are the isothermal entropy change $\Delta S(T)_{\Delta B}$ and the adiabatic (isentropic) temperature change $\Delta T_{\Delta B}(T_f)$. These parameters are defined in analogy with the well known Carnot gas-compression refrigeration cycle. Regarding the magnetocaloric effect this cycle consists of an initial isothermal magnetisation obtained by increasing the magnetic induction B from B_i to B_f which causes the entropy of the system to change by $\Delta S(T)_{\Delta B}$ (for the sake of simplicity we assume that the applied field H can be related to B through $B = \mu_0 H$). The second step is an isentropic demagnetisation (decreasing B from B_i to B_f) which induces an adiabatic temperature change $\Delta T_{\Delta B}(T_f)$ in the sample.

The simplest magnetic system exhibiting this effect is an ideal spin-1/2 paramagnet. In the limit where there are no interactions between the spins in the system and all entropy contributions other than the magnetic configurational entropy of the system are negligible, adiabatic demagnetisation cools the system down to absolute zero. In any real system however the former condition can never be met and it is found that magnetic interactions limit the lowest temperatures attainable via the MCE [1].

While interactions among magnetic moments generally act to the detriment of the MCE in paramagnetic systems they are responsible for a significant MCE in systems where such interactions are strong enough to cause a phase transition between the paramagnetic and a

magnetically ordered state. In systems where the transition is of second order the behaviour of a host of physical properties close to the phase transition temperature T_c is determined by the growth of a characteristic length scale over which coherent fluctuations in the order parameter occur [2]. In zero applied magnetic field this critical behaviour is manifested by the dependence of physical quantities on a reduced temperature ϵ. In the critical region the temperature dependence of physical quantities usually takes on a form dominated by terms of order ϵ^p. The value of p is characteristic of the universality class of the system.

For magnetic systems the magnetisation M is a natural choice for the order parameter. The critical behaviour of M with temperature in zero applied field as the critical phase transition temperature T_c is approached takes on the form

$$\left.\frac{\partial M}{\partial T}\right|_{B=0} \sim \pm\,|\epsilon|^{-\beta} \text{ for } \epsilon = \frac{T - T_c}{T_c} < 0. \tag{1}$$

The critical behaviour of M is suggestive of a significant magnetocaloric effect [3] as is discussed next. The isothermal entropy change in a system as it is magnetised by increasing B from B_i to B_f follows from the well-known Maxwell-relations as

$$\Delta S(T)_{\Delta B} = \frac{1}{\mu_0}\int_{B_i}^{B_f}\left(\frac{\partial M}{\partial T}\right)_B dB. \tag{2}$$

Due to the critical behaviour of M the integrand diverges as $T \to T_c$ and $B \to 0$ (see Eq.1). Therefore large $\Delta S_{\Delta B}$ values are expected in the vicinity of $T = T_c$ due to the critical contribution to the integrand. The critical behaviour of the MCE has been a topic of recent theoretical interest. It is known that the maximum value of $\Delta S(T)_{\Delta B}$ (hereafter referred to as ΔS_{max}) occurs at a temperature T' close to T_c if magnetisation proceeds from $B = 0$ to $B = B_f$ [3]. In the limit where $B_f \to 0$ the temperature T' approaches T_c. It could be shown by Franco et al. (see Ref. [4]) that $\Delta S_{max}(B_f)$ scales with the applied field as $|\Delta S_{max}(B_f)| \sim B_f^n$ and that n is indicative of the universality class of the system.

In this work we turn our attention to the MCE in CeAuGe. It has been noted that lanthanide systems are prime candidates for commercial applications of the MCE due to the large localised magnetic moments associated with the rare earth ions in these compounds. Within this context it may seem counter intuitive to study the MCE in cerium-based intermetallic compounds as the Ce^{3+} free-ion moment ($J = 5/2$) is relatively small when compared to other rare earth ions such as Gd^{3+} ($J = 7/2$). Furthermore in a host of crystalline systems (as is the case with CeAuGe, see Ref.[5]) the degenerate ground state multiplet associated with the free tri-positive Cerium ion is split by the action of an anisotropic crystalline electric field (CEF). In CeAuGe the CEF causes the magnetic ground state to be a doublet (corresponding to $J = 1/2$). The reason for studying the MCE in CeAuGe (and other Cerium systems) however can be seen when the mechanism by which magnetic order is established is investigated.

Magnetic order in lanthanide systems usually occurs due to an indirect exchange mechanism, whereby the magnetic moments localised at lanthanide crystallographic sites polarise the itinerant electrons in the system thereby producing an indirect interaction between neighbouring localised moments. Recent theoretical investigations into such systems has shown that the refrigeration capacity in these systems are enhanced due to the polarisation of conduction electrons [6, 7]. The local interaction between the $4f$-electron of the Ce^{3+}-ion and the conduction electrons in various Ce-based intermetallic systems is strong enough to produce an extensive range of correlated electron phenomena [8]. Given the possibility that such systems may show a significant enhancement in refrigeration capacity, the MCE in Ce-based systems is of conceivable interest. However, a systematic study into the magnetocaloric properties of Ce-based intermetallics is still generally lacking. Here we present a first investigation into the MCE

of CeAuGe, focussing on the critical behaviour of the MCE as well as the refrigeration capacity of the system.

2. Experimental Procedure

The synthesis and characterisation of polycrystalline CeAuGe is discussed in Refs.[5] and [9]. The MCE is calculated indirectly from the measured specific heat reported in the latter two references. The calculation follows the method given in Ref.[10]. The isofield entropy $S_B(T)$ is calculated from the specific heat $C_{p,B}(T)$ measured in constant pressure and applied field as

$$S_B(T') = \int_0^{T'} \frac{C_{p,B}(T)}{T} dT. \tag{3}$$

Eq.3 introduces two significant errors into the results that follow. The first is the numerical error associated with calculating the integral from a discreet set of data points. The second is that the experimental temperature range does not extend down to absolute zero. This implies that the contribution from $\int_0^{T_0} \frac{C_{p,B}(T)}{T} dT$ (where T_0 is the lowest available experimental temperature) needs to be approximated through extrapolation. Both these errors are taken into account in the sections below.

Knowing $S_B(T)$ for a set of applied fields allows the isothermal entropy change $\Delta S_{\Delta B}(T)$ associated with $\Delta B = B_f - B_i$ to be calculated as

$$\Delta S_{\Delta B}(T) = S_{B_f}(T) - S_{B_i}(T) \tag{4}$$

In this work $B_i = 0$ is used throughout in the calculation of the $\Delta S_{\Delta B}(T)$ data reported in the next section. The adiabatic temperature change follows from

$$\Delta T_{\Delta B}(T_f) = T_f - T_i \tag{5}$$

where T_i and T_f have values satisfying the isentropic condition $S_{B_i}(T_i) = S_{B_f}(T_f)$. With regards to $\Delta T_{\Delta B}(T_f)$ it will be assumed that $B_f = 0$ throughout.

3. Results and Discussion

The various isofield entropy curves are calculated from the specific heat of CeAuGe as reported in Refs.[5] and [9]. Shown in figure 1 are the isofield entropy in zero applied field, 1 T and 9 T. The inset shows the magnetic configurational entropy calculated by considering the magnetic contribution to the specific heat. The latter is calculated by using the LaAuGe specific heat reported in [9] as a non-magnetic reference. The saturation value of the magnetic configurational entropy is close to $R \ln 2 = 5.76$ J.mol^{-1}.K^{-1}as is expected for a magnetic doublet. The calculation of the magnetic configurational entropy confirms the correct treatment of the errors inherent to Eq. 3 as noted in the previous section.

$\Delta S_{\Delta B}(T)$ and $\Delta T_{\Delta B}(T_f)$ calculated from $S_B(T)$ are shown in figures 2 and 3. Also shown is the magnetocaloric effect of an ideal $J = 1/2$ mean field ferromagnet with the same transition temperature as CeAuGe. In calculating the MCE of this model system the formalism developed in Ref.[11] was used. A comparison between the MCE in CeAuGe and the ideal mean field ferromagnetic system shows that the characteristic Caret-like shape (see for example Ref. [3]) of $\Delta S_{\Delta B}(T)$ and $\Delta T_{\Delta B}(T_f)$ is much broader for CeAuGe than for the mean field ferromagnetic system. As is discussed below this is an indication of the enhanced refrigeration capacity of CeAuGe. For $\Delta B = 1$ T the extremas in both $\Delta S_{\Delta B}(T)$ and $\Delta T_{\Delta B}(T_f)$ for CeAuGe are close to the corresponding extremas of the mean field system. This is the first indication that the critical behaviour of CeAuGe closely resembles that of a mean field ferromagnet in small applied fields.

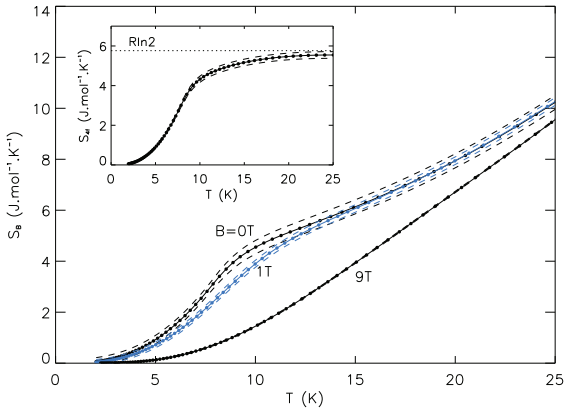


Figure 1. The isofield entropy (filled circles) for CeAuGe calculated from experimental specific heat data in zero field, 1 T and 9 T. Upper and lower bounds are indicated by dashed lines. Inset: Magnetic configurational entropy of the Ce^{3+} $4f$-electrons in CeAuGe in zero applied field.

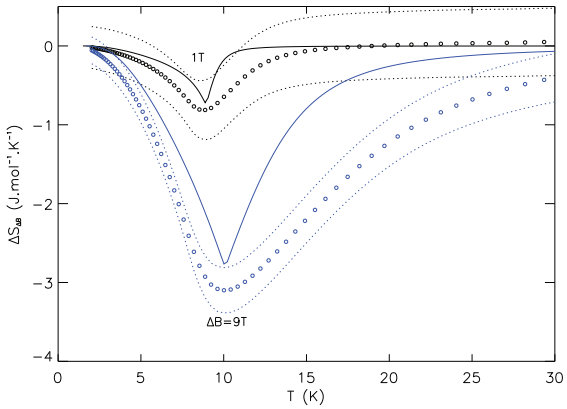


Figure 2. The isothermal entropy change $\Delta S(T)_{\Delta B}$ (open circles) calculated from the isofield entropy curves in figure 1 for $\Delta B = 1$ T and 9 T. Dotted lines show the uncertainty in the calculation. Solid lines show the isothermal entropy change calculated for the corresponding model mean field system.

In order to investigate the critical behaviour further $\Delta S_{max}(B_f)$ for both CeAuGe and the model mean field system are shown in figure 4. $\Delta S_{max}(B_f)$ for the mean field system lies within the margins of error (calculated for CeAuGe) up to $\Delta B = 4$T. For larger fields $\Delta S_{max}(B_f)$ in CeAuGe clearly exceeds that of the mean field system. The behaviour of $\Delta S_{max}(B_f)$ shown here suggests that the critical behaviour of polycrystalline CeAuGe closely resembles that of a mean field ferromagnetic system.

The refrigeration capacity of CeAuGe is calculated from figure 2 as

$$q(B_f) = \int_0^\infty \Delta S_{\Delta B}(T) dT \tag{6}$$

and shown in figure 4. The refrigeration capacity of the mean field system is also shown. We find that the refrigeration capacity of CeAuGe exceeds that of the mean field system by $\sim 40\%$ in all but the lowest applied fields.

4. Conclusion

The MCE in polycrystalline CeAuGe was calculated indirectly from measured specific heat data in various applied fields. It could be shown that the critical behaviour of the MCE in small applied fields is close to the critical behaviour expected in a mean field ferromagnetic system. Importantly, the refrigeration capacity of CeAuGe is shown to be significantly enhanced above that of the ideal mean field system, highlighting the role of local-moment magnetic compositions as model systems for magnetic refrigeration.

Acknowledgements

JLS acknowledges the financial support received through both the University of Johannesburg New Generation Scholarship program, as well as the NRF DoL Scarce Skills bursary program. AMS gratefully acknowledges financial support from the NRF (2072956) and from the UJ-URC.

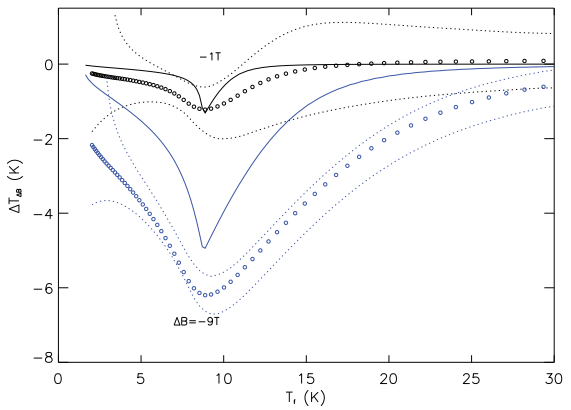


Figure 3. The adiabatic temperature change $\Delta T(T_f)_{\Delta B}$ calculated from the isofield entropy curves in figure 1 for $\Delta B = 1$ T and 9 T (open circles). Dotted lines show the uncertainty in the calculation. Solid lines show the adiabatic temperature change calculated for the corresponding mean field ferromagnetic system.

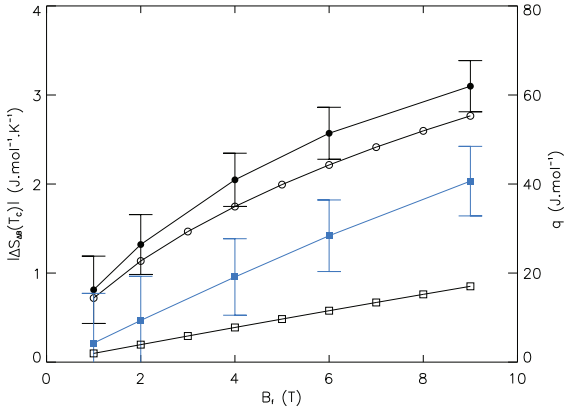


Figure 4. $|\Delta S_{max}(B_f)|$ for $\Delta B = B_f$ between 1 T and 9 T. Filled circles correspond to values calculated from the isofield heat capacities reported in Ref.[9] while the unfilled circles correspond to values calculated for the corresponding mean field ferromagnetic system. Note that the vertical axis on the left of the figure shows values for $|\Delta S_{max}(B_f)|$ in units of J.mol^{-1}.K^{-1}. The refrigeration capacity of CeAuGe (filled squares) is also shown, together with the refrigeration capacity of the mean field ferromagnet (unfilled squares). The vertical axis on the right of the figure shows values of the refrigeration capacity in units of J.mol^{-1}.

References

[1] Blundell S 2001 *Magnetism in Condensed Matter* (Oxford University Press)
[2] Kadanoff L P, Götze W, Hamblen D, Hecht R, Lewis E A S, Palciauskas V V, Rayl M and Swift J 1967 *Reviews of Modern Physics* **39** 395
[3] Pecharsky V K, Gschneidner, Jr K A, Pecharsky O A and Tishin A M 2001 *Physical Review B* **64** 144406
[4] Franco V, Blazquez J S and Conde A 2006 *Applied Physics Letters* **89** 222512
[5] Sondezi-Mhlungu B M, Adroja D T, Strydom A M, Paschen S and Goremychkin E A 2009 *Physica B* **404** 3032
[6] Alfaro F and Bernhard B H 2009 *Physica B* **404** 3066
[7] Bernhard B H and Siqueira M C 2010 *Solid State Communications* **149** 1777
[8] Gegenwart P and Steglich Q Si F 2008 *Nature Physics* **4** 186
[9] Mhlungu B M and Strydom A M 2008 *Physica B* **403** 862
[10] Pecharsky V K and Gschneidner, Jr K A 1999 *Journal of Applied Physics* **86** 565
[11] de Oliveira I G and von Ranke P J 2010 *Physics Reports* **489** 89

Thermodynamic Properties and Inelastic Neutron Scattering of Ferromagnetic CeCuSi

B M Sondezi-Mhlungu[1], A M Strydom [1], D T Adroja[2]

[1]Physics Department, University of Johannesburg, P.O. Box 524, Auckland Park, 2006, South Africa

[2]Rutherford Appleton Laboratory, ISIS, UK

E-mail: bmhlungu@uj.ac.za

Abstract. Cerium-based compounds demonstrate different and interesting phenomena depending on the valence of the Ce ion. The hybridization between Ce-4f electrons and the neighbouring conduction electrons is responsible for the anomalous properties observed in different systems. Depending on the strength of the f-ligand hybridization, one observes in these compounds phenomena such as magnetic ordering, heavy fermion behaviour, Kondo effect and valence fluctuations [1]. The derived local magnetic moments associated with the state of Ce ions order antiferromagnetically in most cases. CeCuGe, CeCuSi and CeAuGe belong to a small family of equiatomic Ce based compounds that order ferromagnetically at low temperatures. The ordering temperatures are T_C = 10 K, 15 K and 10 K for CeCuGe, CeCuSi [2] and CeAuGe [3], respectively. Here we report on the studies of susceptibility, magnetization and specific heat data of CeCuSi. Inelastic neutron scattering studies were also carried out to confirm the presence of crystal electric field effects. The crystal field parameters estimated from inelastic neutron data were used to calculate and plot the heat capacity of CeCuSi.

1. Introduction

The study into the thermodynamic and magnetic-field stability of magnetic ordering, T_C in CeCuSi compound, has been undertaken and was identified and confirmed to exhibit ferromagnetic ground state below T_C = 15 K [2, 4]. CeCuSi is among a few compounds that exhibit ferromagnetic ordering at low temperatures. CePdX (X = P, As, Sb) have been reported to exhibit ferromagnetic ordering at low temperature, (T_C = 17.5 K, for CePdSb; T_C = 4 K for CePdAs and T_C = 5 K for CePdP) [5, 6] having the hexagonal ZrBeSi-type crystal structure for X = P and As; and LiGaGe-type structure for X = Sb. CeCuSi crystallizes in an ordered hexagonal ZrBeSi-type structure (space group P6$_3/mmc$, number 194) [2, 4].

The ferromagnetic nature of the transition in this compound as observed from specific heat (C_p) data is indicated by a λ-type anomaly at low temperatures. At higher temperatures, the magnetic specific heat, C_{4f} (which is accessed by subtracting the data of an isostructural non-magnetic reference compound, LaCuSi) is characterized by a broad Schottky anomaly (at around 80 K) resulting from electronic excitations to crystal electric field (CEF) levels. In this wok, thermodynamic as well as magnetic properties which are supported by inelastic neutron scattering data will be presented.

2. Experimental procedure

Polycrystalline CeCuSi and its corresponding non-magnetic, isostructural counterpart LaCuSi were prepared by arc-melting stoichiometric quantities of the elements (purities in wt.%) Ce and La (4N), Cu (4N5), Si (6N) under-pressure in an ultra-high purity argon gas. This was followed by subsequent annealing of the samples in an evacuated quartz tube at 800^0C for two weeks. The temperature dependent susceptibility was measured using Superconducting Quantum Interference Device (SQUID) magnetometer [7] in small applied magnetic field of 0.003 T between 1.9 K and 400 K. The field dependent magnetization up to 9 T was measured using a Physical Properties Measurement System (PPMS). Specific heat measurements were done using heat pulse relaxation technique PPMS system in the temperature range between 1.9 K and 300 K. Powder X-ray diffraction data and Rietveld profile refinement confirmed that the compounds were homogeneous and single-phased. Inelastic neutron scattering (INS) measurements on both CeCuSi and LaCuSi polycrystalline samples were performed using the time-of-flight HET spectrometer at ISIS Facility, Rutherford Appleton Laboratory at UK with the incident energies $E_i = 14$ and 23 meV between 4.5 K and 15 K.

3. Results and discussion

The temperature dependence of inverse susceptibility data measured over the temperature range of 2 - 400 K in a small applied magnetic field of 0.003 T is exhibited in figure 1. Above 160 K, Curie-Weiss behavior is observed and the effective moment of about 2.52 μ_B (close to the full Ce free-ion value, 2.54 μ_B) was calculated. The paramagnetic temperature θ_p, was fitted to be 20.54 K. At low temperatures, the susceptibility data is characterised by a sharp transition around 15 K, this is associated with ferromagnetic ordering temperature T_C (see inset on the top left corner of figure 1, with the arrow indicating long-range ferromagnetic ordering). The magnetization within the ordered region (bottom right inset in figure 1) tends to saturation at about 1.01 μ_B at very small applied magnetic field. The observed saturation moment is considerably less than the theoretically expected value $g_J J = 2.14$ μ_B for the free Ce^{3+} ion involving the entire six-fold $J = \frac{5}{2}$ multiplet and $g_J = \frac{6}{7}$. This reduction in moment is attributed to the presence of CEF effects in which the crystal field splitting causes a subset of the Ce magnetic levels to freeze out at this temperature as well as the magnetocrystalline anisotropy present in this compound.

The specific heat data, C_p as a function of T measured in zero field in the temperature range 1.9 - 300 K is displayed in figure 2 (black circles). The data for a non-magnetic isostructural monologue, LaCuSi is also shown in the figure (red triangles) and the blue solid line is the least-squares fit according to Debye expression where the Debye temperature, $\theta_D = 244.54$ K was obtained. The low-temperature data is characterised by sharp λ-type anomaly at $T_C = 15$ K which corresponds to the maximum slope in the C_p curve. To access magnetic contribution to the specific heat, C_{4f}, the expression:

$$C_{4f}(T) = C_{pCeCuSi}(T) - [C_e(T) + C_L(T)]_{LaCuSi} \tag{1}$$

was used where $[C_e(T) + C_L(T)]_{LaCuSi}$ is the electronic and lattice contribution to the total C_p assuming that the Debye temperatures, θ_D for both CeCuSi and LaCuSi are the same. A broad Schottky peak is observed at higher temperatures (at about 80 K) of C_{4f} curve, and the Schottky fit for a two-level system (indicated by a blue solid line in the inset) gave the energy separation, Δ of the electronic excitation levels to be 110 K using the expression:

$$C_{4f}(T) = \frac{R\left(\frac{g_0}{g_1}\right)\left(\frac{\Delta}{T}\right)^2 e^{\Delta/T}}{\left[1 + \left(\frac{g_0}{g_1}\right)e^{\Delta/T}\right]^2} \tag{2}$$

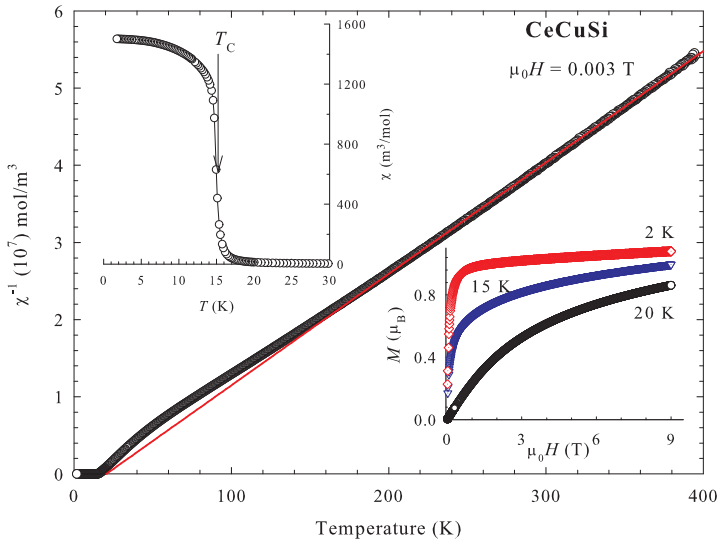


Figure 1. Inverse susceptibility graph with the solid line fitted according to Curie-Weiss expression, the inset on the top left corner shows the enlarged low temperature susceptibility data. The inset on the bottom right corner shows the graph of magnetization vs applied field with three isotherms at 2 K (well within the magnetically ordered region), at 15 K (at transition temperature) and 20 K (in the paramagnetic region).

where R is the gas constant, g_0 and g_1 are the degeneracies of the ground and first excited state, respectively.

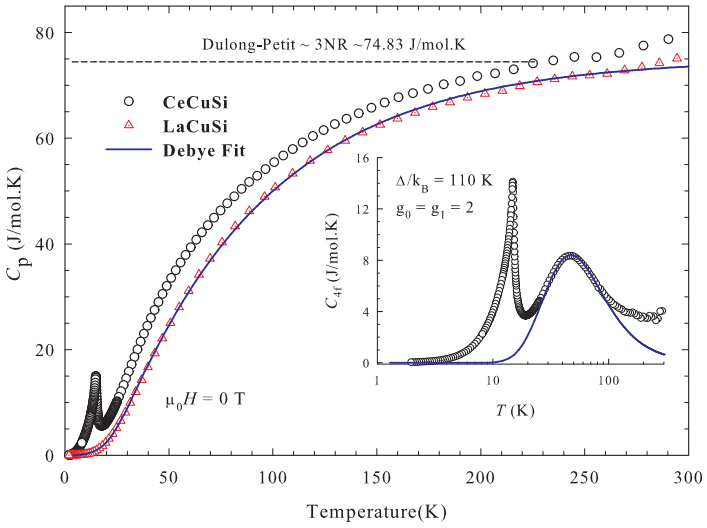


Figure 2. Specific heat data of CeCuSi (circles) and LaCuSi (triangles) measured in zero field. The inset shows the magnetic contribution (C_{4f}) to the C_p. A solid line is the fit to the broad Schottky peak observed at high temperatures.

Inelastic neutron scattering (INS) studies at 15 K confirm the above C_p analyses and reveal

a strong crystal field excitation at 12.5 meV [4]. CEF excitations were analysed based on the crystal field model. The hexagonal structure of CeCuSi has distinguished two different crystallographic sites for the Cu and Si atoms that give three-fold rotation symmetry around the c-axis. The details of the calcutions are reported in ref. [4].

The obtained crystal field parameters were used to calculate specific heat. The calculated C_p along with the experimental C_{4f} are shown in figure 3. The calculated C_p is smaller than the experimental C_{4f}, this may be attributed to the presence of the correlations between Ce-4f moments [8].

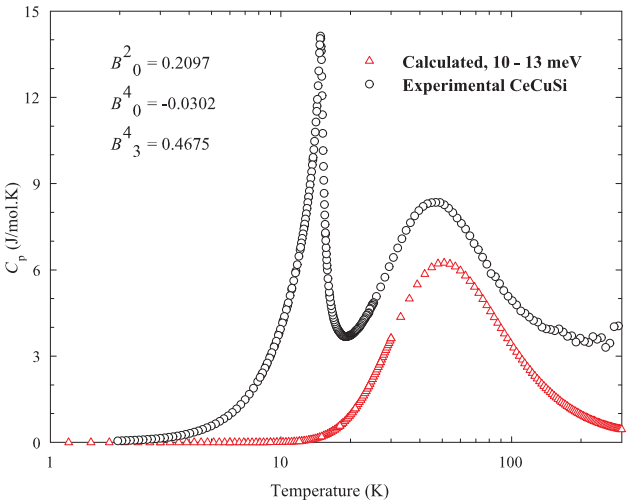


Figure 3. Experimental magnetic specific heat data of CeCuSi (circles) and calculated specific heat dat (red triangles) measured in energy range of 10 - 13 meV.

4. Conclusion

The inverse susceptibility data in CeCuSi confirms a paramagnetic to ferromagnetic (PM-FM) transition around 15 K. The high temperature data follows Curie-Weiss law. The deviation from linear behaviour at decreased temperature is an indication of crystal electric field (CEF), which were confirmed by INS data where an excitations was observed around 12.5 meV. Magnetization as a function of applied magnetic field within the ordered region (2 K) is characterised by a long range spontaneous magnetization which tends to saturate around 1.0 μ_B at very low applied magnetic field. Magnetization within paramagnetic region (20 K), shows much reduced magnetization. However, it is observed that this data does not exhibit complete linearity, this observation is common in low paramagnetic data and the observation can be attributed to thermal fluctuations as well as short range correlations which might still be dominant in this temperature region [9]. CEF effect were also observed in specific heat data, where the magnetic contribution to specific heat is characterised by a broad Schottky peak at higher temperatures. The fit to the Schottky peak gave an energy separation of 110 K, indicating that the ground state and the first excited doublet are not very close to each other. Inelastic neutron scattering measurements confirmed the presence of CEF excitations observed at 12.5 meV. The magnetic scattering at various temperatures when following crystal field model produced parameters which were used to calculate the theoretical specific heat data. The calculated C_{4f} data coincides with the experimental data and the difference in height around the broad Schottky peak is due to the correlations between the Ce-4f atoms.

Acknowledgements
B.M. Sondezi-Mhlungu acknowledges financial assistance provided by the National Research Foundation (NRF) of South Africa as well as the support provided by university research council (URC), UJ Science Faculty and Physics Department. A M Strydom thanks the SA-NRF grant number 2072956.

References

[1] Fujita T,Suzuki T, Nishigori S, Takabatake T, Fujii H and Sakarai J 1992 *J. of Magn. Magn. Mater* **108** 35
[2] Yang F, Kuang J P, Li J, Bruck E, Nakotte H, de Boer F R, Wu X, Li Z and Wang Y 1991 *J. Appl. Phys.* **69** 4705
[3] Pöttgen R 1996 *J. Magn. Magn. Mater* **152** 196
[4] Sondezi-Mhlungu B M, Adroja D T, Strydom A M, Paschen S and Goremychkin E A 2009 *Physica B.* **404** 3032
[5] Trovarelli O, Sereni J G, Schemerber G and Kappler J P 1994 *Phys. Rev. B.* **49** 15179
[6] Katoh K, Takabatake T, Ochiai A, Uesawa A and Suzuki T 1997 *Physica B.* **230** 159
[7] Boulevard L *Quantum Design, Inc. San Diego, USA* **6325** 92121
[8] Adroja D T, Rainford B D, Malik S K, Takeya H, Gschneidner K A and Pecharsky V K 1999 *Journal of Alloys and Compounds* **288** 7
[9] Singh N K, Pecharsky V K and Gschneidner K A 2008 *Phys. Rev. B.* **77** 054414

Critical behavior near the ferromagnetic Curie phase transition in CeCuGe

B M Sondezi-Mhlungu and A M Strydom

Physics Department, University of Johannesburg,P.O. Box 524, Auckland Park, 2006,South Africa

E-mail: bmhlungu@uj.ac.za

Abstract. Isothermal magnetization of a hexagonal, highly ordered ferromagnetic CeCuGe was measured in order to study the critical scaling behavior in this polycrystalline compound around the ordering temperature. From the analyses of the magnetization data, T_C was confirmed by following modified Arrot plot and Kouvel-Fischer techniques. The critical exponents, β, γ and δ obtained from the fit of the spontaneous magnetization, initial susceptibility and isothermal magnetization revealed that the system behaves as Mean-Field around the ordering temperature.

1. Introduction

A large number of ternary equiatomic CeTX compounds where T is a transition metal and X is a metal atom exhibit various ground states. Among them only a few exhibit ferromagnetic ordering at low temperatures. CePdX (X = P, As, Sb) have been reported to exhibit ferromagnetic ordering at low temperature, ($T_{\rm C}$ = 17.5 K, for CePdSb; $T_{\rm C}$ = 4 K for CePdAs and $T_{\rm C}$ = 5 K for CePdP) [1, 2] having the hexagonal ZrBeSi-type crystal structure for X = P and As; and LiGaGe-type structure for X = Sb. Rare earth containing, particularly Ce-based compounds show very interesting magnetic properties resulting from a strong hybridization of Ce-4f electrons with the conduction electrons. Although a fairly large number of studies have been done on these compounds, there are still more questions to be considered and answered. Ce-based compounds are stable under normal condition and can be driven to unstable state by application of hydrostatic pressure and applied magnetic field. CeCuGe is another ternary equiatomic intermetallic compound to be earlier reported to order ferromagnetically at low temperature, $T_{\rm C}$ = 10 K [3, 4]. The compound is highly ordered and was confirmed to crystallize in hexagonal ZrBeSi-type structure, (space group P6$_3/mmc$, number 194). The ordering temperature $T_{\rm C}$ = 10 K, has been observed through specific heat, C_p [4, 5] magnetization, $M(T, H)$ and resistivity, $\rho(T)$ [3, 6] data. The λ-type anomaly associated with ordering temperature has been observed in the $C(T)$ versus T curve [4, 7]. This temperature dependent magnetic susceptibility curve is characterized by a sharp increase in the magnetization around $T_{\rm C}$; and the resistivity curve is characterised by a point of inflection at the transition.

This focus of this particular report is on the critical exponents of CeCuGe around ferromagnetic phase transition and the subsequent classification of the system. Inorder to determine critical exponents close to ferromagnetic transition temperature;(10 K), a number of isotherms (M versus

H; $B = \mu_0 H$ and H will be used as the approximation of the applied magnetic field throughout the text) around the transition temperature, 10 K are measured.

2. Experimental procedure

Polycrystalline CeCuGe sample was prepared in an arc furnace under ultra high-purity argon atmosphere. Stoichiometric amounts of the elements (purities in wt. %) Ce (99.99), Cu (99.995) and Ge (99.9999) were used. The ingots were remelted several times to promote homogeneity. Powder X-ray diffraction measurement performed at room temperature confirmed that the material was homogeneous and single phase. Rietveld refinement profile according to ZrBeSi-type structure (space group $P6_3/mmc$, number 194) was performed and the lattice parameters of the sample were confirmed [6, 3, 5]. A SQUID magnetometer was used to measure magnetization in the temperature range of 2 - 400 K in the presence of applied magnetic field of 0.0005 T. In addition, a number of magnetization isotherms (magnetization as a function of applied magnetic field) were measured for the calculation of critical exponents.

3. Results and discussion

As a first evaluation and the confirmation of the ferromagnetic Curie transition temperature T_C, the temperature dependence of the magnetization (M vs T) of CeCuGe cooled in applied magnetic field $H = 0.0005$ T is shown in figure 1. The graph is characterised by a strong increase of M vs T at the ferromagnetic ordering temperature. The maximum slope corresponding to the ordering temperature is indicated by an arrow at about 10 K. In the vicinity of this transition, the existence of diverging correlation length $\xi = \xi \left|1 - \frac{T}{T_C}\right|^{-\nu}$ leads to universal scaling laws for spontaneous magnetization, M_s and susceptibility, χ. These are characterised by the set of critical exponents, β, which corresponds to the spontaneous magnetization below T_{C}, γ, which corresponds to the inverse initial susceptibility, $\chi_0^{-1}(T)$ above T_{C} and δ which corresponds to the critical magnetization isotherm [8] at T_{C}.

These critical exponents in the vicinity of critical temperature are defined by the expressions:

$$M_s(T) = M_0(-t)^{\beta}, \quad t < 0 \tag{1}$$

$$\chi_0^{-1}(T) = (h_0/M_0)(t)^{\gamma}, \quad t > 0 \tag{2}$$

$$M = DH^1/\delta, \quad t = 0 \tag{3}$$

where t is a reduced temperature ($\frac{T-T_C}{T_C}$) and M_0, h_0 and D are critical amplitudes [9, 8, 10, 11]. In order to determine critical exponents in and eventually define the behavior of the system around T_{C}, modified Arrot plots were plotted whereby $M^{1/\beta}$ vs $(H/M)^{1/\gamma}$ is shown in figure 2. Only few isotherms are shown for the sake of clarity. The region of scaling is indicated by a solid line in each isotherm shown in figure 2.

From the intercepts of the fitted $M^{1/\beta}$ vs $(H/M)^{1/\gamma}$ straight lines (from figure 2), spontaneous magnetization, M_{s} were determined from the y-axis and initial susceptibility, χ_0^{-1} were extrapolated from the x-axis intersection. The values obtained from extrapolating the data through scaling profile were plotted as a function of temperature. The scaling procedure for the isotherms that showed almost parallel lines resulted in the critical exponents $\beta = 0.736$ and $\gamma = 0.930$ associated with spontaneous magnetization $M_{\mathrm{s}}(T)$ and the initial inverse susceptibility; $\chi_0^{-1}(T)$, respectively. The corresponding δ value can be obtained from Widon scaling relation [12]; $\delta = 1 + \gamma/\beta$ which gives $\delta = 2.265$.

During the fitting process the transition temperature, T_{C} was also a fitting parameter. The transition temperature according to the fit of the $T < T_{\mathrm{C}}$ data yields the temperature value of

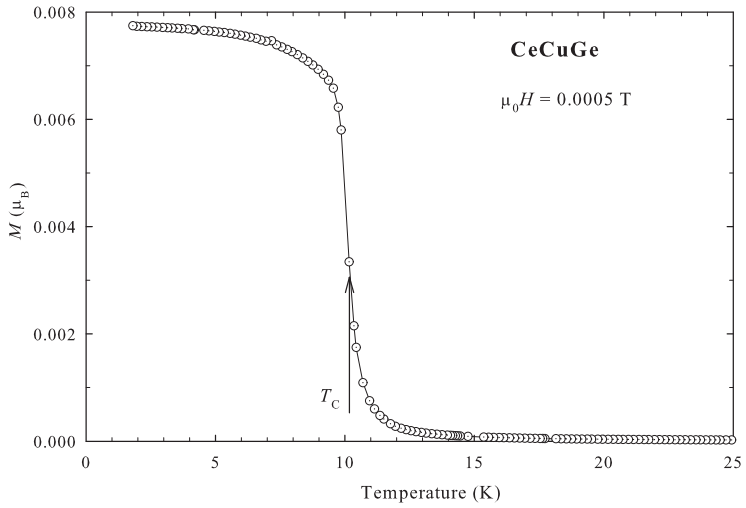


Figure 1. Magnetization versus temperature of CeCuGe measured in 0.0005 T.

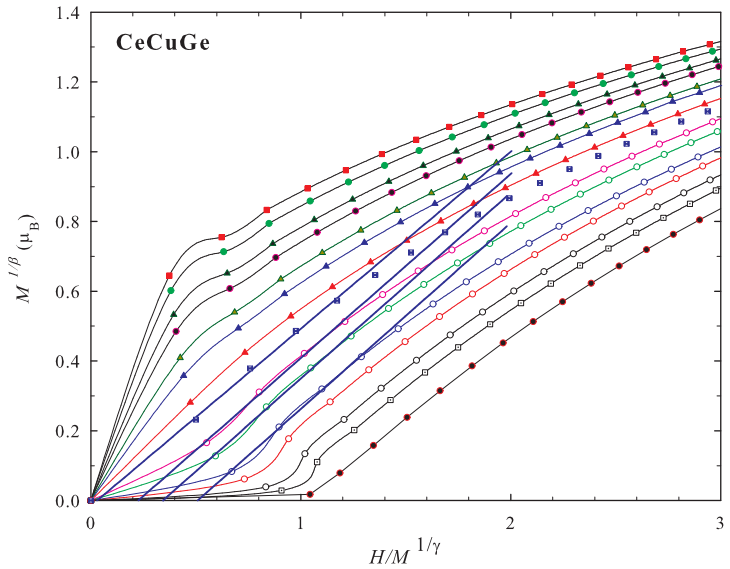


Figure 2. Modified Arrot plot; $M^{1/\beta}$ vs $H/M^{1/\gamma}$ of CeCuGe isotherms. Data points were depopulated for clarity purposes.

10.081 K whilst the fit of $T > T_C$ data yields 10.128 K (see figure 3). From the fitting parameters obtained in this step, the exponents T_C, β and γ are computed using Kouvel-Fischer method. Although $M_s(T)$ and $\chi_0^{-1}(T)$ were used to extract the critical exponents obtained above, new values of T_C, β and γ can be determined with high precision following the equations:

$$Y(T) = -M_s \left(\frac{\partial M_s}{\partial T} \right)^{-1} = -\frac{T - T_C}{\beta} \tag{4}$$

$$X(T) = \chi_0^{-1} \left(\frac{\partial \chi^{-1}}{\partial T} \right) = \frac{T - T_C}{\gamma} \tag{5}$$

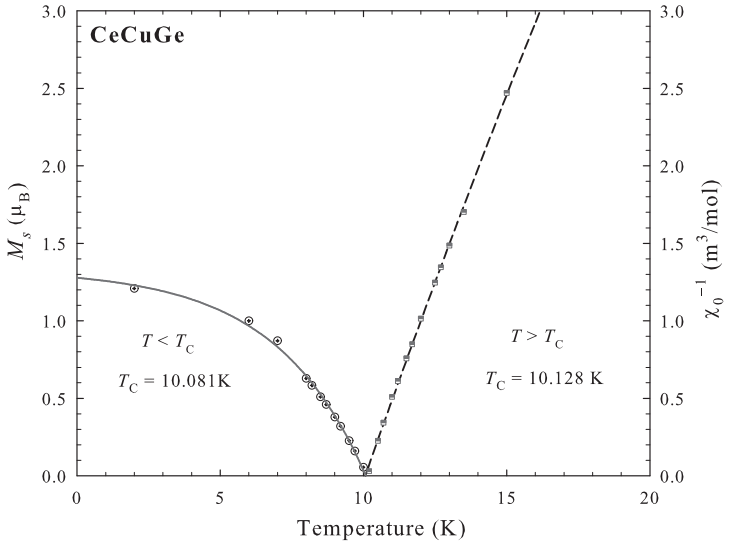


Figure 3. The plots of spontaneous magnetization, M_s and initial susceptibility, χ_0^{-1} as a function of temperature (T). The fits of the power laws, for the respective data are shown as a dashed line and the critical values are labeled in the graph.

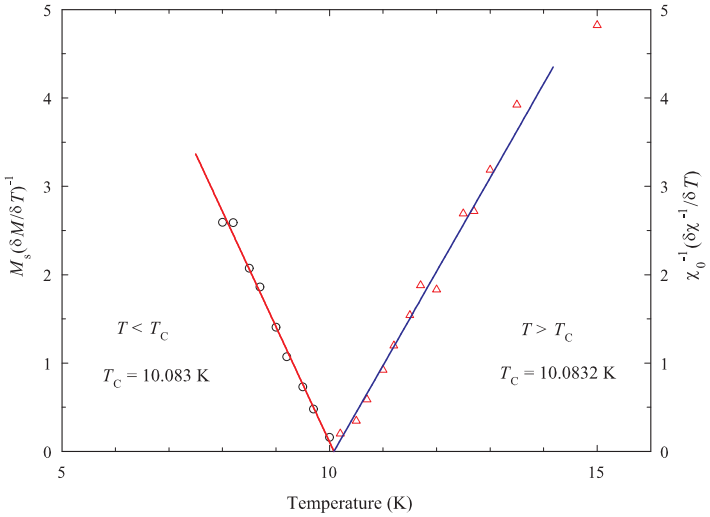


Figure 4. The plots of $Y(T)$ and $X(T)$ vs T and the linear fits according to Kouvel-Fischer method are indicated by solid lines.

The Kouvel-Fischer method suggests that the straight line plots (see figure 4 of $Y(T)$ and $X(T)$ vs T) gives the lines with the gradients -$\frac{1}{\beta}$ and $\frac{1}{\gamma}$. The extrapolation of these lines to the T-axis gives the value of T_C [9]. These linear fits resulted in $\beta = 0.637 \pm 0.01$ and $\gamma = 0.962 \pm 0.01$ and $T_C = 10.083 \pm 0.1$ K obtained from both the fit of $Y(T)$ and $X(T)$ vs T data. Computing for the critical exponent δ from the expression; $\delta = 1 + \gamma/\beta$ the value $\delta = 2.509 \pm 0.01$ is obtained.

Table 1. The table of critical exponents and T_C value for CeCuGe obtained from Kouvel-Fischer scaling method. Theoretical critical exponents of difference models are also listed in the table.

	β	γ	δ	$T_C(K)$	References
CeCuGe	0.637	0.962	2.509	10.083	This work
MFT	0.5	1.0	3.0	-	[13]
3D Ising Model	0.325	1.24	4.82	-	[13]
3D HM	0.365	1.336	4.60	-	[13]

4. Conclusion

The magnetization data measured in CeCuGe indicated a paramagnetic to ferromagnetic (PM-FM) transition around 10 K. This transition was confirmed by modified Arrot plot and Kouvel-Fischer techniques. The critical exponents around the critical temperature T_C for this ferromagnetic compound from both methods converged to similar values which allowed the system to be classified as Mean Field. The ordering temperature obtained after the fitting models were used gives confirms 10.083 K. The critical exponents obtained in this work indicate that the system can be classified as Mean-Field ferromagnet.

Acknowledgements

B.M. Sondezi-Mhlungu acknowledges financial assistance provided by the National Research Foundation (NRF) of South Africa as well as the support provided by university research council (URC), UJ Science Faculty and Physics Department. A M Strydom thanks the SA-NRF grant number 2072956.

References

[1] Trovarelli O, Sereni J G, Schemerber G and Kappler J P 1994 *Phys. Rev. B.* **49** 15179
[2] Katoh K, Takabatake T, Ochiai A, Uesawa A and Suzuki T 1997 *Physica B.* **230** 159
[3] Oner Y, Kamer O, Ross J H, Lue C S and Kuo Y K 2005 *Solid State Communications* **136** 533
[4] Yang F, Kuang J P, Li J, Bruck E, Nakotte H, de Boer F R, Wu X, Li Z and Wang Y 1991 *J. Appl. Phys.* **69** 4705
[5] Sondezi-Mhlungu B M, Adroja D T, Strydom A M, Paschen S and Goremychkin E A 2009 *Physica B* **404** 3032
[6] Chevalier B, Pasturel M, Bobet J L, Weill F, Decourt R and Etourneau I, 2004 *Journal of Solid State Chemistry* **177** 752
[7] Sondezi-Mhlungu B M, Adroja D T and Strydom A M 2010 *Journal of Condensed Physics Series* **32** 7209
[8] Mohan C V, Seeger M, Kronmuller H, Murugaraj P and Maier J 1998 *Journal of Magnetism and Magnetic Materials.* **183** 348
[9] Said M R, Hamam Y A, Abu-Aljayaresh I and Mahmood S 1999 *Journal of Magnetism and Magnetic Materials* **195** 679
[10] Khiem N V, Phong P T, Bau L V, Nam D N H, Hong L V and Phuc N X 2009 *Journal of Magnetism and Magnetic Materials* **321** 2027
[11] Mukherjee S, Raychaudhuri P and Nigam A K 2000 *Physical Review B.* **61** 8651
[12] Eugene S H 1971 *Introduction to Phase Transitions and Critical Phenomena* (New York: Oxford University Press) p 182
[13] Bagnuls C and Bervillier C 1985 *Phys. Rev. B.* **32** 7209

Low-lying magnetism in strongly correlated $CeRh_2Sn_2$

AM Strydom

Physics Department, University of Johannesburg, PO Box 524, Auckland Park 2006, South Africa

E-mail: amstrydom@uj.ac.za

Abstract. $CeRh_2Sn_2$ is a heavy-fermion compound which acquires a very high effective electron mass at low temperatures. The enhanced electronic density of states at the Fermi energy classifies it as a strongly correlated electron system. Here we investigate the previously reported magnetic ordering occurring close to $T_N \simeq 0.4$ K using specific heat and magnetic susceptibility of a polycrystalline sample of $CeRh_2Sn_2$. $C_P(T)$ measurements extending down to 0.05 K and well below T_N enable the calculation of entropy, which provides a new explanation of the origin of the massive specific heat just above the phase transition.

1. Introduction

The simultaneous magnetic Kondo interaction which tends to screen or renormalize a well-localized magnetic spin, and the cooperative behaviour among a lattice of spins has been a topic of extensive research over a number of decades. More recently, the notion of quantum criticality at the verge of a $T = 0$ magnetic phase transition has brought a resurgence of experimental and theoretical interest to this field. Here, we investigate physical properties of the ternary intermetallic compound $CeRh_2Sn_2$ which belongs to the strongly correlated electron class of systems. The magnetism in $CeRh_2Sn_2$ originates from the single localized $4f-$electron per Ce^{3+}ion.

The formation of $CeRh_2Sn_2$ has been known as a member of the Ce-based series of compounds CeM_2Sn_2 since the crystallographic report of Selsane *et al* [1]. M is one of the transition elements Ni, Cu, Rh, Pd, Ir, or Pt. Subsequent work [2] has revealed antiferromagnetic ordering among all of these compounds with Néel temperatures ranging from $T_N = 4$ K in $CeIr_2Sn_2$ to $T_N = 0.4$ K in $CeRh_2Sn_2$. The reported magnetic susceptibilities [2] returned systematic local-moment behaviour towards higher temperatures which confirmed Ce as the principal magnetic species. Moreover, a cerium $4f-$electron based magnetic Kondo interaction is prevalent in this series [3, 4, 5, 6] and hence a description of the ground state has to account for long-range antiferromagnetic coupling between Ce spins as well as the competing on-site Kondo coupling of Ce spins with conduction electrons. Beyermann *et al* [7] noted in particular the giant heat capacity of these compounds at low temperatures. At the time, $CePt_2Sn_2$ was one of the heaviest electron systems known. The large effective mass comes about as a result of mass renormalization due to the Kondo interaction. From a metallic energy band point of view it is useful to describe the enhanced effective electron mass in terms of a high spectral weight acquired right at the Fermi energy $\mathbb{E}_F$ as a result of the sharp $4f-$electron resonance. The $4f-$electron level is situated somewhat below $\mathbb{E}_F$ in the free-ion case but due to bonding in the solid state

this level may be shifted upwards closer to $\mathbb{E}_F$. Admixing of the extremely sharp $4f-$electron resonance with conduction band states close to $\mathbb{E}_F$ which are involved in, for instance, electronic heat capacity and electronic transport processes, means that the electronic behaviour in these physical properties will reflect the very highly correlated states of the electrons. $CeRh_2Sn_2$, the title compound of this study, has the lowest Néel point $T_N = 0.4$ K in the series. Previous heat capacity studies [7] performed down to $T \simeq 0.3$ K revealed the peak in $C_P(T)/T$ associated with the phase transition but were inconclusive as to the electronic behaviour below T_N. Here we extend heat capacity measurements to much lower temperatures. These we compliment with data of magnetic susceptibility at low temperatures and in various applied magnetic fields in an effort to clarify the ground state of this heavy fermion system.

2. Experimental

Polycrystalline samples of $CeRh_2Sn_2$ and $LaRh_2Sn_2$ have been synthesized using starting elements of purity (wt.-%) Ce, La, and Rh (99.99) and Sn (99.999). The syntheses involved direct arc melting in an *in situ* purified argon atmosphere followed by prolonged heat treatment at 800°C for one week. A powder x-ray diffraction survey was done to ascertain the phase purity of synthesized samples. No unreacted elements or spurious phases were detected, enabling the identification of all coherent Bragg peaks according to the expected crystal structure - see figure 1. Table 1 lists the lattice parameters obtained from a full-profile simulation of the spectrum between $2\theta = 20°$ and 75°. There is a 1.75 % unit-cell volume contraction evident from $LaRh_2Sn_2$ to $CeRh_2Sn_2$, which is in line with the well-known lanthanide contraction.

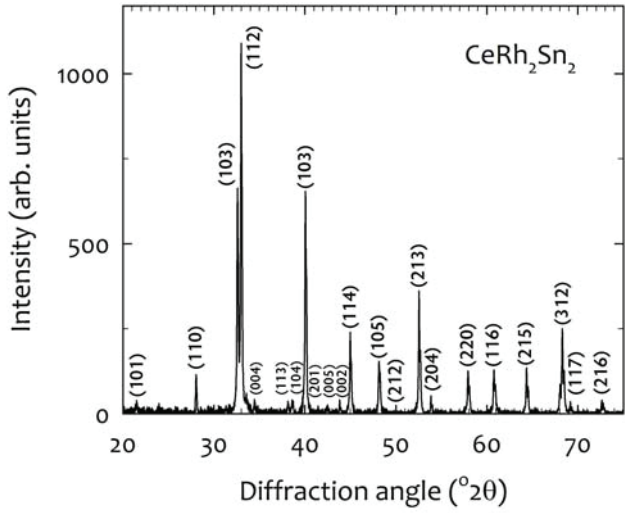


Figure 1. Powder x-ray diffractogram for the representative compound $CeRh_2Sn_2$. The diffraction peaks have been labelled using Miller indices according to the expected $CaBe_2Ge_2$-type structure [2], tetragonal space group $P4/nmm$. The crystallographic unit cell comprises two formula units.

3. Results

A summary of our magnetic susceptibility results collected on $CeRh_2Sn_2$ are shown in figures 2 and 3. Towards higher temperatures, the magnetic susceptibility $\chi(T)$ obtained in a small field of 0.01 T is found to follow the Curie-Weiss law $\chi(T) = (N_A/3k_B)\left[\mu_{eff}^2/(T-\theta_P)\right]$ down to as low as about 10 K. The magnetic fit parameters obtained for the susceptibility of $CeRh_2Sn_2$ are $\mu_{eff} = 2.26(1)\ \mu_B$ for the effective moment and a Weiss temperature $\theta_P = -5.53(6)$ K. The effective moment value is slightly less than the full free-ion value of Ce^{3+}, an observation which may point to Kondo moment screening in this compound. The negative sign of the Weiss temperature suggests a net *anti*ferromagnetic coupling between Ce spins. Magnetization

Table 1. Lattice parameter values obtained from profile fitting of powder x-ray diffraction spectra of $CeRh_2Sn_2$ and $LaRh_2Sn_2$ with diffraction spectrau simulations based on the tetragonal $P4/nmm$ space group. Unit-cell sites have been taken from Atomic occupations have been taken from the work of Selsane *et al* [1].

Compound	a (Å)	c (Å)	V (Å^3)
$LaRh_2Sn_2$	4.519(1)	10.471(1)	213.84(1)
$CeRh_2Sn_2$	4.492(1)	10.413(1)	210.09(1)

isotherms (figure 3) show linear-in-field $M(B)$ behaviour at small fields up to about 1 T, confirming that a field of 0.01 T used for $\chi(T)$ measurements (figure 2) is in fact representative of the intrinsic $(\partial M/\partial B)_{B\to 0}$ behaviour of the magnetic susceptibility. Towards low temperatures however $M(B)$ is found to become strongly curved and at our lowest accessible temperature of 0.46 K for this particular measurement a tendency towards $M(B)$ saturation is noted, reaching $B \simeq 1.3$ μ_B/Ce at the high-field limit of 7 T. Taking $g_J = 6/7$ and the full six-fold multiplet for $J = 5/2$ appropriate for Ce^{3+} yields a theoretical saturation magnetization of $g_J J = 2.14$ μ_B per Ce ion. Our measured magnetization at 7 T thus extracts only about 60% of the full moment. A likely reason for this moment deficiency is magnetocrystalline anisotropy, according to which only a part (typically one third) of the moment measured on a polycrystalline sample corresponds to that of the magnetically easy direction, while the crystallographic directions that are much harder to magnetize make up the major part of the gauged magnetization.

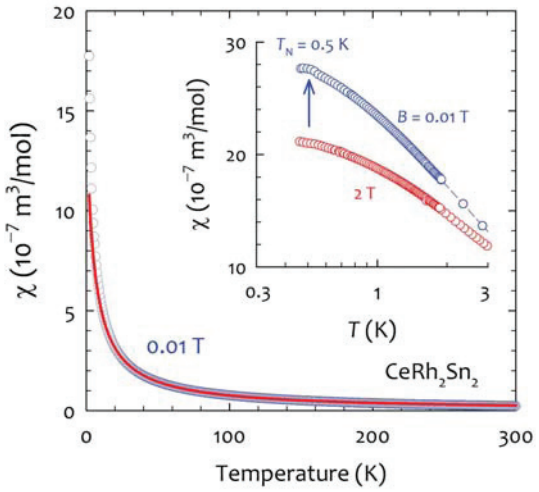


Figure 2. (main panel) Magnetic susceptibility $\chi(T)$ of $CeRh_2Sn_2$ measured in a small field of 0.01 T over a wide temperature range. The solid line shows adherence of the data to the Curie-Weiss law. (inset) Low-temperature $\chi(T)$ measured in two different fields.

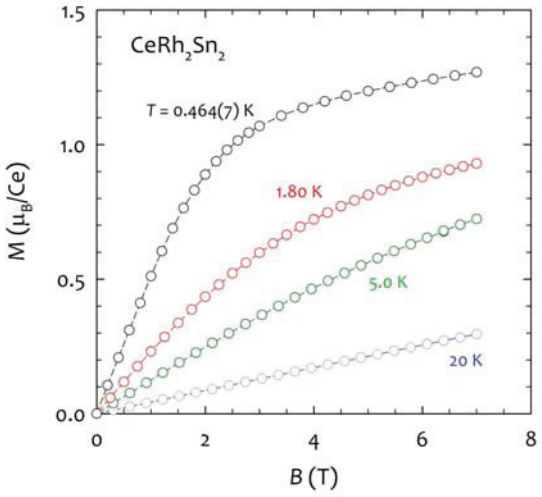


Figure 3. Field-dependent magnetization $M(B)$ of $CeRh_2Sn_2$ at a number of successively lowered temperatures. $M(B)$ curvature becomes progressively severe towards lower temperatures and echoes the lowered susceptibility in higher fields shown in figure 2 (left, inset).

The specific heat $C_P(T)$ of $CeRh_2Sn_2$, see figure 4, reveals a large anomaly associated with the magnetic phase transition through a peak reaching $C_P = 3.16$ J/mol^2 at $T_N = 0.53$ K. The peak height is severely affected by applied fields and in 6 T the transition is no longer visible. Comparing the data collected in zero and in 2 T for data cast in the form $C_P(T)/T$, see figure 4(b), reveals how a massive amount of entropy becomes displaced towards higher temperatures as the magnetic ordering is turned into spin-disorder related entropy. Figure 4(b) inset depicts an estimation of the Sommerfield coefficient $\gamma = C_P(T)/T - \beta T^2$ in a $C_P(T)/T$ *vs* T^2 plot. We obtain $\gamma = 0.23(5)$ J/mol·K^2. This value represents a considerable enhancement over a regular metallic-like value ~ 5 mJ/mol·K^2. A relation that is often applied to Kondo systems to connect the Sommerfield coefficient $\gamma = C_P(T)/T - \beta T^2$ with the Kondo scale T_K is given by $\gamma T_K = R \ln 2$ and returns for $CeRh_2Sn_2$ a Kondo temperature of $T_K \sim 25$ K. It has to be noted however that the γ and T_K values estimated here were drawn from the *total* specific heat of $CeRh_2Sn_2$. The major anticipated nonmagnetic contribution to $C_P(T)$ ought to originate in the lattice or phononic heat capacity, but at temperatures as low as 1 K and below the nonmagnetic component may reasonably be expected to be small and having an imperceptible temperature dependence when compared to that of the magnetic compound $CeRh_2Sn_2$.

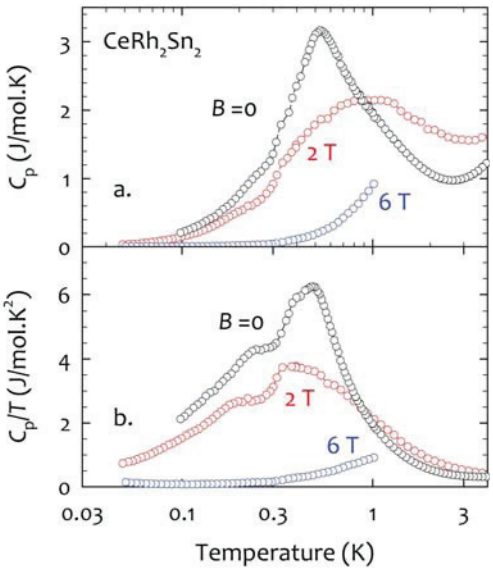


Figure 4. (a) Semi-log plot of specific heat $C_P(T)$ of $CeRh_2Sn_2$ in different applied magnetic fields. (b) Semi-log plot of specific heat cast in the form $C_P(T)/T$ at the same applied fields as in (a). In a field amounting to $B = 6$ T (blue) the phase transition anomaly becomes entirely inconspicuous.

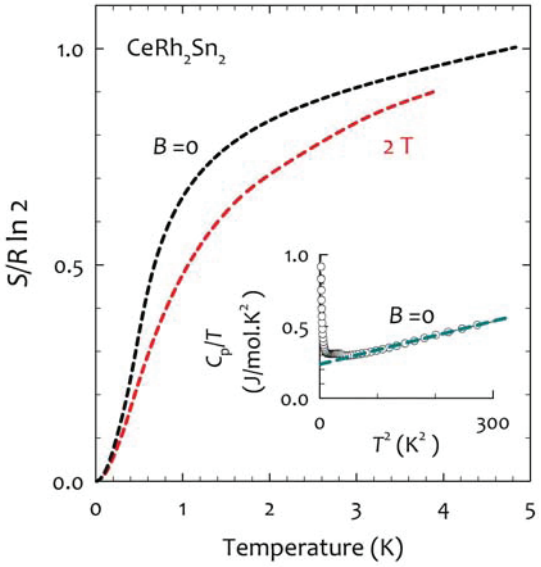


Figure 5. Entropy $S(T)$ of $CeRh_2Sn_2$ calculated from $C_P(T)/T$ data (see figure 4) in applied fields of zero and $B = 2$ T. Units of S are in terms of the entropy of a doublet magnetic state $R \ln 2$. (inset) $C_P(T)/T$ *vs* T^2 of $CeRh_2Sn_2$ to estimate the electronic Sommerfeld coefficient γ at $T \to 0$.

The entropy of $CeRh_2Sn_2$ has been calculated using $S(T) = \int_0^{T'} C_P(T)/T \, dT$ - see figure 5. Entropy S is shown in units of $R \ln 2$ to facilitate comparison with the expected doubly degenerate ground state of Ce^{3+}. At a temperature corresponding to the maximum $T_{max} = 0.53$ K noted in figure 4, the calculated entropy amounts to only about 37% of that of the full doublet

out of which magnetic ordering is established. The entropy plot of figure 5 furthermore shows that the full doublet entropy is only recovered at about ~ 5 K, which is well above $T_{\rm N}$. There is clearly a considerable amount of entropy unaccounted for right at the phase transition, and this spin entropy is evidently only released well above $T_{\rm N}$. This observation is interpreted in terms of short-range correlations which precursor the long-range ordering at $T_{\rm N}$. Above 5 K, the system is expected to gain a total further entropy equaling $R\ln 6 - R\ln 2$ according to two further doublets beyond the ground state.

4. Conclusions

$CeRh_2Sn_2$ reveals stable local-moment behaviour near and below room temperature. The magnetic susceptibility is commensurate with the trivalent magnetic state of Ce, as well as a net antiferromagnetic exchange between Ce spins. At low temperatures the spin system becomes strongly susceptible to the influence of an applied magnetic field when thermal energy has been quenched. Here, the exchange coupling becomes dominant and eventually results in long-range magnetic ordering at $T_{\rm N} \simeq 0.53$ K. Our specific heat measurements revealed that a major part of the ground state spin entropy resides in fact somewhat above $T_{\rm N}$, which questions the previous interpretation of the heavy-electron state to which the specific heat enhancement was ascribed [7]. Future studies will need to track the magnetic ordering by magnetic susceptibility through $T_{\rm N}$. ac-Susceptibility measurements are in progress in order to quantify the $T > T_{\rm N}$ correlations, as well as studies to determine the phononic contribution to $C_{\rm P}(T)$ of $CeRh_2Sn_2$.

5. Acknowledgments

The author thanks the SA-NRF (2072956), the Faculty of Science and the Central Research Committee of the University of Johannesburg for financial assistance. D Britz is thanked for assistance in sample synthesis.

References

[1] Selsane M, Lebail M, Hamdaoui N, Kappler JP, Noel H, Achard JC, and Godart C 1990 *Physica* B **163** 213.
[2] Selsane M, Achard JC, Godart C, Beyermann WP, Hundley MF, Canfield PC, Smith JL, Thompson JD, Hamdaoui N, and Kappler JP 1991 *Eur. J. Solid State Inorg. Chem.* **28** 567, *ibid* Erratum 1991 **28**.
[3] Beyermann WP, Hundley MF, Canfield PC, Godart C, Selsane M, Fisk Z, Smith JL, and Thompson JD 1991 *Physica* B **171** 373.
[4] Mentink SAM, Nieuwenhuys GJ, Mydosh JA, Buschow KHJ, and du Plessis P de V 1995 *J. Magn. Magn. Mater.* **140-144** 887.
[5] Takabatake T, Teshima F, Fujii H, Nishigori S, Suzuki T, Fujita T, Yamaguchi Y, and Sakurai J 1990 *J. Magn. Magn. Mater.* **90 & 91** 474.
[6] Beyermann WP, Hundley MF, Canfield PC, Thompson JD, Fisk Z, Smith JL, Selsane M, Godart C, and Latroche M 1991 *Phys. Rev. Lett.* **66** 3289.
[7] Beyermann WP, Hundley MF, Canfield PC, Thompson JD, Latroche M, Godart C, Selsane M, Fisk Z, and Smith JL 1991 *Phys. Rev.* B **43** 13130.

Thermal and electronic transport studies of the Kondo energy scale in the heavy-fermion system $CeCu_{5-x}Al_x$

A M Strydom and D Britz

Physics Department, University of Johannesburg, PO Box 524, Auckland Park 2006, South Africa

E-mail: amstrydom@uj.ac.za

Abstract. The binary compound $CeCu_5$ is an archetypal heavy-fermion Kondo lattice. Strong hybridization effects between the localized 4f-electron spins of the Ce and degenerate conduction electrons result in strong magnetic moment renormalization, but antiferromagnetic ordering sets in below $T_N = 3.9$ K. Here we report on results of the electronic (electrical resistivity) and thermal transport (thermoelectric power) studies in the pseudo-binary system $CeCu_{5-x}Al_x$, in order to explore commonalities in the way these two physical properties are affected by the Kondo interaction.

1. Introduction

Ce-based compounds are known for a variety of complex magnetic phenomena. The typically localized 4f-electron spin of the magnetic Ce^{3+} ion in quantum-mechanical contact with degenerate conduction electrons produces the moment screening associated with the Kondo interaction. A lattice framework of 4f spins, on the other hand, is generally susceptible to the long-range RKKY interaction, which may result in antiferromagnetic order. The delicate nature of the ground state especially in Ce and Yb compounds, where the 4f electron might be suitably disposed toward hybridization effects, are determined by these competing magnetic interactions. These magnetic interactions may generally be adjusted by an experimental variable. This has been achieved in this work by doping the parent compound, on the normally non-magnetic Cu site forming the pseudo-binary $CeCu_{5-x}Al_x$ compound. This leaves the Ce framework of magnetic centres invariant, to first order.

$CeCu_5$ orders antiferromagnetically (AFM) at a Néel point $T_N = 3.9$ K [1]. A Kondo scale $T_K = 2.2$ K determined from the effective paramagnetic moment [2], places the on-site Kondo interaction and the long range AFM interaction at competing energy levels. The introduction of Al (donating 3 valence electrons) into the $CeCu_5$ would result in the augmentation of the density of states close to the Fermi energy, which is also the location of the magnetic and hybridized 4f electrons. This is attested to by the enhanced electronic specific heat of $CeCu_5$, $\gamma = 100$ mJ.mol^{-1}K^{-2} [3], which reflects its heavy-fermion character. Substituting Cu by Al furthermore achieves a decrease in the AFM ordering, up to the point in $CeCu_3Al_2$ where T_N becomes undetectable down to the lowest measured temperatures [3].

2. Experimental

Polycrystalline samples were prepared by ultra-high purity Argon arc melting. Stoichiometric quantities of the starting elements with purities in wt% were Ce: 99.99, Cu: 99.995 and Al: 99.9999. The samples were melted 4 times and flipped after each melting process to promote homogeneity. The mass loss for the melting process was confined to within 0.2 % on average, which ensured that the compounds were still stoichiometric.

Measurements of electrical resistivity, $\rho(T, B)$, were performed on all compounds, using a standard four-probe dc method between 20 and 2 K and in applied magnetic fields up to 8 T. Samples for $\rho(T, B)$ studies were bar shaped and of typical dimensions 6 $\times$ 1 $\times$ 1 mm^3. Samples for thermoelectric power, $S(T)$, studies were bar shaped and of typical dimensions 6 $\times$ 1.5 $\times$ 1.5 mm^3. The magnetoresistance and thermoelectric power were measured with a Physical Properties Measurement System (Quantum Design, San Diego), which provides stable and controllable sample temperature environment and high resolution voltage detection.

3. Results

The isothermal electrical resistivity in applied magnetic fields of up to 8 T showed no signs of hysteresis during increasing and decreasing field runs. The experimental results were found to conform with the results of the Bethe ansatz calculations of the Coqblin Schrieffer model as given by Schlottmann for the single-ion Kondo magnetoresistivity. The predicted relation for the $j = 1/2$ integer-valence limit is [4]:

$$\frac{\rho(B=0)}{\rho(B)} = \frac{1}{2j+1}\sin^2\left(\frac{\pi \mathrm{n_f}}{2\mathrm{j}+1}\right)\sum_{\mathrm{l}=0}^{2\mathrm{j}}\sin^{-2}(\pi \mathrm{n_l}) \tag{1}$$

Figure 1 is a representative example of a least-squares (LSQ) fit of equation 1 (solid lines) to the experimental magnetoresistance (MR) data (symbols). In the Schlottmann formalism the MR is completely determined by a single characteristic field B^*, which follows a linear-in-T relationship [5]:

$$B^*(T) = B^*(0) + \frac{k_\mathrm{B}T}{g\mu_\mathrm{K}} = \frac{k_\mathrm{B}(T_\mathrm{K}+T)}{g\mu_\mathrm{K}} \tag{2}$$

where μ_K is the renormalized Kondo moment, after screening has occurred, and $B^*(0) = k_\mathrm{B}T_\mathrm{K}/g\mu_\mathrm{K}$. A LSQ fit of $B^*(T)$ vs T gives the Kondo temperature and Kondo moment (MR data for $CeCu_3Al_2$ was obtained from [6]).

The fit results of our magnetoresistivity study are summarised in Table 1:

Table 1. Data obtained from the two fitting procedures discussed in the text.

Compound	T_K^MR(K)	$\mu_\mathrm{K}^\mathrm{MR}(\mu_\mathrm{B})$	T_K^S(K)	N_f
$CeCu_4Al$	2.7(2)	0.0681(9)	2.2(1)	0.195(3)
$CeCu_{3.5}Al_{1.5}$	7.2(2)	0.098(1)	4.9(2)	0.454(6)
$CeCu_3Al_2$	11.0(5)	0.1(1)	10.5(2)	0.945(2)
$CeCu_{2.9}Al_{2.1}$	7.4(2)	0.089(1)	12.0(3)	1.10(1)
$CeCu_{2.8}Al_{2.2}$	8.9(2)	0.115(2)	13.2(3)	1.24(1)
$CeCu_{2.7}Al_{2.3}$	5.3(2)	0.131(1)	12.4(3)	1.17(1)
$CeCu_{2.6}Al_{2.4}$	3.6(2)	0.143(2)	9.2(3)	0.659(8)

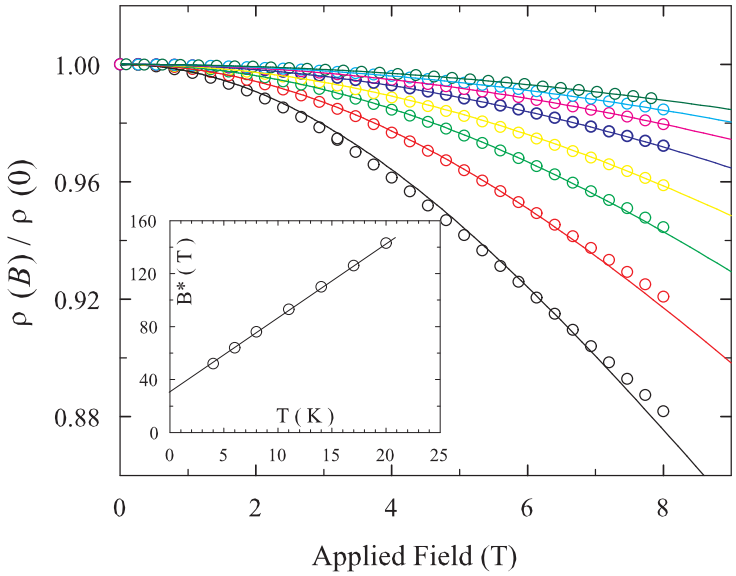


Figure 1. The magnetoresistance isotherms from 2-20 K for the representative compound $CeCu_{2.7}Al_{2.3}$ where the symbols correspond to the experimental data and the lines correspond to the Schlottmann expression for the Bethe ansatz solution to the single-ion Kondo solution, Eq. 1. Inset: B^* obtained from fits in main panel follows a linear-in-T behaviour, which was fitted with Eq. 2

The values obtained for μ_K are far below the expected free-ion value for Ce^{3+} ($\mu_{eff} = 2.54\mu_B$), which illustrates the extreme moment renormalization that occurs as a result of the Kondo effect in these systems. This analysis has been used to obtain the Kondo temperatures for many other systems: CePtIn [7], $CeCu_3Ga_2$ [3], $CeCu_5In_{0.6}$(Al or Ga)$_{0.4}$ [8] for example. The maximum in T_K at $x = 2$ could be caused by the non-Fermi liquid ground state formed by the suppression of T_N by the Kondo interaction.
In the following, we discuss a representative measurement of the thermoelectric power $S(T)$ for $CeCu_{2.8}Al_{2.2}$ see Figure 2. A pronounced peak is found in $S(T)$ at $T_{max} = 25$ K. The positive peak is due to the 4f energy level being slightly above the Fermi energy, causing a sharp change in the density of electronic states near the Fermi energy $\left(\frac{\partial N(\epsilon)}{\partial \epsilon}|_{\epsilon_F}\right)$, which in turn is proportional to the electronic thermoelectric power. $S(T)$ is characterised by a phenomenological approach developed for intermediate valence (IV) systems [9, 10]:

$$S(T) = \frac{2}{3}\pi^2 \frac{k_B}{e} \frac{T \cdot T_K}{(\pi^2/3)T^2 + (1 + \pi^2/N_f^2)T_K^2} \tag{3}$$

Here $N_f = 2J + 1$ is the degeneracy of the unperturbed electronic multiplet of Ce^{3+} and e is the fundamental electron charge. A LSQ fit to the thermoelectric power Figure 2 was used to determine the values of T_K^S and N_f for the dilution compounds in the $CeCu_{5-x}Al_x$ series. The results are summarized in Table 1.
The reason for the deviation of the fit at $T > T_{max}$ may be found in the action of the crystal-electric field splitting of the 6-fold $J = 5/2$ multiplet of the Ce^{3+} ion. Scattering of charge carriers from higher excited 4f-electron states can be expected to contribute higher-order corrections to the calculated scattering obtained from flat conduction bands [11]. The crystal field

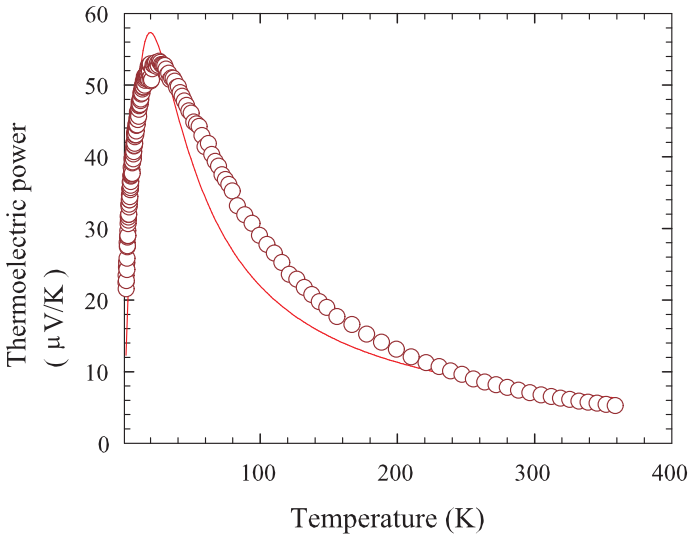


Figure 2. Phenomenological fit (solid line) to the thermoelectric power $S(T)$ of $CeCu_{2.8}Al_{2.2}$.

level scheme (Δ_{CEF}) for $CeCu_4Al$ was found to be 0-65-85 K [12].

4. Conclusion

The series $CeCu_{5-x}Al_x$ for $x \in \{1.0, 1.5, 2.0, 2.1, 2.2, 2.3, 2.4\}$ has been found to follow the single-ion Kondo-type electronic scattering behaviour around 4 K. The Kondo temperatures and associated Kondo moments for these systems can be determined within the Bethe ansatz model of the magnetoresistance, which provides an exact value of T_K. The phenomenological model [9, 10] developed for intermediate valence systems is also applicable to Kondo systems, for the data corresponded to the data obtained from the exact single-ion theory applied MR. Therefore the Kondo temperature can be estimated from the thermoelectric power data provided that the underlying CEF splitting is not large enough to produce multiple peaks in $S(T)$. Our results provide a firmer ground for calculating and comparing Kondo dynamics through magnetoresistance and thermoelectric power data in Kondo systems such as $CeCu_5$.

A.M. Strydom thanks the NRF (2072956) and the URC of the University of Johannesburg for financial support. D. Britz acknowledges financial assistance from the Next Generation Scholarship program of the University of Johannesburg.

5. Reference

[1] Bauer E, Gignoux D, Schmitt D and Winzer K 1987 *J. Magn. Magn. Mater.* **69** 158
[2] Bauer E, Kohlmann J and Winzer K 1990 *Physica B* **163** 188
[3] Bauer E 1991 *Adv. Phys.* **40** 417
[4] Schlottmann P 1983 *Z. Phys B* **51** 223
[5] Andrei N, Furuya K and Lowenstein J 1983 *Rev. Mod. Phys.* **55** 331
[6] Strydom A 2010 *Phys. Status Solidi B* **247** 713
[7] Regal F, Plessis P and Strydom A 2009 *J. Phys.: Condens. Matter* **21** 046008
[8] Tchoknté M, Plessis P and Strydom A 2007 *J. Magn. Magn. Mater.* **310** e33
[9] Koterlyn M, Babych O and Koterlyn G 2001 *J. Alloy Comp.* **325** 6
[10] Koterlyn M, Yasnitskii R, Koterlyn G and Morokhivskii B 2003 *J. Alloy Comp.* **348** 52
[11] Zlatić V, Horvatic B, Milat I, Coqblin B, Czycholl G and Grenzebach C 2003 *Phys. Rev. B* **68** 104432
[12] Bauer E, Pillmayr N, Müller H, Kohlmann J and Winzer K 1990 *J. Magn. Magn. Mater.* **90** 441

Synthesis and characterization of the ZnO nanoparticles and the polyvinyl pyrrolidone (PVP) encapsulated ZnO nanoparticles

[1]M.A. Tshabalala, [1]B.F. Dejene[*] and [2]H.C. Swart[*]

[1]Physics Department, University of the Free State, Private Bag X13, Phuthaditjhaba, 9866, South Africa

[2]Physics Department, University of the Free State, P. O. Box 339, Bloemfontein, 9300, South Africa

[*]Corresponding author: Dejenebf@qwa.ufs.ac.za or SwartHC@ufs.ac.za

Abstract ZnO and Polyvinyl pyrrolidone (PVP) capped ZnO nanoparticles were synthesised using the sol-gel method at low temperature with ethanol as a solvent, zinc acetate as a precursor and methanol as the cleaning agent. The influence of the PVP on the morphology, structure and the optical properties of the ZnO nanoparticles were investigated. The effect of the addition of different masses of the PVP during the synthesis on the ZnO emission peaks was systematically monitored. PVP is utilized to cap the ZnO nanoparticles from a zinc acetate precursor at low temperature, since the PVP can be easily removed by burning the solid products. The photoluminescence characterization of the ZnO nanostructures exhibited a broad emission in the visible range with maximum intensity peaks at 449 and/or 530 nm, this was influenced by the addition of different molar masses of the PVP. The scanning electron microscopy images of ZnO and PVP capped ZnO has showed the presence of the agglomerated ZnO particles which could be due to the agglomeration of smaller particles. The x-ray diffraction spectra for ZnO nanoparticles show the entire peaks corresponding to the various planes of wurtzite ZnO, indicating a single phase. The absorption edges of these PVP encapsulated ZnO did not show any shifting. The absorption spectra of the ZnO showed slight shifts with reference to the various masses of PVP.

1. Introduction

As a semiconductor material, zinc oxide spans a wide range of applications from solar cells and chemical sensors to electrical, acoustic and luminescent devices [1]. ZnO is also a versatile direct-band gap semiconductor with a wide band gap of 3.37 eV, and its large exciton binding energy (60 meV) makes the exciton state stable even at room temperature. Furthermore, it is an environmentally friendly material and shows a broad luminescence emission spectrum in the blue yellow region [2]. ZnO appears to be a major candidate for room temperature optoelectronic applications such as efficient short-wavelength ultraviolet light-emitting diodes and laser diodes [3]. Various processes have been employed to synthesize the ZnO nanostructures, such as electrochemical deposition [4], hydrothermal [5], sputter deposition technique [6], and vapour method [7]. From all these methods, we have chosen to use the sol-gel method because of its reliability, inexpensiveness and most of all simplicity. The crystalline quality of ZnO films is determined not only by the growth processes, but also by the do-pants, impurities, surfactants and the surface modifiers such as polymer matrices used.

The nanoparticles which are produced by this route show good optical properties, but this can be achieved only by good control of the size and the morphology of the particles [8]. Several workers have used capping agents such as poly vinyl pyridine (PVP), poly ethylene glycol (PEG) etc to stop particle agglomeration and obtained nanoparticles of size less than 5 nm and these polymers produce spherical nanoparticles in shape because of the property of surface tension [9]. PVP is utilized to make ZnO nanorods from a zinc acetate precursor at a low temperature, since the PVP can easily be removed by burning the solid products [10]. In this investigation we concentrating on the encapsulated ZnO nanoparticles and the effect that PVP have on the optical properties and the morphology and structure of the ZnO powders.

2. Experimental

The solutions of ZnO were prepared by dissolving 5.508 g of Zinc Acetate in 300 mL of ethanol with different masses of PVP. This was repeated six times with the following different molar masses; 0.5, 0.75, 1.0, 1.5, 2.0 and 2.5 g. The mixtures were magnetically stirred for about 24 hours at a temperature of 80 °C, this was to let the mixture mix thoroughly. The mixtures where let on the stirrer until the gel-like of the zinc acetate and PVP was obtained. It was then washed using methanol to remove the impurities. Then the gel was calcined in a furnace at 150 °C for 2 hours. The X-ray diffraction (XRD) patterns were recorded to characterize the phase and crystal structure of the nanoparticles using a D8 Bruker Advanced AXS GmbH X-ray diffractometer (XRD), room temperature photoluminescence (PL) of the samples was measured, using a He–Cd laser (325 nm) as excitation sources, the morphology of the nanoparticles was observed by a Shimadzu Superscan SSX-550 system scanning electron microscope (SEM) operated at 20 kV equipped with energy dispersive X-ray spectroscopy (EDS). Optical absorption was performed on a Perkin Elmer Lamb 950 UV-visible spectrophotometer (UV).

3. Results and Discussion

3.1. Structure

In this part, we studied the role of PVP molar masses in controlling the size and shape of the PVP encapsulated ZnO nanoparticles. Figure 1(a) represents the SEM image of the PVP encapsulated ZnO nanoparticles. The figure shows the chunk of particles which probable consist out of agglomerated smaller particles in the PVP matrix, the agglomeration is due to the aggregating or overlapping of smaller particles. However, small agglomerates on some of the large particles are clearly visible at high magnifications. Therefore it is concluded that the chosen PVP/Zn^{2+} molar ratio range has influence on the morphology of the ZnO samples. Fig. 1(b) is the EDS for the encapsulated ZnO nanoparticles. It gives out the elements that exist in the sample. All expected elements (Zn, O) were detected and the adventitious carbon is also to be seen. Figure 2 (a) and (b) are the powder x-ray diffraction pattern (XRD) of the ZnO nanoparticles and the 0.5 g PVP encapsulated ZnO nanoparticles. The ZnO nanoparticles pattern showed the standard diffraction pattern of wurtzite hexagonal ZnO. Fig. 2(b) shows that the XRD peak positions shifted as well as the peaks have broadened. This confirms that the PVP has an effect on the structure of the ZnO nanoparticles. The crystallite size was calculated from the peaks width using the Debye-Scherer formula

$$t = \frac{0.9\lambda}{B\cos\theta}$$

Where λ is the x-ray wavelength (1.5406Å), Θ is the Bragg diffraction angle, t is the diameter of the crystallites, and B is the peak width at half maximum (FWHM). The particle sizes of the encapsulated PVP ZnO and ZnO nanoparticles were calculated were found to be 13 nm and 42 nm.

The XRD peaks in Fig. 2(a) gave lattice parameters a = 3.25 Å and c = 5.20 Å and that of Fig. 2(b) are a = 3.25 Å and c = 4.88 Å.

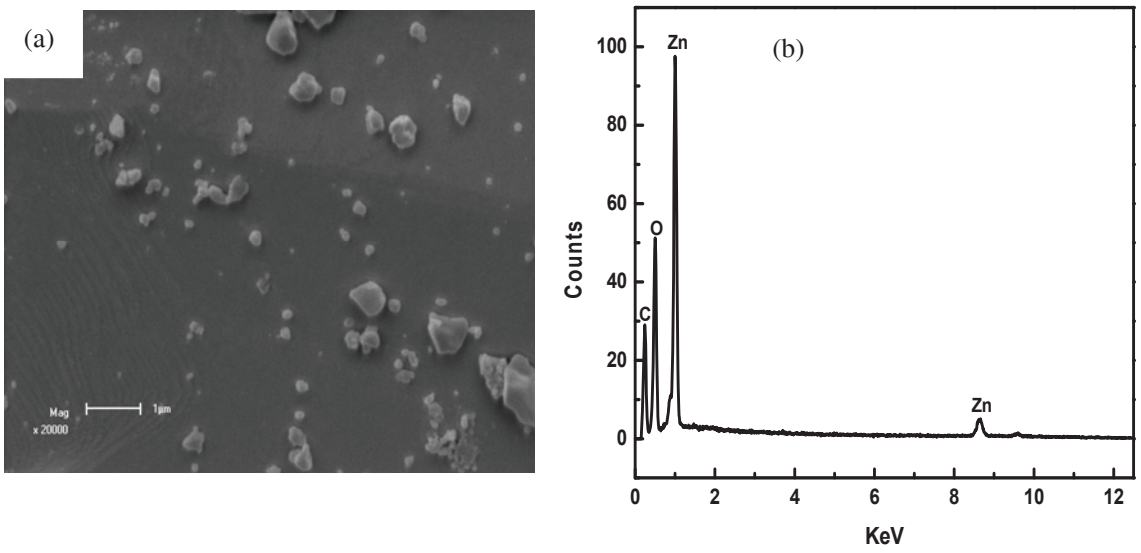


Figure 1: (a) SEM image of the ZnO (b) EDS spectra for the ZnO nanoparticles capsulated in PVP which confirms the presence of the different elements.

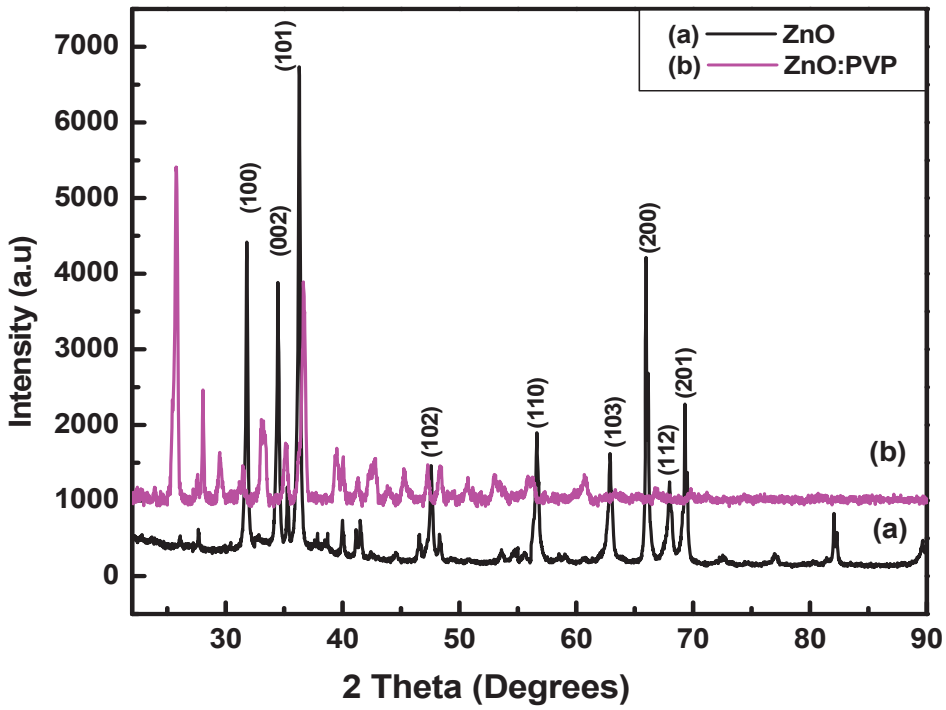


Figure 2: XRD patterns of (a) ZnO nanoparticles and (b) PVP encapsulated ZnO nanoparticles.

3.2. Optical properties

The photoluminescence spectra of the seven different concentrations of PVP samples, which were recorded with the excitation wavelength set at 325 nm, as shown in Figure 3. The two emission peaks, one a blue emission peak around 448 nm and another green emission peak around 532 nm, were observed for all samples except for the sample with no PVP. The sample without PVP shows only one violet emission peak around 403 nm. Comparing the maximum intensity and peak shape of the six samples with the one without PVP, it is observed that PVP has a clear effect on the ZnO nanoparticles. The relative intensities of these two emission band vary drastically with the molar ratio of Zn:PVP. The addition of the PVP has brought an increase in the intensity with various molar masses. Particle agglomeration of the nanoparticles was prevented by the addition of the PVP and due to the properties such as surface tension a shift and increase in PL peak position and intensity were obtained.

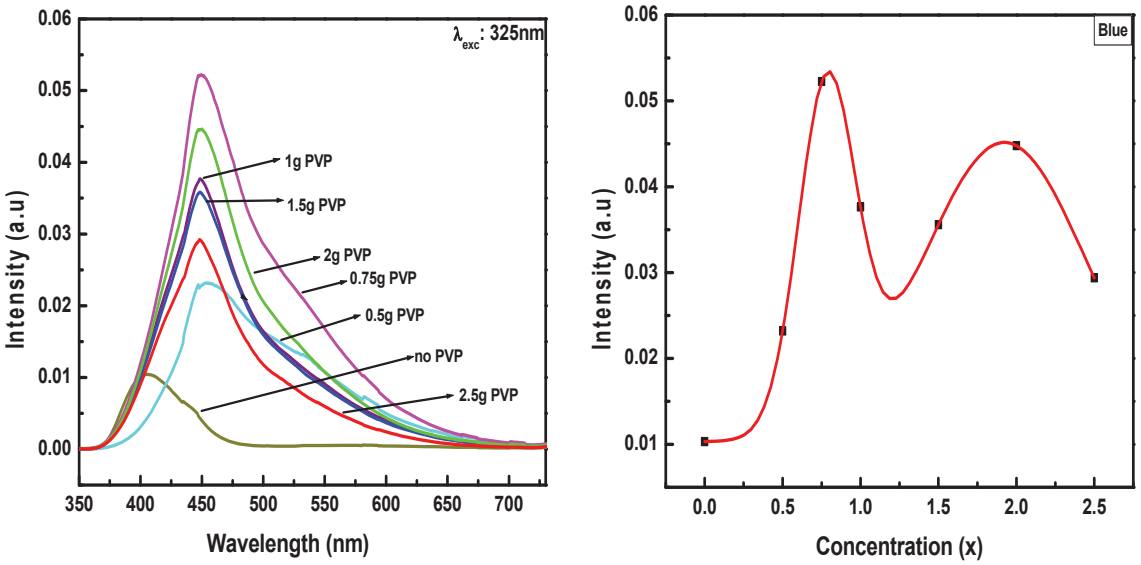


Figure 3: (a) PL emission spectra of the ZnO with different masses of PVP (b) PL emission intensity versus various masses of the PVP encapsulated ZnO nanoparticles.

Fig. 3(b) shows the intensity versus the concentration or different masses of PVP which shows that the 0.75g of PVP has the highest intensity followed by the 2.0 g of PVP. The absorption spectra of the PVP encapsulated ZnO nanoparticles with different masses of PVP are shown in Fig. 4. The absorption peaks of PVP encapsulated ZnO are obtained at the wavelength of 272 nm and the other one at around 358 nm.

4. Conclusion

In this investigation, the synthesis of PVP encapsulated ZnO nanoparticles with different molar mass of PVP was reported. This was done with the sol-gel method. The characterization was done using XRD, SEM, PL, and UV-VIS absorption. The PL covered the wavelength range of 350-750nm.

Acknowledgement

The financial support of the National Research Foundation (NRF) for funding the project and the University of Free State Physics (Physics department) for the research techniques used in this study.

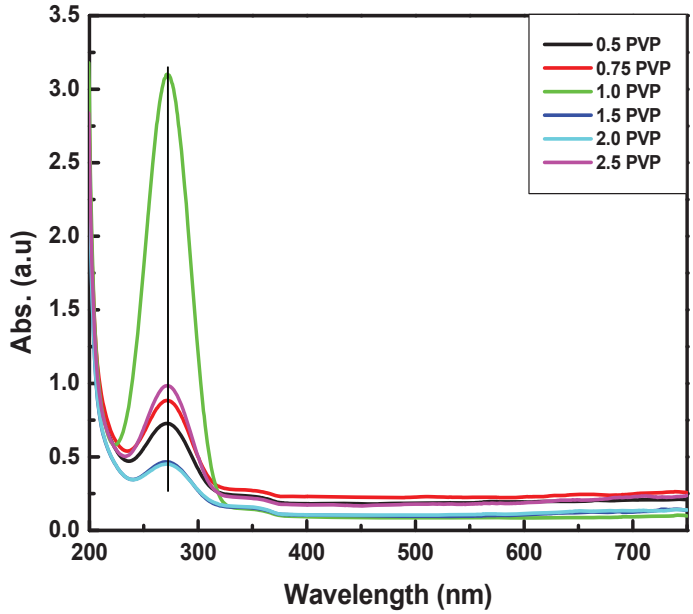


Figure 4: UV-visible absorbance spectra of the PVP encapsulated ZnO nanoparticles.

References

[1] Yang C, Ping Y, Jiannong W, Weikun G, and Wong G K L 2000 *Chem. Mater*. **12** 2268
[2] Yang Y, Li Y, Fu S, and Xiao H 2008 *J. Phys. Chem. C* **112** 10553
[3] Kenanakis G, Androulidaki M, Koudoumas E, Savvakis C, Katsarakis N 2007 *Superlattices and Microstructures* **42** 474
[4] Li G R, Qu D L, Zhao W X, Tong Y X, J. 2007 *Mater. Chem*. **9** 1661
[5] Tam K H, Cheung C K, Leung Y H, Djurišić A B, Ling C C, Beling C D, Fung S, Kwok W M, Chan W K, Phillips D L, Ding L and Geet W K 2006 *J. Phys. Chem. B* **110 (42)** 20865
[6] Chiou W T, Wu W Y, Ting J M 2003 *Diamond Relat. Mater.* **12** 1841
[7] Zhang Y, Yu K, Jiang D, Zhu Z, Geng H, Luo L 2005 *Appl. Surf. Sci.* **242** 212
[8] Vafaee M, Sasani G.M 2007 *Materials Letters* **61** 3265
[9] Atul G, Bhatti H S, Kumar D, Verma N K, Tandon R. P 2006 Digest *Journal of Nanomaterials and Biostructures* **1**(1) 1
[10] Eskandari M, Ahmadi V, Ahmadi S H 2009 *Physica B* **404** 1924

Enhanced green emission from UV down-converting Ce^{3+} and Tb^{3+} co-activated $ZnAl_2O_4$ phosphor

K.G. Tshabalala[1], S-H. Cho[2], J-K. Park[2], I.M. Nagpure[1], J.H. Neethling[3], J.R. Botha[3], H.C. Swart[1] and O.M. Ntwaeaborwa[1,4]

[1]Department of Physics, University of the Free State, Bloemfontein, ZA 9300, South Africa

[2]Nano-Materials Research Center, Korea Institute of Science and Technology, 39-1 Hawolkok, Seoul 136-791, South Korea

[3]Department of Physics, Nelson Mandela Metropolitan University, Port Elizabeth, ZA 6031, South Africa

E-mail: ntwaeab@ufs.ac.za

Abstract. The Ce^{3+} and Tb^{3+} co-activated $ZnAl_2O_4$ nanocrystal phosphors were successfully prepared by a solution combustion method, using urea as a fuel. X-ray diffraction results confirmed the formation of a cubic spinel structure of $ZnAl_2O_4$. Distorted hexagonal and irregular platelet-like particles were observed from the SEM images of undoped $ZnAl_2O_4$ and Ce^{3+}-doped $ZnAl_2O_4$ respectively. The microstructural environment of aluminium ions (Al^{3+}) were shown to change with heat treatment, as observed from the Fourier transform infrared spectra. The photoluminescent data demonstrated enhanced green emission from Tb^{3+} due to energy transfer from Ce^{3+} to Tb^{3+} in the $ZnAl_2O_4$ host by a down-conversion process.

1. Introduction

Zinc aluminate ($ZnAl_2O_4$) is a wide band gap semiconductor (E_g = ~3.8 eV) and it belongs to the class of inorganic materials called spinels. It has a closed-packed face-centered-cubic structure with *Fd***3***m* space group symmetry [1,2]. In a polycrystalline form, $ZnAl_2O_4$ is found to be highly reflective in the UV regime (300 nm) of the electromagnetic spectrum. It has attracted considerable interest among researchers for a variety of applications including catalysis, ceramics and optoeletronics. As a catalyst, $ZnAl_2O_4$ spinel is widely used in reactions such as cracking, dehydration, hydrogenation and dehydrogenation [3]. In this study, $ZnAl_2O_4$ was used as a host matrix of Ce^{3+} and Tb^{3+} ions to prepare an efficient green emitting phosphor that can be used as a UV down-converting layer to improve the absorption efficiency of the conventional silicon (Si) photovoltaic (PV) cells. Down-converted green emission, as a result of energy transfer from Ce^{3+} to Tb^{3+}, was observed when $ZnAl_2O_4$:Ce^{3+},Tb^{3+} powders were excited by a 325 nm HeCd laser.

2. Experimental procedure

Nanocrystals of $ZnAl_2O_4$ co-doped with nominal concentrations of 0.5 mol% Ce^{3+}– 1 mol% Tb^{3+} , 0.75 mol% Ce^{3+}– 1 mol% Tb^{3+}, 1 mol% Ce^{3+}– 0.5 mol%Tb^{3+} and 1 mol%Ce^{3+}– 0.75 mol% Tb^{3+} were prepared by a solution combustion method as described in refs [4,5]. The equivalence (oxidizer:fuel) ratio was calculated based on oxidizing (*O*) and fuel (*F*) valences of the reactants, keeping *O/F* = 1, as reported previously [6]. Zinc nitrate and aluminum nitrate were used as oxidizers, urea (CH_4N_2O) was used as fuel and both cerium and terbium nitrates were used as dopant precusors. All the samples prepared were annealed in a reducing H_2 atmosphere at 700°C for 4 h. The structure and morphology were analyzed using X-ray diffraction (XRD), Scanning electron microscopy (SEM) respectively. The

[4] Corresponding author: email: ntwaeab@ufs.ac.za (OM Ntwaeaborwa)

stretching mode frequencies were determined using Bruker Tensor 27 FTIR spectrometer. A photoluminescence system consisting of a 325 nm HeCd laser, SPEX 1870 0.5 m monochromator and a photomultiplier tube detector was used to record photoluminescence (PL) spectra.

3. Results and discussion

3.1. X-ray diffraction

The XRD patterns of the $ZnAl_2O_4$: Ce^{3+} powders shown in figure 1 indexed well to a pure and highly crystalline face-centered cubic spinel $ZnAl_2O_4$ consistent with the JCPDS file No. 05-0669. The spectra were recorded from as prepared and post-preparation annealed samples. The fact that the patterns were almost identical suggests that highly crystalline $ZnAl_2O_4$ can be obtained using the combustion method even without post-preparation annealing.

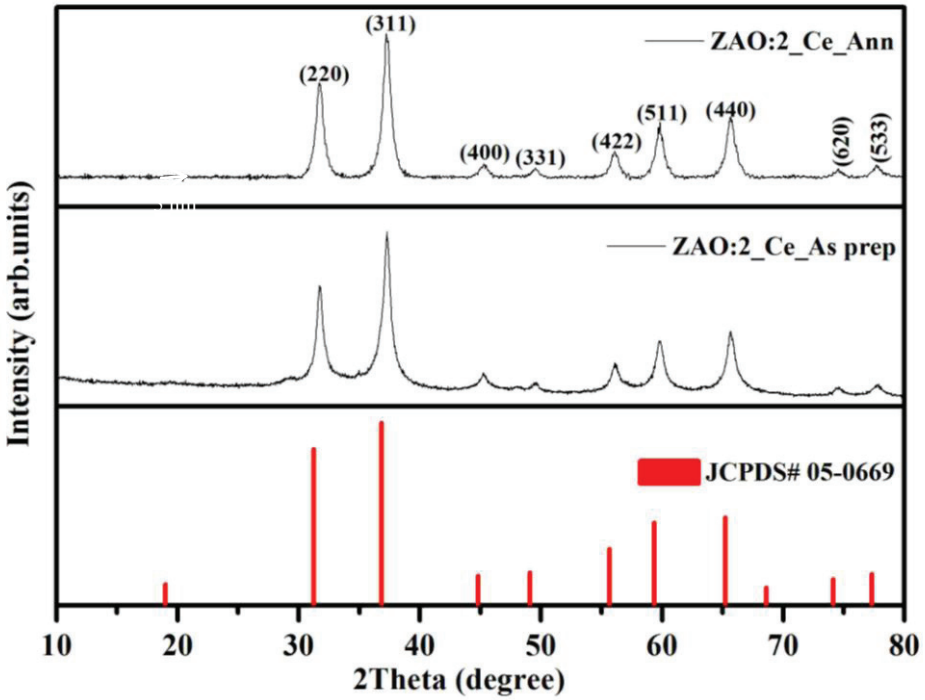


Figure 1. Room Temperature XRD pattern of $ZnAl_2O_4$:2 mol% Ce^{3+} (as-prepared) and $ZnAl_2O_4$:2 mol% Ce^{3+} (reduced in 4%H_2) at 700°C for 4h**.**

3.2. Fourier transform infrared spectroscopy

The FT-IR spectra of the (a) as-prepared $ZnAl_2O_4$: 0.75 mol%Ce^{3+}, 1 mol% Tb^{3+} and (b) $ZnAl_2O_4$: 0.75 mol%Ce^{3+}, 1 mol% Tb^{3+} reduced in H_2 atmosphere at 700°C for 4 h are shown in figure 2. The bands at low energy (400−1000 cm $^{-1}$) are related to Al— O stretching mode with peaks at 825, 698 and 582 cm $^{-1}$ in (a) and 680, 582, 495 and 553 cm $^{-1}$ in (b). These bands are the characteristic of zinc aluminate spinel structure [7]. Generally, the bands from the as prepared sample at low energy values (500 - 1000 cm $^{-1}$) were less intense than similar bands from the annealed sample. The bands at 495, 553, 582, 680 and 698 cm $^{-1}$ from both samples are assigned to the stretching modes of AlO_6 (octahedral site) [8],[9]. A shoulder related to Al^{3+} in a tetrahedral coordination is observed around 825 cm $^{-1}$ [10] in the as-prepared sample, suggesting partial inversion of the spinel structure. The as-prepared $ZnAl_2O_4$:Ce, Tb sample in spectrum in figure 2 (a) exhibits strong vibration modes at 1360 – 1570 cm $^{-1}$ that can be assigned to the groups originating from the organic compounds [11], and these bands were less intense in figure 2 (b) probably due to heat treatment. Furthermore, the bands at 3408,

3352, 1631 and 1632 cm $^{-1}$ can be assigned to the vibration mode of carbon containing groups and to the deformation vibration of water molecules [12]

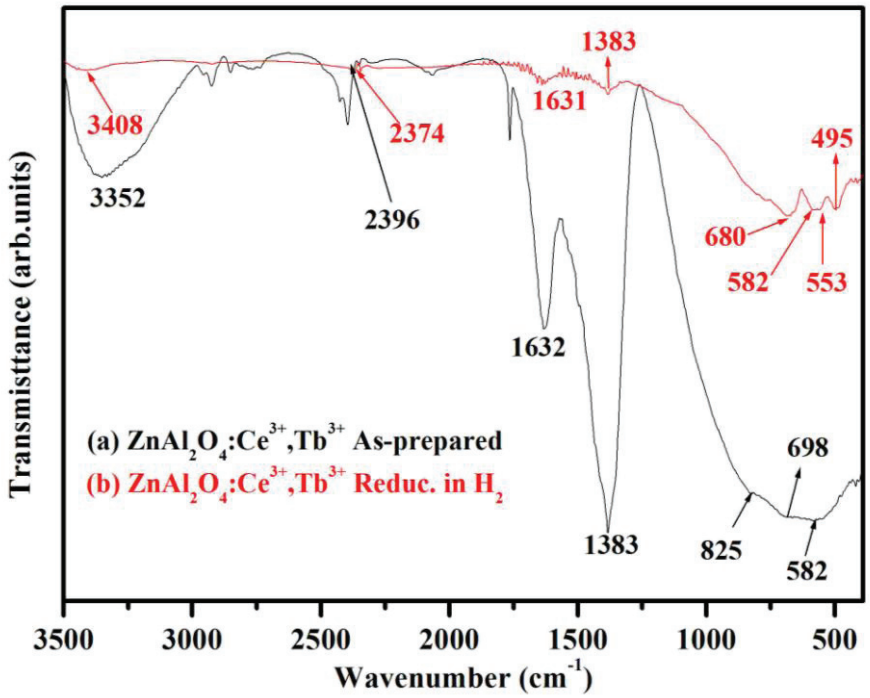


Figure 2. The FTIR spectra of the (a) as-prepared $ZnA_{l2}O_4$: 0.75 mol%Ce^{3+},1 mol% Tb^{3+} and $ZnAl_2O_4$: 0.75 mol% Ce^{3+}, 1 mol% Tb^{3+} reduced in H_2 atmosphere at 700°C for 4 h.

3.3. High resolution scanning electron microscopy (SEM)

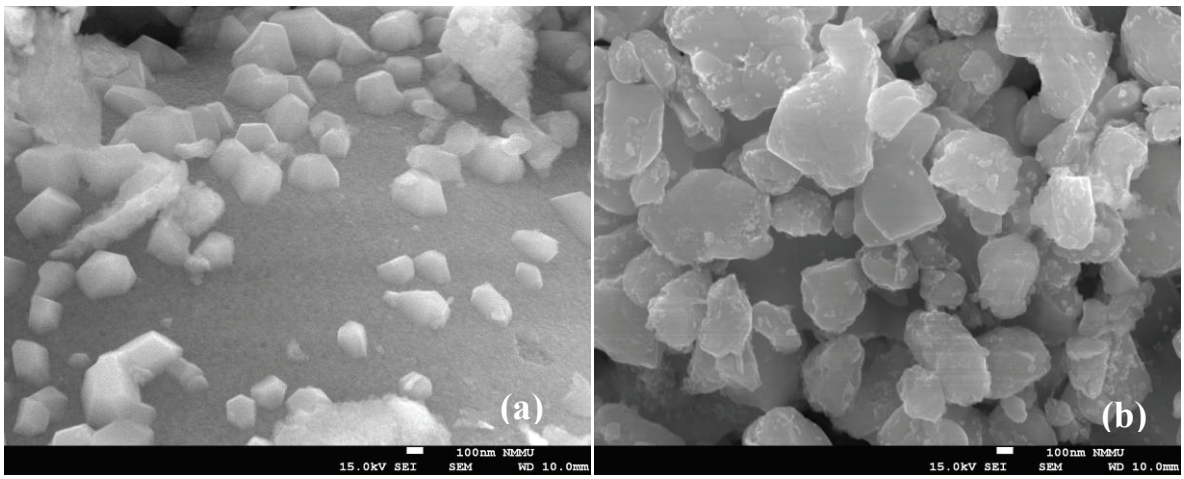


Figure 3. SEM images of the (a) $ZnAl_2O_4$ host and (b) $ZnAl_2O_4$:1mol% Ce^{3+} samples.

The SEM images in figure 3, provides the general morphology of the (a) $ZnAl_2O_4$ host and (b) $ZnAl_2O_4$:1mol% Ce^{3+} respectively. As shown in figure 3 (a), the $ZnAl_2O_4$ host was made up of particles with distorted hexagonal edges and corner angles and the well known characteristic platelet-like particles (figure 3(b)) of the combustion method where obtained after incorporating Ce^{3+} ions.

3.4. Photoluminescence studies

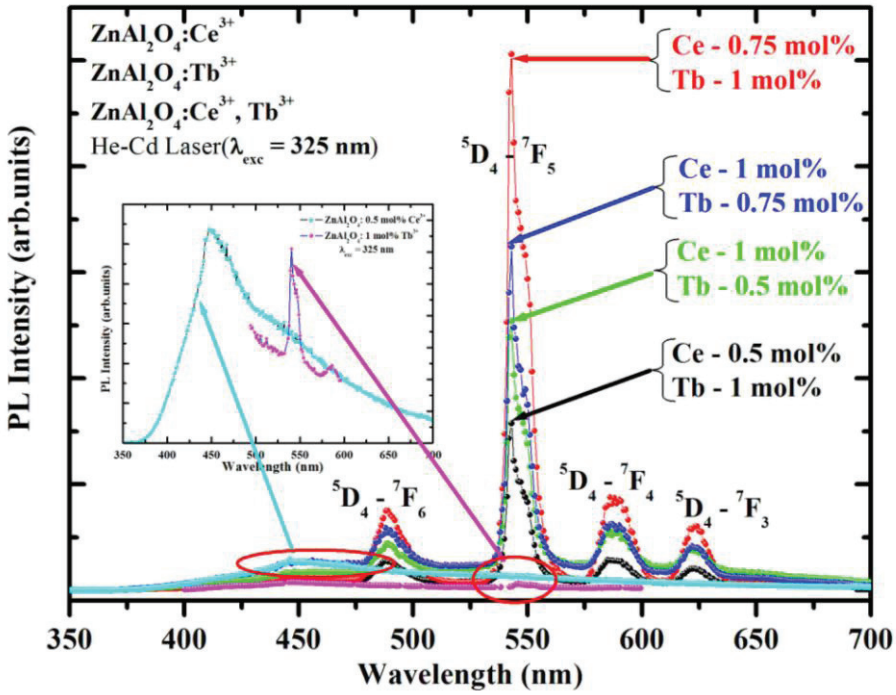


Figure 4. Emission spectra from $ZnAl_2O_4$: Ce^{3+}, Tb^{3+} with different concentrations and annealed in hydrogen atmosphere (He-Cd laser λ_{exc} = 325 nm). The inset is the emission spectra of $ZnAl_2O_4$: 0.5 mol%Ce^{3+} (cyan) and $ZnAl_2O_4$: 1 mol% Tb^{3+} (magenta) respectively.

The emission spectra of $ZnAl_2O_4$:Ce^{3+}, Tb^{3+} nanocrystal powders with different concentrations of Ce^{3+} and Tb^{3+} are presented in figure 4. The spectra were recorded when the powders were excited with a 325 nm HeCd laser in air at room temperature The inset of figure 4 shows the emission spectra of $ZnAl_2O_4$: 0.5 mol%Ce^{3+} (cyan) and $ZnAl_2O_4$: 1 mol% Tb^{3+} (magenta). The Ce^{3+} singly doped $ZnAl_2O_4$ gave blue broad band emission at 447 nm while the Tb^{3+} singly doped $ZnAl_2O_4$ gave green line emission at 543 nm as shown in the inset of figure 4. These emissions are associated with the $5d \rightarrow 4f$ and $^5D_4 \rightarrow {}^7F_5$ transitions of Ce^{3+} and Tb^{3+} respectively. The green line emission of Tb^{3+} at 543 nm was enhanced considerably and the blue Ce^{3+} emission was suppressed when different concentrations of Ce^{3+} and Tb^{3+} were incorporated simultaneously in the $ZnAl_2O_4$ host. The enhancement of the green line emission was maximized when 1 mol% of Tb^{3+} was co-activated with 0.75 mol% of Ce^{3+}. These results suggest that Ce^{3+} absorbed the UV excitation energy and transferred it non-radiatively to Tb^{3+} enhancing its green emission at 543 nm. Energy transfer from Ce^{3+} to Tb^{3+} was most probably by phonon-mediated process as previously reported [13]. The act of absorbing high energy (UV) photons and a subsequent emission of low energy (visible) photons is referred to as down-conversion. As previously reported, UV down-converting phosphors can be used as coatings to improve absorption efficiency of Si PV solar cells [14]. Similary, as a potential UV down-converting phosphor, $ZnAl_2O_4$:Ce,Tb was also evaluated for possible application as coating to improve the absorption efficiency of Si PV cells.

4. Conclusions

The green emitting $ZnAl_2O_4:Ce^{3+},Tb^{3+}$ phosphor was successfully prepared by the solution combustion method. The phosphor crystallinity indexed well to a pure face-centered cubic $ZnAl_2O_4$ with a spinel structure. No reflections attributable to other impurity phases such as ZnO, Al_3O_2 or CeO_2 are observed in XRD patterns of both the as prepared and reduced samples. The SEM data showed that the samples were made up of either distorted hexagons or platelet-like particles. The green emission was enhanced by energy transferred from Ce^{3+} to Tb^{3+} and this is mainly on the excitation cross-section of Ce^{3+} that is proportional to the f-d radiative transition of Ce^{3+} and the concentrations of both Ce^{3+} and Tb^{3+}.

Acknowledgements

This work is supported by the South African National Research Foundation (NRF), National Research Foundation of Korea, Korea Institute of Science and Technology (KIST) and Nanomaterials Cluster fund of the University of the Free State. The Centre for High Resolution Transmission Electron Microscopy at the Nelson Mandela Metropolitan University is gratefully acknowledged for the SEM results. The research supported by the South African Chairs Initiative of the Department of the Science and Technology and National Research Foundation for the maintenance of the He–Cd laser at the Nelson Mandela Metropolitan University is gratefully acknowledged for the PL results.

References

[1] Hill R J, Craig J R and Gibbs G V 1979 *Phys. Chem. Minerals.* **4** 317
[2] Sampath S K and Cordaro J F 1998 *J. Am. Ceram. Soc.* **81** 649
[3] Nabarawy T El, Attia A A and Alaya 1995 *Mater. Lett*. **24** 319-25
[4] Patil K C, M S Hegde, Rattan T and Aruna S T 2008 *Chemistry on Nanocrystalline Oxide Materials, Combustion Synthesis, Properties and Application.* (Singapore:World Scientific Publishing) p. 42
[5] Pitale S S, Kumar V, Nagpure I M, Ntwaeaborwa O M and Swart H C 2011 *Appl. Surf. Sci.* **257** 3298-3306
[6] Jain S R, Adiga K C and Pai Vernekar V R 1981 *Combustion Flame.* **40** 71
[7] Zhu Z, Li X, Zhao Q, Liu S, Hu X and Chen 2011*Mater Lett.* **65** 197
[8] Maaza D, Vallino M and Busca G 1992 *J. Am. Ceram. Soc.* **75** 1929
[9] McMillan P and Piriou B 1982 *J Non-Cryst Solids*. **53**279
[10] Da Silva A., Goncalves A and Davolos M 2009 *J. Sol-Gel Technol.* **49** 103
[11] Staszak W, Zawadzki M and Okal J 2010 *J. Alloys Comp* **492** 500 – 7
[12] Gammard A, Babaot O, Jousseaucne B, Rascle M, Toupance T and Campet G 2000 *Chem Mater* **12** 3419 - 26
[13] Ntwaeaborwa O M, Swart H C, Kroon R E, Holloway P H and Botha R J 2007 *J. Phys. Chem. Sol.* **67** 1753 -94
[14] Chung P, Chung H-H and Holloway P H 2007 *J. Vac. Sci. Technol* **A 25**(1) 61- 66

On the orbital rehybridization in tetrahedral amorphous carbon

Aniekan Magnus Ukpong

Department of Physics, University of Pretoria, South Africa

aniekan.ukpong@up.ac.za

Abstract. The pi-orbital axis vector theory is used in this study to investigate the phenomenon of rehybridization of the carbon bonding in tetrahedral amorphous carbon. It is shown that the correlated pi - pi* orbitals is locally entangled due to the competition between the on-site Coulomb interaction and the hybridization. A microscopic model is proposed for the reduction in strain in the material due to hydrogenation, based on the $sp^3 \rightarrow sp^2$ conversion reaction. It is demonstrated that the activation energy for this conversion is dependent on the Urbach energy, and therefore on the structural disorder. The implication of the results on the manifestation of strain in similar, carbon-based, materials is discussed.

1. Introduction

Solid-state quantum information processing is only achievable if there is control of the properties of the functional material. One basic resource for controlling the electronic degrees of freedom in such systems is the quantum entanglement between electronic states [1-3]. Entanglement is a measure of the correlation in quantum mechanical systems. In the entangled state, the best possible knowledge of the whole system does not necessarily imply the best possible knowledge of its constituents [4]. Over the nearly eight decades since its introduction, the concept of quantum entanglement has significantly changed in meaning, and its nature and interpretation may still continue to change [5]. In this study, its interpretation in terms of the anisotropic coupling between atomic orbitals allows for the characterization and tunability of the physical properties of materials. This is directly applicable to the selection of suitable absorptive materials for bolometer applications. For instance, although amorphous silicon (a-Si) [6] or vanadium dioxide (VO_2) [7] films are the commonly used photo-responsive materials in bolometers operated within the infra-red/ultra-violet (IV/UV) range, other materials can equally be used. In particular, the use of tetrahedral amorphous carbon (ta-C) films in this role cannot be overruled due to similarities between the local network structures in carbon and silicon. However, the suitability of ta-C for such application depends strongly on the ability to tune its optical properties. This can be achieved by controlling the density of dangling bond defects and the structural disorder in the material. Amorphous carbon therefore constitutes an attractive subject of fundamental investigation since the hybridization state of the carbon atoms is expected to play an important role in determining the resulting physical properties [8]. In our recent molecular dynamics simulations of typical ta-C networks for instance, the network stress was found to relax in the presence of hydrogen although the exact mechanism for the relaxation remained unclear [9]. This paper clarifies the mechanism of stress relaxation in ta-C networks by investigating the role of hydrogen on stereo-electronic factors and the resulting physical properties of the material.

2. Theory and computations

First principles molecular dynamics calculations were performed to implement the melt-quenching procedure of simulating tetrahedral amorphous carbon (ta-C) from diamond-like carbon. The calculations were performed on a 64-atom supercell using the VASP electronic structure code [10-13]. This allows for self-consistent solutions to the Kohn-Sham equations using the projector augmented wave (PAW) method [14,15], and a plane wave basis set. Kinetic energy cutoff of 500 eV was used for expanding the plane wave basis set of the electron wave functions. Atomic positions were relaxed on the fly using forces derived from Hellman-Feynman theorem. The convergence limit of 10^{-3} eV was set for the determination of electronic energies. The structures were relaxed until the forces reduced to 0.03 eV/Å in each case. Hydrogen atoms were selectively incorporated to the optimized ta-C structure to produce ta-C:H structures at saturation levels corresponding to 15, 30, 45 and 60 at % H. These levels of hydrogen concentration were deliberately chosen to maintain consistency with the bulk structures generated in previous tight-binding molecular dynamics (TBMD) simulations using a 512-atom supercell [9]. The electronic structure was calculated within the local density approximation (LDA) to density functional theory (DFT), and the optoelectronic properties were investigated in terms of stimulated optical inter-band transitions. Level 1 [16] of the pi-orbital axis vector (POAV) theory was used within the σ-π orbital separability approximation to calculate the effect of hydrogen on π-orbital curvature [17-20], and the rehybridization sp^{2+n}, where n characterizes the degree of the rehybridization [21,22].

3. Results

The effect of hydrogen on the stereo-electronic properties of tetrahedral amorphous carbon (ta-C:H) is presented in Table 1. The error bars quoted in Table 1 were determined as the standard error in the mean of ten repeated calculations of the same data. The strain in the hydrogenated networks is normalised to the strain in the hydrogen-free structure. The presence of a substantial fraction of sp^2-hybridzed bonds, even at 0% H, implies that a fraction of π-electrons will always be present in structure due to formation of localised domains of aromatic or aliphatic carbon chains. Table 1 shows that the ta-C:H networks which contain low hydrogen concentrations are relatively highly strained. From the hydrogen/sp^2/sp^3 ternary phase diagram [8], such networks contain olefinic chains (ta-C:H) in which the sp^2-hybridised carbon sites tend to cluster. On the other hand, ta-C:H networks that contain high H concentrations exhibit significantly low strains and tend to contain aromatic carbon chains. Hence, the physical properties of amorphous carbon thin films are determined by the level of H-concentration in the networks. The average interaction between two π-orbitals in these structures was determined using resonance theory [23] for the 64-atom supercell. The dependence of the strength of the inter-orbital interaction $V_{pp\pi}(1)$ with changing levels of H-saturation is also shown in Table 1.

Table 1. Effect of hydrogen concentration on orbital mixing factor, degree of rehybridization, π-orbital curvature, inter-orbital strength and σ-bond strain in ta-C:H.

H content (%)	Orbital mixing factor, k	Rehybridization, n	π-orbital curvature (θ^2)	π-π interaction (eV)	σ-bond strain
0	0.61 ± 0.01	0.62 ± 0.01	0.69 ± 0.10	-1.55 ± 0.02	1.00 ± 0.02
15	0.40 ± 0.04	0.54 ± 0.03	0.84 ± 0.08	-1.83 ± 0.03	0.79 ± 0.05
30	0.33 ± 0.05	0.43 ± 0.02	1.06 ± 0.12	-2.03 ± 0.05	0.52 ± 0.01
45	0.27 ± 0.02	0.32 ± 0.01	1.17 ± 0.10	-2.57 ± 0.01	0.32 ± 0.03
60	0.20 ± 0.03	0.29 ± 0.02	1.22 ± 0.05	-3.06 ± 0.12	0.18 ± 0.02

Figure 1(a) and (b) shows the orbital angular momentum resolved density of states (DOS). The Fermi level is aligned to correspond to 0 eV. Fig. 1(a) shows that *s*-state DOS is dominated by valence electron states. The number of conduction electron states in the interval 0 - 5.0 eV is reduced when the structure is saturated with 15% hydrogen, and reduces even further at higher concentrations of

hydrogen. Figure 1(b) shows the π-state DOS at different hydrogen saturation levels. The 1*s* state of hydrogen induces the perturbation of the valence band states in this case, even though it is not a core electron state. The DOS profiles show significant perturbations as the level of H saturation changes. The perturbation of the valence band results in the promotion of valence electron states into the conduction band as more hydrogen is incorporated. This effect is mostly seen at valence band energies above -15 eV. Comparisons of the conduction band states between 0 and 5 eV for structures containing 0 and 60 at. % H show that the number of electron states is neither increased nor decreased even after saturation with 60% H. The number of valence band states within this energy interval is depleted.

The π-electron DOS (see Fig. 1(b)) predicts metallic ta-C:H systems while the *s*-electron DOS (see Fig. 1(a)) gives an opposing effect at all levels of hydrogen saturation. The relative reduction in the density of bonding states (denoted as π-orbitals) at different levels of H-saturation, and the corresponding changes in the density of anti-bonding states (denoted as π*-orbitals) suggests that an anisotropic coupling exists between the π-orbitals. In terms of the quantum nonseparability paradigm [24, 25], the anisotropic coupling indicates the local entanglement of the π-orbitals. The bonding (π) and anti-bonding (π*) states are simultaneously influenced by H-saturation irrespective of the concentration level. Figure 2(a) shows the total density of states for ta-C structures, which contain 0, 30 and 60 at. % H. The dynamical transfer of spectral weights from *s*- and *p*-electron states to the total density of states is therefore crucial in maintaining the semiconducting state expected in ta-C:H. Hence, the dynamics of the entangled system was investigated further by treating the orbital mixing parameter *k* as an anisotropic coupling constant to measure the evolution of orbital curvature and the degree of rehybridization as a function of H-saturation.

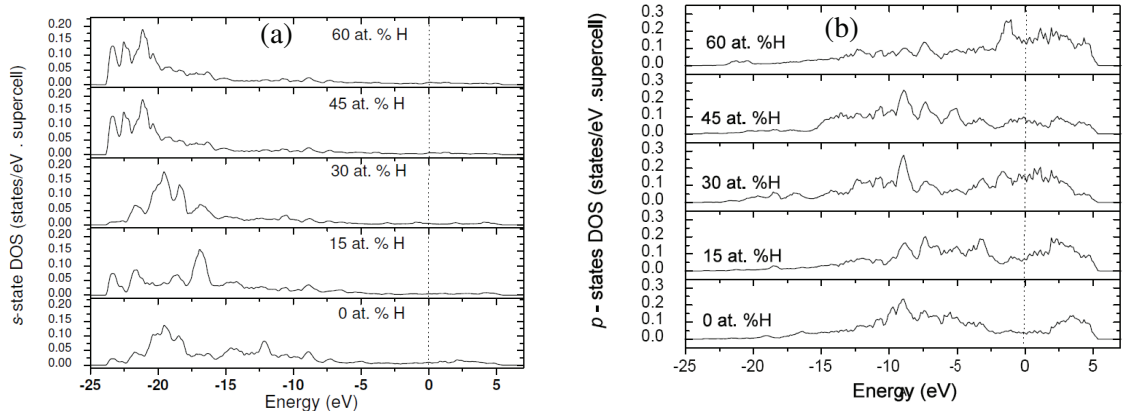


Figure 1. Contributions of *s*-electron states (a), and π-electron states (b), to the total density of states.

Optical inter-band transitions may occur due to excitations from extended to localized states and vice versa. In Figure 2(a), the conduction and valence band tails decay exponentially into the band gap. These Urbach-like band tails [26] suggest that optical transitions in ta-C:H can be described with a relationship of the form $\alpha(E) \sim \exp(\frac{E}{U_E})$ between absorption coefficient α and the photon energy E, where U_E is the Urbach energy [27]. The Urbach energy relates the structural disorder to photoconductive properties of a semiconductor. Figure 2(b) shows the dependence of U_E on the sp^3-site faction in ta-C:H networks. The Urbach energy increases with increasing sp^3 site fraction and reaches a maximum when the sp^3 fraction is ~62 %. Further increase in sp^3 site fraction beyond this point results in a steady decrease in U_E. The dependence of U_E on sp^3-content shows an approximate linear relationship up to 50%. The sp^3 site fraction in ta-C:H is known to decrease with increased level of H-saturation and results in increased disorder in second neighbour positions [9]. This implies that the sp^2/sp^3 ratio scales with the U_E. The sudden increase in U_E at 60% and the resulting deviation from linear trend line dependence beyond 50% suggests a phase transition towards the more ordered structural networks expected at low hydrogen concentrations. The present LDA calculations show that U_E saturates at a

slightly lower sp^3-site fraction compared to the 65% predicted in empirical tight binding calculations of Mathioudakis *et al*. [28]. Nevertheless, the agreement between these two fundamentally dissimilar calculations - calculations performed within the LDA approximation to DFT, and those based on environment-dependent tight-binding (EDTB) method - shows that the physical and chemical properties of carbon-based materials are accurately described within phenomenological potentials.

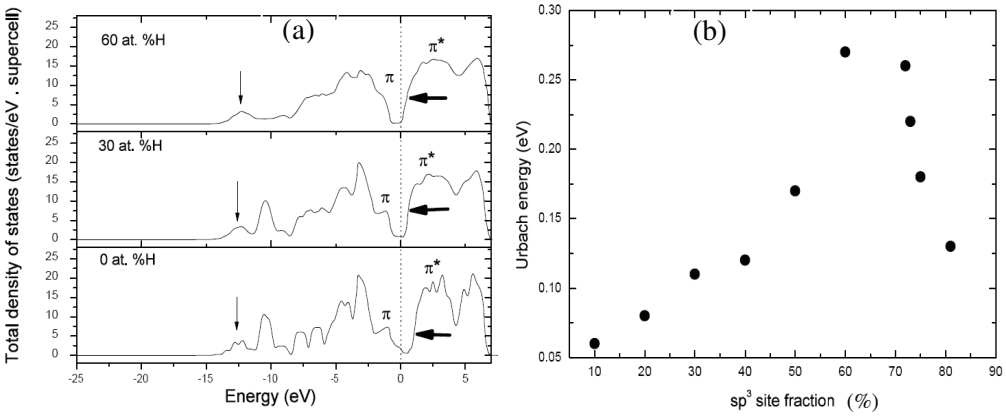


Figure 2. Total DOS in ta-C:H showing the effect of hydrogen saturation level on bonding and anti-bonding states (a), and the dependence of Urbach energy on sp^3 site fraction (b). The Urbach energy is the width of the exponential band tails due to states localized within the band gap of a semiconductor.

Results of isothermal annealing experiments do not show any changes in the local bonding structure of *ta*-C:H materials below 600 °C [29]. This indicates that the $sp^3 \rightarrow sp^2$ bond conversion reaction does not pass through a transitional (or intermediate) reaction phase once it is activated. It is therefore assumed that the observed changes in local bonding structure in Fig. 2(b) can be modelled using the first-order reaction kinetics. Numerical modelling of the hydrogen dynamics as a random walk to simulate the structural evolution with temperature shows activated hydrogen dynamics at temperatures between 570 K and 1180 K. This observation agrees with the experimentally observed deuterium diffusion in ta-C and a-C films [30]. The diffusivity of hydrogen is 0.018 (± 0.002) cm^2/s at 600 K. From this simulation, the activation energy for the hydrogen diffusion was derived based on the assumption of Arrhenius-like diffusion. The activation energy was obtained by plotting the hydrogen diffusion coefficient against the simulated annealing temperature as shown in Figure 3. An interpolation of the temperature and H-content that yields the maximum sp^3 site-fraction of 62% and Urbach energy U_E of Fig. 2(b) gives the activation energy as $\Delta H = 1.14$ meV for the bond conversion reaction.

4. Discussion

The π-orbital curvature is generally small in ta-C:H materials but its effect on orbital rehybridization is nontrivial. The curvature varies within the range $0.69\pm0.10 < \theta < 1.22\pm0.05$ under the influence of H-saturation. It is found that the degree of rehybridization decreases rapidly towards a limiting value of 0.3 as the levels of H concentration is increased. The saturation of the rehybridization n at a small, non-zero, value at 60 % H content indicates that full graphitic sp^2-bonding is unachievable by hydrogenation alone. In addition, the magnitude of the π-orbital curvature saturates at a limiting value of 1.27 at high H-saturation limit. However, irrespective of relative orientations of π-orbitals, the local geometries of C-clusters only achieve full closure when the π-orbital curvature is 2.42. Such local geometries are only found in Buckminster fullerenes such as C_{60}, carbon nanotubes, and cyclic hydrocarbons where the conservation of curvature is well known [11]. Schmalz *et al*. [14] showed that in such curved C-structures, the strain on the σ-bonds scales linearly with orbital curvature and the π-orbital rehybridization n. The strain in ta-C:H networks exhibits a similar behaviour, and therefore

considered to arise from deviations from planar graphitic geometry due to misorientation of some π-orbitals, and due to the σ-orbitals being oriented in the tangential direction to the planar surface. The imperfect alignment of pairs of bond-forming orbitals also leads to the rehybridization of the π-orbitals. The rehybridization is necessary to bring the two nearby orbitals into closer alignment. The above analysis, of the role of hydrogen in increasing the π-orbital curvature in ta-C:H, is based on the assumption that planar graphite layer ($\theta = 0$) is strain-free. The evolution of bond strain in ta-C:H networks has therefore been gauged relative the graphite surface.

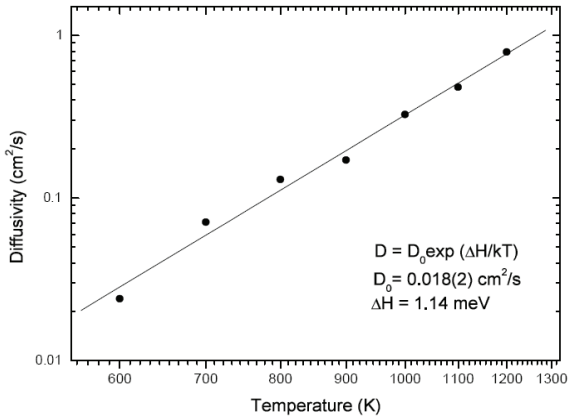


Figure 3. Arrhenius plot of hydrogen diffusion coefficient as a function of temperature on a logaritmic scale. The straight line indicates the best-fit trendline between ln*D* and the inverse temperature 1/*kT*.

The 1*s* state of hydrogen is found to induce the entanglement of π-electron states by introducing the nonlocal correlation between the π-bonding and anti-bonding state. By locally switching on the inter-orbital interactions $V_{pp\pi}(1)$, we have explicitly included the effects of the on-site Coulomb interactions between electrons on two interacting π-orbitals. The inter-orbital interaction must therefore be in competition with the strength of the local hybridization of the local carbon bonding in order to maintain the semiconducting state in ta-C:H. Due to the local π-orbital entanglement, further saturation with hydrogen does not lead to appreciable increases in the number of conduction band states. The correlated nature of the bonding (π) and anti-bonding (π*) electron states arises due to the local competition between inter-orbital interactions and sp^{2+n} hybridization. The effect of the 1*s* state of hydrogen on carbon bonding and anti-bonding orbital is non-negligible. The total density of states shows that π* (anti-bonding) electron states in the bottom of the conduction band are effectively shifted towards lower energies as the concentration of hydrogen increases (see Fig. 2(a)). Although the defect-induced mid-gap levels are substantially eliminated upon saturation with hydrogen, the downward shift in energy levels leads to substantially reduced energy gaps, as H-concentration increases, giving rise to exponential conduction and valence band tails.

The changes in π-orbital curvature and the rehybridization and corresponding reduction in bond strain, shown in Table 1, suggest that the experimentally observed bond conversion reaction above 600 °C [29] is consistent with the systematic increase in orbital coupling constant. The total H-content in typical thin films depends on the deposition temperature and the amount of bonded-hydrogen reduces when the annealing temperature is high enough to induce spontaneous hydrogen diffusion. The H-evolution can be determined experimentally by characterizing the post-annealing hydrogen content in a thin film after annealing using materials characterization techniques such as nuclear reaction profiling or IR spectroscopy. It therefore follows as corollary that the observed bond conversion is denoted by the continuous decrease (and increase) in sp^3 (and sp^2) site-fraction as the hydrogen saturation level increases. Since the rehybridization necessary to reduce strain also depends on the concentration of hydrogen, it is argued that the activation energy of the $sp^3 \rightarrow sp^2$ bond conversion must also depend on the residual network disorder at the onset of the reaction, and therefore on the Urbach energy.

5. Conclusion

The π-orbital axis vector theory has been applied to the study of the changes in ta-C:H networks under variable H-saturation levels. The mechanism of H-mediated curvature-induced rehybridization of π-orbitals is found to lead to the relaxation of strain in the material. From the analysis of the orbital-resolved density of states, it is found that the bonding (π) and anti-bonding (π*) electronic states are locally entangled. The correlated nature of the π-orbitals is interpreted in terms of the anisotropic coupling of atomic orbitals, and shown to occur due to competition between the on-site, inter-orbital, Coulomb-like interaction and the orbital hybridization of the carbon atoms. Using simulated inter-band optical transitions, the total density of states and the orbital rehybridization, it is shown that the activation energy for the $sp^3 \rightarrow sp^2$ bond conversion depends on the Urbach energy and therefore on the network disorder.

Acknowledgments

The author acknowledges financial support from the University of Pretoria under E2020 Project 5.

References

[1] Huang Z and Kais S 2005 *Chem. Phys. Lett*. **413** 1
[2] Samuelsson P and Verdozzi C 2007 *Phys. Rev. B* **75** 132405
[3] Craco L 2008 *Phys. Rev. B* **77** 125122
[4] Schrödinger E 1935 *Naturwissenschaften* **23** 807
[5] Bruß D 2002 *J. Math. Phys*. **43** 4237
[6] Liu X M, Fang H J and Liu L T 2007 *Microelec. J.* **38** 735
[7] Rajendra Kumar R T, Karunagaran B, Mangalaraj D, Narayandass S K, Manoravi P, Joseph M, Gopal V, Madaria R K and Sing J P 2003 *Mat. Res. Bull* **38** 1235
[8] Robertson J 2002 *Mat. Sci. and Eng*. R **37** 129
[9] Ukpong A M 2010 *Mol. Phys.* **108** 1607
[10] Kresse G and Hafner J 1993 *Phys. Rev*. B **47** 558
[11] Kresse G and Hafner J 1994 *Phys. Rev*. B **49** 14251
[12] Kresse G and Furthmuller J 1996 *Phys. Rev.* B **54** 11169
[13] Kresse G and Furthmuller J 1996 *Comput. Mater. Sci.* **6** 11169
[14] Blochl P E 1994 *Phys. Rev.* B **50** 17953
[15] Kresse G and Joubert J 1999 *Phys. Rev*. B **59** 1758
[16] Haddon R C 1987 *J. Phys. Chem.* **91** 3719
[17] Haddon R C 1988 *Acc. Chem. Res.* **21** 243
[18] Haddon R C 1993 *Science* **261** 1545
[19] Haddon R C 1986 *Chem. Phys. Lett.* **125** 231
[20] Schmalz T G, Seitz W A, Klein D J and Hite G E 1988 *J. Am. Chem. Soc*. **110** 1113
[21] Kleiner A and Eggert S 2001 *Phys. Rev.* B **64** 113402
[22] Yu O, Jing-Cui P, Hui W and Zhi-Hua P 2008 *Chinese Physics* B **17** 1674
[23] Harrison W A 1989 *Electronic structure and the properties of solids: the physics of the chemical bond* (New York: Dover)
[24] Miller W H 1976 *Acc. Chem. Res*. **9** 306
[25] Healey R 2009 *Holism and nonseparability in physics*, ed E N Zalta (Stanford: Encyclopedia of Philosophy)
[26] Pan Y, Inam F, Zhang M and Drabold D A 2008 *Phys. Rev. Lett.* **100** 206403
[27] Fanchini G and Tagliaferro A 2004 *Diam. Relat. Mater.* **13** 1402
[28] Mathioudakis C, Kopidakis G, Patsalas P and Kelires P C 2007 *Diam. Relat. Mater.* **16** 1788
[29] Grierson D S, Sumant A V, Konicek A R, Friedmann T A, Sullivan J P and Carpick R W 2010 *J. Appl. Phys.* **107** 033523
[30] Kröger H, Ronning C, Hofsäss H, Neumaier P, Bergmaier A, Görgens L and Dollinger G 2003 *Diam. Relat. Mater.* **12** 2042

Synthesis and Characterization of Structural and Luminescent properties of long afterglow $CaAl_2O_4$: Eu^{2+}, Nd^{3+}, Dy^{3+} phosphors by solution – combustion technique.

[1]A.H Wako*, [1]F.B Dejene*, and [2]H.C. Swart

[1]Department of Physics , University of the Free State, QwaQwa Campus, Private Bag X13, Phuthaditjhaba 9866, South Africa.

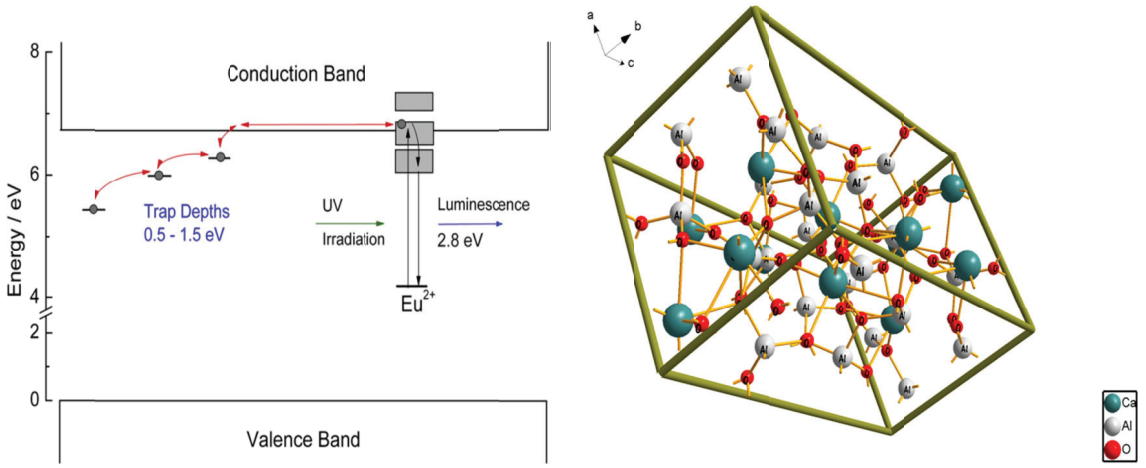
[2]Physics Department, University of the Free State, P. O. Box 339, Bloemfontein, 9300, South Africa

Email: wakoa@qwa.ufs.ac.za or dejenebf@qwa.ufs.ac.za

Abstract. Long afterglow calcium aluminate phosphors doped with Eu^{2+} and co-doped with Nd^{3+} and/or Dy^{3+} ($CaAl_2O_4$:Eu^{2+},Nd^{3+},Dy^{3+}) were fabricated by urea-nitrate solution – combustion method at 500 °C. The effects of varying concentrations of Ca: Al and co-dopants molar ratio (Nd^{3+}:Dy^{3+}) on the structural and luminescent properties of the phosphor were investigated. It was observed that Ca : Al mass ratios greatly affect the crystalline structure of the material. The results of the X-ray diffraction (XRD) analysis reveal that the formation of several crystalline phases depends on the ratios of the host material. The XRD peaks show the presence of other phases such as $Ca_3Al_2O_6$ and $CaAl_4O_7$ but the predominant phase formed was that of $CaAl_2O_4$. However it was found that the crystalline structure is generally not affected by the variation of the co-dopants concentration. PL studies revealed a general rise in intensity with an increase in the mole ratio of Ca : Al . The highest PL intensity was observed with 0.7% Ca. The luminescent intensities vary from each other when co-doped with various proportions of Nd^{3+} and Dy^{3+}.

1. Introduction

Aluminates doped with rare earth ions have advantages of being highly stable, bright and versatile in industrial processes that are appropriate for lighting and display devices [1, 2]. Unlike other phosphors so far reported in the literature [3-5], in this study rare earth metal ions, Europium (Eu^{2+}), Dysprosium (Dy^{3+}) and Neodymium (Nd^{3+}) and the host $CaAl_2O_4$ are added together to improve the luminescent properties of the phosphor. The solution-combustion synthesis method also has the merits of low temperature, low cost and time saving. Earlier studies explained the mechanism of persistent luminescence as being caused by the interplay of the rare earth ions with the host band structure and the lattice defects (Fig. 1(a)). The ability of the rare earth species to trap electrons/holes can be predicted from the positions of the 4f and 5d levels of the Eu^{2+} ion and the other R^{3+} ions in the host lattice structure [6], but the energy storage and luminescent mechanisms cannot be explained from the level locations alone [1]. The radiative transitions are highly affected by the crystal field components because the 5d orbital is exposed to the surrounding ion. Thus the structure of the host crystals strongly influences the wavelength of a maximum emission peak. This explains why for instance, Eu^{2+} -doped Ca and Sr aluminates show blue and green emissions, respectively [7]. $CaAl_2O_4$ has a tridymite type monoclinic structure (space group: P21/n (No. 14), Z: 12) [8], where [AlO_4] tetrahedral construct a three-dimensional framework. Each aluminium ion is bonded to two oxygen ions [9]. The structure consists of channels built up of rings formed by six corner- sharing AlO_4 tetrahedra and Ca^{2+} cations situated within the channels (Fig. 1(b)).



(a) (b)

Figure 1: (a). Persistent luminescence mechanism for $CaAl_2O_4:Eu^{2+}$, R^{3+} materials. (b) Structure of monoclinic $CaAl_2O_4$.

There are three different Ca^{2+} sites in the $CaAl_2O_4$ lattice: one (Ca_3) is nine-coordinated and two of them (Ca_1 and Ca_2) are six-coordinated [10]. Doped Eu^{2+} ions prefer nine-coordinated Ca^{2+} sites (rCa: 1.18 Å) to two six-coordinated ones (rCa: 1.0 Å), because larger spaces are necessary for the substitution of Eu^{2+} ions (rEu: 1.30 Å) due to the difference in ionic size. The average Ca–O distances 2.42 Å and 2.43 Å for the six-coordinated Ca^{2+} ions are clearly shorter compared to that of the nine-coordinated ones (2.78 Å) [11]. By doping with Eu^{2+} ions, $CaAl_2O_4$ powders exhibit a strong blue emission around 440 nm when excited with 320 nm [12]. $CaO–Al_2O_3$ systems easily degenerate into multi-phased compounds with slight deviation from the exact stoichiometric composition and firing temperature which makes it difficult to obtain $CaAl_2O_4$ single phase. So the synthesis conditions, such as host elements, dopant concentration and temperature are very important to the formation of single phase. However, in spite of many previous studies, a thorough investigation has so far not been carried out. This study focuses on the blue emitting persistent luminescence, $CaAl_2O_4:Eu^{2+},Nd^{3+},Dy^{3+}$ phosphor of which $CaAl_2O_4:Eu^{2+},Nd^{3+}$ is already in commercial use.

2. Experimental

2.1 Synthesis

$CaAl_2O_4:Eu^{2+},Nd^{3+},Dy^{3+}$ phosphors were synthesized using the solution - combustion method. The starting raw materials used in the experiment include various proportions of analytical pure grade $Ca(NO_3)_2.4H_2O$, $Al(NO_3)_3.9H_2O$, $Eu(NO_3)_3.5H_2O$, $Nd(NO_3)_3$, $Dy(NO_3)_3$, urea (CH_4N_2O) and boric acid (H_3BO_3). The raw materials were weighed according to the chemical composition of $CaAl_2O_4$: Eu^{2+}, Nd^{3+},Dy^{3+}, dissolved in 10 ml of de-ionized water and thoroughly mixed using a magnetic stirrer for 15 minutes without heating to obtain a uniform solution. In the first group, eight samples were prepared without flux. The Al concentration was kept constant at 1.5 mass % for all the prepared solutions. Varying concentrations; 0.1, 0.4, 0.5, 0.6, 0.7, 0.8, 1.0, and 1.5 mass % of Ca were taken. The second set consist of five samples used to study the influence of variation of co-dopants, Nd^{3+} and Dy^{3+} molar ratio on the structural and luminescent properties of the $CaAl_2O_4:Eu^{2+}$ phosphor. The samples were mixed in the mass ratios 0:1, 0.25:0.75, 0.5: 0.5, 0.75:0.25 and 1:0 of Nd^{3+}: Dy^{3+}. The solutions were then poured into China crucibles and placed one at a time in a muffle furnace pre-heated at 500^0C. Combustion time was 5–6 min per sample. White voluminous foam was obtained by

combusting the mixture at temperatures of 400–500°C. Initially, the solution boiled and underwent dehydration, followed by decomposition releasing large amounts of gases (oxides of carbon, nitrogen and ammonia). Then, spontaneous ignition and smoldering occurred which gradually led to an explosion with enormous swelling. When taken out of the muffle furnace and cooled down, voluminous foam was obtained, which was a mixture of the incomplete combustion compound $CaAl_4O_7$ and the complete combustion compound $CaAl_2O_4$. The voluminous foam was milled to obtain the fine, white powders. The powders were stored in transparent sample glass bottles for characterization.

2.2 Characterization

The synthesized products were characterized by X-ray diffraction (XRD) using a Bruker D8 X-ray diffractometer operating at 40 kV and 4 mA using Cu Kα = 0.15406 nm for elemental analysis. The morphologies were investigated using a Shimadzu model ZU SSX-550 Super scan Scanning Electron Microscope (SEM) and an Energy Dispersive X-ray Spectrometer (EDS). The decay curves, excitation and emission spectra were measured at room temperature using a Cary Eclipse fluorescence spectrophotometer model: LS- 55 with a built-in 150W xenon flash lamp operating at excitation of 325 nm.

3. Results and discussion

3.1 The influence of the Ca: Al mass ratio on structure.

Scanning electron microscopy (SEM) study was carried out to investigate the surface morphology and crystalline sizes of the synthesized phosphor powder. Figure 2 shows the representative SEM micrographs taken for the 0.4% Ca sample and 0.8% Ca sample respectively. As can be seen, the surfaces display many cracks and pores formed by the escaping gases during the combustion reaction. The material also presented agglomeration as can be observed in Fig.2 (b). This tends to reduce the luminescent properties when the sample is ground.

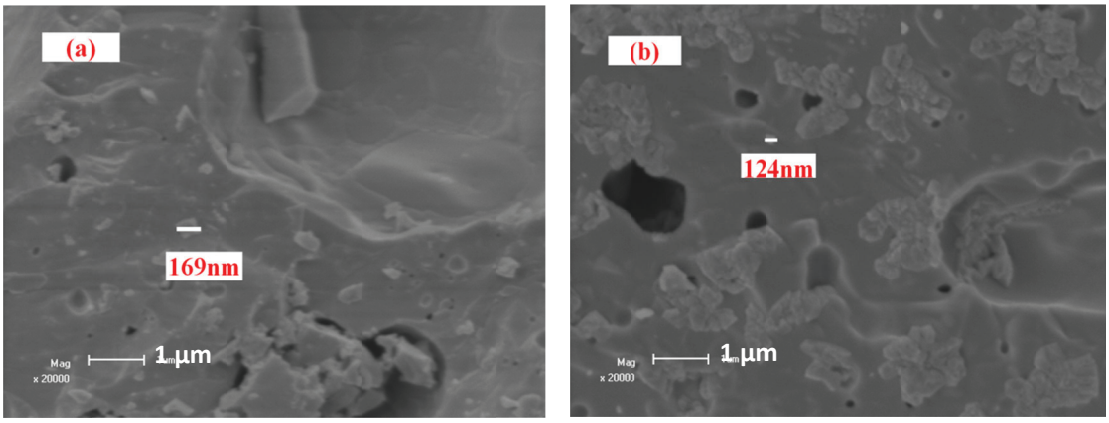


Figure 2: SEM micrographs for (a) 0.4% Ca sample and (b) 0.8% Ca sample respectively.

Figure 3 shows the powder XRD pattern for the $CaAl_2O_4:Eu^{2+}$, Nd^{3+}, Dy^{3+}. Several phases are easily observable. Analysis of these diffractive peaks reveal that 0.5% Ca and 0.6% Ca favour pure monoclinic phase matching with the JCPDS data file (no. 23-1036) while higher concentrations of calcium; 0.8% Ca, 1.0% Ca and 1.5% Ca seem to favour the formation of cubic phase. Weak diffraction peaks of $CaAl_4O_7$ compound formed due to incomplete combustion were also observed. The calculated lattice parameters for the monoclinic crystal system were a=8.6942Å, b=8.0930Å, c=15.2097Å and β=90.17° while for the cubic system was a=11.9868Å.

3.2 The effects of Ca : Al mass ratio on photoluminescence properties.

Figure 4 (a) shows the excitation spectrum of the phosphor. There is an excitation peak at around 325 nm in the emission spectrum, λ_{em}= 440 nm which is in conformity with the absorption of the host $CaAl_2O_4$. It's known that $CaAl_2O_4$ has a broad band excitation spectrum ranging from 250 to 340 nm. This means that in a broad range ($\lambda_{ex} \leq$360 nm), $CaAl_2O_4:Eu^{2+}$, Dy^{3+}, Nd^{3+} phosphor can be activated to produce long persistence. As can be seen in Fig. 4(b), the emission spectrum at λ_{ex} = 325 nm shows a broad band from the $4f^65d^1$ to the $4f^7$ configuration of the Eu^{2+} ions at about 440 nm, which is consistent with the literature.[25]. Energy levels are located between excited state $4f^65d^1$ and ground state $4f^7$ of Eu^{2+}, and there is the energy gap (ET) between the energy levels and the excited state energy level. When the material is excited, some of the Eu^{2+} excitation electrons get stored in the trap energy level through the relaxation process. When the excitation stops, the electrons stored up in the trap energy level obtain energy through thermal agitation at room temperature, overcome the energy level gap between the trap energy level and the excited state energy level, return to the excited state, then transit back to the ground state and emit light energy.

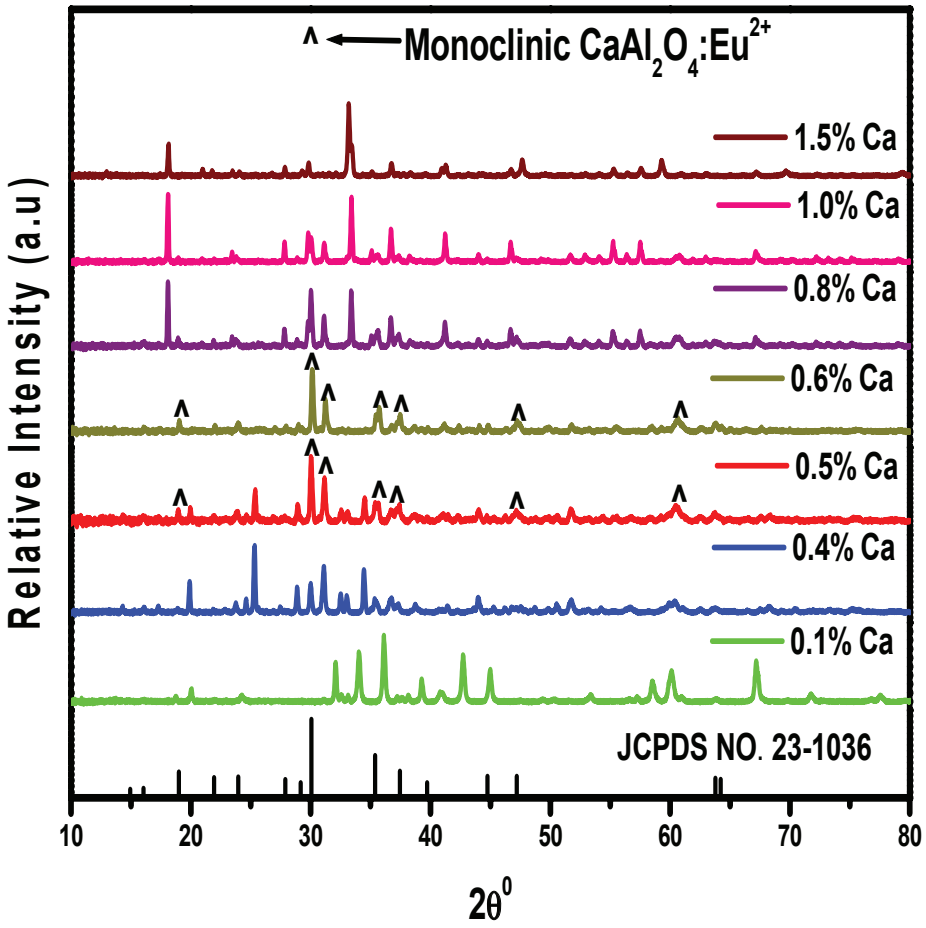


Figure 3: Monoclinic phase of $CaAl_2O_4:Eu^{2+}$,Nd^{3+},Dy^{3+} phosphors.

From the decay spectrum of figure 4(c),It can be seen that the samples show quite long time especially when the powder was efficiently activated by using 0.7% Ca. In general all the PL results show 0.7% Ca as the optimal concentration for the excitation, emission and decay characteristics. Additional emission peaks observed between 525 nm and 675 nm are attributed to the unreduced Eu^{3+} ions in the host lattice.

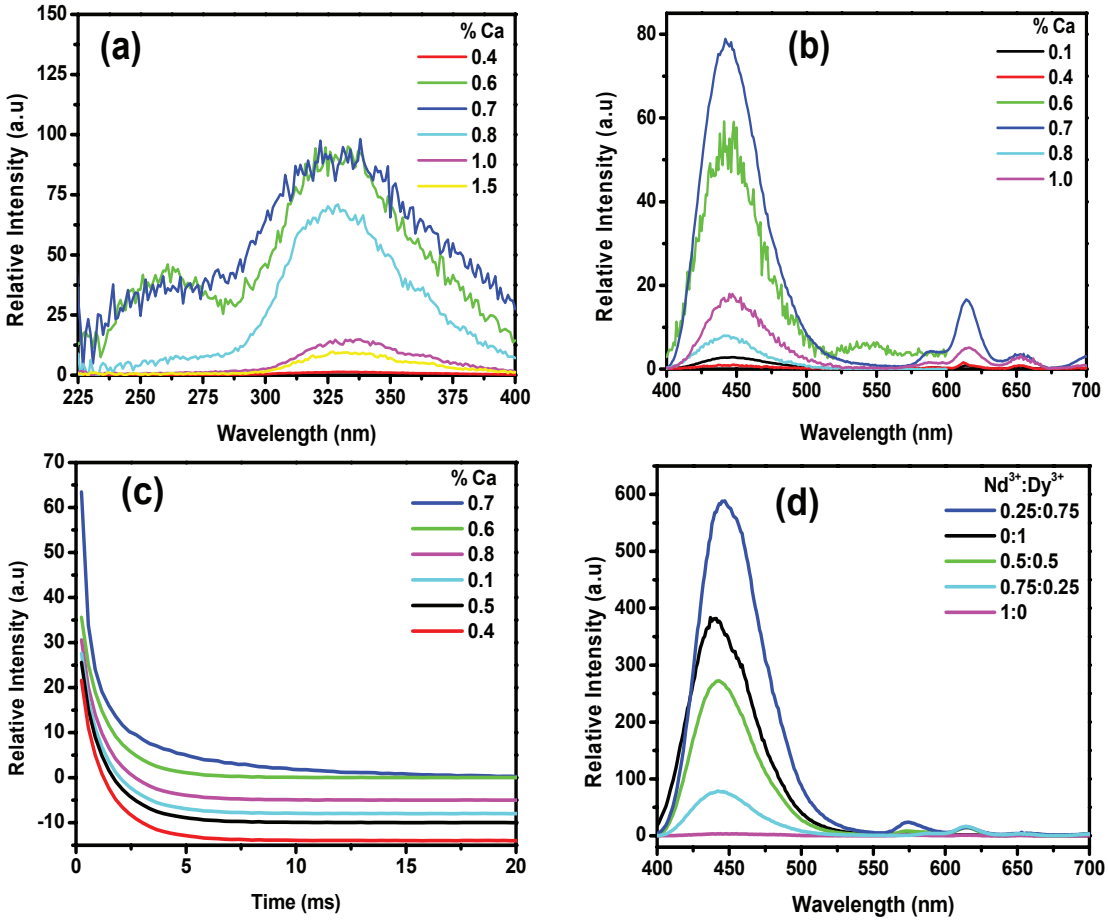


Figure.4: (a) Excitation, (b) emission (c) decay characteristics of the phosphors dependence on Ca: Al mass ratio and (d) emission intensities of the varying Nd^{3+} : Dy^{3+}co-dopant ratio obtained at room temperature.

3.3 The effects of co-dopants mass ratio (Nd^{3+}:Dy^{3+}) on the photoluminescence properties.

Figure 4(d) shows the emission intensities of the $CaAl_2O_4$:Eu^{2+},Nd^{3+},Dy^{3+} phosphors while the phosphors are excited at a wavelength of 325 nm. Within the scope of content ratio 0.25:0.75, Nd^{3+} and Dy^{3+} ions seem to improve the luminescent properties of the phosphors $CaAl_2O_4$:Eu^{2+}, Nd^{3+}, Dy^{3+}. The peak wavelength of the phosphor does not vary with the doped Nd^{3+}:Dy^{3+}. This implies that the crystal field, which affects the 5d electron states of Eu^{2+}, is not significantly changed by the variations in the co-dopants. When the phosphor is doped with Nd^{3+} and Dy^{3+}, Ca^{2+} ions are replaced by Nd^{3+} and Dy^{3+} ions. But Ca^{2+} ions and Nd^{3+}/Dy^{3+} ions are different resulting in unequal replacement. This

causes more trap energy level and the depth of the energy level becomes deeper. Otherwise, Nd^{3+} seems to possess the correct affinity toward to the electrons more than Dy^{3+} in the energy level.
When the concentration of Eu^{2+} ion is fixed, the number of electron-hole pairs is fixed. As the concentration of Nd^{3+} ion increases in a range, the number of the Nd^{3+} traps increases, and the traps capture more holes. However, if the doped amount of Nd^{3+} is too large, it may result in concentration quenching and may lower the luminescent effect, as can be observed from Fig.4(d), when the value of Nd^{3+}: Dy^{3+} was 1:0, the emission intensity was relatively low for this phosphor.

4. Conclusion

The phosphor $CaAl_2O_4:Eu^{2+}$, Nd^{3+}, Dy^{3+} can be prepared by solution-combustion method. The influences of the quantity of mixed Nd^{3+}, Dy^{3+} and varying concentrations of Ca: Al on the phosphor were studied. The analytical results indicate that the broad emitted band of the $CaAl_2O_4:Eu^{2+}$, Nd^{3+}, Dy^{3+} is observed in the blue region (λ_{max} = 440 nm) due to transitions from the $4f^65d^1$ to the $4f^7$ configuration of the Eu^{2+} ion. It may serve as a promising material for use as a lamp phosphor in the blue region. Furthermore, the solution- combustion method is cost-effective, saves energy and time.

Acknowledgements

The authors wish to convey their sincere gratitude to the National Research Foundation (NRF) for funding the project and the University of Free State Physics department for the research techniques used in this study.

References

[1] Justel T, Krupa JC and Wiechert DU 2001 *J. Lumin*. **93** 179
[2] Zhou LY, Huang J L, Yi LH, Gong ML and Shi J X 2009 *J. Rare Earths* **27** 54
[3] Hölsä J, Jungner H, Lastusaari M and Niittykoski J 2001 *J. Alloys Compd*. **323** 326
[4] Nag A and Kutty TR N 2003 *J. Alloys Compd*. **354** 221
[5] Lu Y, Li Y, Xiong Y, Wang D, Yin Q. 2004 *J.Microelectron.* **35** 379
[6] Dorenbos P, 2005 *J. Electrochem. Soc*. **152** H107
[7] Holsa J, Jungner H, Lastusaari M and Niittykoski J, 2001 *J Alloys Compd*. **323–324**, 326
[8] Hörkner W and Müller-Buschbaum HK1976 *J. Inorg. Nucl. Chem*. **38** 983
[9] Ravichandran D, Johnson ST, Erdei S, Roy R and White WB 1999 *Displays* **19** 197
[10] Aitasalo T, Deren P, Holsa J, Jungner H, Lastusaari M, Niittykoski J and Strek W 2004 *Radiat. Meas*. **38** 515
[11] Hölsä J, Laamanen T, Lastusaari M, Malkamäki M, Welter E and Zajac DA 2010 *Spectrochimica Acta* **B 65** 301
[12] Lin Y, Tang Z, Zhang Z and Nan C 2003 *J. Eur. Ceram. Soc*. **23** 175

Division B – Nuclear, Particle and Radiation Physics

A Century of Nuclear Theory: Severely Condensed

Ken Amos[1,2]

[1] School of Physics, University of Melbourne, Victoria 3010, Australia
[2] Department of Physics, University of Johannesburg, Auckland Park, 2006 South Africa

E-mail: amos@unimelb.edu.au

Abstract. Defining the underlying structure of matter has been a holy grail for science. The discovery of the 'ultimate' nature of matter often has been seen as offering the ability to explain most everything else in the universe. During the twentieth century, nuclear physics became a science which promised to reveal many of these secrets. Through theory and experiment, nuclear scientists charted the underlying structure of matter and gave new insights into the fundamental nature of things. Over the last century, progress in nuclear physics has relied on a symbiotic relationship between theory and experiment. One cannot isolate developments in one from those in the other. However, as a full account of the developments in nuclear theory over the century would take multiple volumes, this article can be no more than a scant selection.

1. Introduction

Over the last hundred years nuclear physics has flourished. It is still an active and interesting field of study today. Through theory and experiment, nuclear scientists not only probed the underlying structure of matter but discovered a plethora of applications of the physics. Perhaps the most commonly known applications of nuclear physics are nuclear power generation and nuclear weapons technology. But the research has provided application in many other fields, such as magnetic resonance imaging, ion implantation in materials engineering, and radio-nuclide dating in geology and archaeology.

Through the years, nuclear physics delved ever more deeply into the atom with larger and larger machines constituting what came to be known as 'Big Science' - a characteristic form of late twentieth-century science. But there were spin-offs. The ability to produce radioactive ion beams for experimentation has enabled the nuclear landscape to be vastly increased with new physics, especially with nuclear systems at and beyond the nucleon drip-lines.

As noted in the abstract, I have had to be super selective of what I cover. I begin with some history, though a very narrowly focussed history, of the first half-century of developments in nuclear theory. In the second half-century there was such a rapid growth of ideas, models of structure and of reactions (and of numbers of nuclear physicists), coupled with the effects of the computer revolution, that it is impossible herein to give even a very narrowly focussed history. So in Sec. 3, there is an overview of some questions posed, and of the diversity of methods used, in three prime areas of study. Then, in Sec. 4, the optical model is discussed and some results of its application to a modern form are given. Finally, in Sec. 5, some general comments are made on a few of the current and future studies in the field.

2. The first half-century

During the years from 1911 into the 1960's, nuclear theories established the content, and basic structures, of nuclei. Many of the mechanisms involved in their interactions also were defined. Later, with the computer revolution, evaluations became feasible so that very detailed studies of structure and reactions could be made.

A hundred years ago, Rutherford suggested to Geiger and Marsden to look in their experiment for back scattering of α particles from target foils. To their great surprise, such were observed. Using classical scattering in a Coulomb field to interpret this, Rutherford established a maximal size to the nucleus which fostered the new idea of the structure of atoms with a small but massive nucleus at its center. Later, in 1919, his group transmuted one element into another for the first time. In their experiment nitrogen was converted into oxygen through the nuclear reaction $\alpha+{}^{14}\mathrm{N} \rightarrow {}^{17}\mathrm{O}+p$. Thereby protons were identified as being constituents of the nucleus. Radiation studies also established that nuclei could emit both β-rays and α-particles, so it was an obvious view that protons, electrons and α-particles were the building blocks of a nucleus. But there were problems even then with that view. In 1921 Rutherford theorised about the existence of neutrons to somehow compensate for the repelling effect of the positive charges of protons by causing an attractive nuclear force to keep nuclei from breaking apart. They had to wait until 1932 for the existence of neutrons to be demonstrated by Chadwick.

Prior to Chadwick's discovery, most studies on the nucleus related to ascertaining their gross properties, e.g. their size, mass, energy systematics, etc., and of the characteristics of spontaneous and induced radiations from them. Reaction data were limited to processes that could be obtained using nuclear projectiles (α's mostly) from radioactive sources. However, during preceding years, great developments were made in Quantum Theory (notably of atomic structure) that laid the foundation for much in nuclear theory as is known today. Notably, mindful of Heisenberg's uncertainty principle, that it was impossible to determine simultaneously pairs of exact 'dynamical variables' of a system, Pauli found that an electron in an atom can be identified by four distinct quantum numbers. In 1926, Gamow was the first to apply quantum mechanics to deal with a nuclear physics question. He used the concept of tunnelling to explain α-emission from nuclei. Later (1928-29) he used quantum mechanics to re-analyse the Rutherford (α, p) data, ascertaining that nuclei had a critical radius given by $R = 1.21A^{1/3}$. Nuclear volumes then are proportional to A. This was a result of note. Nuclear volumes proportional to the atomic number meant that nuclei would 'saturate', i.e. to have a constant central nucleon density. This also implied that the nucleon-nucleon force is short ranged.

A most important development in Quantum Theory was the work of Dirac in formulating a relativistic equation to describe fermions, and concomitantly, antiparticles. Fermi was to use this later to explain β-decay. Those two theories, coupled with Segré's later discovery of the anti-proton, arguably may be the genesis of the modern topic of particle physics.

The real start to theories of the nucleus began only with Chadwick's discovery of the neutron. For even growth of the liquid drop model, a theory that should have been attributed to Gamow, awaited the idea of two types and two spin values of fluids in the drop. Heisenberg, in 1932, was the first to state categorically that nuclei consisted of protons and neutrons without any constituent electrons. He introduced the concept of isospin, deeming the neutron to be a fermion. Of great importance, the discovery of the neutron meant that one could finally calculate the binding energy fractions of each nucleus by comparing the nuclear mass with the sum of those of the constituent protons and neutrons. Differences between nuclear masses were calculated in this way and, when nuclear reactions were measured, they were found to agree to high accuracy with Einstein's mass-energy equivalence formula.

In 1935, Yukawa proposed the first significant theory of the strong force to explain how the nucleus is held together. In the Yukawa interaction a virtual particle, later called a meson, mediated a force between all nucleons (protons and neutrons). This force explained why nuclei

did not disintegrate under the influence of proton repulsion and it also gave an explanation of why the attractive strong force had a very short range of interaction.

Electrons do not exist in their own right within the nucleus, though the process of internal conversion revealed that the atomic K-shell electron could be captured. The energy systematics were established to show that not all nuclei would be stable. Unstable nuclei may decay by hadronic-particle emission (α, p, n) or it may β-decay with emission of an electron or a positron. After one of these decays the resultant nucleus may be left in an excited state, from which it may cascade by γ-decay to its ground state. The β-decay of nuclei reveals a most important aspect, the existence of the weak force of nature. Electrons, protons, and neutrons are fermions and β-decay (electron) was supposed to be the disintegration of a neutron to a proton and an electron. If that were all, then spin would not be conserved. The neutrino is required. Postulated by Pauli (1931) and used by Fermi (1933), the neutrino allowed preservation of conservation laws in β-decay and explained the quite distinctive structure of the observed energy distribution of electrons. Eventually, β-decay would lead to the concept of parity violation and a re-examination of the symmetries of physical laws which often have been taken for granted.

The period 1936-60 was the golden age of nuclear theory with specification of diverse models of nuclear structure and of nuclear reaction theories that are the bases of study to this day. Nuclear theory, along with the numbers of its practitioners, burgeoned. Many textbooks were written during this period with that of Blatt and Weisskopf being a 'bible'. A look at the contents of this 850 page tome shows clearly just how intense had been the research activity into nuclear theory to the date of its publication (1952). Even so, there is little in that book about the shell model; a concept that underpins a vast range of studies now.

By 1960, the computer revolution was in full swing. One only has to consider what exists today to see this as a revolution. As examples, consider what size of evaluations can be done and how quickly, and consider the development the Internet and social networking systems and how they have affected societies. Nowadays it is almost a prime necessity to have a computer in every house and on every desk.

3. The second half-century

From the discovery of the neutron, basic aims of nuclear theory were to form a picture of the structure of the nucleus, to understand how nuclei interact, and to ascertain what results from those interactions. Answers were sought to three basic questions.

3.1. What is the force between two nucleons?

General properties of nuclei indicate that the internucleon (NN) force has to be

- strong: The binding energies of nuclei are large and the nucleus is quite small and compact.
- short ranged: Rutherford's analysis (of α scattering) showed that nuclear forces would have no noticeable effect at ranges of a few times the nuclear diameter.
- basically attractive: Since nuclei exist the components had to 'stick'
- repulsive at short range: If it were otherwise, heavy nuclei would collapse catastrophically.

The combination of repulsive character inside an attractive one for the short ranged NN force meant that nuclei saturate. Nuclei then increase in size with mass (A), nuclear radii vary as $A^{1/3}$, and nuclear binding energies per nucleon tend to a constant value of ~ 8 MeV/A. Note also that the exchange nature of the force specified by Heisenberg contributes to forming saturation. Thus more than just the exchange of a single pion is required in defining the NN force.

3.2. How can nuclear structures be described?

Understanding the structure of a nucleus is one of the central challenges in nuclear physics. Essentially it is a quantum mechanical problem of a few to fairly many strongly interacting

fermions. However, to this day, precise knowledge of the interactions between the constituent nucleons is not totally settled. There are many questions. For example, given that a nucleus has protons and neutrons confined within quite a small volume, does the two-nucleon interaction vary from the free space form when the two are embedded in nuclear matter? Also are there three and more nucleon forces at play? Those vagaries, as well as the diversity of the properties of nuclear systems, and the serious difficulties of finding 'exact' solutions to the many-fermion Schrödinger equations, has lead to development of many disparate models for nuclear structure. Some that have been developed are:

- Collective models: The liquid drop model, quantised rotors and vibrators, vibrating potential model, and other harmonic vibration models.
- The shell model: Standard shell model, deformed (Nilsson) shell model, random phase approximation (RPA), quasi-RPA, and the continuum shell model.
- Mean field theories (nonrelativistic and relativistic): Hartree-Fock (HF), Hartree-Fock-Bogoliubov (HFB), time dependent HF methods.
- Cluster models: α-cluster models, cluster-orbital model, generator coordinate methods.
- Few body models: Fadde'ev three-body, four-body (AGS), quantum Monte Carlo models.
- Group theoretic models: Wigner supermultiplet theory, SU(3) and other group classifications of states, interacting boson models.

Attempts have been made to unify structure models, but they have had limited success.

3.3. How can one describe nuclear reactions?

There many types of nuclear reactions that theories seek to describe. Examples of several common, or otherwise notable, type are:

- Fusion and capture reactions: Two light nuclei join to form a heavier one either as a final system or as a temporary 'compound' that breaks up with one or more particles ejected.
- Fission reactions: A very heavy nucleus spontaneously, or after absorbing additional light particles (often neutrons), splits into two or more pieces.
- Spallation: A nucleus hit by another with sufficient energy and momentum to knock out several small fragments, or to be smashed into many.
- Direct reactions: A projectile transfers energy and momentum and (possibly) picks up or loses nucleons to the target nucleus in a single quick (10^{-21} sec.) event.
- Compound nuclear reactions: A low-energy projectile undergoes many collisions and is absorbed forming a nucleus with too much energy to be fully bound. On a time scale of about 10^{-19} sec., particles are 'boiled' off, often they are neutrons. The compound exists until enough energy concentrates on a nucleon (or cluster of nucleons) to escape.

There were almost as many reaction theories developed through these years. Some are

- Hauser-Feshbach: Used in analyses of low-energy compound nuclear reactions.
- $\mathcal{R}$-matrix theory: Used to classify resonance structures; a central feature of large modern data-analysis codes.
- Transport theories: Used to analyse spallation, heavy ion fragmentation data especially with relativistic heavy ion collisions. They include evaporation models, the Boltzmann-Uehling-Uhlenbeck (BUU) model and its relativistic form.
- Direct reactions: The optical model, distorted wave approximation (DWA), and the multi-step Feshbach-Kerman-Koonin (FKK) theory

A key ingredient for many analyses is the optical model which is now considered in more detail assuning nucleons as projectiles, though much applies for other particles.

4. The optical model

The optical potential plays a vital role in the history of nuclear reaction theories. The concept of a complex optical potential as a single particle representation of the interaction between two nuclei dates at least to the study by Bethe in 1940 of neutron-nucleus cross sections. All early optical potentials were phenomenological. Studies of that phenomenology proceeded apace thereafter, culminating in attempts to prescribe global forms for all target masses and for projectile energies up to $\sim$ 1 GeV. Phenomenological and semi-phenomenological optical potentials are still used to interpret elastic scattering data as well as to define the distorted waves required in DWA analyses of non-elastic processes.

4.1. A formal derivation of the optical potential

Whatever approach is used (Lippmann-Schwinger equations, time dependent or stationary Schrödinger equations) to specify the optical potential acting between two nuclei, the requirement is to project the equations for the state vector onto the 'elastic channel subspace'. An elegant way of doing this was given by Feshbach in 1958. His formalism divides the Hilbert space using two projection operators P and Q. P projects onto the elastic channel and Q on to all others. Thus there is an algebra: $P^2 = P, Q^2 = Q, P + Q = 1, PQ = QP = 0$ and $Q\left|\Psi_{gs}\right\rangle = 0$. With these projection operators, the Schrödinger equation segments as

$$(E - H_{PP})P\left|\Psi^+\right\rangle = H_{PQ}Q\left|\Psi^+\right\rangle \quad ; \quad (E - H_{QQ})Q\left|\Psi^+\right\rangle = H_{QP}Q\left|\Psi^+\right\rangle ,$$

where $H_{XY} = XHY$. Using the second equation to eliminate $Q\left|\Psi^+\right\rangle$ from the first gives

$$\left(E - H_{PP} - H_{PQ}[E - H_{QQ} + i\epsilon]^{-1}H_{QP}\right) P\left|\Psi^+\right\rangle = 0 ,$$

where outgoing wave boundary conditions are assumed. The Feshbach formalism reduces the many nucleon problem to an effective one body one by invoking explicitly the ground state and transitions to other channels from the ground state. With $H = H_0 + V,\ H_{PQ}(\equiv H_{QP}) = V_{PQ}$,

$$\left(E - H_0 - \left\langle\Phi_{gs}\left|V\right|\Phi_{gs}\right\rangle - \left\langle\Phi_{gs}\left|VG^{(+)}_{QQ}V\right|\Phi_{gs}\right\rangle\right)\ \left|\chi^+\right\rangle = 0 .$$

Thereby the relative motion, *distorted wave*, function $\left|\chi^+\right\rangle$ for elastic scattering is defined. The intermediate state propagator is complex due to pole contributions. It contains the whole complex spectrum of many body excitations of the projectile and target nucleons in bound as well as in continuum states through $G^{(+)}_{QQ}$. Thus the optical potential is identified by

$$U = U_{OM}(E) = \left\langle\Phi_{gs}\left|V\right|\Phi_{gs}\right\rangle + \left\langle\Phi_{gs}\left|VG^{(+)}_{QQ}V\right|\Phi_{gs}\right\rangle \quad \text{where} \ \ G^{(+)}_{QQ} = [E - H_{QQ} + i\epsilon]^{-1} .$$

In general it is non-local, energy dependent, and if the incident energy is larger than the first excited state threshold, complex due to the second term. That is so even if the leading term is real, local, and energy independent. The second term is the dynamic polarisation potential.

For incident energies in the range of the discrete excited states of a target (usually less than a few MeV), specific coupled-channel effects are paramount. For higher energies, ones coinciding with a dense continuum in the spectrum of the target, the leading interaction can be taken as a sum of pairwise interactions between the projectile and each and every bound nucleon.

With regard to the latter condition, initially it was suppose that, for sufficiently high incident energies, those NA interactions would be ascertained from free NN scattering. Bethe in 1958 showed that the cross section and polarization from the scattering of 310 MeV protons from ^{12}C at forward scattering angles are consistent with that conjecture. Then, in 1959, Kerman, McManus, and Thaler (KMT) developed the Watson multiple scattering approach expressing the NA optical potential by a series expansion in terms of free NN scattering amplitudes. Those formulations define an effective interaction between projectile and the target nucleons.

4.2. The g-folding model

This model of the optical potential is that which my colleagues and I use today. It is based upon a non-relativistic multiple scattering theory but with the NN scattering amplitudes modified from the free NN values. Those modifications are caused by the two nucleons interacting within the nuclear medium and are due to *Pauli blocking* and *mean field effects* for both projectile and bound state nucleons. The effects are taken into account in the Brueckner-Bethe-Goldstone (BBG) equations for infinite nuclear matter systems, solutions of which are termed NN g-matrices. Also of importance is the antisymmetrization of incident nucleon with each of the A nucleons in the target. That leads to direct and knock-out (exchange) components to scattering. The effect of the exchange terms is not small at any energy and they are a source of non-locality. With non-local NA optical potentials formed using the Melbourne (effective) NN g-matrices,

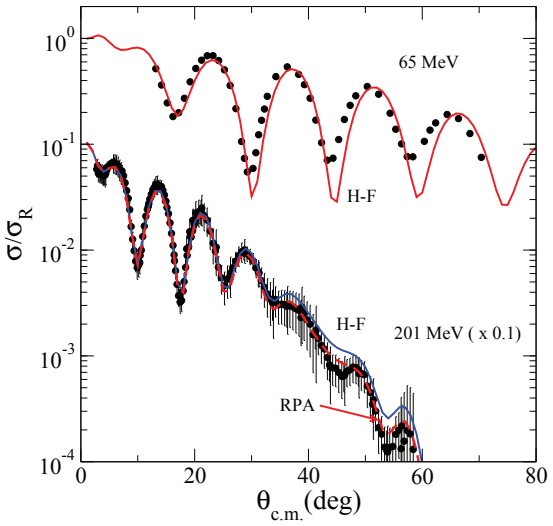


Figure 1. p-^{208}Pb elastic cross sections

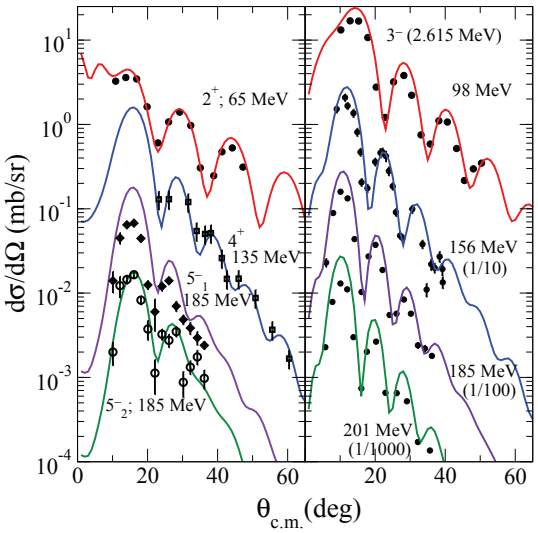


Figure 2. p-^{208}Pb inelastic cross sections

and two structure models for the ground state of ^{208}Pb, Dupuis *et al.* [1] obtained the results given in Fig. 1 The structure models from which the required. one-body density matrix elements (OBDME) were obtained are as indicated. Clearly the matches with data are very good and can be used to discriminate between the model structures. Results of similar quality were found for spin observables. In particular, such data analyses lead to the belief that, in its ground state, ^{208}Pb has a neutron skin thickness of $\sim$ 0.16 fm; a value consistent with that found from a recent parity violating electron scattering experiment.

The relative motion wave functions (the distorted waves), whatever was chosen for the optical model potential, are considered the appropriate ones to use in other reaction evaluations provided elastic scattering data are well fit. While such is an essential condition, it is not a sufficient one because elastic scattering cross section evaluations require knowledge only of a set of phase shifts. Those are quantities determined from the asymptotic properties of the wave functions. The veracity of those wave functions throughout the nuclear volume is not guaranteed. For a sufficient condition, more information about an optical potential is needed. Use of the relative motion wave functions in a successful fit to other scattering data, such as of inelastic scattering data, could suffice. This may be so, but a predictive application, i.e. one without additional parameterisations or scale adjustments, is needed since, often, the DWA is used to analyze data.

The best one can do at present is to use a formulation of the optical potential that accounts as best possible for physical laws. Notably one should take into account the Pauli principle

and treat the ensuing non-localities exactly. Only then might analyses of elastic and inelastic proton scattering data be used as good tests of structure. The results of M. Dupuis *et al.* [2] are shown in Fig. 2. They were found using the DWA with distorted waves of the g-folding optical model potentials, with the OBDME for a transition as given by the same model of structure. The Melbourne effective NN interaction was used as the transition operator. In the left-hand panel, cross-section data are compared with DWA results for excitations of different spin states in ^{208}Pb, while in the right-hand panel, results for excitation of the 3^- state are compared with data taken at different energies.

The OBDME required in the scattering calculations (elastic and inelastic) were obtained from self-consistent RPA calculations where self-consistency means that the same interaction (the D1S interaction) was used to determine the mean field single particle states, as well as being the residual interaction in RPA calculations. As Dupuis *et al.* [1] noted in their conclusions, 'excitations of high spin states up to the 12^+_1 state have also been well explained with the model and so, proton scattering has been shown to be a means to precisely investigate the structure description of heavy spherical nuclei. The goal has been achieved because no phenomenological input or arbitrary renormalization process enters the microscopic model analyses. Only with that condition can unambiguous conclusions be drawn about the structure of target nuclei.''

5. The future very briefly

The development of nuclear theories of structure and reaction has been enormous over the last few decades. Research continues in the physics of relativistic heavy-ion collisions, in nuclear astrophysics processes, in the structure of new and exotic nuclear systems and their interactions, and in the new aspects of strangeness with hypernuclei, to name a few.

Nuclear structure theories that are in use and being developed are based upon no-core shell models, upon Green's function Monte Carlo models with three-nucleon forces included, upon diverse mean field models some being relativistic, and various cluster models such as the antisymmetrized molecular cluster model.

Nuclear reaction theories also have progressed. These, in particular, result from the development in the last few decades of heavy ion facilities. Those facilities seem to be one of two categories; the first to produce relativistic heavy ion beams whose collision with targets, in particular it is hoped, will give evidence of a quark-gluon plasma, while the other group is used to produce radioactive ion beams (RIBs) for use in delineating the structure and reaction properties of nuclei off the line of stability, and indeed beyond the nucleon drip-lines.

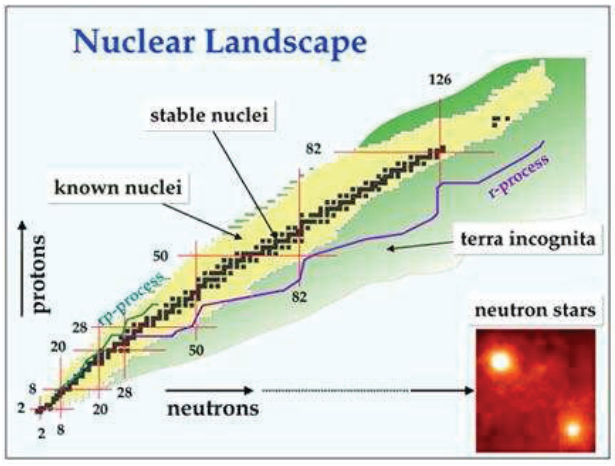


Figure 3. A stylised modern chart of nuclei

The scope is evident in Fig. 3, which is a stylised modern chart of the nuclides. Therein

- *'terra incognita'* edges are the nucleon drip-lines
- most nuclei $\beta^{\pm}$ decay
- some are particle emissive (p, n, α)
- nuclei near to the drip lines have weakly bound nucleons
- the general lines of the astrophysics r- and rp- processes are shown.
- Not shown, there is a third axis, of strangeness giving hypernuclei.

Studies of exotic nuclei are of great interest today. There are two prime methods of making RIBs. The first, the isotopic separation on-line (ISOL), is appropriate to make low energy RIBs, e.g. ≤ 10 MeV/A. For energies > 10 MeV/A, RIBs can be made using in-flight fragmentation.

As an example, consider the radioactive nucleus, ^{6}He, which has two weakly bound neutrons. The β-decaying system is known to be Borromean in that, under stimulus, it will break readily into an α plus those two neutrons. RIBs of ^{6}He have been formed and the scattering from hydrogen targets measured. Those data are compared with g-folding and DWA results in Figs 4 and 5. The elastic scattering results shown in Fig. 4 compare well (magnitude and shape) with

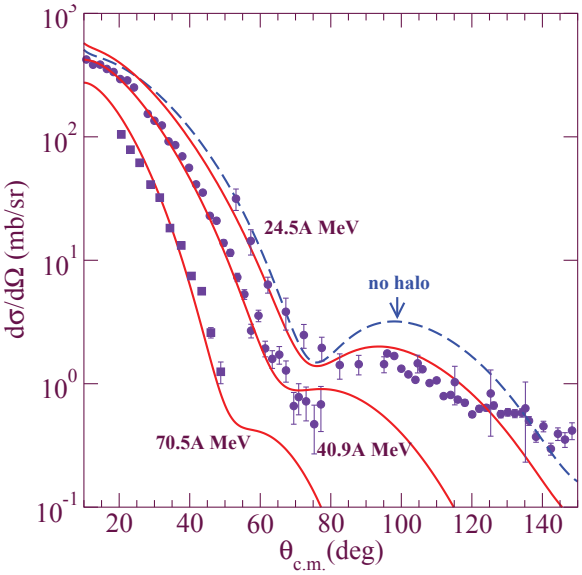


Figure 4. ^{6}He-p elastic scattering for three incident energies and a neutron halo density

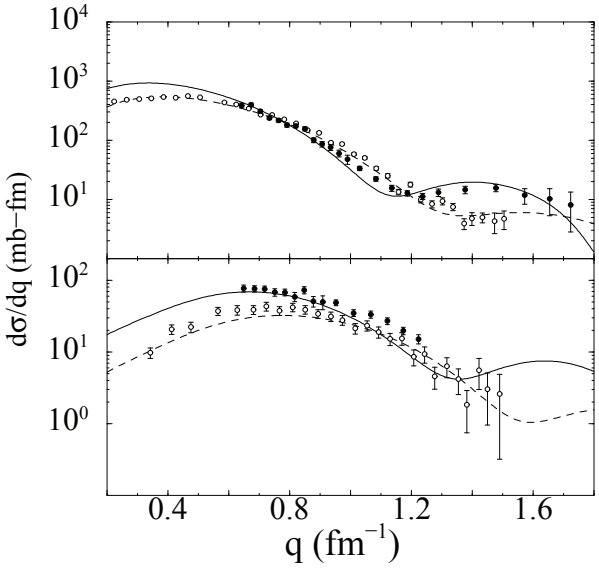


Figure 5. ^{6}He-p elastic (top) and inelastic scattering to the 2_1^+ state (bottom)

calculations made using large space no-core shell model wave functions with single particle states chosen to give ^{6}He an extended neutron halo. When wave functions are used to give it neutron skin (no halo), the match is quite imperfect.

With elastic scattering the halo character affects the results for high momentum (large angle scattering) as is shown by the results compared with data taken at 24.5A MeV (filled circles) and at 40.9A MeV (open circles) in the top segment of Fig. 5. In the bottom segment DWA results at those incident energies are compared with data from the inelastic excitation of the 2_1^+ (1.8 MeV) state in ^{6}He. In this case it is the smaller momentum transfer results that are most affected by the halo prescription reflecting the reduction in strength of wave functions in the surface region to have the extended neutron distribution that is the halo. These momentum transfer diagrams also show that the energy dependence of the effective NN interaction used in the DWA calculations is important since the structure used in all calculations was fixed.

Acknowledgments

Material in this presentation has been taken from many books, papers, and articles; so many that I have not given a list of references. I do reference the two recent papers from which figures were taken. The nuclear landscape figure was taken from a Google search and is one of many in the literature.

[1] Dupuis M, Karataglidis S, Bauge E, Delaroche J P and Gogny D 2006 *Phys. Rev. C* **73** 014605
[2] Dupuis M, Karataglidis S, Bauge E, Delaroche J P and Gogny D 2008 *Phys. Lett.* **B665** 152

Quasifree alpha cluster knockout studies

A A Cowley[1,2]

[1]Department of Physics, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa
[2]iThemba Laboratory for Accelerator Based Sciences, P O Box 722, Somerset West 7129, South Africa

Email: aac@sun.ac.za

Abstract. Cluster-like structures in the shell-model description of the ground state of nuclei can be conveniently studied by means of knockout reactions. Of these the (p, $p\alpha$) reaction is perhaps the simplest, especially from the viewpoint of the tractability of theoretical calculations used to interpret experimental results. Analyzing power angular distributions, which are simple ratios of cross sections, are investigated, as these are expected to be very sensitive to details of the reaction mechanism. The distorted wave impulse approximation (DWIA) is a versatile theory which is applicable to knockout reactions. Fortunately its results appear to be reasonably insensitive to uncertainties in the exact ingredients, such as distorting optical potentials, which are obtained from unrelated elastic scattering studies. It is shown that surprisingly simple approximations in the DWIA hold for α–cluster knockout from light nuclei. Furthermore, results for a medium-mass nuclear target such as ^{40}Ca are also consistent with expectation if the appropriate distorting optical potentials for the outgoing α–particle are employed in the DWIA formulation.

1. Introduction

Spectroscopic factors for pickup, stripping and knockout of α–clusters from and to the ground state of atomic nuclei appear to be in reasonable agreement with shell model estimates, for example as shown by Chung *et al.* [1]. Knockout, such as (p, $p\alpha$) reactions to the ground state of the final residual nucleus, offer a convenient experimental technique to study such cluster-like subsets of normal shell model wave functions. The virtue of knockout experiments is that the intrinsic three-body kinematics in the exit channel can be selected in such a way that the nuclear α-cluster structure of the target nucleus can be studied separately from the two-body projectile-cluster interaction. This feature is in strong contrast with either pickup or stripping reactions, for which the nuclear structure is convoluted with the reaction part, which complicates interpretation of experimental results. More specifically, in knockout experiments the angles and energies of the observed outgoing light ejectiles can be varied in such a way that the residual nucleus always remains at rest. This so-called quasi-free angle pair setup allows the proton-α two-body interaction to be studied and to be directly compared with free scattering of protons from ^{4}He. Alternatively, the geometry can be adjusted to keep the two-body kinematic condition fixed in order to investigate the momentum distribution of α–clusters in the

ground state of the target nucleus. In this work the emphasis will be on the two-body aspect of the knockout reaction.

A schematic diagram of the knockout process is shown in figure 1. The upper vertex in the diagram represents the projectile α-cluster interaction, and the lower vertex accounts for the cluster structure of the target nucleus. Of course, in a theoretical treatment distorted waves, which are required to reproduce elastic scattering, are used to take interactions in the incident and outgoing channels into account (not explicitly shown in the diagram).

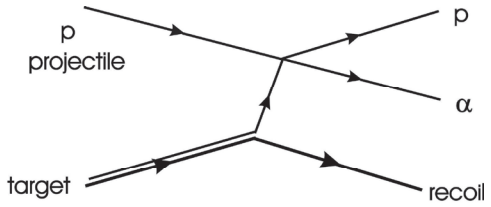


Figure 1: Schematic diagram of the knockout process.

An angular distribution of the analyzing power, which represents a measure of the left-right asymmetry experienced in the scattering of a spin-polarized projectile, for the (p,$p\alpha$) reaction on ^{12}C [2] will be compared with the results of the target nucleus ^{40}Ca [3]. The coincident analyzing power distribution, which is strongly influenced by the collision of the projectile with the α–cluster bound in the target ^{12}C, retains the characteristic features of the corresponding observables of ^{4}He(p,p)^{4}He elastic scattering at the same incident energy to a remarkable extent. This is consistent with a distorted wave impulse approximation [4] (DWIA) calculation. It should be noted that projectile-cluster optical model potentials [2], which give a reasonably good account of experimental cross section [5] and analyzing power distributions [6] of free ^{4}He(p,p)^{4}He scattering at an incident energy of 100 MeV, are used for the two-body system in our DWIA calculations.

In strong contrast with the simple results of the light-mass target, the analyzing power distribution of knockout from ^{40}Ca is profoundly affected by the heavier mass of the spectator part of the target. This results in an induced asymmetry, which is evidently caused by an increased distortion affecting the outgoing wave functions. Nevertheless, the DWIA theory provides an excellent description of the observed analyzing power distributions also for knockout α–cluster knockout from the target nucleus ^{40}Ca if care is taken to use a distorting potential for the outgoing α–particle which reproduces elastic scattering well, as will be shown in the next section.

Cross sections of the (p,$p\alpha$) knockout reaction have been investigated relatively frequently in the past, but because the analyzing power should be more sensitive to details of the reaction mechanism, it is the observable which is studied in this work. A relevant review of clustering in general, including knockout reactions, is provided by Hodgson and Běták [7].

It should be mentioned that in the DWIA theory those ingredients such as distorting potentials are, as usual, obtained from elastic scattering experiments and are not treated as free parameters. Of course, these parameters are of variable reliability. For example, the α–projectile optical potentials extracted from elastic scattering are known to suffer from discrete ambiguities at low incident energies, but fortunately guidance [8] may be obtained from higher-energy experiments. Various global parameter sets that reproduce elastic scattering of protons from target nuclei very well, over a large mass and incident energy range, are fortunately freely available.

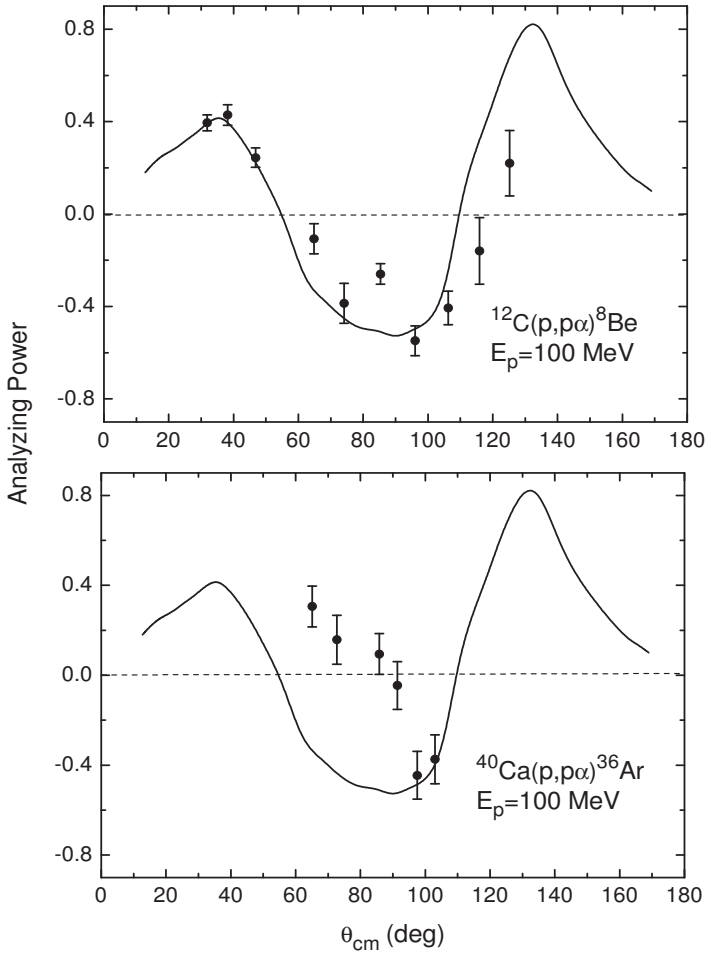


Figure 2: Analyzing power distributions as a function of the two-body p-α centre-of-mass scattering angle for the (p,pα) reaction on ^{12}C and ^{40}Ca at an incident energy of 100 MeV. Experimental values are shown with statistical error bars. Appropriate kinematics is selected for zero recoil momentum of the heavy residual nucleus. The impulse approximation implies that the struck cluster is at rest in the laboratory co-ordinate system. The curves represent a smooth line drawn through experimental analyzing power angular distributions for ^{4}He(p,p)^{4}He elastic scattering at the same incident energy from reference [6].

2. Results and discussion

The analyzing power angular distributions as a function of the two-body p-α centre-of-mass scattering angle for the (p,pα) reaction on ^{12}C and ^{40}Ca at an incident energy of 100 MeV are shown in figure 2. The experimental values are compared with results measured for elastic scattering of protons from ^{4}He at the same incident energy. Clearly the two experimental distributions are very similar for ^{12}C, as would be expected on simplistic grounds. Furthermore explicit DWIA calculations demonstrate that this relationship holds simply because the spectator part of the target nucleus, ^{8}Be, does not influence the knockout reaction to an appreciable extent [2].

Unfortunately angular limitations of the detectors used to measure the ^{40}Ca(p,pα)^{36}Ar knockout reaction (magnetic spectrometer in coincidence with a Si-Ge telescope) prevented measurements over a wider angular range. Results for the ^{40}Ca(p,pα)^{36}Ar reaction clearly differ extensively from the analyzing power of free elastic scattering. In fact, over this restricted range the free elastic scattering distribution has mostly a different sign from the knockout data.

Initial calculations [3] with the DWIA failed to reveal the reason for this discrepancy and it suggested that the analyzing power data of the knockout reaction should follow the trend of free scattering. These earlier calculations [3] relied mainly on an optical potential for the outgoing α–particle which was derived by Carey *et al*. [8] for use in (*p*,*p*α) cross section distributions over a large mass-range of target nuclei. Because, as was mentioned before, α–particle potentials are not as well-established as proton global potentials, Carey *et al.* used a procedure, which was sound for the purpose of generating a global set of α–particle potentials from the available literature that describes

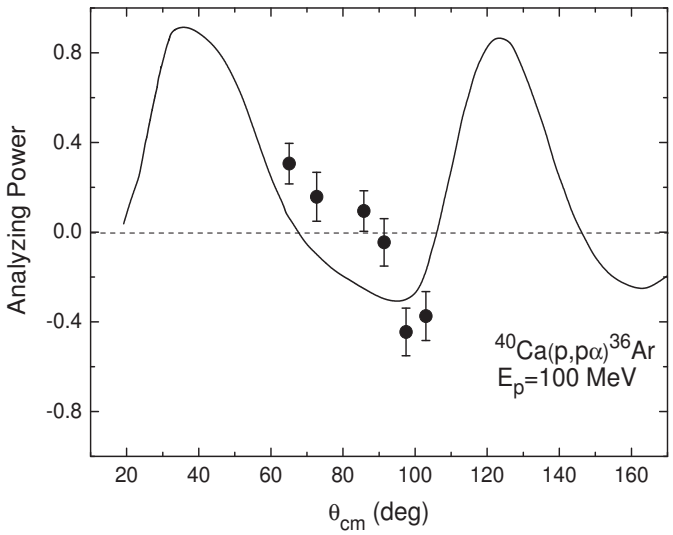


Figure 3: Analyzing power distribution for the reaction ^{40}Ca(*p*,*p*α)^{36}Ar at an incident energy of 100 MeV. Also see caption to figure 1. The curve is a prediction of the DWIA as described in the text.

the average target mass and incident energy dependence very well. However, for the present application to the target nucleus ^{40}Ca, it turns out that elastic scattering of α–particles from the residual nucleus ^{36}Ar is subject to the so-called anomalous large angle (ALAS) effect, first observed by Gruhn and Wall [9]. Use in the DWIA of a distorting potential that describes the ^{36}Ar(α,α)^{36}Ar elastic scattering correctly (as opposed to the Carey global potential which predicts elastic scattering cross sections that differ from experimental values for ^{36}Ar by up to two orders of magnitude at backward angles) results in the theoretical prediction shown in figure 3. Of course, due to the fact that the projectile-cluster interaction (upper vertex in figure 1) is nevertheless affected by the distorted waves associated with the lower vertex, the correct treatment of distortions becomes crucial, especially for this fairly massive target-recoil system.

The α–particle optical model potential parameters employed in the DWIA calculation displayed in figure 3 are from the work of Reidemeister *et al*. [10]. The α–particle potential used was the Woods-Saxon squared version of the form factor for both the real as well as the imaginary parts of the optical potential, with parameters from [10].

The DWIA result of figure 3 is evidently a considerable improvement on the curve shown in figure 2 for ^{40}Ca. For example, the difference between the DWIA and the first data point (figure 3) is less than 0.2, whereas figure 2 gives a discrepancy of 0.8 (-0.4 as opposed to 0.4; a difference in sign, as was mentioned before!).

Clearly, further improvement is required, but guidelines to achieve this are not clear. It would not be meaningful, for example, to merely adjust the parameters further to get best agreement with the experimental distribution. Nevertheless, it is apparent that the introduction of a more appropriate α-

particle optical model potential improves the DWIA calculation of the analyzing power from something that is far out of line to a distribution which is comparable to the experimental values. In other words, the results which are shown in figure 3 are encouraging.

It should be mentioned that Carey *et al.* [8] did indeed explore alternative α-particle optical model parameterizations of Woods-Saxon shape, but they found only a very small sensitivity to the cross section. Of course, the present study indicates that the analyzing power is much more sensitive to this ingredient. In fact, if we compare a DWIA calculation using the appropriate parameter with the cross section measurement of Carey *et al.* [8], a similar insensitivity as encountered by them is observed. The only effect is that the spectroscopic factor, which is extracted by means of a normalization of the theory to the experimental cross sections, changes by a modest 10%. This change in the DWIA results is well within the differences found by Carey *et al*. Furthermore, the modification of the shape of the cross section energy distribution is insignificant.

3. Summary and conclusion

The (p,$p\alpha$) knockout reaction at an incident energy of 100 MeV to the ground state of the residual nucleus on the target ^{12}C displays an analyzing power distribution which follows the trend of elastic scattering of protons from ^{4}He. This resemblance of quasifree knockout to free scattering is in agreement with a DWIA prediction. The knockout α–cluster analyzing power angular distribution for the target nucleus ^{40}Ca, however, shows a significant deviation from a free interaction between a proton projectile and ^{4}He. Only when an optical model parameter set, which reproduces the elastic scattering between the residual nucleus ^{36}Ar and the emerging α–particle, does the DWIA predict the experimental distribution reasonably accurately.

The mere fact that ^{40}Ca is so much heavier than ^{12}C is probably not the only, or even the crucial difference between these two cases that accounts for the observed behaviour of the analyzing power distributions. It appears that it is of more importance that in the former situation the DWIA theory requires a distorted wave in the α–^{36}Ar outgoing channel that gives an accurate account of anomalous large angle elastic scattering.

It would be useful to investigate the (p,$p\alpha$) knockout reaction for other adjacent medium-mass target nuclei, most of which do not involve an outgoing system that is subject to anomalous elastic scattering. The basic formulation of the DWIA appears to be sound, but more refined analysis would require a better understanding of the distorting optical model parameters for especially the α–particle. Consequently there is a need for further experimental as well as theoretical development.

Acknowledgment

This work was performed with funding from the National Research Foundation (NRF) of South Africa. The financial support was essential for the completion of this study. Numerous discussions with my collaborators, Justin Mabiala, Siegfried Förtsch, Deon Steyn and Retief Neveling regarding the significance of experimental results upon which this report is based, are gratefully acknowledged.

References

[1] Chung W, Van Hienen J, Wildenthal W H and Bennett C L 1978 *Phys. Lett.* B **79** 381
[2] Mabiala J, Cowley A A, Förtsch S V, Buthelezi E Z, Neveling R, Smit F D, Steyn G F and Van Zyl J J 2009 *Phys. Rev.* C **79** 054612
Cowley A A, Mabiala J, Buthelezi E Z, Förtsch S V, Neveling R, Smit F D, Steyn G F and Van Zyl J J 2009 Europhys. Lett. **85** 22001
[3] Neveling R, *et al.* 2008 *Phys. Rev*. C **77** 037601
[4] Chant N S and Roos P G 1977 *Phys. Rev.* C **15** 57
Chant N S and Roos P G 1983 *Phys. Rev.* C **27** 1060
[5] Goldstein N P, Held A and Stairs D G 1970 *Can. J. Phys.* **48** 2629

[6] Nadasen A, Roos P G, Mack D, Ciangaru G, Rees L, Schwandt P, Kwiatkowski K, Warner R E and Cowley A A 1983 *Indiana University Cyclotron Facility Scientific and Technical Report* p. 13 (unpublished)
[7] Hodgson P E and Běták E 2003 *Phys. Rep.* **374** 1
[8] Carey T A, Roos P G, Chant N S, Nadasen A and Chen H L 1984 *Phys. Rev.* C **29** 1273
[9] Gruhn C R and Wall N S 1966 *Nucl. Phys.* **81** 161
[10] Reidemeister G, Ohkubo S and Michel F 1990 *Phys. Rev.* C **41** 63

Low-lying states in ^{17}C from a multi-channel algebraic scattering theory

S. Karataglidis

Department of Physics, University of Johannesburg, P.O. Box 524 Auckland Park, 2006, South Africa

E-mail: stevenka@uj.ac.za

K. Amos

Department of Physics, University of Johannesburg, P.O. Box 524 Auckland Park, 2006, South Africa, and, School of Physics, University of Melbourne, Victoria 3010, Australia

L. Canton

Istituto Nazionale di Fisica Nucleare, Sezione di Padova, e Dipartimento di Fisica dell'Università di Padova, via Marzolo 8, Padova I-35131, Italia

J. P. Svenne

Department of Physics and Astronomy, University of Manitoba, and Winnipeg Institute for Theoretical Physics, Winnipeg, Manitoba, Canada R3T 2N2

D. van der Knijff

School of Physics and ITS Research Services, University of Melbourne, Victoria 3010, Australia

P. R. Fraser

Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, 04510 México, D.F., Mexico

Abstract. The structure of ^{17}C is used to define a nuclear interaction that, when used in a multichannel algebraic scattering theory for the $n+\ ^{16}$C system, gives a credible definition of the (compound) excitation spectra. When couplings to the low-lying collective excitations of the ^{16}C-core are taken into account, both sub-threshold and resonant states about the $n+\ ^{16}$C threshold are found.

The spectra of radioactive nuclei, especially of masses at or just beyond a drip line, are most intriguing. To date, details of their spectra are at best poorly known. Few, if any, excited states have been identified. Likewise, the spin-parities of known states have not been, or are uncertainly, assigned. Nowadays, opportunities exist to investigate spectra of such exotic systems using isotope separator on-line facilities with which production of radioactive ion beams having energies typically $0.1A$ to $10A$ MeV is possible. Reactions using these beams with higher

incident energy can be, and have been, used to study the structure of the radioactive ions as well. However, it is the low energy domain that interests us as we wish to consider structures of compound systems formed by amalgamation of the beam ion and a nucleon.

With light mass systems having charge number π and neutron number ν, there is often the possibility to link the structures of mirror systems. Usually there is a reasonably well known spectrum of a nucleus (${}^{A+1}_{\pi=Z}X_{\nu=N+1}$), which we treat as a compound of a neutron (n) with ${}^{A}_{\pi=Z}X_{\nu=N}$ to define a chosen nuclear model interaction. With that model, assuming charge invariance of the nuclear force and adjusting for Coulomb effects, the spectrum of the mirror, ${}^{A+1}_{\pi=N+1}Q_{\nu=Z}$, may be predicted. This has been done [1], for example, for the mass-7 isobars with multichannel algebraic scattering (MCAS) [2] evaluations of the spectra of the compound systems; ^{7}Li (as $n+{}^6$Li), ^{7}Be (as $p+{}^6$Li), ^{7}He (as $n+{}^6$He), and ^{7}B (as $p+{}^6$Be). Also, the approach predicted a spectrum for the (particle unstable) nucleus ^{15}F when treated as a compound of $p+{}^{14}$O, characteristics of which were later observed. In that study [3] the nuclear interaction was set by an analysis of the mirror system ^{15}C treated as $n+{}^{14}$C. It was found that key requirements for obtaining resonance states of ^{15}F were, a) the Coulomb barrier, which is essential in recreating the resonance aspect of the observed spin $\frac{1}{2}^+$ ground state, b) coupling of the extra core proton to distinct states of ^{14}O and c) consideration of the Pauli principle within a coupled-channel collective model prescription. The latter two features ensured a credible sequence of spin-parity values, a very good fit to the high-quality elastic scattering cross sections found at that time [4, 5], and prediction of a set of narrow resonances only a few MeV above the (two) known ones. Narrow resonances in the region of that excitation energy, subsequently have been observed [6]. (Note: There is an error in Ref. [6] relating to citation of the results of our earlier study [3]. In their table I, references 6 and 7 have been reversed in both the table caption and the header row.)

Recently, Timofeyuk and Descouvemont [7] used the same philosophy of fixing the nuclear aspect by a two center cluster model of ^{17}C, treated as $n+{}^{16}$C, to then find a spectrum of ^{17}Na as $p+{}^{16}$Ne. They expect there to be narrow resonances in the low excitation spectrum of the particle unstable isotope of Na with very broad ones above that. Those results were a spur to us to use the MCAS approach as a complementary study of that exotic nucleus. Thus, in this paper, we consider the nucleus ^{17}C ($n+{}^{16}$C) given that the low excitation spectrum of ^{17}C has been found experimentally [8, 9, 10] and microscopic models of that nucleus' structure have been proposed [8, 9, 11, 7]. An MCAS analysis of the low excitation spectrum of ^{17}C has been made before [11]. However, the results found were from an (overly) simple, two-channel, evaluation. In that study, we also considered a distorted wave approximation analysis of inelastic scattering data of 70A MeV ^{17}C ions from hydrogen targets. Those DWA evaluations were made using no-core shell model wave functions and the complete set of results allowed us to pose some constraints upon the structure of ^{17}C.

However, inadequacies remained due, in part, to simplifications in the previous MCAS evaluations. Experiments [9] now have suggested a number of spin-parity assignments, especially of the three closely spaced sub-threshold states of ^{17}C, which current microscopic (shell) models fail to match adequately. It has been suggested [7] that coupling of a neutron to states at about 4 MeV excitation in ^{16}C is needed to improve the results. That coupling was not included in our previous study [11] and so we present herein results of three state MCAS evaluations in which coupling to the 4^+ state in the mass-16 nuclei is added.

We have made MCAS calculations of the $n+{}^{16}$C system using three states in ^{16}C; the 0^+ (ground), 2^+ (1.766 MeV), and 4^+ (4.142 MeV) states. We presume that coupling to the other states (presumed $0^+_2, 2^+_2$, 3^+) in the spectrum between 3 and 5 MeV excitation is not strong and that the rotational model, as defined previously [2], suffices in seeking the spectrum of ^{17}C. The parameter values required to get the results displayed in Fig. 1 are listed in Table 1. The potential parameter values and Pauli blocking/hindrance weights are quite similar to the

Table 1. The parameter values used to define the channel coupling properties of the $n + {}^{16}$C system. Energy units are MeV, length units are fm.

V_0	V_{ll}	V_{ls}	V_{Is}
-37.0	-2.0	9.0	1.6
R	a	β_2	β_4
2.9	0.8	0.33	0.1
state in ^{16}C	$(1s_{\frac{1}{2}}, 1p_{\frac{3}{2}}, 1p_{\frac{1}{2}})$	OPP λ_{lj} $1d_{\frac{5}{2}}$	$2s_{\frac{1}{2}}$
0^+ (0.000)	10^6	2.7	0.0
2^+ (1.766)	10^6	2.7	0.0
4^+ (4.142)	10^6	0.0	2.0

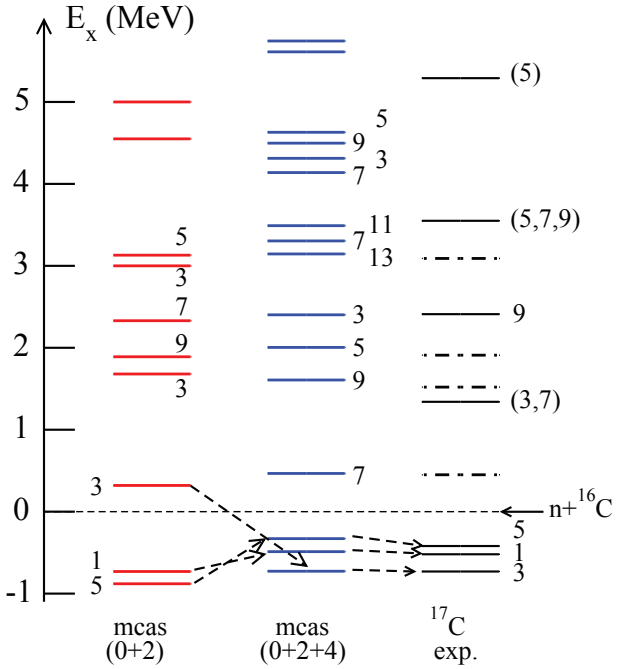


Figure 1. The spectrum of ^{17}C compared with the MCAS results.

set used previously [11], now with inclusion of a small hexadecapole deformation to link the 4^+ state to the ground in first order, and with some Pauli hindrance of the $2s_{\frac{1}{2}}$ orbit in the connections to the 4^+ target state.

The spectrum that results is shown in Fig. 1 and labelled 'mcas(0+2+4)'. It is compared with the spectrum found previously from the 2-state MCAS evaluation ('mcas(0+2)'), and with the experimentally-known one that is a combination of states listed in Table 4 of Ref. [9] and in Fig. 5 of Ref. [12]. The energies of that tabled spectrum have been adjusted by −0.729 MeV; the value of the $n + {}^{16}$C threshold in ^{17}C. Some states specified by Raimann *et al.* [12] are displayed by the dash-dot lines. Only positive parity states are known in the low excitation spectrum and (twice) their spins are given as the integers associated with the individual levels shown.

Clearly the three known subthreshold (n+^{16}C) states are now matched well in energy and spin-parity by the new MCAS results. The other known and uncertain spin-parity states also have matching MCAS partners in proximity of their excitation energies. Additionally, the

uncertain states from Raimann *et al.* [12] seem to have possible matches with the first low-lying state above threshold expected to be a $\frac{7}{2}^+$ resonance. The MCAS spectrum has a number of aspects in common with that shown in Fig. 1 of Ref. [7]. Besides the three closely spaced and weakly bound sub-threshold states, there is a group of $\frac{3}{2}^+$, $\frac{5}{2}^+$, $\frac{7}{2}^+$, and $\frac{9}{2}^+$ states in the region around 3 MeV excitation and a second group in the region of 5 MeV excitation. There is also a higher excited $\frac{1}{2}^+$ state found with both calculations above 6 MeV excitation, notable by being very broad (width $\Gamma = 5.6$ MeV with MCAS). The third $\frac{3}{2}^+$ state in the MCAS spectrum (centroid at 5.04 MeV excitation) is also very broad ($\Gamma = 4$ MeV). Such half-widths are typical of a single-particle potential resonance. Our MCAS result shows more states, including two very narrow ones of spin-parity $\frac{13}{2}^+$ and $\frac{11}{2}^+$ at 3.82 and 4.16 MeV excitation respectively.

Table 2. Spectra for ^{17}C. Units are MeV.

	Ref. [9]		Ref. [12]	MCAS		
J^π	E	Γ	E	J^π	E	Γ
$\frac{3}{2}^+$	0.00	—	0.000	$\frac{3}{2}^+$	0.00	—
$\frac{1}{2}^+$	0.21	—	0.292	$\frac{1}{2}^+$	0.16	—
$\frac{5}{2}^+$	0.31	—	0.295	$\frac{5}{2}^+$	0.22	—
			(1.18)	$\frac{7}{2}^+$	1.17	10^{-8}
$(\frac{3}{2}^+ - \frac{7}{2}^+)$	2.06	0.25	(2.25)	$\frac{9}{2}^+$	2.31	$3\text{x}10^{-5}$
			(2.64)	$\frac{5}{2}^+$	2.70	0.018
$\frac{9}{2}^+$	3.10	0.10		$\frac{3}{2}^+$	3.01	0.096
			(3.82)	$\frac{7}{2}^+$	4.03	$3\text{x}10^{-5}$
$(\frac{5}{2}^+, \frac{7}{2}^+, \frac{9}{2}^+)$	4.25	0.14		$\frac{11}{2}^+$	4.24	$2\text{x}10^{-5}$
				$\frac{7}{2}^+$	4.88	0.010
				$\frac{3}{2}^+$	4.94	1.78

Specifics of the states in the ^{17}C spectrum are listed in Table 2. Columns 1 to 3 display values of spin-parity J^π, excitation energy (or centroid) E, and half-widths $\frac{1}{2}\Gamma$, of resonances ascertained in a study [9] of three neutron transfer cross sections for ^{12}C scattering from ^{14}C. The excitation energies of states shown in Fig. 5 of Ref. [12] are listed in column 4, while the MCAS results are displayed in columns 5, 6, and 7. The three nucleon transfer reaction widths in general do not match those from the MCAS evaluation but the two sets are quite different; the latter being single nucleon removal values.

The resonant states found using MCAS in the region of 3 MeV excitation have widths that agree well with most of the matching ones from the MCM evaluation [7]. Widths quoted in Ref. [7] are $\frac{7}{2}^+\Big|_1$ (10^{-12} MeV), $\frac{9}{2}^+\Big|_1$ (10^{-6} MeV), and $\frac{5}{2}^+\Big|_2$ (0.015 MeV). The $\frac{3}{2}^+\Big|_2$ resonance half-width of 0.265 MeV is larger than the MCAS value of 0.102 MeV.

Now, it is said that ^{17}C has a 'peculiar' structure connected perhaps with the neutron separation energy from the ground state being only 0.728 MeV. That is typical of a halo nucleus. Indeed, the channel-coupling interaction we require to give the spectrum of ^{17}C reflects that with diffuseness being large. As well, features of this interaction are just as identified [7] as being required in a two-body potential model of this system, *viz. "The bound ^{17}C spectrum cannot be understood in the two-body potential model with deformation and the 2^+ excitation of the ^{16}C core either, if standard sets of potentials are used. An ℓ-dependent and nonstandard spin-orbit $n + {}^{16}C$ potential must be used for these purposes."* That is illustrated in Fig. 2 which shows

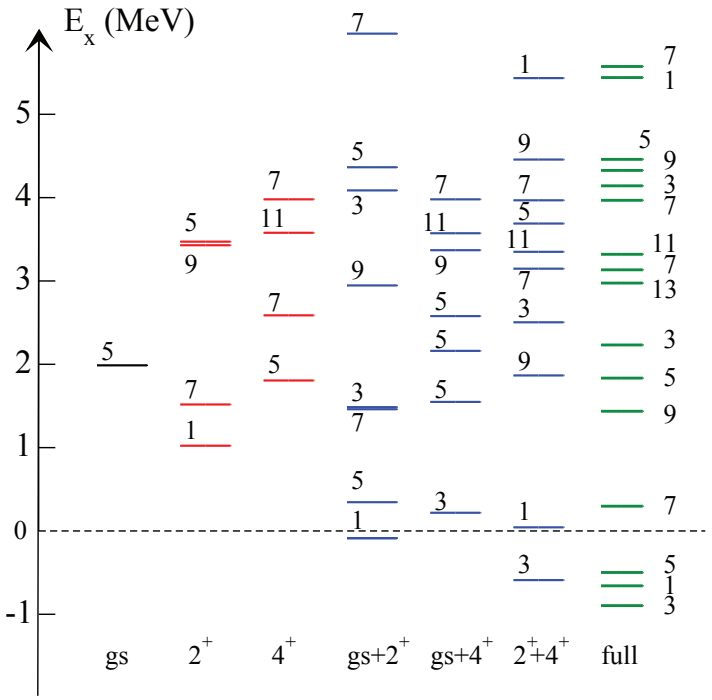


Figure 2. Spectra from MCAS evaluations restricting couplings to single- and two-channels of the three compared with the full three-channel spectrum. Again the energy scale is set with the $n + {}^{17}$C threshold as zero, to emphasise the sub-threshold from resonance states in each evaluation.

how spin states arise from underlying components of the coupled-channel approach. Calculations have been made for each subdivision of the full three target state coupled-channel problem. The numerals indicate two times the spin values. The first three columns from the left give the states found when each state alone is considered as a single channel problem. Obviously, within the searched energy range (to 6 MeV), the nucleon on the ground state gives only a single state from adding a $d_{\frac{5}{2}}$ neutron. The addition of a neutron (probably into a $d_{\frac{5}{2}}$ single-particle state) gives the set shown for the single channels of the 2^+ and 4^+. The order shown is due to the spin and angle dependent features used in the base interaction. The existence of the second $\frac{7}{2}^+$ state when a neutron is added to the isolated 4^+ state may be reflecting addition in a $d_{\frac{3}{2}}$ state.

The next three columns are the spectra found when two of the three ^{16}C states are allowed in the coupling. The results found coupling the ground and 2^+ states is very similar to what was published in an earlier paper [11]; differences reflecting changed interaction parameter values. The inclusion of the 4^+ state, whether in a 2- or the full 3-state coupling study, leads to a richer spectrum and shifts the order of states. Of note, only evaluations in which the 2^+ state is included give a low lying $\frac{1}{2}^+$ state. However, the three-state coupling markedly moves the energy value from that found otherwise.

The key feature is finding the three closely spaced sub-threshold states and in this spin order. Significant changes of the parameters, particularly with the deformations, radius and diffuseness to smaller values, cause such packing to be lost. The coupling of the 4^+ state in ^{16}C with the ground and 2^+ states causes notable changes to the predicted spectrum; changes that better align with (so far) experimentally defined states in ^{17}C, a few of which have been assigned spins. Clearly we predict many other (resonance) states, but whether they exist or can be found if they do, remains an open question. It is most unlikely that any direct $n + {}^{16}$C experiment will ever

be done so one must rely on some surrogate approach, such as ^{16}C(d, p), or some study that identifies neutron emissions from ^{17}C.

We have described the low-energy spectrum of ^{17}C, as a neutron coupled to ^{16}C, using the MCAS framework, finding correspondence with the experimental spectrum. The results from the present calculation are a distinct improvement on the earlier results. Adding a Coulomb potential to the system would give the spectrum of the mirror nucleus ^{17}Na; work is currently under way to determine that spectrum to compare to the cluster model results.

References

[1] Canton L, Pisent G, Amos K, Karataglidis S, Svenne J P, and van der Knijff D 2006 *Phys. Rev. C* **74** 064605
[2] Amos K, Canton L, Pisent G, Svenne J P, and van der Knijff D 2003 *Nucl. Phys. A* **728** 65
[3] Canton L, Pisent G, Svenne J P, Amos K, and Karataglidis S 2006 *Phys. Rev. Letts.* **96** 072502
[4] Goldberg V Z *et al.* 2004 *Phys. Rev. C* **69** 031302(R)
[5] Guo F Q *et al.* 2005 *Phys. Rev. C* **72** 034312
[6] Mukha I *et al.* 2009 *Phys. Rev. C* **79** 061301
[7] Timofeyuk N K and Descouvemont P 2010 *Phys. Rev. C* **81** 051301
[8] Elekes Z *et al.* 2005 *Phys. Lett. B* **614** 174
[9] Bohlen H G *et al.* 2007 *Eur. J. A* **31** 279
[10] Satou Y *et al.* 2008 *Phys. Lett. B* **660** 320
[11] Karataglidis S, Amos K, Fraser P, Canton L and Svenne J P 2008 *Nucl. Phys. A* **813** 235
[12] Raimann G *et al.* 1996 *Phys. Rev. C* **53** 453

Characterization of cluster states in ^{16}O with the (p,t) reaction

J Mabiala[1,2], E Z Buthelezi[1], A A Cowley[1,2], S V Förtsch[1], M Freer[3], T T Ibrahim[1,2], J P Mira[1], R Neveling[1], P Papka[1,2], F D Smit[1], G F Steyn[1], J A Swartz[1,2], I Usman[1], J J Van Zyl[2] and S M Wyngaardt[2]

[1] iThemba Laboratory for Accelerator Based Sciences, Somerset West 7129, South Africa
[2] Department of Physics, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa
[3] School of Physics and Astronomy, University of Birmingham, Birmingham, B15 2TT, United Kingdom

E-mail: mabiala@sun.ac.za, justinm@aims.ac.za

Abstract. Several theoretical cluster calculations predict the existence of the 0_6^+ state in ^{16}O, located at 15.1 MeV. This state is considered as a very good candidate for the equivalent of the Hoyle state in ^{12}C. In order to investigate this possibility a high energy resolution measurement of the ^{16}O spectrum in coincidence with the ^{16}O decay products was proposed, using a (p,t) reaction at zero degrees for a proton energy of 200 MeV.

1. Physics motivation

The excited states of light nuclei lying near and above the particle-decay threshold play an important role in the nucleosynthesis in stars. In 1954, Hoyle showed that the observed amount of carbon in the cosmos could be made in stars only if there was an excited state in carbon with a particular spin and parity, 0^+, and a particular energy of about 7.65 MeV that enhances the fusion of three α-particles. This state was found experimentally [1] and lies just 375 keV above the three α-decay threshold. The energy of this state is not well reproduced by the no-core shell model [2]. However, many theoretical calculations that use microscopic α-cluster models indicate that the 0_2^+ state in ^{12}C has a loosely bound three-α cluster structure [3,4,5].

The ^{8}Be nucleus has the simplest form of α-particle clustering in nuclei that is found in its ground state [6]. Aside from this nucleus, there are many other light systems which are proposed to exhibit cluster-like properties most likely in excited states with a low density structure. The sixth 0^+ state in ^{16}O located 660 keV above the four-α decay threshold at E_x=15.1 MeV ($\Gamma = 166$ keV) has been identified as the best candidate for the equivalent of the Hoyle state in ^{12}C [7]. This state should be found to have a well developed four-α substructure with a rather low density.

An experiment was recently proposed to populate the 15.1 MeV 0^+ state in ^{16}O using the ^{18}O$(p,t)^{16}$O reaction in order to study its decay properties. Due to the density of states in the region of excitation energy around 15 MeV in ^{16}O the triton measurement was performed with the high energy resolution K600 magnetic spectrometer. This paper reports on the preliminary analysis of data acquired with targets containing ^{6}Li, ^{7}Li, ^{12}C, ^{16}O and ^{18}O. At incident energies

around 200 MeV, very little is known about the behaviour of (p,t) cross sections across this target mass range. An investigation was therefore carried out to verify whether the distorted-wave Born approximation (DWBA) approach, which works moderately well at lower energies for these targets, can be extended to higher energies.

2. Experimental procedure

The measurements were performed with a 10-13 nA beam of 200 MeV protons provided by the separated sector cyclotron (SSC) accelerator at iThemba LABS. The beam impinged on various metallic oxide targets that were composed of different combinations ^{6}Li, ^{7}Li, ^{12}C, ^{16}O and ^{18}O. The decay products of the excited recoil nuclei were detected at angles close to 180 degrees by means of two 50 mm $\times$ 50 mm double sided silicon-strip detectors (DSSSD), each approximately 300 μm thick. The outgoing tritons emitted at zero degrees were identified in the K600 magnetic spectrometer focal plane detector system, which consists of vertical drift chambers (VDC). Each VDC has two wireplanes, a U- and X wireplane, both with a 4 mm wire pitch that allows the determination of horizontal and vertical positions and angles in the focal plane. Two plastic scintillators mounted behind the wire planes allowed particle identification and time of flight measurement. The beam was operated in the momentum dispersion matched mode and a typical energy resolution of about 45 keV was achieved. We have initially focused on precise measurement of ^{16}O energy spectrum using the K600 events as the master trigger. The segmented silicon detectors were used in slave mode in order to collect the decay products of different excited recoil nuclei. The measured triton energy spectrum is displayed in Fig.1. This spectrum was used to identify the different excited states of residual nuclei. The ^{14}O and ^{10}C excited states are strongly populated whereas ^{16}O excited states are weakly populated. However, the 15.1 MeV 0^+ is not clearly seen. This is due to the low grade ^{18}O enriched material used in the target manufacturing process and also to lower cross sections. Energy calibration was achieved by comparing the relative positions of known ^{10}C discrete states in the various targets. Absolute cross sections were extracted from measured yields, beam currents, target thicknesses, and defined spectrometer acceptance solid angle.

Charged particles from breakup nuclei were successfully detected in coincidence with the ejectile tritons. In Fig. 2, correlated structures are observed with the ^{14}O* $\rightarrow$p+^{13}N binary breakup identified in the two clear loci.

3. Preliminary results and outlook

Theoretical cross sections were calculated in the framework of a zero-range distorted-wave Born approximation (DWBA), which was successfully used in (p,t) reaction studies on ^{208}Pb and ^{116}Sn nuclei at an incident energy close to 200 MeV [8]. These calculations were performed by implementing the code DWUCK4 [9]. Microscopic form factors, built up from the single-neutron form factors generated in a standard Woods-Saxon potential, were used by the code. The Woods-Saxon potential was adjusted to give a binding energy equal to one half two-neutron separation energy.

The experimental (p,t) cross section $(d\sigma/d\Omega)_{exp}$ is related to that calculated by the DWUCK4 code $(d\sigma/d\Omega)_{DWUCK}$ by the the expression

$$(d\sigma/d\Omega)_{exp} = \frac{9.72 D_0^2 C^2 S}{2J+1} \left[(d\sigma/d\Omega)_{DWUCK}\right] \varepsilon \ ,$$

where C is the Clebsch-Gordan coefficient that accounts for the isospin coupling of the residual nucleus to that of the transferred neutron pair to yield the isospin of the target nucleus. For the (p,t) reaction, C^2 is equal to unity. The S factor represents the two-neutron spectroscopic factor and is set to $(2J+1)$ for neutron closed-shell nuclei. For partially filled subshells, the S factor is equal to $(2J+1)V_{j1}^2 V_{j2}^2$, where V_{j1}^2 and V_{j2}^2 are the occupation probabilities for

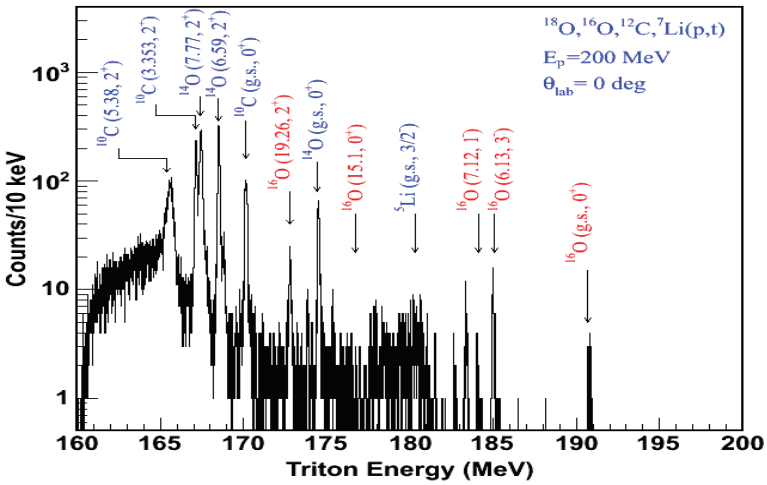


Figure 1. Measured triton energy spectrum at E_{lab}=200 MeV and θ=0°. The excitation energies of the levels together with their corresponding spins and parities for the different residual nuclei are labelled. The expected position of the 15 MeV 0^+ state is also indicated.

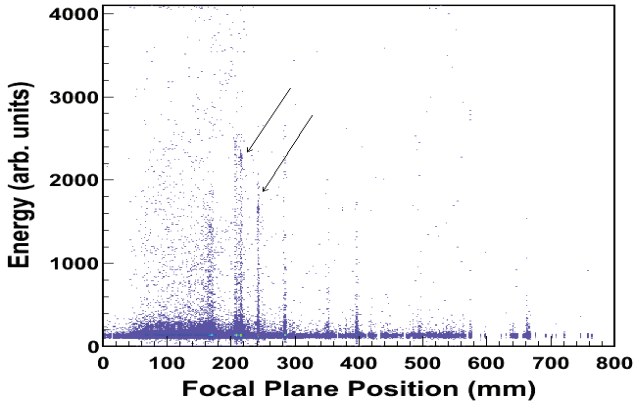


Figure 2. A two-dimensional plot showing correlation of the light charged particles with the tritons from excited states in ^{14}O. The arrows indicate the $^{14}O^* \rightarrow p+^{13}N$ binary breakup loci.

each neutron (j_1, j_2). These two quantities are taken from one-neutron pickup experiments [10,11,12] and shell-model calculations [13]. The ε parameter is called the enhancement factor which characterizes the ratio of experimental to DWBA cross sections. It measures the relative strenghts of various (p,t) transitions. D_0^2 is the zero-range normalization constant related to the normalization of the triton wave function. The value of 6.6 in 10^4 fm^3 MeV^2 units was adopted as in Ref. [8]. Global proton and triton optical model potential parameters were used in DWBA analysis [14,15]. The experimental cross sections along with the theoretical cross sections and the enhancement factors are listed in Table 1.

The enhancement factors (ε) extracted from data are typically less than the ideal value of unity. DWBA calculations consistently overpredict the magnitude of the cross sections. This discrepancy can be ascribed to a bad choice of optical potential parameters or to inconsistent spectroscopic factors. The finite-range effect that arises from an increase of the momentum

Table 1. Extracted cross section values of strongly populated states compared with theoretical predictions. The errors on the extracted cross section values are based primarily on statistics while the uncertainty on the target thicknesses was estimated to be less than 15 %.

Levels of ^{16}O						
$E_x(MeV)$	J^π	$\sigma_{exp}(\mu b/sr)$	$\sigma_{DWBA}(\mu b/sr)$	C^2S	ε	Pickup configuration
0	0^+	1.6±0.6	54.18	0.68	0.01	$(1d_{5/2})^2$
6.13	3^-	6.5±1.2	10.59	0.41	0.61	$(1p_{1/2} \times 1d_{5/2})$
7.12	1^-	1.8±0.7	10.20	0.44	0.18	$(1p_{1/2} \times 2s_{1/2})$
19.26	2^+	14.0±1.7	41.28	0.25	0.34	$(1d_{5/2})^2$
Levels of ^{14}O						
$E_x(MeV)$	J^π	$\sigma_{exp}(\mu b/sr)$	$\sigma_{DWBA}(\mu b/sr)$	C^2S	ε	Pickup configuration
0	0^+	3.0±0.3	26.14	1	0.11	$(1p_{1/2})^2$
6.59	2^+	15.9±0.7	64.88	1	0.25	$(1p_{3/2} \times 1p_{1/2})$
7.77	2^+	22.7±0.9	75.62	1	0.30	$(1p_{3/2} \times 1p_{1/2})$
Levels of ^{10}C						
$E_x(MeV)$	J^π	$\sigma_{exp}(\mu b/sr)$	$\sigma_{DWBA}(\mu b/sr)$	C^2S	ε	Pickup configuration
0	0^+	17.1±0.9	64.70	0.39	0.26	$(1p_{3/2})^2$
3.353	2^+	28.2±1.1	169.7	1.95	0.17	$(1p_{3/2})^2$
5.38	2^+	41.9±1.3	205.6	1.95	0.20	$(1p_{3/2})^2$

transfer to the triton at higher bombarding energies may also be of importance. Further theoretical investigation remains the object of a future study.

This collaboration will, within a few months, repeat the measurement with a better ^{18}O target, i.e. made from higher grade ^{18}O enriched material. As a consequence of the improved target and the allocation of more beam time, we will be able to collect more statistics in the excitation energy region of interest. This will hopefully allow the measurement of breakup particles emanating from the 0^+_6 state in ^{16}O.

References

[1] Cook C W, Fowler W A, Lauritsen C C and Lauritsen T 1957 *Phys. Rev.* **107** 508.
[2] Navrátil J P, Vary J P and Barrett B R 2000 *Phys. Rev. Lett.* **84** 5728.
[3] Schuck P, Funaki Y, Yamada T, Horiuchi H, Rpke G and Tohsaki A 2009 *AIP Conf. Proc.* **1165** 39.
[4] Ohkubo S and Hirabayashi Y 2010 *Phys. Lett. B* **684** 127.
[5] Funaki Y *et al.* 2009 *Phys. Rev. C* **80** 064326.
[6] Wiringa R B *et al.* 2000 *Phys. Rev. C* **62** 014001.
[7] Funaki Y *et al.* 2008 *Journal of Physics: Conference Series* **111** 012012.
[8] Gerlic E *et al.* 1989 *Phys. Rev. C* **39** 2190.
[9] Kunz P D 1978 *Computer code DWUCK4, University of Colorado.*
[10] Roos P G *et al.* 1975 *Nucl. Phys. A* **255** 187.
[11] Mairle G *et al.* 1977 *Nucl. Phys. A* **280** 97.
[12] Pignanelli M *et al.* 1973 *Phys. Rev. C* **8** 2120.
[13] Zuker A P, Buck B and McGrory J B 1968 *Phys. Rev. Lett.* **21** 39.
[14] Koning A J and Delaroche J P 2003 *Nucl. Phys. A* **713** 231.
[15] Pang D Y *et al.* 2009 *Phys. Rev. C* **79** 024615.

Study of rare modes of "Collinear Cluster Tri-Partition" of ^{252}Cf(sf)

V D Malaza[1], N M Jacobs[1], D V Kamanin[2], Yu V Pyatkov[2, 3], A A Aleksandrov[2], I A Aleksandrova[2], N A Kondratyev[2], E A Kuznetsova[2], O V Strekalovsky[2], A O Strekalovsky[2] and VE Zhuchko[2]

[1] Stellenbosch University, Faculty of Military Science, Military Academy, Saldanha 7395, South Africa
[2] Joint Institute for Nuclear Research, 141980 Dubna, Russia
[3] National Nuclear Research University MEPHI, 115409, Moscow, Russia

E-mail: a_mala@ma2.sun.ac.za

Abstract. In our previous experiments we have observed multiple manifestation of a new ternary decay of low excited nuclei called “Collinear Cluster Tri-partition” (CCT). A special mode of this new type of ternary decay was observed based on the double magic ^{132}Sn cluster. A pre-scission configuration which brings a question whether the observed double magic ^{132}Sn can be replaced by double magic ^{208}Pb is being studied. A specially designed experimental setup has been put in place to give better statistics and more precise time-of-flight measurements. An improved calibration procedure is presented.

1. Introduction

In our previous experiments [1-4] we identified multiple manifestations of a new ternary decay of low excited nuclei called "Collinear Cluster Tri-partition" (CCT) due to the features of the process observed. The unusual decay channel was revealed both in the framework of "missing mass" method, meaning that only two decay partners were detected, the other one was missing [1,2]. The third partner is determined by subtracting the sum of the observed two masses from the total mass of ^{252}Cf. In this way all three decay partners are accounted for [3,4]. Recently a specific CCT mode was observed based on the double magic ^{132}Sn cluster. The mass-mass distribution of the events selected by velocities and energies is shown in figure 1.

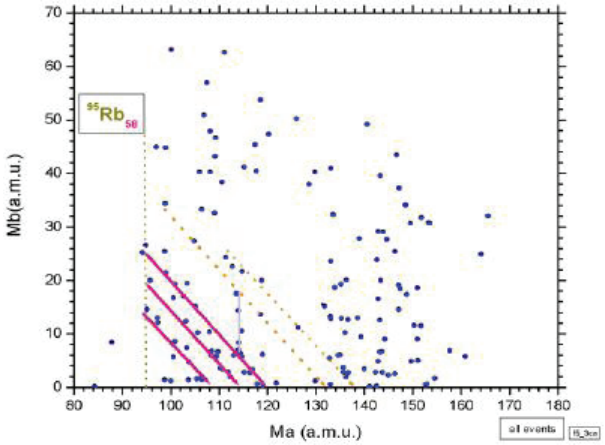


Figure 1. Mass-mass distribution of fragments selected by velocities and energies

The tilted red lines in the figure correspond to the missing magic clusters of ^{132}Sn and ^{144}Ba. This missing cluster can be clearly seen in the mass spectrum in figure 2 which is the projection along these lines.

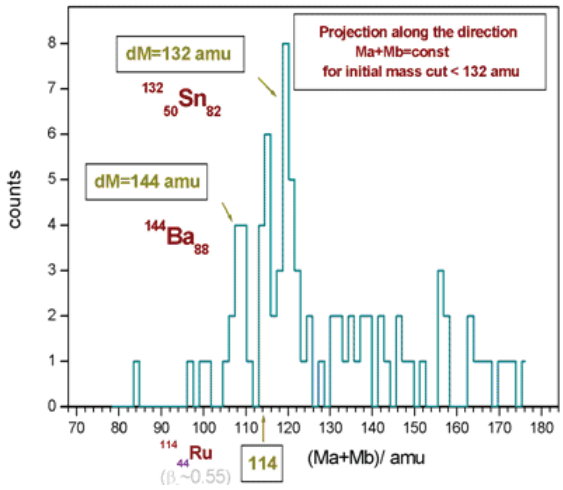


Figure 2. Mass spectrum for the structures marked by red lines in figure 1.

Pre-scission configuration which presumably gives rise to the mode under discussion is shown in figure 3. As we can see from the figure, the Sn cluster can "move" as a whole along the cylinder-like configuration that consists of residual nucleons. Two light fragments accompanying this cluster marked by M1 and M2 were actually detected in previous experiments. The value of M2 lies between 0 amu and the difference between the initial mass of ^{252}Cf and the detected fragments. M1 cannot assume any value less than 95amu (deformed magic ^{95}Rb) due to the fact that experimentally the Sn cluster is always observed and the detected mass of the third fragment suggest that the lowest value of M1 should be 95amu.

2. Motivation

The main aim of our study is to investigate the pre-scission configuration which is almost collinear. The question that arises is whether ^{132}Sn can be replaced by double magic ^{208}Pb in this mode. Theoretical indications of such a mode were obtained in [5] and are illustrated in figure 4 which

shows the potential energy of the fissioning nucleus of ^{252}Cf corresponding to the bottoms of the potential valleys as a function of parameter Q proportional to the quadrupole moment of the system.

The panels depict the shapes of the fissioning system at the points marked by arrows. Both in valley of mass-asymmetrical shapes (4) and mass-symmetrical shapes (5) the system consist of pairs of magic clusters (Sn/Ni, Sn/Ge and so on) and nucleons left over forming a "neck" between the clusters [6]. Valley 1 in figure 4 is due to the preformation of double magic ^{208}Pb. Evolution of the nuclear shape in this mode is presented above the figure.

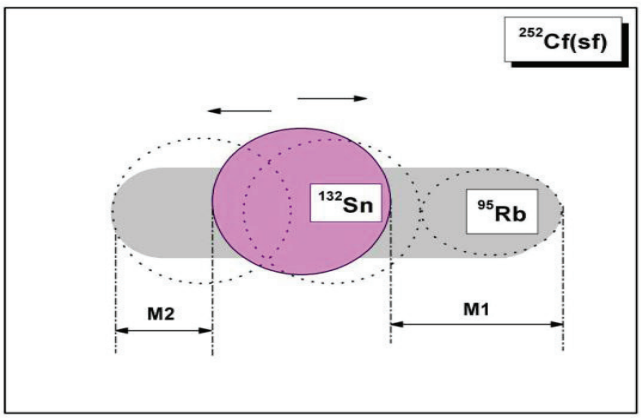


Figure 3. Schematic pre-scission configuration of the CCT mode based on the double magic ^{132}Sn cluster.

Coming back to the question whether ^{132}Sn can be replaced by double magic ^{208}Pb, if so, this will lead to a new type of lead radioactivity. Searching for such a mode is one of the goals of our forthcoming experiment, which will require better statistics and more precise time-of-flight measurements.

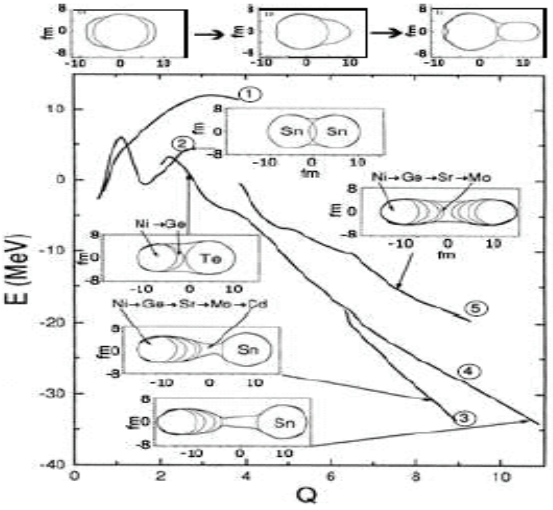


Figure 4. The bottoms of the fission valleys as a function of parameter Q (proportional to the quadruple moment) for ^{252}Cf nucleus.

3. Experimental problems

As we can see in figure 1 the masses of fragments defining the modes under investigation differ radically that is, one is very light while the second one is very heavy. Therefore we have a problem involving the method of measuring the correct energy and time-of-flight of heavy ions in the wide

range of energies and masses using PIN diodes as “stop” detectors (as it is a case from our previous experiments).

The problem we have is due to the negative influence of the known “plasma delay” in registering time-of-flight of fission fragments and the “pulse-height defect” (PHD) in registering energy of fission fragments using semiconductor detectors. In order to exclude the influence of “plasma delay” for timing of the fragments, three micro-channel plate (MCP) timing detectors shown in figure 5 will be used. The first MCP will be used to deliver a start signal and the other two MCPs will be used for stop signal in both arms. This setup will enable us to measure the influence of plasma delay by analyzing the difference between the time signal from MCP detectors and PIN diodes.

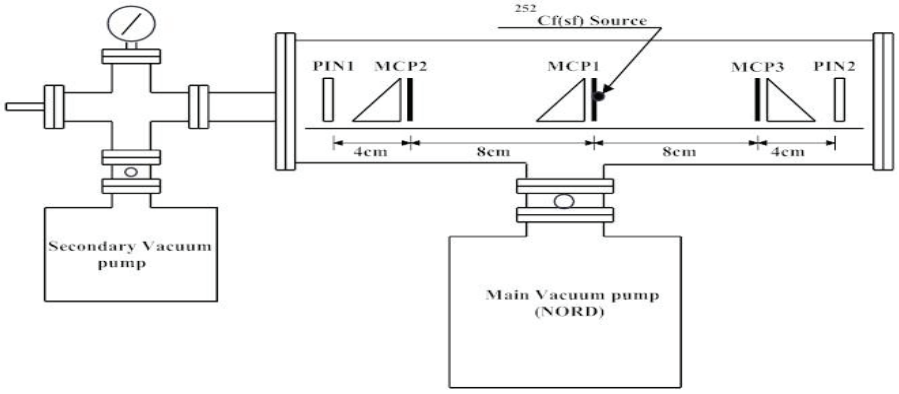


Figure 5. Schematic view of our setup.

4. Calibration and reconstruction of FF masses

The calibration process comprises of a selection of raw data and performs a direct transformation of energy in channels to energy in Mega-electron Volts. This first step of processing is referred to as “a first approximation” because in this approach the PHD is only approximated roughly. In the second step we perform a so called "true calibration" where we take the PHD into account and reconstruct the masses of the fission fragments.

4.1. “The First Approximation” in the FF mass reconstruction

In “first approximation” the energy in channels *E[ch]* is converted to the energy in MeV, *E[MeV]* according to the following equation:

$$E_i[\mathrm{MeV}] = C \cdot \exp(-\frac{E_i[\mathrm{ch}]}{D}) + E_0 \qquad (1)$$

The values of *C, D,* and E_0 are determined by using the known positions for the energy peak of light and heavy fragment and the natural alpha peak from ^{252}Cf with Eα = 6.118MeV. The subscript *i* in equation (1) shows that each event is processed individually. The time in channels *T[ch]* is converted according to the following equation to the time in nanoseconds *T[ns]*:

$$T_i[\mathrm{ns}] = A \cdot T_i[ch] + B \qquad (2)$$

The values of *A* and *B* are determined by using the known velocities $V_{L,H}^{ref}$ of light and heavy fragment from literature and the flight path of fission fragments. Once the values of *A* and *B* are obtained we apply equation (2) to our raw data to calculate time in nanoseconds and we use the time to calculate the velocity in centimeters per nanoseconds as follows:

$$V_i[cm/ns] = \frac{L_{TOF}}{T_i[ns]} \tag{3}$$

Equation (1) and (3) allows us to calculate the mass of the FF as follows:

$$M_i[amu] = \frac{1.9297 E_i[MeV]}{\left(V_i[cm/ns]\right)^2} \tag{4}$$

4.2. True calibration and reconstruction of the FF masses

The true calibration and reconstruction of FF masses is quiet a complicated task due to the influence of pulse-height defect (PHD). The channel number of energy in which we register the fission fragment depends on the energy of the fission fragment as well as on the PHD. But on the other hand, the PHD depends on the mass and the kinetic energy of the registered fragment. To combine together the calculation of true energy and reconstruction of fission fragments masses we use a specially designed procedure presented in [6].

The energy E in MeV, of the registered fission fragment is defined as the sum of the detected energy E_{det} and the pulse-height defect denoted by *R(M,E)*:

$$E = E_{\text{det}} + R(M, E) \tag{5}$$

Where the detected energy of fission fragments is given by:

$$E_{\text{det}}[\text{MeV}] = E[\text{ch}] \cdot E_{gr} + E_0 \tag{6}$$

Where E_{gr} and E_0 are calibration parameters calculated experimentally by using a high precision pulse generator and the natural alphas from ^{252}Cf source. The expression for the pulse-height defect in equation (5) was proposed by *Mulgin* and his colleagues [7] as given by the following empirical expression:

$$R(M, E) = \frac{\lambda \cdot \text{E}}{1 + \varphi \cdot \dfrac{E}{M^2}} + \alpha \cdot ME + \beta \cdot E \tag{7}$$

Where $\{\lambda, \varphi, \alpha, \beta\}$ are parameters for the true calibration. In addition we know that:

$$E = \frac{M \cdot V^2}{1.9297} \tag{8}$$

Where E is the energy of the FF in *MeV*, M is the mass of the FF in *amu* and V is the velocity of the FF in *cm/ns*. The velocity for this purpose is calculated using the parameters obtained from time calibration. From the above equations, (5, 6, 7 and 8) we can calculate the mass of the fission fragment provided the parameters $\{\lambda, \varphi, \alpha, \beta\}$ are known. It is worth noting that the numerical values for the parameters $\{\lambda, \varphi, \alpha, \beta\}$ proposed in [7] make it impossible to reconstruct the mass M_{TE} for the FF. In order to find the correct values of the parameters $\{\lambda, \varphi, \alpha, \beta\}$ a special iterative procedure has been designed. This procedure consists of obtaining a solution of the following equation analytically:

$$G(\{\lambda, \varphi, \alpha, \beta\}, M) = 0 \tag{9}$$

Where G depends on the parameters $\{\lambda, \varphi, \alpha, \beta\}$ and the mass. To obtain the solution of equation (9) above, we combine equation (5), (7), and (8) as follows:

$$\frac{MV^2}{k} = E_{\text{det}} + \frac{\lambda \cdot \frac{MV^2}{k}}{1 + \varphi \cdot \frac{V^2}{Mk}} + \alpha \cdot \frac{M^2V^2}{k} + \beta \cdot \frac{MV^2}{k} \tag{10}$$

Where $k = 1.9297$ and we obtain the following third order equation:

$$M^3 + aM^2 + bM + c = 0 \tag{11}$$

Where

$$a = \frac{\varphi V^2}{k} + \frac{\beta + \lambda - 1}{\alpha}, b = \frac{kE_{\text{det}}}{\alpha V^2} + \frac{\varphi V^2}{\alpha k}(\beta - 1), c = \frac{\varphi E_{\text{det}}}{\alpha}$$

Using the above procedure we process each event individually based on the current values of $\{\lambda, \varphi, \alpha, \beta\}$ and select the mass of the fission fragment from the calculated values in equation (11). The mass is selected from the condition $M_{TE} \in [1\text{ amu}, 252\text{ amu}]$. After processing an amount of data a mass spectrum is obtained.

The procedure uses the MINUIT package to minimize the following criterion function by changing the parameters $\{\lambda, \varphi, \alpha, \beta\}$:

$$F = [(< ML_T > - < ML >)^2 + (< MH_T > - < MH >)^2] + \mu \sum_{M_{TE}} \frac{(Y(M_{TE}) - Y_T(M_{TE}))^2}{Y(M_{TE})} \tag{12}$$

Where μ is a free parameter that is chosen by the user and it is used as an input parameter to the MUNUIT minimization procedure. This parameter plays a role of specific relative weight of the second term in the criterion function F. The values $< ML >$ and $< MH >$ are average masses of light and heavy fragments calculated from the experimental mass spectrum $Y(M_{TE})$. In the above equation the known values from literature are denote by "T".

5. Conclusion

In our previous experiments the "first approximation" approach presented above was used for the purpose of energy calibration and reconstruction of fission fragment masses. The "true calibration" is the improved version for the calculation of the fragment mass, and will be used for the calibration and reconstruction of the fission fragment masses in our current experiment. It is worth noting that the code for calibration is based on the iterative procedure using the MINUIT package to minimize equation (12).

References

[1] Pyatkov Yu V et al. 2007 *Romanian Reports in Physics* **59** p 388
[2] Kamanin D V, Pyatkov Yu V, Tyukavkin A N and Kopatch Yu N 2008 *Int. Journal of Modern Physics* V **17** p 2250
[3] Pyatkov Yu V et al. 2010 *European Physics Journal* A**45** p 29
[4] Kamanin DV et al. 2010 *Physics of Particles and Nuclei Letters* **7** p 121
[5] Pashkevich V et al. 2009 *Int. Journal of Modern Physics* E **18** p 907
[6] Pyatkov Yu V et al. 2007 *Proc. ISINN-14* (Dubna, May 24-27, 2006) p 134
[7] Mulgin S et al. 1997 *NIM* A 388 p 254

Radon escape from mine tailings dams

J N Ongori[1], R Lindsay[1], R T Newman[2] and P P Maleka[2]

[1]Department of Physics, University of the Western Cape Private Bag X17 7537 Bellville South Africa
[2]Environmental Radioactivity Laboratory, Physics Group, iThemba Laboratory for Accelerator Based Sciences (LABS), P. O. Box 722, 7129 Somerset West South Africa

E-Mail: rlindsay@uwc.ac.za

Abstract.
Gold mine tailings dumps contain considerable amounts of ^{226}Ra and have therefore been identified as large sources of radon. Gamma ray measurements at the field (Kloof mine dump) were done using the MEDUSA detector system while an HPGe was used for laboratory based measurements. Field activity concentrations were analyzed using Full Spectrum Analysis. Thirty four samples were collected from the field, analyzed and were found to contain an average of 308 ±12 Bq/Kg, 255±12 Bq/Kg and 18 ±1 Bq/Kg for ^{238}U, ^{40}K and ^{232}Th respectively. The link between MEDUSA and HPGE was established through normalisation factors and after normalising the activity concentrations and correcting for moisture content, the normalised radon flux was computed to be 0.105±0.015 Bq m^{-2} s^{-1}.

1. Introduction

In the Gauteng province, South Africa where mining activities take place, gold is extracted from rocks which are crushed and processed. After extraction, the waste is disposed in tailing dumps. The tailings usually contain naturally radioactive elements mainly uranium (^{238}U), potassium (^{40}K) and thorium (^{232}Th). ^{238}U decays into several products (see Figure 1) and among them is the radioactive gas known as radon (^{222}Rn).

Radon further decays to form short lived radioactive particles which enter the body through inhalation and can adhere to the lining of the lung. The deposited particles decay emitting alpha radiation which has a potential to damage cells in the lung [1]. Therefore radon exhalation from these tailing dumps poses a health concern.

The discussion that follows highlights a novel technique [7] for mapping radon exhalation from Kloof mine dump near Carletonville by using the MEDUSA (**M**ulti **E**lement **D**etector for **U**nderwater **S**ediment **A**ctivity) γ-ray detection system [4] in the field in conjunction with a laboratory based hyper-pure germanium (HPGe) detector. The dump is approximately 2 square kilometres and it has been inactive for over seven years.

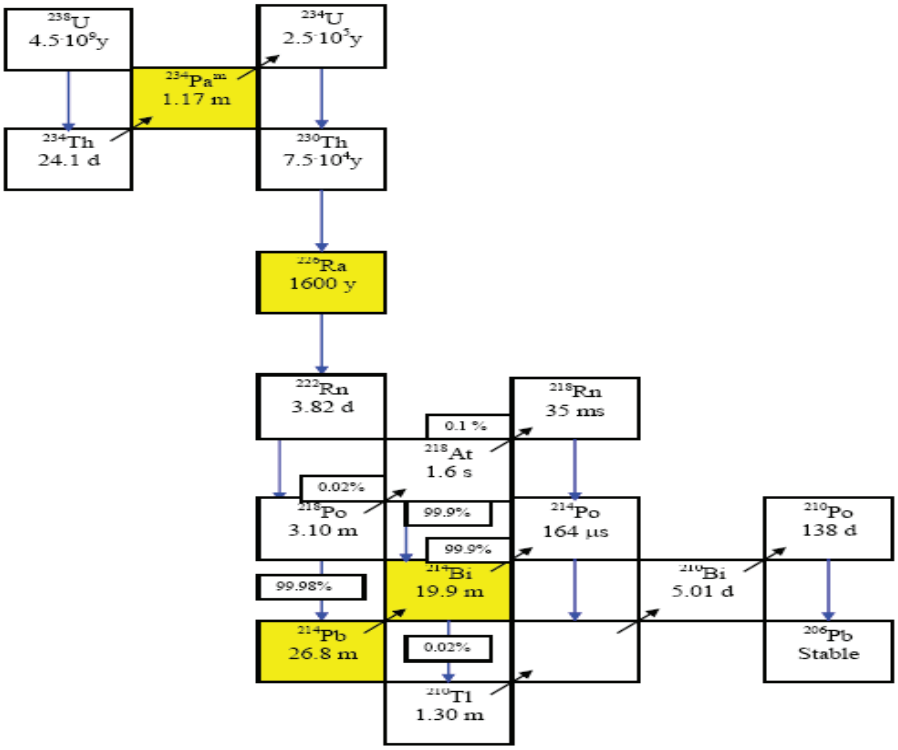


Figure 1. A schematic illustration of the uranium decay series. The half-life of each radionuclide in the series is indicated in years (y), days (d), minutes (m) and seconds (s). The yellow boxes represents the γ-ray emitters. The vertical arrow (blue) indicate α-decay while upward, slanted arrows (black) indicate β-decay.

2. Radon exhalation

Radon generated from the decay of radium in mineral grains migrates and is released to the atmosphere. Radon exhalation studies have been conducted and among the methods that have been used include the flow through method, the can technique and the accumulator [6]:

Unlike the mentioned techniques above MEDUSA technology which is based on gamma ray spectrometry will be used to map the radon exhalation from Kloof mine dump. This technology was developed in the Netherlands by the Nuclear Geophysics Division of Groningen. The MEDUSA system incorporates a gamma-ray detector (CsI (Na)), ALADIN box, MPA (MEDUSA Post Analysis) software tool and global positioning system (GPS). The CsI (Na) crystal with 70 mm diameter and 150 mm length was integrated into the system mainly because of its stability in light output at variable temperatures. The ALADIN box contains the data acquisition system and MPA software tool for gamma-ray spectra analysis. The gamma-ray intensities are interfaced on the surveyed area from the GPS.

Previously, the system has been used to map sand/mud ratios in rivers and dispersion of dredge spoil dumped on the sea floor [3; 9]. Furthermore, the system has been used in doing airborne surveys at iThemba LABS [5] and also at a mine dam [7]. The detector system was mounted 0.55 m off the ground on the front of a 4× 4 vehicle and accessible parts of the dam were transversed at 2 m/s (~7.2 km/h) with the spectra recorded every 2 s as also the GPS location logged-in for the given data.

A map with radiometric data was used to select locations where MEDUSA stationary measurements were done for 30-60 minutes and thereafter five samples were collected, one sample

just beneath the detector and four more samples roughly 80 cm from the detector at north, south, east and west to be analyzed using the HPGe detector.
The acquired MEDUSA spectra were analyzed using full spectrum analysis [4] to extract radiometric information such as total counts (Figure 2a) and then activity concentrations of ^{238}U, ^{232}Th and ^{40}K (Figure 3a). These maps indicate accessible parts of the dam which are in colour while the inaccessible parts are in grey and beside them are their corresponding interpolated maps using Golden Software Surfer 8.

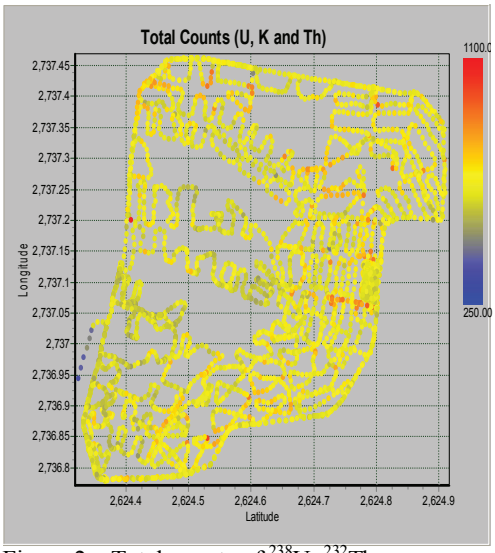


Figure 2a. Total counts of ^{238}U, ^{232}Th and ^{40}K. Inaccessible parts on the map indicated by grey

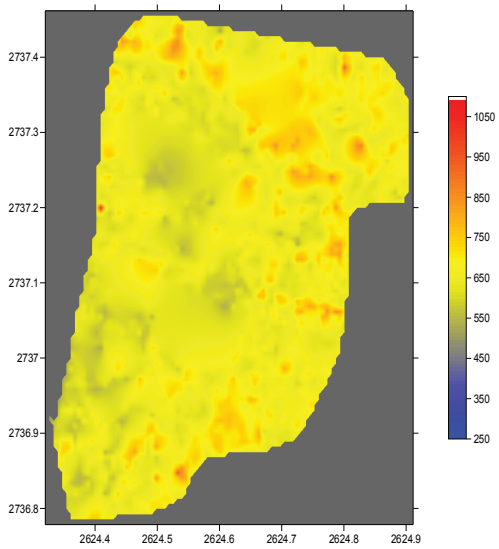


Figure 2b. Interpolated total counts of ^{238}U, ^{232}Th and ^{40}K using Golden Software Surfer 8

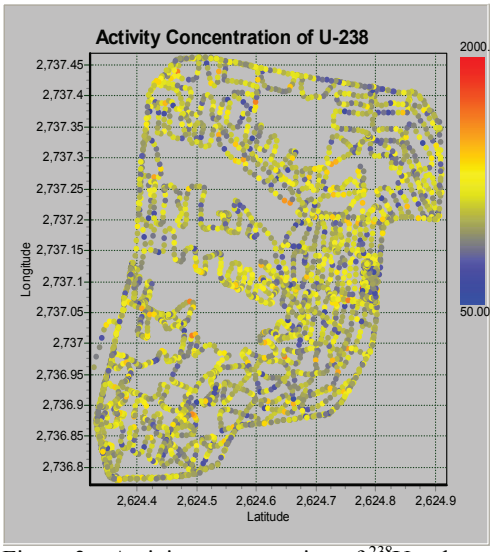


Figure 3a. Activity concentration of ^{238}U only after Full Spectrum Analysis

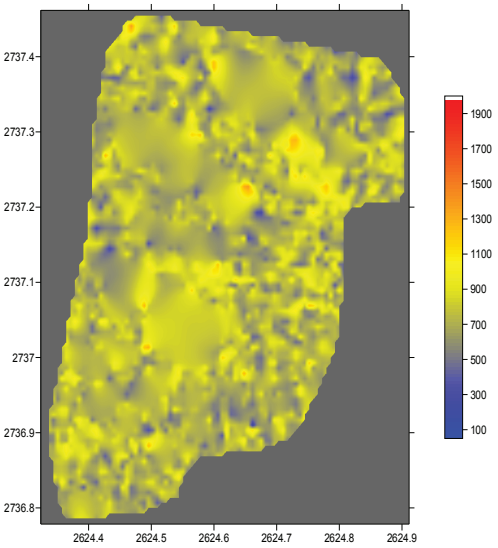


Figure 3b. The corresponding interpolated activity concentration of ^{238}U only using Golden Software Surfer 8

Thirty four samples collected from the mine dump were further analyzed using the laboratory based HPGe detector. This detector is a Canberra GC4520, p-type detector with 45% relative efficiency and 2 keV FWHM resolution at 1.33 MeV [8]. The detector is encased in a 10 cm thick lead castle fitted with a 2.0 mm thick copper inner lining in order to reduce the background in the sample spectra. Standard nuclear electronics are used to process the detector signals. The pulses from the amplifier are collected and sorted by the ATOMKI Palmtop software multi-channel analyzer installed on a desktop PC. The HPGe detector system is energy and efficiency calibrated regularly with certified ^{232}Th, ^{238}U, ^{40}K ore (International Atomic Energy Agency/RGTh-1, RGU-1, RGK-1, respectively) and a liquid cocktail reference source (^{60}Co, ^{137}Cs and ^{152}Eu) from the National Metrology Institute of South Africa.

The soil samples were dried, sieved, sealed in Marinelli beakers and then stored for three weeks before counting. Radionuclides present in soil samples were identified and activity concentration (A) calculated using the following expression;

$$A(Bq/kg) = \frac{counts}{Br \cdot \varepsilon \cdot t \cdot m} \quad (1)$$

where counts represent the full-energy peak content corrected for continuum contribution, Br is the branching ratio (that is the percentage of the decay of the nuclide that will proceed via the emission of a particular gamma-ray), ε the detection efficiency, t the live time (in s) for spectrum acquisition and m (in kg) the sample mass. The counts for a given full-energy peak from the spectrum is determined by manually setting a region of interest (ROI) around it. The ATOMKI Palmtop MCA software uses an algorithm to calculate the net area (counts) associated with the ROI.

The average activity concentrations for the thirty four samples were determined to be: 308 ±12 Bq/Kg, 255±12 Bq/Kg and 18 ±1 Bq/Kg for ^{238}U, ^{40}K and ^{232}Th respectively.

In order to work out radon exhalation from the tailing dump a link between field based MEDUSA and the laboratory based HPGe was established. The normalisation factor between MEDUSA data and HPGe results was calculated using this expression:

$$N_f = \frac{AC_{HPGe}}{AC_{MEDUSA}} \quad (2)$$

where N_f represent the normalisation factor, AC_{HPGe} is the activity concentration measured by the HPGe and AC_{MEDUSA} is the activity concentration measured by the MEDUSA γ-ray detector system.

MEDUSA activity concentrations in equation 2 refers to the concentration extracted from measurements done while the detector was stationary for a period of 30-60 minutes, whereas the HPGe concentration refers to the average concentration of the five samples picked beneath and around the MEDUSA detector after the measurement was completed. These stationary measurements were done at four different spots spanning the mine dam and the average normalisation factors were 0.54±0.05, 0.36±0.03 and 0.21±0.01 for ^{238}U, ^{40}K and ^{232}Th respectively. The normalisation factors were specifically utilized to convert MEDUSA concentration into absolute concentrations.

Since ^{40}K and ^{232}Th activity concentrations are not or hardly affected by radon exhalation, their concentrations will be independent of the radon loss and therefore the ratio of ^{238}U and ^{40}K and/or ^{232}Th is a measure of the radon exhalation at a particular location which is determined using the following expressions [7].

$$F = 1 - \left[\frac{(^{238}U_{Field} / ^{40}K_{Field})}{(^{238}U_{Lab} / ^{40}K_{Lab})} \right] \quad (3)$$

where, F is the radon flux, $^{238}U_{Field}$ and $^{40}K_{Field}$ represents the activity concentrations of uranium and potassium respectively, which are measured using MEDUSA in the field, $^{238}U_{Lab}$ and $^{40}K_{Lab}$ represents the activity concentrations of uranium and potassium respectively, which are measured using the HPGe in the laboratory.
In a similar way thorium activity concentrations in combination with uranium activity concentrations were used to extract the radon flux as shown in the expression below:

$$F = 1 - \left[\frac{(^{238}U_{Field} / ^{232}Th_{Field})}{(^{238}U_{Lab} / ^{232}Th_{Lab})} \right] \quad (4)$$

where, $^{232}Th_{Field}$ and $^{232}Th_{Lab}$ represents the activity concentrations of thorium, which are measured using MEDUSA at the field and the HPGe in the laboratory respectively.

Finally the flux obtained by equation 3 and 4 is normalised and the average flux computed for Kloof mine dump was 0.105±0.015 Bq m^{-2} s^{-1}. Figure 4 shows a normalised flux which was computed and for parts which were inaccessible interpolation was done using Golden Software Surfer 8. The computed range of flux was 0.02-0.26 Bq m^{-2} s^{-1}.

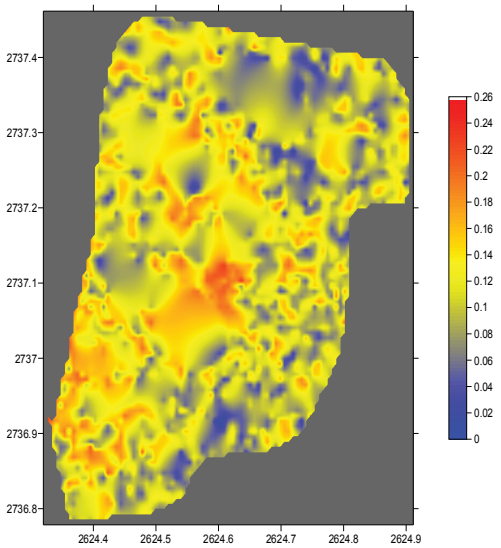


Figure 4 Map of interpolated flux for Kloof mine dump

3. Conclusion

A gamma ray technique based on the MEDUSA detector technology has been used to determine the radon exhalation from a mine dump. The analysis demonstrates a novel way which has been practically applied to determine radon exhalation from a large area.

References

[1] BEIR (VI) 1999 http://www.nap.edu/catalog/5499.html

[2] De Groot A V, Van der Graaf E R, De Meijer R J and Maucec M. 2009 *Nuclear Instruments and Methods in Physics Research A* **600** 519-523

[3] De Meijer R J, 1998 *Journal of Geochemical Exploration*, **62** 81-103

[4] Hendriks P H G M, Limburg J and De Meijer R J 2001 *Journal of Environmental Radioactivity* **53** 365-380

[5] Hlatshwayo I N, Lindsay R, Ndwandwe O M and Newman R T 2009 *Radioprotection*, vol. **44** 825–830

[6] International Atomic Energy Agency 1992 *Technical Report* Series No. **333**

[7] Lindsay R, De Meijer R J, Maleka P P, Newman R T, Motlhabane T G K and De Villiers D 2004 *Nuclear Instruments and Methods in Physics Research B* **213** 775-778

[8] Newman R T, Lindsay R, Maphoto K P, Mlwilo N A, Mohanty A K, Roux D G, De Meijer R J and Hlatshwayo I N 2008 *Applied Radiation and Isotopes* **66** 677–1066

[9] Venema L B and De Meijer R J 2001 *Journal of Environmental Radioactivity* **55** 221-239

Division C – Lasers, Optics and Spectroscopy

Cross phase modulation induced depolarization of a probe signal and its impact on polarization mode dispersion compensators

R R G Gamatham, T B Gibbon and A W R Leitch
Physics Department, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa.

E-mail: romeo.gamatham@live.nmmu.ac.za

Abstract. Polarization mode dispersion continues to pose a threat to high speed optical networks threatening to corrupt the transmitted data. One solution is to deploy polarization mode dispersion compensators (PMDCs). This article considers the composite problem of polarization mode dispersion (PMD) and cross phase modulation (XPM) in wavelength division multiplexing (WDM) networks. The degree of polarization (DOP) is a prominent indicator of the level of PMD within an optical fibre link to PMDCs, however in a two channel wavelength division multiplexing (WDM) system XPM may degrade the DOP within the link to a notable extent misleading PMDCs. Potentially leading to a worsening of the PMD within the fibre link. This article presents experimental results showing the degradation of the average DOP of a probe signal in the presence of a high intensity pump signal modulated at a bit rate of 10.3 Gbps. The continuous wave probe signal reaches a minimum DOP of 0.61 for a power ratio of 22 dB between pump and probe.

1. Introduction

Nonlinear effects are easily observed in single mode optical fibres because of the fibre small spot size and extreme low loss [1]. At high optical intensities in fibres the refractive index becomes a function of the intensity of the optical signal. This is known as the Kerr nonlinearity. This leads to the optical Kerr effect where the nonlinear phase shift induced by an intense high power pump changes the characteristics of the probe beam [1]. Cross phase modulation (XPM) refers to a nonlinear effect where a pump beam with high intensity changes the phase of a low power probe beam. Polarization mode dispersion broadens the input pulse which will lead to bit errors in the transmitted data. A potential solution is to insert a polarization mode dispersion compensator in the fibre link. Real time PMD compensation (PMDC) techniques require fast monitoring of a PMD related parameter from which the PMD can be determined [2]. Polarization mode dispersion compensators (PMDCs) monitor the link PMD in an indirect manner where a popular monitoring technique tracks the degree of polarization (DOP) of a signal in the link [2]. Popular PMDC techniques are the feed-back and feed-forward PMDC techniques. Both these techniques make use of the DOP to track the current PMD value of the link. In [3] Collins and Boivin showed the rotation of a continuous wave pump and continuous wave probe signal around a common pivot. In 2004 Lin and Agrawal introduced a model of XPM for the case of a pump and a probe. The model is used to look at the temporal and spectral polarization effects occurring inside an optical fiber. In 2003 Bononi *et al.* presented a novel analytical

model of XPM for a pump and probe system [5]. It is pointed out that the rotation angle is no longer constant in time when one of the signals is intensity modulated [5].

2. Theoretical background

This article considers theoretical models from [5] where the Manakov equation is used to describe the nonlinear evolution of the optical waves. The propagation equations of the pump and probe wave propagating along the link can be written in Stokes space as

$$\frac{\partial \vec{s}(z,t)}{\partial z} = \frac{8}{9}\gamma e^{-\alpha z}\left(\vec{p}\left(z,t-d_{sp}z\right)\times \vec{s}(z,t)\right) \tag{1}$$

and

$$\frac{\partial \vec{p}(z,t)}{\partial z} = \frac{8}{9}\gamma e^{-\alpha z}\left(\vec{s}(z,t)\times \vec{p}\left(z,t-d_{sp}z\right)\right), \tag{2}$$

where vectors s and p represent the probe and pump waves respectively, γ the nonlinear coefficient, z and t the propagation length and time, α the attenuation and d_{sp} the walk-off. The walk-off is a product of chromatic dispersion (CD) and the wavelength spacing. From equation (1) and (2) it is evident that the pump and probe waves modify each other over the propagation distance. It can also be concluded that the rate of change of the probe and pump Stokes vectors over distance is power dependent and depends on their relative orientation in Stokes space as well.

Bononi *et al.* present a Carousel model in [5] which takes finite walk-off and an intensity modulated pump signal into account [5]. In the case of continuous waves the rotation angle is time independent and becomes a function of the pivot power, sum of pump and probe, and fibre length [5]. The Carousel model presents the time dependent rotation angle as

$$\Psi(z,t) = \frac{8}{9}\gamma P_m \int_0^z e^{-\alpha z'} p(t-d_{sp}z')dz' \text{ ,} \tag{3}$$

where P_m is the pivot power [5]. Finally the DOP is related to the swing of the rotation angle expressed as

$$DOP = \sqrt{1-\sin^2\theta_s\left[1-\left\langle\cos\Delta\Psi(t)\right\rangle^2-\left\langle\sin\Delta\Psi(t)\right\rangle^2\right]} \tag{4}$$

which simplifies to

$$DOP = \sqrt{1-\sin^2\theta_s\left[1-\left(\left\langle\cos\Delta\Psi(t)\right\rangle\right)^2\right]} \tag{5}$$

for a periodically modulated pump. The relative angle between the pump and probe is denoted as θ_s.

An explicit expression relating the DOP to the change in a rotation angle is derived [5]. This model is able to predict the XPM induced depolarization in a two channel transmission system [5]. The time variant rotation angle is a direct result of a modulated pump signal.

3. Experimental

A continuous wave probe and an intensity modulated pump are multiplexed and transmitted over 24 km of dispersion shifted fibre (DSF) as shown in figure 1. The probe wavelength is set at 1551.72 nm while the power was kept fixed at 3 dBm. The pump was intensity modulated at 10.3 Gbps with a PRBS pattern repetition rate of 2^7-1. Correct modulation of the pump signal is monitored using a scope after demultiplexing. The probe is isolated using a 50 GHz and a narrow band filter. The pivot and probe signal are both monitored by a polarimeter.

After aligning the pump and probe SOP by using the sum of the input SOP as a guide the

polarization controller (PC) shown in figure 1 was used to rotate the SOP of the pump by 360° on the Poincaré sphere. In this way the relative angle between the pump and probe was varied between zero and 360°. The probe power was kept fixed at 3 dBm and the pump power varied from 13 dBm to 25 dBm while rotating the pump SOP for each power ratio.

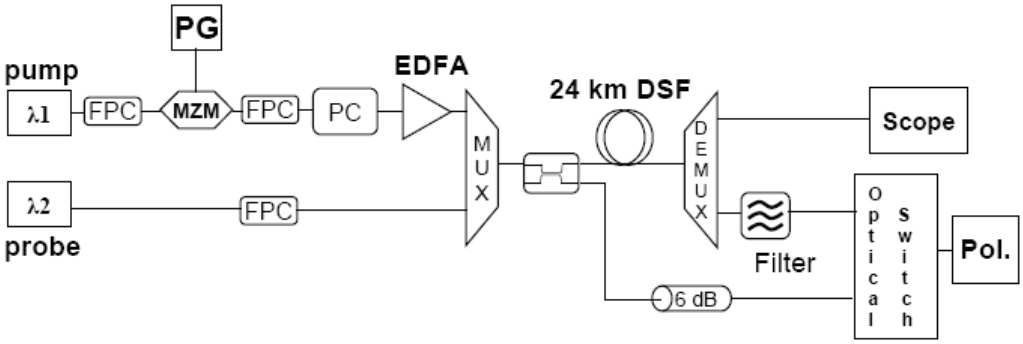


Figure 1. NRZ modulated pump and probe spaced at 50 GHz multiplexed and transmitted through 24 km of DSF.

To ensure good filtering of the probe signal the probe signal was demultiplexed using a demultiplexer as well as a narrow band filter. The probe SOP and the pivot (sum of probe and pump) where recorded using an optical switch and polarimeter.

4. Results and discussion: Degradation of average probe DOP

According to equation (5) the DOP of the probe signal in the presence of a high intensity pump signal which is intensity modulated has two minima at pump and probe relative angles of 90° and 270° through a scan of 360°. This characteristic is clearly demonstrated in figure 2 which is a plot of equation (5).

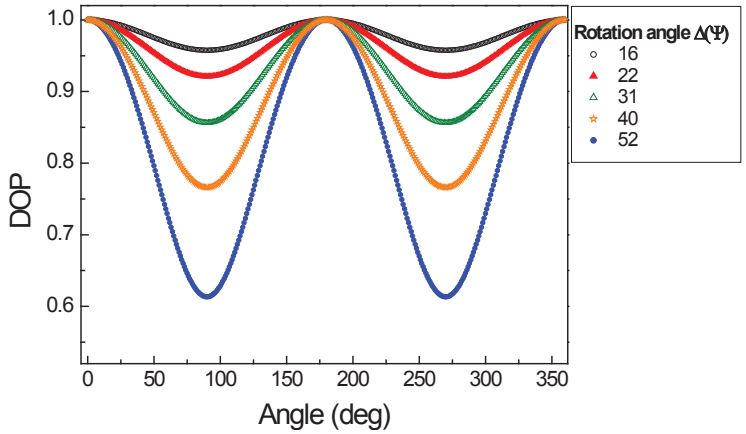


Figure 2. Probe DOP against the relative angle θs according to equation 5.

Further the average DOP decreases with an increase in the rotation angle of the probe signal which is also shown in figure 2. A bigger rotation angle corresponds to a greater power ratio between the pump and the probe signal.

The rotation angles in figure 2 correspond to the power ratios set during the experiment; 15 dB (16°), 17 dB (22°), 19 dB (31°), 22 dB (52°). Figure 3 shows the experimental results of the average probe DOP as the pump is rotated through 360° for pump powers from 13 dBm to 22 dBm. As predicted by the theory there are two minima in the spectrum at 90° and 270°. From the perspective of XPM there is maximum interaction between the pump and probe when the two signals have an orthogonal orientation with respect to each other.

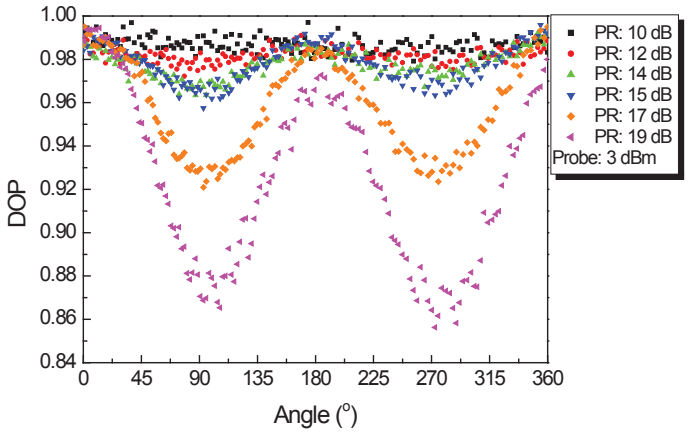


Figure 3. Transmitted probe DOP plotted as a function of the relative angle between pump and probe for different pump power levels.

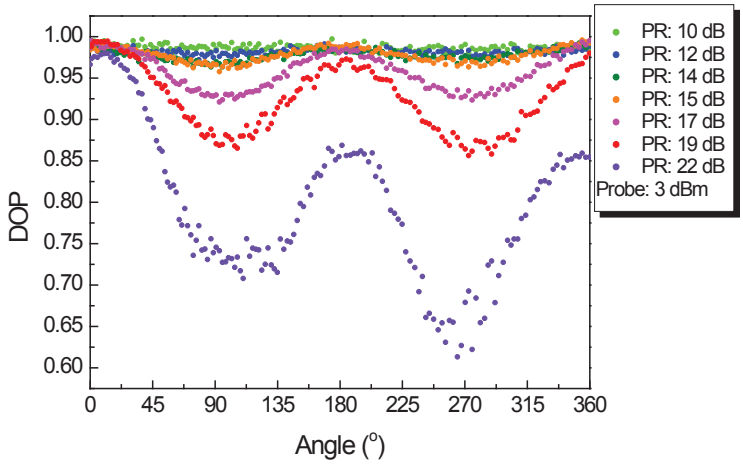


Figure 4. Transmitted probe DOP for pump power level up to 25 dBm.

When the probe and intensity modulated pump are oriented orthogonal with respect to each other then XPM changes the SOP of the probe signal as the power of the pump signal increases. When the power of the pump signal decreases the probe SOP retraces its path back to its original position. Hence the SOP of the probe signal is modulated by the intensity modulated pump signal via XPM at the pump bit rate of 10.3 Gbps. Polarimeters record the average DOP which in this case will be the sum of the states of polarization (SOPs) in time. The larger the power ratio between the probe and pump signal the

bigger the influence of the pump signal on the probe signal, increasing the swing of the SOP of the probe signal decreasing the probe DOP. This effect is very clear in figure 3 and 4 where a 22 dB power ratio translates to a rotation angle of 52^{o} according to equation (5).

5. Conclusion

This article successfully reinforces via experimental results the work in [3] proving that the average output probe DOP is a function of the relative angle, rotation angle and the power ratio between the pump and probe signal for the case of a modulated pump. PMDCs monitor the average DOP, hence if the source of the average DOP is from XPM this will be misleading and will lead to the compensator worsening the PMD within the link.

References

[1] Agrawal G P 1989 *Nonlinear Fiber Optics* (USA San Diego: Academic Press)
[2] Song Y W, Motaghian S M R, Pan Z and Willner A E, 2005 Efficient DOP monitoring of WDM channels for simultaneous PMD compensation, *Optics Communications* **255** 225-29
[3] Collings B C and Boivin L, 2000 Nonlinear Polarization Evolution Induced by Cross-Phase Modulation and Its Impact on Transmission systems, *Photon. Technol. Lett.* **12** 1582–84
[4] Lin Q and Agrawal, G P, 2004 Vector Theory of Cross-Phase Modulation: Role of Nonlinear Polarization Rotation, *Photon. Technol. Lett.* **40** 958–64
[5] Bononi A, Vannucci A, Orlandini A, Corbel E, Lanne S and Bigo S, 2003 Degree of Polarization Degradation Due to Cross-Phase Modulation and Its Impact on Polarization –Mode Dispersion Compensators, *J. Lightwave Technol.* **21** 1903–13

Quantum cryptography for satellite communication

M Mariola[1], A Mirza[1] and F Petruccione[1,2]

[1]University of KwaZulu-Natal, Westville Campus, Durban, South Africa,
[2]National Institute for Theoretical Physics, South Africa

E-mail: mmspazio@libero.it

Abstract. This project shows the possibility to use a Quantum key distribution system for aerospace applications. The project includes the possibility to use the radio signal of a public channel to synchronize the Quantum bit from Alice (transmitter) to Bob (receiver). We also use the radio signal for tracking.

1. Introduction

The research project consists of an aerospace system protecting the communication through Quantum Key Distribution (QKD) cryptography. The key is transmitted and received as polarized photons. There are many kinds of transmission protocols [1, 2, 3], but it is important to send and receive the photons with the same synchronization from the transmitter and receiver systems. At present this kind of communication is possible in optical fibre and in free space where the transmitter and receiver are stationary [4]. The stationary systems do not undergo the Doppler effect and it is possible to synchronize the signals through GPS signals. For system that move of high speed the Doppler effect is not negligible. Generally, we call the receiver system Alice and the transmitter system Bob.

The major challanges are the tracking for the laser of Alice and the sensor of Bob and the synchronization of the quantum bit of the key. This paper is a feasibility study of a satellite QKD system. When the satellite is ready Alice sends Bob the key (We are sure that Eve is not in the workshop). Once Alice is at the site of the spacecraft, the password is sent in order to be recognized. As soon as Alice receives the password, she sends Bob the answer. If any problem arises Alice requests from Bob (or vice versa) that he shifts some bits of the old key and repeat the process (Eve does not know the key). If Alice and Bob are certain that Eve does not intercept the key, they send another key and start the communication from the public channel. Figure 1 shows a flow-chart of the safe protocol:

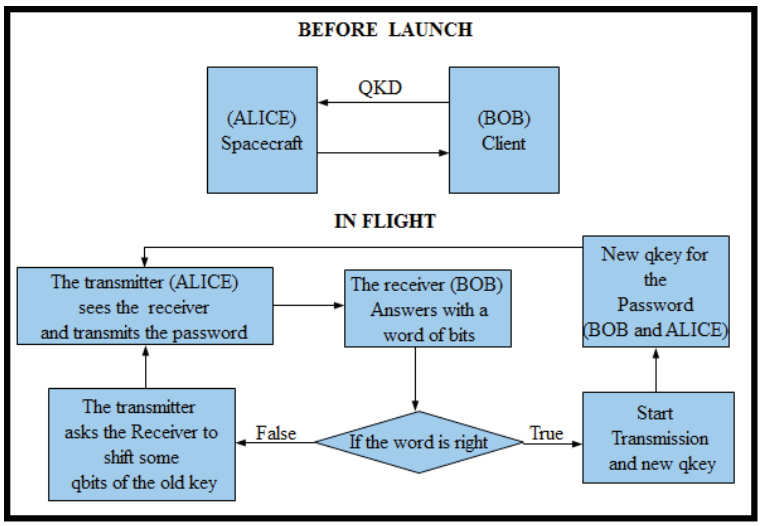


Figure 1. This figure represents the algorithm to recognize Alice to Bob and vice versa during the flight.

The encrypted password is transmitted on a public channel because Eve does not know the key.

2. Synchronization

Generally, Alice sends the qubit with random polarization. Bob does not know which polarization Alice is using, but it is necessary that he knows the time. For this reason timing is necessary for the process of synchronization between Alice and Bob. In the stationary frame system it is possible to use the GPS signal, but if we want to use this signal it is necessary to know the position of any satellite. There are, however, no such autonomous systems. We propose to synchronize Alice and Bob with a signal from Alice. The channel of the signal of synchronization can be optics or radio. The optic channel exhibits the problem of noise and if an external body is in the middle of the channel, the system loses synchronization. With the second system we do not have the problem of an external body and can use the public channel to synchronize, send information and track. To design the synchronization we must build a computational model to see the difference in propagation time between the laser and the radio signals. The frequency of the radio channel is 1.2 GHz. The variation, Δf, of the radio frequency for Doppler effect is shown in equation (1):

$$\Delta f = \frac{1}{\lambda} v_s \cos\chi, \tag{1}$$

where v_s is a tangent orbit velocity, χ is the angle between velocity vector and the direction of the Earth station, λ the wavelength. The order of variation of the frequency is 30 kHz. The variation of the frequency is independent from the shape of the waves front. If the transmitter and receiver choose the clock to be a ratio of the frequency of the radio carrier signal, we can take the clock from the radio signal that Bob receives. The system uses the phase modulation (BPSK) to send information and synchronization. The receiver recovers the carrier and takes the clock from it. When Alice receives the correct encrypted password from Bob, they begin to transmit. It is necessary to add a delay times due to different propagation time between laser and radio signal transmission. The distance between Alice and Bob in terms of the radio signal is longer if compared to the laser path because the atmosphere causes refraction and the

final trajectory is a curve. For the same reason the tracking angle of the laser and aerial can be different from each other. We can have this kind of problem in the highest layers of the troposphere and ionosphere. In the following paragraph we show the refractive index in the atmosphere. In case of a difference in time of the propagation between the laser and radio signal we can exploit the time of persistence of the laser (red line in figure 2). In this mode when we have a clock impulse the system takes the value of the quantum bit. This is shown in figure 2:

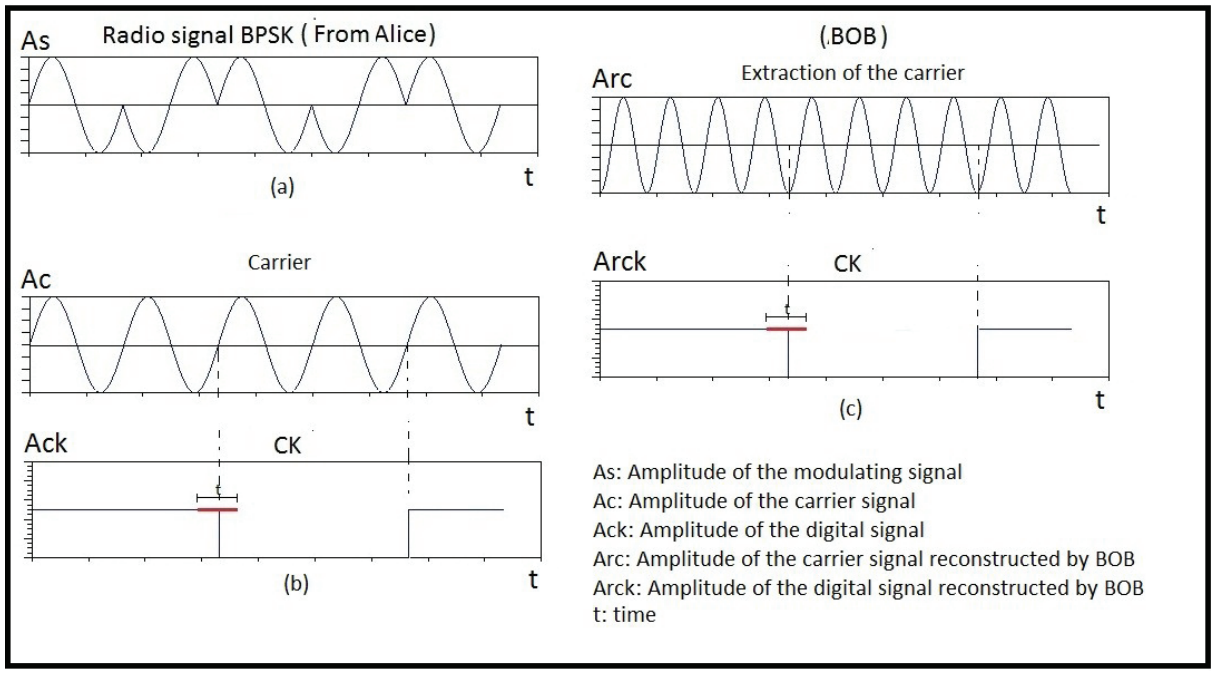


Figure 2. Clock reconstruction: (a) BPSK signal; (b) Clock signal as a fraction of the carrier signal; (c) Reconstruction of the carrier signal from Bob and the clock as its fraction

2.1. Refractive index

The trajectory of signals are dependent on the atmospheric conditions. The state parameter of atmospheric gas like pressure, temperature and humidity changes in function of the altitude and consequently also changes the refractive index [5]. The trajectory will be a curve and the angle of tracking will be different from the real position of the satellite. We make an estimation of the refractive index from the average annual atmospheric conditions. The refractive index formula is [6]:

$$n = 1 + N10^{-6}, \tag{2}$$

where:

$$N = \frac{77.6}{T}\left(P + 4810\frac{e}{T}\right), \tag{3}$$

and P and T, the pressure and temperature, respectively, are functions of the altitude. The behaviour of these state parameters is calculated by the Rec. ITU-R 835-3 and for the vapour pressure we have [7]:

$$e = \frac{\rho T}{216.7}. \tag{4}$$

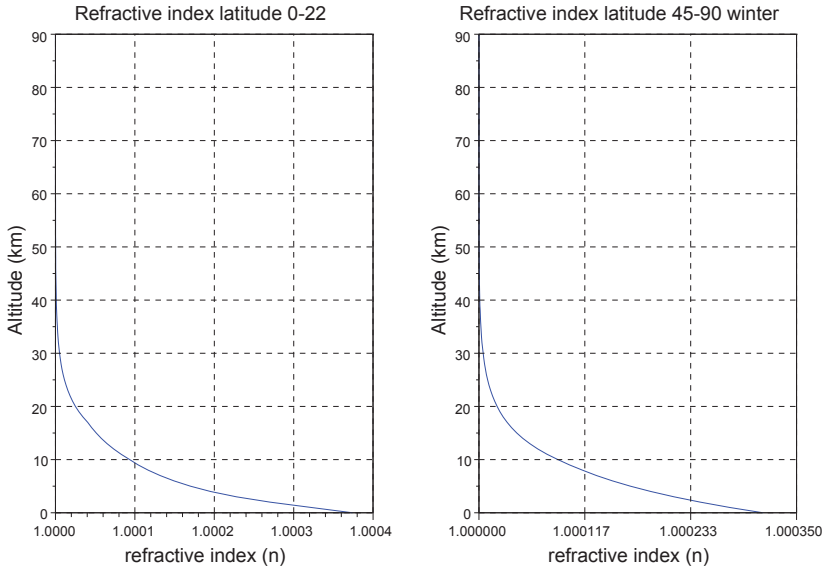


Figure 3. The graph shows that the index of refraction is a function of altitude, latitude and season. All variables are calculated by the simulation software when we know the orbit.

Equation (2) can be used only for altitudes between 0 km to 50 km. Over 50 km of altitude, the solar activity ionizes the gas leading to presence of free charge. Calling gyrofrequency f_b, critical frequency f_c and the frequency of satellite f, the formulas of the refractive index in the ionosphere are shown in equations (5) and (6) [5]:

$$n_{os}^2 = 1 - \frac{f_c^2/f^2}{1 + (f_b/f)\cos\theta}, \tag{5}$$

$$n_{xd}^2 = 1 - \frac{f_c^2/f^2}{1 - (f_b/f)\cos\theta}, \tag{6}$$

where n_{os}^2 and n_{xd}^2 are the refractive index for the characteristic waves propagation in respect to the angle of terrestrial magnetic field θ. The refractive index depends on solar activity and for that we have periodical changes every eleven years and also changes to the refractive index of day and night. We can make an estimation about the electronic charge in the atmosphere with the ionograms from different institutes of research. In the first approximation we use the data from latitude of Rome and obtain the diagrams in figure 4 [8]. The gyrofrequency at 90 km in altitude [9] can be determined from figure 5:

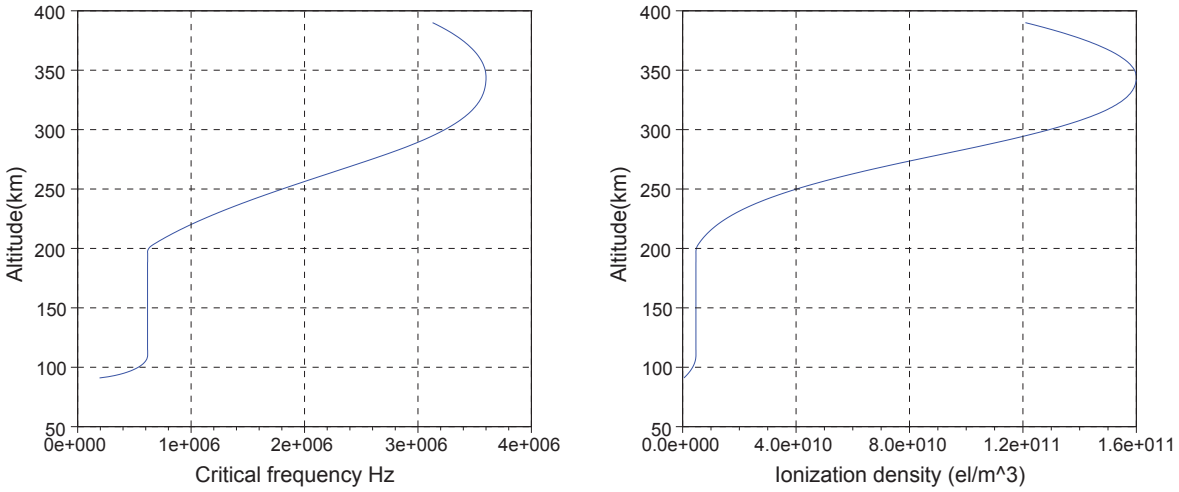


Figure 4. Reconstruction of the ionogram of the ionosphere from experimental data gathered in Rome . These trends will be used in our simulation program.

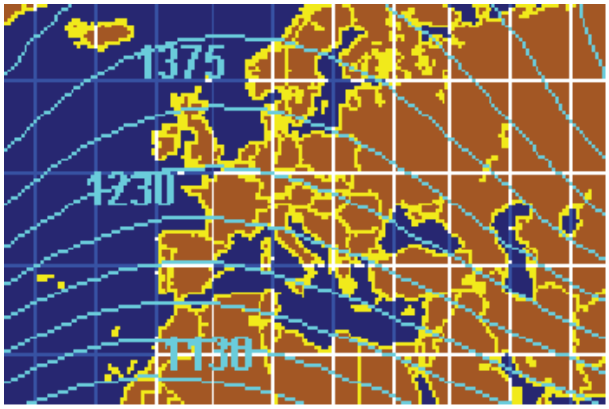


Figure 5. Map of the gyrofrequency (kHz) for 90 km in altitude: This chart is used to calculate the first approximation feasibility simulation of the proposed system. The magnetic field will be modeled in our simulation program, using the precision for each altitude.

With these frequencies and for low orbit we can estimate that the radio signal and laser have the same path because in the lowest layer of atmosphere the refractive index depends on the thermodynamic conditions and in the low ionosphere the refractive indexes are approximatively one. If the clock from the radio signal has some delay time with respect to the laser we can compensate by minimizing the bit rate of the quantum key or using a higher radiofrequency. For the highest layers of atmosphere the gas molecules can have a distance comparable with the wavelength of the laser. In this case the path between the laser and the radio signal is different because of the higher concentrations of ionized gas in the highest layers of ionosphere. The refractive index for the laser and radio signal is different in the ionosphere. We assume that the path of the laser is a straight line of sight and that the radio signal follows the trajectory spanned out by equations (5) and (6).

3. Tracking

Using the public channel and a minimum of two aerials it is possible to build a tracking system for a satellite or balloon with an Earth station and vice versa. The conceptual scheme is shown in the figure 6:

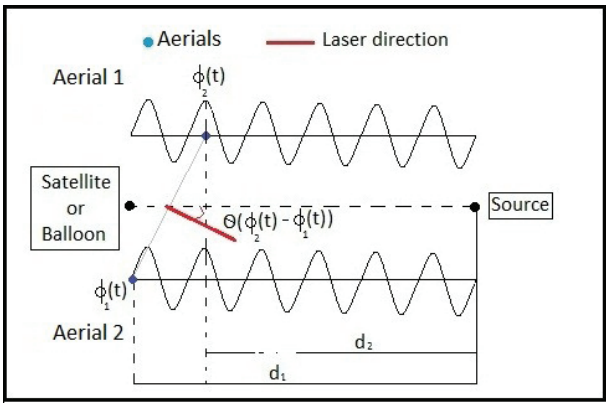


Figure 6. Principle diagram of the tracking system from the difference of the phase: Aerial 1 receives the signal with phase ϕ_1 and Aerial 2 with phase ϕ_2 . From the difference of the two phases we know the angular position of the satellite or balloon. For that we must use two different frequencies for Alice and Bob in transmission and reception.

4. Conclusion

From approximating the refractive index it is evident that the radio signal and laser signal follow very similar paths for altitudes of up to 50 km. For this reason we can synchronize Alice and Bob by the radio signal, because they do not have much difference with regards to the time of propagation. For low orbit, where the radio and laser signal are inclined to approach one another. We can use the radio signal for the tracking system. When the computational model is complete it will be possible to refine the project from this first approximation.

Acknowledgments

This work is based on research supported by the South African Research Chair Initiative of the Department of science and Technology and National Research Foundation.

References

[1] Bennett C and Brassard G 1984 *Quantum cryptography: public key distribution and coin tossing* Proc of IEEE International conference on computer systems and signal processing 175-179
[2] Bennet C 1992 Quantum cryptography using two nonorthogonal states *Phys. Rev. Lett.* **68** 3121-3124
[3] Hughes Richard J and Buttler William T 1999 *Pratical quantum cryptography for secure free-space communications*, Los Alamos, University of California
[4] Ursin R et al.3 June 2007 *Entanglement-based quantum communication over 144 km* Nature physics **3** 481-486
[5] Bernardini A January 2008 Lezioni del corso di sistemi di comunicazione satellitari Vol 1 ed Edizioni Ingegneria 2000 Rome chapter 4 pp.192-199
[6] [online] Rec. ITU-R P.453-8; http://www.catr.cn/radar/itur/201007/P020100714503680970889.pdf
[7] [online] Rec. ITU-R P.835-3; http://www.catr.cn/radar/itur/201007/P020100714463265537648.pdf
[8] [online] http://ionos.ingv.it/Roma/
[9] [online] De Canck M H Ionosphere Properties and Behaviors Part 1 p.5, June 2006; http://www.antennex.com/prop/prop0606/prop0606.pdf

Spectroscopy with a mode-locked femtosecond laser frequency comb

Vela Mbele and Elias Sideras-Haddad

School of Physics, University of the Witwatersrand, 1 Jan Smuts, Braamfontein, Johannesburg

E-mail: vela.mbele@students.wits.ac.za

Abstract. We summarise high-resolution, high-precision spectroscopy experiment where caesium-133 atoms in a vapour cell are excited directly with a femtosecond laser frequency comb. In this experiment the laser beam is split into two, aligned counter-propagating, and focussed to a reasonable waist in the interaction region in the vapour cell, thereby exciting a multitude of low lying transitions allowing the measurement of transition energies and hyperfine coupling coefficients for the $8S_{1/2}$, $9S_{1/2}$ and $7D_{3/2,5/2}$ states.

1. Introduction

The technology awarded the one-half of the 2005 Nobel prize for Physics involving a mode-locked femtosecond laser frequency comb (FLFC) [1] has been gaining in the range of applications over the past years. This technology originally designed to count cycles in optical clocks, now has applications extending over a vast array of research areas including spectroscopy, trace gas detection, signal processing, astrophysics and many others. The basic ideas of using repetitive pulses and mode-locked lasers for high-resolution precision spectroscopy followed soon after the development of the laser itself, in the early 1970s. Stabilization of optical frequency combs [2], in the mid to late 1990's triggered and stimulated this renewed interest and the many spectroscopic applications of mode locked laser's [3, and references therein].

The higher intensities and longer coherence lengths in lasers allowed scientists to perform experiments, that had until then only been restricted to the theoretical domain, including predicted experiments that could not be performed with the lamps (see for an example earlier work in caesium [4]) that were used in earlier years of spectroscopic studies. Experiments which had, before the advent of the laser, been only restricted to theoretical predictions included Doppler-free two photon transitions [5]. The two-photon theory [6] on which these experiments are based, had been predicted long before the lasers were developed (see also, for example, the review [7, and references therein] and the theoretical formalization that followed after the development of the laser [8]).

Precision laser atomic spectroscopy, has and continues to play an important role in fundamental physics studies [9, 10, 11, 12, 13, 14, 15, 16]. Caesium is, and continues to be, an important test-bed for fundamental studies and atomic theory [9]. Hyperfine interactions are sensitive to the details of nuclear structure, electron correlations, relativistic and core polarisation effects. The correlational effects come about because of the degree to which the $P-D$ effects mix into $S-S$ parity violation amplitudes. Generally, experimental considerations usually restrict high precision work to S and P states [13, 16, 17], while measurements in the D

states still remain a challenge because of correlational effects [18] and hence there is a need for more measurements in the D states [19], answered in part in the work presented here.

There is a likelihood that in future, these applications may be taken over by the compact monolithic micro-resonator generated frequency combs [20, 21]. There is, however, no doubt that mode-locked FLFC will continue growing their range of applicability over the next few decades.

2. The femtosecond frequency comb

The femtosecond laser frequency comb used in these measurements is based on the Kerr-lens mode-locked [22] titanium doped sapphire (Ti:Sa) crystal laser. The octave-spanning nature of this laser, with radiation ranging from 530 nm to well above 1200 nm, aids in the self-referencing [2]. The frequency mode structure of the femtosecond laser comb can be understood from mode-locking which is dependent on time/frequency domain structure of mode-locked lasers. Generally, any single pulse will have a spectrum that is the Fourier transform of its envelope function and centered at the optical frequency of its carrier and for any pulse shape, the frequency width of the spectrum will be inversely proportional to the temporal width of the envelope. On the other hand, for pulses separated by a fixed interval, the spectrum is obtained by a Fourier series expansion, yielding a comb of regularly spaced frequencies, where the comb spacing is inversely proportional to time separation between pulses, and the offset is rigidly shifted in dependence of the carrier envelope phase. The spacing between the comb modes is the repetition rate, $f_{\rm rep}$ and the carrier envelope phase dependent shift is the offset frequency, f_0, such that the optical frequency of the n-th of those comb modes is

$$\nu_n = n f_{\rm rep} + f_0, \tag{1}$$

where n ranges from 250 000 to 570 000. The successive slip of the phase with each pulse, $\Delta\phi$, the repetition rate $f_{\rm rep}$ and the offset frequency, f_0, are related by

$$\Delta\phi = 2\pi(f_0/f_{\rm rep}). \tag{2}$$

It is equation 1 that provides the optical frequency comb with the "gearbox" [21] capability which enables the reliable and accurate means of counting optical cycles on the femtosecond (10^{-15} s) time scale, the single most important character in the realization of optical clocks. The accuracy of such clocks is now measured to the 18th digit [23], exceeding the performance of their microwave counterparts by more than an order of magnitude. This development involves the measurement of the energy structure of atoms at an unprecedented level, allowing some of the most precise laboratory tests of the physics governing these systems.

Precision spectroscopy requires that the laser be referenced to a frequency standard. With a low noise microwave standard, the hydrogen maser in the case of this experiment, as a reference for the frequency comb the frequency of each element of the comb can be determined absolutely with a fractional uncertainty at or below $\sim 2 \times 10^{-13}/\sqrt{\tau}$ where τ is the averaging period in seconds or longer. The precision of the optical clocks are fast getting to the $\Delta f/f$ limit [23] and the use of an optical reference would provide ~ 1 Hz optical resolution and fractional uncertainty of a few parts in $\sim 10^{17}$ range. However, for the spectroscopy experiments that we report in this work the microwave maser reference is sufficient and the atomic reference provided by the global positioning system (GPS) or the commercially available beam clocks would also be adequate. Stated alternatively, the linewidth of the comb modes, which are ~ 100 kHz, with the maser reference, could have been a few Hz if the laser comb were referenced to an optical clock. This study therefore serves to re-affirm the need of improved references and hence all laboratories with capabilities should look into developing their own optical standards.

3. The caesium experiment

The experimental set-up block diagram of the caesium experiment is shown in figure 1. Radiation from the laser is sent to a caesium vapour cell via two beam paths. In each route the laser radiation is filtered, by the optical bandpass filters marked F1 and F2, to insure excitation of individual transitions and to eliminate excitation by Doppler broadened co-propagating photons. Without the filters the collected spectrum is complicated, as can be seen in figure 3, and introduction of the filters to isolate one of the states results in the spectrum shown in figure 4. The excitation pathways are shown figure 2. Fluorescence collected from the $7P_{1/2,\,3/2}$ states is detected by a photomultiplier tube placed next to the vapour cell at a right angle to the two counterpropagating beams. We used the bandpass filter marked F3 to isolate this decay while also eliminating the background scattering. Two-photon resonance signals were monitored by measuring the $7P_J \rightarrow 6S_{1/2}(J = 1/2, 3/2)$ cascade fluorescence at 459 and 455 nm, because these wavelengths were outside the range of the laser spectrum. The detector was shielded from background light and the scattered light from the excitation laser beams by a 457.9-nm filter with a 10-nm pass band (F3).

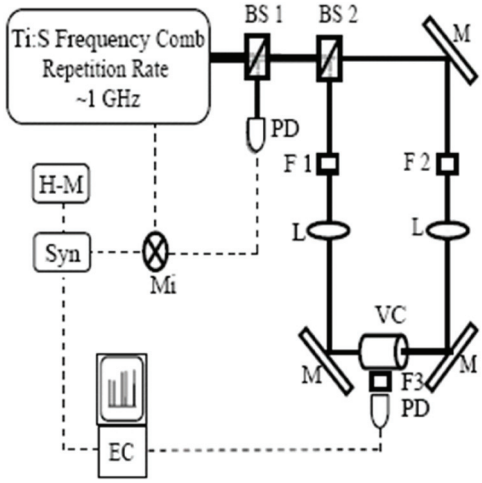


Figure 1. Block diagram of the Cs spectrometre using the femtosecond frequency comb and a Cs vapor cell (VC). The other components are: BS, beamsplitter; M, mirror; F1, F2, F3 bandpass filters; L, lens; PD, photodetector; Mi, Mixer; SYN, frequency synthesizer; H-M, hydrogen maser and EC, denotes the data collection and experimental control computer.

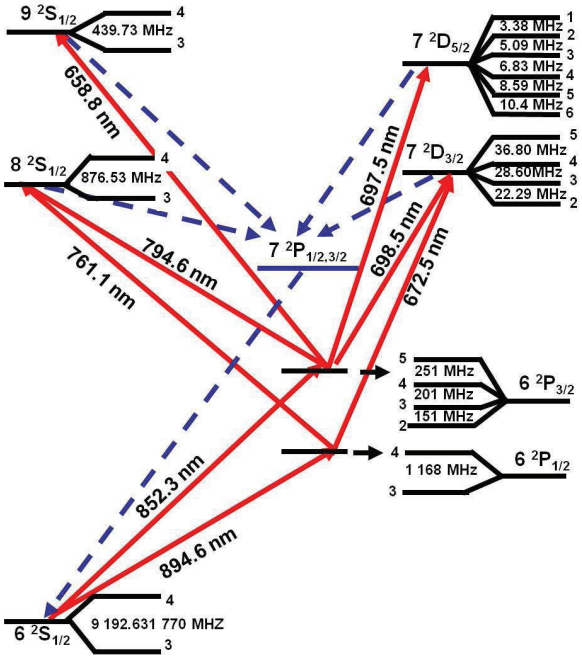


Figure 2. The low-lying energy levels of ^{133}Cs (I = 7/2) that are relevant to this work. The two-photon transition is resonant with the real intermediate state and the atom is subsequently excited to the final state. The excitation pathways are shown in solid red lines along with the transition wavelengths. The decay channels used in the detection are shown in blue dashed arrows. The total angular momentum of the hyperfine components is listed next to each state.

The filters are shown in table 1, where F1 is the filter used in either the D1, $6S_{1/2} \rightarrow 6P_{1/2}$, or D2 , $6S_{1/2} \rightarrow 6P_{3/2}$, transitions in the one arm and F2 the complimentary excitation to the higher state in the other. The laser spectrum had sufficient light intensities to allow access of the $8S_{1/2}$ state through both of the $6P_{1/2,3/2}$ states as the intermediary state. We could only access the $9S_{1/2}$ state through the $6P_{3/2}$ because of lack of sufficient light intensities in the complimentary radiation so as to excite this state through the $6P_{1/2}$ state. Our ability to measure the $8S_{1/2}$ state through both of the intermediate states allowed us to use the $8S_{1/2}$ state to check for repeatability in our measurements by quantifying the agreement in the absolute frequency and the hyperfine coupling constant measurements attained through the two different excitation pathways. We found very good agreement between the two sets of measurements. This agreement also augured well for the validity of our method of analysis.

The data for the $7D_{3/2}$ was only collected through the D1 transition, because the manifold structures of both the $7D_{3/2}$ and $6P_{3/2}$ states resulted in a complicated spectrum. Another complication we had was that the fine-structure splitting between the $7D_{5/2}$ and $7D_{3/2}$ states is so small such that it was impossible to selectively excite both of the states through the $6P_{3/2}$ state, as such the $7D_{3/2}$ state was only excited through the D1 transition, and the $7D_{5/2}$ through the D2 transition. Furthermore, selection rule considerations do not allow electric dipole transitions through the $6P_{1/2}$ to $7D_{5/2}$, because that would have meant $\Delta J = 5/2 - 1/2 = 2$ for the electric dipole transition. We still had a residual $7D_{3/2}$ detected, together with the $7D_{5/2}$. The $7D_{3/2}$ is accessible through both routes. This was because the wavelengths connecting $6P_{3/2}$ to the two $7D_{3/2}$ and $7D_{5/2}$ states are 698 nm and 697 nm, respectively. This made it impossible to completely separate the $7D_{3/2}$ contribution from the $7D_{5/2}$ using the bandpass filters available to us. However, the matrix elements for the $6P_{3/2} \rightarrow 7D_{5/2}$ were far stronger than the corresponding $6P_{3/2} \rightarrow 7D_{3/2}$ matrix elements. This fact, coupled with the lower transmission at 698 nm of the bandpass filter we used to excite the complementary transition allowed us to extract $6P_{3/2} \rightarrow 7D_{5/2}$.

Table 1. Optical filters and combinations thereof used as F1 and F2 in figure 1. Wavelengths are in nanometers, followed by transmission bandwidth (nm), l.p. indicates long-pass and s.p. short-pass edge filters.

Final	Intermediate	Filter 1	Filter 2
$8S_{1/2}$	$6P_{1/2}$	890 (10)	755(40)/715 l.p.
$8S_{1/2}$	$6P_{3/2}$	850 (10)	800 s.p./780 l.p.
$9S_{1/2}$	$6P_{3/2}$	850 (10)	657.9 (10)
$7D_{3/2}$	$6P_{1/2}$	890 (10)	670 (30)
$7D_{5/2}$	$6P_{3/2}$	850 (10)	700 (25)

The probability that the atom will make a transition according to the second-order time-dependent perturbation theory is given by [24]:

$$P(6S_{\frac{1}{2}} : n''L''_{J''}) \propto \frac{1}{2F+1} \frac{I_1 I_2}{\left[\omega_{6S_{\frac{1}{2}}F:n''L''_{J''}F''} - (\omega_1 + \vec{k_1}.\vec{v}) - (\omega_2 + \vec{k_2}.\vec{v})\right]^2 + \left(\frac{\gamma_{n''L''}}{2}\right)^2}$$
$$\times \sum_{M_F, M'_F} \left| \sum_{F', M''_F} \frac{\langle n''L''_{J''}F''M''_F|\hat{e}_2.\vec{d}|6P'_{J'}F'M'_F\rangle\langle 6P'_{J'}F'M'_F|\hat{e}_1.\vec{d}|6S_{\frac{1}{2}}FM_F\rangle}{\omega_{6S_{\frac{1}{2}}F:6P'_JF'} - \left(\omega_1 + \vec{k_1}.\vec{v}\right) - i\gamma_{6P_{J'}}} \right|^2 . \quad (3)$$

In equation 3, $\vec{v}$ is the atomic velocity, $\vec{d}$ the electric dipole operator, M_F, M'_F and M''_F are the projections of F, F' and F'' along the quantization axis, $\hat{e_1}$ is the unit vector in the polarisation

direction of the beam with wave vector $\vec{k_1}$ and $\hat{e_2}$ is the same for $\vec{k_2}$. We use equation 3 to reproduce the spectra, with the resulting spectra being used in the extraction of the absolute frequencies and the hyperfine coupling coefficients, by fitting the peak frequencies from the experimental and the calculated spectra. We designate the peak centers in the repetition rate frequencies of the experimental spectra, $f^{\rm expt}$, and the theoretical spectra, $f^{\rm calc}$. For each peak we then evaluate the χ^2 function, as a normalized difference between the calculated interpolating function and the experimental data,

$$\chi^2(\Delta\nu_{6{\rm S}-{\rm n}''{\rm L}''_{\rm J}}, \Delta{\rm A}_{{\rm n}''{\rm L}''}, \Delta{\rm B}_{{\rm n}''{\rm L}''}) = \sum_i \left(\frac{f_i^{\rm expt} - f_i^{\rm calc}\left(\Delta\nu_{6{\rm S}-{\rm n}''{\rm L}''_{\rm J}}, \Delta{\rm A}_{{\rm n}''{\rm L}''}, \Delta{\rm B}_{{\rm n}''{\rm L}''}\right)}{\sigma_i^{expt}} \right)^2, \quad (4)$$

Table 2 shows all the frequencies calculated and inferred from these measurements.

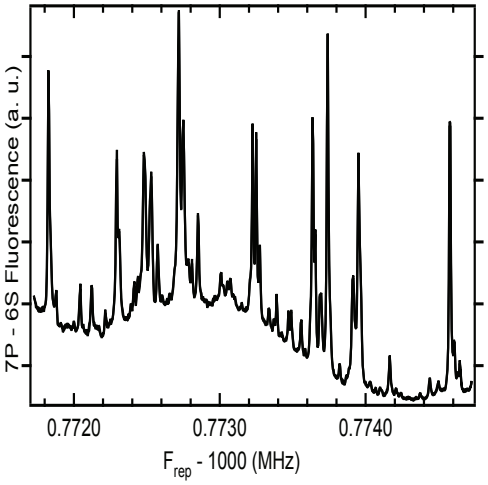


Figure 3. Caesium fluorescence from $7{\rm P}_{1/2,3/2} \rightarrow 6{\rm S}_{1/2}$ decay. To collect the spectrum the laser cavity is scanned without the filters (F1 and F2, Fig. 1) which we use to restrict the laser's spectral bandwidth.

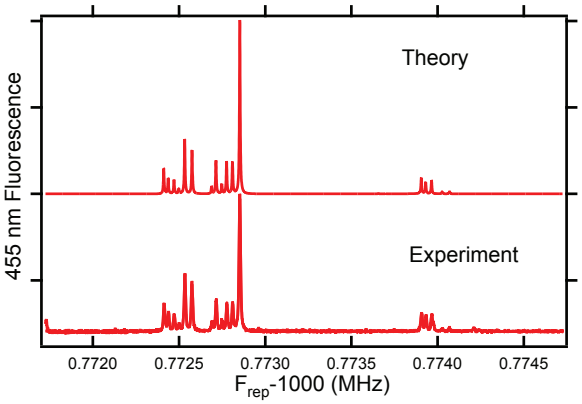


Figure 4. Spectrum of the $6{\rm S}_{1/2} - 7{\rm D}_{3/2}$ transition via $6{\rm P}_{1/2}$ intermediate state. The lower of the two traces is the experimental data and the other the modeled spectrum. The plot shows the fluorescence intensity from the 7P to 6S decay as a function of the laser repetition rate, $f_{\rm rep}$.

We find the centre of gravity frequency of 729 009 798.863(29) MHz and the hyperfine coupling constant of 219.133(17) MHz for the excitation through the the D1 transition and 729 009 798.86(19) MHz with 219.14(11) MHz for the excitation through the D2 transition. These are consistent while agreeing quite well with a recent measurement [25] where they found parameters of 729 009 799.020(26) MHz and 219.125(4) MHz. All the other measurements represent a two order improvement in the centre of gravity frequency[26] while representing an order of improvement in the hyperfine coupling constants [27]. The results for all of the transitions are summarized in table 2.

4. Conclusion

This technique has potential metrological applications for secondary frequency standards at optical wavelengths. At a more fundamental level, the hyperfine coupling measurements are

Table 2. The single photon transition frequencies, $\nu_{6S_{1/2}\rightarrow n''L''_{J''}}$, hyperfine coupling coefficients, $A_{n''L''_{J''}}$ and $B_{n''L''_{J''}}$, with the complementary transition frequencies, $\nu_{6P_{1/2}\rightarrow n''L''_{J''}}$ and $\nu_{6P_{3/2}\rightarrow n''L''_{J''}}$ and $\nu_{FF''}$ measured in this work. All numbers are in MHz.

Parameter	$8S_{1/2}$	$9S_{1/2}$	$7D_{3/2}$	$7D_{5/2}$
$\nu_{6S_{1/2}\rightarrow n''L''_{J''}}$	729 009 798.86(15)	806 761 363.14(10)	780 894 762.250(64)	781 522 153.68(16)
$A_{n''L''_{J''}}$	219.133(75)	109.932(50)	+7.386(18)	-1.717(15)
$B_{n''L''_{J''}}$			-0.182(163)	-0.182(517)
$\nu_{6P_{1/2}\rightarrow n''L''_{J''}}$	393 893 750.11(15)	471 645 314.63(10)	445 778 713.502(64)	446 406 104.93(16)
$\nu_{6P_{3/2}\rightarrow n''L''_{J''}}$	377 284 080.39(15)	455 035 644.91(10)	429 169 043.778(64)	429 796 435.21(16)
ν_{31}				781 527 343.75(36)
ν_{32}			780 899 883.152(163)	781 527 340.37(25)
ν_{33}	729 014 476.67(22)	806 766 286.89(15)	780 899 905.440(97)	781 527 335.28(19)
ν_{34}	729 015 353.20(20)	806 766 726.62(13)	780 899 935.036(99)	781 527 328.45(23)
ν_{35}			780 899 971.836(121)	781 527 319.86(22)
ν_{42}			780 890 690.521(163)	781 518 147.74(25)
ν_{43}	729 005 284.03(22)	806 757 094.26(15)	780 890 712.809(97)	781 518 142.65(19)
ν_{44}	729 006 160.57(20)	806 757 533.98(13)	780 890 742.405(100)	781 518 135.82(23)
ν_{45}			780 890 779.205(121)	781 518 127.23(22)
ν_{46}				781 518 116.83(24)

valuable in improving atomic structure calculations. The measurement accuracy of multiple transitions at significantly different wavelengths rivals that of cw laser spectroscopy techniques and demonstrates the versatility of DFCS.

Acknowledgements
We acknowledge the assistance of Dr Scott Diddams and his femtosecond laser frequency comb laboratory staff at NIST in Boulder, Colorado for allowing these experiments to be performed in his laboratory and helping us set up the experiment.

References

[1] John L. Hall 2006 *Rev. Mod. Phys.* **78(4)** 1279; Theodor W. Hänsch 2006 *Rev. Mod. Phys.* **78(4)** 1297
[2] D. J. Jones, S. A. Diddams, J. K. Ranka, A. Stentz, R. S. Windeler, J. L. Hall, and S. T. Cundiff 2000 *Science* **288** 635
[3] M. C. Stowe, M. J. Thorpe, A. Péer, J. Ye, J. E. Stalnaker, V. Gerginov, and S. A. Diddams 2008 *Adv. At. Mol. Opt. Phys.* **55** 1; S. A. Diddams 2010 *J. Opt. Soc. Am. B*, **27** B51
[4] H. Kleiman and K. W. Meissner 1959 *TECHNICAL REPORT No. 1* 1; H. Kleiman 1962 *J. Opt. Soc. Am.* **52(4)** 441
[5] F. Biraben, B. Cagnac, and G. Grynberg 1974 *Phys. Rev. Lett.* **32** 643; F. Biraben, B. Cagnac, and G. Grynberg 1974 *Phys. Lett. A* **48** 469; M.D. Levenson and N. Bloembergen 1974 *Phys. Rev. Lett.* **32** 645; T.W. Hänsch, K. Harvey and G. Meisel 1974 *Opt. Comm.* **11** 50
[6] M. Göppert-Mayer 1931 *Annalen der Physik* **9(2)** 272
[7] N. Bloembergen 1982 *Rev. Mod. Phys.* **54** 685
[8] L. S. Vasilenko, V. P Chebotaev, and A.V. Shishaev 1970 *JETP Lett.* **12** 113
[9] J. Guena, M. Lintz, and M-A. Bouchiat 2005 *J. Opt. Soc. Am. B* **22** 21; C. S. Wood, S. C. Bennett, D. Cho, B. P. Masterson, J. L. Roberts, C. E. Tanner, and C. E. Wieman 1997 *Science* **275** 1759
[10] S. C. Bennett and C. E. Wieman 1999 *Phys. Rev. Lett.* **82** 2484
[11] S. A. Murthy, D. Krause, Z. L. Li, and L. R. Hunter 1989 *Phys. Rev. Lett.* **63** 965; H. S. Nataraj, B. K. Sahoo, B. P. Das, and D. Murkherjee 2008 *Phys. Rev. Lett.* **101** 033002
[12] A. Wicht, J. M. Hensley, E. Sarajlic, and S. Chu 2002 *Phys. Scr.*, **T102** 82; G. Gabrielse, D. D. Hanneke, T. Kinoshita, M. Nio, and B. Odom 2006 *Phys. Rev. Lett.* **97** 030802
[13] V. Gerginov, K. Calkins, C. E. Tanner, J. J. MacFerran, S. A. Diddams, A. Bartels, and L. Hollberg 2006 *Phys. Rev. A* **73** 032504
[14] Peter J. Mohr, Barry N. Taylor, and David B. Newell 2008 *Rev. Mod. Phys.* **80** 633; Ulrich D. Jentschura, Peter J. Mohr, Joseph N. Tan, and Benedikt J. Wundt 2008 *Phys. Rev. Lett.* **100** 160404; B. J. Wundt and U. D. Jentschura 2010 *J. Phys. B: At. Mol. Opt. Phys.* **343** 115002
[15] E. Peik, B. Lipphardt, H. Schnatz, T. Schneider, Chr. Tamm, and S. G. Karshenboim 2004 *Phys. Rev. Lett.* **93** 170801; Chr. Tamm, T. Schneider, and E. Peik 2004 ed S. G. Karshenboim and E. Peik *Lecture Notes in Physics:Astrophysics, Clocks and Fundamental Constants* chapter 15, (Springer, Berlin) p 247; T. M. Fortier, N. Ashby, J. C. Bergquist, M. J. Delaney, S. A. Diddams, T. P. Heavner, L. Hollberg, W. M. Itano, S. R. Jefferts, K. Kim, et al. 2007 *Phys. Rev. Lett.* **98** 070801
[16] V. Gerginov, C. E. Tanner, and A. Derevianko 2003 *Phys. Rev. Lett.* **91** 072501
[17] A. Kortyna, N. A. Masluk, and T. Bragdon 2006 *Phys. Rev. A* **74** 022503; M. S. Safronova, W. R. Johnson, and A. Derevankio 1999 *Phys. Rev. A* **60** 4476; V. Gerginov, C. E. Tanner, S. A. Diddams, A. Bartels, and L. Hollberg 2004 *Phys. Rev. A* **70** 042505
[18] V. A. Dzuba, V. V. Flambaum, and J. S. M. Ginges 2001 *Phys. Rev. A* **63** 062101
[19] M. S. Safronova and Charles W. Clark 2004 *Phys. Rev. A* **69** 040501
[20] P. Del'Haye, A. Schliesser, O. Arcizet, T. Wilken, R. Holzwarth, and T. J. Kippenberg 2007 *Nature* **450** 1214
[21] T. J. Kippenberg, R. Holzwarth, and S. A. Diddams 2011 *Science* **332** 555
[22] A. Bartels 2005 ed J. Ye and S. T. Cundiff *Femtosecond Optical Frequency Comb Technology: Principle, Operation and Application* chapter 3 (Springer, New York) p 78
[23] W. H. Oskay, W. M. Itano, and J. C. Bergquist 2005 *Phys. Rev. Lett.* **94** 163001; T. Rosenband, D. B. Hume, P. O. Schmidt, C. W. Chou, A. Brusch, L. Lorini, W. H. Oskay, R. E. Drullinger, T. M. Fortier, J. E Stalnaker, et al. 2008 *Science* **319** 1808
[24] W Demtröder 1998 *Laser Spectroscopy: Basic Concepts and Instrumentation.* (Springer, New York)
[25] P. Fendel, S. D. Bergeson, T. Udem, and T. W. Hänsch 2007 *Optics Letters* **32** 701
[26] K.-H. Weber and Craig J. Sansonetti 1987 *Phys. Rev. A* **35** 4650

[27] J. Farley, P. Tsekeris, and R. Gupta 1977 *Phys. Rev. A.* **15** 1530; R. Gupta, W. Happer, L. K. Lam, and S. Svanberg 1973 *Phys. Rev. A* **8** 2792; M. Auzinsh, K. Bluss, R. Ferber, F. Gahbauer, A. Jarmola, M. S. Safronova, U. I. Safronova, and M. Tamanis. 2007 *Phys. Rev. A* **75** 022502; A. Kortyna, V. Fiore, and J. Farrar 2008 *Phys. Rev. A* **77** 062505

Efficiency of Tm^{3+}-doped silica triple clad fiber laser

W I Ndebeka, A M Heidt, H Schwoerer and E G Rohwer
Laser Research Institute, Physics Department, University of Stellenbosch, South Africa
E-mail: ndebeka@sun.ac.za

Abstract. We present measurements of the slope efficiency and the pump power at threshold of a Tm^{3+}-doped silica triple clad fiber laser emitting at 2.02 μm using different cooling techniques. The slope efficiency of 53.6 % was obtained at a temperature of 0 °C with a maximum output power of 5 W for 19 W of absorbed power at the pump wavelength of 800 nm and 9.9 W threshold. In a slightly different setup, the output power could be increased to 10 W for an absorbed pump power of 32 W.

1. Introduction

The continuous search for high power light sources has led scientists to look for a suitable device to generate high power laser output. Recent advances in fiber technology have produced demonstrations of fiber lasers that can generate output powers of several kilowatts. To date, the highest power devices have been based on ytterbium (Yb^{3+})-doped silica fibers that operate in the wavelength region centered around 1080 nm. Lasers in this wavelength region pose serious hazard for eyes since their beams are invisible but the power can be imaged onto the retina. Currently the power scaling of thulium-(Tm)-doped fiber lasers, emitting in the so-called eye-safe wavelength region around 2 μm, is of particular interest [1]-[4]. Tm-doping is especially interesting for high power laser operation due to the possibility of cross relaxation process (figure 1), in which two excited ions can be produced with only one pump photon [5]-[6].

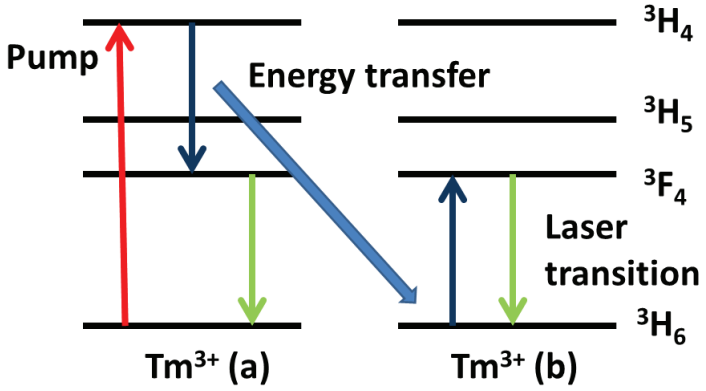


Figure 1. Energetic diagram of Tm^{3+} doped fiber between two Thulium ions.

Thus, in theory one can obtain a pump efficiency of 2 [1], which increases the power scaling possibilities due to reduced heat extraction problems. In this work, we investigate

the efficiency of newly developed triple-clad Tm doped fiber manufactured by the Canadian company CorActive. A fiber laser was constructed, at the Laser Research Institute (LRI), to investigate the efficiency of the fiber, and different cooling techniques were used to investigate the occurrence of the cross-relaxation process and the results are presented and discussed.

2. Experiments

The Tm^{3+}-doped silica fiber had a Tm^{3+} concentration of 5 % by weight, an inner core diameter of 18.1 μm (effective NA = 0.11) surrounded by a second core with 48.5 μm diameter. The cladding had 403 μm diameter. The specified absorption coefficient was 2.75 dBm^{-1} at 800 nm. Experiments were conducted using a pump diode laser (JOLD-75-CPXF-2P W) which is specified to provide 75 W maximum output power at 805 nm. A Coherent FieldMaster power meter was used to measure the output power. A schematic diagram of the laser setup is shown in figure 2. The output beam from the diode is collimated by a lens (f = 15 mm), reflected with two dichroic mirrors, M_1 and M_2 (HR@ 800 nm, HT@ 2 μm). The beam is then focused inside the Tm-doped fiber (NA = 0.11, length = 4 m) with a second lens (f = 8 mm). Feedback for laser oscillations was provided by a plane gold mirror (M_3) which had a high reflectivity at 2 μm. During the experiments, the fiber was required to be immersed in water in order to control the temperature to achieve maximum cooling efficiency. The temperature of 0 °C was achieved by using ice water. The spectral output was recorded using a 0.5 m monochromator with a 300 lines per mm grating blazed for 2 μm and an InGaAs detector.

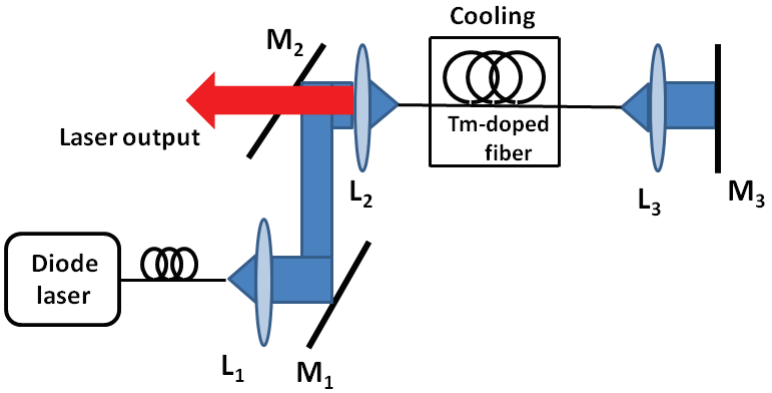


Figure 2. Typical setup used for the characterization of the Tm-doped fiber; M_1 and M_2 are dichroic mirrors (HR @ 800 nm, HT @ 2 μm)and M_3 is a plane gold mirrors and L_1 (f = 15 mm) is a coated lens, L_2 (f = 8 mm) is an uncoated lens and L_3 (f = 8 mm) is a collimated lens.

3. Results

The measured output power of the diode-pumped Tm-doped fiber against the absorbed pump power at room temperature (25 °C) and ice water (0 °C) is shown in figure 3. For these measurements the absorbed power was limited to 19 W. An output power of 4.7 W was obtained at 25 °C with a slope efficiency of 49.3 % and threshold of 9.9 W absorbed input power. At 0 °C the slope efficiency increased to 53.6 % and the threshold dropped to 7.8 W and an output power of 5.5 W was obtained for 19 W of absorbed pump power. A typical recorded output spectrum for 4.7 W output at 25 °C is shown in figure 4. We observe a spectral output with multiple wavelengths around 2.02 μm. The input power of the diode laser was increased to 32 W of absorbed pump power and an output power of 10 W was obtained at 0 °C. The output power was limited by thermal damage inside the fiber and the optics.

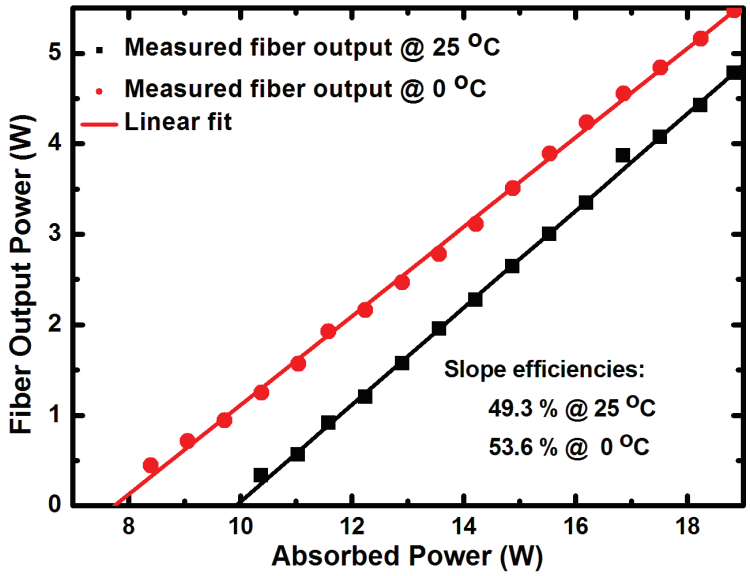


Figure 3. Measured slope efficiency of the Tm^{3+}-doped fiber at 25 °C and 0 °C.

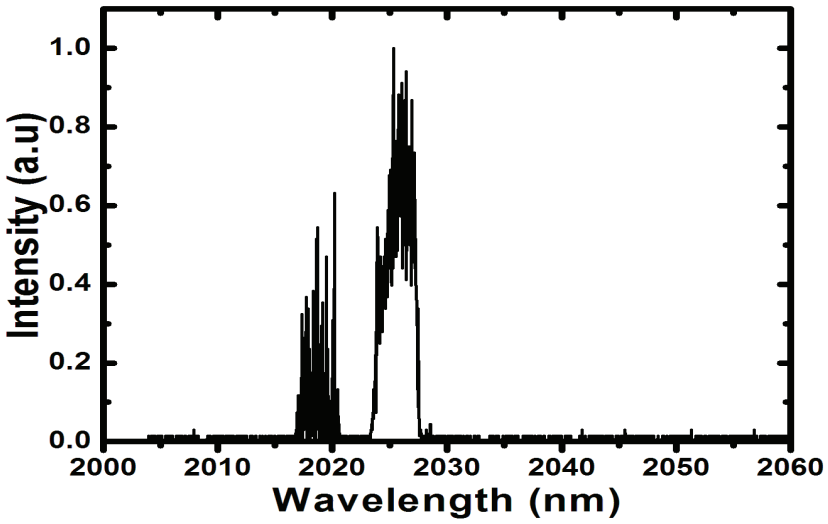


Figure 4. The measured spectral output of the Tm^{3+}-doped silica fiber at 25 °C.

4. Discussion

In the experiments the fiber was immersed in water to facilitate heat dissipation. Cooling the fiber improved the slope efficiency of the fiber and shifted the output towards longer wavelengths. This can be explained by considering the Boltzmann distribution of the population of the Stark levels of the ground state ($^3H_6$). Operating the fiber laser at lower temperatures reduces the population of the higher Stark levels in the ground state while increasing the lower Stark level population according to the Boltzmann distribution. This increases population inversion for the longer wavelength transitions relative to the shorter wavelength transitions [7], thus increasing the slope efficiency of the longer wavelength laser action.

5. Conclusion

We have have produced a continuous wave (CW) laser with optimal output power of $\sim 4\,\mathrm{W}$ at room temperature, $\sim 5\,\mathrm{W}$ in ice water (for $\sim 19\,\mathrm{W}$ of absorbed power) from a diode-pumped Tm^{3+}-doped silica fiber laser that operated at a slope efficiency of 49.3 % at 25 °C and 53.6 % at 0 °C, and a spectral output with multiple wavelengths around 2.02 μm. An output power of 10 W was obtained in high power pumping regime (for $\sim 32\,\mathrm{W}$ of absorbed power).

Acknowledgments

We acknowledge the University Stellenbosch, the African Institute for Mathematical Sciences (AIMS), the African Laser Centre (ALC), the CSIR National Laser Centre and the DAAD for funding this project, CorActive for the thulium doped silica fiber used and the Laser Research Institute (LRI) for the facilities and the equipment used.

References

[1] Scott C, Frith G and Samson B 2006 *Developments in thulium-doped fiber offer high powers. SPIE Newsroom*
[2] Moulton P F, Rines G A, Slobodtchikov E V, Wall K F, Samson B and Carter A L G 2009 *IEEE Journ. of Selected Topics in Quant. Electron.* **15** 85
[3] Hayward R A, Clarkson W A, Turner P F, Nilson J, Grudinin A B and Hanna D C 2000 *Electron. Lett.* **36** 711
[4] Jackson S D and Mossman 1992 *Appl. Opt.* **42** 2702
[5] Jackson S D 2004 *Opt. Commun.* B **230** 197
[6] Jackson S D and King T A 1998 *Opt. Lett.* **23** 1462
[7] Frith G, Lancaster D G and Jackson S D 2005 *Electron. Lett* **41** 687

Evaluation of WC-9Co-4Cr laser surface alloyed coatings on stainless steel

B A Obadele[1], P A Olubambi[1], S L Pityana[2] and A P I Popoola[1]
[1] Department of Chemical and Metallurgical Engineering, Tshwane University of Technology, P.M.B. X680, Pretoria, South Africa, 0001.
[2] Center for Scientific and Industrial Research- National Laser Centre, P.O.Box 395, BLD 46F, Pretoria, South Africa, 0001.

E-mail: abiodunob@tut.ac.za

Abstract. In order to examine the microstructure and hardness behaviour of WC cermets, coatings have been obtained by laser surface alloying technique. WC-9Co-4Cr particulate was injected onto the surface of AISI 304L austenitic stainless steel (ASS) under different laser processing variables. The morphologies and microstructures of the composite coatings were investigated using optical microscopy and high resolution field emission scanning electron microscopy (FESEM) equipped with energy dispersive spectrometer (EDS), while the phase changed were observed using x-ray diffraction (XRD). The surface hardness was determined using the Vickers microhardness tester. The decomposition of WC-9Co-4Cr into W_2C, C and W is as a result of low heat of formation of WC and low affinity of tungsten for Carbon. Free Co and C in the meltpool formed intermetallic phases such as Co_6W_6C and $M_{23}C_6$ (M=Fe, Cr, W). A considerable increase in hardness value of the matrix 246 $H_{v0.1}$ compared to the coating 1331 $H_{v0.1}$ was achieved when alloying was carried out at 2.0 kW laser power and scan speed of 0.6 m/min.

1. Introduction

Tungsten carbide (WC) belongs to the group of advanced ceramic materials with great industrial importance and well known as hardfacing material with Co or Ni alloys as binders. Apart from high hardness, WC has a unique set of properties; high melting point, wear resistance, good thermal shock resistance, thermal conductivity and good resistance to oxidation [1-3].

However, WC cermets melt and dissolve in melt pool formed by laser irradiation of the surface layer of the substrate due to a low free formation enthalpy of 38.5 kJ/mol. According to Nerz et al, [4] depending on time and temperature WC decarburize to form W_2C, and later free tungsten and carbon. The carbon can react with atmospheric oxygen to form CO/CO_2, which could be trapped in the melt pool during rapid solidification. As a result of this, pores are formed in the composite coating [2] and this could limit the wide-scale industrial recognition of this composite.

Laser alloying gives a perfect adhesion to the interface of the bulk steel and the coating with homogeneous microstructure due to strong marangoni convection caused by surface tension gradients and high cooling rates [5, 6]. With optimum laser processing parameters, a reliable coating that is free of cracks and pores can been produced on the matrix. In this paper, WC-9Co-4Cr composite coatings were prepared on 304L ASS the laser processing parameters.

2. Experimental methods

The specimen investigated was 304L stainless steel with dimension of 64x40x4 mm. The reinforced powder was a pure agglomerated and sintered WC-9Co-4Cr cermet as shown in figure 1.

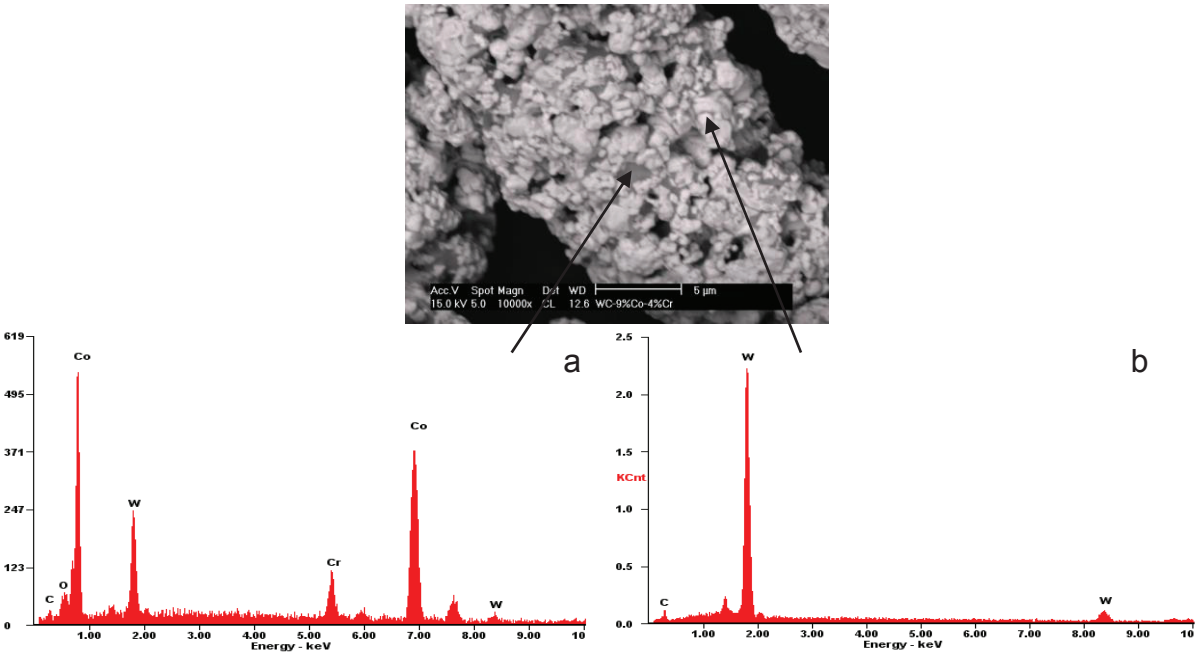


Figure 1. SEM image and EDS of WC-9Co-4Cr particle showing the (a) Co-Cr binder and (b) WC.

The average particle size was approximately 26 µm. The surface melting operation was conducted using a 4.4 kW Nd:YAG laser. The specimen surface was shielded by argon gas flowing at 2 L/min during the laser scanning to prevent oxidation. The laser beam was focused to a 3 mm diameter size while the laser power used was 2.0 kW and the scanning speed was varied from 0.6 to 1.2 m/min.

After laser injection, the coatings were sectioned, mounted and polished for metallographic examination as well as hardness testing. XRD of the coated surface was performed on a PW1710 Philips diffractometer, using monochromatic Cu Kα radiation at 40 kV and 20 mA while the phases were identified using X pert High Score Plus software. The microstructures and the distribution of hard phases were examined using a field emission scanning electron microscope (FESEM) equipped with energy dispersive spectrometer (EDS). Microhardness profiles of the cross section from the coated zone to the matrix at 100 µm intervals were measured using an EMCO TEST Durascan microhardness tester at a load of 100 g.

3. Results and discussion

Figure 2 shows the SEM micrographs of the specimen reinforced with WC-9Co-4Cr at laser power of 2.0 kW and varying scan speeds of 0.6 to 1.2 m/min. No pores and cracks were present from the SEM images. It could be seen that undissolved WC cermets are present at the top surface of the coatings and this vary for all the coatings with specimen 1 alloyed at laser power of 2.0 kW and scan speed of 0.6 m/min having most of such undissolved carbides . This could be attributed to the volume of carbides injected and the coated surface exposure to the atmosphere. At low scan speed, more carbides are injected into the melt pool and due to high cooling rate, the carbides at the top solidifies before being dissolved. High wettability

of WC and strong marangoni convection also aids the uniform distribution of dissolved carbides in the melt pool and segregation of the undissolved carbides at the coating surface as can be seen in figure 2.

Melting, dissolution and resolidification of WC took place to form WC, W_2C, Co_6W_6C and $M_{23}C_6$ (M=Fe, Cr, W) as indicated in the XRD analysis in figure 3. The EDS point analysis in figure 4 also confirms the presence of W, C, Co, Cr elements in the matrix.

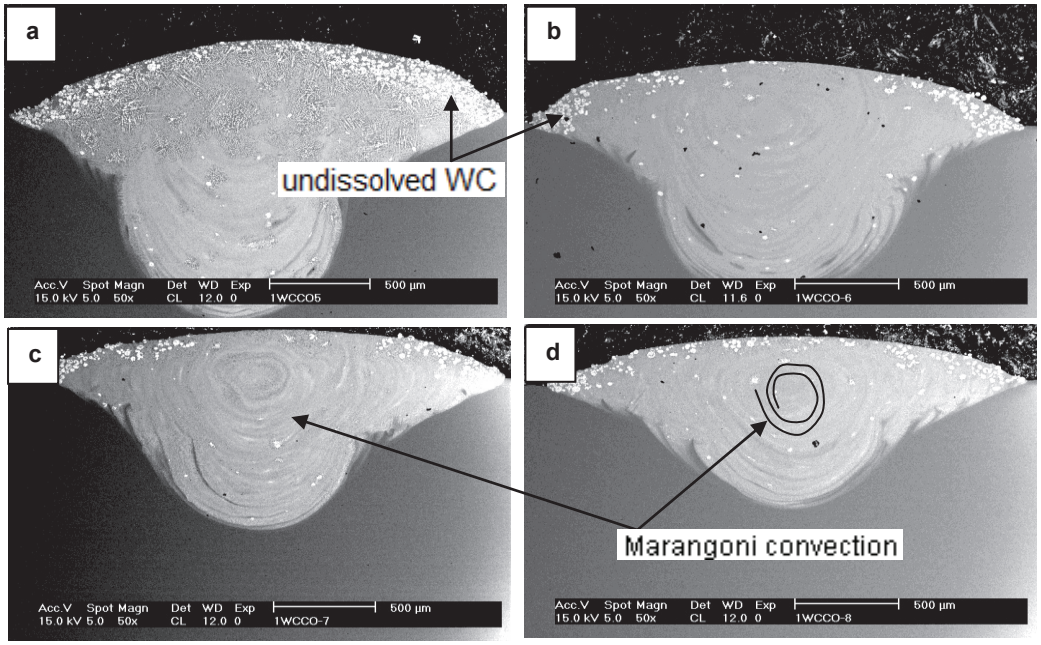


Figure 2. SEM micrograph of 304L stainless steel alloyed with WC-Co-Cr at laser power of 2.0 kW and scanning speed of (a) 0.6, (b) 0.8 (c) 1.0 and (d) 1.2 m/min

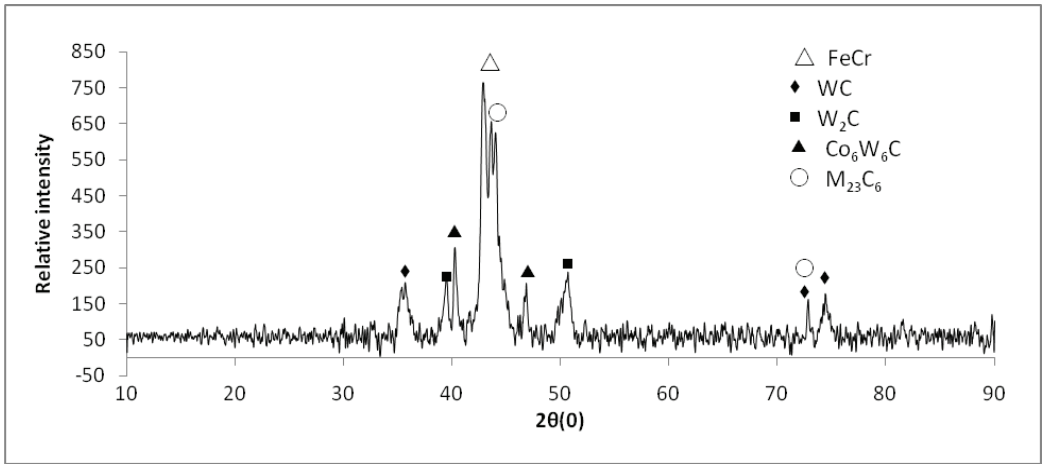


Figure 3. XRD analysis of specimen 1at laser power of 2.0 kW and scan speed of 0.6 m/min.

Microhardness measurement was performed on the cross section of the specimens from the coated layer to the matrix as shown in figure 5a. Specimen 1 at laser power of 2.0 kW and scan speed of 0.6 m/min gave the highest hardness of 1331 $Hv_{0.1}$. The coating thicknesses for the specimens were found to be approximately 1750, 1260, 1140 and 1055 µm respectively as shown in figure 5b. The variation in hardness and thickness is dependent on the speed of the laser beam. At lower scan speed, the laser beam irradiates the specimen longer, thus large and deep meltpool forms.

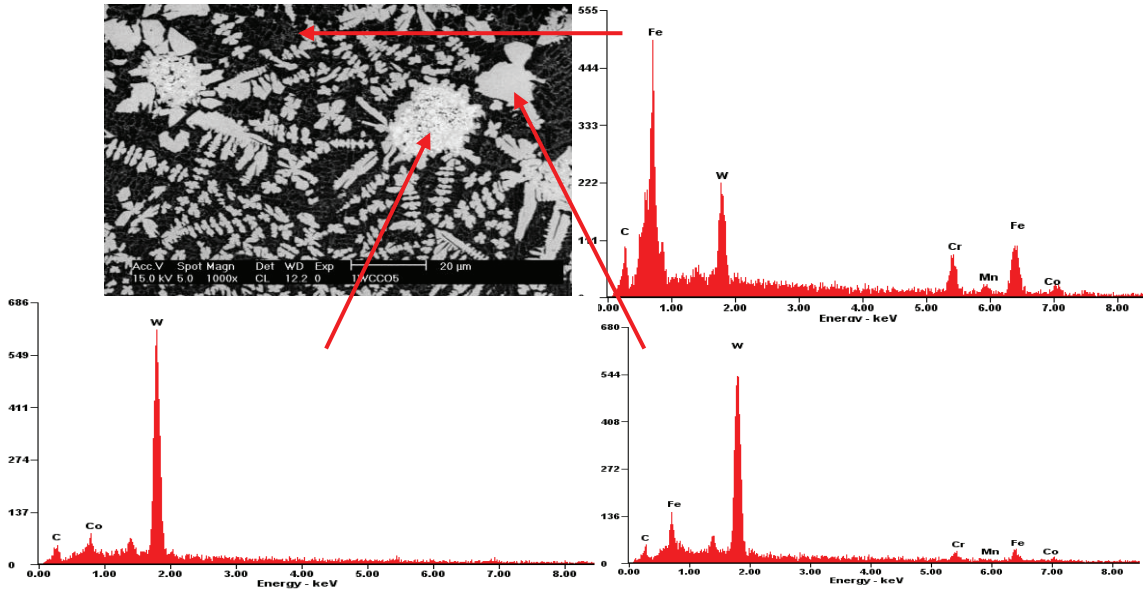


Figure 4. SEM image and EDS point analysis at laser power of 2.0 kW and scan speed of 0.6 m/min

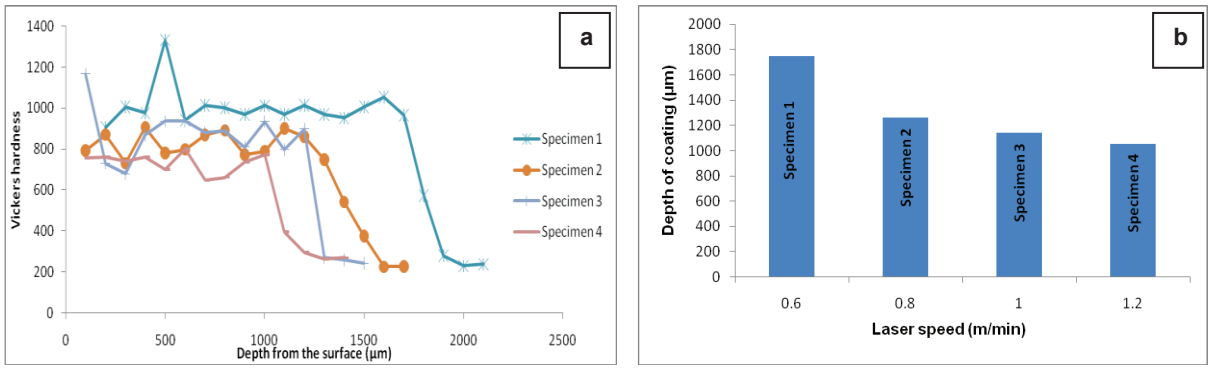


Figure 5. (a) Microhardness profiles of the alloyed layers and (b) Depth of coated zone as a function of laser speed.

4. Conclusion

It has been shown that 304L ASS matrix composite can be prepared by laser surface alloying without pores and cracks. Laser surface alloying with WC cermet powder result in an increase in hardness of 304L ASS. At laser power of 2.0 kW and scan speed of 0.6 m/min, hardness value was about 5 times that of the matrix. This is attributed to the formation of hard phases such as WC, W_2C and Co_6W_6C.

Acknowledgements

The authors gratefully acknowledge the financial support from the African Laser Centre (ALC) (LHEAB01 Task 300), Tshwane University of Technology and laser time from the National Laser Centre-CSIR.

References

[1] Wu P, Zhou C Z and Tang X N 1995 *Surf. Coat. Technol.* **73** 111–114
[2] Wu P, Du H M, Chen X L, Li Z Q, Bai H L and Jiang E Y 2004 *Wear* **257** 142–147
[3] Zhou S, Zeng X, Hu Q and Huang Y 2008 *Appl. Surf. Sci.* **255** 1646–1653
[4] Nerz T C, Nerz J E, Kushner B A and Riggs W L 1993 *Surf. Eng.* **9** 3–10
[5] Xiao-Hu Y and Xi C 2002 *Chin. Phys Lett.* **19** 788–790
[6] Gugliotti M, Baptista M S, Politi M J, Silverstein T P and Slater C D 2004 *J. Chem. Edu.* **81** 824–826
[7] Cheng F T, Lo K H and Man H C 2003 *Surf. Coat. Technol.* **172** 308–31

Polarisation encoded quantum key distribution in fibre

S Pillay[1], A R Mirza[1, 2] and F Petruccione[1, 2, 3]

[1] Quantum Research Group, School of Physics, University of KwaZulu-Natal, Private Bag X54001, Durban 4000, South Africa
[2]QZN Technology, Innovation Centre, Howard College Campus, University of KwaZulu-Natal, South Africa
[3]National Institute for Theoretical Physics, South Africa

E-mail: 206507614@ukzn.ac.za

Abstract. Quantum Key Distribution (QKD) employs the laws of quantum mechanics for the purpose of cryptography. Using QKD, two parties, commonly known as Alice and Bob, are able to share a random symmetric key that may be used to encrypt a message. Any form of eavesdropping requires the measurement of the properties of the quantum data carrier. Due to its quantum nature, this observation induces a disturbance into the system that is detectable by Alice and Bob. Thus, any compromise in the key distribution process will result in the detection of an eavesdropper and the discarding of the key. The BB84 protocol exploits the properties of a quantum two-level system (qubit), such as the polarisation of single photons. Due to the current Information and Communications Technology (ICT) infrastructure, QKD protocols are generally implemented using a fibre optical link as the quantum channel between Alice and Bob. The challenge of such an implementation however, is that polarisation of light is altered when passed through a fibre due to intrinsic birefringence. If the fibre is not disturbed, the change in the State of Polarisation (SOP) will be unique and constant. This can be compensated at the receiving end by rotating the SOP of each photon by a predetermined value with the use of a polarisation controller prior to the measurement of the photons. For a robust implementation, the polarisation controller must be automated in order to correct for changes in SOP in real time. Wavelength Division Multiplexing and Time Division Multiplexing are discussed as methods for testing the changes in SOP before compensation.

1. Introduction

The security of data transmission has become more important to society in recent years. Reliance on technology for global communication has made data transmission vulnerable to interception. Cryptographic protocols secure such data by using algorithms and a unique key to encode the data during transmission. Classical methods of passing a key between two parties may include physically couriering the key or electronically transmitting the key via a mathematical algorithm [1]. However this method is bound by technology and is susceptible to attacks. Quantum Key Distribution (QKD) relies on the laws of physics, and not the complexity of a numerical algorithm, to ensure the security of the key [2, 3]. Hence, increased computing power of an eavesdropper (Eve) will not render QKD obsolete [4]. Eve may attempt to copy or measure a quantum state, however, both these attacks violate the laws of quantum mechanics. Eve cannot copy the quantum state being passed between the transmitter (Alice) and the receiver (Bob) due to the `No Cloning' theorem [5]. Eve is also not able to implement a `man-in-the-middle' attack by measuring the quantum state and passing the same particle

to Bob. This would violate Heisenberg's uncertainty principle [6]. This shows that the presence of Eve introduces errors in the transmission of the key between Alice and Bob. The information shared between Alice and Bob is in the form of a quantum two-level system (qubit). The qubit state is shown as a linear superposition of two basis states, denoted by $|0\rangle$ and $|1\rangle$. As an example, the information of the qubit may be encoded into the state of polarisation of a single photon. In this case, $|0\rangle$ may represent a vertically polarised state and $|1\rangle$ may represent a horizontally polarised state [7].

2. The BB84 protocol

The first QKD protocol, known as the BB84 protocol, was developed by Bennet and Brassard in 1984 [7]. The protocol utilises two non-orthogonal polarisation bases, e.g. the rectilinear basis and the diagonal basis. For this example, the resulting qubits are the four available States of Polarisation (SOP): vertical, horizontal, right diagonal and left diagonal. Alice randomly chooses a state for a single photon and transmits it to Bob. Bob randomly chooses which basis to measure this photon in. If Bob chooses the correct basis, he will measure the correct state of polarisation with 100% probability. If Bob chooses the wrong basis, he could obtain either of the possible measurements, with a 50% probability. After transmission of the qubit stream, Alice and Bob announce the basis that they chose for each photon and keep only the measurements for which they used the same basis. The other measurements are discarded. Alice and Bob should now have a string of identical bits, called the sifted key. Table 1 shows an example of how Alice and Bob may establish a sifted key using the BB84 protocol. The sifted key established in this protocol must then undergo error correction and privacy amplification in order to minimize any information leaked to Eve through the implementational limitations. If the error rate is found to be above an acceptable threshold, due to the apparatus and background noise, Alice and Bob have detected Eve and they then discard the key [8].

3. A fibre quantum channel

The BB84 protocol can be implemented over a fibre optical channel. A fibre optic cable propagates light through the processes of total internal reflection and waveguide refraction [9]. A fibre link provides a waveguide, hence is independent of a line of sight connection between Alice and Bob. However, a fibre link is only suitable for shorter distances [10, 11]. This is because impurities in the fibre absorb photons and the qubits become too weak to measure after a length of 200 km [12]. Furthermore, a fibre optic cable is not able to maintain the SOP of light that is transmitted through it due to the birefringent effects of fibre [13]. Birefringence occurs due to asymmetries in the fibre optic cable, due to impurities in the fibre or manufacturing errors [14]. The asymmetry causes refractive differences for orthogonal components of the SOP of light. Therefore, the component aligned with the fast axis of the fibre is transmitted faster than the other component [9]. This causes the SOP of light to be rotated as it is transmitted through fibre. If the fibre is bent or subject to environmental stresses, such as heating or movement, the birefringent effects will vary with time [15]. It is clear that birefringent effects must be auto-compensated in order to successfully implement the BB84 protocol over a fibre channel.

4. Proof of principle

The setup shown in figure 1 served as a proof of principle that compensating for the birefringent effects of fibre will allow for successful fibre-based QKD. The laser pulses were attenuated to represent the power of a single photon. The power was calculated in accordance to the wavelength of the laser pulses using equation (1),

$$P = \frac{hc}{\lambda t} \tag{1}$$

Table 1: An example of establishing a sifted key using the BB84 protocol.

BB84 Protocol								
Alice sends	↑	↖	→	↖	↗	↗	↖	→
Bob's basis	+	x	x	+	x	x	+	+
Bob measures	↑	↖	↗	→	↗	↗	↑	→
Chosen bits	↑	↖			↗	↗		→
Bit values of the sifted key	0	1			0	0		1

where h is Plank's constant, c is the speed of light and t is the period of each laser pulse. The photons were then randomly polarised with a polarisation state generator (PSG). The laser and PSG served as Alice in the experimental setup. The photons were then transmitted through a fixed 1000 m length of fibre and the SOP of each photon was rotated by a manual polarisation controller (PC) to counteract the SOP rotation within this fibre. A half wave plate rotated the photons further before they were transmitted to a polarisation beam splitter (PBS). This was to counteract birefringent effects caused by the fibre housed with the PBS. The half wave plate also served as a means to change the measurement basis of the PBS. After the PBS, the photons were directed to one of two single photon detectors. The PC, half wave plate, PBS and single photon detectors served as Bob in this experimental setup. Figure 2 shows detection rates taken from the single photon detectors. For an incoming vertical SOP, the detector measuring the horizontal SOP had a greater detection rate before compensation. After compensation, the vertical SOP signal was at a maximum and the horizontal SOP signal was reduced to a minimum. These results show an effective compensation for a fixed fibre.

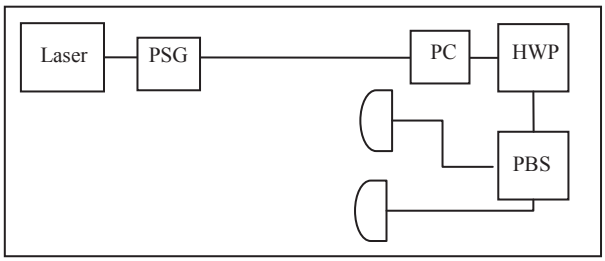


Figure 1: The experimental setup used as a proof of principle to show the importance of compensating for birefringence in fibre.

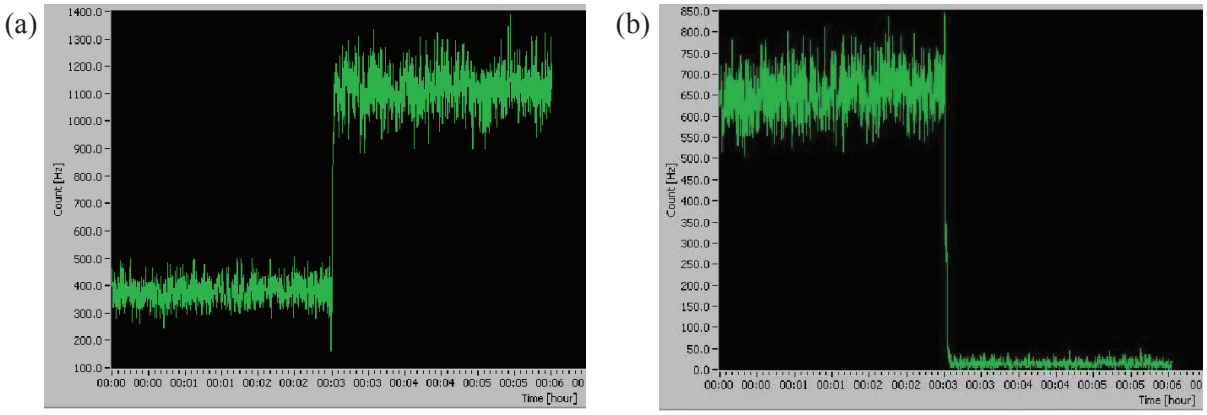


Figure 2: Single photon detection rates of (a) the vertical SOP signal and (b) the horizontal SOP signal. Rotating the initially vertically polarised signal back to its original state brings the vertical SOP signal to a maximum and reduces the horizontal SOP signal to a minimum.

5. Testing the changes in SOP

In order to efficiently compensate for birefringent effects in a disturbed fibre, the changes in the SOP of the photons must be measured so that an automated system may be developed. This can be done by sending test pulses through the fibre to indicate what degree of correction is needed. Time Division Multiplexing (TDM) or Wavelength Division Multiplexing (WDM) can be used to deploy test pulses into the system.

5.1. Wavelength division multiplexing

WDM utilises a calibration signal with a wavelength close to that of the single photons. This can be done in real time and is more effective for generating higher key rates [13, 16]. A disadvantage of WDM is the wavelength-dependence of birefringent effects [17]. Measurements were done to monitor the difference in the birefringent effects of fibre for different wavelengths. The experimental setup is shown in figure 3. Two laser channels, each of a different wavelength, were first given an initial SOP using polarisation controllers, PC1 and PC2. The channels were multiplexed and transmitted through a fibre optic cable of 1.69 km. Both the channels were passed through PC3 so that the same effects could be applied to both wavelengths. The channels were then demultiplexed and measured on separate polarimeters, POL1 and POL2.

The wavelength of λ_1 was fixed at 1550.12 nm and the difference in wavelength between the two channels was set to multiples of 0.4 nm. The settings of PC3 were gradually adjusted and ideally, both wavelengths should undergo the same changes. However, this small difference in wavelength can still result in significant differences in the change in SOP for the two signals [17]. Figure 4 shows the difference in angles of inclination and azimuth of the SOP of both channels with regards to the Poincaré sphere. As seen in this figure, the relative angles between the channels were not constant. Similar tests were carried out for various channel spacing and for a longer fibre length of 24.70 km. As before, the changes in SOP of the two channels were not the same.

A proposed method for auto-compensating the change in SOP caused by the fibre is to replace the PC shown in figure 1 with an SOP locker to return the test pulse to its predetermined state. The single photons would then undergo the same rotation, returning them to their original states as well. This principle was tested using the manual polarisation controller, PC3. Both channels were set to the same SOP and the fibre optic cable between Alice and Bob was disturbed. PC3 was used to return one of the channels to its original state and it was found that the second channel was not always corrected simultaneously. This shows that when using WDM, compensating the test signal will not ensure that the single photon signal will also be compensated.

5.2. Time division multiplexing

For TDM, the key transmission stops periodically and test pulses are transmitted through the fibre [18]. This can be time consuming, and the birefringent effects may vary before the next test is done. An SOP stability test was done by transmitting a polarised laser source through a length of fibre and

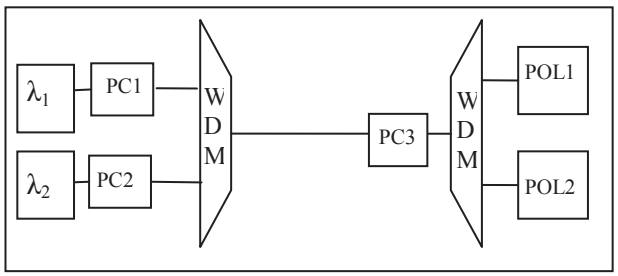


Figure 3: The experimental setup used to monitor the difference in the change in SOP for different wavelengths when subjected to the same conditions.

measuring the SOP with a polarimeter for 16 hours. Changes in SOP measured during this test were due to environmental changes, such as a temperature gradient. The results are shown in figure 5. These results were used to determine the period for a 1° change in the angles of azimuth and inclination in this environment. On average, these periods were 57.97 sec and 184.25 sec, respectively. Since orthogonal SOP's are located 180° apart on the Poincaré sphere, 180° can be considered as a 100% difference in information for the purpose of QKD. It is important to maintain a quantum bit error rate that does not exceed 1% so that errors caused by Eve are distinguishable from errors caused by the equipment and quantum channel. The average period for the angle of azimuth to change by 1% was 147.45 sec and the corresponding period for the angle of inclination was 503.72 sec. Therefore, when using TDM to test for the changes in SOP in this environment, the test signal must be deployed at least every 147 sec in order to ensure an acceptable quantum bit error rate. The duration of the test signal is dependent on the resolution of the compensator used for the system.

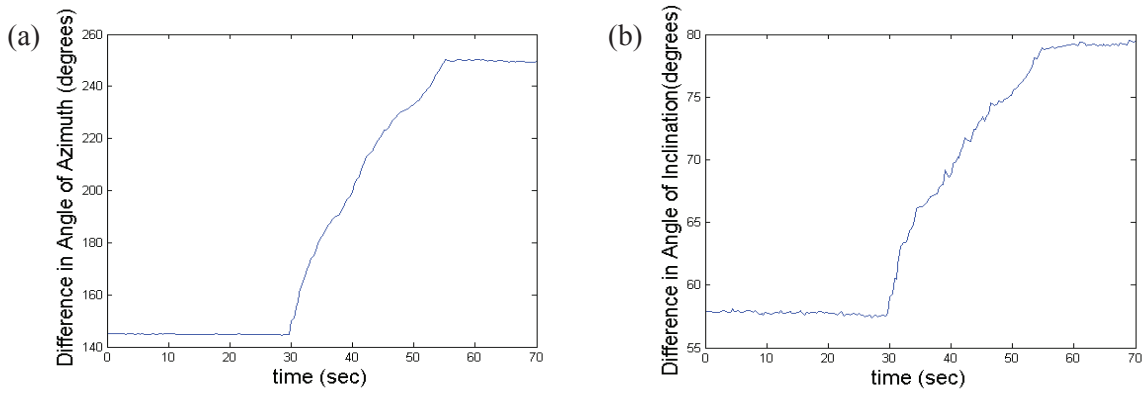


Figure 4: An illustration of the difference in the angle of (a) azimuth and (b) inclination between laser sources of different wavelengths. The wavelength spacing was 0.4 nm. The manual polarisation controller was gradually adjusted between the times of 30 sec and 55 sec. During this period, the difference of the angles of azimuth and inclination between the channels vary greatly.

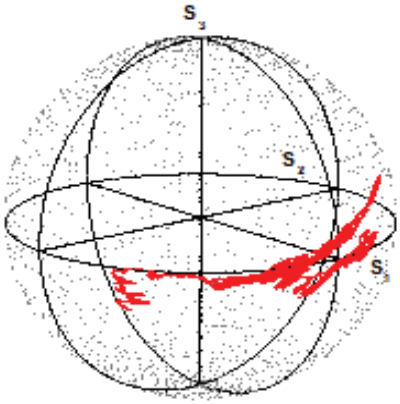


Figure 5: A plot of the SOP of a laser source over 16 hours. All changes in SOP occur naturally with changes in the surrounding environment.

6. Conclusion

The BB84 protocol utilises the SOP of single photons as qubits. This protocol can be implemented over a fibre quantum channel but the SOP of the single photons cannot be maintained during transmission from Alice to Bob. This challenge can be overcome by auto-compensating for the birefringent effects. In order to monitor the changes in SOP, TDM or WDM can be used to deploy test signals into the system. An effective compensation system must have both speed and accuracy. WDM does not allow for total accuracy due to the wavelength dependence of birefringence. TDM does allow

for an accurate compensator but for this environment, test pulses had to be deployed every 147 sec in order to ensure a quantum bit error rate of below 1%. Effectively compensating for the birefringent effects in fibre will allow a realisation of polarisation encoded QKD over a fibre channel.

Acknowledgements
This work is based on research supported by the South African Research Chair Initiative of the Department of Science and Technology and National Research Foundation. S.P. would like to thank Tim Gibbon and Jesman Changundega of the Optical Fibre Research Group at NMMU for their assistance with the experiments.

References

[1] Diffie W and Hellman M E 1976 New Directions in Cryptography *IEEE Trans. on Info. Theory* **IT-22** 644-654
[2] idQuantique [Online] http://www.idquantique.com/images/stories/PDF/clavis2-quantum-keydistribution/ clavis2-whitepaper.pdf.
[3] Buttler W T, Hughes R J, Kwiat P G, Luther G G, Morgan G L, Nordholt J E, Peterson C G and Simmons C M, Free-space Quantum Key Distribution 1998 *Phys. Rev. Lett*. **57** 2379-2382
[4] SECOQC [Online] http://www.secoqc.net.
[5] Wootters W K and Zurek W H 1982 A Single Quantum Cannot Be Cloned *Nature* **299** 802-803
[6] Krane K 1996 *Modern Physics* (New York: John Wiley& Sons)
[7] Bennet C and Brassard G 1984 Quantum cryptography:Public Key Distribution and Coin Tossing *Proc. of IEEE International Conference on Computers Systems and Signal Processing* 175-179
[8] Capraro I 2008 Advanced Techniques in Free Space Quantum Communication PhD Thesis University of Padua 33-34
[9] Hecht E 2001 *Optics* (Reading, MA: Addison-Wesley)
[10] Hubel H, Vanner M R, Lederer T, Blauensteiner B, Lorunser T, Poppe A and Zeilinger A 2007 High-fidelity Transmission of Polarization Encoded Qubits from an Entangled Source over 100 km of Fiber *Optics Express* **15** 7853-62
[11] Hughes R J, Morgan G L and Peterson C G, 2000 Quantum Key Distribution Over a 48 km Optical Fibre Network *Journal of Modern Optics* **47** 533-547
[12] Gisin N, Ribordy G, Tittel W and Zbinden H 2002 Quantum Cryptography *Review of Modern Physics* **74**, 145
[13] Xavier G B, Walenta N, Vilela de Faria G, Temporo G P, Gisin N, Zbinden H and von der Weid J P 2009 Experimental Polarization Encoded Quantum Key Distribution over Optical Fibres with Real-time Continuous Birefringence Compensation *New Journal of Physic***11** 045015
[14] Luna Technologies [Online] www.lunatechnologies.com/files/22pmdweb.pdf
[15] OZ Optics [Online] www.ozoptics.com/ALLNEW_PDF/APN0005.pdf
[16] Xavier G B, Vilela de Faria G, Temporo G P and von der Weid J P 2008 Full polarization control for fiber optical quantum communication systems using polarization encoding *Optics Exress* **16** 1867-1873
[17] Chen J, Wu G, Xu L, Gu X, Wu E and Zeng H 2009 Stable Quantum Key Distribution with Active Polarization Control Based on Time-division Multiplexing *New Journal of Physics* **11** 065004
[18] Ma L, Xu H and Tang X 2006 Polarization recovery and auto-compensation in Quantum Key Distribution Network *Proc. of SPIE* **6305** 630513

Magneto-optical trapping: first step towards BEC

M Semonyo[1], M Rehn[1] and F Petruccione[1, 2]
[1]School of Physics, University of KwaZulu-Natal, University Road, Westville, Durban.
[2]National Institute for Theoretical Physics, University of KwaZulu-Natal, University Road, Westville, Durban.

E-mail: semonyo@ukzn.ac.za, petruccione@ukzn.ac.za

Abstract. Experimental demonstration of Bose-Einstein condensation (BEC) is one of the most exciting developments in physics research in recent years. Its success has made significant contributions to our understanding of atomic and molecular physics, precision metrology and future technologies in quantum optics and quantum information science. We discuss briefly an experiment for magneto-optical trapping of Rubidium atoms constructed in our laboratory as a preliminary step towards obtaining a BEC. We put emphasis on the three main parts making up the experiment, namely: the lasers, the magnetic fields and the vacuum system.

1. Introduction

Laser cooling and trapping of neutral atoms has become an important area of research in atomic and molecular physics. The first demonstration using neutral atoms was made by Phillips and Metcalf in 1982 using sodium [1]. They slowed down the atoms in an atomic beam using a counter-propagating laser beam and a spatially varying magnetic field. This was followed by an important breakthrough of Steven Chu *et al.* [2] who used three pairs of counter-propagating laser beams to create a 3D optical trap. The atoms were cooled at the intersection of the beams resulting in what is known as 'optical molasses'. They managed to cool them to a temperature of 240 micro-Kelvin. The temperature was measured using the time-of-flight technique. This first optical trap was followed by the addition of anti-Helmholtz coils to generate a quadrupolar magnetic field whose centre occurs at the intersection of the beam to create the most commonly used trap today, the magneto-optical trap (MOT).

The MOT has served as an important step towards achieving a Bose-Einstein condensate (BEC). The theory behind BEC was first explained by Bose in 1924 for photons and added upon by Einstein in 1925 for a system of bosons with mass. A BEC is a state of matter in which the de Broglie wavelength of the atoms becomes larger than the inter-atomic separation and all atoms are said to be in the ground energy state. To obtain a BEC, sufficiently low temperatures and high densities of the atoms must be attained. In the past sixteen years since the first creation of BEC many techniques have been studied for example magneto-optical trapping followed by evaporative cooling in a magnetic trap [3, 4] and an all-optical trap [5]. In this paper we discuss the achievement of trapped ^{87}Rb atoms in a home-built vapour cell as a preliminary step towards BEC.

2. Experimental setup

The experimental setup consists of three main sections: lasers, the MOT and the vacuum system.

2.1 Lasers.

We use external cavity diode lasers (ECDL) from Moglabs operating at 780 nm in a Littrow configuration. The first diffracted order from a grating is fed back into the laser diode to provide frequency selective optical feedback and the zeroth order diffracted light forms the output. The output beam then goes through an optical isolator to prevent back scattering into the diode and through various optics (waveplates, beamsplitters and mirrors) and a fraction of the beam is directed to a

saturation absorption spectroscopy (SAS) set-up as illustrated in figure 1. SAS is used to lock the lasers at the right frequencies.

In SAS, a relatively weak probe beam is sent through a Rubidium vapour cell at room temperature and is incident on a photodiode. In addition a counter-propagating pump beam of stronger intensity is sent across the probe and transmittance of the probe beam is recorded. Atoms moving in the direction opposite to the propagation of the pump beam see a blue shift in the frequency of pump beam and see a red shift in the frequency of the probe beam. Stationary atoms however see both the pump and probe beams at the same unshifted frequency and if the intensity of the pump beam is high enough, the ground state will be significantly depleted causing a decrease in the probe beam absorption. The plot of absorption versus laser frequency will result in a Doppler broadened profile with a saturated absorption dip also called Lamb dip. To record the spectrum in repump laser, we have two probe beams, one without the pump beam and one with. We first record the transmission of probe beam without pump beam, to get a Doppler broadened spectrum. After this we record the probe-pump beam combination to get hyperfine features in the spectrum. Subtraction of the two gives us a Doppler free spectrum which is used to lock onto the correct frequency.

For the cooling laser SAS is done with the probe and pump beam on the same optical path. The rubidium vapour cell used in the cooling laser SAS has a copper coil that is used to Zeeman-shift some peaks in the spectrum to facilitate the necessary detuning of the cooling beams.

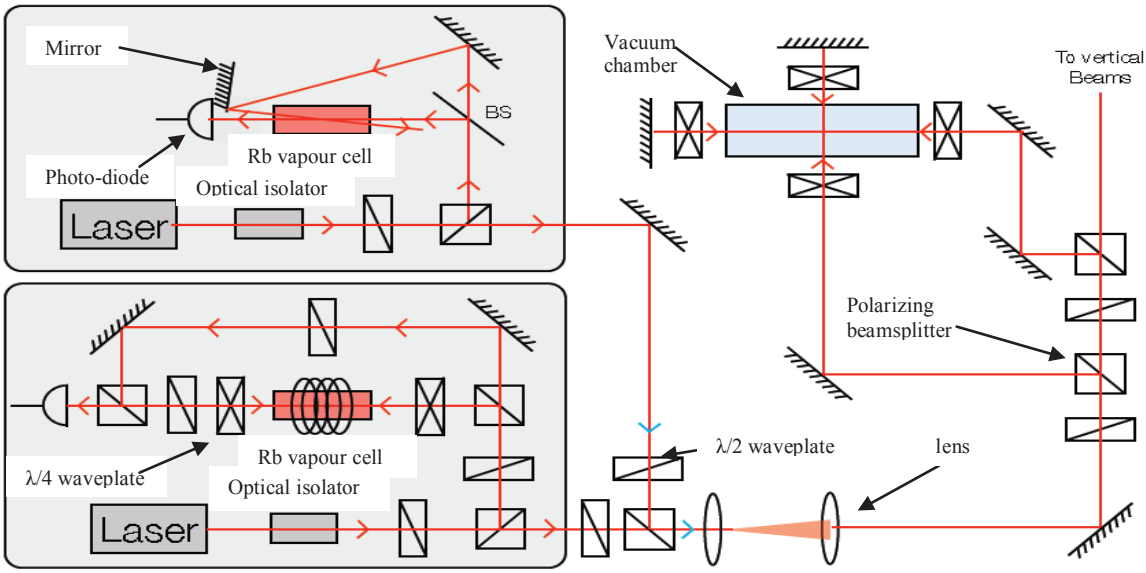


Figure 1. Schematic illustration of the experimental layout. The two boxed areas show the cooling and repump lasers and the saturation spectroscopy set-up. The top laser is used as a repump and the bottom laser is the cooling laser.

2.2 The magneto optical trap.

The trap consists of three counter-propagating laser beams slightly red detuned from the atomic transition. They intersect at the centre of a quadrupole magnetic field generated by a pair of coils in an anti-Helmholtz configuration. The laser beams are circularly polarized. At the centre, the magnetic field is zero and it grows linearly in all directions resulting in energies of the Zeeman sublevels of the excited state increasing linearly. The selection rules for electronic transitions allow for the magnetic field to dictate how the laser beams exert force on the atoms ultimately resulting in trapping. The principle of the MOT is illustrated in figure 2.

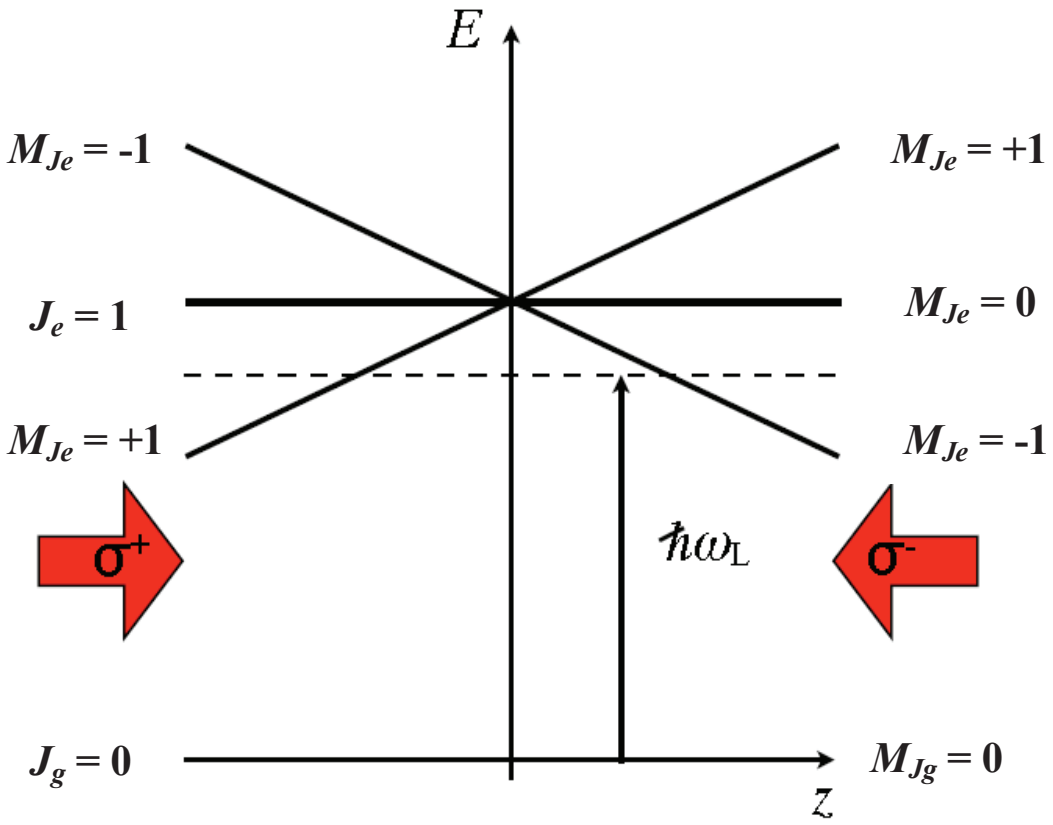


Figure 2. Schematic illustrating the operation of the MOT. For simplicity a J = 0 to J = 1 transition is shown, with M_{Jg} and M_{Je} referring to the excited and ground states respectively. The frequency of the laser is ω_L. The magnetic field splits the excited state energy levels away from the origin resulting in atoms preferring to absorb light travelling toward the trap centre thus creating a restoring force on the atoms.

Rubidium has two isotopes namely, ^{87}Rb and ^{85}Rb with excited states $5^2P_{1/2}$ and $5^2P_{3/2}$ and ground state $5^2S_{1/2}$. For our experiment we use ^{87}Rb whose ground state has hyperfine levels split into F = 1 and 2. The excited state has hyperfine splitting of F= 3, 2, 1 and 0. The energy level structure of ^{87}Rb is shown in figure 3. The cooling transition is between the F = 2 ground state and the F = 3 excited state. During this cycle, some of the atoms are transitioned to the F = 1 ground state, and are thus lost to the cooling cycle. To prevent this, a second 'repump' laser, resonant with the F=1 ground state to the F=2 excited state, is used to pump the atoms into the F = 2 of the excited state.

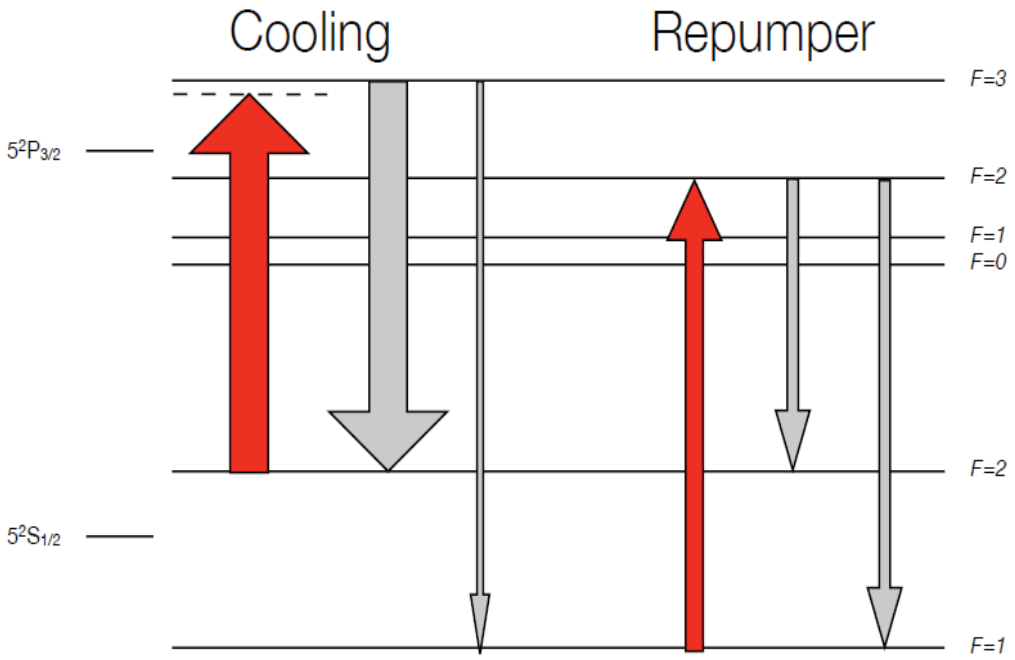


Figure 3. The energy levels of ^{87}Rb showing the cooling and repump transitions. The cooling laser is slightly detuned from the atomic transition. There is a small probability that some of the atoms may decay into the F =1 state where they would them require the repump laser to excite them into the F = 2 of the excited state.

2.3 The vacuum system.

The vacuum system consists of a single chamber where the atoms are dispensed as a vapour at room temperature from a Rb getter source from Alvatec. The chamber is home built from standard glass and is glued together to make a rectangular cell. The cell is placed at the centre of the anti-Helmholtz coils that are 9 cm apart with gradient up to 10 Gauss/cm. The current through the coils is sourced from a standard power supply. To maintain the high vacuum needed (10^{-7}mbar) the chamber is continuously pumped by a 10 litre/sec ion pump (Gamma Vacuum) with additional pumping provided by turbo-molecular pump usually activated at the beginning of the evacuation process. The entire vacuum system is mounted on an optical table. In figure 4 we show the cloud of atoms trapped in the MOT in our laboratory. The image was obtained using a CCD camera.

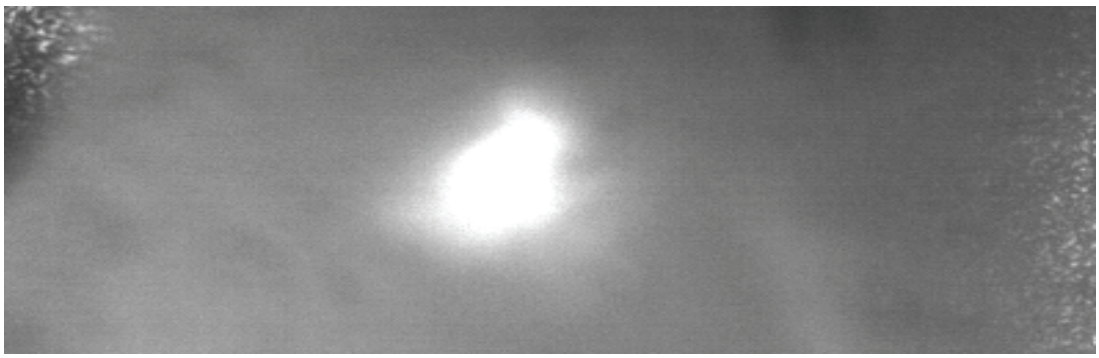


Figure 4. Image of the cloud of trapped ^{87}Rb as seen by a CCD camera connected to a computer. The actual diameter of the cloud is close to 2 mm.

3. Summary and results

The MOT was designed and built entirely in our laboratory with a home built glass cell for a vacuum chamber. We have set it up as a proof of principle and gained a working knowledge of the experiment to be able to take it further towards the realization of a BEC. Several improvements are planned, such as replacing the glass cell, implementation of acousto-optic modulators for precise frequency control, developing diagnostic system and designing a double MOT system for BEC.

Acknowledgements

This work is based on research supported by the South African Research Chair Initiative of the Department of Science and Technology and National Research Foundation.

References

[1] Phillips W D and H Metcalf 1982 *Phys. Rev. Lett.* **48** 596
[2] Chu S, Hollberg L, Bjorkholm J E, Cable A and A Ashkin 1985 *Phys. Rev. Lett.* **55** 48
[3] W Ketterle an N J van Druten 1996 *Adv. At. Mol. Opt. Phys* **181** 181-236
[4] Anderson M H, Ensher J R, Matthews M R, Wieman C E and Cornell E A 1995 *Science* **269** 198-201
[5] Kinoshita T, Wenger T and Weiss D S 2005 *Phys. Rev.* A **71** 011602(R)

Phase transition in hydrothermal synthesized VO_2(M) nano-crystals: an X-ray diffraction study

A Simo[1,2], R Madjoe[2], L Edomwonyi-Otu[1,3] , A Gibaud[4] and M Maaza[1,5]

[1] Nanosciences African Network, Nanoscience Laboratories, Materials Research Dept., iThemba LABS-National Research Foundation, Po Box 722, Somerset West 7129, South Africa

[2] Physics Dept., University of Western Cape, Belleville, South Africa

[3] Chemical Engineering Dept, Ahmadu Bello University, Zaria, Nigeria. 870001

[4] Laboratoire de Physique de l'Etat Condense, Dept. De Physique, Universite du Maine, UMR CNRS 6087, Le Mans- France

[5] Faculty of Sciences, Pretoria-Tshwane University of Technology, Private Bag X 680, Pretoria, South Africa.

E-mail: simo@tlabs.ac.za

Abstract. We report on the phase transition of VO_2(M) nano-particles synthesized by hydrothermal processing. The study was carried out by X-ray diffraction with a focus on the major Bragg peaks located at 37.1 and 36.9°. The reversible crystallographic transition from monoclinic to tetragonal phase at about 67°C is close to the bulk value. It was observed that the dynamic of the tetragonal and monoclinic one are complementary. This structural phase transition studies were complemented by elemental composition, selected area electron diffraction and Fourier transform attenuated total reflection infrared spectroscopy.

1. Introduction

The unusual properties of the metallic state near the insulating transition have been most extensively studied in d-electron systems. Among these materials, the Vanadium dioxide (VO_2) has intrigued researchers for five decades since its discovery by Morin in 1959, especially with regard to its thermochromic properties in the Infrared spectral range [1]. This latter peculiar behaviour determines the ability of VO_2 based coatings to regulate the infrared radiations and so acting as an intelligent reversible window filter. In addition, it has potential applications in thermal sensing and switching devices [2]. VO_2 is considered as an archetypical system with a conductivity change of several orders at the critical temperature T_c (341.1 K) accompanied by a transition in lattice structure from a monoclinic to a tetragonal phase and a dramatic increase of the infrared modulation [3]. Various micro/nanostructures of VO_2 such as nanoribbons, nanobelts, nanorods, nanowires, hollow microspheres by facile template-free process have been prepared [4]. Few reports are available to date on direct synthesis of pure VO_2 (M) micro-nano radiative platelets with high crystallinity without thermal annealing due to the abscence of propagation of the transition across the grain boundaries and the dissipation of the intrinsic stress. Grain size is widely believed to determine the structure-dependent hysteresis curve in the phase transition [5]. Figure 1 presents structural information during VO_2 metal insulator transition (MIT) with an increase in symmetry both in the VO_6 octahedron and in the V atomic chains where the V-V pairs undergo the elongation, shortening and twist from the

zigzag-type to a linear chain. Different techniques have been used to investigate the kinetics of the transition via electron –correlation-driven Mott transition, structure-driven Peierls transition or the cooperation of the two mechanisms [3] and synchrotron X-ray scattering in conjunction with real space analysis showing that the transition from low temperature monoclinic to high temperature tetragonal phase occurs in a first order manner with coexistence of the two phases [6]. It has been shown that the phase of electronic property occur simultaneously suggesting for the MIT of VO_2 the existence of a cooperative mechanism of a structural driven and electron correlation mechanisms.

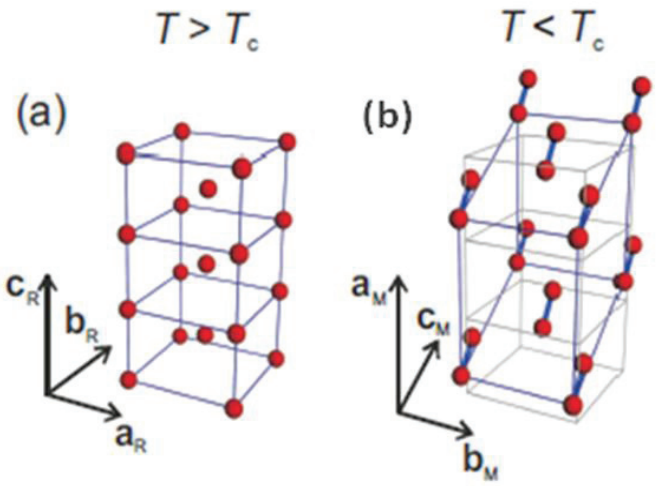


Figure 1: (a) the metallic high temperature tetragonal rutile form of VO_2 with a single V-V distance and (b) the insulating low-temperature monoclinic form showing dimerized chains of alterning short and long V-V distances along the c-axis [6].

This contribution is highlighted in the present study via in-situ x-ray diffraction versus temperature within a thermal range of 25-110°C on highly crystalline VO_2(M) nano-crystals. It has been reproducibly observed that such hydrothermal synthesized nano-crystals exhibit an average transition temperature at around ~67°C equivalent to the bulk value pointing out that the investigated snow-flake type nano-crystals could be stress/strain free.

2. Experimental techniques

All the reagents were of analytical grade and used without further purification. 0.75ml of Sulfuric acid H_2SO_4 (Kimix, 98%) followed by the drop-wise addition of 0.25ml of $N_2H_2{\cdot}2H_2O$ (Merck) were added into an aqueous suspension (10 mL) containing 0.45g of V_2O_5 (Alfa Aesar). After being warmed at 95˚C while stirring, the solution changes from yellow to green, then green to blue, characteristic of the presence of V^{+4} ions in the solution, finally blue to gray depending on the concentration of NaOH used to stabilize the dissolved precipitates. Hydrothermal synthesis was carried out in a Teflon-lined autoclave at ~230˚C for ~48h. Then the content was air-cooled at room temperature followed by the filtration of the formed precipitates. The final black product was washed thorougthly with water and ethanol for the exfoliation of bulk layered V_2O_5 and then dried at about 60˚C for 12h in an oven.

3. Results and discussion

3.1 Morphological studies and elemental analysis

The morphology of the synthesized crystals was observed by a scanning electron microscopy Nova NanoSEM 230 equipped with an elemental EDAX system. As typically reported on figure 1, the surface morphology of the samples exhibit snowflake radiative platelike aggregates in 3-D structures,

with an anisotropic orientation in shape. This was found to be correlated to the V_2O_5 initial concentration and the high surface-energy liable with chemical reactions of the medium. The correlation is due to the deviation from the minimum free energy configuration. The state of thermodynamic equilibrium is restored by condensation of nuclei of the reaction product [7]. Particles size and microstructure are affected by supersaturation, nucleation, growth rates, colloidal stability and recrystallization. In the present study, the hydrothermolysis for the aqueous solvent, enhancing the crystal growth for a better crystallization with narrow particle size distribution and giving nanostructures particules with large surface area, favorises the agglomeration via the attraction of the Van Der Waals forces and the tendency of the system to minimize the total interfacial energy. The dimension of the crystallites is in the range of ~400 nm to 3µm. The elemental composition of the material acquired from the X-rays emitted is shown in figure 3. The snow-flake crystals consist of ~35.24 and ~64.76 wt% of oxygen and vanadium respectively demonstrating that the stoichiometry of the compound is almost VO_2.

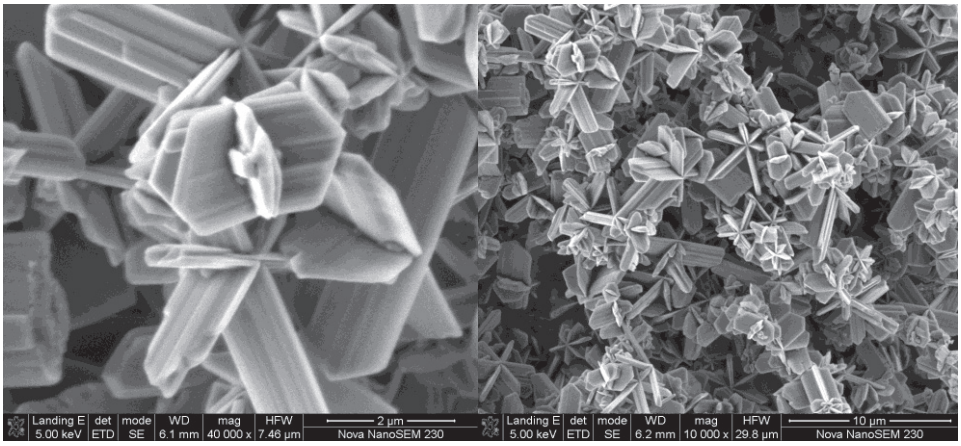


Figure 2: High and low magnification scanning electron microscopy images of a typical VO_2 synthesized crystal.

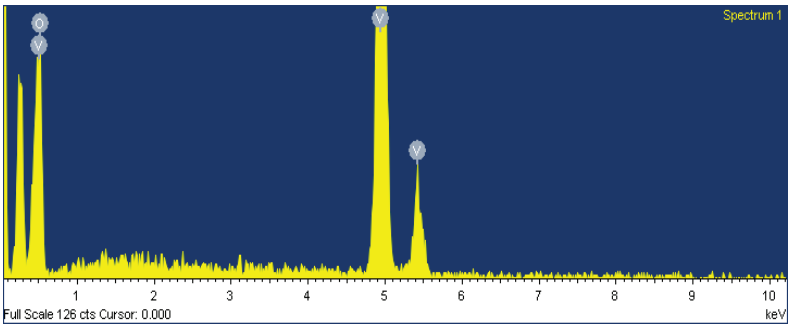


Figure 3: Elemental composition of a typical VO_2 synthesized crystal.

3.2 Vibrational spectroscopy studies

Chemical and kinetic information related to band absorption for multiple internal reflections have been performed by attenuated IR total reflection spectroscopy using a Perkin Elmer Spectrum 1000 FTIR-ATR spectrometer. Initially and as a reference, V_2O_5 powder was investigated. Its IR spectrum exhibited the 1000.703 cm^{-1} and 782.392 cm^{-1} bands, characteristic of the intermediate oxidation state V^{+5} to V^{+4} of V=O bond. The snow-flake synthesized nano-crystals exhibited four bands; at 840.65cm^{-1} describing the coupled vibration V=O and V-O-V (figure 4). This concurs with the transition from VO_2 (amorphous) to VO_2 (M). The 420 cm^{-1} band is a weak vibration of the absorption band of V-O bond while 544.5cm^{-1} is assigned to the V-O-V octahedron bending modes. In accordance to the previous IR studies of Sorapong et al [8], such a spectrum is a characteristic of pure VO_2(M) phase.

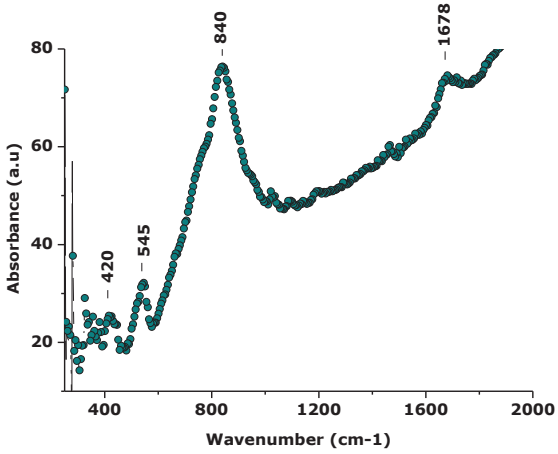


Figure 4: Typical room temperature ATR-FTIR of the snow-flake like synthesized nano-crystals.

3.3 Phase transition and in-situ x-rays diffraction studies

The crystalline structure was determined by X-ray diffractometry in a θ-2θ mode with CuKa 1 (AXS Bruker, λ=1.54056Å). Figure 5 reveals sharp peaks and intense diffraction which demonstrate that the sample is well crystallized without any additional phases and any presence of the reference. All peaks are indexed as VO_2 (M) according to a JCPDS card 00-043-1051 with the lattice constants a, b and c are of 5.75170 Å, 4.53780 Å and 5.38250 Å respectively, β=122.64˚. XRD analysis shows that VO_2 (M) with a space group of $P2_1/c$ has a strong preferential orientation along (011) plane which agrees with Shidong et al's investigations [9]. The different peaks are indexed as: (011) with a preferred orientation, (200) the 2nd main Bragg peak (002), (012), (210), (-302), (102), (211) the 3rd main Bragg peak. To investigate the Mott phase transition of the synthesized snow-flake like crystals which exhibit a priori pure VO_2(M) phase, in-situ x-rays diffraction was conducted on powder samples. The heating of the samples during the x-ray measurement was performed using a Peltier thermoelectric heat pump with a regulation module of ~0.1°C. The ability to automate the scattering intensity and particle size trend measurements is a major advantage in many applications in this set up. The change of particles size is due to the supersaturation showing predominant influence on the morphology of precipitates giving three dimensional superlattices (nanoplates) denoted Ostwald ripening.

Processes as aggregation, solubilisation, sedimentation and change in molecular conformation can be followed by the scattering of the intensity of the samples as a function of temperature. Figure 6 depicts the evolution of the major Bragg peaks such as the one with a preferential orientation i.e (011) one versus temperature. More accurately, and in the range of 50°C to 75°C , the intensity of two

different diffraction peaks at 37.1° corresponding to (200) monoclinic and its equivalent at 36.9° were followed.

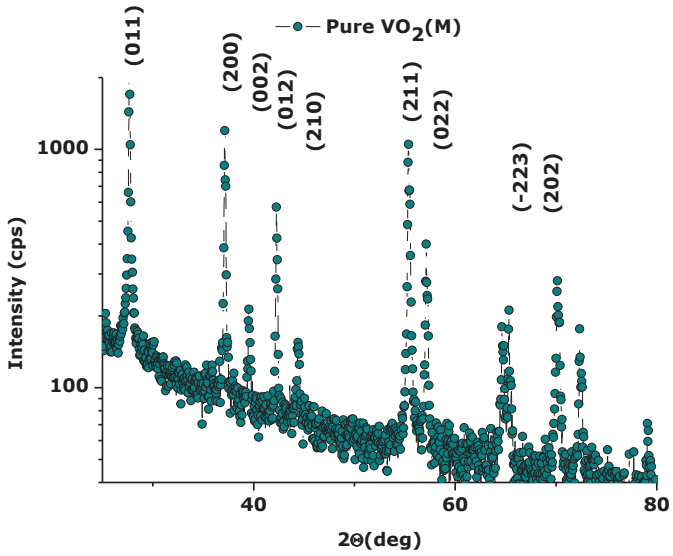


Figure 5: Typical room temperature indexed X-rays diffraction of the snow-flake like synthesized nano-crystals.

Figure 6a reports the isoline 3-D view of the X-ray spectra showing the shifting of the entire transition monoclinic-tetragonal with temperature. The evolution is continuous at the operational time scale. While there is net decrease of the monoclinic Bragg peak intensity (situated at about 37.1°), there is a steady increase of the intensity of the Bragg peak corresponding to the tetragonal peak located at about 2θ of 36.9°. This trend concurs with Joyeeta et al's predictions [10]; the area under the monoclinic peak decreases as the tetragonal fraction grows. Both phases coexist in equal fractions at around 64°C as substantiated by figure 6b. The VO_2 (M) nano-crystals are entirely tetragonal at 70°C. Naturally, the Bragg peak shift during the transition is attributed to thermal expansion and VO_2 lattice. One can notice that the transition seems to start as early as 55°C.

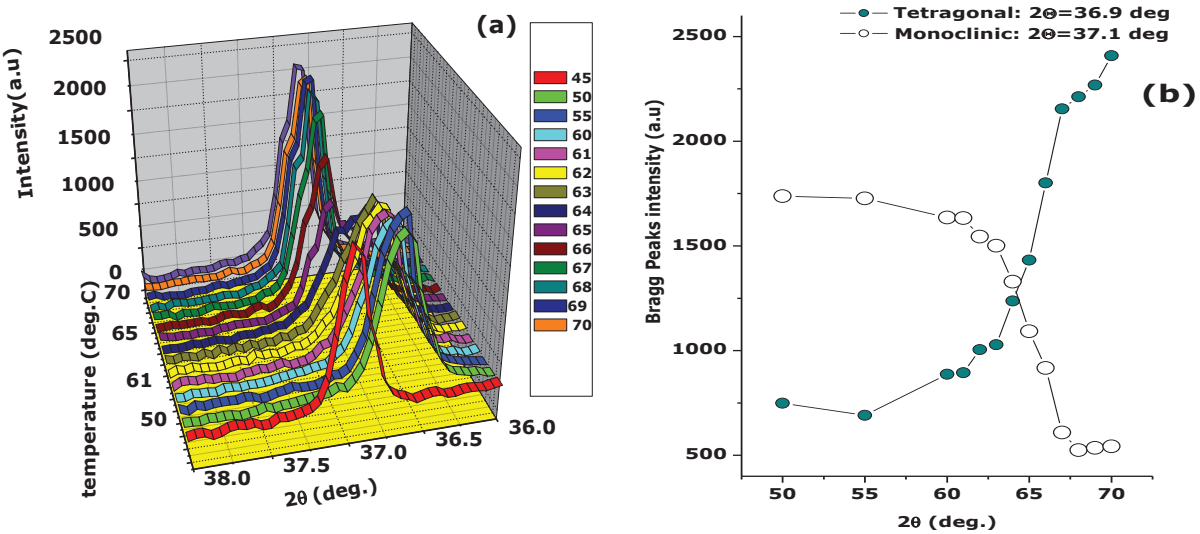


Figure 6: Thermodiffraction of Bragg peaks at 36.9° and 37.1°: isoline 3-D view of X-ray spectra (a) and evolution of the monoclinic and tetragonal phases versus temperature (b).

At around 64°C, the intensity of both monoclinic and tetragonal main Bragg peaks are quasi equal. Thermal evolution of the intensity of the 37.1° Bragg peak reflection increases continuously between 55 and 70°C due to the rearrangement of the structure while the monoclinic one seems to stabilize at about 68°C as per reported on figure 6b.

4. Conclusions

In summary, snow-flakes aggregated VO_2(M) nano-crystals have been demonstrated to be synthesized by hydrothermal process. Their phase transition from monoclinic to tetragonal phase with temperature was followed using in-situ x-rays diffraction. It was found that the crystallographic phase transition induced thermally is not sharp and ultrafast in the case of investigated VO_2(M) hydrothermally engineered nano-crystals. This is due to the conditions synthesis as concentration of the reagents, pH, temperature, surfactant and mechanical agitation influencing the kinetics and the thermodynamics rate of the reaction**.** It was clearly observed that the tetragonal phase grows in favor of the monoclinic one.

Acknowledgements

This work was sponsored by the African Laser Centre (ALC) in collaboration with iThemba LABS-MRD/ National Research Foundation of South Africa, Tshwane University of Technology and the Council for Scientific and Industrial Research in South Africa, as well as the Abdus Salam ICTP-Trieste., Italy.

References

[1] Morin F J 1959 *Phys. Rev. Lett.* **3** 34
[2] Kana J B, Ndjaka J M, Ngom B D, Fasasi A Y, Nemraoui O, Nemutudi R, Knoesen D and Maaza M 2010 *Optical Materials* **32** 739-742
[3] Yao T, Zhang X, Sun Z, Liu S, Huang Y, Xie Y, Wu C, Yuan X, Zhang W, Wu Z, Pan G, Hu F, Wu L, Liu Q and Wei S 2010 *Phys. Rev. Lett*. **105** 226405
[4] Xin-Ping Y, Jiao C, Jian-Min L, Pan-Pan Z, Xiang L and Zhong X 2010 *Materials Letters* **64** 278-280
[5] Suh J Y, Lopez R, Feldman L C and Haglund R F Jr 2004 *J.Appl.Phys.* **96** 1209-1213
[6] Corr S A, Shoemaker D P, Melot B C and Seshadri R 2010 *Cond-Mat* **1** 958
[7] KOCH C C, *Nanostructured Materials*, William Andrew-Interscience, Norwich NY, 2007
[8] Sorapong P, Yoshikazu S, Athapol K, Sommai P and Susumu Y 2005 *J Solid State Chem.* **178** 2152-2158
[9] Shidong J, Feng Z, Ping J 2011 *Materials Letters* **65** 708-711
[10] Nag J, Haglund R F, Payzant E and More K 2010 *Cond-Mat* **1** 3876

Division D – Astrophysics and Space Science

Searching for signatures of nearby sources of cosmic rays in their local chemical composition

D. Bisschoff[1], I. Büsching[12] and M. S. Potgieter[1]

[1] Centre for Space Research, North-West University, Potchefstroom Campus, Private Bag X6001, Potchefstroom 2520, South Africa

[2] Institut für Theoretische Physik IV, Ruhr-Universität Bochum, Bochum 44780, Germany

E-mail: 20056950@nwu.ac.za

Abstract. The direct evidence for the acceleration of hadronic cosmic rays at supernova remnants (SNR) underlined the need for a 3D time dependent treatment of the propagation of galactic cosmic rays (CRs). Full 3D time dependent calculations of the propagation of CRs have shown that if CRs indeed originate from SNR, transient point-like sources, the flux of the CR primary component measured at Earth depends strongly on the local source history, whereas the secondary component shows little to no variation due to nearby sources. However, due to their simplicity, steady state, rotational symmetric models (2D) of CR propagation are still widely used, although they cannot properly take into account the history of local point sources, but rather assume smeared-out sources. We make use of the fact that 2D models don't decribe CR primaries and secondaries equally well, to look for evidence of a nearby point source. Doing a parameter study, one may expect different best fit values looking at the primary and secondary CR components separately, as it is unlikely that the source history mimicked by the 2D models coincides with the real local source history. For this study, we adapted the 2D version of the GALPROP code to a compute-cluster environment and perform parameter studies comparing CR spectra with mainly primary and secondary CR data separately. The first results of these studies are presented and recommendations for further such studies given.

1. Introduction

The discovery of direct evidence for the acceleration of highly energetic hadronic particles at the shell supernova remnant RXJ1713.7-3946 [1] supports the origin of hadronic cosmic rays (CRs) in supernova remnants (SNR). This finding also emphasized the need of a 3D, time-dependent treatment of the galactic CR propagation, as only those models are able to properly describe transient, point-like sources like SNR. Using time-dependent calculations taking into account all three spatial dimensions, it has been shown that the flux of the CR primary component measured at Earth strongly depends on the local source history, given that the bulk of the galactic CRs originate in transient, point-like sources [2] like supernova and their remnants. These 3D time dependent calculations also showed that the secondary CR component is not affected by local point-like sources. This leads to the conclusion that the widely used steady state, rotational symmetric models (2D models)[1] of CR propagation [3, 4, 5, 6] may not be adequate to describe

[1] For CR propagation, the Galaxy can be described as a cylinder with a radius of $\approx 20\,\mathrm{kpc}$ and a height of $\approx 8\,\mathrm{kpc}$, including the Galactic halo, in which CRs have a finite chance to return to the Galactic disk. Assuming symmetry in azimuth leads to 2D models that only depend on Galactocentric radius and height. Dropping the time dependence leads to a steady state model.

the CR primary component originating from transient, point sources, but are well suited to model the secondary component.

Thus, 2D models and to some extent also 'leaky box' models[2] [7], the simplist propagation models for the Galaxy [8], are sufficient to model the local flux of these nuclei. Steady state, rotational symmetric models (2D models) assume smeared-out sources, which do not necessarily result in the same local CR flux as the real local sources, leading to a poor description of the CR primary component. When working with 2D models, concentrating on secondary, tertiary and higher CR nuclei may thus yield a better description of the galactic CR propagation, as the local flux of these nuclei doesn't depend on the local source history.

2. Method

Time-dependent calculations taking into account all three spatial dimensions are still numerically too involved for large parameter studies, we thus adopt for our current work the 2D version of the GALPROP[3] code [3, 4, 5] for an extensive parameter study. The GALPROP code has been designed by A. Strong and I. Moskalenko [3] for the propagation of CRs on either a 2D or 3D spatial grid and is aimed to enable simultaneous predictions of all relevant observations to CR physics; including, but not limited to, CR nuclei. As shown above, the 2D version is a valid approach for secondary CR component, but may fail to correctly compute the flux of the CR primary component.

To use a steady state, 2D propagation model for our search for signatures of nearby point sources of CR in their local chemical composition, we make use of the fact these models are well suited to describe the CR secondary component, but, due to their failure to adequately describe transient, point sources, do a poor job describing primary CR.

Therefore, we divide the existing CR data into three components according to the fraction of secondary nuclei they contain. The fraction of secondaries and primaries in the isotopes that make up each species are added up seperately. The addition is weighted according to the known abundances of the isotopes in a species when detected at Earth, integrated over all energies. Using this method to differentiate between CR species, all the CR data and local interstellar spectra (LIS) can then be divided into one of three component groups: Primary CRs, Mixed CRs and Secondary CRs:

- Primary component: secondary fraction <30%
- Mixed component: secondary fraction >30%,<70%
- Secondary component: secondary fraction >70%

3. Calculations

For the parameter study results presented here we used the plain diffusion model built into GALPROP. This model was chosen because it doesn't take reacceleration into account which simplifies the model and reduces the number of free parameters to consider. The plain diffusion model was used in cylindrical coordinates with two spatial dimensions, the Galactocentric radius r and the height above the Galactic plane z, with symmetry in the angular dimension.

We scanned the parameter space given in table 1, where k_0 is the magnitude of the diffusion coefficient at particle rigidity 4 GV, δ the spectral index of the diffusion coefficient and α that of the sources. The halo height above Galactic plane was fixed to 4 kpc, also we did not take into account effects due to a Galactic wind. To avoid having the spectral index of the diffusion coefficient below 4 GV as an additional free parameter and also to minimize the impact of our crude description of solar modulation, we only consider data with rigidities above 4 GV.

[2] The leaky box model doesn't take into account any spatial information and treats the Galaxy as a box containing CRs, their sources and interstellar matter. CRs have a finite, energy dependent chance to escape out of this box.
[3] http://galprop.stanford.edu/web_galprop/galprop_home.html

Table 1. Parameter space considered.

parameter	min	max	unit
k_0	0.50	5.0	$10^{28}\,\mathrm{cm^2 s^{-1}}$
δ	0.1	1.0	
α	1.50	3.50	

A total of 30720 models were calculated and the calculations were performed on the institutional cluster of the North-West University in Potchefstroom using a MPI[4] code to run the models in parallel.

The full nuclear reaction network was solved over all isotopes implemented in GALPROP. Thus both primary and secondary CR species were run at the same time. The LIS calculated with GALPROP were then compared to data by method of a χ^2 test in order to find the best fit parameter set. The data was taken from the CR database[5] maintained by A. Strong and I. Moskalenko [9]. For each model, we calculated the χ^2 value for each entry in the database when compared to the corresponding calculated LIS value after the temporal variation of the modulation during a solar cycle was taken into account. At energies below $\approx$10 GeV the effect of solar modulation has to be taken into account. This was done by using the force field approximation [10] and a time-dependent modulation parameter, obtained from proton data from different epochs in the solar cycle.

The resulting χ^2 values were then added up for each of the three CR components separately.

4. Results

The first results of our calculations are presented in Figures 1 to 3, where we show contour plots of best χ^2 values over the parameter range considered for the CR primary, secondary and mixed component, respectively. Our best fit parameters for the three components are given in Table 2. These best fit parameters are marked on the contour plots for easier comparison of the relative locations. Looking at Table 2 and Figures 1 and 2, the different locations of the minimum χ^2

Table 2. Best fit values for the secondary, primary and mixed component.

parameter	secondary	primary	mixed	unit
k_0	1.92831	2.86808	1.02168	$10^{28}\,\mathrm{cm^2 s^{-1}}$
δ	0.767742	0.10000	0.10000	
α	2.20968	2.66129	2.79032	

for primary and secondary CR component in the k_0-α, α-δ, and k_0-δ planes is apparent. The χ^2 contours are also quite different, thus the calculated spectra of the three components show different sensitivities to the model parameters.

Not surprisingly, the plots for the mixed component resemble somewhat a superposition of the corresponding secondary and primary plots. The high χ^2 values for models with α not in the range $2.0 < \alpha < 3.0$ indicate that values outside this range can be disregarded.

In our calculations, the primary and secondary components of the galactic CR seem to favour different regions in the scanned parameter space. As mentioned in the introduction, a possible

[4] Message Passing Interface: http://mpi-forum.org/

[5] http://www.mpe.mpg.de/∼aws/propagate.html

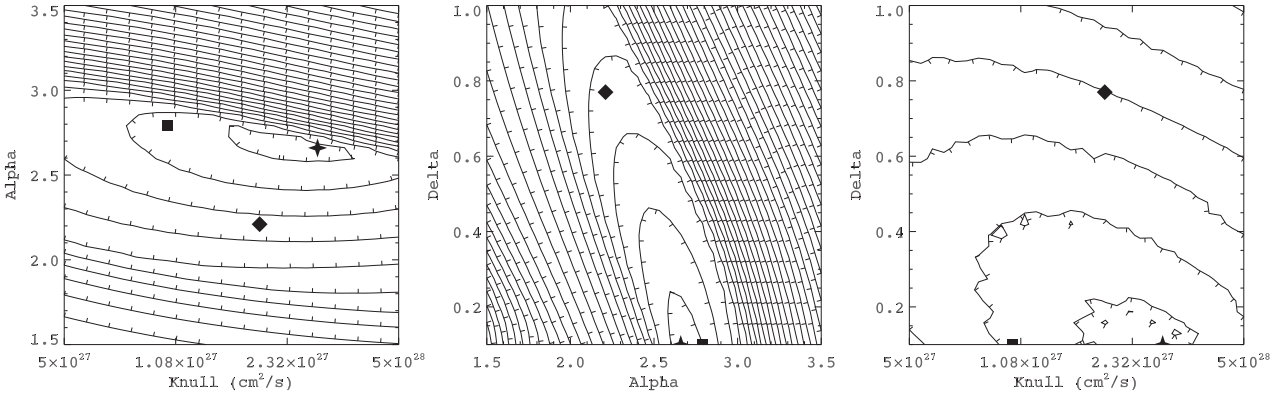


Figure 1. χ^2 distribution for the primary CR component in the $k_0-\alpha$ (left) $\alpha-\delta$ (middle) and $k_0-\delta$ (right) plane. Minimum value in each plane is marked by a 4-point star. The minimums for the other two components are marked for comparison, a diamond for the secondary component and a square for the mixed component.

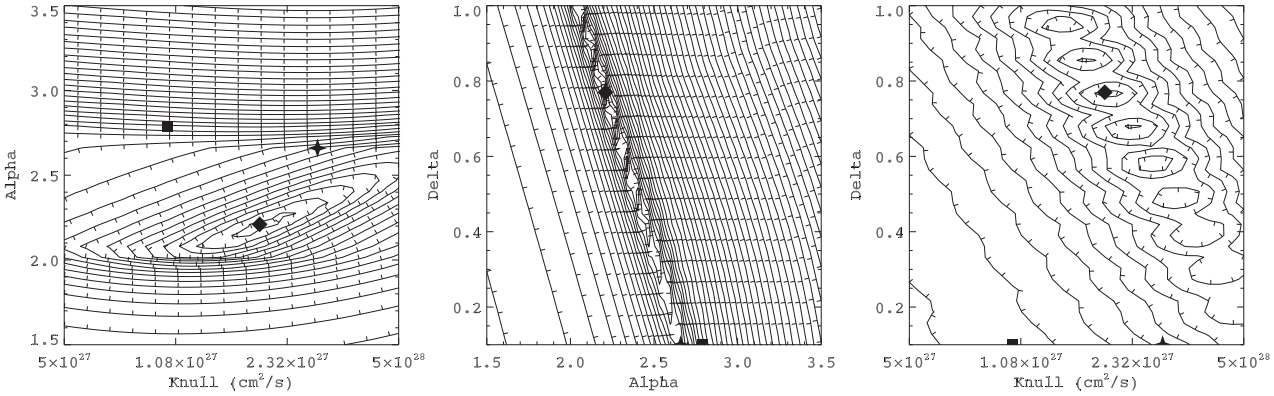


Figure 2. Same as Figure 1, but the χ^2 distribution is for the secondary CR component.

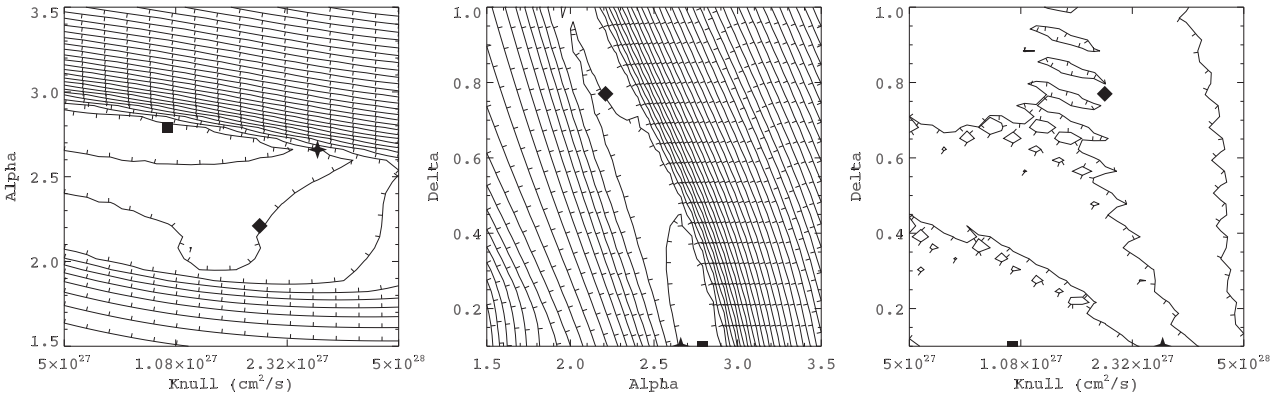


Figure 3. Same as Figure 1, but the χ^2 distribution is for the secondary CR component.

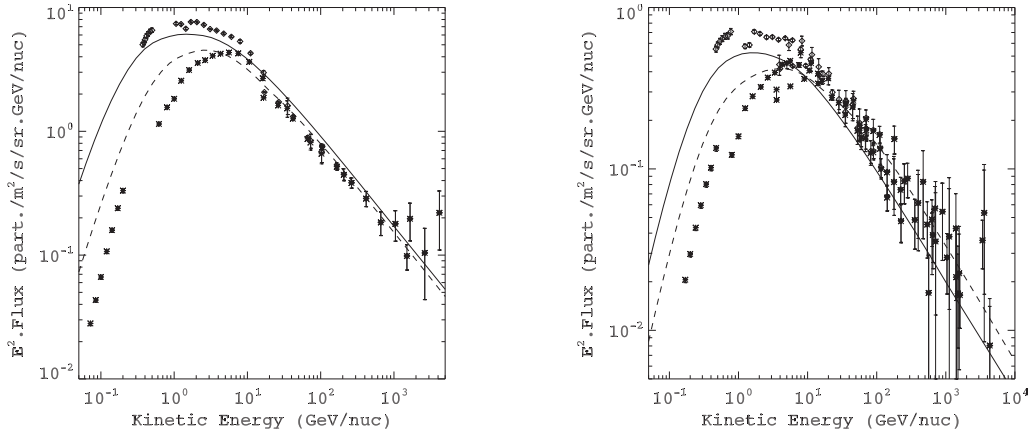


Figure 4. LIS for Primary CR component species Carbon and Iron. The solid line is the LIS for this study and the dashed line is the LIS from Ptuskin *et al* [6]. Experimental data is marked with stars and the data with the effect of solar modulation removed is marked with diamonds.

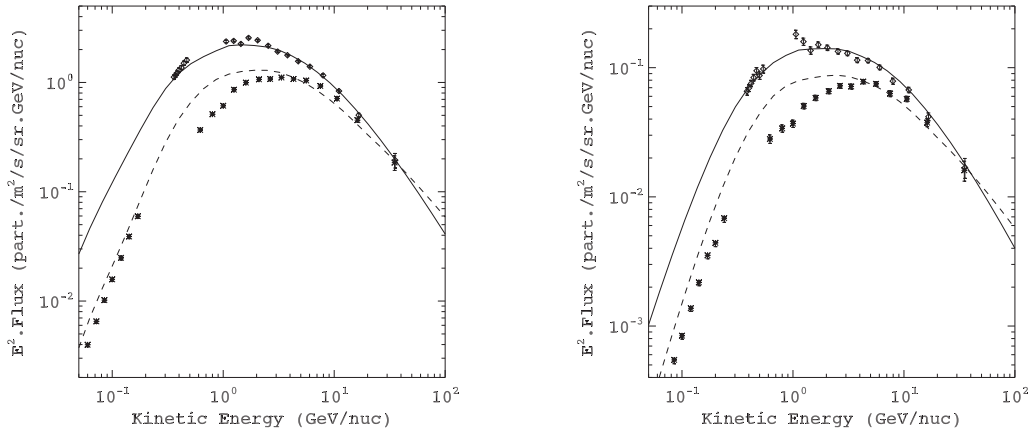


Figure 5. Same as Figure 4 but the LIS for Secondary CR component species Boron and Fluorine is shown.

explanation of our findings is that 2D models are indeed incapable of correctly describing the local CR sources.

The LIS produced by the best fit models with parameters listed in Table 2, are plotted in Figures 4 and 5. These figures show the LIS for selected CR species of the primary and secondary component groups. LIS obtained by Ptuskin *et al* [10] using the GALPROP code (but with a different analysis than the study conducted here) are also shown for comparison, because their parameter choices where used as a starting point for this parameter study. The LIS shown are Carbon and Iron for the primaries; Boron and Fluorine for the secondaries; and finally Nitrogen and Sodium for the mixed group. The experimental data and the corresponding demodulated data above 4GeV used to calculate the χ^2 values are also shown.

All the Ptuskin LIS presented are much lower than those LIS obtained in this study at energies below 10 GeV. For the LIS shown the Ptuskin LIS do correspond to the obtained LIS at higher energies. The LIS for the primary CR species lie within the trend displayed by the data, with Iron lying in the lower part of the trend. The LIS for the secondary CR species show good fits at all energies. Small deviations in the fit of any one CR species in a group are to be expected due to the fact that all the calculated LIS of the CR species in a component group

were simultaneously fitted to the data. The individual fitting of a CR species LIS may thus be lower or higher than expected to fit the data points, but for the whole group the χ^2 value is still a minimum value.

The data points are also inconsistent between different experiments for CR species due to systematic errors especially for the primary component. This results a wider spread of data points and thus larger χ^2 values for species such as Iron, eventhough the LIS can bee seen to lie within the trend displayed by the data. This mutual exclusion by the experimental datapoints is due to using as many different sets of data as possible. Different experiments are not always consistent in measuring LIS and makes fitting the calculated LIS to the data difficult for such large data sets using the χ^2 test. The mixed component shows similarly large χ^2 values eventhough this component's experimental data doesn't suffer from the same inconsistencies and suggests that the wide spread of data points might not be a determining factor.

5. Summary

We performed a parameter study using the steady state, 2D plain diffusion model of the public available GALPROP code. Looking at the CR primary and secondary components separately, we found that these components favour different best fit values for the magnitude of the diffusion coefficeint k_0, the source spectral index α and the spectral index of the energy dependence of the diffusion coefficient δ. This result is expected in case there is a significant contribution of a nearby point source to the local CR flux, as the steady state 2D plain diffusion model is a good model for the CR secondary component but a poor model for the CR primary component due to its failure to adequately describe transient point sources.

Although this finding needs further investigation, we suggest it is an indication that the primary CR flux indeed contains a component injected by a nearby point source, indicating the presence of a nearby CR accelerator.

The differences between the LIS obtained in this study and the Ptuskin LIS, which are produced by the standard GALPROP parameter set, can possibly be attributed to dependence of the fitting on the data sets. Using different data sets or excluding data from certain experiments will have a significant effect on the best fit LIS found. Also, the method of including modulation is important, as choosing a modulation parameter for the force field model can be done arbitrarily.

Future studies could include a better modulation implementation such as a 2D drift model. A next step in this line of study would be chosing one set of data from a reliable experiment and also including other parameters, such as halo height and galactic wind, in the parameter study. Taking reacceleration into account can also be considered.

Acknowledgements

The authors wish to thank the South African National Research Foundation (NRF) for partial financial support.

References

[1] Aharonian F, *et al* 2004 *Nature (London)* **432** 75
[2] Büsching I, *et al* 2005 *Astrophys. J.* **619** 314
[3] Strong A W and Moskalenko I V 1998 *Astrophys. J.* **509** 212
[4] Moskalenko I V, Strong A W, Mashnik S G and Ormes J F 2003 *Astrophys. J.* **586** 1050
[5] Strong A W, Moskalenko I V and Reimer O 2000 *Astrophys. J.* **537** 763
[6] Ptuskin V S, Moskalenko I V, Jones F C, Strong A W and Zirakashvili V N 2006 *Astrophys. J.* **642** 902
[7] Strong A W, Moskalenko I V and Ptuskin V S 2007 *Ann. Rev. Nucl. Part. Sci.* **57** 285
[8] Ptuskin V S, Strelnikova O N and Sveshnikova L G 2009 *Astrop. Phys.* **31** 284
[9] Strong A W and Moskalenko I V 2009, Paper ID 0626, 31st ICRC, Lodz, Poland (arXiv:0907.0565v1)
[10] Caballero-Lopez R A, and Moraal H 2004 *J. Geophys. Res.* **619** A01101

Evaluation of satellites LAGEOS I and II; general relativistic accelerations in the Schwarzschild field of Earth

L Combrinck

HartRAO, PO Box 443, Krugersdorp, 1740, South Africa

E-mail: ludwig@hartrao.ac.za

Abstract. Solutions of General Relativity Theory (GRT) pertaining to space geodesy are weak field, slow motion approximations. These approximations are valid as the gravitational field, in which the solutions are calculated, has a potential of small magnitude and the velocities involved for any of the satellites are much smaller than the velocity of light. The basic effects of GRT on space geodetic measurements are related to how GRT affects the observables and dynamics of satellite orbits. We evaluate the GRT accelerations ascribed to the Schwarzschild field, as well as those due to Lense-Thirring precession (frame dragging) and de Sitter precession. The Shapiro delay for a laser pulse as applied to Satellite LAser Ranging (SLR) is calculated and the importance of including this GRT correction in the range corrections for SLR is described in terms of evaluation of the Observed-Computed residuals.

1. Introduction

The work of Einstein [1] indicates that the geometrical properties of spacetime and the spacetime metric g are dependent and determined to some extent, through the field equation by the distribution of mass-energy and mass-energy currents in space. In this work, the interpretation of the origin of inertia is seen as local inertial forces resulting from rotations and accelerations relative to local inertial frames. These inertial (local) frames are influenced (dragged) by the distribution of mass-energy and mass-energy currents. In space geodesy, there are four main techniques, VLBI (Very Long Baseline Interferometry), GNSS (Global Navigation Satellite Systems), SLR/LLR (Satellite or Lunar Laser Ranging) and DORIS (Doppler Orbitography and Radiopositioning Integrated by Satellite). Analyses of the data have to be done within the full framework of a post-Newtonian formalism and must therefore consider GRT and its effects on signal propagation and observables, such as laser pulse travel time [2]. Consideration must also be given to GRT implications on reference and time frames, as well as on solar body ephemerides. Geodetic software in general follows the recommendations of the International Earth Rotation Service (IERS), to a certain level; see McCarthy and Petit [3], although there may be differences due to processing strategy. Solutions of GRT as utilised in space geodesy are weak field, slow motion approximations and are valid due to fact that the velocities v involved for any of the satellites are much smaller than the speed of light c and that the gravitational field, in which the solutions are calculated, has a potential U of small magnitude. Thus $U/c^2 \ll 1$ and $v/c \ll 1$. Therefore Einstein's field equations may be linearised and be expressed in a form analogous to that of Maxwell's equations in electromagnetism.

2. Satellite Laser Ranging

There are several GRT effects which need to be considered when utilising SLR data. These effects directly influence the accuracy of the observations, although the magnitudes of the effects differ.

2.1. Shapiro delay

Raw satellite laser ranging data (as determined by each laser pulse return) are converted to a normal point (NP), made up of a number of raw data points. Using the speed of light and factoring in some additional corrections on can calculate the Normal Point Range (NPR), through the range equation

$$NPR_i = \left(\frac{NPtof_i}{1\times 10^{12}} \times c \right) \Big/ 2 - \Delta a_i + \Delta CoM_i - \Delta R_{b_i} - \Delta GR_i - \Delta \varepsilon_i . \tag{1}$$

In (1), the main observable is the normal point time-of-flight (in picoseconds) $NPtof_i$ at a certain time and c is the velocity of light (299,792,458.0 $\mathrm{m.s^{-1}}$). The calculated range found in (1) must be be corrected for the delay due to the atmosphere Δa_i, a satellite dependent centre-of-mass correction ΔCoM_i, the SLR station range-bias ΔR_{b_i}, a relativistic correction (Shapiro delay) ΔGR_j and other $(\Delta \varepsilon_i)$ errors. If one uses LAGEOS as an example, the Shapiro delay correction is about 7 mm. According to McCarthy and Petit [3]

$$t_2 - t_1 = \frac{\left| \vec{x}_2(t_2) - \vec{x}_1(t_1) \right|}{c} + \sum_J \frac{(1+\gamma) GM_J}{c^3} \ln \left(\frac{r_{J1} + r_{J2} + \rho}{r_{J1} + r_{J2} - \rho} \right). \tag{2}$$

In (2), the Parameterized Post-Newtonian parameter (PPN) γ should equal unity if GRT is valid, $t_2 - t_1$ is the total time delay when a laser pulse is emitted from the SLR station (coordinate x_1) at time t_1 and the return pulse is received at the SLR station (coordinate x_2) at time t_2. In (2) the range defined by $\rho = \left| \overrightarrow{x_2} - \overrightarrow{x_1} \right|$ is the uncorrected (for GRT) range. Furthermore, $r_{J1} = \left| \overrightarrow{x_1} - \overrightarrow{x_J} \right|$ and $r_{J2} = \left| \overrightarrow{x_2} - \overrightarrow{x_J} \right|$. Formulation (2) was first derived by Holdridge [4]. The Shapiro correction for space curvature is described by the last term in (2). Determination of the relativistically uncorrected range ρ involves an iterative solution of two light-time equations for the uplink and downlink path [5, 6].

2.2. GRT accelerations

Accelerations due to GRT can be separated from those that are purely Newtonian. The perturbing acceleration can be written as

$$\ddot{\vec{r}} = -\frac{GM_\oplus}{r^3} \vec{r} + \vec{f}, \tag{3}$$

where the total perturbing force $\vec{f}$ consists of additional forces which include gravitational forces from the Sun, Moon, planets and the earth, the sun radiation pressure and the effects of GRT.

According to McCarthy and Petit [3] the relativistic correction to the acceleration of a satellite in Earth orbit is

$$\begin{aligned} \Delta \ddot{\vec{r}} = {} & \frac{GM_\oplus}{c^2 r^3} \left\{ \left[2(\beta+\gamma) \frac{GM_\oplus}{r} - \gamma \left(\dot{\vec{r}} \cdot \dot{\vec{r}} \right) \right] \vec{r} + 2(1+\gamma) \left(\vec{r} \cdot \dot{\vec{r}} \right) \dot{\vec{r}} \right\} + \\ & (1+\gamma) \frac{GM_\oplus}{c^2 r^3} \left[\frac{3}{r^2} \left(\vec{r} \times \dot{\vec{r}} \right) \left(\vec{r} \cdot \vec{J} \right) + \left(\dot{\vec{r}} \times \vec{J} \right) \right] + \left\{ (1+2\gamma) \left[\dot{\vec{R}} \times \left(\frac{-GM_s \vec{R}}{c^2 R^3} \right) \right] \times \dot{\vec{r}} \right\}. \end{aligned} \tag{4}$$

Equation (4) follows the formalism of Damour and Soffel [7]. In (4), c is the velocity of light, the PPN parameters β, γ equal unity if general relativity is valid and $\vec{R}$ is the position of the Earth relative to the Sun. The position of the satellite relative to the Earth is $\vec{r}$. Angular momentum per unit mass of the Earth is described by $\left| \vec{J} \right| \cong 9.8\times 10^8\ \mathrm{m^2 s^{-1}}$, $GM_\oplus$ is the gravitational coefficient of the Earth and GM_s is the gravitational coefficient of the Sun. These accelerations are small, but must be included for accurate orbit determination purposes as there are long term periodic and

secular effects [8] on the satellite orbit. Precession of perigee results from the Schwarzschild effect, de Sitter precession can lead to long-period variations of some orbital elements $(\Omega, \omega, \mathrm{M})$ and Lense-Thirring precession leads to secular rates [9] in the orbital elements Ω and ω.

2.2.1. Schwarzschild field contribution to acceleration

According to Hugentobler [10], if we set $GM = n^2a^3$ the Schwarzschild acceleration for a circular orbit can be written as

$$\Delta\ddot{\vec{r}}_S = \frac{GM}{c^2a^3}\left[4\frac{GM}{a} - v^2\right]\vec{r} = \frac{GM}{c^2a^3}\left[4\frac{GM}{a} - \frac{GM}{a}\right]\vec{r} = 3\frac{(GM)^2}{c^2a^4}\cdot\vec{r}. \tag{5}$$

The radial component of the Schwarzschild acceleration can then be written as

$$R = \vec{e}_R \cdot \Delta\ddot{\vec{r}}_S = 3\frac{(GM)^2}{c^2a^3}, \tag{6}$$

where the alongtrack and crosstrack accelerations equal zero as

$$S = \vec{e}_S \cdot \Delta\ddot{\vec{r}} = 0 \text{ and } W = \vec{e}_W \cdot \Delta\ddot{\vec{r}} = 0. \tag{7}$$

Taking into account the sign convention of the gravitational acceleration in (3), the radial component of the Schwarzschild field imparts an outward acceleration, reducing the Newtonian component, changing GM so that

$$\vec{e}_R \cdot \ddot{\vec{r}}_{tot} = -\frac{GM}{a^2} + R = -\frac{GM}{a^2}\left(1 - \frac{Ra^2}{GM}\right) = -\frac{GM'}{a^2}. \tag{8}$$

Using Kepler's Third Law and holding the period T of the satellite motion unchanged and $a' = a + \Delta a$ then $a'^3n^2 = GM'$ and to first order,

$$\Delta a = -\frac{1}{3}\frac{Ra^3}{GM}. \tag{9}$$

Substituting R from (6) into (9) one can write [10]

$$\Delta a = -\frac{GM}{c^2} = -4.43 \text{ mm}. \tag{10}$$

The Schwarzschild acceleration on a satellite in orbit is thus a radial, in–plane effect, with zero magnitude effects in the alongtrack and crosstrack directions.

2.2.2. Lense-Thirring precession

The earth's angular momentum J plays an important role in the generation of the Lense-Thirring effect. In the case of LAGEOS, the effect is very small (~31 mas.yr^{-1}) and is therefore very difficult to measure [11]. Following Hugentobler [10] the angular momentum vector can be written as $\vec{J} = J(0, \sin i, \cos i)$. In the case of a circular orbit therefore

$$\vec{r} \cdot \vec{J} = aJ\sin i\sin\nu \text{ and } \vec{r} \times \vec{J} = anJ(\cos\nu\cos i, \sin\nu\cos i, -\sin\nu\sin i). \tag{11}$$

The radial, alongtrack and crosstrack components are

$$R = \Delta\ddot{\vec{r}}_{LT} \cdot \vec{e}_R = 2\frac{GM}{c^2a^4}\vec{r} \cdot (\vec{r} \times \vec{J}) = 2\frac{GM}{c^2a^4}\vec{J} \cdot (\vec{r} \times \vec{r}) = 2\frac{GM}{c^2a^2}nJ\cos i, \tag{12}$$

$$S = \Delta\ddot{\vec{r}}_{LT} \cdot \vec{e}_S = 0, \tag{13}$$

$$W = \Delta\ddot{\vec{r}}_{LT} \cdot \vec{e}_W = 2\frac{GM}{c^2a^3}\left[\frac{3}{a^2}a^2naJ\sin i\sin\nu - anJ\sin i\sin\nu\right] = 4\frac{GM}{c^2a^2}nJ\sin i\sin\nu, \tag{14}$$

$$\Delta a = -\frac{2}{3}\frac{anJ}{c^2}\cos i \propto a^{-½}. \tag{15}$$

As a result of the constant radial acceleration the semi-major axis has a reduction given by (15). If the orbit is perpendicular to the equatorial plane, the $\cos i$ factor in (15) ensures that there is no radial acceleration or change in semi-major axis. Semi-major axis changes are directly proportional to the square root of the semi-major axis, therefore higher orbit satellites experience a smaller effect. Setting $u = \nu$ (true anomaly is undefined for circular orbits as they have no periapsis) in the case of a circular orbit,

$$\dot{\Omega}_{LT} = \frac{\sin u}{na\sin i}\cdot W = 4\frac{GM}{c^2 a^3 J\sin^2 u} = 2\frac{GM}{c^2 a^3 J\left(1-\cos 2u\right)}. \tag{16}$$

The precession of the node is not dependent on the inclination i of the orbit as it is a frame precession effect and the precession is inversely related to a^3, so that it decreases very rapidly.

2.2.3. de Sitter precession

In curved space-time, a parallel transported (local inertial) reference frame undergoes precession with respect to distant stars [10]. For a satellite the precession rate (19.2 mas.yr^{-1} in the vicinity of Earth) is independent of the distance from the Earth. Consequently the acceleration (compare (4)) is

$$\Delta\ddot{\vec{r}}_{dS} = -2\left[-\frac{3}{2}\frac{GM_S}{c^2R^3}\vec{R}\times\dot{\vec{R}}\right]\times\vec{r}. \tag{17}$$

Equation (17) has the form of a Coriolis term. The Coriolis acceleration of a satellite in orbit around the Earth due to geodesic precession is $\ddot{\vec{r}} = 2\vec{\Omega}\times\dot{r}$, where the angular velocity vector due to geodesic precession [12] is

$$\vec{\Omega} = \frac{GM_S\left(\gamma+\frac{1}{2}\right)}{c^2R^3}\left(\vec{R}\times\dot{\vec{R}}\right). \tag{18}$$

The formulation of (18) (instead of (17)) is useful if the PPN parameter γ needs to be included in an estimation process and is equivalent to the third term of (4). This is implemented in the analysis software at HartRAO. In (17) M_S and R are respectively the mass of the Sun and the distance to the Sun. The precession $\vec{\omega}_{dS}$ vector represents the precession vector of the global frame relative to the local inertial frame and points to the southern ecliptic. If the frame is redefined by keeping the z-axis the same but rotating the x-axis towards the ascending node with respect to the ecliptic plane [10], the precession vector of the global frame can be written as

$$\vec{\omega}_{dS} = -\frac{3}{2}\frac{GM_S}{c^2R^3}\vec{R}\times\dot{\vec{R}} = -\frac{3}{2}\frac{GM_S}{c^2R}\sqrt{1-e_S^2}\,n_S\left(0,\sin\beta,\cos\beta\right). \tag{19}$$

In (19) the mean motion of Earth around the Sun is denoted by n_S and the inclination of the orbital plane to the ecliptic is represented by β. The inclination is in the range $i-\varepsilon$ and $i+\varepsilon$ and can be calculated using

$$\cos\beta = \cos\varepsilon\cos i + \sin\varepsilon\sin i\cos\Omega, \tag{20}$$

with the obliquity of the ecliptic $\varepsilon = 23^{\circ}.26'$. In the frame, as defined for a circular orbit, the relation can be written as

$$\left(\vec{R}\times\dot{\vec{R}}\right)\times\ddot{\vec{r}} = R^2 n_S\sqrt{1-e_S^2}\,an\left(-\cos\beta\cos\nu,-\cos\beta\sin\nu,\sin\beta\sin\nu\right) \tag{21}$$

and consequently [10] the perturbing accelerations are

$$R = \Delta\ddot{\vec{r}}_{dS}\cdot\vec{e}_R = -3\frac{GM_S}{c^2R}n_S\sqrt{1-e_S^2}\,an\cos\beta, \tag{22}$$

$$S = \Delta\ddot{\vec{r}}_{dS}\cdot\vec{e}_S = 0,\quad W = \Delta\ddot{\vec{r}}_{dS}\cdot\vec{e}_W = 3\frac{GM_S}{c^2R}n_S\sqrt{1-e_S^2}\,an\sin\beta\sin\nu. \tag{23}$$

Considering the $\cos\beta$ factor in the radial component, orbits which are perpendicular to the ecliptic plane experience no change in semimajor axis and no radial acceleration, as the Coriolis acceleration is perpendicular to the orbital plane. The negative radial acceleration increases the semimajor axis by

$$\Delta a = +\frac{GM_S}{c^2}\frac{a}{R}\frac{n_S}{n}\sqrt{1-e_S^2}\cos\beta \propto a^{5/2}. \tag{24}$$

Setting $u = \nu$ the precession of the ascending node with respect to the ecliptic plane is calculated as

$$\dot{\Omega}_{dS} = \frac{\sin u}{na\sin\beta}\cdot W = 3\frac{GM_S}{c^2R}n_S\sqrt{1-e_S^2}\sin^2 u = 3\frac{GM_S}{c^2R}n_S\sqrt{1-e_S^2}\,\mathrm{s}\left(1-\cos 2u\right). \tag{25}$$

3. Results

I find the following representative values for the three terms of (4) split into R (radial), N (normal) and T (tangential) components and the Shapiro delay, utilising SLR data from LAGEOS 1 as:

- First term, the nonlinear Schwarzschild field of the Earth $\left(\approx 1.9\times10^{-9}\ \mathrm{ms}^{-2}\right)$
  $R = 1.9\times10^{-9}\ \mathrm{m.s}^{-2},\ T = -2\times10^{-14}\ \mathrm{m.s}^{-2},\ N = 0\ \mathrm{m.s}^{-2}$
- Second term, Lense-Thirring precession (frame dragging) $\left(\approx 2.0\times10^{-11}\ \mathrm{ms}^{-2}\right)$
  $R = -4.6\times10^{-12}\ \mathrm{m.s}^{-2},\ T = -1.5\times10^{-15}\ \mathrm{m.s}^{-2},\ N = 1.2\times10\ \mathrm{m.s}^{-11}$
- Third term, de Sitter (geodesic) precession $\left(\approx 0.9\times10^{-11}\ \mathrm{ms}^{-2}\right)$
  $R = 6.4\times10^{-12}\ \mathrm{m.s}^{-2},\ T = 8\times10^{-15}\ \mathrm{m.s}^{-2},\ N = 4\times10^{-11}\ \mathrm{m.s}^{-2}$
- Shapiro delay ~ 6.7 mm

The very low values in the 10^{-14} to 10^{-15} range confirm the zero values of the formulations of Hugentobler [10]. Testing GRT using SLR is very demanding however and efforts are found in the literature which describes various approaches, see for instance Ciufolini and Pavlis [13]. The gravitomagnetic effects are embedded in the total acceleration (relativistic plus Newtonian) as experienced by satellites and SLR data have been used to evaluate the effect of frame dragging by estimating the advance of perigee due to frame dragging by Ciufolini and Pavlis [14] and Luchessi and Peron [15]. Efforts also include the direct evaluation of PPN parameters γ and β as a test of GRT utilising SLR data by Combrinck [16].

4. Conclusions

Implementation of (4) in a complete solution (utilising SLR processing software developed by the author) taking into account the complex dynamics of the crust of the earth as it responds to earth-tide, pole-tide, atmospheric loading, the variation of the gravity field due to earth and pole tide, and incorporating gravitational accelerations from the moon, sun and planets, as well as solar and earth-reflected radiation, indicate that in order to minimise the observed- minus-computed residuals, GRT must be taken into account. It is clear that all terms of (4) need to be taken into account for the purposes of precise orbit determination utilising SLR data.

Acknowledgements

This work is based on research supported by the National Research Foundation under grant IFR2011041500034. The comments of an anonymous reviewer and F Frescura contributed to the readability of this work and their comments are appreciated.

References

[1] Einstein A 1920 *Relativity: The Special and General Theory* (New York: Henry Holt and Company)

[2] Müller J Soffel M and Klioner S A 2008 Geodesy and relativity *J. of Geodesy* **82** 133-145

[3] McCarthy D D and Petit G 2003 *IERS Conventions (2003) (IERS Technical Note; 32)* Frankfurt am Maim: Verlag des Bundesamts für Kartographie und Geodäsie 72-84

[4] Holdridge D 1967 An Alternate Expression for Light Time Using General Relativity, *Space Programs Summary* 37-48 **III** NASA 2-4

[5] Montenbruck O and Gill E 2001 *Satellite orbits: models, methods and applications* (Berlin: Springer-Verlag)

[6] Combrinck L 2010 Satellite Laser Ranging *Sciences of Geodesy I, (Advances and Future Direction)* ed G Xu (Berlin: Springer) chapter 9 pp 301-336

[7] Damour T Soffel M and Xu C 1994 General-relativistic celestial mechanics IV. Theory of satellite motion *Phys. Rev. D* **49** 618–635

[8] Huang C and Liu L 1992 Analytical solutions to the four post-Newtonian effects in a near-Earth satellite orbit *Celestial Mechanics and Dynamical Astronomy* **53** 293-307

[9] Ciufolini I and Wheeler J A 1995 *Gravitation and Inertia* (Princeton: Princeton University Press)

[10] Hugentobler U 2008 Unpublished notes available at ftp://maia.usno.navy.mil/conv2010/chapter10/add_info/

[11] Ciufolini I 1986 Measurement of the Lense-Thirring drag on high-altitude laser- ranged artificial satellites *Phys. Rev.Lett*. **56** 278–281

[12] Moyer T D 2000 *DESCANSO Book Series Deep Space Communications and Navigation Series,* ed Joseph H. Yuen Vol 2, Formulation for Observed and Computed Values of Deep Space Network Data Types for Navigation

[13] Ciufolini I and Pavlis E 2003 The Gravitomagnetic Field and its Measurement with the Lageos Satellites in *Nonlinear Gravitodynamics: The Lense-Thirring Effect* eds. Remo Ruffini and Costantino Sigismondi World Scientific Publishing Co. pp 185-200

[14] Ciufolini I and E C Pavlis 2004 A confirmation of the general relativistic prediction of the Lense-Thirring effect *Nature* **431** pp 958-960

[15] Lucchesi D M and Peron R 2010 Accurate Measurement in the Field of the Earth of the General-Relativistic Precession of the LAGEOS II Pericenter and New Constraints on Non-Newtonian Gravity, Phys. Rev. Lett. **105**, 231103.

[16] Combrinck L 2011 Testing General Relativity Theory through the estimation of PPN parameters γ and β using Satellite Laser Ranging data, in press, SA Journal of Geology

Optical study of PMS stars in the southern hemisphere high mass star forming region RCW 34

R.J. Czanik, D.J. Van der Walt, S.I. Loubser

Centre for Space Research, North-West University, Potchefstroom campus, Potchefstroom 2520, South Africa

E-mail: 20175132@nwu.ac.za

Abstract. RCW 34 is an HII region in the constellation of Vela which is only visible in the Southern hemisphere. Two recent studies which were conducted in the NIR on the region indicated that there is a strong presence of pre-main sequence (PMS) [1] stars in the region. A photometric and spectroscopic study of data that was obtained in 2002 showed that there are PMS stars in a wide $7' \times 7'$ field surrounding the HII region of which two exhibit an Hα emission line. A follow up spectroscopic study done in 2011 found 4 more stars that exhibit an Hα emission line.

1. Introduction

Stars are the primary visible constituents of galaxies. To get a better understanding of the structure and evolution of galaxies a fundamental understanding for the lifespan of stars is required. The only stage of stellar evolution which is not well understood is their formation which occur in molecular clouds.

A theoretical foundation for star formation involves an understanding on the dynamics of molecular cloud collapse when its mass exceeds the critical Jean's mass, and how a starless part of the molecular cloud contracts due to radiative cooling under the influence of gravity.
If a molecular cloud collapses in on itself the infalling gas will gain angular momentum and form rotating discs. Hartman [1] states that if these discs have enough angular momentum they will accrete matter. The compressed matter in the centres of these accretion discs results in the formation of protostars. The first accretion disc model was proposed by Lynden-Bell [2] which described most of the properties of T Tauri stars. These are lower mass PMS stars characterized by variable luminosity and an infrared excess due to dust emission from an accretion disc. Lynden-Bell[2] can explain a strong blue spectrum that spikes in the ultra violet by means of dust moving from an elderly disc towards the star's atmosphere. Matter from the accertion disc which is left over after the star has entered the main sequence may form planets.

A star forming region that is not often mentioned in the literature is RCW 34 (α=08:56:28.1;δ=-43:05;28) which was first catalogued in Rodgers *et al.* [4]. A stacked image of RCW 34 is shown in figure 1 Although this region has been studied by a number of researchers, it was only recently that a small number of new studies revisited this region. One of them is the near-infrared (NIR) study by De Villiers[5]. Using data obtained with the 1.4 m IRSF telescope

[1] PMS stars are very young stars which accretes matter through gravitational collapse and have not yet stared core hydrogen burning

Figure 1: A deep-field colour-composite of the B, V, and R filter images. The total exposure time of the image is 8210 seconds. The field of view for this image is $7' \times 7'$.

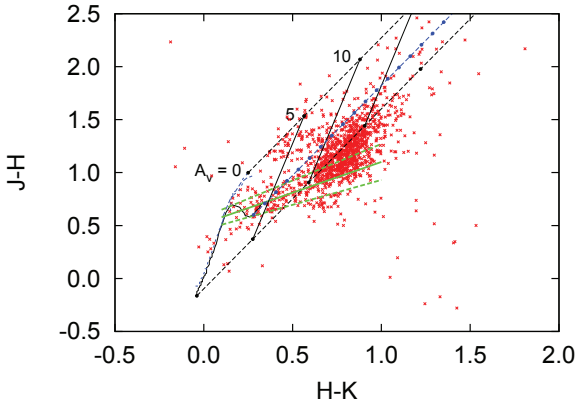


Figure 2: NIR two-colour diagram for the 1283 stars that were discovered by De Villiers. The green lines are the empirical T Tauri locus and it's upper and lower limits and given by Meyer *et al.*[3]

at Sutherland a total of 1283 objects were found in the J, H and K bands in a field of $7' \times 7'$ centred on the exciting star of the HII cloud. Using their NIR colours in a J-H against H-K two-colour diagram, De Villiers found that a group of nearly 800 objects cluster near and above the empirical classical T Tauri locus of Meyer *et al.*[3] as shown in figure 2. The formula for the empirical classical T Tauri locus is:

$$(J-H)_{CTTS} = 0.58 \pm 0.11 \times (H-K)_{CTTS} + 0.52 \pm 0.06 \quad (1)$$

The T Tauri locus is an empirical formula which was discovered by Meyer *et al.* [3]. The locus is a narrow band where T Tauri class stars appear on a NIR two colour diagram due to excess emission from the associated accretion disc.

This group might be reddened classical T Tauri type stars[2]. A smaller group of objects seem to cluster above the main-sequence branch, which might be reddened weak line T Tauri[3] type stars.

2. Observation and data reduction

The photometric and spectroscopic data were obtained in February of 2002 by D.J. van der Walt. The photometric data were collected with Dandicam on the 40″ telescope. The spectroscopic data were collected with grating number 6 which had a resolution of 2Åand a usable range of 1600 Åand the longslit spectrograph on the 74″ telescope at the SAAO in Sutherland. The follow up spectroscopic data were collected with grating number 7 which had a resolution of 5 Åand a usable range of 3200 Åand the longslit spectrograph on the 74″ telescope during February of 2011.

[2] Classical T Tauri (CTTS) type stars are T Tauri class stars where the accretion disc has a very strong emission in the NIR.

[3] Weak line T Tauri (WTTS) stars are pre-main-sequence stars of which the accretion disc has stopped contributing matter to the protostar and has receded from the centre of the formation region.

The photometric images were stacked in IRAF with a sum algorithm, resulting in a total exposure time of U: 960, B: 3110, V: 2850, R: 2250 and I: 2250 seconds. The spectroscopic images were stacked after aperture extraction using a sum algorithm in IRAF.

The spectroscopic observations done is 2002 was of 12 sources in the blue with peak wavelength of $\lambda = 4500$ Å. 22 sources were observed in the red region which has a peak wavelength of $\lambda = 5900$ Å.

A very low number of sources matched in the U, B, V, R and I. It was decided to only use the data for the R and I where 184 sources matched up. The field of view of the NIR study was a few arcseconds smaller than the field of view for the optical study so that only 172 sources from the 184 could be matched to their NIR sources. It was decided that only sources with an apparent R magnitude of 17.5 would be practically observable with the 74″ telescope, the list was narrowed to 76 candidates. From the 76 sources a total of 38 were observed. Those with $M_R < 16$ had an exposure time of 1200 seconds while if $M_R > 16$ an exposure time of 2400 seconds was used. The sequence in which the 38 sources were observed was by starting with the one with the highest H-K value.

The photometric and spectroscopic data were reduced with the IRAF software package. A collection of 25 standard E-region stars were used for the photometric calibration. A standard star was not observed in the spectroscopic study of 2002 so flux calibration was not possible, thus a second degree polynomial was fitted through each spectrum and subtracted from the spectrum profile so that the continuum was flat and the emission features were normalised relative to the strongest feature on the profile. The standard star that was used for flux calibration in the spectroscopic dataset obtained in 2011 is LTT-3864.

3. Results and discussion

3.1. Photometry

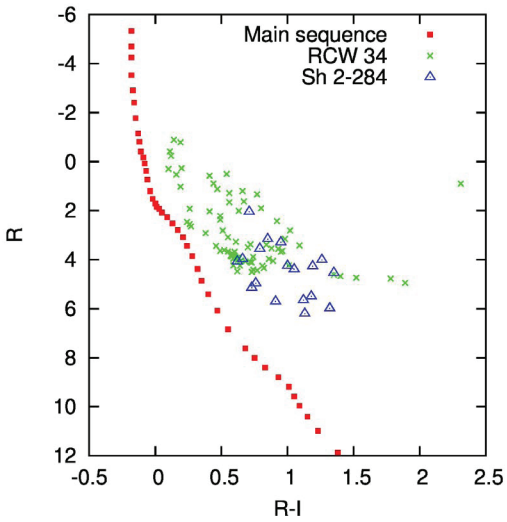


Figure 3: Colour-magnitude for the 76 candidate stars with isochrones and the 23 PMS stars observed by Cusano *et al.* [8].

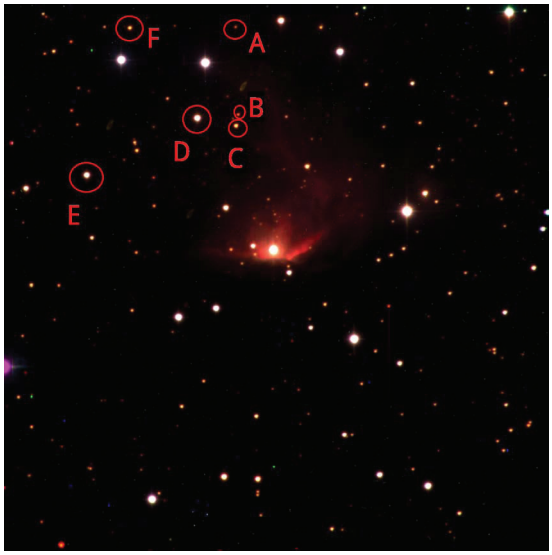


Figure 4: The 6 stars that showed an Hα emission line.

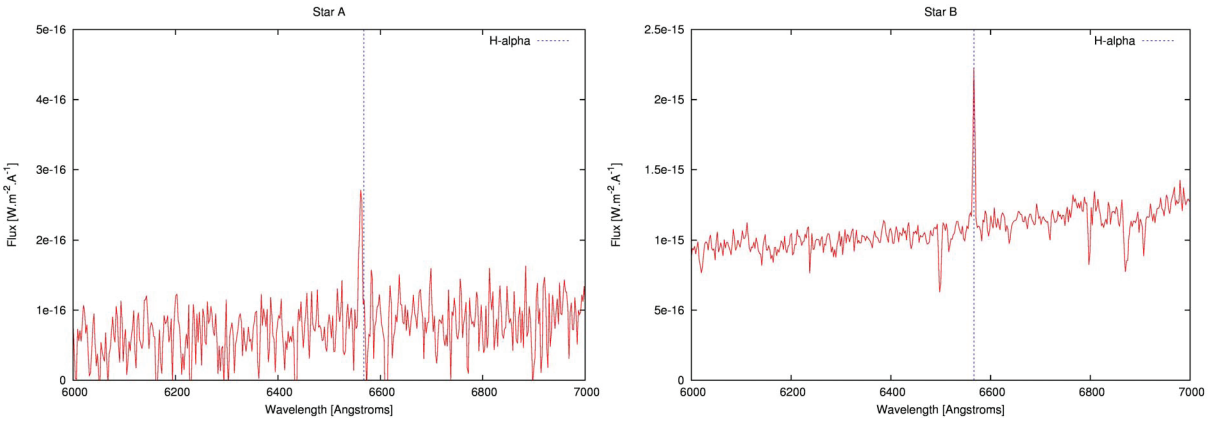
The sources were dereddend using $A_V = 4.2$ as measured for the HII region by Hedary-Malayeri *et al.*[6]. The distance modulus was calculated for the distance of 2.5 kpc as measured by Bik *et al.*[7]. A control population of 23 PMS stars as observer by Cusano *et al.*[8], were dereddened for a distance of 4 kpc and each star for its individual value of A_V. All of the

observed sources were dereddened using the interstellar extinction law as given by Rieke and Lebofsky[9].

The colour-magnitude diagram shown in figure 3 shows the population observed by Cusano *et al.* of which they determined a maximum age of 10 Myrs. A metallicity of $Z = 0.02$ was used in the construction of the main sequence branch for an age of 10 Myrs. The population for RCW 34 seems to be in the same age range as the comparison population because both populations cluster above the main sequence branch. The molecular cloud associated with RCW 34 lies towards the top of figure 1 so that the extinction may be higher for stars in this region of than those close to the HII region. Pagani *et al.*[10] shows that the molecular cloud spans the entire top half of figure 1, with the densest part of the cloud almost next to the dissociation region of the cloud. The dissociation region is the bright fringe of the HII region where the molecular hydrogen get disband by the infalling radiation from the activating star and then the separated hydrogen gets ionized by high energy photons. No stars in the population of RCW 34 shown in figure 3 have an absolute magnitude fainter than $M_R = 5$. This can be due to incorrect calculation of extinction for individual stars. It can also be due to the chosen limit of an apparent magnitude of $M_R = 17.5$ for the candidates.

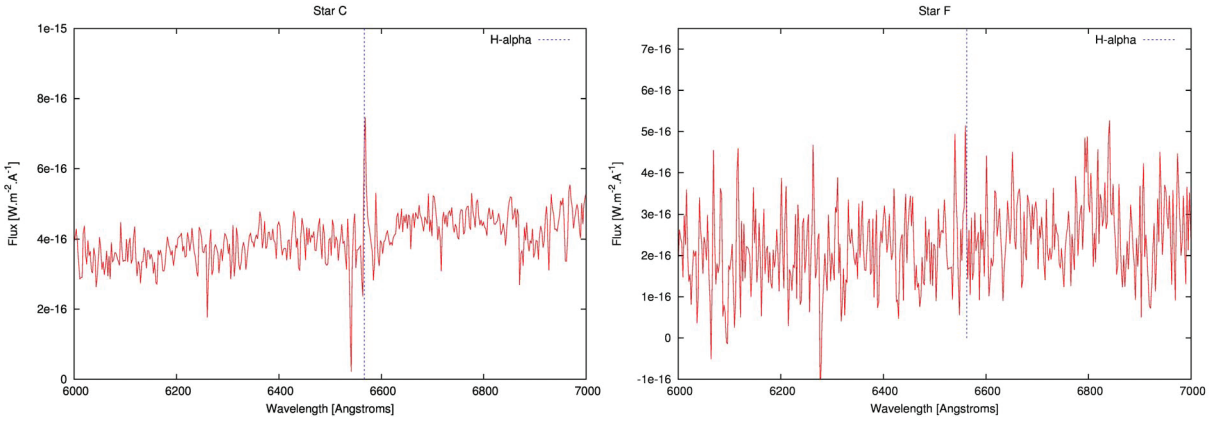
3.2. Spectroscopy

A set of 38 stars were observed during 2011 from which 4 showed an Hα emission line shown in figure 5a, 5b, 6a and 6b. Out of the set observed in 2002 two were identified to have an Hα emission line shown in figure 7b and 7a. The Hα emission line is due to high energy photons that is radiated from the proto star which is absorbed by the matter in the accretion disc so that the accreting hydrogen becomes very hot and excited and then emits light at a wavelength $\lambda = 6562.18$Åwhen an electron falls from the second energy level to the ground state. The position of each star is given in figure 4. Stars A and F lies far from the molecular cloud and the HII region. Star B and C lies in an open crescent area where the visible excited HII gas has been blown away. Star D and E are in the same region as the other four stars. All six stars are in the upper left corner of the field of view which according to Pagani *et al.*[10] has the lowest density of molecular gas in the entire field of view.



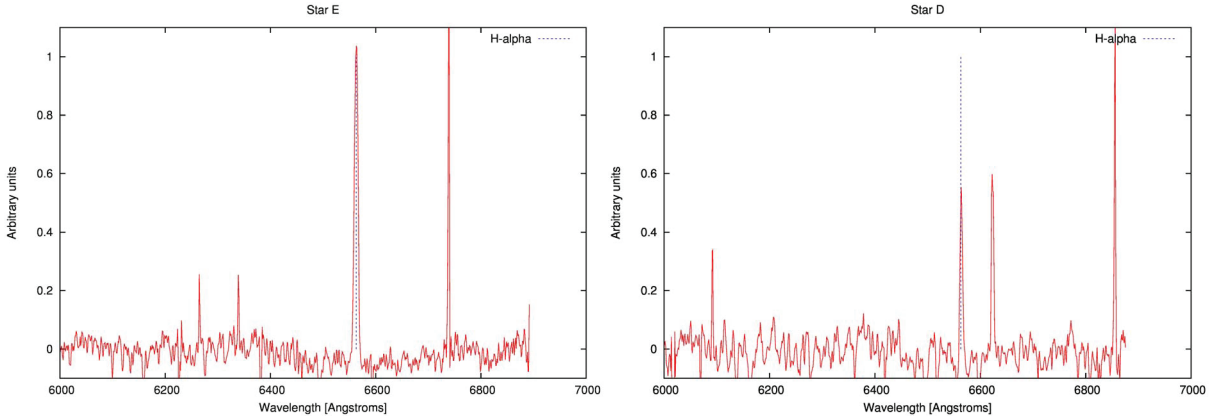
(a) Star A (b) Star B

Figure 5



(a) Star C (b) Star F

Figure 6



(a) Star E (b) Star D

Figure 7

4. Conclusion

The colour magnitude diagram in figure 3 shows that the population of 76 objects in RCW 34 are PMS stars due to their position relative to the isochrones. An accurate estimation of most of the stars' age cannot be made without more precise knowledge of extinction in the region. The position of the stars that showed Hα emission are in an open, bubble shaped part of the clouds. The visible excited HII cloud makes a clearing towards the top half of figure 1 in the form of a large arch surrounding the 6 Hα emitters. If one looks at the images given in Pagani *et al.*[10] one sees that the molecular cloud has a very similar arch shape. These objects can be seen in the lower density bubble shaped region.

References

[1] Hartman L 1998 *Accretion processes in star formation* 1st ed (*Cambridge Astrophysical series* vol 32) (Cambridge University press) ISBN 0 521 43507 2
[2] Lynden-Bell D; Pringle J 1974 *Monthly notices of the Royal Astronomical Society* **168** 603–637
[3] Meyer M R, Calvet N and Hillenbrand L A 1997 *Astronomical Journal* **114** 288–300

[4] Rodgers A W, Campbell C T and Whiteoak J B 1960 *Monthly notices of the Royal Astronomical Society* **121** 103– 8
[5] de Villiers H 2009 *A near-infrared study of the Southern high mass star forming region RCW 34* Master's thesis North-West University, Potchefstroom campus
[6] Cusano F *et al.* 2011 *Monthly notices of the Royal Astronomical Society* **410** 227– 40 (*Preprint* 1009.0435)
[7] Heydari-Malayeri M 1988 *Astronomy and Astrophysics* **202** 240– 52
[8] Bik A *et al.* 2010 *The Astrophysical Journal* **713** 883– 99
[9] Rieke G H and Lebofsky M J 1985 *The Astrophysical Journal* **288** 618– 21
[10] Pagani L, Heydari-Malayeri M and Castets A 1993 *Astronomy and Astrophysics* **275** 573– 89

Pulsating B stars in the LMC

CA Engelbrecht[1] & FAM Frescura[2]

[1]Department of Physics, University of Johannesburg, PO Box 524, Auckland Park 2006, Johannesburg. [2]School of Physics, University of the Witwatersrand, Private Bag 3, WITS 2050, Johannesburg.

E-mail: chrise@uj.ac.za

Abstract. The first discovery of Beta Cephei (BCep) pulsators in the LMC was announced nine years ago. This was a remarkable discovery, since theoretical analyses of pulsational stability had previously predicted that early B main-sequence stars with metallicities lower than Z = 0.01 should not pulsate at all. Following this announcement, and subsequent announcements of the discovery of 92 BCep candidates in the LMC, more detailed studies adopting a variety of opacity calculations and metal mixtures indicated that BCep pulsations could be explained in low-metallicity environments after all. In order to ascertain the nature of these pulsations, multi-colour photometry of sufficient precision is required. We have obtained 4 weeks of UBVI photometry on two fields in the LMC that surround stars which have been identified as strong Beta Cephei candidates from OGLE data. We report on pulsations detected in two stars taken from these fields.

1. Pulsating B stars

Early B stars are very likely to become progenitors of supernovae that will produce a neutron star as the core remnant. A precise understanding of early B stars is essential for a precise understanding of the formation and character of the lower-mass population of core-collapse supernovae. The degree of uncertainty in our present descriptions of stellar structure and evolution is illustrated in a recent paper by Pietrzynski et al. [1], where the mass value of a classical Cepheid variable, determined from traditional evolutionary models, was shown to be approximately 20% in error. Equally dramatic improvements in our understanding of the detailed structure and evolutionary timescales of B stars are possible. Such improvements will have a significant impact on our understanding of the formation of supernovae, and also on our understanding of the formation and character of pulsars. Supernovae play an important role in the study of a wide variety of astrophysical and cosmological problems, while pulsars are set to play a key role in the science programmes of MeerKAT and the SKA.

Asteroseismology is currently the most powerful tool available to solve many outstanding questions in stellar structure and evolution [2], since the oscillation frequencies of a star depend very sensitively on the details of its structure and on its size. Early B stars show dominant pulsations with periods in the range 3 – 8 hours. Many of the pulsation modes have relatively low amplitudes, requiring extensive observation to raise the signal-to-noise ratio to sufficient levels. Early B stars that display such pulsations are classified as Beta Cephei stars (after the prototype). Two of the most important questions about early B stars that invite an asteroseismological answer are concerned with the *instability strip* associated with Beta Cephei-type pulsations: i) Why have only a few O-type stars

been identified as Beta Cephei-type pulsators? ii) What are the exact locations of the edges of the instability strip? The project being reported on is concerned with the latter of these two questions.

2. Metallicity and pulsation

Almost a full century after the original discovery of Beta Cephei-type variability, the driving mechanism for the pulsations was finally tracked down to the presence of a sharp, localised increase in opacity in the temperature zone where the iron-group elements are appropriately ionised. Naturally then, the extent of pulsation mode driving that can occur in an early B star will depend on its metal content. Until fairly recently, theoretical analyses of pulsational stability for early B stars had predicted that BCep pulsations would not occur in stars with metallicities lower than Z = 0.01 [3].

Pigulski and Kolaczkowski [4] announced the first discovery of Beta Cephei (BCep) pulsators in the LMC in 2002. A few years later, Kolaczkowski and Pigulski identified 92 BCep candidates in the LMC [5]. These discoveries were unexpected, since the average metallicity of stars in the LMC is well established at a value of [Fe/H] = -0.03 (with little dispersion around this value [6]), which corresponds to Z = 0.008 [7], lower than the metallicity threshold quoted above. Following the discovery of BCep stars in the LMC, more detailed studies adopting a variety of opacity calculations and metal mixtures indicated that BCep pulsations could be explained in low-metallicity stars after all [8].

3. The observing project

Precise studies of the pulsation properties of the newly-discovered BCep stars in the LMC are required to put the new modifications in the stellar models to the test. Multi-colour photometry across the optical spectrum is the most widely-used technique for identifying the non-radial degree ℓ of the pulsations (an important parameter that needs to be determined). Furthermore, the largest possible number of observations should be obtained, to raise the signal-to-noise ratio of the data high enough that a useful number of pulsation modes can be detected.

We have embarked on a multi-year project to meet these demands for a selection of the BCep candidates that have been announced. The first season of observations was conducted on the 1.0-m telescope at the Sutherland site of the South African Astronomical Observatory (SAAO). The first two target fields for the project contain two of the only three BCep stars that have been identified in the literature to date [4]. We conducted UBVRI photometry on these two fields for two fortnights, using the STE4 camera, and covering a total time base of 48.2 days. This allows for a frequency resolution of 0.03 cycles per day in the detection of periodic signals, using the criterion of Loumos and Deeming [9].

Typical examples of the two fields as they were captured through the B filter on the STE4 camera are displayed in Figure 1:

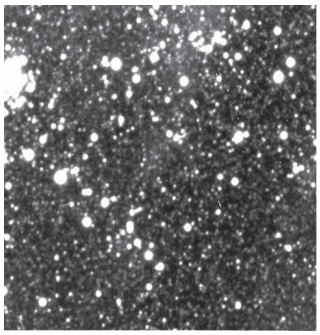


Figure 1(a). The field "LMC1".

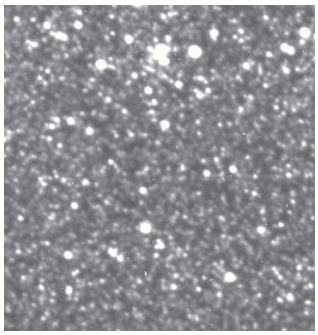


Figure 1 (b). The field "LMC2".

4. Results

Approximately 200 measurements in each of U, B, V, R and I were obtained on each of the two selected fields. The frames depicted in Figure 1 clearly show that there are dozens of stars in each field that can be explored for pulsation signatures. As a first step to ascertaining the quality of the output from this 4-week observing allocation, we compared our results to those obtained by Pigulski and Kolaczkowski [4] for the two stars that they identified in these fields: OGLE 051841.98−691051.9 (alias LMC1) and OGLE 052809.21−694432.1 (alias LMC2).

Table 1 displays frequencies and amplitudes found by Pigulski and Kolaczkowski (abbreviated as PK[4]) in the I filter and in our work in the B filter respectively:

Table 1. Comparison of frequency analysis for [4] and for this work.

Star		PK [4] (I) frequencies (c/d)	amplitudes (mmag)	Our work: (B) frequenccies (c/d)	amplitudes (mmag)
LMC1	f1	**5.179046**	5.1	**-**	
	f2	**3.495009**	3.9	**2.52 (alias?)**	14.9
	f3	**2.005502**	4.5	**2.01**	42.1
	f4	**1.684015**	3.1	**1.71**	20.0
	f5	**3.816132**	2.3	**3.76**	27.9
	f6	**-**		**5.41**	22.7
LMC2	f1	**3.497909**	12.4	**3.50**	17.3
	f2	**3.685343**	4.0	**-**	

For LMC1, we confirm all of the frequencies quoted in [4] (except for f1) if 2.52 c/d is regarded as a 1-day alias of 3.50 c/d (note that the resolution of our dataset is 0.03 c/d). Instead of f1 we find f6, differing from f1 by 0.2 c/d. It is interesting to note the amplitude ratios $A_I : A_B$ for the correlated modes, albeit from different observing programs, as displayed in Table 2:

Table 2. Amplitude ratios $A_I : A_B$ observed in LMC1, computed from our data (B) and the data reported in [4] (I).

Mode	$A_I : A_B$
f2	0.26
f3	0.11
f4	0.16
f5	0.08

As expected for an early B star, the amplitudes in I are considerably smaller than the amplitudes in B. However, the amplitudes in I are far *weaker* relative to the amplitudes in B than any current models predict. We are comparing amplitudes obtained from vastly different observing programs, which is far from ideal. A proper comparison of pulsation amplitudes across the optical spectrum can only be done once our data have been processed in all five colours.

From the perspective of stellar structure considerations, the relatively low values of the BCep frequencies detected make sense, since the BCep instability strip shrinks out to higher-mass stars as

metallicity decreases (for example, see [8]) and we expect BCep pulsations to be excited only in high-mass (i.e earlier than B2-B3) stars in the LMC's low-metallicity environment. Pulsations in higher-mass stars will indeed have longer periods. This expectation is borne out for LMC1 by the colour-magnitude plot displayed in Figure 2 (taken from reference [4]). LMC1 is the star represented by the filled circle right at the top of the diagram. The calibrating scale (for luminosity class V stars) presented in the diagram suggests that LMC1 might even be a late O star.

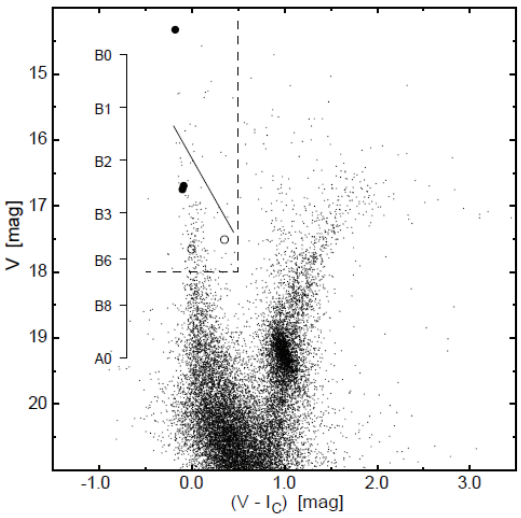


Figure 2. Colour-magnitude plot of stars in the LMC. From Pigulski & Kolaczkowski [4]

LMC2 appears to have a much more limited pulsation spectrum than LMC1. This suggests that LMC2 is perhaps right on the edge of the low-metallicity instability strip. It is represented by the lower of the two overlapping filled circles shown in Figure 2, corresponding to a B2.5 dwarf on the calibrating scale. Therefore, if it is tracing an edge of the instability strip, it has to be the *lower* edge. Interestingly, the $A_I : A_B$ ratio for the single mode we detected in LMC2 is 0.72, in line with theoretical predictions for a low-ℓ mode. The discrepancy between the amplitude ratios seen in LMC1 and LMC2 respectively warrants further investigation on our part.

5. Conclusions

We have shown that a 4-week multi-colour observing programme on a 1.0-m class telescope allows us to accurately glean multiple periods out of a BCep star in the LMC. We observed two fields in this first stage of the project; the signal-to-noise ratio could be improved by a factor of more than 50% by limiting a 4-week programme to one field only.

We have identified a number of multi-periodic variable stars in these two fields, besides the test cases of LMC1 and LMC2. It is envisaged that a programme covering 5 fields in total will deliver precise data on twenty-odd BCep stars – a sufficiently large sample to make significant impacts on the modelling of B stars of low metallicity.

Acknowledgments

The National Research Foundation and the University of Johannesburg's Central Research Fund as well as its Faculty of Science are thanked for providing financial support for this project.

The Director of SAAO is thanked for awarding sufficient observing time on the 1.0-m telescope in support of this project.

Dr LA Balona is thanked for valuable discussions.

References
[1] Pietrzynski G et al. 2010 **Nature (London) 468** 542
[2] Christensen-Dalsgaard J 2010 **Astronomische Nachrichten 331** 866
Guzik J 2010 **Astrophysics and Space Science**, in press
Roxburgh I 2010 **Astrophysics and Space Science 328** 3
[3] Pamyatnykh A 1999 **Acta Astronomica 49** 199
[4] Pigulski A and Kolaczkowski Z 2002 **Astron Astrophys 388** 88
[5] Kolaczkowski Z and Pigulski A 2006 **Memorie della Società Astronomica Italiana**. **77** 336
[6] Luck, RE et al. 1998 **Astron J 115** 605
[7] Bertelli G et al. 1994 **Astron Astrophys Suppl Ser 106** 275
[8] Miglio A et al. 2007 **Mon Not R Astron Soc 375** L21
MiglioA et al. 2007 **Communications in Asteroseismology 151** 48
Zdravkov T and Pamyatnykh A 2008 **J Phys Conference Series 118** 012079
[9] Loumos GL and Deeming TJ 1978 **Astrophysics and Space Science 56** 285

Lie derivatives in stellar perturbation theory

F A M Frescura[1] and C A Engelbrecht[2]

[1]School of Physics, University of the Witwatersrand, Private Bag 3, WITS 2050, South Africa
[2]Department of Physics, University of Johannesburg, P O Box 524, AUCKLAND PARK 2006, South Africa

E-mail: fabio.frescura@wits.ac.za

Abstract. Perturbation theory uses Lagrangian techniques that require vector fields to be compared at finitely separated points. This theory can be generalised to the strong gravitational field regime in one of two ways, using either covariant or Lie derivatives. In this paper, we argue that methods based on the Lie derivative are more useful. The Lie derivative provides a clear picture of how the deformation of the fluid flow takes place. We also propose a framework for fluid perturbation theory that provides a natural way to discuss large perturbations.

1. Introduction

Perturbation techniques in fluid dynamics were developed principally to study the stability of solutions of the equations of motion of a fluid [1]. In general, the study of stability requires only infinitesimal deviations from the known solution to be considered. In situations where a flow is known to be stable, the same techniques used in the study of stability can be extended to investigate the properties of flows that differ by only small amounts from a given flow and to deduce approximate solutions for them.

In this paper, we propose a general framework for developing fluid perturbation theory in a way that permits generalisation to arbitrary spacetimes and allows for flow perturbations of arbitrary size. Our proposal is an extension of the work of Friedman and Schutz (see reference [4] and the references therein). The key to their work lies in replacing the covariant derivative by the Lie derivative. The former can be defined only on spacetimes on which a connection is defined. The latter does not require any special structure, and may defined whenever a flow is present. We argue, therefore, that methods based on the Lie derivative are more natural and more useful in perturbation theory than those based on the covariant derivative. Furthermore, the Lie derivative provides a clear picture of how the deformation of the flow takes place. It also provides a means for searching for differential equations that define deformations of some specified kind.

In Section 1, we consider the different available definitions of Lagrangian displacement. In Section 2, we show why the elementary definition of it is unsuitable for use in advanced studies, discuss the alternative definition in terms of covariant derivatives, and show that the most general definition available is in terms of the Lie derivative. We outline a general mathematical framework for extending perturbation theory in Section 3. The framework proposed is not limited to small perturbations, but can be extended to perturbations of arbitrary size. In Section 4, we review the concepts of Lie transport and of the Lie derivative. Finally, in Section 5, we justify the definition of Lagrangian difference in terms of the Lie derivative and use this

to define a derivative that measures rate of change with respect to deformation. The work presented here is still in its infancy, but preliminary results suggest that it is worth pursuing.

2. Lagrangian variation

Two types of variation are used in perturbation theory, called respectively Eulerian and Lagrangian [1]. The Eulerian description of fluids is in terms of time-dependent fields defined in three dimensional space. The Eulerian variation measures by how much a given field of the perturbed flow differs from the corresponding field in the unperturbed flow. To determine the Eulerian variation, one inspects the value of the field for the perturbed flow at a given point $\vec{x}$ in space at time t and compares it with the value of the corresponding field for the unperturbed flow *at the same space point* at time t. To express this analytically, denote the field for the perturbed and unperturbed flows generically by the symbols Q and Q_0 respectively. The fields Q and Q_0 might be scalar, vector or tensor fields. Then the Eulerian difference for the field Q at position $\vec{x}$ at time t is defined to be

$$\delta Q(\vec{x},t) \quad = \quad Q(\vec{x},t) - Q_0(\vec{x},t). \tag{1}$$

The difference δQ is of two fields defined at the same space point, and so is well defined irrespective of the nature of the field Q. Elementary problems in perturbation theory can be solved exclusively in terms of δQ, which is easy to use mathematically, and is relatively easy to interpret physically. More advanced problems, however, require us to identify given material elements in the fluid and to track and compare their properties in the perturbed and unperturbed flows. In general, position $\vec{x}$ is occupied at time t in the two flows by *different* material elements of the fluid, so δQ cannot be used directly to compare how the variable Q differs in the two flows for one and the same material element.

To track a given fluid element, we need to switch from the Eulerian description of the fluid to the Lagrangian. In the Lagrangian description, each material point of the fluid is assigned, once and for all, a fixed set (a^1, a^2, a^3) of coordinates and its motion through space is then tracked by means of its trajectory

$$x^i \quad = \quad F^i(\vec{a},t). \tag{2}$$

The function $\vec{F}$ is called the *flow*. It delivers the position $\vec{x}$ in space of the particle at time t when its material coordinates $\vec{a}$ and the value of t are used as inputs into the function $\vec{F}$. Denote the perturbed and unperturbed flows by $\vec{F}$ and $\vec{F}_0$ respectively. The Lagrangian difference ΔQ in the value of a fluid variable Q is defined in elementary treatments by

$$\Delta Q \quad = \quad Q(\vec{F}(\vec{a},t),t) - Q_0(\vec{F}_0(\vec{a},t),t). \tag{3}$$

Using the same value of $\vec{a}$ in both arguments guarantees that we are inspecting the value of Q for the same material element in both flows.

This definition of ΔQ is problematic. It requires us to compare the values of the fields Q and Q_0 at *different* points in space. This poses no difficulties when Q and Q_0 are scalar fields. However, in general, vector and tensor fields such as fluid velocity, stress and rate of strain, at different space points cannot be compared. To effect such a comparison, we need a method for 'copying' the field Q to the same space point used to evaluate Q_0. We can then form the difference between the 'copy' of Q and Q_0 at the same space point $\vec{x} = F_0(\vec{a},t)$. This problem does not occur when the fluid flow takes place in an Euclidian space equipped with Cartesian coordinates, the setting normally assumed for perturbation problems. However, it rears its head the moment we choose to use non-Cartesian coordinates, or to consider flows in spaces other than Euclidian.

Tensor fields at different locations may be compared by one of two methods. The one used most commonly relies on the presence in the space of a connection and of an associated covariant derivative [2]. The 'remote' field Q can then be copied by parallel transport to the position where Q_0 is measured and compared to it. The advantage of using this method is that it provides the most obvious generalisation of the elementary definition of Lagrangian displacement, given in equation (3). Comparison of tensor fields at different points in an Euclidian space relies on the property of distant parallelism, which is a special property of its connection. This property is obscured when we use Cartesian coordinates, 'hidden in plain sight' as it were, but forces itself into our consciousness when we change to curvilinear coordinates, manifesting itself through the path independence of parallel transport. Distant parallelism means that we are able to parallel-transport $Q(\vec{x}', t)$ unambiguously from position $\vec{x}'$ to position $\vec{x}$ and there to construct an unique 'copy' $\widetilde{Q}(\vec{x}, t)$ of $Q(\vec{x}', t)$. Using this unique copy, we can then form the difference

$$\Delta Q \quad = \quad \widetilde{Q}(\vec{x}, t) - Q_0(\vec{x}, t). \tag{4}$$

The quantity ΔQ is unambiguously defined by virtue of the path independence of parallel transport in Euclidian space, and can thus be used to replace definition (3) of the Lagrangian difference. This definition leads to the approximate relation

$$\Delta Q \quad = \quad \delta Q + \nabla_{\vec{\xi}}\, Q \tag{5}$$

where $\vec{\xi}$ is the Lagrangian displacement (defined in the next section), used by many authors. See for example references [1] and [2].

Many spaces of interest, however, do not possess the property of distant parallelism, even when they possess a connection, so definition (4) of the Lagrangian difference also fails. In a general connected manifold, parallel transport is in general path dependent. The quantity $\widetilde{Q}(\vec{x}, t)$ is not uniquely defined, making definition (4) empty of content. The only exception to this occurs when the points $F(\vec{a}, t)$ and $F_0(\vec{a}, t)$ are infinitesimally close. This means that we can continue to use definition (4) provided that we are interested only in perturbed flows that differ only infinitesimally from the unperturbed one.

There is an alternative method available for comparing tensor fields at different positions. This method does not rely on the existence of a connection on the space, is available in all manifolds, and does not require the perturbation to be small. It is therefore suitable not only for the discussion of stability of flows in general spaces, but can also be extended to deal with large amplitude perturbations. In this method, the concept of Lie transport replaces that of parallel transport, and the Lie derivative that of the covariant derivative.

3. A model for perturbations

Rather than consider a single perturbed flow $\vec{F}$, which would force us to work only with finite differences in the flow variables, it is simpler to work with a one-parameter family $\vec{F}_\lambda$ of perturbed flows. One way to visualise these is by constructing a 1-dimensional continuum of spacetimes M_λ. We can then regard each perturbed flow F_λ as taking place on a separate sheet M_λ of these stacked spacetimes, with the flow on each sheet differing infinitesimally from one on the next sheet. Another way to visualise this is by regarding all of the flows as defined on a single spacetime M, which is the canonical projection of the spacetime stack in the direction of λ. This projection, applied to a single perturbed flow, is visualised by Lynden-Bell and Ostriker as a 'ghostly flow' accompanying the perturbed flow in much the same way as a residual image accompanies the real image on a poor quality television screen.

The 1-parameter family of perturbed flows can be regarded as 'distortions' of the unperturbed flow. Were we, at fixed time t, to allow the parameter λ to run through values from 0 to

some finite value, each material element of the flow would trace out a curve in space. This three parameter family of curves defines the *distortion flow*, $G(\vec{x}, t, \lambda)$. This flow in turn defines a vector field $\vec{\eta}(\vec{x}, t, \lambda)$. Knowledge of $\vec{\eta}$ would allow us to deduce the distortion flow G by integration, and hence to construct fully the perturbed flow corresponding to any given parameter value λ. The problem of perturbation then reduces to that of finding equations that determine the distortion field $\vec{\eta}(\vec{x}, t, \lambda)$. The form and nature of these equations depend not only on the equations of motion of the fluid, but also on the constraints that the perturbations are assumed to obey.

The Lagrangian displacement field $\vec{\xi}(\vec{x}, t)$ of standard perturbation theory is related to the field $\vec{\eta}(\vec{x}, t, \lambda)$ at the parameter value $\lambda = 0$. If we confine ourselves to small perturbations, we can express the distortion flow approximately in the form

$$G^i(\vec{x}, t, \lambda) \quad = \quad G^i(\vec{x}, t, 0) + \lambda \, \frac{\partial G}{\partial \lambda}(\vec{x}, t, 0) + \cdots = x^i + \lambda \eta^i(\vec{x}, t) \tag{6}$$

so that $\vec{\xi}(\vec{x}, t) = \lambda \, \vec{\eta}(\vec{x}, t, 0)$. We do not develop this general scheme further in this paper. We now return to the definition of the Lagrange difference in a general space.

4. The Lie derivative

Lie transport is the convective transport of objects by means of a flow [3]. Any point in the fluid will convect in time t to a new position in space. Any curve consisting of fluid points will also convect to a new position. A vector can always be constructed as the tangent to a given curve that passes through its point of attachment. The convected curve will have a tangent vector at the convective image of the attachment point. This vector is defined to be the convective image of the given vector and is said to be *Lie transported* by the flow. The Lie transport of a tensor is defined inductively from that of a vector by contracting it with the requisite number of vectors (and covectors) to form an invariant. The Lie transport of the tensor is then that tensor at the transported point that yields the same value of the invariant by contraction with the transported vectors as is yielded by the original tensor with the original vectors. If T is now taken to be a time independent tensor field rather than a single tensor, its Lie transported image can be compared at any point on a flow line with the tensor defined by the field at that point. The Lie derivative measures the rate at which the tensor of the field differs from its Lie transported version at the same point.

Formally, this may be seen as follows. Let $F(\vec{x}, t)$ be a flow. Denote by ϕ_t the diffeomorphims defined by F. Thus, $\phi_t(\vec{x}) = F(\vec{x}, t)$. The fluid point which at time $t = 0$ occupies position $\vec{x}$, will be found at time t at position $F(\vec{x}, t)$. The field T at this position has value $T(F(\vec{x}, t)) = T(\phi_t(\vec{x}))$, while at its original position it has value $T(\vec{x})$. Lie transport of $T(\vec{x})$ to position $F(\vec{x}, t)$ produces a tensor $\widetilde{T}$ at that position. It can be shown (see reference [3], p 90, Theorem 2.2.20 and the discussion preceding it) that $\widetilde{T}$ is given by $T\phi_t(T(\vec{x}))$, where $T\phi_t$ is the derivative of the mapping ϕ_t. Since this is a value at the point $F(\vec{x}, t)$, it can be compared directly to that of the field T at that point. The difference in these values is a well defined quantity, called the *Lie difference*. For small values of t, it can be used to define a differential quotient and, in the limit as $t \to 0$, it defines the *Lie derivative*. It is worth noting that we have not needed to introduce any additional structure on the space in the definition of any of these operations, and so this method of forming differences and derivatives is available on all types of manifold.

The above operations are easy to visualise physically. However, to get a cleaner mathematical definition, it is better to perform the above procedures in reverse. Instead of convecting $T(\vec{x})$ forwards with the flow to obtain its Lie transport $\widetilde{T}$, reverse the flow in time and Lie transport the value $T(\phi_t(\vec{x}))$ backwards to $\vec{x}$. Then compare this transported value with $T(\vec{x})$ as reference.

Mathematically, the forward-transport and backward-transport procedures are equivalent and yield the same result. The time-reversed transport is called 'pull-back' and yields the field $T\phi_t^{-1}(T(\phi_t(\vec{x})))$. It is denoted more briefly by $\phi_t^* T(x)$. The Lie derivative is denoted by $\mathcal{L}_{\vec{X}} T$, where $\vec{X}$ is the generating vector field for the flow F (that is, the velocity field of the flow), and is thus defined by

$$\mathcal{L}_{\vec{X}} T(\vec{x}) \;=\; \lim_{\varepsilon \to 0} \frac{\phi_\varepsilon^* T(\vec{x}) - \phi_0^* T(\vec{x})}{\varepsilon}. \tag{7}$$

Since ϕ_0 is the identity, we have $\phi_0^* T(\vec{x}) = T(\vec{x})$. All fields in the above definition are evaluated at the same point $\vec{x}$ in space, so we can omit arguments without fear of ambiguity to write

$$\mathcal{L}_{\vec{X}} T \;=\; \lim_{\varepsilon \to 0} \frac{\phi_\varepsilon^* T - T}{\varepsilon} \tag{8}$$

Note that the vector field $\vec{X}$ that generates the flow can be *any* vector field and need not be the velocity field of some actual fluid flow. For example, it could be the vector field $\vec{\xi}$ that generates perturbations or deformations of a given unperturbed flow.

5. Lagrangian difference

Suppose T_λ is a tensor field associated with the perturbed flow $F_\lambda(\vec{x}, t)$. We want to compare the value of T at the fluid point which in the unperturbed flow occupies position $\vec{x}$ at time t, with the value of the field T_λ at the position $G(\vec{x}, t, \lambda)$ occupied by the same fluid point at the same time. In a general space, the only means by which this may be done is by Lie transporting the value $T_\lambda(G(\vec{x}, t, \lambda), t)$ backwards along the deformation flow G to position $\vec{x}$. This can be done via pull back and yields the value $TG_\lambda^{-1}(T_\lambda(G(\vec{x}, t, \lambda), t)) = G_\lambda^* T_\lambda(\vec{x}, t)$. This is the value of the tensor field $G_\lambda^* T_\lambda$ at $\vec{x}$, so it can be compared with the value $T_0(\vec{x}, t)$ of the field T_0. The difference

$$\Delta T \;=\; G_\lambda^* T_\lambda(\vec{x}, t) - T_0(\vec{x}, t) \tag{9}$$

is thus a well defined quantity. Since the 'classical' notion of a Lagrangian difference is undefined on a general manifold, and since the difference defined in the above equation encapsulates as closely as possible the spirit of the 'classical' notion of a Lagrangian difference, we may take this to be the correct generalisation of that concept to an arbitrary manifold [4].

To turn this definition into an useful mathematical operation, convert the above difference into a derivative. This is done as follows. First note that

$$\begin{aligned} \Delta T &= G_\lambda^* T_\lambda(\vec{x}, t) - G_0^* T_\lambda(\vec{x}, t) + G_0^* T_\lambda(\vec{x}, t) - T_0(\vec{x}, t) \\ &= \left[G_\lambda^* T_\lambda - T_\lambda\right](\vec{x}, t) + T_\lambda(\vec{x}, t) - T_0(\vec{x}, t). \end{aligned} \tag{10}$$

The last two terms are the Eulerian difference δT, while the first two are related to the Lie derivative. To first order in λ, we have

$$\Delta T \;=\; \lambda\, \mathcal{L}_{\vec{\eta}} T_\lambda(\vec{x}, t) + \delta T(\vec{x}, t). \tag{11}$$

Writing $\vec{\xi} = \lambda \vec{\eta}$ gives us the formula

$$\Delta T \;=\; \mathcal{L}_{\vec{\xi}} T_\lambda + \delta T \tag{12}$$

used by Friedman and Schutz. They advocate the use of this definition with the claim that it is 'somewhat more natural' than that defined in equation (5). They might have claimed more:

this is the only definition of the Lagrangian difference that can be used universally, irrespective of the nature of the manifold on which the flow occurs. We define also the derivative

$$\left.\frac{d}{d\lambda}T_\lambda\right|_{\lambda=0}(\vec{x},t) = \lim_{\lambda\to 0}\frac{\Delta T}{\lambda} \quad = \quad \mathcal{L}_{\vec{\eta}}T(\vec{x},t,0) + \frac{\partial T}{\partial\lambda}(\vec{x},t,0) \tag{13}$$

which measures the rate at which the field T of the unperturbed flow changes per unit deformation as we move to the perturbed flow. Used in conjunction with the equations of motion of the fluid reformulated in terms of the Lie and exterior derivatives, the above scheme provides a method for the extension of perturbation techniques to encompass also large perturbations. This may provide an useful method for dealing with large amplitude oscillations in stars.

References

[1] Cox J P 1980 *Theory of Stellar Pulsation* (Princeton, NJ: Princeton University Press)
[2] Lynden-Bell D and Ostriker J P 1967 *Mon Not R Astron Soc* **136** 293
[3] Abraham R and Marsden J E 1985 *Foundations of Mechanics* (Redwood City, Ca: Addison-Wesley Publishing Company, Inc)
[4] Friedman J L and Schutz B 1978 *Astrophys J* **221** 937

An African VLBI network of radio telescopes

M J Gaylard[1], M F Bietenholz[1,2], L Combrinck[1], R S Booth[3,4], S J Buchner[1], B L Fanaroff[3], G C MacLeod[5], G D Nicolson[1], J F H Quick[1], P Stronkhorst[1], T L Venkatasubramani[3]

[1] HartRAO, P. O. Box 443, Krugersdorp 1740, South Africa
[2] Department of Physics and Astronomy, York University, Toronto, Ontario M3J 1P3, Canada
[3] SKA South Africa, 17 Baker St, Rosebank, Johannesburg, South Africa
[4] Onsala Space Observatory, SE-439 92 Onsala, Sweden
[5] Department of Science and Technology, Private Bag X894, Pretoria 0001, South Africa

E-mail: mike@hartrao.ac.za

Abstract. The advent of international wideband communication by optical fibre has produced a revolution in communications and the use of the internet. Many African countries are now connected to undersea fibre linking them to other African countries and to other continents. Previously international communication was by microwave links through geostationary satellites. These are becoming redundant in some countries as optical fibre takes over, as this provides 1000 times the bandwidth of the satellite links.

In the 1970's and 1980's some two dozen large (30m diameter class) antennas were built in various African countries to provide the satellite links. Twenty six are currently known in 19 countries. As these antennas become redundant, the possibility exists to convert them for radio astronomy at a cost of roughly one tenth that of a new antenna of similar size.

HartRAO, SKA South Africa and the South African Department of Science and Technology (DST) have started exploring this possibility with some of the African countries.

1. Introduction

New undersea fibre optic cables provide communication bandwidths a thousand times greater than radio links via communications satellites. Large antennas at satellite Earth stations are becoming redundant as a result. This trend has now reached Africa with the advent of fibre optic cables down the East and West coasts of the continent.

More than two dozen large Satellite Earth Station antennas built in the 1970's and 80's exist in Africa. Twenty six are currently known in 19 countries. Antennas that are redundant could be refitted for radio astronomy at comparatively small cost.

Africa currently has only one antenna equipped to participate in global radio astronomy using Very Long Baseline Interferometry (VLBI) for high resolution imaging, this being the 26m ex-NASA antenna at Hartebeesthoek in South Africa. The window of opportunity to create an African network of VLBI-capable radio telescopes from redundant large satellite antennas exists now. This African VLBI Network (AVN or AfVN) would initially work with existing VLBI arrays such as the European VLBI Network (EVN), but could operate independently if sufficient antennas become available.

Such a network would bring new science opportunities to participating countries on a short time scale, enable participation in SKA pathfinder development and science, and would

help create the environment for bringing the Square Kilometre Array (SKA) to Africa. It would provide a ready made VLBI extension for the MeerKAT radio telescope currently being developed as an SKA precursor.

2. Intelsat Satellite Earth Stations

2.1. Intelsat Standard A Satellite Earth Station antennas

From the 1960's, communication via satellites orbiting the Earth was introduced to carry voice, data and TV signals, to supplement undersea cables. The radio bands allocated for this were 5.925 − 6.425 GHz for uplink and 3.700 − 4.200 GHz for downlink. These are within the frequency range known as "C-band". Intelsat defines Earth Stations for this band as Standard A, B, F or H depending on their technical characteristics [1]. Initially a Standard A antenna had to be at least 30m in diameter, and the antennas built in Africa from 1970 to 1985 are this size. From 1985 new technology enabled satellite transmitter power to increase and the required size was greatly reduced.

2.2. Large Satellite Earth Station antennas in Africa

SA SKA partner countries identified with large antennas are South Africa (3 antennas), Ghana, Kenya, Madagascar and Zambia. Other African countries in which large antennas have been located are Algeria (2), Benin, Cameroon (2), Congo Peoples Republic, Egypt (2), Ethiopia, Malawi, Morocco, Niger, Nigeria (3), Senegal, Tunisia, Uganda and Zimbabwe. These two groups are shown in Fig. 1. In addition, Gabon probably has one and one is reported in the Congo Democratic Republic. The antenna in Mozambique was dismantled and probably also the one in Togo. South Africa's SKA partners in Africa are Botswana, Ghana, Kenya, Madagascar, Mauritius, Mozambique, Namibia and Zambia.

2.3. Large Satellite Earth Station antennas outside Africa converted for radio astronomy

Operational converted antennas are the 30m Ceduna antenna in Australia [2] and the 32m Yamaguchi antenna in Japan [3]. The Warkworth 30m antenna in New Zealand was handed over for conversion on 19 November 2010 [4]. In Peru the Sicaya 32m antenna is being converted with assistance from the Yamaguchi team and saw first light in March 2011 [5]. Three antennas at the Goonhilly Downs station in the UK decommissioned in 2008 are proposed for conversion for use with e-Merlin and the EVN [6]. On 10 May 2011 it was announced that the 32m Elfordstown antenna outside Cork in Ireland is to be converted [7].

3. Astronomical applications for redundant large antennas in Africa

3.1. Astronomical Very Long Baseline Interferometry (VLBI)

Widely separated radio telescopes operating together create a virtual telescope equal in size to the project distance, or baseline, between the telescopes. The angular resolution of a telescope depends on its size, so the larger the separation, the better the resolution. Practically, to create good images a number of telescopes are needed, separated by small, medium and large distances.

The HartRAO 26m radio telescope is valuable in providing long baselines to radio telescope arrays in on other continents, e.g. Europe (European VLBI Network - EVN) and Australia (Australia Telescope Long Baseline Array - AT-LBA), and thus high angular resolution imaging. However, the large distance between the South African telescope and the others makes for less than ideal imaging. It would be helpful to have antennas filling the gap between SA and Europe and SA and Australia (for example) and the image quality would be substantially improved. They would be a powerful addition to the radio telescope arrays on other continents.

To create an interferometer able to operate independently to produce images, a minimum of four antennas are needed to provide both phase and amplitude closure. Thus a minimal African

Figure 1. Locations of identified large Satellite Earth Station antennas in Africa

VLBI Network able to operate independently of the radio telescopes on other continents could be formed if four suitable antennas were brought into use on the continent and on neighbouring islands. Science capability would improve substantially if more large antennas could join the network. With four antennas, 50% of the phase information and 33% of the amplitude information can be recovered; with ten antennas these rise to 80% and 78% respectively [8].

Astronomical objects suitable for study by VLBI are those that are radio-bright and of small angular size. Objects of very large physical size meet this requirement if they are quite distant, and radio galaxies and quasars are examples. Radio-bright supernovae in external galaxies are good examples of objects whose evolution can be studied with VLBI. Masers in star-forming regions in the Milky Way are examples of nearby bright, compact sources. Methanol masers at 6.668 GHz and 12.178 GHz are of particular interest currently. Repeated mapping would help in investigating the causes of their variabilitys, especially those found to be showing regular variations. Measurement of their annual parallaxes by repeated VLBI observations enables their distances to be determined, and thus the locations of the spiral arms in the Milky Way, where these occur, to be measured accurately. This is only being done by northern VLBI arrays currently, leaving the fourth quadrant of the Milky Way, in the far southern skies, so far unmapped by this method.

Distances to pulsars can be obtained by astrometric VLBI. Transient sources and gamma-ray bursters are all potential targets. The on-going study of southern hemisphere radio galaxies (TANAMI project) would be enhanced with extra telescopes able to see the southern skies. Microquasars in the Milky Way are targets, as are radio-loud interacting binary stars such as Circinus X-1.

Spacecraft in interplanetary space and at other planets that are equipped with 8.4 GHz

frequency-stable transmitters are being used for precise spacecraft position determination and for study of the interplanetary medium. This technique is known as the Planetary Radio Interferometry and Doppler Experiment (PRIDE). The HartRAO 26m telescope, and others, are regularly observing the Venus Express (VEX) spacecraft in orbit around Venus for this purpose. The Russian Phobos-Grunt mission to Mars' moon Phobos is due for launch late in 2011 and will use this technique. Converted satellite antennas with 8.4 GHz receivers in Africa would be good for this owing to their collecting area and location near the equator.

Antennas equipped with dual 2.3+8.4 GHz receivers would be able to participate in geodetic and astrometric VLBI with the current generation of radio telescopes carrying out this research. This technique provides a very precise absolute location for the telescope, which can be transferred to co-located relative position measuring systems such as the global navigation satellite systems GPS and GLONASS, and the laser ranging and DORIS systems used to obtain precise measurents of satellite orbits.

3.2. Single-dish astronomy

The only VLBI array that runs essentially continuously is the US Very Long Baseline Array (VLBA). The EVN now runs much more frequently than the traditional three four-week sessions per year of the past, with monthly e-VLBI (real-time data streaming to the correlator) and Target Of Opportunity (TOO) VLBI. However there would still be substantial non-VLBI time available for operation of each telescope as a stand-alone instrument.

The time available for single-dish astronomy would be valuable for student training purposes and for selected research projects. Techniques that could be developed would be radiometry - measuring the brightness of broad-band emission sources, spectroscopy - measuring emission and absorption line strengths, and pulsar timing - measuring the arrival times of pulses from neutron stars. Establishing a capability for unattended queue-scheduled single-dish observing (as on the HartRAO 26m telescope) would greatly assist in permitting time series to be built up of the behaviour of scientifically interesting variable sources that can be observed with these techniques.

4. Towards the African VLBI Network

4.1. Identifying candidate antennas

Intelsat documentation and internet mapping services using high resolution photography of the Earth's surface enabled the current existence of the large antennas to be confirmed. Postage stamps commemorating their inauguration were another valuable source of information. Government level contact has been made with some of the countries by DST to further investigate the possibility of conversion. Ghana was one of those approached, and was quick to seize the opportunity. This opportunity is a direct result of the African SKA Project.

4.2. The Kuntunse Satellite Earth Station in Ghana

The Ghanaian government and Vodafone, the owners of the Kuntunse Satellite Earth Station outside Accra, have agreed to convert the 32m antenna at the station for radio astronomy. This antenna has been out of service for about ten years; an adjacent 16m antenna is carrying the remaining satellite traffic. A HartRAO team and Prof. Michael Jones from Oxford University visited Kuntunse in March 2011 for an initial feasability investigation, which indicated that it was in a suitable condition to proceed (Fig. 2). A HartRAO/SKA South Africa team visited in May to assist in starting the rehabilitation.

This antenna is located roughly half way between South Africa and Europe. Fig. 3 shows how it fills the gap in UV coverage for VLBI of the EVN with Hartebeesthoek. It is 6^{o} north of the equator and so it can see the entire plane of the Milky Way and almost the entire sky.

Figure 2. The Kuntunse 32m antenna photographed in March 2011.

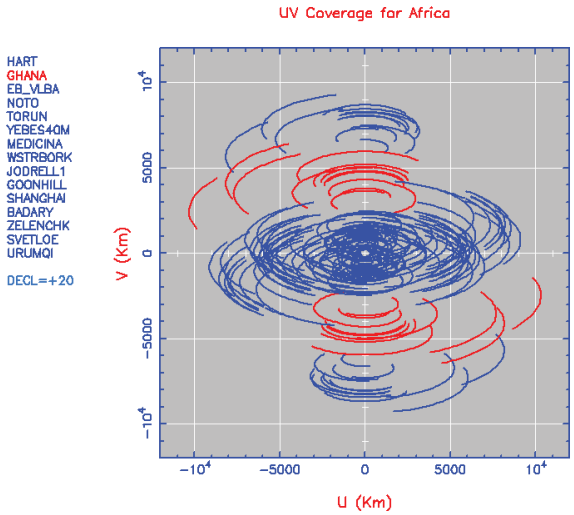


Figure 3. VLBI UV diagram for a source at +20° declination. The red tracks show how the Kuntunse antenna improves the UV coverage by filling the gap between Europe and South Africa.

4.3. The antenna conversion process

The drive systems of out of service antennas will generally need some rectification, and this is in process with the Kuntunse antenna.

More challenging is the development of radio astronomy receiver systems. The initial aim is to build receivers for operation in radio astronomy bands within the original Intelsat band range, possibly re-using the existing microwave feed horn. The most commonly used frequency band for VLBI in the EVN is 4.8 - 5.1 GHz; it falls between the current transmit and receive bands of the antenna and should not be difficult to implement with the existing feed. The 6.7 GHz methanol maser line lies a little above the feed design upper frequency and tests will be needed to see if it works at this frequency. Wide (octave) band feeds and multi-band feeds are now commonplace. A feed covering 4.5 to 9 GHz, either continuously or only in the relevant radio astronomy bands, would include the two bands already discussed and the 8.4 GHz VLBI band and would permit a wide range of science to be done.

Various options exist for developing capability outside these bands. The antenna design is the so-called "beam waveguide" in which large diameter pipes with 45° mirrors pass the signal to a focus in the room below the antenna, where the receiver is located. Most of the African antennas are of this design. It comes with some advantages and disadvantages compared to the Cassegrain design widely used on radio telescopes of this size, but the technical issues are solvable.

5. Summary

The potential for converting obsolete large antennas for radio astronomy has been recognised World-wide. The opportunity exists for African countries to re-use these expensive assets that are becoming redundant to promote science development on the continent, at relatively low cost.

References

[1] *Intelsat Earth Station Standards (IESS) Document IESS-207 (Rev. 4)* 2005 (Luxembourg: Intelsat)

[2] McCulloch P M, Ellingsen S P, Jauncey D L, Carter S J B, Cimo G, Lovell J E J and Dodson R G 2005 *Astron J* **129** 2034

[3] Fujisawa K, Mashiyama H, Shimoikura T and Kawaguchi N 2002 *Proc. IAU 8th Asian-Pacific Regional Meeting, vol II* ed S Ikeuchi, J Hearnshaw and T Hanawa (Astron Soc Japan) p 3

[4] Mansfield I 2010 *Telecom New Zealand Converts Satellite Dish into Radio Telescope* http://www.-cellular-news.com/story/46534.php

[5] Finn E 2011 *Peru: Interstellar contact made in Huancayo observatory* http://www.livinginperu.com/-news-14300-environmentnature-peru-interstellar-contact-made-huancayo-observatory

[6] Heywood I, Klockner H-R, Beswick R, Garrington S T, Hatchell J, Hoare M G, Jarvis M J, Jones I, Muxlow T W B and Rawlings S 2010 *Astronomy with Megastructures: Joint Science with E-ELT and SKA* ed I Hook, D Rigolpoulou, S Rawlings and A Karastergiou (*Preprint* http://-arxiv.org/abs/1103.1214)

[7] Fitzpatrick L 2011 *32m Dish To Get New Life As Deep Space Radio Telescope* http://www.-nationalspacecentre.eu/#/release-deep-space-radio-tel/3845815

[8] Pearson T J and Readhead A C S 1984 *Annu Rev Astron Astrophys* **22** 97

The neutron superfluid and a possible application to neutron star interiors

G J Kemp[1], C A Engelbrecht[1] and F A M Frescura[2]
[1]University of Johannesburg, [2] University of Witwatersrand
E-mail: garry@kemp.za.org

Abstract. Feynman, in his work on superfluid helium, attempted both qualitative and quantitative descriptions of superfluid helium's excitation spectrum. We briefly review Feynman's approach to superfluid helium and assess the applicability of this theory to neutron star interiors. We find that Feynman's calculations for superfluid helium may conceivably be applied to the neutron superfluid in a neutron star.

1. Introduction

London proposed that the onset of superfluidity in liquid helium was a result of Bose-Einstein condensation [1], [2]. Tisza then developed this idea into his so-called two-fluid model [2], [3]. Landau, however, critisised this view-point. He argued that atoms in the condensate would still collide with excited atoms generating friction amongst the atoms in the condensate, thus destroying the superfluidity [4]. Starting from quantum hydrodynamics, Landau proposed his own two-fluid model for liquid helium below the λ-point $T_\lambda \approx 2.19\,\mathrm{K}$ - the temperature at which the phase transition takes place. Thus, there were two opposing view-points regarding the quantum liquid - that of London and that of Landau.

An atomic theory of liquid helium below the λ-point was successfully tackled by Feynman [5], [6], [7] and [8]. His theory describing the transition as well as the excitation spectrum agreed both qualitatively and quantitatively with experiment. For example, his path-integral approach to liquid helium, modeled as a system of non-interacting bosons, predicts a transition in the system roughly where the observed transition occurs.

The exact nature of neutron star interiors is not yet well understood. It is believed that neutron star interiors contain neutron superfluid [9], [10]. Furthermore, it is believed that the superfluid plays an important role in phenomena observed in neutron stars [11]. We pursue a theoretical study of neutron superfluids with the aim of achieving a better understanding neutron star interiors.

In section 2, we review Feynman's path-integral treatment of liquid helium. We do not review his arguments for the transition being attributed to the statistics of the system however. In section 3, we review Feynman's approach to predicting the shape of the excitation spectrum as well as the energy gap Δ. We conclude with a short discussion of how this calculation may be reproduced for the neutron superfluid of a neutron star.

2. The transition

We begin by writing down the partition function for an interacting system of N bosons in the path-integral approach to quantum mechanics. Define imaginary time by $\tilde{t} = it/\hbar$, where t is real time. In this work, we refer to $\tilde{t}$ simply as time but it should be borne in mind that $\tilde{t}$ is not real time.

Denote the particle coordinates at time $\tilde{t}$ by $\mathbf{r}_i(\tilde{t})$ where i labels the particle and runs from 1 to N. Suppose that the initial and final configurations of the system at times $\tilde{t} = 0$ and $\tilde{t}$ respectively, are identical and given by $(\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_N)$, where $\mathbf{z}_i$ denote the coordinates of ith boson at these times: $\mathbf{r}_i(0) = \mathbf{z}_i$ and $\mathbf{r}_i(\tilde{t}) = \mathbf{z}_i$. The interaction forces are two-body forces where there is a weak attraction at the average atomic distance and a strong repulsion when the atoms approach closer than 2.6 Å. He^4 atoms obey Bose-Einstein statistics. Thus, we cannot distinguish between the configuration $\mathbf{r}_i(\tilde{t}) = \mathbf{z}_i$ and a configuration in which the atoms have been permuted amongst the positions. Denote by P a permutation of atoms amongst the $\mathbf{z}_i$ positions. Then, configurations $\mathbf{r}_i(\tilde{t}) = \mathbf{z}_i$ and $\mathbf{r}_i(\tilde{t}) = P\mathbf{z}_i$ are indistinguishable. Thus, the exact partition function for the system is

$$Q = \frac{1}{N!}\sum_P \int d^N\mathbf{z} \int_{tr_P} \mathcal{D}^N[\mathbf{r}(u)] \exp\left\{ -\int_0^{\tilde{t}} \left[\frac{m}{2\hbar^2} \sum_i \left(\frac{d\mathbf{r}_i}{du} \right)^2 + \sum_{ij} V(\mathbf{r}_i - \mathbf{r}_j) \right] du \right\} \quad (1)$$

where $\int_{tr_P}$ means that we integrate over all particle trajectories for a given permutation P and we sum over all possible permutations P. The atoms are then modeled as hard impenetrable spheres preventing the atoms from overlapping. We approximate the atoms as free particles each with an effective mass m'. We therefore make the following approximation for the integrand in the exponential

$$\frac{m}{2\hbar^2} \sum_i \left(\frac{d\mathbf{r}_i}{du} \right)^2 + \sum_{ij} V(\mathbf{r}_i - \mathbf{r}_j) \approx \frac{m'}{2\hbar^2} \sum_i \left(\frac{d\mathbf{r}_i}{du} \right)^2 . \quad (2)$$

The integrals over u and $\mathbf{r}(u)$ are evaluated to give

$$\int_{tr_P} \mathcal{D}^N[\mathbf{r}(u)] \exp\left\{ -\int_0^{\tilde{t}} \left[\frac{m'}{2\hbar^2} \sum_i \left(\frac{d\mathbf{r}_i}{du} \right)^2 \right] du \right\} = \left(\frac{m'}{2\pi\tilde{t}\hbar^2} \right)^{3N/2} \exp\left[-\frac{m'}{2\tilde{t}\hbar^2} \sum_i (\mathbf{z}_i - P\mathbf{z}_i)^2 \right] . \quad (3)$$

The partition function in this model therefore is approximated by [5]:

$$Q = \frac{K_\beta}{N!} \sum_P \int \left(\frac{m'}{2\pi\tilde{t}\hbar^2} \right)^{3N/2} \exp\left[-\frac{m'}{2\tilde{t}\hbar^2} \sum_i (\mathbf{z}_i - P\mathbf{z}_i)^2 \right] \rho(\mathbf{z}_1, ..., \mathbf{z}_N)\, d^N\mathbf{z} \quad (4)$$

where $\rho(\mathbf{z}_1, ..., \mathbf{z}_N)$ is a weighting factor for the configuration $(\mathbf{z}_1, ..., \mathbf{z}_N)$, K_β ($\beta = 1/k_BT$ with k_B being Boltzmann's constant and T, temperature) is some function of β. Although K_β is discussed in detail in [5] it is unimportant for this discussion. As a weighting factor, ρ should be nearly zero if atoms overlap and larger if the atoms are well spaced from each other. Thus, for low temperatures, i.e., large β, ρ is interpreted as the probability density of the ground state wavefunction. Approximately

$$\rho(\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_N) = \begin{cases} 1 & |\mathbf{z}_i - \mathbf{z}_j| > b \\ 0 & |\mathbf{z}_i - \mathbf{z}_j| < b, \end{cases} \quad (5)$$

where b is the diameter of impenetrable atoms of a classical gas.

It may be shown that the partition function in (4) approaches the correct limit for large and small values of β [5].
Feynman then shows that (4) describing a non-interacting system of helium atoms of mass m' obeying Bose-Einstein statistics is capable of predicting the observed transition from normal fluid to superfluid in liquid helium. Thus, Feynman, from first principles, showed that London's view-point is essentially correct; the observed transition at 2.19 K is indeed a result of the Bose statistics of the system. From the approximations made, however, the partition function (4) is not able to predict the order of the transition nor the temperature dependence of the specific heat of He I as the transition is approached.

3. The excitation spectrum

3.1. The lowest excitations

The lowest excited states of liquid helium involve large groups of atoms or long wavelengths, i.e., compression waves [6]. Suppose that there is a small spike in the density of the fluid in some large region along with a rarefaction adjacent to this region. For equilibrium to establish itself, atoms must move from the high density region to the low density region until the densities are equal. Establishing this equilibrium involves the motion of a relatively large mass of fluid which can then have low kinetic energy. If c is the speed of sound in the liquid, k a wavenumber of these waves and $\omega = ck$ the frequency, then the excitation energy $\hbar ck$ can be arbitrarily small. The first excited states above the ground state are therefore phonons. This conclusion is not exclusively applicable to liquid helium and may apply equally well to the lowest excited states of a neutron superfluid.

3.2. The minimum in the spectrum

For atoms obeying Bose statistics, motions on an atomic scale require a minimum energy of excitation above the ground state [7]. The nature of these excited states are speculated to be a small rotating ring of atoms or the rapid motion of an atom through the liquid. The latter excitation involves other atoms in front moving out of its way and closing in behind it. The last possibility is the "excitation of an atom in the local cage formed around it by its neighbors" [7]. For atoms obeying Bose statistics, the wavefunction describing each of the suggested excitations is of the form [7]:

$$\psi = \sum_i f(\mathbf{r}_i)\phi \tag{6}$$

where $f(\mathbf{r})$ is some function of the position, ϕ is the ground state wave function of the system and the sum is taken over all particles in the fluid. The form of function $f(\mathbf{r})$ is found through the variational principle. The Hamiltonian of the system is taken as

$$\hat{H} = -\frac{\hbar^2}{2m}\sum_i \nabla_i^2 + V - E_0 \tag{7}$$

where V is the potential energy and E_0 is the ground state energy. E_0 is subtracted because we are only interested in energies above the ground state. In the variational principle, we select the function that minimizes the quantity

$$E = \frac{\int \psi^* \hat{H}\psi \, d^N\mathbf{r}}{\int \psi^*\psi \, d^N\mathbf{r}}. \tag{8}$$

The solution is found to be

$$f(\mathbf{r}) = e^{i\boldsymbol{k}\cdot\mathbf{r}}, \tag{9}$$

with energy

$$E(k) = \frac{\hbar^2 k^2}{2mS(k)} \tag{10}$$

where $S(k)$, known as the *structure factor*, is the Fourier transform of the liquid's correlation function:

$$S(\mathbf{k}) = \int p(\mathbf{r}) e^{i\boldsymbol{k}\cdot\mathbf{r}}\, d^3\mathbf{r}. \tag{11}$$

The quantity $p(\mathbf{r})d^3\mathbf{r}$ is the probability of finding a helium atom within a volume $d^3\mathbf{r}$ containing the point $\mathbf{r}$ if another atom is known to be at $\mathbf{r} = 0$.

The structure factor may be determined from neutron scattering experiments. Starting from $k = 0$, $S(k)$ rises linearly as $\hbar k/2mc$ implying that $E(k)$ rises as $\hbar kc$. Thus, with the wavefunction (6), Feynman is able to reproduce exactly the phonon part of the spectrum. $S(k)$ then reaches a maximum near $k_0 = 2\,\text{Å}^{-1}$ thus producing a minimum in $E(k)$ for this k value. Beyond this value, $S(k)$ falls to approach unity while undergoing small oscillations. Thus, eventually $E(k)$ will increase according to $\hbar^2 k^2/2m$.

The form of the wavefunction (6) was determined purely on the basis of the particles obeying Bose-Einstein statistics. Thus, its form is not specific to liquid helium and would apply equally well to any other liquid whose atoms obeyed the same statistics. The only quantity thus far that is liquid helium specific is the structure factor $S(k)$. For liquid helium, the presence of the minimum near k_0, predicted by Feynman's calculation, agrees with experiment and with Landau's theory of liquid helium. In this region, we may therefore write $E(k)$ in the form

$$E(k) = \Delta + \frac{\hbar^2}{2\mu}(k - k_0)^2, \tag{12}$$

where μ is a constant determining the curvature of the spectrum and Δ is the energy gap - the minimum energy needed for the excitation. The excitations in this region of the spectrum correspond to Landau's rotons. Qualitatively, (10) thus displays the same behaviour as Landau's curve (see [12]) and is exact for the phonon part of the spectrum. However, while Feynman's curve exhibits a minimum where Landau's does (around k_0), his Δ value is almost double that of Landau's. Landau obtains

$$\frac{\Delta}{k_B} = 8.65 \pm 0.04\,\text{K} \tag{13}$$

(see [13]) while this fraction computed from (10) is 19.1 K [8]. It is suggested that the error in Δ is due to the inaccuracy of the trial wave function ψ, even though it was argued to be a relatively good first approximation. One could lower the value of Δ, for instance, by using a new trial wave function describing the physics of the system more accurately.

The new proposed trial wave function is [8]:

$$\psi = \phi \sum_i \exp(i\boldsymbol{k}\cdot\mathbf{r}_i) \exp\left[i \sum_{j\neq i} g(\mathbf{r}_{ji})\right]. \tag{14}$$

This ψ differs from that in (6) by the exponential factor containing the function $g(\mathbf{r})$. It is argued in [8] that the energy in the variational principle (10) would be lowered if the atoms executed some pattern of flow around the excitation as it flowed through the liquid. The pattern discussed is similar to the "backflow" discussed earlier. The additional exponential factor $\exp[i\sum_{j\neq i} g(\mathbf{r}_{ji})]$

is meant to incorporate a "backflow" as the excitation moves through the liquid. The function g may be determined by solving the differential equation that results from the variation of (8) after the substitution of ψ.

Let $E_1(k)$ and $E_2(k)$ represent the energy spectra found by substituting the old and new wave functions, (6) and (14) respectively, into (10). Values of the ratio $E_2(k)/E_1(k)$ for a range of k values are tabulated in Table II in [8]. There also is a column comparing $E_2(k)/k_B$ with k and it is estimated in [8] that $k = 1.85\,\text{Å}^{-1}$ corresponds to a minimum of 12.00 K - a dramatic improvement over 19, 1 K. If additional minor approximations are made in minimizing E in (10), the energy $E_2(k)/k_B$ at $k = 1.85\,\text{Å}^{-1}$ is lowered further to 11.5 K.

Recall that for small k, the spectrum $E_1(k)$ resulting from (6) was exact. Thus, there is nothing for the new wave function to improve upon in this region. If this new theory is to be correct, $E_2(k)$ must reduce to $E_1(k)$ as k becomes small and reaches the phonon part of the spectrum. It is verified in [8] that this is indeed the case.

The actual calculation of (10) with (14) is long and difficult. In the interest of simplifying the calculation, (14) was approximated by

$$\psi \approx \phi \sum_i \exp(i\boldsymbol{k} \cdot \mathbf{r}_i)\left[1 + \sum_{j \neq i} g(\mathbf{r}_{ji})\right]. \tag{15}$$

This makes the integrals resulting from the variational principle tractable. The quantities used in this calculation which depend on the liquid helium system are the structure factor, the correlation functions and the probability density functions from one, $\rho(\mathbf{r}_1)$, up to four particles $\rho(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4)$. With the approximations made during the calculations, $\Delta/k_B = 11.5\,\text{K}$ is considered in [8] to be accurate to within 0.6%.

4. Discussion

The form of the wavefunctions proposed could apply to any system of bosons in Feynman's model, not only liquid helium. Consider a superfluid composed of paired neutrons in the $^1S_0$ channel. Here, the paired neutrons are the bosons. The same initial wavefunction (6) may be proposed to describe an excitation within the fluid. The mass in the spectrum (10) would be modified to $2m_n$ where m_n is the neutron mass. $S(k)$ in (10) would also be modified to that of the neutron superfluid. The correlation functions, contained in the structure factors, and density functions were the only liquid helium specific properties used in these calculations - the main requirement was that the system obeyed Bose-Einstein statistics. It is therefore conceivable that the calculations presented in [7] and [8] may be reproduced once the corresponding properties for the neutron superfluid are known. We plan to investigate this further in future work.

References

[1] London F 1938 *Nature* **141,** 643–4
[2] London F 1964 *Superfluids, Macroscopic Theory of Superfluid Helium* (*Graduate student series in physics* vol 2) (Dover Publications) 58-64
[3] Tisza L 1938 *Nature* **141,** 913
[4] Landau L 1941 *J. Phys. U.S.S.R.* **5,** 71
[5] Feynman R P 1953 *Phys. Rev.* **91,** 1291–1301
[6] Feynman R P 1953 *Phys. Rev.* **91,** 1301–8
[7] Feynman R P 1954 *Phys. Rev.* **94,** 262–77
[8] Feynman R P and Cohen M 1956 *Phys. Rev.* **102,** 1189–1204
[9] Link B K and Epstein R I 1991 *aj,* Part 1 **373,** 592–603
[10] Shternin P S and Yakovlev D G 2011 *Mon. Not. Roy. Astron. Soc* **412,** 108–12
[11] Anderson P W and Itoh N 1975 *Nature* **256,** 25–7
[12] Landau L 1947 *J. Phys. U.S.S.R.* **11,** 91
[13] Pathria R K 1996 *Statistical Mechanics* 2nd ed (Elsevier Butterworth-Heinemann) 168-74

Daytime modelling of VLF radio waves over land and sea, comparison with data from DEMETER Satellite

S. G. Meyer[1,2], A. B. Collier[1,2], C. J. Rodger[3]

[1] SANSA Space Science, Hermanus, South Africa

[2] School of Physics, University of KwaZulu-Natal, Durban, South Africa

[3] Department of Physics, University of Otago, Dunedin, New Zealand

stphngmeyer@gmail.com, collierab@gmail.com, crodger@physics.otago.ac.nz

Abstract. Very Low Frequency (VLF) radio waves propagate with little attenuation within the Earth-ionosphere waveguide. Perturbations of the lower ionosphere produce a modification of the geometry of the waveguide, resulting in a disruption of the VLF propagation conditions. A model based on the Wait's mode theory is developed to investigate temporal and spatial changes in ionospheric conditions. As VLF waves propagating from a transmitter reflect off the lower ionosphere, a portion of the energy leaks up into space leaving a 'fingerprint' of the modal structure of the fields at the reflection height. Simulations are compared to averaged data taken over a year from the DEMETER satellite over the NWC transmitter in North-West Australia to test the validity of the model.

1. Introduction

VLF radio waves can be used as an extremely useful probe of the lower ionosphere since it is at too low an altitude for direct satellite observation and too high for balloons. Rocket borne instruments can provide excellent information, but this is a once off, localised and expensive method of gathering information. Around the world there exist a number of narrowband VLF transmitters which serve as excellent scientific tools to study radio wave propagation. As the radio waves travel vast distances while suffering little attenuation, the signal observed at a receiver allows investigation of the properties of the waveguide along the propagation path.

For VLF waves, the D- and E-regions of the ionosphere are of interest. The D-region exists only during the day at an altitude of around 70km and arises mostly due to Lyman-α radiation from the Sun ionising NO. The E-region is the lowest region during the night and starts around 85-90km [1,2]. Electromagnetic wave propagation in a waveguide can be explained in terms of waveguide modes that occur when reflections of the waves off the boundaries of the waveguide interfere with each other. The solar influence on the ionosphere has a diurnal and seasonal dependence resulting in the waveguide height, and modal interference pattern, also having similar periodic variance.

2. Method

The approach used in these simulations is based on Wait's mode theory description of waveguide propagation and uses Wait's ionospheric parameters, height (H' in km) and sharpness (β in km^{-1}) [3,4]. The model makes certain assumptions and approximations such as flat Earth, constant propagation conditions along path and treats transmitters as vertical electric dipoles. Input parameters include ionospheric (H', β), ground (surface conductivity and dielectric constant, σ_g and ε_g), ambient magnetic field strength and direction ($\mathbf{B_0}$), electron-neutral collision frequency (ν) and wave frequency (f), bearing and maximum number of modes (n_{max}). The electron density profile (N_e) depends only on height, and is calculated using Wait's parameters as,

$$N_e(z) = 1.43\times10^{13} exp[-0.15H' + (\beta - 0.15)(z - H')]m^{-3}. \quad (1)$$

The other important height dependent quantity is the electron neutral collision frequency, which takes the form:

$$\nu(z) = 1.816\times10^{11} exp[-0.15z]s^{-1}. \quad (2)$$

The electron density is used to determine the plasma frequency (ω_p) while the ambient magnetic field is used to calculate the electron gyro-frequency (ω_B). These two frequencies, along with the electron-neutral collision frequency represent the three fundamental parameters used in the Appleton-Hartree equation for a cold magnetised plasma. These parameters are used to find the height at which the VLF waves are reflected as well as the refractive index for the frequency in question at the reflection height.

At high altitudes (> 100 km), the effect from electron-neutral collisions is stronger than the Debye shielding effect produced by the electron density. As altitude decreases, the electron-neutral collisions become less dominant and the plasma frequency increases. At a certain height the Debye shielding caused by the applied electric field on the plasma becomes strong enough that the electron-neutral collisions no longer suppress the shielding effect which now blocks the waves from propagating further. This height is then taken as the reflection height, or upper boundary of the waveguide. Initially assuming that the ionosphere has a reflection coefficient of -1, the first n_{max} eigenangles are calculated. These angles along with the ambient magnetic field direction and bearing of the wave are then used to calculate the angle between the magnetic field and each mode of propagation which is needed to find the complex refractive index that the wave will experience when reflecting off the ionosphere. Recalculation of the eigenangles is then done using the complex refractive index and the initial angles in a perturbative method. The eigenangles essentially contain all necessary information to calculate the electric fields at any distance from the transmitter. An additional height variability function, f_n, allows for the calculation of the fields at any height in the waveguide. The electric fields can then be calculated at any point along a path at any height in the waveguide with the use of a zero-order Hankel function of 2nd kind and the height function [3].

$$E_z(\rho, z) = \frac{\mu\omega Ids}{2\mathrm{h}} \sum_{n=1}^{n=n_{max}} sin^2\theta_n H_0^{(2)}(ksin\theta_n\rho) f_n(z) \quad (3)$$

3. Data and Results

Even though the bulk of the VLF energy remains in the waveguide, a portion of it does leak up into space and should allow satellites with radio receivers to observe the modal interference pattern around a transmitter. To test the validity of this model, simulation results are compared to satellite data from the DEMETER satellite which is in a sun-synchronous orbit and always passes over a given location at a fixed local time. Figure 1 shows the observed modal interference pattern observed while passing over the NWC transmitter in Australia which operates at 19.8 kHz.

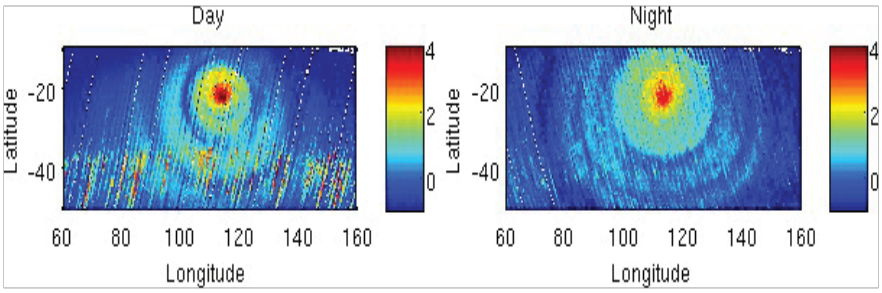


Figure 1: Averaged data over the year 2005 for day (10:30 LT) and night (22:30 LT). Each row and column correspond to 0.5° latitude and longitude respectively.

For convenience sake, to compare data to simulation results, horizontal sections three rows wide were extracted from the data ranging just over 5000km long. Three rows were taken to minimise the effects of missing data points and to create a smoother curve. One advantage of using the NWC transmitter is that it is on the North-West coast of Australia meaning that westward propagation is sea only and eastward is almost entirely land, making it appropriate for comparison with this model which does not take mode conversion into account at the interface of different waveguide slabs.

The height of the ionosphere has a seasonal dependence and changes throughout the year. This means that the observed data cannot be explained very well in terms of one waveguide height, but rather the average of a few heights. Figure 2 shows the variation between night and day as well as eastward versus westward propagation. Eastward propagation tends to show lower field strengths than westward which is expected as this is over land which is a far poorer conductor than sea water. The modal pattern is quite clear for the daytime propagation whereas the night time data does not show such a clear pattern. This can be attributed to the higher waveguide height during darkness that leads to a modal interference pattern which is less well defined. Due to the lower attenuation at night, there will be more of a contribution from higher order modes that would also reduce the clarity of the modal structure. For this reason, only comparison with daytime data will be made.

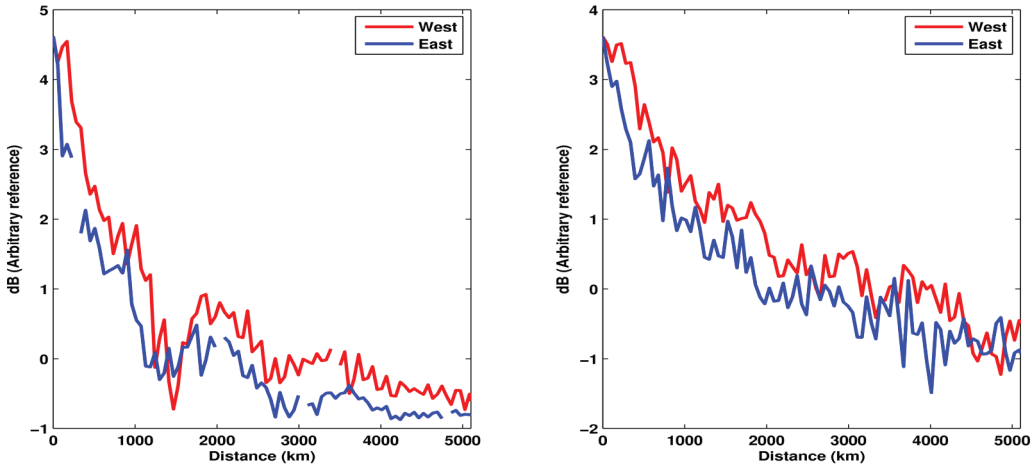


Figure 2: Observed field strength for west- and eastward propagation from NWC transmitter.

To find the best comparison between data and simulation, a minimisation of the error between the two was performed. Bearing in mind the model is working on certain assumptions, the minimisation was performed within a certain region that excluded the area close (< 1017 km) to the

transmitter and the area past a certain point (> 4068 km). The first region was excluded from the comparison since in this region the ground wave and evanescent modes can contribute to the field which are not included in the simulation. As for the far field region, this was excluded due to the flat Earth approximation which is good at short and intermediate distances but can become inaccurate at distances that are close to the radius of the Earth.

The minimisation process was done using the MATLAB optimisation toolbox and by searching for different ionospheric heights and a global adjustment or gain that was added to appropriately compare data and simulation results. Various values of β were tested manually, the value that gave the best results with a sets of 3, 4, 5 and 6 heights was selected and had a value of 0.35 km^{-1}, falling between normal day time summer and winter values. Up to 9 heights were used but it was seen that using 4 or 5 heights was sufficient if looking at propagation in one direction only and adding more values only increased the time taken to run the minimisation process with no improvement in error. When looking to minimise the average error for the two directions, it was found that more heights were needed to deliver satisfactory results. Figure 3 shows the results when a single set of values was used for both directions and the average error in the region of interest was minimised.

It should be noted that another factor that can play a significant role is the conductivity of the surface of the lower boundary of the waveguide. For westward propagation over water, the conductivity was taken to be 4 S/m [5], and the ground was given the value of 4.5×10^{-3} S/m for the eastward propagation. Ground conductivities of dry land are generally between 10^{-2} S/m and 10^{-4} S/m and have been reported to be around 10^{-3} S/m in this area [6]. With the use of this value in the simulation, the results did not fit the data well so it was adjusted to the slightly higher value of 4.5×10^{-3} S/m.

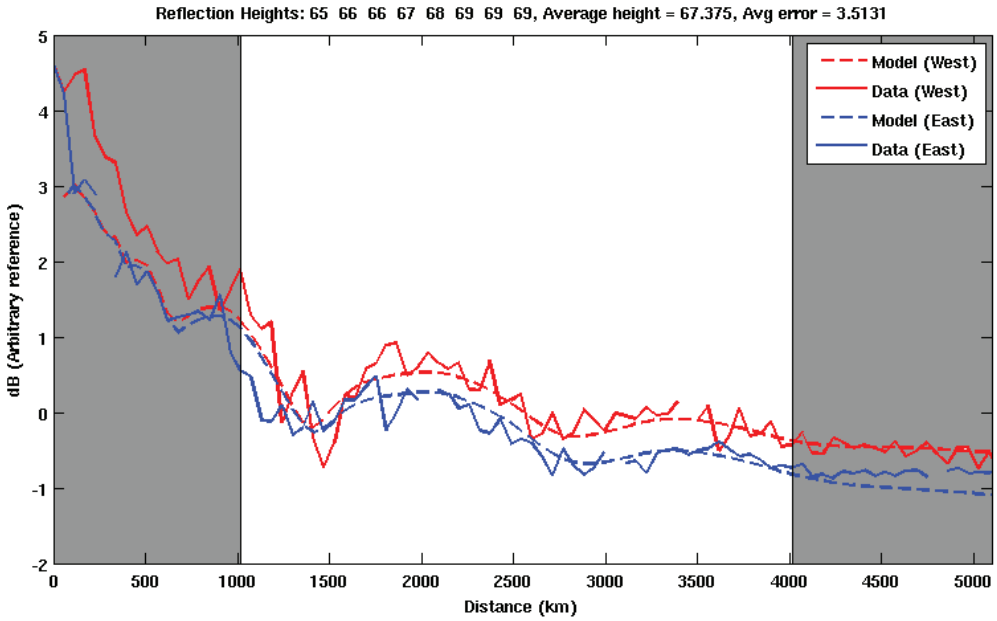


Figure 3: Comparison of simulated results with data. One set of heights used for both directions.

4. Conclusion and Discussion

The average height of the reflection layer around NWC is found to be around 68 km with heights ranging from 65 km to 69 km for summer and winter respectively. These are reasonable results but are somewhat lower than the results published in a study done in 2009 that reported heights of 70 km and 72 km around NWC [5]. This disagreement could be due to the fact that the method of calculating the reflection height and the use of Wait's parameters differs to the methods used here. There is definite space for future improvement to the model which could further increase its validity.

References

1. N.R. Thomson and W.M. McRae (2009), Nighttime ionospheric D region: Equatorial and nonequatorial, *J. Geophys Res.,* Vol.114*,* A08305, doi:10.1029/2008JA014001.

2. M.A. Clilverd, N.R. Thomson and C.J. Rodger (1999), Sunrise effects on VLF signals propagating over a long north-south path, *Radio Science,* Vol.34, pages 939-948, doi:10.1029/1999RS900052.

3. J.R. Wait, *Electromagnetic Waves in Stratified media,* Pergamon Press, 1962.

4. J.R. Wait and K.P. Spies, Characteristics of the Earth-ionosphere waveguide for VLF radio waves, *National Bureau of Standards, Technical Note 300*, 30 December 1964.

5. N.R. Thomson (2010), Daytime tropical D region parameters from short path VLF phase and amplitude, *J. Geophys Res.,* Vol.115*,* A09313, doi:10.1029/2010JA015355.

6. N.R. Thomson (1993), Experimental daytime VLF ionospheric parameters, *J. Atmos and Terr Phys,*Vol.55, No 2, pages 173 – 184, doi:10.1016/0021-9169(93)90122-F.

The effect of rotation on the shape of stars

P R Nicol[1], F A M Frescura[2] and C A Engelbrecht[1]

[1]Department of Physics, University of Johannesburg, P O Box 524, AUCKLAND PARK 2006, South Africa.

[2]School of Physics, University of the Witwatersrand, Private Bag 3, WITS 2050, South Africa.

E-mail: Nicol3.Peter@gmail.com

Abstract. We adapt a recent work by Zahn *et al.* in which they explore the degree of oblateness of rapidly rotating stars. The model they develop incorporates the effect of the quadrupolar moment of the gravitational potential produced by the deformation of the star. We apply their methods to polytropes. We consider a range of polytropic indices and investigate the effect of differential rotation on the shape of the star for a variety of rotational profiles, including the case of uniform rotation. We find that the inclusion of the quadrupolar moment allows for deformations larger than those predicted by the classical Roche model.

1. Introduction

The shape assumed by a rapidly rotating star has been the subject of investigation for many years. Recent advances in long baseline interferometry now make it possible to measure the degree of flattening of some nearby stars. Domiciano de Souza *et al.* [1] recently announced that the Be star Achernar is oblate in shape with a ratio of the major to minor axes of 1.56. This value was later revised to 1.41 after the discovery that Achernar emits a jet [2]. This is the highest measured value in any star. This unexpectedly high value raises challenging problems for current models. An early model that predicts the shape of a star rotating at critical velocity is due to Roche. In this model the entire mass of the star is treated as if it were concentrated at its centre. The maximum flattening ratio predicted by this model is 1.5. Recently, Zahn *et al.* [2] attempted to predict the measured value for the flattening of Achernar by modifying the Roche model. In their model the mass is no longer concentrated in the centre of the star but is now contained in an oblate spheroid. This adds a quadrupolar term to the Roche theory. They illustrated the impact of inclusion of the quadrupolar moment by considering a $7M_{\odot}$ star in several different stages of evolution and undergoing three types of rotation: uniform, differential and shellular. They were able to obtain flattening ratios of greater than 1.5.

In this paper we explore the effects of the inclusion of the quadrupolar moment in the calculation. We examine its effect in the simpler case of the polytrope. We used a polytrope model in the absence of a more realistic model. This study is therefore preparatory and will be modified once a full stellar model is available. We examine the flattening ratio for different polytropic indices and for angular velocities ranging from zero to critical. The critical angular velocity is defined to be that angular velocity at which the effective gravitational force acting on the material at the equator at the surface of the star is zero. In Section 2, we summarise the principal results needed from the theory of polytropes. In Section 3, we review the Roche theory for stellar shape, calculate the first order perturbation of the gravitational potential by a linear

perturbation method first described by Sweet [3], and solve the perturbed Poisson equation for the gravitational potential. This yields a value for the quadrupolar moment of the polytrope which we use to calculate the flattening ratios for a range of rotational profiles. The results are reported in Section 4.

2. Polytropes

A polytrope is defined [4] to be a star in which the pressure P is proportional to a power of the density ρ,

$$P = K\rho^{\alpha}$$

where K is a constant, and α is called the polytropic power. It can be shown that this condition is equivalent to the assumption that the stellar material undergoes a process in which its specific heat has constant value [4]. The equation governing the structure of a non-rotating polytrope is the Lane-Emden equation, given by

$$\frac{1}{\xi^2}\frac{d}{d\xi}\left(\xi^2\frac{d\theta}{d\xi}\right) = -\theta^n \tag{1}$$

where θ and ξ are dimensionless variables related to the density and the radius respectively, and $\alpha = 1 + 1/n$. n is called the polytropic index. The boundary conditions for equation (1) are $\theta = 1$ and $d\theta/d\xi = 0$ at $\xi = 0$. The radius of the star is given by

$$R = \left[\frac{(n+1)K}{4\pi G}\lambda^{(1/n)-1}\right]^{1/2}\xi_1$$

where ξ_1 is the first zero of the solution to equation (1), and λ is the central density of the polytrope. The mass M_* of the star is given by

$$M_* = -4\pi\left[\frac{(n+1)K}{4\pi G}\lambda^{(1/n)-1}\right]^{3/2}\xi^2\,\frac{d\theta}{d\xi}$$

evaluated at $\xi = \xi_1$.

3. Quadrupolar moment

3.1. Flattening of the star

Consider a non-rotating star of mass M and radius R. The stellar surface can be taken to coincide with a level surface of suitably chosen low pressure [2]. Since a polytrope is a barotropic system, isobaric surfaces coincide with isopycnic surfaces, and also with level surfaces of the effective gravitational potential. Suppose now that this star rotates with constant angular velocity Ω. Choose the origin at the centre of the star, and the z-axis in the direction of the axis of rotation. The effective potential, as a function of the spherical coordinates r, θ to first approximation is given by

$$\Phi_{\text{eff}} = -\frac{GM(r)}{r} - \frac{1}{2}\Omega^2 r^2\sin^2\theta$$

where $M(r)$ is the mass contained in a sphere of radius r in the non-rotating case. To determine the value Φ_{eff} on the surface equipotential, note that at the pole the centrifugal force is zero and $r = R_P$, so that at the surface $\Phi_{\text{eff}} = -GM/R_P$. Therefore, at the stellar equator, we have

$$\frac{GM}{R_E} + \frac{1}{2}\Omega^2 R_E^2 = \frac{GM}{R_P}$$

or

$$1 + \frac{1}{2}\frac{\Omega^2 R_E^3}{GM} = \frac{R_E}{R_P} \tag{2}$$

Equation (2) can be used to calculate the ratio of equatorial to polar radii for given angular velocity. The critical angular velocity Ω_{crit} is defined to be that angular velocity at which the star must rotate in order for the effective gravity at a point on its surface at its equator to be zero. This gives,

$$\Omega_{\text{crit}}^2 = \frac{GM}{R_E^3}$$

3.2. The Quadrupolar correction

In the Roche model, it is assumed that the gravitational potential of the star is the same as that which would be set up were its entire mass concentrated at its centre. This is a good first approximation. However, rotation causes redistribution of the stellar mass making the star oblate and causing the mass distribution to deviate from spherical symmetry. The gravitational potential outside the star may, therefore, be expanded as a multipole series of the form

$$\phi(r,\theta) = -\frac{GM}{r}\left[1 - \sum_{l=2}^{\infty}\left(\frac{R_0}{r}\right)^l J_l P_l(\cos\theta)\right] \tag{3}$$

where R_0 is the radius of the spherically symmetric reference model, J_l is a dimensionless constant that measures the degree of oblateness, $P_l(\cos\theta)$ is the Legendre polynomial of degree l. Assume symmetry about the equatorial plane. Then only the even Legendre polynomials are required in the expansion.

3.3. First Order Perturbation of the Gravitational Potential

Following Zahn [2], we use a generalised first order perturbation method first described by Sweet [3] to derive an expression for the distortion of the equipotential surfaces by rotation. The equations of hydrostatic equilibrium in the rotating frame, assuming azimuthal symmetry, are given by

$$\frac{\partial P'}{\partial r} = \rho'\frac{\partial \phi'}{\partial r} + \rho' f_r \tag{4}$$

$$\frac{\partial P'}{\partial \theta} = \rho'\frac{\partial \phi'}{\partial \theta} + \rho' f_\theta \tag{5}$$

where $\rho' f_r$ and $\rho' f_\theta$ are the components of the centrifugal force, ϕ' is Sweet's gravitational potential (which is the negative of the potential ordinarily used), P' is the hydrostatic pressure, and ρ' is the density of the material. Primes are used here to denote values in the perturbed star, while unprimed quantities denote values in the unperturbed star: that is

$$\begin{aligned} P' &= P + \delta P \\ \rho' &= \rho + \delta\rho \\ \phi' &= \phi + \delta\phi \end{aligned}$$

Differentiating (4) with respect to θ and (5) with respect to r, one can eliminate P' from equations (4) and (5) to obtain, correct to the first order,

$$\frac{\partial \rho'}{\partial \theta} = -\frac{\chi}{g}\frac{d\rho}{dr} + \frac{1}{g}\left(\rho\frac{\partial f_r}{\partial \theta} - \frac{\partial}{\partial r}(r\rho f_\theta)\right) \tag{6}$$

where g is the local gravitational acceleration and $\chi = \partial\phi'/\partial\theta$. Differentiating Poisson's equation

$$\nabla^2\phi = -4\pi G\rho$$

with respect to θ (note the difference of convention in [3] for the gravitational potential), we eliminate $\partial\rho'/\partial\theta$ from equation (6) to obtain

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\chi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial}{\partial\theta}\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left[\chi\sin\theta\right]\right) = \frac{4\pi G}{g}\frac{d\rho}{dr}\chi + \frac{4\pi G}{g}\left[\frac{\partial}{\partial r}\left(r\rho f_\theta - \rho\frac{\partial f_r}{\partial\theta}\right)\right]. \quad (7)$$

The form of the left hand side of equation (7) suggests that we expand χ in terms of spherical functions. Expanding χ and the centrifugal acceleration in terms of Legendre polynomials one obtains

$$\begin{aligned} f_r &= \Omega^2 r\sin^2\theta = \sum_{l=0}^{\infty} a_l(r)P_l(\cos\theta) \\ f_\theta &= \Omega^2 r\sin\theta\cos\theta = -\sum_{l=0}^{\infty} b_l(r)\frac{d}{d\theta}P_l(\cos\theta) \\ \frac{d\phi}{d\theta} &= \sum_{l=0}^{\infty} c_l\frac{d}{d\theta}P_l(\cos\theta) \end{aligned}$$

Now

$$a_l(r) = \alpha_l r\Omega^2 \quad (8)$$

$$b_l(r) = \beta_l r\Omega^2 \quad (9)$$

where α_l and β_l are constants. Substituting equations (8) and (9) into equation (7) gives

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dc_l}{dr}\right) - \left[\frac{l(l+1)}{r^2} + \frac{4\pi G}{g}\frac{d\rho}{dr}\right]c_l = \frac{4\pi G}{g}\left[\frac{d}{dr}(r\rho b_l) + \rho a_l\right]. \quad (10)$$

Equation (10) is known as the perturbed Poisson equation. To ensure regularity of the functions c_l at at the origin and at the surface of the star, the boundary conditions required for $l \geq 1$ are [5],

$$\begin{aligned} c_l(0) &= 0 \\ (l+1)c_l(R) &= -\frac{dc_l}{dr}(R) \end{aligned}$$

3.4. Re-scaled Perturbed Poisson Equation

To simplify integration of equation (10), we rescale it. Define $x = r/R_0$, $h(x) = \Omega^2(x)/\Omega_s^2$ and $\phi_l = c_l/\Omega_s^2R_0^2$, where Ω_s is the surface value of the angular velocity and R_0 is the radius of the spherically symmetric non-rotating reference star. Then, using equations (8) and (9), equation (10) becomes

$$\frac{1}{x^2}\frac{d}{dx}\left(x^2\frac{d\phi_l}{dx}\right) - \left[\frac{l(l+1)}{x^2} + \frac{4\pi GR_0}{g}\frac{d\rho}{dx}\right]\phi_l = \frac{4\pi GR_0}{g}\left[\frac{d}{dx}(x^2\rho\beta_l h(x)) + \rho\alpha_l h(x)\right] \quad (11)$$

The domain of integration is $x = [0, 1]$. From equation (3), for even values of l, one obtains

$$c_l = \frac{GM}{r}J_l$$

and since $c_l = \phi_l\Omega_s^2R_0^2$, at the surface we get

$$J_l = \left(\frac{\Omega_s^2R_0^3}{GM}\right)\phi_l(1). \quad (12)$$

3.5. Quadrupolar Moment

We are interested only in J_2, the first correction to the gravitational potential. For $l = 2$, we get $\alpha_2 = -2/3$ and $\beta_2 = 1/3$. Equation (11) thus becomes

$$\frac{1}{x^2}\frac{d}{dx}\left(x^2\frac{d\phi_2}{dx}\right) - \left[\frac{6}{x^2} + \frac{4\pi G R_0}{g}\frac{d\rho}{dx}\right]\phi_2 = \frac{4\pi G R_0}{3g}x^2\frac{d}{dx}(\rho\ h(x)) \tag{13}$$

and equation (12) yields

$$J_2 = \left(\frac{\Omega_s^2 R_0^3}{GM}\right)\phi_2(1) \tag{14}$$

3.6. Quadrupolar Correction to the Flattening of the Star

Inserting the quadrupolar correction into equation (2), and substituting expression (14) for J_2, we obtain

$$1 + \frac{1}{2}(1 + \phi_2(1))\frac{\Omega^2 R_E^3}{GM} = \frac{R_E}{R_P}$$

4. Discussion and Results

We solved equation (1) numerically for various values of n, to obtain ρ, $d\rho/dx$, M and R_0. For the central density, we used the value 1×10^{15} kg m^{-3}, and took the value of K to be 2.76×10^9 kg $^{-1/n}$ m$^{-(2n+3)/n}$ s^{-2}. We then used these quantities to solve numerically the perturbed Poisson equation (13) and thus calculate the quadrupolar moment of the associated perturbed gravitational potential. We used the rotation profile given by

$$h(x) = \frac{1 + a}{1 + ax^2}$$

The value $a = 0$, corresponds to solid body rotation, while non zero values yield differential rotation. We calculated the cases corresponding to values $a = 0, 1, 2, 3, 4$. Figure 1 shows a plot of the flattening ratio as a function of Ω/Ω_{crit} for a polytrope with index $n = 1$. The lowest curve on the figure corresponds to $a = 0$, the next one up corresponds to $a = 1$ and so on. This is the general shape of the curves for different differential rotation profiles. Figure 2 shows an enlargement of the region for the range $0.8 \leq \Omega/\Omega_{crit} \leq 1$, where the flattening ratio reaches its largest values. The panels in Figure 2 represent, respectively, polytropes of indices $n = 1, 2, 3, 4$. The polytropic index n need not be restricted to integer values, but should lie in the range $0 < n < 5$. Values $n \geq 5$ lead to solutions where the stellar material extends to infinity. As can be seen from the plots in Figure 2, the separation of the curves for increasing a decreases with increasing index n. Note also that greatest degree of flattening occurs for the lowest polytropic index. This can be attributed to the fact that the polytropes of lower index are less centrally condensed. Based on the results of our investigation, it is apparent that the classical flattening ratio of 1.5 derived by Roche can be exceeded by correctly including the quadrupolar moment developed by the star as a result of the deformation of its shape.

The investigations reported in this paper need to be repeated using more realistic stellar models. We also need to investigate the inclusion of higher order moments of the gravitational potential.

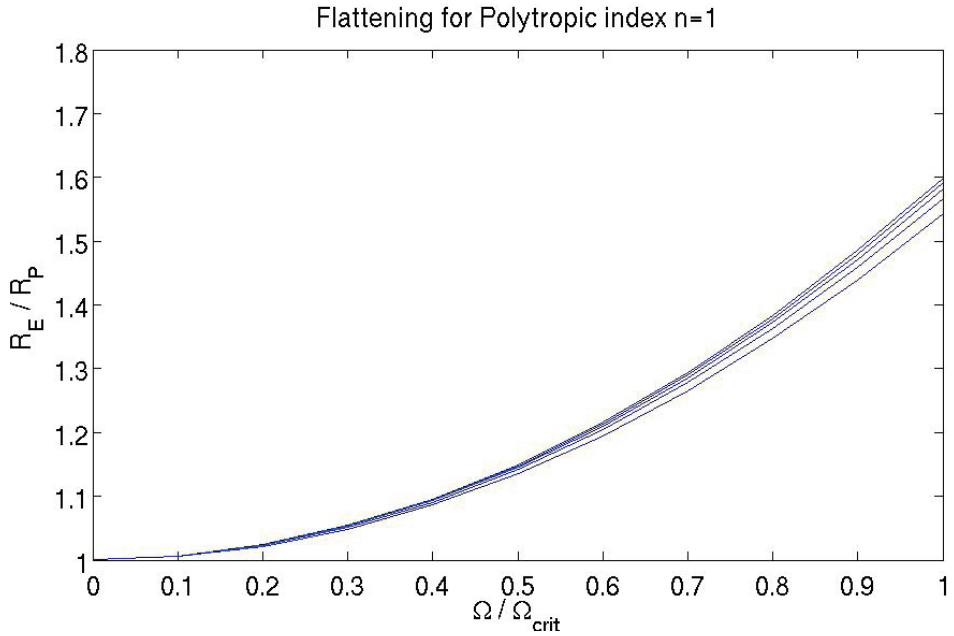


Figure 1. Flattening of polytrope with index n=1.

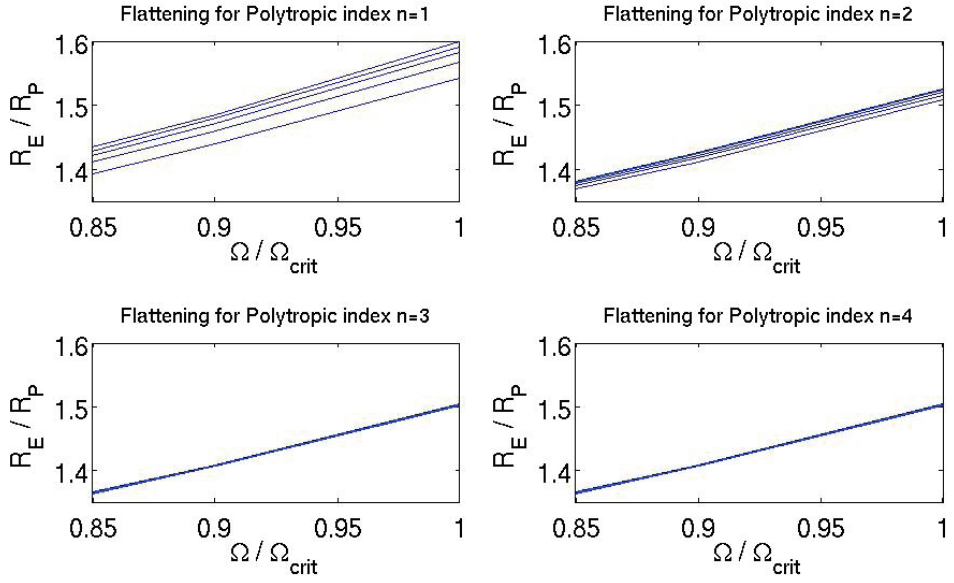


Figure 2. Flattening of polytropes with indices from n=1 to n=4, from left to right, top to bottom.

References

[1] Domiciano de Souza A, Kervella P, Jankov S, Abe L, Vakili F, di Folco E, Paresce F 2003 *Astron. Astrophys.* **407** 47
[2] Zahn J-P, Ranc C, and Morel P 2010 *Astron. Astrophys.* **517** A7
[3] Sweet P A 1950 *Mon. Not. R. Astron. Soc.* **110** 548
[4] Chandrasekhar S 1958 *An Introduction to the Study of Stellar Structure* (New York: Dover) p 87-88
[5] Mathis S, Zahn J-P 2004 *Astron. Astrophys.* **425** 229

Statistical analysis of outer electron radiation belt dropouts: geosynchronous and low earth orbit responses during stream interfaces

Olakunle Ogunjobi[1], Andrew. B. Collier[1,2] and Craig J. Rodger[3]

[1]School of Physics, University of KwaZulu-Natal, Private Bag X54001, Durban 4000, South Africa

[2]SANSA Space Science, P. O. Box 32, Hermanus 7200, South Africa

[3]Department of Physics, University of Otago, Dunedin 9054, New Zealand

E-mail: olakunle.ukzn@gmail.com

Abstract. The OMNI-2 data set enables a correlation study of solar wind and geomagnetic parameters, allowing the stream interfaces events (SIs) to be examined. A superposed epoch analysis of these events was performed to determine the threshold levels of IMF Bz and other geophysical parameters. Based on energy, temporal and spatial characteristics, statistical analysis of electron flux data from LANL-SOPA and NOAA- POES satellites were used to study outer zone electron dropouts and precipitation. The deepest minima of electron flux was observed after the impact of a SI, which coincided with the time of slower-decaying peaks of electron precipitation. Result obtained in this work suggest that the mechanism causing the precipitation could also be responsible for the observed electron flux dropout during Stream Interference.

1. Introduction

Earth′s magnetic field traps energetic particles in the radiation belts. The radiation belts consist of an inner belt (1.1<L<2.5), an outer belt (3<L<9), and an intermediate region with depleted energetic fluxes known as the slot-region (2.5<L<3.0). The outer radiation belt, is highly dynamic, with daily changes in electron flux. Without violation of any of the three adiabatic invariants, trapped radiation-belt particles would indefinitely undergo three types of periodic motion: gyration, mirroring between the northern and southern hemispheres, and drift motion. The loss processes in the outer radiation belt have become a topic of interest over the last decades; yet the principal mechanism depopulating this belt is unknown [1]. However, it has been shown recently that the outer radiation belt's energetic particle fluxes can undergo dramatic temporal variations as response to stream interface events (SIs). The bimodal structure of solar wind, that is, two basic populations: slow and fast provide the platform for SI occurrence. The slow solar wind comes from the vicinity of streamer belt and always at the leading edge of the fast solar wind which originates from coronal holes. The region that separates dense, slow-moving plasma from fast, less dense plasma is a stream interface. The subset of the 66 cleaned SI events identified by [2] are herein used in our statistical studies. One major characteristics of SI events on arrival at earth is the deflection of solar wind azimuthal velocity (GSE-Vy) as shown in figure 1 for the 66 events. In this paper, the focus is primarily on outer electron radiation belt dropouts

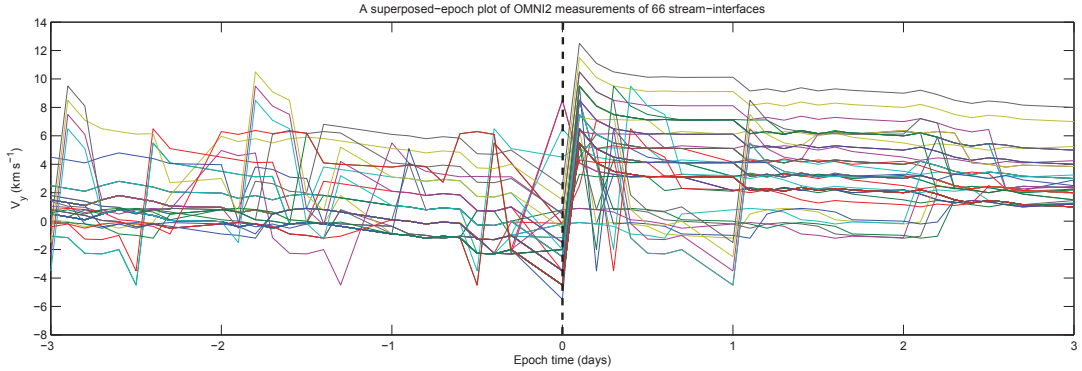


Figure 1. The Solar Wind Azimuthal Velocity (GSE-Vy) for each of the 66 SIs.

due to SI at Geosynchronous Earth Orbit (GEO) and Low Earth Orbit (LEO). In session 2, the data set use as well as the techniques employed in analysing the data is described. In session 3,the implication of the events are explored by analysing some geophysical parameters and solar wind conditions. Session 4 analysed the statistical responses to dropouts during SI events at GEO. In session 5, the statistical relationship between trapped electron flux levels and degree of precipitation at LEO are also analysed and lastly, we conclude of on the varying statistical analyses.

2. Data description and techniques

Using the events in [2], solar wind and geomagnetic parameters were obtained from Omni-2 data [3]. We cleaned the events by removing those with solar proton contamination before use. The energetic electron dropouts were measured using data from the multispacecraft Synchronous Orbit Particle Analyser (LANL/SOPA) located at GEO with nominal data coverage of 24-hours daily. At every 10.24 seconds, SOPA instruments measure differential fluxes of electrons from 50kev to greater than 1.5 MeV in 12 channels. The SOPA telescopes are actively controlled such that the spin axis of the satellite points continuously toward the centre of the Earth. The SOPA flux measurement used here were from three electron channels (225-315 keV, 315-500 keV and 0.75-1.1 MeV) from the LANL spacecraft. For every operational spacecraft, flux data were normalised to obtain the same yearly averaged logarithm [4]. Thus, we have obtained the logarithmic average by summation of log-fluxes divided by the number of satellites per time.

Additionally, NOAA/POES (15-18) serve as proxy for a wider range of L values and magnetic local time. The orbits of these satellites are polar and Sun-synchronous at an altitude of about 850 km and have a period of ∼100 minutes. The Medium Energy Proton and Electron Detector (MEPED) on each satellite provides directional (0° telescopes for precipitating and 90° for trapped) measurements of energetic electrons but with proton contamination [5]. In order to reduce contamination, we applied a first order correction, $\propto +15° < \propto_{LC}$ where $\propto$ is the local particle pitch angle at the satellite, $\propto_{LC}$ is the edge of the loss cone, while the arbitrary opening angle is 15. suggested by [6]. The L used is from the IGRF model for the event period. For clarity, we present a superposed average of > 30 keV electron flux at 0° (90°) that is, precipitating (trapped) channels across three spatial cuts (L= 4.02, 5.02 and 6.02) using POES 15 and 17 (2030-0630 MLT) and POES 16 and 18 (0730-1930 MLT). With POES satellites, our superposed epoch is organised across all MLT such that at any given MLT, one of the POES satellites provide coverage.

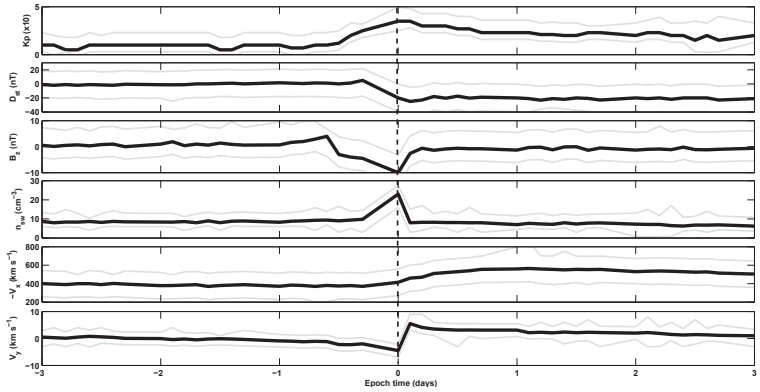


Figure 2. Plots of superposed properties of SIs used: in each panel the grey lines are the upper and lower percentiles while the black line indicates the average, with the zero epoch taken as the time of Vy reversal. From top (first-sixth) is Kp, Dst (nT), GSM Bz (nT), proton number density, solar wind radial velocity (-Vx GSE) and solar wind azimuthal velocity (GSE-Vy).

For all acquired data, the reference time of our superposed epoch analysis is set to SI onset which is defined as the reverse point from westward to eastward flow of solar wind azimuthal velocity. The data extracted at this reference time were 3-hourly averaged and for robustness, we also show the dispersion of geophysical parameters and equitorial electron flux observations by providing upper and lower percentiles at each time.

3. SIs: geophysical observations

Figure 2 presents various superposed geomagnetic and solar wind conditions around the time of SI used in our study. In Figure 2 (first panel), the superposed average of Kp as a measure of magnetospheric convection is seen to be at peak of a 3+ at approximately 3 hours after the arrival of SIs and returns to the pre event level in 0.5 days. Figure 2 (second panel), displays measurement of ring current, Dst (nT) shows a positive value of about 15 nT prior to the arrival of the SI indicating weakly magnetospheric convection pressure and reverses to storm of $\sim$ -30 nT at SI arrival implying small strengthen ring current. Dst perturbation continues through a day although at relatively low level as can be seen. Also in figure 2, third panel, is the superposed average of Bz- component (GSM coordinate) which prior to the arrival of SI was distinctly in the Northward peak of 10 nT and reversed southward close to the arrival of SI. The threshold value of GSM-Bz of less than -10 nT persists for about 3 hours at SI arrival. Figure 2 (fourth panel), shows the superposed plot of proton number density with sharp peak at the arrival of SI reaching approximately 19 cm^{-3} on averaged scale indicating difference of dynamic-pressure as fast wind overtakes slow wind. Prior to this abrupt increase, proton number density is found at relatively low level. The solar wind radial velocity (GSE -Vx) is presented in figure 2 (fifth panel) showing greater than 300 km/s before SI arrival. Just at SI onset, -Vx is elevated to a range of greater than 500 km/s for 3 days after SI implying that high-speed- stream driven storms last for several days. In Figure 2 (sixth) the GSE y-component (solar wind azimuthal velocity) is shown. Close to day 0, Vy is at the level of around -9 km/sec on average but reverses to a point of approximately +10 km/s on the average scale indicating arrival of SI at Earth. As can be seen, the Vy through 0.1 day maintains a steady flow at the passage of SI.

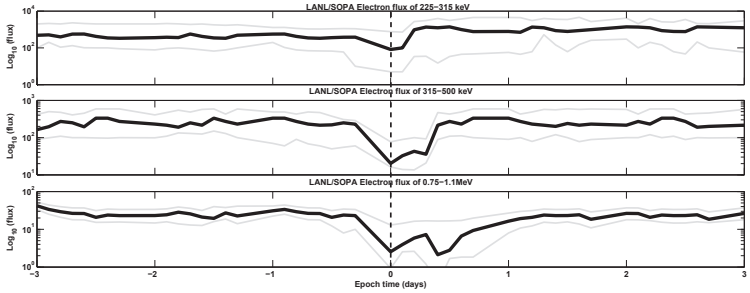


Figure 3. Superposed epoch plots of electron flux as measured by combined LANL SOPA and GOES satellites: (top) is 225 to 315 keV, (middle) is 315 to 500 keV, (bottom) is the 0.75 to 1.1 MeV and (fourth) is GOES > 2 MeV. The dash vertical line indicates arrival of SI. Format is as in Figure 2.

4. Equitorial observation of electron flux dropout

In this session, we present dropouts as measured from GEO. Observations at GEO can also be compared with Global Positioning Sysytem (GPS) taken at L* = 4 - 4.5 when considering adiabatic effects. It will be reasonable to say there is less mixing of adiabatic effects in our analyses since sets of SIs used here do not follow strong enhancement in ring current. Figure 2 presents an overview of superposed average of the logarithm of 225 - 315 keV, 315 - 500 keV and 0.75 - 1.1 MeV electron flux as measured by LANL/SOPA multiple satellites around GEO. In Figure 3 (top), electrons of 225 - 315 keV is plotted showing gradual dropout of electron flux from -1 (day) and drops to minimum at approximately 3 hours after the arrival of SI and recovers quickly to pre-event level in less than a day. In Figure 3 (middle) is the electron flux of 315 - 500 keV and as can be seen there is also gradual dropout prior the arrival of SI but electron flux dropout is about 0.78 magnitude deeper than in 225 - 315 keV channel at arrival of SI. The recovery is seen to be slightly higher than the pre-event level for more than three days. Figure 3 (bottom) displays electrons of 0.75 - 1.1 MeV, there is also observed gradual dropout from -0.5 day till 0.8 day after the arrival of SI. The dropout magnitude is seen to be about 1.55 and 0.11 order differ from 225 - 315 keV and 315-500 keV channel respectively while the recovery level is higher than pre-event level more than 3 days.

5. Estimate of > 30 keV electron flux from LEO

Figure 4 displays POES > 30 keV flux at L of 4.02, 5.02 and 6.02. In Figure 4 (top), electron flux at L= 4.02 is plotted, as can be seen there is a gap of more than factor of 5 difference between trapped and precipitating flux preceding and few hours following the arrival of SI which indicates no substantial precipitation then. The narrow peak is seen on arrival of SI reaching about $7x10^3$ flux. Figure 4 (middle) is the plot of > 30 KeV at L = 5.02 with more than factor of 5 gap between trapped and precipitating electron flux. The peak with slower decaying character is observed in both the trapped and precipitating electron flux at SI arrival. Precipitating electron flux is about $10x10^5$ at SI arrival. Figure 4 (bottom) shows > 30 keV electron flux at L= 6.02 both the trapped and precipitating electron flux showing systematic variation. Precipitating electron flux as can be seen reaches about $15x10^4$ on SI arrival. Similar to L = 5.02, the simultaneous slower decaying peak of both trapped and precipitating electron flux began approximately 0800 MLT to around 1400 MLT on the time of SI as can be seen,

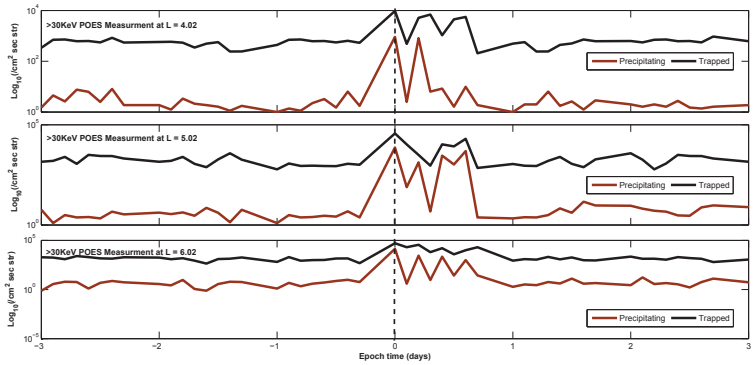


Figure 4. Superposed epoch plot of > 30 keV electron flux measured by NOAA/POES (15-18) satellite. The black line in each plot indicates measurements from 90° telescope (trapped flux) while the red line indicates measurements from 0° telescope (precipitating flux). The dash vertical line indicates arrival of SI (top) is the >30 KeV at L= 4.02, (middle) is L= 5.02 while (bottom) is fluxes at L= 6.02.

there is almost same lower level variation generally maintained between 1500-0500 MLT.

6. Summary and conclusion

Superposed epoch analysis shows that these set of events occured during southward GSM-Bz. This confirms compression of magnetosphere in slow solar wind [2] during the events. Also, the peak of magnetospheric convection coincided with the arrival of SI.

Across GEO, the tendency for dropouts begin during compressed slow solar wind. It is clear that there is time delays in minimum flux at higher energy with respect to lower energy band after the arrival of SI. This analysis corresponds to the GPS observation of L* = 4.5 reported by [2], who found dropout to be energy and L* dependent. Apparently, the cause of the dropouts could arise through inward diffusion from the plasmasheet, or may represent the increased population of ring current electrons via increased magnetospheric convection.

POES 15-18 orbiting LEO shows that more than 30 keV precipitating electron flux level is about a factor of 5 lower than trapped flux precceding and about a day following the time of the event. Consistent with some previous observations, e.g [7, 6], there is tendency in precipitation to follow the trend of trapped flux. Our statistical analysis revealed an unstructured abrupt peak (with slower decaying character) in precipitating electron flux on arrival of the SI which coincided with time of deepest minimum of observed dropouts. Analyis also reveals L = 5.02 and L = 6.02 as region with higer precipitation of electron flux. These observations could be infer to explanation of some loss mechanisms under high speed stream-interference.

Acknowledgments

We thank R. Friedel for providing LANL/SOPA data. Research was supported by the SANSA Space Science research program. We also thank providers of NOAA/POES data in NSSDC, and the OMNI-2 Data sets.

References

[1] Friedel R H W, Reeves G and Obara T 2002 *Journal of Atmospheric and Solar-Terrestrial Physics* **64** 265
[2] Morley S K, Friedel R H W, Spanswick E L, Reeves G D, Steinberg J T, Koller J, Cayton T and Noveroske E 2010 *Proceedings of Royal Society* **466** 3329–3350
[3] Omni data URL `http://omniweb.gsfc.nasa.gov`

[4] Belian R D, Gisler G, Cayton T and Christensen R 1992 *Journal of Geophysical Research* **97** 16897
[5] Rodger C J, Clilverd M A, Green J C and Lam M M 2010 *Journal of Geophysical Research* **115** A04202
[6] Lam M M, Horne R B, Meredith N P, Glauert S, Moffat-Griffin T and Green J C 2010 *Journal of Geophysical Research* **115** A00F08
[7] Clilverd M A, Rodger C J, Gamble R J, Ulich T, Raita A, Seppälä A, Green J C, Thomson N R, Sauvaud J A and Parrot M 2010 *Journal of Geophysical Research* **115** A12304

Pulsar-like hard X-ray emission in AE Aquarii

B Oruru and P J Meintjes

Department of Physics, University of the Free State, P. O. Box 339, Bloemfontein 9300, ZA

E-mail: OruruB@ufs.ac.za, MeintjPJ@ufs.ac.za

Abstract. Spectral analyses have revealed that the X-ray emission in AE Aquarii displays both thermal and non-thermal characteristics. The recently detected non-thermal hard X-ray component above 10 keV is pulsed with the 33 s spin period of the white dwarf, and displays the tell-tale signature of synchrotron emission of high energy electrons in the white dwarf magnetosphere. It is proposed that these electrons are accelerated by large field aligned potentials of $V > 10^{12}$ V, a process common in the environments of fast rotating neutron stars, or pulsars. This places AE Aquarii in a unique category with respect to other members of the cataclysmic variables.

1. Introduction

The low-mass non-eclipsing close binary system AE Aquarii (hereafter AE Aqr) has traditionally been classified as a nova-like variable [1, 2, 3]. The peculiar high variability has been explained in terms of a magnetic white dwarf accreting gas from an orbiting disc surrounding it, which is fed from the mass transfer from the orbiting secondary star. Based upon most of the observable properties, AE Aqr was then considered as a member of the DQ Herculis sub-class of magnetic cataclysmic variables, where a fast rotating magnetized white dwarf is accreting material from an accretion disc [4]. However, the lack of double-peak emission lines, the tell-tale signature of emission originating in a disc, combined with the high line variability, suggest the absence of a disc. The current mass transfer rate $\dot{M} \sim 10^{17}$ g s^{-1} [5] from the secondary star is also not able to penetrate the fast rotating magnetosphere deep enough to form a disc, and is most probably ejected from the binary system through a magnetospheric propeller process [5, 6].

AE Aqr shows orbital modulation at a period of 9.88 h [1, 4], which is rather long for a typical cataclysmic variable. The white dwarf is spinning at $\sim$ 33 s [7, 8] and has a surface magnetic field of $B_* \sim 10^6$ G [10]. The white dwarf is reportedly spinning down at a period rate of $\dot{P} \geqslant 5.642{\times}10^{-14}$ s s^{-1} [7, 9], resulting in an inferred spin-down power of the order of $P_{\rm sd} \sim 10^{34}$ erg s^{-1}, believed to be the driving mechanism behind the magnetospheric propeller process, as well as particle acceleration and non-thermal emission [5, 10].

AE Aqr has been detected in almost the entire electromagnetic spectrum, in which coherent oscillations at the white dwarf spin period have been observed [4, 11, 7, 12, 13]. The most unique characteristic of AE Aqr is perhaps its rapid flaring in almost all wavelengths [14]. In radio and TeV γ-rays, AE Aqr shows up as a transient non-thermal source, and could rather be compared with Cyg X-3 (a microquasar) than with any of the presently known cataclysmic variables [8]. In the remaining parts of the spectrum, the emission is predominantly thermal. However, the recent Suzaku X-ray detection [15, 16] has shown that the high energy component of the X-ray emission ($E_\gamma \geq 10$ keV) displays a non-thermal spectral nature, with a photon power-law index

of $\Gamma = 1.16$ similar to the photon power-law index observed in young rotation-powered pulsars in the energy range between 2-10 keV [16]. Based on the observed characteristics of the hard X-ray emission (listed in §2), a model is proposed to explain it within the framework of a pulsar-like mechanism, driven by the spin-down power released in the propeller process. This would place AE Aqr in the same category as rotation-powered pulsars driven by the loss of rotational kinetic energy of a rapidly rotating, highly magnetized neutron star.

2. Possible models for X-ray emission

The observed X-ray features of AE Aqr make it quite different from most cataclysmic variables, which are mainly accretion driven. The spectrum are mainly soft [17, 18], the inferred plasma density of $n_e \sim 10^{11}$ cm^{-3} [6] is a few orders less than the estimate in the postshock accretion column of cataclysmic variables [18], and the inferred linear scale of $l_{\rm p} \geqslant 2\times 10^{10}$ cm [6, 18] implies that the emission is probably not coming from the surface of the white dwarf. The non-thermal nature of the observed hard X-ray component could indicate a different mechanism, driven by the loss of rotational kinetic energy from the fast rotating white dwarf [19]. In this section, possible emission models are investigated to constrain the observed thermal and non-thermal X-ray emission in AE Aqr.

2.1. Soft X-rays: Magnetospheric accretion

Accretion in AE Aqr may be compared to those in X-ray pulsars, where rapidly rotating magnetized neutron stars accrete from their companions [20]. This can occur either in the accretor or propeller phases, depending on the rotational period and surface magnetic field of the primary star [21]. In the accretor phase, most of the matter is channeled onto the magnetic poles and falls onto the surface creating small heated regions where energy is radiated [22]. The scenario corresponds to the flow of a mixture of gas and radiation inside a magnetic pipe, sealed with respect to the gas but transparent to the radiation [23]. The typical accretion luminosity applicable to most white dwarfs in cataclysmic variables in general, is given by

$$L_{\rm acc} = \frac{GM_*\dot{M}}{R_*} \simeq 1.3\times 10^{34}\left(\frac{M_*}{M_\odot}\right)\dot{M}_{17}R_{*,9}^{-1}\ {\rm erg\ s^{-1}}, \tag{1}$$

where M_* is the mass of the white dwarf in Solar mass units, $R_{*,9}$ its radius in the units of 10^9 cm, and $\dot{M}_{17}$ is the typical mass accretion rate in the units of 10^{17} g s^{-1} respectively. The estimate of the temperature of a black body radiating this luminosity is

$$T_{\rm eff} = \left(\frac{L_{\rm acc}}{A_{\rm cap}\sigma_{\rm SB}}\right)^{1/4} \simeq 3.27\times 10^5\left(\frac{M_*}{M_\odot}\right)^{1/4}\dot{M}_{17}^{\frac{1}{4}}R_{*,9}^{-1}P_{33}^{\frac{1}{4}}\ {\rm K}, \tag{2}$$

where $A_{\rm cap} = \pi(\Omega_* R_*/{\rm c})R_*^2$, is the area of the polar cap, Ω_* is the angular frequency of the white dwarf, $\sigma_{\rm SB}$ is the Stefan-Boltzmann constant, and P_{33} is the spin period, normalized to the 33 s spin period of the white dwarf in AE Aqr. However, the observed X-ray luminosity of AE Aqr is $L_{\rm X} \sim 10^{31}$ erg s^{-1} [12, 17], 3 orders of magnitude less than $L_{\rm acc}$, and the observed temperature of $T_{\rm X} > 10^6$ K [11] is an order of magnitude greater than $T_{\rm eff}$. Therefore, it is unlikely that the X-ray emission is the result of accretion onto the surface of the white dwarf.

If a fraction α of the accretion power liberated at the magnetospheric radius, i.e. the Alfvén radius, is converted to plasma heating, the resultant heating rate is

$$\dot{E}_{\rm th} = \frac{\alpha GM_*\dot{M}}{R_{\rm M}} \simeq 10^{31}\left(\frac{\alpha}{0.01}\right)\left(\frac{M_*}{M_\odot}\right)^{8/7}\dot{M}_{17}^{\frac{9}{7}}\mu_{33}^{-\frac{4}{7}}\ {\rm erg\ s^{-1}}, \tag{3}$$

where $R_{\rm M}$ is the Alfvén radius,

$$R_{\rm M} = \left(\frac{\mu^2}{2\dot{M}(2GM_*)^{1/2}}\right)^{2/7} \simeq 2\times10^{10}\left(\frac{M_*}{M_\odot}\right)^{-1/7}\dot{M}_{17}^{-\frac{2}{7}}\mu_{33}^{\frac{4}{7}}\ {\rm cm}, \tag{4}$$

with μ_{33} the magnetic moment of the white dwarf in the units of 10^{33} G cm^3 and mass transfer rates of the order of $\sim 10^{17}$ g s^{-1}, applicable to AE Aqr and most cataclysmic variables. Hence, $\dot{E}_{\rm th} \rightarrow L_{\rm X}$ for $\alpha \rightarrow 0.01$. It can be shown that this dissipation rate of gravitational potential energy at the magnetospheric radius, i.e. the propeller zone, will heat the plasma to temperatures $kT_{\rm X} \leq 10\,{\rm keV}$, sufficient to explain the soft X-ray emission in AE Aqr [17].

2.2. Hard X-rays: Pulsar model

The energetics of particle acceleration and accompanying synchrotron emission of accelerated electrons are investigated as a possible model to explain the non-thermal power-law nature of the hard X-ray emission (above 10 keV) in AE Aqr. Non-thermal X-ray emission in pulsars has been associated with synchrotron radiation from relativistic electrons accelerated within the magnetosphere of the fast rotating neutron star [24]. This pulsar-like process is particularly attractive for AE Aqr since it contains a rapidly rotating, highly magnetic white dwarf, the key factors behind the energetics of pulsar emission [25]. Also, for AE Aqr, the ratio of the observed hard X-ray luminosity to the spin-down power lies in the range 0.01-0.1 %, which is the same as observed from young rotation-powered neutron stars in the 2-10 KeV range [16, 26].

The process of particle acceleration in rapidly rotating magnetospheres have been discussed extensively [28, 29, 30, 18]. The lack of substantial accretion onto the surface of the white dwarf may result in the white dwarf rotating in a region of low particle density, which is required to support the production of electric fields directed along the rotating magnetic fields. Close to the surface, the electric field force ($e\vec{E}$) will cause charged particles to flow away, forming a magnetosphere whose particle density is given by

$$n_{\rm GJ} = \frac{\mathbf{\Omega}_*\cdot\mathbf{B}}{2\pi ec} \simeq 2\times10^3\left(\frac{R_*}{r}\right)^3 B_{*,6}P_{33}^{-1}\ {\rm cm}^{-3}, \tag{5}$$

where $n_{\rm GJ}$ is the so-called Goldreich-Julian particle density allowing the existence of electric fields in a plasma [27], B is the magnetic field strength at a distance $r > R_*$ (B scales as r^{-3}), and $B_{*,6}$ is the field strength at the surface of the white dwarf in the units of 10^6 G. The electric field component $E_\parallel$ can be evaluated for $s > R_*$, where $r = R_* + s$ [31]:

$$E_\parallel = E_{\rm AS}^\parallel\sqrt{2R_*/r} \simeq 10^3 P_{33}^{-\frac{5}{2}}\mu_{33}r_{l,11}^{-1/2}\ {\rm V\ m}^{-1}, \tag{6}$$

where

$$E_{\rm AS}^\parallel = \frac{1}{8\sqrt{3}}\left(\frac{\Omega_* R_*}{c}\right)^{5/2}B_*, \tag{7}$$

with the field being evaluated in the vicinity of the light cylinder radius ($R_{\rm lc} \sim 1.6\times10^{11}$ cm). A thermal plasma exposed to an electric field in excess of the so-called Dreicer field, $E_{\rm D}$ $\sim 2\times10^{-10}(n_{\rm e}/T_{\rm eff})$ statvolts cm^{-1} [32, 10], will accelerate freely without being hindered by particle-particle collisions. For AE Aqr,

$$E_{\rm D} \sim 0.1 n_{e,11}T_{eff,7}^{-1}\ {\rm V\ m}^{-1}, \tag{8}$$

where the particle density and temperature are expressed in units of $10^{11}\,{\rm cm}^{-3}$ and $10^7\,{\rm K}$ respectively. One can see that $\delta = \frac{E_\parallel}{E_D} \sim 10^4$. Thus, the electric fields along the magnetic fields are large enough to effectively accelerate the electrons to high energy.

In the region of the polar cap, the electric potential is

$$V_{\rm pc}(r) = \int_{R_*}^{r} E_{\parallel} ds \simeq 2 \times 10^{11} P_{33}^{-\frac{5}{2}} \mu_{33} R_{*,9}^{1/2} \left[\left(\frac{r}{R_*} \right)^{1/2} - 1 \right] \ {\rm V}, \tag{9}$$

and close to the light cylinder radius where $r \sim R_{\rm lc} \simeq c/\Omega_*$,

$$V(R_{\rm lc}) \simeq 3 \times 10^{12} P_{33}^{-2} \mu_{33} \ {\rm V}. \tag{10}$$

Figure 1 shows the variation of electric potential with radius. Close to the surface of the white dwarf, $V = 0$ and particle acceleration is not possible. Far outside, the electric potential is large

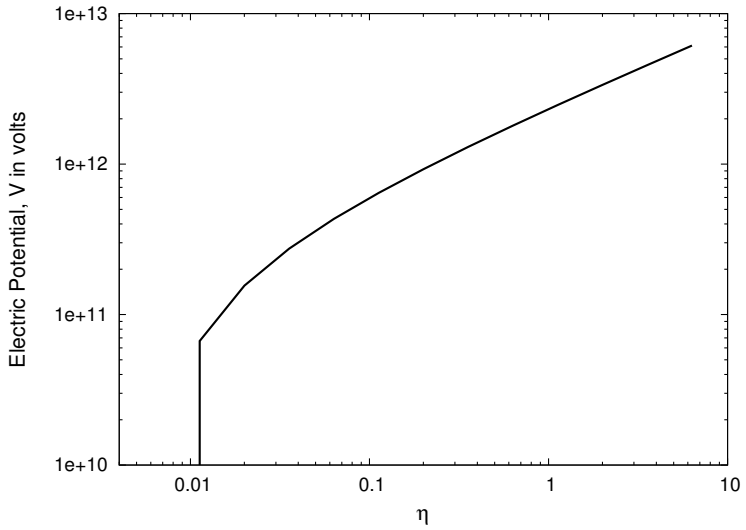


Figure 1. Electric field potential (V) as a function of radius (where $\eta = r/R_{\rm lc}$).

enough to accelerate the electron population. A particle accelerated in $V_{\rm pc}(r)$ has energy,

$$\varepsilon_{\rm p} = eV_{\rm pc}(r) \simeq 2 \times 10^{11} P_{33}^{-\frac{5}{2}} \mu_{33} R_{*,9}^{1/2} \left[\left(\frac{r}{R_*} \right)^{1/2} - 1 \right] \ {\rm eV}, \tag{11}$$

hence a Lorentz factor of

$$\gamma_{\rm p} = \frac{\varepsilon_{\rm p}}{m_e c^2} \simeq 4 \times 10^5 P_{33}^{-\frac{5}{2}} \mu_{33} R_{*,9}^{1/2} \left[\left(\frac{r}{R_*} \right)^{1/2} - 1 \right]. \tag{12}$$

As the electrons are accelerated, they also experience energy losses due to synchrotron radiation and inverse Compton (IC) scattering [33]. In strong magnetospheres, synchrotron radiation dominates significantly, and the energy loss rate of an accelerated electron is

$$L_{\rm syn} = \frac{4}{3} c \sigma_{\rm T} U_{\rm B} \gamma_{\rm p}^2, \tag{13}$$

where $\sigma_{\rm T} \sim 6.65 \times 10^{-25}$ cm^2, is the Thomson cross section, and the magnetospheric energy density $U_{\rm B}$ is given by

$$U_{\rm B} = \frac{B^2}{8\pi} \simeq 10^{12} \left(\frac{B_*}{10^6 \ {\rm G}} \right)^2 \left(\frac{R_*}{r} \right)^6 \ {\rm erg \ cm^{-3}}, \tag{14}$$

where $B = B_*(R_*/r)^3$, is the magnetic field at radial distance r from the centre of the white dwarf. The energy radiation rate from the white dwarf magnetosphere is then

$$L_{\rm rad} = n(r)V'P_{\rm syn}, \tag{15}$$

where $n(r)$ is the accelerated particle density and V' is the volume of the emission region, which is the cylindrical shell bounded by the light cylinder radius and the radial distance r:

$$V' = \pi(r^2 - R_{\rm lc}^2)(2R_*) \simeq 1.6 \times 10^{32}(\eta^2 - 1)R_{*,9}\left(\frac{P}{33\ {\rm s}}\right)^2\ {\rm cm}^3, \tag{16}$$

where $\eta = r/R_{\rm lc}$. The volume V' obtained above is a lower limit, the upper limit is obtained by considering a spherical shell. It can then be shown that

$$L_{\rm rad} \approx 3 \times 10^{40} n(r)(\eta^2 - 1)P_{33}^{-3}B_{*,6}^2\mu_{33}^2R_{*,9}^2\left(\frac{R_*}{r}\right)^6\left[\left(\frac{r}{R_*}\right)^{1/2} - 1\right]^2\ {\rm erg\ s}^{-1}. \tag{17}$$

It can also be shown that, in the region of the polar cap, particles gain energy at the rate,

$$P_{\rm gain} = \varepsilon_{\rm p}\dot{N} \simeq 2 \times 10^{26} n(r)P_{33}^{-\frac{7}{2}}\mu_{33}R_{*,9}^{\frac{7}{2}}\left[\left(\frac{r}{R_*}\right)^{1/2} - 1\right]\ {\rm erg\ s}^{-1}, \tag{18}$$

where $\dot{N} \simeq n(r)c\Delta s$, is the flux of relativistic particles, and $\Delta s = \pi(\Omega R/c)R^2$, is the area of the polar cap on the surface of the white dwarf about which the particles escape.

In the magnetospheric field closed to the white dwarf, energy losses will inhibit the acceleration process. Effective acceleration is anticipated to occur in regions where the synchrotron losses are less dominant than the acceleration. To constrain the lower boundary of the acceleration zone, the radial distance (r) is determined by equating the power in relativistic particles (Eq. 18) to the total synchrotron luminosity (Eq. 17). One then obtains,

$$(\eta^2 - 1)\left(\frac{R_*}{r}\right)^6\left[\left(\frac{r}{R_*}\right)^{1/2} - 1\right] \leqslant 7 \times 10^{-15}. \tag{19}$$

Considering that $R \ll r$, it can be shown that the acceleration dominates synchrotron losses in regions corresponding to $r \geqslant 6{\times}10^{11}$ cm, which translates to $\eta = r/R_{\rm lc} \geqslant 4$. This implies that effective acceleration occurs at distances significantly outside the propeller zone.

Using Hubble Space Telescope (HST) UV spectroscopic measurements, the interblob plasma density in the propeller zone, i.e. $r \sim 10^{10}$ cm [5, 10], has been constrained between $10^6\ {\rm cm}^{-3} \leqslant n_{\rm p} \leqslant 10^{11}\ {\rm cm}^{-3}$ [34, 10]. For the model synchrotron power (Eq. 17) to match the observed hard X-ray luminosity of $L_{\rm HX} \leqslant 5{\times}10^{30}\ {\rm erg\ s}^{-1}$ [15, 16], the relativistic electron number density in the emission region is $n(r) \sim 4{\times}10^3\ {\rm cm}^{-3}$, which is comparable to the estimated Goldreich-Julian value (Eq. 5) and orders of magnitude below the plasma density in the propeller zone closer to the white dwarf. It is encouraging to note that only a minute fraction $\eta \rightarrow 0.004$ of the lower limit of the thermal plasma density in the propeller zone, closer to the white dwarf, accelerated to high energy, is sufficient to explain the non-thermal hard X-ray synchrotron emission in AE Aqr.

3. Summary

It was shown that the thermal X-ray emission in AE Aqr is not the result of accretion onto the surface of the white dwarf but possibly the dissipation rate of gravitational potential energy in the propeller zone. The non-thermal hard X-ray emission has been accounted for in terms of a pulsar-like mechanism, for which synchrotron radiation of accelerated electrons takes place just outside the radius of the light cylinder of the white dwarf. The pulsar-like process will explain successfully the pulsed nature of the hard X-ray emission, as well as the observed power-law spectrum, which is also associated with the synchrotron X-ray emission from fast rotation-powered pulsars. This makes AE Aqr unique among the cataclysmic variables.

Acknowledgments

This study has been funded by the African Square Kilometre Array Project, and the National Research Foundation of South Africa. The authors would like to thank the anonymous referees for the suggestions and positive criticism.

References

[1] Joy A H 1954 *Astrophys. J.* **120** 377–83
[2] Crawford J A and Craft R P 1956 *Astrophys. J.* **123** 44–54
[3] de Jager O C 1991 *Astrophys. J.* **378** 286–92
[4] Patterson J 1979 *Astrophys. J.* **234** 978–92
[5] Wynn G A, King R A and Horne K 1997 *Mon. Not. R. Astron. Soc.* **286** 436–46
[6] Itoh K, Okada S, Ishida M and Kunieda H 2006 *Astrophys. J.* **639** 397–04
[7] de Jager O C, Meintjes P J, O'Donoghue D and Robinson E L 1994 *Mon. Not. R. Astron. Soc.* **267** 577–88
[8] Ikhsanov N K 1997 *Astron. Astrophys.* **325** 1045–54
[9] Mauche C W 2006 *Mon. Not. R. Astron. Soc.* **369** 1983–87
[10] Meintjes P J and de Jager O C 2000 *Mon. Not. R. Astron. Soc.* **311** 611–20
[11] Patterson J, Branch D, Chincarini G and Robinson E L 1980 *Astrophys. J. Lett.* **240** 133–36
[12] Eracleous M, Horne K, Robinson E L, Zhang E, Marsh T R and Wood J H 1994 *Astrophys. J.* **433** 313–31
[13] Meintjes P J, de Jager O C, Raubenheimer B C, Nel H I, North A R, Buckley D A H and Koen C 1994 *Astrophys. J.* **434** 292–05
[14] Ikhsanov N K, Neustroev V V and Beskrovnaya N G 2004 *Astron. Astrophys.* **421** 1131–42
[15] Terada Y, Ishida M, Mukai K, Dotani T, Makishima K, Naik S, Hayashi T, Okada S, Nakamura R and Enoto T 2008 *Advances in Space Research* **41** 512–17
[16] Terada Y, Hayashi T, Ishida M, Mukai K, Dotani T, Okada S, Nakamura R, Naik S, Bamba A and Makishima K 2008 *Publ. Astron. Soc. Japan* **60** 387–97
[17] Choi C-S, Dotani T and Agrawal P C 1999 *Astrophys. J.* **525** 399–06
[18] Ikhsanov N K and Biermann P L 2006 *Astron. Astrophys.* **445** 305–12
[19] Cheng K S, Gil J and Zhang L 1998 *Astrophys. J. Lett.* **493** 35–38
[20] Bildsten L *et al.* 1997, *Astrophys. J. Suppl. Ser.* **113** 367–08
[21] Lipunov V M 1992 *Astrophysics of Neutron Stars* (Berlin: Springer-Verlag) p 82
[22] Becker P A and Wolff M T 2007 *Astrophys. J.* **654** 435–57
[23] Bednarek W 2009 *Astron. Astrophys.* **495** 919–29
[24] Cheng K S, Taam R E and Wang W 2006 *Astrophys. J.* **641** 427–37
[25] Ikhsanov N K 1998 *Astron. Astrophys.* **338** 521–26
[26] Becker W and Trümper J 1997 *Astron. Astrophys.* **326** 682–92
[27] Goldreich P and Julian W 1969 *Astrophys. J.* **175** 869–80
[28] Leung P P, Cheng K S and Fung P C W 1993 *Astrophys. J.* **410** 309–14
[29] Usov V V 1988 *Pis'ma Astron. Zh.* **14** 606–09
[30] Usov V V 1993 *Astrophys. J.* **410** 761–63
[31] Arons J and Scharlemann E T 1979 *Astrophys. J.* **231** 854–79
[32] Dreicer H 1959 *Phys. Rev.* **115** 238–49
[33] Bednarek W and Pabich J 2011 *Mon. Not. R. Astron. Soc.* **411** 1701–06
[34] Eracleous M and Horne K 1996 *Astrophys. J.* **471** 427–46

X-ray timing and spectral analyses of the unusual magnetic cataclysmic variable AE Aquarii

B Oruru and P J Meintjes

Department of Physics, University of the Free State, P. O. Box 339, Bloemfontein 9300, ZA

E-mail: OruruB@ufs.ac.za, MeintjPJ@ufs.ac.za

Abstract. Analyses of the X-ray light curves and spectra of AE Aquarii from data observed contemporaneously with Chandra and Swift X-ray satellites are presented. The X-ray emission in the 0.3-10 keV range is predominantly soft and characterised by flares and emission lines, contrary to the hard X-ray emission above 10 keV, which shows a non-thermal nature, possible synchrotron emission of high energy electrons. The soft X-ray spectra, below 10 keV, can be reproduced by multi-component thermal emission models, and the time-averaged X-ray luminosity is calculated to be $\sim 10^{31}$ erg s^{-1}. It is shown that the X-ray emission is pulsed at a period consistent with the 33 s spin period of the white dwarf, and its source is linked with the interaction between the white dwarf magnetosphere and the accretion flow from the red star (or secondary) companion.

1. Introduction

AE Aquarii (AE Aqr) consists of a fast rotating magnetized white dwarf orbiting a late-type K3-5 main sequence companion [1, 2, 3]. The binary period is 9.88 h [4, 6]. Although the system has been classified as a DQ Her-type cataclysmic variable, usually associated with the presence of an accretion disc, the observed properties of the system do not conform to this [7, 3]. The absence of double-peaked emission lines, usually associated with the presence of an accretion disc, combined with the rapid variation in the line intensities forthcoming from Hubble Space Telescope (HST) UV spectroscopy, associated with collisional excitation and high velocity, cast serious doubt whether an accretion disc is present in the system [5]. These phenomena have been explained successfully in terms of the fast rotating white dwarf expelling the mass flow from the secondary star from the system, i.e. the so-called magnetospheric propeller effect [7].

AE Aqr has been detected in almost all wavelength bands [8], from radio, through optical to TeV γ-rays [9, 1, 10, 11]. Coherent oscillations at 33 s were first observed by Patterson in the optical light [1], and later confirmed in UV and X-ray wavelengths [12, 13, 14]. AE Aqr is a relatively bright X-ray source ($L_{\rm X} \sim 10^{31}$ erg s^{-1} [3]), and coupled with its close proximity to the Earth ($D \sim 100$ pc [4, 6]), it has been observed on a regular basis [12, 15, 3, 16, 17]. Previous analyses have reported that the soft X-ray emission below 10 keV is predominantly thermal, but a recent detection using the Suzaku satellite has reported a non-thermal hard X-ray power-law component above 10 keV [17], probably the result of synchrotron radiation from high energy electrons accelerated by huge potentials in the fast rotating white dwarf magnetosphere, i.e. a pulsar-like process. The reported photon index of $\Gamma = 1.16$ correlates well with the observed photon power-law indices observed in young rotation-powered pulsars at similar energies [17].

This paper will focus on the soft X-ray ($E_X \leq 10$ keV) characteristics of AE Aqr based on the analyses of data from recent observations with Chandra and Swift.

2. Observations of AE Aqr

The Chandra X-ray satellite [18] observed AE Aqr on August 30, 2005 for ~ 80 ks (ObsID 5431 [16]), using the Advanced CCD Imaging Spectrometer (ACIS) detector and the High Energy Transmission Grating. Standard data processing was done at the Chandra X-ray Center, and

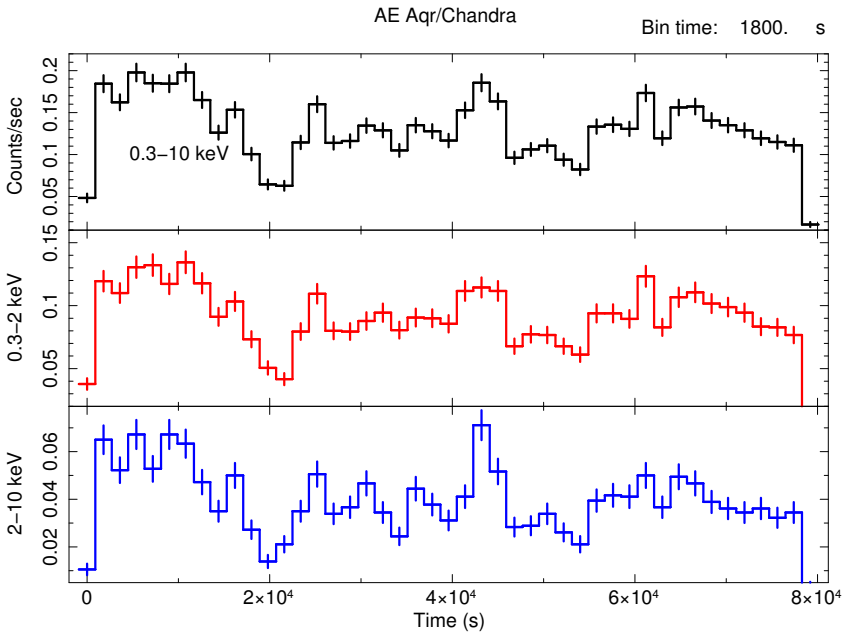


Figure 1. Background subtracted light curve of AE Aquarii from Chandra.

data was acquired through the High Energy Astrophysics Science Archive Research Center (HEASARC) on-line service, provided by the NASA's Goddard Space Flight Center (GSFC). Data reduction and analyses were done using the Chandra Interactive Analysis of Observations software (ciao version 4.2).

The Swift gamma-ray burst explorer [19] observed AE Aqr (target ID = 30295) between August 30 and September 2, 2005 for a total duration of ~ 10.5 ks, as a pre-planned target [20]. Data were archived in September 2005, and standard processing was done later at the UK Swift Science Data Centre. The results presented in this paper are based on data collected when the X-ray telescope (XRT) on-board Swift was operating in the photon counting mode, in which full imaging and spectroscopic resolutions are retained, but timing resolution is limited to 2.5 s [21]. Data was acquired through the HEASARC on-line service. Data reduction and analyses were done using the High-energy Astrophysics software (heasoft version 6.11). The spectra were created using the on-demand software, developed by the Swift team [20].

3. Results

3.1. Light curves

Figures 1 shows the background subtracted light curves of AE Aqr from Chandra data. The figure reveals the well established highly variable nature of the source (i.e. seen in almost all wavelengths), which is also characterised by flares [9, 1]. Most of the X-ray counts are obtained

in the lower energy range (0.3-2 keV). Similar characteristics were also seen in the light curves obtained from Swift XRT data.

3.2. Spectra

Figure 2 shows the unfolded spectrum of AE Aqr from Swift XRT data, fitted with a three-temperature vmekal model. The model temperatures obtained are $0.4^{+0.07}_{-0.07}$ keV, $0.92^{+0.12}_{-0.10}$ keV and $3.95^{+0.94}_{-0.42}$ keV respectively. The spectrum shows emission line features, most of which are in the lower energy range (0.3-2 keV). The unabsorbed X-ray flux obtained in the 0.5-10 keV

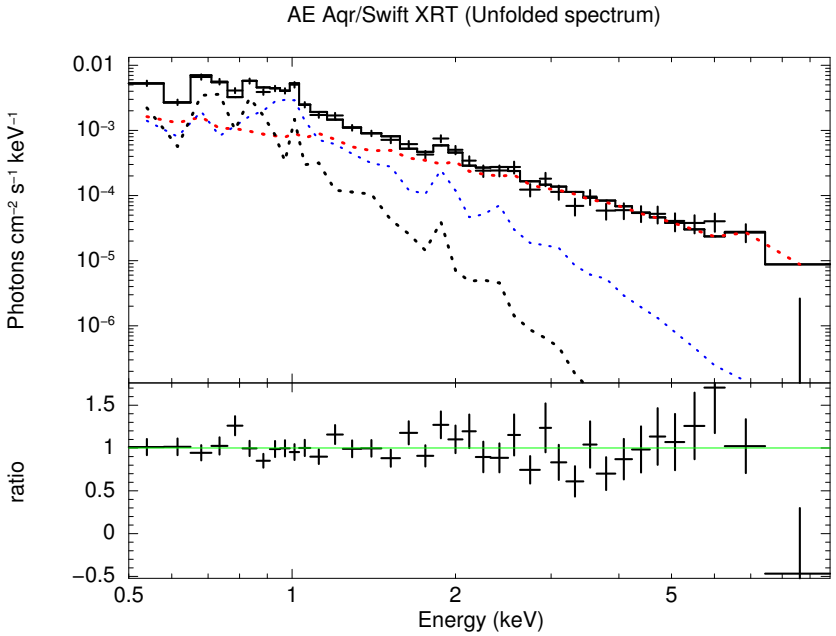


Figure 2. Spectrum of AE Aqr from Swift XRT fitted with a three-temperature vmekal model.

is $7.98^{+0.29}_{-0.44}\times\ 10^{-12}$ erg cm^{-2} s^{-1}, which, for an approximate source distance of 100 pc [4, 6], translates to a luminosity of $L_{\rm X} = 9.03^{+0.23}_{-0.50} \times 10^{30}$ erg s^{-1}.

Figure 3 shows the background subtracted spectrum of AE Aqr from Chandra data, for the medium energy grating (MEG) and diffraction order m = -1. The spectrum was fitted with a three-temperature vmekal model. The model temperatures are $0.24^{+0.02}_{-0.06}$ keV, $0.73^{+0.02}_{-0.07}$ keV and $3.85^{+0.20}_{-0.17}$ keV respectively. Noticeable are the emission lines in the lower energy range of the spectrum, with some of the prominent lines labeled. Similar features were also seen in the spectra (not included in this paper) for the diffraction order $m = +1$, and for the high energy grating (HEG) arm respectively. It has been shown that the conversion of gravitational potential energy of the mass transfer flow from the secondary star towards the region of closest approach, where the plasma is ejected by the propeller, can comfortably drive the total X-ray luminosity assuming a conversion efficiency of only ~ 1 % [15]. This process can also heat the plasma up to temperatures $E_{\rm th} \leq 10$ keV, which is sufficient to drive the emission line spectrum [15].

3.3. Pulse timing

A search for periodicities in the barycentric corrected light curves was carried out using epoch folding method, where the light curves were folded with large number of periods around a guess value (period), and the best period found by chi squared maximization. In this analysis, the

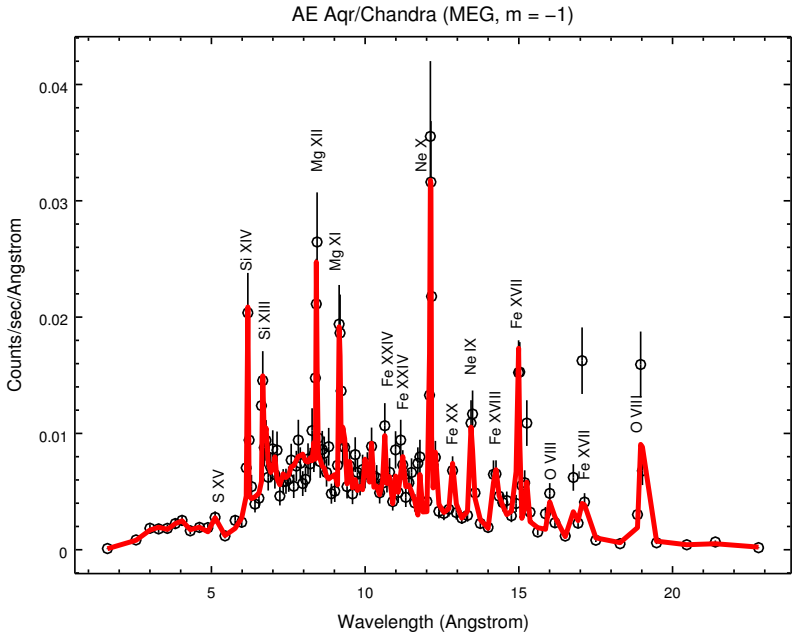


Figure 3. Spectrum of AE Aqr from Chandra fitted with a three-temperature vmekal model.

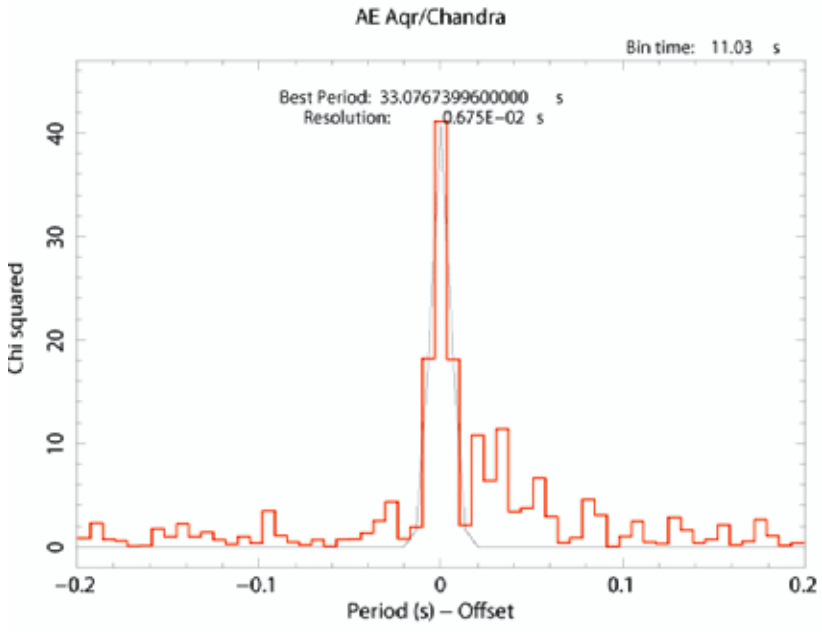


Figure 4. X-ray pulse period determination from Chandra data. Details in the text.

ephemeris for the pulsed emission, obtained from a 14 year baseline of optical observations [13], were used to fold the data. Figure 4 shows the pulse period determined for the Chandra data. The default Fourier period resolution (FPR), $P^2/2T \sim 6.75 \times 10^{-3}$ s was used, where P is the

folding period and T is the length of observation ($\sim$ 80 ks). The resulting peak was fitted with a gaussian, which was used to determine the actual pulse period, $P_{\rm pulse}$ = 33.0767$\pm$0.0068 s. Similarly, the pulse period obtained for the Swift XRT data is 33.076$\pm$0.052 s. The pulse periods obtained for both data sets (i.e. Chandra and Swift XRT) are perfectly reconcilable with the spin period of the white dwarf determined earlier [13]. Using the period obtained from the longer Chandra data set, the pulse profiles were determined using the epoch of BJD 2453673.5 [17]. These profiles are shown in Figures 5 and 6 for Chandra and Swift XRT data respectively.

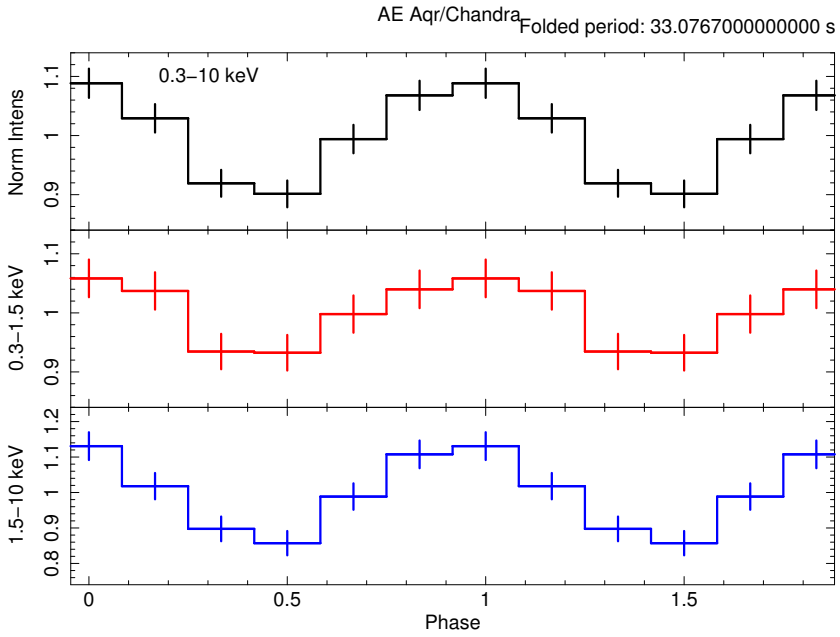


Figure 5. Pulse profile of AE Aqr from Chandra data. Phase 0 corresponds to BJD 2453673.5.

4. Discussion

The light curves of AE Aqr show that the source is highly variable and characterised by flares. Since AE Aqr is a non-eclipsing binary system, the variability is most likely the result of the interaction between the mass transfer from the secondary star a very efficient magnetospheric propeller process. Flares could be the result of sudden increase in the emission due to sporadic mass accretion onto the surface of the white dwarf. The spectra show a number of emission line features, and can be fitted with multi-component thermal emission models, suggesting temperatures in the X-ray emission region below 10 keV. A period and pulse analyses show that the X-ray emission is pulsed at the spin period of the white dwarf.

Acknowledgments

This work made use of data supplied by the Chandra X-ray Center, and the UK Swift Science Data Centre at the University of Leicester. The research has been funded by the African Square Kilometre Array Project, and the National Research Foundation of South Africa. The authors would like to thank the anonymous referees for constructive criticism.

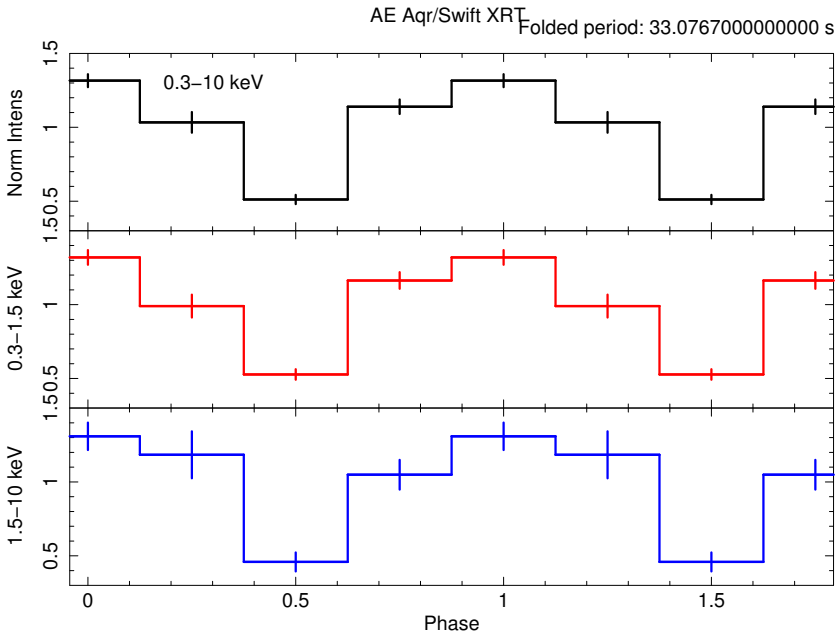


Figure 6. Pulse profile of AE Aqr from Swift XRT data. Phase 0 corresponds to BJD 2453673.5.

References

[1] Patterson J 1979 *Astrophys. J.* **234** 978–92
[2] Welsh W F, Horne K and Gomer R 1995 *Mon. Not. R. Astron. Soc.* **275** 649–70
[3] Itoh K, Okada S, Ishida M and Kunieda H 2006 *Astrophys. J.* **639** 397–04
[4] Welsh W F, Horne K and Gomer R 1993 *Astrophys. J. Lett.* **410** 39–42
[5] Eracleous M and Horne K 1996 *Astrophys. J.* **471** 427–46
[6] Welsh W F, Horne K and Oke J B 1993 *Astrophys. J.* **406** 229–39
[7] Wynn G A, King R A and Horne K 1997 *Mon. Not. R. Astron. Soc.* **286** 436–46
[8] de Jager O C 1991 *Astrophys. J.* **378** 286–92
[9] Bookbinder J A and Lamb D Q 1987 *Astrophys. J. Lett.* **323** 131–35
[10] Chadwick P M *et al.* 1995 *Astroparticle Physics* **4** 99–11
[11] Meintjes P J, de Jager O C, Raubenheimer B C, Nel H I, North A R, Buckley D A H and Koen C 1994 *Astrophys. J.* **434** 292–05
[12] Patterson J, Branch D, Chincarini G and Robinson E L 1980 *Astrophys. J. Lett.* **240** 133–36
[13] de Jager O C, Meintjes P J, O'Donoghue D and Robinson E L 1994 *Mon. Not. R. Astron. Soc.* **267** 577–88
[14] Eracleous M, Horne K, Robinson E L, Zhang E, Marsh T R and Wood J H 1994 *Astrophys. J.* **433** 313–31
[15] Choi C-S, Dotani T and Agrawal P C 1999 *Astrophys. J.* **525** 399–06
[16] Mauche C W 2006 *Mon. Not. R. Astron. Soc.* **369** 1983–87
[17] Terada Y, Hayashi T, Ishida M, Mukai K, Dotani T, Okada S, Nakamura R, Naik S, Bamba A and Makishima K 2008 *Publ. Astron. Soc. Japan* **60** 387–97
[18] Brissenden R J 2001 *Astron. Soc. Pac. Conf. Ser.* **238** 22–31
[19] Gehrels N *et al.* 2004 *Astrophys. J.* **611** 1005–20
[20] Evans P A *et al.* 2009 *Mon. Not. R. Astron. Soc.* **397** 1177–01
[21] Burrows D N *et al.* 2005 *Space Sci. Rev.* **120** 165–95

Atomic processes in gaseous nebulae

A Prozesky

Department of Mathematical Sciences, University of South Africa, Preller Street, Pretoria, 0003

E-mail: prozea@unisa.ac.za

Abstract. The atomic physics relevant to gaseous nebulae is critically examined using modelling software with particular emphasis on radio recombination lines (RRLs). The theoretical spectral line intensities can be deduced if we know the population structure of the bound electrons in the gas under non-thermal equilibrium conditions. The population structure of hydrogen is solved for various environments using a capture-collision-cascade (C^3) model that incorporates an ambient radiation field. Effects of an ambient radiation field on the population structure are examined and processes that are stimulated by a radiation field are included in the model. This has been done as a preliminary investigation to extend the model to a photoionization code.

1. Introduction

The understanding of the physics of ionized gases is crucial to many subjects in astronomy. Emission spectra can be observed from numerous astronomical objects. The study of spectral lines in these spectra has yielded valuable information regarding the most elementary atomic processes occurring in the Universe and has proven to be an essential tool in astronomy.

Gaseous nebulae are often permeated by an external radiation field, generally from a nearby star or stars and the cosmic microwave background radiation (CMBR). To model the theoretical spectrum that we would expect to see from a specific nebula, it is necessary to have a detailed knowledge of the atomic processes occurring within the nebula.

A special class of spectral lines results from transitions between highly excited atomic states. These arise when an electron is captured by an ion into an energy level with a large principal quantum number n. If the downward cascading electron makes transitions between levels with small energy differences, it can produce photons that are in the radio regime. Spectral lines that result from this process that are in the radio regime are referred to as radio recombination lines (RRLs).

A model for hydrogen plasmas that is applicable to a large portion of the electromagnetic spectrum, in particular to the radio regime, has been developed. Hitherto, models for gaseous nebulae have focused on transitions at optical and infra-red wavelengths, while the importance of processes involving levels with large principal quantum numbers has not always been recognised.

The model simulates the influences of the ionizing radiation, free particle temperature and density on the excited level population structure of hydrogen. From this, the expected spectral line intensities can be calculated. The model considers an unbounded pure hydrogen plasma permeated by an external radiation field. Stimulated processes are important in the Rayleigh-Jeans limit, so it is necessary to investigate the influence of the incident radiation field on the

population structure.

The model was coded in the programming language C using the MinGW compiler on a Windows XP platform. The code was written completely independently using algorithms described in literature and no existing codes were copied.

2. Atomic Level Populations

The level populations of states within an atom or ion would follow a Boltzmann distribution if the system were in thermodynamic equilibrium (TE). Menzel [1] introduced a correction factor b_n to compensate for the degree of departure from TE of the level population. In this scheme, the Saha-Boltzmann equation becomes

$$N_n = b_n N_e N_i \left(\frac{h^2}{2\pi m_e k_B T_e} \right)^{3/2} n^2 \exp\left(\frac{\chi_n}{k_B T} \right) . \tag{1}$$

The b_n factors are called departure coefficients and b_n equal to unity indicates strict TE. The level populations of a gas described by a specific electron temperature T_e and electron density N_e are solved if the departure coefficients for all levels are known.

The condition of statistical balance can be used to set up balance equations of the processes affecting the populations of each energy level of the atoms. These balance equations are coupled and need to be solved simultaneously to give the departure coefficients.

Equating all the atomic processes filling and emptying level n, gives the statistical balance equation

$$\begin{aligned} & N_e N_i \left(\alpha_n^r + \alpha_n^s + N_e C_{i,n}\right) + \sum_{m>n} N_m \left(A_{mn} + N_e C_{mn} + B_{mn} J_\nu\right) + \sum_{k<n} N_k \left(N_e C_{kn} + B_{kn} J_\nu\right) \\ = \quad & N_n \left[\alpha_n^p + N_e C_{n,i} + \sum_{k<n} \left(A_{nk} + N_e C_{nk} + B_{nk} J_\nu\right) + \sum_{m>n} \left(N_e C_{nm} + B_{nm} J_\nu\right) \right] . \end{aligned} \tag{2}$$

The left-hand side contains all processes that populate level n. The terms represent radiative recombination, stimulated recombination, three-body recombination, spontaneous emission, collisional de-excitation, stimulated emission, collisional excitation and absorption, respectively.

The right-hand side includes all processes that depopulate level n. The terms represent photoionization, collisional ionization, spontaneous emission, collisional de-excitation, stimulated emission, collisional excitation and absorption, respectively.

J_ν is the mean intensity of the incident radiation field. In this work the electron number density N_e and the ion number density N_i are decoupled from the bound level populations N_n, and are taken as constant throughout the plasma. Terms involving J_ν are not included in the previous models described in Brocklehurst [2], Smits [3] or Storey & Hummer [4].

The number density N_i of atoms in level i can be substituted in the rate equation (2), using equation (1), to yield a form of the rate equation that depends on the departure coefficients. This form of the rate equation can be written in matrix form as

$$\mathbf{a} = \mathbf{b} \cdot \mathbf{X}$$

where the row vector $\mathbf{a}$ has components

$$\mathrm{a}_n = \left(\frac{2\pi m_e k_B T}{h^2} \right)^{3/2} \left(\alpha_n^r + \alpha_n^s + N_e C_{i,n}\right) \tag{3}$$

and the row vector $\mathbf{b}$ has components b_n, $n = 1, 2, 3 \ldots$.

The diagonal entries of the matrix **X** represent all the processes depopulating level n and the non-diagonal matrix elements represent all the processes populating level n. All the processes involving downward transitions into n from higher levels are above the diagonal and the processes populating n from lower levels are below the diagonal.

The departure coefficients b_n can be obtained by inverting the matrix **X** and multiplying it with vector **a** from the left. The elements of **a** and **X** depend on the electron temperature T_e, the density N_e, the external radiation field J_ν, and the rates of the individual atomic processes.

3. Atomic calculations

3.1. Bound-bound radiative processes

The Einstein A-values were computed using the expression given by Brocklehurst [5]. The explicit formula for the bound-bound matrix element of an atomic transition [6] was used.

For very high energy levels ($n > 500$), the expression in [6] is not appropriate due to the large number of terms occurring in the hypergeometric series. Instead, the approximation given by Brocklehurst [2] was used with the bound-bound Gaunt factor as given by Baker & Menzel [7].

The rate coefficients for absorption and stimulated emission were calculated using the Einstein relations.

3.2. Bound-free radiative processes

A method to compute the bound-free radial matrix elements that is based on a set of recurrence relations, satisfied by the exact matrix elements for hydrogenic atoms or ions, has been described by Burgess [8]. These simple recurrence relations allow for fast computing to very high accuracy.

The expression given in [8] for the radiative recombination coefficients α_n^r, using a Maxwellian distribution with a temperature T_e for the free electron velocities, was used. A Gaussian integration method was used to evaluate the integral in the expression for the radiative recombination. This method does the integration over an arbitrary interval, so that smaller intervals can be used for energies close to the ionization threshold when the integrand varies rapidly. A number of five-point Gaussian integrations is done starting with an interval size of $h = 10^{-4}n^{-1}$. The interval size is doubled after every five-point integration and the procedure is terminated when the sum of the integrals are accurate up to six significant digits.

The photoionization cross-section to level n for an hydrogenic atom or ion was calculated using the formula of Burgess & Seaton [9]. From this, the cross-section for stimulated emission σ_n^s was calculated using the Einstein-Milne relations.

For a plasma in an ambient radiation field with mean intensity J_ν, the rate of photoionizations is given by

$$\alpha_n^p = \int_{\chi_n/h}^{\infty} a_n^p(\nu)\,\frac{4\pi J_\nu}{h\nu}\,d\nu \tag{4}$$

where $a_n^p(\nu)$ is the photoionization cross-section from level n for a photon with frequency ν, and χ_n is the ionization potential of level n.

The stimulated emission coefficient α_n^s is found by averaging the stimulated emission cross-section $\sigma_n^s(v)$ over the velocity distribution and accounting for the stimulating radiation field. The stimulated emission coefficient is given by

$$\alpha_n^s = \int_{\chi_n/h}^{\infty} \frac{4\pi J_\nu}{h\nu}\sigma_n^s(v)\,f(v)\frac{h}{m}\,d\nu\,. \tag{5}$$

The integration involved in the calculations of the photoionization and stimulated emission coefficients was handled using a Gaussian quadrature scheme as described above for radiative recombination.

3.3. Collisional processes

The semi-empirical formulae of Vriens & Smeets [10] were used to calculate the collisional rate coefficients for collisional bound-bound and bound-free transitions. Because the values are valid over a wider range of temperature, these were used in favour of the more commonly used formulae of Gee et al. [11]. Vriens & Smeets [10] claim that their values agree within 5 to 20 % with those of Gee et al. [11].

4. Numerical methods

4.1. Transition rates close to the ionization limit

In principle, an atom has an infinite number of energy levels and thus the solution of the population structure of hydrogen requires the solution of an infinite number of coupled equations represented by equation (2). Since the atoms discussed here are not in isolation, there are physical considerations, like the density, that limit the actual number of states in which an electron can be found. Therefore, an upper cut-off n_* was introduced for the highest n level for which the rate equation will be solved explicitly. The details of how n_* was determined can be found in [12].

4.2. Matrix condensation

Burgess & Summers [13] introduced a matrix condensation technique based on Lagrange interpolation which has been used by a number of authors. The technique condenses the sizable matrix $\mathbf{X}$ to a much smaller matrix, which can be readily inverted. Because the departure coefficients vary smoothly and slowly with n, the condensation technique can be applied. The method is presented in detail in [12].

5. Hydrogen Population Structure

The departure coefficients were calculated using codes developed based on the theory and techniques discussed in the preceding sections. Results were in good agreement with other capture-collision-cascade (C^3) models The results of this work for an n-method C^3 model agree on average within 0.5 % with the results of Brocklehurst [2] and 0.3 % with that of Smits [3].

The problem of whether or not to include Lyman transitions in the calculations when determining departure coefficients was investigated. The mean free path of Lyman photons from the different energy levels was investigated under various environmental conditions by calculating the extinction coefficient of the photons explicitly. It was concluded that the escape probability of Lyman photons is negligibly small for conditions found in astronomical plasmas that produce a recombination spectrum. Therefore, it is always appropriate to assume Case B of Baker & Menzel [7] when determining departure coefficients.

Departure coefficients for C^3 models permeated by a radiation field were also calculated. In general it is important to take the ionizing radiation into account, but this cannot be described by recombination theory alone and it is necessary to do radiative transfer calculations. The radiation changes as it propagates through the gas and the gas itself will absorb photons at specific frequencies as well as emit a diffuse radiation field. The b_n problem will then depend on many more parameters than for a C^3 model, in particular it will be geometry-dependent. In this work, the effects of a radiation field on the departure coefficients was investigated, but the amplification of lines by stimulated processes was not considered.

The mean intensity of the external radiation field is represented by a blackbody spectrum with temperature T_r, multiplied by a dilution factor W that depends on geometrical considerations of the system. For a single ionizing star with a radius of R, the dilution factor at a distance d (with $d \gg R$) from the star will be given by $W \approx 0.25\,(R/d)^2$. For an O star with $R = 18R_\odot$, the dilution factor will be $W \sim 10^{-16}$ at a distance of 10 pc and $W \sim 10^{-14}$ at 1 pc.

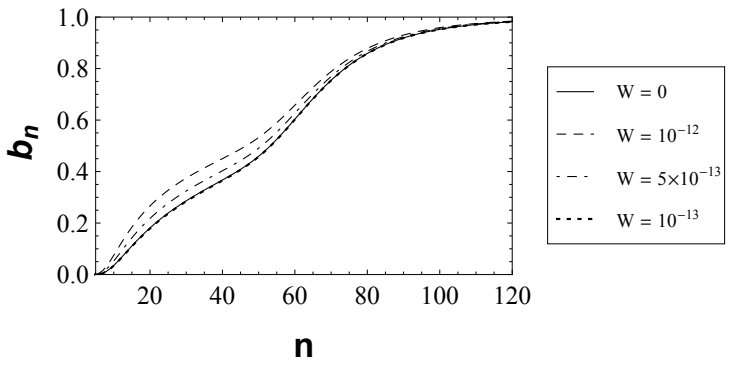


Figure 1. Effect on the population structure of hydrogen of a diluted blackbody radiation field with temperature $T_r = 40\,000\,\mathrm{K}$ with various dilution factors. Parameters for the nebula were taken as $T_e = 10^3\,\mathrm{K}$ with $N_e = 10^3\,\mathrm{cm}^{-3}$, assuming Case B. The solid line shows the population structure if no radiation field is present.

Figure 1 shows the effect that a diluted blackbody field has on the population structure of hydrogen for a plasma with $T_e = 10^3\,\mathrm{K}$ and $N_e = 10^3\,\mathrm{cm}^{-3}$. The radiation temperature was taken as $T_r = 40\,000\,\mathrm{K}$ to represent an O star. For a relatively dense radiation field ($W = 10^{-12}$, 5×10^{-13}), the departure coefficients are larger than they are if no radiation field were present. For less dense fields, the departure coefficients are lowered by the radiation field, as can be seen for the case $W = 10^{-13}$ in figure 1. As the density of the field decreases, the b_n values approach the $W = 0$ case asymptotically from below.

It was found that the effect that a diluted blackbody field has on the population structure of hydrogen for a plasma can be pronounced. For the example described in figure 1, the relative change on the b_n values can be as much as 45% for a relatively dense radiation field ($W = 10^{-12}$).

Even though the CMBR has a blackbody spectrum of only $\sim 3\,\mathrm{K}$, it is very dense and hence affects the population structure of hydrogen atoms. The result of an undiluted 3 K blackbody radiation field on the population structure of hydrogen was examined for 16 environments with parameters $10\,\mathrm{cm}^{-3} < N_e < 10^4\,\mathrm{cm}^{-3}$ and $300\,\mathrm{K} < T_e < 20\,000\,\mathrm{K}$. The effects are most noticeable in cool clouds with low electron densities. The 3 K blackbody radiation field affects the population structures most typically for $50 < n < 150$, increasing the value of the departure coefficients slightly for these levels. In general the b_n values were altered on average by 0.2 % and at most by an average of 0.6 %.

The influence of free-free continuum on the departure coefficients was also investigated and was found to be minimal. This is consistent with the work of [14], who found that the effects of free-free radiation on the b_n values are negligible.

6. Conclusion

A comprehensive model for calculating the n-method populations of a pure hydrogen plasma has been presented. It has been assumed that the nebula is homogeneous, unbounded and permeated by a constant radiation field. Departure coefficients for bound energy levels were computed by accounting for all radiative and collisional processes, bound-bound and bound-free, via all possible transition routes. The model is comparable with the most definitive models available at present [15].

The solution for the departure coefficients presented here considers only distinct energy levels of the atoms, known as an n-model. A more complete description resolves the momentum states and a departure coefficient calculated for every angular momentum state, called an nl-model. In

this model it has been assumed that the angular momentum states are populated according to their statistical weights. This n-model will serve as the basis for future studies using nl-models.

The code is valid for a larger range of temperatures than any of the current models. Because the parameters for astronomical plasmas vary greatly, this broadens the code's potential for modelling a variety of astronomical environments. Nova shells with electron temperatures $T_e < 1\,000\,\mathrm{K}$ have been detected [16], while supernova remnants have temperatures of about $10^6\,\mathrm{K}$.

The results from the model developed in this work were compared to previous calculations and found to be in good agreement. Discrepancies are small and can be explained by the different methods used to calculate atomic transition probabilities and to handle numerical and computational challenges. One of the objectives of this project was to check for systematic errors in previous calculation by independently developing code. No such errors were found.

A preliminary investigation of the effects of an external radiation field on the departure coefficients was done. In principle the external radiation field should be included in the calculations of b_n values, but the incorporation of a radiation field into such a model is not trivial and it necessitates that the geometry of the system be taken into account. It was found that the ionizing radiation from a nearby star can have a significant effect on the departure coefficients of high n levels, as can the CMBR. The influence of free-free emission on the b_n values was found to be insignificant.

Acknowledgments
The financial assistance of the South African Square Kilometre Array Project towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

References

[1] Menzel D H 1937 *Astrophys. J.* **85** 330
[2] Brocklehurst M 1970 *Mon. Not. R. Astron. Soc.* **148** 417
[3] Smits D P 1991 *Mon. Not. R. Astron. Soc.* **248** 193–216
[4] Storey P J and Hummer D G 1995 *Mon. Not. R. Astron. Soc.* **272** 41–8
[5] Brocklehurst M 1971 *Mon. Not. R. Astron. Soc.* **153** 471–90
[6] Gordon W 1929 *Ann. der Phys.* **2** 1031
[7] Baker J G and Menzel D H 1938 *Astrophys. J.* **88** 52
[8] Burgess A 1965 *Mem. R. Astron. Soc.* **69** 1–17
[9] Burgess A and Seaton M J 1960 *Mon. Not. R. Astron. Soc.* **120** 121
[10] Vriens L and Smeets A H M 1980 *Phys. Rev. A* **22** 940–51
[11] Gee C S, Percival L C, Lodge J G and Richards D 1976 *Mon. Not. R. Astron. Soc.* **175** 209–16
[12] Prozesky A 2011 MSc dissertation University of South Africa
[13] Burgess A and Summers H P 1969 *Astrophys. J.* **157** 1007
[14] Dyson J E 1969 *Astrophys. J.* **155** 47
[15] Salem M and Brocklehurst M 1979 *Astrophys. J. Suppl. Ser.* **39** 633–51
[16] Williams R E, Woolf N J, Hege E K, Moore R L and Kopriva D A 1978 *Astrophys. J.* **224** 171–73

Genetic algorithms in astronomy and astrophysics

Vinesh Rajpaul[1,2]

[1] Astrophysics, Cosmology and Gravity Centre (ACGC), Department of Astronomy, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa
[2] South African Astronomical Observatory, PO Box 9, Observatory, 7935, South Africa

E-mail: vinesh.rajpaul@uct.ac.za

Abstract. Genetic algorithms (GAs) emulate the process of biological evolution, in a computational setting, in order to generate good solutions to difficult search and optimisation problems. GA-based optimisers tend to be extremely robust and versatile compared to most traditional techniques used to solve optimisation problems. This paper provides a very brief introduction to GAs and outlines their utility in astronomy and astrophysics.

1. Introduction

Many interesting mathematical problems can be reformulated as global optimisation problems; the solution of systems of algebraic or even differential equations, for example, can be cast quite naturally in terms of optimisation. The same holds true for the all-important inverse problems that are ubiquitous in the physical sciences, i.e. problems where one seeks to transform experimental data into model parameters in order infer properties of the physical systems being studied (a simple example: choosing parameters to minimise a χ^2-statistic when fitting a Voigt-profile to a spectral line).

The goal of a global optimisation problem is, given a so-called *cost function*[1] $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$, to try to find a point (more generally, a set of points) $\vec{x}^* \in \Omega$ such that:

$$\forall \vec{x} \in \Omega : f(\vec{x}) \geqslant f(\vec{x}^*); \tag{1}$$

$f(\vec{x}^*)$ is called the global minimum[2]. A local minimum, $f(\hat{\vec{x}})$, is defined by the condition:

$$\forall \vec{x} \in \Omega, \exists \delta > 0 : \left\| \vec{x} - \hat{\vec{x}} \right\| < \delta \Rightarrow f(\vec{x}) \geqslant f(\hat{\vec{x}}). \tag{2}$$

Whereas finding an arbitrary local minimum of a function is a relatively straightforward task, especially if one has a good "first guess" – extremely efficient techniques exist to solve such local optimisation problems – finding *global* minima is a far more challenging problem. Real-world cost functions tend to be nonlinear, discontinuous and/or hugely multimodal, and there is no foolproof approach to locating their global minima.

Most established approaches – whether deterministic, stochastic or (meta)heuristic – to solving global optimisation problems yield excellent results on a limited class of problems,

[1] Depending on the context, cost functions are also referred to as *energy functions* or *objective functions*.
[2] Maximisation of $f(\vec{x})$ is, of course, equivalent to minimisation of $g(\vec{x}) := -f(\vec{x})$.

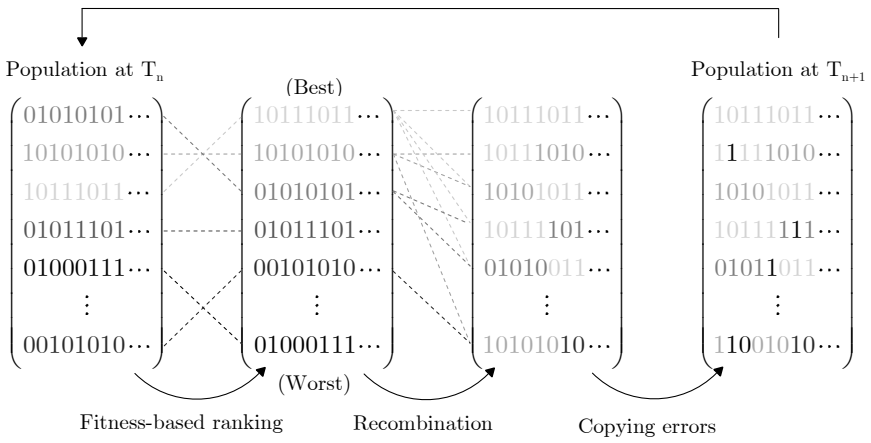


Figure 1. Schematic to illustrate the workings of a simple binary-coded genetic algorithm. Each bit represents a gene; here the genes of high-fitness solutions are given lighter colours.

but have drawbacks that tend to cripple them when faced with certain (reasonably) difficult problems. For example, they might get stuck too easily in local minima, they might be thwarted by discontinuous functions or they might be too slow to be of practical value when faced with enormous search spaces [1].

Evolutionary algorithms, inspired by biological evolution, are metaheuristic optimisation algorithms that tend to yield "good enough" results on a very wide range of (even extremely difficult) optimisation problems. So-called *genetic algorithms* form one of the most successful subsets, and certainly the most popular subset, of evolutionary algorithms.

2. Genetic algorithms: the basic idea

Genetic algorithms, or GAs for short, draw inspiration from population genetics (and, like all evolutionary algorithms, from evolutionary biology in general) and they incorporate, in a computational setting, notions such as natural selection/survival of the fittest, genetic recombination, inheritance and mutation. The first GA-based optimiser was proposed in the mid-1970s [2], and since then many modifications and improvements to the basic algorithm have been developed, including mechanisms without any direct biological analogues [3].

In spite of the rich variety of their potential incarnations, most GAs share a basic working scheme: they start with a population of many candidate solutions (called individuals or phenotypes), associate with each solution an encoded version of the phenotype (called a chromosome, genotype or an individual's genetic material) and also a measure of the solution's fitness (quality). This fitness function is often simply the additive inverse of the cost function to be optimised. Then, by repeated application of "genetic operators" mainly at the genotypic level, they cause the population as a whole to increase in phenotypic fitness, i.e. they cause the solutions to evolve towards optimality.

A typical (though simplistic and by no means general or optimal) working scheme for a genetic algorithm is as follows:

1. construct a random initial population of genotypes;
2. decode the genotypes and evaluate their phenotypic fitness; if the fittest phenotype matches the user-defined target fitness (or other termination criterion), `break`, otherwise continue;
3. produce offspring by randomly selecting and recombining genetic material from the current population, favouring individuals with high phenotypic fitness;

4. introduce, with some low probability, random changes (copying errors) into the genetic material of the offspring;
5. replace low-fitness members of the old population with the offspring created in the previous step, and `goto` step 2.

The selective recombination of the population's genetic material exploits good solutions to build even better ones, and the random mutations serve to inject entirely new and potentially favourable material into the gene pool that could not be obtained simply by recombining the genetic material of existing individuals.

Fig. 1 illustrates the working scheme of a simple GA where the solutions are encoded as binary strings[3]. It may be shown that given enough time, and subject to a few reasonable assumptions, a GA will always converge to the global optimum of a cost function[4] [4, 6].

3. GAs: pros and cons

Relative to more conventional optimisation algorithms, GA-based optimisers offer a number of striking advantages, some of which are outlined below.

Robustness. GA-based optimisers can handle – with aplomb – problems with multimodal or low-contrast objective functions, multiple objectives and/or problems where the parameter spaces have a very high dimensionality [1].

Simplicity. In order to solve a given optimisation problem, an "off-the-shelf" GA requires only a single, unambiguous measure of the quality (fitness) of candidate solutions. They do not require, for example, gradients or Hessian matrices, the computation of which might be prohibitively difficult or impossible in some problems. Moreover the ideas underpinning GAs are intuitively accessible and it is a relatively easy task to develop a working GA from scratch.

Speed. Apart from the intrinsically high speed with which GAs tend to explore large parameter spaces [4], they are embarrassingly parallel: very little effort is required to transform a serial GA-implementation to a parallel implementation. Thus they are well-suited to exploiting high-performance hardware (multi-core workstations, graphics processing units, clusters etc.).

Versatility. A single GA-based optimiser can be expected to yield "good enough" results on a very wide class of problems – from a problem as simple as fitting a three-parameter Gaussian to some data, to one as complex as choosing a molecular configuration to minimise a Buckingham potential with hundreds of parameters – and it is easy to incorporate problem-specific knowledge into a GA-based solver. The widespread adoption of GAs in fields such as engineering, chemistry, biology and economics bears testimony to their great versatility [3].

To illustrate the great robustness and versatility of a typical GA, consider the following cost function proposed by Charbonneau [7]:

$$f(x, y; n) = -[16x(1-x)y(1-y)\sin(n\pi x)\sin(n\pi y)]^2, \tag{3}$$

where $x, y \in [0, 1]$ and $n \in \{2k+1; \forall k \in \mathbb{N}\}$. For $n = 13$, say, it may be shown that $f(x, y)$ has 169 (in general n^2) local minima on its domain, only one of which is the global minimum; moreover, the minima are separated by steep walls and there is little contrast between many of the minima (see fig. 2).

Fig. 3 illustrates how a GA-based optimiser fared on the (rather challenging) $n = 13$ problem: in 10000 trials, the algorithm converged to the global minimum every single time, with the minimum location determined to a median accuracy of about one part in a billion after only

[3] Most early GAs encoded solutions as binary strings, both for the sake of simplicity and supposed theoretical optimality; a large body of empirical evidence, however, indicates that it is preferable to work directly with floating-point representations of solutions when solving numerical optimisation problems [1, 3, 4, 5].

[4] Of course this knowledge is of little practical value; of more importance is the *rate* of convergence to the global optimum, though unfortunately with GAs this rate is highly problem-dependent and difficult to estimate *a priori*.

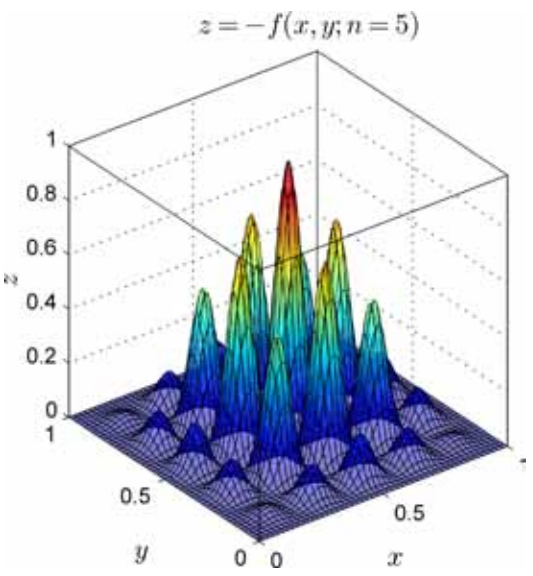


Figure 2. Surface plot of the function $-f(x, y; n)$ defined by equation 3, for the case $n = 5$. In this case there are $n^2 = 25$ local optima on the domain $x, y \in [0, 1]$.

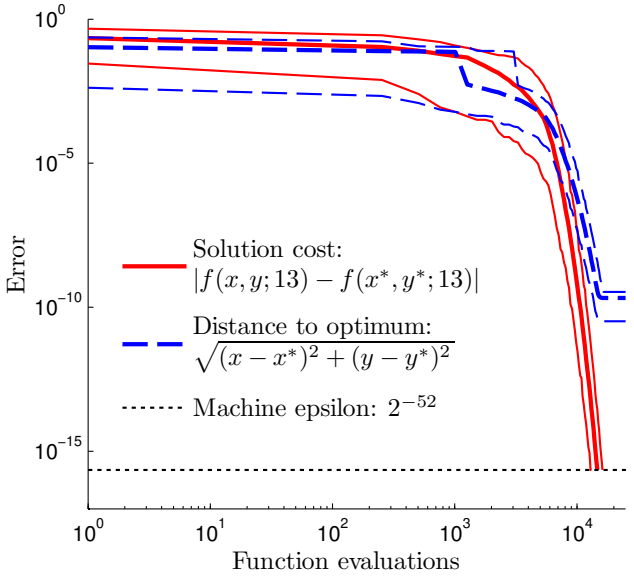


Figure 3. Performance of a GA-based optimiser applied to the $n = 13$ case of $f(x, y; n)$; the thick lines denote median performance in 10000 trials, and the thin lines, upper and lower 3σ limits.

$\sim 10^4$ function evaluations (or a fraction of a second on a modern workstation). For comparison, a blind random search would require $\sim 10^{18}$ evaluations to guarantee similar accuracy!

Although this performance is impressive in its own right, it is worth emphasising that it took mere minutes to adapt an existing GA-based optimiser[5] to solve this problem, and that the algorithm control parameters were not optimised in any way for this new problem.

An obvious question arises: *why* do GAs work as well as they do? This topic is far beyond the scope of this survey paper but suffice it to say that a universally-accepted explanation has not yet been developed. Holland's famous *Schema Theorem* has long been touted as providing an explanation for GAs' success [2], although more recently it has become apparent that this theorem provides insight only into the workings of simplistic GAs; and even then, it is not clear whether the assumptions underlying the theorem are tenable [8, 9].

Despite all their attractive features, GAs also have their share of disadvantages (more or less in accordance with Wolpert and Macready's famous "no free lunch" theorem [10]). GAs might be called "Jacks of all problems, but masters of none": optimising a GA's performance on a given problem is often difficult or impossible, and in order to achieve near-optimal performance it is usually necessary to hybridise a GA with problem-specific heuristics. For example, they tend to be better at locating than at fine-tuning solutions: once a GA is in the vicinity of a global optimum, it is usually a good idea to let a local optimiser take over [1]. GAs can be inefficient on simple problems where the computational expense of applying the genetic operators outweighs that of evaluating the function to be optimised; conversely, on problems where each cost function evaluation is extremely expensive – for example, where each evaluation requires a long simulation to be run – a GA-based (or indeed any) forward modelling approach could be impractical.

Finally the (currently) limited theoretical understanding of GAs is regarded by some, quite

[5] This GA, coded by the author, used floating-point encoding, dynamically-adjusted mutation rates and tournament-style selection of reproducing partners.

understandably, as a drawback and this might explain their relatively slow uptake in the physical sciences [7].

4. Applications: astronomy and astrophysics

This section presents a sample of the numerous and diverse applications that genetic algorithms have found in astronomy and astrophysics. For brevity's sake, only one or two short but representative examples have been drawn from different subfields.

Astrophysical dynamics. Wahde and Donner developed a method for reliably determining the orbital parameters of interacting galaxies and applied their method to both artificial and real data [11]. Their method is based on a GA that searches very efficiently through the large space of possible orbits; indeed, the authors argue that GAs are ideally suited for investigations of tidally interacting galaxies, where large multimodal spaces must be searched in order to constrain a large number of model parameters. Cantó *et al.* devised an interesting variant of the canonical GA which they applied successfully to various problems, including the challenging task of finding the orbital parameters of the planets orbiting 55 Cancri, based on radial velocity measurements of the aforesaid stellar system [12].

Physical and observational cosmology. Although Monte Carlo methods seem to predominate in cosmology, GAs have already found a number of applications in the field. To mention just a few: Nesseris and Shafieloo used GAs to reconstruct the expansion history of the universe in a model-independent manner and thence, in conjunction with the so-called *Om statistic*, they derived a null test on the cosmological constant model ΛCDM [13]; via GAs, Allanach *et al.* were able to answer some important questions related to the discrimination of SUSY-breaking models, and in particular to quantify the measurements necessary to tell different SUSY-breaking scenarios apart [14]; and Bogdanos and Nesseris used GAs to analyse Type Ia SNe data and to extract model-independent constraints on the evolution of the dark energy equation of state [15]. The latter authors note that as a non-parametric method, GAs provide a convenient model-independent platform for cosmological data analysis that can minimise bias due to premature choice of e.g. a dark energy model.

Gravitational lens modelling. Gravitational microlensing is an ideal technique for probing the galactic population of faint or dark objects such as substellar objects, stellar remnants and exoplanets. Though very successful, theoretical microlensing models tend to be complex and their associated inverse modelling problems are notoriously difficult. The author of this paper has recently been developing GAs to speed up this difficult modelling, with a view to being able to model ongoing events approximately in real time (i.e. on a timescale of minutes rather than weeks or months!) and thereby to facilitate better-informed observations and thus more useful observational data. Results of this work are expected to be published in early 2012. As another example, Liesenborgs *et al.* presented a GA-based, non-parametric technique for inferring the projected lensing-mass distributions in strongly lensed systems [16].

Stellar spectrum fitting. Performing fits to stellar spectra is a nontrivial but important undertaking; from fitted models one can infer a veritable multitude of stellar properties. Baier *et al.* were able to combine radiative transfer codes with a GA to produce an automated procedure for fitting the dust spectra of AGB stars. Their GA-based routine dramatically improved extant fits made with more traditional methods and provided a quantitative platform from which to compare different models [17]. In a similar vein, Mokiem *et al.* used a parallelised GA as the basis for an autonomous fitter of spectra of massive stars with stellar winds [18].

Stellar structure modelling. Metcalfe and Charbonneau implemented a highly-parallelised and distributed GA to determine the globally optimal parameters for stellar models. The efficient, parallel exploration of parameter space made possible by their GA-based optimisation led to some important results in the field of white dwarf astroseismology, including the unexpected resolution of a then-puzzling discrepancy between stellar evolution model and astroseismic

inferences of He-layer masses in DBV white dwarfs [19].

Telescope scheduling. Autonomous telescope scheduling is a difficult task that requires dynamic adjustment of numerous observational constraints whilst trying to ensure the efficient achievement of many different scientific objectives. Kubanek developed an easy-to-implement yet robust approach to a robotic telescope scheduling problem, based on a GA that seeks out Pareto-optimal solutions (telescope schedules) [20].

5. Conclusions

This paper introduced genetic algorithms, mentioned some of their strengths (and weaknesses) and finally illustrated their utility in astronomy and astrophysics. For those who would like to learn more about GAs or other evolutionary algorithms, there are many fine books and papers on the subject: to mention just a few, Michalewicz's book [4] gives an excellent introduction with a theoretical leaning, and Haupt's book provides an equally good though more "hands-on" treatment [3]. Goldberg's seminal book [21], one of the most widely-cited works in all of computer science, serves as an outstanding tutorial-style reference, and finally Charbonneau's note [1] provides a straightforward discussion of how standard statistical methods can be used to construct confidence intervals for GA-estimated model parameters.

The author of this paper would welcome correspondence from anyone who would like to discuss evolutionary algorithms, perhaps with a view to applying them in their own work.

Acknowledgments

The author is grateful to the University of Cape Town and the National Research Foundation for the provision of financial support, and also to the anonymous referees for their helpful comments.

References

[1] Charbonneau P 2002 *An Introduction to Genetic Algorithms for Numerical Optimization* (Boulder, CO: National Center for Atmospheric Research)
[2] Holland J H 1975 *Adaptation in Natural and Artificial Systems* (Ann Arbor, MI: The University of Michigan Press)
[3] Haupt R L and Haupt S E 2004 *Practical Genetic Algorithms* (Toronto: Wiley-Interscience)
[4] Michalewicz Z 1996 *Genetic Algoritms + Data Structures = Evolution Programs* (Berlin, AL: Springer)
[5] Wright A H 1991 *Foundations of Genetic Algorithms* ed Rawlins G J (San Mateo, CA: Morgan Kaufmann) pp 205–18
[6] Eiben A E, Aarts E H L and van Hee K M 1991 *Proc. 1st Workshop on Parallel Problem Solving from Nature* (London: Springer-Verlag) pp 4–12
[7] Charbonneau P 1995 *Astrophys. J. Suppl. Ser.* **101** 309–34
[8] Syswerda G 1989 *Proc. 3rd Int. Conf. on Genetic Algorithms* ed Schaffer D J (San Mateo, California: Morgan Kaufmann Publishers, Inc.) pp 2–9
[9] Wright A H, Vose M D and Rowe J E 2003 *Proc. 2003 Int. Conf. on Genetic and Evolutionary Computation: Part II* (Berlin, Heidelberg: Springer-Verlag) pp 1505–17
[10] Wolpert D H and Macready W G 1997 *IEEE Trans. Evol. Comp.* **1** 67–82
[11] Wahde M and Donner K J 2001 *Astron. Astrophys.* **379** 115–24
[12] Cantó J, Curiel S and Martínez-Gómez E 2009 *Astron. Astrophys.* **501** 1259–68
[13] Nesseris S and Shafieloo A 2010 *Mon. Not. R. Astron. Soc.* **408** 1879–85
[14] Allanach B C, Grellscheid D and Quevedo F 2004 *J. High Energy Phys.* **7** JHEP07(2004)069
[15] Bogdanos C and Nesseris S 2009 *J. Cosmol. Astropart. Phys.* **5** JCAP05(2009)006
[16] Liesenborgs J, De Rijcke S and Dejonghe H 2006 *Mon. Not. R. Astron. Soc.* **367** 1209–16
[17] Baier A, Kerschbaum F and Lebzelter T 2010 *Astron. Astrophys.* **516** A45
[18] Mokiem M R, de Koter A, Puls J, Herrero A, Najarro F and Villamariz M R 2005 *Astron. Astrophys.* **441** 711–33
[19] Metcalfe T S, Nather R E and Winget D E 2000 *Astrophys. J.* **545** 974–81
[20] Kubanek P 2008 *Genetic Algorithm for Robotic Telescope Scheduling* (Granada: University of Granada Press)
[21] Goldberg D E 1989 *Genetic Algorithms in Search, Optimization and Machine Learning* 1st ed (Boston, MA: Addison-Wesley Longman Publishing Co., Inc.)

Atmospheric turbidity over Soweto

T Sethabane, H Winkler

Dept. Physics, University of Johannesburg, PO Box 524, Auckland Park, 2006

hwinkler@uj.ac.za

Abstract. We have analysed broad-band global solar irradiation measurements collected at the Soweto campus of the University of Johannesburg over the two-year period 1998-1999. We estimate the direct and diffuse irradiation using radiation models, and utilise the results to determine the atmospheric turbidity for all cloud-free days. From this we calculate aerosol densities and the solar radiation potential for the location. We further evaluate the applicability of some irradiation models, incorporating molecular absorption, Rayleigh and Mie scattering, to this part of Gauteng.

1. Introduction

Solar radiation is abundant in South Africa and has been identified as an important source of energy. Optimal use of this resource will mitigate South Africa´s energy shortage and lessen its dependence on coal. The need for accurately quantifying and understanding solar irradiation has never been greater, particularly in the energy intensive Gauteng province, where a proliferation of smaller solar energy devices is likely to occur in the coming years.

There have been several attempts to quantify the solar radiation potential in South Africa [1, 2, 3]. These were, however, generally based on the interpolation of values between rather distant measurement sites, or the result of satellite image analysis that is often affected by calibration uncertainties. This work proposes to develop a tool for obtaining better irradiation estimates in parts of Gauteng, through the analysis of hitherto unutilized solar data.

2. Theory

The amount of solar radiation reaching a solar device on the Earth´s surface critically depends on the orientation of the device relative to the Sun´s position and the effect of the atmosphere on the radiation. The latter strongly depends on the wavelength of the incoming light.

Cloud cover constitutes the most obvious example of an atmospheric effect, but the degree of aerosol loading and specific gas concentrations (e.g. ozone, water vapour) are also important. But even a 'clean' atmosphere will scatter some of the solar radiation through the Rayleigh scattering mechanism. The multitude and erratic nature of all these atmospheric effects makes it extremely difficult to devise models that replicate solar irradiation accurately. Most studies therefore use empirical models instead. These models are usually based on the theory developed by Linke [4].

2.1. Some basic principles

The airmass m quantifies the amount of atmosphere that light from the Sun needs to traverse to reach the light collection area. It is measured in terms of the amount of atmospheric particles traversed until sea level is reached by a vertical ray of light (i.e. for such a ray, $m = 1$). Let us first define the quantity

$$\mu = \sec z = (\cos z)^{-1} \ , \tag{1}$$

where z is the zenith angle, i.e. the angle between the solar position and the zenith.

If we neglect Earth curvature effects and refraction by the atmosphere (which only become significant when z approaches 90°), then the airmass may be estimated by $m \sim p\mu$, where p is the relative pressure at the site (in units of standard atmospheric pressure).

The global radiation G refers to the total radiation recorded on a horizontal surface. That consists of the direct radiation I received from the direction of the Sun in addition to the so-called diffuse radiation D, i.e. the light from the rest of the sky (effectively scattered sunlight). The direct radiation needs to be multiplied by cos z to account for the misalignment between the solar beam and the horizontal collecting surface. Thus

$$G = I \cos z + D \cong I/\mu + D \ . \tag{2}$$

The diffuse radiation depends on the solar position, the angular displacement of each sky patch from the Sun, and the aerosol type, particle size and concentration. This function is generally extremely complex to model, and a variety of approximations have been suggested. In view of the limited information available from our data, we chose to model this component using the hitherto unused formulation

$$D = cm^{-b} + a \ , \tag{3}$$

where a corresponds to the global radiation recorded immediately before sunrise and immediately after sunset. The power law term cm^{-b} was used in view of its ability to reasonably describe the diffuse radiation curve, which reaches its maximum when the Sun is at its highest, and declines as it approaches the horizons. The constant c is then expected to be an indicator of the atmospheric particle concentration.

2.2. Turbidity

Direct solar radiation is subject to attenuation as it passes through the atmosphere. Monochromatic radiation would experience an exponential weakening, where the exponent term is proportional to the traversed atmosphere (i.e. the airmass) and the so-called opacity (which is directly proportional to the particle concentration). In the case of broad-band irradiation (i.e. all visible irradiation irrespective of wavelength), however, a correction term δ is required to characterise deviations from the pure exponential form caused by the wavelength dependence of opacity. Using empirical data, Linke [4] found that the broad-band direct radiation is well described by the relation

$$I = I_0 \exp(-m\Lambda\delta(m)) \ . \tag{4}$$

The function $\delta(m)$ represents the direct radiation attenuation for an aerosol-free atmosphere. It has been modeled by a variety of researchers, with Ineichen and Perez [5] concluding that the Kasten [6] formulation

$$\delta(m) = (9.4 + 0.9m)^{-1} \tag{5}$$

was most accurate near m = 2. Λ is referred to as the Linke turbidity factor, and amounts to a descriptor of the effective number of aerosol-free atmospheres that would yield the measured attenuation. So Λ would be 1 for a completely clean atmosphere, but can be as high as 10 for very polluted skies.

Because of the dependence of the opacity on wavelength the Linke turbidity is itself slightly airmass-dependent. Hence Λ_2, which is the Linke turbidity at airmass 2, has been suggested as a standard descriptor of turbidity. In order to then make the turbidity compatible with $\delta(m)$, an additional factor is inserted into the exponential function, determined to be 0.8662 [7]. Hence

$$I = I_0 \exp\left(-0.8662\Lambda_2 \delta(m) m\right)$$

$$\Rightarrow \ln I = \ln I_0 - 0.8662\Lambda_2 \delta(m) m \ . \tag{6}$$

If we now apply equation 5, and calculate the derivative of equation 6, we find that

$$\frac{\partial}{\partial m}(\ln I) = -0.8662\Lambda_2 \left((9.4 + 0.9m)^{-1} - 0.9(9.4 + 0.9m)^{-2} m\right)$$

$$\Rightarrow \frac{\partial}{\partial m}(\ln I) = \frac{-8.14\Lambda_2}{(9.4 + 0.9m)^2} \ . \tag{7}$$

It then follows that

$$\Lambda_2 = -15.406 \times \left(\frac{\partial(\ln I)}{\partial m}\right)_{m=2} . \tag{8}$$

3. Measurements and calculations

3.1. Instrumentation and data description

Measurements were carried out at the University of Johannesburg´s Soweto campus throughout most of 1998 and 1999. Broad-band global irradiation measurements were obtained with a LI-200SA pyranometer mounted horizontally. Readings were collected once a minute on an ongoing basis.

The site is at an altitude of 1640 m above sea level. It is occasionally affected by smoke from domestic fires, or even by dust from gold mine dumps at the periphery of Soweto. We adopted a mean atmospheric pressure of 0.823 p_0 for the site, i.e. $p = 0.823$

3.2. Procedure for estimating the Linke turbidity

On a number of days (referred to hereafter as the calibration days) the diffuse radiation was determined periodically by shading the detector for a few minutes. An example of the light curve on such a day is illustrated in Figure 1 below.

A function of the form $D = cm^{-b} + a$ was then fitted to the diffuse readings on the calibration days. The direct irradiation I was thereafter determined after subtracting the fitted function D from the global measurements G.

In view of equation 6, ln I is expected to be a smooth function of m. Indeed, in the range $2 < m < 6$ it is always very well fitted by a polynomial of order 3. In particular, we find that the slope of this polynomial function matches equation 7 particularly well at $m = 2$. We are therefore able to calculate Λ_2 directly by substituting the value of the derivative of the polynomial function at $m = 2$ into equation 8.

We finally determine the relationship between the power law coefficient c and Λ_2, as well as between b and c, by plotting these values obtained on the calibration days against each other. We find that the first of these relationships is well fitted by the expression $c = 16.6 \times \Lambda_2$. We also note an interdependence between b and c, which we approximate by the linear relationship $c = 103.2b + 21.9$.

Therefore equation 3 becomes

$$D = 16.6\Lambda_2 m^{-0.161\Lambda_2 + 0.212} + a \ . \tag{9}$$

a, the sky brightness with the Sun just below the horizon, may be measured directly for each day. So on a typical day when only measurements of G are made, D and I can be estimated with equation 9 if we are able to approximate Λ_2.

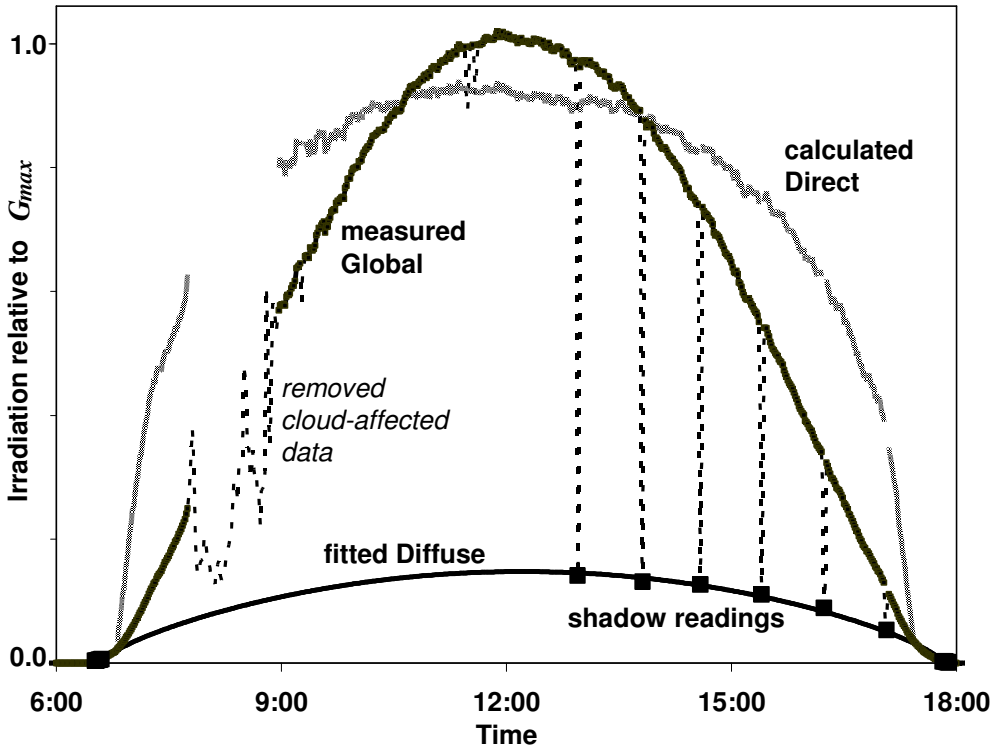


Figure 1. Illustrative example (17 August 1998): The measured irradiation data shows a period of cloud in the morning and diffuse radiation measurements (by casting a shadow on the detector) in the afternoon. The global curve was obtained by removing cloud and shadow affected points from the measured data. The diffuse curve is a fit through the shadow points (according to equation 3) and the direct curve was calculated from the global and diffuse curve by equation 2.

3.3. Daily turbidity measurements

Therefore the following iterative process was used for calculating the daily turbidity. The data for each day was first checked visually to assess the level of cloud cover and cloud interference. Daily turbidity was then estimated for all days with irradiation curves that were largely unaffected by cloud in the airmass span $2 < m < 6$. An initial estimate of Λ_2 is made, and the corresponding diffuse irradiation is then determined using equation 9. The diffuse irradiation estimate is then subtracted from the global measurements to calculate the estimated direct irradiation. As earlier, a polynomial fit to the ln I vs m relationship is used to compute the derivative of that function at $m = 2$, from which a revised estimate of Λ_2 is obtained through equation 8. This process is repeated a few times until the Λ_2 estimates converge.

In this manner, Linke turbidities were calculated for all largely clear days up to two times (for the morning and for the afternoon). We found that these ranged from $\sim 2 < \Lambda_2 < \sim 5$, and the monthly averages were determined to be as shown in Figure 2.

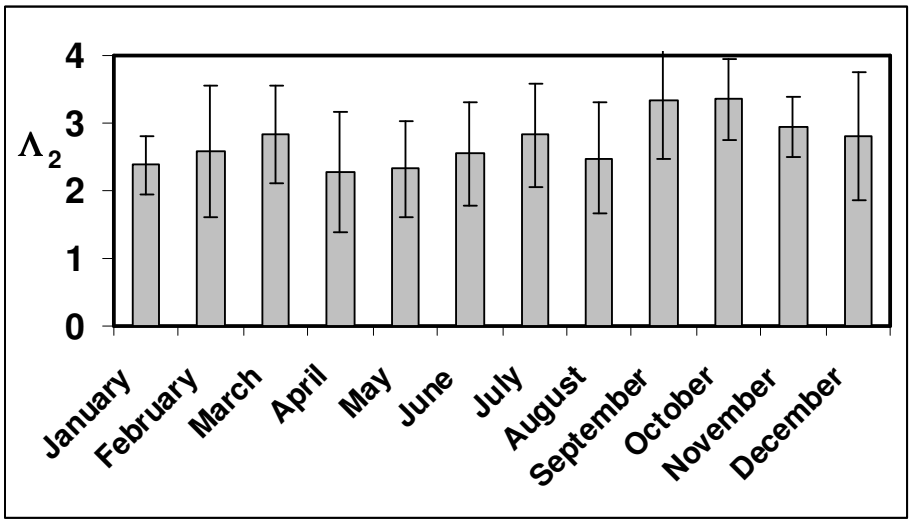


Figure 2. Monthly Linke turbidity coefficient averages for the period 1998-1999

4. Discussion

4.1. Monthly turbidity averages

Figure 2 highlights a pattern typical for the Highveld region of South Africa. Turbidity increases towards the end of the winter dry season. During this time fires are common and smoke only dissipates slowly, leading to enhanced aerosol concentrations over the whole of the southern African sub-continent [8]. We point out that the summer average turbidities were calculated from relatively few days (as totally clear days are rare over that period), and are therefore less certain. We also note that quite a few winter days with extreme turbidity caused by smoke from domestic fires would not have been represented in the calculation of the averages, as smoke may mimic clouds. The true local winter average turbidities will therefore be higher than indicated here.

4.2. Data validation and application to other sites

Satellite data-based models of the solar irradiation also enable the estimation of the average turbidity through the use of the PVGIS package [9]. These are provided in table 1 together with the corresponding values found from our measurements.

Table 1. Average airmass $m = 2$ Linke turbidities for (I) the 1998-1999 Soweto measurements and (II) derived through PVGIS. The last row indicates the difference between the two.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year
$\Lambda_{2,I}$	2.39	2.58	2.83	2.28	2.32	2.55	2.82	2.48	3.33	3.35	2.94	2.80	2.72
$\Lambda_{2,II}$	2.8	2.6	2.7	2.9	2.6	2.6	2.8	2.6	3.8	3.0	3.0	2.8	2.8
$\Delta\Lambda_2$	–0.4	+0.0	+0.1	–0.6	–0.3	+0.0	+0.0	–0.1	–0.5	+0.4	–0.1	+0.0	–0.1

When we compare the PVGIS monthly averages for our site to the values displayed in figure 2, we find that they are similar, with the slightly lower overall average being explained by the omission of data from exceptionally turbid days. This suggests that, even though one of these data sets refers to 2 years only, the satellite-based turbidity estimates and the method employed in this work are

compatible, at least in conditions applicable to Gauteng. The analysis should however be repeated with data from other sites to confirm that the results are indeed equivalent.

4.3. An atmospheric model for Gauteng?

We conclude that, outside episodes of sub-continental aerosol enhancement due to biomass burning, turbidity over Gauteng is low, at least compared to sites in Europe, where Λ_2 is usually above 3. This low turbidity may in part be related to the high altitude. Thus aerosols play a less significant role in the attenuation of solar radiation. A comparatively simple Rayleigh scattering model with a minor Mie scattering component for the aerosols adequately describes such conditions. Local deviations from this are however expected in the vicinity or downwind from areas characterized by excessive domestic fires, mine dumps or industrial areas. Here more complex models incorporating air flow and aerosol scattering characteristics will be required.

Acknowledgements

The study made use of data obtained from the Photovoltaic Graphical Information Systems (PVGIS) internet resource, which is maintained by the European Research Commission.

References

[1] Eberhard A 1990 *A Solar Radiation Handbook for Southern Africa*, Elan Press, Cape Town
[2] Power H C and Willmott C J 2001 *International Journal of Climatology* **21** 579
[3] Tsubo M and Walker S 2003 *S. Afr. J. Sci.* **99** 360
[4] Linke F 1922 *Beitr. Phys. fr. Atmos.* **10** 91
[5] Ineichen P and Perez R 2002 *Solar Energy* **73** 151
[6] Kasten F 1980 *Meteor. Runds.* **33** 124
[7] Rigollier C, Bauer O and Wald L 2000 *Solar Energy* **68** 33
[8] Swap R J, Annegarn H J, Suttles T, King M D, Platnick S, Privette J L and Scholes R J 2003 *J. Geophysical Research* **108**, D13, 1
[9] European Research Commission, *Photovoltaic Graphical Information Systems (PVGIS)*, http://re.jrc.ec.europa.eu/pvgis/

Constraining viewing geometries of pulsars with single-peaked γ-ray profiles using a multiwavelength approach

A S Seyffert[1], C Venter[1], T J Johnson[2,3] and A K Harding[2]

[1] Centre for Space Research, North-West University, Potchefstroom Campus, Private Bag X6001, Potchefstroom 2520, South Africa

[2] Astrophysics Science Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

[3] Department of Physics, University of Maryland, College Park, MD 20742, USA

E-mail: 20126999@nwu.ac.za

Abstract. Since the launch of the Large Area Telescope (LAT) on board the *Fermi* spacecraft in June 2008, the number of observed γ-ray pulsars has increased dramatically. A large number of these are also observed at radio frequencies. Constraints on the viewing geometries of 5 of 6 γ-ray pulsars exhibiting single-peaked γ-ray profiles were derived using high-quality radio polarization data [1]. We obtain independent constraints on the viewing geometries of 6 by using a geometric emission code to model the *Fermi* LAT and radio light curves (LCs). We find fits for the magnetic inclination and observer angles by searching the solution space by eye. Our results are generally consistent with those previously obtained [1], although we do find small differences in some cases. We will indicate how the γ-ray and radio pulse shapes as well as their relative phase lags lead to constraints in the solution space. Values for the flux correction factor (f_Ω) corresponding to the fits are also derived (with errors).

1. Introduction

We study 6 pulsars (J0631+1036, J0659+1414, J0742−2822, J1420−6048, J1509−5850, and J1718−3825) detected by *Fermi* LAT, all exhibiting single-peaked pulse profiles within current statistics at energies > 0.1 GeV. These pulsars are also detected in the radio band. Both of these properties aid in constraining the possible solution space for the respective pulsar geometries, as derived from predictions of their light curves (LCs).

A study of the pulsars' geometric parameters, the inclination and observer angles α and ζ, has been performed [1] using the radio polarization and LC data. They constrained the solution space for α and β of 5 of the individual pulsars, with the impact angle $\beta = \zeta - \alpha$, using fits to these radio data as well as predictions for the value of the half opening angle, ρ, of the radio beam derived from the radio pulse width (e.g., [2]).

The aim of this study is to obtain similar constraints on α and ζ of the same pulsars using an independent, multiwavelength approach. We use a geometric pulsar emission code (e.g., [3]) to model the *Fermi* LAT and radio LCs, and fit the predicted radio and γ-ray profiles to the data concurrently, thereby significantly constraining α and ζ. This also allows us to test the consistency of the various approaches used to infer the pulsar geometry.

2. Model

We use an idealized picture of the pulsar system, wherein the magnetic field has a retarded dipole structure [4] and the γ-ray emission originates in regions of the magnetosphere (referred to as 'gaps') where the local charge density is sufficiently lower than the Goldreich-Julian charge density [5]. These gaps facilitate particle acceleration and radiation. We assume that there are constant-emissivity emission layers embedded within the gaps in the pulsar's corotating frame. The location and geometry of these emission layers determine the shape of the γ-ray LCs, and there exist multiple models for the geometry of the magnetosphere describing different possible gap configurations.

We included two models for the γ-ray emission regions in this study, namely the Outer Gap (OG) and Two-Pole Caustic (TPC) models. In both the OG and TPC models (the Slot Gap model [6] may serve as physical basis for the latter), emission is produced by accelerated charged particles moving within narrow gaps along the last open magnetic field lines. In the OG model [7], radiation originates above the null charge surface (where the Goldreich-Julian charge density changes sign) and *interior to* the last open magnetic field lines. The TPC gap starts at the stellar surface and extends *along* the last open field lines up to near the light cylinder[1], where the corotation speed approaches the speed of light [8]. The special relativistic effects of aberration and time-of-flight delays, which become important in regions far from the stellar surface (especially near the light cylinder), together with the curvature of the magnetic field lines, cause the radiation to form caustics (accumulated emission in narrow phase bands). These caustics are detected as peaks in the observed γ-ray LC [8, 9].

We used an empirical radio cone model [10], where the beam diameter, width, and altitude are functions of the pulsar period P, its time derivative $\dot{P}$, and the frequency of observation ν, in order to obtain predictions for the radio LCs. This is different from the approach taken by [1], as we have a different prescription for the radio emission altitude and ρ, and we do not use the polarization data.

3. Method

Using the above geometric models, we generated LCs as a function of α and ζ, keeping the gap widths and extents constant. Due to the large size of the (α,ζ)-space at a 1° resolution, it is impractical to search blindly for the optimum LC by eye. We therefore started by generating a so-called atlas of LCs with a 10° resolution for each pulsar and first identified candidate LC fits. We then performed refined searches in the regions of (α,ζ)-space around those candidate fits at a 5° and later 1° resolution. This approach of generating all the LCs in a candidate region and comparing them directly, allowed us to obtain constraints on α and ζ. We then inferred contours for α and ζ for each pulsar and for each geometric model. We lastly used these contours to find the corresponding values of f_Ω (with errors; see Section 3.2) predicted by the OG and TPC models.

3.1. Obtaining the contours

To obtain an (α,ζ)-contour delineating the region of plausible LC fits, we start at the candidate solution (5° resolution) and first move away from it in steps of 1° in α. At each step we then step through ζ until the upper and lower bounds of the contour are located for the fixed α. This procedure is repeated until no further acceptable LCs are found. As an example, the red regions **A** and **B** in the left panel of Figure 1 indicate the resulting contours obtained in this manner for PSR J1509−5850 for the TPC case.

[1] The slight difference in transverse polar position of the gaps with respect to the magnetic axis (as motivated by physical models) will result in a systematic shift in best-fit (α,ζ)-contours between the two models. See Section 4.

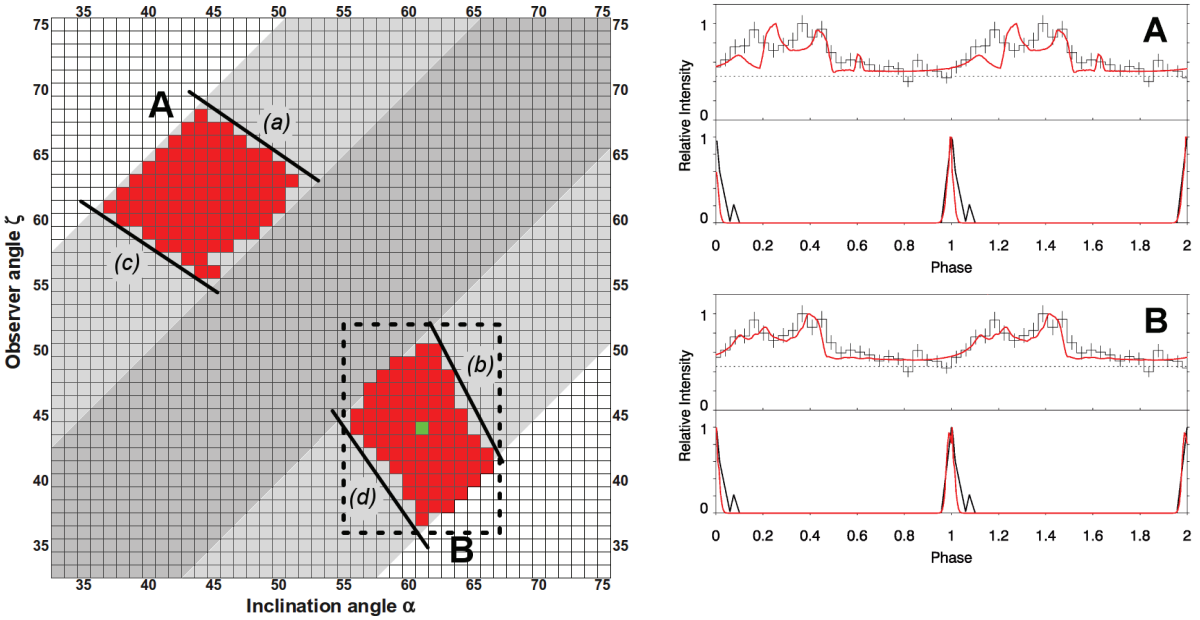


Figure 1. *Left:* The (α,ζ)-contours (regions **A** and **B**) for radio and γ-ray LC fits of PSR J1509−5850 for the TPC model (*red*). The dark grey and light grey regions correspond to double-peaked and single-peaked radio LCs respectively. The solid lines indicate approximately the locations of the boundaries to the contours due to: (a) the relative phase lag between the radio and γ-ray peaks being too large, (b) the bridge emission's relative intensity being too low, and (c,d) the γ-ray peak being too narrow. The dashed line (bounded area) indicates the resulting errors on α and ζ, with the *green* block indicating our best fit. *Right:* Representative LC fits corresponding to the contours **A** and **B**.

Oftentimes we find two disjoint solution contours, with one of the two sometimes producing a better LC fit. The example shown here is one of these cases, as can be seen from the representative LCs shown in Figure 1. Contour **B** yields a considerably better fit than contour **A**, and thus we can ignore contour **A** in favour of contour **B** as indicated by the dotted line in Figure 1. The values of α and ζ reported in Table 1 are those at the centre of this dotted box, while the errors on these values are conservatively chosen to encompass the dimensions of the box. In this case the resulting values are $\alpha = (61^\circ \pm 5^\circ)$ and $\zeta = (44^\circ \pm 7^\circ)$.

3.2. Finding f_Ω

It is important to be able to convert the observed energy flux of a pulsar to its all-sky luminosity. The flux correction factor f_Ω is used for this purpose. It is a highly model-dependent parameter, and allows us to determine what fraction of the all-sky luminosity we are observing from a pulsar if the geometry is known (α and ζ), for a given model. It is also crucial in deriving the γ-ray efficiency of a pulsar, and is defined as [11]

$$f_\Omega(\alpha, \zeta_E) = \frac{\iint F_\gamma(\alpha, \zeta, \phi) \sin\zeta d\zeta d\phi}{2 \int F_\gamma(\alpha, \zeta_E, \phi)\, d\phi}, \quad (1)$$

with F_γ being the photon flux per solid angle ('intensity'), and ζ_E the Earth line-of-sight. The value of f_Ω is typically taken to be 1, meaning that the observed energy flux is assumed to be equal to the average energy flux over the entire sky. Due to the OG model predicting emission over a relatively smaller region of phase space, its associated f_Ω values are typically smaller

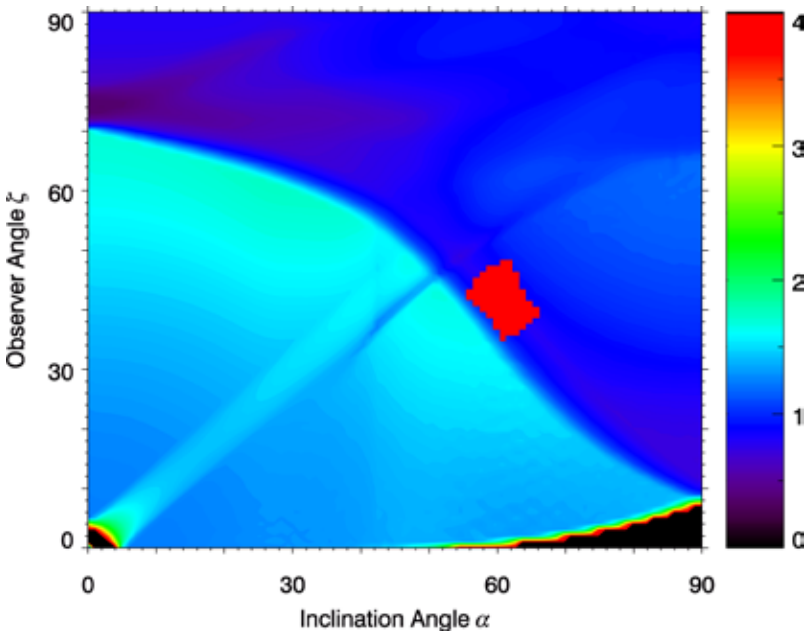


Figure 2. Contour plot of f_Ω as function of α and ζ implied by the TPC model for PSR J1509-5850. The *red* area indicates the best-fit (α, ζ)-contour for the radio and γ-ray LCs.

than those of the TPC model for the same α and ζ. (Note, we assume that $F_\gamma(\zeta, \phi)/F_{\gamma,\mathrm{tot}}$ is approximately equal to the ratio of observed energy flux vs. total energy flux.)

We can now use the (α,ζ)-contours obtained by the method discussed in Section 3.1 to constrain the value of f_Ω by computing $f_\Omega(\alpha, \zeta)$ and overplotting those contours (see Figure 2).

4. Results

We infer values for α and ζ for each pulsar with typical errors of $\sim 5^\circ$ (see Table 1). The tabulated values are the average values of α, ζ, and β implied by the solution contours, while the errors are chosen conservatively so as to include the full (non-rectangular) contour. We note that once a solution is found at a particular α and ζ, it is worthwhile to study the LC at the position $\alpha' = \zeta$ and $\zeta' = \alpha$ (i.e., when the angles are interchanged). Such a complementary solution may provide a good fit in some cases (see the alternative solutions listed in Table 1), due to the symmetry of the model as well the symmetry in pulsar geometry (i.e., $|\beta|$ remaining constant under this transformation).

As mentioned in Section 2, we see a shift in best-fit contours, with the TPC model consistently yielding smaller α and ζ values compared to those obtained using the OG model due to a different gap geometry. This is because similar phaseplots are obtained when a larger α is chosen for the OG model than for the TPC model. Reproduction of the radio pulse shape then necessitates a larger value of ζ for the OG fit, since the contour is constrained to appear in one of the grey regions (as determined by peak multiplicity; see Figure 1). The net effect of this is that the OG contours are found farther from the origin at $\alpha = \zeta = 0$ along the grey bands than the corresponding TPC contours.

Our results for f_Ω for each of the pulsars (and models) are given in Table 1.

5. Discussion and Conclusions

The good constraints on α and ζ, obtained using the method described above, emphasizes the merit of a multiwavelength approach. For example, the light grey regions in Figure 1 indicate single-peaked radio profile LC solutions. These would represent typical constraints one would be able to derive for α and ζ when considering only the radio profile shapes. Similarly, when only considering the (single-peaked) γ-ray LCs, one would find relatively large (α,ζ)-contours [11]. The requirement of fitting both the radio and γ-ray profile shape, as well as their relative phase lag, results in much smaller contours, as shown in Figure 1.

Generally, our best-fit (α,ζ) compare favourably with those inferred by [1] for the first three pulsars (see Table 1). However, comparison is hampered by uncertainties in estimating the half opening angle, ρ, which sensitively influences the optimal solutions obtained by [1]. Even a small error of 5° on ρ leads to relatively large errors on the allowed α, so that our best fits would then be included in their inferred parameter ranges. Direct comparison of our results with those of [1] for PSR J1420−6048 and PSR J1718−3825 is not feasible due to different approaches for estimating the radio pulse width, W. We only model the most prominent radio peak, while [1] measures W as the width of the total two-peaked profile. Implementing a factor 2 difference in W will change the ρ-contours and thus their best-fit solution considerably, improving the agreement between the two approaches. Comparison with [1] is not possible for PSR J1509-5850 as the lack of polarization data inhibited inference of a best-fit (α,β) by [1] in this case.

The next step for this kind of study is to apply a mathematically rigorous method of determining the best-fit LC solutions as a function of many free model parameters to the six pulsars modelled in this paper. An example is the Markov chain Monte Carlo method which has been successfully applied to millisecond pulsar LCs [12].

Table 1. Values for α, ζ, β, and f_Ω derived for 6 pulsars for both the OG and TPC models. The *'s indicate alternate best-fit solutions upon interchange of α and ζ.

Pulsar	Model	α (°)	ζ (°)	β (°)	f_Ω
J0631+1036	OG	74 ± 5	67 ± 4	$-\ 6 \pm 2$	0.93 ± 0.06
	TPC	71 ± 6	66 ± 7	$-\ 5 \pm 3$	1.04 ± 0.04
J0659+1414	OG	59 ± 3	48 ± 3	-12 ± 5	1.16 ± 0.53
	TPC	50 ± 4	39 ± 4	-13 ± 6	1.64 ± 0.04
	TPC*	38 ± 1	50 ± 4	11 ± 4	1.63 ± 0.05
J0742−2822	OG	86 ± 3	71 ± 5	-16 ± 6	0.99 ± 0.10
	OG*	71 ± 6	86 ± 4	16 ± 6	0.81 ± 0.09
	TPC	64 ± 8	80 ± 4	15 ± 6	0.88 ± 0.41
J1420−6048	OG	67 ± 5	45 ± 7	-22 ± 9	0.77 ± 0.13
	TPC	64 ± 6	43 ± 8	-21 ± 9	0.90 ± 0.10
	TPC*	42 ± 5	63 ± 5	21 ± 9	0.77 ± 0.06
J1509−5850	OG	66 ± 4	50 ± 7	-18 ± 8	0.77 ± 0.11
	TPC	61 ± 5	44 ± 7	-18 ± 8	0.89 ± 0.10
J1718−3825	OG	67 ± 6	48 ± 6	-19 ± 8	0.76 ± 0.12
	TPC	61 ± 5	43 ± 6	-19 ± 8	0.86 ± 0.07
	TPC*	42 ± 6	62 ± 5	19 ± 7	0.83 ± 0.10

Acknowledgments
CV is supported by the South African National Research Foundation. AKH acknowledges support from the NASA Astrophysics Theory Program. CV, TJJ, and AKH acknowledge support from the *Fermi* Guest Investigator Program as well as fruitful discussions with Patrick Weltevrede.

References
[1] Weltevrede P et al. 2010 *Astrophys. J.* **708** 1426–41
[2] Gil J, Gronkowski P and Rudnicki W 1984 *Astron. Astrophys.* **132** 312–16
[3] Venter C, Harding A K and Guillemot L 2009 *Astrophys. J.* **707** 800–22
[4] Deutsch A J 1955 *Ann. d'Astrophys.* **18** 1–10
[5] Goldreich P and Julian W H 1969 *Astrophys. J.* **157** 869–80
[6] Arons J 1983 *Astrophys. J.* **266** 215–41
[7] Cheng K S, Ho C and Ruderman M 1986 *Astrophys. J.* **300** 500–21
[8] Dyks J and Rudak B 2003 *Astrophys. J.* **598** 1201–6
[9] Morini M 1983 *Mon. Not. R. Astron. Soc.* **202** 495–510
[10] Story S A, Gonthier P L and Harding A K 2007 *Astrophys. J.* **671** 713–26
[11] Watters K P, Romani R W, Weltevrede P and Johnston S 2009 *Astrophys. J.* **695** 1289–1301
[12] Johnson T J, Harding A K and Venter C 2011, in preparation

Constraining the phase space for chameleon dark energy

Muzikayise E. Sikhonde and Amanda Weltman

Astronomy, Cosmology and Gravity Centre, Department of Mathematics and Applied Mathematics, University of Cape Town, Private Bag, Rondebosch, South Africa, 7700

E-mail: skhmuz002@gmail.com, awelti@gmail.com

Abstract. A number of solutions to the dark energy problem have been proposed in literature, the simplest is the cosmological constant Λ. The cosmological constant lacks theoretical explanation for its extremely small value, thus dark energy is more generally modelled as quintessence scalar field rolling down a flat potential. For the quintessence scalar field to be evolving on cosmological scales to day its mass must be of order H_0 , which is the present value of the Hubble constant. A scalar field ϕ whose mass varies with the background energy density was proposed by Khoury and Weltman. This scalar field can evolve cosmologically while having coupling to different matter fields of order unity. Such a scalar field also couples to photons in the presence of an external magnetic field via the ϕF^2 interaction, where F stands for the electromagnetic field strength tensor. The chameleon-photon coupling of this nature causes a conversion of photons to light Chameleon(ϕ) particles and vice versa.

1. Introduction

One of the most intriguing discoveries in modern cosmology is the accelerating expansion of the universe. This acceleration is said to be caused by a fluid component of the universe which has negative pressure, this form of energy density is called dark energy (DE). Cosmological observations such as supernovae luminosity distance measurements [1], and the cosmic microwave background anisotropy [2] suggests that DE forms about 70% of the energy density of the universe. A number of solutions to the dark energy problem have been proposed in literature, the simplest is the cosmological constant Λ. The cosmological constant lacks theoretical explanation for its extremely small value, thus dark energy is more generally modelled as quintessence scalar field rolling down a flat potential [3]. For the quintessence scalar field to be evolving on cosmological scales to day its mass must be of order $H_0 \sim 10^{-33}$eV, with its Compton wavelength $\lambda_c \sim 4$Gpc.

A scalar field ϕ whose mass varies with the background energy density was proposed by Khoury and Weltman [4]. This scalar field can evolve cosmologically while having coupling β_m to different matter fields of order unity. On Earth for example, the mass of ϕ is sufficiently large i.e $\mathcal{O}(1\text{mm}^{-1})$, while on the field its Compton wavelength is typically hundreds of astronomical units (AU). As a result ϕ evades equivalence principle(EP) tests and fifth force constraints from laboratory experiments. The dependence of this scalar fields mass on its environment has given it the name "chameleon scalar field".

Such a scalar field also couples to photons in the presence of an external magnetic field via the ϕF^2 interaction term, where F stands for the electromagnetic field strength tensor [5]. The chameleon-photon coupling of this nature causes a conversion of photons to light chameleon particles and vice versa. These effects are similar to those of Axion-like particles(ALPs) which

interact with light. This conversion needs two photons and one scalar particle in order to take place, as a result it happens when a photon or a scalar field passes through an external electromagnetic field [6].

This paper is organised in the following way, section 2 we discuss the chameleon scalar field theory, in section 3 we look at how to calculate the chameleon-photon system oscillation probability, in section 4 it's conclusions, in section 5 it's future work and section 6 it's acknowledgements followed by references.

2. Chameleon theory

We consider the following chameleon action S adapted from [7],

$$S = \int d^4x\sqrt{-g}\left(\frac{1}{2}M_{Pl}^2R - \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi) - \frac{e^{\beta_\gamma\phi/M_{Pl}}}{4}F^{\mu\nu}F_{\mu\nu}\right) + \int d^4x\left(\mathcal{L}_m(e^{2\beta_m^i\phi/M_{Pl}}g_{\mu\nu},\psi_m^i)\right) \quad (1)$$

where S is the sum of the normalised vacuum Einstein-Hilbert action(S_{EH}) by equation (2), in which R is the Ricci scalar, $M_{Pl} = (8\pi G)^{-1/2}$ is the reduced Planck mass and g is the determinant of the metric tensor $g_{\mu\nu}$ [8].

$$S_{EH} = \int dx^4\sqrt{-g}\frac{1}{2}M_{Pl}^2R, \quad (2)$$

The second and the third terms give the scalar field ϕ action, with a potential $V(\phi)$ which is of a runaway form, the potential of this form is required to have the following characteristics given in [4].

$$\lim_{\phi\to+\infty}V(\phi) = 0, \quad \lim_{\phi\to+\infty}\frac{V_{,\phi}}{V} = 0, \quad \lim_{\phi\to+\infty}\frac{V_{,\phi\phi}}{V_{,\phi}} = 0...,$$

as well as

$$\lim_{\phi\to0}V(\phi) = \infty, \quad \lim_{\phi\to0}\frac{V_{,\phi}}{V} = \infty, \quad \lim_{\phi\to0}\frac{V_{,\phi\phi}}{V_{,\phi}} = \infty...,$$

The scalar field action is as follows,

$$S_\phi = \int dx^4\sqrt{-g}\left(-\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi)\right), \quad (3)$$

and finally we have

$$\mathcal{L}_\gamma = -\frac{e^{\beta_\gamma\phi/M_{Pl}}}{4}F^{\mu\nu}F_{\mu\nu}, \quad (4)$$

and

$$\mathcal{L}_m(e^{2\beta_m^i\phi/M_{Pl}}g_{\mu\nu},\psi_m^i) \quad (5)$$

which are the photon-scalar and matter-scalar interaction Lagrangians respectively. The ψ_m^i are matter fields which couple to ϕ by a conformal coupling of the form $g_{\mu\nu}^i = e^{2\beta_m^i\phi/M_{Pl}}g_{\mu\nu}$ [4], in which β_m^i and β_γ are dimensionless matter-chameleon and photon-chameleon coupling constants respectively. To obtain the equation of motion for ϕ we vary the action S with respect to ϕ, baring in mind that $g_{\mu\nu}^i = e^{2\beta_m^i\phi/M_{Pl}}g_{\mu\nu}$, where i stands for different matter fields, this gives the following equation of motion.

$$\Box\phi = V_{,\phi}(\phi) + \frac{\beta_\gamma e^{\beta_\gamma\phi/M_{Pl}}}{4M_{Pl}}F^{\mu\nu}F_{\mu\nu} + \frac{\beta_m^i}{M_{Pl}}e^{\beta_m^i\phi/M_{Pl}}\rho_m^i \quad (6)$$

Equation (6) can be written in a Klein-Gordon form as $\Box\phi = -\frac{\partial V_{eff}}{\partial\phi}$ where V_{eff} is the effective chameleon potential shown below.

$$V_{eff}(\vec{x},\phi) = V(\phi) + e^{\frac{\beta_m \phi}{M_{Pl}}}\rho_m + e^{\frac{\beta_\gamma \phi}{M_{Pl}}}\rho_\gamma \tag{7}$$

in which $\rho_\gamma = (F^{\mu\nu}F_{\mu\nu})/4 = (|\vec{B}|^2 - |\vec{E}|^2)/2$ is the electromagnetic field Lagrangian density, and ρ_m is the matter density. We assume the fiducial exponential potential used in chameleon dark energy theories given by the following equation [7] and $\kappa > 0$.

$$V(\phi) = M_\Lambda^4 \exp\left[\kappa\left(\frac{M_\Lambda}{\phi}\right)^n\right] \tag{8}$$

which we can approximate by an inverse power law potential if we assume that $M_\Lambda/\phi \ll 1$.

$$V(\phi) \approx M_\Lambda^4\left[1 + \kappa\left(\frac{M_\Lambda}{\phi}\right)^n\right] \tag{9}$$

The effective mass of small fluctuations about the minimum of V_{eff} is given by $m_{eff}(\phi) = \sqrt{\frac{\partial^2 V_{eff}}{\partial\phi^2}}$, which gives

$$m_{eff}(\phi_{min}) = \sqrt{n(n+1)\kappa^{-\left(\frac{1}{n+1}\right)}M_\Lambda^{n+4}\left(\frac{\beta_m\rho_m + \frac{\beta_\gamma B^2}{2}}{nM_\Lambda^{n+4}M_{Pl}}\right)^{\frac{n+2}{n+1}} + \frac{\beta_m^2\rho_m}{M_{Pl}^2} + \frac{\beta_\gamma^2 B^2}{2M_{Pl}^2}}. \tag{10}$$

When we vary the action in (1) with respect to A^μ which is the electromagnetic 4-potential, it can be shown that the photon fields equation of motion is (here $A^\mu = (\psi, \vec{A})$ [7])

$$\frac{\beta_\gamma}{M_{Pl}}\vec{\nabla}\phi \times (\vec{\nabla}\times\vec{A}) + \Box\vec{A} = 0. \tag{11}$$

3. Oscillation probability for the chameleon-photon system

3.1. homogeneous magnetic fields

The Chameleon-Photon oscillation Probability ($P_{\gamma\leftrightarrow\phi}$) is derived by considering the Chameleon-Photon system shown below, which is obtained from equations (6) and (11) adapted from a paper by Raffelt and Stodolsky (1988) [5], where B is the magnetic field strength, ω is the energy, ω_p is the electron plasma frequency of the medium, m_ϕ is the chameleon mass and z is the direction of propagation. The photon and the chameleon fields are indicated by $\vec{\Psi}_\gamma$ and Ψ_ϕ respectively. The photon-chameleon coupling factor is given by $\frac{1}{M} = \frac{\beta_\gamma}{M_{Pl}}$.

$$\left[\omega^2 + \partial_z^2 + \begin{pmatrix} -\omega_p^2 & \frac{B\omega}{M} \\ \frac{B\omega}{M} & -m_\phi^2 \end{pmatrix}\right]\begin{bmatrix} \vec{\Psi}_\gamma \\ \Psi_\phi \end{bmatrix} = 0 \tag{12}$$

The system in equation(12) can be solved by diagonalizing the matrix below

$$A = \begin{pmatrix} \Delta_\gamma & \Delta_M \\ \Delta_M & \Delta_\phi \end{pmatrix} \tag{13}$$

with $\Delta_\gamma = -\omega_p^2$, $\Delta_M = \frac{B\omega}{M}$ and $\Delta_\phi = -m_\phi^2$. Therefore A can be diagonalized by a rotation matrix

$$P = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \tag{14}$$

such that $A = PDP^{-1}$, with

$$D = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}, \tag{15}$$

where $\lambda_{\pm}$ are the eigenvalues of the matrix A given by

$$\lambda_{\pm} = -\frac{(\omega_p^2 + m_\phi^2)}{2} \pm \frac{\sqrt{(\omega_p^2 - m_\phi^2)^2 + 4B^2\omega^2/M^2}}{2} \tag{16}$$

$$= \Lambda \pm \Omega \tag{17}$$

Now we define

$$k_{\pm} = \sqrt{\omega^2 + \lambda_{\pm}} \tag{18}$$

Consider a plane wave solution in the primed fields traveling in the $z-$direction of the following form

$$\begin{bmatrix} \vec{\Psi}'_\gamma(z,t) \\ \Psi'_\phi(z,t) \end{bmatrix} = \begin{pmatrix} e^{i(\omega t - k_+ z)} & 0 \\ 0 & e^{i(\omega t - k_- z)} \end{pmatrix} \begin{bmatrix} \vec{\Psi}'_\gamma(0,t) \\ \Psi'_\phi(0,t) \end{bmatrix} \tag{19}$$

We let the transfer matrix $\mathcal{T}(z,t)$ be calculated as follows

$$\mathcal{T}(z,t) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} e^{i(\omega t - k_+ z)} & 0 \\ 0 & e^{i(\omega t - k_- z)} \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

$$= e^{i\omega t} \begin{pmatrix} [e^{-ik_+ z}\cos^2\theta + e^{-ik_- z}\sin^2\theta] & \frac{1}{2}\sin 2\theta[e^{-ik_+ z} - e^{-ik_- z}] \\ \frac{1}{2}\sin 2\theta[e^{-ik_+ z} - e^{-ik_- z}] & [e^{-ik_+ z}\sin^2\theta + e^{-ik_- z}\cos^2\theta] \end{pmatrix} \tag{20}$$

For propagation we require $k_+^2 > 0$ and $k_-^2 > 0$ such that the chameleon-photon oscillation probability becomes

$$P_{\gamma\leftrightarrow\phi} = \frac{1}{2}\sin^2 2\theta\,[1 - \cos([\Delta k]\, z)] \tag{21}$$

where $\Delta k = k_- - k_+$, and when substituting back for $\sin^2 2\theta$ we obtain the probability as follows;

$$P_{\gamma\leftrightarrow\phi} = \frac{4B^2\omega^2\sin^2([\Delta k]\, z/2)}{M^2(m_\phi^2 - \omega_p^2)^2 + 4B^2\omega^2}. \tag{22}$$

3.2. Inhomogeneous magnetic fields

Considering the perturbative solution of the Schrödinger type equation outlined in [5], we can rewrite equation (12) as follows;

$$i\partial_z\vec{\Psi}(z) = (\mathcal{H}_0(z) + \mathcal{H}_1(z))\vec{\Psi}(z) \tag{23}$$

where $\vec{\Psi}(z) = (\vec{\Psi}_{\gamma\|}(z), \Psi_\phi(z))$,

$$\mathcal{H}_0(z) = \begin{pmatrix} \omega - \frac{\omega_p^2(z)}{2\omega} & 0 \\ 0 & \omega - \frac{m_\phi^2(z)}{2\omega} \end{pmatrix} \tag{24}$$

and

$$\mathcal{H}_1(z) = \begin{pmatrix} 0 & \frac{B(z)}{2M} \\ \frac{B(z)}{2M} & 0 \end{pmatrix} \tag{25}$$

The uncoupled system is given by $\frac{B(z)}{2M} \to 0$, this system has a solution of the form

$$\vec{\Psi}(z) = \mathcal{U}(z)\vec{\Psi}(z_0) \tag{26}$$

where $\mathcal{U}(z) = \exp\left(-i\int_{z_0}^{z}\mathcal{H}_0(z')dz'\right)$. We define the interaction transformation by means of a unitary transformation $\vec{\Psi}_{int}(z) = \mathcal{U}^\dagger(z)\vec{\Psi}(z)$, then by differentiating both sides with respect to

z and using the fact that $\mathcal{U}^{\dagger}(z)\mathcal{H}_0(z)\mathcal{U}(z) = \mathcal{H}_0(z)$, we obtain the following equation for the interacting fields.

$$i\partial_z \vec{\Psi}_{int}(z) = \mathcal{H}_{int}(z)\vec{\Psi}_{int}(z) \tag{27}$$

with $\mathcal{H}_{int}(z) = \mathcal{U}^{\dagger}(z)\mathcal{H}_1(z)\mathcal{U}(z)$, the solution to equation (27) can be found order buy order from the integral equation

$$\vec{\Psi}_{int}(z) = \vec{\Psi}_{int}(z_0) - i\int_{z_0}^{z} \mathcal{H}_{int}(z')\vec{\Psi}_{int}(z')dz' \tag{28}$$

the zeroth-order solution is $\vec{\Psi}^{(0)}_{int}(z) = \vec{\Psi}_{int}(z_0)$, substituting the zeroth-order solution back in equation (28) we can obtain the first-order solution to be

$$\begin{aligned}\vec{\Psi}^{(1)}_{int}(z) &= \vec{\Psi}_{int}(z_0) - i\int_{z_0}^{z} \mathcal{H}_{int}(z')dz'\vec{\Psi}_{int}(z_0) \\ &= \vec{\Psi}_{int}(z_0)\left(1 - i\int_{z_0}^{z} \mathcal{H}_{int}(z')dz'\right)\end{aligned} \tag{29}$$

by substituting equation (29) into the right hand side of equation (28) we obtain the second-order solution as

$$\vec{\Psi}^{(2)}_{int}(z) = \vec{\Psi}_{int}(z_0)\left(1 - i\int_{z_0}^{z} \mathcal{H}_{int}(z')dz' - \int_{z_0}^{z} dz' \int_{z_0}^{z'} dz''\mathcal{H}_{int}(z')\mathcal{H}_{int}(z'')\right). \tag{30}$$

The chameleon-photon oscillation probability form the first-order solution is

$$\begin{aligned}P_{\gamma\leftrightarrow\phi} &= \left|\int_{z_0}^{z} \mathcal{H}_{int}(z')dz'\right|^2 \\ &= \left|\int_{z_0}^{z} \frac{B(z')}{2M}\exp\left(i\int_{z_0}^{z'} \frac{\omega_p^2(z'') - m_\phi^2(z'')}{2\omega}dz''\right)dz'\right|^2\end{aligned} \tag{31}$$

where

$$\mathcal{H}_{int}(z') = \begin{pmatrix} 0 & \frac{B(z')}{2M}\exp\left(-i\int_{z_0}^{z'} \frac{\omega_p^2(z'') - m_\phi^2(z'')}{2\omega}dz''\right) \\ \frac{B(z')}{2M}\exp\left(i\int_{z_0}^{z'} \frac{\omega_p^2(z'') - m_\phi^2(z'')}{2\omega}dz''\right) & 0 \end{pmatrix} \tag{32}$$

4. Conclusion

It is evident from equation (22) that we can calculate the chameleon-photon probability in a region with homogeneous magnetic field, this probability is a function of the coupling constant β_γ, the magnetic field strength B and the propagation distance z. On the other hand we can also calculate the oscillation probability when photons transverse a region with inhomogeneous magnetic fields using equation (31), one can constrain the coupling constant β_γ by predicting the effects of chameleon particle to photon oscillation on the fluxes of astrophysical objects e.g. pulsars and comparing them with observations.

5. Future work

We need to calculate the oscillation probability at different photon frequencies and coupling constant β_γ, then predict the shape of the spectra in order to compare with spectra of astrophysical objects.

Acknowledgments
I would like to thank my supervisor Dr. A. Weltman for her guidance and for proposing this project, and the NRF/SKA scholarship for their funding to make this research possible and the University of Cape Town where most of this work was done.

References
[1] Riess A, et al. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. Astronomical Journal. 1998 September;116(3):1009–1038.
[2] Spergel D, et al. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. The Astrophysical Journal Supplement Series. 2003 September;148:175–194.
[3] Ratra B, Peebles PJE. Cosmological consequences of a rolling homogeneous scalar field. Phy Rev D. 1988 June;37(12):3406–3427.
[4] Khoury J, Weltman A. Chameleon Cosmology. Phys Rev. 2004 Dec;D69(044026).
[5] Raffelt G, Stodolsky L. Mixing of photons with low mass particles. Phy Rev D. 1988 March;37(5):1237–1249.
[6] Burrage C, Davis A, Shaw DJ. Detecting Chameleons: The Astronomical Polarization Produced by Chameleon-like Scalar Fields. Phy Rev D. 2009 Feb;79.
[7] A Upadhye, J H Steffen, Weltman A. Constraining chameleon field theories using the GammeV afterglow experiments; 2009. [hep-ph] arXiv:0911.3906v1.
[8] Waterhouse TP. An Introduction to Chameleon Gravity. arXiv:astro-ph/0611816v1. 2006 Nov;.

Eclipsing contact binary stars - period analysis using SuperWASP data

P L Skelton

Department of Mathematical Sciences, University of South Africa, Pretoria 0003, South Africa

E-mail: skeltpl@unisa.ac.za

Abstract. Some eclipsing contact binary stars of the W UMa-type are known to undergo changes in orbital period. These changes can be as a result of the light travel time effect if the contact binary is a member of a multiple stellar system or due to some intrinsic phenomena that are poorly understood at this stage. Analysing systems that appear to be undergoing changes in orbital period may shed some light on the important physical processes that occur in close binary systems. Pilecki *et al.* searched through the All Sky Automated Survey (ASAS) database for semi-detached and contact binary systems with high period change rates. They present 31 interacting binaries whose periods either increased (10) or decreased (21) in a five year interval of observations. Using data from the Wide Angle Search for Planets (SuperWASP) project, it has been possible to do period analyses using O−C diagrams which provide a more reliable measure of period change. Results of the analysis of two of the contact binaries studied by Pilecki *et al.* are presented. For both stars, the analysis results indicate that the classification of the stars as high period change rate systems is correct.

1. Introduction

W Ursae Majoris (W UMa)-type variable stars are contact eclipsing binary stars with orbital periods of less than 1d and mass ratios between 0.08 and 0.8 [1; 2]. The component stars of a W UMa-type variable star range in spectral type from mid-A to late-K and each is assumed to be a main sequence star. The structure and evolution of these systems is poorly understood [3–5]. The traditional view is that a W UMa-type star forms from a close, detached binary that loses angular momentum through magnetic stellar winds. Evolution through angular momentum losses due to mass lost in magnetized winds [6–8], thermal relaxation oscillations due to mass exchange between the two components [9–11] or a combination of both processes, are possible ways these systems evolve over time. An alternative formation mechanism is based on observations which suggest that many W UMa-type stars are members of triple or multiple star systems. Hendry and Mochnacki [12] suggest that a tertiary component (or multiple components) may play a vital part in the formation of a W UMa-type star by removing angular momentum from the system.

Some W UMa-type stars are known to display changes in their orbital period. These changes are sometimes complex, with intervals of constant period interrupted with intervals where the period increases or decreases. The orbital period changes observed in contact binaries are either short-term, with a time scale of less than a decade, or long-term, which occur on a thermal time scale [13]. Typical period change rates for contact binaries are between 10^{-7} and 10^{-8} d yr^{-1}.

Changes in orbital period can be due either to a redistribution of the matter between the stars or when angular momentum is gained or lost by the system [2]. The period of a W UMa-type star can also be reduced by magnetic braking through a magnetized stellar wind. Applegate [14] proposed a mechanism that relates changes in orbital period to changes in the distribution of angular momentum in the active star, which in turn varies the oblateness of the star producing variations in luminosity. The Applegate mechanism is yet to be reliably confirmed. Luminosity changes on parts of the binary star's photosphere can also lead to shifts in the minima times, thus mimicking period changes. Unlike a real change in orbital period, these changes do not accumulate over time.

Period changes observed in binary stars do not have to be intrinsic phenomena. Light travel time effects are produced when a binary system is a member of a triple or multiple stellar system. The presence of a third body (or more) can be inferred from periodic variations in an O−C diagram. Clearly, analysing systems that appear to be undergoing changes in orbital period may shed some light on the formation of these stars and the important physical processes that occur in close binary systems.

Situated at the Las Campanas Observatory in Chile, the All Sky Automated Survey (ASAS) is a project that was set up in 1996 to detect and monitor the variability of stars between 8^{th} and 12^{th} mag south of declination $+28^{\circ}$. The ASAS consists of two wide-field telescopes and all stars are observed once per one to three nights, with observations made in the standard V-band and I-band filters. Although complete light curves of variable stars are never observed by the ASAS, information such as the period and reference time of minimum or maximum brightness are provided so that phase-magnitude diagrams can be created. The project has already detected over 50,000 variable stars, a large portion of which have not been classified previously as variable stars. Of the 50,000 stars, over 5000 have been classified as contact binary systems.

Pilecki *et al.* [15] searched the ASAS data for semi-detached and contact binary systems with high period change rates. For a star to be selected for their analysis, the authors imposed the following selection criteria: a minimum of 300 observations of high quality, random distribution of the observations and orbital periods shorter than 10 d. Of the stars that fulfilled these criteria, the authors imposed a final constraint of a high signal-to-noise ratio, ending up with a total of 1711 stars of which 576 are semi-detached and 1135 contact stars. Using a local scatter reduction method, the authors find 31 binaries (22 contact and 9 semi-detached) whose periods either increased (10) or decreased (21) in a five year interval of observations. Of the 22 contact binaries, the orbital period was found to have increased for 5 and decreased for 17. The period change rates of the 31 binaries are around 10^{-5} d yr^{-1}. The authors suggest mass transfer as one mechanism that can produce the observed period change rates. If the orbital periods of the systems are changing, and if the period change rates are high, it should be detectable in an O−C diagram.

SuperWASP is an exoplanet survey programme, designed to detect exoplanets via transit events. SuperWASP consists of two observatories, SuperWASP-North situated on the island of La Palma and SuperWASP-South, situated at the South African Astronomical Observatory. These robotic observatories each consist of eight wide-angle cameras with a total field of view of 482 deg^2. Because of the observing procedure of SuperWASP, the project has seasons of photometric data for many variable stars. For variable stars with short periods, like W UMa-type stars, complete light curves are often obtained during a night's observing. The complete light curves can be used for photometric modelling of the systems as well as for period analyses.

Using SuperWASP data, a period analysis was performed for the contact binary stars ASAS 002449−2744.3 and ASAS 002821−2904.1. The results of the analyses are presented.

2. Period Analysis

2.1. Method

For the SuperWASP data, times of minimum brightness were determined using a method similar to the Kwee & van Woerden method [16]. Due to the geometry of a W UMa-type variable star, the minima should be symmetrical in shape. By fitting a second-order polynomial of the form $V(t) = at^2 + bt + c$ to a minimum, the time of minimum brightness can be calculated using the parameters of the best fit polynomial, that is, $T_{min} = -b/2a$. The error of the minimum is calculated using the covariance matrix entries.

The minima times were used in conjuction with the second-order ephemeris parameters determined by Pilecki *et al.* and a SuperWASP minimum to construct an O−C diagram. Both a first and second-order least squares analysis was performed. The period change rate is given by $dP/dt = (2k/P)$, where k and P are obtained from the second-order ephemeris

$$T_{calc} = T_0 + PE + kE^2$$

and E is the cycle number.

2.2. Results

2.2.1. ASAS 002449−2744.3 This system is listed in the ASAS database with a period $P = 0.31367$ d. Pilecki *et al.* find $P = 0.313661$ d and $dP/dt = -2.3 \times 10^{-6}$ d yr^{-1}. A total of 44 minima were obtained from the SuperWASP data. The O−C diagram shown in Fig. 1 was obtained using

$$T_{calc} = \text{HJD } 2454296.6348 + 0.313661E - 9.9 \times 10^{-10}E^2$$

The O−C residuals suggest that the period and k value of the system are not correct.

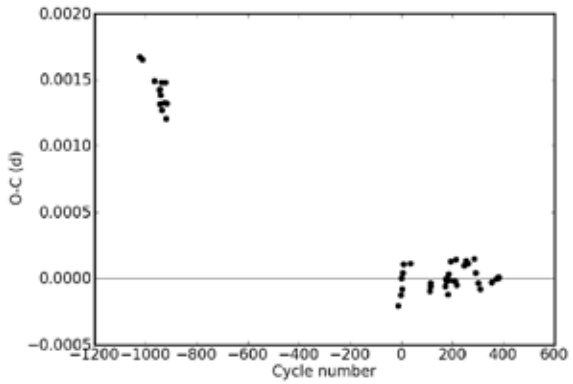


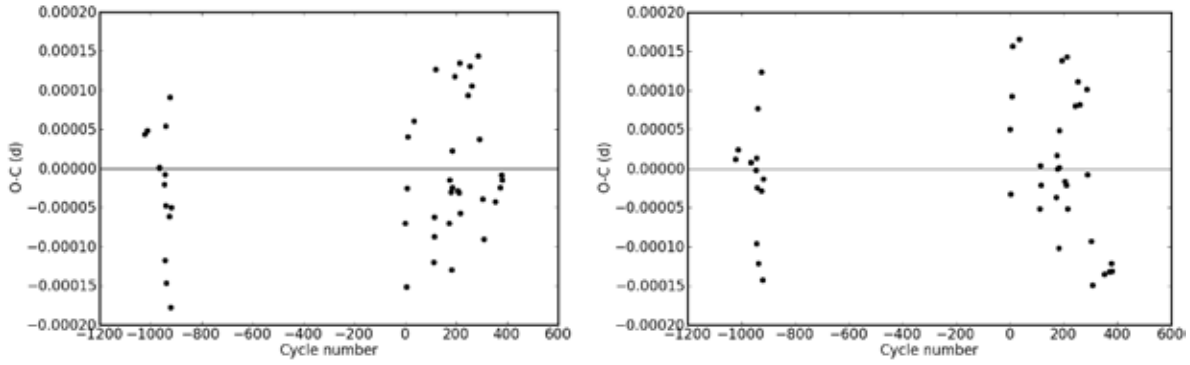
Figure 1. O−C diagram for ASAS 002449−2744.3 obtained using the Pilecki *et al.* period and k value.

The linear ephemeris obtained from the regression analysis is

$$T_{calc} = \text{HJD } 2454296.63490(2) + 0.31366049(4)E$$

with the uncertainties listed in brackets. The O−C diagram for this ephemeris is shown in Fig. 2(a). The second-order ephemeris

$$T_{calc} = \text{HJD } 2454296.63478(4) + 0.3136609(2)E + 5.5(2.0) \times 10^{-10}E^2$$



(a) (b)

Figure 2. O−C diagrams for ASAS 002449 − 2744.3 obtained using the linear ephemeris, (a), and the second-order ephemeris (b). The second-order ephemeris produces a lower residuals value compared to the linear ephemeris.

produces a lower O−C residuals value compared to the linear ephemeris. The O−C diagram obtained using this ephemeris is shown in Fig. 2(b).

The value of 5.5×10^{-10} corresponds to a period change rate $dP/dt = 1.3 \times 10^{-6}$ d yr^{-1}. The direction of the period change is different to that determined by Pilecki *et al.* indicating that the orbital period is increasing and not decreasing. The values for the orbital period obtained from the linear and second-order regressions are shorter than the ASAS period, while the period obtained from the second-order regression is very close to the Pilecki *et al.* value.

2.2.2. ASAS 002821−2904.1 Pilecki *et al.* determined a period $P = 0.269892$ d, shorter than the ASAS period $P = 0.269896$ d, and a period change rate $dP/dt = -2.3 \times 10^{-6}$ d yr^{-1}. Fig. 3 shows the O−C diagram for the star obtained using the ephemeris

$$T_{calc} = \text{HJD } 2454003.3140 + 0.269892E - 9.9 \times 10^{-10}E^2$$

The O−C residuals suggest that the k value and the period P are not correct.

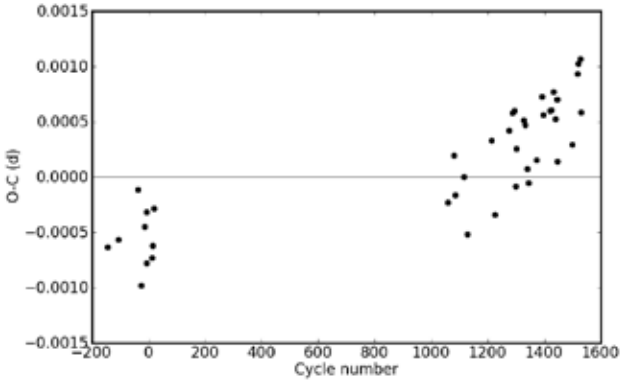
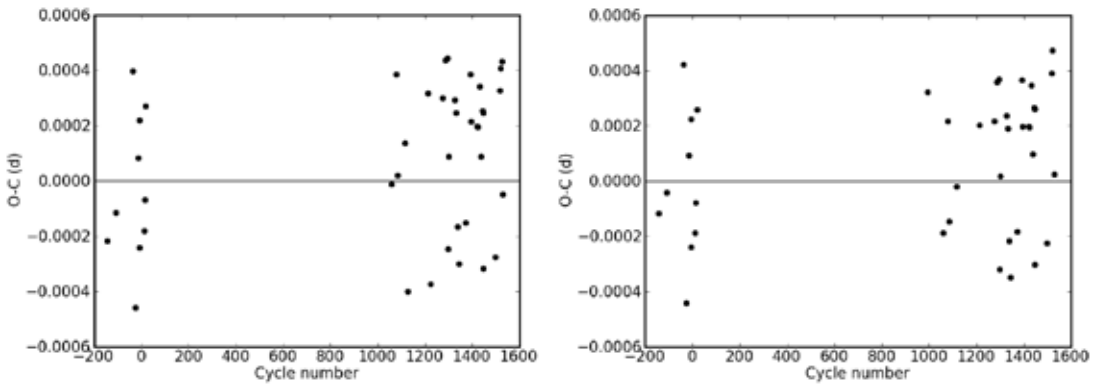


Figure 3. O−C residuals for ASAS 002821−2904.1 obtained using the Pilecki *et al.* period and k value.



(a) (b)

Figure 4. O−C diagrams for ASAS 002821−2904.1 obtained using the linear ephemeris, (a), and the second-order ephemeris, (b). The residuals value for the second-order ephemeris is lower than the linear ephemeris value.

A total of 39 minima were obtained from the SuperWASP data. The linear ephemeris obtained from the regression analysis is

$$T_{calc} = \text{HJD } 2454003.3135(1) + 0.26989126(9)E$$

The O−C diagram obtained using the linear ephemeris is shown in Fig. 4(a) and Fig. 4(b) shows the O−C residuals obtained using the second-order ephemeris

$$T_{calc} = \text{HJD } 2454003.3135(1) + 0.2698919(6)E - 4.5(4.0) \times 10^{-10}E^2$$

From a visual inspection, the O−C diagrams obtained using the linear and second-order ephemerides appear similar. The O−C residuals value for the second-order ephemeris is lower than the linear ephemeris residuals. The period change rate $dP/dt = -1.2 \times 10^{-6}$ d yr^{-1} and has the same direction of change as the value determined by Pilecki *et al.*, but the corresponding uncertainty value, 1.0×10^{-6} d yr^{-1}, is large. As shown in table 1, the periods returned by the linear and second-order regressions are shorter than the ASAS period. Although the period obtained from the second-order regression is similar to the Pilecki *et al.* period, more data are needed to confirm the validity of a second-order ephemeris.

Table 1. Comparison between the Pilecki *et al.* and SuperWASP values for P and dP/dt for both ASAS stars.

ASAS ID	Pilecki P (d)	SuperWASP P (d)	Pilecki dP/dt ($\times 10^{-6}$ d yr^{-1})	SuperWASP dP/dt ($\times 10^{-6}$ d yr^{-1})
002449 − 2744.3	0.313661	0.3136609	-2.3(0.5)	1.3(0.5)
002821 − 2904.1	0.269892	0.2698919	-2.3(0.5)	-1.2(1.0)

3. Discussion

Pilecki *et al.* searched the ASAS database and identified 22 contact binaries with high period change rates. Two of these stars, namely ASAS 002449−2744.3 and ASAS 002821−2904.1, were

selected for period analysis. The periods obtained from the regression analyses are shorter than the listed ASAS periods while the periods obtained from the second-order regressions are the same as the values determined by Pilecki *et al.*. For both stars, a second-order ephemeris is found to produce a lower O−C residuals value compared to the linear ephemeris value, suggesting that the orbital periods of the systems are changing. The calculated period change rates are larger than the typical values found for contact binaries, confirming the classification of these two stars as high period change rate systems.

For ASAS 002449−2744.3, the period change rate dP/dt obtained using the SuperWASP data is 1.3×10^{-6} d yr^{-1}, compared to -2.3×10^{-6} d yr^{-1} determined by Pilecki *et al.*. The direction of the period change is different. Tests are being performed to determine if the orbital period was decreasing while the ASAS data were obtained. This would explain the difference in the direction of the change between the Pilecki *et al.* value and the value determined from the SuperWASP data. The star will be monitored to see if the direction of the orbital period change reverses. For ASAS 002821−2904.1, the value of the period change rate is -1.2×10^{-6} d yr^{-1} compared to the Pilecki *et al.* value of -2.3×10^{-6} d yr^{-1}. The calculated dP/dt uncertainty, 1.0×10^{-6} d yr^{-1}, is large compared to the period change rate. More data are needed to confirm the validity of using a second-order ephemeris for ASAS 002821−2904.1.

For both stars, there is a gap of just under a year between the last ASAS observation in the data used by Pilecki *et al.* and the start of the SuperWASP observations. The results of the analyses are based on two seasons of SuperWASP data and extra data will confirm the changing period, the direction of the period change and possibly identify the cause of the period change. The SuperWASP data have proven to be extremely useful for the period analysis. The advent of exoplanet survey programmes such as SuperWASP and Kepler provide extra data for binary star research. These survey programmes are likely to have long lifespans and they can provide a long baseline of data. This will be particularly useful for detecting and monitoring orbital period changes in contact binary systems.

Acknowledgments
This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France and the SAO/NASA Astrophysics Data Stream.

References

[1] Shore S N 1994 *Saas-Fee Advanced Course 22: Interacting Binaries* ed Shore S N, Livio M and van den Heuvel E P J
[2] Hilditch R W 2001 *An Introduction to Close Binary Stars* (Cambridge University Press)
[3] Moss D L and Whelan J A J 1970 *Mon. Not. R. Astron. Soc* **149** 147
[4] Lapasset E and Claria J J 1986 *Astron. Astrophys.* **161** 264
[5] Paczynski B 1997 *Variables Stars and the Astrophysical Returns of the Microlensing Surveys* ed Ferlet R, Maillard J P and Raban B p 357
[6] Vilhu O 1982 *Astron. Astrophys.* **109** 17
[7] van't Veer F and Maceroni C 1989 *Astron. Astrophys.* **220** 128
[8] Stepien K 1995 *Mon. Not. R. Astron. Soc* **274** 1019
[9] Lucy L B 1976 *Astrophys. J.* **205** 208
[10] Wang J M 1999 *Astron. J.* **118** 1845
[11] Qian S 2003 *Mon. Not. R. Astron. Soc* **342** 1260
[12] Hendry P D and Mochnacki S W 1998 *Astrophys. J.* **504** 978
[13] Kalimeris A, Rovithis-Livaniou H and Rovithis P 1994 *Astron. Astrophys.* **282** 775–786
[14] Applegate J H 1992 *Astrophys. J.* **385** 621
[15] Pilecki B, Fabrycky D and Poleski R 2007 *Mon. Not. R. Astron. Soc* **378** 757
[16] Kwee K K and van Woerden H 1956 *Bull. Astron. Inst. Neth* **12** 327

What can we learn from phase alignment of γ-ray and radio pulsar light curves?

C Venter[1], T J Johnson[2,3] and A K Harding[2]

[1] Centre for Space Research, North-West University, Potchefstroom Campus, Private Bag X6001, Potchefstroom 2520, South Africa

[2] Astrophysics Science Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

[3] Department of Physics, University of Maryland, College Park, MD 20742, USA

E-mail: Christo.Venter@nwu.ac.za

Abstract. The *Fermi* Large Area Telescope (LAT) has revolutionized high-energy (HE) astronomy, and is making enormous contributions particularly to γ-ray pulsar science. As a result of the many new pulsar discoveries, the γ-ray pulsar population is now approaching 100. Some very famous millisecond pulsars (MSPs) have also been detected: J1939+2134 (B1937+21), the first MSP ever discovered, as well as J1959+2048 (B1957+20), the first black widow pulsar system. These, along with other MSPs such as PSR J0034−0534 and J2214+3000, are rare among the pulsar population in that they exhibit nearly phase-aligned radio and γ-ray light curves (LCs). Traditionally, pulsar LCs have been modelled using standard HE models in conjunction with low-altitude conal beam radio models. However, a different approach is needed to account for phase-aligned LCs. We explored two scenarios: one where both the radio and γ-ray emission originate in the outer magnetosphere, and one where the emission comes from near the polar caps (PCs) on the stellar surface. We find best-fit LCs using a Markov chain Monte Carlo (MCMC) technique for the first class of models. The first scenario seems to be somewhat preferred, as is also hinted at by the radio polarization data. This implies that the phase-aligned LCs are possibly of caustic origin produced in the outer magnetosphere, in contrast to the usual lower-altitude conal beam radio models. Lastly, we constrain the emission altitudes with typical uncertainties of $\sim 10\%$ of the light cylinder radius. The modelled pulsars are members of a third γ-ray MSP subclass, in addition to two others with non-aligned radio and γ-ray LCs.

1. Introduction

For quite some time, there were only a handful of pulsars detected in the γ-ray waveband [1]. This small sample already provided a starting point for spectral and light curve (LC) modelling as well as preliminary population studies. However, the launch of the *Fermi* Large Area Telescope (LAT) [2] heralded an exciting new era, particularly for high-energy (HE) pulsar physics.

1.1. γ-ray pulsars galore

Following the first *Fermi* Catalog which included about 1 500 γ-ray sources [3], the second *Fermi* Catalog is in now production, containing nearly 1 900 γ-ray sources [4, 5] including almost 100 γ-ray pulsars [6, 7]. About a third of these are millisecond pulsars (MSPs), some having been found by observing non-variable, unassociated *Fermi* sources at high latitudes with

radio telescopes and searching for pulsed radio signals [8, 9, 10]. Some of these MSPs exhibit the unusual phenomenon that the peaks of their γ-ray and radio LCs occur at the same observer phase ('longitude' ϕ), i.e., they are phase-aligned.

1.2. Traditional models

Two classes of models have been invoked to explain HE pulsar radiation. The first class, low-altitude polar cap (PC) models [11, 12], assume that primary electrons are accelerated above the neutron star surface, and that magnetic pair production of curvature radiation or inverse-Compton-scattered γ-rays occurs in the intense B-fields close to the stellar surface. In the case that pair creation is suppressed along the last open field lines, which mark the boundary between the corotating closed field line zone and the active open one, a slot gap (SG) [13] may form, corresponding to a two-pole caustic (TPC) geometry [14] which extends from the stellar surface up to near the light cylinder. The second class of models are the outer gap (OG) models [15, 16], which assume the production of HE radiation along the last open field lines above the so-called null charge surface, where the Goldreich-Julian charge density becomes zero. In both classes of outer-magnetospheric models (SG and OG), the HE pulse profiles are the result of the formation of *caustics* – the accumulation of photons in narrow phase bands due to a combination of special relativistic effects and magnetic field line curvature. Also, the narrow gaps in these models require abundant production of electron-positron pairs which will be able to screen the accelerating E-field.

For the case of MSPs, LCs and spectra have been modelled [17, 18, 19] using a pair-starved polar cap (PSPC) model [20]. The latter is similar to the traditional PC model, but in this case the number of pairs is not sufficient to screen the accelerating E-field, such that the accelerating region above the star is pair-starved. MSP spectra and energetics have also been modelled using an OG model [21, 22], while LC modelling using an annular gap model can furthermore reproduce the salient features of the γ-ray LCs [23].

The outer magnetospheric models usually invoke relatively low-altitude 'core' and 'conal' radio beams centred on the magnetic axis, in addition to the HE radiation being produced in extended regions reaching up to the light cylinder. The difference in the location of the γ-ray and radio emission regions thus implies that the corresponding pulse profiles will have non-zero lags (phase offsets) between the γ-ray and radio LCs.

1.3. Phase-aligned radio and γ-ray light curves: a new MSP subclass

The LCs of the first 8 *Fermi*-detected γ-ray MSPs [24] have been modelled [25], yielding two distinct MSP subclasses: those whose LCs are well fit by a standard OG or TPC model, and those whose LCs are well fit by a PSPC model. These fits are mutually exclusive. Importantly, this unexpectedly implied that there is copious pair production taking place even in MSP magnetospheres with their characteristic low B-fields.

A third class of MSPs emerged with the discovery of PSR J0034−0534 [26], which exhibited (nearly) phase-aligned radio and γ-ray LCs. Prior to this detection, such behaviour has only been observed for the Crab pulsar [27]. This phenomenon has now also been seen for PSR J1939+2134, PSR J1959+2048 [28], and PSR J2214+3000 [10]. Some other MSPs, including PSR B1821−24 and PSR J0737−3039A, may also have phase-aligned radio and γ-ray LCs, on the basis that their radio and X-ray profiles are phase-aligned [29, 30].

2. Geometric models predicting phase-aligned light curves

The non-zero radio-to-γ lags of MSPs with phase-aligned LCs disqualify the application of standard models, since the phase alignment implies co-located emission regions. We consider two model scenarios, one where the γ-rays and radio both occur in extended, high-altitude regions, and one where they originate at low altitudes. (For more details, see [31].)

2.1. High-altitude models

We use the same framework as previously [25], except that the radio emission region is now extended in altitude, and not assumed to be coming from a radio cone at a single altitiude. We furthermore limit the minimum and maximum radii of the radio and γ-ray emission regions, and assume constant emissivity as a function of altitude. These models are therefore called altitude-limited OG (alOG) and TPC (alTPC) models. Technically, the traditional OG model is a specific instance of an alOG model, and similar for the traditional TPC and alTPC models.

2.2. Low-altitude models

The so-called low-altitude Slot Gap (laSG) models represent an alternative possibility assuming a non-caustic origin of the emission. These are actually very-low-altitude geometric SG models resembling a hollow-cone beam close to the stellar surface. We modulate the emissivity as motivated by detailed radiation models, where the emissivity rises and falls exponentially along the B-field lines, peaking at a distance of about one or two stellar radii. Additionally, we also investigated the case where the emissivity is constant.

3. Finding optimal LC solutions in multidimensional phase space

The models described above involve multiple free parameters describing the emission region extent and width, as well as the pulsar geometry (inclination and observer angles α and ζ). In order to pick statistically the best-fit parameters when comparing to the *Fermi* LC data, we have developed a Markov chain Monte Carlo (MCMC) maximum likelihood procedure [32] and applied this to the alOG and alTPC models for three MSPs as indicated below (see Table 1). The γ-ray LCs are fit using Poisson likelihood while the radio LCs are fit using a χ^2 statistic. These two values are then combined. Using this method, we can also derive uncertainties on the minimum and maximum altitudes of the emission regions.

4. Results

4.1. The alOG and alTPC models

Increasing the lower limit of the emission region's altitude while keeping the upper limit fixed has the effect of growing the PC size (i.e., deleting a ring of radiation around the PC in the phaseplot, which indicates relative intensity per solid angle vs. ζ and ϕ). This suppresses low-level off-peak emission, especially for the alTPC case, so that the peaks become sharper. Conversely, decreasing the maximum radius while keeping the lower one fixed constrains the emission to a ring-like structure around the PC. As a result, the relative peak heights change in the corresponding LCs, and low-level features may become more prominent.

Figure 1 shows an example of best alOG and alTPC LC fits for PSR J1939+2134. The best-fit values for α and ζ, for the three MSPs we have modelled, are summarized in Table 1,

Table 1. Best-fit pulsar geometries for different models.

	PSR J0034−0534		PSR J1939+2134		PSR J1959+2048	
Model	α	ζ	α	ζ	α	ζ
alOG	12°	69°	84°	84°	16°	82°
alTPC	30°	70°	75°	80°	43°	44°
laSG1	10°	34°	30°	32°	20°	43°
laSG2	10°	37°	35°	25°	25°	45°

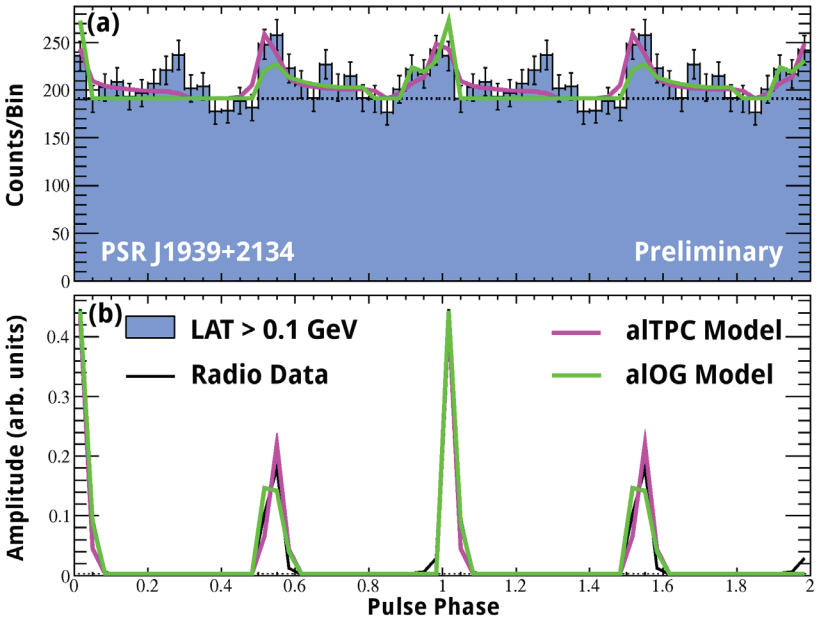


Figure 1. LC fits for PSR J1939+2134 using alTPC and alOG models. Panel (a) shows the γ-ray data [28], while panel (b) shows the radio data.

having typical errors of $\sim 5 - 10°$. Best-fit emission altitudes and gap widths are presented elsewhere [31].

4.2. The laSG models

It is interesting that in the laSG models the leading peak is wider while the trailing peak is sharper, even for the low-altitude emission considered here. This means that the caustic effects start to appear, given the large corotation speed of MSPs already near their surface. By including emission from higher altitudes, the peaks become broader. Also, the peak phase separation and width may be altered by choosing different values for the minimum and maximum emission altitudes. The constant-emissivity assumption leads to block-shaped LCs, and is therefore not considered viable. For the modulated-emissivity case, the peak widths may be fine-tuned using fading parameters, while the peak separation depends sensitively on ζ. This class of models shows less variation in profile shape than the high-altitude ones, and there are also multiple combinations of the free parameters that give very similar LCs, so a 'best solution' is probably not unique. The different best-fit solutions usually are a trade-off between best fits for γ-ray vs. radio LCs. For this reason, we indicate two similar laSG solutions (labeled 'laSG1' and 'laSG2', differing only in fading parameters which set the emissivity fading properties) in Table 1.

5. Discussion

Pulsars with radio and γ-ray LCs which are non-aligned in phase may be modelled by γ-ray emission regions extended over a large range of altitudes (OG / TPC / PSPC models) in conjunction with a conal radio beam at relatively lower emission altitudes. However, phase-aligned LCs require co-located γ-ray and radio emission regions. We studied the viability of reproducing such phase-aligned LCs using both high-altitude and low-altitude geometric models. We found that both classes of models could produce LC fits that can capture the most prominent features of the pulse profiles.

5.1. Model constraints

We have found best fits for the two angles describing the pulsar geometry, α and ζ, as indicated in Table 1. These have typical errors of $\sim 5-10^{\circ}$. In addition, we were able to limit the emission geometry for the high-altitude models, and found that the typical errors on the best-fit emission altitudes are $\sim$ 10% of the light cylinder radius. Furthermore, the maximum radio emission altitude seems to be better constrained than the γ-ray one. Since the low-altitude models can produce very similar LCs for similar values of its free parameters, we regard the inferred values of these free parameters when applying the model to the data as less robust than the case of the altitude-limited models.

5.2. Which model is preferred?

While we used the MCMC technique to search for the optimal set of parameters that produced the closest match to the LC data for the case of the altitude-limited models, we have not yet implemented this technique for the laSG models. Since the parameter space has not been fully explored for the latter class of models, we cannot really claim to have found the best fit for the alSG models. It is therefore difficult to quantitatively favour one class of models above the other. However, we did calculate the likelihood of the manually-selected best-fit laSG LCs as well as the MCMC-selected best-fit LCs of the altitude-limited models, and found that the altitude-limited models give better fits than the laSG ones, while the alTPC models provide slightly better LC fits than the alOG models. This implies that the phase-aligned γ-ray and radio LCs are most probably of caustic origin, produced in the outer magnetosphere, and the radio emission is most likely originating near the light cylinder (see also [33]).

5.3. Does causticity imply phase-alignment?

By studying the visibility of two samples of γ-ray and radio pulsars, it was concluded [34] that radio and γ-ray beams must have comparable sky coverage, especially for pulsars with large spin-down luminosities. This implies that the radio emission should originate in very wide beams at a significant fraction of the light cylinder, motivating studies of high-altitude caustic radio emission. However, this scenario is limited in application to pulsars having nearly phase-aligned LCs (with phase differences of up to ~ 0.2), unless the radio emission region is significantly offset from the γ-ray one. The bulk of the γ-ray pulsar population exhibits quite large ($\sim 0.1-0.5$) radio-to-γ-ray phase lags, so that caustic radio emission is probably is not ubiquitous in young pulsars. Furthermore, the γ-ray profiles are usually double-peaked, while the radio ones are single-peaked. Unless the radio emissivity has a strong altitudinal or azimuthal dependence, both radio and γ-ray LCs would be mostly double-peaked if they result from caustic emission, which is not observed. On the other hand, radio caustics may be more common for short-period MSPs, as there are many more examples of MSPs with phase-aligned LCs. Detailed population studies involving both young and old pulsars will provide more quantitative answers to this question.

5.4. The potential of polarization

The rotating vector model [35] generally provides a good description of radio polarization data (position angle as function of phase). However, this model assumes a static dipole B-field with no rotation effects included, and is probably not valid for fast-spinning MSPs which are expected to have significant B-field distortion. Caustic emission models predict rapid position angle swings with phase, coupled with depolarization [36], since the emission from a large range of altitudes and B-field orientations is compressed into a narrow phase interval to form the peaks. Rapid changes in the polarization angle and low levels of linear and circular polarization near the peak phases have indeed been seen for the MSPs modelled in this paper, and may be indicative of caustic effects. Polarization signatures will therefore be important to help discriminate between

LCs produced by caustic emission (such as occurs in alOG / alTPC models) or not (e.g., in the laSG model).

6. Conclusions

Future studies should include development of full radiation models which will be able to reproduce both the multiwavelength LC shapes, polarization properties, as well as the energy-dependent behaviour of the spectra of the γ-ray pulsars. Ways in which to induce abundant pair creation in low-B MSP magnetospheres, such as using offset-dipole [37] or multipole [21] B-fields, will need to be investigated.

Acknowledgments

CV is supported by the South African National Research Foundation. AKH acknowledges support from the NASA Astrophysics Theory Program. CV, TJJ, and AKH acknowledge support from the *Fermi* Guest Investigator Program as well as fruitful discussions with Dick Manchester and Matthew Kerr.

References

[1] Thompson D J 2004 *Cosmic Gamma-Ray Sources* vol 304, ed K S Cheng and G E Romero (Dordrecht: Kluwer Academic Publishers) pp 149–68
[2] Atwood W B et al. 2009 *Astrophys. J.* **697** 1071–102
[3] Abdo A A et al. 2010 *Astrophys. J. Suppl. Ser.* **188** 405–36
[4] Burnett T H 2011 *Proc. 3rd Fermi Symp. (Rome)*, in press
[5] Thompson D J 2011 *Proc. 3rd Fermi Symp. (Rome)*, in press
[6] Abdo A A et al. 2010 *Astrophys. J. Suppl. Ser.* **187** 460–94
[7] Smith D A 2011 *Proc. 3rd Fermi Symp. (Rome)*, in press
[8] Cognard I et al. 2011 *Astrophys. J.* **732** 47–57
[9] Keith M J et al. 2011 *Mon. Not. R. Astron. Soc.* **414** 1292–300
[10] Ransom S M et al. 2011 *Astrophys. J.* **727** L16–21
[11] Daugherty J K and Harding A K 1982 *Astrophys. J.* **252** 337–47
[12] Daugherty J K and Harding A K 1996 *Astrophys. J.* **458** 278–92
[13] Arons J 1983 *Astrophys. J.* **266** 215–41
[14] Dyks J and Rudak B 2003 *Astrophys. J.* **598** 1201–6
[15] Cheng K S, Ho C and Ruderman M 1986 *Astrophys. J.* **300** 500–21
[16] Romani R W 1996 *Astrophys. J.* **470** 469–78
[17] Frąckowiak M and Rudak B 2005 *Adv. Space Res.* **35** 1152–7
[18] Harding A K, Usov V V and Muslimov A G 2005 *Astrophys. J.* **622** 531–43
[19] Venter C and De Jager O C 2005 *Astrophys. J.* **619** L167–70
[20] Muslimov A G and Harding A K 2004 *Astrophys. J.* **617** 471–9
[21] Zhang L and Cheng K S 2003 *Astron. Astrophys.* **398** 639–46
[22] Zhang L, Fang J and Chen S B 2007 *Astrophys. J.* **666** 1165–73
[23] Du Y J, Qiao G J, Han J L, Lee K J and Xu R X 2010 *Mon. Not. R. Astron. Soc.* **406** 2671–77
[24] Abdo A A et al. 2009 *Science* **325** 848–52
[25] Venter C, Harding A K and Guillemot L 2009 *Astrophys. J.* **707** 800–22
[26] Abdo A A et al. 2010 *Astrophys. J.* **712** 957–63
[27] Kniffen D A, Hartman R C, Thompson D J, Bignami G F and Fichtel C E 1974, *Nature* **251** 397–9
[28] Guillemot L et al. 2011 *Astrophys. J.*, in press
[29] Backer D C and Sallmen S T 1997 *Astron. J.* **114** 1539–43
[30] Chatterjee S, Gaensler B M, Melatos A, Brisken W F and Stappers B W 2007 *Astrophys. J.* **670** 1301–6
[31] Venter C, Johnson T J and Harding A K 2011 *Astrophys. J.* in press
[32] Johnson T J, Harding A K and Venter C 2011, in preparation
[33] Manchester R N 2005 *Astrophys. Space Science* **297** 101–8
[34] Ravi V, Manchester R N and Hobbs G 2010 *Astrophys. J.* **716** L85–89
[35] Radhakrishnan V and Cooke D J 1969 *Astrophys. J.* **3** L225–9
[36] Dyks J, Harding A K and Rudak B 2004 *Astrophys. J.* **606** 1125–42
[37] Harding A K and Muslimov A G 2011 *Astrophys. J.* **726** L10–14

Campaign for vicarious calibration of SumbandilaSat in Argentina

L M Vhengani, D Griffith and M Lysko

Council for Scientific and Industrial Research, Defence Peace Safety and Security, Pretoria, South Africa

E-mail: Lvhengani@csir.co.za

Abstract. The importance of calibrating satellite imagers has been explained in literature such as K Arai (2007) and K J Thome (2001). Calibration of satellite sensors (imagers) is crucial for data consistency, reliability and comparability. To perform a meaningful analysis of a satellite image, the Digital Numbers (DNs) of the image are first converted to absolute radiance by using the sensor-specific radiometric calibration coefficients. Satellite imagers are calibrated pre-launch and for continuous assessment, they are also calibrated post-launch. Various post-launch techniques exist including cross-sensor, solar, lunar and vicarious calibration. Vicarious calibration relies on *in-situ* measurements of surface reflectance and atmospheric transmittance to estimate Top-Of-Atmosphere (TOA) spectral radiance. A vicarious calibration field campaign was executed in Argentina to support monitoring of the radiometric response of the multispectral imager aboard SumbandilaSat. Results obtained using two Radiative Transfer Codes (RTCs) MODTRAN and 6SV are presented.

1. Introduction

Satellite images are key to most earth observation studies. The importance of a satellite sensor calibration has been explained in literature such as K Arai (2007) and K J Thome (2001). In gist, calibration of a satellite sensor is crucial for data consistency, reliability and comparability. To perform meaningful analysis of a satellite image, the Digital Numbers (DNs) of the image are first converted to absolute radiance by using the sensor specific radiometric calibration coefficients [3]. The derived radiance is expected to be comparable to the radiances derived with alternate and similarly specified sensors.

Satellite sensors are calibrated before launch to determine pre-flight radiometric calibration coefficients. However, the process of launching is accompanied by extreme vibrations and thermal fluctuations which may change the characteristics of the sensor. In addition, the characteristics of the sensor may change while the sensor is on-orbit. This change may be due to detector out-gassing and the deterioration of electronic components. This means that there is a need for continuous calibration while the sensor is on-orbit. There are various methods of on-orbit calibration, including cross sensor calibration, the use of on-board calibration instruments, vicarious calibration and lunar calibration [4].

SumbandilaSat does not have any on-board calibration instruments. In addition, after launch SumbandilaSat experienced a failure of some components, which compromised the functionality of its attitude control system. This means that it would be difficult and risky to perform lunar calibration. Currently, the only options available for continuous calibration of SumbandilaSat are the cross sensor

calibration method and the vicarious calibration method. In this paper, the focus is on vicarious calibration. Vicarious calibration is a calibration method that is independent of the pre-launch calibrations [5].

The first step of any vicarious calibration campaign is the selection of a calibration site. Calibration sites are characterized using reference information such as reflectance data and atmospheric conditions during the calibration campaign. Scott *et. al.* (1996) gives a summary of the criteria for selecting a radiometric calibration site. A calibration site must be suitably large, homogenous and cloud-free to allow fine ground characterization [6]. Once a suitable site is chosen, reflectance and atmospheric properties of the site are measured at the time of satellite overpass. The measured data are used as inputs into a Radiative Transfer Code (RTC) to estimate the at Top-Of-Atmosphere (TOA) spectral radiance. The average DNs of the calibration site are then computed from the satellite image acquired during the satellite overpass. Depending on the DN-to-radiance conversion model of the sensor, the calibration coefficients are then calculated. The simplest sensor model for converting DN to radiance is shown in equation (1).

$$L_\lambda = CC_\lambda \times DN_\lambda - O_\lambda \tag{1}$$

where CC_λ is the calibration coefficient, DN_λ is the average digital count of the sample site on the satellite image, L_λ is the TOA radiance and O_λ is the offset or dark signal of the band situated at wavelength λ. O_λ can be obtained using various methods, including observing deep space, using camera shutters or from onboard calibrators [5].

The aim of this study is to compute the TOA spectral radiance using two different RTCs. The surface spectral reflectance and atmospheric characteristics used as inputs were measured at two sites with relatively homogeneous ground reflectance targets in Argentina during near coincident SumbandilaSat overpasses.

2. Radiative transfer codes

Optical remote sensing sensors measure spectral radiance reflected from the surface of the earth. The source of the radiance in the shortwave spectrum is the sun. The strength and characteristics of this reflected energy depends on the characteristics of the surface reflectance, however, it is also affected by the gaseous absorption and the scattering by molecules and aerosols in the atmosphere.

RTCs are used to compute the scattering and absorption effects of the atmosphere as well as the TOA spectral radiance. There are a number RTCs available but the focus in this study will be on the **MOD**erate resolution atmospheric **TRAN**smittance (MODTRAN) code and the Second Simulation of a Satellite Signal in the Solar Spectrum Vector code (6SV). These two RTCs are the most commonly used codes within the Committee on Earth Observation Satellite (CEOS)-Working Group on Calibration and Validation (WGCV).

MODTRAN was developed by the Geophysics Division of the Air Force Research Laboratory (AFRL) and their partners. It was developed in the late 1980s using the FORTRAN computer language. Currently the latest version is MODTRAN 5, but MODTRAN 4 is still widely used and has been available to the public since January 2000. 6S was developed in 1997 and is an improved version of 5S (Simulation of the Satellite Signal in the Solar Spectrum), developed by the Laboratoire d'Optique Atmosphérique [7]. 6SV is a vector version of the 6S RTC for the clear sky Earth atmosphere under the plane parallel assumption. 6SV is a vector code, as opposed to a scalar code such as the current MODTRAN generation and takes full account of the polarization state of the Electro-Magnetic radiation field.

Polarization turns out to be an important consideration when Rayleigh scattering dominates or when retrieving TOA reflectance, particularly for low reflectance targets such as ocean or dark vegetation. For vicarious calibration of land imagers one is usually dealing with a high reflectance target with viewing angles near nadir and relatively high Solar Zenith Angle (SZA). In these circumstances, the difference between a scalar and a vector code should be minimal. However, it is always good practice to use at least two RTCs and to compare the results. MODTRAN and 6SV differ

greatly in the nature and format of their inputs, their capabilities and their outputs as well as availability and licensing.

3. Methodology

Field campaigns were held on the 19th and the 24th of October, 2010. The campaigns were executed at Barreal Blanco playa and Salar de Arizaro in Argentina. *In-situ* surface reflectance and atmospheric characteristic were measured in close to the times of the SumbandilaSat overpasses. The spectral reflectance at the two sites is plotted in Figure 1. These plots show that the two sites have relatively uniform (flat) spectral reflectance in the wavelength region of 600 nm to 900 nm.

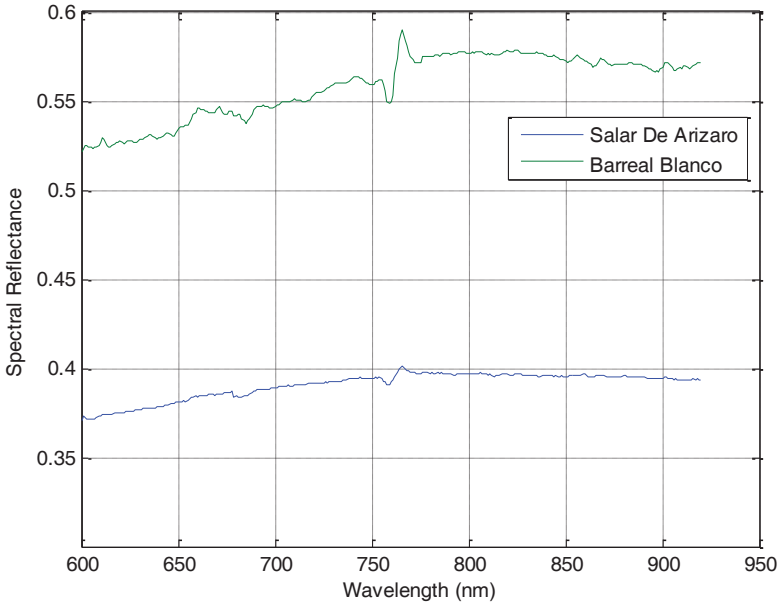


Figure 1. Surface Spectral Reflectance of Salar de Arizaro and Barreal Blanco.

Pre-processed *in-situ* measurements were used as inputs into MODTRAN and 6SV in order to compute the TOA spectral radiance. The computed TOA radiances for the two sites are given in the plots in Figure 2. The equivalent TOA spectral radiances are computed as shown in equation (2).

$$L_{spectralj} = \frac{\int_{\lambda_1}^{\lambda_2} S_j(\lambda) L_{TOA}(\lambda) d\lambda}{\int_{\lambda_1}^{\lambda_2} S_j(\lambda) d\lambda} \tag{2}$$

Where $S_j(\lambda)$ is the spectral response function or sensitivity of band j, L_{TOA} is the TOA spectral radiance and λ_1 and λ_2 are the spectral range of band j. The results of this computation are shown in Figure 1.

4. Results and discussions

It can be observed in Figure 1 that the surface spectral reflectance values of the two sites are different. Any differences between the TOA radiance values over the two sites may be attributed to the difference in surface spectral reflectance, differences in atmospheric conditions or SZA at the time of overpass.

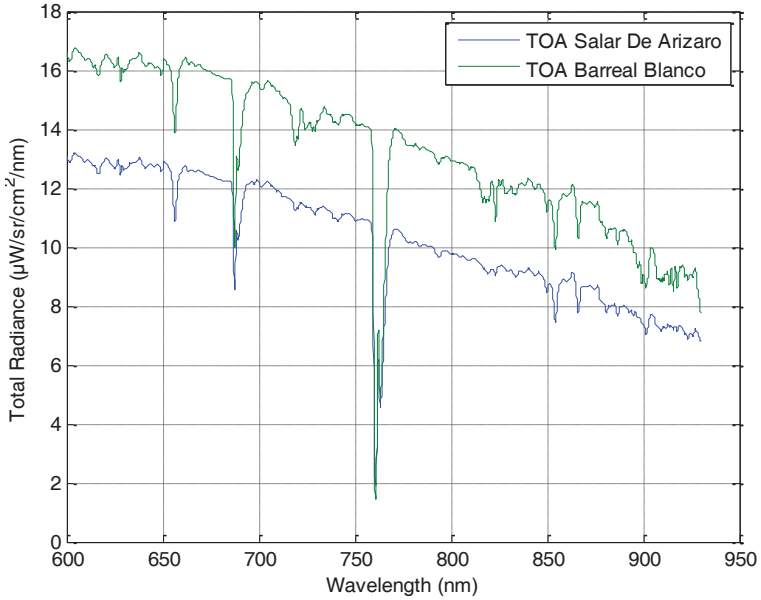


Figure 2. Top of Atmosphere Radiance for Salar De Arizaro and Barreal Blanco.

It can be observed in Table 1 that 6SV and MODTRAN results are consistent with each other for the respective sites. The TOA spectral radiance values shown in Table 1 have units of $Wm^{-2}sr^{-1}\mu m^{-1}$. The percent difference of the results shown in Table 2 yielded from MODTRAN and 6SV range between 0.14 % and 2.3 %. These results are not in agreement with the value of 0.5 % as discussed in Thome (2004).

Table 1. Top of Atmosphere Spectral Radiance in $Wm^{-2}sr^{-1}\mu m^{-1}$.

	Salar De Arizaro		Barreal Blanco	
Red	6SV	MODTRAN	6SV	MODTRAN
Red	122.571	123.819	157.711	157.933
Red Edge	114.021	116.646	146.552	147.986
NIR	95.644	97.168	125.490	126.246

A likely cause of not achieving the expected 0.5 % agreement in results is the difference in aerosol models between the two codes. Refinement of the inputs to the two codes which control aerosol spectral characteristics is expected to improve the agreement. MODTRAN and 6SV permit detailed control of aerosol characteristics with user-defined aerosol options. Usually there is insufficient measurement data to completely specify the aerosols and assumptions have to be made.

Table 2. Percent Difference of 6SV and MODTRAN Results.

Band	Salar De Arizaro (%)	Barreal Blanco (%)
Red	1.013	0.141
Red Edge	2.276	0.974
NIR	1.581	0.601

5. Conclusions

There is an expectation that the calibration coefficients for the SumbandilaSat imager would have changed after launch, due to the effects of launch and due to the extreme nature of the space environment. For quality assurance, the results obtained from this Argentina campaign need to be verified. One method of verification is 'cross-sensor calibration'. A cross-sensor calibration refers to comparing two images, of the same target, acquired almost simultaneously with the uncalibrated, and a well calibrated sensor. One critical point to consider during cross-sensor calibration is the difference in spectral response of the two sensors as discussed in Thome (2004).

In terms of the vicarious calibration campaign in Argentina, the surface spectral reflectance measurements that were taken over the ground targets did not extend over all spectral bands in the SumbadilaSat imager. This shortfall had to be compensated for by extrapolation of the surface reflectance values in order to run the RTCs. This extrapolation may have introduced an uncertainty, which could not be quantified. It is therefore recommended that surface reflectance should always be measured in wavelength ranges that fully encompass the spectral response functions of the satellite's imager.

At the time of writing this paper, the bias or dark current of SumbandilaSat's sensor was not yet known. The dark current, once determined, must be subtracted from the average DNs before making any comparison.

The recommendation is that the results shown in this report be verified during the second calibration campaign, to be held in the second or third quarter of 2011. The protocol for the forthcoming campaigns will be similar to the protocol followed in CEOS-WGCV intercomparison campaigns that were held in Turkey in 2009 and 2010.

The first iteration of calibrating SumbandilaSat imager was a learning exercise. The lessons learned in this exercise will be highly valuable during the subsequent iterations.

The methodology employed in this study was adequate to obtain consistent results between MODTRAN and 6SV to within a few percent. In order to achieve best agreement between multiple RTCs it is necessary to pay detailed attention to the inputs, in particular the inputs that control aerosol characteristics. Forthcoming campaigns will make use of more detailed atmospheric measurements and hence better quality RTC inputs.

Acknowledgments

Authors would like to thank Satellite Application Centre for SumbandilaSat image acquisition and Mr Azwitamisi Mudau for Atmospheric and Reflectance data collection. This document is the result of a research effort funded by the Department of Science and Technology (DST) in terms of Order DST/CON-0067-2010.

References

[1] Arai K 2007 Vicarious calibration of the solar reflection channels of radiometers on-board satellites through the field campaigns with measurements of refractive index and size distribution of aerosols *Advances in Space Research* **39** 13-19

[2] Thome K J 2001 Absolute calibration of Landsat 7 ETM+ using the reflectance-based method *Remote Sensing of Environment* **78** 27–38

[3] Chander G, Markham B L, Helder D L 2009 Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO1-ALI sensors. *Remote Sensing of Environment* **113** 893–903

[4] Keiffer H H, Wildey R L 1985 Absolute calibration of Landsat instruments using the Moon *Photogrammetric Engineering and Remote Sensing* **51** 1391–1393

[5] Thome K, Helder D, Aaron D A and Dewald J 2004 Landsat-5 TM and Landsat-7 ETM+ absolute radiometric calibration using the reflectance-base method *IEEE Transactions On Geoscience and Remote Sensing* **42** 2777-2785

[6] Scott K P, Thome K and Brownlee M R 1996 Evaluation of Railroad Valley Playa for use in vicarious calibration, Multispectral imaging for terrestrial applications *Proceedings of SPIE* **2818** 158-166, Denver, Colorado

[7] Vermote E F, Tantré D, Deuzé J L, Herman M and Morcrette J J 1997 Second Simulation of the Satellite Signal in the Solar Spectrum, 6S: An overview *IEEE Transactions On Geoscience and Remote Sensing* **35** 675-686

[8] Thome K 2004 In-flight intersensor radiometric calibration using vicarious approaches. Post-launch calibration of satellite sensors. Morain, S. A. and Budge, A. M. London, Taylor and Francis, 95-102

Brightest cluster galaxies – single or composite stellar populations?

D N Viljoen[1] and S I Loubser[1]

[1] Centre for Space Research, North-West University, Potchefstroom, 2520, South-Africa

E-mail: 20569513@nwu.ac.za

Abstract. This paper is devoted to the study of the stellar populations in brightest cluster galaxies (BCGs), more specifically determining whether the star formation histories (SFHs) of these galaxies can accurately be represented by single stellar populations (SSP) or composite stellar populations (CSP). High signal-to-noise ratio, longslit spectra for a sample of 41 galaxies were obtained on the Gemini telescopes. The stellar populations were studied by using the software package *ULySS*. Each galaxy was fitted against a SSP and 2 and 3 SSP models, separately. For the 2 and 3 SSP fits, the stellar components of the galaxies were divided into groups by age: a young, intermediate and an old component. Each SSP component is characterised by their ages and metallicities ([Fe/H]). A series of 500 Monte-Carlo simulations were then performed to aid in the selection of the most probable SFH of the BCGs. Then the χ^2-maps were drawn to assist in understanding the structure of the parameter space. During the fits, two models, Pegase.HR (P.HR) and the Vazdekis-MILES (V/M) models, were used as independent stellar population models. We found that the stellar population model, P.HR represented the SFHs of these BCGs more accurately. The P.HR model revealed that the CSP provided the best fit for some galaxies, indicating that at least some of these BCGs have a more complex evolution and formation than first thought.

1. Introduction

BCGs are the most massive and luminous galaxies in a cluster [1]. However, recent literature, for example [2] (hereafter Paper 2), gives a more detailed definition in which the BCGs are described as the central dominant galaxies in clusters, with a typical mass of $\sim 10^{13} M_{\odot}$ [3]. These BCGs are well aligned with the host cluster galaxy distribution [1], which consequently implies that BCGs are located at the bottom of the gravitational potential well. This in turn implies that the origin of the BCGs is closely related to the formation of the host cluster because it is widely accepted that the stars have settled to the bottom of this potential well.

Some BCGs are also called cD galaxies, which indicates that they have an extended stellar halo. This subclass of BCGs is the most luminous, and contains the largest stellar mass of any of the galaxies. They are significantly brighter than $L \sim 10^{10} L_{\odot}$ [4] and their galaxies are almost always the brightest galaxies in the clusters. The unique and distinctive properties of BCGs create a challenge for astronomers studying galaxy formation.

The formation history of the most massive galaxies in clusters and the role the cluster environment plays in galaxy formation, are important but poorly understood. It is difficult to find the progenitors of these early-type galaxies by means of direct observations, so astronomers

have used an alternative approach by studying the SFHs of large samples of nearby galaxies in terms of the properties, i.e. their ages and [Fe/H], of these stellar populations.

Stellar populations can be analysed to provide important information about the age and chemical composition of galaxies, thus providing us with knowledge about their formation and evolution. In this paper stellar populations were investigated and used to learn more about the age, [Fe/H] and consequently the formation and SFH of BCGs.

When using the popular Lambda Cold Dark Matter (ΛCDM) model of structure formation [5], the central cluster galaxies, more specifically BCGs, form naturally when a cluster collapses along the filaments. By using numerical simulations, it was shown that the BCGs in clusters with redshifts higher than $z \sim 1$ experienced a significant number of mergers [6]. It was found that the stars will end up in a BCG formed at high redshifts (with 50 per cent of the stars already formed $\sim$ 12.5 Gyr ago), yet BCGs continue to assemble mass at much lower redshifts (with 50 per cent of the mass assembling after $z \sim 0.5$) [7]. The nature of these mergers is dissipationless and, therefore, no new stars are formed during this process. According to these simulations, we therefore expect to see evidence of dissipationless mergers and old stellar populations in these galaxies.

Cooling flows are not fully understood, may play a role in the formation of the stellar populations of BCGs, and are typically found at low redshifts [8]. The model was contradicted due to the lack of widespread detection of iron lines [9], which are expected from cluster gas cooling below 1-2 keV as observed by XMM-*Newton* observations of cool-core clusters. Thus, BCG formation is not a consequence of cooling flows [9]. It was found that star formation in these cool-core clusters was likely but that it would take place at a much reduced rate [10]. The gas fuelling these star formations is a source of mystery and controversy.

Little is known about the properties of the stellar populations of BCGs [9]. A study was recently undertaken in which 625 brightest group and cluster galaxies were taken from the Sloan Digital Sky Survey (SDSS); their stellar population properties were compared with those of elliptical galaxies with the same mass [11] (hereafter called VDL). This study can be regarded as a point of reference in the investigation of stellar populations in BCGs. However, the VDL study did not include any spatial information about the BCGs. The merger history of a galaxy determines the kinematical and stellar population properties and these properties can be used in an analysis of the formation history of those galaxies.

No star formation took place during dissipationless mergers in the formation of BCGs [7]. Hence, only old stellar populations in these galaxies are expected. However, several studies reported examples of recent or ongoing star formation in BCGs hosted by cooling-flow clusters [10, 12, 13]. Hence, the question arises of whether the SFH of BCGs can be represented accurately by SSP models, as is widely assumed, or whether a CSP model is needed?

In this paper, we report on the central stellar populations of the galaxy sample, more specifically whether the SFHs of these stellar populations are more complex than first thought, i.e. is a SSP model sufficient enough when the SFH of a galaxy is reconstructed?

2. Method

2.1. Sample and Observations

The literature presents different definitions on how cD galaxies can be described or classified and due to these discrepancies the galaxies in the sample, as described in [9], may not all be classified as cD galaxies. However, all the galaxies in this sample are the dominant galaxies which were located closest to the X-ray peaks in the centres of the respective clusters (thus for a small fraction these galaxies might not strictly be the brightest galaxy in the cluster). We called these central cluster galaxies BCGs to concur to the recent definitions given in, for example [7].

A more detailed description regarding the sample selection and observations is given in [9]. In summary: these 40 galaxies were classified as cD either in the NASA/IPAC Extragalactic

Data base (NED)[1], in the morphological classification or in the notes of previous observations and/or have surface brightness profiles breaking the de Vaucouleurs $r^{\frac{1}{4}}$ law. In addition, NGC 4946, an ordinary elliptical, were also observed with the same observational set-up.

High signal-to-noise ratio, longslit spectra for the sample were observed on the Gemini telescopes during an extensive observational campaign from July 2006 to January 2008 [9]. Out of this sample, 31 galaxies lacked emission line and 10 galaxies showed weak emission lines. The spectra of the galaxies with weak emission lines were analysed with the software package *GANDALF* [14] to accurately separates the stellar and emission line contributions to the observed spectra.

2.2. ULySS - Université de Lyon Spectroscopic analysis Software

The software package *ULySS*[2] was used in this study. *ULySS* has the advantage of using full spectral fitting as opposed to fitting specific parts of the spectrum (so-called indices) as used up to now [15]. This software can be used to determine (i) the stellar atmospheric parameters and (ii) star formation and metal-enriched history of galaxies. This software can be used to fit SSPs or CSPs.

Studies conducted on stellar population synthesis models concluded that the P.HR, with the Elodie 3.1 stellar library, and V/M stellar population models are trustworthy and consistent (see [15], and references there in). These two models were used in this analysis and expanded further by defining additional input parameters, for example the instrumental line-of-sight velocity broadening, the velocity dispersion, the error spectra and the redshift of the galaxy analysed. A wavelength range of [3800, 6000] Å was chosen for all the galaxies in order to include all the important spectral feature elements, for example Hβ and Fe5270.

We incorporated a line spread function (LSF) to match the resolutions of the models and the observed spectrum as these are different. In *ULySS*, the relative LSF between the models and observation has to be determined and then this relative LSF has to be injected into the models. We used stars that were observed with the same instrumental configuration as the galaxy sample (see [9]). The stars HD 004037, HD 069267, HD 140283 and HD 184406 were used. The relative LSF was calculated by comparing the composite spectrum of these respective stars with that extracted from the P.HR model and then this LSF was injected into the P.HR model to generate the resolution-matched P.HR model by using the LSF convulsion function of *ULySS*. This resolution-matched P.HR model was then used in the further analysis. The same procedure was followed to generate the resolution-matched V/M model.

Here we note that NGC 6173 was excluded from further analysis because the models provided a visibly poor fit to the observed spectrum.

Each galaxy spectrum was fitted against a SSP and CSP. *ULySS* can be used to reconstruct a detailed SFH by directly fitting a positive linear combination of many SSPs, generally limited to 2 to 4 epochs of star formation. However, such an approach would be unstable because of the degeneracies and the finite quality of the models and observations [16]. To avoid these degeneracies, we started with simple physical assumptions - the presence of an old stellar population (OSP) and then the time axis was divided into intervals by setting limits on two or more intervals. Therefore the CSP was studied in terms of three epochs: a OSP, an intermediate-age stellar population (ISP) and a young stellar population (YSP).

We chose the following three different boxes, which were generated by setting limits on the ages and by leaving the [Fe/H] free [16]. When using the P.HR model, we choose the three boxes as follows: (i) For the OSP box: the age boundaries were fixed to [12, 20] Gyr. (ii) For the ISP box: the age boundaries were fixed to [4, 12] Gyr. (iii) For the YSP box: the age boundaries

[1] `http://nedwww.ipac.caltech.edu/`.

[2] `http://ulyss.univ-lyon1.fr/`.

were fixed to [0.01, 1] Gyr. When using the V/M model, we choose the three boxes as follows: (i) For the OSP box: the age boundaries were fixed to [12, 17.78] Gyr. (ii) For the ISP box: the age boundaries were fixed to [4, 12] Gyr. (iii) For the YSP box: the age boundaries were fixed to [0.1, 1] Gyr. This was necessary because the upper and lower limits of the respective models are different.

During the fits, the three boxes are suitable for the galaxy sample, except for several galaxies that have results reaching to the boundaries of at least one of the three boxes. For these cases, we changed the upper or lower age boundaries of the boxes until the final results converged. This method is also justified by the fact that whenever four or more components were fitted the weights of the additional components were zero.

The χ^2-statistical test assisted in choosing the most probable fit, more specifically whether the BCG can be presented more accurately by a SSP or CSP. These stellar population components are characterised by their ages and [Fe/H]. A series of 500 Monte-Carlo simulations were then performed, checking the residual of the fits and assessing the relevance of the solutions, aiding the selection of the most probable SFH of the BCGs. After this, the χ^2-maps were drawn to assist in understanding the structure of the parameter space.

3. Results and Discussions

Table 1 summarises the stellar components which provided the best fit for the galaxies (emission and non-emission lines) as determined by the P.HR and V/M models. For the non-emission line galaxies, the P.HR model determined that the results for the composite populations provided a better fit as indicated by the higher combined fractions for the two and three components than for the 1 SSP while the V/M model indicated that the single stellar population provided the best fit. On the other hand, the results of both models indicated that for the emission line galaxies the single stellar population provided the best fit. The total fraction of the 1 SSPs are higher than the total fraction of either the 2 and 3 SSPs.

Table 1. The fraction of stellar populations of the galaxies as determined by the P.HR and V/M models.

Component		P.HR	V/M
1 SSP	Non-emission	14/40	19/40
	Emission	7/40	6/40
	Total	21/40	25/40
2 SSPs	Non-emission	14/40	12/40
	Emission	2/40	2/4
	Total	16/40	14/40
3 SSPs	Non-emission	3/40	—
	Emission	—	1/40
	Total	3/40	1/40

Figures 1 and 2 illustrate two examples of where a CSP and SSP provided the best fit respectively. Figure 1 shows the solutions obtained by the P.HR and V/M models for GSC555700266. By comparing the two figures, it follows that the results, obtained by both models, that the two components provided the best fit which indicates that this BCG is most likely comprised out of composite populations. The CSP fit, determined by the P.HR model is the most accurate representation of the SFH of this BCG because the χ^2-value (indicated by

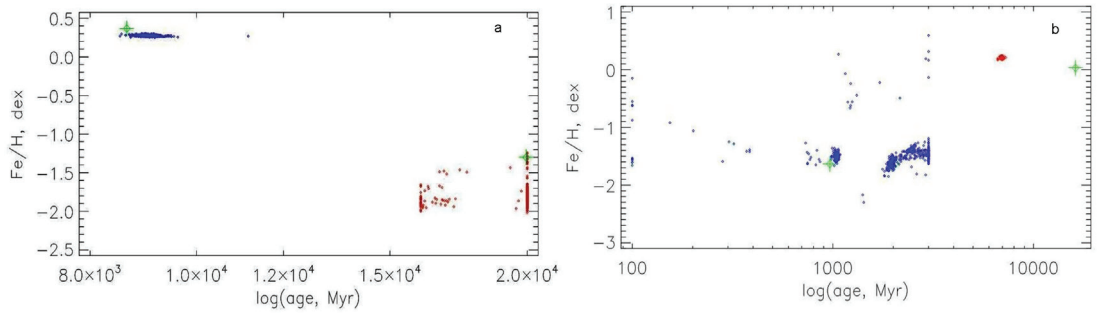


Figure 1. Stellar populations of GSC555700266 as determined by the (a) P.HR and (b) V/M models. The blue dots represent the YSP while the red dots indicate the OSP. The local minima (best fit) is indicated by the green cross.

the green cross) of each component converges with the spread of the components. This is not the case in 1(b).

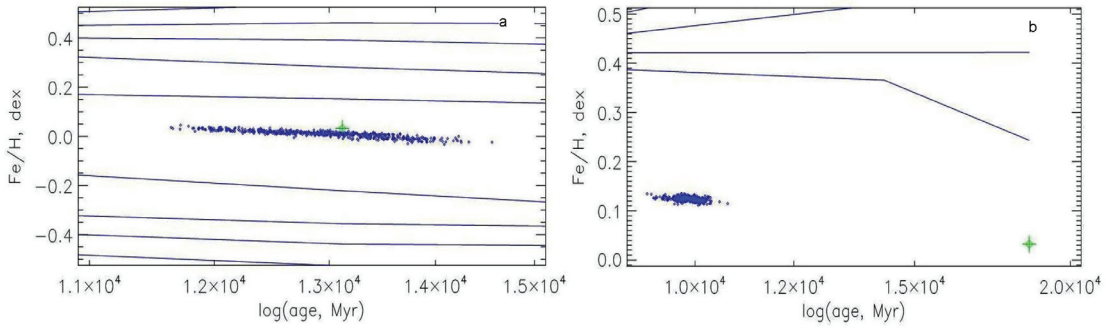


Figure 2. Stellar population of NGC 7597 as determined by the (a) P.HR and (b) V/M models. The local minima (best fit) is indicated by the green cross.

Figure 2 shows the solutions obtained by the P.HR and V/M models for NGC 7597. It is clear from the figures that the results indicated that one component provided the best fit. By comparing 2(a) and 2(b), the results are represented more accurately by the P.HR model because the χ^2-value (indicated by the green cross) converges with the spread of the component. Once again this is not the case for 2(b).

Table 2 summarises the average values for the ages and [Fe/H] of each stellar population of the galaxy sample as determined by the P.HR stellar population model. Only the results obtained by the P.HR model is shown here because the results for the stellar populations are more accurately represented by this model (see discussion of Figures 1 and 2). For the non-emission line galaxies, the 2 SSPs contains an ISP and OSP while the 3 SSP has a YSP, ISP and OSP. For the emission line galaxies, the 2 SSPs contains an ISP and OSP while there where no galaxies that showed that the 3 SSPs provided the best fit.

4. Conclusions

Between the two models tested, namely the P.HR and V/M stellar population models, the P.HR model gave a better representation of the SFHs of the BCG sample. An example is shown in Figures 1(a) and 2(a) where the local minimum, more specifically the χ^2-value (indicated by

Table 2. The average values for the ages and [Fe/H] of each stellar population as determined by the P.HR model for the galaxy sample. The errors indicated here is the standard error on the average values.

	Component	Non-emission Age (Gyr)	Emission Age (Gyr)	Total Age (Gyr)	Non-emission [Fe/H] (dex)	Emission [Fe/H] (dex)	Total [Fe/H] (dex)
1 SSP		9.768 ± 1.142	5.993 ± 0.401	8.509 ± 0.860	0.174 ± 0.046	0.158 ± 0.047	0.169 ± 0.034
2 SSP	First	2.493 ± 0.916	5.056 ± 1.958	2.813 ± 0.847	-0.851 ± 0.234	0.542 ± 0.157	-0.677 ± 0.236
	Second	17.268 ± 0.539	13.251 ± 2.076	16.765 ± 0.611	0.018 ± 0.163	-1.201 ± 0.823	-0.135 ± 0.191
3 SSP	First	0.674 ± 0.279	—	0.674 ± 0.279	-0.914 ± 0.382	—	-0.914 ± 0.382
	Second	9.442 ± 2.213	—	9.442 ± 2.213	0.055 ± 0.378	—	0.055 ± 0.378
	Third	16.970 ± 1.010	—	16.970 ± 1.010	-0.128 ± 0.266	—	-0.128 ± 0.266

green cross in Figures 1 and 2), of the each stellar component converges with the spread of the stellar populations. Hence only the results of this particular model will be interpreted in further analysis.

Hence, the stellar population model, P.HR, determined that the best fit for some of the galaxies is given by a CSP, indicating that at least some of these galaxies have a more complex evolution and formation than first thought. The large variety of stellar population ages and [Fe/H] points to a more diverse formation history for these BCGs.

Acknowledgments

The financial assistance of the South African Square Kilometre Array Project towards this research is hereby acknowledged. Opinions expressed and conclusions presented, are those of the authors and are not necessarily to be attributed to the National Research Foundation.

References

[1] Fabian A C 1994 *Annu. Rev. Astron. Astrophys.* **32** 277-318
[2] Loubser S I, Sánchez-Blázquez P, Sansom A E and Soechting I K 2009 *Mon. Not. R. Astron. Soc.* **398** 133-56
[3] Katayama H, Hayashida K I, Takahara F and Fujita Y 2003 *Astrophys. J.* **585** 687-93
[4] Sparke L S and Gallagher J S 2007 *Galaxies in the Universe: An Introduction* (Cambridge University Press: New York) pp 242-9
[5] Dubinski J 1998 *Astrophys. J.* **502** 141-9
[6] Gao L, Loeb A, Peebles P J E, White S D M and Jenkins A 2004 *Astrophys. J.* **614** 17-25
[7] De Lucia G and Blaizot J 2007 *Mon. Not. R. Astron. Soc.* **375** 2-14
[8] Chen Y, Reiprich T H, Bohringer H, Ikebe Y and Zhang Y Y 2007 *Astron. Astrophys.* **466** 805-12
[9] Loubser S I, Sánchez-Blázquez P, Sansom A E, Soechting I K and Bromage G E 2008 *Mon. Not. R. Astron. Soc.* **391** 1009-28
[10] Bildfell C, Hoekstra H, Babul A and Mahdavi A 2008 *Mon. Not. R. Astron. Soc.* **389** 1637-54
[11] Von der Linden A, Best P N, Kauffmann G and White S D M 2007 *Mon. Not. R. Astron. Soc.* **379** 867-93
[12] Crawford C S, Allen S W, Ebeling H, Edge A C and Fabian A C 1999 *Mon. Not. R. Astron. Soc.* **306** 857-96
[13] McNamara B R, Rafferty D A, Bîrzan L, Steiner J, Wise M W, Nulsen P E J, Carilli C L, Ryan R and Sharma M 2006 *Astrophys. J.* **648** 164-75
[14] Sarzi M, Falcón-Barroso J, Davies R L, Bacon R, Bureau M, Cappellari M, de Zeeuw P T, Emsellem E, Fathi K, Krajnović D, Kuntschner H, McDermid R M and Peletier R F 2006 *Mon. Not. R. Astron. Soc.* **366** 1151-200
[15] Koleva M, Bouchard A, Prugniel P and Wu Y 2009 *Astron. Astrophys.* **501** 1269-79
[16] Du W, Luo A L, Prugniel Ph, Liang Y C and Zhao Y H 2010 *Mon. Not. R. Astron. Soc.* **409** 567-81

The unusual Seyfert Markarian 926 – a link to the LINERs?

H Winkler

Dept. Physics, University of Johannesburg, PO Box 524, Auckland Park, 2006

hwinkler@uj.ac.za

Abstract. Markarian 926 (= MCG-2-58-22) is one of the earliest Seyfert galaxies identified. At discovery it was one of the most luminous nearby active galactic nuclei (AGN), with strong, wide broad lines. In the late 1980's it started fading, eventually settling at barely ~10% of its recorded peak luminosity. The luminosity decrease was accompanied by significant spectral changes, with the broad-line component now much weaker and highly asymmetrical. Low ionisation narrow lines, however, are now unusually strong, more typical of the AGN class referred to as LINERs rather than a Seyfert. This peculiar low-luminosity phase spectrum has remained relatively constant over the last decade. The paper discusses the spectral and other characteristics of Markarian 926 and compares these to standard Seyfert models. It furthermore investigates Markarian 926's relationship to the LINERs, and whether SALT observations of this object could shed light on the interrelationship between Seyferts and LINERs.

1. Introduction

Quasars, the most luminous objects in the universe, are widely believed to be powered by matter sucked onto an accretion disk surrounding a large black hole. The range in luminosity of these black hole-accretion disk systems is however yet to be accurately determined. In this regard, a class of active galactic nuclei referred to as the LINERs (Low Ionisation Nuclear Emission-line Regions) has yet to be fully understood [1]. LINERs are defined by the comparative weakness of their high-excitation spectral lines. Initially believed to be unrelated to the generally more luminous Seyfert galaxies and quasars, there is now a lot of evidence suggesting that LINER spectra can also be produced by photoionization from an accretion disk.

Markarian 926 (Mkn 926) was identified as a broad line Seyfert galaxy by Ward et al [2] during an investigation of bright x-ray sources. That work remarked about the asymmetric Balmer lines with broad line widths of over 20000 km.s^{-1} (FWZI), and the absence of Fe II bands. At that time only few Seyferts were known, and it was therefore not noticed that low-ionisation spectral features such as the [O I] 6300 Å line were rather strong compared to other members of that class.

2. Spectral and luminosity variations

2.1. Spectral variations

Numerous spectra have been recorded of Mkn 926 since its discovery spectrum, in 1978 [3], 1984 [4], 1987-1988 [5], 1990-1997 [6] and 2004-2005 [7]. Mkn 926 was also observed spectroscopically during the Sloane Digital Sky Survey (SDSS) in 2004 [8]. While already bright at discovery, Mkn 926´s luminosity was to increase by the mid-1980´s. It appears to have achieved its greatest nuclear brightness around 1984, when the upper spectrum in Figure 1 was recorded by Morris and Ward [4],

before generally declining in the late 1980´s. A weaker maximum was reached in August 1991 before the nuclear luminosity faded to a minimum in August 1994 [6]. Thereafter the spectrum maintained an appearance similar to the SDSS spectrum in Figure 1, although small changes in the broad line profile have seemingly been ongoing [7].

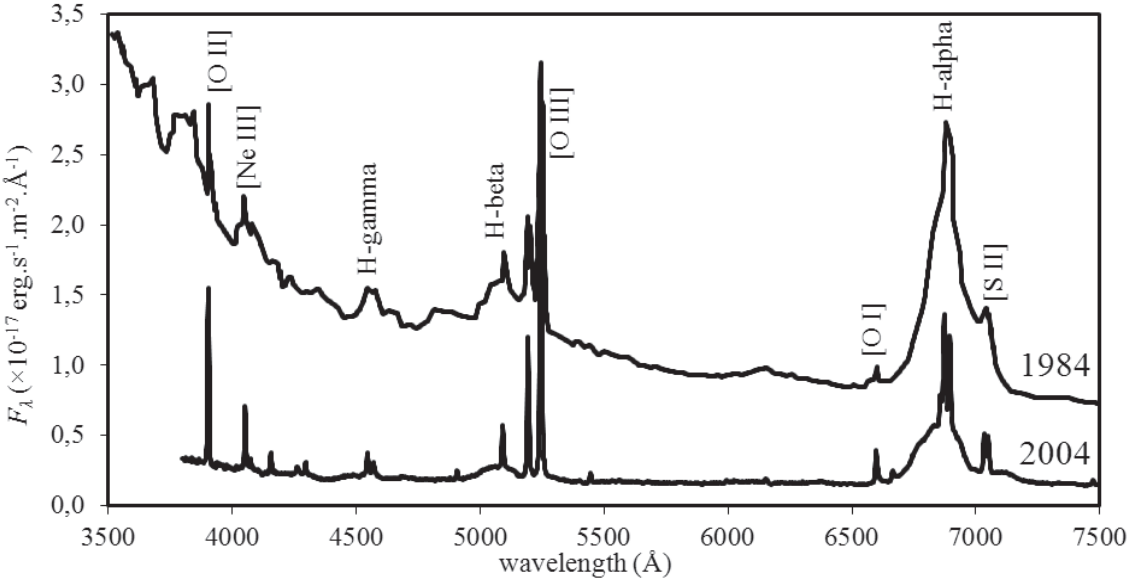


Figure 1. Flux calibrated optical spectra of Mkn 926 in 1984 (from Morris and Ward [4]) and in 2004 (from SDSS [8]). Differences in slit width and orientation are not expected to affect the nuclear flux in any significant way.

2.2. Photometric variations

Early *V*-band measurements suggest that the magnitude of Mkn 926 was approximately $V = 14$ mag (adjusted to 20˝ aperture diameter) in the early 1980´s [9], and then brightened to a peak of $V \sim 13.7$ mag [10]. A large set of visual photometric data recorded at the South African Astronomical Observatory (SAAO) [11, 12, 13] trace the fading of Mkn 926 from 1986. These data are complemented by SAAO infrared data [14] spanning the period 1980-1998, with one early measurement in December 1980 followed by reasonably regular points after July 1984. Light curves for the optical *V*-band and infrared *J*-band are displayed in Figure 2. Note how these light curves replicate the spectroscopic maxima around 1984 and 1991, as well as the long-term minimum that started in 1994.

3. Spectral characteristics

Spectral line strengths were measured in the 2004 SDSS spectrum that is characteristic of the low-luminosity epoch. The broad lines are difficult to measure accurately due to their comparative weakness, extreme width and asymmetry. The total (broad plus narrow components) H-beta line at 4861.33 Å corresponds to 0.70 times the line strength of the narrow [O III] 5006.84 Å line. This ratio had been 2.07 in 1984 [4]. The ratio of the total 6562.80 Å H-alpha line to H-beta was measured to be 5.08. The relative strengths of the narrow lines are in turn listed in Table 1.

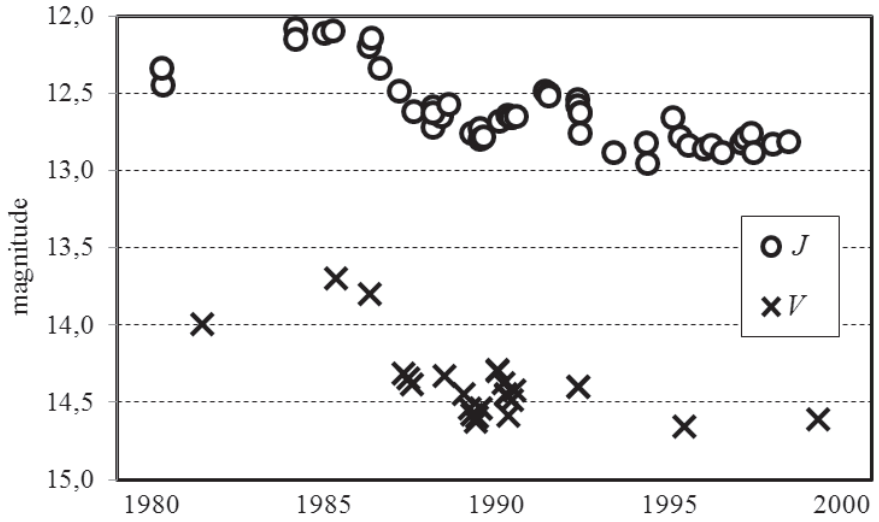


Figure 2. Infrared (*J*) and optical (*V*) light curves of Mkn 926 for the period 1980-2000. Data sources [9, 10, 11, 12, 13, 14]

Table 1. Narrow line strengths relative to the [O III] 5006.84 Å emission line

Ion	λ_0 (Å)	*I*/*I*(5007)
[O II]	3727[a]	0.372
[Ne III]	3868.74	0.109
[S II]	4072[a]	0.021
H I	4340.46	0.042
[O III]	4363.21	0.028
He II	4685.68	0.020
H I	4861.33	0.097
[O III]	4958.91	0.318
[N I]	5199[a]	0.020
He I	5875.62	0.011
[Fe VII]	6086.92	0.021
[O I]	6300.30	0.096
H I	6562.80	0.358
[N II]	6583.45	0.290
[S II]	6716.44	0.120
[S II]	6730.82	0.120

[a]representative wavelength for blended lines.

3.1. Broad line profiles

The H-alpha and H-beta line profiles are compared in Figure 3. The blue-shifted peak at –2000 km.s^{-1} and the bump near +10000 km.s^{-1} are similarly represented in both profiles. A weaker bump is also visible at –5000 km.s^{-1}.

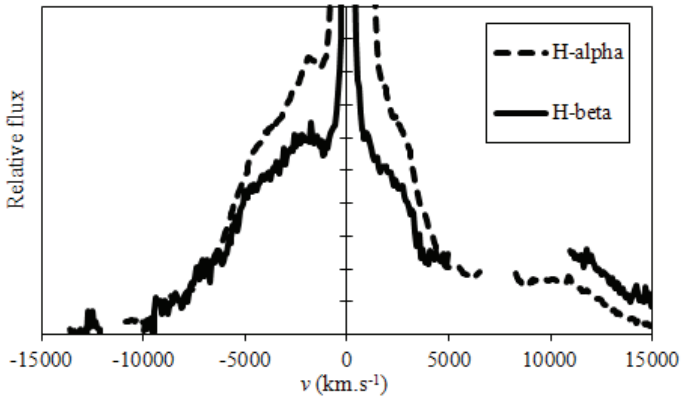


Figure 3. Comparison of H-alpha and H-beta broad line profiles. The H-beta line vertical scale was increased to ease the comparison between the two profiles

3.2. Narrow line profiles

Whittle [15] and Busko and Steiner [16] noticed [O III] asymmetry towards the red, virtually the only such object in their samples. This asymmetry is still noticeable in the SDSS spectrum, indicating that the basic structure of the narrow line region remains thus far largely unaffected by the nuclear variability.

3.3. Reddening

The ratio of H-alpha to H-beta is known to be of the order of 3, the exact value being dependent on the applicable excitation and recombination model. Using standard reddening formulas (e.g. Osterbrock and Ferland, 2006 [17]) leads to extinction estimates of $A_V = 0.62$ for the narrow line region and $A_V \sim 1.6$ for the broad lines. It is further highlighted that the red bump is comparatively more pronounced in H-beta, indicating that the part of the emission line region responsible for this emission is less reddened.

3.4. Gas conditions

The measured emission line ratio I([S II] 6717)/I([S II] 6731) of 1 corresponds to free electron density of $N_e \sim 10^{2.5}$ cm^{-3} ~ 300 cm^{-3} in the narrow line region (Osterbrock and Ferland [17]). The line ratio (I([O III] 4959)+I([O III] 5007))/I([O III] 4363) in turn determines the electron temperature of the region where the [O III] lines form. Applying the formula given in [17] implies that $T_e \sim 19{,}000$ K in the O^{++} region. These values are typical of Seyfert galaxies.

4. Discussion

4.1. Double-peaked emission line AGN

The strikingly wide, asymmetric Balmer-line profiles and associated bumps are characteristic of a group of AGN extensively studied by Eracleous, Halpern and co-workers [18, 19]. This type of AGN includes bright quasars as well as fainter objects with a more hidden nuclear core. The list of objects even includes LINERs and radio galaxies, such as Pictor A. These studies were able to explain the profile changes through warped elliptical accretion disk models as well as through disks with spiral arm-like irregularities. The detailed study of spectral variations in Mkn 926 by Kollatschny and Zetzl

[7] does not refer to these works, but highlights a very similar broad-line behaviour. Thus Mkn 926 is a new member of this class of AGN. It is a particularly important example of the class due to its comparative proximity, and because of its documented history of extreme luminosity changes.

4.2. Seyfert or LINER?

The high-ionisation [O III] and [Ne III] lines have been comparatively strong throughout the spectral evolution, and thus Mkn 926 is technically not a LINER (which would require I([O II] 3727)/I([O III] 5007) > 1 and I([O I] 6300)/I([O III] 5007) > 1/3 [1]). The [O I] and [S II] lines are however more prominent than in most Seyferts and quasars, especially when compared to the H-alpha-[N II] complex during the faint stage. As mentioned, the broad line profile and variations also resemble some LINERs. It is plausible that the light echo from the bright phase in the mid-1980´s is still crossing much of the narrow-line region, and that therefore the [O III] lines still appear bright. If the broad line region stays faint then it should be possible to monitor the eventual fading and profile changes of these lines, which in turn will offer a unique opportunity to map the narrow line region of a Seyfert galaxy.

4.3. Potential future work

The light collecting potential and high resolution spectroscopic capacity of the Southern African Large Telescope (SALT) could be used to characterize the narrow line shapes at different ionisations, and thus unravel the density and temperature gradients in the outer parts of the nucleus. As the outer nuclear region reflects and reacts to the reverberated and time-delayed behaviour of the central power source, such measurements could shed light on the behaviour of the nucleus before the discovery of the Mkn 926's AGN nature. SALT data would also highlight the latest changes in the broad line profiles. In conclusion, regular monitoring of the optical luminosity would indicate whether the luminosity decline is permanent or whether bright phases are a common occurrence.

Acknowledgements

This paper utilized data from the Sloan Digital Sky Survey (SDSS). Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

References

[1] Filippenko A V 2003 *ASP Conference Series* **290** 369
[2] Ward M J, Wilson A S, Penston M V, Elvis M, Maccacaro T and Tritton K P 1978 *Astrophys. J.* **223** 788
[3] Durret F and Bergeron J 1988 *Astron. Astrophys. Suppl.* **75** 273
[4] Morris S L and Ward M J 1988 *Monthly Notices Roy. Astron. Soc.* **230** 639
[5] Winkler H 1992 *Monthly Notices Roy. Astron. Soc.* **257** 677
[6] Kollatschny W, Zetzl M and Dietrich M 2006 *Astron. Astrophys.* **454** 459

[7] Kollatschny W and Zetzl M 2010 *Astron. Astrophys.* **522** 36
[8] Abazajian K, Adelman J, Agueros M, et al. 2003 *Astron J.* **126** 2081
[9] Mallama A D 1983 J. A. Variable Star Observers **12** 69
[10] Hamuy M and Maza J 1987 *Astron. Astrophys. Suppl. Series* **68** 383
[11] Winkler H, Glass I S, van Wyk F, Marang F, Spencer Jones J H, Buckley D A H and Sekiguchi K 1992 *Monthly Notices Roy. Astron. Soc.* **257** 659
[12] Winkler H 1997 *Monthly Notices Roy. Astron. Soc.* **292** 273
[13] Winkler H unpublished
[14] Glass I S 2004 *Monthly Notices Roy. Astron. Soc.* **350** 1049
[15] Whittle M 1985 *Monthly Notices Roy. Astron. Soc.* **213** 1
[16] Busko I C and Steiner J E 1988 *Monthly Notices Roy. Astron. Soc.* **232** 525
[17] Osterbrock D E, Ferland G J 2006 *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei* 2nd Ed. (University Science Books, Sausalito, California)
[18] Eracleous M and Halpern J P 2003 *Astrophys. J.* **599** 886
[19] Lewis K T, Eracleous M and Storchi-Bergmann T 2010 *Astrophys. J. Suppl.* **187** 416

Division E – Physics Education

Contrasts between student and examiner perceptions of the nature of assessment tasks

D Clerk, D Naidoo and S Shrivastava

School of Physics, University of the Witwatersrand, Private Bag 3, WITS 2050

douglas.clerk@wits.ac.za

Abstract: For several years an on-going study has examined student performance in relation to the various *types* of assessment task in first year physics examinations. Four basic task types: "routine operations", "novel problems", "intuitive - interpretive questions" and "bookwork" are recognised. Average student performance has usually been strongest for routine operations, variable for bookwork, weak for novel problems and weakest for intuitive – interpretive questions. The purpose of the present study is to examine the question: Do students experience an assessment task the way the examiner intended? – i.e. if a question was intended by the examiner as a routine operation, do the students experience it as a routine operation or as something else, such as a novel problem? Examiners (n = 4) who set, and students (n = 177) who wrote a physics test in May 2011 were asked to classify the test questions according to the typology above. The resulting student and staff classifications were then compared. For most students there was no significant mismatch with the staff classification, although there was some disagreement on an individual level. Also found: a given assessment task can comprise a mix of several of the four types.

1. Introduction:

Previous studies [1 - 3] have investigated the performance of students in relation to assessment task type and have attempted to find reasons why student performance is weaker in heuristic than in algorithmic problem solving. Additional data now enables us to further develop our question typology and to answer an intriguing question: how reliable is the judgement of an examiner concerning the type of question experienced by the student – i.e. to what extent do students experience assessment tasks in ways *un*intended by the examiner?

An online literature search failed to provide a suitable ready-made typology – for example the Rhodes' typology [4] was considered but although it was promising, it was not deemed suitable for our purposes. While it is *'extremely well suited for use in Socratic dialogues'* (p 10) it is intended for content-based questions during the ***learning*** process with the objective of teaching the students to ask their own questions. Hence, it does not provide a comprehensive set of answers to the question: "Faced with a given ***examination*** question, what exactly does the student need to ***do***, when attempting to generate the required answer?" The Rhodes' typology [4] goes some way to giving a list of possible answers to this question but, apart from not being intended for examination questions, it omits at least one important operation: solving the novel problem [5] and hence seems not fully suited to our purpose. Hence we set about devising our own typology, as follows: A given examination question can require the student to perform a variety of possible basic operations, these are summarised below:

Table 1: Question Typology.

Type:	Abbreviation	Description
"Bookwork"	**BW**	Material (excluding formulae) has to be recalled from memory and presented to answer the question.
"Intuitive - Interpretive"	**INT**	Given information is used to produce a reasoned, possibly non-numerical answer – such as an explanation or a prediction or a justification, etc.
"Routine Operation"	**RO**	Familiar, well practiced (mathematical) procedures are used to produce the answer.
"Novel Problem"	**NP**	The student uses a procedure that (s)he has never used before, or else ***in a way*** (s)he has never used it before, to produce the answer.

This typology is at present very much a work in progress and further developments – such as, the possible inclusion of additional categories, or perhaps the subdivision of the existing categories - will in time be forthcoming.

The question currently under consideration concerns the validity of making statements about student performance in the various question types, such as: "students are good at routine operations" or "students are weak at novel problem solving". There are two possible complications: Firstly, a student might experience a given question in a different way to what was intended by the examiner – e.g. an examiner might have intended a question to be a routine operation. A correct answer would normally be interpreted to imply that a student has proven him or herself capable of performing that routine operation. But what if a significant fraction of the students experienced that question as a novel problem instead - simply because they had not prepared sufficiently well to be able to use the usual algorithm? At least some of those answering the question correctly could have been proving that they were able to solve a novel problem, which would mask the fact that they were not able to perform the routine operation. How would the examiner know the difference? Secondly, a given question might involve more than one (possibly all four) of the categories above. For example, a question which is basically intended as intuitive – interpretive might require the student to recall some piece of material - in which case there is also an element of 'bookwork' involved. A student who cannot – through poor preparation – recall a key law, definition or other key piece of information, might *for that reason* be unable to produce an answer to what was intended as an intuitive – interpretive question. How would we know *what* to conclude about that student's strengths and weaknesses and hence, what form of remediation might be appropriate? It is thus important to know to what degree students writing a test or examination experience the questions as intended by the examiner. An analysis of a first year physics examination at mid-year in 2010 yielded the following:

Table 2: Student performance by question type: mid-year 2010.

Question Type: (as deemed by examiner)	Percentage contributed to the total mark	Average student mark
Bookwork	6%	77%
Intuitive - interpretive	4%	18%
Routine Operation	69%	59%
Novel Problem	21%	40%

These results might possibly tempt one to the following conclusions:

- Students performed the best at bookwork.
- Student performance at novel problem solving was clearly lower than at routine operations.
- The marks were lowest – dramatically so – for the intuitive / interpretive questions.

But are these conclusions reasonable? Certainly, 77% *looks* different from 18%, even without a statistical test of significance but what about 59% as opposed to 40%? Caution dictates that the two questions mentioned above should be addressed before concluding anything from such data.

To attempt to answer these two questions a class test written in May this year (i.e. 2011) was classified according to question type by: the examiners who set the questions, by other teaching staff who are familiar with the course, and by the students who had written the test. We then set about comparing staff opinion with student opinion. The course involved was a first year physics course for engineering students. Participants were not randomly selected. This was because:

- The number of suitable staff available was small (n = 4) – we used *all* available staff with familiarity with the course – i.e. who would have the necessary insight; from having taught on the course - to be able to make a reasonable judgement on question type. The staff participants were thus *purposively* selected.
- The student respondents (n = 177) were students attending the "academic development" tutorial sessions during the week following the writing of the test and were thus effectively self-selected. It should be noted that this group of students represents a range of performance levels from extremely poor to excellent.

Each participant was asked to rate each of the questions in the test as shown in the example (question 2 from May 2011 test paper) below:

A telephoto camera lens consists of a converging lens with a focal length of 300 mm and a diverging lens with a focal length of – 100 mm, separated by 275 mm. When a photograph is taken, light from a distant object must pass first through the converging lens, then through the diverging lens and then must form a real image on the sensor. How far from the diverging lens must the sensor be situated to form a focussed image?

Question type:	Featured:				
	Not at all	Slightly	Moderately	Strongly	Exclusively
Memory / Bookwork:					
Intuitive / Interpretive:					
Routine operation:					
Novel problem:					

Figure 1: Selection from May test with evaluation grid for survey.

This question was selected for discussion here because it yielded the worst average performance of the test: 8%. From the examiner's perspective this was essentially a routine operation and as this question type has in the past usually yielded good average performances from the students (see table 2), the exceptionally low score actually achieved is puzzling to say the least. Perhaps the students didn't experience this question as a routine operation, but as something else.

The respondents were asked to consider each question and then to mark the blocks in the grid according to their perception of the question. For example a student who remembered this question as nothing but a routine operation would have marked the block under 'exclusively' in the 'routine operation' row and marked all other types as 'not at all'. If the student felt that there was also an element of, say, interpretation and intuition involved, then the respondent would have marked a non-zero response in this row as well. The responses were given numerical values as follows:

Table 3: Numerical equivalents for responses used in typology survey.

Featuring:	Occurrence Rating
Not at all	0
Slightly	1
Moderately	2
Strongly	3
Exclusively	4

The occurrence ratings provided by the respondents were entered into a spreadsheet for numerical analysis, in which we attempted to assess to what extent the students experienced the test questions as the examiners had intended.

2. Findings:

Three caveats:

- The respondents were not randomly selected and -
- the numerical values used (see Table 3 above) are ***codes***, not measurements and -
- were based on highly subjective decisions on the part of the respondents.

Hence, we should be cautious about reading too much into the data. Nevertheless, some useful interim conclusions can tentatively be drawn:

2.1 Individual questions

Question 2 (discussed above) yielded the following result:

Table 4: Average occurrence ratings for question 2.

Question 2:	Occurrence Rating		Conflated occurrence rating	
Question Type:	Staff Average	Student Average	Staff Average	Student Average
Memory / Bookwork:	1.0	1.3	3	3
Intuitive / Interpretive:	2.0	1.7		
Routine operation:	3.0	1.4	3.5	2.7
Novel problem:	0.5	1.3		

The students' average rating for 'routine operation' was less than half that of the staff; also, the rating for 'novel problem' is higher for students than for staff. As past studies [3] have suggested that students are generally better at routine operations than at novel problem-solving, perhaps this could explain the low average performance. However, this might be unwarranted. Firstly, with an average performance of only 8% for this question, surely there would have been a *much* higher discrepancy between the staff and student ratings than there is? Second, most of the apparent discrepancies can be made to disappear by as simple an operation as rounding off the average ratings.

2.2 The test as a whole

Further analysis involved answering the question: "what percentage of each category of respondents gave each of the question *types* a given occurrence rating across the test as a whole?" The answer to this question is summarised below:

Table 5: Average overall occurrence ratings for staff and students.

Type	Staff					Students				
Memory / Bookwork:	18%	13%	25%	10%	35%	27%	11%	15%	20%	27%
Intuitive / Interpretive:	38%	18%	25%	20%	0%	41%	12%	17%	21%	9%
Routine operation:	45%	10%	3%	23%	20%	44%	12%	17%	17%	11%
Novel problem:	80%	5%	13%	3%	0%	70%	10%	7%	8%	5%
Occurrence rating:	Not at all	Slightly	Moderately	Strongly	Exclusively	Not at all	Slightly	Moderately	Strongly	Exclusively

The data summarised here seem to show that mostly, the students agree with the staff on what a given test question amounts to by way of question type. Hence, while there are individual differences between staff and students – i.e. there *are* individual students who, for example, perceive as a novel problem a question which the examiner had intended as a routine operation – the overall situation is that the majority of the students agree with the staff consensus about what type of question they are answering. Compounding the situation is the fact that many questions were of mixed type making a clean separation between types difficult to justify. Conflation of question types frequently confused by students (see table 4) shows a good correlation between the opinions of staff and students, as shown in Figure 2. To produce the plot below, we conflated

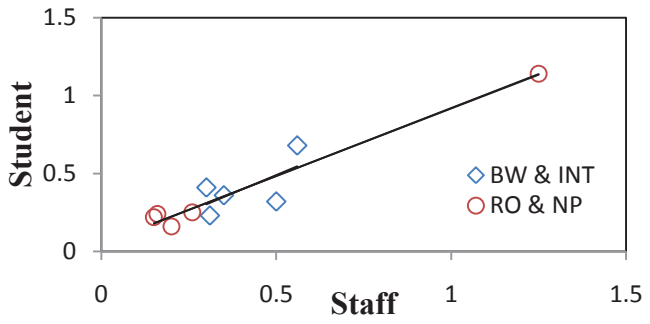


Figure 2: Scatter plot of staff versus student ratings for conflated question types. (Diamond: Bookwork (BW) and Interpretive questions (INT); Circle: Routine Operations (RO) and Novel Problems (NP)).

For each of these two conflated categories student ratings were plotted against staff ratings to produce the scatter plot above. The resulting positive gradient in the trend-line indicates that the staff and student ratings are positively correlated.

2.3 Conclusions

Caution is still needed when drawing conclusions about students' ability to perform according to question type as, apart from the three caveats mentioned above, some examination questions involve the student in more than one category from the typology. In these cases we need to be cautious about what we conclude from students' average performance in a given question. However, it does seem that while there will be exceptions, we can with reasonable confidence regard our own judgement about question types to be reliable. If an examiner deems a question to be a routine operation, the majority of the students can be expected to experience it as a routine operation as well. Pending further data, we suggest that examiners nevertheless be cautious when dealing remedially with individual students as there *are* exceptions – some students *do* experience assessment tasks in ways not intended by the examiner.

References:

[1] Clerk D and Naidoo D 2009 *Proceedings of the Seventeenth Annual Meeting of the Southern African Association for Research in Mathematics, Science and Technology Education* 72-76

[2] Clerk D and Naidoo D 2010 *Proceedings of the Eighteenth Annual Meeting of the Southern African Association for Research in Mathematics, Science and Technology Education 142-149*

[3] Clerk D and Naidoo D 2010 *55th Annual Conference of the South African Institute of Physics*

[4] Wenning C J, Holbrook T W and Stankevitz J 2006 *Journal of Physics Teacher Education Online* **4**(1) 10 – 13

[5] Martinez M 1998 *Phi Delta Kappan* **79** 605-609

Visualizing Physics: Good Vibrations

D Fish[1], **T R Anderson**[2]

[1] Director: Unizul Science Centre, University of Zululand, KwaDlangezwa, 3886
[2] Head, Science Education Research Group (SERG), School of Biochemistry, Genetics and Microbiology, University of KwaZulu-Natal (Pietermaritzburg)

E-mail: thefish@iafrica.com

Abstract. Science Centre around the world routinely present science shows to visitors, but seldom evaluate their effectiveness, or adapt them for different audiences. After running a Science Centre for 20 years, the author is investigating perceptions of Physics in school visitors to Unizul Science Centre. Visitors typically fall into three distinct groups with very different needs and responses: those coming from: rural schools, township schools and urban schools. In this study, pupils were exposed to an interactive Science Show: Good Vibrations, presenting the basics of sound and waves through musical instruments, and then asked various questions about the show. The conclusions reached after analysis will be used to inform best practice when dealing with different groups of visitors with different needs.

1. Introduction and Methodology

This paper is based on a pilot study conducted during the "Science Unlimited" festival at the Royal Showgrounds in Pietermaritzburg in October 2010. A 45 minute science show was presented to various audiences (both primary and high schools) of about 250 pupils, but for the survey pupils were chosen from high school only, and comprised:

- Urban Boys School (very privileged) – 21 Grade 9 pupils
- Township Co-educational School – 19 Grade 11 pupils
- Rural Co-educational School – 18 Grade 10 pupils

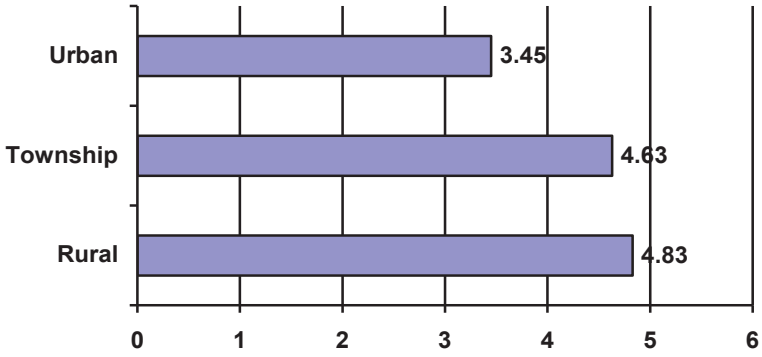
After watching the show, pupils were asked to respond to questions requiring both written and graphical answers. From these, conclusions are drawn as to the ability of these pupils to understand sound both visually and conceptually.

The show: "Good vibrations", is a dynamic, highly interactive science show. It covers the science of sound, presenting it through the medium of music and musical instruments. It does this in an integrated, multimedia fashion, using instruments, simulations, graphics and video, volunteers, sounds and practical applications.

2. Discussion of Survey Results

2.1 Questions dealing with impressions of the show, novelty and links

2.1.1 Question 1: "How much did you enjoy the show on a scale of 1 (awful) to 5 (excellent)?"

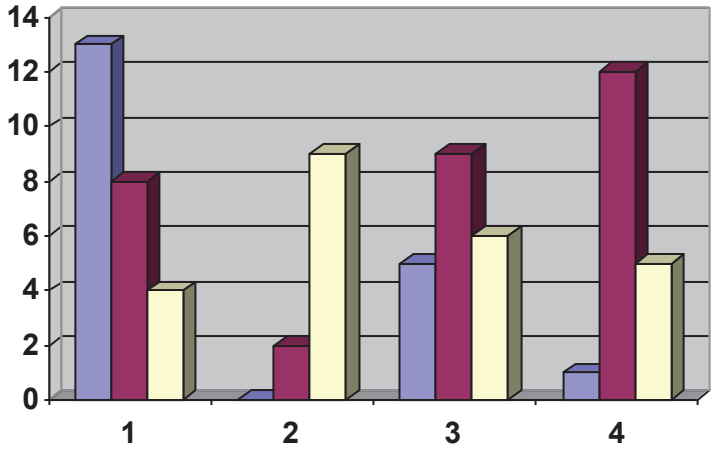


School-type vs. Enjoyment (1-5)

The more privileged urban school rated the show lower, perhaps because they had higher expectations of it (rather than that they enjoyed it less). This also seems to be born out in the choice of words for answering question 2 below. Township and urban schools rated the show almost equally highly.

2.1.2: Question 2: "At home tonight, what will you tell your friends/ family about from the show?"
These answers were classified into four groups:

1. Positive comments about the show (although the language used here differed significantly)
2. Comments of what was learnt, gratitude
3. Intention to pass on some information or learning to family
4. Information/ excitement about the different instruments seen



Number of Comments vs. School-type (in answer-groups)

Again it is fascinating to see the different expectations of the show born out in the responses.

- Urban pupils mostly (62%) reported that they would tell family/ friends that the show had been good (1), but none said they would report how much they had learnt (2)! Only 5 expressed an intention to teach something of what they had learned to people at home, or pass on the knowledge (3). And only 1 wanted to share his excitement at all the different instruments seen

(4), indicating that this was probably not a new or novel experience. ("..that it was interesting and filled with interesting facts that I learned")

- Township pupils to a lesser extent (42%) reported that they would tell family/ friends that the show had been good (1), and only 2 said they would report how much they had learnt (2)! But here almost half expressed an intention to teach something of what they had learned to people at home, or pass on the knowledge (3). The most common response was that 12 of them (63%) wanted to share their excitement at all the different instruments seen (4), indicating that this was definitely a new or novel experience, and had made a big impact. ("I will teach them that different musical instruments have different pitch and there wavelength is not the same")
- Rural pupils: very few (22 %) reported that they would tell family/ friends that the show had been good (1), but half of them (50 %) said they would report how much they had learnt (2)! One third of them expressed an intention to teach something of what they had learned to people at home, or pass on the knowledge (3). And 5 (28%) wanted to share their excitement at all the different instruments seen (4). ("I will tell them that I have learned about kinds of flute and learned types of guitars")

It is also interesting to break down group 1 further, and look at the different language used in expressing positive comments about the show:

- Urban: interesting (x10), good, cool, OK
- Township: awesome, fantastic, nice, fun, amazing, new
- Rural: wonderful, surprising, important, practical

The words were peculiar to each school group with almost no overlap! Urban pupils focused on the show being interesting (10 responses) but this word was not used even once by either of the other groups. Township pupils focussed on fun, cool words. Rural pupils seemed to see the whole experience in a far more serious light, mentioning that the presentation was important and practical. These different expectations should definitely be kept in mind when presenting shows to these different groups.

2.1.3 Question 3: "What surprised you most about the show?": There was a fair bit of overlap between the next three questions (3, 4 and 5), indicating that they probably need to be more clearly stated to get more useful answers. It is interesting to note that Urban schools mentioned being surprised by more technological aspects (the data studio software and ultrasound) while rural schools were surprised at the notion of making musical instruments from waste or simple things – something one could expect them to be more used to! Both urban and township pupils mentioned instruments by name (or description) in response, whereas not one rural pupil did! The notion that one needs vibrations for sound was surprising to township and rural schools, but not to urban pupils. This would suggest a presentation focusing more on the technology for urban schools, and on instruments – especially simple, recycled ones, for township and rural schools, in order to get the greatest novelty or surprise effect. Some sample comments:

U: "That music is scientific"
"The way a flute bends waves of sound"
T: "Is that noise and sound are too different things"
"Is that before the sound come out the vibration comes first"
R: "… to see different types of instruments and how they work"
"… that if the [ ] is long the sound is low"

2.1.4 Question 4: "What new thing did you learn today?": Again comparisons are very interesting: T and R schools noted learning the *new* fact that vibration is needed to make music or sound, whereas no U pupil mentioned it. Both groups also equally reported learning anew that different instruments produce different sounds and the key concept: "long is low, short is high", neither of which was new to the U group. These are fairly basic or lower level points, perhaps indicating that one should adopt a

lower entry level for these groups (especially considering that the average age of these groups is higher than the U group). The R group echoed the Q3 results in highlighting different or new instruments. It is always debatable how many instruments to use in the show. Probably there are too many at present, but the R group especially seemed to be very excited at this opportunity to see them (possibly once in a lifetime). Once again (as in Q3) the R group did not mention range of hearing (ultrasonic sound) as being new (or surprising in Q3), possibly because they did not fully grasp the concept. The same goes for the ultrasonic cellphone ring. This could be further probed through appropriate content questions. Comments were very similar to Q3.

2.1.5 Question 5: "What thing you saw today reminded you of something you already knew?" : It is always good to tie a presentation to pupils' prior experiences, and this music show is ideally suited for that, as everyone has some experience of music and musical instruments. The guitar and piano were clearly the best known, and should be kept in the show no matter how mundane they may seem as they provide a common reference point across all three school groups. I was very interested that the R school seemed more familiar with the piano, and the U school with the vuvuzela – quite the opposite of what I would have expected. Again the majority of the R school's responses had to do with musical instruments, showing that this is a clear contact point for them.

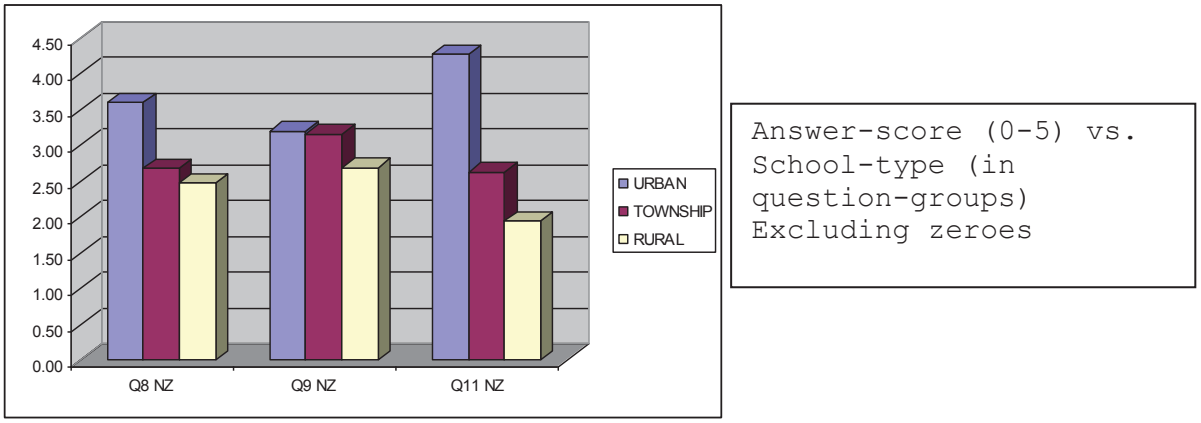
I was interested that the T and R schools mentioned the show reminding them of graphs, even specifying trig graphs from Maths. It is very pleasing to see this link being made, but this may have been obscured by other pupils just mentioning an easier answer like a familiar musical instrument. A dedicated probe would reveal interesting answers here. Perhaps: "Did the show remind you of something you have covered in another class at school?"

2.2: Conceptual Questions testing understanding of sound and waves

Answers for these three questions (8, 9 and 11) were tested against model answers (provided by myself as subject expert) and classified according to a 6 point scale – see table below. This scale is based on a *four* point scale suggest in Falk and Gillespie (2009) [1]. I felt the need to widen the scale more as these questions essentially had an answer (yes or no) *and* an explanation.

SCORE	ANSWER
0	Did not answer
1	Gave an incorrect answer
2	Gave correct answer but not explained (or incorrect reason)
3	Gave correct answer, used correct words but without understanding
4	Gave correct answer with good but incomplete explanation
5	Gave correct answer with correct explanation (but not necessarily perfect!)

Answers varied greatly and were generally not very good. In many cases pupils used correct technical terms but not as part of coherent explanations: "It is because of vibrations". Language is clearly a huge problem for the rural and (to a lesser extent) township schools, and it was clear that they were struggling both to understand the question and to state their answers. In many cases rural pupils simply restated the question in their answer, ie: Why are men's voices deeper?" was answered "It is because they are deeper!" The results of all three groups for all three questions are graphed below:



From these graphs there is a clear decrease in answering ability from U, through T to R. This is most marked in Question 11, which surprised me, as this question was explicitly dealt with in the show! There are many questions which arise from this result. Are the R schools learning much (as they expressed in Q2), but simply unable to relate it in a foreign language, exam scenario? Clearly three questions to 58 pupils is not nearly enough to answer these questions fully, and different types of questions should be used to exclude language and other effects.

It must also be born in mind though that the best results came from the youngest students in the lowest grade, as we had:

- Urban Grade 9 Average age 15.0 3 question average (excl 0's) 3.69
- Township Gr 11 Average age 17.1 3 question average (excl 0's) 2.82
- Rural Grade 10 Average Age: 15.9 3 question average (excl 0's) 2.37

In addition these U pupils would probably NOT have covered sound yet in class (as it is taught in Grade 10), while the other two groups should have covered it. There was little variation of any importance in the answering of the specific questions, and where a school group did seem to "get it" (the U school with Q11), they seemed to "get it" pretty much right across the group.

To be sure of presenting at an appropriate level for the pupils in terms of their knowledge and cognitive grasp of sound, it is clear that allowances will have to be made for rural groups. The presentation for them should probably contain more repetition and less depth to ensure that key concepts are grasped. With U schools one can safely assume a higher starting point, and can safely delve into topics in more depth, further stimulating the interest which they expressed in Q2.

2.3 Conceptual questions: detailed explanation and drawing or visualization (6, 7, 10)

2.3.1 Question 7: This asked pupils to complete a drawing showing how one man hears another man shouting. The results were fascinating, with not a single pupil producing a correct picture of sound waves (along the lines of ripples in a pond). Many chose not to answer this question, and those who did drew a representation (sine wave) of a wave rather than the reality. By far the most common choice was a basic sine wave representation (26 out of 58 pupils – almost half). 12 pupils didn't answer the question – coming especially from the township school. I was intrigued at how many of them drew loops, with the wave doubling back on itself. Pictures resembling bouncing were also interesting, perhaps representing a picture they have seen before of a ball bouncing along the ground. A few also drew multiple waves together. Two pupils drew sound waves which arrived at the ear but travelled only half or less of the distance, not emanating from the shouter! At this stage one can only speculate at the reason for these drawings, but it would be good to have the opportunity in a follow up interview to explore further the reasons for these visualizations.

3. Conclusions

Despite having performed this show for over 10 years to many thousands of pupils, this brief study has made me rethink many of the assumptions I had been making. Science pupils in South Africa are at a very low level of knowledge and understanding (TIMSS) and one must assume very little prior knowledge when presenting a show to them.

Urban pupils appear to have very high expectations of a science show and are consequently harder to please. There is less which is novel to them, and they seem to be most enticed by what especially interests them. They expressed no particular intention to pass on the knowledge or experience to family or friends, perhaps because they came from educated families. Surprise or novelty seemed to come mostly from the technological aspects of the show. Urban schools clearly started from a better conceptual base, so one could present at a higher level, and possibly extend the show into greater detail (e.g. tackling harmonics). The Urban schools also benefited from good English language capability, which obviously aids understanding.

Township pupils had lower expectations of the show and were consequently easier to please! They seemed to like the aspects of the show which they thought were "awesome or amazing". They seemed to be intrigued by the instruments and expressed a strong intention to share this excitement with family and friends. Township pupils differed from urban pupils in expressing surprise or novelty at the link between sound and vibrations. Their conceptual understanding was much weaker than urban schools, and not much different from rural! A large number of township pupils did not make drawings when requested to do so. It would be interesting to probe further if this is lack of ability to draw, or a feeling that it is childish or beneath them?

Rural pupils rated the show very highly and clearly saw the occasion as very serious and important (judging by their choice of descriptive words). They clearly saw the show as a profound learning experience and expressed the intention to tell friends and family about this and to pass on some of their knowledge. This has positive implications for our HIV/Aids show where we plan to use pupils as ambassadors into their communities. The rural pupils definitely related better to the instruments than to the computer simulations and it would appear that concrete practical equipment will be more effective with these groups. Rural pupils were very weak conceptually and shows should focus on reinforcing basic concepts and not focussing too much on counter-intuitive or surprising facts, as suggested by Lee (2003) [2] Language also makes understanding more of a problem.

These brief insights make a strong case for science show presenters to be regularly involved in action research on the impact of their shows, and best practice for different school groups. Sadler [3] has shown that this retrospection leads to improvement in both practice and impact.

4. References

[1] Falk JH and Gillespie KL 2009 *Visitor Studies*, 12(2), 112-132
[2] Lee O 2003 *Teachers College Record* Vol 105 No. 3 pp 465-489
[3] Sadler W 2004 *Masters Thesis*

Active Learning in the Third Year Statistical Physics Module at the University of the Witwatersrand

Jonathan M. Keartland
School of Physics, WITS 2050, South Africa
E-mail: Jonathan.Keartland@wits.ac.za

Abstract. Active learning techniques were used for the first time in the Statistical Physics module presented to third year students at the University of the Witwatersrand in the first semester of 2009. The reaction of the students, recorded in a formal Student Assessment of Lecturer Performance, administered by the Centre for Teaching and Learning Development at the University, was overwhelmingly positive. In addition, the final results of the students were excellent. This paper reports on the methods employed, the results of the student survey and the module results. These results are compared with the results of the same group of students in the second year Thermal Physics module, taught by the same lecturer using more traditional methods of instruction, during the second semester of 2008. The employment of innovative teaching strategies makes a significant difference to the confidence and attitude of the students, and an increase in the level of performance of the majority of students is apparent. Active learning is now the standard method of instruction in both the second and third year Thermal and Statistical Physics modules, and details of the materials developed over the last three years will be provided.

1. Introduction

A recent review of the research into Active Learning techniques [1] examines the evidence for the effectiveness of active learning. Four types of active learning are identified, these being active learning, collaborative learning, cooperative learning and problem-based learning. In essence, the types of teaching interventions that fall under the heading of active learning involve changing the approach of the lecturer from the traditional model, where a lecturer imparts knowledge and the students in the lecture absorb the information. The introduction of active learning techniques have, in the past, lead to a number of beneficial outcomes. Prince quotes several studies which claim improved student attitudes, an improvement in skills development and retention of knowledge. He also cautions that empirical evidence of improvement in student performance may be less convincing than expected. Prince also points out that the adoption of active learning within a discipline is often treated with scepticism by those teachers within the discipline who regard changing the approach to teaching as following fashionable educational trends, and not based on hard empirical evidence.

Felder *et al* [2] make a stronger case for the adoption of teaching methods that they, and the works that they cite, believe facilitate learning. These include the publication of clear instructional objectives for each course, establishing of the relevance of course material, balancing concrete and abstract information, promoting active learning, using co-operative learning, giving challenging but fair tests and examinations and conveying concern about the students' learning.

They also point out that the techniques that they advocate can be implemented without large investments in infrastructure (such as computers). This has a particular resonance in South African Universities, where funding for infrastructure may be limited.

There have been many studies of active learning teachniques applied to large first year undergraduate classes in Physics and Engineering (see for example the review by Thacker [3]). It has been less common to find these techniques used in upper level Physics [4]. This is perhaps related to the fact that these classes are usually much smaller than first year classes, and that the courses will often be taught by senior staff who may be more inclined to use methods that they have found effective during the course of long careers.

This paper reports on an attempt to introduce active learning techniques in the third year Statistical Physics module at the University of the Witwatersrand (WITS) during 2009. It will be shown empirically that the introduction of active learning has a profound effect on student attitudes and motivation, and that the results of the students in the module improved when compared to their results in the Thermal Physics module presented in the previous semester (at the end of 2008). Active learning methods are now utilized in both of these modules, and the methods adopted continue to play a positive role in facilitating effective learning.

2. Teaching Strategies

The teaching strategies employed in the first instance were inspired by a workshop given by Richard Felder at WITS in December 2008. I had, over the years, attempted to bring a more interactive approach to lecturing, and met with some success. As Table 1 in the next section shows, the students in the second year Thermal Physics module recognized my intention, but obviously did not feel comfortable in becoming actively involved in the lecture. Despite asking the obvious 'Any questions?' at the end of each lecture example, derivation and lecture, I received very little feedback. There was also an obvious drop in student attention during the course of each 45 minute lecture. The methods described by Felder in the workshop, and in his publications, offered me the possibility of improving communication with students in demanding upper level courses. I introduced active learning explicitly during the first lecture of 2009, so the students were fully aware that they were a test group. It should be noted that the module curriculum was not altered significantly from the Statistical Physics module presented in 2008.

2.1. Lecture Activities and Co-operative Learning

Felder and Silverman [5] point out that different students have different learning styles, and that many students have problems dealing with abstract information. Moreover, it is clear [1] that the attention of students in a lecture drifts away from the lecture material after approximately 20 minutes. Lecture activities allow the lecturer to break the monotony by introducing a set activity. These activities varied. Sometimes they were applications of material presented in the previous lecture. Other activities included the derivations of some of the fundamental equations, and yet others were used to acquaint the students with new material (e.g. adiabatic demagnetization). Students were encouraged (but not forced) to discuss the material with their neighbour in the lecture theatre, thus promoting co-operative learning. After a suitable period of time (depending on the complexity of the activity) we discussed the solution to the problem. The class were told that they were responsible for providing a clear solution to all lecture activities, and that these activities would form the basis for test and exam questions.

2.2. Minute paper and responses

Minute papers (see Felder *et al* [2]) were administered at least once a week (usually after the end of the only double lecture). The students were asked to anonymously write down the most important points made in the double lecture, and the least clear point(s) on a pro forma handout. Before the next class I prepared a table showing the frequency of reponses to both questions,

and commented on the students' identification of the most important points. It was gratifying to note that the majority of students were proficient at recognizing the most important points. I responded to each of the 'unclear' points at the start of the next lecture, whether it was a verbal comment or part of a prepared set of display slides. Often these responses gave me the opportunity to present an alternative representation of the material associated with the 'unclear' points. As Table 1 shows, the students appreciated that I took these minute papers seriously. The minute paper responses were posted on the internet following the lecture in which they were presented, and they proved to be a valuable revision resource, according to the students.

2.3. Conceptual Questions

Both Thermal Physics and Statistical Physics offer the opportunity to probe the students' understanding of difficult concepts (e.g. entropy, adiabatic processes). During lectures I would invite students to discuss conceptual questions in pairs. This often elicited some lively debate.

2.4. Active tutorials

The students were encouraged to prepare for tutorial sessions by attempting all of the set problems on the tutorial sheet. Volunteers were called for to present solutions on the board, and if no-one volunteered then I selected one of the students. This allowed the students the opportunity to see how their colleagues solved a particular problem, and there was generally more discussion compared to when I presented the solution and asked for comments and questions, as I had in previous years.

3. Effect on Active Learning Strategies on the students

Assessment of the success of the measures described in the previous section were two-fold. Firstly, I employed the service offered by the Centre for Teaching and Learning Development (CLTD) at WITS and administered two formal Student Surveys of Lecturer Performance (SELP). The first of these was, at the time, intended purely for promotional and teaching development purposes. The subject of this survey was my teaching of the second year Thermal Physics module in the second semester of 2008. Following the introduction of active learning to the third year Statistical Physics module in 2009 I ran another survey with exactly the same format, and I also asked the students to comment explicity on the introduction of the new teaching style.

3.1. Comparison of Student Evaluations

SELP at the CLTD follow a format that will be familiar to many involved in education research. A positive assertion of lecturer performance is made, and the students indicate whether they (A) strongly agree, (B) agree, (C) are neutral, (D) disagree or (E) strongly disagree. Each individual response is scored as follows: (A) 10, (B) 7.5, (C) 5, (D) 2.5 and (E) 0, and an average is then obtained for each item from all responses. A SELP is voluntary, but these must be presented by those seeking confirmation or promotion at WITS. The efficacy of SELP surveys is often challenged by university teachers, as described by Felder [6]. He points out that this is often done without providing any evidence that they are not effective. The results of SELP surveys can be used to improve and refine teaching practice, and the results of such surveys are likely to be significant. Complementary initiatives, such as peer reviews, will strengthen their effectiveness.

The results of pre- and post-intervention SELP on essentially the same cohort of students are tabulated in Table 1. The overall results show a positive response to active learning. Several items have been highlighted in *italics* (average score decreases from 2008 to 2009) and in **bold**. The bold items indicate results that I regard as particularly significant. Perhaps the most significant is the large relative increase in the number of students who felt more comfortable

in participating in the class (although there was a small decrease in the score for the item 'Encourages audience participation'). Active learning appears to have had the effect of improving the students' attitude to several items considered to be positive for effective learning. The results also show an appreciation of the extra work involved in treating the feedback from the students with the respect it deserves ('Gets feedback on understanding') and in establishing the relevance of the course material ('Digressions made which add interest' and 'Links lecture to other parts of the course').

Table 1. The results of two SELP administered by the WITS CLTD during the modules presented in 2008 and 2009. The student samples are for second and third year Physics modules on related material, taught by the same lecturer, before and after the introduction of active learning techniques. The highlighted items in the table are discussed in the text.

Abbreviated Assertion	2008	2009	% Change
Makes clear the purpose of the lecture	8.96	9.29	+ 3.7
Stimulates interest in the subject	**8.08**	**9.11**	**+ 12.7**
Always well prepared for class	**9.24**	**10.00**	**+ 8.2**
Available for consultation outside lectures	8.66	9.29	+ 7.3
Encourages audience participation	*9.43*	*9.29*	*- 1.5*
Communicates effectively	8.47	8.93	+ 0.46
Chooses and organizes material well	8.08	8.93	+ 10.5
Pitches lectures at the appropriate level	8.66	9.11	+ 5.2
I gained understanding of concepts	7.70	8.58	+ 11.4
Motivated to read/do extra work	7.70	8.04	+ 4.4
Shows thorough subject knowledge	9.24	9.47	+ 2.5
Clear, understandable explanations	**8.08**	**9.11**	**+ 12.7**
Grasp of my level of knowledge is realistic	8.08	8.40	+ 4.0
Gets feedback on understanding	**7.70**	**9.65**	**+ 25.3**
Summarizes main points effectively	8.08	8.93	+ 10.5
Writes legibly on the board	*9.24*	*8.93*	*- 3.7*
Links lecture to other parts of the course	**7.89**	**9.47**	**+ 20.0**
Digressions made which add interest	**7.89**	**9.47**	**+ 20.0**
Directives for written work clear	8.47	9.29	+ 9.7
Lecturer has enthusiasm for subject	8.85	9.83	+ 11.1
Approachable for questions	**8.96**	**9.29**	**+ 13.9**
Welcomes different viewpoints	8.08	8.93	+ 10.5
Comfortable about participating	**6.54**	**8.58**	**+ 31.2**
Average	**8.33**	**9.44**	**+ 9.8**

The comments made by the students in the open-ended questions in the second survey, in which they were particularly asked to give their opinion of active learning, were overwhelmingly positive. All students who responded expressed an appreciation for the fact that they were engaged, rather than passive, during lectures and tutorials.

3.2. Comparison of Module Results

There were a total of fourteen students who took both the Thermal Physics module at the end of 2008 and the Statistical Physics module at the beginning of 2009. The final and examination

results for these 'common' students are tabulated in Table 2. With two notable exceptions, the performance of the students has either improved or remained roughly the same following the introduction of active learning. More than 60 % of students improved both their final and examination mark. A comparison of the examination results for 2008 and 2009 also show an improvement (see Table 3) and this improvement is mirrored in the final marks, which show a similar relative improvement. One might expect third year students to perform better than second year students overall, but it should be noted that the results in the other third year year modules were not uniformly excellent during 2009. It is true that the student numbers are small, and so the improvements in performance noted may be statistical fluctuations, it can nevertheless be confidently concluded that active learning techniques do not have a negative effect on the results.

Table 2. Comparison of student performance in 2008 and 2009 for students who took both modules in successive years.

Change in final result	Total mark	Examination mark
+ 20 % to + 30 %	2 (14 %)	2 (14 %)
+ 10 % to + 20 %	3 (21 %)	3 (21 %)
0 % to + 10 %	4 (29 %)	5 (36 %)
- 10 % to 0 %	3 (21 %)	2 (14 %)
- 10 % to - 20 %	1 (7 %)	1 (14 %)
- 20 % to - 30 %	1 (7 %)	1 (14 %)
Improved mark	9 (64 %)	10 (71 %)

Table 3. Comparison of final examination results in 2008 and 2009 for the full complement of students in each year.

	2008	2009
Class average	70.8 %	76.0 %
Pass Rate	89 %	100 %

4. Conclusions

Active learning techniques were introduced in the third year Statistical Physics module offered in the School of Physics at WITS during the first semester of 2009. A comparison of SELP and final module results for largely the same group of students support the conclusion that the intervention has been met with success. The students who attended the third year module presented using active learning initiatives described in this paper not only showed a general improvement in their final results, they also appear to be happier and more confident than they were when more traditional methods were employed.

References

[1] Prince M 2004 *Journal of Engineering Education* **1** 223–231
[2] Felder R M, Woods D R, Stice J E and Rugarcia A 2000 *Chem. Engr. Education* **34** 26–39
[3] Thacker B A 2003 *Rep. Prog. Phys.* **66** 1833–1864
[4] Sorenson C M, McBride D L and Rebello N S 2011 *Am. J. Phys.* **79** 320–325
[5] Felder R M and Silverman L K 1988 *Engr. Education* **78** 674–681
[6] Felder R M 1992 *Chem. Engr. Education* **26** 134–135

è

Simple Pendulum: A first year student's dilemma

Frank Komati[1] and Itumeleng Phage[1,2]

Email: iphage@cut.ac.za and fkomati@cut.ac.za

[1] Natural Science Department, Central University of Technology, Free State, Bloemfontein, South Africa

[2] Natural Science Department, North West University (Potchefstroom), South Africa

1. Abstract

Many first year university students in the Science related fields have a problem in identifying and applying the fundamental mathematical concepts that they have learnt from as far back as Grade 10 in solving some of the physics related problems. The majority of students, mostly from public schools, cannot seem to relate the two fields of study. Previous studies have shown that majority of first year students lack the integrated approach to different scientific disciplines in solving some specific problems and in experimental analysis.

This paper presents the findings of an investigation on the ability of the first year students to use the mathematical concept of the straight line equation and graph, in analysing physics properties, with particular reference to a Simple Pendulum motion experiment. The method employed was experimental. Students were given a simple pendulum experiment to determine the gravitational acceleration of the pendulum by graphical analysis. Two graphical approaches were employed, with both expected to yield the same results. Participants were required to analyse the motion of the pendulum, specifically the variation of the length of the pendulum with the period of oscillation.

The feedback showed a far reaching implications relating to first years' abilities to mathematically analyse a physics concept. These findings serve as a basis for a need to improve the teaching of science and mathematics in the schools, especially the practical approach to teaching and learning science.

2. Introduction

Most first year university students are faced with a dilemma of adjusting from the traditional separation of teaching and learning of mathematics and physical science in schools to the integrated approach to teaching and learning in both disciplines. Graph-related activities take an important place in recent reform efforts in mathematics and science education (Wolff-Michael and Bowen, 2001). This is particularly prevalent in students who come from mostly public and disadvantaged schools, with no practical laboratory experience whatsoever. According to Hung and Jonassen, 2006, students are not able to connect corresponding qualitative and quantitative knowledge of physics to each other as experts do.

This paper seeks to investigate the extent to which first year students are able to manipulate graphical mathematical principles and approach in the analysis and solving physics problems. This is done with particular reference to a Simple Pendulum motion experiment. Unfortunately even students who are good in physics have a difficult time obtaining useful information from graphs that consist of more than a single straight line (McDermott, Rosenquist and Van Zee, 1987). This survey aims to identify

whether the students can relate and use the straight line equation and properties to find the values of gravitational acceleration from the Simple Pendulum experiment.

The research has been conducted over a two year period, 2010 and 2011. The target group being mainly first years who have enrolled for B.Ed (Natural Science) programme at Central University of Technology. The results of this research survey are particularly important in helping to structure the instructional methods of teaching practical physics experiments to first year university students.

3. Method

3.1 Participants

3.1.1 First Sample (2010)

A total of 45 first year students in the Department of Teacher Education, Central University of Technology, Bloemfontein Campus participated.

3.1.2 Second Sample (2011)

A sample of 61 first year physics student in the Department of Teacher Education, Central University of Technology, Bloemfontein Campus participated. Of the 61, 47 (20%) were from private schools and 3 (80%) from mainly rural public schools.

3.2 Instruments

A simple pendulum set with varying length.

3.3 Procedure

The participants were required to set the pendulum swinging and measure the total time (t) that it takes to make 50 complete oscillations. They recorded the results in a table form. The procedure was repeated 10 times while uniformly increasing the length of the pendulum.

3.3.1 First approach:

The students were required to plot two sets of graphs using the data collected, i.e. Period (T) vs length (l) and period squared (T^2) vs length(l). From each graph they must identify the linear relationship between the period and the length of the pendulum given the relation $T^2=4\pi^2\, l./g$, and hence find the value of g.

From the graphs and relation $T^2=4\pi^2\, l./g$, the students were required to find and identify the slope of the graph as well as the y-intercept and hence find the value of the gravitational acceleration g.

3.3.2 Second approach

Using the same set of data, students were asked to plot the graph of $logT$ against $logl$. Given the relation $logT = nlogl + logk$, where $k = 2\pi/\sqrt{g}$, students were asked to find the slope n, $logk$ and k from the graph. From the information obtained above, the value of g had to be calculated.

4. Results

4.1 First approach:

In figure 1, of the 61 students, 8 (13%) managed to give a linear relation between period squared (T^2) and length(l) and hence the straight line graph (bar 1). 53 (87%) said there exist a linear relation between period (T) and length (l) (bar 2). 54 (89%) could not find and relate the slope to $4\pi^2./g$ (bar 3). 54 (89%) could not relate that y-intercept = 0, and hence the graph must start at the origin (bar 4). 56 (92%) could not find the value of gravitational acceleration as required (bar 5).

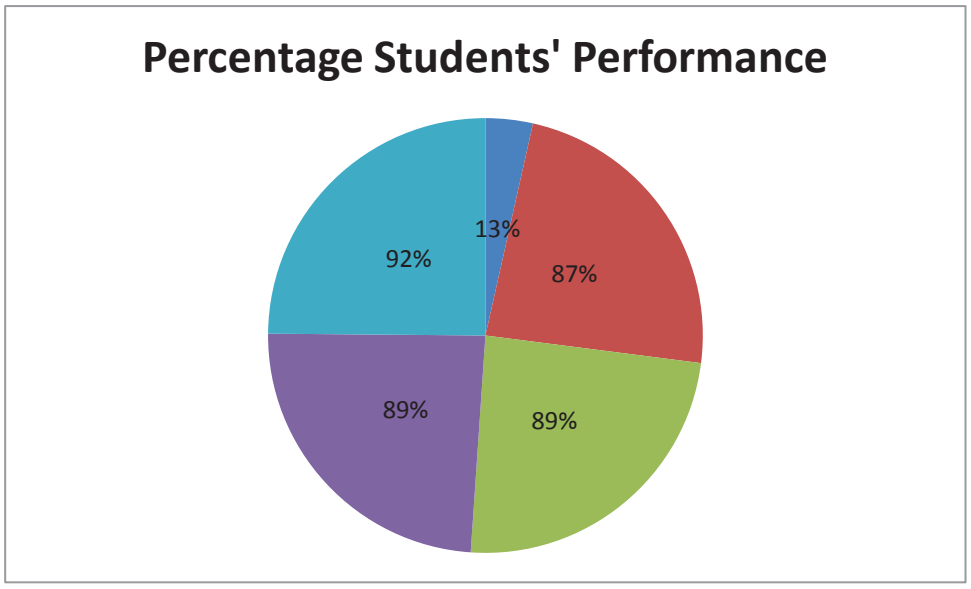


Figure 1: Percentages of students' performance.

4.2 Second approach

15% (bar 1) of the students managed to find and calculate the value of *n*, the slope. 8% (bar 2) were able to extrapolate the value of *logk* directly from the graph and hence calculate *k* and *g*. Because of the concepts of logs, students could not relate the logs to the straight line. On the other hand, only 12% (bar 3)of the participants managed to obtain a straight line graph.

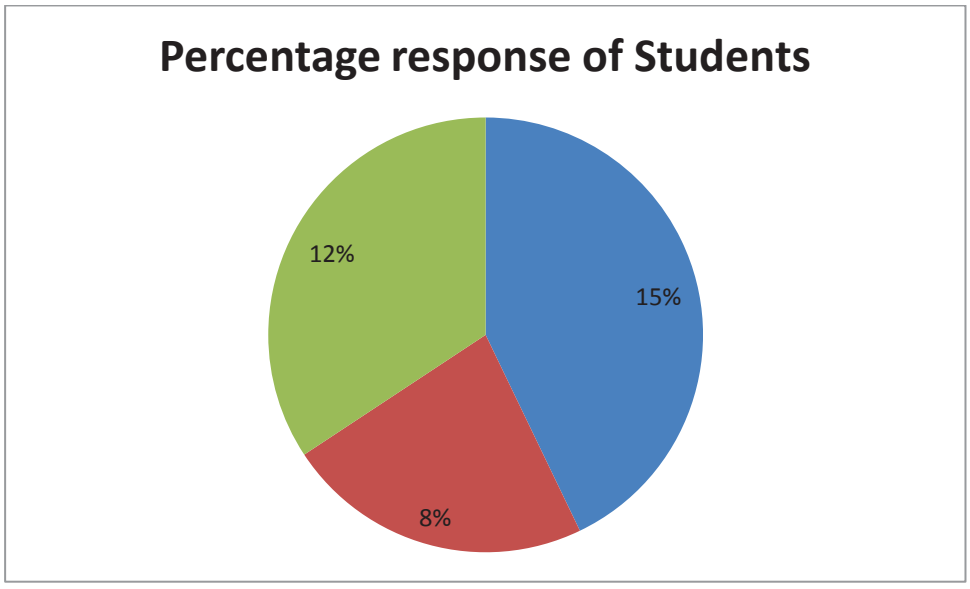


Figure 2: Sample percentage versus students' performance

5. Conclusion

Of 61 participants, 90% could not graphically analyse the pendulum motion. There is no cohesion between the mathematical concept of a straight line and data analysis of the motion of the pendulum. The study shows that there is a need to redefine the curriculum to accommodate first year students as far as graphical analysis and mathematical manipulation of data is concerned in the first part of the term. Emphasis should be placed on the need for competency in graphical analysis.

Practical work was generally effective in getting students to do what is intended with physical objects, but much less effective in getting them to use the intended scientific ideas to guide their actions and reflect upon the data they collect. (Abrahams and Millar, 2008) and make links between Science/Physics and Mathematics. Therefore for as long as laboratory experiments are not conducted in schools, the problem will always persists.

6. References

6.1. Palmquist, B.C. (2001). Assessing motion graphs. *The Science Teacher;* Nov 2001; 68, 8; Research Library pp. 75

6.2. Abrahams, I. and Millar, R. (2008). Does Practical Work Really Work? A study of the effectiveness of practical work as a teaching and learning method in school science. *International Journal of Science Education,Vol. 30, No. 14, 17 November 2008, pp. 1945–1969*

6.3. Hung, W. and Jonassen, D.H. (2006). Conceptual Understanding of Causal Reasoning in Physics. *International Journal of Science Education Vol. 28, No. 13, 27 October 2006, pp. 1601–1621*

6.4. Wolff-Michael, R. and Bowen, G.M. (2001). Professionals Read Graphs: A Semiotic Analysis. Journal for Research in Mathematics Education 2001, Vol. 32, No. 2, pp 159-194

6.5. Gaigher, E., Rogan, J. M., and Braun, M. W. H. (2007). Exploring the Development of Conceptual Understanding through Structured Problem-solving in Physics. *International Journal of Science Education, Vol. 29, No. 9, pp. 1089–1110*.

6.6. McDermott, L. C., Rosenquist, M. L. and Van Zee, E. H. (1987). Student difficulties and connecting graphs and physics: Examples from kinematics. *American Journal of Physics, Vol. 55, No. 6, pp. 503-513.*

"Seeing is believing": Visual perceptions and the learning of kinematics

Miriam Lemmer

School of Physical and Chemical Sciences, North-West University, Potchefstroom, South Africa

miriam.lemmer@nwu.ac.za

Abstract. The proverb "seeing is believing" is generally applied in science education today. Many of the contemporary teaching strategies engage the learners in activities that guide them in making their own observations. Still, problems regarding the learning of physics, especially the prevalence of alternative conceptions, remain a headache in physics education research. The study reported here investigated learners' visual perceptions as possible cause for the existence, persistence and transfer of the intuitive conception in kinematics that DiSessa (1993) called changes-take-time. An example of this conception is the assertion that a cannon ball keeps on accelerating after the shell has been launched. A focus group discussion and questionnaire were used in the investigation. Grade 9 and 10 physical science learners participated in the study. The results showed that limitations in visual perceptions (such as differences in real and perceived velocities and changes in velocities) seem to contribute to their intuitive concepts.

1. Motivation and problem statement

Our senses are the only means to obtain information about the world around us. Every second a multitude of stimuli reaches the sensory register of the brain. According to the information processing model [1], selective perception takes place in the sensory register and only the selected stimuli are transformed to neural information. In the working memory the new information is combined with existing knowledge that has been extracted from the long-term memory. If the information is found to be useful and also consolidated through repeated use, stable knowledge structures are formed in the long-term memory.

Learning of concepts consequently depends on the new information that is attended to, as well as the existing knowledge that is cued. While physics education research paid much attention to existing knowledge (especially misconceptions), perceptions formed from new information have not often been studied. In chemistry education visual perception is of utmost importance due to the problem that macroscopic observations are explained in terms of microscopic entities [2]. For example, a table does not look or feel as consisting out of very small vibrating particles with spaces in between. Visual perceptions can also account for the finding that learners often attribute macroscopic properties to microscopic particles, e.g. they think that the particles of solids are hard while liquid particles are soft and fluid.

This study focused on the role of visual perceptions in the learning of kinematics. In the teaching of kinematics at school and first year university level, the motion of well-known everyday life objects such as balls and cars is often used in examples and problems. As long as the mass of an object is perceived to be concentrated in the centre of mass, the basic concepts and laws can be applied. Still, misconceptions regarding force and motion are amongst the most reported ones [3]. These misconceptions seem to be independent of culture [4] and are very resistant to change [5], even with

contemporary teaching strategies. One of the misconceptions found in kinematics is the so-called changes-take-time perception [6]. Accordingly, learners assume that an object maintains its acceleration or velocity after it has been launched. This perception seems to be related to the well-known misconception force-as-mover according to which learners always attribute a force in the direction of motion. Another well-documented problem in the learning of kinematics is that learners tend to confuse related concepts such as position, velocity and acceleration [7]. Since learners daily observe the motion of rolling balls, projectiles, etc. the question arises as to whether their visual perceptions can cause the formation of these misconceptions and oppose their understanding of the scientific interpretation of motion. The following research questions were investigated:

(1) What reasons do learners give for having the changes-take-time perception?
(2) What role does visual perception play in the formation of this perception?
(3) What are consequences for the teaching and learning of physics?

2. Theoretical framework

A variety of causes may be ascribed to the occurrence of alternative conceptions, e.g. that textbooks contain them [8], teachers themselves hold them [9] or do not attend to learners' intuitive ideas. Learners may also focus on contextual features of an event and do not observe what educators intend them to [3]. For example, when a force is applied to an object, the learners may attend to features of the person performing the action (e.g. a strong male exerts a large force) while, for the scientist, the actor is of no importance since the force is related to the acceleration of the object's motion. Furthermore, in kinematics scientists differentiate between constant velocity and accelerated motion, while everyday observations is usually just concerned with whether the object moves or stands still.

Apart from the effect of selective perception on the formation of conceptions, limitations of our sensual perceptions can also influence what new information reaches the working memory. For example, we cannot see very small particles such as atoms or molecules with an unaided eye. There are also limitations to the sounds that the human ear can detect (namely between 20 Hz and 20 kHz). Furthermore, two sounds in the hearing range cannot be distinguished when their frequencies are too near [10]. In his standard work on visual perception, Brown [11] reported factors that cause differences in perceived (phenomenal) and real velocities of objects, which include:

1. Perceived velocity decreases with increase in distance from the observer.
2. Increase in the size of the moving object decreases the perceived velocity.
3. The direction of the movement relative to the observer affects the perceived velocity, e.g. vertical movements are phenomenally faster than horizontal movements, while movements diagonally between horizontal and vertical fall phenomenally between these two.
4. Movements observed while fixating a stationary point are phenomenally faster then when the eye follows the moving object.

3. Research design

A qualitative and a quantitative research strategy were combined in the investigation. In a focus group discussion with five Grade 9 learners of a local school, it was qualitatively determined whether they hold the changes-take-time perception or not, and what reasons they give for or against it. This was followed by a questionnaire that quantitatively determined the occurrence of and reasons for the changes-take-time perception amongst 208 Grade 10 learners from 14 local schools.

4. Results and discussion of results

4.1. Focus group discussion

In the focus group discussion the learners' perceptions of the change in speed in different contexts were investigated. Three out of the five learners (A, B and C) showed the changes-take-time perception, while the other two learners (D and E) held more scientific ideas.

The interviewer (I) started by pushing and releasing a small ball to roll on a long table.

I: (demonstrates while talking): When I am pushing the ball, how does its speed change?
A & B (together): It accelerates
I: And after I've removed my hand?
E: It slows down
D: And then stops

B: No, it will keep on accelerating
D: Ma'am … (laughs) … You pushed it. Then it's on its own and moving and then slows down because you're not pushing it anymore.

The learners were then asked to roll balls on the floor and observe the changes in speed.

I to learners A, B & C: Do you still have the idea that it goes faster after release?
A: Yes
B: No, it's constant
A: Yes, it stays constant … until it stops.
I: Why do you think will it stay constant?
A: I think that the surface that we use right now is flat, so there's nothing interrupting it. It just goes.
I: Why does it stop later?
B: (Thinking) … Ma'am …
(The three learners discuss the problem with each other but cannot provide any answer.)

The flat surface of the floor was used as explanation for the constant velocity motion of the ball, but it could not account for the observation that the ball eventually stops. Still, learners A, B and C did not alter their intuitive perception. They were also not sure whether the ball accelerates or moves with constant velocity. The changes-take-time perception as well as confusion of the concepts acceleration and velocity may thus be due to limitations in visual perceptions. From own experiences one cannot clearly see if and how the velocity of a moving object changes after it has been accelerated by some force. For example, a deceleration of -2 $m.s^{-2}$ cannot be observed if the velocity decreases in 1 s from 24 to 22 $m.s^{-1}$, but it can be clearly detected when the velocity decreases from 4 $m.s^{-1}$ to 2 $m.s^{-1}$ in 1 s, i.e. with the same deceleration at a lower velocity.

The learners treated vertical motion differently from horizontal motion. This is in agreement with the finding of Brown [11] that the direction of movement relative to the observer affects the perceived velocity. Learner A differentiated between the up- and downwards motions of a ball that was thrown vertically upwards:

A: In think that when you throw the ball up, say it goes with 50 km/h, then when it comes down, it may be 70 km/h, because gravity is pulling it down and gravity is going to be a larger force than for throwing it up.

For the first time in the discussion, a learner referred to a "force for throwing it up" which may indicate the force-as-mover perception. Later on learners B and C also reasoned than something inside the ball is pushing it:

C: There cannot be only gravity. There must be something inside the ball.
I: What is inside the ball?
C: Potential energy
B: Air – there is air inside this ball. It is pushing the ball. A force is something that pushes – It's a push or a pull.

The learners with the changes-take-time perception easily adapted the force-as-mover conception when they needed it. Learner B cued existing physics knowledge to validate this conception. Later in the discussion learners D and E combined the idea of an inside force with the retarding force to explain why an object slows down after it has been released. This result is in accord with the finding of Hammer [12] that learners initially reasoned phenomenologically, e.g. "moving objects slow down and stop". Only when asked to explain the motion, the learners begin to look for causes and activate a force-as-mover conception.

After discussions regarding a ball rolling along the table and in free-fall, the motion of a ball on a ramp that was inclined upwards, downwards and horizontal was used to further probe the learners' reasoning. Upon the question of the interviewer of why the ball rolls down the ramp, learner E immediate said "the slope" and learner D added "the steepness of the slope." The other learners agreed and no one referred to the force of gravity. For the ball rolling on a horizontal ramp the learners maintained their ideas of the previous discussions. The changes-take-time learners added a new explanation for their assertion that the ball will move with a constant speed on the ramp, which was expressed by learner A as follows:

A: The ball doesn't stop on the ramp; therefore it will be constant all the way. If it rolls from here to there (shows to the other side of the classroom), it is going to stop. But if you give us

10 cm (shows a distance of about 30 cm between her hands), it will move with the same speed.

Learners A and B agreed, but D and E reacted as follows:

D: But the speed of the ball in the beginning and end of the 10 cm won't be the same. At the end of the ramp it will be slower than at the beginning

I: Why do you say so?

D: Because it can't move on its own.

E: Something that goes uninterrupted until it stops shows that it was decreasing. It doesn't go constant and ...gup… it stops (moves her hand horizontally in the air and suddenly drops it). It goe-e-es (hand gradually slows down) and then it stops.

In this discussion both learners A and E referred to observations of the ball moving on the horizontal ramp. For learner A the speed of the ball is determined by whether you observe that it stops or not, while learner E referred to her observation that the ball's speed does not change suddenly. The different observations that they focused on therefore supported their perceptions. Learners with intuitive conceptions seem to use contextual features as explanations, while the more science-orientated learners tend to use logical reasoning.

Second questionnaire:

From the focus group discussion different reasons for the perception changes-take-time emerged, namely some "inside" force and situational features such as the "flatness" of the floor or the short length of the ramp. The learners who did not show the changes-take-time perception, argued against the idea that the external force causing the motion maintains it, but later in the discussion accepted an "inside" force. A short questionnaire was used to investigate which reasons occurred more frequently among a larger number of learners.

The questionnaire comprised of three questions and was answered by 208 Grade 10 learners from 14 schools in the district of the researcher. The first question was a multiple choice item on the change in speed of a ball after a hockey player hit it from rest. Then the learners had to choose reasons for their responses from a given list to which they were urged to add more reasons. They also had to indicate which of the chosen reasons they felt most sure about. In the third question they were asked to draw force vectors at four positions of a ball that rolls over a hockey field after it has been hit.

The result of item 1 of the questionnaire showed that the largest percentage of learners believed that the hockey ball either accelerates or moves at a constant velocity after it has been hit, i.e. they displayed the changes-take-time perception. The reasons for their responses (item 2) were categorized as in table 1 with examples of each category given in table 2.

Table 1. Results of item 2: Number and percentages of learners who gave certain reasons for their perceptions regarding the speed of a hockey ball after it has been hit.

Reasons for answers to item 1	Number	Percentage
Visual perception	132	25.3
Force-as-mover	135	25.9
Energy	188	36.1
Situational features	50	9.6
Opposing forces	16	3.1

Approximately 25 % of the reasons given by the learners were based on their visual perceptions and a similar percentage on the force-as-mover misconception. The largest percentage (36 %) was obtained for the idea of energy transfer. This can be understood in terms of the emphasis that is laid on the concepts of energy and change in the preceding years of study in the South African schools. About 10 % of answers referred to situational features such as the smoothness (or grassiness) of the field, while only 3 % of learners referred to forces that oppose the forward motion of the ball. Since the learners could chose as many of the reasons as they wanted, they were also asked which one they were most sure of. Their responses yielded very much the same percentages as the percentages for the different options in table 1. They were thus most sure about the transfer of energy, followed by the changes-take-time and force-as-mover intuitive conceptions.

Table 2. Examples of reasons given by learners.

Reason	Examples
Visual perception	I can see the different velocities of the ball. I play hockey and knows how the ball moves after being hit. You can see how hard the ball is hit ball is hit.
Force-as-mover	The force of the hockey stick helps the ball's velocity to be constant. The ball has been hit with a large force. The force of the stick still acts on the ball after it has been hit.
Energy	The hockey stick transferred energy to the ball. The ball uses the energy transferred to it which is eventually used up and the ball stops. Energy gets lost after a while and you cannot transfer any more energy.
Situational features	The ball moves away fast / rolls far. The ball is in motion. The surface is smooth. / There is grass on the hockey field.
Opposing force	It stops because of the wind. Friction and/or gravity oppose the ball's force. When the ball touches the ground the speed decreases.

In the third question of the questionnaire, the vast majority of learners (88%) included a force vector in the direction of motion. Only about 20 % of learners showed an opposing force which was in many cases only used towards the end of the motion. It is interesting to note that less than a quarter of the 88% of learners that used a force vector in the direction of motion indicated the force-as-mover misconception in their reasons in item 2. It seems that they did not initially connect the force-as-mover misconception to their perceptions of the change in speed of the ball, but that it was brought forward when they were prompted to draw the force vectors.

5. Conclusions

The results indicate that the proverb "seeing is believing" is relevant in the learning and teaching of kinematics. Visual perceptions seem to play a significant role in the formation of the changes-take-time intuitive perception. This perception is probably established by repeated observations of moving objects, attention to situational features and limitations in seeing changes in the speed of fast-moving objects. When unattended, syncretism with science concepts may occur and robust misconceptions such as the force-as-mover conception may be formed.

In the science classroom limitations in visual perceptions regarding the speed of moving objects can be compensated for by technology such as stroboscopic photos, ticker-timer experiments and computer simulations. By focusing learners' attention and providing repeated experiences in a variety of contexts the kinematics concepts can be formed and established.

References

[1] Redish E F 2004 A theoretical framework for physics education research: Modelling student thinking Eds. E F Redish and M Vicentini Research on physics education Proc. of the Int. School Phys "Enrico Fermi" Course CLVI Villa Monastero 15–25 July 2003 (Amsterdam: IOS press) pp 1-63

Woolfolk A 2010 *Educational Psychology* 11th edition (New Jersey: Pearson Education International)

[2] Crespo M A G and Pozo J I 2004 Relationships between everyday knowledge and scientific knowledge: Understanding how matter changes *Int. J. of Sc. Educ.* **26** 1325-1343

[3] Palmer D 1997 The effect of context on students' reasoning about forces *Int. J. of Sc. Educ* **19** 681-696
[4] Thijs G D and Van den Berg E 1995 Cultural factors in the origin and remediation of alternative conceptions in physics *Sc. & Educ.* **4** 317-347
[5] Kim E and Pak S J 2002 Learners do not overcome conceptual difficulties after solving 1000 traditional problems *Am. J. Phys.* **70** 759–65
[6] DiSessa A A 1993 Toward an epistemology of physics *Cogn. & Instr* **10** 105-225
[7] Shaffer P S and McDermott L C 2005 A research-based approach to improving student understanding of the vector nature of kinematical concepts *Am. J. Phys.* **73** 921-931
[8] Lemmer M, Edwards J M and Rapule S 2008 Educators' selection and evaluation of natural sciences textbooks *S.A. J. Educ.* **28** 175-187
[9] Kriek J and Basson I 2009 The effect of physics computer simulations on teachers' misconceptions of DC circuits Proc. SAARMSTE Conf. January 2009 Rhodes University Grahamstown
[10] White R T 1989 *Learning Science* (Oxford: Basil Blackwell)
[11] Brown J F 1931 The visual perception of velocity *Psychologishe Forschung* **14** 199-232 Reprinted in: Ed I M Spigel 1965. Readings in the study of visually perceived movement, (New York: Harper & Row) pp 64 – 107
[12] Hammer D 1996 Misconceptions or p-prims: How may alternative perspectives of cognitive structure influence instructional perceptions and intentions? *J. Learn. Sc* **5** 97-127

Learners' understanding of ammeter and voltmeter in direct current schematic circuits

R T Mautjana [1, 2] and T T Netshisaulu[2]

[1]NMISA, Private Bag X 34, Lynnwood Ridge, 0040

[2] Physics and Geology Department, University of Limpopo, Private Bag X1106, Sovenga, 0727

tmautjana@nmisa.org

Abstract. This study explores learners' knowledge of measuring devices in a direct current resistive circuit. It also investigates learners' understanding of how ammeters and voltmeters are connected. Ammeters and voltmeters are used to determine the behaviour of a circuit by indicating the readings of current and voltage, respectively. To an educator, an ammeter does not change the characteristics of a circuit; it is therefore an invisible component. However, learners view the roles of ammeters differently. A questionnaire was administered to grade 12 learners and structured group interviews were used for additional evidence. The results indicated that learners lacked the basic understanding of the role played by measuring devices. It was also evident that because of a lack of practical experience with real circuits, some learners did not know how measuring devices should be connected. There were also findings that indicates learners have difficulties with concept differentiation.

1. Introduction

The present study investigates the understanding of measuring devices and the role they play in electrical circuits. An in-depth knowledge of measuring devices is essential, if learners were to develop conceptual understanding on how electrical circuits work. After secondary education, learners are expected to have developed conceptual models of key concepts in electrical circuits such as current and potential difference [5]. The measuring devices help to visualize abstract concepts that cannot be physically investigated. When connecting devices, such as ammeters and voltmeters to a circuit, their roles should be taken into consideration. Their presence does not adversely affect the circuit's normal operation [6]. However, learners are expected to appreciate the effect of changes made by adding measuring devices to a circuit either in parallel or in series. In addition, they should be able to connect measuring devices correctly and provide explanations in terms of scientific principles [9]. However, numerous studies on current, potential difference and brightness of bulbs have shown that learners still have difficulties and misunderstandings after systematic instruction [7]. This study focuse only on learners' knowledge of the use of ammeter and voltmeter in a direct current resistive circuit. The study investigates learners' understanding of key concepts in electrical circuits, specific roles of measuring devices and the learners' general understanding of how measuring devices operate within electrical circuits.

2. Research Methods

2.1 Instruments

A questionnaire with closed and open-ended questions was administered to grade 12 secondary school learners. Most of the items were drawn from diagnostic instrument on conceptual difficulties developed by Engelhardt and Beichner [4] and the researcher added a part where learners had to provide reasons for their options. In addition, rephrasing of some questions was done to clarify the questions. Structured group interviews were used to support the questionnaire data. Due to limited space in this paper, only quantitative data is presented. To achieve content validity, questionnaires were submitted to two Physics and one English (Extended Degree Programme) lecturers. Furthermore, questions were formulated in simple language for clarity and ease of understanding.

2.2 Sampling

A total of 137 grade 12 learners comprising of males and females from two public schools (township and rural in Mankweng circuit of the Limpopo participated in this study. Cluster sampling technique [2] was used where the researcher–selected two schools**.** The selected schools did not have laboratories or apparatus to perform experiments. A relatively larger percentage (56%) of the respondents came from the rural school whilst 44% were from the township school. The average age group for the sample ranged from 18-20 years old. Learners had completed their electricity topic for their school curriculum**.**

2.3 Data analysis

The collected data was analyzed by counting the frequencies of selected options from each multiple-choice question. Reasons provided were analyzed by extracting patterns and or themes emerging from the responses. Responses were then categorized and frequencies were counted for each category.

3. Results and Discussions

3.1 Role of ammeter

An ideal ammeter has zero internal resistance, for it to drop no voltage as possible as charged particles flow through it. However, real ammeters have as little resistance as practically possible [5]. To an educator an ammeter does not change the characteristics of the circuit; it is therefore an invisible component used to indicate current readings in a circuit [9]. However, learners view the roles of ammeters differently.

Learners were requested to indicate what they thought to be the role of an ammeter in a circuit and the results are summarised in figure 1 below.

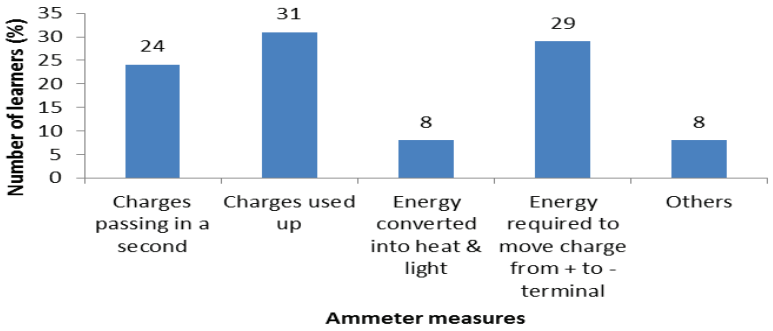


Figure 1. Learners' responses about the role of an ammeter in a circuit

Figure 1 shows that majority of learners (60%) held the well documented misconceptions about

electricity concepts [1, 9]. Of the given misconceptions, 31% of the learners indicated that the ammeter measured the charges which were used-up by circuit components (like bulbs, resistance, etc.). Learners thought that current decreases as charges move through circuit elements. In addition, 29% of the learners indicated that ammeter measured the energy required to move a charge from one point to the other. This view was also evident from the interviews with learners from both schools. This misconcept could be attributed to concept confusion, not relating concepts properly (namely, potential difference, current and resistance) [1, 9].Only 24% presented scientifically acceptable responses; therefore, the role of the device is not understood.

3.2 Role of voltmeter.

Voltmeter is connected between two points because a real voltmeter has the highest resistance possible to prevent draw of current from the circuit [6, 5]. Learners were requested to indicate their beliefs about the role of a voltmeter in a circuit and the results are summarised in figure 2 below.

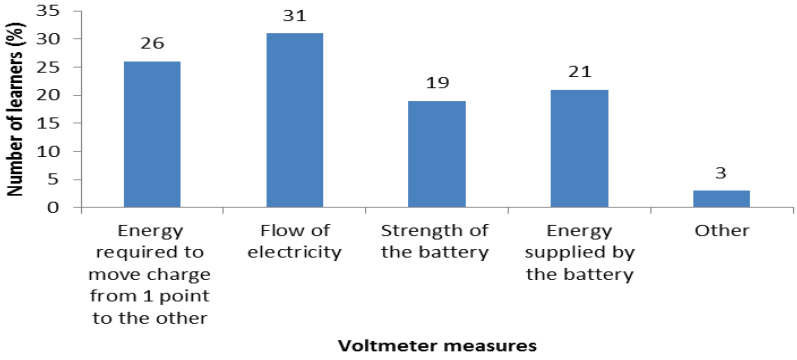


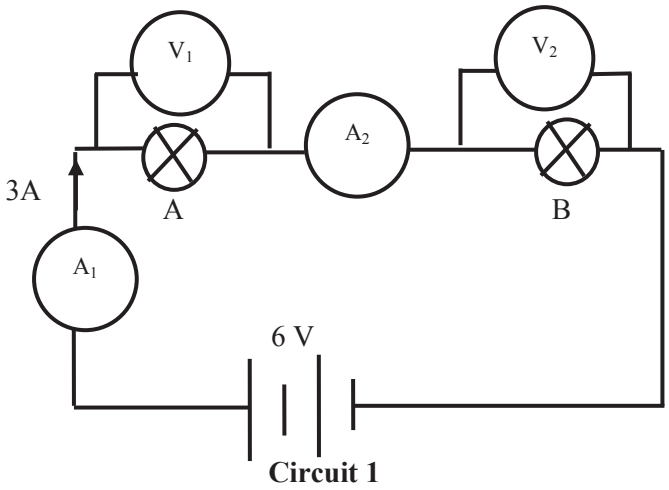
Figure 2. The role of a voltmeter in a circuit

Figure 2 depicts that majority of learners (71%) held misconceptions about electricity. Of these, 31% of them thought that the voltmeter measured the electricity flowing through the conductor. Learners refer to the substance flowing within the wires as electricity.Literature indicate that current is a flow of charge. This indicates learners were taking electricity to be the same as current. These learners thought voltmeter which is connected across components is used to measure flow of charge. These learners had difficulty with concepts differentiation [1, 3]. The use of the term electricity can be traced to Africans' everyday languages where electricity, power and current are used interchangeably [7]. This finding is consistent with literature..

About one in five of the learners thought that the voltmeter was measuring the energy supplied by the battery to the circuit and the interview with learners confirmed this view. Only 26% of learners indicated a correct option. It was apparent that learners did not understand the role played by voltmeter in a circuit. The latter could be attributed to learners' misunderstanding of the role of the voltage in a circuit, as alluded by Gilbert [5].

3.3 Voltmeter and ammeter in a series circuit

Learners were presented with a schematic circuit diagram (see circuit 1) below.



Circuit 1

The circuit consisted of two identical dry cells, two identical light bulbs and two ammeters all in series and two voltmeters across each bulb. Ammeters and voltmeters were labelled A_1, A_2, V_1 and V_2, respectively. The light bulbs were labelled A and B as shown in circuit 1. The total current and voltage of the circuit were presented as 3A and 6V respectively.

3.3.1 *Using voltmeters correctly in circuits*. Learners were asked to identify and to indicate the expected readings on the two voltmeters V_1 and V_2. The results are presented below in figure 3.

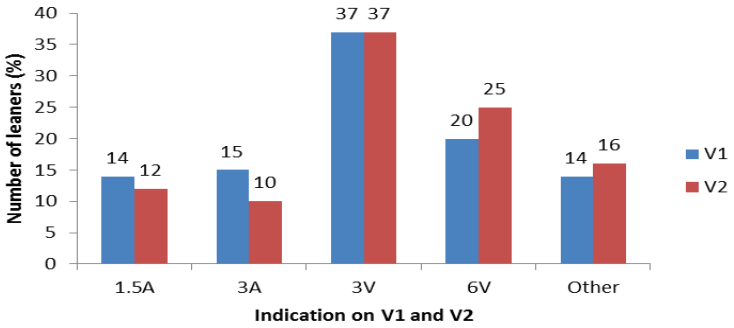


Figure 3. Learners' responses to readings on voltmeters V_1 and V_2

Figure 3 shows that 57% and 62% correctly identified V_1 and V_2 to be voltmers respectively (shown by 3V and 6V). This suggests that majority knew that voltmeter had to be connected in parallel in electric circuits. On the other hand, 30% and 22% of the learners thought V_1 and V_2 were ammeters (1.5A and 3A). These learners seemed to have difficulties with how measuring devices should be connected in a circuit. Only 37% presented correct indications for both V_1 and V_2. Theselearners correctly applied the rule that V_1 and V_2 add to equal voltage from the battery. This finding confirms literature that learners who fail to apply the basic rules had difficulty in understanding the concept [1, 5, 7].

3.3.2 *Using ammeter correctly in circuits*. Learners were also required to identify A_2 (ammeter) and to indicate the current through it. They were expected to apply the conservation of current rule in a series circuit. The results about A_2 are presented in figure 4 below.

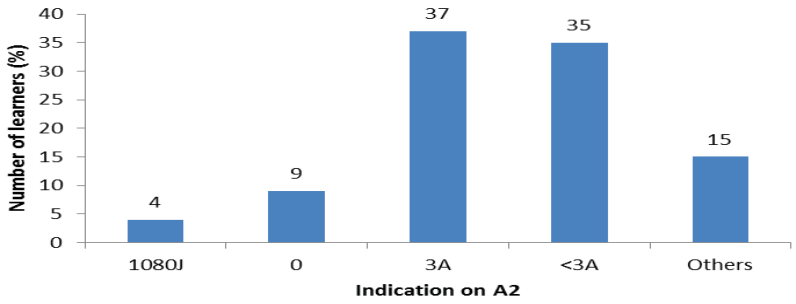
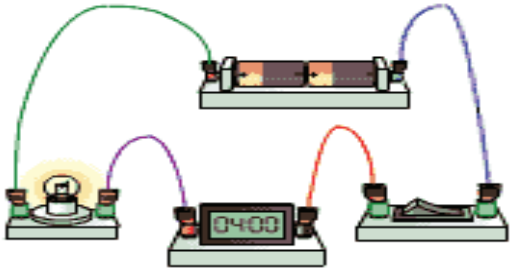


Figure 4. Learners' responses about the indication on A_2

Figure 4 indicates that majority of learners (72%) correctly identified A_2 to be an ammeter (indications 3A and <3A). This suggests majority of learners knew how ammeter should be connected in a circuit. However, only 37% indicated the correct reading meter A_2 would have indicated. The result imply these learners understood the conservation of current in a series circuit. However, 35% thought ammeter A_2 would indicate less current. This result confirms the observations made in figure 1 where 31% of learners preferred the current consumption model. Some studies indicated that learners believed that an ammeter consume current for it to function [8]. Although, majority of leaners seems to know that ammeter is connected in series, results suggest that most learners do not understand the current conservation.

The learners were also presented with a pictorial representation of a realistic circuit (as in circuit 2 below), which they had to evaluate to identify the corresponding schematic circuit diagrams.



Circuit 2. Representation of a realistic circuit.

Figure 5 shows learners' responses to matching realistic circuit to a schematic diagram.

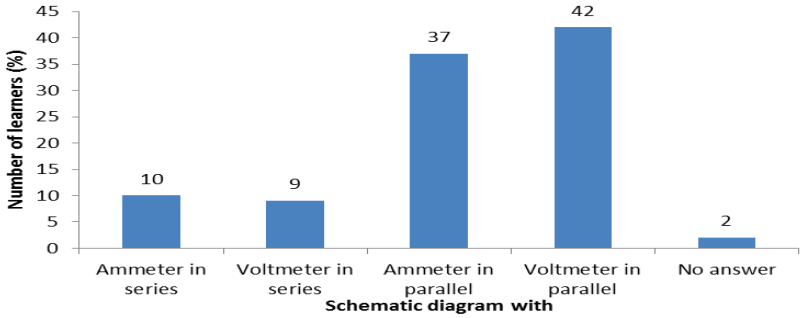


Figure 5. Learners responses to matching realistic circuit to a schematic diagram

Figure 5 depicts that 51% of the learners opted for circuit diagrams with voltmeter connected either in parellel or in series. During the interview, learners were asked to identify the components. Majority of learners positively identified battery, bulb and switch however, almost all failed to identify the

ammeter. The following is an example of a learners who could not identify ammeter, the learner said "*...not sure whether there is voltmeter or not*". Some learners (46%) chose options which had their measuring devices connected incorrectly, for example, ammeter in parallel and voltmeter in series. Only 10% provided scientifically accepteble responses. Carstensen et al. [1] indicated that learners who cannot differentiate between voltage and current tend to struggle in connecting their measuring instruments.

Dzama [3] found that learners were unlikely to make mistakes of how measuring devices are connected if they had practical experiences. Engelhardt et al [4] found that learners had difficulties in translating from schematic circuits to real circuits not vice versa. However, this study showed that learners had difficulty in translating from realistic circuit to a schematic circuit. The difficulty could be attributed to not understanding the concepts, lack of practical experience with electrical circuits and not understanding the role played by measuring devices .

4. Conclusions and Recommendations

This study investigated learners' knowledge about the role played by the ammeter and voltmeter in an electrical circuit. The results indicated that majority of learners knew how measuring devices should be connected in a circuit but lacked the basic understanding of the role played by these devices. There were also findings such as current consuption and difficulty with concept differentiation which are well documented in literature. It also emerged that learners had difficulty in translating from a realistic circuit to schematic circuit due to lack of practical experience with electrical circuits, concept confusion and not understanding the role played by the measuring devices. It is therefore necessary for educators to take into considerations the role played by measuring devices in planning their instructions.

References

[1] Carstensen A and Bernhard J 2007 Threshold concepts and keys to the portal of understanding: some examples from Electrical Engineering. *Threshold Concepts within the Disciplines*, (Sense Publishers) p 1-12

[2] Cohen L, Manion L and Morrison K 2007 Research Methods in Education. 6th Edition. (Routledge London), part 4 and 5.

[3] Dzama E 1997 Towards locating sources of alternative conceptions about Electricity among learners in developing countries. Focus on the learner. Proceedings of the 5th annual meeting of the Southern African Association for Research in Mathematics and Science Education. University of the Witwatersrand, Johannesburg. 22-26 January p 105-110

[4] Engelhardt P V and Beichner R 2004 Students understanding of direct current resistive electrical circuits *Am J. Phys*. **72** (1) p 98-114

[5] Gilbert W 2009 Experiment 7 DC Circuits and Instruments "Look for knowledge not in books but in things themselves" http://www.owlnet.rice.edu/~phys126/labs/expt07 p 37-41 [Accessed 21/09/2009]

[6] Habash R W Y 2002 ELG2331 Lab Manual - Experiment 1: DC Circuits http://www.site.uottawa.ca/~rhabash/ELG2331Exp1.pdf [Accessed 21/09/2009]

[7] KÜÇÜközer H and Kocakülah S 2007 Secondary school students' misconceptions about simple electrical circuits. *Journal of Turkish Science Education*, **4**(1) p 100-115

[8] Planinic M, Boone WJ, Krsnik R and Beilfusss M L 2006 Exploring Alternative Conceptions from Newtonian Dynamics and Simple DC Circuits: Links between Item Difficulty and Item Confidence. *Journal of research in science teaching*. **43**(2) p150-171.

[9] Taber K S, de Trafford T and Quail T 2006 Conceptual resources for constructing the concepts of Electricity: The role of models and analogies and imagination. *Physics Education.* **41**(2) 155-160

The pre-knowledge of the physics I students on vectors

M R Mhlongo, M E Sithole and A N Moloi

Physics Department, University of Limpopo, Medunsa Campus. RSA

mrebecca@ul.ac.za

Abstract. The purpose of this study was to investigate the pre-knowledge of the Physics I students on vectors. The study was conducted on 234 Physics I students from the University of Limpopo (Medunsa campus). The sample was divided into four groups, where the first group (n = 119) did their grade 12 in 2010, the second group (n = 46) did their grade 12 prior 2010, the third group (n = 42) did Foundation Physics in 2010 and the fourth group (n = 27) were those repeating Physics I. An ex post facto research design was chosen for the study whereby all the groups were given a vector test at the beginning. The test was divided into three questions, testing definitions, classifications, drawing and interpretations of graphs. Results showed a less difference between average percentage of the groups whereby the group before 2010 (47.4%), foundation group (47.3%), the repeaters (44.3%) and those who did grade 12 in 2010 (44.1%).

1. Introduction

A vector is a quantity with magnitude and direction that is represented geometrically by a directed line segment, i.e., an arrow [1]. Vector concepts and calculation methods are the core of the physics curriculum, underlying most topics covered in introductory university physics courses. [5] emphasized that the vector nature of forces, fields, and kinematical quantities requires that students have a good grasp of basic vector concepts if they are to be successful in mastering even introductory-level physics. The primary concept of Newtonian mechanics is force, and forces are vectors, they should be added using vector addition to determine the net force along the axis of motion. It appears that most of students' knowledge in the university level is brought from high school physics [6]. There are considerations that the idea of vector has different meanings in different contexts and therefore it is not easily transferable from one context to another.

Unlike scalar quantities such as temperature, mass and time, the mathematical manipulation of vectors is somewhat more complicated. For a typical introductory mechanics course the topics to be studied include areas such as kinematics, dynamics, Newton's Law of motion, work and energy, impulse and momentum, and rotational motion. In order to reach a sound understanding of the concepts presented in these topics a basic understanding of vector algebra is also needed [2].

There are suggestions in the literature that a qualitative approach to teaching would help students to learn. The Hestenes and Wells Mechanics Baseline Test look at the directional aspects of kinematic vectors and at the superposition of force vectors and the lowest scores were reported in the questions that required an understanding of vector properties [3]. [7] reported on students learning difficulties related to basic vector operations as employed in introductory physics courses.

608 SA Institute of Physics 2011 ISBN: 978-1-86888-688-3

British study [8] indicates that, "in spite of clear instruction in procedures for vector addition, many students 'forget' to draw the final side of the triangle (the resultant) when finding a vector sum, or have difficulties when vectors are in non-standard positions (i.e., crossing one another or pointing at the same point)". In the study [6] found that ¼ of students who had completed a calculus-based physics course and ½ of students who had completed an algebra-based physics course could not add vectors in two dimensions.

All these studies suggest that students seem to have their own ways of conceptualizing the vector concept. Despite most students' previous exposure to vector concepts, there are still misconceptions and difficulties related to vectors. These studies suggest that instructors in introductory physics courses must give explicit consideration to students' familiarity with and learning of vectors. The question asked in this study is what is it that students know about vectors before doing the introductory physics at the university? To answer the question, we have developed a *vector test* which was moderated by two external assessors who are experts in the field, and given it to University of Limpopo (Medunsa campus) physics I students at the beginning of the semester.

The aim of the study was to investigate the students' pre-knowledge of vectors. The goal of the developed vector test was to check if the students possess the basic knowledge of vectors that will allow them to understand kinematics or Newtonian mechanics. The test was designed to measure students' knowledge on basic level of vectors, it contained questions based on addition of vectors algebraically and graphically, classification as vectors or scalars, expressing vectors in terms of magnitude and direction, and using trigonometric functions to calculate vectors.

2. Methods

The test was divided into three sections. The first section was a multiple choice questions consisting of twenty questions with four options to choose from. The second section was to classify quantities as vectors or scalars. It consisted of ten questions. The third section was calculations, where magnitude and direction were determined both algebraically and graphically. In the graphical representation, the graph paper was used and the graphs were expected to be drawn according to scale.

The test was administered to 234 physics I students from the University of Limpopo (Medunsa campus). The sample was divided into four groups, where the first group (n = 119) passed their grade 12 in 2010, the second group (n = 46) passed their grade 12 prior 2010, the third group (n = 42) did Foundation Physics in 2010 and the fourth group (n = 27) were those repeating Physics I. The students who participated in this study were not doing the same courses. Some were doing mathematics, others were intending to do physics as their major course, and majority were just doing it as a non-major course.

3. Results

The test was administered before the students could attend physics classes at the beginning of the year. The knowledge tested in this study was the knowledge acquired from the previous levels. The average score of the whole class was 45.8 % for the entire test. The groups did not show much difference in terms of their averages. Figure 1 below shows the average performance of the individual groups. The foundation group and the prior 2010 groups scored 47.3% and 47.4% respectively whereas the matric 2010 and the repeaters scored 44.3% and 44.1% respectively.

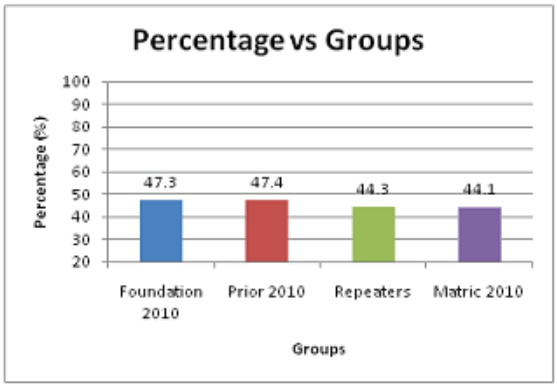


Figure 1: Percentages for groups

Table 1 summarizes the responses on multiple choice questions for each group.

Table 1: Students' response of question 1

Question 1	Classification	Matric 2010 **n=119** Response %	Prior 2010 **n=46** Response %	Foundation 2010 **n=42** Response %	Repeaters **n=27** Response %
1.1	Wrong	0	0	0	0
	Correct	100	100	100	100
	Not attempt	0	0	0	0
1.2	Wrong	5	0	4.8	0
	Correct	95	100	95.2	96.3
	Not attempt	0	0	0	3.7
1.3	Wrong	99.2	100	100	96.3
	Correct	0.84	0	0	3.7
	Not attempt	0	0	0	0
1.4	Wrong	32.8	43.5	71.4	44.4
	Correct	65.5	56.5	28.6	55.6
	Not attempt	1.7	0	0	0
1.5	Wrong	58	52.2	47.6	74.1
	Correct	40.3	45.6	47.6	25.9
	Not attempt	1.7	2.2	4.8	0
1.6	Wrong	12.6	8.7	14.3	7.4
	Correct	87.4	91.3	85.7	92.6
	Not attempt	0	0	0	0
1.7	Wrong	17.6	80.4	76.2	81.5
	Correct	81.5	19.6	21.4	14.8
	Not attempt	0.8	0	2.4	3.7
1.8	Wrong	91.6	91.3	100	92.6
	Correct	8.4	8.7	0	7.4
	Not attempt	0	0	0	0
1.9	Wrong	16.8	15.2	40.5	14.8
	Correct	83.2	84.8	59.5	85.2
	Not attempt	0	0	0	0
1.10	Wrong	31.1	34.8	61.9	74.1
	Correct	68.9	65.2	35.7	25.9
	Not attempt	0	0	2.4	0
1.11	Wrong	41.2	45.7	35.7	25.9
	Correct	58.8	50	64.3	74.1
	Not attempt	0	4.3	0	0

610 SA Institute of Physics 2011 ISBN: 978-1-86888-688-3

1.12	Wrong	8.4	12	7.1	3.7
	Correct	45	54.3	50	46.3
	Not attempt	47	33.7	43	50
1.13	Wrong	47.1	56.5	45.2	29.6
	Correct	52.9	43.5	54.8	66.7
	Not attempt	0	0	0	3.7
1.14	Wrong	68.1	52.2	50	51.9
	Correct	31.9	47.8	50	48.1
	Not attempt	0	0	0	0
1.15	Wrong	82.4	73.9	69	77.8
	Correct	17.6	26.1	31	18.5
	Not attempt	0	0	0	3.7
1.16	Wrong	64.7	54.3	59.5	40.7
	Correct	35.3	45.6	35.7	59.3
	Not attempt	0	0	4.8	0
1.17	Wrong	63.9	63	59.5	59.3
	Correct	36.1	37	40.5	40.7
	Not attempt	0	0	0	0
1.18	Wrong	41.2	23.9	52.4	44.4
	Correct	58.8	76.1	47.6	55.6
	Not attempt	0	0	0	0
1.19	Wrong	14.3	8.7	7.1	14.8
	Correct	85.7	91.3	90.5	85.2
	Not attempt	0	0	2.4	0
1.20	Wrong	31.9	26.1	57.1	37
	Correct	67.2	73.9	42.9	63
	Not attempt	0.8	0	0	0

The questions were classified as correct if the student got the correct answer, wrong if the student attempted the question but got it wrong and not attempted if the student left an empty space. The questions 1.1 and 1.2 from table 1 were asking for the definition of a scalar and a vector. It shows that all the groups managed to score very well on the definitions. The whole class were able to define a scalar irrespective of the group they are in.

For question 1.3, the correct answer was not among the given ones, meaning the answer was ***none of the above***. The two groups (foundation and the prior 2010) got it wrong. Only 3% of the repeaters managed to get it correct. This indicates that students do not work out the multiple choice questions, they just assumed that the correct answer is there and they pick up anyone.

Questions 1.4 and 1.5 were asking about determining the resultant of two vectors. Above 70% of the foundation group scored 1.4 wrong, while above 70% of the repeaters scored 1.5 wrong. It indicates that there is a problem with the knowledge brought forward by the two groups who were exposed to the university physics before in determining the magnitude and direction of two vectors. The matric 2010 group was the highest in scoring question 1.5 with 58%.

Question 1.8 was a true or false statement asking them to choose the statement which is not true. More than 90% of the students got it wrong. Most of them chose the statement which is true. It indicates that students do not read questions carefully. Majority of the students from all the groups did not attempt question 1.12 in full. The question had more than one correct answer. Most of the students gave one answer instead of two, the reason might be students thought that multiple choice questions have only one correct answer.

Students were given different drawings in question 1.15, they had to interpret the drawing and give the corresponding correct drawing based on the statement given. 82% of the matric 2010 group got it

wrong. This question indicated that it is difficult for students to relate a statement to a drawing or to interpret the drawings. 69% of foundation group got it wrong. It indicates the difficulty of the question, because this group had the opportunity of going through foundation level. Similar trend was observed with repeaters, who studied vectors during their first attempt; only 33% got it correct. Question 2 was classification. Students' responses are shown in table 2 below.

Table 2: Students' response to question 2.

Question 2	**Classification**	**Matric 2010 n=119** Response %	**Prior 2010 n=46** Response %	**Foundation 2010 n=42** Response %	**Repeaters n=27** Response %
2.1	Wrong	27.7	21.7	33.3	25.9
	Correct	69.7	76.1	61.9	70.4
	Not attempt	2.5	2.2	4.8	3.7
2.2	Wrong	25.2	37	31	29.6
	Correct	71.4	60.9	64.3	66.7
	Not attempt	3.4	2.2	4.8	3.7
2.3	Wrong	17.6	30.4	19	22.2
	Correct	79.8	67.4	76.2	74
	Not attempt	2.5	2.2	4.8	3.7
2.4	Wrong	16	15.2	9.5	8.3
	Correct	81.5	82.6	85.7	88.9
	Not attempt	2.5	2.2	4.8	3.7
2.5	Wrong	30.3	15.2	35.7	11.1
	Correct	67.2	82.6	59.5	85.2
	Not attempt	2.5	2.2	4.8	3.7
2.6	Wrong	8.4	4.3	4.8	11.1
	Correct	89.1	93.5	90.5	85.2
	Not attempt	2.5	2.2	4.8	3.7
2.7	Wrong	30.3	30.4	38.1	37
	Correct	69.7	69.6	61.9	63
	Not attempt	0	0	0	0
2.8	Wrong	63	58.7	45.2	81.5
	Correct	36.1	41.3	52.3	18.5
	Not attempt	0.8	0	2.4	0
2.9	Wrong	61.3	63	42.9	63
	Correct	37.8	37	57.1	37
	Not attempt	0.8	0	0	0
2.10	Wrong	18.5	26.1	28.6	22.2
	Correct	81.5	73.9	71.4	77.7
	Not attempt	0	0	0	0

Questions 2.1 to 2.6 were quantities, from which students had to classify them as either a vector or a scalar, for example 10N. Most of the students got them correct. For questions 2.7 to 2.10, the questions were in a statement form. The students had to read the sentence and decide whether it is a scalar or a vector. Students scored low marks for questions 2.7 – 2.10 as compared to question 2.1 – 2.6. It means it is difficult for students to read sentences and classify them as compared to classifying one word. Table 3 below shows the results of question 3. Question 3 was graph drawing based on calculations. Question 3 was rated 30 marks, and the number in the parentheses is the total mark for each section.

612 SA Institute of Physics 2011 ISBN: 978-1-86888-688-3

Table 3*: Students' response of question 3.*

Question 3 (30)	**Matric 2010 n=119**	**Prior 2010 n=46**	**Foundation 2010 n=42**	**Repeaters n=27**
3.1 (6)	0.5	0.8	2	1
3.2 (8)	1.1	1.4	1	1
3.3 (16)	3.2	3.8	5	3
Average	4.8 = **16%**	6 = **20%**	8 = **26.7%**	5 = **16.7%**

Foundation group was the highest in question 3.1 and 3.3 whereas the repeaters were the lowest in question 3.2 and 3.3. Question 3.1 students were expected to use parallelogram method, question 3.2 students were expected to resolve a vector in its components and calculate magnitude and direction, and question 3.3 students were expected to calculate the resultant of two vectors in vector component form and calculate its magnitude and finally represent them graphically. All groups scored below 30%, which indicates that students lack some knowledge, especially graphical representation of vectors.

4. Conclusion

This study showed that all students move to the next level with certain knowledge of vectors from the previous levels. There are some questions that students who passed their matric in 2010 performed better than those who went through foundation course and those who are repeating the same level. It can be concluded that high school physics gives some background on the topic. The prior 2010 group, are the students who were not studying during 2010 and they are not the product of OBE system. Their performance was good as compared to all the groups. It can be concluded that the education system plays a role in students' prior knowledge, since some concepts are not covered in the new system.

[5] and [4] have documented student difficulties with both algebraic and graphical aspects of vector concepts among students in introductory physics courses at several institutions. Those difficulties were similar to what we discovered in this study. Most students seem to lack a clear understanding of what is meant by vector direction. Many students are confused about the head-to-tail method and parallelogram addition rules.

This study recommends that, some instructional materials have to be developed and the amount of time given to vector concepts be increased because the vectors are the core of mechanics sections in introductory physics.

References

[1] Barnett R and Ziegler M 1994 *Analytic Trigonometry with Applications* (Boston: PWS Publishing).

[2] Cataloglu E 2006 Open source software in teaching physics: A case study on vector algebra and visual representations. *The Turkish Online Journal of Educational Technology*. **5(1)**, 68-74.

[3] Hestenes D and Wells M 1992 Mechanics baseline test. *The Physics Teacher*. **30**, 159.

[4] Kanim S E 1999 An investigation of student difficulties in qualitative and quantitative problem solving: Examples from electric circuits and electrostatics, Ph.D. dissertation, University of Washington.

[5] Knight R.D 1995 Vector knowledge of beginning physics students. *The Physics Teacher*.**33** 74–78.

[6] Nguyen N and Meltzer D 2003 Initial understanding of vector concepts among students in introductory physics. *American Journal of Physics 71***(6)** 630-638.

[7] Ortiz L.G, Heron P.R.L, Shaffer P.S and Mc-Dermott L.C. 2001 Identifying student reasoning difficulties with the mathematical formalism of rotational mechanics, *AAPT Announcer* **31**(4) 103.

[8] Poynter A and Tall D 2005 Relating Theories to Practice in the Teaching of Mathematics.

The transition from high-school Physics to first-year Physics: How well prepared are our students?

P Molefe and B M Sondezi-Mhlungu

University of Johannesburg (UJ), P.O.Box 524, Auckland Park, 2006

Email: pmolefe@uj.ac.za

Abstract: South Africa is currently experiencing an increasing in number of students pursuing studies in Science, Engineering and Technology (SET), however low retention rate in universities is still an outstanding problem. Most of students' drop-out occur during the first semester of their firt year of university studies.The current National Senior Certificate (NSC) curriculum has shown an increase in number of students obtaining minimum requirements for university entry. However, this is not translated into the performance of the students by the end of the first semester during the first year of university studies.

The students entry-level preparedness became an issue of concern and investigations in this regard were undertaken. The students' high-school NSC results, administered diagnostic test (test written by students during orientation period) and first year first were analyzed. The results obtained gave some recommendations and suggestions as to the method to follow in trying to deal with the current status in our universities.

1. Introduction

South Africa as a developing country amongst other Southern African countries, having a relatively standard of living, still has the lowest proportion of its university graduates in SET. This is the current status despite the new Department of Education's (DoE) science curriculum in an endeavour to increase the number of SET graduates as demanded by the country.

The low numbers of SET graduates might be seen as the local problem, but it resembles the situation in much larger and resource rich countries such as the United States. The research investigations over the last 15 years in academic success and persistence within the engineering programs has been identified by French et al (2005), as being linked to a declining interest in engineering amongst graduating high school students and low completion rates by students entering US universities as engineering majors. Besterfield-Scare, et al, (1997) has seen the first year as critical for both academic and retention of engineering students.

In South Africa there is an increase in number of students pursuing studies in SET and relatively low first semester completion rates by students entering first year university studies. With the evidence given by DoE in 2010, the number of graduates in SET demanded by the country is small as compared to other fields of studies. At University of Johannesburg (UJ), the first semester of the first year of study is one of the major factors for early drop-out from university, since a pass in this semester is a pre-requisite to enroll for second semester.

Since the introduction of National Senior Certificate (NSC) in 2008, more students are passing grade 12 with university entry (Admission Point Score (APS)) required. Within the past two years South African universities witnessed an overflow of students registering for first year physics. This course at UJ is a key module in SET-related degrees, as it is not only chosen by students who are

614 SA Institute of Physics 2011 ISBN: 978-1-86888-688-3

doing Bachelor of Science degree (BSc), but is also a requirement for studies in other degrees such as Bachelor of Science (BSc LE)-Life Environmental Sciences and BSc Eng (Mechanical, Civil, Electrical and Engineering with IT). However, irrespective of the increment in enrolment rate, attrition rate has been frequently reported to be high. The problem has been pronounced since 2009, the year after the first group of learners graduated from NSC curriculum and admitted in universities.

It is known that strong academic background, achievement of good grade, and academic motivation are needed for students to persist in their science studies. It is realized that if physics educators are to respond to country's calls for an increased number of SET graduates, a large portion of that increase is likely to come from a more diverse range of students, including women, as well as students with lower entry-level qualifications in mathematics and physical science. It is also realized that while these students may initially struggle academically they may well have the potential and motivation to make a significant contribution to the SET profession, provided appropriate academic and social support systems, remedial and "catch-up" courses are provided.

For first year courses, the diversity of students' academic backgrounds is a continuous challenge, particularly in courses where a level of pre-requisite knowledge is assumed. In South Africa (SA), Senior Certificate Examination (SEC) was a well-established indicator for University entrance for many years, providing a certain measure of surety about levels of pre-requisite knowledge based on higher and standard grades curricula. The current situation, with introduced high school qualification (NSC) has been the subject of debate, challenging previous assumptions about the commonality of prior knowledge. It was seen essential to identify the academic preparedness of current students in their first year studies in order to enable the curriculum, assessment and teaching methods to respond effectively, but it had also been perceived that the new and more diverse entry qualifications introduced over the past two years had made this identification task very difficult.

The dual objectives of identifying students with the ability to succeed in SET, and ensuring their academic success and retention to graduate were the motivations for the project from which results for this paper were drawn, with the specific objectives:

- to analyze educational background of the students in first year physics,
- to determine the educational achievement of the students in first year physics and
- to analyze the assessment results.

2. Methodology

In 2011, UJ registered approximately 1182 first year physics students from various disciplines of SET degrees, of which 626 students enrolled for a three year degree and 556 students spread between life sciences and four-year degree. The students entry-level preparedness was investigated by analysis of their high-school NSC results, administration of a diagnostic test and first year first semester marks. The key factors to success in this module were investigated by comparing high-school results and final first year results. APS methods are selections tools as well as predictors of academic success. These results together with the diagnostic test completed by students during the orientation week, including the overall performance from the first semester of 2011, will be studied and analyzed in this work. The observed data was fitted with Gaussian expression ($y = a\exp^{[-0.5(x-\bar{x}/b)]}$; where a is the maximum number of students that obtained an average mark, $\bar{x}$ is the average mark of the entire students and b is the parameter that indicates the goodness of fit) which is characterized by symmetric "bell curve" shape that quickly falls off towards plus/minus infinity.

As a further step in gathering information about the student's level of preparedness, a diagnostic test (written by 547 students, three year degree students) was administered during the orientation week

for the first year physics course. This 20 minutes assessment consisted of 10 questions. All questions were multiple-choice worth 3 marks each, having four possible options to chose from. They covered motions (graphs), vectors, force exerted on two parallel conductors, electric circuit, projectile motion, Newton first law, momentum and Newton's law of Universal Gravitational force. These questions were based on high school physics concept.

3. Results and discussion

3.1. Entry Level Preparedness

A bar graph of the 2011 enrolment of first year physics students (three year degree and other science related degrees mentioned above) based on APS, is presented in figure 1. It is important to mention that the minimum percentage of 60 is required for acceptance into the three year degree. It is indicated from the graph that approximately 53% (sum of the median value of all data above 60%) of students qualified for the course, and the remaining 47% were absorbed by life sciences and four-year degree. A Gaussian fit (red solid line in the figure) reveals an almost symmetric distribution in the data, which averages around 52.5%. This average percentage is in agreement with the number of students enrolled for a three year degree.

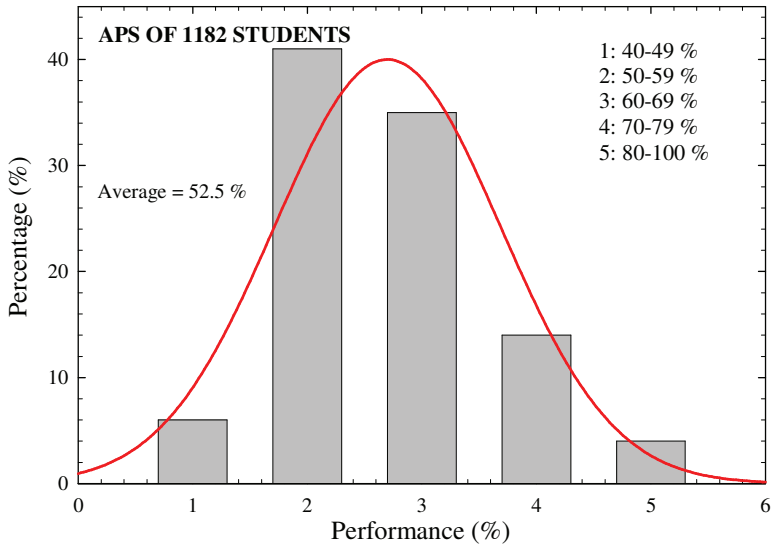


Figure 1: A bar graph showing enrolment of first year physics students based on APS 2011.

Figure 2 shows the results of the diagnostic test written by 547 students enrolled for a three year degree. The average mark obtained by most students was 25.9%, which is clearly a very disappointing result. They appeared to indicate that most of the students either had not understood or had forgotten much of the basic physics they had covered in their last three years of high school. This is also exhibited by the Gaussian fit (red solid line in the figure), which skews more towards the left hand side, as an indication of poor performance.

616 SA Institute of Physics 2011 ISBN: 978-1-86888-688-3

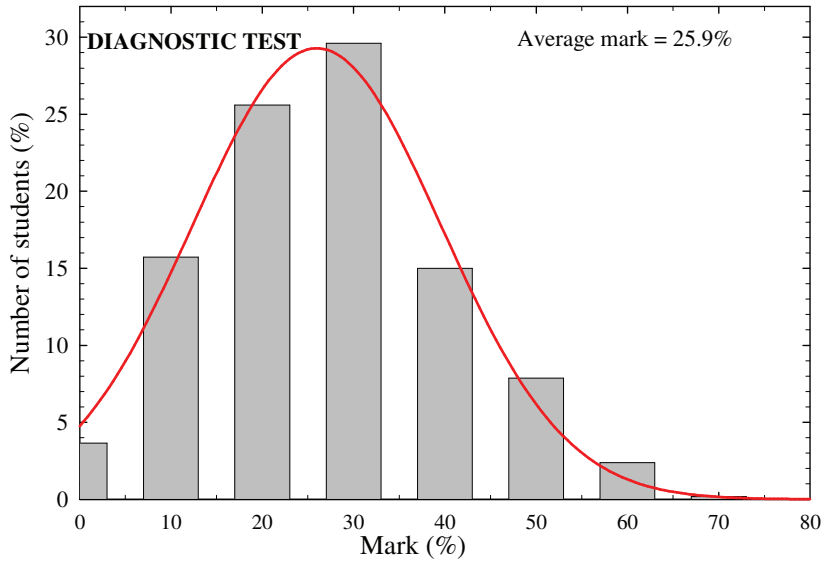


Figure 2: Performance of the 2011 diagnostic test.

3.2. First semester assessment results

Figure 3 below shows the achievement performance of student during their first semester of university studies. It is observed that only 32% (accumulated by adding all the students who obtained at least 50% in their semester's assessment) passed their first semester. This analysis provides evidence that APS alone cannot be solely used as a tool to assess the readiness of the students in studying first year physics. These results correlate more with the diagnostic test as compared to APS where the average mark of about 25% is obtained in both cases.

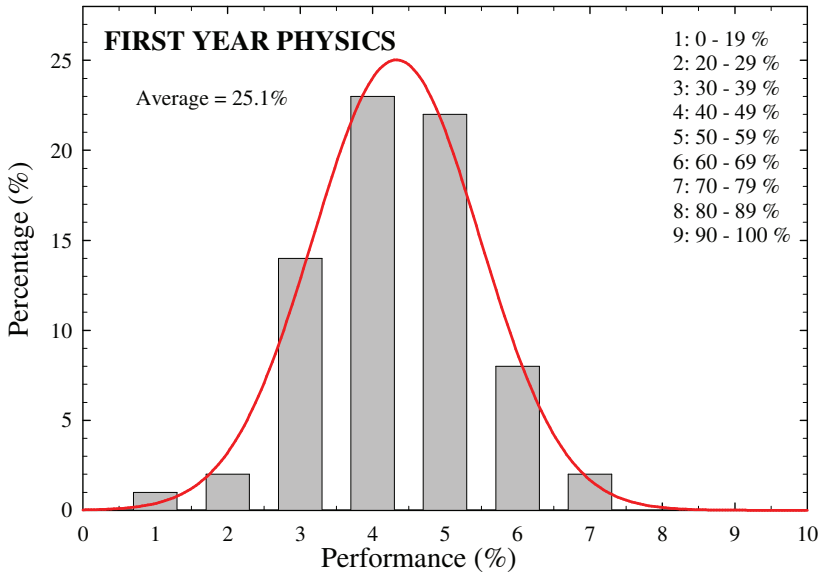


Figure 3: Overall performance of 2011 cohort by the end of the first semester.

4. Conclusion and recommendations

Analysis of the comparison between APS and the first semester module results indicate that the most significant determinant of examination success in first year physics was whether students had thoroughly mastered the NSC physics module rather than the physical science module (physics and chemistry) studied during their high school studies.

A significant difference is observed between the diagnostic test results and APS. Almost 50% difference is evident, which raises concerns on the reliability of APS. This diagnostic test was valuable as a "wake-up call" and led to behavioural changes to students and to some modification of teaching methods. It may also be worth doing the diagnostic test to combine it with the pre-course questionnaire probing students' confidence in their high school physics preparation, thus allowing a better gauge of the incoming students cohort.

The results of such test should not be used as a predictor of success, but as a guide to teaching, assessment and to motivate students. This can be of help to first year lecturers to know what students they expect from high school and what measures can be taken to improve retention in the first year class.

Acknowledgements

P Molefe acknowledges the assistance of B M Sondezi-Mhlungu and the support from the department of Physics at UJ.

References

[1] Department of Education. (2010). Education measurement, assessment and public examination: report on the national senior certificate examination results. A reference

[2] Besterfield-Sacre, M., Atman, C.J. and Shuman, L.J (1997). Characteristics of freshman egineering students: Models for determining student attrition and success in egineering. *Jounal of Engineering Educaion*. 86, pp. 139-149.

[3] Aboma, O, (2009). Predicting first year university students' academic success: Electronic Journal of Research in education Psychology, 7(3), 1053-1072. 2009

[4] Smaill, C, Godfrey, E, and Rowe, G.B. (2007). The transition from final-year high school Physics and mathematics to first-year Electrical Engineering

Instructional Design Principles applied to Physics laboratory and tutorial courses

D. Naidoo and D. Clerk

School of Physics, University of the Witwatersrand, Private Bag 3, WITS, 2050

Deena.Naidoo@wits.ac.za

Abstract. This study examines the effect on the learning experiences of students of changes to the running of laboratory and tutorial sessions of the first year auxiliary physics course at the University of the Witwatersrand, between 2001 and 2010. This research is underpinned by the principles of indirect interactive instructional skills and experiential learning. Two concerns drove this study: the first being a mismatch between the marks awarded to students for their laboratory reports, and their subsequent performance in formal practical tests. The second was the students' poor engagement – and hence performance in problem solving - during tutorial sessions. Several possible contributing factors were identified and changes were implemented in an attempt to improve the learning experienced by the students. Three sources of data suggest that the changes have had a positive effect on the learning of physics by the students in this course.

1. Introduction

The Physics I Auxiliary course is a non-continuing course aimed predominantly at Biology students - the class size varied between 300 and 600 students. The non-calculus curriculum has a balance between theoretical content and problem solving and is believed to be typical of algebra-based physics courses worldwide. The course is delivered by means of lectures, laboratory sessions and tutorials. Each student receives 180 minutes of lectures per week spread over four sessions, one 45 minute tutorial session per week and a three hour laboratory session every two weeks. In the alternate weeks, an additional one-hour tutorial session is held, followed by a one hour session in the computer aided laboratory (CAL) for underperforming students. In the CAL, the student can access interactive programs, with simulations and animations and can write either practice tests – with the aid of a tutor - or else formal tests for which marks are accumulated.

***Laboratory Sessions*:** From 2001 – 2005, each student performed ten exercises during the course of the year. The laboratory manual contained a comprehensive description with detailed, explicit instructions on how to perform each exercise. The students recorded their data by filling in the blanks. Each laboratory session began with a 20 to 30 minute preparatory talk, given by the laboratory demonstrators, outlining in detail the various aspects of the experiment to be conducted. Most of the balance of the 3 hour session was dedicated to conducting measurements, data analysis and interpretation, the plotting of graphs and report writing. During these processes a fair amount of assistance was given by the demonstrators. In the final half hour of the session, the laboratory reports (defined as laboratory work within text to follow) were submitted by the students to the

demonstrators who were required to complete the assessment of the reports by the end of each session.

***Concerns*:** Several authors [1,2] maintain that 'cookbook' activities are generally uninteresting, unchallenging and not engaging to the students. In addition to this criticism, we note that students were seldom well prepared for laboratory sessions – perhaps feeling that there was no need, given the detailed 'recipe' they had to follow. This led to a 'robotic' performance of the exercises with little understanding of the underlying principles - to compound this, laboratory demonstrators often gave too much assistance during the exercises. Also, sessions frequently ran overtime. Laboratory reports were poorly assessed because of the inadequate time allowed and as a result, laboratory report marks tended to be much higher than the corresponding laboratory test marks.

***Tutorial sessions*:** Tutorials are arguably the primary learning activity of any first year Physics course, in particular for problem solving. In 1999, Heller and Heller [3] emphasised the importance of cooperative group work in solving text-rich problems. The guided discovery approach developed by Redish [4] and the Technology Enabled Active Learning program designed by Belcher [1] both highlight the importance of small group learning activities. Although at Wits we have long employed "small" group tutorial sessions, we continually seek to improve the effectiveness of learning activities. Prior to 2006, students were expected to prepare approximately 10 to 12 tutorial exercises on a particular section for a session. The tutorial questions were designed to complement the lecture notes and the lecture examples. The first 35 minutes of a session was a question and answer (Q&A) session in which the tutor helped students and gave feedback on their efforts. A tutorial test based on this section was then given during the last 10 minutes of the session. The tutorial group sizes varied between 25 and 35 students. Tutors were mainly postgraduate students who were supplied with solutions to tutorial questions together with guidelines of how to conduct the tutorial.

***Concerns*:** Interactive learning was fairly non-existent in these sessions which generally resembled "recitation sessions" where tutors and to lesser extent students solved problems on the blackboard. It was also felt that groups were overlarge, students and tutors alike were inadequately prepared and there was a lack of expected interaction between students and tutors. Also there was no proper training for tutors who tended to give inadequate feedback to the students. In addition there were too many multi-level questions in the tutorial exercises.

1.2. Research question and Interventions

Will the implementation of simple innovative procedural changes to the running of practical and tutorial sessions result in improvements in the learning of physics by first year students?

***Practical Sessions*:** From 2006, the following changes were implemented: greater pressure was placed on students to prepare for the laboratory. During the first 30 to 45 minutes of the session, students were required to explore the apparatus and to attempt to set it up. This allowed them to 'play' with – and thus become familiar with - the equipment and hence to improve their confidence. Demonstrators were instructed not to assist the students with this process unless they ran into difficulty. They also monitored the progress of students during this part of the session and encouraged student engagement by mini-question and answer sessions. After this students conducted their measurements and wrote their reports. These were submitted at the end of the laboratory session and demonstrators had two weeks to assess the reports. This longer timescale for the assessment of reports resulted in reports been critically marked with substantial feedback and comments which had been lacking before. Also, the assessment of laboratory reports was

monitored by the laboratory co-ordinator and the standard of the laboratory test matched that of previous years' tests.

Tutorial Course: began with a reduction of the group size to a maximum of twenty during the fourth block of 2006. The tutorial sheets were revised to reduce the amount of conceptual overlap and duplication. This resulted in a smaller number of exercises (5 to 8) with escalating levels of difficulty. Preparation of selected material was made compulsory and tutors were required to provide feedback on the selected problems. A "floating tutor" was appointed to monitor individual students' preparation. A full set of tutorial solutions was posted on the notice board after completion of each section of the syllabus. Tutors also attended regular meetings with the course coordinator at which concerns could be aired and roles clarified. A spot test is given during the last 10 minutes of the session to monitor individual performance on a weekly basis.

1.3. Findings

Below is a brief summary of the qualitative and quantitative data obtained from student's performances and surveys based on changes implemented in the tutorial and laboratory courses.

Practical Sessions: Figure 1 shows a comparison of marks (%) accumulated for both the laboratory work and corresponding laboratory tests over a ten year period.

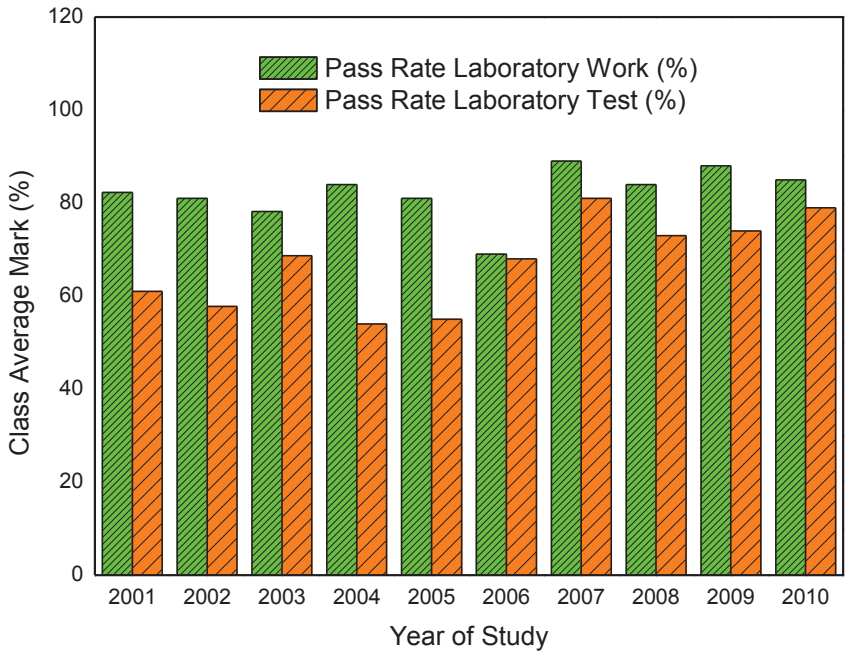


Figure 1: Comparison of laboratory report and test marks between 2001 and 2010.

During the period 2001 and 2005, the data shows on average approximately a 20% mismatch between the laboratory work and tests marks. This difference prompted the authors to investigate the reasons for the discrepancy in results. After careful consideration of a number of concerns highlighted above, it was tentatively concluded that a proper understanding of the aims of laboratory course was not fully achieved. In year 2006, the changes to the operation of the laboratory activities discussed above seem to show a better correlation between the laboratory work and test marks. Therefore these minor changes of indirect interactive instructional skills though far from ideal seems to be effective and has contributed to the improved experiential learning by students.

The validity of our results was tested by conducting a survey with the assistance of the Centre of Learning, Teaching and Development (CLTD) unit of the University in 2006. A questionnaire comprising four open ended questions was designed to ascertain whether there was evidence of improved learning. This survey was aimed at the laboratory demonstrators as the new students in 2006 could not have had any experience of the old dispensation. The responses of the demonstrators were quite strongly positive as illustrated below.

What impact has the modification of the laboratory structure had at these sessions?

- ✓ *Students are strongly encouraged to prepare (read in advance) for the laboratory prior to attending a session unlike other first year courses where the demonstrators give lengthy introductory talks.*
- ✓ *Students now understand that laboratory demonstrators are there to guide them instead of undertaking the majority of the laboratory work.*
- ✓ *They have realised that independent and group work is important when discussing and sharing experimental ideas.*
- ✓ *Their learning skills have improved which is reflected by their marks.*
- ✓ *Students have become more independent.*

How are the students responding to this format compared to the structure of other first year courses?

- ✓ *At the beginning of the year, it was not easy but they adapted and shown keen interest in the course.*
- ✓ *Students became very anxious about the apparatus and experiments and arrived early for classes.*
- ✓ *They realised that they were required to do extra work prior to the laboratory session.*

How has this change impacted on your demonstration skills?

- ✓ *This has improved my "co-operative learning", problem identification, leadership and group work skills.*
- ✓ *I have developed a new way to test student's skills and my own skills.*
- ✓ *This model allows us a new method to assess student's understanding which improves the relationships between student and laboratory demonstrator.*

To what do you attribute the positive changes in the delivery of the course?

- ✓ *In the past, demonstrators used to do most of the laboratory work and students were lazy. Now students prepare well, understand the material, and are able to work independently.*
- ✓ *Students ask more questions during the sessions.*
- ✓ *The laboratory is now full of activity, more interaction between students and demonstrators.*
- ✓ *The structure of course is clear; students have a lot of flexibility in learning of equipment and measurements either independently or as a group.*
- ✓ *I feel that the change in structure, the small group sizes (less noise), and interactive classes has resulted in a good environmental for experimental work.*

Tutorial Sessions: Table 1 shows a summary of student responses based on a CLTD questionnaire which was administered to 96 students to assess their reaction to the changes.

Table 1: Student responses based on CLTD questionnaire.

Probe Questions	Positive Responses	Negative Responses	Neutral
How has the change in the tutorial structure (compared to the way it was done before) impacted on your work?	85	2	7
How have the tutors improved their delivery of feedback to tutorial questions?	70	16	10

A summary of the qualitative analysis of their responses is given below:

"*The majority of students appreciated the enforced preparation prior to tutorials which enables them to obtain answers to their problem questions. They comment that the new system allows them to focus better and that much more is achieved in tutorial sessions. There is a sense that students enjoy taking some responsibility for their learning experience. It seems there are different standards of teaching skills amongst the tutors and also some communication and language issues. Overall the new system seems to be popular and well received by students*."

1.4. Conclusions and Implications

There seems to be evidence of improved learning in both the laboratory and tutorial components of this course - the students are more confident and better able to solve problems. The mismatch between their laboratory report marks and their laboratory test scores has been reduced and their performance in the laboratory test has improved. Some ideas for the future include a redesign of the laboratory manual is envisaged with the aim of introducing new fundamental experimental work which students can relate to everyday experiences. It will also be useful to conduct ongoing tutorial and laboratory course evaluations to monitor the success of innovative modifications made to a particular component(s) of the course. The postgraduate students who serve as tutors have limited teaching experience. Therefore, it is imperative that future modifications to these courses must include well defined tutor training programs to develop their teaching skills – especially methods to promote interactive learning and hence broaden student participation. A further improvement could be to extend the time for tutorial sessions to 90 minutes but this would require major restructuring of the overall curriculum.

Acknowledgements

The authors wish to express their thanks to the following staff members of the University of the Witwatersrand: (a) Ms Carol Cooper of the Centre of Learning, Teaching and Development (CLTD) unit for the preparation of the open ended questionnaires and the qualitative analysis of these surveys; (b) Professor John Carter for designing and developing the Physics I Auxiliary course over the last two decades; (c) Professor Barry Cole for restructuring the second semester lecture notes; (d) Mr John Crawford for sharing his expertise and continued input in making the course a success and (e) All the lecturers who contributed to improvement of course.

References

[1] Belcher J W 2001 *Studio Physics at MIT. MIT Physics Annual*.
[2] Hanson D and Wolfskill T 2000 *Journal of Chemical Education*, **77(1)** 120-130.
[3] Heller P and Heller K 1999 *University of Minnesota.*
[4] Redish E F 1999 *American Journal of Physics* **67** 562-573.

National Curriculum Statement achievement levels – Can they serve as a measure of science students' preparedness for university study?

Sam Ramaila, Padmanabhan Nair, Leelakrishna Reddy

Physics Department
University of Johannesburg
Johannesburg, South Africa
E-mail: samr@uj.ac.za

Abstract. The preparedness of matriculants for university studies has been a subject of intense debate and interrogation at various levels of public discourse. In addition, the standardization of assessment outcomes by the Quality Assurance Council in General and Further Education and Training (Umalusi) has of late become a highly contentious issue. It is in recognition of these considerations that we undertook to investigate the extent to which Grade 12 achievement levels in Mathematics, Physical Science and English can serve as a measure of science students' preparedness for university study. Analysis of data collected strongly suggests the existence of the disparity between the National Curriculum Statement achievement levels and the subsequent academic performance of the first-year Electrical Engineering (National Diploma Programme) students at the University of Johannesburg. Some of the underlying critical factors that may have a significant bearing on the aforementioned scenario were also investigated.

1. Background and contextualization

The implementation of the National Curriculum Statement (NCS) was largely characterised by a complex mix of fanfare and uncertainty. Developments in this regard posed critical challenges for teachers and various stakeholders with varied interest in education. Assessment is an inherent core component of the basic structure of the National Curriculum Statement and the achievement levels are increasingly used as a measure of learner performance and competence. The NCS achievement levels appear to present an unprecedented dilemma for tertiary institutions in South Africa when it comes to admission criteria and the subsequent academic performance of the students. This serves as a unit of analysis which prompted this research endeavour. The critical elements that constitute the disparity between the NCS achievement levels and the subsequent academic performance of the students form the nexus of the analysis in this respect.

At another level, tertiary institutions in South Africa and elsewhere are under severe pressure to produce skilled personnel through appropriate and relevant academic programmes. The relevance of academic programmes has been a highly contested issue over the years as it was felt that some of these programmes are out of sync with the needs of the labour market. To this end, there was a strong push for institutions to undertake far-reaching and strategic programme realignment in order to yield the critical human capital needed by the economy. Consistent with this trend, the throughput rate has largely been used as an essential monitoring tool to signify academic success. Yet, the manner in which it is determined makes for an interesting reading especially if one reflects on the ambiguity in relation to the determination of the throughput rate and the extent to which the National Curriculum Statement achievement levels complicates the situation.

2. Umalusi

Umalusi is a quality assurance council in the general and further education and training bands of the national qualifications framework (NQF) whose broad mandate is to ensure that the provision of education and training is carried out in accordance with expected standards of quality [1]. The council is guided by the General and Further Education Act of 2001 [2]. In terms of the clarification of roles, it is important to point out that the South African Qualifications Authority (SAQA) is entrusted with the responsibility to oversee the development of the NQF and establishment of Education and Training Quality Assurers (ETQAs). More specifically, Umalusi is tasked with ensuring quality in respect of levels of literacy and numeracy, meeting educational targets, raising standards in provision, linking and accessing higher education, preparation for the labour market, and social and economic development.

3. Standardization of assessment outcomes

Umalusi defines standardization as a process used to mitigate the effect of factors other than the learners' knowledge and aptitude on the learners' performance. However, the standardization process has been given a somewhat political interpretation by a considerable number of education commentators in South Africa as recent developments attested. In fact, education commentators in South Africa and indeed members of civil society appear to be extremely pessimistic about the purpose of the standardization process. The report released by Umalusi revealed that marks in some matric papers written during 2010 were adjusted upwards and others downwards. Ironically, upward adjustment affected key subjects such as Mathematics, Physical Science and English. Consistent with this trend, research conducted using the Stellenbosch University Access Test [3] revealed that grade inflation occurred particularly in the results of the lower performance group. Why should the standardization of assessment outcomes be a political terrain despite the reasons advanced by Umalusi for embarking on this exercise? Is it really an exercise in futility or an established law? This matter should be genuinely and honestly debated without compromising both the integrity and the quality of the assessment outcomes. In addition, South Africa cannot afford to replace the wheel while the car is moving.

4. National Curriculum Statement achievement levels

While the South African curriculum may be regarded to be on par with curricula elsewhere, it is the quality and integrity of the achievement levels that have always been a cause for great concern. Do these achievement levels really serve any purpose if indeed they are inflated? This question lies at the core of this research. It is a known fact that the National Curriculum Statement is characterised by a seven-point rating scale used to rate assessment for Grades 7-12 [4].

Achievement levels falling within 30%-39% band are always questionable in terms of required competence for undertaking tertiary studies. This appears to be a critical band in view of the findings of the research conducted [3] as pointed out elsewhere in this article. Are the achievement levels falling within this percentage band taken into account for purposes of unnecessarily inflating the overall matriculation pass rate so that it looks politically correct and acceptable? This key aspect should be fully and consciously debated as part of a progressive intellectual discourse in order to engender innovative solutions that will benefit the South African education system. However, the appropriateness of the context within which this debate should occur cannot be over-exaggerated.

5. Research design

Data collection involved the administration of a questionnaire to a group of first-year Electrical Engineering (National Diploma Programme) students ($n = 176$) at the University of Johannesburg, South Africa. The design largely employed quantitative analysis techniques.

6. Focus on the analysis of data

Data collected was quantitatively analysed to provide valuable insights into the research question as stipulated. The first phase of the analysis focused on the NCS achievement levels in Mathematics, Physical Science and English in Grade 12. Figures 1 below depicts the students' NCS achievement levels in respect of Grade 12 Mathematics.

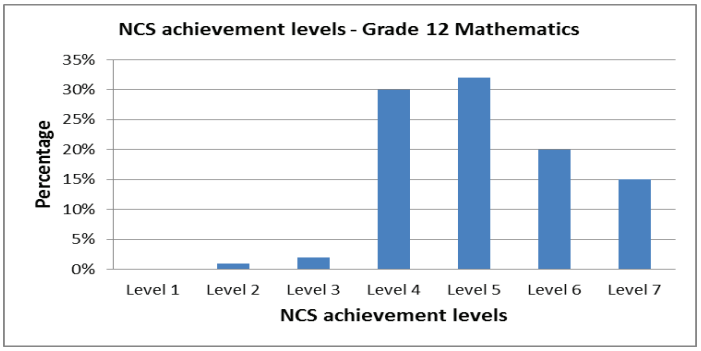


Figure 1: NCS achievement levels - Grade 12 Mathematics.

Of the 176 students in first-year Electrical Engineering (National Diploma Programme) a substantial number appear to have achieved level 4 (30%) and level 5 (32%) in Mathematics during the 2010 National Senior Certificate Examination. The number of students who achieved level 6 (20%) and level 7 (15%) is not pleasing. Schools have been urged to produce excellent results in Mathematics in order to pave the way for the learners to pursue higher education studies and subsequently professional careers in Mathematics, Science and Engineering.

Figure 2 below reflects the students' NCS achievement levels in Grade 12 Physical Science. The picture depicted in Figure 2 is not fundamentally different from the picture in Figure 1 in terms of the NCS achievement levels. The majority of the students appear to have achieved level 4 (38%) and level 5 (31%). This also points to the need for appropriate interventions to be put in place to generate the required excellence as an envisaged policy imperative.

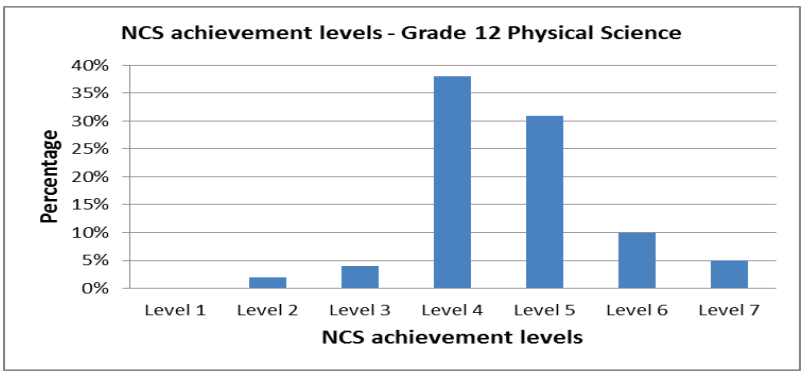


Figure 2: NCS achievement levels - Grade 12 Physical Science.

It is interesting to note that the picture presented in Figure 3 in relation to students' NCS achievement levels for English is not necessarily different from Mathematics and Physical Science. Can this be attributed to the fact that English is not the "mother tongue" for the group of students in this regard? If that is the case, will the performance of the students improve dramatically if they are taught Mathematics and Physical Science in their "mother tongue"? These are some of the critical questions that need to be broadly answered through similar research endeavours.

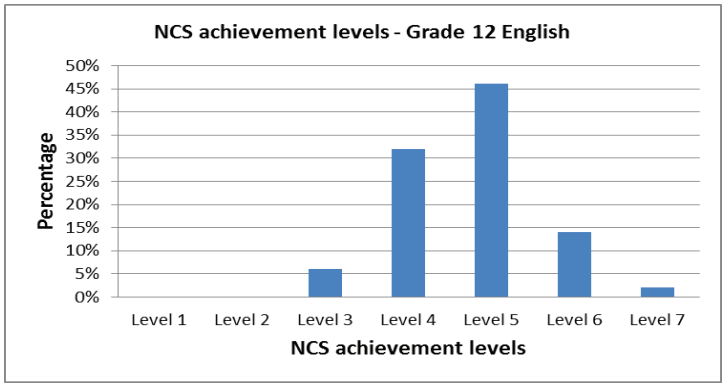


Figure 3: NCS achievement levels - Grade 12 English.

It is important to point out that the second phase of the analysis focused on the Physical Science syllabus coverage for Grade 11. Suffice to say that the syllabus coverage for Mathematics was not particularly investigated as the research undertaking was largely influenced by the authors' field of expertise. The picture obtained seems to suggest that some themes are covered at the expense of others. For instance, Longitudinal waves and Sound appear to have been covered the most as opposed to Geometrical optics and the Physics of music. The picture obtained in relation to the coverage of Matter and Materials as one of the key knowledge areas was quite disturbing. This may be attributed to the teachers' lack of competence and expertise in teaching this knowledge area.

While Doppler Effect appeared to be given prominence in terms of the coverage for Sound, Waves and Light in Grade 12, 2D and 3D Wave-fronts and Wave nature of matter appeared to be problematic in this regard. The coverage for Electricity and Magnetism in Grade 12 does not compare favourably with the coverage in Grade 11. The Photoelectric effect appeared to have been given adequate attention at the expense of other themes in terms of the coverage of Matter and Materials in Grade 12.

Figure 4 below reflects an unsatisfactory overall performance of the students in the three tests written during the first semester of 2011. The first test was based on the topics, Atomic structure, Wave nature of matter and Photoelectric effect. Test 2 was based on Vectors. While the students indicated that the section on Vectors was extensively covered at school level, this does not seem to translate into acceptable performance on this topic at tertiary level. Test 3 was a major test based on all the topics covered in Test 1 and Test 2 and Nuclear Physics as an additional topic.

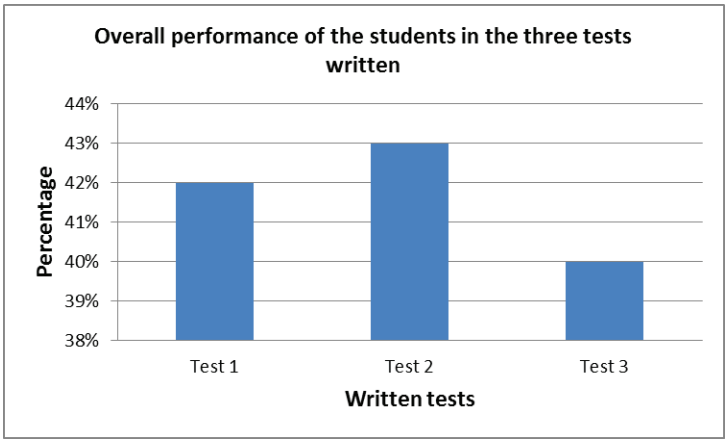


Figure 4: Average pass rates of the three tests written during the first semester of 2011.

The performance analysis for Test 1 in the respective percentage bands revealed that performance marks of a substantial number of students fell within the 0%-29% band. There was a marginal improvement in terms of the performance marks for Test 2 falling within the 30%-39% band in comparison with Test 1. This is still a worrying trend given the fact that this band essentially represents a low performance range. The performance analysis for Test 3 in terms of percentage bands was not drastically different from the analysis for Test 1 and Test 2.

7. Recommendations

Within the context of this research endeavour, various education stakeholders in South Africa are urged to make a profound strategic contribution towards the creation of a new social order that is in line with frontiers of human development. Concerted efforts in this regard must admittedly be targeted and directed towards the realisation of the quest for zero-defect in relation to the South African education system. The education authorities in this country may have to treat the problems associated with NCS delivery as an educational emergency. In the light of the aforementioned considerations, we put forward the following recommendations pertaining to NCS assessment outcomes:

- NCS examination papers must assess the acquisition of core skills such as performance of complex procedures and problem-solving to a significant extent by considerably improving the quality of the assessment items.
- NCS assessment opportunities must provide appropriate platforms for intellectual growth and adequate preparation for university studies as a direct consequence of quality teaching and learning.
- The quality of the assessment items must be responsive to the critical needs of learners in terms of cognitive demands and level of difficulty particularly in relation to Mathematics and Physical Science by putting more emphasis on testing of insight and inculcating critical thinking skills.
- NCS assessment opportunities must, by their very nature, serve to bridge the huge gap between school and university in order to facilitate a smooth migration between these educational levels.
- Epoch-making and tangible efforts are required to address the gross educational wastage emanating from a heavy reliance on throughput rates which do not necessarily provide a true reflection of academic success at university.

8. Conclusion

The disparity between the NCS achievement levels and the academic performance of the students in the subsequent higher education sector seems to be a pervasive problem which requires concerted efforts to address.

Acknowledgement

The authors express gratitude to colleagues in the physics department of the University of Johannesburg for providing assistance with the collection of data for this research.

References

[1] Umalusi - Council for Quality Assurance in General and Further Education and Training. (2000-2011). Pretoria.

[2] Government Gazette. (2001). *General and Further Training Quality Assurance Act*, 2001.

[3] Nel, C. & Kistner, L. (2009). The National Senior Certificate: Implications for access to higher education. *South African Journal of Higher Education*, **23**(5), 953-973.

[4] Department of Education (DoE). (2001). *Draft Revised National Curriculum Statement.* Pretoria: Government Printer.

Lightning – Scientific knowledge versus mythological beliefs

Sam Ramaila, Padmanabhan Nair, Leelakrishna Reddy

Physics Department
University of Johannesburg
Johannesburg, South Africa
E-mail: samr@uj.ac.za

Abstract. Lightning has more recently wreaked havoc during the rainy season in several parts of South Africa and this prompted the physics community under the auspices of the South African Institute of Physics (SAIP) to devise meaningful strategies for promoting public awareness. In response to this call, first-year Chemical Engineering (National Diploma Programme) students at the University of Johannesburg responded to a carefully designed questionnaire that seeks to probe students' level of understanding of lightning. Analysis of responses reveals lack of scientific understanding of lightning as a natural phenomenon. Amongst others, this can to some degree be attributed to superstitious or cultural beliefs.

1. Background and contextualization

Research on students' alternative conceptions produced a vast array of substantial literature over many decades. It is a known fact to the science education community that students undertake physics studies with their own understandings of the physical world. Students' inadequate understanding of natural phenomena has been widely reported in various settings in an attempt to meaningfully and comprehensively deal with alternative conceptions. These firmly established beliefs have the potential to persist as lingering suspicions in student's mind and can hinder further learning [1].

This research was largely triggered by rather interesting responses obtained from a diagnostic questionnaire administered to a group of first-year Chemical Engineering [National Diploma Programme (NDP)] students at the University of Johannesburg, South Africa. More specifically, the questionnaire administered seeks to probe students' level of understanding of lightning as a natural phenomenon. In a similar vein, researchers employed various multimedia interventions to probe students' understanding of lightning [2] & [3]. In particular, computer-based multimedia learning environments consisting of animated pictures and narrated words were employed for improving students' understanding [2]. Cognitive theory and multimedia design principles were employed to augment findings from computer-based multimedia learning environments [3]. This approach involves the design of multimedia learning environments coupled with reciprocal relation between cognitive theory and educational practice.

2. Theoretical framework

This research is underpinned by the Worldview Theory as the underlying theoretical framework. The Worldview Theory provides a non-rational foundation for thought, emotion, and behaviour and also provides a person with presuppositions [4]. In terms of the Worldview Theory, a person sitting in a science classroom is not just a science student but a thinking human being who sees the world in terms of a variety of other contexts influenced by gender, ethnicity, religion and so forth [5]. The implication of this intellectual discourse is that knowledge is then viewed as depending on a reasonably large number of different concepts, each refined through use, example, and experiences and consequently the intellectual picture of the world inside the student's mind includes prior conceptions or beliefs about the natural world. The Worldview Theory sees a person as having presuppositions about what the world is really like and what constitutes valid and important knowledge about the world [4]. These

presuppositions are regarded as views that a person holds about natural phenomena which include *commonsense*, *alternative frameworks*, *indigenous beliefs*, *misconceptions*, and *valid science* [4]. This in itself provides the justification for using the Worldview Theory as the underlying theoretical framework in this research.

3. Research design

As alluded to earlier, data collection involved the administration of a diagnostic questionnaire to a group of first-year Chemical Engineering (NDP) students (*n = 157*) at the University of Johannesburg, South Africa. The design largely employed quantitative analysis techniques.

4. Focus on the analysis of responses

It is important to point out upfront that students provided varied responses to the diagnostic questionnaire administered as detailed below.

1. *What causes static electricity?*

   A. *Static electricity is caused by stationary charges*
   B. *Static electricity is caused by deficiency of charges*
   C. *Static electricity is a balance between positive and negative charges*

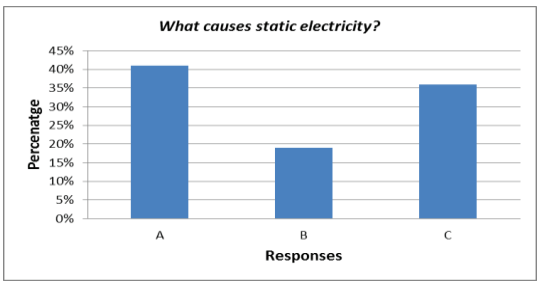


Figure1: Response pattern for item 1.

Item 1 on the diagnostic questionnaire reflects on the cause of static electricity. In terms of the responses, 41% of the students indicated that "*static electricity is caused by stationary charges"*. 19% of the students indicated that "*static electricity is caused by deficiency of charges"*. While this is not a scientifically correct response, it does seem to highlight some measure of alternative conceptions exhibited by the students. A further 36% of the students indicated that "*static electricity is a balance between positive and negative charges"*. This seems to suggest that students cannot differentiate between the terms *"static"* and *"balance"* in relation to the interaction of electrical charges. The interplay between language barriers and scientific understanding appears to be of critical importance in this respect. Both cultural and language concerns should be taken into account in learning situations [6 & 7].

2. *What causes lightning?*

   A. *A massive flow of electric current between the clouds and the ground following separation of charges*
   B. *Witchcraft*
   C. *Thunder*

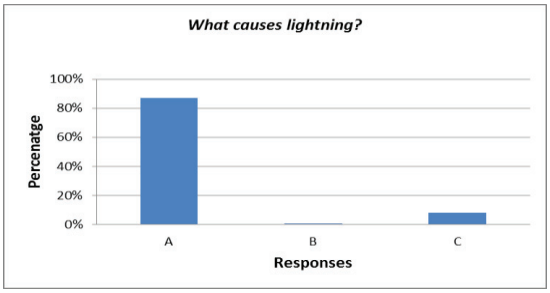


Figure 2: Response pattern for item 2.

Item 2 on the diagnostic questionnaire reflects on the causes of lightning. A vast majority of the students (87%) indicated that lightning is caused by "*a massive flow of electric current between the clouds and the ground following separation of charges".* Additional responses highlighted *"witchcraft"* (1%) and *"thunder"* (8%) as causes of lightning. This seems to suggest students' incoherent understanding of lightning and thunder as natural phenomena.

3. *How is lightning related to static electricity?*

   A. *Lightning is a form of static electricity*
   B. *No relation*
   C. *Both lightning and static electricity do not happen because of the attraction between the opposite charges*

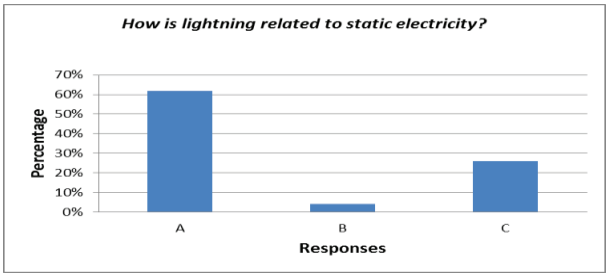


Figure 3: Response pattern for item 3.

Item 3 on the diagnostic questionnaire seeks to establish students' understanding of the relationship between lightning and static electricity. While a substantial number of respondents (62%) indicated that *"lightning is a form of static electricity",* a further 26% indicated that "*both lightning and static electricity do not happen because of the attraction between the opposite charges".* These responses paint a somewhat gloomy picture in terms of students' understanding of the relationship between lightning and static electricity.

4. *What causes a spark?*

   A. *Electrons moving across the atmosphere and heating up the air*
   B. *Electrons moving back and forth across the atmosphere*
   C. *Moisture in the atmosphere*

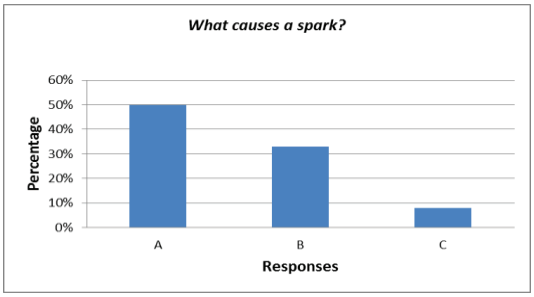


Figure 4: Response pattern for item 4.

Item 4 on the diagnostic questionnaire reflects on the cause of a spark. Some respondents (50%) indicated that a spark is caused by "*electrons moving across the atmosphere and heating up the air"*. A further 33% indicated that a spark is caused by "*electrons moving back and forth across the atmosphere"* while 8% indicated that a spark is caused by "*moisture in the atmosphere"*. For students to think that a spark is caused by "*electrons moving back and forth across the atmosphere"* and "*moisture in the atmosphere"* seems to suggest a lack of coherent scientific understanding of this natural phenomenon.

5. *How does lightning differ from a spark?*

   A. *Lightning occurs in summer while sparks occur in winter*
   B. *Lightning is a big spark*
   C. *Lightning makes noise but sparks don't*

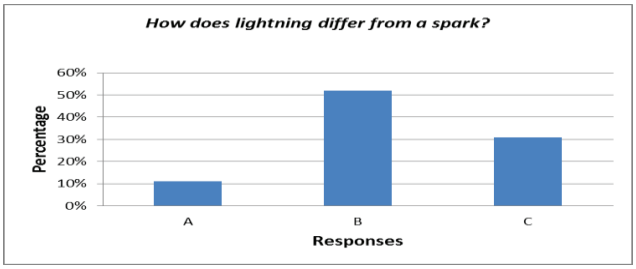


Figure 5: Response pattern for item 5.

Item 5 on the diagnostic questionnaire seeks to establish students' understanding of the difference between lightning and spark. In response to this item, 52% indicated that *"lightning is a big spark"*. Other respondents (11%) and (31%) indicated that *"lightning occurs in summer while sparks occur in winter"* and *"lightning makes noise but sparks don't*", respectively. It is interesting to note that some respondents seem to attribute the difference between lightning and spark to seasonal changes which clearly has no scientific basis in relation to the phenomena in question.

6. *What causes electrons to jump across the atmosphere and cause a spark?*

   A. *The force from a large accumulation of positive charge on the other side of the atmosphere*
   B. *Force of gravity*
   C. *Heating up of the atmosphere*

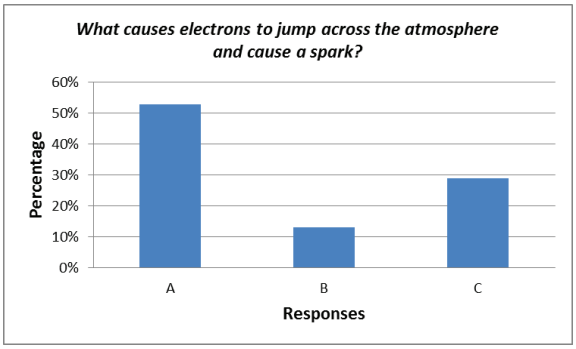


Figure 6: Response pattern for item 6.

Responses to item 6 provided an interesting pattern in the sense that 53% of the respondents expressed the view that "*the force from a large accumulation of positive charge on the other side of the atmosphere"* is what causes electrons to jump across the atmosphere and cause a spark. While 29% of the respondents opted for "*heating up of the atmosphere"* which is scientifically sound, a further 13% opted for *"force of gravity"*. The responses in this regard seem to suggest lack of coherent understanding of the nature of electrostatic force and the force of gravity.

7. *What causes thunder?*

   A. *Static electricity popping your eardrums*
   B. *The rapid expansion of air*
   C. *The actions of the rain gods*

The response pattern to this item seems to suggest that the majority of the students (79%) understand that thunder is caused by "*the rapid expansion of air"*. However, there are others (17%) and (4%) who hold the view that thunder is caused by *"static electricity popping your eardrums"* and *" the actions of the rain gods*", respectively. This may partly be attributed to religious or mythical beliefs as explained earlier.

8. *How do you minimize the risk of injury from lightning?*

   A. *Going outdoors and standing under trees and near tall buildings*
   B. *Staying indoors*
   C. *Using electrical appliances*

There appeared to be a fair understanding of the minimization of the risk of injury from lightning on the part of the respondents. However, fewer respondents (9%) still hold the view that the risk of injury from lightning can be minimized by *"going outdoors and standing under trees and near tall buildings"* and *"using electrical appliances"* which is not in line with precautionary safety measures.

9. *Why should we be away from water when lightning strikes?*

   A. *Lightning conducts through water*
   B. *Lightning does not conduct through water*
   C. *Water attracts lightning?*

While the majority of the respondents (73%) seem to understand that staying away from water when lightning strikes is an important precautionary measure in a scientific sense, there are others (22%) who do not particularly understand the scientific basis for adhering to this

safety measure. An understanding of the electrolytic nature of water as a polar compound appears to be inadequate if not lacking.

10. Is there any connection between lightning and witchcraft?

A. Yes
B. No

Most respondents (68%) expressed the view that there is "*no connection between lightning and witchcraft"*. However, some 27% of the respondents hold the view that there *"is connection between lightning and witchcraft"*. This seems to suggest a view or belief deeply rooted in religious or mythical teachings.

5. Recommendations

Myths associated with natural phenomena such as lightning among communities can be addressed through educating the younger generation as it appears to be extremely difficult to eradicate preconceived beliefs among the older generation. The provision of proper education based on established and generally accepted scientific principles to younger generation can produce a chain reaction which might turn out to be beneficial to older generation in the long term. One of the possible ways of achieving the above objective is by broadening the operational agenda and scope of the Physics Education Division of the South African Institute of Physics (SAIP) for purposes of playing a meaningful and an influential role. It is imperative for funding agencies to make financial resources available for the achievement of this noble goal. Through the provision of these much anticipated financial resources, there could be sessions in future SAIP conferences specifically dedicated to educational seminars, practical demonstrations and exhibits involving natural phenomena such as lightning, rainbow, eclipse, echoes, typhoons and so forth.

6. Conclusion

Analysis of students' questionnaire responses in this regard suggests incoherence and fragmentation in relation to scientific understanding of lightning as a natural phenomenon. There is an urgent need to embark on an intensive campaign to educate communities about the nature of lightning and related precautionary safety measures possibly with well-informed students from university as educators.

Acknowledgement

The authors express gratitude to colleagues in the physics department of the University of Johannesburg for providing assistance with the collection of data for this research.

References

[1] McDermott, L**. (**1991) 'What we teach and what is learned – Closing the gap'. *American Journal of Physics*, **59** (4), 301–315.

[2] Mayer, R & Moreno R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, **12**, 107-119.

[3] Mayer, R. (2002). Cognitive theory and the design of multimedia instruction: An example of the two-way street between cognition and instruction. *New Directions for Teaching and Learning*, **89**, 55-71.

[4] Cobern, W.W. (1996). Worldview Theory and conceptual change in science education. *Science Education*, **80**(5), 579-610.

[5] Shipman, H.L., Brickhouse, N.W., & Dagher, Z. (2002). Changes in student views of religion and science in a college Astronomy Course. *Science Education*, **86**(4), 526-547.

[6] Hallberg, D. (2010). Socioculture and cognitivist perspectives on language and communication barriers in learning. *International Journal of Human and Social Sciences*, **5**(8), 533-542.

[7] De Voogd, G. (1998). Computer levers power sharing: Multicultural students' styles of participation and knowledge. *Computers & Education*, **31**, 351-364.

Foundation Provision: Any Difference in Student's Performance?

B M Sondezi-Mhlungu; P Molefe
Physics Department, University of Johannesburg, P.O. Box 524, Auckland Park, 2006,South Africa
E-mail: bmhlungu@uj.ac.za

Abstract. Science, Engineering and Technology (SET) have become essential for the development of the country in all its facets. It has been observed over the past few years that students entering university for degree studies in their respective fields experience considerable level of difficulty. Inadequate understanding of basic mathematical and physical science concepts has been identified as a critical deficiency characterizing the performance of SET first year university students. As a norm, the intake to Higher Education Institutes (HEI's) is usually based on grade 12 results. Admission to three year (mainstream) programme or a four year extended programme is determined on the basis of Admission Point Scores (APS) set by respective institutions. Analysis of students' performance in the first year of their four year programme, at the University of Johannesburg (UJ), painted a bleak picture about the students' capabilities in tackling their first year work. This led to the introduction of the Foundation Provision Programme (FPP) which serves as a precursor to the normal first year curriculum. This programme serves to adequately develop the foundational competencies necessary for students to embark on successful first year physics studies and beyond. The sample in this study comprised of a group of freshmen from grade 12 registered for various degrees in SET. All these groups receive physics tuition in one lecture class. The FPP was implemented for the first time at UJ in 2010 and the results obtained were compared to the results of students' performance obtained from previous years.

1. Introduction

Different institutions of higher learning have different criteria of accepting students into their system. Although APS has been a standard norm, some institutions require tests to be written by the students before they can be admitted for their respective degrees. These tests have been used as means of screening the students and thereby providing the departments with the information as to whether the students can be admitted for a certain desired course or degree. In many institutions the culture of writing National Benchmark Tests (NBT) is re-introduced after some time where students were solely admitted based on their grade 12 results. From 2005, it turned out that the results obtained by the students in grade 12 were not necessarily a true reflection of the students' capabilities in their first year of university experience. These were not the necessary tools that could predict the students' performance especially in their physics lectures. Different approaches in trying to help the students cope with the amount of work to be conceptualized and applied thereof have been employed over the years. In 2005 UJ introduced SET programme in order to give the proper foundation in Maths,

Applied Maths, Physics and Chemistry needed by the students to undertake their respective degrees. Students would do their first year of SET and then be transferred into the first year mainstream degree, during their second year of university study. The curriculum has been continually re-visited and an extended degree was introduced in 2007 whereby the normal three-year degree was extended into four years, specifically designed to address the fundamentals of Mathematics and Physical Science for students who had lower than minimum required APS to be admitted into a three-year program. The extended programme splits the first year curriculum over two years. The aim of which was to help with the simulation of the physics concepts within the students more effectively. This has been an exercise full of unexpected results, since the performance did not improve as initially predicted.

Although the first year syllabus was still going to be done over two years, the student's performance at the end of the year was not satisfactory. This was evident in the academic results obtained at the end of 2009. From the beginning of 2010, FPP was introduced; that is, the first term of year one was dedicated to basic mathematics concepts, as these were identified as problem areas for most of the students coming fresh from high school. In 2011 intensity in the level of presenting FPP was increased in terms of time spent (from a term to a semester) in this module. This paper presents the observations and analysis of the physics students' performance after being taken through this programme.

2. Methodology

Students enrolled for any of SET extended degrees having Physics as one of the courses were taught FPP, which comprises of basic maths, (FOIL rule, BODMAS rule, Fractions, Brackets, Linear and quadratic equations, Straight line graphs, Simultaneous equations, Limits, Basic Trigonometry), vectors and forces was done prior to starting with first year syllabus. The other physics topics such as, motion in one and two dimension, forces and Newton's laws of motion, work and energy, impulse and momentum, waves and sound and the basics of electricity were part of FPP syllabus.
FPP was implemented for the first time at UJ in 2009. The performance of students after the first semester was observed and analyzed. It was perceived that the students enrolling for an extended degree were not well prepared or ready to undertake physics at a university level. A questionnaire was designed and given to students in their 10^{th} week of the first semester. These were aimed to establish the level of students' preparedness from their high school physics subject. Only selected questions are shown in table 1, for the purpose of this study. The responses received were considered as possible indicators likelihood that the students would succeed in their first year of foundation physics programme.

3. Results and Discussion

After a lot of changes have occurred in different sectors of South Africa, including some noticeable changes in education system, the first group of students under new system, National Senior Certificate (NSC) matriculated in 2008. This is the group that was accepted in HEI's in 2009, UJ also had a high number of students who registered after passing their grade 12 national examinations. The performance of this group of students during their first year university physics experience was of great interest, and the results are shown in figure 1. Approximately 82.5% of students obtained less than 50% (pass mark), whilst 28% obtained an average of 24.5%. This is seen as a low average mark obtained from the exam. The Gaussian fit shows its maximum peak (around 30%) broadening more towards the left (lower performance level), which confirms that most students did not do well in this course. It was observed that the trend of these results carried on to the end-year results (not shown; to be published elsewhere). This very low performance observed from the 2009 cohort is ascribed to the standard of grade 12 examinations written at the end of 2008 academic year.

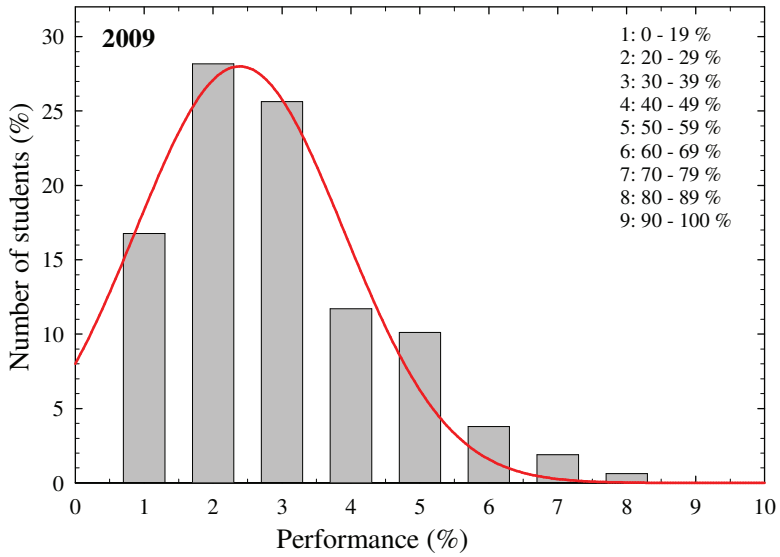


Figure 1: The histogram indicating the performance of students after they wrote their mid-year exam in 2009, and the red solid line is the Gaussian fit to the observed data.

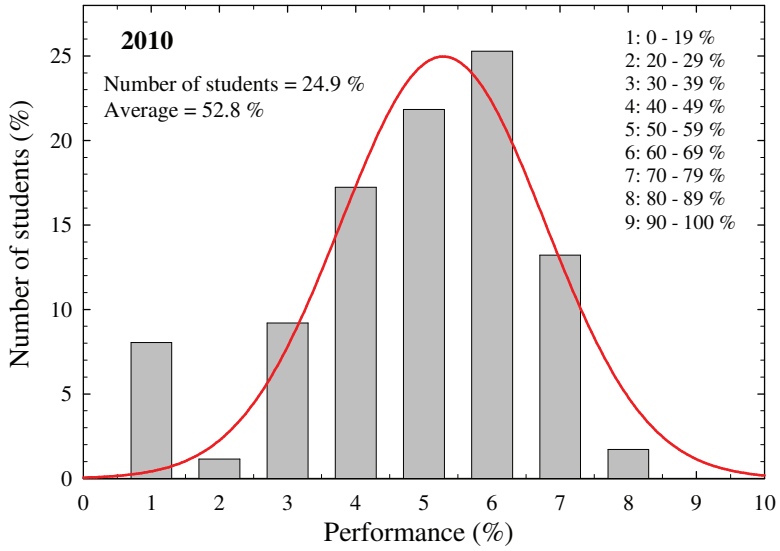


Figure 2: The histogram indicating the performance of students after they wrote their mid-year exam in 2010, and the red solid line is the Gaussian fit to the observed data.

The year 2010 (see figure 2) and 2011 (see figure 3) cohorts had shown an improvement in result, where most students achieved an average of 52.8% and 51.8%; respectively as compared to an average of 24.5% in 2009 cohort. This improvement can be attributed to the introduction of FPP that was taught in the first term (in 2010) and in the first semester (in 2011). The red solid lines in figure 2 and 3 were obtained from the fits of Gaussian distribution. The bell shape curve shows normal distribution which is evidenced by the maximum peak centralised around 50%. It is observed that a reasonable number of students (about 68%) managed to pass the semester in 2011.

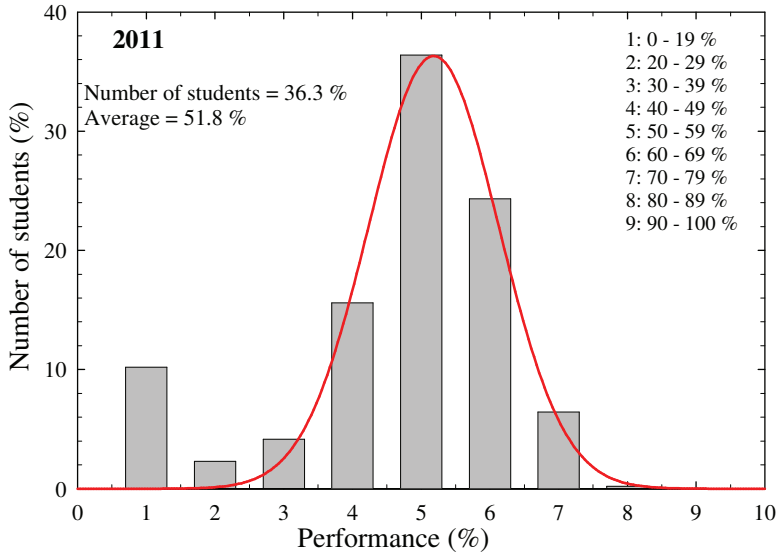


Figure 3: The histogram indicating the performance of students after they wrote their mid-year exam in 2011, and the red solid line is the Gaussian fit to the observed data.

Table 1: A table showing selected questions used for the survey.

Question Number	Question Phrasing
Q1	How well did your experiences in grade 12 Physics and Mathematics prepare you for this module? Make a cross (X) on the relevant box.
Q4	Are there any areas of FPP in which you had to "unlearn" concepts you had previously learnt at high school? _____. If so, please specify the concept(s).
Q6	Are there any areas of FPP that repeated material that you had already thoroughly mastered at high school? _____. If so, please elaborate.
Q7	Was the introduction of Basic Mathematics helpful in understanding the Physics topics studied in this module? Make a cross (X) on the relevant box.

Selected questions from the questionnaire are given in table 1, and the responses from students are presented in bar graphs. Figure 4 shows the number of responses received from students against responses. Students had to answer whether there were any concepts learnt from high school which needed to be unlearnt. The responses to this question indicated that 70.4% of them needed not to unlearn their high school concepts, whereas only 20.6% spotted some misconceptions which needed to be corrected and unlearnt. This high percentage in the response of Q4 is in agreement with the response of Q6, which needed to establish whether there are concepts that they thoroughly mastered from high school. Almost 60% of the students indicated that they had previously mastered (from high school) most of the concepts given in FPP. If students had a good foundation from their high school, it is expected that they should be able to undertake first year university physics more effectively.

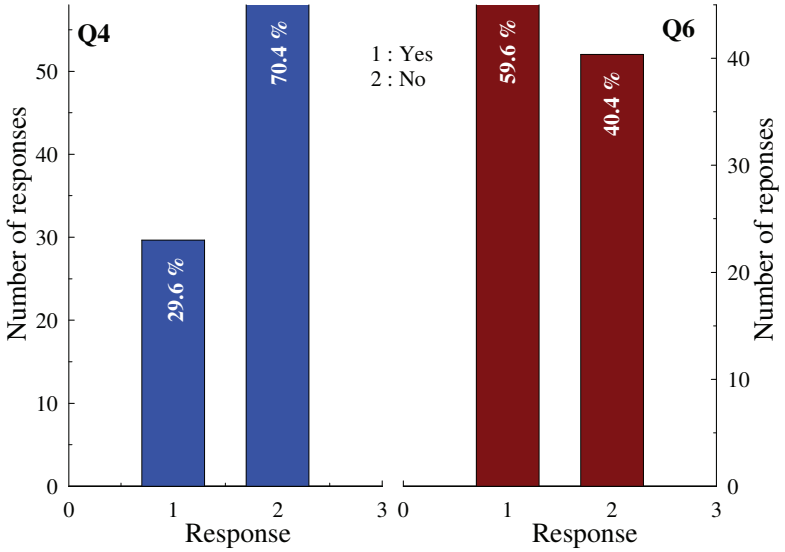


Figure 4: The histograms showing the responses of students for question 4 and 6.

In response to Q1 (see figure 5) where it had to be established how well grade 12 prepared the students for university; almost 81% responded that grade 12 prepared them well. This overwhelming response supports Q4 and Q6's response, implying that if the students were well prepared in their grade 12 (Q1), there was no need to unlearn concepts when they were exposed to FPP (Q4) due to the fact that they thoroughly mastered most of the concepts from high school (Q6), then it is expected that the pass rate in this module should be reasonably high. However, the response of Q6 was a bit confusing in a sense that, if the responses of the questions discussed above are positive (grade 12 Physics and Mathematics prepared them very well, there was no need to unlearn the pre-conceived concepts, which they have thoroughly mastered), then the response of Q7 gives a very important point to be taken care of in presenting this module. Most students' response (almost 86%) reveals that the introduction of Basic Mathematics was helpful in understanding the Physics topics studied in this module (FPP). This is confirmed by their 2011 mid-year results where an average of 51% was achieved.

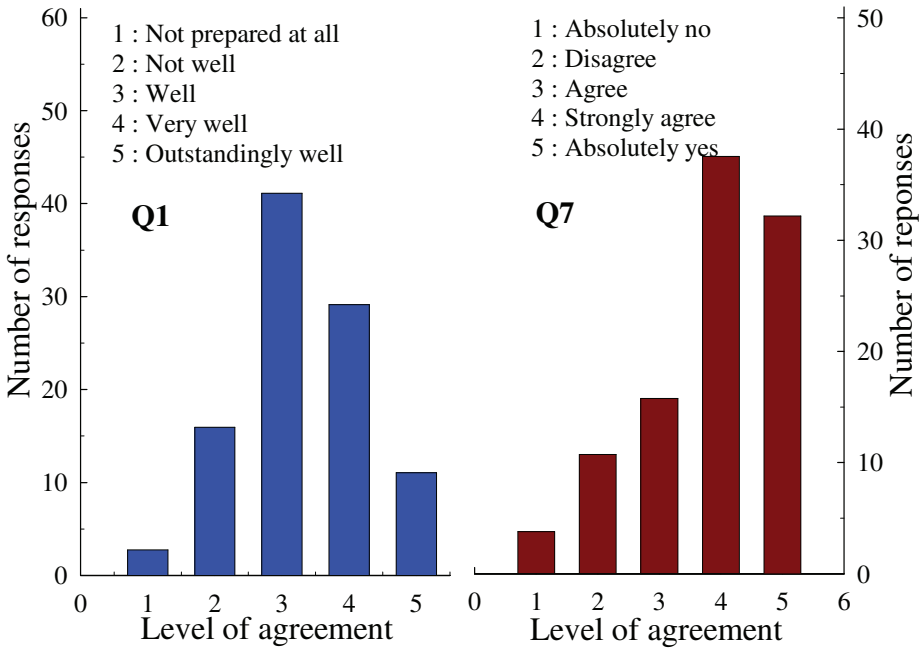


Figure 5: The histograms showing the responses of students for question 1 and 7.

4. Conclusion

The observation at hand indicates a great improvement in students' performance due to the introduction of Foundation Provision Programme. Further analysis is ongoing as to establish the reasons behind the constant average mark of the 2010 and 2011 cohorts, despite the difference in time spent in FPP 2010 and FPP 2011. From this outcome, it is evident that the introduction of FPP played an important role in improving the results of the students. The introduction or the addition of Basic Mathematics as a topic in FPP module serves as a tool in understanding Physics and should be an integral component of the module, for the maximum benefit of the student.

Acknowledgements

B M Sondezi-Mhlungu acknowledges the assistance of P Molefe and the support from the department of Physics at UJ.

References

[1] Department of Education. (2010). Education measurement, assessment and public examination: report on the national senior certificate examination results.
[2] Tinto, V, (1993). Promoting student retention through classroom practice.

Visualising and quantifying students' academic disappointment in science courses

Anitta Thomas[1] and André E Botha[2]

[1]School of Computing, P.O. Box 392, University of South Africa, Pretoria, 0003

[2]Department of Physics, P.O. Box 392, University of South Africa, Pretoria, 0003

E-mail: thomaa@unisa.ac.za; bothaae@unisa.ac.za

Abstract. Based on existing hypotheses that correlate student satisfaction and final grades, a working hypothesis was made for Science courses: student academic disappointment in a course is dependent on the similarity of their performances in various grade components. Based on this hypothesis, a new way of quantifying student disappointment was presented along with a unique visualisation to facilitate interpretation of student performances and disappointments in a course. The visualisation was used to explore some of the consequences of the hypothesis using the assessment data gathered at the University of Victoria, Canada. The visualisation was further applied as a tool to assist in two examples of academic activities: evaluating teaching effectiveness and increasing pass rates.

1. Introduction

In recent years student satisfaction surveys have become increasingly popular in universities [1], because the satisfaction of students is considered one of the key indicators of the performance of a university [2]. The view of higher education as a service where the satisfaction of students is an indicator of the quality of the service provided, may have helped to solve some of the challenges universities faced; such as, high student dropout rates, decreasing student enrolment, and the allocation of public funds based on the successful completion of programmes by students [3]. However, many of these challenges remain unsolved. Hence, it is important to understand the factors that contribute to student satisfaction, or equivalently to student disappointment, and to develop the ability to assess student satisfaction accurately.

Considering all the contributing factors, student satisfaction in general is a complex concept to measure and understand [4]. Usually it is measured, on a scale from one to five, through student satisfaction questionnaires that include questions about a wide variety of aspects which are thought to contribute to student satisfaction. In principle, a student's whole educational experience may contribute to his or her overall satisfaction with a university. However, in this paper, we make the distinction between satisfaction that is dependent on either academic or non-academic aspects of university life. We then focus on understanding aspects that contribute to a student's academic satisfaction with a course.

In the past a large number of studies, aimed at understanding the various aspects that contribute to academic satisfaction, focussed on the relationship between the final grades (expected or obtained) and the student satisfaction in the course [5-10]. This debate was already active since the early 1970s without producing consensus between the researchers regarding the reported results [10]. Due to the

disparity in the results some studies explored additional aspects that could influence and explain the correlation between final grades and student satisfaction; such as, student motivation, teaching effectiveness, classroom size, and the type of courses taken by the students [7, 9].

Studies focusing on grades as a contributing factor to student satisfaction have produced disagreeing results: some show positive correlation between grades and satisfaction [5-10], while others show insignificant or no correlation [11-12]. It is worth noting that in order to ensure anonymity, most of the aforementioned studies compared student satisfaction to the anticipated final grades, and not to the actual grades. In the study reported Holmes [6], students were asked to complete the course survey just after their final grades were made available. The final grades were distorted in such a way that half the students obtained one grade less that their actual grade. Immediately after collecting the course surveys students were informed of their correct grades. By this method the study confirmed the idea that high grades need not necessarily result in high student satisfaction, but rather that it is the difference between the actual grade and the expected grade that is the most important contributing factor.

Guided by the idea that satisfied students are the ones whose actual grades match their expected grades, and by induction from our own teaching experience in the Sciences, we make the following working hypothesis: Students will be least disappointed in a course when they perform equally well in all grade components making up the final mark. Conversely, students will be most disappointed in courses where they perform very differently in two or more grade components. Our hypothesis may be regarded as a natural extension of the postulate made in [6], specifically adapted to the Sciences. Note that, other than grade data, it does not take into account other factors that may also contribute to student satisfaction, such as those which were mentioned in [7,9]. In a typical Science course, for example, the various 'grade components' would include the grades obtained for (i) the practical or laboratory work, (ii) assignments, (iii) a midterm examination, and (iv) the final examination. In most universities the weighted sum of grade components (i) to (iii) contribute roughly the same as (iv).

As a first step towards testing the above hypothesis we have developed an approach to quantify student disappointment based on the differences in the marks obtained by a student in various grade components, taking into account the weight of each grade component in the final grade. We have then developed a new visualisation technique that can clearly display how students perform in the various grade components of a Science course. The visualisation is constructed in such a way as to give instructors a view of grade data which is consistent with the above hypothesis. Based on the calculated student disappointment, the visualisation is also able to indicate disappointment indices of the students.

While this study is based on the ideas from existing research on student satisfaction and grades, it differs in several ways. Firstly, unlike previous studies, the present study considers a new aspect of the important problem of understanding and accurately gauging student satisfaction. The new aspect is related to the observation that grade components in the Sciences are very different to those in the Humanities. Arguably, they are more reflective and measureable than in the Humanities. Secondly, since student satisfaction surveys are not a measure of student learning [4], the working hypothesis that is developed in the present study may be able to provide an indication of true student learning, since it is based on performance (grade) data, rather than opinion. Unlike grade data, the data gathered from student satisfaction surveys can at best provide a measure of perceived satisfaction, which is often not related to how much learning actually took place in a course. Thirdly, a new technique to quantify student disappointment is developed based on the working hypothesis. Lastly, the visualisation technique developed here is also novel. It has been designed specifically for the Sciences in order to be consistent with our working hypothesis.

The layout of this paper is as follows. In Section 2, the technique developed for quantifying student disappointment is presented. The technical aspects of the visualisation are summarised in Section 3. A discussion on how to interpret the data in the visualisation using an ideal distribution of various grade components is included in Section 4. In Section 5, we use the visualisation to analyse real assessment

data from the University of Victoria, Canada. Two of the potential applications of the visualisation are discussed in Section 6, followed by a conclusion in Section 7.

2. Definition of the student disappointment index

Consider a class of size n in which there are m grade components. Assume that the grade data is contained in an array in which the j^{th} grade component for the i^{th} student is stored as a value between zero and one hundred in the entry B_{ij}. The Student Disappointment Index D_i for the i^{th} student is then defined as

$$D_i = \frac{1}{10000}\sum_{j,k=1}^{m}\left(w_j B_{ik} S\left(B_{ik} - B_{ij}\right)\right) \tag{1}$$

where $0 < w_j < 1$ is the weight of the j^{th} grade component towards the final grade and S is a step function which returns the value of its argument when its argument is positive and is zero otherwise. Both j and k are summation indices which run over all the grade components. The factor of 10000 ensures that the index is normalised to unity. Note that D_i is not a measure of success or failure in an academic program, but rather a measure of dissatisfaction with a particular course.

The overall disappointment index for the whole class is defined as the average

$$D = \frac{1}{n}\sum_{i=1}^{n} D_i \tag{2}$$

To illustrate the meaning of the definition in (1), consider a simple case in which a certain course has only two components with w_1=0.01 and w_2=0.99. Suppose a student in this course obtains 96% for component 1 and only 3% for component 2. In this case there is only one non-zero term in the summation for the disappointment index, i.e. it is given by

$$D_i = \frac{1}{10000}\left(0.99 \times 96 \times (96 - 3)\right) = \frac{8838.7}{10000} \approx 0.884$$

Being close to unity, the calculated value indicates that this student is highly likely to be disappointed with the course. The value of D_i depends on the weighting of the grade components. In this case the student is dissatisfied, not because he obtained 3%, but because he obtained 3% for a grade component that was weighted 99%. Had the student obtained the same grades with the weights reversed, the calculated value of D_i would have been 0.009, i.e. highly likely to be satisfied.

3. Basics of the developed visualisation

In order to visualise the working hypothesis and the student disappointment indices, we have devised a new way of plotting grade component data, using the Python programming language [13]. The script which we have written can easily be modified by non-Python users and is currently compatible with data that is formatted in tab separated or comma separated format. Typically such data might be exported from learning management systems such as *Moodle* [14] or *Blackboard* [15].

After reading the data into the array B, the script calculates the individual and class disappointment index according to the definitions in (1) and (2). It then plots the various possible combinations of grade components against one another in a two-dimensional plot by using the transformation

$$x_i = t_i \cos\theta_i\,, \quad y_i = t_i \sin\theta_i\,, \quad \text{with } \theta_i = \frac{\pi}{4}\left(1 + \frac{B_{ij} - B_{ik}}{100}\right) \tag{3}$$

Here t_i is the final overall grade obtained by the i^{th} student and (x_i, y_i) are the Cartesian coordinates for the i^{th} student in the plot of grade component j against k. The Cartesian coordinates for the i^{th} student is plotted as a filled coloured circle, where the colour is determined using the calculated disappointment index of the student. The developed visualisation uses colours from blue through to orange, where blue indicates the lowest possible value of disappointment of 0 (highly satisfied) and orange indicates the highest possible disappointment value of 1 (least satisfied).

4. Fundamental interpretations of plots generated by the developed visualisation

Figure 1 shows an illustration of the visualisation discussed in Section 3 for ideal (fictitious) data. The radial distance from the bottom left corner to a filled circle equals the final grade of the student, while the angular separation away from the diagonal red line is the difference in grade components, as defined by θ_i in equation (3). Two visual cues are added in the grid lines: the red quarter circle indicates a final grade of 50% and the red diagonal line indicates zero difference between the marks obtained in the plotted grade components. One advantage of looking at the data in this way is that it tends to amplify the grade component differences of students with higher overall grades, while reducing the differences for students with lower overall grades. This feature of the plots is based on our frequent observation that high performing students are generally more concerned with even minor differences in their grade component marks.

A number of elementary interpretations are valid for all the plots generated by this visualisation. All plotted circles below the red curve indicate students who failed the module. Students who have done equally well in both the grade components will appear as a filled circle on the red diagonal line. A student who performed differently in the plotted grade components appear on either above or below the red diagonal line, depending on in which grade component they did better and the angular distance from the diagonal indicates the difference in the marks obtained in the grade components plotted.

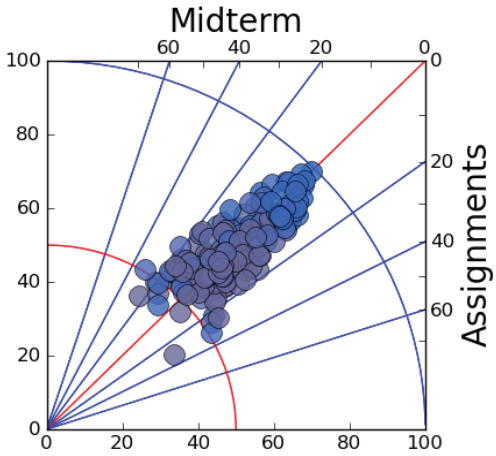


Figure 1: An ideal distribution of grades based on fictitious data for a class size of 157. Notice that the distribution is symmetric about the red diagonal line and, as indicated by the color of the circles, no student has a disappointment index higher than 0.5. Generally, in our experience, it is worth investigating why any student obtains a disappointment index above 0.5, since there are usually obvious reasons. The overall disappointment index for this class is 0.295.

The elementary interpretations described above can be applied to figure 1 to determine the number of students failed in the course and for comparing the general performance of students between *Midterm* and *Assignments before Midterm* grade components. It should also be noted that in figure 1, the distribution of grades occurs symmetrically about the red diagonal line, with a gradual spreading out towards the lower grades that are nearer the 50% overall mark. The symmetry of data is an ideal characteristic in such a plot because it demonstrates that on average students did equally well in both the grade components. Moreover in figure 1, the filled circles are colour coded in variations of blue and grey indicating low disappointment indices attributed to the low overall disappointment index of this class: D=0.295. The combination of symmetry about the diagonal and the low overall disappointment index makes figure 1 an ideal plot.

5. Analysing plots generated using real assessment data

Figure 2 illustrates visualisations of grade data obtained by students in various grade components of a second year Electricity and Magnetism course taught by one of the authors (AEB) at the University of Victoria in 2009. Unlike the ideal plot given in figure 1, plots in figure 2 show asymmetry of data along the diagonal red lines as well as some students with high disappointment indices, which are coloured orange.

From the plots in figure 2, it is evident that the students did not do equally well in various grade components: most students obtained higher grades for assignments than for the final exam and

midterm, performed better in the practical than in assignments, and performed better in the midterm than in the final exam. In this course the grade components were not assessed at an equal standard. For example, the final exam for this course was more difficult, than the assignments students did during the term.

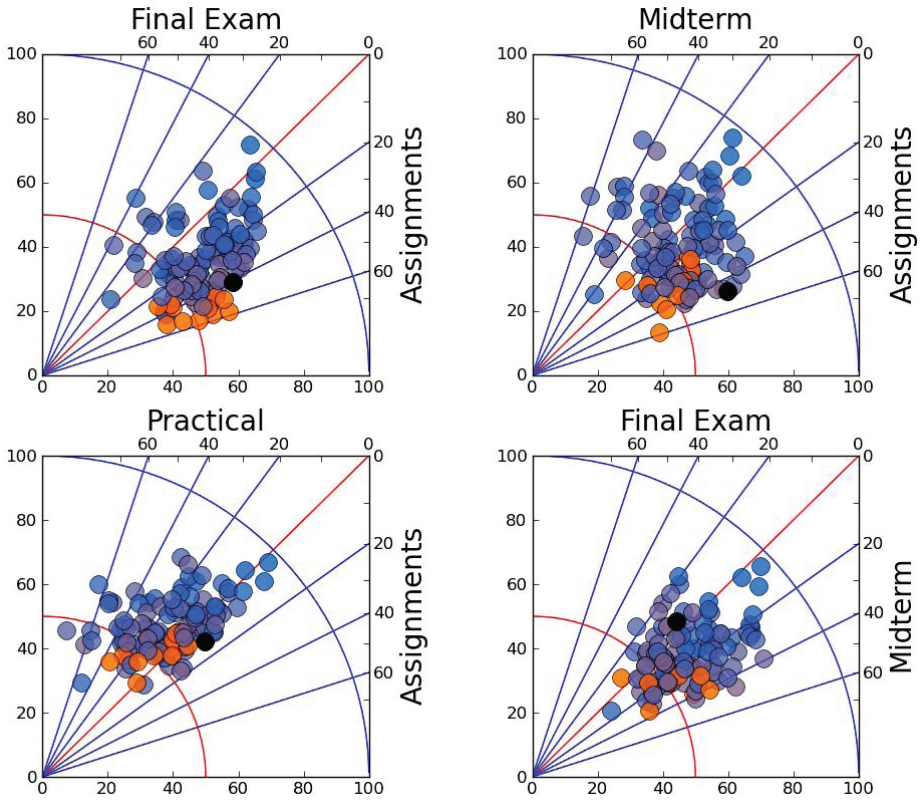


Figure 2: Four out of the six possible visualisations of grade data for 113 second year students who attended an Electricity and Magnetism course. The weights of the grade components were 0.10 for assignments, 0.20 for practical, 0.20 for midterm and 0.50 for the final exam. The overall disappointment index for this class is 0.375, which is higher than the ideal described in figure 1. Orange coloured circles flag students with disappointment indices greater than 0.5. One orange circle was changed to black purely for the purposes of the individual analysis given in the main text.

Due to the disparity in the marks in the various grade components the overall disappointment index is 0.375, and fifteen students have individual disappointment indices greater than 0.5. One of these students with high disappointment index (denoted with a black circle in figure 2) obtained 99, 97, 100, 98, 100, 86, 100 % for the seven assignments respectively, 87% for the practical but only received 50 and 56 % for the midterm and final exam respectively, giving him or her a final grade of 65%. This highlighted student did equally well in the assignments and in practical but failed to perform equally well in the midterm and final exam. After the course it was discovered through student evaluations that many of the students had had access to the assignment solutions (in the form of the instructor solution manual for the textbook) and that the practical component, which had been taught by a different instructor, had been too easy.

6. Potential applications

In addition to being able to view student grades and student disappointments, the visualisation can also be used as a tool to guide other academic activities in Science courses. As examples, two such

activities will be considered: improving the teaching effectiveness of various grade components and adjusting grades to increase the pass rates in a course. Even though increasing pass rates is not always academically justifiable, there are times when it may be necessary. For example, it may be reasonable to normalise the current pass rates with previous years' pass rates or to normalise pass rates of various classes of the same course taught by different instructors when there were differences in the assessment standards.

Based on the visualised performances of students one can more easily evaluate the effectiveness of various assessments in the respective grade components in a course. Consider a scenario where the visualisation of differences in the marks between final exam and assignments clearly demonstrates that students performed poorly in the final examination compared to the assignments. In such a scenario one can investigate whether this difference in performances is due to the difference in the levels of difficulty of these two grade components or if the assignments did not adequately prepare students for aspects that were tested in the exam. In this respect the visualisation has the potential to be used as a tool to reflect upon teaching effectiveness.

The visualisation can also be used to increase pass rates, when there are reasonable grounds for doing so, without increasing the overall disappointment index. The increase is achieved by reducing the differences in student performance within each grade component pair. To reduce the differences systematically, a best fit polynomial curve is constructed for each plot (see black curve in figure 3). Subsequently the lowest of the two grade component marks for each student is *increased* by a small fraction of the corresponding difference between the polynomial curve and the red diagonal line. After all students have been adjusted in this way, the new best fit polynomial for the adjusted data is calculated. The area between this new polynomial and the red diagonal line is now slightly smaller than before, i.e. some of the asymmetry in the data has been removed. At the same time the student disappointment indices have been reduced and, since each student's grade has been increased in the process, there is an increase in pass rate. This process can be repeated until either the desired pass rate is achieved or else until all the asymmetry in the grade component data has been removed, i.e. until all the areas are zero.The step-by-step procedure to increase the pass rate while reducing the overall disappointment index is given in Appendix A.

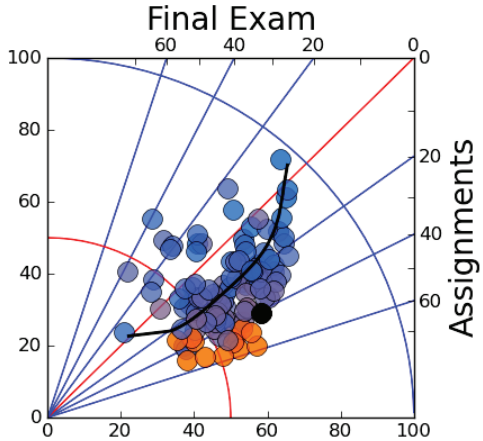


Figure 3: The *Final Exam* and *Assignments* plot in figure 2 is supplemented with a polynomial fit (shown by the black curve). The asymmetry in the data is now more apparent and moreover it is quantified in terms of the area between the polynomial fit and the red diagonal line.

7. Conclusion

To conclude, the working hypothesis made in this preliminary work has allowed us to quantify student academic disappointment and to give brief analyses of student performances in various grade components, through the use of the developed visualisation. Two examples have been provided to illustrate how the method may be used in typical academic activities within the Sciences. A study is currently underway to test the working hypothesis and the results will be published in a future article.

Acknowledgments
This work is based upon research supported by the National Research Foundation (NRF) of South Africa. Any opinion, findings and conclusions or recommendations expressed in this material are those of the authors and therefore the NRF do not accept any liability in regard thereto.

Appendix A. Step by step procedure to increase the pass rate of a class with *m* grade components
In order to describe the procedure we first make the following definitions:

Q = the number of different grade component (GC) pairs. Each pair represents one possible plot.
$I = [I_1, \ldots, I_Q]$ is a list in which each I_i contains the two grades for all students in the i^{th} GC pair.
R is the red diagonal lines in each of the *Q* plots.
F is the list of final total grades.
CP is the pass rate (number of students who passed/total number of students in the class)
DP is the desired pass rate

Step	Description of procedure
1.	Compute the list of polynomial curves $C = [C_1, \ldots, C_Q]$ to best fit the lists of data in *I*.
2.	Compute list of areas $A = [A_1, \ldots, A_Q]$, where A_i represents the area between C_i and *R*.
3.	Find A_h = first maximum area in *A*.
4.	If A_h is zero, end.
5.	Increase grades in I_h by a small fraction of corresponding difference between C_h and *R*.
6.	Compute pass rate *CP* using the modified data in I_h.
7.	If ($CP < DP$) and all areas are not zero then go to step 2, else end.

References

[1] Walker D J and Palmer E 2011 The relationship between student understanding, satisfaction and performance in an Australian engineering programme *Assessment and Evaluation in Higher Education* **36** 157-70
[2] Ansari W E and Moseley L 2011 You get what you measure: assessing and reporting student satisfaction with their health and social care educational experience *Nurse Education Today* **31** 173-8 and references therein.
[3] Elliott KM and Shin D 2002 Student satisfaction: an alternative approach to assessing this important concept *Journal of Higher Education Policy and Management* **24** 197-209
[4] Wiers-Jenssen J, Stensaker B and GrØgaard J B 2002 Student satisfaction: towards an empirical deconstruction of the concept *Quality in Higher Education* **8** 183-95
[5] Feldman K A 1976 Grades and college students' evaluations of their courses and teachers *Research in Higher Education* **4** 69-111
[6] Holmes D S 1972 Effects of grades and disconfirmed grade expectancies on students' evaluations of their instructor *Journal of Educational Psychology* **63** 130-3
[7] Howard G S and Maxwell S E 1980 Correlation between student satisfaction and grades: a case of mistaken causation? *Journal of Educational Psychology* **72** 810-20
[8] Pike G R 1991 The effects of background, coursework, and involvement on students' grades and satisfaction *Research in Higher Education* **32** 15-30 and references therein.
[9] Powell R W 1977 Grades, learning, and student evaluation of instruction *Research in Higher Education* **7** 193-205 and references therein.
[10] Scheurich V, Graham B and Drolette M 1983 Expected grades versus specific evaluations of the teacher as predictors of students' overall evaluation of the teacher *Resource in Higher Education* **19** 159-73 and references therein.

[11] Hildebrand M, Wilson R C and Dienst E R 1971 Evaluating university teaching *Report by the Berkeley Center for Research and Development in Higher Education* (University of California, Berkeley)
[12] Voeks V and French G M 1960 Are Student-ratings of teachers affected by grades?: Report of three studies at the university of Washington *Journal of Higher Education* **31** 330-4
[13] Python Programming Language – Official Website. Available from http://www.python.org/.
[14] Moodle. Available from http://moodle.org/.
[15] Blackboard. Available from http://www.blackboard.com/.

Division F – Applied and Industrial Physics

Raman Characterization of the Phase Transition of Ion Implanted Hexagonal Boron Nitride to cubic Boron Nitride Nanoparticles

E. Aradi, S. R. Naidoo, R. M. Erasmus, T. E. Derry

School of physics, University of the Witwatersrand and DST/NRF Centre of Excellence, Private Bag 3, Wits 2050, Johannesburg, South Africa

E-mail: 395719@students.wits.ac.za

Abstract. We investigated the phase transition of hexagonal boron nitride (h-BN) to the c-BN form by the ion implantation process. Raman Spectroscopy was the major technique used in the characterization of the possible h-BN - c-BN phase transformation. The h-BN samples were implanted with various ions including helium, lithium and boron, at different ion fluences ranging from 1×10^{16} ions/cm^2 to 1×10^{16} ions/cm^2. The effect of varying the ion implantation energies ranging from 40 keV to 160keV were also investigated on samples implanted by helium ions. Finally lithium implanted samples was studied using the 488 nm and 514 nm Raman laser lines. Micro Raman spectroscopy (μ-RS) measurements indicated that indeed ion implantation induced a phase change to the c-BN, evident from the longitudinal optical (LO) peaks at wavenumbers 1300 cm^{-1} and 1304 cm^{-1} appearing in the spectra after implantation. The nature and extent of these new peaks, and their dependence on the different implantation parameters is discussed using the spatial correlation model.

1. Introduction

Due to its structure and bonding, c-BN exhibits excellent properties which make it very important in various applications [1, 2, 3, 4, 5]. It exhibits hardness values of 70 GPa, second only to diamond's 100 GPa [6, 7], as such it is an important material for abrasive processes and as a sintered ceramic for sawing, cutting, drilling or crushing applications. It has high thermodynamic stability at high pressures, with high melting point (3000 K) and thermal conductivity (13 Wcm^{-1}K^{-1}), second to diamond (4273 K) and 20 Wcm^{-1}K^{-1}) respectively. It also has low dielectric constants making it useful in high temperature high frequency microelectronic device fabrication [7, 8]. It is the lightest known group III-V compound with the widest band gap of E_g= 6.0-6.5 eV of all semiconductors [9, 10]. Its high refractive index (n $_{c-BN}$)=2.17) and a high transparency in a wide region of the electromagnetic spectrum enables it to be used in the fabrication of opto-electronics and electron emitting devices [11, 12].

It also surpasses diamond with respect to oxygen stability and chemical inertness at elevated temperatures. Cubic BN's oxidation (1400 K) and phase change temperatures to h-BN (1700 K) are much higher than the oxidation temperature (900 K) and phase change temperature to graphite (1100 K) of diamond [13]. Cubic BN also shows high chemical inertness with molten ferrous materials at high temperatures of up to 1700-1800 K, whereas diamond readily forms iron carbide (Fe_3C_2) at these temperatures. This property makes c-BN very important

in coatings and tribological applications of ferrous materials [14, 15, 16]. It can also be doped as both the p-type and n-type semiconductor making it superior to diamond whose shallow n-type semiconductor is still problematic [8]. This makes it useful in high speed electronic applications [17, 18], for example in the creation of p-n junction diode having a portion of its emission spectrum in the UV region [19, 20].

The excellent properties exhibited by *c*-BN leading to its enormous potential application have motivated researchers towards finding ways of synthesizing it. This research has been ongoing for the past half a century after its discovery in the late 1950's. To date, *c*-BN has been commercially synthesized in large quantities as powder by high pressure, high temperature method (HPHT) by Wentorf Jr [21].

The major setback in the *c*-BN synthesis has been the ability to produce large *c*-BN grains. As such it has been produced as thin films by Ion Beam Assisted Depositions IBAD techniques such as Chemical Vapour Deposition (CVD) and Physical Vapour Deposition (PVD). Current research is being focused development of thin *c*-BN films by role of defects, with *h*-BN as a starting material. Extensive research has been done theoretically [22] and experimentally [23, 24, 25] and concluded that defects induced in *h*-BN led to a phase transition to the *c*-BN phase. The work presented herein focused on optimizing the ion implantation conditions such as ion mass, implantation energies, ion fluence and temperature for *h*-BN to *c*-BN phase change induced by ion implantation.

2. Experiment

2.1. Sample and sample preparation

Hot pressed *h*-BN samples were used as the starting material throughout this work. The samples were supplied by Goodfellow UK Company limited as a 50 mm long and 15 mm in diameter rod. The rod was cut into 2 mm thick slices using a diamond wire saw with a wire thickness of 0.2 mm. The cut samples were then polished by a 1200 Å sand paper, ready for implantation.

2.2. Ion Implantation

All implantations were carried out at the iThemba Labs Gauteng, using the 200-20A2F ion implanter. The implanter consists of the ion source, where ions are produce, the analyzer magnet where ions with the specific m/q ratio are allowed to pass through for implantation, the accelerator which is able to accelerate ions at energies raging from 25 keV to 200 keV and the end station where the sample is located and where the implantation takes place. The implantation experiments were done at room temperature.

The first set of samples were implanted by varying the ion mass using He^+, Li^+ and B^+ at an energy of 150 keV, and fluences ranging from 1×10^{14} ions/cm^2 to 1×10^{16} ions/cm^2. The second set were implanted with helium ions at the fluence of 5×10^{15} ions/cm^2, varying the implantation energy from 40 keV to 160 keV.

2.3. Raman Spectroscopy (RS)

All Raman measurements were performed at the Raman and Luminescence laboratory at the University of the Witwatersrand. The system consists of an argon ion-laser, a Jobin- Yvon T64000 Raman triple grating spectrograph, and an optical microscope with a movable stage. A camera attached to the microscope and the micrometer movement of the stage enables the study of the specimen at various locations of the samples surface.

Characterization of the samples was done before and after implantation in order to determine any structural changes. Measurements on the first two sets of samples were carried out at room temperature. Samples implanted with lithium ion were analyzed using the Ar^+ and Kr^+ laser lines. Cubic boron nitride powder samples were also analyzed by RS.

2.4. unimplanted samples

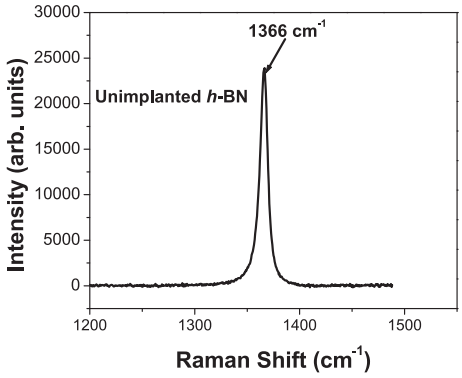


Figure 1. Raman spectra for unimplanted *h*-BN sample.

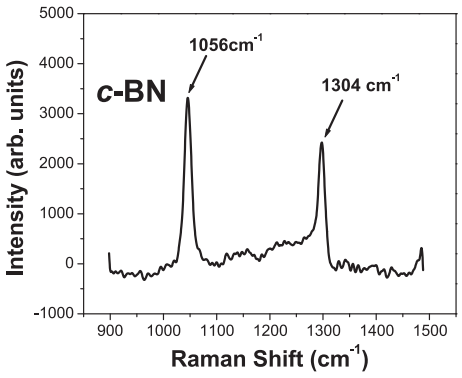


Figure 2. Raman spectra for *c*-BN nanopowder.

Due to the difference in the bonding nature of *h*-BN and *c*-BN, they exhibit different Raman vibrational phonon mode signal. Figure 1 represents the spectrum for *h*-BN sample before implantation. This spectrum indicated a longitudinal optics (LO) $2E_{2g}$ phonon mode peak for *h*-BN at wavenumber 1366cm^{-1} and with FWHM of 8.9 cm^{-1}, also reported by [26]. Figure 2 is a typical Raman signal for *c*-BN nanopowder sample. Two phonon peaks were observed at 1056cm^{-1} and 1304cm^{-1} representing the characteristic transverse optical (TO) and longitudinal optical (LO) modes respectively for the *c*-BN nanocrystal, also reported by [3]. Our observations were focused on the behaviour of the spectrum in figure 1 in comparison to features shown in figure 2.

2.5. Ion Mass Dependence

The ion mass dependence was achieved by implanting the *h*-BN sample at 150 keV, using ions with different atomic masses, i.e. helium at 5×10^{15} ions/cm^2, lithium at 1×10^{15} ions/cm^2, and boron 5×10^{14} ions/cm^2. These doses represented the optimum dose obtained for each ion. Figure 3 shows the spectrum after implantation with the three different ions.

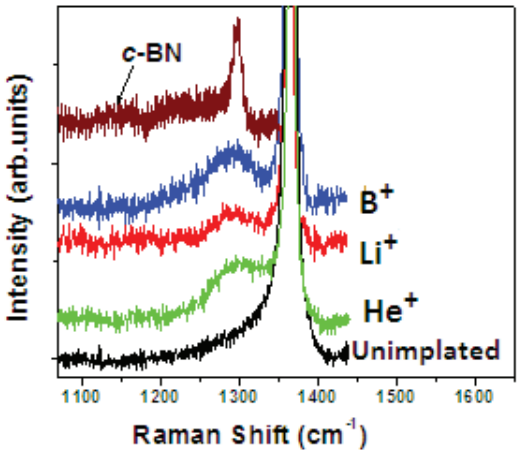


Figure 3. Raman spectra of the implanted sample at higher resolution in comparison to the spectra before implantation.

From figure 3, it is evident that there is a new broad peak appearing after implantation. The wavenumbers vary from 1300 cm^{-1} for helium to 1301 cm^{-1} and 1303 cm^{-1} for lithium and boron respectively. The peaks had FWHM of 60 cm^{-1}, 52 cm^{-1} and 45 cm^{-1} respectively. These peaks are synonymous to the Longitudinal Optical (LO) phonon feature in figure 2 that occur at 1304 cm^{-1}, predicting a possible ion induce phase transition to *c*-BN nanocrystals. The phonon features broadened, became asymmetric and shifted to lower frequencies with decreasing ion mass for the three ions. This phenomenon has been explained in line with Spatial Correlation Model developed by Parayanthal and Pallak [27], where it was proposed that a decrease in the crystal size causes a shift in the wavenumber to lower frequencies as well as a broadening and increase in the asymmetry of a given Raman peak.

The intensities and FWHM of these peaks were dependent on the ion mass. This is evident from the spectrum of samples implanted by B^+ which gave a higher intensity signal with narrower peaks followed by lithium while helium had the lowest intensity at the same conditions. Smaller values of FWHM indicates higher crystallinity of the a given material [28, 29] as such, since boron ions had a higher intensity and small value of FWHM it presents itself as the most suitable candidate in the possible *h*-BN- *c*-BN phase transition. Due to the possible reduced crystalline quality and size, no signal was observed at the for the TO mode of *c*-BN which is normally associated with larger (micron-size) *c*-BN crystals.

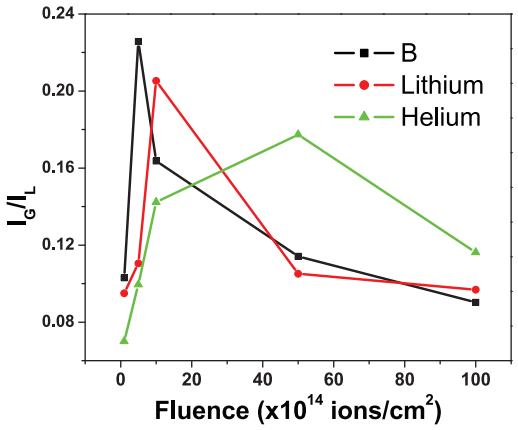


Figure 4. Relationship between normalized Gaussian peak intensity as a function of the ion fluence for boron lithium and helium.

Figure 4 is a graph of the Gaussian peak intensities for the LO broad peak normalized to the principal *h*-BN Lorentzian peak (I_G/I_L) as a function of the ion fluence. The ratio I_G/I_L roughly estimates the amount of *c*-BN present in the sample [8]. There was a clear optimum fluence for each ion which gave the highest I_G/I_L ratio. A greater ratio was observed for boron ions implanted at a at lower fluence of 5×10^{14} ions/cm^2. The other two ions had highest intensity ratios at higher fluence i.e. lithium at 1×10^{15} ions/cm^2 and helium at 5×10^{15} ions/cm^2.

2.6. Implantation Energy Dependence

Energy was varied for samples implanted with helium ions at the optimum dose of 5×10^{15} ions/cm^2. Figure 5 shows the different spectra as obtained using energies from 40 keV to 160 keV. It should be noted that simulations for the damage profile and predictions for the implantation density for these results were carried out using SRIM (Stopping Range of Ions in Matter) program, not presented in this work.

From figure 5, it is observed that at 40 keV, the *c*-BN Raman signal is almost completely faded, and hence the spectrum looks more or less as that of the unimplanted sample. As the energy increases to 80 keV, a weak signal is observed which slightly increases in intensity at 120

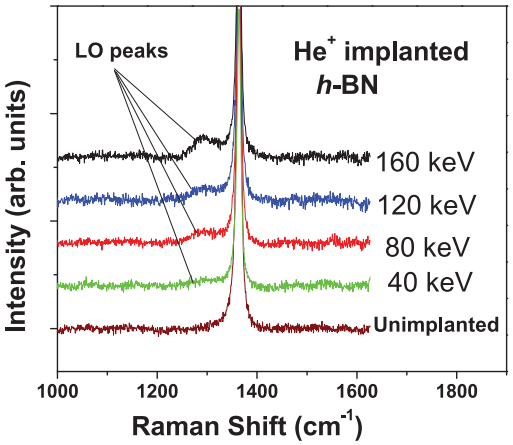


Figure 5. Raman spectrum of samples implanted by helium with various energies.

keV and finally it is more prominent at 160 keV. This energy dependence suggested that a critical amount of damage must accumulate in the implanted layer before the *c*-BN transformation is prominent as the He ions penetrate the *h*-BN samples.

2.7. Raman Laser Dependence

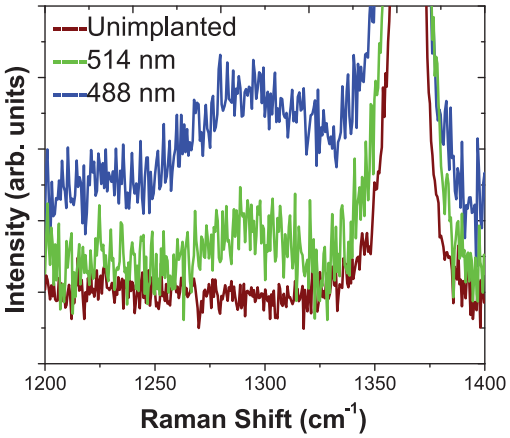


Figure 6. Raman spectrum of samples implanted by lithium analyzed by different laser powers.

The samples that were implanted with lithium ions at the optimum fluence of 1×10^{15} ions/cm^2 were analyzed by two different Raman lines i.e. the 514 nm Ar^+ and the 488 nm Kr^+ line to determine which of the two laser lines was the most effective in *c*-BN analysis, and results shown in figure 6.

The blue Kr^+ 488 nm line gave a stronger LO peak signal intensity with a sufficient signal-to-noise ratio compared to the green 514 Ar^+ nm line . The 488 nm Kr^+ line is well known for the detection of nanosize particles, this also gave a possible explanation for the presence of nanoparticle in the sample.

3. Conclusions

From the results presented herein, it is proposed that ion implantation induced a phase change of *h*-BN to *c*-BN nanocrystals. This is evident from the LO phonon peaks observed in samples

after implantation. The effect of size reduction explains the broadening and the frequency shift from the known c-BN LO mode that occurs at 1306 cm^{-1}. This effect is caused by a possible high density stress caused by defects induced during implantation. Boron ion implantation gave the best Raman signals with its peaks being narrower and symmetrical as compared to lithium and helium. Boron also had the highest peak intensity at much lower fluence followed by lithium and finally helium. The experiments show that there is a high dependence of the possible c-BN nanocrystals production with the implantation energy for He ions. Lighter ions required high energy for the transformation as compared to heavier ones. Finally our experiments show that the laser line in the blue spectral region are most effective in the investigation of c-BN thin films. Future work will involve the use of other characterization techniques such as Transmission Electron Microscopy (TEM), Surface Brillouin Scattering (SBS) and X-ray Diffraction to complement Raman spectroscopy ascertaining this possible phase change.

4. Acknowledgements

Financial support form DST/NRF Centre of excellence is highly appreciated. Many thanks to iThemba Labs Gauteng for ion implantation experiments and to Ronald Machaka,Isaac Motochi, Emmanuel Nshigabigwi and Anna Kozakiewicz.

References

[1] Mårlid B, Larsson K, and Carlsson J. O 1999 *Phys. Rev. B* **60** 16055.
[2] Otfired Madelung 2004 *Springer- Verlag, Berlin Hedenburg* **3** 755.
[3] Zhang W. J, Chong Y. M, Bello I, Lee S. T, 2007 *J. Phys. D Appl. Phy.* **40** 6159.
[4] Mirkarimi P. B, McCarthy K. F, and Medlin D. L 1997 *J. Mater. Sci. Eng* **21** 47.
[5] Erasmus R. M, Comins J. D and Fish M. L 2000 Diam. Relat. Mater. **9** 600.
[6] Ooi N, Rairkar A, Lindsey L, and Adams J. B 2006 *Phys.: Cond. Matter Phys.* **18** 97.
[7] Karlsson J, and Larsson K 2010 *J. Phys. Chem.* **2** 3516.
[8] Reich S, and Ferrari A. C, Arenal R, Loiseau A, Bello I, and Robertson J 2005 *Phys. Rev. B.* **71** 205201.
[9] Guo J, Wang H, Zheng K, Yan H, Yoshimura M 2007 *Electrochemistry communication* **9** 1824.
[10] Samantaray C. B, and Singh R. N 2005 *Int. Mater. Rev.* **50** 313.
[11] Sekkal W, Bouhafs B, Aourag H, and Certier M 1998 *J. Phys. Condens. Matter* **10** 4975.
[12] Oshima C and Nagashima A 1997 *J. Phys. Condens. Matter* **9** 9.
[13] Uchida H, Yamashita M, Hanaki S, and Kurihara C 2008 *Mater. Sci. Engi.: A* **483** 695.
[14] Zhang J, Cui Q, Li X, He Z, Li W, Ma Y, Guan Q, Gao W, and Zou W 2004 Phys. Rev. B, **399** 451.
[15] Nain Z. Y, and Tian Z. G, and Wen Z. P 1999 *Chin. Rev. Lett* **16** 155.
[16] Ronning C, Feldermann H, and Hofsass H 2000 *Diam. Relat. Mater.* **9** 1769.
[17] Yan Y, Zhang S. B, and Al-Jassim M. M 2002 *Phys. Rev. B.* **66** 201401.
[18] Albe Karsten 1997 *Phys. Rev. B.* **55** 6203.
[19] Wang C. X, Yang Y. H and Yang G. W 2004 *J. Phys. Chem. B* . **108** 728.
[20] Solozhenko V. L, Kurakevych O. O, Andrault D, Le Godec Y, and Mezouar M 2009 *Phys. Rev. Lett.* **102** 015506.
[21] Wentorf R. H. J 1957 *J. Chem Phys.* **26** 957.
[22] Mosuang T. E, and Lowther T 2002 *Phys. Rev. B.* **63** 014112.
[23] Hu C, Wu Q. M. J, Shen J, Kotake S, and Y. Suzuki 2002 *Thin Solid Films* **402** 117.
[24] Machaka R, Erasmus R. M, and Derry T. E 2010 *Diam. Relat. Mater* **19** 11331.
[25] Aradi E, Erasmus R. M, Derry T. E 2011 *In Press.*
[26] Nemanich R. J, Solin S. A, Martin R. M 1981 *Phys. Rev. B.* **12** 6348.
[27] Parayanthal P and Pollak F.H 1984. *Phys. Rev. Lett.* **52** 1822.
[28] Kubota Y, Watanabe k, Tsuda O, Taniguchi T 2007 *Sci* **317** 932.
[29] Zhang W. J, and Matsumoto S 2001 *Phys. Rev. Lett.* **63** 073201.

Application of Electroluminescence and Thermal Imaging in Defect Identification in Photovoltaic Modules

J L Crozier*, E E van Dyk and F J Vorster
Nelson Mandela Metropolitan University, Port Elizabeth, South Africa
*Corresponding Author e-mail address: s207094248@live.nmmu.ac.za

Abstract.
In order for solar energy to become an alternative to traditional fossil fuel energy it is important that defects in photovoltaic (PV) modules can be easily identified. Typically a module is characterised by measuring the current-voltage (I-V) characteristics under standard test conditions. In addition to this, electroluminescence (EL) can be used to indentify defects in the module. A cooled Si CCD camera is used to detect the EL signal. Infrared imaging can be used to identify irregular heating patterns which are indicative of defective cells or contacts. In this study both techniques were used in conjunction to assess the defects present in an Edge-defined film-fed (EFG) silicon module and a single crystalline silicon module. Both modules had visible signs of degradation. These techniques identified several defects in the modules. Cells which have cracks due to mechanical damage were quickly identified. The cracked areas appeared dark in the EL image due to the lack of electrical contact to these areas. The damaged cells are visible as a hot spot in the infrared image indicating that these cells are low current producing cells.

1. Introduction

Solar energy is an alternative energy source that has great potential to solve the energy problems in South Africa. However, defects present in photovoltaic (PV) modules can greatly reduce the performance of the module. Characterisation techniques such as current-voltage curves provide information about the short circuit current (Isc), open circuit voltage (Voc) and maximum power (Pmax). However, one is unable to identify the cause of reduced performance from I-V curves. Electroluminescence (EL) is a non-destructive characterisation technique that can evaluate the extent of the degradation while also identifying damaged areas which are not optically visible.

When a silicon solar cell is forward biased carriers are generated in the junction, when these carriers recombine energy is emitted in the form of electroluminescence. The intensity of the EL is proportional to the recombination rate and the minority carrier lifetime and diffusion length. [1] The minority carrier lifetime is the measure of how long a carrier exists before recombination occurs. The minority carrier diffusion length is a measure of how far the carrier can move before it recombines. [2] Both depend on the type and magnitude of recombine processes and relate to the collection efficiency of the cell material. Defects in the cell material that result in a decrease in the recombination rate or collection efficiency can be identified in the EL image. The EL signal intensity is detected using a cooled CCD camera which provides a greyscale spatial representation of defects in the cell. EL imaging is used during the manufacturing process to identify damaged cells before they are laminated into modules and is also used for quality control of modules.

Thermal imaging is used in conjunction to differentiate between intrinsic material defects such as grain boundaries and extrinsic defects such as manufacturing faults, broken contact fingers or cracks.

Variations in the module temperature (hot-spots) indicate poorly performing cells. A hot-spot occurs when that cell is performing worse than the other cells in the series string. It becomes reverse biased and starts to dissipate power and heat up.

2. Experimental

In this study two modules with different visible degradation were examined. The Edge-defined film-fed (EFG) module has 36 cells connected in series. The module is configured with two strings of 18 cells. There is visible delamination around the edges of the module. Previous studies on this module have shown that the encapsulant delaminated and moisture ingress has occurred. [3] The Single crystalline silicon module consists of 44 cells connected in series. The module is configured into two strings of 18 cells. The module is visibly damaged over the majority of the cells. These white speckled variations in the cells may be attributed to degradation in the cell material or the anti-reflective coating.

Electroluminescence was used to non-destructively evaluate the extent of the degradation in these two modules. The modules under investigation were forward biased to a current greater than the short circuit current of the module. The emitted luminescence of a silicon sample has a peak at about 1150 nm and a portion of this peak can be detected by a Silicon CCD camera which has a range of 300-1000 nm. [1] The CCD camera is cooled to -50 °C and has a resolution of 3300 x 2500 pixels. Cooling the CCD chip prevents noise in the signal due to dark currents. The data acquisition time is in the range of 1.5 to 2.5 seconds. The experimental set-up is illustrated in figure 1. The DC power supply forward biases the module by 30 V and 5 A. In order for the electroluminescence to be detected the setup must be placed in a dark room as any other light source will affect the results. The CCD camera is connected to the computer where the Sensovation image processing software, which is supplied with the camera, is used to capture and analyse the images.

An Infrared (IR) camera was used in conjunction with EL images to detect hot spots in the module. The IR camera can measure temperature in a range from 0 to 280°C with approximately 2° uncertainty. The IR camera is a hand-held portable device that can be used for indoor and outdoor temperature measurements. The camera is held some distance away from the module and focussed to acquire a clear image of the temperature variations in the module.

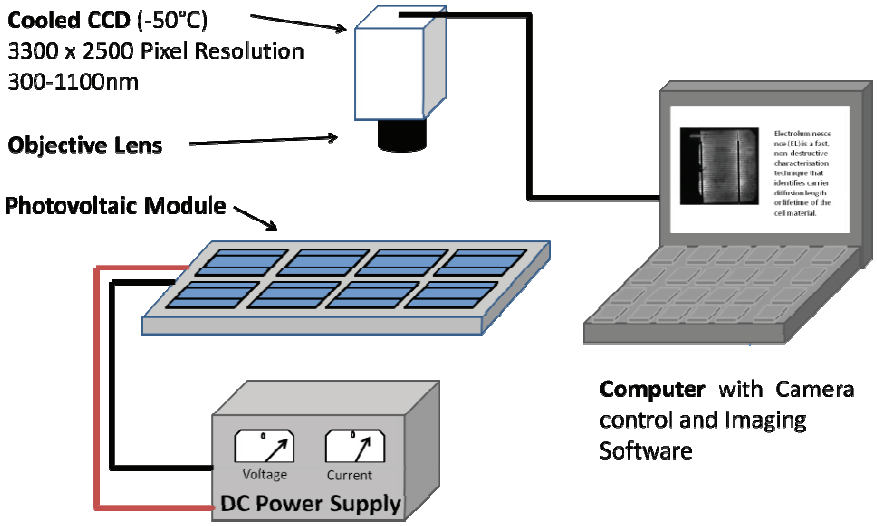


Figure 1. Schematic outline of EL experimental setup.

3. Results and Discussion

3.1. EFG Module.

The EFG module has delaminated allowing air and moisture to infiltrate between the encapsulant around the sides of the module. Figure 2(a) shows the optical image of the EFG module with cells C5 and C9 indicated. Figure 2(b) and 2(c) show the EL images of the C5 and C9 respectively.

Despite there being no optically visible damage to cell C5, the EL image shows that the cell has cracked. This crack decreases the electrical contact to a portion of the cell and thus lowers the intensity of the EL signal. The nature of the EFG growth process results in non-parallel variations in the cell material which are visible in the EL image. The eight back contacts that are present in each EFG cell are visible are areas of lower EL signal intensity. This is due to the high collection efficiency in these areas due to the back contact.

Cell C9 is severely affected by delamination which has spread along the busbar and affects about half of the cell. The EL image shows very low signal intensity over the areas of the cell affected by delamination meaning that the exposure to moisture has degraded the cell material and what little luminescence that is produced is blocked by the encapsulant discolouration.

Figure 3 shows the infrared image of the bottom-right corner of the EFG module under short circuit conditions. The ambient temperature of the room is below 20°C. A hot spot is visible over cell C9 identifying it as a damaged cell. This suggests that cell C9 is a poor current producing cell since it is being reverse biased and dissipating heat. This supports what is seen in the optical and EL images. The cells that are good current producing cells are operating at lower temperatures in the region of 22-24°C.

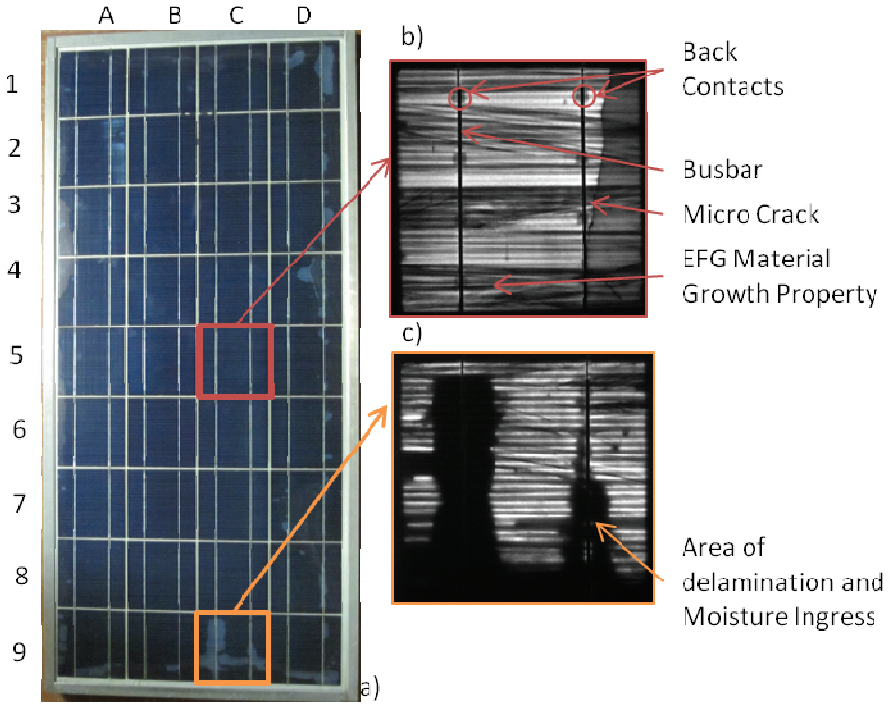


Figure 2. (a) Optical image of EFG module with cells indicated. (b) EL image of cell C5 with defects highlighted, (c) EL image of cell C9 showing the effect of delamination.

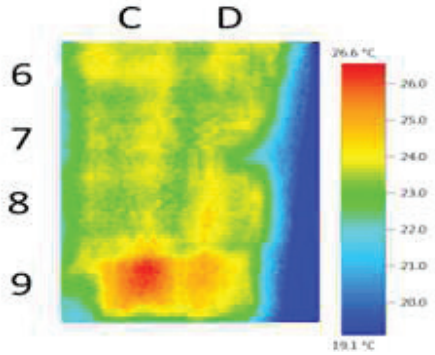


Figure 3. Infrared image of a portion of the EFG module.

3.2. Single Crystalline Module

Figure 4(a) shows an optical of the single crystalline module with the damaged cells highlighted. The infrared image of the single crystalline silicon module under illumination and in short circuit connection is shown in figure 4(b). The EL image of the damaged cells is shown in figure 4(c). The hot-spot indicates poorly performing cells which are reverse biased and heat up. The cells identified as poorly performing cells are A7, B7, B8 and C8. The El images of these cells indicate cracks resulting in large inactive areas. Figure 5(a) shows an optical image of these cells with defects highlighted. Figure 5(b) shows scratches on the back of the module. Figure 5(c) shows a magnified

image of damage on cell C8. Scratches on the back of the module correspond with features observed in the EL image. These cracks most likely occurred due to some mechanical damage. Scratches on the back of the module suggest that the module might have been dropped or hit with some force. The cracks remove portions of the cell from electrical contact and thus the carriers are not injected into this region and thus no luminescence occurs in these areas. These four damaged cells contribute greatly to the decreased performance of the module from a specified power output of 65W to a measure 53W.

The silver grey discolouration present in this module is most likely due to degradation in the silver nitride anti-reflective coating. The areas of discolouration result in an EL signal that varies across the cell. In a undamaged cell the EL signal should be uniform across the cell. Degradation in the anti-reflective coating may not affect the performance of the cell material in that area but rather block the luminescence so that it is not detected from that area. The areas affected by the reflective coating degradation thus affect the photoresponse of the cell material.

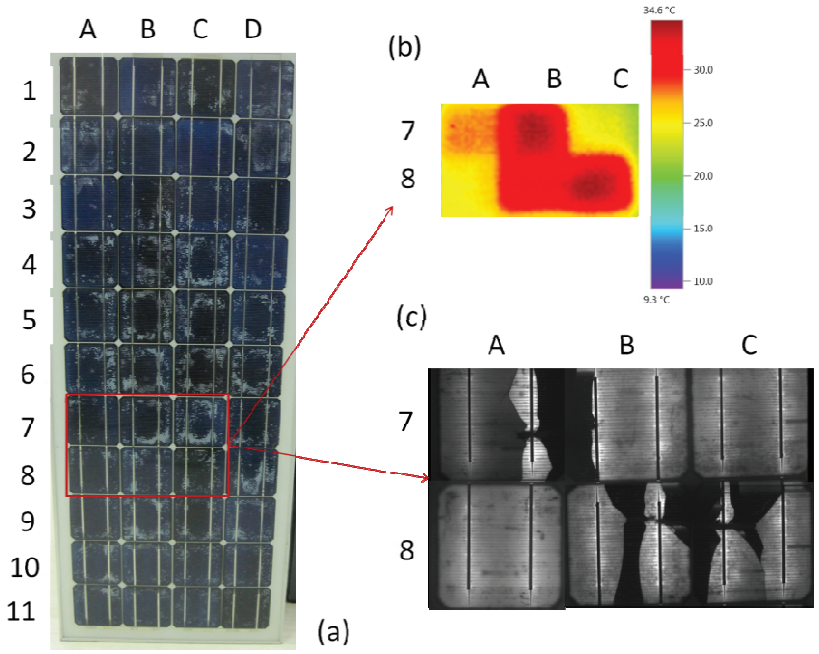


Figure 4. (a) Optical image of the single crystalline silicon module with damaged cells highlighted. (b) Thermal image of portion of module under illuminated, short-circuited conditions. (c) EL image of damaged portion of module

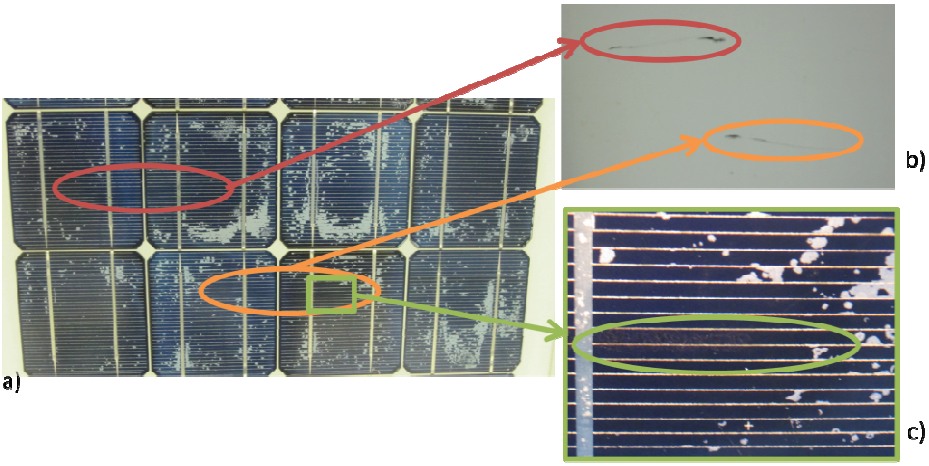


Figure 5. a)Optical image of damaged cells with defects indicated. b) Optical image of back of modules with scratches, c) close-up image of damaged area of cell with visible damage to cell material

4. Conclusions

Electroluminescence and thermal imaging are informative, non-destructive characterisation techniques that allow defects in assembled modules to be detected and quantified. When used together they provide a better idea of the defects present. Extrinsic defects such as cracks and broken fingers are difficult to identify from visual inspection but can be quickly identified in EL images and by hotspots in thermal images. Degradation of the anti-reflective coating in a PV module results is a decrease in the module performance and the effects can be seen in the lower EL signal in these regions. The effects of delamination are visible in the EL image of a module indicating that there is a drop in photoresponse of the affected cell material.

References

[1] Fuyuki T, Kondo H, Yamazaki T, Takahashi Y and Uraoka Y 2005 *Appl. Phys. Lett.* **86** 262108
[2] Green M 1982 *Solar Cells* (Englewood Cliffs: Prentice-Hall Inc)
[3] Van Dyk E E, Chamel J B and Gxasheka A 2005 *Sol. Energ. Mat. Sol C.* **88** 403-411

Comparative efficiency study of a solar trough receiver: hot mirror and selective coating

M C Cyulinyana and P Ferrer

University of the Witwatersrand, School of Physics, South Africa

E-mail: Philippe.Ferrer@wits.ac.za

Abstract. We consider a solar trough system in which cylindrical parabolic mirrors focus the sun's radiation onto a receiver pipe, heating it. Conventionally, the pipe is enclosed in a glass cover under vacuum and the dominant radiation losses are reduced by the use of a selective coating on the receiver pipe. We study the suitability of applying a 'hot mirror' coating on the inside of glass cover instead, which transmits in the visible but reflects well in the infrared. We compare the performance of the 'hot mirror' coating using its optical properties (reflectivity, absorptance and emissivity). It is seen that a hot mirror is a viable alternative, and certainly allows higher temperatures of the working fluid and therefore a higher Carnot efficiency.

1. Introduction

The solar trough is among the most-studied concentrated solar thermal power systems. The largest solar trough power plant in the world, the SEGS (Solar Energy Generating Systems) in California with a peak capacity of 364 MW, has been operational and studied in detail since the mid-eighties [1]. Such systems consist of long, cylindrical trough-shaped parabolic mirrors which concentrate the sun's radiation on the focal line where the receiver unit (collector tube) runs and through which a 'working fluid' circulates and is heated [1, 2] (figure 1). The working fluid is then circulated to a thermal power station where it is converted to electricity. Generally, the receiver unit must be designed in such way that it loses as little heat as possible via radiation, convection and conduction to its surroundings [1]. The receiver consists typically of a glass cover, encapsulating a metal receiver pipe with a vacuum in between. The vacuum minimizes convective losses from the heated receiver pipe to the surroundings [1, 3]. Conduction losses are reduced by minimizing the contact between the receiver pipe and the glass sleeve [4]. Thermal radiation losses are reduced by the use of a selective coating applied to the receiver pipe [5, 6]. Research has been done on different aspects of the receiver unit, such as thermal and material properties [3, 7]. Our research concentrates on the possibility of a substitute for the selective coating and the improvements of selective coating absorptance, emissivity and long-term stability in air.

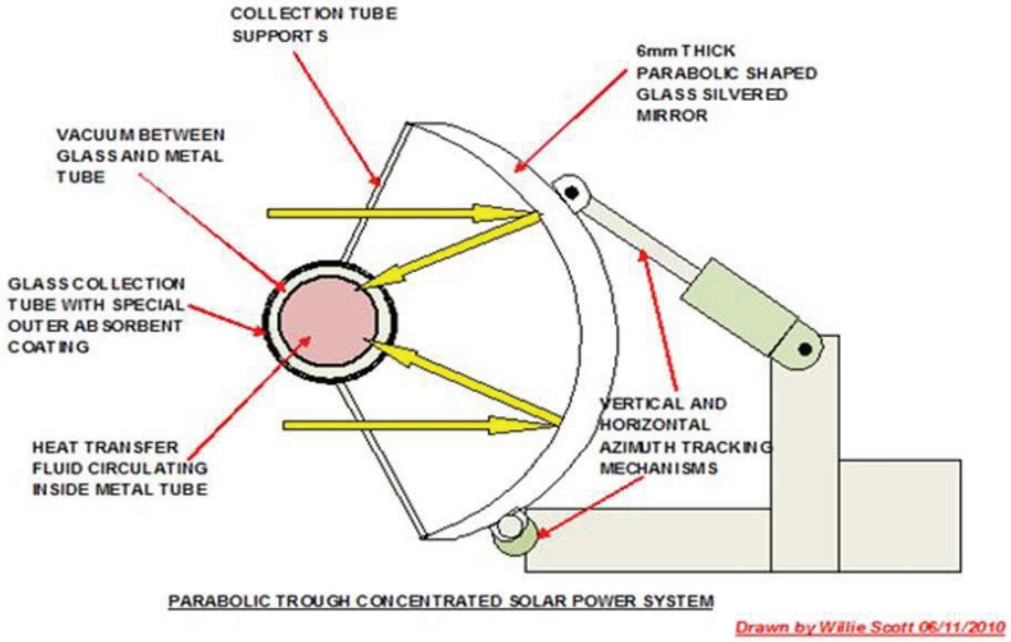


Figure 1: Solar Trough Plant layout: http://japanesestyletattoos.blogspot.com/2011/05/solar-power-plant-layout.html

2. Selective Coating

The selective coating applied on the receiver pipe (i.e. Black Chrome) is a dielectric material which absorbs well in the visible region of spectrum (i.e. sunlight received from the mirrors) but emits very poorly in the Infrared region [3], which is the thermal radiation that is lost from the receiver at its operating temperature (~ 400^{o}C). A selective coating on the receiver pipe will render the system more efficient in terms of the fraction of solar energy absorbed (i.e. α=0.94) by the working fluid.

Numerous works have been published on selective coatings on the receiver pipes and their properties [1, 2]. With this research into selective coatings, we want to create a stable (long-term stability in air), efficient coating over a wide temperature range (0°C to 1000°C) which has the high absorptance and emissivity in Infrared region.

The main weakness of commercial selective coatings is that they will deteriorate thermally at temperatures around 680°K (~400°C) [2, 5, 6, 9, 10], thereby restricting the maximum temperature to which the receiver pipe can be heated, and hence the Carnot or Rankine efficiency of the subsequent heat to electrical conversion [11]. The temperature of the receiver pipe rises with its length, hence the receiver pipe length is restricted to a maximum length, after which thermal breakdown of the selective coating occurs.

The basic idea in this paper is to investigate the theoretical performance of an alternative to selective coating, the “hot mirror”.

3. Hot mirror

A hot mirror (i.e. Indium Tin Oxide (ITO)) refers to a dielectric coating that is designed to reflect the IR region of the spectrum and to transmit the visible. In our model, it is applied to the inside of the glass cover. The hot mirror on the glass cover will reflect IR radiation emitted from the receiver pipe back onto itself, thereby reducing the amount of thermal radiation leaving the receiver unit [3]. The net effect is similar to that of the selective coating. The hot mirror coating also breaks down thermally around 680K [4]. The hot mirror is applied to the glass cover, which is cooler than the receiver pipe during operation (by around 200K, depending on coating). It is therefore possible to sustain higher temperatures in the receiver pipe using a hot mirror instead of a selective coating on the receiver pipe. This additional temperature increase is transferred to the working fluid, and subsequently to the steam power station, where it will improve thermal efficiency [5].

The questions posed in this paper are, first, how a hot mirror system compares to a selective coating system in terms of efficiency of heat transfer into the working fluid; and second, whether a hot mirror system can be used in a temperature region where the selective coating system breaks down (>680K) on the receiver pipe. In order for the plant to continue operating beyond the length (and hence temperature) at which the selective coating thermally decomposes, we suggest that, in this high temperature region of the receiver pipe, the selective coating can be substituted by a hot mirror coating on the glass cover. The plant will therefore use two technologies (selective coating and hot mirror), the selective coating operating in the low temperature region ($< 400^0$C) and the hot mirror coating beyond this temperature, i.e. a hybrid system.

4. Receiver unit analysis

By using the heat transfer mechanism and conservation of energy law we derive the equation of the receiver pipe and the glass cover [3]. To our best knowledge, the receiver unit has not been studied in this context. We could also find no reference to the possibility of a hybrid system, where different types of technologies are used at different temperature ranges, within this context, as being discussed in this paper. Our model describes the heat transfers via radiation between the glass cover, the receiver pipe and the outside, as well as the conductive heat transfer into the working fluid and convective losses to the surroundings. The relevant equations are presented in the following section**.**

4.1. Heat transfer mechanism

4.2. Radiation falling on a surface interacts with it via the mechanisms of reflection (r), transmission (t) and absorption (a = 1 – r – t) in relative proportion [2, 3, 9,] where, "r" and "t" are the reflection and transmission coefficients respectively. In each case, the superscript refers to whether the term of interest applies to visible (v) or Infrared (IR) and the subscript refers to its physical location, being either on the glass cover (g) or the receiver pipe (r).

We did not specifically consider convection and conduction interactions inside the receiver unit, since they are sub-leading mechanisms of heat transfer at high temperatures and can also be assumed to remain similar in both systems. For our calculations we use some approximations [3]. The heating of the receiver unit by the solar radiation was assumed to be uniform. We do not take a temperature gradient along the circumference of the glass cover or receiver pipe into account.

This amounts to assuming good heat conduction of the materials involved. We assumed the glass cover and the receiver to be close enough together so that the majority of the radiation leaving either one is intercepted by the other [3]. We also do not include any heat transfer effects into the working fluid due to the type of flow it displays, laminar or turbulent. We also did not include the abovementioned sub- leading heat transfer mechanisms, such as convection and conduction inside the receiver. Considering all above approximations we get the following equations [3]:

Equations for glass cover [3]:

$$\boldsymbol{AQ^{v} + BA_r\mathcal{E}_r\sigma T_r^4 + CA_g\mathcal{E}_g\sigma T_g^4 - hvA_g(T_g - T_0) = 0} \quad (4.1)$$

Equation for Receiver [3]:

$$\boldsymbol{DQ^{v} + EA_r\mathcal{E}_r\sigma T_r^4 + FA_g\mathcal{E}_g\sigma T_g^4 - \frac{KA_r}{L\left(1+\frac{KA_rK1}{L}\right)}(T_r - T_L) = 0} \quad (4.2)$$

Where A, B, C, D, E and F are in terms of transmission and reflection coefficients, and T_o is the ambient temperature [3]. Detailed form of the coefficients is provided in [3].

The last terms in both equations represents heat loss via convection to the surrounding by the glass cover (4.1), and heat loss to the working fluid via conduction (4.2). Both depend on the local temperature difference linearly. These results have been tested in numerous limits of coefficients and provided the expected results argued on purely physical grounds

4.2 Computer simulation

We have written a programme which simulates the behavior of a solar trough system [3]. Solar radiation is incident on a section of the receiver unit, where it undergoes the interactions described above leading to equations (4.1) and (4.2). The program solves these equations simultaneously for the equilibrium temperature of the glass cover and receiver pipe based on the optical properties of the coating [3]. We used five different reflectivities of ITO coatings, whose coefficients are provided in table 1.

Table1: Reflectivity (R), Transmissivity (T) and Absorptance (A) for five different hot mirror coatings

Reflectivity	Transmissivity	Absorptance
0.75	0.06	0.19
0.80	0.05	0.15
0.85	0.35	0.12
0.90	0.02	0.08
0.95	0.01	0.04

Quantities derived from the program include the equilibrium temperatures of the glass cover, the receiver pipe and the working fluid at incremental points along its length, as well as all heat flows related to them. We used a visual basic program coupled with an excel worksheet. We further checked the results against an independent program in Math CAD.

5. Results and Discussion

5.1 Results

Figure 2 shows the results for the temperatures of glass cover and the receiver pipe for the hot mirror system as a function of length of the solar trough system, with a fluid flowing inside it. The temperature flattens with length due to more thermal radiation being lost to the surrounding, and hence unavailable to heat the receiver and fluid inside. Results for the of heat transfer into the working fluid for the hot mirror system are shown in figure 3.

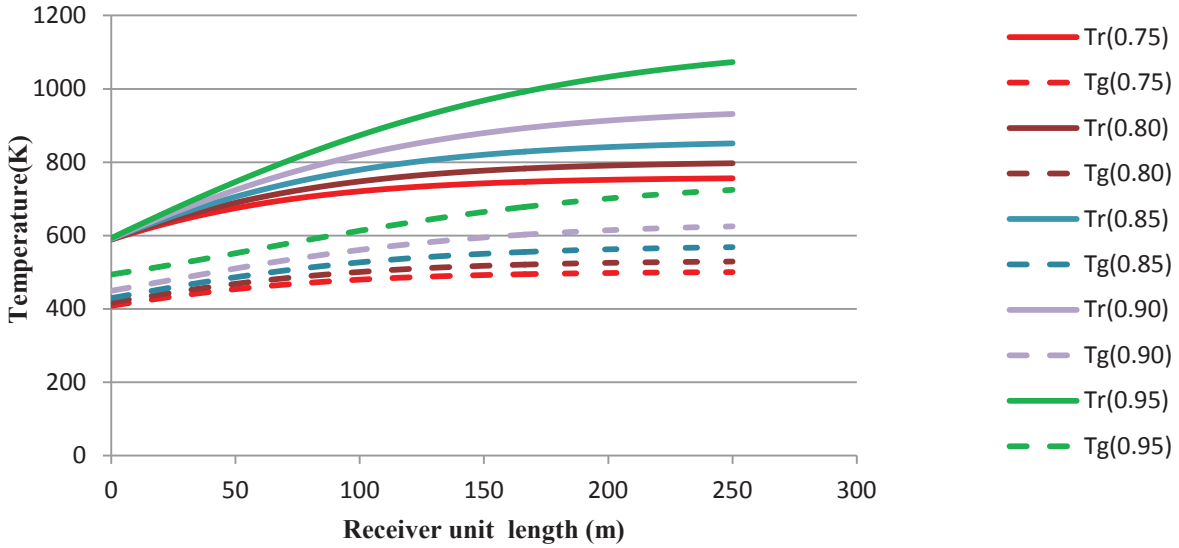


Figure 2: Comparison of the temperatures of glass cover and the receiver pipe for hot mirror system as a function of length of the solar trough system, with a working fluid flowing inside it. Curves are labeled by their IR reflectivity.

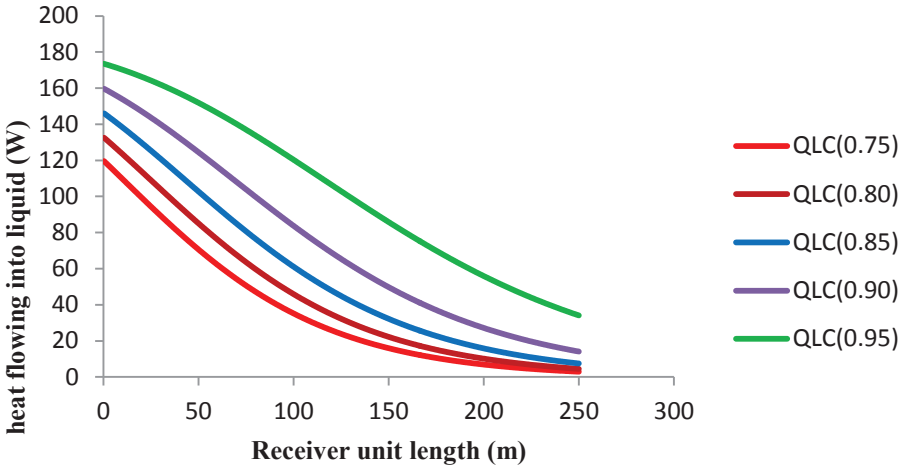


Figure 3: The heat going into the working fluid for hot mirror as a function of receiver unit length.

5.2 Discussion

The simulation was started at a temperature close to where failure of the selective coating is expected (600K). Relative equilibrium temperatures of hot mirror can be compared in this region, and the question of applicability answered.

The selective coating should be used to the temperatures where it is stable (below 680K). Beyond this critical temperature, a substitution to the hot mirror system is advisable.

The glass cover temperature is seen to be much lower than the receiver pipe temperature (by 200K), and this allows the glass cover to be coated with a hot mirror, since the hot mirror coating can operate at these temperatures. This will allow the working fluid to be heated to more elevated temperatures, making the overall system more efficient. The hot mirror coating is stable up to temperatures of about 600K, and the glass cover temperatures exceeds this for T_{glass} (R=0.95) and T_{glass} (R=0.90).

Figure 3 indicates the heat moving into the working fluid (heat into working fluid, QLC) along the pipe unit length. It is seen that the hot mirror perform significantly better in terms of working fluid heat transfer for higher reflectivity in the IR region, which increases the efficiency of a solar trough plant.

6. Conclusions

A set of heat transfer equations was derived, modeling the thermal behavior of a solar trough receiver unit. Radiative heat transfer within the receiver unit was considered, as well as convective losses to the outside, and heat transfer into the liquid. A code was written using the equations, and this was the main source of our results. It was seen that the glass cover temperature was sufficiently low for a hot mirror system for some reflectivities. This allows the construction of solar trough systems with longer receiver pipes of two types: selective coating should be used in the temperature region where it is appropriate, and a hot mirror system at higher temperatures. This will allow the working fluid to reach a greater temperature, and hence better overall plant performance. For this purpose, it would be well worth investigating the hot mirror receiver system in the future.

We compared five different thicknesses of ITO coating with different optical properties. It was seen that the heat transfer was higher for higher IR reflectivity of the hot mirror, and that more heat was transferred into the working fluid. However, very high reflectivities can lead to the temperature of the glass cover to exceed the working temperature of the hot mirror coating, damaging it.

References

[1] Sargent and Lundy 2003 *Assessment of Parabolic Trough and Power Tower Solar Technology Cost and Performance Forecasts,* Subcontractor Report, SL-5641, prepared for the U.S. Department of Energy and National Renewable Energy Laboratory, Chicago, Ill

[2] Twidell J and Weir T 2006 *Renewable Energy Resources*, pp. 115-176. (Taylor & Francis Group, New York)

[3] Cyulinyana MC 2011 *Investigation into applicability of existing renewable energy technologies and possible efficiency increase, case study: Rwanda*, pp.85-122. M.Sc. thesis, University of the Witwatersrand, South Africa.

[4] DURAN. Tubing, Capillary and Rod of Borosilicate Glass3.3, Schott, Glass made of Ideas. Online at: http://www.schott.com/tubing/english/download/schott-tubing_brochure_duran_english.pdf

[5] Kennedy CE and Price H 2005 *Progress in development of high-temperature solar-selective coatings* NREL/CP-520-36997 (Proc. ISEC2005 2005 International Solar Energy Conference August 6-12, 2005, Orlando, Florida USA, ISEC2005-76039)

[6] Kennedy CE 2002 *Review of Mid- to High-Temperature solar Selective Absorber Materials (*NREL/TP-520-31267, Golden, CO: National Renewable Energy Laboratory)
[7] Kaltschimitt M, Streicher W and Wiese A 2007 *Renewable energy technology-economics and environment* (Springer, Berlin, New York)
[8] Academic press. 2009 *Renewable Energy Focus handbook* (Elsevier.)
[9] Kreith F and Kreider JK 1978 *Principles of solar engineering* (pp. 1-33, pp. 85-159, pp. 203-309. Hemisphere Publishing Cooperation)
[10] You Y and Hu EJ 2002 *A medium-temperature solar thermal power system and its efficiency optimization*. (Appl. Thermal Eng. **22**, 357-364)
[11] Ferrer P 2008 *Enhanced efficiency of a parabolic solar trough system through use of secondary radiation concentrator* (S. Afr. J. Sci. **104**.)
[12] Kumashiro Y 2000 *Electric Refractory Materials* (Marcel Dekker Inc., New York)

Protein localization and folding mechanisms revealed by molecular dynamics simulations

James Gumbart

Postdoctoral Fellow, Argonne National Laboratory, Argonne, IL, USA

E-mail: jgumbart@mcs.anl.gov

Abstract. Many nascent proteins, including nearly all membrane and secreted proteins, must traverse a membrane-bound protein-conducting channel prior to their full maturation. This channel, the Sec translocon, is found in all domains of life and possesses the novel ability to direct nascent proteins to the membrane or to the extracellular space, depending on their sequences, often concomitant with their synthesis by the ribosome. By combining atomic structures with cryo-electron microscopy data using the molecular dynamics flexible fitting method, we have developed some of the first views of inactive and active translocons in complex with the ribosome. These views reveal a conserved mode of interaction between translocon and ribosome as well as the roles of specific elements of both in protein localization. We also carried out two-dimensional potential-of-mean-force calculations to explore the structure of the nascent peptide within the translocon environment. The calculations revealed that the translocon exerts a small bias on the peptide towards a helical state. This bias can serve to facilitate, e.g., the insertion of nascent membrane proteins into the otherwise inhospitable lipid bilayer.

1. Introduction

All proteins begin their existence as long polypeptide chains synthesized by the ribosome, a large cytoplasmic protein/RNA complex, after which they fold into specific three-dimensional structures. Many proteins are made in the cytoplasm of the cell, where they remain to carry out their function(s). However, proteins not meant to be localized to the cytoplasm are directed by means of a signal sequence to the membrane, that of the endoplasmic reticulum in eukaryotes and the cytoplasmic membrane in prokaryotes. These proteins include nearly all membrane proteins, for which the first hydrophobic transmembrane (TM) segment often serves as the signal sequence, as well many proteins destined for the ER, the bacterial periplasmic space or outer membrane, or even outside the cell [1, 2].

Nascent proteins can be directed to the translocon after synthesis by the ribosome, but also during formation, processes referred to as post- or co-translational translocation, respectively [3]. Co-translational translocation is similar in all species. When the ribosome begins synthesizing a protein with an N-terminal signal sequence that directs the nascent protein to the membrane, it halts translation until the ribosome/nascent chain complex can be directed to the membrane and there docked with the translocon. Once bound, the ribosome resumes translation and begins delivering the nascent protein into the channel.

After recognition of the signal sequence, the translocon begins its most complex role: placing the pre-protein in the proper location, either into the membrane or on the other side of the membrane. The mechanisms of this dual ability (soluble protein translocation and membrane protein integration) have slowly come into focus thanks in part to the release of a high-resolution

crystal structure of SecYEβ in 2004 [4]. SecYEβ is the core component of the archaeal translocon, a heterotrimeric complex that is homologous to SecYEG in bacteria and the Sec61 complex in eukaryotes. Prior to the release of the crystal structure, most structural data on the translocon came from low resolution cryo-electron microscopy (cryo- EM) maps. These maps typically were determined for ribosome-translocon complexes in which the resolution of the translocon was too low to discern any structure beyond a toroidal shape [5–7]. This shape was presumed to represent 3-4 SecY/Sec61 complexes with the channel proper located at the central depression, between the complexes. However, the crystal structure of a single SecY complex illustrated features suggesting that translocation occurred through a monomer alone [4]. Key features of the channel observed in the structure included a plug and a pore ring proposed to maintain membrane integrity (see Fig. 1). Also seen in the structure was the likely location of a lateral gate for the integration of membrane proteins into the lipid bilayer [4].

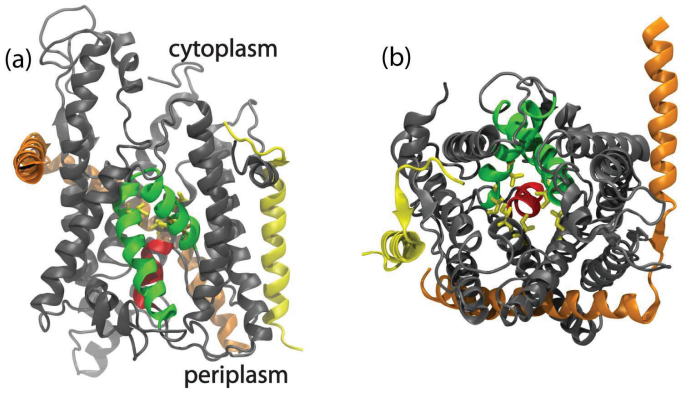


Figure 1. SecY complex. The SecY complex, shown (a) in the membrane plane and (b) from the cytoplasmic side, is composed of three proteins, SecY (grey), forming the channel itself, and two accessory proteins, SecE (orange) and Secβ/SecG (yellow). The plug (red), lateral gate (green), and pore ring (yellow sticks) are also shown.

While the crystal structure suggests a monomer can form the functional channel, significant debate over the possible role and organization of oligomers of the channel remains. As noted previously, multiple cryo-EM maps appear to show as many as four copies of the channel present when bound to the ribosome, although the resolution in the channel region is poor. More recent maps of SecY bound to a ribosome as well as Sec61 bound to a ribosome show only a single copy of the channel, with additional density in the map contributed by detergent and lipids [8, 9]. Experiments suggest, however, that oligomers of the Sec61 complex form a high affinity binding site for ribosomes in eukaryotes, but that their formation occurs only after initial targeting of the ribosome to the membrane by SRP [10].

In contrast to increased binding sites, a larger pore has also been put forward as a reason for channel oligomerization. Fluorescence quenching experiments led to the conclusion that the functional translocon pore is 40-60 Å wide [11]. However, more recent experiments indicate that the channel is incapable of supporting translocation of molecules larger than ∼20 Å in size [12]. Another functional consideration is the need for the channel to support at least modest folding of nascent membrane proteins prior to their insertion. Because an exposed peptide backbone cannot be inserted into the core of the bilayer, secondary structure, i.e., transmembrane helices, need to be formed while still within the translocon [13]. While the requirement of folding would suggest a large pore in the translocon is required, experiments have demonstrated that folding can occur within the confined environment of the ribosome's exit tunnel, which has a diameter of 10-20 Å [14–16].

In this proceeding, recent advancements in both structural and functional aspects of the protein-conducting translocon are discussed. First, the availability of cryo-EM maps of ribosome-channel complexes at unprecedented resolutions combined with novel computational methods permit unambiguous determination of the channel's oligomeric state. This state is found to

be consistently monomeric, across species and at different stages of the translocation process. Further supporting the validity of the resolved structures, numerous conserved interactions modulating formation of the complex are found. Additionally, using free-energy techniques, it is demonstrated that, rather than presenting a hindrance to folding, the monomeric translocon's narrow diameter actually aids protein folding, thus catalyzing the development of membrane proteins.

2. Results

2.1. Structures of ribosome-channel complexes

The stringent requirements of X-ray crystallography, including a high stability of the molecule being imaged during both crystallization and bombardment by high energy photons, greatly constrains the number of functionally relevant biomolecular states that can be resolved at atomic resolutions. This constraint becomes even more apparent for large complexes, such as the ribosome which itself contains over 50 proteins and RNA strands. Despite the challenges faced, however, crystallographers have made significant progress in recent years for ribosomes and ribosome complexes, including capturing it in complex with partners such as EF-Tu [17]. However, the functional relevance of these complexes has been questioned, due in part to the inclusion of unnatural inhibitors required to stabilize them [18].

In contrast to crystallography, cryo-electron microscopy has the ability to visualize large complexes in their native environments and in functionally relevant states. However, this ability comes at a cost of (typically) a resolution insufficient for placing individual atoms. Circumventing this resolution limit, recent methodological advancements have led to the use of multimodal data, combined computationally, to determine atomic structures from cryo-EM maps. One method in particular, molecular dynamics flexible fitting (MDFF), fits atomic-scale structures of individual components in unphysiological states to cryo-EM maps of complexes in the course of a molecular dynamics simulation [19, 20]. This method has been used extensively on ribosome complexes, providing multiple breakthroughs in understanding of tRNA accommodation, protein synthesis, and extrusion of the nascent peptide [21–24].

MDFF has also been utilized to resolve complexes between the ribosome and the translocon. The first cryo-EM map to hint at a monomeric state for the translocon, that of a non-translating ribosome bound to a bacterial SecY complex, was still of too low resolution to visualize the channel within the density, except for those regions directly interacting with the ribosome [8]. Application of MDFF to this map along with subsequent free MD simulations, served to refine these interactions, allowing their identification and classification [21]. Interactions between the ribosome and channel were found to localize to four regions of the SecY channel (see Fig. 2a). The two most prominent regions involve the two cytoplasmic loops of SecY, loops 6/7 and 8/9, which insert directly into the ribosome's polypeptide exit tunnel. Conserved basic residues in these loops bind strongly to the highly charged RNA backbone of the ribosome's 23S [8, 21]. Although it may seem that such insertion would block the tunnel and, thus, prevent exit of the nascent protein, simulated translocation indicates that the tunnel has sufficient width to permit co-existence of SecY's loops and the nascent protein [21]. The other two regions of ribosome-channel interactions involve the accessory protein, SecE, as well as the C-terminus of SecY, both conserved [21].

Although numerous pieces of circumstantial evidence support the conclusions drawn from the first MDFF-determined ribosome-translocon structure, the fact that the experiment used non-translating, and thus non-fiunctional, ribosomes remains. More recent maps of eukaryotic ribosome-translocon complexes [24], however, display translating ribosomes, and, thus, present a test of the proposed nature of the ribosome-channel structure and interface. MDFF applied to a map of a mammalian complex confirmed earlier conclusions: first, that the channel is monomeric and second, that the pattern of interactions involve the same four regions (see Fig. 2b).

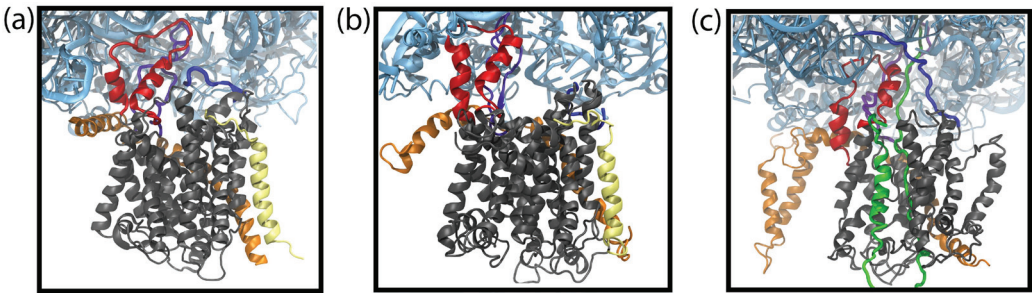


Figure 2. Interactions between ribosome and translocon. In all panels, the ribosome is shown in light blue above, with the channel colored as in Fig. 1. Loops 6/7 and 8/9 are shown in red and purple, respectively. SecY/Sec61's C terminus is colored in dark blue. (a) Non-functional, bacterial ribosome-SecY complex. (b) Mammalian ribosome-Sec61 complex. (c) Bacterial ribosome-SecY-nascent-chain complex. The nascent peptide is colored in green.

Nonetheless, unambiguous density for the nascent polypeptide could not be assigned [21]. This deficiency was remedied by the most completely resolved complex to date, that of a ribosome bound to SecY in the process of inserting a nascent membrane protein [25]. The structure reveals the signal anchor of the nascent protein at SecY's lateral gate, now open, rationalizing decades of biochemical data and proving numerous assertions about the mechanisms of membrane protein insertion. The channel is again monomeric, definitively proving that a monomeric Sec complex is competent for translocation [25].

2.2. Folding of the nascent chain

As noted above, prior to insertion into the membrane, a nascent protein must shield its polar backbone from the nonpolar bilayer core. The MDFF-derived structure of a nascent membrane protein in an intermediate state of insertion, indeed, confirms that the signal anchor, which is in contact with the bilayer, is already helical (see Fig. 2c). Thus, folding must have occurred at some point during its secretion from the ribosome exit tunnel into the channel, despite the narrow environments encountered. Numerous experimental studies have demonstrated that nascent peptides can form secondary structure inside the ribosome and translocon channel [14–16], although the factors governing such formation are not always clear. While it may be assumed that the ability to integrate into the membrane would correlate with folding in the translocon, early experiments indicated this is not always the case. For example, polyVal, a strongly hydrophobic sequence, does not fold in the translocon, at least initially, whereas polyAla does [14]. Based on these results, it was suggested that folding in the channel mimics folding in a purely aqueous environment, rather than a non-polar environment [14].

To determine the energetics of folding within the translocon, two-dimensional potential-of-mean-force (PMF) calculations were carried out on a 10-alanine oligopeptide inside the channel [26]. The two reaction coordinates utilized were the position along the channel axis, z, and the average i, $i+4$ hydrogen bonding distance along the backbone, i.e., a measure of *alpha*-helicity. In this way, the relatively probability of forming a helix at different points in the channel could be determined and correlated with the channel's properties. The channel was restrained in a laterally open conformation, thus exposing the 10-alanine peptide to the bilayer at least part of the time, as may be the case during translocation [27, 28].

Projections of the resulting 2D PMF along different values of the translocation coordinate, z, reveal distinct behaviors of the oligopeptide at different regions of the channel. At all points in the channel, the 10-alanine peptide displays two folding minima, one in a compact, helical state and one in an extended state. These minima mimic those of the peptide in water, confirming

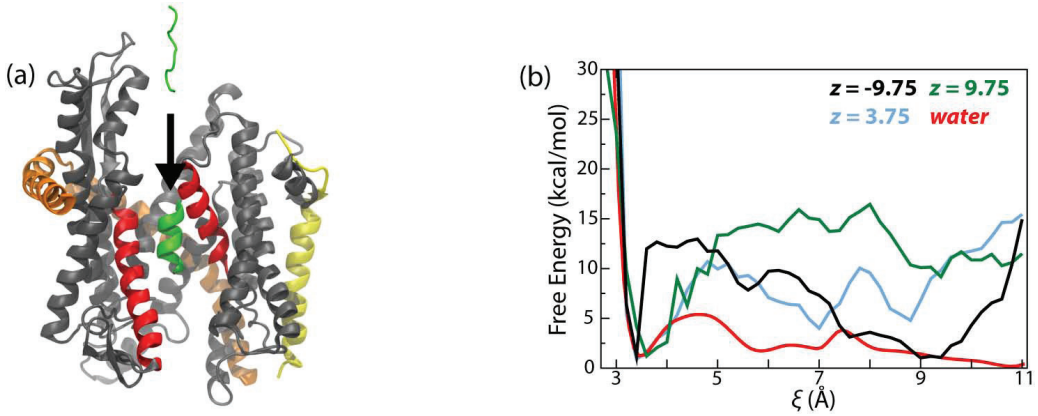


Figure 3. Folding within SecY. (a) SecY complex with its three subunits colored gray, orange, and yellow, respectively. The two helices lining the lateral gate are shown in red. The lateral gate (red) is an open state. The 10-alanine peptide is shown inside the channel as a green helix. (b) Projections of the 2D PMF along different values of z as a function of ξ. The 1D PMF for folding in water is included for comparison. The α-helical state is at ξ=3.4 Å, while the extended states are found in the range ξ=9-11 Å.

the suggestion of Mingarro et al. [14] (see Fig. 3b). However, the free energies of those two minima are shifted with respect to each other depending on channel location, with the helical state being lower in free energy than the extended state within the channel. Additionally, the barrier separating them is much greater in the channel than in water. Thus, it is apparent that although the predominant conformations of 10-alanine reflect those in water, the channel biases it toward the helical state and, furthermore, locks it in that state by slowing transitions to the extended state.

3. Conclusion

Within the last few years, knowledge of the structure and function of the protein-translocation machinery has taken a great leap forward. Beginning with the resolution of the first atomic structure in 2004 [4], new advancements have come at an accelerating pace. Two major areas of such advancements were highlighted in this proceeding, namely determination of the common structure of ribosome-channel complexes across multiple species and also the functional consequences of that structural arrangement. The structures clearly reveal not only a monomeric Sec complex as the functional translocation channel, but also the specific interactions that maintain the complex. Furthermore, 2D PMF calculations demonstrate that the channel exerts a slight helical bias on entering polypeptides, thus enabling subsequent insertion for membrane proteins into the bilayer. Far from being a simple conduit, the translocon is an active player in the localization of nascent proteins. Taken together, these results provide an exceptional starting point for future investigations into membrane protein development, a complete understanding of which is necessary for *ab initio* structure prediction as well as for the design of synthetic proteins.

4. Appendix I

All molecular dynamics simulations were run using the highly parallelized program NAMD [29]. For MDFF simulations, the system of interest is subjected to forces from three primary sources, namely the standard MD force field, a potential derived from the cryo-EM density that drives the structure toward high-density regions of the map, and restraints to maintain secondary structure, which is often not well resolved in the density [19]. For calculation of the 2D PMF, adaptive

biasing forces as implemented in NAMD were used [26, 29, 30]. The translocation reaction coordinate was subdivided into 10-Å-sized windows, which were simulated independently and their resulting biasing forces joined after completion. Further details of specific simulation setups are provided in Gumbart et al. (2009), Frauenfeld et al. (2011), and Gumbart et al. (2011) [21, 25, 26].

Acknowledgments
The author sincerely thanks Klaus Schulten, Chris Chipot, and Roland Beckmann for their contributions to multiple aspects of the work described in this proceeding.

References

[1] Rapoport T A, Goder V, Heinrich S U and Matlack K E 2004 *Trends Cell. Biol.* **14** 568–575
[2] White S H and von Heijne G 2005 *Curr. Opin. Struct. Biol.* **15** 378–386
[3] Rapoport T A 2007 *Nature* **450** 663–669
[4] van den Berg B, Clemons Jr W M, Collinson I, Modis Y, Hartmann E, Harrison S C and Rapoport T A 2004 *Nature* **427** 36–44
[5] Morgan D G, Ménétret J F, Neuhof A, Rapoport T A and Akey C W 2002 *J. Mol. Biol.* **324** 871–886
[6] Beckmann R, Bubeck D, Grassucci R, Penczek P, Verschoor A, Blobel G and Frank J 1997 *Science* **278** 2123–2126
[7] Ménétret J F, Neuhof A, Morgan D G, Plath K, Radermacher M, Rapoport T A and Akey C W 2000 *Mol. Cell* **6** 1219–1232
[8] Ménétret J F, Schaletzky J, Clemons Jr W M, Osborne A R, Skånland S S, Denison C, Gygi S P, Kirkpatrick D S, Park E, Ludtke S J, Rapoport T A and Akey C W 2007 *Mol. Cell* **28** 1083–1092
[9] Ménétret J F, Hegde R S, Agular M, Gygi S P, Park E, Rapoport T A and Akey C W 2008 *Structure* **16** 1126–1137
[10] Schaletzky J and Rapoport T A 2006 *Mol. Biol. Cell* **17** 3860–3869
[11] Hamman B D, Chen J C, Johnson E E and Johnson A E 1997 *Cell* **89** 535–544
[12] Bonardi F, Halza E, Walko M, Plessis F D, Nouwen N, Feringa B L and Driessen A J M 2011 *Proc. Natl. Acad. Sci. USA* **108** 7775–7780
[13] White S H and von Heijne G 2008 *Annu. Rev. Biophys.* **37** 23–42
[14] Mingarro I, Nilsson I, Whitley P and von Heijne G 2000 *BMC Cell Biol.* **1** 3
[15] Lu J and Deutsch C 2005 *Nat. Struct. Mol. Biol.* **12** 1123–1129
[16] Tu L W and Deutsch C 2010 *J. Mol. Biol.* **396** 1346–1360
[17] Schmeing T M, Voorhees R M, Kelley A C, Gao Y G, Murphy F V, Weir J R and Ramakrishnan V 2009 *Science* **326** 688–694
[18] Gumbart J, Schreiner E, Trabuco L G, Chan K Y and Schulten K 2011 *Molecular Machine in Biology* ed Frank J (Cambridge University Press) in press.
[19] Trabuco L G, Villa E, Mitra K, Frank J and Schulten K 2008 *Structure* **16** 673–683
[20] Trabuco L G, Schreiner E, Gumbart J, Hsin J, Villa E and Schulten K 2011 *J. Struct. Biol.* **173** 420–427
[21] Gumbart J, Trabuco L G, Schreiner E, Villa E and Schulten K 2009 *Structure* **17** 1453–1464
[22] Villa E, Sengupta J, Trabuco L G, LeBarron J, Baxter W T, Shaikh T R, Grassucci R A, Nissen P, Ehrenberg M, Schulten K and Frank J 2009 *Proc. Natl. Acad. Sci. USA* **106** 1063–1068
[23] Seidelt B, Innis C A, Wilson D N, Gartmann M, Armache J P, Villa E, Trabuco L G, Becker T, Mielke T, Schulten K, Steitz T A and Beckmann R 2009 *Science* **326** 1412–1415
[24] Becker T, Bhushan S, Jarasch A, Armache J P, Funes S, Jossinet F, Gumbart J, Mielke T, Berninghausen O, Schulten K, Westhof E, Gilmore R, Mandon E C and Beckmann R 2009 *Science* **326** 1369–1373
[25] Frauenfeld J, Gumbart J, van der Sluis E O, Funes S, Gartmann M, Beatrix B, Mielke T, Berninghausen O, Becker T, Schulten K and Beckmann R 2011 *Nat. Struct. Mol. Biol.* **18** 614–621
[26] Gumbart J, Chipot C and Schulten K 2011 *J. Am. Chem. Soc.* **133** 7602–7607
[27] Gumbart J and Schulten K 2007 *Biochemistry* **46** 11147–11157
[28] du Plessis D J F, Berrelkamp G, Nouwen N and Driessen A J M 2009 *J. Biol. Chem.* **284** 15805–15814
[29] Phillips J C, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel R D, Kale L and Schulten K 2005 *J. Comp. Chem.* **26** 1781–1802
[30] Hénin J, Forin G, Chipot C and Klein M L 2010 *J. Chem. Theor. Comp.* **6** 35–47

Laser irradiation: a complementary treatment for wounds

N N Houreld[1] **and H Abrahamse**

Laser Research Centre, Faculty of Health Sciences, University of Johannesburg, P.O. Box 17011, Doornfontein, South Africa, 2028. Tel: +27 (0)11 559-6406

E-mail: nhoureld@uj.ac.za / habrahamse@uj.ac.za

Abstract. *Aim:* Since the invention of the laser, its application in the health sector has been studied and in an attempt to discover effective alternative treatments, Low Level Laser Therapy (LLLT), commonly known as biostimulation or photo-biostimulation, has emerged. This therapy has been successfully used both in *in vitro* and *in vivo* studies in wound healing. Although this therapy is in use worldwide, the mechanism of action is not fully understood. *Methods:* Various cell culture models, such as wounded, diabetic wounded, ischemic and hypoxic have been exposed to visible and infra-red laser light and the effect on cell migration, cell survival, proliferation, cytotoxicity, mitochondrial responses, nitric oxide (NO) release, secondary messenger activation, DNA damage and pro-inflammatory cytokine expression have been studied. *Results:* Laser irradiation at the correct wavelength and fluence has shown to have a positive effect on stressed cells *in vitro*. There is an increase in migration, survival and proliferation, mitochondrial activity, NO release and secondary messenger activation. A decrease in cytotoxicity, DNA damage and pro-inflammatory cytokines is also seen. *Conclusion:* LLLT offers an alternative wound healing therapy. At a biochemical level there is a positive effect on cells, with stressed cells being pushed into cell survival pathways.

1. Introduction

Wound healing involves a series of overlapping and intertwining events all aimed at reversing the loss of structural integrity and is controlled by a wide variety of cells, growth factors, cytokines and enzymes, all of which are released at the wound site [1]. Wound healing is divided into four main events, namely haemostasis, inflammation, proliferation and remodelling. When this sequence of events is disrupted, delayed wound healing ensues. Chronic, slow-to-heal or non-healing wounds are a common complication of diabetes mellitus (DM), particularly on the feet and lower limbs. These diabetic foot ulcers are susceptible to infection and often necessitate lower-limb amputation, which impacts heavily on patients, their families, health care departments and government. Around 20% of all patients with DM who develop foot ulcers require amputation [2]. Amputation occurs 30 times more frequently in diabetic patients than in the general population [3]. The underlying causes of these foot ulcers are thought to be due to micro- and macrovascular disease and advanced glycation end products [4] and are associated with sensory loss. The current treatment protocol is both systemic, treatment of the metabolic condition (diabetes), and local treatment of the diabetic ulcer. Common treatment of the diabetic foot involves debridement, wound care and dressing, antibiotic therapy,

[1] To whom any correspondence should be addressed.

offloading, accommodative orthodontics and wearing the correct footwear. Patients are encouraged to examine their feet daily and should visit their healthcare providers annually for check-ups.

Low Level Laser Therapy (LLLT), also known as phototherapy or photobiomodulation, is the use of low-powered lasers to stimulate cellular metabolism and biochemical processes. LLLT is widely used in dermatology, somatology, dentistry, physiotherapy and veterinary science, and applications include wound healing, pain attenuation, modulation of the immune system and bone and nerve repair. Wavelengths of between 500 – 1,100 nm are commonly used with fluencies around 1-4 J/cm^2 [5]. The choice of wavelength is dependent on the depth of penetration required, and wavelengths of 600 – 700 nm are commonly used to treat superficial wounds (shorter penetration depth), while 780 – 950 nm is used for treatment of deeper tissues [6].

LLLT has been show to stimulate and improve wound healing both *in vitro* and *in vivo*. Caetano and colleagues [7] showed that phototherapy (combined 660 and 890 nm, 3 J/cm^2) not only promoted the healing of large ulcers which did not respond to conventional treatments, but healing was also significantly faster in irradiated groups. Minatel and colleagues [8] also showed that combined 660 and 890 nm promoted healing in diabetic leg ulcers. Granulation and healing rates were higher in irradiated groups. Moore and co-workers [9] showed that LLLT in the visible red range (665 and 675 nm; 10 J/cm^2) stimulated fibroblast and endothelial cell proliferation. Lim and co-workers [10] showed that LLLT with 635 nm reduced intracellular reactive oxygen species (ROS), an important aspect in angiogenesis, and an increase in viability in hypoxic/ischemic endothelial cells. LLLT has also been shown to be effective in the treatment of infected wounds [11,12].

2. Methods

Fibroblasts (WS1, ATCC, CRL-1502) have been used in wound healing studies in a variety of cell models. These include normal, wounded (scratch model), diabetic (continuous growth in additional glucose), diabetic wounded, acidic (pH 6.70), hypoxic (oxygen deprivation) and ischemic (oxygen and foetal bovine serum, FBS, deprivation) and have been previously described [13-16]. Irradiation parameters are shown in table 1, and cellular assays to monitor changes in irradiated cells are shown in table 2. Cells were irradiated in the dark from the top without culture lids (figure 1). A non-irradiated control (0 J/cm^2) was included for each model in each experiment and treated in the same manner, barring irradiation; instead they were placed in a dark box on the bench. Cells were irradiated with 0, 0.5, 2.5, 5, 10 or 16 J/cm^2. Post-irradiation cells were incubated for a variable amount of time and cells detached by trypsinization and re-suspended to a concentration of 1 x 10^5 / 100 μl.

Table 1. Laser parameters.

	Laser			
Parameter	He-Ne	Diode		Nd:YAG
Wavelength (nm)	632.8	830	636	1064
Wave emission	Continuous	Continuous	Continuous	Continuous
Power output (mW)	23	55	95	1,000
Power density (mW/cm^2)	2 - 3	6	11	12.7
Spot size (cm^2)	9.1	9.1	9.1	78.5

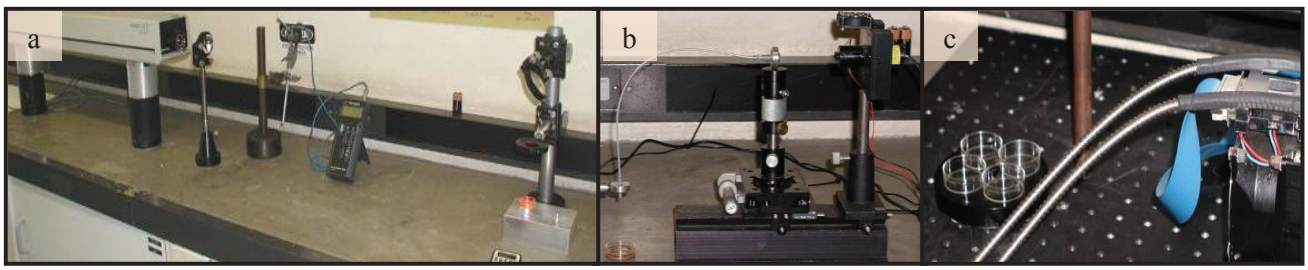


Figure 1. Laser irradiation of cells using a He-Ne laser (a), 830 nm diode (b) and Nd:YAG laser (c).

Table 2. Assays used to measure cellular responses to LLLT**.**

Parameter	Assay	Data Collection
Morphology	Migration, haptotaxis	Light microscopy
Viability	Trypan blue staining	Light microscopy
	Adenosine Triphosphate (ATP)	Luminescence
Proliferation	Optical density (OD)	Spectroscopy (A_{540} nm)
	Basic Fibroblast Growth Factor (bFGF)	ELISA (A_{450} nm)
	Alkaline Phosphatase (ALP)	Colorimetric (A_{405} nm)
	XTT (tetrazolium salt)	Colorimetric (A_{450} nm)
Cytotoxicity	Lactate Dehydrogenase (LDH)	Colorimetric (A_{490} nm)
Apoptosis	Caspase 3/7	Luminescence
DNA Damage	Comet assay	Fluorescent microscopy
	Real Time RT-PCR	Fluorescence
NO	Griess assay	Colorimetric (A_{540} nm)
Cytokines	Interleukin (IL)-6	ELISA ($A_{450/\ 570}$ nm)
	IL-1β	ELISA (A_{450} nm)
	Tumour Necrosis Factor alpha (TNF-α)	ELISA (A_{450} nm)
Secondary messengers	Cyclic Adenosine Monophosphate (cAMP)	Colorimetric (A_{405} nm)
	Intracellular calcium (Ca^{2+})	Colorimetric (A_{612} nm)
Mitochondrial responses	Mitochondrial Membrane Potential (MMP)	Flow Cytometry
	Enzyme kinetics	Spectroscopy

3. Results and Discussion

All results are summarised in table 3.

3.1. He-Ne laser (632.8 nm)

Hawkins and Abrahamse [13,17] showed that when normal and wounded fibroblast cells were irradiated at 632.8 nm, 2.5 J/cm^2 increased cellular proliferation and had no adverse effect on cellular viability, DNA damage or cytotoxicity. When the same cells were irradiated with 5 J/cm^2, cells migrated towards the central scratch and there was an increase in viability and proliferation. On the other hand, 10 and 16 J/cm^2 produced a significant amount of cellular and molecular damage to cells. Morphologically, there was evidence of cellular lysis with no migration and cells detaching from the culture dish. There was a decrease in viability and proliferation and an increase in cytotoxicity and genetic damage. Wounded cells responded better when irradiated once a day on two consecutive days (24 h between irradiations). The effect on proliferation was further validated by incubating cells in the presence of 5 mM hydroxyurea (HU), an inhibitor of proliferation. In the presence of HU, irradiation with 5 J/cm^2 was still able to increase proliferation and migration of wounded cells [18].

Diabetic wounded cells irradiated at 5 J/cm^2 showed significant increases in migration, proliferation and IL-6 (a marker of proliferation and differentiation), while 16 J/cm^2 gave the opposite results [19]. Cells responded better when irradiated on two non-consecutive days (72 h between irradiations) [14,20,21]. There was complete closure of the central scratch when irradiated with 5 J/cm^2. In fact, there was a decrease in DNA damage and membrane damage (LDH released by damaged, compromised cells) as compared to non-irradiated cells. Cells irradiated on two non-consecutive days with 16 J/cm^2 still showed an increase in damage and incomplete wound closure.

The induction of acidosis and hypoxia significantly reduced MMP. When irradiated with 5 J/cm^2, there was an increase in Ca^{2+} which produced an increase in MMP, ATP and cAMP in wounded, acidotic and hypoxic cells, while as expected; a fluence of 16 J/cm^2 produced a decrease in MMP [22]. Thus laser irradiation was shown to restore cellular homeostasis, even in stressed cells (diabetic,

hypoxic, acidotic) by increasing mitochondrial activity which in turn stimulates a variety of other cellular functions.

Table 3. Summary of results.

	Laser (nm)							
	632.8		636		830		1064	
Parameter	5 J/cm^2	16 J/cm^2	5 J/cm^2	16 J/cm^2	5 J/cm^2	16 J/cm^2	5 J/cm^2	16 J/cm^2
Morphology (migration)	↑	↓	↑	↓	↑	↓	↓	↓
Viability	↑	↓	↑	↓	↑	↓	↓	↓
Proliferation	↑	↓	↑	↓	↑	↓	↓	↓
Cytotoxicity	↓	↑	↓	↑	↓	↑	↑	↑
Apoptosis	↓	↑	↓	↑	↓	↑	↑	↑
DNA Damage	↓	↑	↓	↓	↓	↑	↑	↑
NO	n/a	n/a	n/a	n/a	↑	n/a	n/a	n/a
Inflammatory Cytokines	n/a	n/a	↓	n/a	↓	n/a	n/a	n/a
Secondary messengers	↑	↓	n/a	n/a	n/a	n/a	n/a	n/a
Mitochondrial responses	↑	↓	↑	n/a	↑	n/a	n/a	n/a

↑ - Increase ↓ - Decrease

3.2. Diode laser (830 nm)

Wounded cells irradiated with 5 J/cm^2 showed increased migration of cells across the central scratch as compared to unirradiated cells [23]. Cells also showed an increase in proliferation (IL-6).

Diabetic wounded cells irradiated at a longer wavelength in the infra-red range also responded to 830 nm when irradiated with 5 J/cm^2 on two non-consecutive days [24,25]. There was an increase in migration and hastened wound closure; however closure of the wound was incomplete as seen in the same cells irradiated at 632.8 nm, an increase in proliferation and a decrease in cellular and genetic damage was also seen. Cells irradiated with 16 J/cm^2 showed more damage compared to non-irradiated and irradiated (5 J/cm^2) cells. A fluence of 5 J/cm^2 was able to increase NO and ROS, and produce a decrease in the inflammatory cytokine TNF-α and IL-1β [1]. The increase in NO appears to be due to a direct photochemical process since the increase is seen immediately and not at 1 h post-irradiation. Thus, laser irradiation has an anti-inflammatory response on diabetic wounded cells.

3.3. Diode laser (636 nm)

As with irradiation with the He-Ne laser, irradiation at 636 nm with a diode laser produced an increase in DNA damage in wounded cells irradiated with 5 or 16 J/cm^2. However, when cells were left to incubate for 24 h post-irradiation, a significant decrease was seen compared to the same cells incubated for only 1 h [26]. Thus the initial increase in DNA damage is repairable and DNA mechanisms are activated. When real time reverse transcription (RT) polymerase chain reaction (PCR) was used to determine the expression of the DNA repair enzyme N-methylpurine DNA glycosylase (MPG; repairs modified bases including 8-oxyguanine, which is detected by the modified comet assay using enzymes), no significant difference in expression was detected. Thus DNA repair is by a mechanism other than MPG [26].

At 5 J/cm^2 irradiation increased cellular viability and proliferation and decreased apoptosis in hypoxic cells [16]. There was also a decrease in the pro-inflammatory cytokine TNF-α. A decrease in

IL-1β and apoptosis was seen in diabetic wounded cells, as was an increase in viability and proliferation.

3.4. Nd:YAG (1064 nm)

With irradiation using near infrared laser light, both wounded and diabetic wounded cells irradiated at 1064 nm with 5 or 16 J/cm^2 showed a significant decrease in migration, viability and proliferation [23-25,27]. Diabetic wounded cells showed an increase in damage and apoptosis at both fluencies [24,25], while normal wounded cells did not show an increase at 5 J/cm^2 [23].

4. Conclusion

Stressed cells (wounded, diabetic, diabetic wounded, hypoxic, ischemic, acidic) showed significant differences as compared to normal cells, showing that these models are effective to use in wound healing studies, as well as LLLT studies [15]. The effect of LLLT, or any other treatment, may show insignificant effects on normal, non-stressed cells. Thus it is important to induce some sort of stress onto cells when conducting these kinds of experiments, as well as testing these models against normal cells.

LLLT stimulates cells in a dose, wavelength and cell model dependent manner. The higher the dose and longer the wavelength, the more damage is inflicted on cells. The more stressed cells are, the better they respond to the laser irradiation. It is thus important to choose the correct laser parameters, otherwise results may be interpreted as negative or having no effect due to the incorrect factors. Although a dose of 16 J/cm^2 may not appear to be very high, it must be remembered that cells are irradiated within a confined environment. The energy absorbed by the cells cannot be dispersed as occurs *in vitro* [14]. Cells irradiated at 1064 nm showed the most damage compared to cells irradiated at shorter wavelengths. This damage may be due to the generation of ROS above the threshold [24]. This difference may also be due to the different output powers of the different lasers, as lower irradiances have been found to be more effective [25]. Visible and near-infrared light is absorbed by different chromophores, or light absorbing components within the cell. Visible light has been found to directly stimulate the mitochondria, while near-infrared light affects membrane channels [28]. Cells irradiated at 632.8 nm and 830 nm was more effective than the same cells irradiated at 1064 nm. The time of incubation post-irradiation can also play a part in the results seen, and should be chosen with care [27,29] as an increase can initially be seen, but with time can change to a decrease.

LLLT is able to stimulate cells in a positive manner and has been shown to be effective in wound healing *in vivo* [8]. The underlying effects for this healing are complex and intertwined. LLLT stimulates mitochondrial activity [22], which stimulates secondary mechanisms which ultimately leads to normalisation of cellular function. When using the correct parameters, it protects cells from apoptosis and damage [25]. It has an anti-inflammatory effect [1,16] and directs cells into a cell survival pathway.

Acknowledgements

This work was supported by grants from the University of Johannesburg (UJ), National Research Foundation (NRF), Medical Research Council (MRC), National Laser Centre (NLC), and African Laser Centre (ALC). All work was conducted by students and staff in the Laser Research Centre. Lasers were supplied and set up by the NLC.

5. References

[1] Houreld N N, Sekhejane P R and Abrahamse H 2010 *Lasers Surg Med*. **42** 494
[2] Pendsey S P 2010 *Int J Diabetes Dev Ctries* **30** 75
[3] Trautner C, Haastert B, Giani G and Berger M 1996 *Diabetes Care* **19** 1006
[4] Peppa M, Stavroulakis P and Raptis S A 2009 *Wound Repair Regen* **17** 461
[5] Hawkins D, Houreld N and Abrahamse H 2005 *Ann. N.Y. Acad. Sci.***1056** 486

[6] Hamblin M R and Demidova T N 2006 *Mechanisms for Low-Light Therapy:* Proc SPIE (San Jose, California, USA, 22-24 January 2006) ed M R Hamblin and R W Waynant *et al* vol 6140 (Washington; USA) pp 614001-1 – 12

[7] Caetano K S, Frade M A C, Minatel D G, Santana L A and Enwemeka C S 2009 *Photomed Laser Surg.* **27** 111
[8] Minatel D G, Frade M A C, França S C and Enwemeka C S 2009 *Lasers Surg Med.* **41** 433
[9] Moore P, Ridgway T D, Higbee R G, Howard E W and Lucroy M D 2005 *Lasers Surg Med.* **36** 8
[10] Lim W B, Kim J S, Ko Y J, Kwon H, Kim S W, Min H K, Kim O, Choi H R and Kim O K 2011 *Lasers Surg Med*. **43** 344
[11] Santos N R S, Sobrinho B, Almedia P F, Ribeiro A A, Cangussú M C T, dos Santos J N and Pinheiro A L B 2011 *Photomed Laser Surg* **29** 177
[12] Enwemeka C S, Williams D, Enwemeka S K, Hollosi S and Yens D 2009 *Photomed Laser Surg* **27** 221
[13] Hawkins D and Abrahamse H 2005 *Photomed Laser Surg.* **23** 251
[14] Houreld N and Abrahamse H 2007 *Photomed Laser Surg.* **25** 78
[15] Zungu I L, Hawkins Evans D, Houreld N and Abrahamse H 2007 *AJBR* **1** 60
[16] Sekhejane P R, Houreld N N and Abrahamse H 2011 *Photomed Laser Surg.* **29** 521
[17] Hawkins D and Abrahamse H 2006 *Lasers Surg Med.* **38** 74
[18] Zungu I L, Mbene A B, Hawkins Evans D, Houreld N and Abrahamse H 2009 *Lasers Med Sci.* **24** 144
[19] Houreld N and Abrahamse H 2007 *Diabetes Technol Ther.***9** 451
[20] Houreld N and Abrahamse H 2007 *Photomed Laser Surg.* **25** 474
[21] Houreld N and Abrahamse H 2010 *Diabetes Technol Ther.* **12** 971
[22] Zungu I L, Hawkins Evans D and Abrahamse H 2009 *Photochem Photobiol*. **85** 987
[23] Hawkins D and Abrahamse H 2008 Photodermatol, Photoimmunol Photomed **24** 199
[24] Houreld N and Abrahamse H 2007 *Laser Chemistry* Article ID 80536
[25] Houreld N and Abrahamse H 2008 *Lasers Med Sci.* **23** 11
[26] Mbene A B, Houreld N and Abrahamse H 2009 *J Photochem Photobiol B.* **94** 131
[27] Hawkins D and Abrahamse H 2007 Journal Laser Appl. **19** 74
[28] Almeida-Lopes L, Rigau J, Zaˆngaro R A, Guidugli-Neto J and Jaeger M M 2001 *Lasers Surg Med* **29** 179
[29] Hawkins D and Abrahamse H 2007 J Photochem Photobiol B: **88** 147

Resin phantoms as skin simulating layers

Aletta Elizabeth Karsten[1,2], Ann Singh[1], Ivy Mankekolo Ndhundhuma[1]

[1]National Laser Centre, CSIR, P O Box 395, Pretoria, South Africa, [2]Department of Physics, University of Pretoria, Pretoria, 0002, South Africa.

E-mail: akarsten@csir.co.za

Abstract. In order to apply light treatment to skin, the absorption through the outer layers of the skin needs to be considered. Darker skin has a higher concentration of melanin in the epidermis and absorbs more light than fair skin. Ideally the effect of the skin treatment on the outer layers of the skin should be tested on *in vitro* multi layer skin models. This is not always feasible. For this work, phantoms were used together with skin cancer cells to test the effect of outer layer absorption on the efficiency of Photodynamic Therapy (PDT) treatment. Two resin based solid phantoms were prepared to simulate two different skin types. Cells were prepared and PDT treatment were done on cells with and without the phantoms, by keeping the total dose delivered to the cells constant at 4.5 J/cm^2. Cell viability for the cells with the phantoms was less than without the phantoms and the differences are attributed to more uniform light distribution, but this needs to be investigated in more detail. The initial results of the experiments indicate that solid resin based phantoms can be used to optically mimic the effect of the outer skin layers.

1. Introduction

Lasers as both a research and application tool have been established in numerous fields in the last 50 years. The use of lasers in medical applications did not stay behind the other fields. As early as 1964, the ophthalmological applications of lasers were reported on [1].

Human skin is a highly scattering medium and the melanin in the epidermal layer of the skin is a major absorber of light in the visible and near infrared wavelengths. The human skin consists of several layers. In text books on light propagation through tissue [2], four layers are specified: stratum corneum (SC), epidermis, dermis and hypodermis.

The SC also called the horny layer consists of dead cells while the epidermis consists of melanosomes and melanocytes which produce the melanin in the skin that is responsible for skin colour. Most non-invasive laser treatment or diagnostic procedures need to penetrate the outer skin layers (SC and epidermis) to reach the treatment site. The melanin in the epidermis makes it a highly absorbing layer.

Melanin is an optically dense material which absorbs radiation in the visible wavelength range. Melanin is not a single pigment, but is composed of a number of different chromophores with varying optical and physical properties [3].

Darker skin has a higher concentration of melanin in the epidermis and absorbs more light than fair skin. Ideally the effect of the skin treatment on the outer layers of the skin should be tested on *in vitro* multi layer skin models before treatment. This is not always feasible.

PDT is a treatment where a photosensitezer (PS) or drug is applied to the cancer cells and in the presence of light, tuned to the absorption wavelength of the PS, and the oxygen present in the cells,

singlet and triplet oxygen are formed. The singlet and triplet oxygen is considered to be reactive oxygen species (ROS) that are lethal to the cells. The treatment is a localized treatment due to the short lifetimes of the singlet and triplet oxygen. Photosense®, which is a mixture of sulfonated aluminium(III) phthalocyanines, is clinically used in Russia for the treatment of a range of cancers [4].

In this paper the suitability of using solid skin simulating phantoms to mimic the effect of the outer skin layers were investigated by testing the efficacy of PDT treatment on cancer cells with and without the phantoms.

2. Materials and Methods

In this section the phantom preparation, the computer model used to predict the transmission through the sample and the cell work will be discussed.

2.1. Skin simulating phantoms

Skin simulating phantoms were prepared according to the recipe of Firbank [5,6]. Scattering particles (TiO particles) and absorbing particles (Carbon black) were added to an optically clear resin base (Akasel) and a hardener (Aka-cure slow, Akasel). The mixture was poured into plastic containers with a diameter of 30 mm and was allowed to cure at room temperature under a fume hood for 24 h.

The cured solid phantoms were cut into three 1 mm thick discs (A_1(I)-A_1(III)) to have a phantom for each of the 3 wells with cells that will be irradiated. The amount of Carbon black was varied while the TiO concentration was kept constant. The ratios (per weight) were 133:0.25:31000 (TiO: Carbon black:resin) for the A_1 samples and 133:0.75:31000 for the A_3 samples. Both the absorption coefficient, μ_a, and reduced scattering coefficient, μ_s', were measured with the integrating sphere (IS) system described in [7].

2.2. Computer modelling

Modelling the interaction of laser light with human tissue is an area of research in many parts of the world. Accurate modelling allows the researcher to test different extremes without the cost of clinical trials, even though it can never replace the value and importance of clinical trials. A model was developed in the ASAP software environment and tested to determine the accuracy and validity of the model. In the model the optical properties, μ_a, μ_s', refractive index (n=1.4) and the anisotropy (g=0.79), are described for each layer or medium. A light source is specified and 3.1 million photons are traced, making use of a Monte Carlo process, through the system. The system is divided into thin layers (nominally 0.1 mm) for analysis. The number of photons absorbed and passing through each layer is calculated and stored for later analysis. More details can be found in [6].

The optical properties of the phantoms were measured with the IS system, the results were used in the computer model to predict the energy fluence at the bottom part of the model (light that was transmitted through the sample). More detail on this study at λ = 632.8 nm can be found in [8].

2.3. Cell tests

For this paper Photosense®, a commercially available and approved PS for some skin cancers, was obtained from the Chemistry Department of Rhodes University (courtesy of Prof T Nyokong). Stock solutions of 1 mg/ml sensitizer was made up in Dulbecco's modified Eagle's medium, (DMEM) with phenol red and sterilized by filtration using a 0.2 μm filter.

2.3.1. Cell culture

Cultures of human squamous cell carcinoma (SCC) cells (A431 cell-line, obtained from Sigma-Aldrich, South Africa, Cat # 85090402) were grown in the supplemented culture medium (DMEM) with L-glutamine and phenol red, supplemented with 10 % heat-inactivated fetal calf serum (FCS) and 1 % penicillin-streptomycin mixture (5K/5K) obtained from Whitehead Scientific (Pty) Ltd, South Africa, at 37 °C and 5 % CO_2.

2.3.2. Cytotoxicity test

The optimal concentration of Photosense® and diode laser dose were determined by incubating the cells with various concentrations of the PS (10, 20, 40 and 80 µg/ml) for 24 h or irradiating cells with 676 nm diode laser with a power density of 39,1 mW/cm^2 and irradiation at 0.5-5 J/cm^2 for 12,8-127.9 s, respectively. The laser used was built at the NLC from commercially available parts. The laser beam from the diode had a rectangular shape and was shaped through a lens to a square and an iris was used to deliver a round beam on the cells with an area of 1.2 cm^2 and a total power of 49.2 mW.

Cells were seeded into 24-well tissue culture plates at a density of 50 cells/mm^2 per well in 1 ml of supplemented DMEM with phenol red. Cells were allowed to attach for 48 h before being washed twice with 2 ml Dulbecco's phosphate buffered saline (DPBS, from Whitehead Scientific (Pty) Ltd, South Africa,) then photosensitized by the addition of culture medium containing 10, 20, 40 and 80 µg/ml Photosense®. Control cells contained medium without Photosense®. Plates were incubated at 37 °C in 5 % CO_2 in the dark for 24 h, the wells were then washed twice with 2 ml DPBS and the medium replaced with 1 ml of culture medium. The loss of sensitizer from cells to medium was not observed.

Surviving cells were quantified after re-incubation with culture medium with the use of Trypan blue dye exclusion viability assay (Sigma-Aldrich, South Africa) after 48 h.

2.3.3. Phototoxicity test

The effects of the phantoms were determined by incubating the cells with 10 µg/ml Photosense® for 24 h. Cells were seeded into 6-well tissue culture plates at a density of 50 cells/mm^2 in 3 ml of culture medium. Cells were allowed to attach for 48 h before being washed twice with 2 ml DPBS, then photosensitized by the addition of culture medium containing 10 µg/ml Photosense®. Control wells contained medium without Photosense® and 0 J/cm^2 diode laser irradiation (no laser irradiation). Positive control wells contained medium with Photosense® and 0 J/cm^2 diode laser irradiation. Plates were incubated at 37 °C in 5 % CO_2 in the dark for 24 h, the wells were then washed twice with 2 ml DPBS and the culture medium replaced with 3 ml of DPBS. Wells containing experimental cells (cells treated with Photosense®) were irradiated either with a 676 nm diode laser alone or with the diode laser in the presence of phantoms with a power density of 41 mW/cm^2 and irradiation at 4,5 J/cm^2 for 14 min (laser alone) or 32 min in the presence of phantoms, respectively. The DPBS was then replaced with 3 ml of fresh supplemented culture medium and the preparation returned to the incubator for a further 48 h. Surviving cells were quantified after re-incubation with culture medium with the use of Trypan blue dye exclusion viability assay after 48 h.

2.3.4. Changes in cell viability

Change in morphology and viability of cells (controls and experimental) was assessed using an inverted microscope. A trypan blue dye, cell viability assay reagent, based on the principle that live cells possess intact cell membranes that exclude trypan blue was used. The dye exclusion test is used to determine the number of viable cells present in a cell suspension. The assay is performed by mixing 50µl trypan blue reagent and 50 µl of cell suspension then examine cells visually under the microscope to determine whether cells take up or exclude the dye. A viable cell will have a clear cytoplasm whereas a nonviable cell will have a blue cytoplasm. Cells were counted within 3 to 5 min of mixing cells with trypan blue dye.

3. Results

3.1. Phantom results

The optical properties of the two different phantom sets were measured on the IS system. The results of the A_1 samples (very low concentration of Carbon black), could not be used for the computer model due to the inaccuracies of the results. These values falls outside parameters for the IS model.

The values for the A_3 samples are given in Table 1. All the phantoms had a diameter of 30 mm and a thickness of 1 mm.

Table 1: Optical properties of phantom A_3. Samples cut into the 1 mm thick discs.

Phantom name/Property	μ_a (mm^{-1}) (λ = 676 nm)	μ_s (mm^{-1}) (λ = 676 nm)
A_3(I)	0.15	17.89
A_3(II)	0.12	14.86
A_3(III)	0.12	15.35

3.2. Computer model predictions

The results in Table 1 were in used in the computer model, simulating a single phantom layer A_3. The fraction of the light that were not absorbed or scattered in the back-ward direction (transmitted light) was compared with the results from the IS measurements and are tabulated in Table 2. In previous studies [7,8] we showed that the comparison between the IS measurement and the computer model predictions is less than 10% at λ=632.8 nm. For this study we used a wavelength of 676 nm (the absorption wavelength of the Photosense®).

Table 2: Comparison between IS and computer model for the light reflected back from the sample (R) and the light transmitted through the sample (T).

Phantom	R(IS)	T(IS)	R(model)	T(model)	% error in R	% error in T
A_3(I)	0.39	0.29	42.79	21.37	11.04	25.36
A_3(II)	0.38	0.33	41.88	26.12	9.05	21.65
A_3(III)	0.39	0.33	42.19	25.45	9.14	21.78

Due to the inaccuracies in the IS model and therefore the prediction values, the actual transmitted values through the phantoms (all 6) were measured and used to calculate the increased treatment time of the cells to ensure the cells with the phantoms and those without the phantoms receive the same total dose during the PDT treatment.

3.3. Cell tests

3.3.1. Cytotoxicity tests

Cytotoxicity of Photosense®, as measured by quantifying surviving cells using the trypan blue dye exclusion reagent viability assay 48 h after the applied treatment, was found to be significant. Cytotoxicity tests showed that at 80 µg/ml Photosense® only 27% of the cells survived, however, with 10 and 20 µg/ml more than 60 % of cells survived (Figure 1). The higher the concentration of Photosense®, the higher is the loss of cell viability of SCC cells.

Upon irradiation with diode laser at 676 nm, viability of cells, as shown in Figure 2, increased in a dose dependant manner. Accelerated levels of cell growth were observed with light dose at 5 J/cm^2 laser irradiation. The results are in support of previously reported data [9] that laser irradiation at certain fluences and wavelengths can enhance the release of growth factors from cells and stimulate cell proliferation *in vitro*. As another control to prove that the phantoms do not have any adverse effect on the cells, cells were incubated with phantoms for a maximum period of 2 h to evaluate any loss of cell viability with phantoms and no effect on cell viability was observed.

3.3.2. Photodynamic effect

PDT is an established cancer treatment modality that involves the combination of visible light and a PS. Each factor is harmless by itself, but when combined, in the presence of oxygen, can produce lethal cytotoxic agents that can destroy tumour cells [10]. The combination of 10 µg/ml Photosense® with 4,5 J/cm^2 diode laser at 676 nm, was phototoxic to the SCC cells. Upon irradiation of cells with

PS combined with laser in the absence of phantoms (laser only) or PS combined with laser in the presence of phantoms, Phantom A1 and A3 (Figure 3), a more significant decrease in cell viability was observed. Viability studies have shown that phantoms mimic the outer skin layer, hence it takes longer to irradiate cells in the presence of phantoms to achieve phototoxic effect as obtained with laser only.

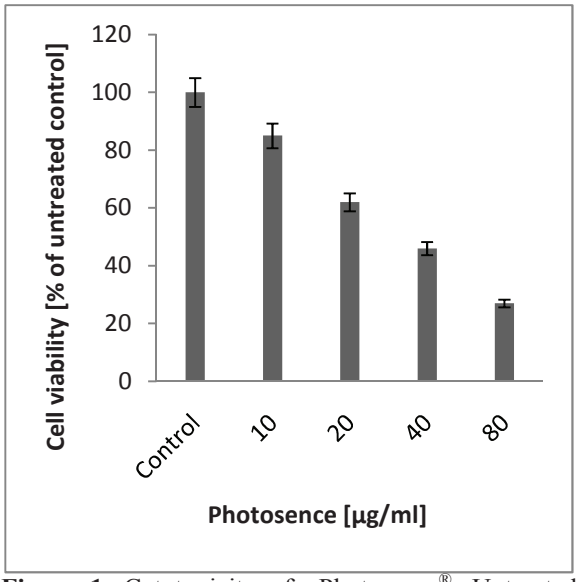


Figure 1. Cytotoxicity of Photosense®. Untreated SCC cells (control; 0 µM of Photosense®) were compared with those treated with 10–80 µg/ml of Photosense®.

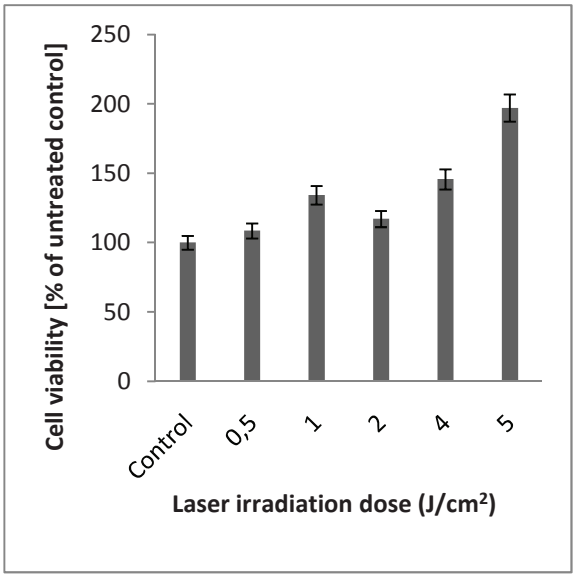


Figure 2. Cytotoxicity test of 676 nm diode laser. Untreated SCC cells (control; 0 J/cm^2) of diode laser were compared with those treated with 0, 5 – 5 J/cm^2.

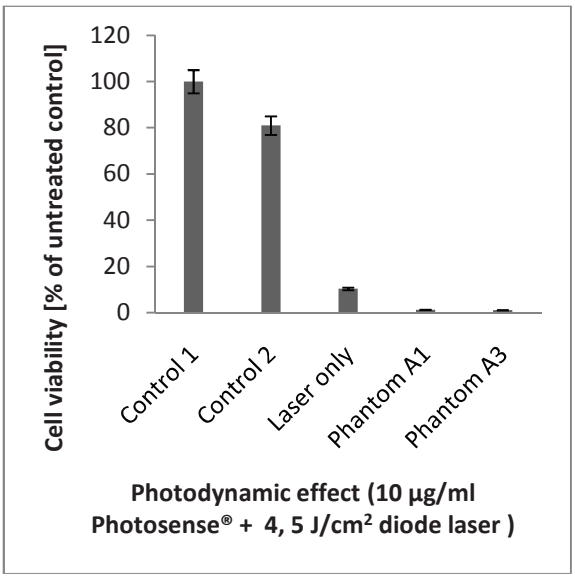


Figure 3. Photodynamic effect. Untreated SCC cells (without Photosense® and un-irradiated, Control 1) were compared with those treated with 10 µg/ml of Photosense® and irradiated with 4,5 J/cm^2. Further control: SCC cells un-irradiated but treated with 10 µg/ml Photosense® (Control 2).

4. Discussion and Conclusions

In this work, the suitability of solid phantoms as a replacement for the outer skin layers was investigated. The phantoms were measured with an IS system to obtain the optical properties. These properties can be used in a computer model to predict the fluence that is transmitted through the model. Even thought the IS model is not as accurate as one would desire, it can be used for predictions, but needs to be refined to allow for data with very low absorption coefficients.

The cell work proved that the phantoms do not have any adverse effect on the cells in the absence of any treatment parameters. Similar results are obtained when the phantom is placed on the cells to mimic the skin layers and the treatment times are adjusted to allow for the same dose of treatment. This is the first experiment to the authors' knowledge where such a system is used to mimic human skin. The experimental procedures will still have to be refined for more conclusive results.

References

[1] Niemz M H 2007 *Laser-Tissue Interactions: Fundamental applications* (Springer)
[2] Tuchin V 2000 *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnostics* (SPIE Press, pp 3-108)
[3] Alaluf S, Heath A, Carter N, Atkins D, Mahalingam H, Barrett K, Kolb R and Smit N 2001, *Pigment Cell Res* **14** 337
[4] Chi-Fung C, Jian-Dong H, Pui-Chi L, Wing-Ping F and Dennis K P 2008, *Organic & Biomolecular Chemistry (Org. Biomol. Chem)* **6** 2173
[5] Firbank M and Delpy D T 1993, *Phys. Med. Biol*. **38** 847
[6] Karsten A E, Singh A and Braun M W 2011, *Lasers Med Sci*, DOI 10.1007/s10103-011-0926-x (published online April 2011).
[7] Singh A, Karsten A E and Dam J S, *Proc. of the World Association of Laser Therapy (WALT)*, 2008
[8] Karsten AE and Singh A 2011, *Proc. 4th International Symposium on Photoelectronic Detection and Imaging (ISPDI 2011*)
[9] Hawkins-Evans, D and Abrahamse, H 2008, *Photodermatology, Photoimmunology & Photomedicine* **24** 199
[10] Sharman, W M, Allen, C M and van Lier J E 1999, *Drug Discovery Today* **4** 507

Enhancement of hydrogen production using biomass gasification process

A. I. Mabuda, N. S. Mamphweli and E. L. Meyer
University of Fort Hare, Institute of Technology
Private bagx1314, Alice, 5700, South Africa

E-mail: amabuda@ufh.ac.za

Abstract. Hydrogen production from biomass holds the greatest promise, since biomass is abundantly available everywhere in the world. However, hydrogen from biomass has major challenges. The yield of hydrogen is low from biomass since the hydrogen content in biomass is low to begin with (approximately 6%) and the energy content is low due to the 40% oxygen content of biomass. A novel gasification method for hydrogen production from carbonaceous materials using a CO_2 sorbent has been widely used. It mainly uses steam as gasification agent. For this study the above method has been adopted to test if it will work for air-blown biomass gasifiers. The main purpose of this project is to enhance the yield of hydrogen from air-blown biomass gasification process. The produced hydrogen will be further separated and purified for fuel cell application. Ultimate and proximate analyses of the biomass material were conducted and the obtained results were used for the simulations in order to determine the efficiency of the gasifier with biomass and biomass/sorbent blends. It was found that the biomass/sorbent blends increase the yield of not only H_2 but also other syngas constituents such as CO leading to enhancement of the gasifier efficiency since it is dependent on the volume of combustible gases.

Introduction

Hydrogen production has required much attention in recent years due to its environmental friendly as an energy carrier. It is produced via steam methane reforming (SMR), coal gasification, oil reforming, and electrolysis. Fossil fuel sources amount to 96% [1] of total H_2 production. Therefore, clean and renewable resources and sustainable pathways are necessary if H_2 is to become a fundamental energy resource for the future [2]. The sources of hydrogen should be renewable and sustainable, efficient and cost effective, convenient and safe. Hydrogen production based on renewable energy sources such as biomass is of great interest since biomass is carbon-neutral. However, biomass has, on average, only 6 wt % of hydrogen, which would make it, in principle, of not much interest for H_2 production [3].

Some additional H_2 can come from the cracking of H_2O if the biomass is also used as a fuel to supply the high amount of heat needed to crack the molecule of H_2O [3]. Since this raw gasification gas contains some CH_4, tar, and light hydrocarbons, when a bed of calcined dolomite (CaO.MgO), limestone (CaO), and/or magnesite (MgO) can be used downstream from the gasifier, the H_2 content in the gasification gas was increased [4].The use of CaO sorbent for enhancing hydrogen yield could be applied to any carbonaceous material such as coal, hydrocarbon or biomass. The product gas through

rich hydrogen also contains other gases such as CO, CH_4 and some hydrocarbons [5]. The following reaction takes place when calcium oxide has been employed:

$$CaO + CO_2 \rightarrow CaCO_3 \quad (1)$$

The mixture of biomass with CaO sorbent loaded to the gasifier absorb the released CO_2 to enhance the water gas shift reaction towards hydrogen production.

$$CO + H_2O \rightarrow CO_2 + H_2 \quad (2)$$

It also provides the necessary energy for the endothermic gasification through the carbonation reaction (equation 1) releasing heat. This method was used for air blown instead of steam gasification. The steam in this case comes from the vaporization of moisture in the feedstock. It is necessary to get rid of as much of this moisture as possible to avoid the production of condensates inside the gasifier during operation and after shut down.

Research methods

Various blends of Calcium oxide and or magnesium oxide with pinewood were prepared. A Carbon Hydrogen, Nitrogen and Sulphur (CHNS) Analyzer was used to determine the elemental composition of the various blends for the purpose of gasification simulation in order to predict the mass and energy balance of the gasifier when fuelled with the various blends. A cone calorimeter (CAL2K model) was used to determine the calorific value of the various blends. A thermogravimetric (TGA) analyzer was used to establish the reaction temperature of the various blends in order to establish a suitable gasification temperature and the thermal stability of the materials. The moisture, ash as well as fixed carbon content of the material was determined from the TGA data. A downdraft gasifier modeling program developed by T.H. Jayah [6] was used to undertake gasification simulation.

Results and discussions

Ultimate and proximate analysis

Figure 1 shows the composition of carbon, hydrogen and oxygen of biomass and biomass/sorbents materials. The results presented in figure 1 were obtained using the CHNS analyser; the oxygen was obtained by difference. The Sulphur and nitrogen content are not presented in the figure because they are of insignificant impact on gasification.

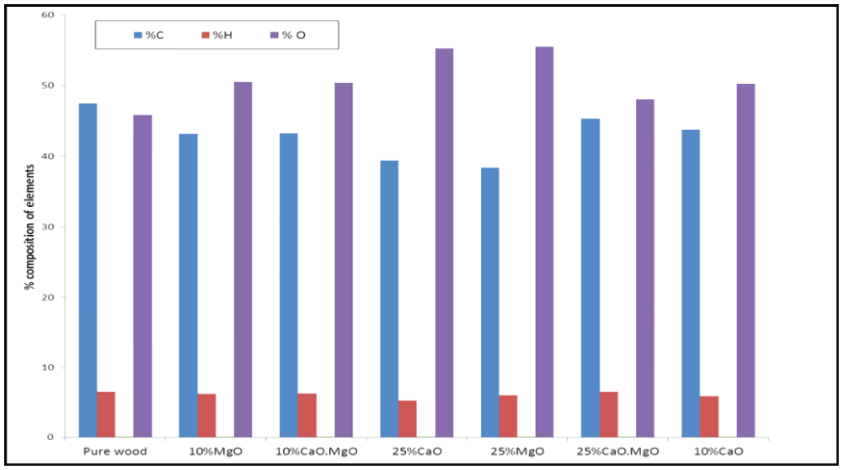


Figure 1. The composition of carbon, hydrogen and oxygen of biomass and biomass/sorbents materials.

The addition of the sorbents did not have a significant impact on the elemental structure of the pinewood. A 5-8% decrease in carbon content for the blends as compared to the pure wood can be seen in figure 1 with a corresponding increase in oxygen content for the blends. The hydrogen content however remained fairly constant in both pure wood and the blends; this should not be confused with the impact of these blends on hydrogen production because the production of hydrogen during gasification is enhanced by chemical reactions as discussed in section 1.

Reaction kinetics

Figure 2 shows the rate of thermal degradation of the various wood/sorbents blends. Figure 3 shows the rate of thermal degradation for a blend of 25% CaO and 75% pinewood. The two graphs were obtained using TGA after calculation of the derivatives.

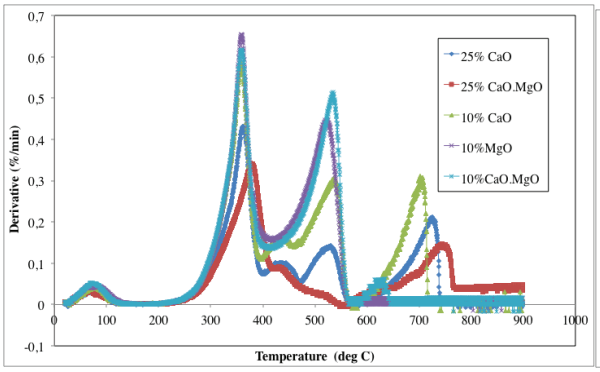


Figure 2. The rate of degradation for various sorbents/wood blends.

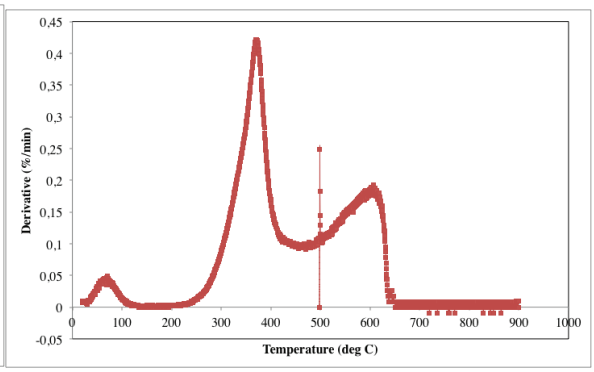


Figure 3. The rate of thermal degradation for 25% MgO and 75% wood blend.

The initial stage of degradation in both figure 2 and 3 involves the vaporization of water at around 100°C. Then the reaction rate picks up at around 250°C when carbonization or devolatization of the material starts to take place resulting in the formation of charcoal. The second stage at about 350 °C is due to dehydration of $Ca(OH)_2$ and at about 650 °C can be attributed to the carbonation of $CaCO_3$. The peaks, especially calcium, are mainly in the form of hydroxides and carbonates, which during heating

in the TGA apparatus decompose releasing H_2O and CO_2. For the MgO and CaO.MgO samples, the decomposition peaks appear at temperatures (250 and 400 ^oC) which indicate an effect of MgO in the mixture as it promotes decarbonation and dehydration at lower temperatures [7]. The two distinct mass loss stages (Fig 3) of CaO, one appeared at the temperature (250 and 400 ^oC) is due to the CaO carbonation reaction (Eq. (1)), through which CaO can absorb the released CO_2 to form $CaCO_3$ product. The peak appears at temperature about 600^{o}C corresponding to the second weight loss. The previous studies showed that this is due to the degradation of carbonates [8].

The rate is faster between 300^{o}C and 600^{o}C for most of the materials and it slows down at higher temperatures above 600^{o}C. This is basically because at that point the remaining charcoal is converted to gas and ash. The charcoal consists of the major part of the carbon and therefore requires more energy to break down over time hence the reaction rate slows down. The carbon content is almost the same for both 25% MgO and 25% CaO blends hence the reaction rates at lower temperature resembles the same pattern. However there is a significant difference in the reaction rates for the two blends at higher temperatures (above 600^{o}C). The main cause of this is still under investigation. It is clear from figures 2 and 3 that the addition of the sorbents results in lower reaction rates, the high concentration of MgO for instance lowered the reaction rate from approximately 0.6%/min to approximately 0.4%/min. This is because of the low volatile matter content in the sorbents as compared to the pinewood.

Figure 4 and 5 show the volume of hydrogen obtained during computer simulation of the gasification of the various sorbents/wood blends, and the possible conversion efficiency that could be achieved when gasifying the various blends. The figures were generated using a downdraft gasifier modelling program mentioned in the methodology section. The input parameters were obtained from the proximate and ultimate analysis of the various blends. The simulations were undertaken assuming the same gasifier operating conditions.

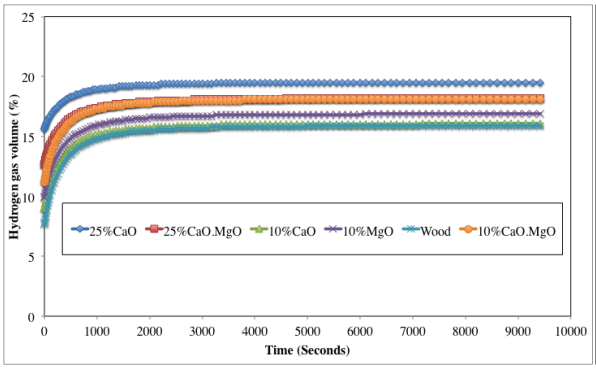


Figure 3. The hydrogen gas volume obtained during computer simulation.

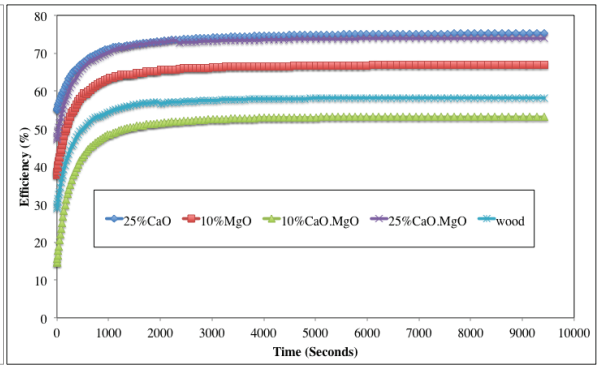


Figure 4. The simulated possible efficiency achieved when gasifying the various blends.

It can be observed that the highest hydrogen content was obtained from a blend with 25% CaO and 75% pinewood followed by the blend with a mixture of 50% CaO and MgO constituting 25% of the CaO/MgO/pinewood blend as well as the blend with a mixture of 50% CaO and 50% MgO constituting 10% of the CaO/MgO/pinewood blend. The 25% CaO blend also resulted in higher gasifier conversion efficiency followed by the blend with a mixture of 50% CaO and MgO constituting 25% of the CaO/MgO/pinewood. This is because of the higher hydrogen content produced by the two blends respectively as indicated in figure 1. The efficiency of the gasifier is dependent on the volume of combustible gases, which are hydrogen, methane and carbon monoxide.

Conclusions

The paper presented ways in which hydrogen production from an air-blown fixed bed downdraft gasifier can be enhanced without using steam as a gasification agent. It demonstrated the possibility of taking advantage of the steam produced during the drying of the feedstock and converts it to useful hydrogen through the water-gas-shift reaction enhanced by the addition of sorbents into the process. Computer simulation results suggested that a blend of 25%CaO/75% pinewood results in higher hydrogen content produced during gasification and enhanced gasification efficiency.

References

[1] Ewan B C, Allen R W K 2005 *Int. J. Hydrogen Energy* **30** 809 – 819.
[2] Florin N H, Harris A T 2008 *Chem. Eng. Sc* **63** 287–316.
[3] Aznar M P, Caballero M A, Corella J, Molina G, and Toledo J M 2006 *Energy & Fuels* **20** 1305-1309.
[4] Delgado J and Aznar M P 1997 *Ind. Eng. Chem. Res.* **36** 1535-1543.
[5] Madhukar R M, Goswami D Y 2007 *Int. J. Hydrogen Energy* **32** 2803 – 2808.
[6] Jayah, T. H., L. Aye, R. J. Fuller and D. F. Stewart 2003 *Biomass and Bioenergy* **25** 459-469.
[7] Alarcón N, García X., Centeno M.A, Ruiz P, Gordon A 2004 *Applied Catalysis A: General* **267** 251–265.
[8] Biagini E., Barontini F., and Tognotti L 2006. *Ind. Eng. Chem. Res., **45*** 4486–4493.

Investigating the thermal performance of a hybrid photovoltaiv solar system

B Mtunzi, EL Meyer and M Simon

University of Fort Hare, Private Bag x1314, Alice 5700, South Africa

bmtunzi@ufh.ac.za

Abstract. Photovoltaic thermal heating systems have gained momentum in the recent years. Many investigations being done are to improve the performance of the module by way of removing the waste heat at the back of the module. In this research water in direct contact with the back of module was used to cool the photovoltaic module through thermosyphon effect. The findings indicated an improved electrical efficiency of the photovoltaic module under water as compared to one under normal operating conditions. The larger implications of the findings indicate heat utilization and an improved performance of the photovoltaic module.

1. Introduction

A photovoltaic (PV) hybrid system produces both electricity and hot water simultaneously. In PV modules, as the module temperature increases, the efficiency of the module drops. In hybrid modules, the heat produced by the module is absorbed by water and in the process decreases the cells' temperatures. In this way, PV efficiency is optimized as the heat is transferred into water for hot water production.

When the sun's rays fall onto the module, part of solar spectrum is converted to electricity and the other part heats the module. Infrared light does not contribute towards the electricity production of the module; instead it heats up the module [1].

It is a well-known fact that some solar cells experience significant efficiency degradation and a decreased lifetime when operated under high temperatures [2]. Under normal operation conditions, the back of the module's temperatures have been found to reach between 40°C and 70°C on sunny summer weathers [3]. Temperature levels of around 40°C have been noted to cause modules to lose up to 7% of their power [4]. In South Africa we have lots of solar radiation throughout the year and this has been found to range from 6000MJ/m2 to 9500MJ/m2 in different provinces in the country [5]. With no cooling, photovoltaic modules' energy production is compromised.

Several researchers have carried out experiments in trying to find out ways of cooling the modules so as to improve the output of the modules. Air, water and glycol have been used as coolants in different research studies. Water or glycol is circulated in copper pipes that are fixed on a copper plate attached to the back of the module. These pipes allow the fluid to circulate, while at the same time providing cooling effect on modules. The silicone gel is usually used to reduce the thermal resistance between the back of the module and the copper plate, thus enhancing heat transfer to cooling water.

Air cooling was used to cool the module [6]. In this research Tiwari et al used a fan to drive air across the module. At low air speeds, the cooling effect on the module was found to be more pronounced while at high air speed the cooling effect was found to be minimal. The heat utilisation of the air was not considered in this research. Air has been noted to have less thermal conductivity as compared to water; hence water would provide a better cooling effect when compared to air.

A photovoltaic hybrid system that involved the use of copper tubes fixed at the back of the module (risers) with water as a coolant was investigated [7]. In this research pumps were used to circulate water in these risers to effect heat transfer by force. Ibrahim et al investigated the cooling effects of different absorber designs on the performance of PV thermal collectors [8]. These researchers had

seven different designs checked and simulations were carried out on all the designs. Based on simulations, the spiral flow design proved to be the best design and was found to give a thermal efficiency of 50.12% and corresponding cell efficiency of 11.98%. The spiral design had tubes wound in such a way that the tubes touched each other allowing cross heat transfers.

The research on use of water as a cooling agent has been based on indirect cooling, where water is not in direct contact with the back of the module. This paper looks at the use of water in direct contact with the PV module. This form of cooling offers the opportunity to remove heat directly from the module with no intervening thermal conduction resistance, other than that of ethyl vinyl acetate (EVA) found at the back of the module. Direct water cooling offers a high heat transfer coefficient which reduces the temperature rise of the PV module. Water has light and heat trapping ability hence the ability to reduce solar-cell surface recombination velocity. A greater uniformity in temperature of the cell is attained in this way hence making them operate efficiently. A direct contact between the back of module and the water was considered and it was hoped that a new solar module cell-cooling method would be established.

1.1. Efficiency of a cell

The efficiency of a cell is defined as the ratio of energy output from the solar cell to input energy from the sun. This has been found to depend on the spectrum and intensity of the incident sunlight as well as the temperature of the solar cell [9]. The efficiency of a solar cell is determined as the fraction of incident power which is converted to electricity and is defined as:

$$\eta = \frac{P_{out}}{P_{in}} \quad (1)$$

Where: $P_{out} = V_{oc} I_{sc}\, FF$ (2)

And Fill factor is given by $FF = \frac{V_{peak}\, I_{peak}}{V_{oc} I_{sc}}$ (3)

$P_{in} = GA_{collector}$, (4)

Therefore the efficiency of the cell is: $\eta = \frac{V_{oc}\, I_{sc}\, FF}{P_{in}}$ (5)

Electrical efficiency of the photo voltaic cells has also been found to follow equation 6 [10].

$$\eta_{el} = \eta_o[1 - \beta(T - 25^oC)] \quad (6)$$

Where: η_o is the efficiency of the module at Standard Test Conditions (STC).

β is the coefficient of temperature and its value is equivalent to 0.0045 /°C for crystalline silicon cells; T is the temperature of the module.

1.2. Maximum Power Point

The maximum power point of a module may be determined through Thevenin and Norton models. According to the Thevenin and Norton equivalent circuits, the open circuit voltage (V_{oc})is obtained from Thevenin model and the short circuit current (I_{sc})of the module from the Norton model. However the solar cells have a load curve as shown in figure 1.

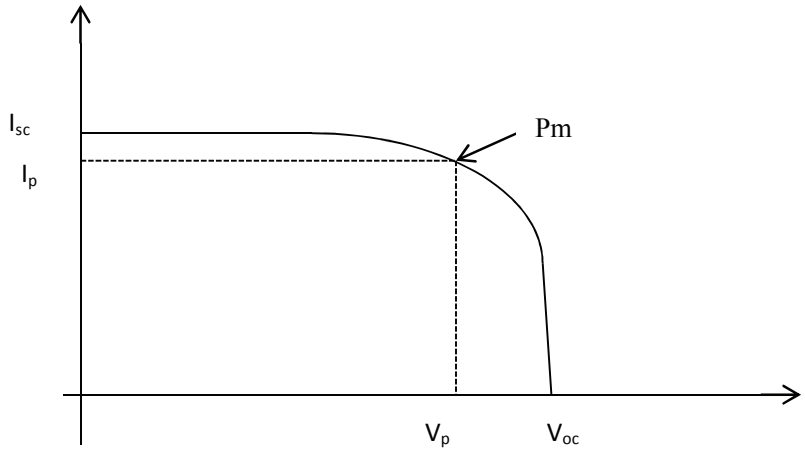


Figure 1: I/V characteristic curve.

The maximum power point of the photovoltaic module is found at Pm if I/V measurements are made on the module using a PVPM system. At maximum point , the current supplied is nearly the short circuit current.

2. The Experimental Method

SikaFlex-11FC sealant was used to provide good electrical-insulation between the module and the water coolant. A perspex box was constructed and fixed at the back of the module. The module with water box was named M2 and the other module M1 was used as a control. M1 was placed under normal operating conditions undergoing natural cooling. Modules used in the investigation were SW80 solar modules. The system setup was fixed on a module rack as shown in figure 1,

Figure 2: System set up, LHS M1 and RHS M2

The solar rays fall onto the module and produce electricity through photoelectric effect and at the same time the rays not used for electricity generation heats up the module. Heat generated at the back of the module is then transferred from the module to water in the container through conduction, radiation and convection. Through the principle of convectional currents, the solar heated water in the perspex box container becomes less dense and moves up to the storage tank. Due to density, cold water in the storage tank, replaces the warm water through the thermosyphon effect. The module gets cooled in the process and improves the electrical efficiency of the module. Current and voltage measurements were made on the module every 10 minutes and the corresponding maximum power point measurements were determined.

The back of the module's average temperature and the current-voltage curves were obtained.

3. Results and Discussion

Initially, the SW80 solar modules' I/V characteristics were measured and recorded to get the benchmark of the modules. Both modules were found to have similar I/V characteristics.

Both modules M1 and M2 were fixed on the PV rack with M2 initially having no water filled at the back of the module. The outdoor I/V measurements were taken and Figure 3 shows the response of the modules.

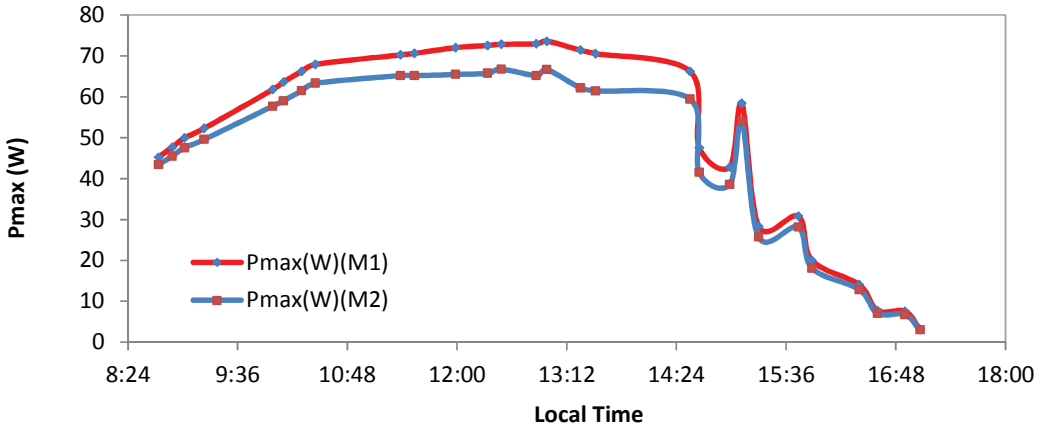


Figure 3: Maximum power output of modules M1 and M2

Using Thevenin-Norton models the maximum power points of the two modules were determined. M2's maximum power output at any given time was found to be lower than that for M1 due to back of module temperatures. The back of the module temperature for M2 was higher than that for M1 and thus M2 had a reduced open circuit voltage and a slightly higher short circuit current as compared to M1. A reduced open circuit voltage on M2 contributed towards reduced maximum power points. M1 had higher maximum power points due to natural cooling from ambient air.

Figure 4 illustrates the maximum power output of the two modules when M2 had water at the back of the module and M1 acting as a control undergoing natural cooling.

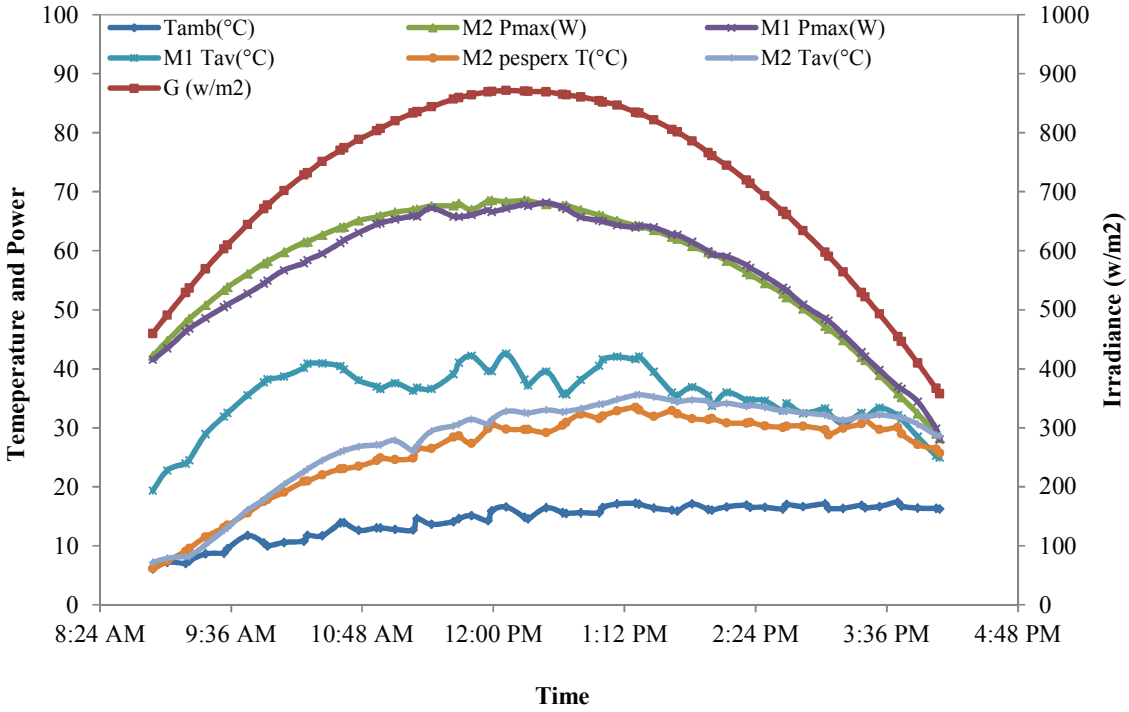


Figure 4: Maximum power out of modules M1 and M2 with M2 water cooled (28/06/2011)

From figure 4, it can be noted that due to water cooling M2 had a higher maximum power output for the better part of the day as compared to M1. Theoretically, when the photovoltaic modules' temperatures rises, the open circuit voltage V_{oc} falls and the short circuit current slightly increases shifting the maximum power point to the lower end and lower temperatures cause V_{oc} to increase. Cooling of module M2 enables the operation of the module at a higher V_{oc} and a slightly lower I_{sc}. This in turn brought about higher maximum power points. Higher power output meant an improved efficiency. Figure 5 illustrates the efficiency curves of the two modules.

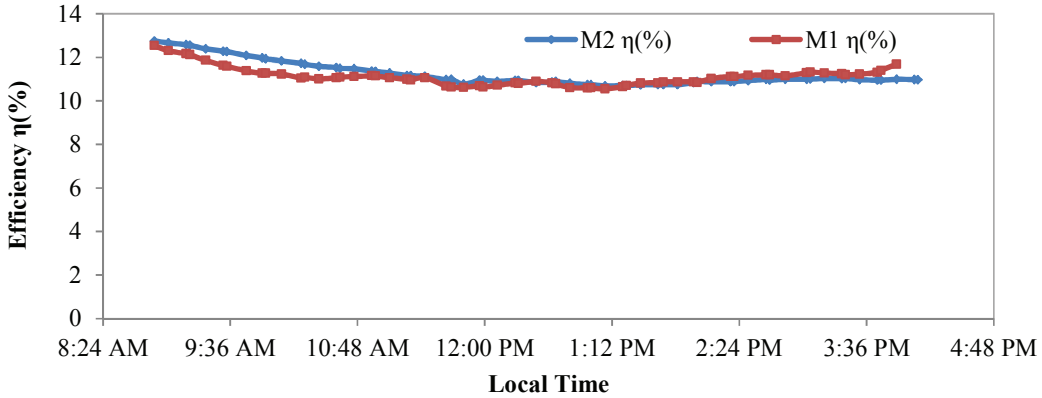


Figure 5: The effect of cooling water on efficiency.

From figure 5 an initial higher efficiency of the water cooled module M2 was noted. Initially an efficiency level of 12.75% was recorded for M2 and that for M1 was found to be 12.24%. A percentage difference of 4.2% between the two modules was noted. This difference was found to be due to the cooling effect of water on M2. Increase in module temperatures brought about a farther drop in efficiency by 1.75% for M2 and 1.24% for M1 and this was noted during the time period from 08:50 to 11:40 am. Both modules' efficiency values then remained at almost 11.00% for the rest of the day. Figure 6 further shows the relationship between the efficiency and irradiance for the two modules M2 and M1. Similarly M2showed a higher efficiency due to cooling effect of water.

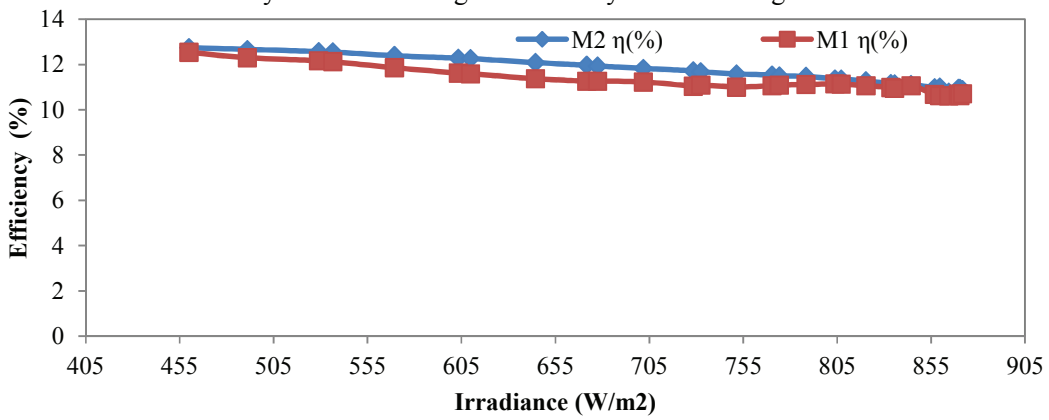


Figure 6: Efficiency of the module as compared to irradiance falling on the module

The parameters for the SW80 modules at solar noon on the 28th of June 2011 were also measured using the PVPM system and the results were as shown in table 1.

Table 1: Parameters of the SW80 modules.

	Voc (V)	Isc (A)	Pmax (W)	FF(%)	η_{ele} (%)
T_{SW80} =25.00°C	21.5	4.82	80.2	77.4	11.14
T_{M2} =33.72°C	20.65	4.81	66.96	67.35	10.77
T_{M1} =41.55°C	20.63	4.91	66.04	65.15	10.62

The measured values showed an average efficiency of approximately 11% for both modules. The efficiency value determined was close to the manufacturer's efficiency value for the SW80 modules. However M2 still showed a higher electrical efficiency as compared to M1 due to cooling effect of water. The module temperature was also found to be lower than that of M1.

The findings were that the average water output module temperature peaked at 31.9°C while the inlet temperature was at approximately 24°C. The ambient temperature was approximately 14.3°C and the irradiance on the tilted surface was 711.18 W/m^2.

Conclusion

Two modules were considered, M2 with water as a coolant and M1 under normal operating conditions was cooled by natural means. The water cooling effect was noted on the performance of the PV module. The water cooled module gave a higher power output as compared to the naturally cooled module for the better part of the day. Also the electrical efficiency of the water cooled module was found to be higher than that of naturally cooled module. This is likely to have a positive contribution towards reduction of the module degradation. Heat utilisation of excess heat at the back of the module was achieved on M2, while that at the back of M1 was lost into thin air. Water as a direct contact coolant on photovoltaic modules has been introduced.

References

[1] Honsberg C and Bowden S PVCD ROM Viewed on 10/06/2011 from http://pveducation.org/pvcdrom/modules/heat-generation-in-PV-module

[2] Machida K, Yamazaki T and Hirasawa T 1997 *J. Solar Energy Materials and Solar cells* **47** 149-153.

[3] Antony F, Durschner C and Remmers K H 2006 *Photovoltaics for Professionals (Berlin:Solar Praxis AG)*

[4] Gxasheka A R, van Dyka E E. and Meyer E L 2005 *J.Renewable Energy* **30** 611–620.

[5] AnnualSolar Radiation map for SouthAfrica Viwed on 07/03/09 http://www.sunpower.co.za/Incoming%20Radiation.jpg

[6] Tiwari A, Sodha M S, Chandra A and Joshi J C 2006. *J. Solar Energy Materials and Solar cells* **90** 175–189

[7] Anderson T.N, Duke M, Morrison G.L and Carson J.K 2009 *J. Solar Energy* **83** 445-455.

[8] Ibrahim A, Othman M.Y, Ruslan M H, Alghoul M A, Yahya M, Zaharim A, and Sopian K 2009 Performance of Photovoltaic Thermal Collector (PVT) With Different Absorbers Design. Viewed on 10/06/2011 http://www.wseas.us/e-library/transactions/environment/2009/31-968.pdf.

[9] Eikelboom J A and Jansen M A (2000). Characterisation of PV Modules of New Generations. viewed on 20/04/2011from http://www.ecn.nl/docs/library/report/2000/c00067.pdf

[10] Skoplaki E and Palyvos J A 2009 On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/power correlations *J. Solar Energy* **83** 614-624

Batch anaerobic co-digestion of cow dung and donkey manure

P Mukumba[1], G Makaka[1], S Mamphweli [2]
University of Fort Hare, [1]Physics Department, [2]Institute of Technology, P Bag X1314, Alice 5700, South Africa

E-mail Address: patrickmukumba@gmail.com

Abstract. Biogas from anaerobic digestion can be a solution to current and future energy needs in South Africa. One option for improving biogas yield of anaerobic digestion of organic matter is co-digestion. Cow dung and donkey manure were co-digested together. The co-digested experiments were conducted using 1.0 m^3 batch biogas field digesters that were designed and constructed. The biogas volume of biogas produced was measured daily by a biogas flow meter. Highest biogas yield was obtained from a mixing ratio of 50% cow dung to 50% donkey manure, however lowest biogas yield was obtained from cow dung. The study revealed that donkey manure has a higher caloric value and produces more biogas with higher methane content than cow dung. The investigation proved that co-digestion is a simple way to optimize biogas production rate.

1. Introduction

South Africa is the most industrialized country in Africa and is highly dependent on conventional fuels, non- renewable sources such as coal, oil and natural gas. This makes the country be one of the largest emitters of greenhouse gases in the world. Coal in South Africa, provides around 75% of fossil fuel demand and accounts for 90% of power generation in the country (Zafa 2010). These non-renewable sources are dwindling and becoming increasing expensive and environmentally damaging.

Biogas from anaerobic digestion can be a solution to current and future energy needs in South Africa. Compared to other renewable energy sources, such as solar and wind power, the biogas can easily be stored and transported. The implementation of biogas technology is not affected by geography of a particular area. Biogas technology involves the use of biogas digesters that are constructed vessel in which animal waste and other bio-degradable materials are broken down by bacteria complete absence of oxygen to produce biogas. The biogas digester is free from theft risks as compared to solar installations. Biogas consists of different component gases, mainly methane (CH_4), carbon dioxide (CO_2), with traces of hydrogen sulphide (H_2S) and hydrogen (H_2) gas (Bajracharya *et al*., 2010). The energy content of biogas is 9.8 kWh/m^3 (Thours 2007).

During digestion 80% of the pathogens and solids are eliminated and more effective liquid fertilizer is created as micro-organisms transform the organic pollutants into dissolved nutrients (Thy *et al*., 2003, Botero and Hernandez, 2005; Lansing *et al.,* 2008). The anaerobic digestion process is divided into four steps, hydrolysis, acidogenesis, acetogenesis and methanogenesis (Davidsson 2007; Leksell 2005).

- **Hydrolysis** is the first stage of anaerobic digestion where insoluble organic compounds such as proteins, fats, lipids and carbohydrates are converted into soluble organic components such as amino acids, fatty acids, monosaccharides, and other simple organic compounds.
- **Acidogenesis or Fermentation** is the second step where soluble compounds produced in the first stage are further degraded resulting in the production of carbon dioxide (CO_2), hydrogen (H_2) gas, organic acids, alcohols and some organic sulphur compounds (Gerardi, 2003).

- **Acetogenesis** is the third stage in which lactic acid, alcohols and glycerol, are converted by the acetogenic micro-organisms into acetic acid, hydrogen and carbon dioxide (Wiese *et al.,* 2009).
- **Methanogenesis** is the last stage where fermentation products such as acetate and hydrogen are converted to methane and carbon dioxide (Schön, 2009).

Biogas production is affected by the following factors: pH-value, temperature, concentration of solids, hydraulic retention time (HRT), volatile solids (VS) and loading rate (LR), inocula, carbon/nitrogen (C/N) ratio, toxicity, ammonium (NH_4) concentration, particle size, water content, stirring rate, and volatile fatty acids.

Efforts have been dedicated in recent years to find ways of improving the performance of digesters by treating different biomass. Co-digestion is the simultaneous digestion of homogenous mixture of two or more substrates (Wu, 2007). The use of co-substrates usually improves the biogas yields from anaerobic digester due positive synergisms established in the digestion medium and the supply of missing nutrients by the co-substrates (Mata- Alvarez *et al.,* 2003).

The aim of this study is to optimize methane content of biogas through co-digestion. Donkey manure and cow dung were used as co-substrates in the study. Current literature survey shows that co-digestion of donkey manure and other organic wastes in anaerobic digestion have not been carried out hence the current study to investigate the effect of co-digestion of cow dung and donkey manure. The results of the study indicate that co-digestion increases biodegradability of substrate and increases the amount of methane produced per unit of volatile solids.

2. Methodology

2.1 Biogas digester construction

A cylindrical batch biogas digester was constructed on a sunny site. The volume of the digester was 1.0 m^3. The biogas digester foundation was filled with concrete. The ratio of cement, sand and aggregate for the concrete was 1:2:3. The concrete was 20cm deep. The concrete was rammed to increase strength. The concrete was allowed to solidify for seven days. Before any construction, the concrete slab was covered with a black plastic sheet to avoid any moisture transfer from the ground into the digester. It was a brick wall digester with double walls. For mortar, the ratio of cement to sand was 1: 3. Dry soil was put between the double walls. This was done to insulate the digester thereby minimizing temperatures fluctuations within the digester. The brick wall was reinforced in order to strengthen the structure. The top dome part of the digester was also reinforced with concrete and steel rods. The digester was plastered. For inside plastering it was one part cement: two parts sand and outside plastering it was one part cement to three parts sand. Epoxy paint was used for painting the inside of the digester. Epoxy paint is one of the best paints on the market that prevents moisture transfer from inside or outside the biogas digester. A designed mechanical stirrer was used for agitation. Figure 1, shows the diagram of the constructed field-batch biogas digester.

2.2 Substrate preparation

Samples of fresh cow dung and donkey manure were obtained from University of Fort Hare Honey dale farm. The wastes were crushed mechanically to ensure homogeneity. The substrates were weighted on a digital scale. The water content for each sample was determined using the recommendation for better biogas production as reported by Ituen *et al*., (2007), that is, a total solid (TS) of 8% in the fermentation slurry.

2.3 Determination of substrate parameters

Total solids (TS), volatile solid (VS), ammonia-nitrogen (NH_4-N), pH, total alkalinity (TA), and chemical oxygen demand (COD) were determined by using the standard methods of the American Public Health Association (ALPHA), 2005. Calorific value (CV) of prepared slurry was measured using a calorimeter (CAL2K-ECO) and pH of slurry was measured by a digital pH meter. pH values were performed at the start, during and end of the biodegradability process. Temperature of the digester slurry was measured during biodegradability test by type K thermocouple thermometers. The following parameters were kept constant during the biogas production process, temperature, stirring rate and substrate mass.

2.4 Biogas analysis

The pressure of the biogas was measured daily by means of a pressure gauge fixed on top on the batch biogas digester. Methane and carbon dioxide contents in the biogas were sensed by Non-Dispersive Infra red sensors. Palladium/Nickel sensors were used for sensing hydrogen and hydrogen sulphide in biogas. CR 1000 data logger processed and stored data from different sensors. The output from data loggers was displayed on a computer screen. The system was powered with dc rechargeable battery that was connected to 20 W photovoltaic modules. Daily biogas production from the digester was measured by a biogas flow meter. The biogas digester was stirred by a mechanical stirrer. The stirring was done once a day for 25minutes. Previous studies noted higher reaction rates with higher agitation frequency (Cubas *et al.,* 2011). Figure 1, is the schematic diagram of the experimental set-up.

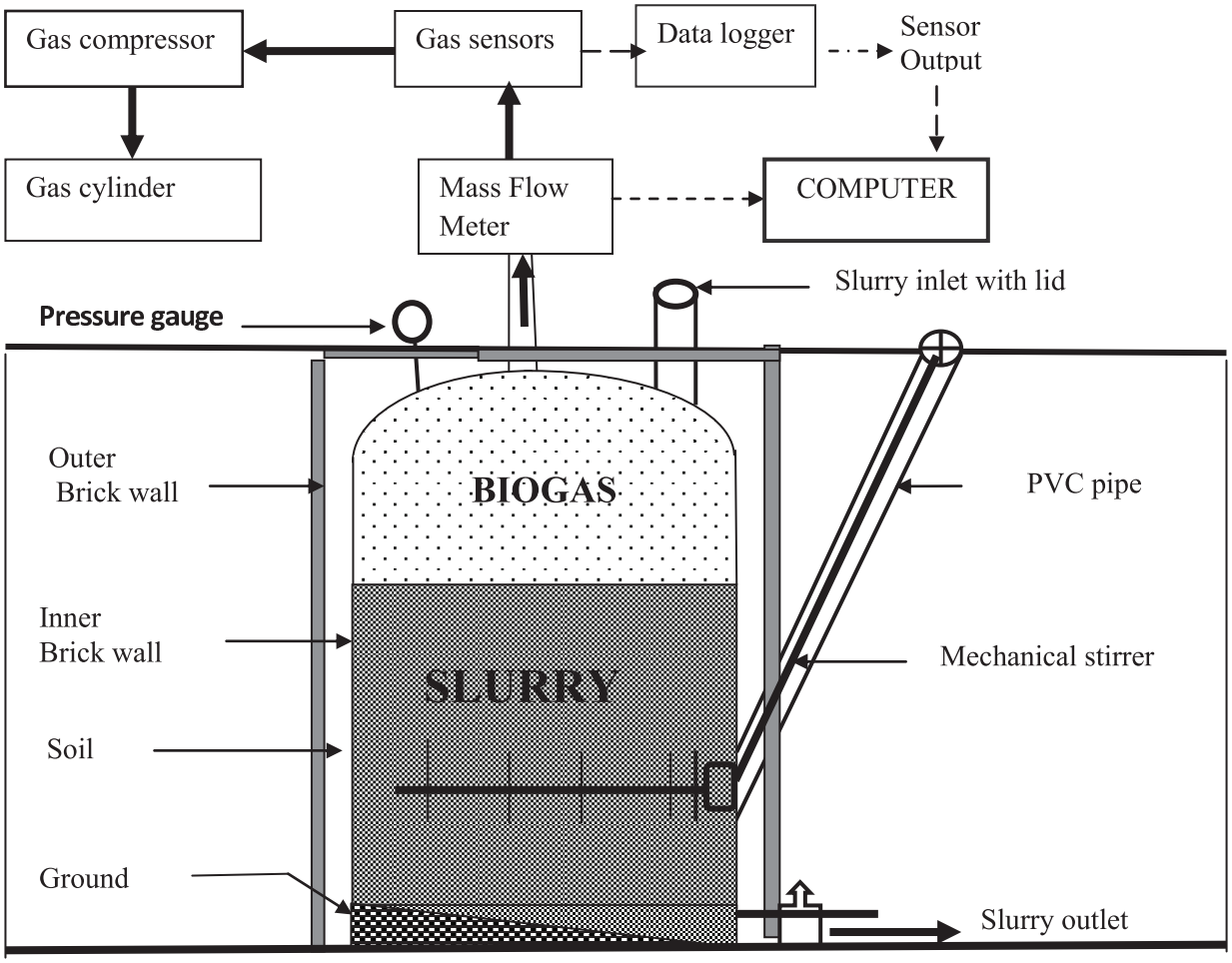


Figure 1. Schematic diagram of the experimental set-up of a batch field biogas digester

3. Results and Discussions

Table 1 shows characteristics of the two samples and Table 2 shows biogas composition for each sample. The composition of biogas was mainly methane, carbon dioxide and hydrogen. There were some traces of hydrogen sulphide. The methane yield for donkey manure was 50-65% and cow dung was 50-60%. Co-digestion of 20% cow dung and 80% donkey manure produced 46-65 % methane, 30% cow dung and 70% donkey manure yielded 45-66% methane and 40% cow dung and 60% donkey manure yielded 56-69% methane. Co-digestion of 50% of cow dung and 50% donkey manure produced 60-75% methane. Donkey manure had more CV, VS and CODS than cow dung, hence more biogas.

Table 1. Substrate characteristics of different samples

Parameter	Cow dung	Donkey dung
TS%	16.83	19.88
TS (mg/l	168160.48	198778.83
VS (mg/l)	117370.89	144189.99
VS/TS%	69.80	72.54
TA(mg/L)	5678-6260	6230-6536
COD(mg/L)	39754-40150	40110-41248
$N\text{-}NH_3$	125-230	135-250
Calorific value (MJ/g)	25.39	29.83
Temperature ^0C	28.0	28.0

Table 2. Biogas composition of different samples

Parameter	Cow dung [CD]	Donkey manure [DM]	20 % CD and 80% DM	30% CD and 70% DM	40% CD and 60% DM	50% CD and 50% DM	60% CD and 40% DM	70% CD and 30% DM	80% CD and 20% DM
pH [slurry]	6.45-7.4	6.58-7.5	6.47-7.6	6.5-7.4	6.4-7.5	6.6-7.7	6.5-7.5	6.48-7.4	6.5-7.5
CH_4%	40-60	50-60	52-65	51-66	56-71	60-75	57-70	54-65	50-64
CO_2%	30-45	35-45	32-40	30-39	25-35	20-30	27-36	31-40	30-42
H_2S%	0	0-0.1	0-0.1	0-0.1	Trace	Trace	Trace	Trace	Trace
H_2%	1-8	1-6	1-6	1-6	1-6	1-6	1-6	1-7	1-7

Figure 2, shows a graph of variation of gas yield for the samples. For all the substrates the gas yield increased with time and then attains a constant value before the biogas production rate drops gradually. However the co-digestion of cow dung and donkey manure attained maximum gas yield on 26^{th} day while for cow dung it was on 30^{th} day and for donkey manure it was on 28^{th} day. From the graph, it can be observed that donkey manure has a higher gas yield (390*l*) as compared to cow dung (365*l*). However, the gas yield was observed to increase greatly if 50% donkey manure was mixed with 50% cow dung. If cow dung is mixed with donkey manure in the ratio 1:1 the gas yield increases by approximately by 56%.

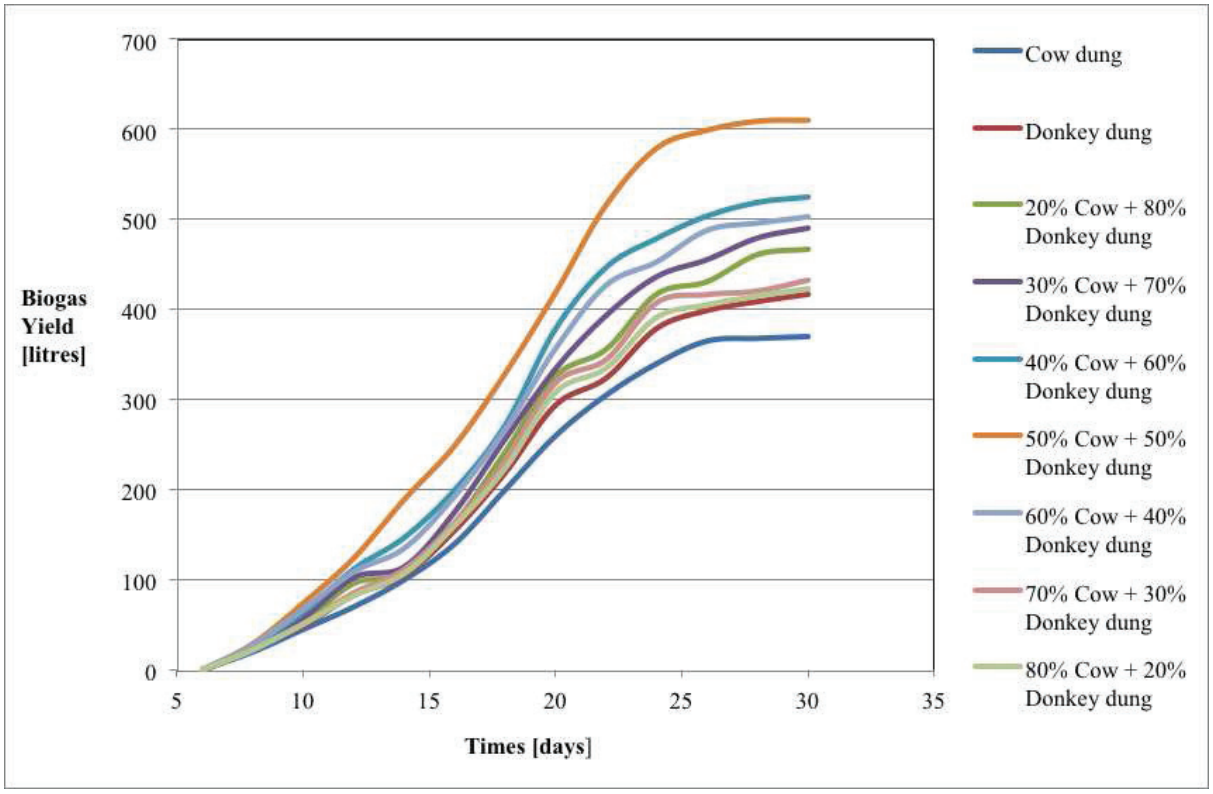


Figure 2. Biogas yield for single and co-substrates.

Regression analysis was used to find the gas yield prediction and equations 1 to 3 gives predictions of the gas yield (Y) for each day (X).

50% Cow and 50% donkey dung:	$Y = 3.725X^2 - 22.95X$	$R^2 = 0.994$
40% Cow and 60% donkey dung:	$Y = 3.137X^2 - 19.44X$	$R^2 = 0.991$
60% Cow and 40% donkey dung:	$Y = 2.939X^2 - 17.94X$	$R^2 = 0.991$
30% Cow and 70% donkey dung:	$Y = 2.709X^2 - 16.87X$	$R^2 = 0.991$
20% Cow and 80% donkey dung:	$Y = 2.518X^2 - 15.65X$	$R^2 = 0.990$
70% Cow and 30% donkey dung:	$Y = 2.620X^2 - 16.50X$	$R^2 = 0.989$
80% Cow and 20% donkey dung:	$Y = 2.498X^2 - 15.72X$	$R^2 = 0.991$
Donkey dung:	$Y = 2.337X^2 - 14.40X$	$R^2 = 0.992$
Cow dung:	$Y = 2.219X^2 - 13.99X$	$R^2 = 0.992$

There was a strong positive relationship between gas production and quantity (%) of co-substrates used ($R^2 = 0.994$) as shown on equation 1. For co-digestion of 50% cow dung and 50% donkey manure gas production was highest between 18-26 days. From 26 day and above biogas production decreased until there was no biogas production because all the food in the biogas digester had been consumed and there was no supply of food for the methanogens. For cow dung only, gas production increased as between 16-28 days. However, cow dung produces less biogas than donkey manure. The highest gas yield from 50% cow dung and 50% donkey manure was attributed by stable pH. The slurry mixture was able buffer its, hence highest biogas yield. Higher biogas yield from donkey manure than cow dung was attributed by high VS, COD, CV and suitable pH. Low biodegradability material in cow dung resulted in low biogas yield. Donkeys process food more efficient than cattle and this is also a contributory factor for higher gas yield in donkey manure. In all

samples the effluent pH fell within the range 6.45 to 6.6 while the final ranged from 7.4 to 7.7 (Table 2). Maximum insulation of the batch biogas digester resulted in negligible temperature variations.

4. Conclusions

Co-digestion of donkey manure and cow dung is highly desirable for increasing methane yield. Biogas production rate was different for different wastes because bacteria involved in the decomposition of the wastes were different. For optimum biogas production, cow dung and donkey manure should be co-digested in the ratio of 1:1 in terms of mass. The co-digestion process improves biogas quantity and quality. The biogas technology can be viable development option for Eastern Cape region of South Africa where plenty of donkeys and cattle are kept. Introduction of inocula in batch digesters has been noted as a way of reducing hydraulic retention time. The study also needs to highlight that there is need to keep more donkeys in rural areas because there have the following studied advantages:

- Take less food compared to cattle
- Have a longer life span (over fifty years)
- Produce more energy per gram than sheep, goats and cattle
- Can be a source of food (meat and milk)
- Produce more biogas with high methane content when co-digested with other wastes.

5. References

[1] American Public Health Association (ALPHA) 2005 *Standard methods for Examination of Water and Waste Water* (Washington DC: WEF)

[2] Bajracharya T R, Hamal G, Thapaliya N and Dhungana A 2010 Purification and compression of Biogas: A research Experience *J. institute of Engineering.* **7** 1-9

[3] Botero R and Hernandez C 2005 *Productive management of waste in tropical cattle farms* (de Costa Rica San Jose)

[4] Cubas S A, Foresti E, Rodrigues J.A, Ratusznei S.M and Zaiat M 2011 Effect of impeller type and stirring frequency on behavior of an AnSBBR in the treatment of low- strength waste water *Bioresour. Techno.* **102** 889-93

[5] Davidsson Å 2007 Increase of biogas production at waste water treatment plant- addition of urban organic waste and pre-treatment of sludge Lund University, Sweden

[6] Gerardi, M H 2003 *The Microbiology of Anaerobic Digesters* (New Jersey: John Wiley and Sons) pp 81-117

[7] Ituen E E, John N M and Bassey B E 2007 Biogas production from organic waste in Akwa Ibom State of Nigeria Appropriate Technologies for Environmental Protection in the Developing World. *ERTEP* **207** 17-19

[8] Lansing S, Botero R B and Martin J F 2008 Waste treatment and biogas quality in small-scale agricultural digesters *Bioresource Techno*. **99** 5881-90

[9] Leksell N 2005 *Present and future digestion capacity of Käppala waste treatment plant a study in a laboratory scale* Käppala wastewater treatment plant (Sweden)

[10] Mata-Alvarez J 2003 *Biomethanation of organic fraction of municipal solid waste* (London: IWA)

[11] Schön M 2009 *Numerical Modelling of anaerobic digestion processes in Agricultural Biogas plants.* Dissertation Innsbruck

[12] Thours A P 2007 Optimization of biogas production at Käppala wastewater treatment plant *Thickening sludge and biogas potential measurement in a laboratory* (Sweden) p 5

[13] Thy S, Preston T R and Ly J 2003 Effect of retention time on biogas production and fertiliser value of biodigester effluent *Livestock Research for Rural Development* **15** 7

[14] Wiese J, GmbH E and Oeynhausen B 2009 *Laboratory analysis and process analysis, biogas monitoring* (Germany) pp 1-12

[15] Wu W 2007 Anaerobic Co-digestion of Biomass for Methane Production Recent Research Achievements pp 1-10

[16] Zafa S 2010 Sustainable Solutions *South Africa's progress in Renewable Energy* (India)

Adipose Derived Stem Cells (ADSCs) and Low Intensity Laser Irradiation (LILI): Potential use in Regenerative Medicine

B Mvula and H Abrahamse[1]

Laser Research Centre, Faculty of Health Sciences, University of Johannesburg, PO Box 17011, Doornfontein 2028, Johannesburg, South Africa

E-mail: habrahamse@uj.ac.za

Abstract. Adipose tissue is highly specialised and a reliable source of adult stem cells. ADSCs can be harvested easily and in large amounts from adipose tissue or lipo-aspirates. Collagenase digestion is one of the procedures used in generating these cells from adipose tissue. ADSCs have been shown to have the differentiation potential to differentiate into other cells including smooth muscle, bone, nerve, heart, cartilage, liver and fat. LILI, a form of phototherapy, involves the application of monochromatic light in the 630 to 905 nm (visible to near infrared) range to biological tissue. Application to a variety of different cell types has been found to enhance cell viability, proliferation and differentiation of ADSCs. The differentiated cells from ADSCs are a major component of cardiovascular, reproductive, urinary, neural and intestinal systems and play a key role in diseases like arterioscelosis, asthma, hypertension and cancer. These cells could be used in regenerative medicine and tissue engineering. This review discusses ADSCs and LILI and their potential use in regenerative medicine.

1. Adipose Derived Stem Cells

Adult stem cells can be isolated from adipose tissue and lipo-aspirates in significant numbers and exhibit stable growth and proliferative kinetics in culture. These cells are termed ADSCs [1]. ADSCs have been isolated from donors with reduced morbidity, they have been multiplied and handled easily compared to bone marrow cells [2]. In the past decade several studies have provided preclinical data on safety and efficacy of these cells, supporting the use of these cells in future clinical applications [3]. Adipose tissue as a source of these cells allows them to be obtained in large quantities without difficulties and at a minimal risk [4]. White adipose tissue can be obtained in large quantities from human tissue and the stem cells residing in it are easily harvested from the tissue with the ability to differentiate into several cell lineages [5,6].

Cells generated from adipogenic origin have also shown to differentiate not only into osteoblasts, chondrocytes, myocytes, cardiomyocytes, fibroblasts and adipocytes but also into vascular lineages such as endothelial, smooth muscle blood cells [7-11]. This is due to the fact that they can release potent angiogenic factors such as leptin and vascular endothelial growth factor [12,13]. In cell transplantation, ADSCs have been found to promote radiological ossification efficiently, 90% of fractures healed eight weeks after surgery and during this process, blood perfusion enhancement through neovasculatisation was observed [14]. These cells have been shown to restore dystrophin expression of Duchenne skeletal-muscle cells *in vitro* [15]. ADSCs may constitute a potential cell based therapeutic alternative for the treatment of pancreatic ductical adenocarcinoma (PDAC) after being found to strongly inhibit human PDAC cell proliferation both *in vivo* and *in vitro* [16].

[1]To whom any correspondence should be addressed

Due to their plasticity and easy harvesting, ADSCs are the preferred stem cells to be used in tissue engineering rather than bone marrow derived stem cells [17,18]. However, ADSCs have certain limiting characteristics due to several factors. First, their capacity is limited in terms of subculturing and secondly, adipose tissue varies in its metabolic activity and as well as its capacity for proliferation and differentiation depending on the location of the depot of tissue and the age and gender of the patient [19,20].

2. Low Intensity Laser Irradiation

LILI, a form of phototherapy, involves the application of monochromatic light to biological tissue to elicit a biomodulative effect within that tissue. LILI is now accepted in many countries and is used in medical and dental practices. It elicits both a photobiostimulative and a photobioinhibitive effect within the irradiated tissue, each of which can be used in a number of therapeutic applications. LILI is not thermic (it does not produce heat) [21], and uses monochromatic light in the 630 to 905 nm range of the electromagnetic spectrum [22]. It stimulates capillary growth, granulation tissue formation and alters cytokine production. Altered keratinocyte motility and fibroblast movement have also been shown following low intensity laser irradiation [23]. These effects aid in the treatment of disorders like acute or chronic tissue hypoxia, destruction of tissues, as well as altered cell metabolism. Studies on LILI and stem cells have shown that low intensity irradiation can alter the metabolism of stem cells, increase adenosine triphosphate (ATP) production and so increase migration [24].

It has been shown that 5 J/cm^2 of laser irradiation at a wavelength of 635 nm positively affects ADSCs by increasing cellular proliferation, viability, and expression of β1-Integrin and Thy-1 (established stem cell markers), [25] and LILI in combination with epidermal growth factor (EGF) enhances the proliferation of ADSCs [26]. Several studies have been conducted on ADSCs identifying the effects of LILI at a cellular and molecular level (Table1).

Table 1. Studies conducted on ADSCs and/or differentiation inducers and/or LILI.

Study	Laser Parameters/ Differentiation Inducer (DI)		Results	References
ADSCs & LILI	5 J/cm^2 635 nm	No DI	Increased viability & proliferation	[25]
ADSCs & LILI	5 J/cm^2 636 nm	EGF	Increased viability & proliferation	[26]
ADSCs & LILI	10 & 15 J/cm^2 830 nm	No DI	Decreased viability & proliferation	[37]
ADSCs	No LILI	Retinoic Acid	Increased viability, proliferation & differentiation	[38]
ADSCs & LILI	5 J/cm^2 636 nm	Retinoic Acid	ADSC differentiation into smooth muscle cells	[38]
ADSCs	No LILI	Angiotensin II, Sphingosylphospho-Rylcholine, TGF-β1	ADSC differentiation into smooth muscle cells	[39]
ADSCs trans-planted in ische-mic mouse limbs	No LILI	No DI	Enhancement of angiogenesis & osteogenesis	[40],[14]
ADSCs	No LILI	Heparin	Differentiation into smooth muscle cells	[41]

Studies have also shown that LILI can increase proliferation of cells, cellular attachment, differentiation and production of transformation growth factor beta1 (TGF-β1) in human osteoblasts cells indicating that, *in vitro,* LILI can modulate the activity of cells and tissues [27]. Fiszerman and Markmann, (2000), discovered that LILI enhances wound healing in chronic diabetic foot ulcers [28]. LILI has been successfully used for pain attenuation and to induce wound healing in non-healing defects [29]. In addition, Abrahamse and co-workers, (2010), demonstrated that increasing the fluence and wavelength caused a decrease in ADSC viability and proliferation in a reciprocal manner.

3. Regenerative Medicine

Tissue engineering and regenerative medicine is a multi-disciplinary science that has evolved in parallel with recent biotechnological advances. It combines biomaterials, growth factors and stem cells to repair organs [30]. Adult stem cells hold great promise for use in tissue repair and regeneration as a novel therapeutic option [31]. This can be done by culturing the cells and differentiating them into the required lineage *in vitro* and then introducing the differentiated cells into the failing organs. Plastic and regenerative surgeons are constantly burdened with the challenge of replacing lost soft tissue. More than 6.2 million individuals received reconstructive plastic surgery procedures in 2002, approximately 70% of them as a result of tumour removal [1]. Elective cosmetic procedures also require the placement of soft tissue implants to restore or improve tissue contour for the purpose of enhancing aesthetic appearance. Conventional soft tissue-grafting procedures have had some clinical success for soft tissue augmentation and reconstruction. However, the need for secondary surgical procedures to harvest autologous tissues and an average of 40-60% reduction in graft volume over time are considered drawbacks of current autologous fat transplantation procedures. It should be possible to overcome these problems with tissue-engineered soft tissue grafts generated from the patient's own adult stem cells [1].

Parkinson's disease, stroke and multiple sclerosis are thought to be caused by a loss of neurons and glial cells. These cells can now be generated from stem cells in culture and can be used to treat the above diseases in human patients through transplantation [32]. Clinical trials for the regeneration of soft tissue, craniofacial tissue and cardiovascular tissue have enrolled a number of patients. Breast reconstruction with ADSCs trials have been reported by Yoshimura and colleages [33]. These cells have also been used to stimulate bone repair in calvarial defects [34]. ADSCs have been used to heal chronic fistulas in Crohns disease [35] and hold great promise for the treatment of cardiovascular diseases [36].

4. Conclusion

More cells with increased differentiation potential are required for the treatment of various regenerative diseases. ADSCs have that differentiation potential and exposing them to LILI increases the proliferation rate. Therefore the novelity of using LILI in conjuction with ADSCs could improve tissue regenerative disease treatment by increasing the number of differentiated specialised cells. Much more research has to be conducted to develop standardise procedures for treatments.

5. References

[1] Choi S S, Cha S M, Lee Y Y, Kwon S W, Park C J and Kim M 2006 *Bioch Biophys Res Comm* **345(2)** 631-637
[2] Ogawa R 2006 *Curr Stem Cell Res Th* **1** 13-20
[3] Tobita M, Orbay H and Mizuno H 2011 *Disc Med* **11(57)** 160-170
[4] Fraser J K, Wulur I, Alfonso Z and Hedrick M 2006 *Trends Biotechnol* **24(4)** 150–154
[5] Zuk P A *et al* 2002 *Mol Biol Cell* **13** 4279-4295
[6] Zuk P A, Zhu M, Mizuno H, Huang J I, Chaudhari S, Lorenzo H P, Benhaim P and Hedrick M H 2001 *Tissue Eng* **7(2)** 211-216
[7] Huang J I, Beanes S R, Zhu M, Lorenz H P, Hedrick M H and Benhaim P 2002 *Plast Reconstr*

Surg **109** 1033–1041
[8] Rodriguez A M *et al* 2005 *J Exp Med* **201** 13 97-1405
[9] Lin Y, Luo E, Chen X, Liu L, Qiao J, Yan Z, Li Z, Tang W, Zheng X and Tian W 2009 *J Cell Mol Med* **9** 929-939
[10] Planat-Benard V, Menard C, Andre M, Puceat M, Perez A, Garcia-Verdugo J M, Pénicaud L and Casteilla L 2004 *Circ Res* **94** 223–229
[11] Planat-Benard V *et al* 2004 *Circul* **109** 656–663
[12] Bouloumie A, Drexler H C and Lafontan M 1998 *Circ Res* **83** 1059-1066
[13] Sierra-Honigmann M R, Nath A K and Murakami C 1998 *Sci* **281** 1683-1686
[14] Shoji T, Li M, Mifune Y, Matsumoto T, Kawamoto A, Kwon S, Kuroda T, Kuroda R, Kurosaka M and Asahara T 2010 *Lab Inv* **90** 637-649
[15] Viera N M, Brandalise V, Zucconi E, Jazedje T, Secco M, Nunes V A, Strauss B E, Vainzof M and Zatz M 2008 *Biol Cell* **100** 231-241
[16] Cousin B *et al* 2009 *PLoS ONE* **4(7)** e6278
[17] Strem B M and Hedrick M H 2005 *Trends Biotechnol* **23** 64-66
[18] Strem B M, Hicok K C, Zhu M, Wulur I, Alfonso Z, Schreiber R E, Frazer J K and Hedrick M H 2005 *Keo Journal of Medicine* **54** 134-141
[19] Tholpady S S, Llull R, Ogle R C, Rubin J P, Futrll J W and Katz A J 2006 *Clin Plast Surg* **33** 55-62
[20] Giorgino F, Laviola L and Erriksson J W 2005 *Acta Physiol Scand* **183(1)** 13-30
[21] Matic M, Lazetic B, Polljacki M, Duran V and Ivkov-Simic M 2003 *Medicinski Pregled* **56(3-4)** 137-141
[22] Stadler L, Lanzafame R, Oskoui P, Zhang R, Coleman J and Whittaker, M 2004 *Photomed Laser Surg* **22(3)** 227-231
[23] Basford J R 1995 *Lasers Surg Med* **16** 331-342
[24] Gasparyan L, Brill G and Makela A 2004 *Laser Florence* 1-7
[25] Mvula B, Mathope T, Moore T and Abrahamse H 2008 *Lasers Med Sci* **23(3)** 277-282
[26] Mvula B, Moore T and Abrahamse H 2010 *Lasers Med Sci* **25** 33-39
[27] Khadra M, Lyngstadans S P, Haanaes H R and Mustafa K 2005 *Biomaterials* **26(170)** 3503-3509
[28] Fiszerman R and Markmann C A 2000 *Diabetes* **49** A393
[29] Hawkins D and Abrahamse H 2005 *Photomed Laser Surg* **23(3)** 251- 259
[30] Butler D L, Goldstein S A and Guilak F 2000 *J Biomech Eng* **122** 570-575
[31] Huh C H, Kim S Y, Cho H J, Kim D S, Lee W H, Kwon S B, Na J I and Park K C 2000 *J Derm Sci* **46(30)** 217-220
[32] Lindvall O and Kokaia Z 2006 *Nature* **441** 1094-1096
[33] Yoshimura et al 2010 *Breast J* **16(2)** 169-175
[34] Lendeckel S, Jodicke A, Christophis P, Heidinger K, Wolff J, Fraser J K, Hedrick M H, Berthold L and Howaldt H P 2004 *J Craniomaxillofac Surg* **32(6)** 370-373
[35] Garcia-Olmo D, Garcia-Arranz M and Herreros D 2008 *Expert Opin Biol Ther* **8(9)** 1417-1423
[36] Bai X and Alt E 2010 *Biochem Biophys Res Commun* **401(3)** 321-326
[37] Abrahamse H, Houreld N N, Muller S and Ndhlovu L 2010 *Med Technol SA* **24(2)** 9-14
[38] De Villiers J A, Houreld N N and Abrahamse H 2011 *Stem Cell Rev Rep* March 5 epub ahead of print DOI 10.1007/s12015-011-9244-8
[39] Harris L J, Abdollahi H, Zhang P, Mcllhenny S, Tulenko T N and DiMuzio P J 2009 *J Surg Res* DOI: 10.1016/j.jss.2009.08.001 ppg 1-9
[40] Nakagami H *et al* 2005 *Arterioscler Thromb Vasc Biol* **25** 2542-2547
[41] Rodriguez L V, Alfonso Z, Zhang R, Leung J, Wu B and Ignarro L J 2006 *PNAS* **103(32)**12167-12172

Transmission electron microscopy of radiation damage by keV multi-implantation in single-crystal diamond.

E.K. Nshingabigwi[1,2], J. H. Neethling[3], C. M. Levitt[1], S. R. Naidoo[1] and T. E. Derry[1]

[1]DST/NRF Centre of Excellence in Strong Materials (CoE-SM) and School of Physics, Wits 2050, University of the Witwatersrand, Wits 2050, Johannesburg, South Africa
[2] Department of Physics, National University of Rwanda, P.O Box 117 Huye, Rwanda
[3]DST/NRF CoE-SM and Physics Department, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa

E-mail: 0413352Y@students.wits.ac.za

Abstract. An understanding of the types of defects produced during the doping/implantation of diamond remains essential for the optimization of high-temperature, high-power electronic applications. Thus this study focuses on investigating the nature of the radiation damage produced during the multi-implantation of carbon ions in synthetic type Ib and natural diamonds, using the CIRA (Cold-Implantation-Rapid-Annealing) routine. The implanted and annealed diamond samples were characterized using transmission electron microscopy. For low fluence implantations, the damaged diamond retains its crystallinity after annealing at 1600 K while implanting using a high fluence, i.e., a fluence above the amorphisation-graphitization threshold, followed by rapid thermal annealing (RTA) at 1600 K, results in a graphite/ amorphous carbon layer close to the surface.

1. Introduction

Diamond with its outstanding and unique physical properties offers the opportunity to be used as semiconductor material in future device technologies. Promising applications are, among others, high speed and high-power electronic devices working under extreme conditions, as e.g. high temperature and harsh chemical environment. With respect to electronic applications, a controlled doping of the material is necessary which is preferably done by ion implantation. This technique allows incorporation of foreign atoms in defined depths and with controlled spatial distribution which is not achievable with other methods. However, the ion implantation process is always associated with the formation of defects which compensate carriers. It is therefore essential to understand the nature of defects produced under various implantation conditions. Previous studies [1, 2, 3] on the ion implantation doping of diamond using the CIRA (Cold Implant Rapid Anneal) technique have shown that a critical fluence of about $5.2\text{x}10^{15}$ ions cm^{-2} exists beyond which the diamond structure transforms into graphite. Below the critical threshold, the damaged diamond anneals to its pristine condition [1].

Similarly, Kalish et al. [4] have studied the nature of damage in ion-implanted and annealed diamond using Raman spectroscopy and reported that below the critical fluence, the damaged

diamond anneals back to diamond, while above the critical threshold, an amorphised layer, mostly sp^2 bonded, converts to graphite upon annealing. Direct observations using X-TEM of keV ion implantation damage in annealed diamond have not, apparently, been attempted due to the difficulty of thin specimen preparation. However, recent studies involving the direct examination of defects in irradiated diamond by transmission electron microscopy [5] using MeV energies, have reported that diamond transforms to an amorphous carbon when a critical threshold is surpassed. The amorphous carbon transforms to a moderately crystalline graphitic region upon annealing for 24 h at 1350 ^{o}C.

2. EXPERIMENTAL DETAILS

Slices of $\sim$ 500 μm thick were cut from a single crystal natural diamond parallel to the {110} plane. The slices were polished down to $\sim$ 40 μm using a scaife. The implantation were carried out at iThemba-LABS (Gauteng), South Africa, using a Varian 200-20A2F ion implanter. The diamond samples were C^+ implanted using ion energies and doses described elsewhere [1] with the dose rate of 1 μA per cm^2 and they were tilted about 7^o to the incident beam direction so as to minimize the effects of channelling. Cross-sectional specimens for TEM observations were prepared as follows: cross-sectional slices of about 40 μm were implanted edge-on, at liquid nitrogen temperature, followed by rapid thermal annealing for 30 min, at 1600 K, in an argon atmosphere. A pair of thin slices was glued edge against edge onto a copper support TEM grid. A 3 mm round disc was laser cut out around the copper grid. The samples were prepared for TEM study by Ar-ion milling using a Gatan Precision Ion Polishing System (PIPS, model 691) at 5 keV at an incident angle of 5^o (See Fig. 1). TEM investigations were carried out at the edge of the hole at the diamond/diamond interface as clearly shown on the 150× magnified image of Fig. 1(b). Bright field images and electron diffraction patterns were recorded in a 200 kV Philips CM20 electron microscope with a point-to-point resolution of 0.27 nm.

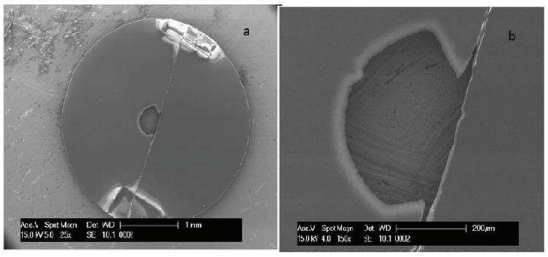


Figure 1. Scanning electron microscope images of the cross-sectional ion milled diamond TEM sample.

3. Results and Discussion

Fig. 2 shows a cross-sectional bright-field TEM image of diamond implanted with carbon ions to a high fluence of 7×10^{15} ions cm^{-2} at liquid nitrogen and rapidly annealed at 1600 K. Microdiffraction analysis of the implanted layer indicated that it was amorphous, however, the region close to the implanted surface displayed signs of nanocrystalline graphite, as evidenced by the 002 arcs of graphite in the selected area diffraction (SAD) pattern on the top right hand side of Fig. 2. The SAD pattern was recorded by placing part of the aperture over the implanted layer only, excluding the single crystalline diamond substrate. This finding suggests a three layered structure, i.e. graphite - amorphous carbon - diamond. The TEM results indicate that the graphite layer extends from the implanted surface to $\sim$ 1000 Å, in agreement with Monte Carlo simulations shown in Fig. 3, where the damage is peaked around $\sim$ 1000 Å, while the amorphous layer runs from 1000 Å to the end of the projected ion range. In the Fig. 2, g and a-C denote nanocrystalline graphite and amorphous-carbon layers, respectively. The SAD pattern on the bottom right hand side of Fig. 2 shows the 111 reflections of diamond. The dark

defect contrast at the interface region between the amorphous implanted layer and diamond substrate is due to the implantation damage at the tail of the projected ion range as shown in Fig. 2.

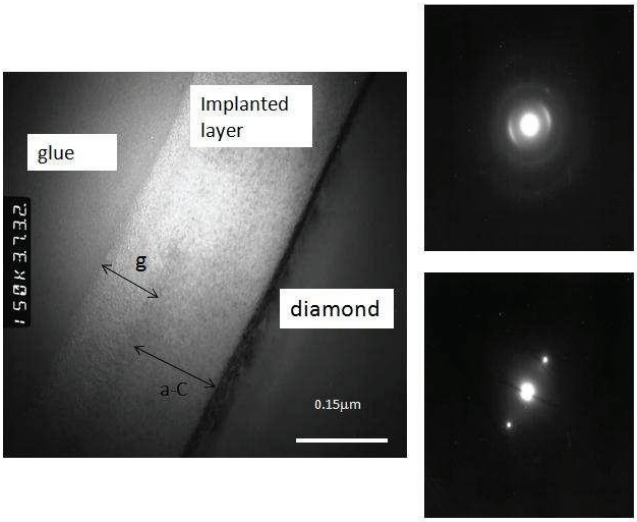


Figure 2. X-sectional bright-field TEM micrograph and SAD patterns of the high fluence ion-implanted diamond surface.

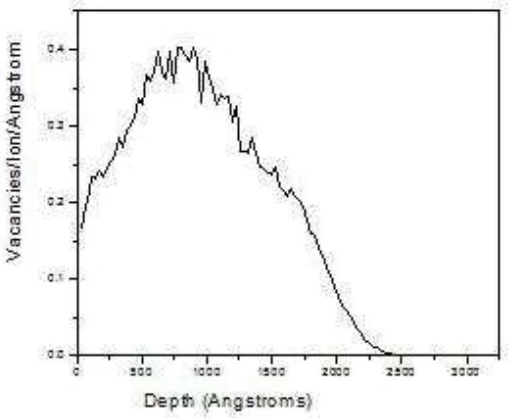


Figure 3. A SRIM-2003 simulation of the damage (vacancy) distribution produced by the carbon ion multi-implantation.

Therefore we can conclude that solid phase epitaxial regrowth of diamond does not occur during annealing near the end of range and no extended defects were observed. In addition, this work presents results which were not achieved in earlier studies [3] although similar implantations and annealing conditions were used in both of them. The explanation is thought to be that the damaged layer which is graphite-like in nature may have been etched off in earlier work after the subsequent cleaning in oxidizing acids. In the present work, the implanted and annealed sample was not cleaned in oxidizing acids which kept intact the damaged layer shown in Fig. 2.

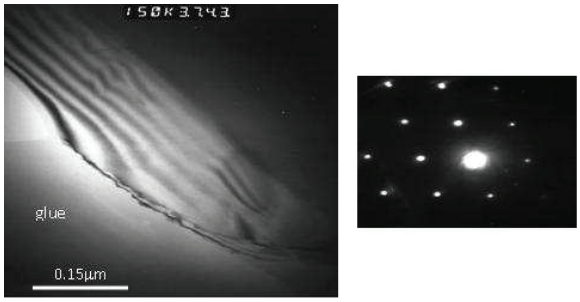


Figure 4. X-sectional bright-field TEM micrograph and SAD patterns of the low fluence ion-implanted diamond surface.

The average projected range predicted by the Monte Carlo program is $\sim$ 2500 Å while the highest level of damage lies at a depth of $\sim$ 1000 Å. The TEM image (See Fig. 4) of a specimen implanted with carbon ions to a lower fluence of 1×10^{15} ions cm^{-2} at liquid nitrogen and rapidly annealed at 1600 K does not exhibit any visible defects due to ion implantation and annealing. The damaged diamond anneals back to crystalline diamond as shown by the corresponding selected area diffraction. The sample has a thin region (result of ion milling) close to the edge showing a number of thickness fringes.

4. Conclusions

XTEM has been successfully used to characterize the radiation damage caused by keV multi-implantation in single crystal diamond. Diamond transforms to a graphite/amorphous-carbon layer when a critical threshold is surpassed whereas the damaged diamond anneals back to crystalline diamond when a fluence below the amorphization/graphitization threshold is used. Solid phase epitaxial recrystallization of diamond does not occur during the annealing near the end of range and no dislocations were observed.

5. Acknowledgements

We would like to thank the staff of iThemba LABS (Gauteng) for the provision of facilities and assistance, Mik Rebak for highly expert diamond polishing. The financial support of the DST/NRF Centre of Excellence in Strong Materials (CoE-SM) towards this research is hereby acknowledged.

References

[1] Spits R A and Derry T E 1994 *Nucl. Instr. and Meth. in Phys. Res.* B **85** 347.
[2] Prins J F 1988 *Phys. Rev.* B **38** 5576.
[3] Derry T E, Nshingabigwi E K, Levitt C M, Neethling J and Naidoo S R 2009 *Nucl. Instr. and Meth. in Phys. Res.* B **267** 2705.
[4] Kalish R, Reznik A, Nugent K W and Prawer S 1999 *Nucl. Instr. and Meth. in Phys. Res.* B **148** 626.
[5] Hickey D P, Jones K S and Elliman R 2009 *Diam & Rel. Mater.* **18** 1353.

Determining optical performance and current generation of a CPV as a function of intensity distribution

R D Schultz, F J Vorster and E E van Dyk

Department of Physics, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa, P.O. BOX 77000, 6031

E-mail: s206029578@live.nmmu.ac.za

Abstract. High concentration photovoltaic (CPV) systems utilise non-imaging optics to concentrate and distribute the solar flux uniformly onto a solar receiver to achieve maximum performance and power output from a CPV module. However, in many cases due to mechanical misalignment, tracker error and imperfections in the optical material, the optimum performance of the module is compromised. A LabVIEW programme employing visualization was used to determine the main contributing factor for current generation, i.e. position and intensity of the solar flux distribution. The topography was determined by multiple raster scans of the optical fibre receiver of a spectrometer in the plane of the reflective secondary's aperture where the cell would be placed. The results showed different currents been generated at different points on the cell surface. These results were put into a CPV cell current-voltage (I-V) characteristic simulating algorithm to extract I-V curve at each point. These were then compared with measured I-V curves obtained from the CPV system. The results showed that there was a non-uniform current density (J_{sc}) distribution due to non-uniform spectral and intensity distribution across the cell surface.

1. Introduction

High concentration photovoltaic (HCPV) systems are a cost effective alternative to flat plate photovoltaic (PV) modules in commercial applications. The advantage is that the amount of semiconductor material used in a CPV module are significantly reduced and replaced by inexpensive manufactured material like glass, poly(methyl methacrylate) (*PMMA*) and aluminum sheeting. These materials are used to concentrate the solar flux onto the device, which increases the irradiance and thus increases the short circuit current (I_{sc}) of PV device.

It is evident that the concentrator systems are dependent on the optical elements to a) collect and concentrate the direct solar flux and b) with the aid of the secondary, distribute the concentrated solar flux uniformly onto the semiconductor device. This will create an area of uniform current-generation over the device. The current production is also affected by the optical alignment, tracker error and absorption from the optical materials.

The current production as a function of intensity distribution of the concentrator system was examined using a custom-build scanning mechanism that raster scanned an optical fibre connected to a spectro-radiometer in the cell plane. In addition, the current-voltage (I-V) characteristics of various cell and module configurations were measured to determine the resulting current produced by the system and placed in a I-V simulator for analysis.

This paper investigates the current production of a Concentrating Triple-Junction cell (CTJ) as a function of intensity distribution.

2. Theory

To determine the current production of the multi-junction device, one must understand the effect of the spectrum distribution on the subcell material of the device. Since the optical system influences the spectrum and intensity distribution of the light incident on the concentrator cell, characterization of the spectrum is crucial for performance analysis.

A concentrator system generally utilises a substantial portion of direct terrestrial solar radiation (AM1.5D) [1] where the solar spectral distribution is due to the absorption by atmospheric molecules such as ozone, water and carbon dioxide [2].

The optical setup used in this study is based on a module design that is similar to the Sandia Baseline III point-focus Fresnel module [3]. This system utilises a Fresnel lens acting as a primary optical element and a truncated reflective pyramid acting as a secondary optical element, which attempts to distribute light uniformly across the cells' surface. Both optical elements have specific spectral losses due to the absorption and reflection properties of the material.

The Concentrating Triple-Junction (CTJ) solar cells used in this study are devices with a stacked structure consisting of three subcells (InGaP, InGaAs and Ge). The subcells are lattice matched to the Ge substrate and arranged in series from the largest to smallest bandgap energy. The spectral response of the subcells corresponds to wavelength ranges of about 300-700 nm, 500-900 nm and 800-1800 nm, respectively [4]. The resulting short circuit current (I_{sc}) produced by the cell is 13.6 mA at one sun (1000 $W.m^{-2}$) [4].

This arrangement of stacked subcells with increasing bandgaps allows for multiple absorptions of the solar spectrum resulting in a bigger portion of the spectrum to be utilised [5]. In the top subcell, absorption of high energy photons occurs, which correspond to the shorter wavelengths within the solar spectrum. This subcell acts as a filter, absorbing the energy within its bandgap range and transmitting the remaining wavelengths to the second and third subcell where a similar process occurs.

The functionality of the triple junction device is very similar to that of conventional series connected single junction solar cells except that in CTJ cells the subcells are connected via tunneling diodes. Since the triple junction subcells are connected in series, the resulting cell voltage (V_{oc}) is the sum of each of the subcell voltages [3,5]. The short circuit current density (J_{sc}) is more complex than the single junction cells where the current density of a single junction at one sun is given by:

$$J_{sc} = \int S(\lambda)\Phi_{inc}(\lambda)\, d\lambda \qquad (1)$$

where S is the spectral response of the material and Φ_{inc} is the incident solar radiation intensity. In a triple-junction solar cell, where the subcells are series connected, the J_{sc} of the whole device is equal to the smallest photo-generated current density ($Min\{J_i\}$) of a subcell.

$$J_{sc} = Min\{J_i\} \qquad (2)$$

Combining equation 1 and 2

$$J_{scc} = C.Min\left\{\int_{\lambda_{mi}}^{\lambda_{mf}} S(\lambda_m)\Phi_{inc}(\lambda_m)\, d\lambda\right\} \qquad (3)$$

The final produced current density under concentration (J_{scc}) is equal the smallest current density produced by one of the series-connected subcells multiplied by the concentration factor C.

To maximize the current produced by a triple junction cell and to avoid any of the subcells becoming mismatched, and hence reverse biased, all three series stacked subcells must produce the same current.

3. Experimental Procedure

In order to measure the intensity and spectral distribution of the concentrated light that is incident on a triple junction cell, a custom-built X-Y raster scanner programmed in LabVIEW was attached to the concentrator test module with the test cell removed. An optical fibre attached to a spectroradiometer was scanned in the optical plane of the test cell while irradiance intensity and spectral measurements as a function of position across the CTJ cell surface were measured. The resulting spectrum at a point as well as the integrated intensity were processed in a Mathematica programme, allowing for visualization of the topographical spectral and intensity distribution across the cell surface. After the completion of the scans, a CTJ cell was placed at the optical aperture of the secondary and current-voltage (I-V) curves were taken for characterization and analysis of the effect of the spectral distribution caused by the optical elements on the CTJ cell. The alignment of the reflective secondary lens was changed to study these effects at a module level.
Spectral measurements were taken at 1X concentration through the Fresnel lens to record the effect on possible current density production.

4.Results and Discussion

Since the spectrum utilised by the concentrator is not standard due to the interaction of the solar flux with the optical material, one must identify and establish where the losses occur.

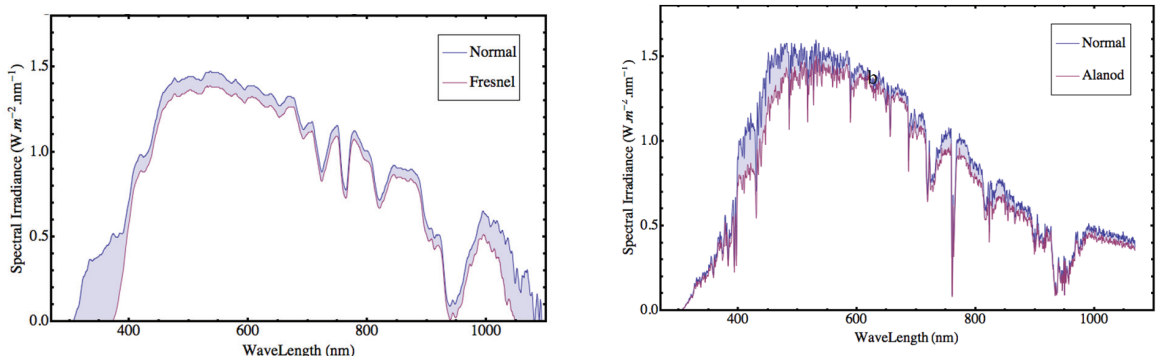


Figure 1: Graphs showing the effect of the Fresnel lens (left) and the Alanod material (right) on the spectrum.

Figure1 shows the effect on the solar spectrum due to the interaction with the concentrator material. Figure 1a shows that the primary refractive optical material (PMMA)[6] has an affinity to absorb the shorter wavelengths as well the longer wavelengths in the infrared region of the spectrum. The total intensity loss from the Fresnel lens due to spectral absorption amounted to about 10%. Figure 1b shows the reflective losses from the reflective secondary optical material (Alanod Aluminum)[7], with a total intensity loss due to reflection amount to about 5%. The effect on the spectrum is much less and seems to be uniform in the visible and infrared region of the solar spectrum. Less reflection occurs in the UV region of the spectrum where the InGaP is active.

The losses have a much more significant effect on the cell than expected. Equation 1 shows that the current density produced is a function of irradiance and the corresponding spectral response range of that material. Any loss in the spectrum will decrease the current production of the cell. This is the

same for flat plate PV. However, a CPV cell has subcells that produce currents independently from each other. They are connected in series and from equation 3 one can see that the subcell, which produces the smallest current, will dictate the current production of the device. The potential current density of each cell was calculated to be 14.33 mA.cm^{-2}, 14.28 mA.cm^{-2} and 19.45 mA.cm^{-2} for the InGaP, InGaAs and Ge, respectively, at 1000 W.m^{-2}.

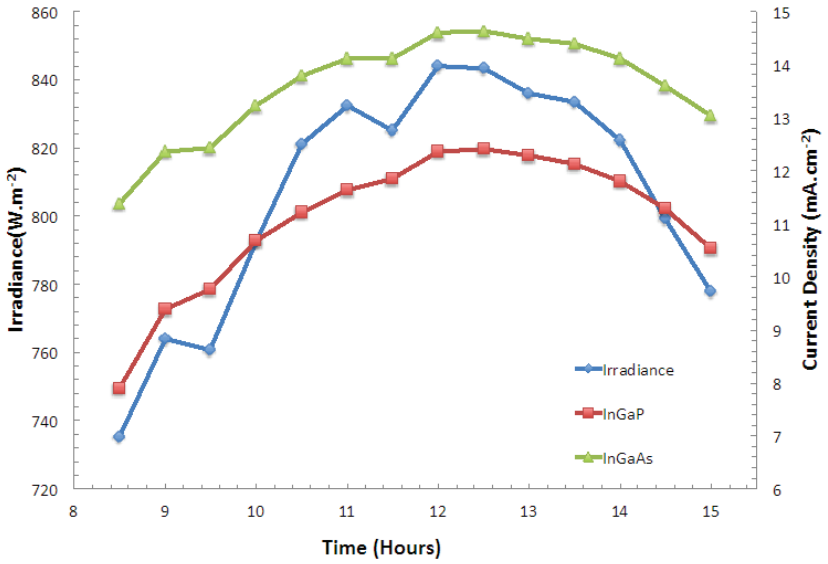


Figure 2: Current density of CTJ subcells at different irradiance measurements.

Figure 2 shows the possible current production of the InGaP and InGaAs subcells at 1X by irradiation through the Fresnel lens material. The theoretical current density values of the subcells are very similar at 1X concentration and since they are irradiance dependent, their increase and/or decrease should be proportional to each other as the device design is optimized to do so. However, the trend shown in figure 2 does not show this. This can be contributed to the fact that the spectrum is not the same throughout the day. Also seen from the figure 2 is that at solar noon there is a difference of about 1 mA.cm^{-2}, which from theory should be relatively the same.

The effect of the absorption of the wavelengths in InGaP region by the optical elements are suspected to be the cause of the decrease in current density as a result of absorption which causes a subcell mismatch. This will decreases the current production of the cell and as the subcells are series connected and will cause mismatch, which could lead to damage of the cell.

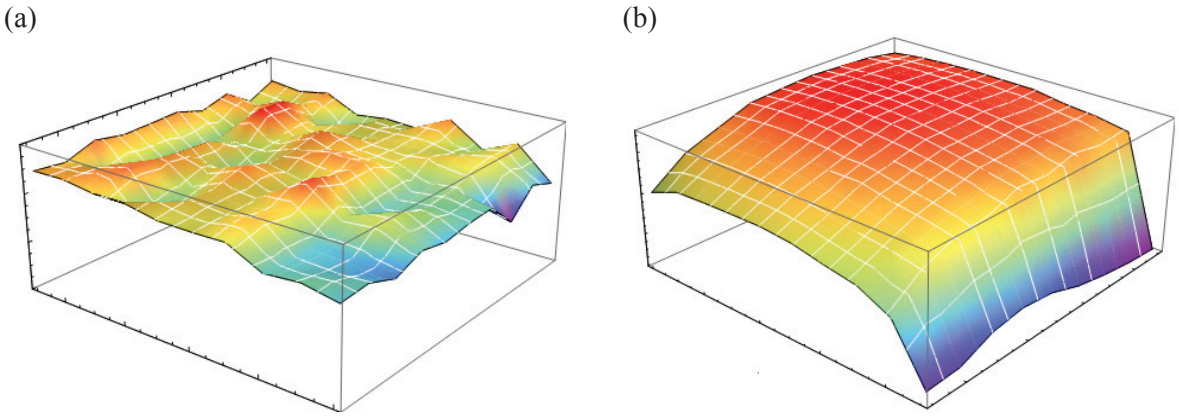


Figure 3: Current density plots of the InGaP subcell as a function of position at (a) 1X and (b) 330X.

Figure 3 shows the current density potential distribution of the current-limiting InGaP top subcell at 1X and 330X concentration, respectively. In figure 3a, one can see that there is a very slight non-uniform distribution pattern at 1X concentration which leads to a small difference in the current generation potential across the cell's surface. The average direct irradiance at 1X was measured to be 852W.m^{-2}. Figure 3b shows the current density distribution at 330X. This current density distribution is less uniform than at 1X and tapers off drastically on the edges. This shape is due to the distribution of the solar flux produced by the optical elements. The optical elements consist of a primary refractive Fresnel lens and a secondary reflecting truncated pyramid structure. The edges showing regions of low current density correspond to regions of low solar intensity.

The current produced for the whole cell at 1X concentration was 12.5mA at 852W.m^{-2} while the current produced by the cell at 330X concentration was 4.38 A at 864 W.m^{-2}. The total current produced by the cell is the integrated current density distribution over the whole cell area shown in figure 3(a) and (b).

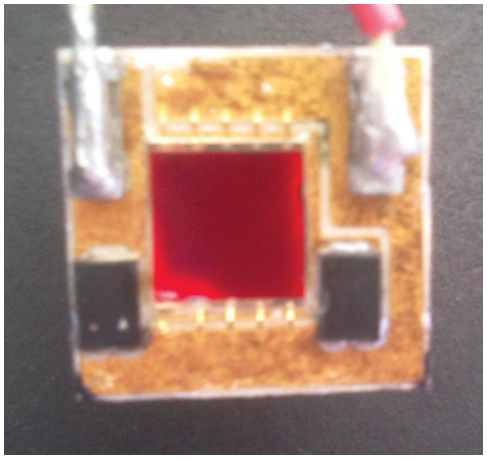


Figure 4: Photograph of CTJ cell that as undergone the describe distribution.

Figure 4 shows the effect on the cell of the current density distribution seen in figure 3b. The cell shown was reverse biased above the V_{oc} of the CTJ cell (2.6V). The bias produced a luminescence pattern in the visible region of the electromagnetic spectrum across the cell. Intense regions of luminescence coincide with areas of high carrier recombination. The intensity of the luminescence pattern seen in figure 4 corresponds well with the current distribution pattern seen in figure 3b. The low luminescence intensity areas shown in figure 4 correspond to regions of low free carrier density in the subcell. It is proposed that a prolonged exposure to a non-uniform high solar flux density may induce damage which leads to deficiencies in the free carrier concentrations of the CTJ device.

5. Conclusion

The results show the effect of the optical performance and current production as a function of the intensity distribution on the performance of a multi-junction solar cell. It is evident that the optical alignment as well as the spectral interaction of the incident solar radiation with the primary refractive and secondary reflective optical material dictates the intensity and spectral distribution of the concentrated solar energy on the CPV cell.

The current density distribution also showed the possible premature failure of the cell due to a non-uniform flux distribution on the cell. To achieve optimum results and performance, the optical system must be carefully designed to allow for the most uniform distribution and least amount of spectral losses. In doing so, one will prevent the occurrence of premature failure from reverse biasing and thus prolonging the lifetime and general performance of the cell as well as the module.

References

[1] Luque A Andreev V 2007 *Concentrator Photovoltaics* (Berlin Heidelberg: Springer-Verlag)
[2] Andreev V Grilikhes V Rumyantsev V 1997 *Photovoltaic conversion of Concentrated Sunlight* (Wiley-VCH)
[3] Markvart T Castaner L 2003 *Practical handbook of Photovoltaics Fundamentals and Applications* (Elsevier)
[4] Emcore specification datasheet, Triple-Junction Solar Cell for Terrestrial Applications, 2008(*www.emcore.com/assets/photovoltaics/CTJ_B_Web.pdf)*
[5] Luque A Hegedus S 2003 *Handbook of Photovoltaic Science and Engineering* (Wiley-VCH)
[6] PMMA specification datasheet.
[7] Alanod solar specification datasheet, MIRO high reflective 95, (http://alanod-solar.com/opencms/opencms/Reflexion/Technische_Info.html)

Inflammatory response of injured diabetic fibroblasts after low intensity laser irradiation at a wavelength of 830 nm

P R Sekhejane, N N Houreld and H Abrahamse[1]

Laser Research Centre, Faculty of Health Sciences, University of Johannesburg, P.O. Box 17011, Doornfontein, South Africa, 2028. Tel: +27 (0)11 559-6406

E-mail: habrahamse@uj.ac.za

Abstract. *Introduction*: Diabetes mellitus (DM) is a chronic disease characterized by impeded glucose metabolism and preceded by diabetic ulcers which are chronic due to deteriorated healing processes. Hypoxia, decreased fibroblast proliferation and impaired growth factors are amongst root factors that contribute to impaired healing. Photostimulation is a non-invasive treatment that utilizes low intensity laser irradiation (LILI) to stimulate appropriate cellular functions. *Materials and Methods*: Human skin fibroblast cells (WS1) were used in this study that consisted of four groups viz. normal, normal wounded, diabetic wounded and hypoxic, each with a non-irradiated control. Wounding was simulated by creating a central scratch using a pipette. A diabetic state was induced by growing cells in media that contained excess glucose to a final concentration of 22.56 mM, and for hypoxic insult, cells were incubated under anaerobic conditions (0% O_2 and 20% CO_2) for 4 h. Cells were then irradiated at a wavelength of 830 nm with 5 J/cm^2 and incubated for 1 or 24 h. Morphological changes were observed by light microscopy; ELISA and flow cytometry were used to determine interleukin (IL)-1β, IL-6 and tumour necrosis factor (TNF)-α as inflammatory markers; and caspase 3/7 for apoptosis was determined by luminescence. *Results*: After a 24 h incubation period the wounded area appeared decreased and hypoxic cells had regained normal morphologic features when irradiated, TNF-α and IL-1β had decreased in irradiated samples, whereas IL-6 was increased. Caspase 3/7 had decreased in irradiated samples at both 1 and 24 h. *Conclusion*: This study demonstrated the beneficial effects of LILI since the results showed significantly reduced inflammatory responses *in vitro* and hastened wound healing particularly under diabetic and hypoxic conditions.

1. Introduction

Diabetes mellitus (DM) is a group of chronic diseases with diverse etiology which are characterized by metabolic changes and chronic hyperglycaemia as a result of insulin deficiency [1]. The incidence of diabetes has increased in the past decade due to increased obesity often due to incorrect nutrition, changes in lifestyle and durability or physical inactivity [2,3]. The consequences of insulin deficiency lead to abnormal lipid, protein and carbohydrate metabolism as well as coronary vascular diseases (CVD) which attributes to 50-80% of deaths and usually affect every organ system in the body including the skin [4].

[1]To whom correspondence should be addressed.

Diabetic patients with extensive duration of metabolic abnormalities are prone to complications such as vascular diseases, stroke and ischemic heart conditions (macrovascular disorders) as well as nephropathy, neuropathy and retinopathy (microvascular disorders) [5,6]. Hypoxia is a decrease or deficiency in oxygen supply to the tissue. Diabetic patients tend to suffer from low vascular oxygen diffusion thereby inducing hypoxic insults [7,8].

Chronic foot ulcers or diabetic wounds, which are lesions resulting from skin breakage, are common in diabetic patients. Diabetic wounds are chronic since the healing process is impaired as determined clinically and experimentally [9]. They are also difficult to treat and usually prone to infection which may lead to gangrene due to ineffective wound closure [10]. Factors such as ischemia/hypoxia, decreased fibroblast proliferation or migration and impaired growth factors contribute to impaired wound healing [11]. Wounding initiates the expression of multi-factorial cytokines and growth factors which have ample activities that render them attractive agents for stimulating tissue repair. Cytokines and growth factors are signalling molecules expressed in response to activating stimuli such as tissue damage. Inflammatory markers such as interleukin-6 (IL-6) and tumour necrosis factor alpha (TNF-α) have been associated with diabetes [12].

Low intensity laser irradiation (LILI) or therapy has been a subject of interest in recent decades and is due to its desirable effects, particularly photo-biostimulation, when applied properly in addition to being non-invasive. It has been documented to be effective in wound closure. Exposure to light photons result in primary effects which are due to photoreception of photons with cytochromes; secondary effects which are induced by primary effects i.e. proliferation; and finally tertiary effects which are indirect because distant cells respond to changes incurred by other cells that have interacted directly with other photons [13,14]. Positive recovery of diabetic ulcers has been reported [15] in addition to reduced inflammatory markers [16]. Kawalec *et al*., [17,18] found that the healing time of diabetic wounds had reduced and both chronic (58%) and acute (100%) wounds completely healed. The aim of our study was to determine the effects of irradiation on wound closure or morphological changes and inflammatory markers *in vitro* using a fluence of 5 J/cm^2 at a wavelength of 830 nm.

2. Methods and Materials

2.1. Cell cultures

Human skin fibroblasts, WS1 (ATCC, CRL-1502) were cultured in minimum essential media (MEM; 32360026) and the media was supplemented with 10% foetal bovine serum (FBS), 1% penicillin-streptomycin (antibiotic) and fungizone (anti-fungal), 2 mM L-glutamine, 0.1 mM non-essential amino acids and 1 mM sodium pyruvate. Cultures were maintained at 37 °C in 5% CO_2 and 85% humidity.

The study consisted of four groups viz. normal (N), normal wounded (NW), diabetic wounded (DW) and hypoxic (H) and each group had a non-irradiated (sham-irradiated) control. Diabetic insult was simulated by maintaining cells in a media containing additional glucose (17 mM) resulting in a final concentration of 22.56 mM [19,20]. Hypoxic insult was achieved by growing cells in FBS free media overnight [21-23] and incubating them in an anaerobic chamber for 4 h with anaerobic gas pack (0% O_2 and 20% CO_2 within 2 h) [24]. A wound was simulated by making a central scratch using a 1 ml pipette.

2.2. Laser irradiation

All laser irradiations were conducted in a dark room. An 830 nm diode laser was used with an output power of 45 mW. The irradiance spot size was 9.1 cm^2. Cells were irradiated for 16 min 50 s in order to receive a fluence of 5 J/cm^2.

2.3. Biological assays

2.3.1. Cellular morphology. Qualitative changes in morphology were observed using an inverted light microscope and images were captured using an Olympus digital camera.

2.3.2. Cytokine analysis. The sandwich type optEIA™ enzyme-linked immunosorbent assay (ELISA) and flow cytometry were used to quantify cytokines IL-1β, IL-6 and TNF-α. Briefly, ELISA plates were coated by specific capture antibody overnight. It was then washed and blocked by assay diluents for inhibiting non-specific binding. Samples and standards were individually pipetted into wells and incubated at room temperature. Following the wash, cells were incubated with biotinylated anti-human monoclonal detection antibody, conjugated to streptavidin-horseradish peroxidase. Subsequent incubation with substrate was performed before the reaction was stopped and absorbance at a wavelength of 450 nm was read.

For flow cytometry, cytometric bead array (CBA) human flex sets were used. Supernatant from samples was mixed with capture beads and incubated in the dark. Thereafter samples were incubated with phycoerythrin (PE) detection reagent and incubated in the dark again. Samples were analyzed with the BD FACSArray bioanalyzer.

2.3.3. Apoptosis. Caspase-Glo 3/7 luminescence assay was used to quantify activated caspase-3 and -7 which play a vital role in the apoptosis pathway in mammalian cells. Briefly, cells were incubated and allowed to attach to the luminescent plate for 3 h. Then, Caspase-Glo 3/7 reagent was added followed by room temperature incubation. Luminescence was determined subsequently.

3. Results

3.1 Morphology

Cells were observed for any changes in morphology or reduction of wounded area after 1 or 24 h (1 h data not shown). The wounded area in irradiated normal wounded and diabetic wounded cells was decreased as compared to non-irradiated cells. Hypoxic cells had lost their characteristic morphology with some cells detaching. However, after irradiating cells had regained their characteristic morphology (Figure 1). Inducing hypoxia in fibroblasts represents the greatest stressor when compared to normal wounded and diabetic wounded cells as this condition has greater and more prolonged effects to the detriment of the cells. This is demonstrated by morphological variation (Figure 1 j and k).

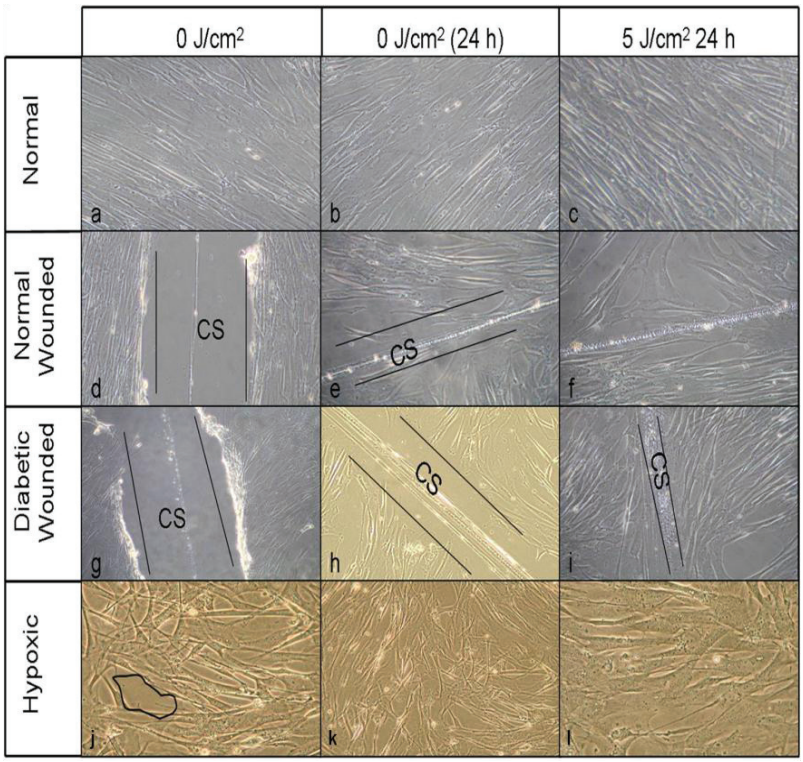


Figure 1. Morphological changes observed under the light microscope, in normal, normal wounded, diabetic wounded and hypoxic cells. Cells were exposed to 5 J/cm^2 at a wavelength of 830 nm and incubated for 24 h at 37 ºC. Migration was rapid in irradiated wounded cells although, in diabetic wounded cells they showed no contact in the central scratch (CS) as compared to normal wounded cells. Non-irradiated normal wounded cells had infiltrated and made contact at the CS, whereas diabetic wounded cells had infiltrated the CS. Spaces (encircled areas) in hypoxic cells were minimized post-irradiation and cells had regained their morphological features common to fibroblasts.

3.2. Cytokine analysis

3.2.1. *ELISA*. TNF-α was significantly decreased after 1 h incubation in irradiated normal, normal wounded ($P \leq 0.01$), diabetic wounded ($P \leq 0.05$) and hypoxic cells ($P \leq 0.01$) in comparison to their respective non-irradiated controls. After 24 h incubation the difference was seen in normal wounded, diabetic wounded ($P \leq 0.05$) and hypoxic cells ($P \leq 0.001$) as shown in Figure 2A. IL-1β had decreased significantly after 1 h incubation in irradiated normal and hypoxic cells as compared to their non-irradiated controls ($P \leq 0.05$). It was noted that after 24 h, IL-1β increased with an increase in severity of stress. Nonetheless, irradiated diabetic wounded and hypoxic cells showed significant IL-1β decrease ($P \leq 0.05$) as shown in Figure 2B. IL-6 was seen increasing with severity of the stress, excluding hypoxic stress; although the increase in irradiated normal, normal wounded and diabetic wounded cells was insignificant. However, a significant increase was seen in irradiated hypoxic cells after 1 h incubation. After 24 h incubation no significant results were observed (Figure 2C).

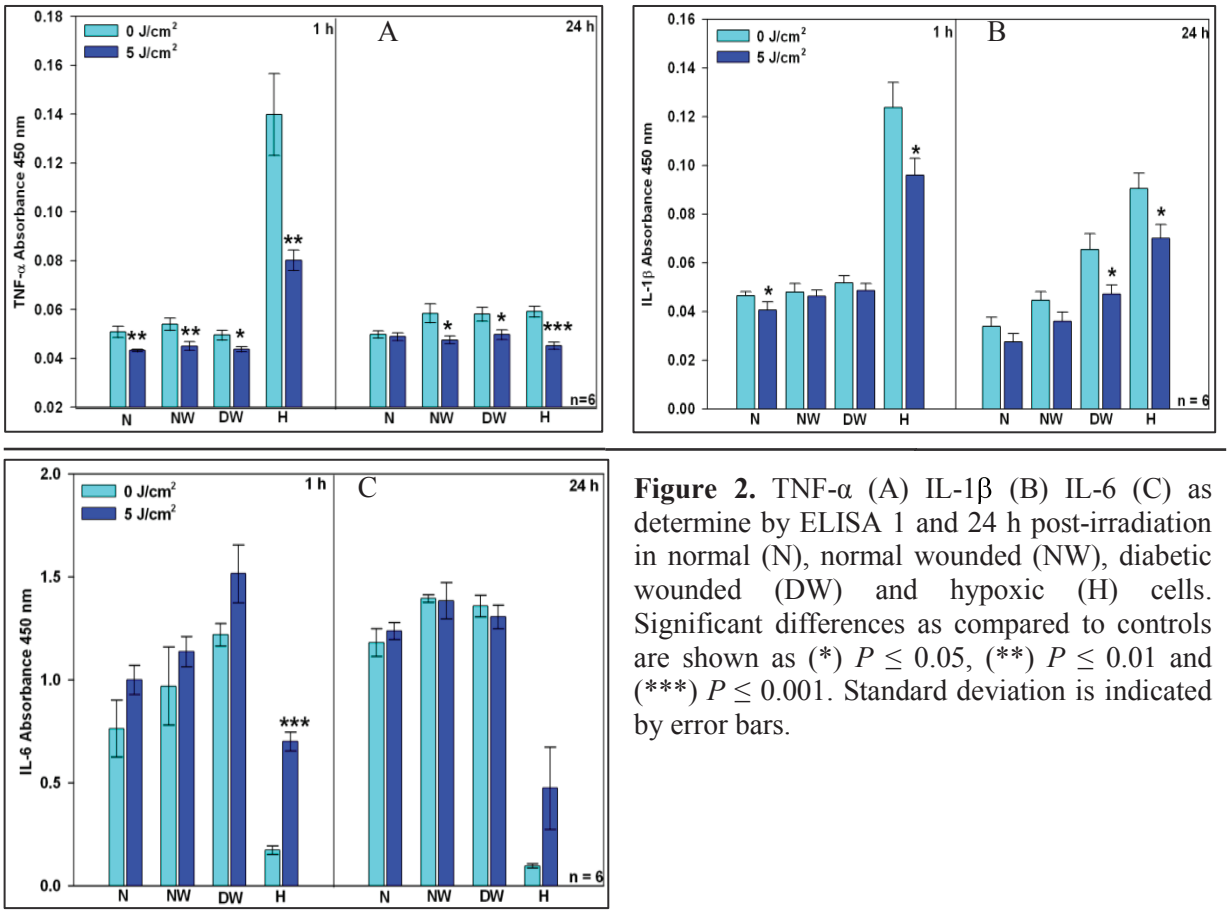


Figure 2. TNF-α (A) IL-1β (B) IL-6 (C) as determine by ELISA 1 and 24 h post-irradiation in normal (N), normal wounded (NW), diabetic wounded (DW) and hypoxic (H) cells. Significant differences as compared to controls are shown as (*) $P \leq 0.05$, (**) $P \leq 0.01$ and (***) $P \leq 0.001$. Standard deviation is indicated by error bars.

3.2.2. *Flow cytometry*. TNF-α in cells irradiated at a wavelength of 830 nm did not produce any significant changes as compared to non-irradiated control cells, except in hypoxic cells ($P \leq 0.05$) after 1 h incubation. After incubation of cells for 24 h post-irradiation, only hypoxic cells had significantly decreased TNF-α compared to their non-irradiated control ($P < 0.01$) in Figure 3A.

IL-1β determination with flow cytometry showed no significant difference when normal, normal wounded, diabetic wounded and hypoxic cells were irradiated and compared to their non-irradiated controls after 1 h incubation. Despite irradiation with 5 J/cm^2 and an incubation of 24 h post-irradiation, there was no significant difference between irradiated and non-irradiated controls, except in diabetic wounded cells ($P \leq 0.01$) as shown in Figure 3B.

Irradiation seemed to increase expression of IL-6 although a significant increase was seen in irradiated normal and normal wounded cells after 1 h incubation ($P \leq 0.01$ and $P \leq 0.05$; respectively). However, IL-6 in both irradiated and non-irradiated hypoxic cells was decreased as compared to other stress models ($P \leq 0.001$). After 24 h incubation, irradiated normal wounded cells had a significantly decreased IL-6 ($P \leq 0.05$) and diabetic wounded cells that received irradiation had a significantly increased IL-6 ($P \leq 0.01$; Figure 3C).

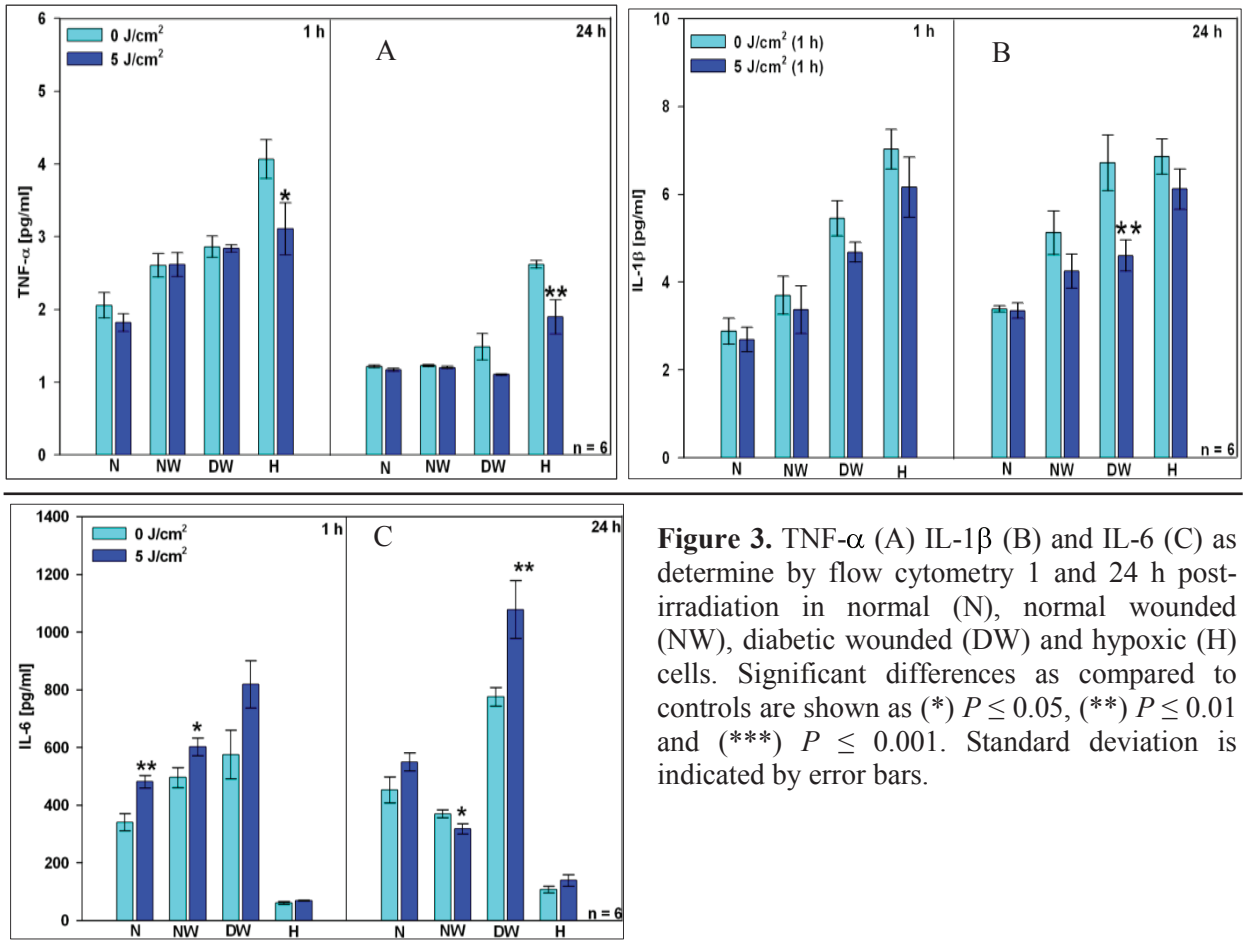


Figure 3. TNF-α (A) IL-1β (B) and IL-6 (C) as determine by flow cytometry 1 and 24 h post-irradiation in normal (N), normal wounded (NW), diabetic wounded (DW) and hypoxic (H) cells. Significant differences as compared to controls are shown as (*) $P \leq 0.05$, (**) $P \leq 0.01$ and (***) $P \leq 0.001$. Standard deviation is indicated by error bars.

3.3. Apoptosis

Caspase 3/7, (Table 1), in irradiated cells was not significantly different as compared to their non-irradiated respective controls after 1 h incubation. Normal cells, despite irradiation, had a significantly decreased caspase 3/7 as compared to both irradiated and non-irradiated stressed models ($P \leq 0.001$). After 24 h incubation, a significant decrease was seen in irradiated normal wounded ($P \leq 0.001$) and diabetic wounded cells ($P \leq 0.01$). Comparison of the two incubation periods showed significantly decreased caspase 3/7 after 24 h incubation in all models ($P < 0.01$).

Table 1. Irradiated and non-irradiated cells were analyzed for caspase 3/7 activity after 1 or 24 h incubation. Significant differences as compared to non-irradiated controls are shown as (**) $P \leq 0.01$ and (***) $P \leq 0.001$.

Cells	0 J/cm^2 (1 h)	5 J/cm^2 (1 h)	0 J/cm^2 (24 h)	5 J/cm^2 (24 h)
Normal	1135.20 ±77.32[a]	1017.37 ±76.72	787.07 ±49.94	683 ±42.14
Normal Wounded	2235.74 ±132.43	2145.12 ±163.84	1699.73 ±108.41	932.75*** ±90.95
Diabetic Wounded	3345.81 ±91.07	3128 ±73.12	1814.97 ±92.90	1387.42** ±100.52
Hypoxic	3353.58 ±132.29	3117.97 ±63.28	2644.98 ±146.00	2432 ±111.28

[a]Standard Error

4. Discussion and Conclusion

Research conducted by Kneebone suggest that LILI stimulates mitochondrial activity by altering the redox state which increases production of ATP and alters cell membrane permeability due to activation of sodium/potassium pumps, which in turn affects calcium ion influx. It is thus believed that biological effects manifest through energy absorbing chromophores such as mitochondrial cytochromes and endogenous porphyrins. Subsequently, it may result in stimulation of various biological responses such as increased proliferation due to increased adenosine triphosphate (ATP), nitric oxide (NO) for modulation of several cytokines [13, 14].

LILI demonstrated effective wound healing as demonstrated by a reduction in the central scratch area and normalization of morphology of cells exposed to hypoxic conditions after treatment with a fluence of 5 J/cm^2. Hopkins *et al.*, [25] also observed similar findings; which were also a time-dependent decrease in wound area. Ricci *et al*., [26] found hypoxic stressed endothelial cells regained their characteristic morphology post-irradiation. Another study by Zungu *et al*., [27] showed that hypoxic cells had a significantly reduced mitochondrial membrane potential (MMP) when compared with normal non-irradiated control cells. Wounded and hypoxic cells irradiated with 5 J/cm^2 showed an increase in mitochondrial responses when compared with non-irradiated cells. TNF-α and IL-1β are potent inducers of inflammation, therefore continuous or over-production of these pro-inflammatory cytokines may lead to delayed wound healing, cytotoxity or cell death. In hypoxic cells, these cytokines were found more elevated as compared to other stress models. Post-irradiation these cytokines were found decreased; suggesting that laser irradiation triggered other biological mediators that contribute to homeostasis or balancing of cellular physiology.

Since diabetic wounds are difficult to heal or close effective, the decrease of these mediators found in irradiated diabetic wounded cells correspond with the reduced wounded area seen post-irradiation. IL-6 is a pleitropic cytokine which has both pro- and anti-inflammatory effects [28] and it is usually expressed in response to or together with IL-1 and TNF-α [29]. An increase seen in irradiated cells after 1 h incubation could suggest that IL-6 is radio-protective and thus plays an anti-inflammatory role since after 24 h incubation it is decreased in some irradiated models i.e. irradiated normal wounded cells (flow cytometry). Ali *et al*., [30] found increased IL-6 in hypoxic endothelial cells, whereas our hypoxic stressed cells exhibited diminished IL-6 expression. It is unclear why IL-6 is under-expressed in these cells, but it could be attributed to its pleiotropic nature as well as the ability to mediate the expression of IL-1β and TNF-α [31]. As it is difficult to rule which effects could have been exerted by the cytokines discussed above since their effects are also dependent on the pathway they follow, caspase 3/7 was determined to establish if cells were surviving or dying. Our findings

suggested that laser irradiation had an anti-apoptotic effect on stressed cells. This also insinuates that the decrease in IL-1 and TNF-α was as a result of cells undergoing an anti-inflammatory state.

In conclusion, this study demonstrated the ability of LILI to be a potential therapeutic modality as it enhanced the recovery of stressed cells *in vitro*.

5. References

[1] Kuzuya T *et al* 2002 *Diabetes Res Clin Pract* **55** 65-85
[2] Heydari I, Radi V, Razmjou S, Amiri A 2010 *Int J Diabetes Mellit* **2** 61-63
[3] Dynyak A K, Dynyak AA and Popova F V 2010 *Med Hypotheses* **74** 1002-1005
[4] Levin M E 1986 *The Skin in Diabetes* ed Jelinek J E (Philadelphia: Lea and Febiger) 73-94
[5] Wild S, Roglic G, Green A, Sicree R and King H 2000 *WHO Report Global Burden of Disease 2000* 1-28
[6] World Health Organization 2006 *Report of a WHO/IDF Consultation* 1-50
[7] Gao W, Ferguson G, Conell P, Walshe T, Murphy R, Birney Y A, O'Brien C 2006 *J Mol Cell Cardiol* **42** 609-619
[8] Simanonok J P *Med Hypotheses* **46** 155-161
[9] Loots M A M, Lamme E N, Mekkes J R, Bos J D and Middelkoop E 1999 *Arch Dermatol Res* **291** 93-99
[10] Nather A, Bee C S, Huak C Y, Chew J L L, Lin C B, Neo S and Sim E Y 2008 *J Diabetes Complicat* **2** 77-82
[11] Brem H and Tomic-Canic M 2007 *The J Clin Invest* **17** 1219-1222
[12] Zozulinska D and Wierusz-Wysocka B 2006 *Diabetes Res Clin Pract* **74S** S12-S16
[13] Kneebone W J 2007a *Dynamic Chiropractic* **25** 1-7
[14] Kneebone W J 2007b *Practical Pain Management* **7** 62-66
[15] Kazemi-Khoo N 2006 *The Foot* **16** 184-187
[16] Safavi S M, Kazemi B, Esmaili M, Fallah A, Modarresi A and Mir M 2008 *Lasers Med Sci* **23** 331-335.
[17] Kawalec J S *et al* 2001 *The Foot* **11** 68-73
[18] Kawalec J S, Pfennigwerth T C, Hetherington V J, Logan J S, Penfield V K, Flauto J A and Shearer P M 2004 *The Foot* **14** 68-71
[19] Houreld N and Abrahamse 2007 *Diabetes Technol Ther* **9** 451-459
[20] McDermott A M, Kern T S and Murphy C J 1998 *Curr Eye Res* **1** 7924-932
[21] Kwon Y B, Lee Y-S and Sohn K-C 2007 *J Dermatol Sci* **46** 91-99
[22] Cuvier C, Jang A and Hill R P 1997 *Clin Exp Metastas* **15** 19-25
[23] Galindo M, Santiago B, Alcami J, Rivero M, Martin-Serrano J and Pablos J L 2001 *Clin Exp Immunol* **123** 36-41
[24] Van Horn K G, Warren K and Baccaglini E J 1997 *J Clin Microbiol* **35** 2170-2173
[25] Hopkins J T, McLoda T A, Seggmiller J G and Baxter G D 2004 *J Athl Training* **39** 223-229
[26] Ricci R, Pazos M C, Borges E and Pacheco-Soares C 2009 *J Photochem Photobiol B* **95** 6-8
[27] Zungu I L, Hawkins Evans D and Abrahamse H 2009 *J Photochem Photobiol* **85** 987-996
[28] Heinrich P C, Behrmann I, Haan S, Hermanns H M, Muller-newen G and Schaper F 2003 *Biochem J* **374** 1-20
[29] Xing Z, Gauldie J, Cox G, Baumann H, Jordana M, Lei X and Achong M K 1998 *J Clin Invest* **101** 311-320
[30] Ali M H, Schlidt S A, Chandel N S, Hynes K L, Schumacker P T and Gewertz B L 1999 *Am J Physiol Lung Cell Mol Physiol* **277** L1057-1065
[31] Yamagishi S, Ohnishi M and Pawankar R 2000 *Nippon Jiblinkoka Gakka* **103** 829-835

Optical coherence tomography as a diagnostic tool

Ann Singh[1,2] **and Aletta Elizabeth Karsten**[1]

[1] Biophotonics Group, National Laser Centre, CSIR, P.O. Box 395, Pretoria, South Africa, 0001
[2] Laser Research Institute, Stellenbosch University, Private bag X1, Matieland, 7602

E-mail: asingh1@csir.co.za

Abstract. Optical Coherence Tomography (OCT) has been used in biomedical applications as a method to non-invasively detect changes occurring in tissue such as the detection of skin cancer The effect of skin tone on detection of skin cancer has however not been studied extensively. Skin simulating phantoms with different optical properties can be used to optimise a system before it can be used on patients. For this reason accurately knowing the optical properties is imperative. An Integrating Sphere system can be used to determine these optical properties. Preliminary results from phantom and human skin indicate that skin tone may influence the data obtained from OCT images.

1. Introduction

OCT is a non-invasive imaging technology based on a Michelson interferometer. It is analogous to ultrasound, however whereas ultrasound uses sound waves, OCT works on light waves, measuring the backscattering properties of tissues. The light reflected off a sample is interfered with a reference wave utilizing a technique known as low-coherence interferometry. An interference signal arises when the optical pathlength difference between the two arms of a Michelson interferometer is within the coherence length of the source. Hence, the axial resolution in OCT is proportional to the temporal coherence length of the source. OCT is an imaging technique that fills the gap between confocal microscopy and ultra-sound imaging. The typical high resolutions offered by OCT are 1-5 μm although the penetration depth is currently limited to about 2-3 mm. It offers a non-destructive, non-contact, repeatable method of in vivo real time imaging at high speeds [1].

Fercher et al. [2] provide a very concise and nice description of the applications of OCT. "At present OCT is used in three different fields of optical imaging, in macroscopic imaging of structures which can be seen by the naked eye or using weak magnifications, in microscopic imaging using magnifications up to the classical limit of microscopic resolution and in endoscopic imaging, using low and medium magnification."

OCT was first demonstrated for retinal imaging [3]. Since then it has been used in biomedical applications such as ophthalmology, gastroenterology, urology, dentistry, cardiology and dermatology among other disciplines [4-9]. Atherosclerosis plaques, which contribute to about 75% of cardiovascular diseases, have been characterised by measuring the backscattering and attenuation coefficients using OCT [4]. Distinguishing breast tumour through computational methods on OCT

images has also been done [5]. Some other studies include successfully monitoring the changes in skin after UV radiation [9] as well as in the evaluation of sun damage and pre-cancer diagnosis [10]. Other than skin cancer OCT has been applied in the field of oncology for detection of tumours in the breast, bladder, brain, gastrointestinal, respiratory, and reproductive tracts, skin, ear, nose, and throat. In many of these studies it has been able to identify the malignancy in the very early stages which often is the key to patient survival. OCT has also been explored as an imaging technique for 3-D cell activities in tissue models with the potential to extend imaging of scaffolds to in vivo applications, such as following grafting of engineered tissues into a host tissue [11].

Of particular importance to us is the use of OCT in dermatology, where OCT has been used to visualize the different layers of the skin such as stratum corneum of glabrous skin in palmoplantar locations, the epidermis and the upper dermis of skin and skin appendages and blood vessels. This has led to it being a powerful technique for skin cancer detection. Furthermore OCT has been shown to be useful in non-invasive monitoring of cutaneous inflammation etc.

The aim of our work is to determine how well our commercially bought OCT System (Swept Source OCT (SS-OCT), Thorlabs) can indentify skin alterations as well as it's sensitivity to the influence of skin tone on such detection.

To begin this investigation we will look at a large range of skin simulating phantoms with known optical properties i.e absorption and reduced scattering coefficients (μ_a and μ'_s). The phantoms are manufactured in-house and measured on an Integrating Sphere (IS) system [12,13]. These phantoms are then imaged using the OCT system and reconstructed to give a 3D representation of the area imaged.

2. Methodology

2.1 Preparation of phantoms

Solid resin phantoms are prepared by adding carbon black and TiO particles to a hardener-resin mixture (Aka-cure, slow, Aka-resin, Akasel) [14,15]. The phantoms are cut into discs of varying thickness. For these experiments discs with a thickness of 2mm were used.

2.2 Determination of Optical Properties

The optical properties of the phantoms are determined by using the Integrating Sphere system. In this particular case, the optical properties at 632.8 nm are determined. The calibration model at this wavelength has been validated against other models, for a He-Ne laser and a white light source(WL-S) for single sphere configurations, and shown to have low prediction errors [12,14,15]. Furthermore currently there are many therapeutic lasers with a red wavelength range. However the optical properties can be extracted for different wavelengths as we use a supercontiniuum white light source (Koheras) with wavelength ranging from 500nm to 1750nm. Shown in figure 1 is the IS setup used for reflectance and transmittance measurements on the phantoms to determine the optical properties. The details of obtaining the properties are described in [12,13,16].

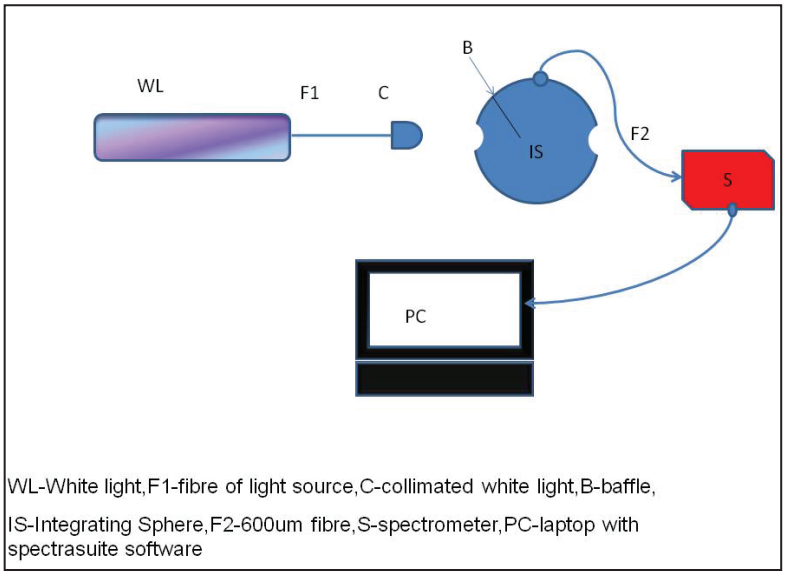


Figure 1. Single Integrating Sphere setup.

2.3 Taking an image

The system used is a Swept Source OCT (SS-OCT) with a resolution of 10-20 μm. The system is a frequency domain system with a centre wavelength of 1325nm. The magnitude and time delay of reflected light is measured and used to construct depth profiles (A-scans). Adjacent A-scans are then synthesized to create an image. The object to be scanned is placed on the path of the microscope. An volume of 10x10x3 mm^3 (x,y,z) was imaged. The contrast and brightness for all images taken was kept the same. 3D reconstructions of these images were then made.

2.4 Samples Imaged

Four 2mm phantoms with different optical properties, a 4.2mm phantom with 1mm holes, a standard pavilion and the index fingers of 2 individuals (two different skin types) were imaged.

2.5 Ethics

Ethics clearance has been obtained for this work. ETHICS REFERENCE NO: N11/01/009.

3 Results

The optical properties of the 4 phantoms were determined using the IS measurements. The measurements were done using a He-Ne laser and the white light. Table 1 shows the optical properties predicted at 632.8nm for both set of measurements. The predictions for the WL-S model are consistently higher than the He-Ne laser.

Table 1: Comparison of the predicted μ_a and μ'_s for the solid phantoms using the different models.

Sample	μ_a He-Ne	μ_a WL-S	μ'_s He-Ne	μ'_s WL-S
A1	0.041	0.05	13.27	14.50
A3	0.11	0.17	10.64	11.90
B1	0.07	0.08	5.47	6.84
B3	0.08	0.12	5.09	6.11

Based on this data these phantoms were shown to have different optical properties. These phantoms and the other objects were then imaged as described in 2.3 and 2.4. The 3D reconstructions of these images are shown in Figures (2)-(11).

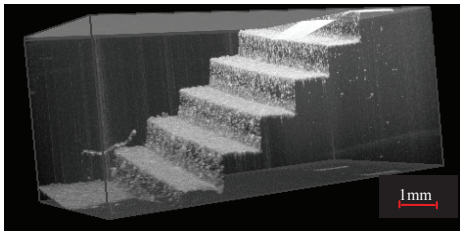


Figure 2. Pavillion.

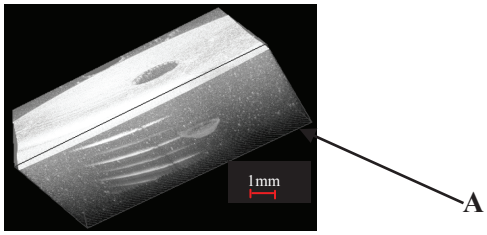


Figure 3. Phantom with holes.

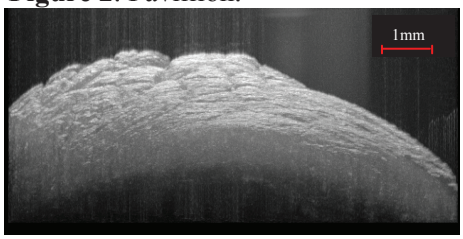


Figure 4. left index top (person A).

Figure 5. left index top (person B).

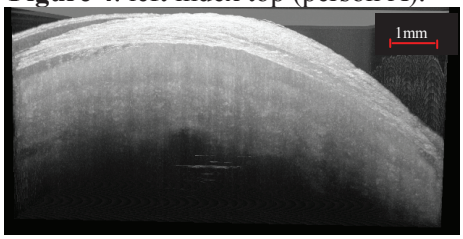


Figure 6. left index bottom (person A).

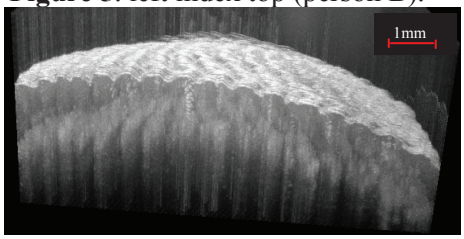


Figure 7. left index bottom (person B).

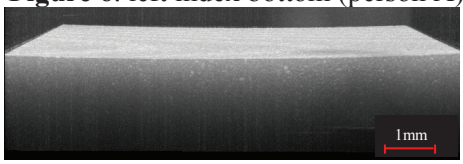


Figure 8. Phantom A1.

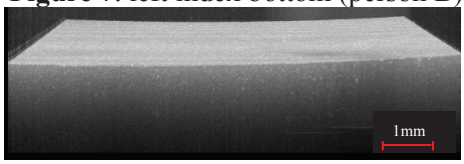


Figure 9. Phantom A3.

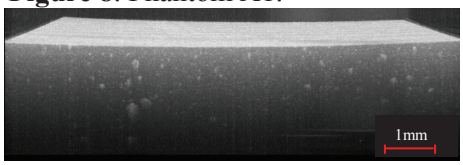


Figure 10. Phantom B1.

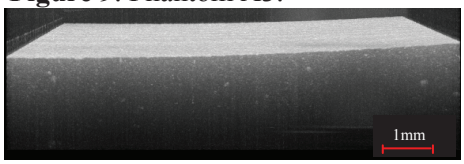


Figure 11. Phantom B3.

Figure 2 clearly shows the steps on a pavilion. Each step size is 0.5mm. This sample is used as a quick test to confirm that the system is working properly. Figure 3 shows the 4.2mm phantom with a hole. Although one can clearly see the hole with the eye, the depth of the hole is not visible to the eye. However reconstructing the image shows us the depth of the hole (point A). Currently the depth has not been evaluated, however this image shows us that it may be possible to view and determine the depth of an object such as a tumour imbedded in skin. Figures 4 and 5 compare the images, taken from the topside, on the left index finger of A and B. Figures 6 and 7 compare the images, taken from the bottom, on the left index finger of A and B. Figures 7 and 5 show more clearly defined layers of the skin than seen in figures 4 and 6. This suggests to us that the skin tone may influence the images

obtained by OCT. However such conclusions can only be made once a study of a larger data set is done. Furthermore one would need to determine whether the differences are due to skin tone or age, skin structure etc. The differences in the images obtained for the same person show that the area imaged also influences the data due to the composition of the skin in the area.

Phantom B1 (Figure 10) appears to show the most detail through the sample and phantom A3(Figure 9) the least. Phantom B1 has the third lowest absorption and scattering properties whereas phantom A3 has the highest absorption and second highest scattering properties. Phantom B3 with the lowest scattering and second highest absorption properties shows more clarity than phantoms A1 and A3. Finally phantom A1 (Figure 8) with the lowest absorption and highest scattering appears to show a little more detail than phantom A3. These results are interesting because OCT is based on the scattering properties of samples. It is the light reflected of the sample that creates the image seen. However these results show that both the absorption and scattering properties influence the images. These are encouraging results but need to be re-evaluated with a much larger study.

4 Conclusions

These preliminary findings support our hypothesis that skin tone may influence the OCT images obtained. It is premature to state that we have proven our hypothesis to be true until the larger study has been completed.

Acknowledgments

We would like to acknowledge Bafana Moya for preparing the phantoms used in these experiments.

References

[1] Levitz D, Thrane L, Frosz M, Andersen P, Andersen C, Andersson-Engels S, Valanciunaite J, Swartling J, and Hansen P 2004 Opt. Express **12** 249
[2] Fercher, AF, Drexler W, Hitzenberger CK and Lasser T 2003 Rep. Prog. Phys. **66** 239
[3] Huang D, Swanson E A, Lin CP, Schuman JS, Stinson WG, Chang W, Hee MR, Flotte T, Gregory K, Puliafito CA and Fujimoto JG 1991 Science **254** 1178
[4] Faber D, van der Meer F, Aalders M and van Leeuwen T 2004 Opt. Express **12** 4353
[5] Zysk AM, Nguyen FT, Oldenburg AL, Marks DL and Boppart SA Oct 2007 J. Biomed. Opt. **12** 051403
[6] Chenyang Xu, Joseph M. Schmitt, Stephane G. Carlier and Renu Virmani 2008 J. Biomed. Opt. **13**(3) 034003
[7] Mujat M, Ferguson RD, Hammer DX, Gittins C and Iftimia N 2009 J. Biomed. Opt. **14** 034040
[8] Mogensen M, Nürnberg B, Forman J, Thomsen J, Thrane L and Jemec G 2009 British Journal of Dermatology **160** 1026
[9] Gambichler T, Boms S, Stucker M, Moussa G, Kreuter A, Sand M, Sand D, Altmeyer P and Hoffmann K 2005 Arch Dermatol Res **297** 218
[10] Korde VR, et al. 2007 Lasers Surg Med **39** 687
[11] Tan W, Oldenburg AL, Norman JJ, Desai TA and Boppart SA 2006 Optics Express **14** 7159
[12] Singh A, Karsten AE 2011 Proc. SPIE 8192, 81924Z-1
[13] Singh A, Karsten AE and Dam JS 2008 Proc. of the World Association of Laser Therapy (WALT), Sun City, South Africa 4
[14] Karsten AE, Singh A and Braun MW (published online April 2011) Lasers Med Sci, DOI 10.1007/s10103-011-0926-x
[15] Karsten AE and Singh A 2011 Proc. of SPIE Vol. **8192** 81924U-1
[16] Singh A, Karsten AE, Mputle I, Chetty A and Naidoo K 2009 Proc. SPIE **7373** 737317 doi:10.1117/12.831882

Indoor temperature predictions in an energy efficient solar house

S Ziuku[1] **and E L Meyer**
Institute of Technology, University of Fort Hare, Private bag X1314, Alice 5700, South Africa.

E-mail: sziuku@ufh.ac.za

Abstract. This paper presents results of long term temperature monitoring in an energy efficient solar house at the University of Fort Hare, South Africa. Measured data was stored by a datalogger every 10 minutes. Formulas for predicting the daily indoor maximum, average and minimum temperatures were developed on the basis of outdoor climatic parameters. Passive solar housing aims to raise and lower indoor temperatures in winter and summer respectively. As a result, analysis of the data and development of predictive formulas of indoor temperature were done separately on part of the winter and summer seasons. The models were then validated against measurements taken in different time periods. Results indicated that indoor maximum, average and minimum temperatures can be predicted on the basis of outdoor temperature. Prediction of maximum indoor temperature was improved by incorporating daily solar irradiance in the formula. It was also revealed that indoor temperatures are affected by outdoor temperatures of the previous three days.

1. Introduction

Desirable thermal performance of energy efficient solar housing is based on two objectives: raising the indoor temperatures in winter using solar radiation as the principal heating source and lowering indoor temperatures in summer through natural ventilation and minimising solar gains [1]. Indoor temperature is influenced by the outdoor weather patterns, thermal absorption, thermal capacity of the house envelope and thermal mass, and presence/absence of heat sources. Human activity also influences the indoor thermal behaviour. Together with wind speed and relative humidity, indoor temperature is often used to determine thermal comfort [2]. The energy efficient solar house was built at the University of Fort Hare and its thermal performance has been monitored since 2009.

The majority of national meteorological stations measure and keep records of daily maximum, minimum and average temperatures among other climatic factors. This data is often published through public media and is made available to researchers and interested parties. This article analyses the influence of outdoor weather factors on indoor temperatures. Measured outdoor weather data was used to develop formulae for predicting indoor maximum, minimum and average temperatures. The objective of the exercise was to develop formulae that use the least amount of measured outdoor data which give a reasonable description of indoor temperatures.

[1] To whom any correspondence should be addressed.

2. Measurements

Outdoor and indoor temperature and relative humidity were measured by a HMP50 probe. Global solar irradiance and wind speed were measured by a Li-Cor pyranometer and wind sentry set respectively. Data acquisition was made using a CR1000 datalogger which recorded average values every ten minutes. Measurements were taken throughout the winter and summer seasons of 2009 and 2010.

3. Indoor temperature sensitivity to outdoor weather

Linear regression was used to analyze the relationship between indoor temperature and each of the measured outdoor variables. For example, the measured outdoor ambient temperature was plotted against the indoor temperature. The best fit linear regression correlation for June 2009 is shown in figure 1. It was observed that both indoor and outdoor temperatures do not exceed 27.00°C while the outdoor temperature gets to subzero values. The outdoor temperature range was 29.24°C (varying from -2.46°C to 26.78°C) while the indoor range was 20.05°C, that is 46% lower.

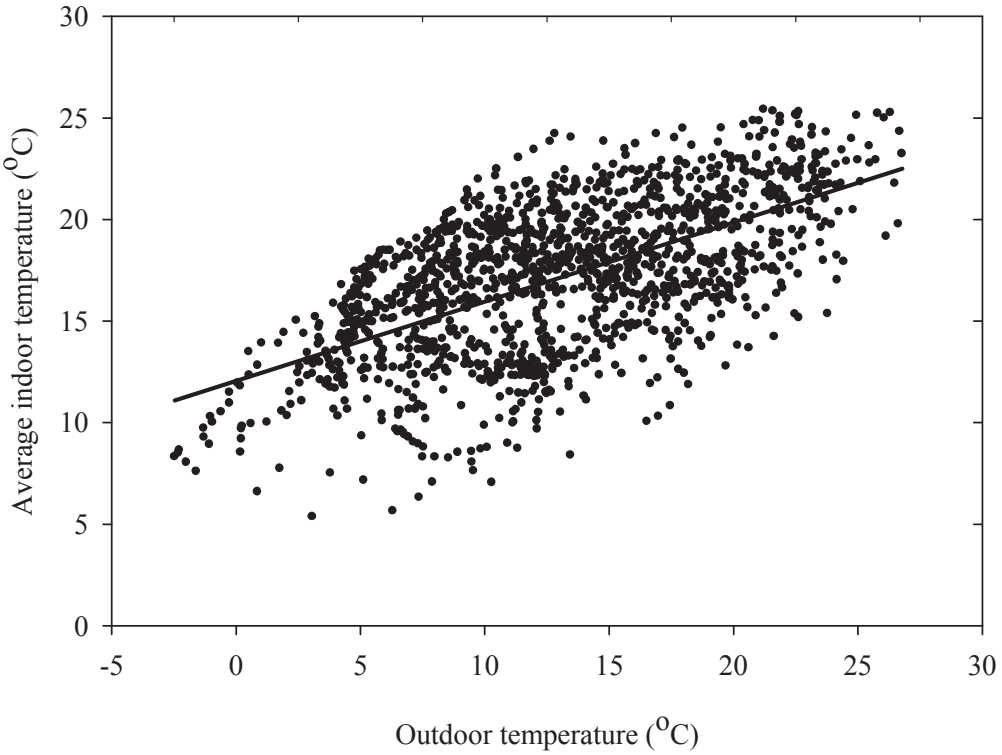


Figure 1. Correlation of indoor and outdoor temperature for June.

The correlation between the indoor and outdoor temperatures was deduced to be:

$$T_{in} = 0.39T_{out} + 12.06 \qquad R^2 = 0.37 \tag{1}$$

The relationship has a correlation coefficient $R = 0.61$ and a coefficient of determination $R^2 = 0.37$. This means that only 37% of the variation of indoor temperature about its mean can be explained using outdoor temperature. This correlation is weak, implying that T_{out} cannot individually be used to predict T_{in}. Other parameters account for the 63% deficit.

The indoor temperature (T_{in}) was also plotted against outdoor relative humidity (RH_{out}), wind speed (W_s), and solar irradiance (G). The results of the correlations are summarized in table 1.

Table 1. Correlations of indoor temperature with outdoor variables

Y-variable	X-variable	R^2	Comment
T_{in}	RH_{out}	0.10	No correlation
T_{in}	W_s	0.03	No correlation
T_{in}	G	0.06	No correlation

It was observed that individual climatic factors cannot independently predict indoor temperature. Combinations of independent climatic factors and the previous day(s) climatic conditions have to be used to model indoor temperature.

4. Indoor maximum, minimum and average temperature predictions

The energy efficient solar house was designed to perform differently in summer and in winter. It follows that the analysis and development of predictive formulae of indoor temperatures was done separately for the summer and winter periods. The measured data for each period was divided into two sub-groups. The first was used to generate the predictive formulae and the second dataset of the subgroup was used for validation as independent data.

Scatter and regression plots of indoor temperatures as dependant variables against outdoor temperatures were performed. Kruger and Givoni [3] and Ogoli [4] reported that indoor climate is influenced by that day's weather conditions and the weather conditions of the previous days. With the first day of the dataset taken to be day *n*, the outdoor temperatures were lagged by one (*n-1*), two (*n-2*) and three (*n-3*) days. The procedure involved analyzing the coefficient of determination R^2, the p-value and standard error of each indoor parameter plotted against various outdoor variables. The R^2 values greater than 0.5 were deemed to signify a valid linear relationship and a p-value less than 0.05 was considered to be statistically significant. Variables which gave correlation relations which were statistically insignificant were neglected. Correlations for $T_{o\text{-}max,\ n\text{-}3}$ and $T_{o\text{-}min,\ n\text{-}3}$ and lower, gave p-values which were statistically insignificant and R^2 values which were trivial. The relationship between indoor temperature and the moving average (*mav*) outdoor temperature of the previous days was also investigated.

The indoor temperature formulae generated for the winter period were:

$$T_{in\_\max,n} = 0.63T_{o\_\max,n} + 0.051T_{o\_\max,n-1} + 0.19T_{o\_mav} + 0.29G + 3.76 \quad (2)$$

where T_o is the outdoor temperature
T_{in} is the indoor temperature
max is maximum
min is minimum, and
mav is moving average.

The correlation coefficient for the generation period was 0.87 and for the validation period was 0.80.

$$T_{in\_ave,n} = 0.28T_{o\_\max,n} + 0.14T_{o\_\max,n-1} + 0.05T_{o\_\max,n-2} + 0.15T_{o\_\min,n} + 0.23T_{o\_mav} + 3.60 \quad (3)$$

The correlation coefficient for the generation period was 0.90 and for the validation period was 0.80.

$$T_{in-\min,n} = 0.20T_{o-\max,n-1} + 0.01T_{o-\max,n-2} + 0.51T_{o-\min,n} + 0.13T_{o-\min,n-1} + 0.11T_{o-mav} + 2.83 \quad (4)$$

The correlation coefficient for the generation period was 0.87 and for the validation period was 0.80.

For the three formulae generated in winter, the indoor maximum temperature is the only one which appears to be influenced by the daily solar irradiance. The daily indoor minimum temperatures are not

affected by the outdoor maximum temperature of the same day. The minima cannot be physically affected by the maxima which were observed to occur about seven hours later. Minimum temperatures generally occurred just after sunrise in winter. Figures 2 and 3 show the measured and computed indoor maximum and minimum temperatures during the winter and summer seasons respectively. Equipment failure caused data loss in the period 12 to 19 July.

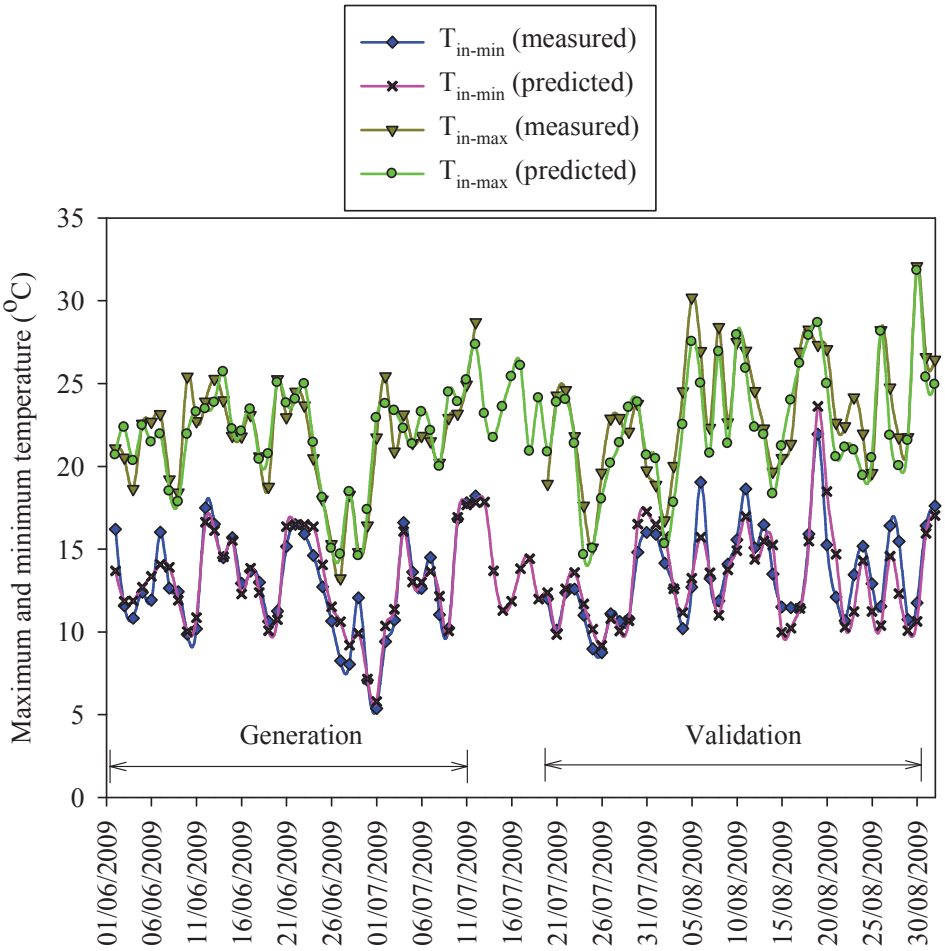


Figure 2. Measured and predicted indoor minimum and maximum temperatures in winter.

The indoor temperature formulae generated for the summer period were:

$$T_{in-\max,\,n} = 0.47T_{o-\max,\,n} + 0.23T_{o-\max,\,n-1} + 0.19T_{o-mav} + 0.19G + 6.37 \quad (5)$$

The correlation coefficient for the generation period was 0.91 and for the validation period was 0.80.

$$T_{in-ave,\,n} = 0.26T_{o-\max,\,n} + 0.29T_{o-\max,\,n-1} + 0.16T_{o-\min,\,n} + 0.14T_{o-mav} + 0.13G + 5.79 \quad (6)$$

The correlation coefficient for the generation period was 0.94 and for the validation period was 0.91.

$$T_{in-\min,\,n} = 0.37T_{o-\max,\,n-1} + 0.05T_{o-\max,\,n-2} + 0.37T_{o-\min,\,n} + 0.13T_{o-mav} + 3.77 \quad (7)$$

The correlation coefficient for the generation period was 0.92 and for the validation period was 0.87.

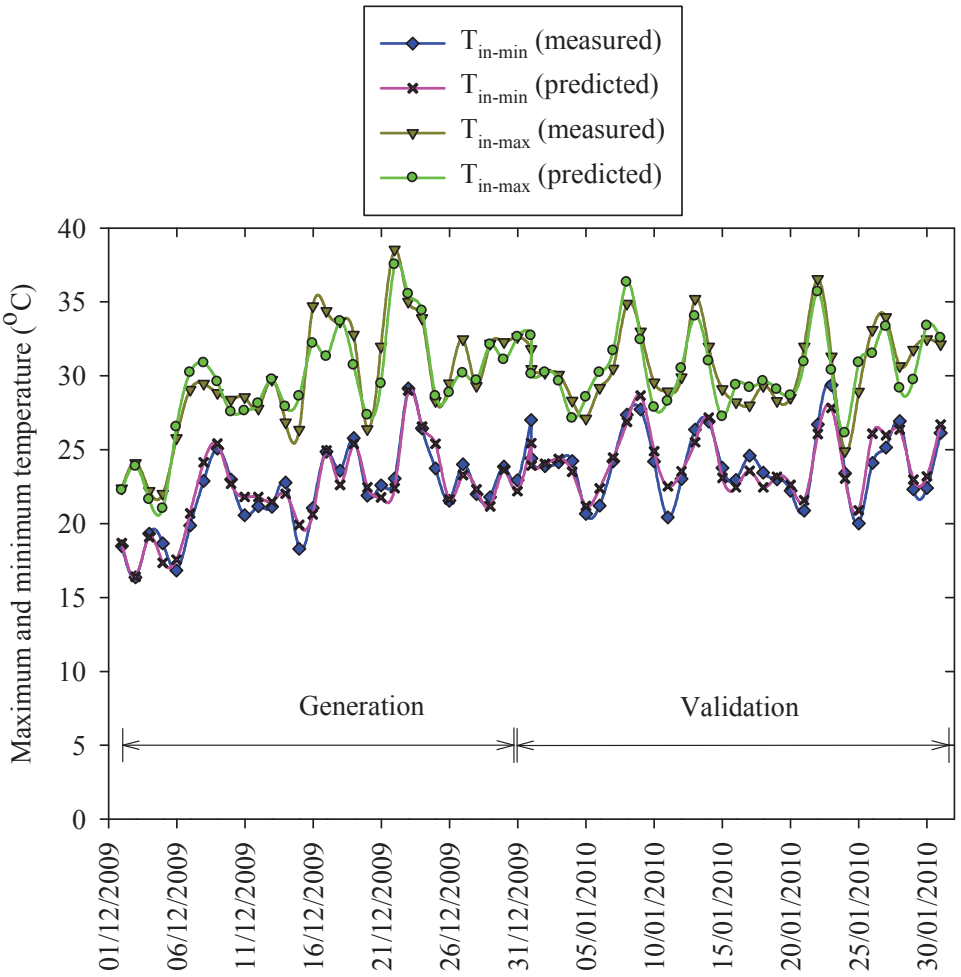


Figure 3. Measured and predicted indoor minimum and maximum temperatures in summer.

The relationship between the average indoor temperature and predicted average temperature is shown graphically in the scatter plot of figure 4 in the winter season.

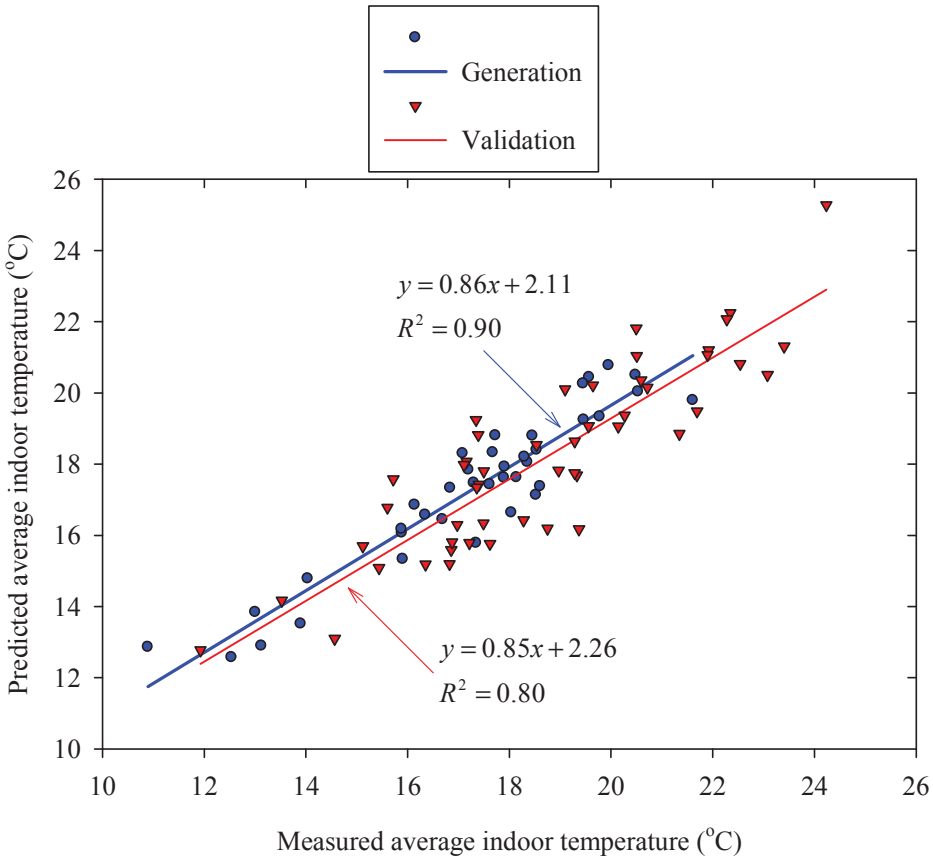


Figure 4. Scatter plot of indoor measured and predicted average temperatures.

Indoor temperatures have been shown to respond directly to outdoor climatic conditions. This can be attributed to the fact that the solar house does not have heating, ventilation and air conditioning equipment. Outdoor temperature has been proven to be a parameter that can serve as a basis for indoor temperature predictions for this type of buildings. In addition, the amount of climatic data required in these formulae is only a fraction of that required to run simulations and computerized models.

5. Conclusion

The developed models have maximum temperatures recorded two days earlier and a three day moving average parameters. Correlations with moving averages with n greater than four gave R^2 values which were trivial. Thus, the thermal mass retains heat received extending to the third previous day. Records of outdoor temperature can be used as a basis for predicting indoor maximum, minimum and average temperature. However, it is envisaged that the developed formulae can be improved by including the thermo-physical properties of the building components.

6. References

[1] Givoni B 1998. *Climate considerations in building and urban design* (John Wiley & Sons Inc.)
[2] De Dear R and Brager G S 2002. Thermal comfort in naturally ventilated buildings: revisions to the ASHRAE Standard 55 *Energy and buildings* **34** 549-561.
[3] Kruger E and Givoni B 2008. Thermal monitoring and indoor temperature predictions in a passive building in an arid environment *Building and environment* **43** 1792-1804.
[4] Ogoli D M 2003. Predicting indoor temperatures in closed buildings with high thermal mass *Energy and buildings* **35** 851-862.

Division G – Theoretical and Computational Physics

Wave–packet scattering off a soliton

A.M.H.H. Abdelhady and H. Weigel

Physics Department, Stellenbosch University, Matieland 7602, South Africa

E-mail: abdelhady@sun.ac.za, weigel@sun.ac.za

Abstract. We investigate the scattering of a wave–packet off a soliton in the (1+1) dimensional kink model. We solve the classical, time-dependent field equation numerically subject to the initial condition that the wave–packet is widely separated from the kink solution at very early times and propagates towards the soliton. After some time the wave-packet interacts with the static soliton and departs from it at later times. At very late times the wave-packet is finally again separated from the soliton. We then extract the scattering matrix from the distorted wave-packet and compare it to the known result from the static scattering calculation. This constitutes a first step towards studying crossing symmetry in soliton models, i.e. in a framework beyond perturbation theory.

1. Introduction

Non–linear field theories often contain non–perturbative solutions with a localized energy density. These solutions are called *solitons* and are quasi–static in the sense that the center and the shape of the energy density have simple time dependences. These soliton solutions [1] possess a wide range of applications in physics. To present an incomplete list, this range covers cosmology [2, 3], particle and nuclear physics [4, 5], as well as condensed matter physics [6]. The prototype soliton model is the kink in one time and one space dimension[1] for the real scalar field ϕ

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\,\partial^\mu\phi - \frac{\lambda}{4}\left(\phi^2 - \frac{M^2}{2\lambda}\right)^2 . \tag{1}$$

Here M is the mass parameter and λ is the coupling constant. The equation of motion reads

$$\ddot{\phi}(x,t) = \phi''(x,t) - \lambda\left(\phi^2(x,t) - \frac{M^2}{2\lambda}\right)\phi(x,t)\,. \tag{2}$$

For simplification we have introduced the partial derivatives $\dot{\phi} = \partial\phi/\partial t$ and $\phi' = \partial\phi/\partial x$. The kink soliton

$$\phi_{\rm K}(x) = \frac{M}{\sqrt{2\lambda}}\tanh\left(\frac{M}{2}x\right) . \tag{3}$$

is a static solution to eq. (2) with boundary conditions at spatial infinity that connect the two vacuum configurations $\phi_0 \equiv \pm\frac{M}{\sqrt{2\lambda}}$, these connections define conserved topological charges [8]. They arise from the spontaneous breaking of the reflection symmetry $\phi \leftrightarrow -\phi$. Time dependent

[1] Soliton solutions in one space dimension have early been reviewed in ref. [7].

solutions can easily be constructed by an appropriate Lorentz transformation [1]. Typically small amplitude fluctuations $\eta(x,t)$ about the solution, $\phi(x,t) = \phi_{\rm K}(x) + \eta(x,t)$ are introduced to quantize this classical configuration. The first non–trivial order of the expansion in powers of the wave–function $\eta(x,t)$ yields a Schrödinger type equation with a potential induced by the kink. Standard techniques [1] yield the momentum dependent phase shift

$$\delta(k) = 2\arctan\left(\frac{3Mk}{2k^2 - M^2}\right) . \tag{4}$$

for the scattering solution. In addition there are two bound state solutions with energies $\sqrt{3}M/2$ and $\sqrt{2}M$. Here we present a calculation for this phase shift that is based on the direct solution of the equation of motion (2) in coordinate space with suitable initial conditions that reflect the physical situation of particle scattering. These methods can then be utilized to study features beyond the small amplitude approximation. Most of the results presented here are part of a more comprehensive paper [9].

2. Wave–packet

We want to simulate the scattering of a particle off the kink. To this end we consider a localized wave–packet that represents the particle far away from the kink. We impose initial conditions such that the wave–packet propagates towards the kink, interacts with it and then moves away from it. Finally we extract the scattering data (in particular the phase shift) from the structure of the wave–packet at a very late time.

We build the wave–packet as a superposition of solutions to the free equation of motion in the small amplitude approximation. These solutions are plane waves with the dispersion relation $\omega_k = \sqrt{k^2 + M^2}$,

$$\eta_{\rm wp}(x) = \int_{-\infty}^{\infty} dk\, A(k)\, {\rm e}^{ikx} . \tag{5}$$

At $t = 0$ the dispersion relation enters only via the velocity of the initial wave–packet

$$\dot\eta_{\rm wp}(x) = -i \int_{-\infty}^{\infty} dk\, \omega_k\, A(k)\, {\rm e}^{ikx} . \tag{6}$$

In the context of our numerical simulations we assume the spectral function in momentum space to be of Gaußian shape

$$A(k) = a_0\, {\rm e}^{-\frac{(k-k_0)^2}{\sigma_k^2}} , \tag{7}$$

where a_0 is the amplitude, k_0 is the average momentum, and σ_k is the width of the distribution. This wave–packet defines the initial configuration

$$\phi(x,0) = \phi_{\rm bg}(x - x_0) + \eta_{\rm wp}(x) \qquad \text{and} \qquad \dot\phi(x,0) = \dot\eta_{\rm wp}(x) \tag{8}$$

for integrating the time–dependent equation of motion (2). The background configuration can be either the trivial vacuum, $\phi_{\rm bg}(x) = \frac{M}{\sqrt{2\lambda}}$ or the kink, $\phi_{\rm bg}(x) = \phi_{\rm K}(x)$ and $x_0 > 0$ denotes the position of its center. The small amplitude approximation can be neatly parameterized by the limit $a_0 \to 0$. For sufficiently large k_0 the wave–packet will travel towards the background and will interact with it within a finite time interval [10].

For small enough a_0 the solutions in the kink background at late times (after the interaction with the kink is completed so that the overlap between the kink and the wave–packet can be safely ignored) are

$$\phi(x,t) = \phi_{\rm K}(x - x_0) + \int_{-\infty}^{\infty} dk\, A(k) \exp\left[i\left(kx - \omega_k t + \delta(k)\right)\right] + \mathcal{O}\left(a_0^2\right) . \tag{9}$$

We may converse this line of argument to extract the phase shift from the solution

$$\eta^{(S)}_{\rm wp}(x,t) = \phi(x,t) - \phi_{\rm K}(x - x_0) \tag{10}$$

to the equation of motion with initial conditions, eq. (8) at very late times $t_f \gg 0$,

$$\mathrm{e}^{i\delta(k)} = \frac{\mathrm{e}^{i\omega_k t_f}}{A(k)} \int_{-\infty}^{\infty} \frac{dx}{2\pi}\, \mathrm{e}^{-ikx} \eta^{(S)}_{\rm wp}(x,t_f)\,. \tag{11}$$

The dependence on t_f will cancel, at least in the small amplitude approximation.

Though the initial configuration is motivated by the small amplitude approximation, we emphasize that our solutions to the time–dependent equation of motion are not restricted to this condition.

3. Numerical results

Before addressing details of the numerical results, we remark that conservation of the total energy

$$E = \int_{-\infty}^{\infty} dx\, \epsilon(x,t) \qquad \text{with} \qquad \epsilon(x,t) = \frac{1}{2}\left(\dot\phi^2(x,t) + \phi'^2(x,t)\right) + \frac{\lambda}{4}\left(\phi^2(x,t) - \frac{M^2}{2\lambda}\right)^2 \tag{12}$$

is the most important criterion for accepting a numerical solution. Note that the above mentioned initial conditions yield a complex field $\phi(x,t)$ and so is the corresponding energy. In eq. (12) we have also made explicit the time–dependent energy density, $\epsilon(x,t)$.

All results presented in this section will be measured with respect to the dimensionless quantities on the right hand side of

$$(x,t) \longrightarrow \frac{(x,t)}{\sqrt{2}\,M} \qquad \text{and} \qquad \phi \longrightarrow \frac{M}{\sqrt{2}\,\lambda}\,\phi\,. \tag{13}$$

Then a plane wave has mass $\sqrt{2}$. In figure 1 we present the physical energy density for the time–dependent solution that is characterized by restricting the initial wave–packet to its real part. To single out the wave–packet we particularly consider $\bar\epsilon(x,t) = \epsilon(x,t) - \frac{1}{2}\left[1 - \tanh^2(x - x_0)\right]^2$, where the subtraction refers to the energy density of the kink. Though we can clearly identify the propagation of the wave–packet, we also recognize an unexpected residual structure at the center of the kink. It represents a displacement of the kink [9]. This displacement increases with the initial amplitude a_0 and corresponds to an attraction shortly before and a repulsion shortly after the interaction with the wave–packet. The displacement can also be observed from the actual configuration in figure 2, where we display the real part of the solution to the equation of motion with complex initial conditions (8) with the kink background. The imaginary part exhibits a similar behavior, merely phase shifted by $\pi/2$. We can clearly identify the three phases of the propagation of the wave–packet: First the propagation of the wave–packet towards the kink, second the interaction as the wave–packet climbs up the kink, and third the distorted wave–packet moving away from the kink. From the field configuration in the third phase we will eventually extract the phase shift via eq. (11). Unfortunately there are numbers of numerical obstacles that we need to overcome first.

To begin with, we have to incorporate the above mentioned displacement of the kink in the integral, eq. (11). This is straightforwardly accomplished by restricting the integration interval to the regime of the wave–packet. For very late times ($t_f > 100$) this regime is clearly separated from the kink. Other obstacles are more cumbersome. Wave–packet components with small

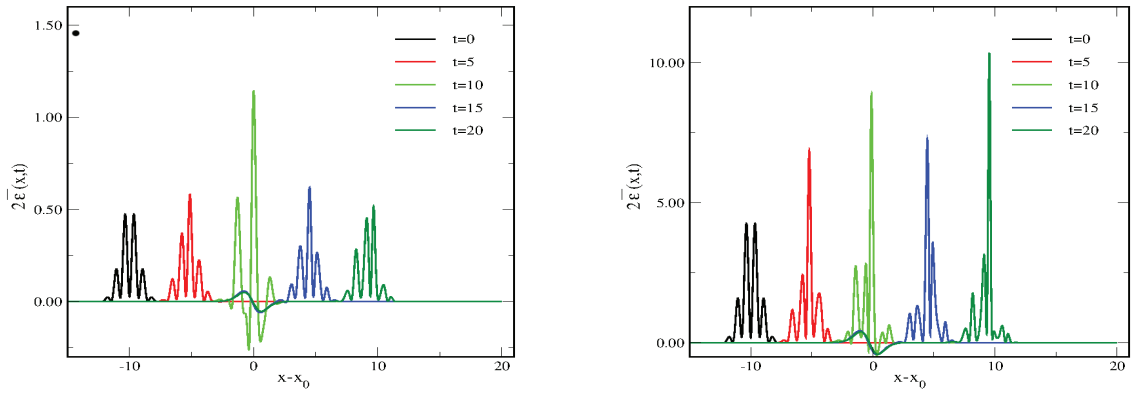


Figure 1. (Color online) Time snapshots of the energy density $\bar{\epsilon}(x,t)$ (in $M^4/8\lambda^2$) of the wave–packet for real initial conditions. We have used $k_0 = 4$ and $\sigma_k = 2$. Left panel: $a = 0.05$, right panel: $a = 0.15$. Note the different scales on the ordinate.

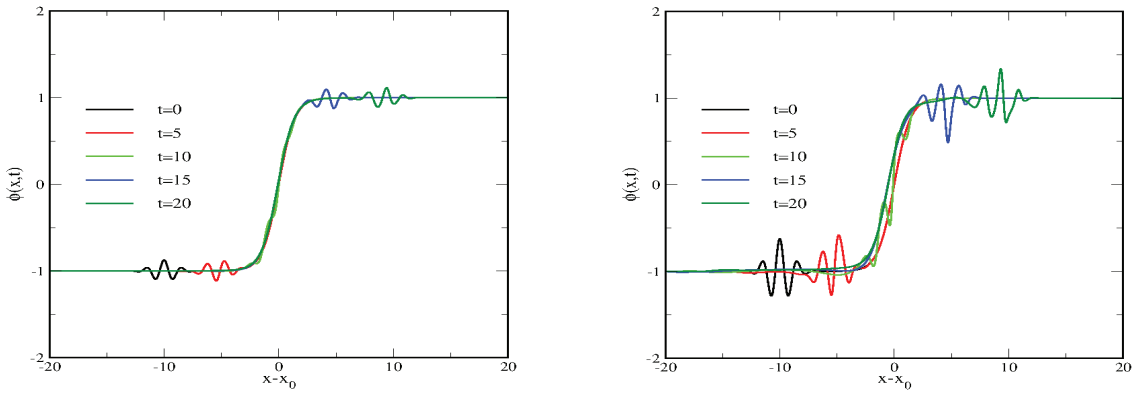


Figure 2. (Color online) Time snapshots of the real part of the field configuration in the kink background.

momenta take a long time to finalize the interaction with the kink. Hence we need to solve the equation of motion for a large interval on the time axis. We consider $t \in [0, 200]$ but also vary the upper limit to ensure stability of the results. However, components with large momenta will propagate a long distance in the same time interval. Hence we also need to consider a large interval in coordinate space. In order to reliably find the Fourier transform in eq. (11) we require a dense grid in coordinate space for large momentum components. This increases the numerical cost additionally. To keep the numerical effort within a manageable range, it is therefore appropriate to split the computation in (at least) two regimes in momentum space. To extract the phase shift for small momenta, we consider a large time interval but a small interval in coordinate space. This leads to unreliable results at large momenta. For that regime we consider a small time interval but a large one in coordinate space; together with a dense grid. At intermediate momenta the two procedures yield identical results. Furthermore we have the freedom to tune the parameters of the wave–packet, k_0 and σ_k to suit the considered regime in momentum space. These regimes of different numerical treatments are indicated in figure 3.

We also display the numerical result for the imaginary part of the right hand side of eq. (11). Its deviation from unity serves as a further test on the numerical accuracy. As expected this occurs for very small and very large momenta. Otherwise the agreement with the analytical

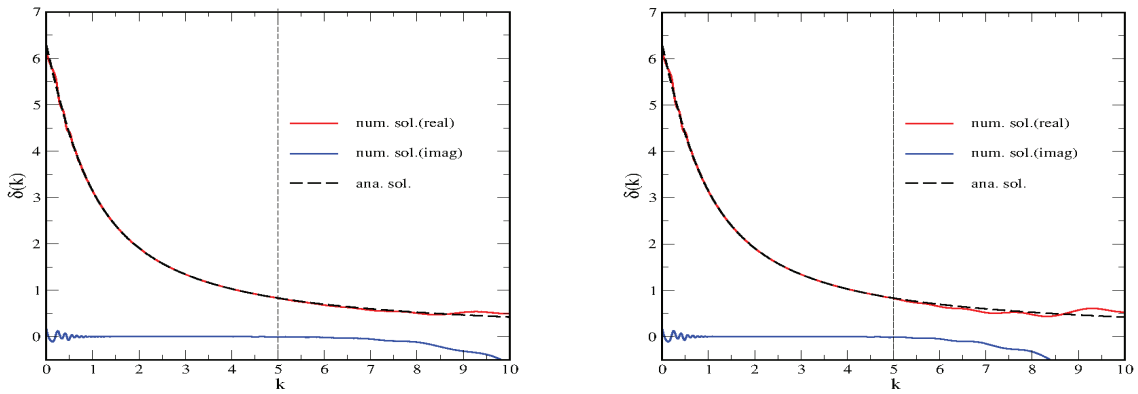


Figure 3. (Color online) Phase shifts extracted from the interaction between the kink and the wave–packet with amplitude $a_0 = 0.05$ (left) and $a_0 = 0.15$ (right). The vertical line indicate different numerical treatments as discuss in the text.

result, eq. (4) is astonishing. Certainly, further segmentation of the momentum axis and optimization in each segment will yield even better agreement. We have obtained the result displayed in figure 3 for $a_0 = 0.05$ in which case the small amplitude approximation is expected to be accurate. Figure 3 shows that we find identical phase shifts also for larger amplitudes, $a_0 \approx 0.15$. Though our computations show deviations from the small amplitude approximation in the imaginary part of the phase shift, we believe them to be short–comings of the numerical procedure and conclude that the scattering data do not exhibit non–linear effects even at moderate amplitudes.

4. A non–linear effect

Above we have already identified the displacement of the kink as a non–linear effect. This does not occur in the small amplitude approximation. However, as we increase the amplitude of the wave–packet the center of the kink gets slightly shifted towards the initial position of the wave–packet. As we further increase the amplitude, the sign of the displacement changes and eventually the kink co–moves with the (dominant piece of the) wave–packet after the interaction. This effect is clearly seen in figure 4 where we display both, the real parts of the subtracted energy density and the field as functions of time. In addition we observe that the wave–packet splits into various pieces that propagate with different velocities. This appears to be an indication of particle production.

5. Conclusion

The phase shift for the scattering of small fluctuations about the kink soliton has been known for quite a while. We have reconstructed this phase shift via the time–dependent simulation of the collision between the kink and a wave–packet of small and moderate amplitude. By choosing a significantly wide initial spectral function this turned out to be possible (by a single integration of the equation of motion), although numerical accuracy can be improved by optimizing both, numerical parameters as well as initial conditions to a particular momentum regime.

The wave–packet approach is more descriptive of the scattering process than the small amplitude approach. In addition this technique captures non–linear effects. Since the initial wave–packet is localized this numerical simulation also preserves the topological charge of the kink.

We have established that the wave–packet drags the kink after the collision for sufficiently

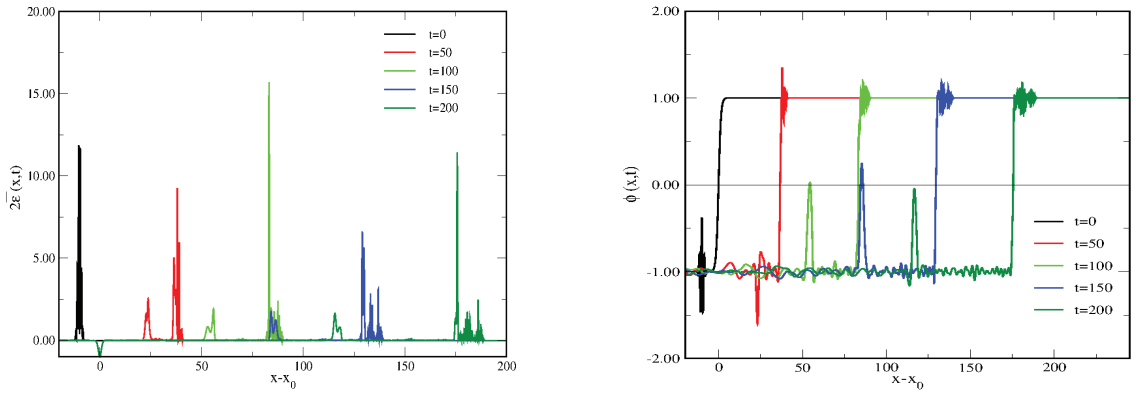


Figure 4. (Color online) Results for real initial conditions using $a_0 = 0.25$. Left panel: subtracted energy density $\bar{\epsilon}(x,t)$, right panel: field configuration $\phi(x,t)$.

large amplitudes. Furthermore the wave-packet splits into several pieces in this non–linear regime. These occur to be so far unknown features of the model.

The successful reproduction of the phase shift for scattering off the kink from the time–dependent equation of motion in configuration space is the first step towards establishing crossing symmetry in topological soliton model. The next step consists of assuming an initial condition that describes colliding kink–antikink configurations [11, 12] and relating the time–dependent field configuration to the one discussed here.

Acknowledgments

One of us (AMHHA) is supported in parts by the African Institute for Mathematical Sciences (AIMS).

References

[1] Rajaraman R 1982 Solitons and Instantons. *North Holland, Amsterdam*
[2] Vilenkin A and Shellard E P S 1994 Cosmic Strings and other Topological Defects, *Cambridge University Press, Cambridge (UK)*
[3] Vachaspati T 2006 Kinks and domain walls: An introduction to classical and quantum solitons, *Cambridge University Press, Cambridge (UK)*
[4] Manton N and Sutcliffe P 2004 Topological Solitons *Cambridge University Press, Cambridge (UK)*
[5] Weigel H 2008 *Lect. Notes Phys.* **743** 1.
[6] Bishop A R and Schneider T 1978 Solitons in Condensed Matter Physics, *Springer Verlag, Berlin*
[7] Scott A C , Chu F Y F and McLaughlin D W 1973 *Proc. IEEE* **61** 1443
[8] Finkelstein D and Rubinstein J 1968 *J. Math. Phys.* **9** 1762
[9] Abdelhady A M H H and Weigel H 2011 *Int. J. Mod. Phys. A* **26** 3625, *arXiv:*1106.3497v1 [nlin.PS]
[10] Hasenfratz W and Klein R 1977 *Physica A* **89** 191
[11] Demidov S V and Levkov D G 2011 *JHEP* **1106** 016, *arXiv*:1103.2133 [hep-th]
[12] Romanczukiewicz T and Shnir Y 2010 *Phys. Rev. Lett.* **105** 081601, *arXiv*:1002.4484 [hep-th]

Density-matrix renormalization group study of the electro-absorption in conjugated polymers

A O Abdelwhab[1], E Jeckelmann[2] and A Artoli[3]

[1] Physics Department, College of Science, Sudan University of Science and Technology, 11113, Khartoum, Sudan
[2] Institut für Theoretische Physik, Leibniz Universität Hannover, D-30167 Hannover, Germany
[3] Faculty of Science-School of Physics, Al Neelain University, 11121 Khartoum Sudan

E-mail: anasomer@sustech.edu, eric.jeckelmann@itp.uni-hannover.de, artoli@neelain.edu.sd

Abstract. Symmetrized and dynamical density-matrix renormalization group methods are used to study the optical properties of the one-dimensional half-filled extended Peierls-Hubbard model of π-conjugated polymers. We have computed the linear optical conductivity spectrum and the electro-absorption difference spectrum at strong, intermediate and weak coupling. In all cases the lowest optical excitation is an exciton whose binding energy and size we have determined. Only in the strong coupling regime we can see a clear separation between the exciton peak and the particle-hole continuum in the linear optical spectrum. However, we can observe the Stark effect under static electric field both at strong and intermediate coupling.

1. Introduction

Linear optical absorption is one of the most-used methods in experimental studies of the dynamical properties of materials such as π-conjugated polymers [1]. The dynamical DMRG (DDMRG) method is an accurate and reliable extension of the density-matrix renormalization group (DMRG) [2, 3] which makes possible the computation of dynamical properties in one-dimensional strongly-correlated electron models [4–7]. In this paper we report a symmetrized [8] and DDMRG study of the optical properties of the one-dimensional extended Peierls-Hubbard (EPH) model [9] which is a generic model for $\pi-$conjugated polymers [10–13].

2. Model and method

The one-dimensional extended Peierls-Hubbard model is often used to describe correlation effects in $\pi-$conjugated polymers. It is defined by the Hamiltonian

$$\begin{aligned}\hat{H} &= -t\sum_{l;\sigma}\left(1-(-1)^l\frac{\delta}{2}\right)\left(\hat{c}^\dagger_{l+1,\sigma}\hat{c}_{l,\sigma}+\hat{c}^\dagger_{l,\sigma}\hat{c}_{l+1,\sigma}\right)\\ &\quad +U\sum_l\left(\hat{n}_{l,\uparrow}-\frac{1}{2}\right)\left(\hat{n}_{l,\downarrow}-\frac{1}{2}\right)\\ &\quad +V\sum_l(\hat{n}_l-1)(\hat{n}_{l+1}-1),\end{aligned} \tag{1}$$

where the first term describes the electron hopping between nearest-neighbor sites, $\hat{c}^{\dagger}_{l,\sigma}$ and $\hat{c}_{l,\sigma}$ represent creation and annihilation operators for electrons with spin $\sigma =\uparrow,\downarrow$ on site l, t and δ are the hopping and dimerization parameters, respectively. The second term describes the on-site Coulomb repulsion with interaction parameter U and $\hat{n}_{l,\sigma}$ is the number operator for spin σ electrons. The third term describes the nearest neighbor repulsion with strength V and $\hat{n}_l = \hat{n}_{l,\uparrow} + \hat{n}_{l,\downarrow}$. Natural units are used: $a_0 = t = e = \hbar = 1$. We consider only the half-filled band case where the number of electrons N equals the number of lattice sites L. Half-filing in the ground state is guaranteed by the particle-hole symmetry of the above Hamiltonian. This Hamiltonian has also a spin-flip symmetry and a spatial reflection symmetry (through the lattice center). Therefore, each eigenstate has a well-defined parity for the charge conjugation ($P_c = \pm 1$) and spin flip ($P_s = \pm 1$) transformations and belongs to one of the two irreducible representations, A_g or B_u, of a one-dimensional lattice reflection symmetry group.

Using the DMRG method we can calculate ground states and thus the charge gap

$$E_{\rm c}(L) = E_0(L, L+1) + E_0(L, L-1) - 2E_0(L, L)\,, \tag{2}$$

where $E_0(L, N)$ is the ground-state energy of the Hamiltonian in Eq. 1 on a L-site lattice with N electrons. For $L \to \infty$, $E_{\rm c}$ gives the energy threshold of the electron-hole excitation continuum.

The optical absorption is proportional to the real part of the linear optical conductivity which is related to the imaginary part of the current-current correlation function by

$$\sigma_1(\omega > 0) = \frac{{\rm Im}\{\chi(\omega > 0)\}}{\omega}\,, \tag{3}$$

$$\chi(\omega > 0) = -\frac{1}{L}\langle 0|\hat{\jmath}\frac{1}{E_0 - \hat{H} + \hbar\omega + i\eta}\hat{\jmath}|0\rangle \tag{4}$$

$$= -\frac{1}{L}\sum_n \frac{|\langle 0|\hat{\jmath}|n\rangle|^2}{\hbar\omega - (E_n - E_0) + i\eta}\,. \tag{5}$$

Here, $\hat{\jmath}$ is the current operator, $\eta = 0^+$, $|0\rangle$ is the ground state, $|n\rangle$ are excited states, and E_0, E_n are their respective energies.

We note that the current operator is invariant under the spin-flip transformation but antisymmetric under charge-conjugation and spatial reflection. Therefore, if the ground state $|0\rangle$ belongs to the symmetry subspace (A_g, P_c, P_s), only excited states $|n\rangle$ belonging to the symmetry subspace $(B_u, -P_c, P_s)$ contribute to the optical conductivity. According to selection rules, the matrix element $\langle 0|\jmath|n\rangle$ vanishes if $|n\rangle$ belongs to another symmetry subspace. The optical gap is the energy difference $E_{\rm opt} = E_n - E_0$ between the lowest optically-allowed eigenstate $|n\rangle$ (i.e., the lowest eigenstate in the symmetry subspace contributing to the linear optical conductivity) and the ground state $|0\rangle$. In this paper we consider only the regime of the EPH model where optically excited states can be described as bound (excitons) or unbound particle-hole pairs. The exciton binding energy is usually defined as the difference $\delta E = E_c(L) - E_{\rm opt}$ between the optical gap $E_{\rm opt}$ and the band edge of the particle-hole continuum $E_c(L)$. We use the symmetrized DMRG method [8] to calculate the lowest optically-allowed excitation and the optical gap. This also makes possible the computation of static correlation functions for this state.

The DDMRG method [4] allows us to calculate dynamical correlation functions, such as the r.h.s. of Eq. 4, very accurately over the full frequency range for fairly large systems (here from $L = 32$ to 120 sites) with open boundary conditions and a *finite* broadening factor η (from $\eta = 0.4$ for 32 sites to $\eta = 0.1$ for 120 sites). Thus DDMRG actually yields

$$\sigma_{\eta;L}(\omega) = \frac{1}{L}\sum_n \frac{|\langle 0|\hat{\jmath}|n\rangle|^2}{E_n - E_0}\frac{\eta}{[\omega - (E_n - E_0)]^2 + \eta^2}\,. \tag{6}$$

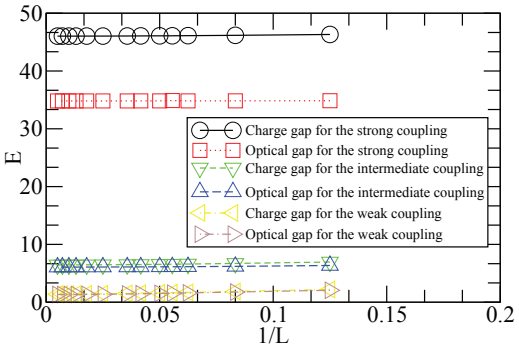


Figure 1. Charge and optical gaps versus the inverse of system lengths for strong, intermediate and weak couplings (see text).

For $\eta \to 0$, $\sigma_{\eta;L}(\omega)$ reduces to $\sigma_1(\omega)$ as defined in Eq. 3. A very useful consistency check of the method is to test various sum rules, relating moments of the function $\sigma_1(\omega)$ to ground-state expectation values, which can be evaluated with great accuracy using a standard DMRG method. For instance, for the Hamiltonian in Eq. 1 with open boundary conditions

$$\int_0^\infty \frac{d\omega}{\pi}\omega\sigma_1(\omega) \quad = \quad \frac{1}{L}\,\langle 0|\hat{\jmath}^2|0\rangle \ . \tag{7}$$

All DMRG methods have a truncation error which is reduced by increasing the number m of retained density matrix eigenstates (for more details, see Ref. [2]). Varying m allows one to compute physical quantities for different truncation errors and thus to obtain error estimates on these quantities. For all numerical results presented in this paper the number m of density-matrix eigenstates ranges between 64 and 256 and DMRG truncation errors are negligible.

3. Linear optical spectrum and excitons

We first investigate the binding energy by calculating the charge gap in Eq. 2 and the optical gap using symmetrized DMRG for three sets of parameters introduced in [13]. The results are presented in Figure 1. For very strong coupling ($\delta = 0.1, U = 50$ and $V = 15$) the optical and charge gaps are well separated and thus the lowest excitation is clearly an exciton with a large binding energy $\delta E = 11.2$. For intermediate coupling ($\delta = 0.1, U = 10$ and $V = 3$) and (relatively) weak coupling ($\delta = 0.1, U = 4$ and $V = 1.5$) the difference between charge gap and optical gap is much smaller but we still obtain finite values $\delta E = 0.37$ and $\delta E = 0.065$ in the thermodynamic limit $L \to \infty$. Thus we expect an exciton with a small binding energy to be present in the optical spectrum in the intermediate and weak coupling cases too.

These binding energy results are consistent with our analysis of the nature of the lowest optical excitation using correlation functions. To measure the exciton size we choose the correlation function for electron-hole excitations [6, 12]:

$$C_{\rm eh}(x) = \left|\left\langle n\left|\hat{P}_{l,l+x} + (-1)^{|x|}\hat{P}_{l+x,l}\right|0\right\rangle\right|^2 \ , \tag{8}$$

where $|0\rangle$ is the ground state, $|n\rangle$ is the excited state under investigation (here the lowest optical excitation), and the operator $\hat{P}_{i,j} = \sum_\sigma \hat{c}^+_{i,\sigma}\hat{c}_{j,\sigma}$ creates an electron at site i and a hole at site j. Obviously, $C_{\rm eh}(x)$ evaluates the importance of an electron-hole pairs with distance x in the excited state $|n\rangle$. This correlation function is shown in figures 2 and 3 for our three sets of couplings. In the exact result for the strong-coupling limit $U > 2V \gg t$, the ground state consists in localized electrons (one on each site) while the lowest optical excitation is a nearest-neighbor pair made of an empty site and a doubly-occupied site with excitation energy $E_{\rm opt} = U - V$ [6, 7]. Thus for our strong-coupling parameters we expect the lowest optical excitation to be a very small exciton. Indeed, we see in figure 2 that $C_{\rm eh}(x)$ vanishes as soon

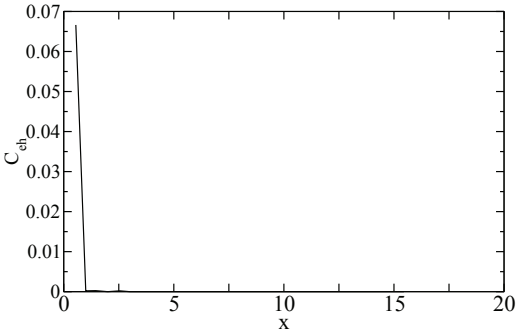


Figure 2. Electron-hole correlation function $C_{\rm eh}(x)$ (see eq. 8) in the strong-coupling case.

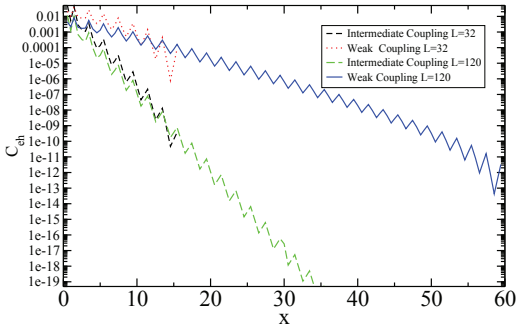


Figure 3. Same as figure 2 for two system sizes $L = 32$ and 120 for the intermediate- and weak-coupling cases.

as $x > 1$. For intermediate and weak couplings we observe that $C_{\rm eh}(x)$ decreases exponentially with x which confirms the presence of a bound state.

Rather than fitting the function $C_{\rm eh}(x)$ to an exponential we define and calculate the average electron-hole distance $\zeta_{\rm eh} = \sum_x C_{\rm eh}(x)|x| / \sum_x C_{\rm eh}(x)$. For strong coupling we have obtained $\zeta_{\rm eh} = 0.56$ (in units of the lattice constant a_0) while for intermediate coupling $\zeta_{\rm eh} = 1.17$, confirming that the excited state consists in a tightly bound electron-hole pair in both cases in qualitative agreement with the exact result in the strong-coupling limit $U > 2V \gg t$ [6, 7]. For weak coupling the exciton size varies with the the system length L but this finite-size effect is negligible for the present discussion. We have obtained $\zeta_{\rm eh} = 3.34$ with 32 sites and $\zeta_{\rm eh} = 4.10$ with 120 sites. Thus the electron-hole pair extends over a few lattice constants. We note that the exciton size increases only by a factor 10 from strong to weak coupling although the binding energy decreases by a factor 200. This confirms that there is no simple relation between an exciton binding energy and its size in correlated systems [6].

Using this analysis of exciton binding energies and sizes we can interpret the optical spectrum of the EPH model. The reduced optical conductivity spectra $\omega\sigma_1(\omega)$ calculated with DDMRG are shown in figures 4, 5, and 6. For strong coupling figure 4 clearly shows a strong exciton peak around $\omega \approx 35$ which accounts for most of the total spectral weight. The position of this peak agrees with the optical gap $E_{\rm opt} = 34.8$ obtained with symmetrized DMRG. The logarithmic scale allows us to see a weak continuum which is well-separated from the exciton peak and starts at the energy given by the charge gap $E_c = 46.0$. For intermediate coupling figure 5

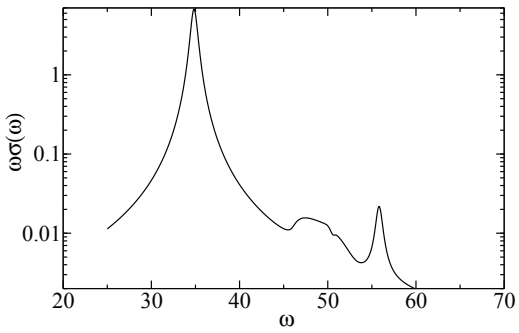


Figure 4. Reduced optical conductivity $\omega\sigma_1(\omega)$ for a 32-site lattice ($\eta = 0.4$) at strong coupling ($U = 50$, $V = 15$).

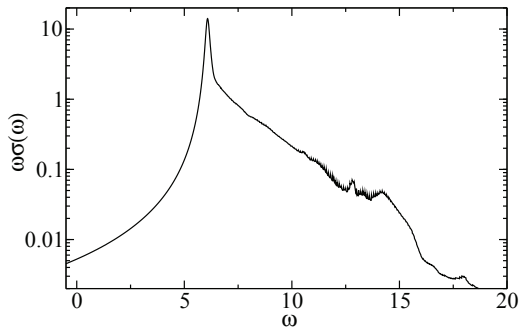


Figure 5. Reduced optical conductivity $\omega\sigma_1(\omega)$ for a 120-site system ($\eta = 0.1$) at intermediate coupling ($U = 10$, $V = 3$).

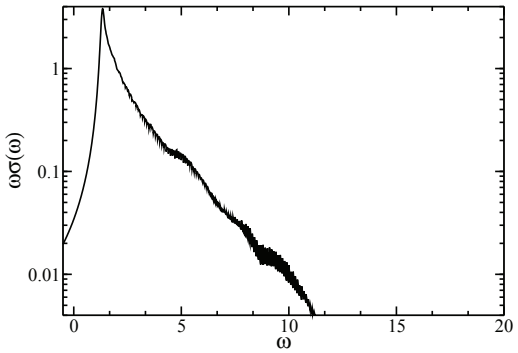


Figure 6. Reduced optical conductivity $\omega\sigma_1(\omega)$ for a 120-site system ($\eta = 0.1$) at weak coupling ($U = 4$, $V = 1.5$).

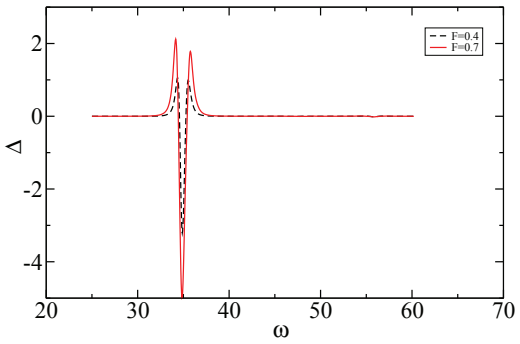


Figure 7. Electro-absorption difference spectrum at strong coupling (32 sites, $\eta = 0.4$) for $F = 0.4$ and $F = 0.7$.

also clearly shows the exciton peak around $\omega \approx 6$ in agreement with the energy predicted by the symmetrized DMRG results for the optical gap $E_{\rm opt} = 6.08$. In that case, however, the continuum has substantial spectral weight and we do not see the separation between the exciton and the continuum because the binding energy $\delta E = 0.37$ is comparable to the broadening $\eta = 0.1$. Finally, in the weak-coupling case figure 6 shows that the continuum is indistinguishable from the exciton peak, which is relatively weak. The maximum around $\omega = 1.4$ is compatible with the optical gap $E_{\rm opt} = 1.33$. However, from the optical spectrum (figure 6) alone one could not decide whether there is an excitonic peak or the continuum has some singularities at the onset. Nevertheless, our analysis of the correlation function Eq. 8 confirms that the lowest optical excitation is an exciton. We emphasize that the symmetrized DMRG results for the lowest optically excited state are always in perfect agreement with the DDMRG results for the linear optical conductivity, confirming the accuracy of both methods.

4. Electro-absorption

In electro-absorption experiments the optical conductivity defined in Eq. 3 is measured in the presence of a constant electric field F. The difference spectrum $\Delta(\omega)$ between the optical conductivity for $F \neq 0$ and for $F = 0$ provides information about the non-linear optical properties of the system including optically-forbidden excited states. We have calculated this difference spectrum in the EPH model using the DDMRG method. To avoid that finite-size effects (i.e., the discrete energy spectrum) dominate the difference spectrum we have used fields from $F = 0.1$ to $F = 0.7$ which are much larger than those used in experiments.

In figure 7 we show the electro-absorption difference spectrum for the strong-coupling case. The left structure (below $\omega \approx 35$) is due to the Stark shift of the $1B_u$ exciton state while the right one (above $\omega \approx 35$) can be explained as the optically-forbidden mA_g exciton state because it is located in the bound-state energy range (i.e., below the continuum onset at $E_c = 46.0$) [13]. We note that these two structures become stronger with higher electric field F. Thus the EPH ground state is stable even at such high field as $F = 0.7$ thanks to its large gap. For the intermediate coupling figure 8 shows a clear splitting and shift for the lower field ($F = 0.1$). Again the low-energy structure (below $\omega \approx 6.2$) is due to the Stark shift of the $1B_u$ exciton. The broad peak and oscillations observed above $\omega \approx 6.2$ are explained by the optically-forbidden mA_g exciton state and excitations in the particle-hole continuum which are known to give rise to an observable oscillatory signal in the difference spectrum [13]. For the higher electric field ($F = 0.5$) one cannot interpret the difference spectrum. The presence of a signal at very low energy $\omega \approx 0.3$ well below the optical gap $E_{\rm opt} = 6.08$ demonstrates that the ground state has

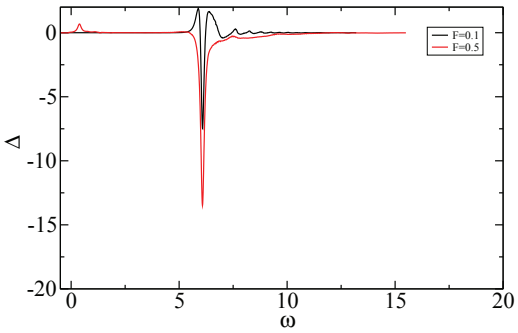


Figure 8. Electro-absorption difference spectrum at intermediate coupling (120 sites, $\eta = 0.1$) for $F = 0.1$ and 0.5.

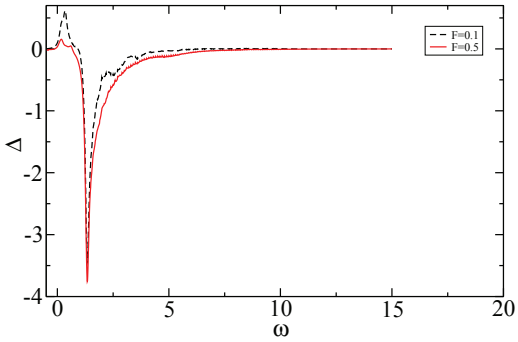


Figure 9. Electro-absorption difference spectrum at weak coupling (120 sites, $\eta = 0.1$) for $F = 0.1$ and 0.5.

been modified by the presence of the electric field. Thus this electric field is already too strong for an insulator with a charge gap $E_c = 6.45$. In the weak coupling figure 9 shows that even a field of $F = 0.1$ already destabilises the ground state of the EPH model with a gap of $E_c = 1.39$. Therefore, we have not been able to gain additional information about the optical properties using electro-absorption in that case.

5. Conclusion

We have investigated the optical properties of the extended Peierls-Hubbard model using various DMRG techniques to compute the exciton binding energy, electron-hole correlations, linear optical spectrum and electro-absorption difference spectrum. In the strong-coupling regime all results confirm that the lowest optical excitation is a small exciton with a high binding energy as predicted by an exact analysis in the strong-coupling limit. In the intermediate-coupling regime we cannot observe the separation between exciton peak and electron-hole continuum in the linear optical spectrum directly because of the DDMRG broadening $\eta > 0$. However, all other results confirm that the lowest optical excitation is a small exciton with a low binding energy. Finally, in the weak-coupling regime we have found a very small but finite binding energy and an extended but finite electron-hole separation which confirm the presence of an excitonic state. However, we could not observe any clear signature of this exciton neither in the linear optical spectrum nor in the electro-absorption spectrum.

References

[1] Kuzmany H 1998 *Solid State Spectroscopy* (Berlin: Springer)
[2] Peschel I, Wang X, Kaulke M and Hallberg K 1999 *Density Matrix Renormalization* (Berlin: Springer)
[3] White S R 1992 *Phys. Rev. Lett* **69** 2863
[4] Jeckelmann E 2002 *Phys. Rev. B* **66** 045114
[5] Jeckelmann E 2008 *Progress of Theoretical Physics Supplement* **176** pp 143–164
[6] Essler F H L, Gebhard F and Jeckelmann E 2001 *Phys. Rev. B* **64** 125119
[7] Jeckelmann E 2003 *Phys. Rev. B* **67** 075106
[8] Ramasesha S, Pati S K, Krishnamurthy H R, Shuai Z and Bredas J L 1996 *Phys. Rev. B* **54** 7598
[9] Hubbard J 1963 *Proc. R. Soc. London* **Ser. A 276** 238
[10] Baerswyl D, Campell D K and Mazumdar S 1992 *Conjugated Conducting Polymers* ed Kiess H (Berlin: Springer)
[11] Shuai Z, Pati S K, Su W P, Bredas J L and Ramasesha S 1997 *Phys. Rev. B* **55** 115368
[12] Rissler J, Gebhard F and Jeckelmann E 2005 *J. Phys.* Condens. Matter **17** 4093
[13] Guo D, Mazumdar S, Dixit S N, Kajzar F, Jarka F, Kawabe Y and Peyghambarian N 1993 *Phys. Rev. B* **48** 1433

Representation of the quantum and classical states of light carrying orbital angular momentum

Humairah Bassa[1] **and Thomas Konrad**[1,2]

[1] School of Physics, University of KwaZulu-Natal, Private Bag X54001, Durban 4000
[2] National Institute of Theoretical Physics, South Africa

E-mail: konradt@ukzn.ac.za

Abstract. Ever since the discovery by Allen *et al.* that beams which carry orbital angular momentum could be easily realized in the laboratory, there have been many classical and quantum optical experiments that utilize this new optical degree of freedom. It is therefore necessary to familiarize oneself with the representation of the quantum and classical states of light in order to explain the results of these experiments. Classical paraxial light is described by a superposition of analytical solutions to the paraxial wave equation. However, due to the well-known analogy between quantum mechanics and paraxial optics we can apply operator algebra to the classical states of a light beam. A short review of the applications of this formalism is given. For the quantum-mechanical description of light, we quantize the paraxial light beams and this leads to a description of the photon in terms of creation operators acting on the vacuum state. The analogy between the quantum and classical states of light is analyzed.

1. Introduction

It is well-known from Maxwell's theory that light carries both energy and an associated momentum, which consists of linear and angular contributions. In 1909, Poynting proved that circularly polarized light carries spin angular momentum, which is related to the polarization of the light beam. This was experimentally verified by Beth, 20 years later, using a quarter-wave plate suspended from a fine quartz fiber. Although the concept of a light beam possessing orbital angular momentum is not new, it was only in 1992 that Allen *et al.* [1] realized that it was possible to easily produce a light beam possessing a well-defined orbital angular momentum in the laboratory. In their seminal work they showed that, within the paraxial approximation, any monochromatic beam with an azimuthal phase dependence of $\exp(il\phi)$ will possess an orbital angular momentum of $l\hbar$ per photon. Since then there have been many experimentalists who utilize this new optical degree of freedom in novel experiments of both quantum and classical optics.

In this article we take the viewpoint that quantum and classical optics offer adequate descriptions of the properties of light in different physical regimes. Operator algebra and Dirac notation can be employed in the quantum description of paraxial light beams and, surprisingly, also in their classical description, as we will review in Section 2. The application of Dirac notation in classical optics is advantageous because it allows one to calculate quantities associated with the field in a very simple way as well as to analyze the behavior of light beams during propagation without going into detailed calculations. In Section 3 we quantize the field of a paraxial light beam in order to obtain a description of a single-photon state in terms of creation operators

acting on the vacuum state. Finally, in Section 4 we discuss the analogy but also the difference between the representation of the classical and quantum states of light.

2. Operator algebra for laser modes

Maxwell's equations can be combined to give scalar wave equations for each component Ψ of the electric and magnetic fields of a light beam of the form,

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\Psi = 0. \tag{1}$$

Our aim is to describe a laser beam which is a collimated, monochromatic light beam whose transverse beam profile, $u(x, y, z)$, varies much more slowly along the z-axis (which is the axis of propagation) than it does along the x- and y-axes. In this case we use the separation of variables $\Psi(x, y, z, t) = u(x, y, z)\exp[i(kz - \omega t)]$ and invoke the paraxial approximation (where we neglect $\frac{\partial^2}{\partial z^2}$, because of the slow variation, but not $2ik\frac{\partial}{\partial z}$) which leads to the paraxial wave equation,

$$2ik\frac{\partial}{\partial z}u(x, y, z) = -\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)u(x, y, z). \tag{2}$$

We notice two properties of this equation that gave Stoler [2] a strong motivation to apply the operator formalism (which is used in quantum mechanics to describe the time-evolution of a quantum state) to the longitudinal propagation of classical light beams. Firstly, the form of the paraxial wave equation (Eq. (2)) is analogous to the two-dimensional time-dependent Schrödinger equation for a free particle with the z coordinate replacing the time variable and the wavelength replacing Planck's constant. Secondly, Eq. (2) has the explicit solution,

$$u(x, y, z) = \exp\left[\frac{i(z - z_0)}{2k}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\right]u(x, y, z_0), \tag{3}$$

which tells us that if we know the transverse profile of the beam at a point z_0 we can use the exponential factor to find the transverse profile of the beam at a different position z. This means that this exponential factor is analogous to the time-evolution operator that describes the time displacement of the wave function of a particle from time t_0 to t. We can therefore transform Eq. (3) into an operator relation in Dirac notation as follows.

In this notation we can represent the transverse beam profile $u(x, y, z)$ as a state vector $|u(z)\rangle$. Such a state vector lies in the Hilbert space $L^2(\mathbb{C}^2)$ and is normalized, obeying the identity,

$$\langle u(z)|u(z)\rangle \equiv \int d^2\mathbf{r}u^*(\mathbf{r}, z)u(\mathbf{r}, z) = 1 \tag{4}$$

where $\mathbf{r} = (x, y)^T$ is the position in the transverse plane. As well as showing the normalization of the state vectors, this equation shows how the inner product is defined on this space. When $|u(z)\rangle$ is normalized for one position z, it is automatically normalized for all other z values [5]. The basis vectors in this space are $|\mathbf{r}\rangle$, which are eigenvectors of the transverse position operator $\hat{\mathbf{r}} = (\hat{\mathbf{x}}, \hat{\mathbf{y}})$ that acts on state vectors in this space. If we take the inner product of the state vector $|u(z)\rangle$ with a basis vector we recover the beam profile,

$$u(x, y, z) = \langle\mathbf{r}|u(z)\rangle. \tag{5}$$

The differential operators $-i\frac{\partial}{\partial x}$ and $-i\frac{\partial}{\partial y}$ which act in the function space containing $u(x, y, z)$ as shown in Eq. (3) are represented by the momentum operators $\hat{\mathbf{p}}_x$ and $\hat{\mathbf{p}}_y$, respectively, in

this Hilbert space. The commutation relations between the position and momentum operators define a Heisenberg algebra and resemble the ones in quantum mechanics, with $\hbar$ set to 1. With these preliminaries, we may now express Eq. (3) in Dirac notation as,

$$|u(z)\rangle = \hat{\mathbf{U}}|u(z_0)\rangle, \tag{6}$$

where the unitary operator $\hat{\mathbf{U}}$, which governs the propagation of the state from position z_0 to z, is given by,

$$U = \exp\left[-\frac{i(z-z_0)}{2k}\hat{\mathbf{p}}^2\right] \tag{7}$$

with $\hat{\mathbf{p}}^2 = \hat{\mathbf{p}}_x^2 + \hat{\mathbf{p}}_y^2$. Please note that even though we are using the operator algebra, which is usually used in quantum mechanics, we are still in the purely classical regime.

The expectation value of an operator, $\hat{\mathbf{O}}$ is defined in the standard way as,

$$\langle O\rangle = \langle u(z)|\hat{\mathbf{O}}|u(z)\rangle = \int d^2\mathbf{r}u^*(\mathbf{r},z)\hat{\mathbf{O}}u(\mathbf{r},z). \tag{8}$$

In this equation the states are z-dependent and the operators are "stationary", which is the Schrödinger picture. We now move to the Heisenberg picture where the operators become z-dependent, so that the evolution of the expectation value of an operator can be expressed as,

$$\langle O\rangle = \langle u(z_0)|\hat{\mathbf{U}}\hat{\mathbf{O}}\hat{\mathbf{U}}|u(z_0)\rangle. \tag{9}$$

A complete orthogonal set of solutions to the paraxial wave equation (Eq. (2)) is given by the Laguerre-Gaussian (LG) modes, $u_{lp}(z)$. This means that the set of LG modes forms a basis for paraxial light beams, so any laser beam can be expanded as a superposition of LG modes. It is well-known that the analytic form of these modes resemble the wave functions of the stationary states of a two-dimensional quantum-mechanical harmonic oscillator. Therefore, Nienhuis and Allen [5] were able to use operator algebra to obtain ladder operators of the harmonic oscillator which allow one to generate higher-order LG modes from the fundamental Gaussian mode.

Let us consider an example given in [3] that shows how this operator formalism simplifies the calculations of quantities associated with the electric field. The full scalar version of the electric field, which includes the time dependence of it, can be written as,

$$\langle \mathbf{r}|E\rangle = E_0\exp[i(kz-\omega t)]\langle \mathbf{r}|u\rangle \tag{10}$$

where E_0 is the amplitude of the field and $u \equiv u(z)$ for convenience. If $u(\rho,\phi,z) = u_0(\rho,z)e^{il\phi}$, which is the form that an LG mode assumes in cylindrical coordinates, then the expectation value of the angular momentum operator is,

$$\langle L_z\rangle = \left\langle E\left|i\frac{\partial}{\partial\phi}\right|E\right\rangle = |E_0|^2 l\langle u|u\rangle. \tag{11}$$

The linear momentum of the state in the direction of propagation, z, is given by,

$$\langle p_z\rangle = \left\langle E\left|i\frac{\partial}{\partial z}\right|E\right\rangle = |E_0|^2 k\langle u|u\rangle, \tag{12}$$

while the energy of the state amounts to,

$$\langle W\rangle = \int\int dxdyE^*(\mathbf{r},t)\left(i\frac{\partial}{\partial t}\right)E(\mathbf{r},t) = |E_0|^2\omega\langle u|u\rangle. \tag{13}$$

Therefore the ratio of the orbital angular momentum carried by the wave to the energy that it carries is,

$$\frac{\langle L_z \rangle}{\langle W \rangle} = \frac{l}{\omega} \tag{14}$$

while the ratio of orbital angular momentum to linear momentum is,

$$\frac{\langle L_z \rangle}{\langle p_z \rangle} = \frac{l}{k}. \tag{15}$$

These relations are valid classically as well as quantum mechanically, since $\hbar$ does not appear in them. Every photon has an energy of $\hbar\omega$ and a linear momentum of $\hbar k$ so for the ratios to be valid on a single-photon level, each photon must possess a well-defined orbital angular momentum of $l\hbar$. Thus we have, in a simple way, arrived at the conclusion reached by Allen *et al.* [1], that every photon in an LG beam carries a quantized orbital angular momentum of $l\hbar$ per photon. Also note that due to Eq. (9) the quantities in Eqs. (11), (12) and (13) are all conserved during free space propagation since the operators in these equations commute with $\hat{\mathbf{p}}^2$ of the propagation operator $\hat{\mathbf{U}}$ [4].

This operator formalism that we have developed is not limited to free space propagation. Indeed many authors [2, 4, 5] have applied it to the propagation of beams through thin lenses. We can implement the effect of such a lens with a focal length f by means of an operator $\hat{\mathbf{T}}_{lens}$ that is a function of the transverse coordinate operators $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ [2]:

$$\hat{\mathbf{T}}_{lens} = \exp\left[-\frac{ik}{2f}(\hat{\mathbf{x}}^2 + \hat{\mathbf{y}}^2)\right]. \tag{16}$$

Let us consider a system which consists of two identical cylindrical lenses orientated along the x-axis with focal length f and separated from each other by a distance $2d$ with free propagation between the lenses. The propagation operator for this system is given by,

$$\hat{\mathbf{U}}_{cyl} = \hat{\mathbf{T}}_{lens}\hat{\mathbf{U}}\hat{\mathbf{T}}_{lens} = \exp\left[-i\frac{k\hat{\mathbf{x}}^2}{2f}\right]\exp\left[-i\frac{d\hat{\mathbf{p}}^2}{k}\right]\exp\left[-i\frac{k\hat{\mathbf{x}}^2}{2f}\right] \tag{17}$$

Using this propagation operator van Enk and Nienhuis [4] were able to prove, in a simple way, that if $d = f/\sqrt{2}$ then this setup converts an LG mode which contains orbital angular momentum into a beam that contains no orbital angular momentum. This setup is known as an Astigmatic Mode Converter, and in [6], Allen *et al.* arrived at this conclusion in a more complex way.

3. Quantum mechanical representation of classical light

The electric $\mathbf{E}$ and magnetic $\mathbf{B}$ fields can be written in terms of a vector potential, $\mathbf{A}$, and it is often easier to solve Maxwell's equations corresponding to this potential rather than those given for the $\mathbf{E}$ and $\mathbf{B}$ fields. The first step for the quantization of the electromagnetic field is to express the vector potential in terms of a basis. In the Coulomb gauge ($\nabla \cdot \mathbf{A} = 0$), the vector potential is usually written in the well-known continuous plane wave expansion. The second step is to convert the complex amplitudes that appear in the plane wave expansion into creation ($\hat{\mathbf{a}}_s^\dagger(\mathbf{k})$) and annihilation ($\hat{\mathbf{a}}_s(\mathbf{k})$) operators which obey the commutation relation $[a_s(\mathbf{k}), a_{s'}^\dagger(\mathbf{k}')] = \delta_{ss'}\delta^{(3)}(\mathbf{k}-\mathbf{k}')$. The vector potential becomes a quantum mechanical operator given by,

$$\hat{\mathbf{A}}(\mathbf{r},t) = \sum_s \int d^3\mathbf{k}\left(\frac{\hbar}{16\pi^3\varepsilon_0 c|\mathbf{k}|}\right)^{\frac{1}{2}}[\boldsymbol{\epsilon}_s(\mathbf{k})\hat{\mathbf{a}}_s(\mathbf{k})e^{i(\mathbf{k}\cdot\mathbf{r}-c|\mathbf{k}|t)} + \text{h.c}]. \tag{18}$$

where $\mathbf{k}$ is the wave-vector, $\boldsymbol{\epsilon}_s$ is the two-dimensional polarization vector, ε_0 is the permittivity of free space and h.c. refers to the hermitian conjugate of the first term. This is an often used quantization of the electromagnetic field, although it is not relativistically invariant. Moreover, the gauge condition ($\nabla \cdot \mathbf{A} = 0$) does not hold as an operator identity and has to be replaced by a weaker condition involving the expectation value. In the Lorentz gauge, this problem has been addressed by Gupta and Bleuler (cp. for example, [7]).

We want to find a representation of the $\mathbf{A}$ field for laser beams and a novel method to obtain this is provided by Calvo *et al.* [9]. They introduce a dispersion relation which allows them to write the wave vector as $\mathbf{k} = q\mathbf{q} + k_0(1-\theta^2)\mathbf{u}_z$ (cf. [8]) where $\mathbf{q}$ is the transverse wave vector and $\theta := \frac{q}{\sqrt{2}k_0}$ is the parameter that governs the degree of paraxiality, which is approximately equal to the divergence angle of the beam [8]. By substituting this into Eq. (18) and invoking the paraxial approximation (i.e. $\theta \ll 1$) they find,

$$A_P(\mathbf{r},t) = \sum_s \int_0^\infty dk_0 \left(\frac{\hbar}{16\pi^3\varepsilon_0 c k_0}\right)^{\frac{1}{2}} \int d^2\mathbf{q}[\boldsymbol{\epsilon_s}\hat{\mathbf{a}}_s(\mathbf{q},k_0)e^{ik_0(z-ct)}e^{i(\mathbf{q}\cdot r_\perp - k_0\theta^2 z)} + \text{h.c.}] \quad (19)$$

where θ^2 which appears in the second phase factor is retained because it is the only relevant paraxial contribution. The transverse basis of LG modes are a well-known orthonormal and complete basis set for paraxial light beams so the vector potential can now be expressed in this basis. By noticing that the structure of the $\mathbf{q}$-integrand in Eq. (19) resembles the paraxial angular spectrum, the second phase factor can be written as,

$$e^{i[\mathbf{q}\cdot\mathbf{r}_\perp - k_0\theta^2 z]} = \sum_{l,p} \mathcal{LG}^*_{l,p}(\mathbf{q}) LG_{l,p}(\mathbf{r}_\perp, z; k_0) \quad (20)$$

where $LG_{l,p}(\mathbf{r}_\perp, z; k_0)$ are the LG modes and $\mathcal{LG}_{l,p}(\mathbf{q})$ are the Fourier transformed profiles at $z = 0$. The proof of this relation and the expressions for these modes can be found in the Appendix of [9]. Thus Eq. (19) can be written as,

$$A_P(\mathbf{r},t) = \sum_{s,l,p} \int_0^\infty dk_0 \left(\frac{\hbar}{16\pi^3\varepsilon_0 c k_0}\right)^{\frac{1}{2}} [\boldsymbol{\epsilon_s}\hat{\mathbf{a}}_{s,l,p}(k_0)e^{ik_0(z-ct)}LG_{l,p}(\mathbf{r}_\perp, z; k_0) + \text{h.c.}] \quad (21)$$

where they introduced the LG mode annihilation operators,

$$\hat{\mathbf{a}}_{s,l,p}(k_0) = \int d^2\mathbf{q}\mathcal{LG}^*_{l,p}(\mathbf{q})\hat{\mathbf{a}}_s(\mathbf{q},k_0), \quad (22)$$

which satisfy the commutation relation $[\hat{\mathbf{a}}_{s,l,p}(k_0), \hat{\mathbf{a}}^\dagger_{s',l',p'}(k_0')] = \delta_{ss'}\delta_{ll'}\delta_{pp'}\delta(k-k_0')$. The Fock state of a photon in a linearly polarized, monochromatic beam in the LG basis is given by,

$$|\psi_s\rangle = \sum_{l,p} f_{l,p}\hat{\mathbf{a}}^\dagger_{s,l,p}(k_0)|0\rangle, \quad (23)$$

where the complex coefficients $f_{l,p}$ satisfy the normalization condition $\sum_{l,p}|f_{l,p}|^2 = 1$ and represent the probability of finding the photon in the state $|lp\rangle$. A general single-photon state of light may involve a superposition with different wave numbers, k_0. For experiments that utilize the orbital angular momentum of light beams and photons it is common to experimentally realize linearly polarized, monochromatic LG beams (with specific l and p values) in the laboratory so we should be able to represent the state of a photon in this mode. Such a state will contain a quantized orbital angular momentum of $l\hbar$ and can be written as,

$$|lp\rangle = \hat{\mathbf{a}}_{s,l,p}(k_0)|0\rangle = \int d^2\mathbf{q}\mathcal{LG}_{lp}(\mathbf{q})\hat{\mathbf{a}}^\dagger_s(\mathbf{q},k_0)|0\rangle. \quad (24)$$

where we have used Eq. (22).

4. Discussion

We have shown how to use operator algebra in order to represent paraxial light beams and to analyze the propagation of these beams, in the framework of classical optics, through free space and thin lenses. The algebra simplifies this kind of analysis and the calculations of quantities associated with the light beam. We then presented the quantization of a paraxial light beam and the representation of a single-mode Fock state in the LG basis.

Although Dirac notation is a standard technique used in quantum mechanics, it is not specific to it; it can be employed for any Hilbert space vectors. Since square integrable functions can be represented as vectors in a Hilbert space it is perfectly acceptable to apply this algebra to the beam profiles of the electric field, as we have done. We must therefore note that by using this algebra we have not transitioned to the quantum level; we are still in the classical regime. Our state vectors represent the transverse beam profile of the electric field which can be measured; they do not represent the wave function of a photon in that mode of the electric field. There does however, exist an analogy between these two physical objects.

Consider the state in Eq. (23) which is a single-mode Fock state in the LG basis. This state represents the creation of a photon or a single excitation with a probability amplitude $f_{l,p}$ in one LG mode u_{lp} with a specific l and p index and vacuum in all other modes (i.e. modes with different l and p values). As shown in [10], this state can be written as $|\psi\rangle = \int \langle \mathbf{r}|\psi\rangle |\mathbf{r}\rangle d^2\mathbf{r}$, where $\psi(\mathbf{r}) = \langle \mathbf{r}|\psi\rangle$ is the transverse wave function of the photon in position representation. The authors adopt operator algebra by denoting the wave function as the ket $|\psi(\mathbf{r}, z)\rangle$ and use the operators we have discussed in Section 2 to propagate this wave function in free space and through thin lenses. We therefore notice that optical elements function analogously on a single-photon level.

Although there is an analogy between the states of classical light beams and single photons, the difference becomes clear during measurement. Note that a general paraxial light beam in classical optics with state vector $|u(z)\rangle$ is a superposition of the LG mode functions, $|u_{lp}(z)\rangle$, since the LG modes form a basis for solutions of the paraxial wave equation of such beams (i.e. $|u(z)\rangle = \sum_{l,p} \alpha_{l,p}|u_{lp}(z)\rangle$). With $|\langle \mathbf{r}|u(z)\rangle|^2 = \sum_{l,p} |\alpha_{l,p}|^2 |\langle \mathbf{r}|u_{lp}(z)\rangle|^2$ we obtain the intensity of the electric field, which is the sum of the intensities $|\alpha_{l,p}|^2$ of each of the LG modes, u_{lp}. This is a measurable and predetermined quantity. However, in quantum optics the square of the single-photon state $|\psi_s|^2$ equals unity and the square of the coefficients $|f_{l,p}|^2$ gives only the probability of finding a photon in the mode u_{lp}.

Acknowledgments

The authors are grateful to National Institute of Theoretical Physics in South Africa which supported this work with a bursary for Humairah Bassa. They would also like to thank Sujit Choudhary for a very helpful conversation.

References

[1] Allen L, Beijersbergen M W, Spreeuw R J C and Woerdman J P 1992 *Phys. Rev.* A **45** 8185
[2] Stoler D 1981 *J. Opt. Soc. Am.* **71** 334
[3] Schechner Y Y and Shamir J 1996 *J. Opt. Soc. Am.* A **13** 967
[4] van Enk S J and Nienhuis G 1992 *Opt. Commun.* **94** 147
[5] Nienhuis G and Allen L 1993 *Phys. Rev.* A **48** 656
[6] Beijersbergen M W, Allen L, van der Veen H E L O and Woerdman J P 1993 *Opt. Commun.* **96** 123132
[7] Mandl F and Shaw G 1984 *Quantum Field Theory* (Chichester: Wiley) p 88
[8] Aiello A and Woerdman J P 2005 *Phys. Rev.* A **72** 60101
[9] Calvo G F, Picón and Bagan E 2006 *Phys. Rev.* A **73** 13805
[10] Tasca D S, Gomes R M, Toscano F, Souto Ribeiro P H and Walborn S P 2011 *Phys. Rev.* A **83** 52325

Optimised periodic and hyperchaotic modes of a triple pendulum

André E Botha[1] **and Guoyuan Qi**[2]

[1]Department of Physics, University of South Africa, P.O. Box 392, Pretoria 0003, South Africa
[2]F'SATI and Department of Electrical Engineering, Tshwane University of Technology, Private Bag X680, Pretoria 0001, South Africa

E-mail: bothaae@unisa.ac.za, qig@tut.ac.za

Abstract. Analytical equations of motion, in the form $\dot{\mathbf{x}} = f(\mathbf{x}, t)$, were derived for a damped harmonically driven triple plane pendulum. This form of the equations displayed the nature of the non-linear coupling and provided a basis for physical interpretation. It also facilitated the derivation of the Jacobian matrix in analytical form, an essential result for the accurate numerical computation of Lyapunov exponents. Sets of optimised initial conditions and parameters were derived by applying Nelder-Mead simplex optimisation to the calculated Lyapunov exponents. As an example of the method, it was used to calculate the initial conditions for a periodic mode of the un-damped pendulum. It was also used to demonstrate that the maximum Lyapunov exponent of the pendulum could be made to vary from zero, for a periodic mode, to above ten, for a hyperchaotic mode. Numerical simulations were coded in Python and used to visualise the results.

1. Introduction

Recently there has been a considerable interest in studying chaotic dynamical systems and exploiting their properties for possible technological applications. In this regard the triple pendulum, which is an easily realisable mechanical system, continues to attract interest. Pendulums of varying complexity, ranging from the textbook example of a simple pendulum to more complicated coupled systems, play an important role in mechanics; mainly, because they illustrate interesting nonlinear dynamical effects in ways that often can be analysed mathematically. Although it is not well known, chaotic dynamics can even be observed in a simple pendulum, when it is externally excited [1].

There have been a number of experimental and theoretical investigations aimed at understanding the stability of human gait (manner of stepping) through the use of inverted pendulum models [2, 3]. Experimental investigations of either simple or coupled electro-mechanically driven pendulums have been undertaken, with the view of developing more precise conditions for the onset of chaos in these systems [4, 5]. On the technological front, a triple pendulum suspension system has been used to seismically isolate optical components on the GEO 600 interferometric gravitational wave detector [6]. This invention has allowed the detector to achieve a seismic noise sensitivity level which is well below the level from thermal noise.

Coupled pendulums with obstacles have been used to model real mechanical systems that exhibit nonlinear phenomena such as resonances, jumps between different system states, various continuous and discontinuous bifurcations, symmetry breaking and crisis bifurcations, pools of

attractions, oscillatory-rotational attractors, etc. In reference [7], for example, it is shown that a triple pendulum model can provide insight into the real, highly-complicated dynamics of a piston connecting-rod crankshaft system.

By definition, hyperchaotic systems are non-linear systems with more than one positive Lyapunov exponent. In simple terms, hyperchaotic systems are more disordered than ordinary chaotic systems, and thus possess more complicated topological structures and dynamics. Although hyperchaos has often in the past been generated through the construction of purely mathematical systems of non-linear equations, it occurs only rarely in nature, with the exceptions of a few well know systems or else the electrical circuit implementations of, what would otherwise be, purely mathematical systems [8].

Even though the triple pendulum has been the subject of an ongoing investigation by Awrejcewicz *et al.* [1, 9, 10] and Kudra [11], the chaotic and hyperchaotic regimes of the pendulum have yet to be explored fully. In previous work, such as [10], the highest two positive Lyapunov exponents were reported to be 0.06 and 0.01 (see table 2 of reference [10]). Since the chaos-hyperchaos transition is defined by when the second Lyapunov exponent becomes positive, Awrejcewicz *et al.* [10] correctly classify the attractor as hyperchaotic; however, it the present work it will be shown that in fact much higher values for the maximum Lyapunov exponent (above ten) are possible through optimisation.

This paper is organised as follows. In section 2, the equations for a triple pendulum are derived in a form that is well suited for the analysis of chaotic dynamics and for the use of optimisation. In section 3, details of the computational implementation of the optimisation calculation and visualisation are presented. Preliminary results for optimised periodic and hyperchaotic modes of the pendulum are presented in section 4. In the concluding section 5, it is noted that optimisation may provide a useful way of controlling the degree of chaos in non-linear dynamical systems. The method developed here for the pendulum could thus find applications in other non-linear systems ranging from biophysics to information technology.

2. Theory and analysis of the governing equations

Figure 1 shows a visualisation of the triple plane pendulum. It was generated from a screen-shot of a three-dimensional animation of the pendulum, created by using the *Visual* module in the *Python* programming language [12]. The pendulum consists of a series of absolutely rigid bars which form the three links of the pendulum (shown in red, green and blue). Additional point-like masses are attached to the bottom of each link (shown in yellow).

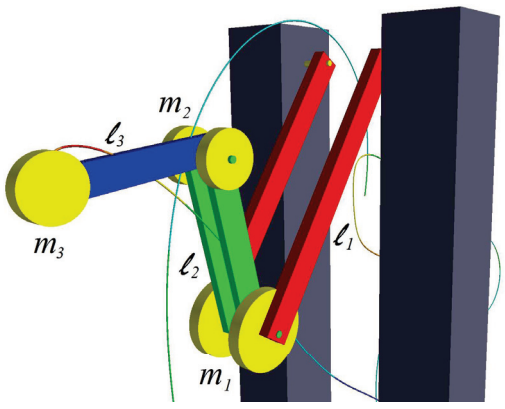


Figure 1. Visualisation of the triple plane pendulum. The pendulum is made of absolutely rigid bars (two of length ℓ_1, two of length ℓ_2 and one of length ℓ_3) to which cylindrical point-like masses are attached (two of mass $\frac{m_1}{2}$, two of mass $\frac{m_2}{2}$ and one of mass m_3). The pendulum is assumed to be under the influence of gravity and in vacuum. The three pivot points may exert viscous damping, with coefficients c_1, c_2 and c_3. Also shown in the figure is the trajectory followed by the centre of m_3.

The governing equations for the triple plane pendulum have been derived in reference [10]. They have the following matrix form:

$$\mathbf{M}(\boldsymbol{\psi})\,\ddot{\boldsymbol{\psi}} + \mathbf{B}(\boldsymbol{\psi})\,\dot{\boldsymbol{\psi}}^2 + \mathbf{C}\dot{\boldsymbol{\psi}} + \mathbf{D}(\boldsymbol{\psi}) = \mathbf{F}(t) \tag{1}$$

Equation (1) is written in the notation of reference [10], in which the angles between the vertical and each link of the pendulum have been denoted by ψ_i $(i = 1, 2, 3)$, with $\boldsymbol{\psi} = (\psi_1, \psi_2, \psi_3)^{\mathrm{T}}$.

Since the kinetic energy of the pendulum T is of the form

$$T = \frac{1}{2}\dot{\boldsymbol{\psi}}^T \mathbf{M}\dot{\boldsymbol{\psi}} \ , \tag{2}$$

where $\mathbf{M}$ is a (real symmetric) inertia matrix [10], the existence of $\mathbf{M}^{-1}$ is guaranteed by the positive definite property of T. Equation (1) can therefore be multiplied through by $\mathbf{M}^{-1}$ without any loss of generality. In the present work, a computer algebra system, *Maple 11*, was used to invert the matrix $\mathbf{M}$ in order to solve for $\ddot{\boldsymbol{\psi}}$ in equation (1). By then making the substitutions $x_1 = \psi_1$, $x_2 = \psi_2$, $x_3 = \psi_3$, $x_4 = \dot{\psi}_1$, $x_5 = \dot{\psi}_2$, and $x_6 = \dot{\psi}_3$, the equivalent first order system:

$$\dot{\mathbf{x}} = f\left(\mathbf{x}, t\right) \tag{3}$$

was obtained. Equations (3) are given explicitly in in Appendix A. Having analytical expressions for the governing equations in the form of (3) presents two important advantages.

Firstly, it can provide insight into the unusual nature of the non-linear coupling in this system. If one performs a power series expansion of the factor $1/D$ on right hand side of equations (A.4) to (A.6), one sees that the masses are coupled by terms involving products of sine and cosine functions (which are bounded by ± 1) with either constants or terms linear or quadratic in x_4 to x_6. Since x_4 to x_6 represent the angular speeds of the pendulum links, the increase in non-linear coupling is fundamentally limited by the total amount of mechanical energy stored in the pendulum. For a fixed amount of energy; however, the maximal coupling will be determined by the values of the physical parameters m_i and ℓ_i. In particular, one notices that all the coupling terms are inversely proportional to one of the three lengths, so that the coupling can be increased indefinitely by reducing one of the lengths. This is not the case for the mass parameters, which dimensionally tend to 'cancel' out in the coupling terms. For this reason changes in the mass parameters, as opposed to changes in length, have less of an effect on the overall coupling strength. This trend has been confirmed by numerical simulations.

Secondly, it is required in order to calculate the Lyapunov exponents accurately. The maximal Lyapunov exponent provides a quantitative measure of a non-linear system's sensitivity to initial conditions. Although a variety of methods have been developed for the accurate computation of Lyapunov exponents [13, 14], they invariably require analytical expressions for all the elements of the Jacobian matrix

$$J_{jk}(\mathbf{x}) = \frac{\partial f_j}{\partial x^k} \quad (j, k = 1, \ldots, 6) \tag{4}$$

In dynamical systems for which the analytical expressions are not know, the Lyapunov exponents can be estimated by, for example, fitting the time series with an analytical function through least square minimisation [15]. Note that it is not sufficient to calculate the Jacobian matrix numerically, because the sensitivity to numerical errors in calculations of this nature is extreme. Therefore, in this work, equations (A.1) to (A.6) were used to obtain the required analytical expressions for the Jacobian matrix.

3. Computational aspects and visualisation

A visualisation of the pendulum was developed in the *Python* programming language [12], using the *Visual* module. In order to improve the efficiency of the numerical calculations, which involve solving the system (3) in real time, the numerically intensive parts of the calculation were first coded as two separate Fortran 90 subroutines. The first subroutine was used to calculate the Lyapunov exponents according to the method of Chen *et al.* [14], and the second to integrate the system of equations by calling the ISML subroutine *dverk.f90*. Note that the first subroutine

makes calls to the second in order to integrate the linearized system of $n(n+1)$ equations ($n = 6$) and returns the Lyapunov exponents in base e, as is the modern convention. A shared library of the two main subroutines was created by using *f2py* [16]. *Scipy.optimize* was imported to perform downhill simplex optimisation via the function *fmin* [17].

4. Results and discussion

In this section two examples of the optimisation technique, applied to equations (A.1) to (A.6), are provided. In both examples the same arbitrarily chosen starting parameter values and initial conditions were used.

In the first, a periodic mode of the pendulum was calculated, with the result shown in figure 2. This mode was obtained by optimising the initial conditions, with the starting parameters fixed.

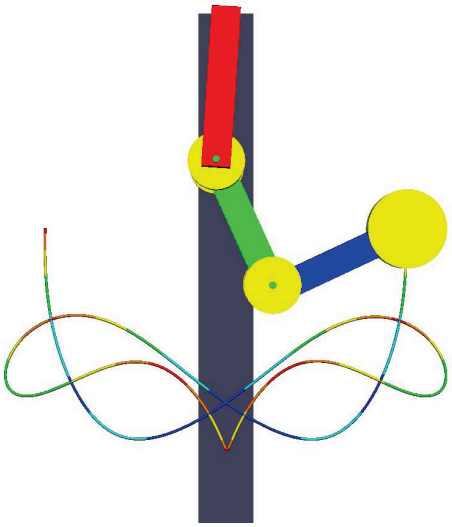


Figure 2. A periodic mode of the un-damped pendulum obtained through optimisation of the initial condition. The parameter values were $\ell_1 = \ell_2 = \ell_3 = 0.1\,\mathrm{m}$ and $m_1 = m_2 = m_3 = 0.1\,\mathrm{kg}$. The starting initial condition was arbitrarily chosen as $x_1 = x_2 = x_3 = 1.5$, with all the initial velocities constrained to zero. The optimised initial condition, which produced the symmetric periodic motion shown, was obtained by demanding $\lambda_{\max} \leq 0.001$. The optimised initial condition was found to be $x_1 = -0.06113$, $x_2 = 0.42713$ and $x_3 = 2.01926$.

After optimisation all the Lyapunov exponents were equal to zero, within the set numerical tolerance of $\lambda_{max} \leq 10^{-3}$ and the corresponding period was 4.38 s.

In the second example a an extreme hyperchaotic mode was calculated, with the result shown in figure 3. This mode was obtained by fixing the initial condition and optimising only the lengths

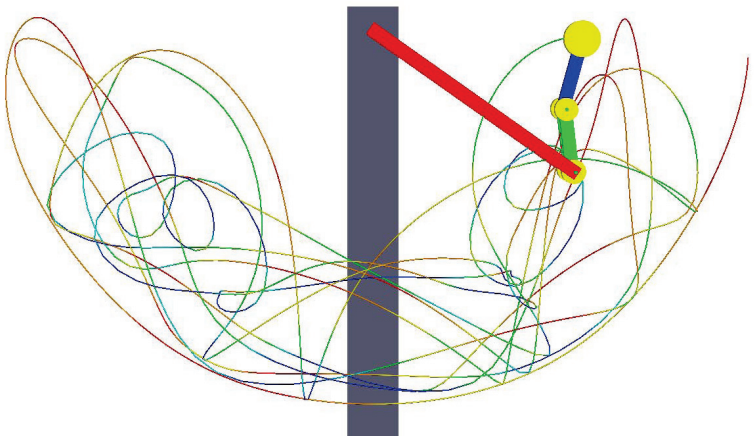


Figure 3. A hyperchaotic mode of the un-damped pendulum obtained through optimisation of the parameters. The initial condition and starting parameter values were the same as in figure 1. Note that, in both figures, one of the two pedestals (shown in gray) has been removed in order to display the trajectory followed by the centre of m_3 more clearly.

and masses. In view of the analysis given just after equation (3), high values of the Lyapunov

exponents were expected for very small lengths. Indeed, in searching for optimised parameter sets for the hyperchoatic modes of the pendulum, it was found that the lengths had to be constrained to remain greater than zero. In this example the lengths were thus arbitrarily constrained to be greater than 0.050 m in order to ensure that the optimised system is physically realisable. Optimisation of all six parameters could then be used to change the Lyapunov exponents from their initial values of $\lambda_1 = 4.58$ and $\lambda_2 = 1.81$ (already hyperchaotic) to a predefined much higher value of (in this example) $\lambda_1 > 10.0 \pm 0.1$. The optimised parameters were found to be $m_1 = 0.106\,\mathrm{kg}$, $m_2 = 0.098\,\mathrm{kg}$, $m_3 = 0.123\,\mathrm{kg}$, $\ell_1 = 0.187\,\mathrm{m}$, $\ell_2 = 0.050\,\mathrm{m}$ and $\ell_3 = 0.051\,\mathrm{m}$. The corresponding Lyapunov exponents were $\lambda_1^{'} = 10.06$, $\lambda_2^{'} = 1.85$ and $\lambda_3^{'} = 0.00$. In other calculations, which are not reported here, it was possible to optimise the parameters and initial conditions simultaneously.

The above results are significant because there are relatively few real physical systems that can be classified as hyperchaotic, with Lyapunov exponents in the range 2-3 or above.

5. Conclusion

The preliminary results reported in the present work show that the triple pendulum system is hyperchaotic, and that its degree of chaos can be controlled through optimisation.

Although the optimisation method presupposes the existence of algebraic expressions for the governing equations, written in the form of equation (3), it has the advantage of providing precise control over the Lyapunov exponents, which characterise the degree of chaos in non-linear dynamical systems. As examples of the method, the initial conditions for periodic motion of a triple pendulum were obtained and the maximal Lyapunov exponent of the same system was enhanced by more than a factor of 2 through optimisation of the parameters alone.

The present article lays the theoretical and numerical foundation for a future in-depth analysis of the hyperchaotic triple pendulum system, which is interesting in its own right. In addition, the optimisation technique employed here may be useful as a means of obtaining precise control over the degree of chaos inherent in other non-linear dynamical system. Since such control is desirable in a wide variety of non-linear applications, which occur in fields as diverse as superconductivity, biophysics and electrical engineering, the method may also be of interest to researchers in a wide variety of such related fields.

Acknowledgments

This work is based upon research supported by the National Research Foundation of South Africa (Nos. IFR2011033100063 and IFR2009090800049) and the Eskom–Trust/Tertiary Education Support Programme of South Africa.

Appendix A. Equations of motion

$$\dot{x}_1 = x_4 \tag{A.1}$$

$$\dot{x}_2 = x_5 \tag{A.2}$$

$$\dot{x}_3 = x_6 \tag{A.3}$$

$$\begin{aligned} \dot{x}_4 \;=\; & \frac{\ell_1 \ell_2^2 \ell_3^2 g}{4D} \Big[\; m_1 m_3^2 \left(\sin\left(x_1 - 2x_2 + 2x_3\right) + \sin\left(x_1 + 2x_2 - 2x_3\right)\right) \\ & -2\left(m_2^2 m_3 + m_2 m_3^2\right) \sin\left(x_1 - 2x_2\right) \; \Big] \end{aligned} \tag{A.4}$$

$$
\begin{aligned}
&-\frac{\ell_1\ell_2^2\ell_3^2 g}{2D} m_3 \left(2m_1m_2 + m_1m_3 + m_2m_3 + m_2^2\right) \sin x_1 \\
&-\frac{\ell_1^2\ell_2^2\ell_3^2}{2D} \left(m_2m_3^2 + m_2^2m_3\right) x_4^2 \sin\left(2x_1 - 2x_2\right) \\
&-\frac{\ell_1\ell_2^3\ell_3^2}{D} \left(m_2m_3^2 + m_2^2m_3\right) x_5^2 \sin\left(x_1 - x_2\right) \\
&-\frac{\ell_1\ell_2^2\ell_3^3}{2D} m_2m_3^2x_6^2 \left(\sin\left(x_1 - x_3\right) + \sin\left(x_1 - 2x_2 + x_3\right)\right) + \chi_1
\end{aligned}
$$

$$
\begin{aligned}
\dot{x}_5 \;=\; & \frac{\ell_1^2\ell_2\ell_3^2 g}{4D} \left(2m_1m_2m_3 + m_1m_3^2 + 2m_2m_3^2 + 2m_2^2m_3\right) \left(\sin\left(2x_1 - x_2\right) - \sin x_2\right) \qquad \text{(A.5)} \\
& -\frac{\ell_1^2\ell_2\ell_3^2 g}{4D} m_1m_3^2 \left(\sin\left(x_2 - 2x_3\right) + \sin\left(2x_1 + x_2 - 2x_3\right)\right) \\
& +\frac{\ell_1^3\ell_2\ell_3^2 m_3}{2D} \Big[\quad \left(2m_1m_2 + m_1m_3 + 2m_2m_3 + 2m_2^2\right) x_4^2 \sin\left(x_1 - x_2\right) \\
& -m_1m_3x_4^2 \sin\left(x_1 + x_2 - 2x_3\right) \quad \Big] \\
& +\frac{\ell_1^2\ell_2^2\ell_3^2}{2D} \left(\left(m_2m_3^2 + m_2^2m_3\right) x_5^2 \sin\left(2x_1 - 2x_2\right) - m_1m_3^2x_5^2 \sin\left(2x_2 - 2x_3\right)\right) \\
& -\frac{\ell_1^2\ell_2\ell_3^3}{2D} \left(\left(2m_1m_3^2 + m_2m_3^2\right) x_6^2 \sin\left(x_2 - x_3\right) - m_2m_3^2x_6^2 \sin\left(2x_1 - x_2 - x_3\right)\right) + \chi_2
\end{aligned}
$$

$$
\begin{aligned}
\dot{x}_6 \;=\; & \frac{\ell_1^2\ell_2^2\ell_3 g}{4D} \left(m_1m_2m_3 + m_1m_3^2\right) \left(\sin\left(2x_1 - x_3\right) + \sin\left(2x_2 - x_3\right)\right) \qquad \text{(A.6)} \\
& -\frac{\ell_1^2\ell_2^2\ell_3 g}{4D} \left(m_1m_2m_3 + m_1m_3^2\right) \left(\sin\left(2x_1 - 2x_2 + x_3\right) + \sin x_3\right) \\
& +\frac{\ell_1^3\ell_2^2\ell_3}{2D} \left(m_1m_2m_3 + m_1m_3^2\right) x_4^2 \left(\sin\left(x_1 - x_3\right) - \sin\left(x_1 - 2x_2 + x_3\right)\right) \\
& +\frac{\ell_1^2\ell_2^3\ell_3}{D} \left(m_1m_2m_3 + m_1m_3^2\right) x_5^2 \sin\left(x_2 - x_3\right) \\
& +\frac{\ell_1^2\ell_2^2\ell_3^2}{2D} m_1m_3^2x_6^2 \sin\left(2x_2 - 2x_3\right) + \chi_3
\end{aligned}
$$

where $g = 9.81\,\mathrm{ms}^{-2}$ is the gravitational acceleration and

$$
\begin{aligned}
D \;=\; \det\left(\mathbf{M}\left(\boldsymbol{\psi}\right)\right) \;=\; & \frac{\ell_1^2\ell_2^2\ell_3^2}{2} m_3 \Big[\quad 2m_1m_2 + m_1m_3 + m_2m_3 + m_2^2 \qquad \text{(A.7)} \\
& -m_2\left(m_2 + m_3\right) \cos\left(2x_1 - 2x_2\right) - m_1m_3 \cos\left(2x_2 - 2x_3\right) \quad \Big]
\end{aligned}
$$

In equations (A.4) to (A.6), χ_1, χ_2 and χ_3 represent terms which are proportional to the damping coefficients c_1, c_2 and c_3. These terms are given by

$$
\begin{aligned}
\chi_1 \;=\; & \frac{\ell_2\ell_3^2 m_3 x_4}{2D} \Big[\quad \ell_2\left(c_1 + c_2\right) \left(2m_2 + m_3 - m_3 \cos\left(2x_2 - 2x_3\right)\right) \qquad \text{(A.8)} \\
& +\ell_1c_2\left(2m_2 + m_3\right) \cos\left(x_1 - x_2\right) - \ell_1c_2m_3 \cos\left(x_1 + x_2 - 2x_3\right) \quad \Big]
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{\ell_2\ell_3 m_3 x_5}{2D}\Big[\ \ell_2\ell_3 c_2\left(2m_2+m_3-m_3\cos\left(2x_2-2x_3\right)\right)\\
&+\ell_1\ell_3\left(c_2+c_3\right)\left(\left(2m_2+m_3\right)\cos\left(x_1-x_2\right)-m_3\cos\left(x_1+x_2-2x_3\right)\right)\\
&-\ell_1\ell_2 c_3\left(m_2+m_3\right)\left(\cos\left(x_1-x_3\right)-\cos\left(x_1-2x_2+x_3\right)\right)\ \Big]\\
&+\frac{\ell_1\ell_2\ell_3 m_3 x_6}{2D}c_3\Big[\ \ell_2\left(m_2+m_3\right)\left(\cos\left(x_1-2x_2+x_3\right)-\cos\left(x_1-x_3\right)\right)\\
&-\ell_3\left(m_3\cos\left(x_1+x_2-2x_3\right)-\left(2m_2+m_3\right)\cos\left(x_1-x_2\right)\right)\ \Big]
\end{aligned}
$$

$$
\begin{aligned}
\chi_2 \;=\; &-\frac{\ell_1\ell_3^2 m_3 x_4}{2D}\Big[\ \ell_1\left(2m_1+2m_2+m_3\right)c_2-\ell_1 m_3 c_2\cos\left(2x_1-2x_3\right) \qquad \text{(A.9)}\\
&+\ell_2\left(2m_2+m_3\right)\left(c_1+c_2\right)\cos\left(x_1-x_2\right)-\ell_2 m_3\left(c_1+c_2\right)\cos\left(x_1+x_2-2x_3\right)\ \Big]\\
&+\frac{\ell_1\ell_3 m_3 x_5}{2D}\Big[\ 2\ell_1\ell_3\left(m_1+m_2\right)c_2+2\ell_1\ell_3\left(m_1+m_2\right)c_3+\ell_1\ell_3 m_3\left(c_2+c_3\right)\\
&+\ell_2\ell_3\left(2m_2+m_3\right)c_2\cos\left(x_1-x_2\right)+\ell_1\ell_2\left(2m_1+m_2+m_3\right)c_3\cos\left(x_2-x_3\right)\\
&-\ell_1\ell_2\left(m_2+m_3\right)c_3\cos\left(2x_1-x_2-x_3\right)-\ell_2\ell_3 c_2 m_3\cos\left(x_1+x_2-2x_3\right)\\
&-\ell_1\ell_3 m_3\left(c_2+c_3\right)\cos\left(2x_1-2x_3\right)\ \Big]\\
&-\frac{\ell_1^2\ell_3 c_3 m_3 x_6}{2D}\Big[\ \ell_3\left(2m_1+2m_2+m_3\right)-\ell_2\left(m_2+m_3\right)\cos\left(2x_1-x_2-x_3\right)\\
&-\ell_3 m_3\cos\left(2x_1-2x_3\right)+\ell_2\left(2m_1+m_2+m_3\right)\cos\left(x_2-x_3\right)\ \Big]
\end{aligned}
$$

and

$$
\begin{aligned}
\chi_3 \;=\; &\frac{\ell_1\ell_2\ell_3 m_3 x_4}{2D}\Big[\ \ell_1\left(2m_1+m_3+m_2\right)c_2\cos\left(x_2-x_3\right) \qquad \text{(A.10)}\\
&-\ell_2\left(m_2+m_3\right)\left(c_1+c_2\right)\cos\left(x_1-x_3\right)-\ell_1\left(m_2+m_3\right)c_2\cos\left(2x_1-x_2-x_3\right)\\
&+\ell_2\left(c_1+c_2\right)\left(m_2+m_3\right)\cos\left(x_1-2x_2+x_3\right)\ \Big]\\
&-\frac{\ell_1\ell_2 x_5}{2D}\Big[\ \ell_1\ell_2\left(2m_1\left(m_2+m_3\right)+\left(m_2+m_3\right)^2\right)c_3-\ell_1\ell_2\left(m_2+m_3\right)^2 c_3\cos\left(2x_1-2x_2\right)\\
&-\ell_1\ell_3 m_3\left(m_2+m_3\right)\left(c_2+c_3\right)\cos\left(2x_1-x_2-x_3\right)\\
&-\ell_2\ell_3 m_3\left(m_2+m_3\right)c_2\cos\left(x_1-x_3\right)\\
&+\ell_1\ell_3 m_3\left(2m_1+m_2+m_3\right)\left(c_2+c_3\right)\cos\left(x_2-x_3\right)\\
&+\ell_2\ell_3 m_3\left(m_2+m_3\right)c_2\cos\left(x_1-2x_2+x_3\right)\ \Big]\\
&+\frac{\ell_1^2\ell_2 c_3 x_6}{2D}\Big[\ \ell_2\left(2m_1\left(m_2+m_3\right)+\left(m_2+m_3\right)^2\right)+\ell_3 m_3\left(2m_1+m_2+m_3\right)\cos\left(x_2-x_3\right)\\
&-\ell_3 m_3\left(m_2+m_3\right)\cos\left(2x_1-x_2-x_3\right)-\ell_2\left(m_2+m_3\right)^2\cos\left(2x_1-2x_2\right)\ \Big]
\end{aligned}
$$

When the upper link of the pendulum is driven by a harmonic force, equations (A.1) to (A.6) must be supplemented by a seventh equation, namely $\dot{x}_7 = \omega$, where ω is the angular frequency of the driving force [10]. The driving force introduces additional terms which are not included here in order to reduce the length of the printed expressions.

References

[1] Awrejcewicz J and Lamarque C H 2003 *World Scientific Series on Nonlinear Science* (*Series A* vol 45) ed Chua L O (Singapore: World Scientific)
[2] Eltohamy K G and Kuo C 1999 *International Journal of Systems Science* **30** 505
[3] Furata T, Tawara T, Okumura Y, Shimizu M and Tomiyama K 2001 *Robotics Autonomous Systems* **37** 81
[4] de Paula A S, Savi M A, and Pereira-Pinto F H I 2006 *Journal of Sound and Vibration* **294** 585
[5] Gitterman M 2010 *The Chaotic Pendulum* (Singapore: World Scientific)
[6] Plissi M V, Torrie C I, Husman M E, Robertson N A, Strain K A, Ward H, Lück H and Hough J 2000 *Review of Scientific Instruments* **71** 2539
[7] Awrejcewicz J and Kudra G 2005 *Nonlinear Analysis* **63** 909
[8] Qi G, van Wyk M A, van Wyk B J and Chen G 2009 *Chaos, Solitons and Fractals* **40** 2544
[9] Awrejcewicz J, Supel B, Lamarque C H, Kudra G, Wasilewski G and Olejnik P 2008 *International Journal of Bifurcation and Chaos* **18** 2883
[10] Awrejcewicz J, Kudra G and Lamarque C H 2004 *International Journal of Bifurcation and Chaos* **14** 4191
[11] Kudra G 2002 *Analysis of Bifurcation and Chaos in Triple Physical Pendulum with Impacts* Ph.D. thesis Technical University of Lódź
[12] Chun W J 2007 *Core Python Programming* 2nd ed (New Jersey: Prentice Hall)
[13] Wolf A, Swift J B, Swinney H L and Vastano J A 1985 *Physica D* **16** 285
[14] Chen Z M, Djidjeli K and Price W G 2006 *Applied Mathematics and Computation* **174** 982
[15] Sattin F 1997 *Computer Physics Communications* **107** 253
[16] Langtangen H P 2004 *Python Scripting for Computational Science* (Berlin: Springer Verlag)
[17] Nelder J A and Mead R 1965 *Computer Journal* **7** 308 Also see *www.scipy.org*

Finite element calculations for molecules with multiple Coulomb centers

Moritz Braun

Department of Physics, University of South Africa, PO Box 392, 0003 UNISA, South Africa

E-mail: braunm@unisa.ac.za

Abstract. A novel and efficient method to mitigate the challenges posed by the Coulomb potentials when solving three-dimensional effective Schrödinger equations for molecules or solids is presented. This approach assumes a product ansatz for the wave function, where the first factor satisfies the cusp condition, while the second factor is assumed to be smooth. This ansatz leads to a regularized non-singular potential and a transformed energy functional. Highly accurate finite element ground state energies obtained with this functional as result of two-and three dimensional finite element calculations for the hydrogen molecular ion are also presented.

1. Introduction

Efficient and reliable methods to solve the effective three-dimensional Schrödinger equation are important for density functional or Hartree Fock methods to calculate the properties of molecules and solids. To judge the accuracy of methods using Gaussian basis functions, pseudo potentials [1] or linearized augmented plane waves(LAPW) [2], a method that promises to be less basis dependent, is desirable. The method of finite elements [3, chapter 10] can serve this purpose, since its convergence can be improved systematically.

Any method to solve an effective Schrödinger equation in three dimensions faces the following challenges due to the Coulomb singularities at the nuclei:

(i) The cusp conditions at the nuclei [4]: $\lim_{r_i \to 0} \overline{\frac{df_C/dr_i}{f_C(r_i)}} = -Z_i$ have to be satisfied.

(ii) The wave functions will have many nodes close to the nuclei.

The competing methods mentioned also have some disadvantages: pseudopotentials have to be determined for each atomic species, the matching conditions for LAPW can become involved and Gaussian basis functions are in principle *unable* to satisfy the cusp conditions.

The method of finite elements has been used in density functional calculations for solids by a number of authors [5, and references therein], but these have in general made use of the pseudo potential approach. Using the method of finite elements in the core region of the atoms will however be very difficult for heavier elements since the wave functions of the core electrons are highly oscillatory, and thus probably restricted to the lighter elements up to the third or fourth period.

In this contribution a novel and efficient method to deal with the above mentioned challenges by regularization of the potential is outlined and results obtained for the ground state energy of the hydrogen molecular ion using the method of finite elements are shown. In addition to

providing a sensitive test on the inherent accuracy of the proposed scheme, calculations for dimers and other simple molecules are also of importance for the spectroscopy of interstellar space [6, and references therein].

The remainder of the contribution is organized as follows: In section 2 the regularization approach is outlined, in section 3 the test case of H_2^+ is dicussed, in section 4 the method of finite elements as used in this contribution is sketched, while in sections 5 and 6 results obtained via three-and two dimensional finite element calculations are given, followed by conclusions.

2. Regularization of the potential

In the following we consider a system consisting of N electrons and A nuclei in the Born-Oppenheimer approximation with the Hamiltonian

$$H = -\sum_{i=1}^{N} \nabla_i^2 + \sum_{i>j=1}^{N} \frac{2}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i=1}^{N}\sum_{j=1}^{A} \frac{2Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} + \sum_{i>j=1}^{A} \frac{2Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|}\,, \tag{1}$$

where we use atomic units, i.e. measuring distances in Bohr radii and energies in Rydberg.

In the simplest case with only one electron or when using either the density-functional or Hartree-Fock methods an (effective) three dimensional Schrödinger equation of the form

$$\left[-\nabla^2 - \sum_{j=1}^{M} \frac{2Z_j}{|\mathbf{r} - \mathbf{R}_j|} + v_{\text{eff}}(\mathbf{r}) + \sum_{i>j=1}^{A} \frac{2Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|}\right] \psi(\mathbf{r}) = E\psi(\mathbf{r})\,, \tag{2}$$

where the density-dependent and possibly non-local effective potential $v_{\text{eff}}(\mathbf{r})$ takes the electron-electron interaction into account, must be solved.

In order to mitigate the Coulomb singularities we use the following product ansatz for the wave function:

$$\Psi(\mathbf{r}) = f_{\text{C}}(\mathbf{r})\Phi(\mathbf{r})\,, \tag{3}$$

where f_{C} satisfies

$$\lim_{r_i \to 0} \frac{\overline{df_{\text{C}}/dr_i}}{f_{\text{C}}(r_i)} = -Z_i\ ,\ r_i = |\mathbf{r} - \mathbf{R}_i| \tag{4}$$

and Φ is expected to be smooth.

Substituting the product ansatz into the standard energy functional for Ψ,i.e.

$$E(\Psi) = \frac{\langle \nabla\Psi | \nabla\Psi \rangle + \langle \Psi | V(\mathbf{r}) | \Psi \rangle}{\langle \Psi | \Psi \rangle}\,, \tag{5}$$

where $V(\mathbf{r})$ includes the potential due to the nuclei, the effective potential and the constant Coulomb potential energy between the nuclei, we obtain a transformed functional for Φ as

$$\tilde{E}(\Phi) = \frac{\langle f_{\text{C}}\nabla\Phi | f_{\text{C}}\nabla\Phi \rangle + \langle f_{\text{C}}\Phi | \tilde{V}(\mathbf{r}) | f_{\text{C}}\Phi \rangle}{\langle f_{\text{C}}\Phi | f_{\text{C}}\Phi \rangle}\,, \tag{6}$$

$$\tilde{V} = V - \frac{\nabla^2 f_{\text{C}}}{f_{\text{C}}}\,. \tag{7}$$

Choosing for f_{C} the form

$$f_{\text{C}}(\mathbf{r}) = 1 + \sum_{i=1}^{A} c_i \exp\left(-2Z_i r_i\right)\ ,\ r_i = |\mathbf{r} - \mathbf{R}_i|\ , \tag{8}$$

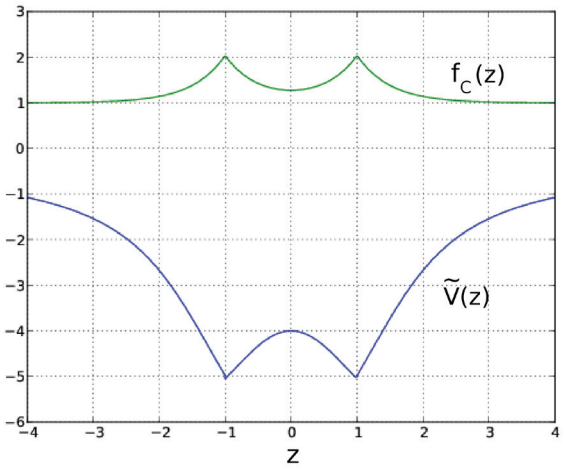


Figure 1. Transformed potential $\tilde{V}$ and cusp factor $f_{\rm C}$ for the hydrogen molecular ion along the axis connecting the protons for an internuclear distance of 2 $a_{\rm Bohr}$.

the c_i can be determined from the following linear system of equations

$$\sum_{j=1}^{A} \left[\delta_{ij} - \nu_{ij} \exp\left(-2Z_i \left|\mathbf{R}_i - \mathbf{R}_j\right|\right)\right] c_j = 1 \ , \ \nu_{ij} = 1 - \delta_{ij} \, , \tag{9}$$

which results from eqn.(7) by requiring the Coulomb singularities to be removed from the potential. The resulting potential has singularities replaced by kinks as shown by the lower curve in Fig. (1) for H_2^+ along the z-axis connecting the nuclei, while the upper curve shows $f_{\rm C}$.

3. The test case H_2^+

Two protons and an electron constitute the simplest possible molecule. In the Born-Oppenheimer approximation the nuclei are considered as fixed. In this section we consider this molecule in the two system of coordinates, namely Cartesian and cylindric, that we will use for our calculations and in elliptic coordinates, which have been used for highly accurate numerical benckmark calculations in the past.

In Cartesian coordinates the Hamiltonian for an internuclear distance of $R = 2a_{\rm Bohr}$, very close to the equilibrium value, reads

$$H = -\nabla^2 - \frac{2}{|\mathbf{r} - \mathbf{R}_1|} - \frac{2}{|\mathbf{r} - \mathbf{R}_2|} + 1 \, , \tag{10}$$

while in cylindrical coordinates ρ,z and ϕ with nuclei on z-axis at $z = \pm 1$ and for $m = 0$, the Hamiltonian is two-dimensional, i.e.

$$H = -\frac{\partial^2}{\partial \rho^2} - \frac{1}{\rho} \frac{\partial}{\partial \rho} - \frac{2}{\sqrt{\rho^2 + (z+1)^2}} - \frac{2}{\sqrt{\rho^2 + (z-1)^2}} + 1 \, . \tag{11}$$

Highly accurate benchmark numerical solutions for ground state energy and wave function were obtained by Peek in the 1960's using elliptic coordinates [7]

$$\lambda \ = \ \frac{1}{2} (r_1 + r_2) \qquad , \ 1 \leq \lambda \leq \infty \, , \tag{12}$$

$$\mu = \frac{1}{2}(r_1 - r_2) \quad , \; -1 \le \mu \le 1\,, \tag{13}$$

$$r_1 = |\mathbf{r} - \mathbf{R}_1| \quad , \; r_2 = |\mathbf{r} - \mathbf{R}_2|\,. \tag{14}$$

leading to the Hamiltonian

$$H = -\frac{1}{\lambda^2 - \mu^2}\left[\sqrt{\lambda^2-1}\frac{\partial}{\partial\lambda}\left(\sqrt{\lambda^2-1}\frac{\partial}{\partial\lambda}\right) + \sqrt{1-\mu^2}\frac{\partial}{\partial\mu}\left(\sqrt{1-\mu^2}\frac{\partial}{\partial\mu}\right)\right] - 4\frac{\lambda}{\lambda^2-\mu^2} + 1\,, \tag{15}$$

which allows separation into two ordinary differential equations.

4. Method of finite elements

The variational problem obtained in section 2 will be solved using the method of finite elements [3]. We divide the region of interest into rectangular (for 2D) or box-shaped (for 3D) *elements* and expand Φ in terms of interpolating polynomials on each element. Applying the variational principle then leads to a symmetric generalized eigen value problem

$$H\mathbf{u} = \lambda U\mathbf{u} \tag{16}$$

with $N \approx 10^4 - 10^6$. Here $\mathbf{u}$ is the coefficient vector of the expansion in the finite element basis set and the matrices H and U are sparse. The lowest eigenpair of the eigen value problem is obtained using the symmetric Jacobi Davidson method (JDSYM) as provided by the PySparse package [8]. This method does not factorize matrices H and U and therefore needs less memory then traditional methods, that use factorization.

5. Calculations in three dimensions

We consider the problem in Cartesian coordinates as described by eqn.(10). Since higher derivatives of ϕ must be discontinous at the nuclei they must be at the corners of elements. The grid for the finite element method is created on the volume $V = [-x_{\max}, x_{\max}]^3$ as the tensor product of seperate x,y and z grids. The grids are compressed quadratically around the nuclei and the number of elements is chosen such, that elements at the nuclei have sides smaller than η. For H_2^+ with the nuclei positioned at $(-1,0,0)$ and $(1,0,0)$ a typical grid is in Fig. (2) in its projection on the xy-plane: The drawback of the above tensor-grid is however that small elements are used in regions, where they are not really needed.

For the calculations interpolation polynomials up to order $p = 6$ order were used and the matrix elements were determined using Gauss-Legendre integration with up to $n_{\mathrm{GL}} = 16$ points in each direction. The code was written in Python [9] with the numerical extension numpy [10] and some Fortran77 linked with f2py [11]. The largest grid had $26 \times 22 \times 22 = 12\,584$ elements. leading to generalized eigen value problem of $N = 2659955$. The resulting sparse matrices were handled using PysSparse and the lowest eigenpair was calculated using JDSYM in about 15000 CPU seconds on the high performance cluster of the University of South Africa, while the time to compute the matrix was approximately 85000 CPU seconds.

In Table 1 the results obtained for the ground state energy for different discretization and integration parameters are given. It is evident, that a reasonable accuracy of 10^{-5} can be achieved using relatively modest resources. The best energy obtained in the last line of the table agrees with the the literature result of $E_{\mathrm{gs}} = -1.2052684289898$ Ryd [7] to 11 digits.

6. Calculations in two dimensions

As mentioned in the previous section the tensor-grid employed for the three-dimensional calculation is computationally not very efficient. The open source project hermes for adaptive h-p-FEM [12] provides libraries which allow for hexaedral grids with hanging nodes. Since

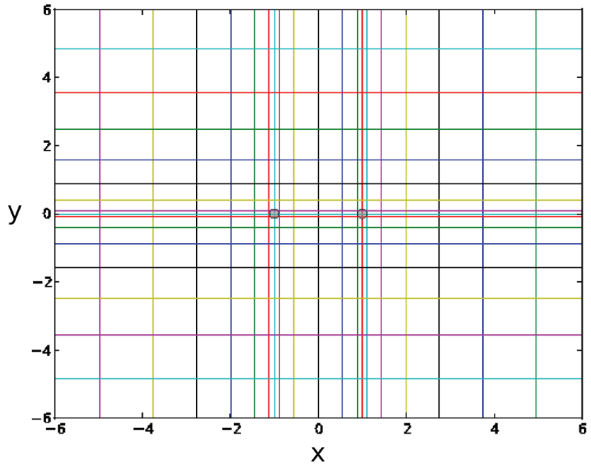


Figure 2. Typical grid for three-dimensional finite element calculation projected on xy-plane.

$x_{\max}$	p	η	n_{GL}	E_{gs}	ΔE	T_{CPU} [s]
8	4	0.50	6	−1.2052629814026	$5.448\ 10^{-6}$	191
8	4	0.25	6	−1.2052675549928	$8.740\ 10^{-7}$	435
8	4	0.25	12	−1.2052675476430	$8.813\ 10^{-7}$	2024
12	6	0.15	12	−1.2052684289614	$2.845\ 10^{-11}$	33665
12	6	0.12	16	−1.2052684289849	$4.92\ 10^{-12}$	107152

Table 1. Ground state energies, their deviations from the value obtained by Peek [7] and run-time in CPU seconds for different grid and integration parameters.

H_2^+ has an axis of symmetry it can also be investigated using cylindrical coordinates and the Hamiltonian given in eqn.(11). As a first step towards improving the grid in three dimensions the calculation was done using the two dimensional version of hermes, namely the C++ library hermes2d.

The transformed potential is then given by

$$\tilde{V} = -\frac{2}{\sqrt{\rho^2+(z+1)^2}} - \frac{2}{\sqrt{\rho^2+(z-1)^2}} - \frac{\nabla^2 f_{\mathrm{C}}(\rho,z)}{f_{\mathrm{C}}(\rho,z)} + 1\,. \tag{17}$$

Using hermes2d and employing the JDSYM eigen-value solver as in section 5, the variational problem defined by eqn.(6) was numerically solved on the domain $(\rho,z) \in [0,16]\times[0,16]$ where symmetry with respect to $z=0$ was used. The grid was refined towards the proton at $z=1$ as shown in Fig. (3), where the hanging nodes [13], that allow for efficient subdivision also of rectangular grids, are evident. The Calculation was performed on an Dell centrino laptop with 4GB of RAM. The most accurate result obtained for $p=6$ and $N=45396$ was $E_{\mathrm{gs}} = -1.2052684289897795$ Ryd, which agrees to 13 digits with the literature value by Peek[7] and T_{CPU} was 217 seconds. The integrals were evaluated with Gauss Legendre Integration of order ord(u) + ord(v) + 8 where u and v are the interpolation polynomials used by hermes2d. The function Φ obtained for these parameters is shown in Fig. (4) together with the grid.

To understand the influence of the factorization on the calculation better a standard finite element calculation without the factorization method was also done, the results of which are given in in Table (2).

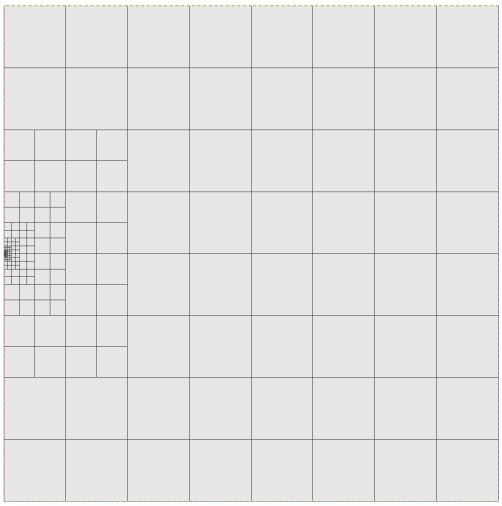


Figure 3. Grid used for region $[0, 2] \times [0, 2]$ in finite element calculation.

Integration order	E	ΔE
$\mathrm{ord}(u) + \mathrm{ord}(v) + 8$	-1.2052684296616842	$-6.718843\ 10^{-10}$
15	-1.2052684292224827	$-2.326829\ 10^{-10}$
20	-1.2052684290596827	$-6.988277\ 10^{-11}$
30	-1.2052684290061166	$-1.63167\ 10^{-11}$
60	-1.2052684289906663	$-8.664\ 10^{-13}$
90	-1.2052684289896383	$1.616\ 10^{-13}$

Table 2. Ground state energies and their errors obtained as result of finite element calculation with $p = 6$, not employing the product ansatz, for a number of different integration orders.

Thus in order to reach the same accuracy many more Gauss-Legendre integration points are required to integrate over the weakly singular term $\rho/\sqrt{\rho^2 + z^2}$. Thus in this case convergent results are still possible, but come at the cost of a much larger integration grid. It must also be noted that *overbinding* occurs if the integration order is insufficient. Thus, comparing the two approaches with and without the factor $f_C(\mathbf{r})$, it is evident that it make the calculations easier also in two dimensions.

The convergence of the discretization error for the finite element method of order p is expected to satisfy $\Delta E \propto h^{2p}$ [14, p. 232]. To investigate the convergence calculations were also done with $p = 3$ and $p = 4$ respectively resulting in the log-log plots shown in Figs. (5) and (6). The slopes are obtained as 5.87 ± 0.07 and 8.22 ± 0.12 respectively, compatible with the expected values of 6 and 8.

7. Conclusions

In this contribution it has been clearly demonstrated, that the singularities due to the Coulomb potentials at the nuclei can be mitigated using the suggested product ansatz for the wave function. The resulting transformed variational problem for H_2^+ has been solved both in Cartesian and cylindrical coordinates using the method of finite elements, yielding highly accurate energies.

After the encouraging results for H_2^+ we plan to extend the three dimensional calculations to

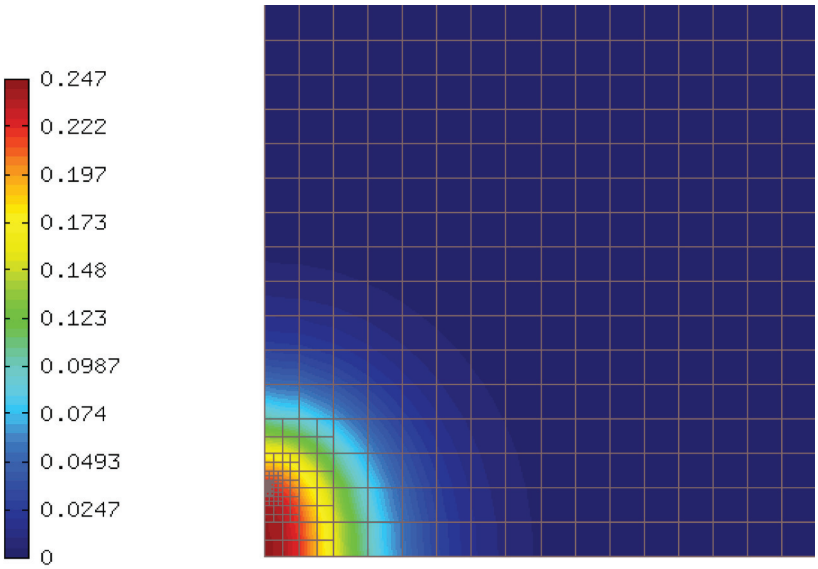


Figure 4. Function Φ for the hydrogen molecular ion on the truncated domain $(\rho, z) \in [0, 8] \times [0, 8]$.

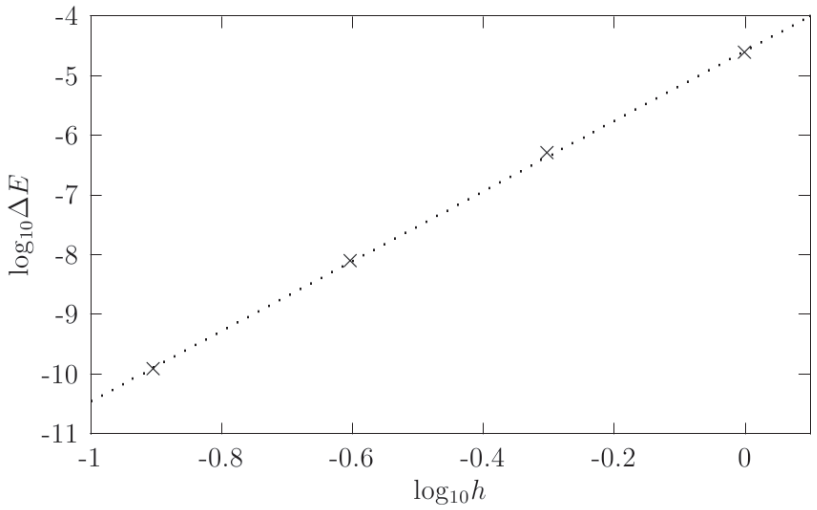


Figure 5. Log-log plot of ΔE versus h for $p = 3$.

larger molecules in the frame work of an all electron density functional approach, also employing the hermes3d framework for adaptive refinement of the grid. An extension to solids, i.e. systems with periodic symmetry, is also envisaged.

Acknowledgments
I would like to thank the hpfem group at the University of Nevada, in particular Pavel Solin, for assistance in making use of the C++ package hermes2d. Financial support by the National Research Foundation (NRF) of South Africa is also acknowledged.

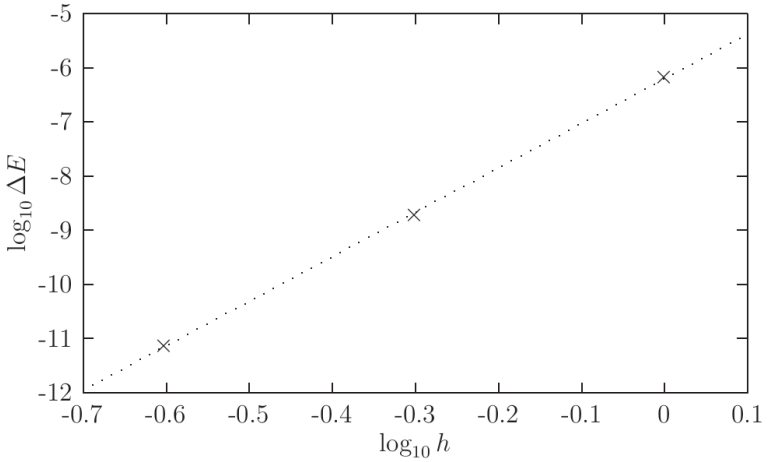


Figure 6. Log-log plot of ΔE versus h for $p = 4$.

References

[1] Troullier N and Martins J 1991 *Phys. Rev. B* **43** 1993–2006
[2] Blaha P, Schwarz K, Madsen G, Vasnicka D and Luitz J 2001 *WIEN2k, An Augmented Plane Wave + Local Orbitals Program for Calculating Crystal Properties* (Technical University Vienna, Austria)
[3] Ram-Mohan L 2002 *Finite Element and Boundary Element Applications in Quantum Mechanics* Oxford texts in applied and engineering mathematics (Oxford University Press, Oxford, UK) ISBN 9780198525226
[4] Kato T 1957 *Commun. Pure Appl. Math* **10** 151–177
[5] Pask J and Sterne P 2005 *Modelling Simul. Mater. Sci. Eng.* **13** 71–96
[6] Oka T and Jagod M F 1993 *Journal of the Chemical Society, Faraday Transactions* **89** 2147–2154
[7] Peek J M 1965 *J. Chem. Phys.* **43** 3004–3006
[8] URL `http://pysparse.sourceforge.net`
[9] URL `http://www.python.org`
[10] URL `http://numpy.scipy.org`
[11] URL `http://www.scipy.org/F2py`
[12] URL `http://www.hpfem.org`
[13] Solin P, Cerveny J and Dolezel I 2008 *Math. Comput. Simul.* **77** 117–132
[14] Strang G and Fix G 1973 *An Analysis of the Finite Element Method* Prentice-Hall series in automatic computation (Prentice-Hall, Upper Saddle River, New Jersey, USA) ISBN 9780130329462

Non-locality without inequality and generalized non-local theory

Sujit K. Choudhary

School of Physics, University of KwaZulu-Natal, Private Bag X54001, Durban 4000, South Africa.

E-mail: choudhary@ukzn.ac.za

Abstract. We find non-local but non-signaling probabilities satisfying the 'nonlocality without inequality' arguments for multiple two-level systems. Maximum probability of success of these arguments is obtained in the framework of a generalized nonlocal theory. Interestingly, for two two-level systems, the probability of success of these arguments converge to a common maximum in this framework. This is in sharp contrast with the quantum case, where for such systems, Cabello's argument succeeds more than that of Hardy's. We also find that the maximum probability of success of Hardy's argument is the same for both the two and three two-level systems in the framework of this more generalized theory.

1. Introduction

There exist correlations between quantum systems which no local-realistic theory can reproduce. This was first shown by Bell by means of an inequality, known as Bell's inequality [1]. Later, Hardy [2] gave an argument which also reveals the non-local character of Quantum Mechanics, but this argument, unlike Bell's argument, does not use inequalities involving expectation values. Afterwards, Cabello [3] introduced another logical structure to prove Bell's theorem without inequality. Though, Cabello's logical structure was originally proposed for establishing nonlocality for three particle states, but it was later exploited to establish nonlocality for a class of two-qubit mixed entangled state [4]. It is noteworthy here that in contrast, there is no two-qubit mixed state which shows Hardy type nonlocality [5] whereas almost all pure entangled states of two-qubits do so (maximally entangled states are the exception)[6, 7]. Likewise, for almost all two-qubit pure entangled states other than maximally entangled one, Cabello's nonlocality argument runs, but, intesestingly, for these states, the maximum probabilty of success of this argument is more than that of the Hardy's [8].

Although no local-realistic theory can reproduce quantum correlations still these correlations cannot be exploited to communicate with a speed greater than that of the light in vacuum. But quantum theory is not the only non-local theory consistent with relativistic causality [9]. Theories which predict non-local correlations and hence permit violation of Bell's inequality but are constrained with the no signalling condition are called Generalized non-local theories (GNLT). In recent years there has been an increasing interest in GNLT [10, 11, 12, 13, 14, 15, 16, 17, 18, 19]. In general, quantum theory has been studied in the background of classical theory which is comparatively restrictive. The new idea is to study quantum theory from outside i.e.,

starting from a general family of theories, and to study properties common to all [13]. This might help in a better understanding of quantum non-locality.

In this paper, we study the non-locality arguments in the framework of GNLT. We find that probability of success of both Hardy's and Cabello's nonlocality arguments converge to a common maximum 0.5 for two two-level systems unlike the quantum case where they are 0.09 [6, 20] and 0.11 (approx) respectively [8]. We, further find that 0.5 is also the maximum probability of success of Hardy's argument for three two-level systems in GNLT, the corresponding qunatum maximum is known to be 0.125 [21].

2. Nonlocality arguments for two two-level systems and the generalized non local theory

In the framework of a generalized probabilistic theory, consider a physical system consisting of two subsystems shared between two far seperated observers, Alice and Bob. Assume that Alice can run the experiments of measuring any one (chosen at random) of the two $\{-1, +1\}$-valued random variables A and $A^{'}$ corresponding to her subsystem whereas Bob can run the experiments of measuring any one (chosen at random) of the two $\{-1, +1\}$-valued random variables B and $B^{'}$ corresponding to his subsystem. Consider now the following four conditions:

$$\left.\begin{array}{lcl} \text{Prob}(A = +1,\ B = +1) & = & 0, \\ \text{Prob}(A^{'} = +1,\ B = -1) & = & 0, \\ \text{Prob}(A = -1,\ B^{'} = +1) & = & 0, \\ \text{Prob}(A^{'} = +1,\ B^{'} = +1) & = & q \end{array}\right\} \text{(with } q > 0). \tag{1}$$

The above four conditions together form the basis of Hardy's nonlocality argument. These four cannot be fulfilled simultaneously in the framework of a local-realistic theory. To see this let us start with the fourth condition in equation (1) which implies that (i) there is a non-zero probability (which is q here) of simultaneous occurrence of $A^{'} = +1$ and $B^{'} = +1$. This will then imply that (according to classical probability theory) (ii) if Alice chooses to perform the measurement of $A^{'}$, there will be a non-zero probability (which should be at least q) of getting the value $+1$, and similarly, (iii) if Bob chooses to perform the measurement of $B^{'}$, there will be a non-zero probability (which should be at least q) of getting the value $+1$. Now the 3rd condition in equation (1) implies that if Bob chooses for the measurement of $B^{'}$ and Alice chooses for A, she is bound to get the value $+1$ whenever Bob gets the value $+1$ (which Bob can indeed get with a non-zero probability, according to (iii)). Similarly the 2nd condition in equation (1) implies that if Alice chooses to the measurement of $A^{'}$ and Bob chooses to perform B, he is bound to get the value $+1$ whenever Alice gets the value $+1$ (which Alice can indeed get with a non-zero probability, according to (ii)). So $A = +1$ is a 'reality' of A while $B = +1$ is also a 'reality' of B, according to EPR[23]. Now, invoking 'locality', $A = +1$ and $B = +1$ is a joint 'reality' of the composite system. And, according to condition (i) (again, using classical probability arguments), the joint probability of occurrence of $A = +1$, $B = +1$ must be at least q. This contradicts the first condition of equation (1). On the contrary almost all pure entangled state of two-qubit (except the maximally entangled one) satisfy the above four conditions simultaneously for suitable choices of observables[7] with a maximum probability of success (q_{max}) equal to 9% (approx) [6, 20].

Cabello's conditions result by replacing the right hand side of the first condition of (1) with a nonzero probability p with $p < q$, and keeping the remaining three conditions intact. As seen above, for local-realistic states to satisfy all the four conditions simultaneously, the joint probability of occurrence of $A = +1$, $B = +1$ must be at least equal to q and thus these conditions are also incompatible with local-realism. But, almost all pure entangled states of two-qubits (except their maximally entangled states) satisfy these with a maximum success probability equal to 11% (approx)[8].

We now study the character of a set of sixteen non-signalling joint probabilities $P(i_m, j_n)$ where $P(i_m, j_n)$ denotes the probability of getting the output as $i_m \in \{+1, -1\}$ and $j_n \in \{+1, -1\}$ in a measurement for $M \in \{A, A^{'}\}$ by Alice and $N \in \{B, B^{'}\}$ by Bob. These sixteen joint probabilities are given below:

$$\begin{array}{l} \text{Prob}(A = +1, B = +1) = p_1,\ \text{Prob}(A = +1, B = -1) = p_2, \\ \text{Prob}(A = -1, B = +1) = p_3,\ \text{Prob}(A = -1, B = -1) = p_4, \\ \\ \text{Prob}(A^{'} = +1, B = +1) = p_5,\ \text{Prob}(A^{'} = +1, B = -1) = p_6, \\ \text{Prob}(A^{'} = -1, B = +1) = p_7,\ \text{Prob}(A^{'} = -1, B = -1) = p_8, \\ \\ \text{Prob}(A = +1, B^{'} = +1) = p_9,\ \text{Prob}(A = +1, B^{'} = -1) = p_{10}, \\ \text{Prob}(A = -1, B^{'} = +1) = p_{11},\ \text{Prob}(A = -1, B^{'} = -1) = p_{12}, \\ \\ \text{Prob}(A^{'} = +1, B^{'} = +1) = p_{13},\ \text{Prob}(A^{'} = +1, B^{'} = -1) = p_{14}, \\ \text{Prob}(A^{'} = -1, B^{'} = +1) = p_{15},\ \text{Prob}(A^{'} = -1, B^{'} = -1) = p_{16}. \end{array} \tag{2}$$

Other than being members of the interval $[0, 1]$, these probabilities must satisfy the normalization condition:

$$p_1 + p_2 + p_3 + p_4 = 1, \tag{3}$$

$$p_5 + p_6 + p_7 + p_8 = 1, \tag{4}$$

$$p_9 + p_{10} + p_{11} + p_{12} = 1, \tag{5}$$

$$p_{13} + p_{14} + p_{15} + p_{16} = 1. \tag{6}$$

The no-signaling constraint implies that if Alice measures for A (or $A^{'}$), the individual probabilities for the outcomes $A = +1$ (or $A^{'} = +1$) and $A = -1$ (or $A^{'} = -1$) must be independent of whether Bob chooses to measure for B or $B^{'}$ and similar should be the case for Bob too. So, the condition for causality to hold is given by:

$$p_1 + p_2 = p_9 + p_{10}, \tag{7}$$

$$p_3 + p_4 = p_{11} + p_{12}, \tag{8}$$

$$p_5 + p_6 = p_{13} + p_{14}, \tag{9}$$

$$p_7 + p_8 = p_{15} + p_{16}, \tag{10}$$

$$p_1 + p_3 = p_5 + p_7, \tag{11}$$

$$p_2 + p_4 = p_6 + p_8, \tag{12}$$

$$p_9 + p_{11} = p_{13} + p_{15}, \tag{13}$$

$$p_{10} + p_{12} = p_{14} + p_{16}. \tag{14}$$

We further assume that these probabilities respect the Hardy-type non-locality conditions:

$$p_1 = 0,\ p_6 = 0,\ p_{11} = 0,\ p_{13} = q. \tag{15}$$

Using equation (15) into equation (7), we get

$$p_2 \geq p_9. \tag{16}$$

Using equation (15) into equation (11), we get

$$p_3 \geq p_5. \tag{17}$$

Using equation (15) into equation (3), we get

$$1 = p_2 + p_3 + p_4 \geq p_2 + p_3 \geq p_5 + p_9, \tag{18}$$

using equations (16) and (17). Using equation (15) into equations (9) and (13), we get

$$p_5 + p_9 = 2q + p_{14} + p_{15} \geq 2q. \tag{19}$$

Using equations (18) and (19), we get

$$1 \geq p_5 + p_9 \geq 2q.$$

Thus we have

$$q \leq \frac{1}{2}. \tag{20}$$

If we now follow the argument, beginning at equation (16) and ending at equation (20), it can be easily shown that for $q = 1/2$, there is a *unique* solution for the above-mentioned sixteen joint probabilities satisfying simultaneously all the conditions (3) to (15):

$$p_2 = p_3 = p_5 = p_8 = p_9 = p_{12} = p_{16} = \frac{1}{2} \text{ and } p_4 = p_7 = p_{10} = p_{14} = p_{15} = 0. \tag{21}$$

Thus we see that above-mentioned sixteen probabilities will be *non-local* as well as *non-signaling* if and only if $0 < q \leq 1/2$.

If, inplace of Hardy's conditions, we put Cabello's conditions $p_1 = p$, $p_6 = 0$, $p_{11} = 0$ and $p_{13} = q$ in equations (3)-(14), we find after a little of algebra that the above nonlocal probability distribution is non-signalling up to $q - p \leq 0.5$ (equality is in the sense'just less than').

Thus as far as maximum probability of success of the Hardy's and the Cabello's arguments are concerned, none has got any practical edge over the other in **GNLT**, whereas in quantum theory probability of success of Cabello's argument is more than that of Hardy's.

3. General non-signaling probabilities satisfying Hardy-type non-locality argument for three two-level systems

Consider a physical system consisting of three subsystems shared among three far apart parties Alice, Bob and Charlie, in the framework of a general probabilistic theory. Assume that Alice, Bob and Charlie can measure one of the two observables X_i, Y_i, where i stands for the 1st (*i.e.*, Alice), 2nd (*i.e.*, Bob), or 3rd (*i.e.*, Charlie) on their respective subsystems. The outcomes of each such measurements can be either up (U) or down (D). We now consider all the sixty four joint probabilities $\text{Prob}(R_1 = j, R'_2 = k, R''_3 = l)$, where $R, R', R'' \in \{X, Y\}$ and $j, k, l \in \{U, D\}$. For the sake of notational simplicity, we will denote X by 0 and Y by 1 and also U by 0 and D by 1. We can thus denote the above-mentioned joint probabilities as $\text{Prob}(i = s_1, j = s_2, k = s_3)$, where $i, j, k \in \{0, 1\}$ and $s_1, s_2, s_3 \in \{0, 1\}$. To make it more readable, we will denote the probability $\text{Prob}(i = s_1, j = s_2, k = s_3)$ by $p_{is_1js_2ks_3}$, where $is_1js_2ks_3$ is the binary representation of the number $32i + 16s_1 + 8j + 4s_2 + 2k + s_3$.

Apart from being the of the interval $[0, 1]$, these sixty four joint probabilities p_0, p_1, ..., p_{63} must satisfy the normalization conditions:

$$\begin{gathered}
p_0 + p_1 + p_4 + p_5 + p_{16} + p_{17} + p_{20} + p_{21} = 1,\ p_2 + p_3 + p_6 + p_7 + p_{18} + p_{19} + p_{22} + p_{23} = 1, \\
p_8 + p_9 + p_{12} + p_{13} + p_{24} + p_{25} + p_{28} + p_{29} = 1,\ p_{10} + p_{11} + p_{14} + p_{15} + p_{26} + p_{27} + p_{30} + p_{31} = 1, \\
p_{32} + p_{33} + p_{36} + p_{37} + p_{48} + p_{49} + p_{52} + p_{53} = 1,\ p_{34} + p_{35} + p_{38} + p_{39} + p_{50} + p_{51} + p_{54} + p_{55} = 1, \\
p_{40} + p_{41} + p_{44} + p_{45} + p_{56} + p_{57} + p_{60} + p_{61} = 1,\ p_{42} + p_{43} + p_{46} + p_{47} + p_{58} + p_{59} + p_{62} + p_{63} = 1.
\end{gathered} \tag{22}$$

The no-signalling condition gives:

$$p_0 + p_1 = p_2 + p_3,\ p_4 + p_5 = p_6 + p_7,\ p_8 + p_9 = p_{10} + p_{11},\ p_{12} + p_{13} = p_{14} + p_{15},$$
$$p_{16} + p_{17} = p_{18} + p_{19},\ p_{20} + p_{21} = p_{22} + p_{23},\ p_{24} + p_{25} = p_{26} + p_{27},\ p_{28} + p_{29} = p_{30} + p_{31},$$
$$p_{32} + p_{33} = p_{34} + p_{35},\ p_{36} + p_{37} = p_{38} + p_{39},\ p_{40} + p_{41} = p_{42} + p_{43},\ p_{44} + p_{45} = p_{46} + p_{47},$$
$$p_{48} + p_{49} = p_{50} + p_{51},\ p_{52} + p_{53} = p_{54} + p_{55}, p_{56} + p_{57} = p_{58} + p_{59},\ p_{60} + p_{61} = p_{62} + p_{63}; \tag{23}$$

$$p_0 + p_4 = p_8 + p_{12},\ p_1 + p_5 = p_9 + p_{13},\ p_2 + p_6 = p_{10} + p_{14},\ p_3 + p_7 = p_{11} + p_{15},$$
$$p_{16} + p_{20} = p_{24} + p_{28},\ p_{17} + p_{21} = p_{25} + p_{29},\ p_{18} + p_{22} = p_{26} + p_{30},\ p_{19} + p_{23} = p_{27} + p_{31},$$
$$p_{32} + p_{36} = p_{40} + p_{44},\ p_{33} + p_{37} = p_{41} + p_{45},\ p_{34} + p_{38} = p_{42} + p_{46},\ p_{35} + p_{39} = p_{43} + p_{47},$$
$$p_{48} + p_{52} = p_{56} + p_{60},\ p_{49} + p_{53} = p_{57} + p_{61},\ p_{50} + p_{54} = p_{58} + p_{62},\ p_{51} + p_{55} = p_{59} + p_{63}; \tag{24}$$

$$p_0 + p_{16} = p_{32} + p_{48},\ p_1 + p_{17} = p_{33} + p_{49},\ p_2 + p_{18} = p_{34} + p_{50},\ p_3 + p_{19} = p_{35} + p_{51},$$
$$p_4 + p_{20} = p_{36} + p_{52},\ p_5 + p_{21} = p_{37} + p_{53},\ p_6 + p_{22} = p_{38} + p_{54},\ p_7 + p_{23} = p_{39} + p_{55},$$
$$p_8 + p_{24} = p_{40} + p_{56},\ p_9 + p_{25} = p_{41} + p_{57},\ p_{10} + p_{26} = p_{42} + p_{58},\ p_{11} + p_{27} = p_{43} + p_{59},$$
$$p_{12} + p_{28} = p_{44} + p_{60},\ p_{13} + p_{29} = p_{45} + p_{61},\ p_{14} + p_{30} = p_{46} + p_{62},\ p_{15} + p_{31} = p_{47} + p_{63}. \tag{25}$$

Given below are the Hardy-type non-locality conditions, which are incompatible with the notion of local-realism, but which all genuinely entangled pure state of three-qubit satisfy [22] with the maximum probabilty of success 12.5% [21] .

$$\begin{aligned} &p_{32} = 0,\ p_8 = 0, \\ &p_2 = 0,\ p_{63} = 0,\ p_0 > 0. \end{aligned} \tag{26}$$

Maximizing p_0 subject to satisfying all the conditions given in equations (22), (23), (24), (25), (26), with the help of *Mathematica*, gives the solution as

$$p_0^{max} = \frac{1}{2}, \tag{27}$$

while rest of the sixty four probabilities are given by[1]

$$p_3 = p_{12} = p_{15} = p_{17} = p_{18} = p_{29} = p_{30} = p_{33} = p_{35} = p_{45} = p_{47} = p_{48} = p_{50} = p_{60} = p_{62} = \frac{1}{2},$$

$$\text{and } p_i = 0 \text{ for all } i \in (\{1, 2, \ldots, 63\} - \{3, 12, 15, 17, 18, 29, 30, 33, 35, 45, 47, 48, 50, 60, 62\}). \tag{28}$$

Thus the maximum probability of success of Hardy's argument for three two-level systems too is more in GNLT than in the quantum theory.

4. Conclusion

In conclusion, we have shown here that in a more general framework of GNLT, the maximum probability of success of Hardy's argument can be enhanced for both the three two-level systems and for two two-level systems. Interestingly, the maximum success probability for both type of systems attains a common value 0.5. The same is the maximum probability of success of Cabello's argument for two two-level system in this more general framework.

Quantum non-locality has attracted much attention since its discovery because it relates quantum mechanics with special relativity. Special relativity forbids sending physical information with a speed greater than that of the light in vacuum. This is reflected in the quantum mechanical joint probabilities appearing both in the violation of Bell's inequality as well as in the fulfillment of Hardy's non-locality conditions, although there is no direct

[1] This solution set is not unique, but maximum probabilty is 0.5, for each such set.

relevance of special relativity in the postulates of non-relativistic quantum mechanics. These quantum mechanical joint probabilities are not only non-local but also non-signaling. The non-local probabilities, coming out from quantum mechanical states can give rise to the maximum violation up to the amount $2\sqrt{2}$ of Bell's inequality, whereas there are non-quantum mechanical non-local joint probabilities which give rise to the maximal possible algebraic violation (namely, 4) of Bell's inequality, without violating the relativistic-causality [9]. Why quantum theory can not provide more than $2\sqrt{2}$ violation of the Bell's inequality? By exploiting the theoretical structure of quantum mechanics it has been shown that a violation greater than $2\sqrt{2}$ will result in signalling in quantum mechanics [24, 25, 26]. We have seen in this paper that the no-signaling constraint cannot restrict the maximum value of the non-zero probability appearing in the Hardy's (Cabello's) argument to 0.09(0.11) all by itself. In a generalized non-signalling theory this value can go up to 0.5. It will be an interesting open question to find what feature of quantum mechanics along with no-signalling condition restricts these to their quantum values.

Acknowledgments
It is pleasure to thank Sibasish Ghosh, Guruprasd Kar, Samir Kunkri, Ramij Rahaman, Anirban Roy and Thomas Konrad for many useful discussions and suggestions.

References
[1] Bell J S 1964 *Physics* **1** 195
[2] Hardy L 1992 *Phys. Rev. Lett.* **68** 2981
[3] Cabello A 2002,*Phys. Rev. A***65** 032108
[4] Liang L M and Li C-Z 2005 *Phys. Lett. A* **335** 371
[5] Kar G 1997 *Phys. Rev. A* **56** 1023
[6] Hardy L 1993 *Phys. Rev. Lett.* **71** 1665
[7] Goldstein S 1994 *Phys. Rev. Lett.* **72** 1951
[8] Kunkri S, Choudhary S K, Ahanj A, Joag P 2006 *Phys. Rev. A* **73** 022346
[9] Popescu S and Rohrlich D 1994 *Fond. Phys.***24** 379
[10] Scarani V, Gisin N, Brunner N, Masanes L, Pino S and Acin A 2006 *Phys. Rev. A* **74** 042239
[11] Barrett J, Hardy L and Kent A 2005 *Phys. Rev. Lett.* **95** 010503
[12] Cerf N J, Gisin N, Massar S and Popescu S 2005 *Phys. Rev. Lett.***94** 220403
[13] Masanes L, Acin A and N. Gisin 2006 *Phys. Rev. A***73** 012112
[14] Barett J, Linden N, Massar S, Pironio S, Popescu S, Roberts D 2005 *Phys. Rev. A***71** 022101
[15] Barrett J and Pironio S 2005 *Phys. Rev. Lett.***95** 140401 .
[16] W. van Dam, (*e-print* quant-ph/0501159)
[17] Wolf S and Wullschleger J (*e-print* quant-ph/0502030)
[18] Short A J, Gisin N and Popescu S 2006 *Quant. Inf. Proc.***5**(2) 131.
[19] Short A J, Popescu S and Gisin N 2006 *Phys. Rev. A***73** 012101
[20] Jordan T F 1994 *Phys. Rev. A* **50** 62.
[21] Choudhary S K, Ghosh S , Kar G , Kunkri S, Rahaman R, Roy A 2010 *Quant. Inf. Comp.* **10** 0859
[22] Ghosh S, Kar G and Sarkar D 1998 *Phys. Lett. A* **243** 249
[23] Einstein A, Podolsky B, Rosen N 1935, *Phys. Rev.* **47** 777
[24] Dieks D 2002 *Phys. Rev. A* **66** 062104
[25] Buhrman H and Massar S 2005 *Phys. Rev. A* **72** 052103
[26] Choudhary S K, Kar G, Kunkri S and Rahaman R 2007 *Phys. Lett. A* **371** 395.

Alternative derivation of the master equation for a particle in an external field subject to continuous measurement

K Garapo[1] **and T Konrad**[1]

[1] School of Physics, University of KwaZulu-Natal, Westville, Durban 4000, South Africa.

E-mail: 209523532@ukzn.ac.za and konradt@ukzn.ac.za

Abstract. The theory of continuous measurement provides a tool to monitor the evolution of the wave function of a single quantum system in real time. We re-derive the master equation in the non-selective regime for the dynamics of the wave function of a particle in an external potential which is subject to continuous measurement of position. In the derivation we view continuous measurement as the limit of a sequence of unsharp position measurements. Unsharp position measurements are achieved by selecting generalised measurement observables, or in mathematical terms, positive operator valued measures (POVM) rather than the standard von Neumann projection operators which are a special class of the sub-class of POVM's called projection valued measures (PVM). We also introduce a commutative algebra that allows us to perform commutative operations with non-commuting position operators. We then deduce the stochastic Ito equations for the selective regime of measurement.

1. Introduction

In 1987, Caves and Milburn [1] suggested a model for the continuous measurement of the position of a quantum system. Their model was based on the theory of continuous quantum measurement as suggested, in 1982, by Barchielli [2] et al. In 1988, Diosi [3] showed that continuous measurement of position in the selective regime can be represented by a certain Ito stochastic master equation. In this paper we take a simplified approach to re-derive the master equations for continuous position measurement in both the selective and non-selective regimes. In our approach, we view continuous position measurement as a sequence of unsharp measurements of position of a quantum system in the following limit;

$$\lim_{\substack{\tau\to 0\\ \sigma\to\infty}} \frac{1}{\sigma^2\tau} = \gamma, \tag{1}$$

where τ is the time interval between two consecutive measurements, σ is the precision parameter of the measurements and γ is a finite quantity called the decoherence rate. For historical reasons we shall refer to this limit as the Barchielli limit. We represent the unsharp measurement of position of a quantum system by generalised position observables.

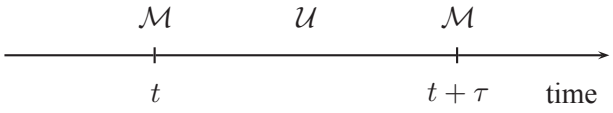


Figure 1. Schematic diagram for the time evolution of a system undergoing a sequence of measurements $\hat{M}$ at time intervals τ. Between two consecutive measurements the closed system evolves unitarily.

2. Non-selective regime

We consider weak position measurements of a system with one spatial degree of freedom x, with outcomes $\bar{x}$, represented by Kraus operators [4]

$$\hat{M}_{\bar{x}} = \frac{1}{\sqrt[4]{2\pi\sigma^2}} \exp\Big(\frac{-(\hat{x}-\bar{x})^2}{4\sigma^2}\Big), \tag{2}$$

and unitaries

$$\hat{U} = \exp\Big(-\frac{i}{\hbar}\hat{H}\tau\Big) = \mathbb{1} - \frac{i}{\hbar}\hat{H}\tau + \mathcal{O}(\tau^2), \tag{3}$$

where $\hat{H}$ is the Hamiltonian of the system, $\hat{x}$ is the position measurement operator and $\hbar = h/2\pi$ (Planck's constant divided by 2π). The effects $\hat{E}_{\bar{x}} = \hat{M}_{\bar{x}}^{\dagger}\hat{M}_{\bar{x}}$ of the measurements are Gaussian. If at time t, the state of the system is represented by the density operator $\hat{\rho}(t)$, then after a time τ, the state of the system is given by (compare figure 1.)

$$\begin{aligned}\hat{\rho}(t+\tau) &= \int_{-\infty}^{\infty} d\bar{x}\, \hat{U}\hat{M}_{\bar{x}}\hat{\rho}(t)\hat{M}_{\bar{x}}^{\dagger}\hat{U}^{\dagger} \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\hat{U}\Big(\int_{-\infty}^{\infty} d\bar{x} \exp\Big(\frac{-(\hat{x}-\bar{x})^2}{4\sigma^2}\Big)\hat{\rho}(t)\exp\Big(\frac{-(\hat{x}-\bar{x})^2}{4\sigma^2}\Big)\Big)\hat{U}^{\dagger}. \end{aligned} \tag{4}$$

Inorder to carry out the integration we introduce a commutative super-algebra [5, 6] with position operators $\hat{x}_L$ and $\hat{x}_R$ which are defined as follows;

$$\hat{x}_L\hat{\rho}(t) \equiv \hat{\rho}(t)\hat{x}_L := \hat{x}\hat{\rho}(t) \tag{5}$$

and

$$\hat{x}_R\hat{\rho}(t) \equiv \hat{\rho}(t)\hat{x}_R := \hat{\rho}(t)\hat{x}. \tag{6}$$

Given that the operator $\hat{x}$ has the following spectral representation,

$$\hat{x} = \int_{-\infty}^{\infty} x\, \hat{P}(dx) \tag{7}$$

where x are position eigenvalues and $\hat{P}(dx) = dx\, |x\rangle\langle x|$ is a projection valued measure. In a similar way we can expand the operators $\hat{x}_L$ and $\hat{x}_R$ as follows;

$$\hat{x}_L = \int_{-\infty}^{\infty} x'\, \hat{P}_L(dx') \text{ and } \hat{x}_R = \int_{-\infty}^{\infty} x''\, \hat{P}_R(dx''). \tag{8}$$

The actions of $\hat{P}_L(dx')$ and $\hat{P}_R(dx'')$ on any arbitrary operator $\hat{A}$ on the Hilbert space $\mathcal{H}$ are defined as follows;

$$\hat{P}_L(dx')\hat{A} \equiv \hat{A}\,\hat{P}_L(dx') := \hat{P}(dx)\,\hat{A}, \tag{9}$$

and

$$\hat{P}_R(dx'')\hat{A} \equiv \hat{A}\,\hat{P}_R(dx'') := \hat{A}\,\hat{P}(dx) \tag{10}$$

respectively. The equations (9) and (10) are consistent with equations (5) and (6). Since the operators $\hat{x}_L$ and $\hat{x}_R$ commute with all operators on the Hilbert space we can rewrite equation (4) as follows;

$$\begin{aligned}\hat{\rho}(t+\tau) =& \frac{1}{\sqrt{2\pi\sigma^2}}\hat{U}\Big(\int_{-\infty}^{\infty} d\bar{x}\exp\Big(\frac{-(\hat{x}_L-\bar{x})^2-(\hat{x}_R-\bar{x})^2}{4\sigma^2}\Big)\Big)\hat{\rho}(t)\hat{U}^\dagger \\ =& \frac{1}{\sqrt{2\pi\sigma^2}}\hat{U}\Big(\int_{-\infty}^{\infty} d\bar{x}\exp\Big(-\frac{1}{2\sigma^2}\bar{x}^2+\frac{\hat{x}_L+\hat{x}_R}{2\sigma^2}\bar{x}-\frac{\hat{x}_L^2+\hat{x}_R^2}{4\sigma^2}\Big)\Big)\hat{\rho}(t)\hat{U}^\dagger. \end{aligned}\tag{11}$$

We shall evaluate the integral in equation (11) as follows;

$$\begin{aligned}&\int_{-\infty}^{\infty} d\bar{x}\exp\Big(-\frac{1}{2\sigma^2}\bar{x}^2+\frac{\hat{x}_L+\hat{x}_R}{2\sigma^2}\bar{x}-\frac{\hat{x}_L^2+\hat{x}_R^2}{4\sigma^2}\Big) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big[\int_{-\infty}^{\infty} d\bar{x}\exp\Big(-\frac{1}{2\sigma^2}\bar{x}^2+\frac{x'+x''}{2\sigma^2}\bar{x}-\frac{(x')^2+(x'')^2}{4\sigma^2}\Big)\Big]\,\hat{P}_L(dx')\,\hat{P}_R(dx'') \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big[\sqrt{2\pi\sigma^2}(1-\frac{1}{8\sigma^2}((x')^2+(x'')^2-2x'x'')+\mathcal{O}(\sigma^{-4}))\Big]\,\hat{P}_L(dx')\,\hat{P}_R(dx'') \\ &= \sqrt{2\pi\sigma^2}\Big(\mathbb{1}-\frac{1}{8\sigma^2}(\hat{x}_L^2+\hat{x}_R^2-2\hat{x}_L\hat{x}_R)+\mathcal{O}(\sigma^{-4})\Big), \end{aligned}\tag{12}$$

where the integration over the possible measurement results $\bar{x}$ has been carried out and the obtained exponential Taylor-expanded. Substituting equations (3) and (12) into equation (11) yields the following;

$$\begin{aligned}\Delta\hat{\rho}(t) =& \hat{\rho}(t)(t+\tau)-\hat{\rho}(t) = (\mathbb{1}-\frac{i}{\hbar}\hat{H}\tau+\mathcal{O}(\tau^2)) \\ &\times(\mathbb{1}-\frac{1}{8\sigma^2}(\hat{x}_L^2+\hat{x}_R^2-2\hat{x}_L\hat{x}_R)+\mathcal{O}(\sigma^{-4}))\hat{\rho}(t)(\mathbb{1}+\frac{i}{\hbar}\hat{H}\tau+\mathcal{O}(\tau^2))-\hat{\rho}(t) \\ &= -\frac{i}{\hbar}[\hat{H},\hat{\rho}(t)]\tau-\frac{1}{8\sigma^2}[\hat{x},[\hat{x},\hat{\rho}(t)]]+\frac{i\tau}{8\hbar\sigma^2}[\hat{H},[\hat{x},[\hat{x},\hat{\rho}(t)]]]+\mathcal{O}(\tau^2)+\mathcal{O}(\sigma^{-4}). \end{aligned}\tag{13}$$

In the Barchielli limit, equation (13) becomes

$$d\hat{\rho}(t) = -\frac{i}{\hbar}[\hat{H},\hat{\rho}(t)]dt-\frac{\gamma}{8}[\hat{x},[\hat{x},\hat{\rho}(t)]]dt. \tag{14}$$

We note that the higher order terms in τ vanish as τ approaches 0. Equation (14) is the master equation of continuous position measurement in the non-selective regime.

3. Selective Regime

While we average the state of the system after measurement over all possible outcomes in the non-selective regime, we have to account for the measurement results in the selective regime. If at a time t the state of the system is represented by the density operator $\hat{\rho}(t)$, then after a time τ the state of the system is given by

$$\hat{\rho}(t+\tau) = \frac{1}{p_{\hat{\rho}}(\bar{x})}\hat{U}\hat{M}_{\bar{x}}\hat{\rho}(t)\hat{M}_{\bar{x}}^{\dagger}\hat{U}^{\dagger}, \tag{15}$$

where $p_{\hat{\rho}}(\bar{x})$ is the probability of obtaining the measurement result $\bar{x}$ given that the state of the system is $\hat{\rho}$. The inverse of the probability is evaluated as follows;

$$\begin{aligned}\frac{1}{p_{\hat{\rho}}(\bar{x})} =& \Big[\mathbf{tr}\Big\{\frac{1}{\sqrt{2\pi\sigma^2}}\exp\Big(-\frac{(\hat{x}-\bar{x})^2}{2\sigma^2}\Big)\hat{\rho}(t)\Big\}\Big]^{-1}\\ \approx& \Big[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\Big(-\frac{\bar{x}^2}{2\sigma^2}\Big)\Big(1+\frac{\bar{x}}{\sigma^2}\langle\hat{x}\rangle_{\hat{\rho}}\Big)\Big]^{-1}\\ \approx& \sqrt{2\pi\sigma^2}\exp\Big(\frac{\bar{x}^2}{2\sigma^2}\Big)\Big(1-\frac{\bar{x}}{\sigma^2}\langle\hat{x}\rangle_{\hat{\rho}}+\frac{\bar{x}^2}{\sigma^4}\langle\hat{x}\rangle_{\hat{\rho}}^2\Big),\end{aligned} \tag{16}$$

and the measurement operator is expanded as follows;

$$\hat{M}_{\bar{x}} = \frac{1}{\sqrt{4\pi\sigma^2}}\exp\Big(-\frac{\bar{x}^2}{4\sigma^2}\Big)\Big(\mathbb{1}-\frac{\hat{x}^3\bar{x}}{8\sigma^4}+\frac{\hat{x}^2\bar{x}^2}{8\sigma^4}+\frac{\hat{x}\bar{x}}{2\sigma^2}+\frac{\hat{x}^4}{32\sigma^4}-\frac{\hat{x}^2}{4\sigma^2}+\mathcal{O}(\sigma^{-6})\Big). \tag{17}$$

In doing both expansions we take note of the fact that except for terms in $\bar{x}^2/\sigma^{-4}$ all terms in σ^{-4} and below vanish in the Barchielli limit. To simplify the evaluation of equation (15), we first evaluate the measurement part.

$$\begin{aligned}&\frac{\hat{M}_{\bar{x}}\hat{\rho}(t)\hat{M}_{\bar{x}}^{\dagger}}{p_{\hat{\rho}}(\bar{x})}\\ =&\Big(1-\frac{\bar{x}}{\sigma^2}\langle\hat{x}\rangle_{\hat{\rho}}+\frac{\bar{x}^2}{\sigma^4}\langle\hat{x}\rangle_{\hat{\rho}}^2\Big)\Big(\mathbb{1}-\frac{\hat{x}^3\bar{x}}{8\sigma^4}+\frac{\hat{x}^2\bar{x}^2}{8\sigma^4}+\frac{\hat{x}\bar{x}}{2\sigma^2}+\frac{\hat{x}^4}{32\sigma^4}-\frac{\hat{x}^2}{4\sigma^2}+\mathcal{O}(\sigma^{-6})\Big)\hat{\rho}(t)\\ &\times\Big(\mathbb{1}-\frac{\hat{x}^3\bar{x}}{8\sigma^4}+\frac{\hat{x}^2\bar{x}^2}{8\sigma^4}+\frac{\hat{x}\bar{x}}{2\sigma^2}+\frac{\hat{x}^4}{32\sigma^4}-\frac{\hat{x}^2}{4\sigma^2}+\mathcal{O}(\sigma^{-6})\Big)\\ =&\Big(1-\frac{\bar{x}}{\sigma^2}\langle\hat{x}\rangle_{\hat{\rho}}+\frac{\bar{x}^2}{\sigma^4}\langle\hat{x}\rangle_{\hat{\rho}}^2\Big)\Big(\hat{\rho}(t)+\frac{\bar{x}}{2\sigma^2}\{\hat{x},\hat{\rho}(t)\}-\frac{1}{8\sigma^2}[\hat{x},[\hat{x},\hat{\rho}(t)]]-\frac{1}{8\sigma^2}\{\hat{x},\{\hat{x},\hat{\rho}(t)\}\}\\ &+\frac{\bar{x}^2}{8\sigma^4}\{\hat{x},\{\hat{x},\hat{\rho}(t)\}\}-\frac{\bar{x}}{8\sigma^4}\{\hat{x}^3,\hat{\rho}(t)\}+\frac{1}{32\sigma^4}\{\hat{x}^4,\hat{\rho}(t)\}\\ &-\frac{\bar{x}}{4\sigma^4}\hat{x}^2\hat{\rho}(t)\hat{x}+\frac{1}{16\sigma^4}\hat{x}^2\hat{\rho}(t)\hat{x}^2+\mathcal{O}(\sigma^{-6})\Big)\\ =&\Big(1-\bar{x}\tau\gamma\langle\hat{x}\rangle_{\hat{\rho}}+\bar{x}^2\tau^2\gamma^2\langle\hat{x}\rangle_{\hat{\rho}}^2\Big)\Big(\hat{\rho}(t)+\frac{\bar{x}\tau\gamma}{2}\{\hat{x},\hat{\rho}(t)\}-\frac{\tau\gamma}{8}[\hat{x},[\hat{x},\hat{\rho}(t)]]\\ &-\frac{\tau\gamma}{8}\{\hat{x},\{\hat{x},\hat{\rho}(t)\}\}+\frac{\bar{x}^2\tau^2\gamma^2}{8}\{\hat{x},\{\hat{x},\hat{\rho}(t)\}\}-\frac{\bar{x}\tau^2\gamma^2}{8}\{\hat{x}^3,\hat{\rho}(t)\}+\frac{\tau^2\gamma^2}{32}\{\hat{x}^4,\hat{\rho}(t)\}\\ &-\frac{\bar{x}\tau^2\gamma^2}{4}\hat{x}^2\hat{\rho}(t)\hat{x}+\frac{\tau^2\gamma^2}{16}\hat{x}^2\hat{\rho}(t)\hat{x}^2+\mathcal{O}(\tau^3)\Big),\end{aligned} \tag{18}$$

in the Barchielli limit. We now take the approach of [3] and [7] and introduce stochastic equations that govern our measurement results $\bar{x}$. From [3] we know that $\bar{x}_t = \langle\hat{x}\rangle_{\hat{\rho}(t)} + \frac{1}{\sqrt{\gamma}}w_t$, where w_t is standard white noise which is defined by $\langle w_t\rangle_{st} = 0$ and $\langle w_t w_s\rangle_{st} = \delta(t-s)$. [3] mentions the mathematical

shortcomings of the equation and introduces the mathematically well behaved quantity Q_t which is the time integrated measurement signal, given by

$$Q_t = \int_0^t \bar{x}_{t'}\, dt' = \int_0^t \left(\langle \hat{x} \rangle_{\hat{\rho}(t')} + \frac{1}{\sqrt{\gamma}} w_{t'} \right) dt'. \tag{19}$$

The integral in equation (19) is generally not easy to evaluate and hence we prefer the differential form,

$$dQ_t = \langle \hat{x} \rangle_{\hat{\rho}(t)} dt + \gamma^{-\frac{1}{2}}\, dW_t, \tag{20}$$

where $W_t = \int_0^t w_{t'} dt'$ is a Wiener process. The Wiener increments dW_t satisfy the following Ito rules; $\langle dW_t \rangle_{st} = 0$, $(dW_t)^2 = dt$, and $(dW_t)^n = 0$ for $n > 2$. In the Barchielli limit, $\bar{x}\tau = dQ_t$. It follows that $\bar{x}^2\tau^2 = (dQ_t)^2 = \frac{1}{\gamma} dt$, after applying Ito rules. Equation (18) thus reduces to

$$\begin{aligned}
&\frac{\hat{M}_{\bar{x}} \hat{\rho}(t) \hat{M}_{\bar{x}}^\dagger}{p_{\hat{\rho}}(\bar{x})} \\
=&\Big(1 - (dQ_t)\gamma \langle \hat{x} \rangle_{\hat{\rho}} + \tau\gamma \langle \hat{x} \rangle_{\hat{\rho}}^2 \Big) \Big(\hat{\rho}(t) + \frac{(dQ_t)\gamma}{2} \{\hat{x}, \hat{\rho}(t)\} - \frac{\tau\gamma}{8} [\hat{x}, [\hat{x}, \hat{\rho}(t)]] \\
&- \frac{\tau\gamma}{8} \{\hat{x}, \{\hat{x}, \hat{\rho}(t)\}\} + \frac{\tau\gamma^2}{8\gamma} \{\hat{x}, \{\hat{x}, \hat{\rho}(t)\}\} \Big) \\
=&\Big(1 - (dQ_t)\gamma \langle \hat{x} \rangle_{\hat{\rho}} + \tau\gamma \langle \hat{x} \rangle_{\hat{\rho}}^2 \Big) \Big(\hat{\rho}(t) + \frac{(dQ_t)\gamma}{2} \{\hat{x}, \hat{\rho}(t)\} - \frac{\tau\gamma}{8} [\hat{x}, [\hat{x}, \hat{\rho}(t)]] \Big) \\
=&\hat{\rho}(t) - \frac{\tau\gamma}{8} [\hat{x}, [\hat{x}, \hat{\rho}(t)]] + \frac{(dQ_t)\gamma}{2} \{\hat{x}, \hat{\rho}(t)\} - (dQ_t)\gamma \hat{\rho}(t) \langle \hat{x} \rangle_{\hat{\rho}} \\
&- \frac{\tau\gamma \langle \hat{x} \rangle_{\hat{\rho}}}{2} \{\hat{x}, \hat{\rho}(t)\} + \tau\gamma \langle \hat{x} \rangle_{\hat{\rho}}^2 \hat{\rho}(t) \\
=&\hat{\rho}(t) - \frac{\tau\gamma}{8} [\hat{x}, [\hat{x}, \hat{\rho}(t)]] + \frac{(dQ_t)\gamma}{2} \{\hat{x} - \langle \hat{x} \rangle_{\hat{\rho}}, \hat{\rho}(t)\} - \frac{\tau\gamma \langle \hat{x} \rangle_{\hat{\rho}}}{2} \{\hat{x} - \langle \hat{x} \rangle_{\hat{\rho}}, \hat{\rho}(t)\} \\
=&\hat{\rho}(t) - \frac{\tau\gamma}{8} [\hat{x}, [\hat{x}, \hat{\rho}(t)]] + \frac{\gamma}{2} \{\hat{x} - \langle \hat{x} \rangle_{\hat{\rho}}, \hat{\rho}(t)\} (dQ_t - \tau \langle \hat{x} \rangle_{\hat{\rho}}).
\end{aligned} \tag{21}$$

From equation (21) it follows that

$$\begin{aligned}
&\hat{\rho}(t + \tau) \\
=&\Big(\mathbb{1} - \frac{i}{\hbar} \hat{H} \tau + \mathcal{O}(\tau^2) \Big) \Big(\hat{\rho}(t) - \frac{\tau\gamma}{8} [\hat{x}, [\hat{x}, \hat{\rho}(t)]] + \frac{\gamma}{2} \{\hat{x} - \langle \hat{x} \rangle_{\hat{\rho}}, \hat{\rho}(t)\} (dQ_t - \tau \langle \hat{x} \rangle_{\hat{\rho}}) \Big) \\
&\times \Big(\mathbb{1} + \frac{i}{\hbar} \hat{H} \tau + \mathcal{O}(\tau^2) \Big) \\
=&\hat{\rho}(t) - \frac{\tau\gamma}{8} [\hat{x}, [\hat{x}, \hat{\rho}(t)]] + \frac{\gamma}{2} \{\hat{x} - \langle \hat{x} \rangle_{\hat{\rho}}, \hat{\rho}(t)\} (dQ_t - \tau \langle \hat{x} \rangle_{\hat{\rho}}) - \frac{i}{\hbar} \hat{H} \tau \hat{\rho}(t) + \hat{\rho}(t) \frac{i}{\hbar} \hat{H} \tau \\
=&\hat{\rho}(t) - \frac{i\tau}{\hbar} [\hat{H}, \hat{\rho}(t)] - \frac{\tau\gamma}{8} [\hat{x}, [\hat{x}, \hat{\rho}(t)]] + \frac{\gamma}{2} \{\hat{x} - \langle \hat{x} \rangle_{\hat{\rho}}, \hat{\rho}(t)\} (dQ_t - \tau \langle \hat{x} \rangle_{\hat{\rho}}),
\end{aligned} \tag{22}$$

and

$$d\hat{\rho}(t) = -\frac{i}{\hbar} [\hat{H}, \hat{\rho}(t)] dt - \frac{\gamma}{8} [\hat{x}, [\hat{x}, \hat{\rho}(t)]] dt + \frac{\gamma}{2} \{\hat{x} - \langle \hat{x} \rangle_{\hat{\rho}}, \hat{\rho}(t)\} (dQ_t - \langle \hat{x} \rangle_{\hat{\rho}} dt). \tag{23}$$

Equation (23) is the master equation for continuous measurement of position in the selective regime. We refer readers who wish to have in-depth understanding of the stochastic calculus we used in this derivation to the book by Gardiner [8].

4. Conclusion

We re-derived the master equations for the continuous measurement of position in both the selective and non-selective regimes. In the derivation of the master equations we applied a commutative super-algebra and the Ito stochastic calculus, which was suggested by Diosi [3]. In contrast to Diosi, our approach is based on combining the Kraus representation of the state change due to measurement with the Ito calculus by expressing the integrated measurement signal Q by means of a Wiener process. This leads to the simplification of the derivation. The derived master equations are important tools to describe state monitoring and control of individual quantum systems.

Acknowledgement

The authors would like to thank Dr. S. K. Choudhary for helpful discussions and remarks.

References

[1] Caves C M and Milburn G J 1987 *Phys. Rev. A* **36** 5543
[2] Barchielli A, Lanz L and Prosperi G 1982 *Il Nuovo Cimento B (1971-1996)* **72**(1) 79
[3] Diosi L 1988 *Phys. Lett. A* **129** 419
[4] Kraus K 1983 *Lecture Notes in Physics* **190**
[5] Audretsch J, Diósi L and Konrad T 2003 *Phys. Rev. A* **68** 034302
[6] Konrad T 2003 Ph.D. thesis Universitat Konstanz
[7] Diosi L, Konrad T, Scherer A and Audretsch J 2006 (*Preprint* `quant-ph/0608097`)
[8] Gardiner C W 1985 *Handbook of stochastic methods for physics, chemistry, and the natural sciences* vol 2nd (Springer-Verlag)

Entanglement dynamics in an oscillating bipartite Gaussian state coupled to reservoirs at different temperatures

A Ghesquière, I Sinayskiy, F Petruccione

Quantum Research Group, School of Physics, University of KwaZulu-Natal and National Institute for Theoretical Physics, Westville Campus, Durban, 4000, South Africa

E-mail: ghesquiere@ukzn.ac.za, sinayskiy@ukzn.ac.za, petruccione@ukzn.ac.za

Abstract. An entangled bipartite Gaussian state is coupled to two thermal reservoirs, one for each particle and a harmonic oscillation is allowed between the two particles. The reservoirs are assumed to have different dynamics and to be coupled to their particles with different coupling strengths. This allows for a realistic situation where a bipartite state may be shared between two parties who "keep" their part in different environments. We solve a master equation previously derived in the non-rotating wave approximation for the system. We show the effect of a variation in the bath temperature on the entanglement, as well as that of the variation in coupling strengths.

1. Introduction

Entanglement is one of quantum mechanics' most useful resources, yet also one of the most fragile; as such, the loss, and in some cases revival, of entanglement in a system coupled to an environment, due to said environment, has been widely studied, see e.g. [1] for an extensive review. As a continuation of previous work [2,3], where the authors studied an initially entangled bipartite system shared between two parties keeping their respective environments at the same temperature, we now examine the case where these temperatures are different. This allows for a somewhat more realistic situation, where an entangled state may be shared between two parties who are not necessarily expected to keep their environment in the same conditions.

To evolve the state, we use a pre-Lindblad master equation, derived in the Non-Rotating Wave (N.R.W.) approximation. The derivation can be found in [2–6]. The N.R.W. master equation is often well suited for systems which are expected to be strongly coupled to their environments. Other pre-Lindblad equations have been derived, using other methods. One may in particular cite that of Caldeira and Leggett [7] who used a path-integrals method, or that of Diósi [8, 9] who obtained an equation valid for low temperatures.

We choose the initial state of the bipartite system to be Gaussian, since there is a formalism [10–13] which allows for a simple study of Gaussian states. These states form a class of continuous variables states which is becoming of growing importance to the field of quantum information processing, both for the ease with which they are manipulated experimentally [14] and the ease with which they can be analytically studied. The latter has led to a variety of studies of systems coupled to heat baths, such as [15–17] to cite but a few.

In the following, we will recall the master equation and solve it for the particular dynamics chosen. We will then discuss some examples of entanglement behaviour.

2. Entanglement dynamics

It was found in [3,18,19] that the dynamics of the entanglement in a bipartite system coupled to an environment is greatly influence by allowing an harmonic interaction between the system's particles. We study a system of two particles of equal mass, each one coupled to its own heat bath; they have coordinates x_1 and x_2, momenta p_1 and p_2; ω_0 denotes the frequency of the oscillation. The overall Hamiltonian reads as

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + \frac{m\omega_0^2}{2}(x_1 - x_2)^2 + \sum_j \left\{ \frac{p_j^2}{2m_j} + \frac{m_j\omega_j^2}{2}(q_j - x_1)^2 \right\} + \sum_k \left\{ \frac{p_k^2}{2m_k} + \frac{m_k\omega_k^2}{2}(q_k - x_2)^2 \right\}. \quad (1)$$

For our study, we take all masses to be equal. The ω's denote the frequencies of the bath oscillators. The initial state is chosen to be the Gaussian state [20, 21]

$$\Psi(x_1, x_2) = \sqrt{\frac{1}{2\pi sd}} e^{-\frac{(x_1-x_2)^2}{4s^2}} e^{-\frac{(x_1+x_2)^2}{16d^2}}, \quad (2)$$

where s and d denote the distance between the particles and the width of the center-of-mass system respectively. With position coupling, the master equation becomes

$$\dot{\rho} = -\frac{\imath}{\hbar}[H_s, \rho] - \frac{\imath\gamma_1}{2\hbar}\left[x_1, [p_1, \rho]_+\right] - \frac{\gamma_1 kT_1}{\hbar^2}[x_1, [x_1, \rho]] - \frac{\imath\gamma_2}{2\hbar}\left[x_2, [p_2, \rho]_+\right] - \frac{\gamma_2 kT_2}{\hbar^2}[x_2, [x_2, \rho]], \quad (3)$$

where T_1 and T_2 are the temperatures of bath 1 and 2 respectively, γ_1 and γ_2 the coupling constants of bath 1 and 2 respectively and $[\,,\,]_+$ represents the anti-commutator. The 1 indices denote particle 1 and bath 1, the 2 indices, particle 2 and bath 2. Writing the density matrix in position representation, $\rho(x_1, x_2; y_1, y_2)$, we get

$$\begin{aligned}\frac{\partial\rho}{\partial t} =& \frac{\imath\hbar}{2m}\left(\frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial x_2^2} - \frac{\partial^2}{\partial y_2^2}\right)\rho \\ &- \frac{\imath m\omega_0^2}{2\hbar}\left((x_1 - x_2)^2 - (y_1 - y_2)^2\right)\rho \\ &- \frac{\gamma_1}{2m}(x_1 - y_1)\left(\frac{\partial}{\partial x_1} - \frac{\partial}{\partial y_1}\right)\rho - \frac{\gamma_1 kT_1}{\hbar^2}(x_1 - y_1)^2\rho \\ &- \frac{\gamma_2}{2m}(x_2 - y_2)\left(\frac{\partial}{\partial x_2} - \frac{\partial}{\partial y_2}\right)\rho - \frac{\gamma_2 kT_2}{\hbar^2}(x_2 - y_2)^2\rho\,. \end{aligned} \quad (4)$$

where k denotes the Boltzmann's constant. To solve this equation, a change of variables is performed $x = u + \hbar z$, $y = u - \hbar z$ and we rename $\rho(\mathbf{x}, \mathbf{y}, t) \to P(\mathbf{u}, \mathbf{z}, t)$, accordingly. Taking the Fourier transform

$$\tilde{P}(\mathbf{q}, \mathbf{z}, t) = \int du_1\, du_2\, P(\mathbf{u}, \mathbf{z}, t) e^{-\imath q_1 u_1 - \imath q_2 u_2}$$

yields the following equation

$$\begin{aligned}&\frac{\partial \tilde{P}}{\partial t}(\mathbf{q},\mathbf{z},t)\\&=\left\{-\frac{1}{2m}\left(q_1\frac{\partial}{\partial z_1}+q_2\frac{\partial}{\partial z_2}\right)-\frac{\gamma_1}{m}\,z_1\,\frac{\partial}{\partial z_1}-\frac{\gamma_2}{m}\,z_2\,\frac{\partial}{\partial z_2}\right.\\&\left.+2m\omega_0^2\left(\frac{\partial}{\partial q_1}-\frac{\partial}{\partial q_2}\right)(z_1-z_2)-4\gamma_1 kT_1\,z_1^2-4\gamma_2 kT_2\,z_2^2\right\}\tilde{P}(\mathbf{q},\mathbf{z},t)\end{aligned}\tag{5}$$

which can be solved to obtain, after some algebra,

$$\begin{aligned}\tilde{P}=\exp\big[&-\mathcal{A}_1q_1^2-\mathcal{A}_2q_2^2-\mathcal{E}q_1q_2-\mathcal{B}_1z_1^2-\mathcal{B}_2z_2^2-\mathcal{D}z_1z_2\\&-\mathcal{C}_{11}z_1q_1-\mathcal{C}_{22}z_2q_2-\mathcal{C}_{12}z_1q_2-\mathcal{C}_{21}z_2q_1\big]\,;\end{aligned}\tag{6}$$

as an example, let us give the explicit expression of the time-dependent coefficient $\mathcal{A}_1$.

$$\begin{aligned}\mathcal{A}_1=&(\epsilon_+ +4k\chi_z^1)\delta_1^{+2}+(\epsilon_+ +4k\chi_z^2)\delta_1^{-2}-(2\epsilon_- -8k\vartheta_1)\delta_1^+\delta_1^-\\&+\left(\frac{\epsilon_+}{4(\epsilon_+^2-\epsilon_-^2)}+4k\chi_q^1\right)\theta_1^{-2}+\left(\frac{\epsilon_+}{4(\epsilon_+^2-\epsilon_-^2)}+4k\chi_q^2\right)\theta_1^{+2}\\&+\left(\frac{\epsilon_-}{2(\epsilon_+^2-\epsilon_-^2)}+8k\vartheta_2\right)\theta_1^-\theta_1^+\\&+8k\left(\delta_1^+\theta_1^-\Lambda_{11}+\delta_1^-\theta_1^+\Lambda_{22}+\delta_1^+\theta_1^+\Lambda_{12}+\delta_1^-\theta_1^-\Lambda_{21}\right)\end{aligned}$$

All greek letters are coefficients, themselves time-dependent and just as bulky.

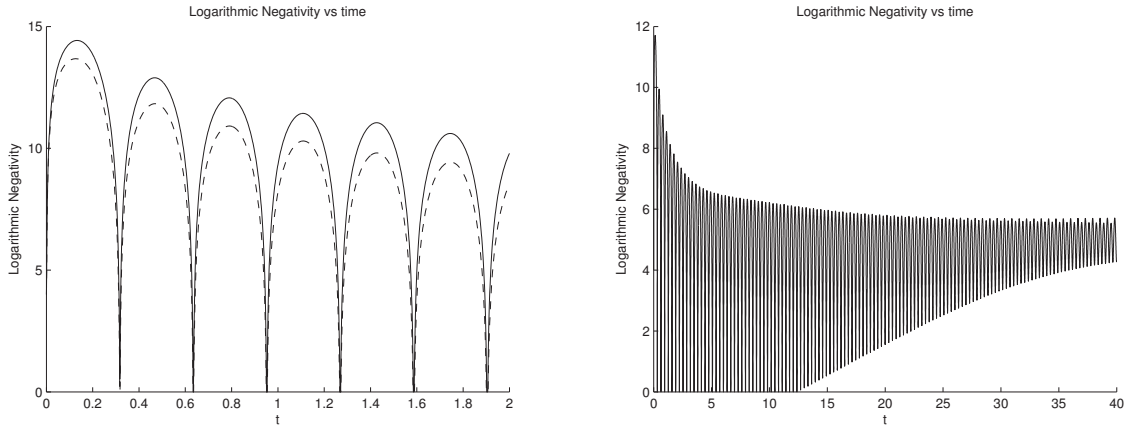
A more detailed solution will be found in [22].

3. Observations and concluding remarks

The following figures illustrate some of the possible evolutions of the entanglement. It should be noted that the behaviours displayed here are obtained in the under-damped case, i.e. when the harmonic potential is greater than either of the coupling constants. We limit our study to this particular class of behaviours since it yields the most striking observations [2,3]. Figure 1a, Figure 2a and Figure 3a illustrate the short time behaviour of the entanglement, whereas Figure 1b, Figure 2b and Figure 3b show the long time behaviour for the same parameters as their counterparts'. It is easily observed at first glance that the behaviour of the entanglement is remarkably similar in all three pairs of figures. The entanglement oscillates with damping oscillations, over the full range of its value until a certain time. After that time, the oscillations do not decrease all the way down to 0; eventually the entanglement oscillates over a stable range. Upon closer examinations of Figure 1a, Figure 2a and Figure 3a, one can see that the range of entanglement is less when the temperatures are higher (Figure 3a). This shows that, as one would expect, the temperatures are crucial to the survival of the entanglement in the system we study. Closer examinations of Figure 1b, Figure 2b and Figure 3b yields similar observations. More precisely, the time at which the oscillations cease to reach the x-axis is longer if the temperatures are equal, but the upper end of the range is lower. This illustrates the difference in the contributions from the baths to the system. To offer a physical interpretation of the behaviour of the entanglement, let us point out that there are two types of entanglement at play here, which compete with each other. The first one is the entanglement resulting from the unitary evolution the system is subjected to and which dies away over time. The second type is the entanglement created by the thermalisation of the system due to its interaction with

the baths, which drive the system to equilibrium. This thermal entanglement is the reason why the entanglement does not disappear completely. Similar behaviour is observed in spin systems [23, 24].

Figure 4 illustrates how increasing or decreasing s, the distance between the particle influences the evolution of entanglement. In particular, larger entanglement is obtained for larger values of s. This suggests that if one was to "squeeze" the system, the entanglement it contained would be greatly reduced. This may be interpreted as follows. If the particles are far apart, the potential energy is quite large and so will be the particles' oscillations around their equilibrium, which may then interfere with each other and destroy the entanglement doing so. On the other hand, if the particles are close together, the potential energy is not so large and the particles then do not oscillate much around their equilibrium.



(a) from $t = 0$ to $t = 2$

(b) from $t = 0$ to $t = 40$

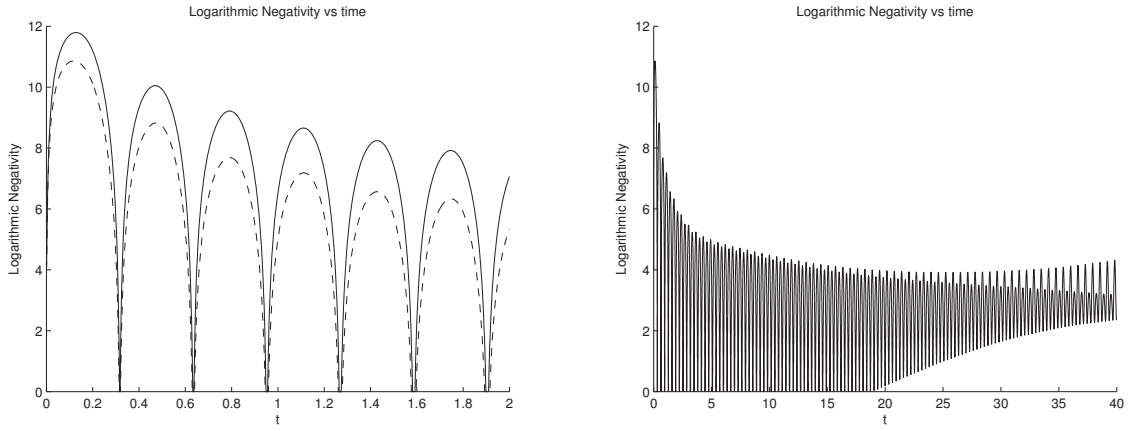
Figure 1: Logarithmic negativity plotted versus time for temperatures $T_1 = 1$ and $T_2 = 1$ with coupling constants $\gamma_2 = 0.8$, full line : $\gamma_1 = 0.1$, dashed line : $\gamma_1 = 0.9$, $s = 1$, $d = 2$, $m = 10$ and $\omega_0 = 7$

Acknowledgements

This work is based on research supported by the South African Research Chair Initiative of the Department of Science and Technology and National Research Foundation.

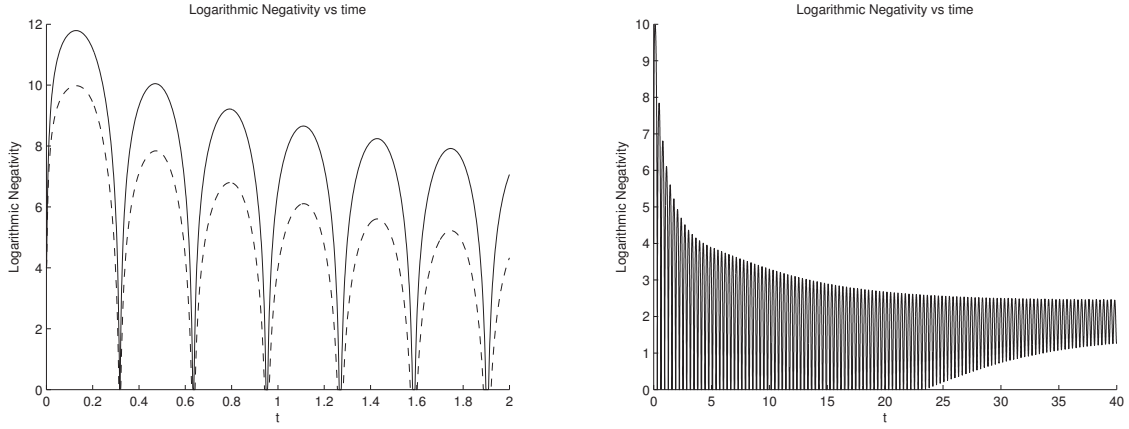
References

[1] Horodecki R, Horodecki P, Horodecki M and Horodecki K 2009 *Reviews of Modern Physics* **81** 865
[2] Ghesquière A 2009 *PhD thesis, available as LANL e-print quant-ph/1105.0564*
[3] Ghesquière A and Dorlas T 2010 *LANL e-print quant-ph/1004.4515 submitted for publication to J. Phys. A*
[4] Gardiner C W and Zoller P 2000 *Quantum Noise* 2nd ed (Springer)
[5] Munro W J and Gardiner C W 1996 *Physical Review A* **53** 4
[6] Breuer H P and Petruccione F 2002 *The Theory of Open Quantum Systems* (Oxford: Oxford Press)
[7] Caldeira A O and Leggett A J 1983 *Physica A* **121** 587
[8] Diòsi 1993 *Physica (Amsterdam)* **199A** 517
[9] Diòsi L 1993 *Europhysics Letters* **22** 1
[10] Plenio M B, Hartley J and Eisert J 2004 *New Journal of Physics* **6** 36
[11] Eisert J and Plenio M B 2003 *International Journal of Quantum Information* **1** 479
[12] Vidal G and Werner R F 2002 *Physical Review A* **65** 3
[13] Anders J 2003 *LANL e-print quant-ph/0610263*



(a) from $t = 0$ to $t = 2$ (b) from $t = 0$ to $t = 40$

Figure 2: Logarithmic negativity plotted versus time for coupling constants $\gamma_1 = 0.79$ and $\gamma_2 = 0.8$ with temperatures $T_2 = 1$, full line : $T_1 = 1$, dashed line : $T_1 = 3$, $s = 1$, $d = 2$, $m = 10$ and $\omega_0 = 7$



(a) from $t = 0$ to $t = 2$ (b) from $t = 0$ to $t = 40$

Figure 3: Logarithmic negativity plotted versus time for coupling constants $\gamma_1 = 0.79$ and $\gamma_2 = 0.8$ with temperatures full line : $T_1 = T_2 = 1$, dashed line : $T_1 = T_2 = 3$, $s = 1$, $d = 2$, $m = 10$ and $\omega_0 = 7$

[14] Ferraro A, Olivares S and Paris M G A 2005 *Gaussian States in Quantum Information* (Napoli Series on Physics and Astrophysics)
[15] Xiang S H, B B S and Song K 2008 *Physical Review A* **78** 052313
[16] Prauzner-Bechcicki J S 2004 *J. Phys. A : Math. Gen.* **37** L173
[17] Serafini A, Paris M G A, Illuminati F and Siena S D 2005 *J. Opt. B : Quantum Semiclass. Opt.* **7** R19
[18] Ficek Z and Tanás R 2006 *Physical Review A* **74** 024304
[19] Ficek Z and Tanás R 2008 *Physical Review A* **77** 054301
[20] Ford G W and O'Connell R F 2008 *Private Communications*
[21] Ford G W, Gao Y and O'Connell R F 2010 *Optics Communications* **283** 831–838
[22] Ghesquière A, Sinayskiy I and Petruccione F 2011 *In preparation*

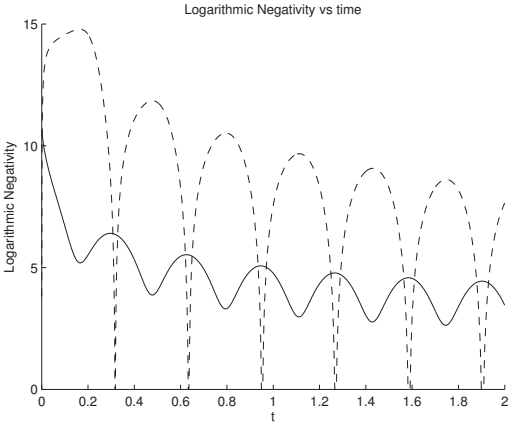


Figure 4: Logarithmic negativity plotted versus time for coupling constants $\gamma_1 = 0.79$, $\gamma_2 = 0.8$, temperatures $T_1 = 1$ and $T_2 = 1$, $d = 2$ and : full line $s = 0.1$, dashed line $s = 15$, $m = 10$ and $\omega_0 = 7$

[23] Sinayskiy I, Burtagh D and Petruccione F 2008 *Physical Review A* **78** 062301
[24] Sinayskiy I, Pumulo N and Petruccione F 2011 *Physics Letters A* **375**(36) 3157–3166

Monte Carlo simulation using GEANT 4 of the EMU μMSR facility at ISIS

JM Hartman and SH Connell

University of Johannesburg, Johannesburg, South Africa

E-mail: hartman.jonathan2@gmail.com

Abstract. A Monte Carlo simulation has been performed for studying Muon Spin Rotation, Resonance and Relaxation (MSR) measurements on the EMU apparatus of the ISIS facility of the Rutherford Appleton Laboratory in the UK. In these measurements, implanted muons which are almost fully polarized act as local probes of their magnetic environment within the particular host material to be studied. The evolution of the polarization of the muon ensemble is monitored by observing the decay positrons. The purpose of the study was to understand the false asymmetry that can be observed for samples which are small compared to the beam lateral dimensions. In this case, a significant number of the muons stop and decay outside of- and downstream of the sample. The GEANT 4 toolkit was used for developing the Monte Carlo simulation. This is a set of C++ libraries developed by CERN to simulate particle physics experiments. However, GEANT 4 can be also be used for other applications that need not have anything to do with particle or high-energy physics. In the simulation, the effect of the external magnetic field on both the incident muon and decay positron trajectories is modeled. It is shown that this leads to a strong field dependence in both the stopping location of muons and the efficiency of the positron detection. The various contributing factors to the systematic field dependent false asymmetry could be quantified. The simulation has been performed in support of the interpretation of experiments on diamond samples where the behaviour of the hydrogen-analogue atom, muonium, is studied.

1. Introduction

The type of MSR measurement under consideration involves the observation of muon ensemble polarization as a function of the longitudinal external magnetic field (LF-μSR). This data can then be decomposed into separate relative contributions from the three states for the muon typically observed in semi-conductors and wide band gap materials [1] . These are the muon in a diamagnetic environment ($f_{\mu_D^+}$), muonium at the tetrahedral interstitial site (f_{Mu_T}) and bond centered muonium ($f_{Mu_{BC}}$). The figure 1 below shows the beam-line, sample position and positron detectors of the EMU facility at ISIS, Rutherford Appleton Laboratory.

The muon polarization is proportional to the measured asymmetry in the count-rates for the forward and backward position detectors,

$$A = \frac{F - \alpha B}{F + \alpha B}$$

with an obvious notation and where the α parameter corrects for efficiency differences between the forward and backward detectors. The asymmetry is clearly field dependent, as seen by a

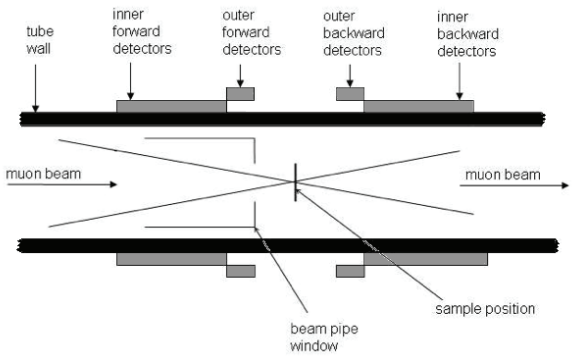


Figure 1. Schematic drawing of the beam-line, sample position and positron detectors of the EMU facility at ISIS, Rutherford Appleton Laboratory. and

measurement shown in figure 2 of the reference silver sample of the same size as the diamonds to be measured.

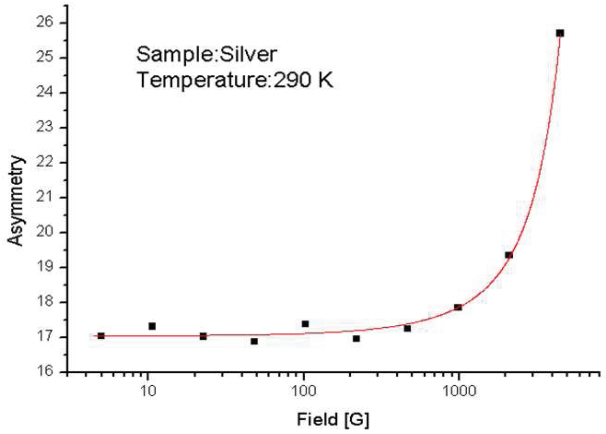


Figure 2. Asymmetry measurement in LF-μSR for a small silver sample, where one expects a constant polarisation as a function of magnetic field due to muons in a diamagnetic environment ($f_{\mu_D^+}$).

The Monte Carlo model should include a description of the material environment in the correct geometry, each muon in the incident muon beam, the interaction of the beam with the sample and the emission of the decay positron, which must be tracked until possible detection in the active material of the forward and backward detectors. This must be done for different external longitudinal dipolar magnetic field strengths, so that the asymmetry can be reconstructed and analyzed for the relative contributions of different effects which may lead to the systematic false asymmetry. In other work, this understanding of the false asymmetry will then be used to analyze the LF-μSR experiments on diamond.

2. Simulation of MUSR

Using the Geant 4 library [2, 3], the application first creates the geometry of the EMU Spectrometer at the ISIS facility. A diagram representing the EMU spectrometer is shown in figure 1. The geometry simulated by the application consists of a tube with the diamond

sample in the center and detectors along the outside of the tube. In the centre and the source of the muons at one end and a stop at the other end. Added to the the setup is also a simulated magnetic field generated by two loops of current, the purpose of which is to create a focusing effect on the muon beams in order to focus them on the diamond. In order to calculate the resultant magnetic field, the field from each of the two loops are calculated separately and then added together vectorially at every point in space. The accuracy of the computed value of the magnetic field was checked by visually comparing the radius of the helical track made by a charged particle in the field to the calculated value made by a pen and paper calculation.

The starting position of the beam is varied randomly within a circular radius of 5 cm at the source of the muon beam. The application was run over a range of diamond sizes, and field strengths on the UJ cluster using a reference swift script (although it could also be run on the SA Grid without swift) in order to determine the asymmetry as a function of diamond size and field strength. 20 Monte Carlo contributions starting with different random number seeds, each sampled with 500000 points were used for each value of diamond size and and field strength. We used internal functions in the Geant 4 library to calculate the random numbers. Geant 4 uses the random number generators from the HEPRandom module from CLHEP. We used GEANT 4's default RANECUE random number engine, which implements a random number generator algorithm with a period of 10^{16} for 32 bit machines and is described in a paper by F. James [4].

Figure 2 shows a graph of the asymmetry of a silver sample used for calibration in a previous experiment using the μSR technique [5, 6] . One would expect the asymmetry of the positrons originating from the sample to be flat due to the symmetry of the magnetic field around the sample. Therefore one of the aims of this simulation is to determine the source of the false asymmetry. In order to parallelize each run across many processes, a single run consisting of many events is broken up into multiple runs each with a smaller number of events but each starting with a different random number seed so that there is still an independent random distribution for each these separate runs. An extra utility was added to be able to kill the secondary tracks depending on whether it originates from the diamond or not. In this way, one would be able to isolate the different possible causes of the false asymmetry in different computer experiments. This can be achieved by plotting graphs of the true asymmetry of just the positrons resulting from muons that interacted with the diamond separately from the false asymmetry from the positrons that did not originate from the diamond. Each run on the cluster outputs a single line of data in a text file. When the application is finished running on the cluster, all the data from each of these output text files can be collected into a single file by the *cat* command. This data text file can be opened in a spreadsheet program, making sure to use a space separated option, in order to remove the unnecessary data and text before being read by a second ROOT [7] application used to plot the data. The ROOT application scans the text file and takes the average of the asymmetry values corresponding to the same values for field strength and diamond size but have different random number seeds. The application then plots a 3D graph of the data using the histogram object in ROOT's library.

In a second computer experiment, the calculations were then repeated except that it was done for a columnated beam of muons and with the number of muons hitting the diamond were counted instead of the asymmetry being calculated. The purpose of this was to check that the simulated magnetic field focuses the beam on the diamond sample as expected. The result of this calculation can then be compared with experimental data. In this case, each run of the application on the cluster still outputs a text file of a line of data like before when the asymmetry is being calculated, except that it outputs an extra variable which is a count of the number muons that interacted with the diamond in that run. This data can be collected into a single text file as before, although it's also read by a modified version of the ROOT application that was previously used to plot asymmetry with respect to the field strength and diamond size.

The modifications to the ROOT application are to take into account the fact that the number of muon interactions with the diamond sample are to be plotted against the field strength and diamond size instead of the asymmetry and that the number of hits calculated with different random number must be added together instead of the average taken. This graph is created using a graph object instead of a histogram object in ROOT's library in order to be able to plot logarithmically spaced data.

3. Results from the Geant 4 simulation

The results are as follows. As shown in figure 3, the total asymmetry forms a slope with respect to the diamond size as well as a small slope that increases with an increase in field strength.

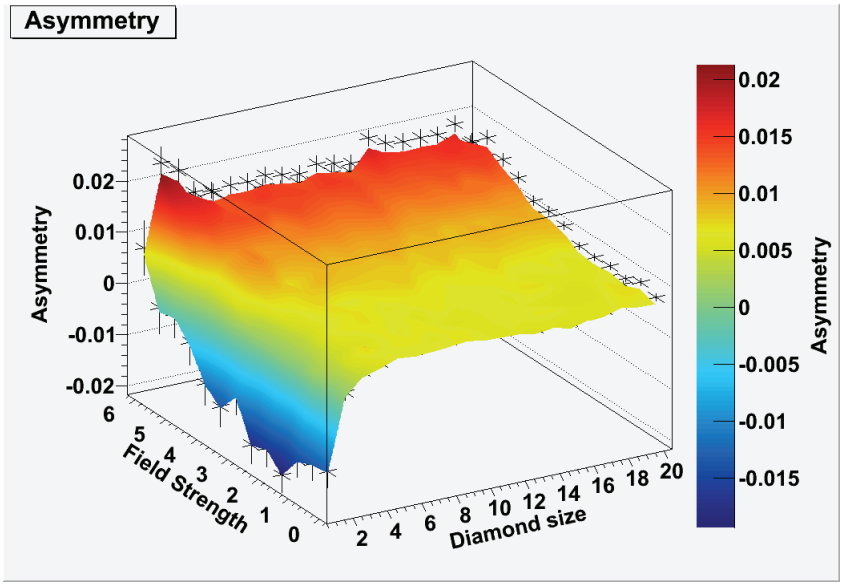


Figure 3. Total asymmetry calculated from the application. This figure is slightly noisy due to statistical fluctuations in the random number generation.

It would be expected that most of the slope with respect to field strength as shown in figure 3 must be due to the false asymmetry from the positrons that originate from the beamstop rather than the diamond sample since the field at the center is symmetrical. This is the case because when only the asymmetry due to the positrons originating from the sample is plotted against diamond size and field strength, then the slope of the graph with respect to field strength is extremely small compared to that in figure 3 as shown in figure 4. A cross section of figure 4 at a diamond size of 5 mm is shown in figure 5.

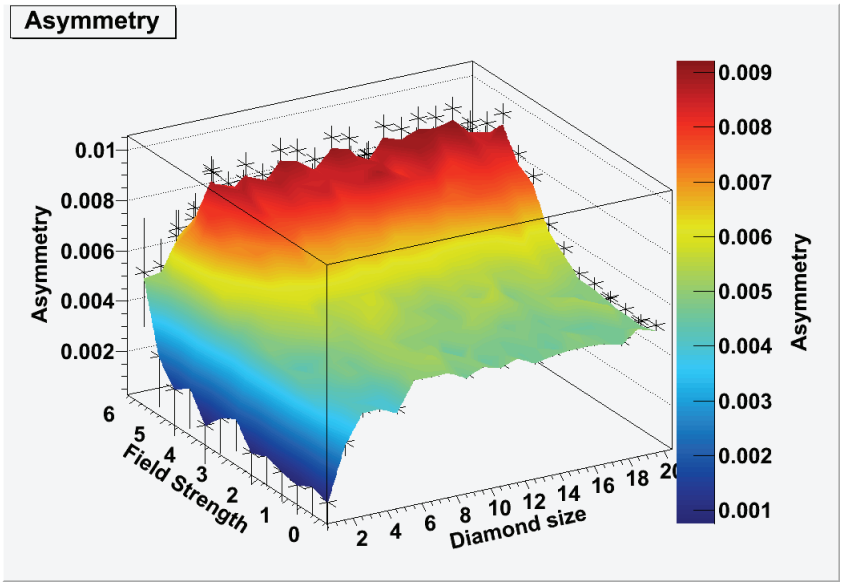


Figure 4. Asymmetry of the positrons originating from the diamond. Again, there are statistical fluctuations in the data.

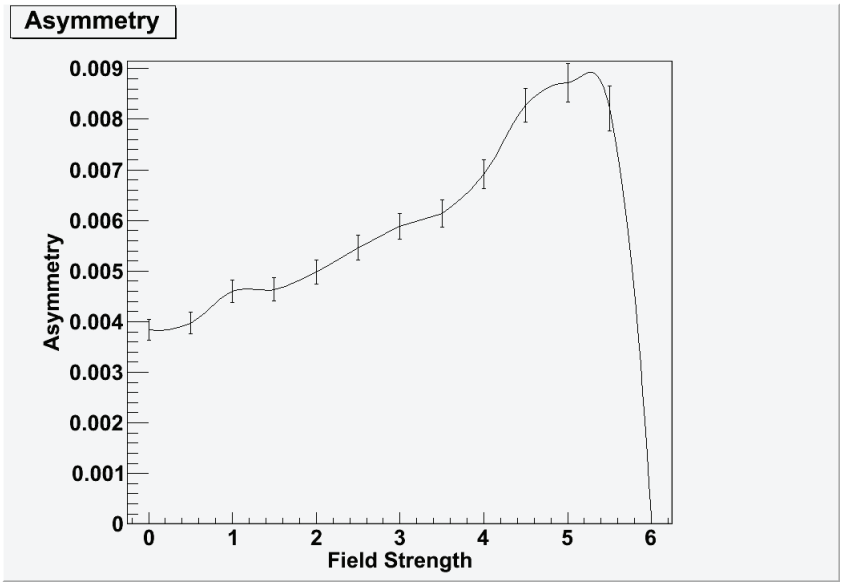


Figure 5. Asymmetry of the positrons originating from the diamond when the diamond size is 5 mm.

.

Figure 6 shows the false asymmetry due to positrons originating from the beam stop.

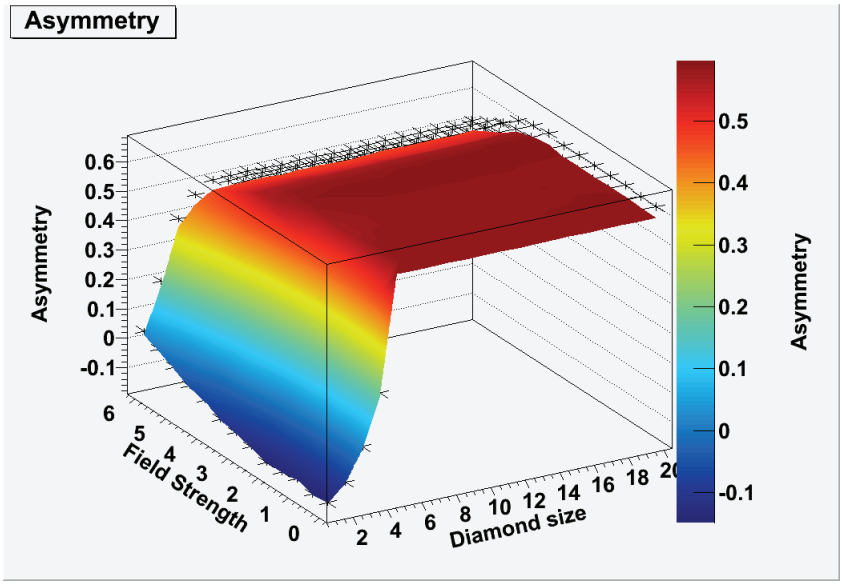


Figure 6. False asymmetry due to the positrons originating from the beam stop.

.

Figure 7 shows the number hits or muon interactions with the diamond with respect to diamond size and field strength. Field strength is plotted on a logarithmic scale. The number of hits increase smoothly with respect to diamond size. The increase in the number of hits with respect to field strength is steeper.

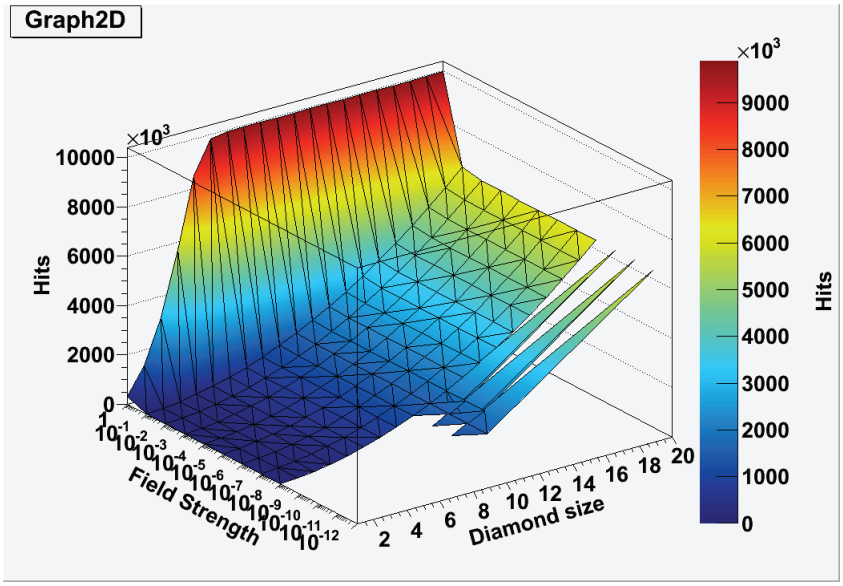


Figure 7. Number of muon interactions with respect to diamond size and field strength, calculated using a columnated muon beam. The triangle objects represent Delaunay triangles which is what ROOT used to render a 3D graph object, as apposed to histogram objects.

.

4. Conclusions and further work

As expected, figures 4 and 6 indicate that most of the asymmetry with respect to field strength as shown in figure 3 is actually the false asymmetry due to positrons originating from the beam stop. One would expect that there would be no asymmetry in figure 4 at all due to the fact that the magnetic field should be completely symmetric about the center of the tube where the diamond sample is placed, however the muons might only interact when they hit the surface of the diamond. Since, in the simulation, the diamond was centered at the center of the tube, the surface where the muons would interact is slightly offset from the center and this is probably the reason for the small asymmetry as shown in figure 4.

The way the number of muon hits increase with field strength and diamond size for the columnated muon beam in figure 7 is exactly what is expected since, qualitatively at least, every cross section of the 3D graph in figure 7 along the y-axis has the same shape as figure 2. It appears as though the application correctly simulates the field dependence of the e and μ trajectories, the field dependence of the flypast, and the detection efficiency and hence the field dependence of the false asymetry [6]. Therefore, in conclusion, it's indeed possible that the simulation described here can be used for advanced analysis of the experimental data produced by those experiments. However, more still needs to be done with regards to the simulation before such advanced analysis can be performed.

In future work, the following goals will be persued:

1. Determine the field dependence of the detector efficiency for positrons originating from the sample.
2. Determine the field dependent distribution of impact positions of flypast muons.
3. Determine detector efficiency for positrons originating from flypast muons.

References

[1] Connell S H, Bharuth-Ram K, Cox S F J and Keartland J M 2011 μsr in diamond *Hyperfine Interactions*, To be published

[2] Allison J, Amako K, Apostolakis J, Araujo H, Dubois P A, Asai M, Barrand G, Capra R, Chauvie S, Chytracek R, Cirrone G A P, Cooperman G, Cosmo G, Cuttone G, Daquino G G, Donszelmann M, Dressel M, Folger G, Foppiano F, Generowicz J, Grichine V, Guatelli S, Gumplinger P, Heikkinen A, Hrivnacova I, Howard A, Incerti S, Ivanchenko V, Johnson T, Jones F, Koi T, Kokoulin R, Kossov M, Kurashige H, Lara V, Larsson S, Lei F, Link O, Longo F, Maire M, Mantero A, Mascialino B, McLaren I, Lorenzo P M, Minamimoto K, Murakami K, Nieminen P, Pandola L, Parlati S, Peralta L, Perl J, Pfeiffer A, Pia M G, Ribon A, Rodrigues P, Russo G, Sadilov S, Santin G, Sasaki T, Smith D, Starkov N, Tanaka S, Tcherniaev E, Tome B, Trindade A, Truscott P, Urban L, Verderi M, Walkden A, Wellisch J P, Williams D C, Wright D and Yoshida H 2006 *IEEE Transactions on Nuclear Science* **53** 270–278 ISSN 0018-9499 URL `http://dx.doi.org/10.1109/TNS.2006.869826`

[3] Agostinelli S, Allison J, Amako K, Apostolakis J, Araujo H, Arce P, Asai M, Axen D, Banerjee S and Barrand G 2003 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** 250–303

[4] James F 1990 *Computer Physics Communications* **60** 329–344 ISSN 00104655 URL `http://dx.doi.org/10.1016/0010-4655(90)90032-V`

[5] Madhuku M, Gxawu D, Connell S H, Machi I Z, Keartland J M, Cox S F J and King P J C 2007 *Hyp. Int.* **177** 27–31

[6] Madhuku M, DGxawu, IZMachi, SHConnell, JMKeartland, SFJCox and PJCKing 2009 *Physica B* **404** 804807

[7] Brun R 1997 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **389** 81–86 ISSN 01689002 URL `http://dx.doi.org/10.1016/S0168-9002(97)00048-X`

The physics of exceptional points

W.D. Heiss

National Institute for Theoretical Physics, Stellenbosch Institute for Advanced Study

E-mail: dieter@physics.sun.ac.za

Abstract. A short resume is given about the nature of exceptional points (EPs) followed by a discussion about their ubiquitous occurrence in a great variety of physical problems. EPs feature in quantum phase transition, quantum chaos, they produce dramatic effects in multichannel scattering, specific time dependence and more. In nuclear physics they are associated with instabilities and affect approximation schemes. EPs could be of interest for weakly bound states such as halos and nuclei along the drip line.

1. Introduction

The topic *exceptional points* (EPs), in particular the physical manifestations of their occurrence in eigenvalue problems, has found increasing interest in the ongoing literature. EPs are spectral singularities that occur generically in eigenvalue problems depending on a parameter [1]. This implies classical as well as quantum mechanical cases. In the simplest case they are studied in two-dimensional matrix problems [2, 3]. They are of physical interest as there is a great variety of physical situations where the singularities explain particular, in some cases dramatic effects [1]. Below we briefly present the formal background followed by a description of the first physical manifestation of the mathematical properties. The subsequent sections are devoted to some of the major physical cases where EPs play a direct role in the understanding of specific phenomena.

2. Exceptional points

For the eigenvalues of a two-dimensional matrix the direct connection of an EP and the phenomenon of level repulsion is easily demonstrated. Consider the problem

$$\begin{aligned} H(\lambda) &= H_0 + H_1(\lambda) = H_0 + \lambda V \\ &= \begin{pmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{pmatrix} + \lambda \begin{pmatrix} \epsilon_1 & \delta \\ \delta & \epsilon_2 \end{pmatrix} \end{aligned} \tag{1}$$

where the parameters ω_k and ϵ_k determine the non-interacting resonance energies $E_k = \omega_k + \lambda\epsilon_k$, $k = 1, 2$. Owing to the interaction invoked by the matrix element δ the two levels do not cross but repel each other. In fact, the levels turn out to be

[1] see also the workshop at Stellenbosch in November 2010: **http://www.nithep.ac.za/2g6.htm**

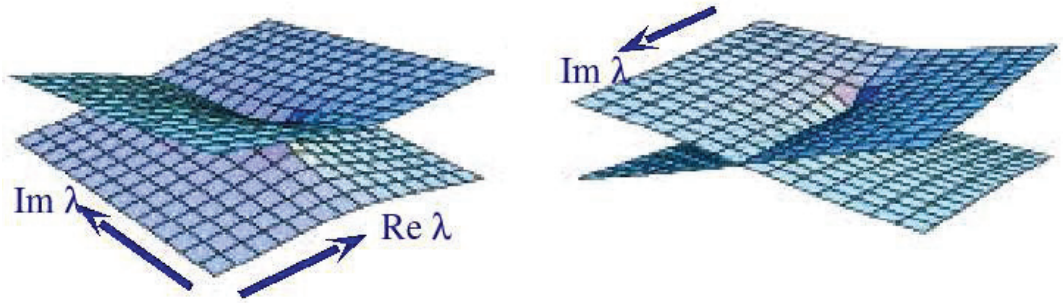


Figure 1. Perspective view of the Riemann sheet structure of two coalescing energy levels in the complex $\lambda-$plane

$$E_{1,2}(\lambda) = \frac{1}{2}(\omega_1+\omega_2+\lambda(\epsilon_1+\epsilon_2)\mp D) \quad (2)$$

$$D = \sqrt{CC(\lambda-EP1)(\lambda-EP2)} \quad (3)$$

$$CC = 4\delta^2+(\epsilon_1-\epsilon_2)^2 \quad (4)$$

and the two levels *coalesce* for complex values of λ in the vicinity of the level repulsion, that is at

$$EP1 = \frac{i(\omega_1-\omega_2)}{-2\delta-i(\epsilon_1-\epsilon_2)} \quad (5)$$

$$EP2 = \frac{i(\omega_1-\omega_2)}{+2\delta-i(\epsilon_1-\epsilon_2)}. \quad (6)$$

We use the term *coalesce* as the pattern is distinctly different from a usual degeneracy encountered for hermitian operators. Note that $H(\lambda)$ is not hermitian for complex values of λ. An approach in the laboratory of an EP therefore requires an open system. The difference between the hermitian and the non-hermitian case is clearly manifested by the occurrence of only *one* eigenvector (instead of the two in the case of a genuine degeneracy). The only one eigenvector is here given by

$$|\phi_{EP1}\rangle = \begin{pmatrix}1\\ i\end{pmatrix} \quad \text{at the EP1, and} \quad (7)$$

$$|\phi_{EP2}\rangle = \begin{pmatrix}1\\ -i\end{pmatrix} \quad \text{at the EP2} \quad (8)$$

independent of parameters. Note that the norm - that is the scalar product $\langle\tilde{\phi}_{EPk}|\phi_{EPk}\rangle, k=1,2$ - vanishes. It is the square root type of behaviour of the eigenvalues - implying an infinite derivative in the variable λ - and the vanishing norm of the likewise coalescing eigenfunctions that invoke specific observable effects.

3. Observable effects

Many cases of specific effects have been reported in the literature during the past ten years. We here can discuss only a few in some detail.

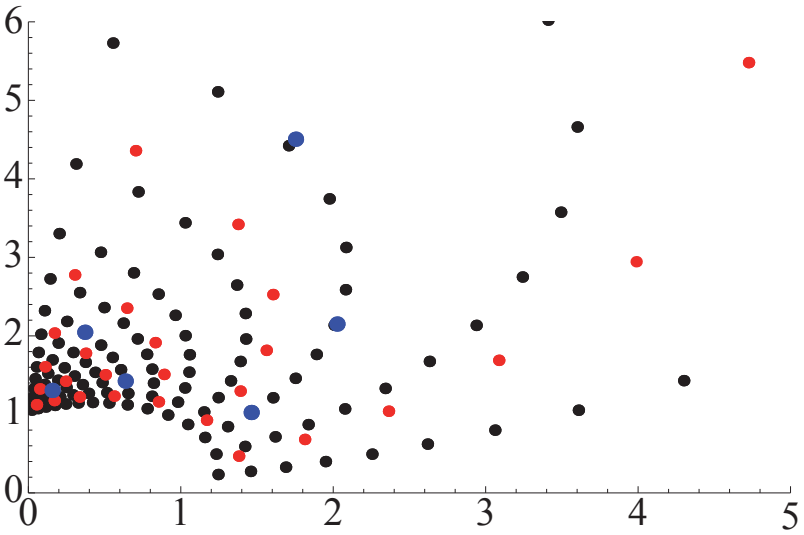


Figure 2. Exceptional points in the complex λ-plane for the Lipkin model with $N = 8$ (blue), $N = 16$ (red) and $N = 32$ (black).

3.1. Microwave cavity

Probably the first time ever the direct encircling of the square root branch point - that is the manifestation of the two Riemann sheets (see Fig.1) - was accomplished with a microwave resonator [4].

The realisation of the complex parameter λ was implemented in the laboratory by two real parameters: (i) the coupling between the two halves of the cavity and (ii) the variation of the one level in one half of the cavity. In the experiment the direct approach of the EP was avoided while the encircling was done at close distance. One encirclement clearly swapped the levels and so did the corresponding wave functions that were measured as well. Moreover, one of the wave function picks up a phase, i.e. after one round one obtains $|\phi_1\rangle \to -|\phi_2\rangle$ and $|\phi_2\rangle \to |\phi_1\rangle$. As a consequence, it needs four rounds for the wave functions to recover the original configuration, in other words one obtains the pattern for subsequent encircling

$$\begin{pmatrix}\phi_1\\ \phi_2\end{pmatrix} \to \begin{pmatrix}-\phi_2\\ \phi_1\end{pmatrix} \to \begin{pmatrix}-\phi_1\\ -\phi_2\end{pmatrix} \to \begin{pmatrix}\phi_2\\ -\phi_1\end{pmatrix} \to \begin{pmatrix}\phi_1\\ \phi_1\end{pmatrix}.$$

This sequence has been predicted and was established experimentally. It confirms a fourth order root for the normalised wave functions (recall: the norm vanishes at the EP, the leading order is $\sim \sqrt{\lambda - \lambda_{EP}}$). Note that the sequence has a chiral property: going clockwise yields a result different from the one going counterclockwise.

This chiral property of the wave function at the EP has been confirmed in a second experiment [5] where the phase difference of $\pi/2$ between the two components (see (7) or (8)) has been confirmed in a direct approach of an EP. For further details see [4, 5].

The same results have been reconfirmed with two coupled electronic circuits [6].

3.2. Quantum phase transitions, chaos

The Lipkin model [7] is a toy model often used to study quantum phase transitions. The interaction of the two level model lifts or lowers a Fermion pair between the two levels. For N particles it can be formulated in terms of the angular momentum operators and reads

$$H(\lambda) = J_z + \frac{\lambda}{N}(J_+^2 + J_-^2) \tag{9}$$

with $J_z, J_\pm$ being the N-dimensional representations of the $SU(2)$ operators. There is a phase transition at $\lambda > 1$ that moves toward $\lambda = 1$ in the thermodynamic limit. The Hamiltonian has an inherent symmetry: even and odd numbers k of the levels E_k do not interact. The phase for $\lambda < 1$ is the 'normal' phase where the symmetry of the problem is preserved by the levels and wave functions. In the 'deformed' phase for $\lambda > 1$ the symmetry is broken in that even and odd k become degenerate. Here the role of the EPs is crucial to bring about the phase change in the spectrum [8]. In Fig.2 the pattern of the EPs is illustrated for low values of N. It is clearly seen how the EPs accumulate for increasing N on the real axis with the tendency to move towards the point $\lambda = 1$. The spectrum remains unaffected by singularities in the region of the normal phase while it is strongly affected around the critical point. For finite temperature these singularities feature in the partition function as is discussed explicitly in [9]. If the model is perturbed the regular pattern of the EPs is destroyed and so is the spectrum accordingly. The onset of chaos [10] is clearly discernible in the region of the phase transition while the model remains robust outside the critical region for sufficiently mild perturbation.

3.3. The role of EPs in approximation schemes

The well known Random Phase Approximation (RPA) used in many body problems yields an effective Hamiltonian that is non-hermitian [11]. As a result, eigenvalues are not necessarily real. Depending on the strength of the, say, particle-hole interaction two real eigenvalues Ω and $-\Omega$ coalesce at $\Omega = 0$ and then move into the complex plane when the interaction is increased. Often this instability point is associated with one or more phase transition of the underlying mean field [12]. It is an EP with all its characteristics: square root branch point in the interaction strength and the vanishing norm of the wave function.

A perturbative approach in shell model calculations can be hampered by singularities associated with *intruder states* [13]. These singularities are EPs where two levels coalesce thus limiting the radius of convergence of the perturbation series.

Recent approaches to model nuclei on the drip line [14] use resonance states to describe the continuum. The coalescence of two resonances can invoke specific physical effects owing to the strong increase of the associated spectroscopic factors being caused by the vanishing norm of the wave functions at the EP.

3.4. The symmetry breaking point for $\mathcal{PT}$-symmetric Hamiltonians

It has been suggested to extent the class of the traditional hermitian Hamiltonians by a specific choice of non-hermitian operators [15]. Hamiltonians that are symmetric under the combined operation of parity and time reversal transformation, i.e. $\mathcal{PT}$-symmetric operators, can have a real spectrum even though the operators can be non-hermitian. It turns out that if the eigenstates preserve the symmetry, the eigenvalues are real, while for symmetry breaking the eigenvalues are complex [16]. The points where this symmetry is broken are the EPs of the problem. In the meantime, while there is plenty of theoretical literature on this subject - specific [17] and more general [18] -, there is beautiful experimental evidence with optical cavities [19], optical lattices [20] and propagation of light [21].

3.5. EPs and Feshbach resonance in atomic/molecular physics

Using Feshbach resonance techniques there are recent proposals for resonant dissociation by lasers of H_2^+ molecules or alkali dimers, where the effects of EPs are expected to feature prominently [22]. Similar in spirit, a Bose-Einstein condensate of neutral atoms with induced electromagnetic attractive $(1/r)$ interaction has been discussed recently as another system allowing a tunable interaction [23]. The critical value - an EP - where the onset of the collapse of the condensate occurs is interpreted as a transition point from separate atoms to the formation of molecules or clusters [24].

3.6. Special effects in multichannel scattering

Depending on a judicious choice of parameters the proximity of EPs can invoke dramatic effects in multichannel scattering such as a sudden increase of the cross section in one channel, even by orders of magnitude. In turn, a second channel is suppressed and can show a resonance curve that deviates substantially from the usual Lorentz shape [25]. Related to this behaviour is the pattern in the time domain [26]. Depending on the initial conditions the wave function displays characteristic features such as very fast decay or the opposite, i.e. very long life time. At the EP the wave function typically has a linear term in time besides the usual exponential behaviour.

4. Summary

The ubiquitous occurrence of EPs in all eigenvalue problems that depend on a parameter can have significant and often dramatic effects on observables in a great variety of physical phenomena. A few decades ago, these singularities appeared as a purely mathematical feature that could cause problems in approximation schemes. It was only about ten years ago that their physical manifestation has been demonstrated in experiments that were basically classical in nature (recall that an EP can be approached in the laboratory only in an open system). At present definite theoretical and experimental proposals are found in the literature relating to atomic and molecular physics, using lasers for triggering and measuring specific transitions. In nuclear physics, where there is now great interest in open systems, that is in nuclei on the drip line, the coalescence of resonance states is expected to produce specific effects such as enhancements of particular reactions.

References

[1] T. Kato, *Perturbation theory of linear operators* Springer, Berlin 1966
[2] W.D. Heiss 1999 *Eur.Phys.J.* D **7** 1
[3] W.D. Heiss 2000 *Phys.Rev.* E **61** 929
[4] C. Dembowski, H.-D. Gräf, H.L.Harney, A.Heine, W.D. Heiss, H. Rehfeld and A.Richter 2001, *Phys.Rev.Lett.* **86**, 787
[5] C. Dembowski, B. Dietz, H.-D. Gräf, H. L. Harney, A. Heine, W.D. Heiss and A. Richter 2003 *Phys.Rev.Lett.* **90**, 034101; 2004 *Phys.Rev.* E **69**, 056216
[6] T. Stehmann, W.D. Heiss and F.G. Scholtz 2004 *J.Phys.* A: Mathematical & General, **37**, 7813
[7] H.J. Lipkin, N. Meshkov and N. Glick 1965, *Nucl.Phys.* A **62**, 188 ; *ibid.* **62**, 199 (1965), *ibid.* **62**, 211 (1965)
[8] W.D. Heiss, F.G. Scholtz and H.B. Geyer 2005 *J.Phys.* A: Mathematical & General, **38**, 1843
[9] P. Cejnar, S. Heinze, M. Macek 2007 *Phys.Rev.Lett.* **99**, 100601
[10] W.D. Heiss and A.L. Sannino 1991 *Phys.Rev.* A **43**, 4159
[11] P. Ring and P. Schuck 1980 *The Nuclear Many Body Problem*, Springer, New York
[12] W.D. Heiss and R.G. Nazmitdinov 2002 *Phys.Rev.* C **65**, 054304
[13] T.H. Schucan and H.A. Weidenmüller 1972 *Ann.Phys.(NY)* **73** 108
[14] N. Michel, W. Nazarewicz, J. Okolowicz, and M. Ploszajczak 2010 *J.Phys.* G: Nucl.Part.Phys. **37**, 064042
[15] C.M. Bender and S. Boettcher 1998 *Phys.Rev.Lett.* **80**, 5243
[16] P. Dorey, C. Dunning and R. Tateo 2001 *J.Phys.* A: Math. Gen. **34** L391
[17] M. Znojil 2006 *J.Phys.* A: Math. Gen. **39**, 441
[18] U. Günter, I. Rotter, and B.F. Samsonov 2007 *J. Phys.* A **40**, 8815, and references therein.
[19] Sang-Bum Lee *etal.* 2009 *Phys.Rev.Lett* **103**, 134101
[20] K.G. Makris, R. El-Ganainy, D.N. Christodoulides and Z.H. Musslimani 2010 *Phys.Rev.* A **81**, 63807
[21] S. Longhi 2010 *Phys.Rev.Lett.* **105**, 13903
[22] R. Lefebvre and O. Atabek 2010 *Eur.Phys.J.* D **56** 317; O. Atabek, R. Lefebvre, M. Lepers, A. Jaouadi, O. Dulieu and V. Kokoouline 2011 *Phy.Rev.Lett.* **106** 173002
[23] I. Papadopoulos, P. Wagner, G. Wunner, and J. Main 2007 *Phys.Rev.* A **76**, 053604
[24] H. Cartarius, J. Main, and G. Wunner 2008 *Phys.Rev.* A **77**, 013618
[25] W.D. Heiss and R. G. Nazmitdinov 2010 *Eur.Phys.J.* D**58**, 53
[26] W.D. Heiss 2010 *Eur.Phys.J.* D**60**, 257

Analytic density functionals

S. Karataglidis

Department of Physics, University of Johannesburg, P. O. Box 524, Auckland Park, 2006, South Africa

E-mail: stevenka@uj.ac.za

B. G. Giraud

Institut de Physique Théorique, Centre d'Etudes Saclay, 91190 Gif-sur-Yvette, France

E-mail: bertrand.giraud@cea.fr

Abstract. A systematic strategy for the calculation of density functionals (DFs) consists in coding informations about the density and the energy into polynomials of the degrees of freedom of wave functions. DFs and Kohn-Sham potentials (KSPs) are then obtained by standard elimination procedures of such degrees of freedom between the polynomials. Numerical examples illustrate the formalism.

1. Introduction

Existence theorems [1] for DFs do not provide directly constructive algorithms. Fortunately, the Kohn-Sham (KS) method [2] spares the construction of a "kinetic functional" and reduces energy and density calculations to the tuning of a local potential, $v_{KS}(r)$. Hence, a considerable amount of work has been dedicated to detailed estimates of electronic correlation energies and the corresponding KSPs, see for instance [3, 4, 5]. Many authors were also concerned with representability and stability questions, see for instance [6] and, for calculations in subspaces, see [7] and [8]. For cases where the mapping between potential and density shows singularities, see [9]. For reviews of the rich multiplicity of derivations of DFs and KS solutions and their properties, we refer to [10] and [11], and, for nuclear physics, to [12].

Local or quasi local approximations use the continuous infinity of values $\rho(\mathbf{r}), \forall \mathbf{r}$, as the parameters of the problem. However, whether for atoms, molecules or nuclei, a finite number of parameters is enough to describe physical situations. For instance, Woods-Saxon nuclear profiles notoriously make good approximations, depending only on a handful of parameters, and it is easy to add a few parameters describing, for example, long tails and/or moderate oscillations of the density. (High frequency oscillations are unlikely, for they might cost large excitation energies.) We can stress here, in particular, the one-dimensional nature of the radial density functional (RDF) theory [13], valid for nuclei and/or atoms, isolated, described by rotationally-invariant Hamiltonians; the constrained density minimization of energy [14] returns isotropic densities, with radial profiles, $\rho(r), 0 \leqslant r < \infty$. The number of parameters to describe a nuclear density, therefore, can be restricted to maybe ~ 10 at most; situations with ~ 20 parameters are a luxury. For molecules, shapes are much more numerous, but a finite, while large number of parameters, truncating a list of multipoles for instance, still makes a reasonable frame. Practical DFs,

therefore, can boil down to *functions* of a finite number of parameters. Functional variations can then be replaced by simple derivatives.

Herein, we show how information about both the density and the energy can be recast into polynomials. This allows eliminations of part of the parameters. Further polynomial manipulations locate energy extrema. Only density parameters are left. The same method gives KSPs. Finally we offer a discussion and conclusion.

2. Formalism for analytic density functional

Consider a basis of n orthonormalized, single-particle states, $\varphi_\alpha(\mathbf{r}\sigma\tau)$, where spin and isospin labels $\sigma\tau$ will be understood. The orthonormalized Slater determinants ϕ_i made out of the φ_α's for N fermions make a finite subspace, of some dimension $\mathcal{N}$, in which eigenstates of the physical Hamiltonian H can be approximated by configuration mixings, $\Psi = \sum_{i=1}^{\mathcal{N}}(C_i + iC'_i)\phi_i$. Here C_i and C'_i are the real and imaginary parts, respectively, of the mixing coefficients, but, in practice, with real matrix elements, $H_{ij} = \langle\phi_i|H|\phi_j\rangle$, of the Hamiltonian H, the imaginary parts C'_i vanish. Both the energy η and the normalization are *quadratic* functions of such coefficients,

$$\eta = \sum_{i,j=1}^{\mathcal{N}} C_i H_{ij} C_j, \;\; \sum_{i=1}^{\mathcal{N}} C_i^2 = 1. \tag{1}$$

Let $a^\dagger_{\mathbf{r}}$ and $a_{\mathbf{r}}$ be the usual creation and annihilation operators at position $\mathbf{r}$. Tabulate the matrix elements $\left\langle \phi_j \left| a^\dagger_{\mathbf{r}} a_{\mathbf{r}} \right| \phi_j \right\rangle$. The density corresponding to Ψ is, again, *quadratic* with respect to the C_i's,

$$\rho(\mathbf{r}) = \sum_{ij} C_i \left\langle \phi_j \left| a^\dagger_{\mathbf{r}} a_{\mathbf{r}} \right| \phi_j \right\rangle C_j, \tag{2}$$

and any parameter that is linear with respect to moments of the density is also a quadratic function of the C_i's.

Let $\{S_\nu(\mathbf{r})\}$, $\nu = 1, \ldots, \infty$, be a complete orthonormal set of "vanishing average" functions. Namely, the two sets of conditions, $\int d\mathbf{r}\, S_\nu(\mathbf{r}) = 0, \forall\nu$, and, $\int d\mathbf{r} S_\mu(\mathbf{r}) S_\nu(\mathbf{r}) = \delta_{\mu\nu}, \forall\mu\nu$, are satisfied. Such sets are easy to find; in the case of one-dimensional problems, including radial ones, they can be implemented by means of orthogonal polynomials [15, 16] and a generalization to more dimensions is easy. Then subtract from ρ some reference density, ρ_0, obtained by some approximation relevant for the N fermions. The difference, $\Delta\rho = \rho - \rho_0$, is of a vanishing average, since, by definition, both ρ and ρ_0 integrate out to N. Then the Fourier coefficients,

$$\Delta_\nu = \int d\mathbf{r}\; S_\nu(\mathbf{r}) \Delta\rho(\mathbf{r}), \tag{3}$$

define ρ, as $\rho = \rho_0 + \sum_{n=1}^{\infty} \Delta_\nu S_\nu$. As already stated, this expansion of ρ can be truncated. at some realistic order $\mathcal{N}'$, lower than the number of independent parameters C_i. The Δ_ν's are *quadratic* in the C_i's,

$$\Delta_\nu = \sum_{ij} C_i \left[\int d\mathbf{r}\; S_\nu(\mathbf{r}) \left\langle \phi_j \left| a^\dagger_{\mathbf{r}} a_{\mathbf{r}} \right| \phi_j \right\rangle \right] C_j - \rho_{0\nu}. \tag{4}$$

Note the auxiliary numbers, $\rho_{0\nu} = \int d\mathbf{r}\, S_\nu(\mathbf{r})\, \rho_0(\mathbf{r})$.

It is then trivial to use the $\mathcal{N}'$ density constraints, Eqs. (4), and the normalization in Eqs. (1), to eliminate, for instance, the last $(\mathcal{N}' + 1)$ coefficients C_i. This leaves a polynomial relation, $\mathcal{R}(\eta, \Delta_1, \ldots, \Delta_{\mathcal{N}'}, C_1, \ldots, C_{\mathcal{N}-\mathcal{N}'-1}) = 0$, between the energy, the density parameters, and the remaining coefficients C_i. Finally, the energy must be minimized with respect to such remaining

coefficients, via still polynomial conditions, $\partial\mathcal{R}/\partial C_i = 0,\ i = 1, \ldots, \mathcal{N} - \mathcal{N}' - 1$. This gives a polynomial relation, $\mathcal{E}(\eta, \Delta_1, \ldots, \Delta_{\mathcal{N}'}) = 0$, between the energy and the density parameters. This polynomial $\mathcal{E}$ is our "algebraic" or analytic DF. It accounts for all contributions to the energy, both without and with correlations, for only matrix elements of the full H are used.

The procedure can be further simplified in the following way. Let $\mathcal{H}$ be the matrix representing the Hamiltonian on an orthonormal basis for a suitable subspace of wave functions, and, similarly, let, for instance, $\mathcal{D}_1, \mathcal{D}_2$ be the matrices representing two constraints selected to parametrize the density, such as, for instance, two among the parameters $(\Delta_\nu + \rho_{0\nu})$. Set the equation, polynomial in all three variables $\varepsilon, \lambda_1, \lambda_2$,

$$P\,(\varepsilon, \lambda_1, \lambda_2) \equiv \det\,(\mathcal{H} - \lambda_1\mathcal{D}_1 - \lambda_2\mathcal{D}_2 - \varepsilon) = 0. \tag{5}$$

Here ε is the free energy, lowest eigenvalue of $(\mathcal{H} - \lambda_1\mathcal{D}_1 - \lambda_2\mathcal{D}_2)$, and the λ's are Lagrange multipliers. It is well known that $\partial\varepsilon/\partial\lambda_i = -D_i,\ i = 1, 2$, where $D_i \equiv \langle\mathcal{D}_i\rangle$ is the expectation value of the corresponding constraint. From Eq. (5) such partial derivatives read, $\partial\varepsilon/\partial\lambda_i = -(\partial P/\partial\lambda_i)/(\partial P/\partial\varepsilon),\ i = 1, 2$, hence two more polynomial relations are obtained,

$$Q_i(D_i, \varepsilon, \lambda_1, \lambda_2) \equiv (\partial P/\partial\varepsilon)\, D_i - (\partial P/\partial\lambda_i) = 0. \tag{6}$$

Replace in Eqs. (5,6) the free energy by its value, $\varepsilon = \eta - \lambda_1 D_1 - \lambda_2 D_2$, in terms of the energy, $\eta \equiv \langle\mathcal{H}\rangle$ and the constraints, D_1, D_2. This creates three polynomials in terms of $\eta, D_1, D_2, \lambda_1, \lambda_2$, out of which λ_1, λ_2 can be eliminated, for a final polynomial equation, $\mathcal{E}(\eta, D_1, D_2) = 0$. This easy Legendre transform generates our "algebraic DF". A generalization to any number of quadratic constraints is trivial. Such algebraic DFs are not open formulae of the form, $\eta = F(D_1, \ldots, D_{\mathcal{N}'})$, but they provide roots for η at any realistic degree of numerical accuracy. Incidentally, they may also give excited energies and/or spurious ones, a well known property [17] of DFs.

3. Toy model

For an illustrative toy model, we consider two fermions only and set the one-body part of H as, $K = -d^2/(2dr_1^2) - d^2/(2dr_2^2) + (r_1^2 + r_2^2)/2$, the sum of two harmonic oscillators, and its two-body part as a translation invariant, separable potential, defined in coordinate representation by,

$$\left\langle r_1 r_2 \left| V \right| r_1' r_2' \right\rangle = -V_0\delta\left[(r_1 + r_2 - r_1' - r_2')/2\right] e^{-[(r_2 - r_1)^2 + (r_2' - r_1')^2]/4}(r_2 - r_1)(r_2' - r_1')/\sqrt{2\pi}. \tag{7}$$

Then, given the first 4 wave functions, $\varphi_0, \ldots, \varphi_3$, of the one-dimensional harmonic oscillator, we create, to prepare a configuration mixing, a basis of 4 negative parity Slater determinants. These read, in a transparent notation, $\{\varphi_0, \varphi_1\}, \{\varphi_0, \varphi_3\}, \{\varphi_2, \varphi_1\}, \{\varphi_2, \varphi_3\}$. We set $V_0 = 3$ for a numerical test. To constrain H, we choose the second moment operator, $r_1^2 + r_2^2$. The matrices representing H and the constraint in the toy subspace read,

$$\mathcal{H} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 7/4 & 3\sqrt{3}/4 & 0 \\ 0 & 3\sqrt{3}/4 & 13/4 & 0 \\ 0 & 0 & 0 & 45/8 \end{bmatrix} \text{ and } \mathcal{D} = \begin{bmatrix} 2 & \sqrt{3/2} & \sqrt{1/2} & 0 \\ \sqrt{3/2} & 4 & 0 & \sqrt{1/2} \\ \sqrt{1/2} & 0 & 4 & \sqrt{3/2} \\ 0 & \sqrt{1/2} & \sqrt{3/2} & 6 \end{bmatrix}. \tag{8}$$

The equations which correspond to Eqs. (5),(6) read,

$$\begin{aligned} P_{\rm toy}(\varepsilon,\lambda) = -360 + 154\varepsilon + 344\varepsilon^2 - 154\varepsilon^3 + 16\varepsilon^4 + 1464\lambda \\ + 1692\varepsilon\lambda - 1636\varepsilon^2\lambda + 256\varepsilon^3\lambda + 725\lambda^2 - 5140\varepsilon\lambda^2 + \\ 1408\varepsilon^2\lambda^2 - 4192\lambda^3 + 3072\varepsilon\lambda^3 + 2064\lambda^4 = 0, \\ Q_{\rm toy}(D,\varepsilon,\lambda) = -1464 - 1692\varepsilon + 1636\varepsilon^2 - 256\varepsilon^3 - 1450\lambda + \\ 10280\varepsilon\lambda - 2816\varepsilon^2\lambda + 12576\lambda^2 - 9216\varepsilon\lambda^2 - 8256\lambda^3 + \\ (154 + 688\varepsilon - 462\varepsilon^2 + 64\varepsilon^3 + 1692\lambda - 3272\varepsilon\lambda + \\ 768\varepsilon^2\lambda - 5140\lambda^2 + 2816\varepsilon\lambda^2 + 3072\lambda^3)D = 0. \quad (9) \end{aligned}$$

Finally, the substitution, $\varepsilon = \eta - \lambda D$, followed by the elimination of λ, generates the desired polynomial equation, $\mathcal{E}_{\rm toy}(\eta, D) = 0$.

We show in Fig. 1 the contour line, $\mathcal{E}_{\rm toy}(\eta, D) = 0$. The ground state is found at the lowest point of the oval envelope, with coordinates, $D = 2, \eta = -1$. The highest and lowest eigenvalues of $\mathcal{H}$ are, 45/8 and -1, and those of $\mathcal{D}$ are, $4 \pm \sqrt{4+\sqrt{15}}$, namely ~ 6.81 and ~ 1.19. This is confirmed by the extremal points, up, down, right and left, of the oval. The inside pattern refers to excited states. The concavity of the lowest part of the envelope and convexity of its highest part are transparent properties of the theory. They generalize for any dimension of the subspace and any number of constraints; we tested this generalization with further toy models. Moreover, when, via embedded subspaces, the dimension $\mathcal{N}$ of the matrices, $\mathcal{H}, \mathcal{D}_i$, grows while H and the constraints are kept the same, a growth of the envelope is found and the bottom of the envelope converges towards a limit, as expected. This gives numerical estimates for an extrapolation of this concave part towards its limit for $\mathcal{N} \to \infty$.

Such concavities should also occur in DF theories with a continuous infinity of constraints. But they are often difficult to verify, and are, therefore, overlooked, although they are an important test of soundness.

A byproduct of the procedure consists of a polynomial relating the potential energy to the constraints. Set the Hamiltonian as, $H = h + V$, with $V = -V_0\mathcal{V}$, where V_0 is an interaction strength and $\mathcal{V}$ gives all details of interaction shapes. Nothing prevents one from considering V_0 as a Lagrange multiplier and obtain, via the polynomial method pushed one step further, a polynomial, $\mathcal{F}(\langle h\rangle, \langle\mathcal{V}\rangle, D_1, \ldots, D_{\mathcal{N}'})$, linking $\langle h\rangle$ to the expectation values of $\mathcal{V}$ and the constraints. A standard result of this Legendre transform is, $\partial\langle h\rangle/\partial\langle\mathcal{V}\rangle = V_0$, *i.e.*,

$$\mathcal{G}(V_0, \langle h\rangle, \langle\mathcal{V}\rangle, D_1, \ldots, D_{\mathcal{N}'}) \equiv (\partial\mathcal{F}/\partial\langle h\rangle)\, V_0 - \partial\mathcal{F}/\partial\langle\mathcal{V}\rangle = 0. \quad (10)$$

Replace, in $\mathcal{F}$ and $\mathcal{G}$, the quantity $\langle h\rangle$ by $\eta + \langle\mathcal{V}\rangle V_0$. Then eliminate η and V_0 between $\mathcal{E}$ and such modified $\mathcal{F}$ and $\mathcal{G}$. This links $\langle\mathcal{V}\rangle$, hence $\langle V\rangle$, to the D_i. It must be stressed here that now $\langle\mathcal{V}\rangle$ should *not* be minimized with respect to the D_i; rather, those D_i values to be used are those that minimize the total energy η.

A similar argument provides the kinetic energy, or any other part of η, in the same context of total energy constrained minimization. Such results are of interest for a detailed analysis of corrections induced by correlations.

An issue which will arise in all future models using this polynomial method is that the final minimization of η must be performed within a convex domain of densities: what conditions selected parameters (moments, local values, etc), satisfy to maintain ρ positive? This question was recently [18] solved by means of the Sturm criterion, for a general class of positive functions having positive Fourier transforms. The criterion gives the number of real roots of a polynomial, and can be used to ensure that a polynomial has no real roots. As seen in the toy model, the detailed structure of the calculation can be a guide to define the physically acceptable domain

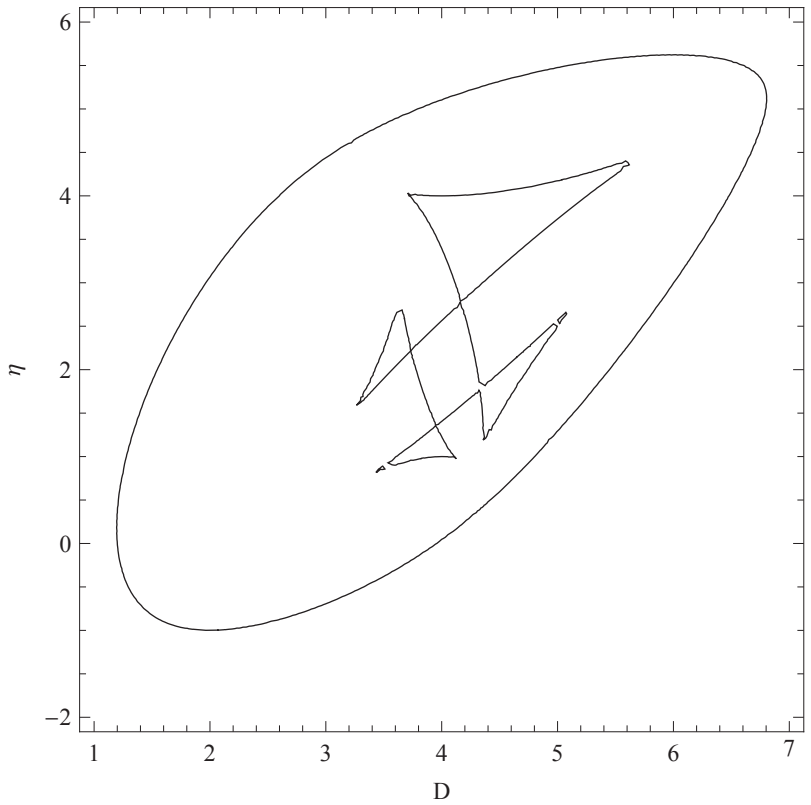


Figure 1. Contour $\mathcal{E}_{\rm toy}(\eta, D) = 0$ for the configuration mixing model with 4×4 matrices, as described in the text.

of parameters. For more subtle questions about the topology of acceptable functional spaces of densities and trial functions, we refer to [19], but will state, without proof, that here with traditional functions (harmonic, Coulomb) and their configuration mixing, the positivity of ρ should be sufficient.

There is also the question of spurious solutions. For more complicated systems, spurious solutions [7, 8] might certainly pop up, but an analysis for their detection remains easy. In particular, for other toy models that we tested, spurious solutions were found to induce values of physical parameters out of their allowed range, and/or even complex values while only real ones are acceptable. We can insist that the final, polynomial equation for the energy, $\mathcal{S}(\eta) = 0$, can only create a finite number of candidate solution branches to be investigated.

A constructive derivation of KSPs is available. For instance, truncate some single particle basis and let $\mathcal{P}$ be the projector upon the resulting, finite dimensional subspace for a system of N fermions, with their Hamiltonian H, or rather now, $\mathcal{P}H\mathcal{P}$. Given the kinetic energy operator T, choose a local potential $w_0(r)$, hence a one-body operator $W_0 = \sum_{i=1}^{N} w_0(r_i)$, hence a one-body Hamiltonian $H_0 = T + W_0$, so that the ground state of $\mathcal{P}H_0\mathcal{P}$, a Slater determinant Φ_0, be non degenerate and providing an approximate density ρ_0 for the system. For any density ρ in the subspace, the integral, $\int \Delta\rho$, of the difference, $\Delta\rho = \rho - \rho_0$, vanishes as already stated. (Here and in the following, the integral sign, $\int$, means $\int r^{d-2}\, dr$ depending on the d-dimensional problem under consideration.) Expand, as already discussed, $\Delta\rho$ in a basis of orthonormal functions $S_\beta(r)$, "constrained by vanishing averages" [15, 16], $\Delta\rho(r) = \sum_{\beta=1}^{\infty} b_\beta\, S_\beta(r)$. Truncate the expansion at some suitable order $\mathcal{N}'$. Again, given a determinant Φ with the parameters $c^\alpha_{n\ell m}$ of its orbitals, or given a correlated state, $\Psi = \sum_q C_q\, \Phi_q$, the constraints, $\Phi \Rightarrow b_\beta$ or $\Psi \Rightarrow b_\beta$,

are polynomials of the parameters. Given H_0, the polynomial method returns a polynomial $\mathcal{K}(\kappa, b_1, \ldots, b_{\mathcal{N}'})$ for a reference functional, such that the lowest root of the equation, $\mathcal{K} = 0$, represents the constrained minimum, $\kappa' = \mathrm{Min}_{\Phi \Rightarrow b_1,\ldots,b_{\mathcal{N}'}} \langle \Phi | H_0 | \Phi \rangle$, for the determinants in the subspace. In the same way, given the full H, the method gives a polynomial $\mathcal{E}(\eta, b_1, \ldots, b_{\mathcal{N}'})$, the lowest η root of which is the constrained minimum, $\eta' = \mathrm{Min}_{\Psi \Rightarrow b_1,\ldots,b_{\mathcal{N}'}} \langle \Psi | H | \Psi \rangle$, for correlated states in the subspace. Then it is trivial to derive from $\mathcal{K}$ and $\mathcal{E}$ a polynomial, $\Omega(\omega; b_1, \ldots, b_{\mathcal{N}'})$, for the difference, $\omega = \eta - \kappa$. The diagonalization of $\mathcal{P}H\mathcal{P}$ then reads,

$$\frac{\partial \kappa}{\partial b_\beta} + \frac{\partial \omega}{\partial b_\beta} = 0, \quad \beta = 1, \ldots, \mathcal{N}'. \tag{11}$$

With the ratio, $v_\beta = -(\partial\Omega/\partial b_\beta)/(\partial\Omega/\partial\omega)$, representing $\partial\omega/\partial b_\beta$, define the one-body, local potential, $v_\Delta(r) = \sum_{\beta=1}^{\mathcal{N}'} v_\beta \, S_\beta(r)$. Let Φ be the ground state of $\mathcal{P}\left[H_0 + \sum_{i=1}^{N} v_\Delta(r_i)\right]\mathcal{P}$. Notice that $\langle \Phi | \mathcal{P} S_\beta \mathcal{P} | \Phi \rangle = \langle \Phi | S_\beta | \Phi \rangle$. Then the energy E of Φ has derivatives,

$$\partial E/\partial v_\beta = \int (\Delta\rho + \rho_0)\, S_\beta = b_\beta + b_{\beta 0}, \tag{12}$$

because of the orthonormality of the S_β's. The numbers, $b_{\beta 0} = \int \rho_0 \, S_\beta$, are easily pretabulated. The quantities, v_β and $(b_\beta + b_{\beta 0})$, are Legendre conjugates, and, moreover, $\partial/\partial(b_\beta + b_{\beta 0}) = \partial/\partial b_\beta$. The conditions, Eqs. (11), read as the diagonalization for a determinant Φ with the same density ρ as that of the eigenstate Ψ of $\mathcal{P}H\mathcal{P}$. The potential, $\mathcal{P}(w_0 + v_\Delta)\mathcal{P}$, is a KSP valid for the subspace, up to the convergence of the truncation with $\mathcal{N}'$ terms.

4. Conclusion

This polynomial method most often uses a very non local parametrization of ρ, that deviates from the quasi-local tradition of the field. In every case, our unconventional parametrization of ρ creates a new zoology of DFs. Nothing of this zoology is known to us, but its interest is obvious, since manipulations of polynomials and properties of their roots, including bounds, are basic subjects. Moreover, extrapolations of polynomials, and criticism of such extrapolations, are easy. The number of *available, exactly solvable* models is *huge*. It is limited only by computational power. For nuclei or atoms, the models will be "radial" [13], somewhat simple. For nuclear physics, our ultimate goal will be to see whether particle number can be used as a constraint, to generate a mass formula. For electrons in molecules or extended systems (metals, thin layers, etc.), however, a necessary algebra of functions of 2 or 3 variables will burden the models. Anyhow, one can always test whether our polynomials from "smaller" models may remain good approximations for "larger" ones, if, for instance, scaling properties can be established. Asymptotic properties of a sequence of "DF polynomials" might guide towards derivations of more traditional DFs. In particular, the polynomial models allow comparisons between the KS and the true kinetic energies of correlated systems. They also provide explicit terms for those correlation energies due to interactions.

In conclusion, this algebraic method simplifies density functional theory into energy minimization under finite numbers of constraints, under very elementary manipulations of polynomials. It retains all essential informations about the density and all components of the energy.

Acknowledgments

SK acknowledges support from the National Research Foundation of South Africa and thanks CEA/Saclay for hospitality during part of this work. BG thanks Rhodes University for its hospitality during part of this work.

References

[1] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
[2] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
[3] Harris J and Jones R O 1974 *J. Phys.* **F 4** 1170
[4] Colonna F and Savin A 1999 *J. Chem. Phys.* **110** 2828
[5] Teale A M, Coriani S, and T. Helgaker 2009 *J. Chem. Phys.* **130** 10411
[6] Chayes J T, Chayes L, and Ruskai M B 1985 *J. Stat. Phys.* **38** 497
[7] Hariman J E 1983 *Phys. Rev.* **A 27** 632
[8] Pino R, Bokanowski O, Ludena E V, and Boada R L 2009 *Theor. Chem. Acc.* **123** 189
[9] Katriel J, Appellof C J, and Davidson E R 1981 *Int. J. Quant. Chem.* **19** 293
[10] Dreizler R M and Gross E K U 1990 *Density Functional Theory*, Springer, Berlin/Heidelberg; see also the references in their review.
[11] Parr R G and Yang W 1995 *Annu. Rev. Phys. Chem.* **46** 701
[12] Drut J E, Furnstahl R J, and Platter L 2010 *Prog. Part. Nucl. Phys.* **64**, 120
[13] Giraud B G 2008 *Phys. Rev.* **C 78** 014307
[14] Levy M 1979 *Proc. Natl. Acad. Sci. USA* **76** 6062; Lieb E H 1983 *Int. J. Quantum Chem.* **24** 243
[15] Giraud B G, Weiguny A, and Wilets L 2005 *Nucl. Phys.* **A 761** 22 (2005); Giraud B G 2005 *J. Phys.* **A 38** 7299 (2005).
[16] Normand J M 2007 *J. Phys.* **A 40** 2341 and 2371.
[17] Perdew J P and Levy M 1985 *Phys. Rev.* **31** 6264
[18] Giraud B G and Peschanski R 2006 *Acta Phys. Pol.* **B 37** 331
[19] Ullrich C A and Kohn W 2003 *Phys. Rev. Lett.* **89** 156401

Towards the unconditional security proof for the coherent-one-way protocol

Mhlambululi Mafu, Adriana Marais, Francesco Petruccione

Quantum Research Group, School of Physics and National Institute for Theoretical Physics, University of KwaZulu-Natal

E-mail: 209526077@ukzn.ac.za, 207521013@ukzn.ac.za and petruccione@ukzn.ac.za

Abstract. Quantum Cryptography, one aspect of which is Quantum Key Distribution, provides the only physically secure and proven method for the transmission of a secret key between two distant parties, Alice and Bob. The goal of QKD is to guarantee that a possible eavesdropper (Eve), with access to the communication channel is unable to obtain useful information about the message. The Coherent-One-Way protocol is one of the most recent practical QKD protocols. However, its unconditional security proof still remains unrealized. We therefore present a necessary condition for the security of the COW protocol. In the proof, we describe Bob's measurements by non-commuting POVM elements which satisfies this condition.

1. Introduction

The Coherent-One-Way (COW) Quantum Key Distribution (QKD) protocol was first proposed by Gisin et al. [1] and belongs to a class of the so-called distributed-phase-reference protocols [2]. Currently no lower bound is known for the unconditional security of this protocol. The existing tools for proving security of protocols against the most general attacks fail to apply to this protocol in a straight forward way. While security proofs for some limited attacks exist, the unconditional security proofs still remain unrealized. This paper presents an improvement in this direction for the COW QKD protocol by presenting a necessary security condition for the most general kind of attacks which is a step towards the unconditional security proof.

Various attacks have been studied for the COW QKD protocol. The security against individual attacks has been derived [3]. Upper bounds for the error rates for the security of COW protocol in the presence of large collective attacks have also been derived [3]. Security against intercept and resend attacks based on unambiguous state discrimination has been shown [4]. Proofs against sequential attacks [5] based on unambiguous state discrimination have also been shown [4]. The collective beam splitting attacks [3] have also been derived [2]. The security against the most general attacks is still elusive because the present tools for proving security of protocols in general cannot be adopted in a straightforward way.

In spite of these proofs against limited examples of kinds of attacks, it still remains unclear how the unconditional security proof can be realized. This is mainly because this class of protocols use coherent sequences of signals which are not symmetric as opposed to qubits in other classes of protocols. Again, these protocols move away from the symbol-per-symbol type of coding [6]. The notation and the formalism to be used to develop a full unconditional security proof is complicated for this class of protocols. Therefore, our goals in this paper is (i) to present

why the COW protocol is useful as a means of distributing a key, (ii) to show how the operation of the COW protocol takes place in the absence of a detailed explanation of the protocol as presented in the original literature [1, 7] and illustrate explicitly how the two parties extract the key and (iii) to provide a notation and formalism that can be a necessary condition for the security of COW QKD protocol.

2. Operation of the COW protocol

Alice prepares states $|\phi_0\rangle$ and $|\phi_1\rangle$ which represent logical states '0' and '1' respectively and decoy states in each of $k = 1, ..., N$ time intervals in a two-pulse sequence consisting of a non-empty and an empty pulse:

$$\begin{aligned} |\phi_0\rangle_k &= |\sqrt{\mu}\rangle_{2k-1}|0\rangle_{2k}, \\ |\phi_1\rangle_k &= |0\rangle_{2k-1}|\sqrt{\mu}\rangle_{2k}, \\ |\text{decoy}\rangle_k &= |\sqrt{\mu}\rangle_{2k-1}|\sqrt{\mu}\rangle_{2k}, \end{aligned} \tag{1}$$

where $2k-1$ and $2k$ label the pulses in the pair. In the case of a small mean photon number, the states $|\phi_0\rangle_k$ and $|\phi_1\rangle_k$ have a large overlap because of their vacuum component and also they poses a phase coherence between any two non-empty pulses with a bit separation. The decoy sequences are used to check for coherence in the data line and they are then going to be discarded in the public discussion. The key is obtained by measuring the time-of-arrival of photons on the data line, detector D_B. The presence of the eavesdropper is checked interferometrically in a monitoring line by randomly measuring the coherence between the succesive non-empty pulses; bit sequences '1-0' or decoy sequences with the interferometer and detectors D_{M1} and D_{M2}. If coherence is broken D_{M2} fires, and an error is recorded.

Bob uses a detector D_B to unambiguously discriminate the non-orthogonal states $|\phi_0\rangle_k$ and $|\phi_1\rangle_k$. Since μ is small Bob doesn't always get a click. But sometimes Bob gets a click in time interval k, and if the click corresponds to the first (second) pulse of the pair, he records a zero (one).

3. Prepare and measure (P&M) scheme

Alice prepares a random sequence of predefined non-orthogonal coherent states $\bigotimes_{k=1}^{N} |\psi(s_k)\rangle$, where each coherent state $|\psi(s_k)\rangle = |\phi_0\rangle_k$ or $|\phi_1\rangle_k$ is defined according to Equation 1. These states are sent to Bob through an untrusted quantum channel. On the receiving side, Bob performs a POVM on the signal he receives.

We divide each time slot k into two and label with integers j, such that $j = 1, ..., 2N$ then, according to Figure 1, the signal entering Bob's interferometer in path '0' after each time interval j can be described in terms of the creation operators $\hat{a_0}^{\dagger j}$ and the outgoing paths, $\hat{a_3}^{\dagger j}$, $\hat{a_7}^{\dagger j}$ and $\hat{a_8}^{\dagger j}$. In order to describe the signals entering Bob's interferometer, we follow the same approach used by A. Marais et al. [8], since these protocols belong to the same class. The total action of the interferometer is derived to be

$$\hat{a_0}^{\dagger j} \rightarrow \tfrac{1}{2\sqrt{2}}(\hat{a_7}^{\dagger j} - e^{i\phi_3}\hat{a_8}^{\dagger j} + 2\hat{a_3}^{\dagger j} + \hat{a_7}^{\dagger (j+1)} + e^{i\phi_3}\hat{a_8}^{\dagger (j+1)}), \tag{2}$$

where the subscripts '0,. . .,8' refer to paths as labeled in Figure 1, and $\phi_1 + \phi_2 + \phi_3 = \phi_{\triangle t}$ are phase shifts associated with symmetric BS1, BS2, BS3 and the time delay, respectively.

When Alice prepares a $|\phi_0\rangle$, the input state is transformed to the output state as follows

$$\begin{aligned} |\phi_0\rangle_k &= |\sqrt{\mu}\rangle_0^{(j-1)}|0\rangle_0^j \\ &\xrightarrow{I} |\tfrac{\sqrt{\mu}}{2\sqrt{2}}\rangle_7^{j-1}| - e^{i\phi_3}\tfrac{\sqrt{\mu}}{2\sqrt{2}}\rangle_8^{j-1}|\tfrac{\sqrt{\mu}}{\sqrt{2}}\rangle_3^{j-1}|\tfrac{\sqrt{\mu}}{2\sqrt{2}}\rangle_7^{j}|e^{i\phi_3}\tfrac{\sqrt{\mu}}{2\sqrt{2}}\rangle_8^{j}. \end{aligned} \tag{3}$$

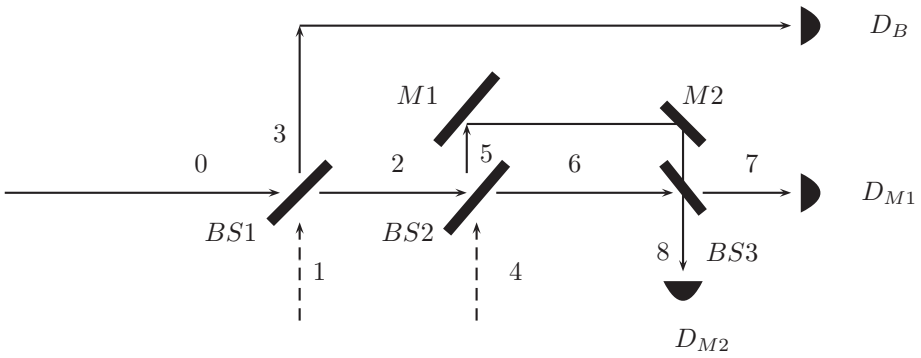


Figure 1. Schematic diagram of the QKD system for realizing the COW protocol. Key extraction takes place along the data-line, with detector D_B. Coherence between successive non-empty pulses is checked interferometrically in the monitoring line with detectors D_{M1} and D_{M2}. BS1, BS2 and BS3 are considered to be symmetric beamsplitters, M1 and M2 are mirrors and '0-8' are paths along which the coherent pulses travel.

$k=6$	$k=5$	$k=4$	$k=3$	$k=2$	$k=1$	
	1	0	0	0	1	$\leftarrow c$
$\|\mu\rangle_{12}\|\mu\rangle_{11}$	$\|0\rangle_{10}\|\mu\rangle_9$	$\|\mu\rangle_8\|0\rangle_7$	$\|\mu\rangle_6\|0\rangle_5$	$\|\mu\rangle_4\|0\rangle_3$	$\|0\rangle_2\|\mu\rangle_1$	$\leftarrow b$
decoy	1	0	0	0	1	$\leftarrow a$

Table 1. An example of the implementation of the COW protocol for $k = 1, \ldots, 6$ where k labels pairs; a is the logical bit recorded by Bob; b represents the states received at detector D_B and c is the bit value sent by Alice.

Here, Bob gets a click in D_B which corresponds to a click in time slot $j-1$ with $p_{\text{click}}=1-e^{-\mu/8}$, which is the first of the slots constituting interval k, and records a '0'. Since D_7 and D_8 click with equal probability in slots $j-1$ and j, there is no test for coherence from $|0\rangle_k$ above.

When Alice prepares a $|\phi_1\rangle$, the output state is of the form

$$\begin{aligned} |\phi_1\rangle_k &= |0\rangle_0^{(j-1)}|\sqrt{\mu}\rangle_0^j \\ &\xrightarrow{I} |\tfrac{\sqrt{\mu}}{\sqrt{2}}\rangle_3^j|\tfrac{\sqrt{\mu}}{2\sqrt{2}}\rangle_7^j|\tfrac{-\sqrt{\mu}}{2\sqrt{2}}\rangle_8^j|\tfrac{\sqrt{\mu}}{2\sqrt{2}}\rangle_7^{(j+1)}|\tfrac{\sqrt{\mu}}{2\sqrt{2}}\rangle_8^{(j+1)}. \end{aligned} \tag{4}$$

Here, Bob gets a click in slot j in D_B with $p_{\text{click}} = 1 - e^{-\mu/8}$ and records a '1' for the time interval k. Again, this follows for each of $k = 1, ..., N$ intervals, Bob records a '0'('1') when he gets a click in slot $j = 2k - 1(j = 2k)$.

In order to check for coherence in the data line, Alice prepares and sends decoy states to Bob. A loss of coherence reveals the presence of an eavesdropper, and the key is simply discarded and no information will be lost during that time interval. When Alice prepares a decoy state, the output is of the form

$$|\text{decoy}\rangle_k = |\sqrt{\mu}\rangle^{(j-1)}|\sqrt{\mu}\rangle^j, \tag{5}$$

but states formed from $|\phi_1\rangle_k|\phi_0\rangle_{k+1}$ can also be used for the channel estimation, i.e., $|\phi_1\rangle_k|\phi_0\rangle_{k+1} = |0\rangle_0^{(j-1)}|\sqrt{\mu}\rangle_0^j|\sqrt{\mu}\rangle_0^{(j+1)}|0\rangle_0^{(j+2)}$. The state $|\sqrt{\mu}\rangle_0^t|\sqrt{\mu}\rangle_0^{(t+1)}$ transforms the interferometer as follows

$$|\sqrt{\mu}\rangle_0^t|\sqrt{\mu}\rangle_0^{(t+1)} \xrightarrow{I} |\tfrac{\sqrt{\mu}}{\sqrt{2}}\rangle_3^t|\tfrac{\sqrt{\mu}}{2\sqrt{2}}\rangle_7^t|\tfrac{-\sqrt{\mu}}{2\sqrt{2}}\rangle_8^t|\tfrac{\sqrt{\mu}}{\sqrt{2}}\rangle_3^{(t+1)}|\tfrac{\sqrt{\mu}}{\sqrt{2}}\rangle_7^{(t+1)}|0\rangle_8^{(t+1)}|\tfrac{\sqrt{\mu}}{\sqrt{2}}\rangle_7^{(t+2)}|\tfrac{\sqrt{\mu}}{\sqrt{2}}\rangle_8^{(t+2)}. \tag{6}$$

So, it can be seen that decoy states and $|\phi_1\rangle_k|\phi_0\rangle_{k+1}$ sequences do not contribute to the key since here D_B has a probability to click for both slots j in the pair k, so that Bob learns no key bit. But, if D_{M2} clicks, this is an indication of a loss of coherence since if the consecutive non-empty pulses have a constant relative phase, this detector has zero probability of clicking as seen above.

Table 1 shows how the bits sent by Alice correspond to the sent states. To obtain the bit value, Bob has to distinguish unambiguously between the two non-orthogonal states given in Equation (1), that arrive at his detector. According to Table 1, Alice can also send decoy states. Alternatively, checks for coherence can be done between two consecutive non-empty pulses for example across the pair in $k = 4$ and $k = 5$. Based on these states Bob can record each respective bit as shown in the example depicted by the same Table 1.

4. Bob's measurements

We exploit the mathematical convenience of POVM's [9, 10] as a tool for describing Bob's measurement statistics. Since Bob has a probability of detecting one or more photons (a click) or vacuum (no click) in each of his detectors in $2N$ time intervals, there are 2^{6N} possible measurement outcomes corresponding to 2^{6N} POVM elements.

The projectors constituting Bob's measurement in the time intervals $j \in \{1, ..., 2N\}$ are written as

$$\begin{aligned}
G_1 &= |0\rangle\langle 0|, \\
G_2 &= \sum_{n=1}^{\infty} |n\rangle_3^1\langle n| \otimes |0\rangle\langle 0|, \\
G_3 &= |0\rangle_3^1\langle 0| \otimes \sum_{n=1}^{\infty} |n\rangle_7^1\langle n| \otimes |0\rangle\langle 0|, \\
G_4 &= |0\rangle_3^1\langle 0| \otimes |0\rangle_7^1\langle 0| \otimes \sum_{n=1}^{\infty} |n\rangle_8^1\langle n| \otimes |0\rangle\langle 0|, \\
G_5 &= |0\rangle_3^1\langle 0| \otimes |0\rangle_7^1\langle 0| \otimes |0\rangle_8^1\langle 0| \otimes \sum_{n=1}^{\infty} |n\rangle_3^2\langle n| \otimes |0\rangle\langle 0|, \\
&\vdots \\
G_{2^{6N}} &= \sum_{n=1}^{\infty} |n\rangle\langle n|. \qquad (7)
\end{aligned}$$

The $G_{i's}$ are projectors onto the basis of photon number states, $|n\rangle$. They reperesent all possible outcomes for an implementation of the COW protocol with signals sent in $2N$ time intervals. The projectors G_1 and $G_{2^{6N}}$ represent an implementation of the protocol when Bob measures vacuum and one or more photons respectively, in all the time intervals.

The action of Bob's beamsplitter BS1, together with the interferometer are represented by the operator $\mathcal{U}$ which maps the incoming state in path '0' to the outgoing states in paths '3', '7' and '8'. Now the POVM's E_j (where $E_j = {}_4\langle 0|_1\langle 0|\mathcal{U}^\dagger G_j \mathcal{U}|0\rangle_1|0\rangle_4$) are the operators that act only on the states in path '0'. The expectation value with respect to the vacuum in path '1' reduces the action of the operator $\mathcal{U}^\dagger G_j \mathcal{U}$ to the subspace of the states in path '0', similar to the partial trace.

The effect that corresponds to a click on Bob's detector D_B in $j = 1; k = 1$ and vacuum in

all other slots is given by

$$\begin{aligned} E_2 &= {}_{1,4}\langle 0|\mathcal{U}^\dagger G_2 \mathcal{U}|0\rangle_{1,4} \\ &= \sum_{n=1}^{\infty} \frac{1}{2^n n!} (\hat{a}_0^{\dagger 1})^n |0\rangle\langle 0| (\hat{a}_0^1)^n. \end{aligned} \tag{8}$$

Similarly a click in D_B in $j = 3; k = 2$ and vacuum in all other slots becomes

$$\begin{aligned} E_3 &= {}_{1,4}\langle 0|\mathcal{U}^\dagger G_3 \mathcal{U}|0\rangle_{1,4} \\ &= \sum_{m=1}^{\infty} \frac{1}{2^m m!} (\hat{a}_0^{\dagger 3})^m |0\rangle\langle 0| (\hat{a}_0^3)^m. \end{aligned} \tag{9}$$

The commutator is then given by

$$\begin{aligned} [E_2, E_3] &= \sum_{n=1}^{\infty} \frac{1}{2^n n!} (\hat{a}_0^{\dagger 1})^n |0\rangle\langle 0| (\hat{a}_0^1)^n \sum_{m=1}^{\infty} \frac{1}{2^m m!} (\hat{a}_0^{\dagger 3})^m |0\rangle\langle 0| (\hat{a}_0^3)^m \\ &- \sum_{m=1}^{\infty} \frac{1}{2^m m!} (\hat{a}_0^{\dagger 3})^m |0\rangle\langle 0| (\hat{a}_0^3)^m \sum_{n=1}^{\infty} \frac{1}{2^n n!} (\hat{a}_0^{\dagger 1})^n |0\rangle\langle 0| (\hat{a}_0^1)^n \\ &= 0, \end{aligned} \tag{10}$$

since $\langle 0|(\hat{a}_0^1)^n (\hat{a}_0^3)^m|0\rangle = 0$, i.e., the time intervals $j = 1$ and $j = 3$ belong to different Hilbert spaces. So without clicks for checks of coherence, it means that there is no security. Therefore, when we describe Bob's measurements by commuting operators, an eavesdropper could measure an observable which commutes with all of Bob's measurements, thus remaining undetected. Therefore, it is important that some of the POVM elements describing Bob's measurements must be non-commuting. This can be shown in the proof below. If we consider a click in D_B in time interval $j = 2$ and a click in D_{M1} in time interval $j = 3$, and vacuum everywhere, we have

$$\begin{aligned} E_4 &= {}_{1,4}\langle 0|\mathcal{U}^\dagger G_4 \mathcal{U}|0\rangle_{1,4} \\ &= \sum_{n=1}^{\infty} \frac{1}{2^n n!} (\hat{a}_0^{\dagger 2})^n |0\rangle\langle 0| (\hat{a}_0^2)^n, \\ E_5 &= {}_{1,4}\langle 0|\mathcal{U}^\dagger G_5 \mathcal{U}|0\rangle_{1,4} \\ &= \sum_{m=1}^{\infty} \frac{1}{8^m m!} (\hat{a}_0^{\dagger 2} + \hat{a}_0^{\dagger 3})^m |0\rangle\langle 0| (\hat{a}_0^2 + \hat{a}_0^3)^m. \end{aligned} \tag{11}$$

The commutator is then given by

$$\begin{aligned} [E_4, E_5] &= \sum_{n=1}^{\infty} \frac{1}{2^n n!} (\hat{a}_0^{\dagger 2})^n |0\rangle\langle 0| (\hat{a}_0^2)^n \sum_{m=1}^{\infty} \frac{1}{8^m m!} (\hat{a}_0^{\dagger 2} + \hat{a}_0^{\dagger 3})^m |0\rangle\langle 0| (\hat{a}_0^2 + \hat{a}_0^3)^m \\ &- \sum_{m=1}^{\infty} \frac{1}{8^m m!} (\hat{a}_0^{\dagger 2} + \hat{a}_0^{\dagger 3})^m |0\rangle\langle 0| (\hat{a}_0^2 + \hat{a}_0^3)^m \sum_{n=1}^{\infty} \frac{1}{2^n n!} (\hat{a}_0^{\dagger 2})^n |0\rangle\langle 0| (\hat{a}_0^2)^n. \\ &= \sum_{n=1}^{\infty} \frac{1}{2^n n!} (\hat{a}_0^{\dagger 2})^n |0\rangle n! \sum_{n=1}^{\infty} \frac{1}{8^n} \langle 0| (\hat{a}_0^2 + \hat{a}_0^3)^n \\ &- \sum_{n=1}^{\infty} \frac{1}{8^n n!} \langle 0| (\hat{a}_0^{\dagger 2} + \hat{a}_0^{\dagger 3})^n |0\rangle n! \sum_{n=1}^{\infty} \frac{1}{2^n} \langle 0| (\hat{a}_0^{\dagger 2})^n \end{aligned} \tag{12}$$

$$
\begin{aligned}
&= \sum_{n=1}^{\infty}\sum_{k=0}^{n}\binom{n}{k}\frac{1}{2^n 8^n n!}\{(\hat{a}_0^{\dagger 2})^n|0\rangle\langle 0|(\hat{a}_0^2)^{n-k}(\hat{a}_0^3)^k - (\hat{a}_0^{\dagger 2})^{n-k}(\hat{a}_0^{\dagger 3})^k|0\rangle\langle 0|(\hat{a}_0^2)^n\} \\
&= \sqrt{n!}\sqrt{k!}\sqrt{(n-k)!}\left(|n\rangle_0^2\langle n-k|_0^2\langle k|_0^3 - |n-k\rangle_0^2|k\rangle_0^3\langle n|_0^2\right) \\
&\neq 0. \qquad (13)
\end{aligned}
$$

Since the operators $\hat{a}_0^{(\dagger)2}, \hat{a}_0^{(\dagger)3}$ act on different Hilbert spaces, the matrix elements do not cancel. Therefore, it has been shown that there exist non-commuting POVM elements in Bob's measurements, hence a precondition for the security of the COW protocol has been shown to be met.

5. Conclusion

From the above calculation, we can recognize that there exist non-commuting POVM elements in Bob's measurement in the P&M version of the COW protocol. Thus the COW protocol has been proven to satisfy an important necessary condition for security and such a description is an essential step in a potential proof for the most general kind of attacks.

Acknowledgments

This work is based on research supported by the South African Research Chair Initiative of the Department of Science and Technology and National Research Foundation.

References

[1] Stucki D, Brunner N, Gisin N, Scarani V and Zbinden H 2005 *Applied Physics Letters* **87** 194108
[2] Scarani V, Bechmann-Pasquinucci H, Cerf N, Dušek M, Lütkenhaus N and Peev M 2009 *Reviews of Modern Physics* **81** 1301
[3] Branciard C, Gisin N and Scarani V 2008 *New Journal of Physics* **10** 013031
[4] Branciard C, Gisin N, Lütkenhaus N and Scarani V 2006 *arXiv:0609090*
[5] Tsurumaru T 2007 *Physical Review A* **75** 062319
[6] Stucki D, Walenta N, Vannel F, Thew R, Gisin N, Zbinden H, Gray S, Towery C and Ten S 2009 *New Journal of Physics* **11** 075003
[7] Stucki D, Fasel S, Gisin N, Thoma Y and Zbinden H 2007 *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* vol 6583 p 18
[8] Marais A, Konrad T and Petruccione F 2010 *Journal of Physics A: Mathematical and Theoretical* **43** 305302
[9] Nielsen M, Chuang I and Grover L 2002 *American Journal of Physics* **70** 558
[10] Audretsch J 2007 *Entangled systems: New Directions in Quantum Physics* (Wiley-VCH)

Many-boson quantum walks on graphs with shared coins

Dibwe Pierrot Musumbu[1], Ilya Sinayskiy[1,2] and Francesco Petruccione[1,2]

[1]Quantum Research Group, University of KwaZulu-Natal, Westville-Campus Durban KwaZulu-Natal South Africa,

[2] National Institute for Theoretical Physics, University of KwaZulu-Natal, Westville-Campus, Durban, South Africa.

E-mail: musumbu@ukzn.ac.za, sinayskiy@ukzn.ac.za, petruccione@ukzn.ac.za

Abstract. Quantum walks of particles obeying Bose statistics are introduced. In such quantum walks the conditional shift operation is performed with a single coin tossing for the whole lattice. An explicit form for the transition probabilities in a single step is derived. This allows to describe the evolution of an arbitrary state and an arbitrary number of steps. This model easily embraces the concepts of the join probability, the counting statistics and the high order correlations. It also presents the computational challenges arising from the exponential increase in the number of basis states entering into the lattice state as a function of the number of quantum walkers and the number of steps. Possible solutions are proposed in some applications of the model to quantum walks on finite graphs.

1. Introduction

After the introduction of quantum walks by Aharonov et al. [1], scores of studies have developed the theory of both discrete-time and continuous-time quantum walks. Quantum walks have revealed strong advantages over their classical counterparts: It has been shown that quantum walks spread quadratically faster compared to their classical counterparts on the line [2, 3]. While the quadratic speed up is also present in search algorithms with discrete quantum walks [4], it is even exponentiated using continuous-time quantum walks in the path searching algorithm over certain types of graphs [5].

Our goal is to make the simplest extension of the single particle quantum random formalism to the case of many indistinguishable particles. Considering an initial state with precisely located walkers, the system evolves to a certain particle distribution. The development of such a formalism could extend our understanding of the concepts of quantum coherence, interference and entanglement. In the following we consider many-particle quantum walks with a shared coin. In other words, we associate a single coin to the whole underlying lattice.

The paper is structured as follows: In the next Section we introduce the mathematical formulation of quantum walks for identical particles. We will use a conventional Hadamard operation, but we will introduce a generalized conditional shift operator. In Section 3 we show how to realize the steps of the random walk and derive a recursion relation between amplitudes. In Section 4 we present the results of simulations of shared coin quantum walks for identical

particles. We present results for the particle density distribution, the counting statistics, the Fano-Mandel factor for each vertex occupation number and apply this scheme in the evaluation of the high order correlation coefficients. The last section contains our conclusions and gives a short outlook.

2. Graph state

We consider a $M+1$-vertices's graph. Suppose that this graph is populated by N particles. For example if the graph is the square in Figure 1 b), the initial configuration is

$$\left|\Psi^{[0]}_{\text{GRAPH}}\right\rangle = \frac{1}{\sqrt{2}}\left\{\left|\text{L}\{0;2;1;0\}\right\rangle + \left|\text{R}\{0;2;1;0\}\right\rangle\right\}. \tag{1}$$

where the upper index represent the discrete time step. Taking into account that the number of particles is conserved, the initial 3 particle state on the 4-node graph, will reduce the Hilbert space to 20 possible independent state vectors. $|\text{L}\rangle$ and $|\text{R}\rangle$ are chiralities of the coins. In general

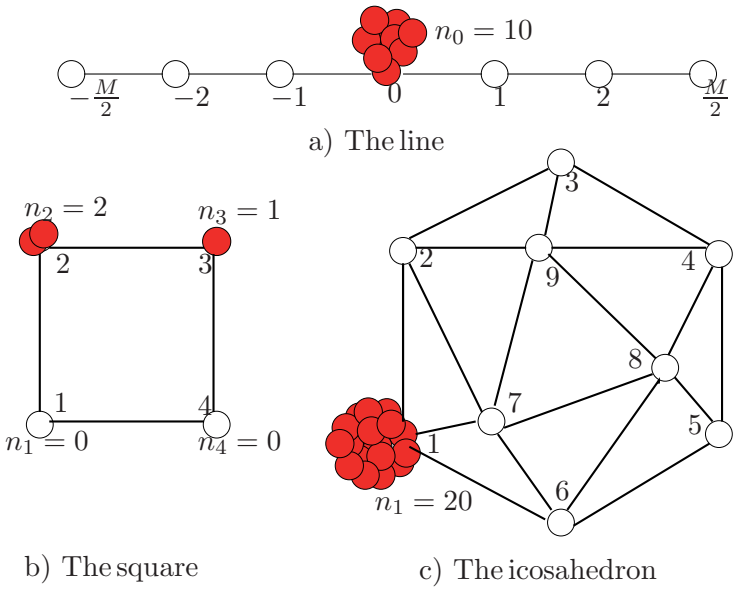


Figure 1. a) the $M+1$-vertices line with $n_0 = 10$ particles, b) the square graph with $n_2 = 2$ and $n_3 = 1$. c) the icosahedron with $n_1 = 20$ particles.

we have

$$\left|\Psi^{[0]}_{\text{GRAPH}}\right\rangle = \sum_{\ell=1}^{20}\sum_{k=1}^{2} c^0_{k\,\ell}\left|v_k\{n^0_{1\,\ell}; n^0_{2\,\ell}; n^0_{3\,\ell}; n^0_{4\,\ell}\}\right\rangle. \tag{2}$$

The coefficients $c^0_{k\,\ell}$s are the weights of configurations. The vertices occupation numbers representation for the system square-graph-3-particles (figure 1 b)) is given by

$$\left|\Psi^{[0]}_{\text{GRAPH}}\right\rangle = \sum_{n^0_1=0}^{3}\sum_{n^0_2=0}^{3}\sum_{n^0_3=0}^{3}\sum_{n^0_4=0}^{3}\sum_{k=1}^{2} C_{\{n^0_1;n^0_2;n^0_3;n^0_4\}}\left|v_k\{n^0_1; n^0_2; n^0_3; n^0_4\}\right\rangle, \tag{3}$$

where the $C_{\{n_1^0;n_2^0;n_3^0;n_4^0\}}$s are products of different types of correlations involved in the selection of occupation numbers over each specific vertex. In general we have

$$\left|\Psi^{[r]}_{\text{GRAPH}}\right\rangle = \sum_{n^r_{-\frac{M}{2}}=0}^{N} ... \sum_{n^r_{\alpha}=0}^{N} ... \sum_{n^r_{\frac{M}{2}}=0}^{N} \sum_{k=1}^{d} C_{k\{n^r_{-\frac{M}{2}};...;n^r_{\alpha};...;n^r_{\frac{M}{2}}\}} \left|v_k\left\{n^r_{-\frac{M}{2}};...;n^r_{\alpha};...;n^r_{\frac{M}{2}}\right\}\right\rangle, \quad (4)$$

$$\left|\Psi^{[r]}_{\text{GRAPH}}\right\rangle = \sum_{\ell=1}^{\text{D}} \sum_{k=1}^{d} c^r_{k\,\ell} \left|v_k\{...;n^r_{\alpha\,\ell};...\}\right\rangle \quad \alpha \in \left(-\frac{M}{2},...,\frac{M}{2}\right). \quad (5)$$

The upper index r indicates of the time step of the quantum walks. $\left|\Psi^{[r]}_{\text{GRAPH}}\right\rangle \in \mathcal{H} = \mathcal{H}_C \otimes \mathcal{H}_V$; [7], with $\mathcal{H}_C$ the coins Hilbert space and $\mathcal{H}_V$ the position Hilbert space.

3. Generalised conditional shift operator

Each step of the coined quantum walks consists of two operations. The first operation is the coin tossing [8] operation performed using the Hadamard gate

$$\mathbf{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (6)$$

Considering the degrees

$$|\text{L}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } |\text{R}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (7)$$

of the coin Hilbert space, the coin tossing respectively transforms $|\text{L}\rangle$ and $|\text{R}\rangle$ as

$$\mathbf{H}|\text{L}\rangle = \frac{1}{\sqrt{2}}(|\text{L}\rangle + |\text{R}\rangle) \text{ and } \mathbf{H}|\text{R}\rangle = \frac{1}{\sqrt{2}}(|\text{L}\rangle - |\text{R}\rangle). \quad (8)$$

The second operation is the particle shifting operation, where the shift is performed using the operator

$$\mathbf{S} = |\text{L}\rangle\langle\text{L}| \otimes \sum_{\eta\in\mathbb{Z}} |\eta-1\rangle\langle\eta| + |\text{R}\rangle\langle\text{R}| \otimes \sum_{\eta\in\mathbb{Z}} |\eta+1\rangle\langle\eta|. \quad (9)$$

In the single particle quantum walk case, the projector $|\eta+1\rangle\langle\eta| \equiv |\eta+1\rangle\langle 0| \cdot |0\rangle\langle\eta|$ corresponds to the product $\hat{\mathbf{a}}^\dagger_{\eta+1}\hat{\mathbf{a}}_\eta$ of quantum creation and annihilation operators. We define $\mathbf{H}_d$, the coin's tossing operator, as

$$\mathbf{H}_d = \left(h_{k\,\jmath}\right) \quad j,\, k \in (1,2,3,...,d) \text{ and } h_{k\,\jmath} \text{ are roots of unity;} \quad (10)$$

and a d-order Hadamard operator as

$$\mathbf{H}_d|v_k\rangle = \sum_{\jmath=1}^{d} h_{k\,\jmath}|v_\jmath\rangle. \quad (11)$$

The shifting operation combines the graph connectivity and the quantum creation and annihilation operation. Since particles can only be shifted between adjacent vertices, the adjacency matrix intervenes during the shifting operation. In general, it is defined as

$$\mathbf{A} = \left(a_{\mu\,\eta}\right) \quad \mu,\, \eta \in (1,2,3,...,M+1); \quad (12)$$

$$a_{\mu\,\eta} = \begin{cases} 1, \text{ if the vertices } \eta \text{ and } \mu \text{ are connected,} \\ 0, \text{ otherwise.} \end{cases} \quad (13)$$

For the graph in Figure 1 b), $\mathbf{A}$ is a 4×4 matrix with mirror reflection symmetry such that it splits into

$$\mathbf{A} = \mathbf{A}_{\mathrm{L}} + \mathbf{A}_{\mathrm{L}}^{\mathrm{T}}, \tag{14}$$

$$\text{with} \qquad \mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \text{ and } \mathbf{A}_{\mathrm{L}} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{15}$$

On the other side, let us consider

$$\mathcal{C} = (\hat{\mathbf{a}}_1^\dagger, \hat{\mathbf{a}}_2^\dagger, \hat{\mathbf{a}}_3^\dagger, \hat{\mathbf{a}}_4^\dagger). \tag{16}$$

The $\hat{\mathbf{a}}_\eta$s and $\hat{\mathbf{a}}_\mu^\dagger$s are respectively boson annihilation and creation operators obeying:

$$\left[\hat{\mathbf{a}}_\eta, \hat{\mathbf{a}}_\mu^\dagger\right] = \delta_{\eta\,\mu} \text{ and } \left[\hat{\mathbf{a}}_\eta, \hat{\mathbf{a}}_\mu\right] = \left[\hat{\mathbf{a}}_\eta^\dagger, \hat{\mathbf{a}}_\mu^\dagger\right] = 0. \tag{17}$$

The conditional shift operator for the square graph Figure 1 b) is given by

$$\mathbf{S} = \mathcal{C}\mathbf{A}_{\mathrm{L}}\mathcal{C}^\dagger \otimes |\mathrm{L}\rangle\langle\mathrm{L}| + \mathcal{C}\mathbf{A}_{\mathrm{L}}^{\mathrm{T}}\mathcal{C}^\dagger \otimes |\mathrm{R}\rangle\langle\mathrm{R}|. \tag{18}$$

The conditional shift operator (18) combined with the Hadamard operator (10) defines the single step quantum walk operator **SH**. Consider the graph state $\left|\Psi_{\mathrm{GRAPH}}^{[r]}\right\rangle$ (5) at time r. One time step later this state evolves to

$$\left|\Psi_{\mathrm{GRAPH}}^{[r+1]}\right\rangle = (\mathbf{SH})\left|\Psi_{\mathrm{GRAPH}}^{[r]}\right\rangle = \frac{1}{\left\langle\Psi_{\mathrm{GRAPH}}^{[r+1]}\middle|\Psi_{\mathrm{GRAPH}}^{[r+1]}\right\rangle} \sum_{s=1}^{\mathrm{D}} \sum_{i=1}^{d} c_{i\,s}^{r+1} \left|v_i\{\ldots; n_{\alpha\,s}^{r+1}; \ldots\}\right\rangle, \tag{19}$$

where

$$n_{\alpha\,s}^{r+1} = n_{\alpha\,\ell}^{r} - \delta_{\alpha\,\eta} + \delta_{\alpha\,\mu} \tag{20}$$

and

$$c_{i\,s}^{r+1} = \sum_{\mu=1}^{M+1} \sum_{\eta=1}^{M+1} \sum_{k=1}^{d} a_{\mu\,\eta}^{\mathrm{T}(k)} h_{k\,\jmath} c_{k\,\ell}^{r} \sqrt{(n_{\alpha\,\ell}^{r})^{\delta_{\alpha\,\eta}} (n_{\alpha\,\ell}^{r} - \delta_{\alpha\,\eta} + \delta_{\alpha\,\mu})^{\delta_{\alpha\,\mu}}}. \tag{21}$$

As an example, applying the above to the initial state (2) of the square graph (figure 1 b)), it follows that

$$\begin{aligned} \left|\Psi_{\mathrm{GRAPH}}^{[1]}\right\rangle = \frac{1}{2\sqrt{5}} \Big(& (1+i)\sqrt{2}\left|v_1\{1;1;1;0\}\right\rangle + (1+i)\sqrt{3}\left|v_1\{0;3;0;0\}\right\rangle \\ & + (1-i)2\left|v_2\{0;1;2;0\}\right\rangle + (1-i)\left|v_2\{0;2;0;1\}\right\rangle \Big). \end{aligned} \tag{22}$$

4. Results and conclusions

The probability of having a specific configuration at the step r is given by

$$P_\ell^r = \sum_{j=1}^{d} c_{j\,\ell}^r \left(c_{j\,\ell}^r\right)^*. \tag{23}$$

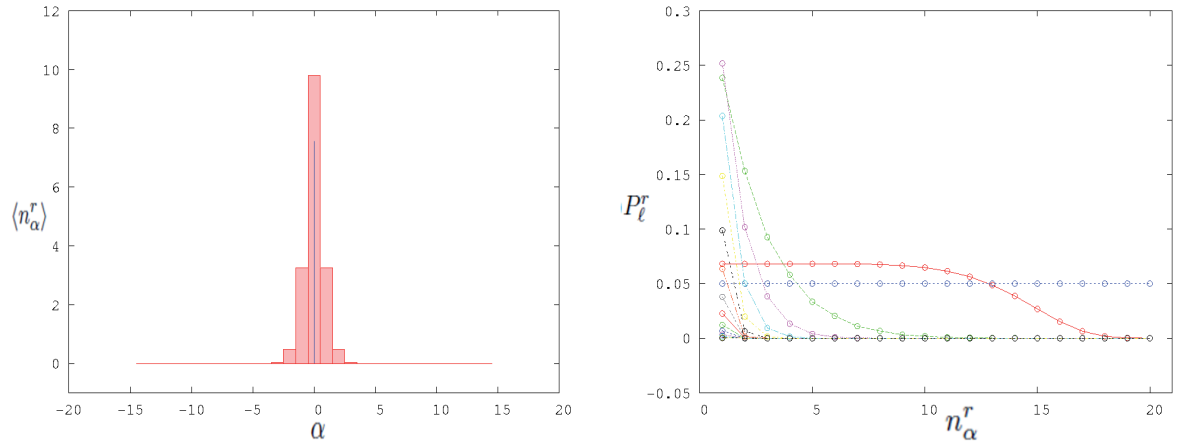


Figure 2. The figure on the left hand side represents the particle density distribution. The solid blue line is the initial distribution and the reds bars are the distribution after 28 steps. The figure on the right hand represents The vertices wise counting statistics for the shared coins many-particle quantum walks for the coherent state $|z\rangle = \frac{1}{\mathcal{N}} \sum_{n=0}^{20} \frac{z^n}{\sqrt{n!}} |n\rangle$. The horizontal blue line represents the counting at step $r = 0$ and all others are after $r = 28$. The particles counting statistics on vertex $\alpha = 0$ shows an increase in the counting of $n_\alpha^r \leq 13$.

This probability is the joint probability of the occupation numbers $n_{\alpha\,\ell}^r$ over the configuration ℓ. The qth-moment of the occupation number on a vertex α is given by:

$$\left\langle (n_\alpha^r)^q \right\rangle = \sum_{\ell=1}^{\mathrm{D}} P_\ell^r (n_{\alpha\,\ell}^r)^q. \tag{24}$$

The counting statistics is directly performed vertex by vertex and step by step. The shared coin's many-particle quantum walks permit the direct accessibility to all elements of the basis entering in the formation of the graph state. We record the number of occupants as they enter in the formation of the occupation numbers and the frequencies of the different occupation numbers.

The Fano-Mandel factor Q_α^r defines the deviation of the counting statistics from Poissonian statistics. It is given by

$$Q_\alpha^r = \frac{\langle (n_\alpha^r)^2 \rangle - \langle n_\alpha^r \rangle^2}{\langle n_\alpha^r \rangle}. \tag{25}$$

If $Q_\alpha^r = 1$, the vertex α counting statistics corresponds to the Poisson distribution. For $Q_\alpha^r < 1$, the vertex α counting statistics corresponds to sub-Poissonian distribution which indicates particles bunching on the vertex. And $Q_\alpha^r > 1$ the vertex α counting statistics corresponds to super-Poissonian distribution which indicates particles anti-bunching on the vertex.

The study of many-particle quantum walks faces the challenge of dealing with an exponential increase in the number of parameters needed for the control of the walkers. In the shared coin many-particle quantum walks, this problem is related to the number $\mathrm{D}(r)$ of configurations effectively entering in the formation of the graph state. So far we have recorded attempts with 8 quantum walkers in continuous or discrete quantum walks. In the discrete many-particle quantum walks the highest number of steps is around $50 = 2 \times 25$ [10], where the study is performed over 8 quantum walkers. In general for an open type of graphs $\mathrm{D}(r)$ increases exponentially (figure 4 right). When we consider finite graphs the expansion slows

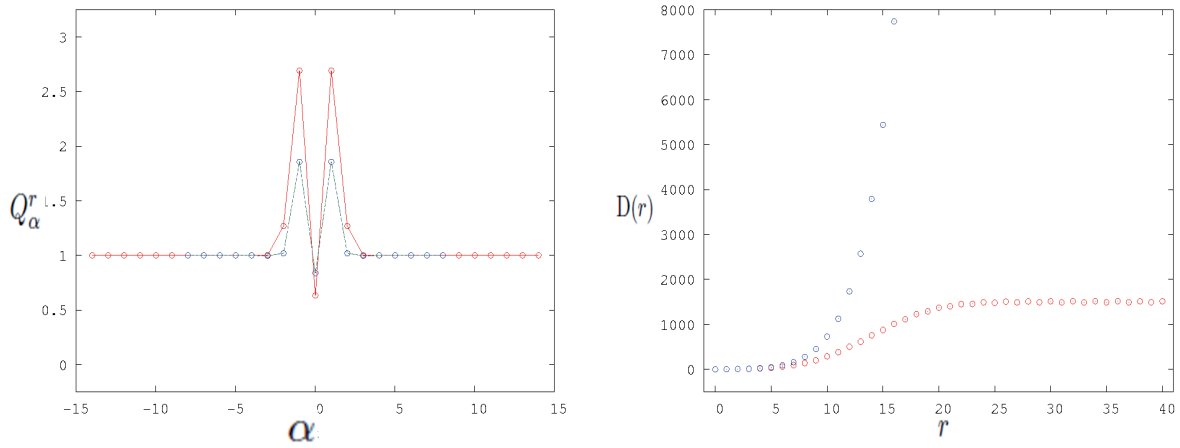


Figure 3. The figure on the left hand side represents the vertices's Fano Q^r_α factor for 16 steps (blue lines) and 28 steps (red color). The figure on the right hand side is the comparison between the numbers D(r) of the configurations effectively involved in the formation of the graph state $\left|\Psi^{[r]}_{\text{GRAPH}}\right\rangle$. The blue dots are for the line while the reds are a 20 vertices cycle both with 20 quantum walkers.

down significantly and therefore this model can be used in the study of many-particle quantum walks on finite graphs. In the case finite graphs the number D(r) converges to an upper bound (figure 4 right) after a certain number of steps.

In conclusion, we have presented a simple generalization to many particles of a quantum random walk. The method is based on the use of a shared coin. Although, such shared coin many-particle quantum walks are computationally challenging, quite efficient ways of performing the counting statistics are available.

Acknowledgments

This work is based on research supported by the South African Research Chair Initiative of the Department of Science and Technology and National Research Foundation.

References

[1] Aharonov Y, Davidovich L and Zagury N 1993 *Phys. Rev.* A **48** 2 1687-1690
[2] A. Nayak, A. Vishwanath 2000 *quant-ph/0010117*
[3] Travaglione B C and Milburn G J 2002 *Phys. Rev.* A **65** 032310
[4] D. Aharonov, A. Ambainis, J. Kempe and U. Vazirani 2001 *Proceedings of the thirty-third annual ACM symposium on Theory of computing* 50-59
[5] Childs A and Goldstone J 2004 *Phys. Rev.* A **70**, 022314
[6] Greiner M, Mandel O, Esslinger T, Hansch T W and Block I 2002 *Nature* **415** 39-44.
[7] Dibwe Pierrot Musumbu, Ilya Sinayskiy, and Francesco Petruccione work in progress.
[8] Julia Kempe 2003 *Contemporary Physics* **44** 307-327.
[9] A. Ambainis, J. Kempe, and A.Rivosh:2005 *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms* 1099-1108
[10] Klaus Mayer, Malte C. Tichy, Florian Mintert, Thomas Konrad, Andreas Buchleitner: 2011 *Phys. Rev.* A **83**, 062307

Few-nucleon systems studies with extended antisymmetrized molecular dynamics

G J Rampho and S A Sofianos

Department of Physics, University of South Africa, Pretoria 0003, South Africa

E-mail: rampho@science.unisa.ac.za

Abstract. Ground-state properties of three- and four-nucleon systems are studied with the angular momentum and parity projected antisymmetrized molecular dynamics. The Hamiltonian is constructed with realistic nucleon-nucleon interactions. The calculated ground-state energies, root-mean-square radii, and magnetic dipole moments are compared with the experimental results. In overall, the ground-state properties of the light nuclei considered are satisfactorily described.

1. Introduction

The structure of few-nucleon systems has been the subject of numerous theoretical and experimental investigations and reviews on a variety of their aspects can be found in the literature (see, for example, Refs. [1, 2]). The simplicity of these systems allows for an exact description and rigorous solutions of the corresponding dynamical equations can be obtained and thus accurate wave functions for bound states can now be constructed using realistic Hamiltonians. An illustration of the level of accuracy in describing ground state properties of the four-nucleon system by seven different state-of-the-art methods is given in Ref. [3]. However, the practical application of these methods become quite complicated as the number of particles involved increases and thus going beyond the $A = 4$ systems using exact methods still is a challenge.

In recent years, a very promising method, namely, the Antisymmetrized Molecular Dynamics (AMD), has been used to study the properties of A-particle systems. The AMD approach [4] is developed from the Time-Dependent Cluster Model [5] describing fermionic systems. This approach combines Fermi-Dirac statistics with quantum mechanics to treat the motion of the A particles [6]. Although the model is not fully quantum mechanical and does not assume a shell structure for the system, improved AMD wave functions are, nowadays, shown to give good predictions of the properties of few-body systems [7, 8]. In the present work the parity projected and angular momentum projected version of the AMD [9] is employed.

In Sect. 2 the general formalism of the AMD approach is summarized and the construction of the wave function, the equations of motion of the variable parameters, and the variational technique used are briefly outlined. Theoretical predictions of the ground state properties of three-nucleon and four-nucleon systems are presented in Sect. 3 while the conclusions drawn are given in Sect. 4.

2. The AMD formalism

The AMD wave function describing a system of A nucleons is constructed as a Slater determinant

$$\Psi_{AMD}(\vec{\mathrm{S}}) = \frac{1}{\sqrt{A!}} \det[\,\phi_j(\alpha,\vec{s}_i),\chi_j(\vec{\sigma}_i),\xi_j(\vec{\tau}_i)\,] \tag{1}$$

where ϕ, χ and ξ are, respectively, the spatial, spin and isospin components of the single-particle wave functions. The $\vec{s}_i$ are complex variational parameters, $\vec{\mathrm{S}} \equiv \{\vec{s}_1,\vec{s}_2,\vec{s}_3,\ldots,\vec{s}_A\}$, while α is a real constant width parameter. The single nucleon wave functions are given by [10]

$$\psi_i(\vec{r}_j) = \left(\frac{2\,\alpha}{\pi}\right)^{2/4} \exp\left[-\alpha\left(\vec{r}_j - \frac{\vec{s}_i(t)}{\sqrt{\alpha}}\right)^2 + \frac{1}{2}\,\vec{s}_i^{\,2}(t)\right] \otimes \chi_i \otimes \xi_i \tag{2}$$

where $\chi_i \otimes \xi_i \equiv \{\, N\otimes\uparrow \text{ or} N\otimes\downarrow \,\}$ are the fixed spin-isospin states of the i-th nucleon, which represent nucleon with spin-up or spin-down. The width parameter is a free parameter and is common for all Gaussians terms.

A wave function with a definite parity π, a total angular momentum J, and angular momentum projection M, is constructed from the AMD wave function as

$$\Psi^{J\pi}_{MK}(\vec{\mathrm{S}}) = \frac{1}{2}\,P^J_{MK}(\Omega)\,[\,1 \pm P^\pi\,]\,\Psi_{AMD}(\vec{\mathrm{S}}) \tag{3}$$

where $P^J_{MK}(\Omega)$ is the angular momentum projection operator, P^π the parity projection operator. The angular momentum projection operator is defined by [11]

$$P^J_{MK}(\Omega) = \frac{2\,J+1}{8\,\pi^2}\int \mathrm{d}\Omega\, D^{J*}_{MK}(\Omega)\,\hat{R}(\Omega) \tag{4}$$

where $D^J_{MK}(\Omega)$ is the Wigner D-function, $\hat{R}(\Omega)$ the rotation operator and $\Omega \equiv \{\alpha,\beta,\gamma\}$ represents the Euler rotation angles.

The time-dependent variational principle [10]

$$\delta\int_{t_1}^{t_2} \frac{\langle\,\Psi(\vec{S})\,|\,\mathrm{i}\hbar\,\frac{\partial}{\partial t} - H\,|\,\Psi(\vec{S})\,\rangle}{\langle\,\Psi(\vec{S})\,|\,\Psi(\vec{S})\,\rangle}\,\mathrm{d}t = 0 \tag{5}$$

with the constraints

$$\delta\Psi(t_1) = \delta\Psi(t_2) = \delta\Psi^*(t_1) = \delta\Psi^*(t_2) = 0. \tag{6}$$

is used to determine the dynamical equations for the variational parameters. The resulting equations can be transformed into the form [12]

$$\frac{\mathrm{d}\vec{s}_i}{\mathrm{d}t} = -b\,\frac{\partial E_0^{J\pm}(\vec{S},\vec{S}^*)}{\partial\vec{s}_i^{\,*}} \qquad , \qquad \frac{\mathrm{d}\vec{s}_i^{\,*}}{\mathrm{d}t} = -b\,\frac{\partial E_0^{J\pm}(\vec{S},\vec{S}^*)}{\partial\vec{s}_i} \tag{7}$$

where b is an arbitrary positive real constant and $E_0^{J\pm}(\vec{S},\vec{S}^*)$,

$$E_0^{J\pm}(\vec{S},\vec{S}^*) = \frac{\langle\,\Psi^{J\pm}_{MK}(\vec{\mathrm{S}})|\,H\,|\Psi^{J\pm}_{MK}(\vec{\mathrm{S}})\,\rangle}{\langle\,\Psi^{J\pm}_{MK}(\vec{\mathrm{S}})|\Psi^{J\pm}_{MK}(\vec{\mathrm{S}})\,\rangle}. \tag{8}$$

is the variational energy determined from the Hamiltonian H of the nucleus given by

$$H = -\sum_i \frac{\hbar^2}{2\,m_i}\,\nabla_i^2 + \frac{1}{2}\sum_{i\neq j}\Big[\,V_{NN}(\vec{r}_{ij}) + V_C(\vec{r}_{ij})\,\Big] \tag{9}$$

where m_i is the mass of nucleon i, $V_{NN}(\vec{r})$ the two-body nuclear potential, $V_C(\vec{r})$ the Coulomb potential and $\vec{r}$ the relative position vector of the interacting nucleons. In this work the AV4′ NN potential with the $V_{C1}(\vec{r})$ Coulomb component is used [13]. The evaluation of the energy expectation values is explained in Ref. [14].

Table 1. AMD results for the ground-state energies E_0, rms radii $\langle r^2 \rangle^{1/2}$, and magnetic moments μ of the three- and four-nucleon systems. The experimental values are taken from reference [15].

	E_0 (MeV)		$\langle r^2 \rangle^{1/2}$ (fm)		μ (μ_N)	
	AMD	EXP	AMD	EXP	AMD	EXP
^{3}H $\left(\frac{1}{2}^+\right)$	-8.95	-8.48	1.33	1.60	2.769	2.979
^{3}He$\left(\frac{1}{2}^+\right)$	-8.61	-7.72	1.33	1.77	-1.847	-2.128
^{4}He$\left(0^+\right)$	-23.04	-28.30	1.16	1.47	0.000	

3. Ground state properties

The "frictional cooling" equations (7) were solved with the values $b \approx 30/\hbar$ and $\alpha = 0.12$. The α is chosen to reproduce, to reasonable accuracy, the ground state energies of light nuclei using only the parity-projected wave functions. The variational energy for the three- and four-nucleon systems was calculated using Eq. (8) and the results obtained are compared with experimental data in Table 1. The root-mean-square (rms) radii of the nuclei were calculate using the expression

$$\left\langle r^2 \right\rangle_{MK} = \frac{1}{A} \frac{\langle \Psi^{J\pm}_{MK} | \sum_{i=1}^{A} [\vec{r}_i - \vec{R}]^2 | \Psi^{J\pm}_{MK} \rangle}{\langle \Psi^{J\pm}_{MK} | \Psi^{J\pm}_{MK} \rangle} \tag{10}$$

where A is the number of nucleons in the nucleus and $\vec{R}$ the center-of-mass of the nucleus. The results obtained are also presented in Table 1.

As can be seen, the binding energy found for the ^{3}H system overestimates the experimental energy by 5 % and for the ^{3}He system by 12 % while . In contrast, the ^{4}He results are lower than the experimental value by $\sim$ 19 %, a result which is in line with other calculations in the field using the same rank in the potential and without the use of three-body forces. The rms radii obtained for the ^{3}H and ^{3}He systems are lower than the experimental values by $\sim$ 16 % and $\sim$ 22 % less, respectively. Similar results are obtained for the ^{4}He system where the calculated rms radius is underestimated by $\sim$ 21 %.

We turn now our attention to the magnetic moment of a nucleon which is given (in nuclear magnetons) by [16]

$$\mu = g_\ell \langle \vec{\ell} \rangle + g_s \langle \vec{s} \rangle \tag{11}$$

where $\langle \vec{\ell} \rangle$ ($\langle \vec{s} \rangle$) is the expectation value of the orbital (spin) angular momentum and g_ℓ (g_s) the orbital (spin) g-factor of the nucleon. The nucleon g-factors are constants, the values of which are [16]

$$g_\ell = \begin{cases} 1 & \text{for proton} \\ 0 & \text{for neutron} \end{cases} \quad : \quad g_s = \begin{cases} 5.585695 & \text{for proton} \\ -3.826085 & \text{for neutron} \end{cases} \tag{12}$$

The magnetic moment of the nuclei $\vec{\mu}_A$ is calculated from

$$\vec{\mu}^{\pm}_{MK} = \frac{\langle \Psi^{J\pm}_{MK} | \sum_{i=1}^{A} [\, g_\ell \, \vec{\ell}_i + g_s \, \vec{s}_i \,] | \Psi^{J\pm}_{MK} \rangle}{\langle \Psi^{J\pm}_{MK} | \Psi^{J\pm}_{MK} \rangle} \, . \tag{13}$$

The values of the magnitude $\mu^{\pm}_{MK} = |\vec{\mu}^{\pm}_{MK}|$ for the three- and four-nucleon systems are given in Table 1. In general, the AMD approach reproduces the experimental values for the magnetic moment of the nuclei quite satisfactorily.

4. Conclusions

To test the suitability of the AMD model in nuclear structure studies, the angular momentum and parity projected AMD wave function was used to calculate the binding energies, rms radii, and the magnetic moments for ^{3}H, ^{3}He, and ^{4}He few-nucleon systems. The nuclear Hamiltonian is constructed from the Argonne AV4′ nucleon-nucleon potential that includes also the Coulomb interaction.

Comparison with the experimental data revealed that the reproduction of the ground-state properties of light nuclei is quite satisfactory. The discrepancies observed can be attributed to reasons not related to the AMD. These include i) the omission of mixed-symmetric states (for three-body) ii) the use of a limited rank for the Argonne AV18 potential, and iii) the omission of three-nucleon forces. As far as the magnetic moment is concerned, the inclusion of relativistic corrections to the magnetic moment operator, are also expected to contribute to the reduction of the discrepancy between theory and experiment.

It short, we demonstrated that it is possible to construct a variational AMD wave function using realistic nucleon-nucleon potentials. The inclusion of three-nucleon forces in the Hamiltonian is also possible and this is expected to reduce the overall discrepancy between the experimental observation and theoretical AMD predictions of the nuclear properties. It should be noted that in nuclear systems described with spherical wave functions, like the ^{4}He nucleus, spatial rotations are not expected to introduce modifications to the results obtained with the parity-projected wave functions and therefore the use of the AMD method should be accurate.

Acknowledgments

This work is part of a PhD thesis submitted to the University of South Africa. The authors would like to thank Prof. S. Oryu (Tokyo University of Science, JAPAN) for useful discussions.

References

[1] Golak J, Skibiński R, Witala H, Glöckle W, Nogga A and Kamada H 2005 *Phys. Rep.* **C 415** 89
[2] Carlson J, Schiavilla R 1998 *Rev. Mod. Phys.* **70** 743
[3] Kamada H *et al* 2001 *Phys. Rev.* C **64** 044001
[4] Horiuchi H 1991 *Nucl. Phys.* **A 522** 257c
[5] Caurier E, Grammaticos B and Sami T 1982 *Phys. Lett.* **B 109** 150
[6] Feldmeier H 1990 *Nucl. Phys.* **A 515** 147
[7] Togashi T, Kat$\bar{o}$ K 2007 *Prog. Theor. Phys.* **117** 189
[8] Watanabe T and Oryu S 2006 *Prog. Theor. Phys.* **116** 429
[9] Kanada-En'yo Y, Horiuchi H and Ono A 1995 *Phys. Rev.* **C 52** 628
[10] Ono A, Horiuchi H, Maruyama T and Ohnishi A 1992 *Prog. Theor. Phys.* **87** 1185
[11] Peierls R E and Yoccoz J 1957 *Proc. Phys. Soc.* **A 70** 381
[12] Ono A and Horiuchi H 2004 *Prog. Part. Nucl. Phys.* **53** 501
[13] Wiringa R B and Pieper S C 2002 *Phys. Rev. Lett.* **89** 182501-1
[14] Rampho G J 2011 *Few-Body System* **50**, 467
[15] Suzuki Y, Horiuchi W, Orabi M and Arai K 2008 *Few-Body Systems.* **42** 33
[16] Wong S S M 1998 *Introductory Nuclear Physics*, 2nd Edition
(John Wiley and Sons Inc., New York)

Electromagnetic form factors of ^{3}H and ^{3}He systems

G J Rampho[1], S A Sofianos[1], S Oryu[2] and T Watanabe[2]

[1] Department of Physics, University of South Africa, Pretoria 0003, South Africa
[2] Department of Physics, Tokyo University of Science, Noda, Chiba 278-8510, Japan

E-mail: rampho@science.unisa.ac.za

Abstract. The angular-momentum-projected and parity-projected Antisymmetrized Molecular Dynamics (AMD) is used to analyze the charge and magnetic form factors of ^{3}H and ^{3}He systems. Non-relativistic nuclear charge and current operators with some relativistic corrections are employed. The Hamiltonian for the nuclear systems is constructed using a realistic nucleon-nucleon potential. The results obtained are in fairly good agreement with the experimental data and therefore the AMD method is a very promising method for use in calculations for electromagnetic form factors of a general A-body nuclear system.

1. Introduction

Charge and magnetic form factors are often used to test model wave functions for the quantum A-body nuclear systems. In such tests one faces the question of a simultaneous accurate description of the form factors and the static properties of the systems. This, of course, presupposes that the model used is able to generate accurately enough, the bound and scattering wave functions needed. In few-nucleon systems this can be achieved using microscopic methods, as for example, those using the Faddeev-type formalisms or methods based on Hyperspherical Harmonics expansion, to extract wave functions for realistic Hamiltonians [1]. These methods are considered, to all practical purposes, to be exact.

This work focuses on the use of a microscopic semi-quantum mechanical method, namely, the AMD method, to study the electromagnetic form factors for the ^{3}H and ^{3}He systems. The AMD approach [2] was developed from the Time-Dependent Cluster Model [3] used in studies of fermionic systems. This approach combines Fermi-Dirac statistics with quantum mechanical assumptions to treat the motion of particles in an A-body system [4] and was previously used to study the dynamics of heavy-ion collisions [5]. It was also used to explain clustering in nuclei as well as angular distributions of scattered protons in proton-nucleus scattering [6]. Furthermore, improved AMD wave functions were constructed that give good predictions in few-body systems [7, 8]. In this work the parity projected and angular momentum projected AMD version [9] is employed to extract the wave functions needed in the calculations.

In Sect. 2 the general formalism of the AMD approach is summarized and the construction of the wave function, the equations of motion of the variable parameters, and the variational technique used are briefly outlined. Results and illustrations of the application of AMD to three-nucleon systems are presented in Sect. 3 while Sect. 4 is devoted to the charge and magnetic form factors, respectively, which are to extract information about ground-state charge and magnetization distributions in the nuclei. Conclusions are drawn in Sect. 5.

2. The AMD formalism

The AMD wave function describing a nuclear system of A nucleons is constructed as a Slater determinant

$$\Psi_{AMD}(\vec{\mathrm{S}}) \;=\; \frac{1}{\sqrt{A!}}\det[\,\phi_j(\alpha,\vec{s}_i),\chi_j(\vec{\sigma}_i),\xi_j(\vec{\tau}_i)\,] \tag{1}$$

where ϕ, χ and ξ are, respectively, the spatial, spin and isospin components of the single-particle wave functions. The spatial components are given in Ref. [5] as Gaussians. The $\vec{s}_i$ are complex variational parameters, $\vec{\mathrm{S}} \equiv \{\vec{s}_1,\vec{s}_2,\vec{s}_3,\ldots,\vec{s}_A\}$ while α is a real constant width parameters The width parameter is a free parameter and taken as common for all the Gaussian functions. The spin-isospin components are fixed in nucleon spin-up or spin-down. A wave function with definite parity(π) and total angular momentum (J) with the angular momentum projection (M) is constructed from the AMD wave function as

$$\Psi^{J\pi}_{MK}(\vec{\mathrm{S}}) \;=\; \frac{1}{2}\,P^{J}_{MK}(\Omega)\,[\,1\pm P^{\pi}\,]\,\Psi_{AMD}(\vec{\mathrm{S}}) \tag{2}$$

where P^{π} is the parity projection operator and $P^{J}_{MK}(\Omega)$ the angular momentum projection operator defined by [10]

$$P^{J}_{MK}(\Omega) \;=\; \frac{2\,J+1}{8\,\pi^2}\int \mathrm{d}\Omega\, D^{J*}_{MK}(\Omega)\,\hat{R}(\Omega). \tag{3}$$

Here $D^{J}_{MK}(\Omega)$ represents the Wigner D-function, $\hat{R}(\Omega)$ the rotation operator and $\Omega \equiv \{\alpha,\beta,\gamma\}$ the Euler rotation angles.

The time-dependent variational principle [5]

$$\delta\int_{t_1}^{t_2}\frac{\langle\,\Psi(\vec{S})\,|\,\mathrm{i}\hbar\,\frac{\partial}{\partial t}-H\,|\,\Psi(\vec{S})\,\rangle}{\langle\,\Psi(\vec{S})\,|\,\Psi(\vec{S})\,\rangle}\,\mathrm{d}t=0 \tag{4}$$

with constraints

$$\delta\Psi(t_1)=\delta\Psi(t_2)=\delta\Psi^{*}(t_1)=\delta\Psi^{*}(t_2)=0\;\;. \tag{5}$$

is used to determine the dynamical equations for the variational parameters. The resulting equations can be transformed to the form [11]

$$\frac{\mathrm{d}\vec{s}_i}{\mathrm{d}t}=-\mu\,\frac{\partial E^{J\pm}_0(\vec{S},\vec{S}^{*})}{\partial\vec{s}^{*}_i}\qquad,\qquad\frac{\mathrm{d}\vec{s}^{*}_i}{\mathrm{d}t}=-\mu\,\frac{\partial E^{J\pm}_0(\vec{S},\vec{S}^{*})}{\partial\vec{s}_i} \tag{6}$$

where μ is an arbitrary positive real constant and

$$E^{J\pm}_0(\vec{S},\vec{S}^{*}) \;=\; \frac{\langle\,\Psi^{J\pm}_{MK}(\vec{\mathrm{S}})|\,H\,|\Psi^{J\pm}_{MK}(\vec{\mathrm{S}})\,\rangle}{\langle\,\Psi^{J\pm}_{MK}(\vec{\mathrm{S}})|\Psi^{J\pm}_{MK}(\vec{\mathrm{S}})\,\rangle}\;\;. \tag{7}$$

the variational energy of the nucleus. Solving these equations minimizes $E^{J\pi}_0$ and determines the variational parameters. The Hamiltonian of the system is given by

$$H \;=\; -\sum_i\frac{\hbar^2}{2\,m_i}\,\nabla^2_i+\frac{1}{2}\sum_{i\neq j}\Big[\,V_{NN}(\vec{r}_{ij})+V_C(\vec{r}_{ij})\,\Big] \tag{8}$$

where m_i is the mass of nucleon i, V_{NN} the two-body nuclear potential and V_C the Coulomb potential. In this work the AV4' NN potential with the $V_{C1}(r)$ Coulomb component is used [12]. The evaluation of the components of the energy expectation values is explained in Ref. [13].

3. Charge form factor

In elastic electron-nucleus scattering, the charge distribution in the nucleus is inferred from the induced electric transitions in the nucleus. The charge form factor is the expectation value of the nuclear charge operator. For a nucleus in an initial state $\mid \Psi^{J_i\pi_i} \rangle$ the charge form factor is given by

$$F_{\rm ch}(\vec{q}) = \frac{1}{Z} \frac{\langle \Psi^{J_i\pi_i} \mid \rho(\vec{q}) \mid \Psi^{J_i\pi_i} \rangle}{\langle \Psi^{J_i\pi_i} \mid \Psi^{J_i\pi_i} \rangle} \tag{9}$$

where Ze is the total charge on and $\rho(\vec{q})$ the charge operator of the nucleus with $\vec{q}$ being the momentum transferred to the nucleus by the electron.

In the plane wave impulse approximation (PWIA) the nuclear charge operator is formed by the superposition of the individual nucleon charge operators and is given by [13]

$$\rho(\vec{q}) = \sum_{k=1}^{A} \left[\frac{q}{Q} G_{Ek}^N(Q^2) - \frac{2\, G_{Mk}^N(Q^2) - G_{Ek}^N(Q^2)}{4\, m_N^2 \sqrt{1+\tau}} \, {\rm i}\vec{\sigma}_k \cdot \vec{q} \times \vec{p}_k \right] \exp(\, {\rm i}\, \vec{q} \cdot \vec{r}_k \,) \tag{10}$$

where $\tau = Q^2/4\, M_N^2$, $Q^2 = q^2 - \omega^2$, $\omega = \sqrt{q^2 + m_N^2} - m_N$ and G_E^N (G_M^N) the nucleon Sachs electric (magnetic) form factor. For the Sachs form factors the phenomenological parameterization derived in Ref. [14] is adopted. The transitions are between states of definite angular momentum. The general multipole analysis of nuclear charge form factors is given by [15]

$$F_{\rm ch}(\vec{q}) = \sqrt{4\,\pi} \sum_{L=0}^{\leq 2J} \langle JJL0|JJ\rangle \, F_L^{\rho}(q) \, Y_{L0}^*(\hat{\mathbf{q}}) \tag{11}$$

where $Y_{LM}^*(\hat{\mathbf{q}})$ are the spherical harmonics, L the nuclear orbital angular momentum and $\langle JJL0|JJ\rangle$ the Clebsch-Gordan coefficients. The summation is over even values of L only. For three-nucleon systems $J_i = \frac{1}{2}$. The intrinsic charge form factor is obtained by dividing the calculated charge form factor by the contributions of the center-of-mass [13].

Using the above definitions, the ground-state charge form factors of the ^{3}H and ^{3}He nuclei are calculated within the PWIA approximation. In this approximation the nucleons inside the target nucleus are assumed non-interacting with one another during the interaction with the electron [17]. This means that the electron interacts with independent nucleons inside the nucleus. The results obtained are displayed in Fig. 1 and Fig. 2. It is noted that the charge form factors are normalized such that $F_{\rm ch}(0) = 1$. As can be seen in these figures, for low momentum transfers, up to the first diffraction minimum, the AMD gives a good description, albeit it slightly overestimates the experimental data. Beyond the first diffraction minimum the results are lower than the data. The first diffraction minimum for the ^{3}H and ^{3}He nuclei are consistent with the predictions of other theoretical models obtained with various nucleon-nucleon potentials [18]. It should be noted that the overestimation of the position of the diffraction minimum indicates an underestimation of the nuclear charge radius.

4. Magnetic form factors

Magnetization density distribution in nuclei are determined from magnetic transitions involving transverse nuclear currents. In the final state of the nucleus is denoted by $|\Psi_{M_f K_f}^{J_f\pi_f}\rangle$, with normalization $\mathcal{N}_{M_f K_f}^{J_f\pi_f}$, the nuclear magnetic form factor is calculated as

$$F_{\rm mag}(\vec{q}) = \frac{1}{\mu_A} \frac{\langle \Psi_{M_f K_f}^{J_f\pi_f} \mid \vec{\mu}(\vec{q}) \mid \Psi_{M_i K_i}^{J_i\pi_i} \rangle}{\sqrt{\mathcal{N}_{M_f K_f}^{J_f\pi_f} \, \mathcal{N}_{M_i K_i}^{J_i\pi_i}}} \tag{12}$$

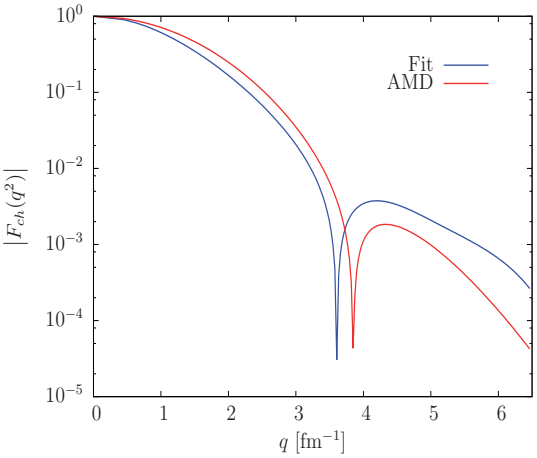


Figure 1. The AMD charge form factor of the ^{3}H nucleus compared with the best fit of the experimental data [16].

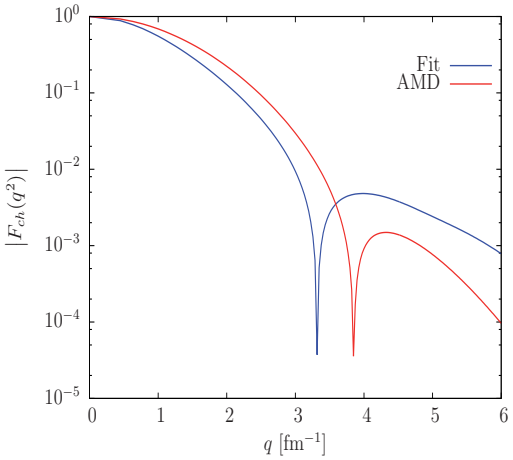


Figure 2. The AMD charge form factor of the ^{3}He nucleus compared with the best fit of the experimental data [16].

where $\vec{\mu}(\vec{q})$ is the magnetization density operator and μ_A the nuclear magnetic dipole moment. The PWIA transverse nuclear magnetization density operator is given by [19]

$$\vec{\mu}(\vec{q}) = \frac{Q}{2\,m_p\,q}\sum_{k=1}^{A}\Big[\,G^N_{Ek}(Q^2)\,\vec{\ell}_k - \mathrm{i}\,G^N_{Mk}(Q^2)\,\vec{q}\times\vec{\sigma}\,\Big]\exp(\,\mathrm{i}\,\vec{q}\cdot\vec{r}_k\,) \tag{13}$$

where $\vec{\ell}_N$ is the nucleon orbital angular momentum. The multipole expansion of nuclear magnetic form factor has the form [15]

$$F_{\rm mag}(\vec{q}) = \frac{\sqrt{4\,\pi}}{\langle JJ10|JJ\rangle}\sum_{L=0}^{\leq 2J}\langle JJL0|JJ\rangle\Big[\,F^{\mu}_{LL-1}(q)\,\mathbf{Y}^{0*}_{LL-1}(\hat{\mathbf{q}}) + F^{\mu}_{LL+1}(q)\,\mathbf{Y}^{0*}_{LL+1}(\hat{\mathbf{q}})\,\Big] \tag{14}$$

where the summation is over odd values of L,

$$\mathbf{Y}^{0*}_{LM}(\hat{\mathbf{q}}) = \sum_m \langle Mm1\,{-m}|L0\rangle\,Y_{Mm}(\hat{\mathbf{q}})\,\hat{\mathbf{e}}_m\,. \tag{15}$$

are the vector spherical harmonics and $\hat{\mathbf{e}}_m$ spherical unit vectors.

The general form of the nuclear magnetic transition multipole operator can be derived as in Ref. [20] for a given nuclear current operator. The intrinsic magnetic form factor of the systems is obtained by factoring out the contributions of the center-of-mass from Eq. (14). The results obtained for the magnetic form factors for three-nucleon systems are displayed in Fig. 3 and Fig. 4 and compared with the best fit to the experimental data [16]. Similarly to the charge form factors, the magnetic form factors are normalized such that $F_{\rm mag}(0) = 1$. As can be seen, the AMD form factors reproduce the experimental data reasonably well at low momentum transfer, for both systems. The position of the first diffraction minimum is underestimated. However, at momentum transfer greater than the diffraction minimum the AMD results are unsatisfactory and are not consistent with theoretical results obtained with other approaches [21].

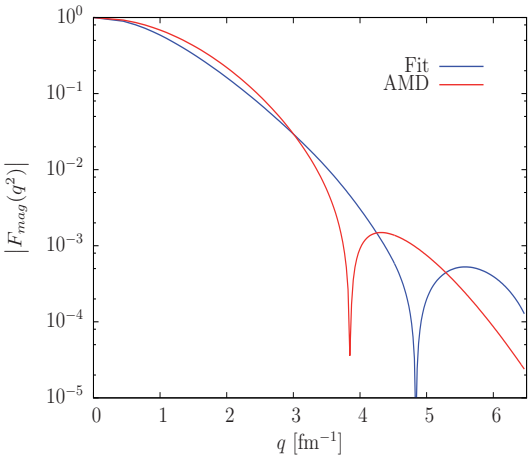


Figure 3. The AMD magnetic form factor of the ^{3}H nucleus compared with the best fit of the experimental data [16].

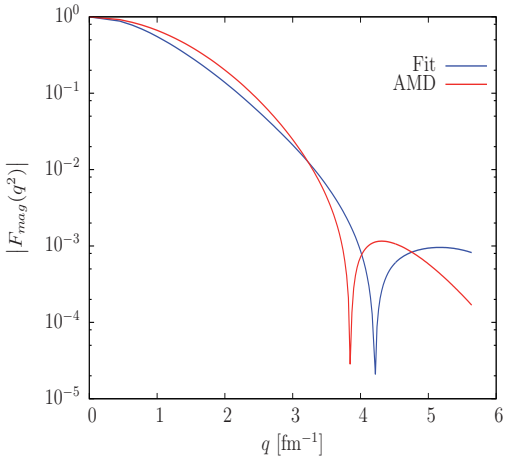


Figure 4. The AMD magnetic form factor of the ^{3}He nucleus compared with the best fit of the experimental data [16].

5. Conclusions

We employ the angular-momentum-projected and parity-projected antisymmetrized molecular dynamics to calculate the charge and magnetic form factors for the ^{3}H and ^{3}He systems. The charge monopole and the magnetic dipole transitions in the nuclei were determined from elastic electron scattering. In overall, the results obtained, within the AMD and PWIA approximation, reproduce the general behavior of the experimental form factors. For both, the charge and magnetic form factors, and for momentum transfer below the first diffraction minimum the reproduction is fairly good. However, beyond the first diffraction minimum the results for the charge form factor are lower than the data while the results for the the magnetic form factor are less satisfactory. The deviations of the theoretical results from experimental data can be minimize once better wave functions can be employed. These can be constructed by using a more complete realistic Hamiltonian, three-body forces, and relativistic corrections in the electromagnetic operators used. It should be noted that these results are consistent with other results obtained by competing theoretical models.

In conclusion, the results indicate that the AMD method is a very promising method in calculating electromagnetic form factors of the general A-body nuclear system.

Acknowledgments

This work is part of a PhD thesis submitted to the University of South Africa.

References

[1] Kamada H *et al*, 2001 *Phys. Rev.* C**64** 044001
[2] Horiuchi H, 1991 *Nucl. Phys.* **A 522** 257c
[3] Caurier E, Grammaticos B, and Sami T, 1982 *Phys. Lett.* **B 109** 150
[4] Feldmeier H, 1990 *Nucl. Phys.* **A 515** 147
[5] Ono A, Horiuchi H, Maruyama T and Ohnishi A, 1992 *Prog. Theor. Phys.* **87** 1185
[6] Tanaka E I, Ono A, Horiuchi H, Maruyama T, and Engel A, 1995 *Phys. Rev.* C **52** 316
[7] Togashi T, Kat$\bar{o}$ K 2007 *Prog. Theor. Phys.* **117** 189
[8] Watanabe T and Oryu S, 2006 *Prog. Theor. Phys.* **116** 429
[9] Kanada-En'yo Y, Horiuchi H, and Ono A, 1995 *Phys. Rev.* C**52** 628
[10] Peierls R E and Yoccoz J, 1957 *Proc. Phys. Soc.* **A 70** 381
[11] Ono A and Horiuchi H, 2004 *Prog. Part. Nucl. Phys.* **53** 501
[12] Wiringa R B and Pieper S C, 2002 *Phys. Rev. Lett.* **89** 182501-1

[13] Rampho G J, 2010 *PhD Thesis*, University of South Africa, unpublished.
[14] Friedrich J and Walcher Th, 2003 *Eur. Phys. J.* **A 17** 607
[15] Überall H, 1971 *Electron Scattering from Complex Nuclei*, Part A (Academic Press, New York)
[16] Amroun A, Breton V, Cavedon J -M, Frois B, Goutte D, F. P. Juster F P, Leconte Ph, Martino J, Mizuno Y, Phan X -H, Platchkov S K, Sick I, Williamson S, 1994 *Nucl. Phys.* **A 579**, 596
[17] Chew G F and Wick G C, 1952 *Phys. Rev.* **85** 636
[18] Kloet W M and Tjon J A, 1974 *Phys. Lett.* **B 49** 419
[19] Ciofi degli Atti C, 1980 *Prog. Part. Nucl. Phys.* **3** 163
[20] Willey R S, 1963 *Nucl. Phys.* **40** 529
[21] Hadjimichael E, Goulard B, and Bornais R, 1983 *Phys. Rev.* C**27** 831

Unravelling of open quantum random walks

Ilya Sinayskiy and Francesco Petruccione

Quantum Research Group and National Institute for Theoretical Physics, School of Physics, University of KwaZulu-Natal, Durban, South Africa, 4001

E-mail: sinayskiy@ukzn.ac.za, petruccione@ukzn.ac.za

Abstract. Recently, in joint work of S. Attal, C. Sabot, F. Petruccione and I. Sinayskiy the concept of Open Quantum Random Walks was introduced, by taking into account dissipation and decoherence that occur in open quantum systems. Open quantum random walks are formulated in terms of discrete completely positive maps for the density matrix. These walks are simulated efficiently with the help of quantum trajectories. Here we report the unravelling of some simple open quantum random walk.

1. Introduction

It is well known that random walks are a very useful mathematical tool, which found successful applications in physics, computer science, economics and biology. The Unitary Quantum Random Walk (UQRW) [1, 2] generalized this concept to the quantum domain [3] and widely applied in quantum computing [4]. Recently, experimental realizations of UQRW have been reported [5, 6, 7, 8]. For the last few years attempts were made to take into account decoherence and dissipation in the quantum walks formalism [9]. But, in all these approaches decoherence and dissipation is treated as a modification of the unitary quantum walk scheme, the destructive influence of which needs to be minimized and eliminated. The general framework of quantum stochastic walks was proposed [10], which incorporates unitary and non-unitary effects of the quantum Markovian dynamics.

Recently, a formalism for discrete time quantum walks was introduced [11], that naturally includes the behavior of open quantum systems. The formalism suggested is similar to the formalism of quantum Markov chains [12] and rests upon the implementation of appropriate completely positive maps [13].

In this paper, we will briefly review the concept of the OQRWs and for a particular example we will show the simulation of an OQRW in terms of quantum trajectories.

2. Open quantum random walks

The dynamics of a walker with internal degrees of freedom in the Hilbert space $\mathcal{H}$ on the nodes (i, j) (i, j are elements of a finite or infinite countable set) of a graph is defined in terms of operators $B^i_j \in \mathcal{H}$. These operators describe the transformation in the internal degree of freedom of the walker due to the jump from node j to node i. The basic idea of the OQRW is to assume that for each j

$$\sum_i B^{i\dagger}_j B^i_j = I,$$

where $B_i^j \in \mathcal{H}$ and I denotes the identity in the appropriate space. This constraint has to be understood in the following way: the sum of all the effects leaving site j is I. The Hilbert space of states specified by the set of nodes will be denoted by $\mathcal{K}$ and will be assumed to have a basis $|i\rangle$. Obviously, this construction is a natural generalization of the "classical" Markov chain concept. In order to describe not only the change in the internal degrees of freedom of the walker, but also the transitions from node j to node i it is convenient to introduce the operators $M_j^i = B_j^i \otimes |i\rangle\langle j|$, which satisfy $\sum_{i,j} M_j^{i\dagger} M_j^i = I$. The OQRW can now be defined in terms of the following completely positive map on $\mathcal{H} \otimes \mathcal{K}$

$$\mathcal{M}(\rho) = \sum_i \sum_j M_j^i \, \rho \, M_j^{i\dagger}.$$

If we assume the initial density matrix of the system to be of the form

$$\rho = \sum_i \rho_i \otimes |i\rangle\langle i|,$$

with $\sum_i \mathrm{Tr}\rho_i = 1$, the iteration formula for the OQRW from step n to step $n+1$ can be expressed as

$$\rho^{[n+1]} = \mathcal{M}(\rho^{[n]}) = \sum_i \rho_i^{[n+1]} \otimes |i\rangle\langle i|,$$

where $\rho_i^{[n+1]} = \sum_j B_j^i \rho_j^{[n]} B_j^{i\dagger}$.

3. Unravelling of the OQRW in terms of quantum trajectories

To introduce the quantum trajectories formalism we start by considering a particular case of initial state of the system, namely, a walker that is localized at a single site,

$$\rho_0 = \rho \otimes |i\rangle\langle i|.$$

After one step the state of the walker will be given by

$$\rho_1 = \sum_j (B_i^j \rho B_i^{j\dagger}) \otimes |j\rangle\langle j| \, .$$

The probability to find the walker at the site j is given by $p_j = \mathrm{tr}(B_i^j \rho B_i^{j\dagger})$. If one performs measurements of the position of the walker at site j the state of the walker reads

$$\frac{1}{p_j}(B_i^j \rho B_i^{j\dagger}) \otimes |j\rangle\langle j| \, .$$

Repetition of this procedure gives rise to a classical Markov chain, valued in the set of states of the form $\rho \otimes |i\rangle\langle i|$. One can see that this procedure on average will simulate a master equation driven by $\mathcal{M}$:

$$\mathcal{E}[\rho_{n+1}] = \sum_j p_j \frac{1}{p_j}(B_i^j \rho_n B_i^{j\dagger}) \otimes |j\rangle\langle j| = \sum_j (B_i^j \rho_n B_i^{j\dagger}) \otimes |j\rangle\langle j| = \mathcal{M}(\rho_n) \, .$$

In particular, if the initial state of the system is the pure state $\rho = |\phi\rangle\langle\phi| \otimes |i\rangle\langle i|$, then the system remains in a pure state. It is easy to see that, any initial pure state $|\phi\rangle \otimes |i\rangle$ will jump randomly to one of the states

$$\frac{1}{\sqrt{p_i^j}} B_i^j |\phi\rangle \otimes |j\rangle$$

with probability

$$p_i^j = ||B_i^j|\phi\rangle||^2 .$$

We have a classical Markov chain valued in the space of wave functions of the form $|\phi\rangle \otimes |i\rangle$. On average, this random walk simulates the master equation driven by $\mathcal{M}$.

As illustration of this unravelling we consider an open quantum random walk on a line. In this case the transition matrices B_i^j are chosen to be,

$$B_i^{i-1} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix}, \quad B_i^{i+1} = \begin{pmatrix} 0 & \frac{1}{2} \\ 0 & 0 \end{pmatrix},$$

and the initial state of the walker is

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}(|10\rangle - |01\rangle) \otimes |0\rangle.$$

The results of the simulations are presented in figure 1. Figure 1(a),(b) and (c) we show three different quantum trajectories. One can clearly see the different qualitative character of some of them. The average over 1000 realizations is shown in the figure 1(d).

More work is in progress in applying the open quantum random walk formalism to quantum computing and quantum state transfer problems.

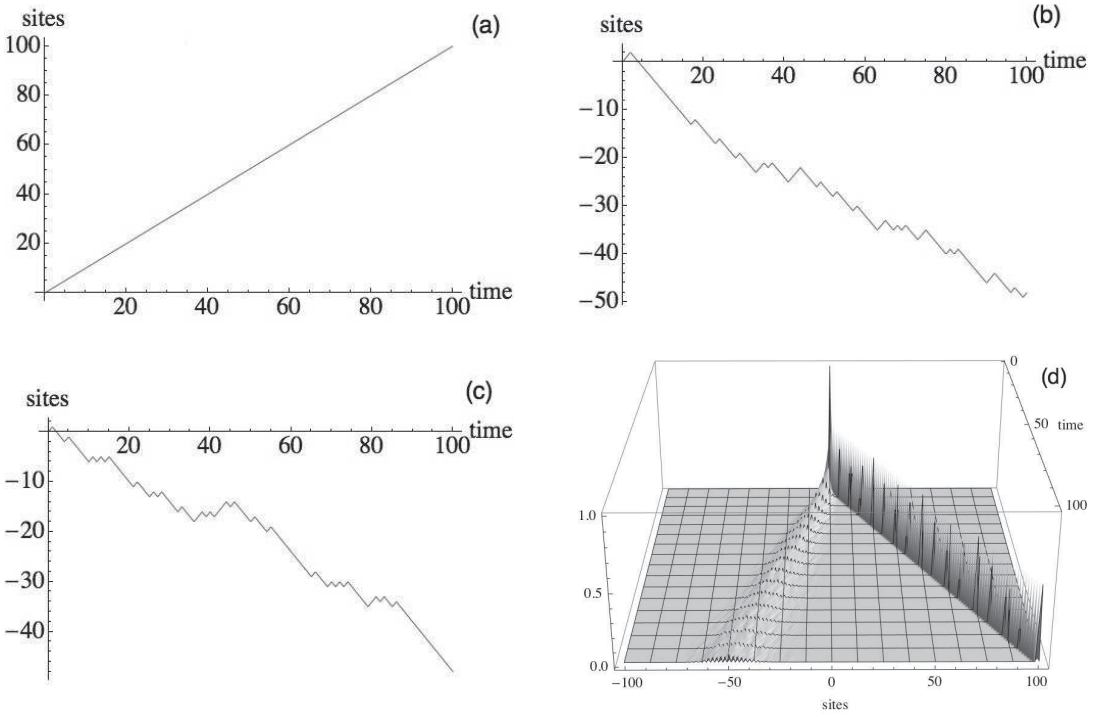


Figure 1. Simulation of the OQRW in term of quantum trajectories. The examples of the trajectories are shown on the curves (a)-(c), curve shows (d) average over 1000 trajectories

Acknowledgments

This work is based upon research supported by the South African Research Chair Initiative of the Department of Science and Technology and National Research Foundation.

References

[1] Aharonov Y, Davidovich L and Zagury N 1993 *Phys. Rev.* A **48** 1687
[2] Farhi E and Gutmann S 1998 *Phys. Rev.* A **58** 915
[3] Kempe J 2003 *Contemp. Phys.* **44** 307
[4] Aharonov D, Ambainis A, Kempe J and Vazirani U 2001 *in Proceedings of the 33rd ACM Symposium on Theory of Computing* 50
Konno N 2008 *Quantum Walks in "Quantum Potential Theory" LNM* 309
Childs A, Farhi E and Gutmann S 2002 *Quant. Info. Proc.* **1** 35
Watrous J 2001 *J. Comput. Syst. Sci.* **62** 376
Shenvi N, Kempe J and Whaley KB 2003 *Phys. Rev.* A **67** 052307
[5] Karski M, Föster L, Choi J-M, Steffen A, Alt W, Meschede D and Widera A 2009 *Science* **325** 174
[6] Perets H B, Lahini Y, Pozzi F, Sorel M, Morandotti R and Silberberg Y 2008 *Phys. Rev. Lett.* **100** 170506
[7] Schmitz H, Matjeschk R, Schneider Ch, Glueckert J, Enderlein M, Huber T and Schaetz T 2009 *Phys. Rev. Lett.* **103** 090504
Zähringer F, Kirchmair G, Gerritsma R, Solano E, Blatt R and Roos C F 2010 *Phys. Rev. Lett.* **104** 100503
[8] Broome M A, Fedrizzi A, Lanyon B P, Kassal I, Aspuru-Guzik A and White A G 2010 *Phys. Rev. Lett.* **104** 153602
[9] Kendon V 2007 *Math. Struct. Comput. Sci.* **17** 1169
Brun T A, Carteret H A and Ambainis A 2003 *Phys. Rev. Lett.* **91** 130602
Srikanth R, Banerjee S and Chandrashekar C M 2010 *Phys. Rev.* A **81** 062123
[10] Whitfield J D, Rodríguez-Rosario C A and Aspuru-Guzik A 2010 *Phys. Rev.* A **81** 022323
[11] Attal S, Petruccione F, Sabot C and Sinayskiy I 2011 Open Quantum Random Walks (in preparation)
Attal S, Petruccione F, Sabot C and Sinayskiy I 2011 Open Quantum Random Walks on Graphs (in preparation)
[12] Gudder S 2010 *Found. Phys.* **40** N9-10 1566
[13] Breuer H-P and Petruccione F 2002 *The Theory of Open Quantum Systems* (Oxford: Oxford University Press)

A simple model of decoherence-assisted transport

Ilya Sinayskiy, Adriana Marais and Francesco Petruccione

Quantum Research Group and National Institute for Theoretical Physics, School of Physics, University of KwaZulu-Natal, Durban, 4001, South Africa

E-mail: sinayskiy@ukzn.ac.za, 207521013@ukzn.ac.za, petruccione@ukzn.ac.za

Abstract. The dynamics of a dimer under the influence of decoherence are studied. The model consists of a dimer with one level in contact with a single bosonic mode. An exact analytical expression for the density matrix of the total system is obtained. It is shown that there exist well-defined ranges of parameters for which decoherent interaction with just a single bosonic mode assists energy transfer in the dimer system. This simple and analytically solvable model for energy transfer in a dimer system is easily extendible to more complex quantum networks, and more complex environmental models.

1. Introduction

The transfer of energy and information in quantum networks plays an important role for quantum communication and quantum computation. In realistic physical situations, unavoidable interaction with the environment leads to decoherence and dissipation, processes typically associated with a destruction of coherence [1]. However, recently ultrafast spectroscopic techniques have been claimed to reveal long-lasting quantum coherence in biological systems, including in the photosynthetic light-harvesting complexes of a species of green sulphur bacteria [2,3], a species of purple bacteria [4], and two species of marine cryptophyte algae [5].

The efficiency of the energy transfer through the network of chromophores together with the evidence for quantum coherence has led to much activity proposing mechanisms for environment-assisted excitonic energy transport in quantum networks, including under the broad headings of noise-assisted transport [6-9] and oscillation-enhanced transport [10-13]. The possibility that quantum entanglement may enhance transport has also been discussed [14-16].

We show that even in the most simple analytically solvable model of a dimer coupled to a single bosonic mode, there exist well-defined ranges of parameters for which decoherent interaction with the environment assists energy transfer in the dimer system.

2. Model

In this paper we are going to study the simplest electronic energy transfer system, the dimer. The Hamiltonian of the dimer is given by $H_D = \varepsilon_1|1\rangle\langle 1| + \varepsilon_2|2\rangle\langle 2| + J(|1\rangle\langle 2| + |2\rangle\langle 1|)$, where ε_i are the energy levels of the dimer and J is the amplitude of transition. It is well-known that if the initial excitation is in level 1, then the maximum probability of transition to level 2 ($P_{1\to 2}(t)$) will be given by $\max(P_{1\to 2}) = J^2/(J^2 + \Delta^2)$, where Δ is half of the energy difference between the energy levels of the dimer ($\Delta = (\varepsilon_2 - \varepsilon_1)/2$). This means that only in the case $\varepsilon_1 = \varepsilon_2$

we can say that the excitation is transferred with certainty ($P_{1\to 2}(t_0) = 1$) at the time $t_0 = \pi/2J$.

The aim of this paper is to show that in the general case ($\varepsilon_1 \neq \varepsilon_2$), even for the most simple analytically solvable model consisting of a dimer with one level in contact with a single bosonic mode, decoherence enhances energy transfer.

The Hamiltonian of the total system has the form

$$H = H_D + H_{BM} + H_I,$$

where in the Hamiltonian for the dimer H_D given above, we take $\varepsilon_2 > \varepsilon_1$, and the Hamiltonians of the the bosonic mode and the interaction are given by

$$\begin{aligned} H_{BM} &= \omega a^\dagger a, \text{ and} \\ H_I &= \gamma a^\dagger a |1\rangle\langle 1|, \end{aligned}$$

respectively.

The projectors $|i\rangle\langle j|$ are conveniently represented in terms of the Pauli matrices as

$$|1\rangle\langle 1| = \frac{\mathbb{1}_2 - \sigma_z}{2}, \quad |2\rangle\langle 2| = \frac{\mathbb{1}_2 + \sigma_z}{2}, \quad |2\rangle\langle 1| = \sigma^+.$$

Thus, the total Hamiltonian can be written in the following form:

$$H = \frac{\varepsilon_2 - \varepsilon_1}{2}\sigma_z + J\sigma_x + \omega a^\dagger a - \frac{\gamma}{2}a^\dagger a\sigma_z + \frac{\varepsilon_2 + \varepsilon_1}{2}\mathbb{1}_2 + \frac{\gamma}{2}a^\dagger a.$$

It is easy to find the evolution operator of the system, which is given by

$$U(t) = \sum_{n=0}^{\infty} \mathrm{e}^{-it[(\omega+\frac{\gamma}{2})n + \frac{\varepsilon_1+\varepsilon_2}{2}]} \begin{pmatrix} \cos\Omega_n t - \frac{i\Delta_n}{\Omega_n}\sin\Omega_n t & -i\frac{J}{\Omega_n}\sin\Omega_n t \\ -i\frac{J}{\Omega_n}\sin\Omega_n t & \cos\Omega_n t + \frac{i\Delta_n}{\Omega_n}\sin\Omega_n t \end{pmatrix} \otimes |n\rangle\langle n|,$$

where $\Omega_n = \sqrt{J^2 + \Delta_n^2}$ and $\Delta_n = (\varepsilon_2 - \varepsilon_1)/2 - \gamma n/2$.

Since we are interested in the probability of transition $P_{1\to 2}(t)$ in the dimer, we assume the dimer is initially in the state $|\psi_D^0\rangle = |1\rangle$, while the initial state of the bosonic mode is the thermal state

$$\rho_{\mathrm{BM}}^0 = \frac{1}{Z}\mathrm{e}^{-\beta\omega a^\dagger a} = (1 - \mathrm{e}^{-\beta\omega})\mathrm{e}^{-\beta\omega a^\dagger a}.$$

The state of the dimer at time t is then given by

$$\rho_D(t) = \mathrm{tr}_{\mathrm{BM}}(U(t)|1\rangle\langle 1| \otimes \rho_{\mathrm{BM}}^0 U^\dagger(t)).$$

The probability of transition is then found to be

$$\begin{aligned} P_{1\to 2}(t) &= \mathrm{tr}_{\mathrm{BM}}(\langle 2|U(t)|1\rangle \rho_{BM}^0 \langle 1|U^\dagger(t)|2\rangle) \\ &= \sum_{n=0}^{\infty}(1 - \mathrm{e}^{-\beta\omega})\frac{J^2}{\Omega_n^2}\sin^2\Omega_n t \mathrm{e}^{-\beta\omega n} \\ &= (1 - \mathrm{e}^{-\beta\omega})\sum_{n=0}^{\infty}\frac{J^2}{J^2 + (\frac{\varepsilon_2-\varepsilon_1}{2} - \frac{\gamma n}{2})^2}\sin^2\left(t\sqrt{J^2 + (\frac{\varepsilon_2 - \varepsilon_1}{2} - \frac{\gamma n}{2})^2}\right)\mathrm{e}^{-\beta\omega n}. \end{aligned}$$

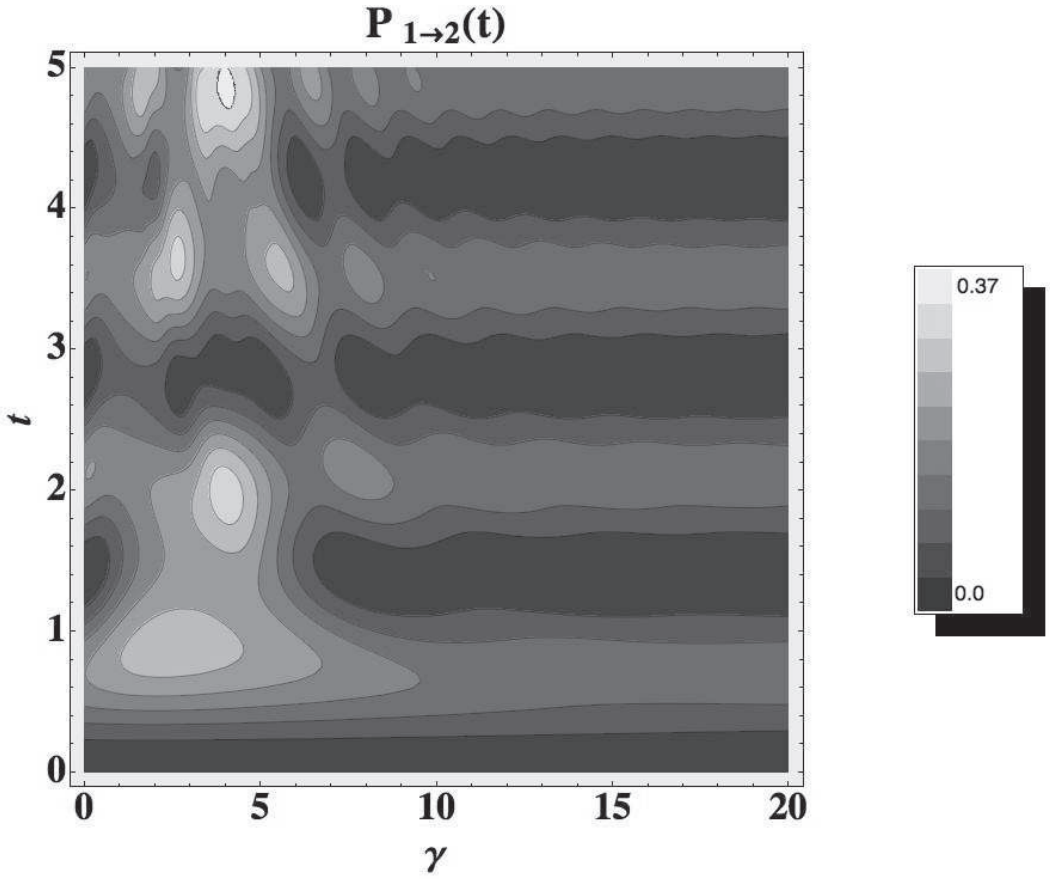


Figure 1. Probability of transition $P_{1\to 2}(t)$ in the dimer system coupled to just one bosonic mode as a function of time and coupling constant γ to the bosonic mode. The parameters are chosen to be $J = 1$, $\Delta = (\varepsilon_2 - \varepsilon_1)/2 = 2$, $\beta = 1$ and $\omega = 1$.

In figure 1 and figure 2 the probability of transition in the dimer system as a function of time and coupling constant to the single bosonic mode is presented. One can see that the maximum of the probability of transition is achieved for the non-zero interaction with a decoherent environment. This means that in the general case of non-zero temperature, decoherence as a result of coupling to even a single bosonic mode assists energy transfer in the dimer system. Note that in figure 2 the temperature is higher by a factor of 10, which means that the effect is present in a wide range of temperatures.

3. Conclusion

In conclusion, we have studied a dimer system in contact with a decoherent environment consisting of just a single bosonic mode. We have shown that there exist parameter ranges for which interaction with the bosonic mode assists energy transfer in a dimer. These results motivate further study of more complicated energy transfer networks and more complicated models of environments.

Acknowledgments

This work is based upon research supported by the South African Research Chair Initiative of the Department of Science and Technology and National Research Foundation.

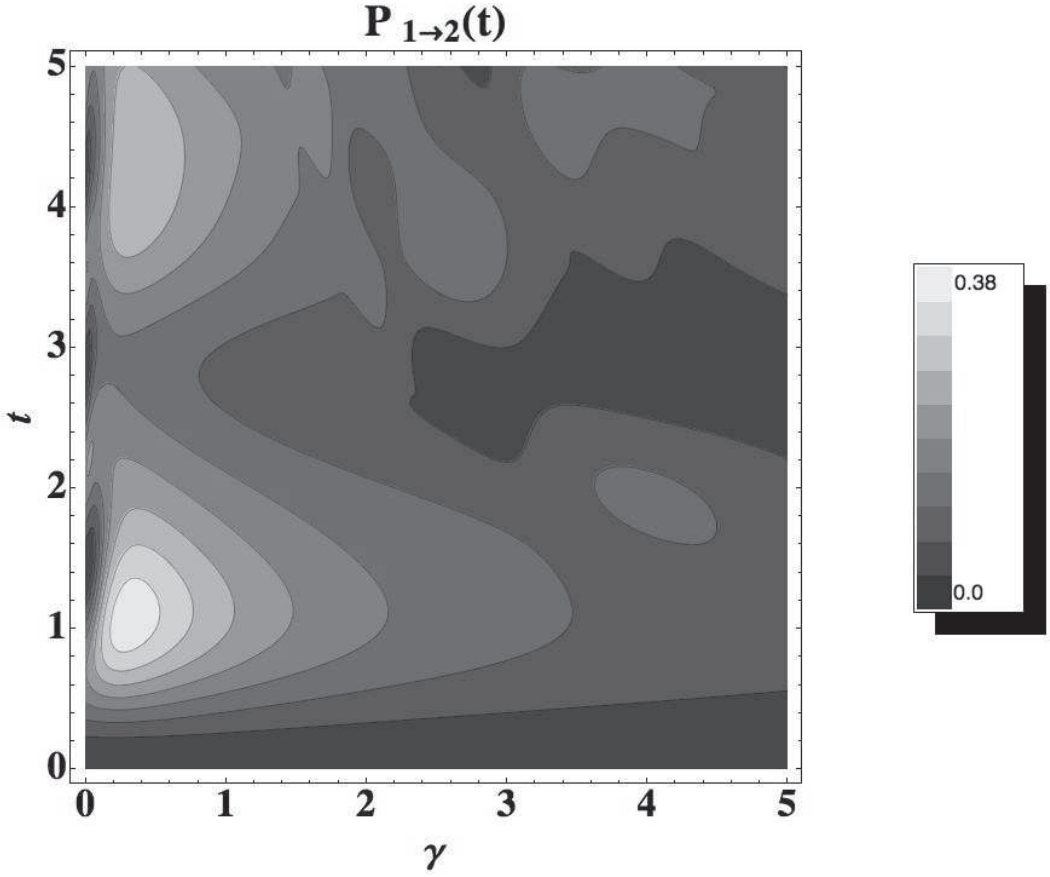


Figure 2. Probability of transition $P_{1\to 2}(t)$ in the dimer system coupled to just one bosonic mode as a function of time and coupling constant γ to the bosonic mode. The parameters are the same as in Fig. 1, except with higher temperature, namely $\beta = 0.1$.

References

[1] Breuer H-P and Petruccione F 2002 *The Theory of Open Quantum Systems* (Oxford: Oxford University Press)
[2] Engel G S, Calhoun T R, Read E L, Ahn T-K, Mancal T, Cheng Y-C, Blankenship R E and Fleming G R 2007 *Nature* **446** 782
[3] Panitchayangkoon G, Hayes D, Fransted K A, Caram J R, Harel E, Wen J, Blankenship R E and Engel G S 2010 *Proc. Natl. Acad. Sci. USA* **107** 12766
[4] Lee H, Cheng Y-C and Fleming G R 2007 *Science* **316** 1462
[5] Collini E, Wong C Y, Wilk K E, Curmi P M G, Brumer P and Scholes G D 2010 *Nature* **463** 08811
[6] Grabert H, Weiss U and Hänggi P 1984, *Phys. Rev. Lett.* **52** 2193
Martinis J M and Grabert H 1988, *Phys. Rev.* B **38** 2371
Weiss U 2008 *Quantum Dissipative Systems* 3rd ed. (Singapore: World Scientific Press)
[7] Mohseni M, Rebentrost P, Lloyd S and Aspuru-Guzik A 2008 *J. Chem. Phys.* **129** 174106
[8] Plenio M B and Huelga S F 2008 *New J. Phys.* **10** 113019
[9] Rebentrost P, Mohseni M, Kassal I, Lloyd S and Aspuru-Guzik A 2009 *New J. Phys.* **11** 033003
[10] Vedral V and Farrow T 2010 *e-print arXiv* 1006.3775
[11] Semio F L, Furuya K and Milburn G J 2010 *New J. Phys.* **12** 083033
[12] Lloyd S and Mohseni M 2010 *New J. Phys.* **12** 075020
[13] Asadian A, Tiersch M, Guerreschi G G, Cai J, Popescu S and Briegel H J 2010 *New J. Phys.* **12** 075019
[14] Sarovar M, Ishizaki A, Fleming G R and Whaley K B 2010 *Nature Physics* **6** 462
[15] Scholak T, de Melo F, Wellens T, Mintert F and Buchleitner A 2011 *Phys. Rev.* E **83** 021912
[16] Caruso F, Chin A W, Datta A, Huelga S F and Plenio M B 2010 *Phys. Rev.* A **81** 062346