

COVER PAGE

List of Participants

Sn.	Name	Country	Sn.	Name	Country
1	Herbert Fleischner	Austria	41	Mohamed Eltom	Sudan
2	Eleonora Ciriza	Italy	42	Mpinganzima Lydie	Sweden
3	Andrew Masibayi	Kenya	43	Pravina Gajjar	Sweden
4	Ben Owino Obiero	Kenya	44	Bjrn Textorius	Sweden
5	Benard Kipchumba	Kenya	45	Olof Svensson	Sweden
6	Bernard Nzimbi	Kenya	46	Bengt Ove Turesson	Sweden
7	Damian Maing	Kenya	47	Christer Kiselman,	Sweden
8	David Angwenyi	Kenya	48	Fredrik Berntsson	Sweden
9	Emma Anyika	Kenya	49	Leif Abrahamsson	Sweden
10	George Muhua	Kenya	50	Paul Vanderlin	Sweden
11	Idah Orowe	Kenya	51	Peter Sundin,	Sweden
12	Isaac Kipchirchir	Kenya	52	Rikard Bogvad	Sweden
13	Ivivi Mwaniki	Kenya	53	Vitalij Tjatyрко	Sweden
14	Jamen H. We're	Kenya	54	Agnes Joseph	Tanzania
15	James Okwoyo	Kenya	55	Akili Babi	Tanzania
16	John Nderitu	Kenya	56	Alex Xavery Matofali	Tanzania
17	Josephine Wairimu	Kenya	57	Allen Mushi	Tanzania
18	Kevin Oketch	Kenya	58	Anna Mwanjoka	Tanzania
19	Kithela Mille	Kenya	59	Beatha Ngonyani	Tanzania
20	Lydia Musiga	Kenya	60	Charles Mahera	Tanzania
21	Masinde Wamalwa	Kenya	61	Christian Alphonse	Tanzania
22	Ongaro Nyang'au	Kenya	62	E. S. Massawe	Tanzania
23	Patrick Weke	Kenya	63	Edith Luhanga	Tanzania
24	Philip Ngare	Kenya	64	Egbert Mujuni	Tanzania
25	Rachel Mbogo	Kenya	65	Emaline Joseph	Tanzania
26	Richard Simwa	Kenya	66	Emmanuel Evarest	Tanzania
27	Waihenya Kamau	Kenya	67	Eunice Mureithi	Tanzania
28	Wycliff Nyang'era	Kenya	68	Greyson. Kakiko	Tanzania
29	Fanja Rakotondrajao	Madagascar	69	Isambi Sailon Mbalawata	Finland
30	Bernard O. Ikhimwin	Nigeria	70	John Mwaonanji	Tanzania
31	Ewaen K.y Osawaru	Nigeria	71	Jonas P. Senzige	Tanzania
32	Cassien Habyarimana	Rwanda	72	Josepha V. Itambu	Tanzania
33	Desire Karangwa	Rwanda	73	Judith Pande	Tanzania
34	Froduald Minani	Rwanda	74	Lilian Olengeile	Tanzania
35	Gahirima Michael.	Rwanda	75	Mashaka Mkandawile	Tanzania
36	Isidore Mahara	Rwanda	76	Moses Mwale	Tanzania
37	E Nshimyumuremyi	Rwanda	77	Mpele James	Tanzania
38	Thomas Bizimana	Rwanda	78	Mpeshe, Saul C.	Tanzania
39	Abdou Sene	Senegal	79	Augustino Isdory	Tanzania
40	Oluwole Daniel Makinde	S. Africa	80	Muhaya Kagemulo	Tanzania

Sn.	Name	Country	Sn.	Name	Country
81	Mussa Ally	Tanzania	104	Mukalazi Herbert	Uganda
82	Pitos Seleka	Tanzania	106	Nanyondo Josephine	Uganda
83	Said Sima	Tanzania	107	Ndagire Majorine	Uganda
84	Santosh Kumar	Tanzania	108	Opio Ismai	Uganda
85	Shaban Nyimvua	Tanzania	109	Sanyu Shaban	Uganda
86	Silas S. Mirau	Tanzania	111	Tumuramye Fred Kacumita	Uganda
87	Soud Khalfa Mohamed	Tanzania	112	Tushemerirwe Phionah	Uganda
88	Sylvester E. Rugeihyamu	Tanzania	113	Tushemerirwe Phionah	Uganda
89	Theresia Bonifasi	Tanzania	114	Vincent Ssembatya	Uganda
90	Theresia Marijani	Tanzania	115	Walakira David Ddumba	Uganda
91	Uledi A. Ngulo	Tanzania	116	Wandera Ogana	Uganda
92	Yaw Nkansah-Gyekye	Tanzania	117	David Henwood	UK
93	Anguzu Collins	Uganda	118	admanabhan Seshaiyer	USA
94	Arop M.	Uganda	119	Alasford M. Ngwengwe	Zambia
95	Buletwenda Charles	Uganda	120	Anthony Moses Mwale	Zambia
96	John Mango	Uganda	121	Ilwale Kwalombota	Zambia
97	Juma Kasoz	iUganda	122	Issac. D. Tembo	Zambia
98	Kitayimbwa Mulindwa John	Uganda	123	John Musonda	Zambia
99	Kito Luliro Silas	Uganda	124	Mbokoma Mainza	Zambia
100	Likiso Remo Winniefred	Uganda	125	Mervis Kikonko Shamalambo	Zambia
101	Livingstone Luboobi	Uganda	126	Mubanga Lombe	Zambia
102	Masette Simon	Uganda	127	Trevor Chilombo Chimpinde	Zambia
103	Mirumbe Geoffrey Ismail	Uganda			

Forewords

The East African Universities Mathematics Programme (EAUMP) is a collaboration project between Eastern Africa Universities and International Science Programme (ISP) of Uppsala University, Sweden. The project started in 2002, and the currently participating universities in the region are University of Dar es Salaam, University of Nairobi, Makerere University, National University of Rwanda, Kigali Institute of Technology and University of Zambia. The main objective of the programme is to promote cooperation and exchange of ideas in mathematical research and teaching of mathematics and to stimulate communication between mathematicians in the Eastern African Region and beyond.

The first EAUMP Conference was held in Nairobi, Kenya, from 18th March to 21st March 2003. Due to the success of the conference it was decided to hold such a conference regularly. The Department of Mathematics of the University of Dar es Salaam agreed to hold the 2nd EAUMP conference to celebrate 10th anniversary of the programme.

The proceedings, which follow, consist of speeches, papers and abstracts presented at the 2nd EAUMP Conference, held at The Nelson Mandela African Institute of Science and Technology, Arusha, Tanzania from 22nd to 25th August 2012. More than 125 participants from about 10 countries attended the conference. The conference program was comprised of 6 invited plenary lectures, and more than 45 contributed talks were presented and discussed.

The aims of the conference were:

- To stimulate regional and international collaboration in research and training.
- To provide a forum for interaction of African Mathematicians and others from the developed countries for research experience.
- To introduce African Mathematicians from the region to some fundamental techniques and recent developments in these fields, thus forming research collaborations.
- To update the knowledge of African Mathematicians, particularly lecturers and M.Sc./Ph.D. students who are stationed at home, to start pursuing these areas as research interest.

The success of the conference could not have been registered without concerted effort from the Local organizing committee in the Department of Mathematics and the EAUMP coordinators committee. I would therefore like to extend my heartfelt thanks to the following.

Local Organizing Committee

- | | |
|--------------------------------|-------------------|
| • Dr. Egbert Mujuni | Chairperson |
| • Dr. Sylvester. E. Rugeihyamu | EAUMP coordinator |
| • Dr. Eunice Mureithi | Member |
| • Dr. Theresia Marijani | Secretary |
| • Mr. Emmanuel Evarest | Member |

EAUMP Coordinating Committee

- Dr. Sylvester E. Rugeihyamu (*Coordinator of University of Dar es Salaam*)
- Prof. Patrick G. O. Weke (*Coordinator of University of Nairobi*)
- Dr. Juma Kasozi (*Coordinator of Makerere University*)
- Dr. Isaac Tembo (*Coordinator of University of Zambia*)
- Dr. Isidore Mahara (*Coordinator of National University of Rwanda*)
- Mr. Michael Gahirima (*Coordinator of Kigali Institute of Science and Technology*)
- Dr. John M. Mango (*Inter Network Coordinator, Makerere University*)

Acknowledgement

Without finances, not much could have been attained. I would like extend my thanks following associations, agencies and institutions for their generous support.

- International Science Programme (ISP) of Uppsala University, Sweden.
- TWAS-The Academy Science of the Developing World.
- The International Mathematical Union-Committee for Developing Countries (IMU CDC).
- NORAD (Through NOMA Project in Mathematics Department, UDSM).
- The German Academic Exchange Service (DAAD).
- University of Dar es Salaam Gender Centre.
- University of Dar es Salaam, Directorate of Research.
- Tanzania Communications Regulatory Authority (TCRA).
- Tanzania Commission for Science and Technology (COSTECH).

Prof. E. S. Massawe

Head, Mathematics Department, UDSM and Overall EAUMP Coordinator

Contents

List of Participants	ii
Forewords	iv
SPEECHES	1
Prof. E. S. Massawe: Welcome Speech	1
Dr. John M. Mango: Eastern Africa Universities Mathematics program (EAUMP - Network <i>Origin, Operation, Achievements the Future and Challenges.</i>	3
Prof. Rwekaza Mukandala: Welcome Speech	8
Prof. Makame Mnyaa Mbarawa: OPENING SPEECH	9
PAPERS	11
M. E. A. El Tom: A Proposed Research Agenda in Mathematics Education in Africa	11
Paul Vaderlind: Mathematical Competitions for Gifted Students: Organization and Training	23
Patrick G. O. Weke: Estimation of IBNR Claims Reserves Using Linear Models . .	26
O. D. Makinde: Heat Transfer Analysis of a Convecting and Radiating Two Step Reactive Slab	43
Mervis Kikonko: Non-Definite Sturm-Liouville Problems Two Turning Points . . .	52
S. A. Sima, M. Ali and I. Campbell: The Three Layers Maize Crop Optimal Distri- bution Network in Tanzania	61
Santosh Kumar: A Survey of the Development of Fixed Point Theory	72
Livingstone S. Luboobi: Epediomological Modelling at Macro and Micro Levels: The Case of HIV/AIDS	81
Christian Baruka Alphonse: Analysis of Cell Phone User's Loyalty in Tanzania Using Markov Chains	87
Soud K. Mohamed: Derivatives over Certain Finite Rings	98
Eleonora Ciriza: Bifurcation results on symplectic manifolds	108
John Musonda: Three Systems of Orthogonal Polynomials and Associated Operators	120
ABSTRACTS	155
Christer Kiselman: Asymptotic Properties of The Delannoy Numbers and Similar Arrays	155
Vitalij A. Chatyrko: The (Dis)connectedness of Products in the Box Topology	155

Frerik Berntsson: Identification of Coefficients in Parabolic Equations Using Measurements on the Boundary	155
Padmanabhan Seshaiyer: Multidisciplinary Research in Mathematical Sciences With Applications to Real World Problems in Biological, Bio-Inspired and Engineering Systems	156
Wandera Ogana: Epidemic Potential for Malaria in Epidemiological Zones in Kenya	156
Fanja Rakotondrajao,: How to Manipulate Derangements	156
Abdou Sene: The Dynamics of Populations in Wetlands	157
Isambi Sailon Mbalawata and Simo Sarkka: Adaptive Markov Chain Monte Carlo Using Variational Bayesian Adaptive Kalman Filter	157
Patrick G. O. Weke: Linear Estimation of Location and Scale Parameters for Logistic Distribution Based on Consecutive Order Statistics	158
Lydia Musiga: A Stochastic Model for Planning a Compartmental Education System and Supply of Manpower	158
Emma Anyika and Patrick Weke: Financial Sector Performance Enhancers .	158
Mashaka Mkandawile: Estimating the List Size Using Bipartite Graph for Colouring Problems	159
Emaline Joseph, Kgosimore and Teresia Marijani: Mathematical Modeling of Pneumonia Transmission Dynamics	159
David Ddumba Walakira: Hydrodynamics of Shallow Water Equations: A Case Study of Lake Victoria	159
J.W. Mwaonanji: Boundary Layer Flow Over a Moving at Surface With Temperature Dependent Viscosity	160
Eunice Mureithi: Absolute-Convective Instability of Mixed Forced-Free Convection Boundary Layers	160
Moses Mwale: Optimal Premium Policy of an Insurance Firm With Delay . . .	160
G.I. Mirumbe, Vincent Ssembatya, Rikard Bgvad and Jan Erik Bjork: On the Coexistence of Distributional and Rational Solutions for Ordinary Differential Equations With Polynomial Coefficients	161
Philip Ngare: On Modelling and Pricing Index Linked Catastrophe Derivatives	161
Lusungu Mbiliri, Charles Mahera, and Sure Mataramvura: Optimal Portfolio Management When Stocks are Driven by Mean Reverting Processes	162
Egbert Mujuni: On Hub Number of Hypercube and Grid Graphs	162
Vincent A Ssembatya: Fixed Points of Homeomorphisms of Knaster Continua	162
Isaac Daniel Tembo: Continuity of Inversion in the Algebra of Locality - Measurable Operators	163

Herbert Fleischner: Uniquely Hamiltonian Graphs	163
Kitayimbwa M. John, Joseph Y. T. Mugisha and Robert A. Saenz: The Role of Backward Mutations on the Within Host Dynamics Of HIV-1 .	163
Kipchirchir, I. C.: Comparative Study of the Distributions Used To Model Dispersion	164
Theresia Marijani: A Within Host Model of Blood Stage Malaria	164
Wilson Mahera Charles: Application of Stochastic Differential Equations to Model Dispersion of Pollutants in Shallow Water	165
R. W. Mbogo, L. Luboobi, J. W. Odhiambo: Stochastic Model for In-Host HIV Virus Dynamics With Therapeutic Intervention	165
F. Berntsson, V. Kozlov, L. Mpinganzima, and B.O. Turesson: An Alter- nating Iterative Procedure for the Cauchy Problem for the Helmholtz Equation	165

THE 2ND EASTERN AFRICA UNIVERSITIES MATHEMATICS PROGRAMME (EAUMP) CONFERENCE

The Nelson Mandela African Institute of Science and Technology, Arusha, Tanzania
August 22nd - 25th, 2012

Welcome Speech

by

Prof. E. S. Massawe

Head, Mathematics Department, University of Dar Es Salaam

Guest of Honour, Minister of Communication, Science & Technology, Tanzania, Hon. Prof. Makame Mnyaa Mbarawa (MP),

The Vice Chancellor of the University of Dar es Salaam, Professor Rwekaza Mkandala

The Vice Chancellor Nelson Mandela African Institute of Science and Technology, Professor Burton Mwamila

Head of ISP and Director of Chemistry Program, Professor Peter Sundin

Director of Mathematics Program, Professor Leif Abrahamsson

Delegates from ISP

Distinguished guests and visitors,

Dear Participants,

Ladies and Gentlemen,

On behalf of the Department of mathematics, University of Dar es Salaam and on my own behalf, I wish to take this opportunity to welcome you all the invited guests and participants and especially you the guest of honour, the Minister of Communication, Science & Technology, Tanzania, Hon. Prof. Makame Mnyaa Mbarawa, to this important Congress. Please do feel at home.

Guest of Honour

The Eastern Africa Universities Mathematics Programme (EAUMP) Network was established in 2002 to further the mathematical sciences in the Eastern Africa Region. The main objective of the Network is to promote cooperation and exchange of ideas in mathematical research and teaching of mathematics and to stimulate communication between mathematicians in the Eastern Africa Region and beyond. The Network, since its foundation, has been organizing schools and workshops and conferences. One of the objectives of these workshops and conferences is to bring together researchers from various branches of mathematics and related fields, and to simulate intersection and cooperation.

Guest of Honour

EAUMP is a non-political and non-profit making Network devoted to the promotion of research, teaching and learning of mathematics at all levels. We are very proud of this because recent years have seen unprecedented growth of interest in the application of mathematical ideas and techniques to problems in Science and Technology in industry.

Guest of Honour

EAUMP is aware of the important role of the mathematics researchers in promoting the subject. In this conference, we shall have a series of presentations and group discussions in the areas of pure mathematics, Financial Mathematics, Epidemiology, Mathematics for the Industry, Theoretical Fluid Dynamics, Statistics, Mathematics Education, Computer Science and Theoretical Physics. The program will include keynote speakers, esteemed researchers and regular presentations.

Guest of Honour

EAUMP was found in 2002. This year we are celebrating the 10th Anniversary of the EAUMP Network. Performance of EAUMP in the last 10 years gives one confidence that the EAUMP will survive the next 10 years and beyond as an important and active Network. Schools and Conferences of this type will have to continue. Schools and Conferences of EAUMP are the lifeline of the Network as is the case of most professional organizations

Guest of Honour

Finally I will like to say that, our motto is "We build for the Future". The future success of EAUMP will depend on the continued cooperation and commitment of all the members of EAUMP and other stakeholders.

Guest of Honour

Allow me on behalf of all the EAUMP members to thank all those who in various ways have supported our Conference and specifically the International Science Programme (ISP) of Uppsala University, Sweden , The Ministry of Communication, Science and Technology, Tanzania, Commission for Science and Technology (COSTECH), Tanzania, The Academi Science of the Developing World (TWAS), The European Mathematical Society - Committee for Developing Countries (EMS-DC), The German Academic Exchange Service (DAAD), The University of Dar es Salaam, The University of Dar es Salaam Gender Centre, the University of Dar es Salaam Directorate of Research, NORAD through NOMA Project, Tanzania Communications Regulatory Authority (TCRA) and All nodes of the EAUMP network

I would also like to thank the local organizing committee and all those behind the scene for the excellent job done in terms of making us stay in Arusha happily.

Once again, you are all warmly welcome.

Thank you.

Eastern Africa Universities Mathematics program (EAUMP - Network Origin, Operation, Achievements the Future and Challenges.

by

Dr. John M. Mango

EAUMP Coordinator and Inter-Network Cooperation

Today is a great and memorable day for EAUMP

- In June 1995, SIDA/SAREC and Uppsala University organized a conference on 'Donor support to development oriented research in Basic Sciences'.
- In March 1999 a conference was organized in Arusha, Tanzania with the aim of addressing the regional challenges.
- In 2001 SIDA/SAREC organized the 1st International conference in Mathematics in Africa South of the Sahara. It was during this conference that the poor state of Mathematics in the Eastern African region was reported. This gave birth to EAUMP in 2002 to try and address the Challenges of the time. It is interesting to note that some of the challenges are still existing though at a reduced level.

The Key People who participated in the initial stages 2001/2002

- Prof. Leif Abrahamson -Uppsala University in Sweden.
- Dr. C. Baruka Alphonse -University of Dar es Salaam.
- Prof. V. Masanja -University of Dar es Salaam.
- Prof. John W. Odhiambo -University of Nairobi.
- Prof. Wandera Ogana -University of Nairobi.
- Dr. Vincent Ssembatya -Makerere University.
- Prof. Livingstone Luboobi -Makerere University.
- Dr Fabbian Nabugoomu, Makerere University.

Objectives of the EAUMP Network

- Enhancement of postgraduate training with special emphasis to PhD training.
- Establishing and strengthening collaborative research in Mathematics.
- Strengthening the collaborating Mathematics departments.
- Development of resources for the collaborating Mathematics Departments.

Membership of the Network

- University of Dar-es-Salaam, Tanzania, (Since 2002)
- Makerere University, Uganda, (Since 2002)
- University of Nairobi, Kenya. (Since 2002)
- National University of Rwanda (NUR) and Kigali Institute of Science and Technology (KIST), joined in August 2008.
- University of Zambia, joined in April 2009.
- NB: University of Addis Ababa, University of Khartoum and Nelson Mandela African Institute for Science and Technology have expressed interest to join the Network.

Coordination Structure

- ISP Mathematics Director–**Prof. Leif Abrahamsson**
- EAUMP Advisory Board
- Overall Coordinator–**Prof Estomih Massawe** (Dar-Main Coordinating office for now)
- Inter Network Coordinator– **Dr John Mango Magero**
- School of Mathematics, University of Nairobi Coordinator– **Prof Patrick Weke**
- Makerere University, Department of Mathematics Coordinator-**Dr Juma Kasozi**
- University of Dar es Salaam, Department of Mathematics Coordinator-**Dr Sylvester Rugeihyamu**
- Kigali Institute of Science and Technology (KIST), Department of Applied Mathematics Coordinator-**Mr Gahirima Michael**
- National University of Rwanda (NUR), Department of Applied Mathematics Coordinator-**Dr Mahara Isidore**
- University of Zambia, Department of Mathematics Coordinator-**Dr Isaac Tembo.**

Sources of funding for the Network

- ISP-International Science Program (Over 95% of EAUMP activities are sponsored by ISP), based at the University of Uppsala, Sweden.
- ICTP-International Centre for Theoretical Physics in Italy.
- AMMSI-Millennium Science Initiative
- LMS- London Mathematical Society

- DAAD
- IMU/ CDC- International Mathematical Union through its Commission for developing Countries
- TWAS-The Third World Academy of Sciences
- The other sources of sponsorship are the local universities.

Major Achievements of the Network since 2002

- Capacity building through Ph.D training (6 completed and 12 ongoing)- All are members of staff eg Egbert Mujuni.
- Capacity building through Postdoc (4 awarded in 2011)- All are members of staff.
- Capacity building through M.Sc. training (more than 50 have benefited)- Some are members of staff, Some doing Ph.Ds. Staff exchange in the region.
- Research visits by Cooperating Scientists (From Sweden, Italy, USA,) eg Paul, Rikard, Fanja Eleonara, Ramadas etc Equipment (Computers, projectors etc).
- Books and Journals (Subscribed to some Journals, obtained books and ebooks).
- Publications (Increased volume of publication in refereed journals).
- Conferences/Workshops/Schools for graduate students and researchers/lecturers. The Schools are organized to cover areas of mathematics where the region is most disadvantaged)- Over 300 different M.sc and some Ph.Ds have attended and benefitted from the EAUMP Schools and Conferences.
- Research projects.
- Established/identified potential of member departments.

Challenges

- Low funding and yet in this region of the world we are not short of interested students to do Masters and Ph.Ds in Mathematics.
- Insufficient local manpower to teach and supervise
- Understaffing in Departments of member universities.
- Low interest of Ph.D students in Pure Mathematics.

The future of EAUMP Network

- The poor state of mathematics in the region is now improved by ISP intervention. The present state needs to be improved further through continued cooperation with ISP and other organizations.
- There is great need for more capacity building in the member Departments through Ph.D, PostDoc and M.Sc training.
- When resources allow, the EAUMP network will be extended to other Universities in the region. There are smaller universities in the region where capacity building in Mathematics is of urgent need.
- We need to use the Network to help reduce the problem of brain drain. From the present experience, students who register in their local universities for their graduate training under the sandwich mode, have settled, and are teaching/working in their local/regional Universities.
- We plan to hold a Conference in each financial cycle (3 years as was the original plan) so that our graduate students, staff and academics outside the region will gather to share research experiences through paper and poster presentations.
- Strengthen the fundraising drive for the network and research cooperation with other networks through the newly created office of **Inter Network Cooperation**. Apart from the usual funding from ISP, ICTP and AMMSI/LMS for schools and conferences, this year using the new office we have been able to secure funds from CDC, TWAS and DAAD and this has supported a total of 13 persons (regional speakers and DAAD Alumni).
- Improve on the way we transport our student participants to EAUMP Schools and Conferences. In the recent past we have lost students in road accidents while travelling to attend EAUMP Schools.
- We request for more support from our local Universities and Governments.

Just a comment

In some discussion at the 2012 European Congress of Mathematicians, one Professor criticized the Scandevian Sandwich training mode of Sida and ISP type practiced in Africa. The professors proposal was that Sida and ISP sends money to South African Universities for capacity building of the SIDA/ISP collaborating Universities in Africa so that the training of the sandwich students takes place in South African Universities and not Swedish Universities. As EAUMP, we are strongly opposed to the idea in that;

- (i) South Africa is still interested in our PhD products/graduates and Sweden is not for they have more than enough. The SIDA/ISP collaboration with African Universities is for capacity building in the collaborating Universities and not in South Africa, Europe, USA etc. It is clear that South Africa has offered some of our graduates well paying positions and these have not come back to meet the objective of the training
- (ii) All the sandwich PhD students trained so far under SIDA/ISP have remained and are serving their home Universities. A case example is the SIDA Makerere Bi-Lateral programmes since 2002 which has trained over 200 PhDs mostly in the hard Sciences like

Engineering, Medicine, Agriculture etc and all these are stationed and serving Makerere University.

(iii) The cost of ISP sandwich PhD training is cheap and affordable.

We also recognize and appreciate the contribution of South African Universities in capacity building of regional Universities and we hope to continue collaborating with them but not to substitute the SIDA/ISP collaboration. As EAUMP, we remain grateful to our sponsors, we promise to work and achieve the set objectives as we also look forward to continued support of the Network by ISP and other organizations.

THANK YOU

THE 2ND EASTERN AFRICA UNIVERSITIES MATHEMATICS PROGRAMME (EAUMP) CONFERENCE

**The Nelson Mandela African Institute of Science and Technology, Arusha, Tanzania
August 22nd - 25th, 2012**

Welcome Speech

by

Prof. Rwekaza Mukandala

Vice Chancellor, University of Dar Es Salaam

Guest of Honour, The Minister of Communication, Science & Technology, Tanzania, Hon. Prof. Makame Mnyaa Mbarawa (MP),

Congress Participants,

Ladies and Gentlemen,

On Behalf of the entire University of Dar es Salaam and on my behalf, I wish to take this opportunity to welcome all the invited guests and participants to this second Eastern Africa Universities Mathematics Programme Congress (EAUMP). Please do feel at home.

The University of Dar es Salaam is proud to host this second EAUMP Congress. I am informed that the Network of EAUMP started on 2002, earlier than in most other regions in Sub-Saharan African region. This programme is unique and flexible since it has led to close collaboration between the participating departments in the network. All indications have shown that there is now more interaction among members of departments of Mathematics in the region and Mathematicians from Sweden and other areas.

When the Department of Mathematics of the University of Dar es Salaam indicated to me that University of Dar es Salaam has been honoured to host the 2nd EAUMP Congress we welcomed the initiative.

Congress of this nature complements the status of our respected and oldest Institutions in Africa. We also know that congresses of this nature are a forum for dissemination of information and for forging meaningful cooperation and collaboration in research and teaching. Our Universities in the region encourages collaboration among scholars of same discipline and also encourages inert-disciplinary arrangements.

At this juncture I wish to pay glowing tribute to the Swedish Universities, in particular Uppsala University through Sida for their commitment in the Development and Education in our region. We in the developing countries are very grateful for the support that Sida has extended to our Universities for collaborative research with scientists at similar Swedish institutions. We have developed capacity and competence in teaching and research.

To you participants of the congress; I wish you productive deliberations. Your contributions will go a long way in promoting the subject of Mathematics.

I now like to take this opportunity to invite our Guest of Honour, Hon. Prof. Makame to address you and officially open the Congress.

THE 2ND EASTERN AFRICA UNIVERSITIES MATHEMATICS PROGRAMME (EAUMP) CONFERENCE

The Nelson Mandela African Institute of Science and Technology, Arusha, Tanzania
August 22nd - 25th, 2012

OPENING SPEECH

by

Prof. Makame Mnyaa Mbarawa

Minister Of Communication, Science & Technology, Tanzania

*The Chairperson of the EAUMP Conference Organising Committee,
Distinguished guests,
Distinguished Conference Participants,
Ladies and Gentlemen,*

It is a great honour and pleasure for me to participate in this special activity of the Eastern Africa University Mathematics Programme Network. This meeting of the EAUMP is significant not only to Mathematicians in higher learning institutions but also to all people who understand the value and role of mathematical Sciences in our everyday life and work. That is why I consider this opportunity to interact with members of this Network a significant one and quite enriching. I must therefore thank the organizing committee for inviting me to participate in this opening session and therefore allowing me time to have a glimpse at some on the professional concerns of mathematicians as reflected in the agenda for this meeting.

I take this opportunity to welcome you all to the Nelson Mandela African Institute of Science and Technology and to the EAUMP conference on particular. It is my sincere hope that you will find this venue a convenience place for the kinds of activities scheduled for this conference. This is the most favourable season for this part of Tanzania. Those of you coming from warmer regions may therefore find this to be the best time of the year to visit Arusha. I am however confident that, in the course of your stay, each one of you will find a memorable aspect of life and places in this town.

Chairperson

I am informed that during this conference, research papers on various topics in mathematics and mathematical sciences will be presented by experts in the field. I have no doubt that the papers to be presented originate from concerted research effort, and that this conference therefore serves as an avenue for the dissemination of the findings of recent research. Yet, while sharing of ideas and research findings among yourselves is in itself a sufficiently noble activity, a lot more will be gained if your deliberations ultimately find a place in professional publications. I hope this is indeed what you plan to do with the papers to be presented here.

Chairperson

I wish to relate to the significant of Mathematical sciences in human experience and development. It is common knowledge that Mathematical reasoning occupies a core position in the foundation of scientific and technological developments that have characterized the entire history of humanity. It is no wonder, therefore, that mathematics is known as the queen of science and technology. But we also know that at the very elementary level, Mathematics is used in

measurements, commerce, engineering, and as a daily language of comparison. I am told, and I have no reason to doubt the fact that, at the most and the more sophisticated level, mathematics is used as a tool to understand the universe. The whole of Information Technology, so we are informed, is basically the mathematics of wave transmission. It is in view of this profound significance of the discipline of mathematics that I revere the work being done by your Network in advancing the frontiers of knowledge in this field. I urge you to maintain vigour and rigour in researching the various topical issues of our day and in improving the public rendering of the nature and role of Mathematics in our lives.

Chairperson

It is gratifying that EAUMP is a regional Network of scientists, and that it has functioned for 10 years. I congratulate you for being one of the oldest and vibrant professional organizations in our region. I also congratulate you for the excellent tradition you have instituted of holding your workshops in the various countries of the region rather than having them conveniently hosted by one country. This is surely a virtue for other regional Networks to emulate.

Chairperson

I am informed that EAUMP was found in 2002, implying that the Network today is 10 years old, and that since then it has held several workshops. I must commend EAUMP for maintaining a strong and stable momentum for 10 years. I strongly join hands with you chairperson that;performance of the EAUMP in the last 10 years gives one confidence that EAUMP will survive the next 10 years and beyond as an important and active Programme.

Chairperson

I understand that in the last workshops, participants drew a list of recommendations or action points. It would be interesting to explore the extent to which those have been implemented. While I am not sure it is in your interest to engage in this kind of exercise at this point in time, I am quite convinced that this would be a useful thing to do. In same vein, I may go a step further and propose that you revisit all the major recommendations made in previous workshops with a view to assessing the impact they have had on the development of Mathematics and mathematical Sciences in the region.

Chairperson

I am sure this opening session is not meant for long speeches. I therefore wish to end my remarks by wishing you very productive deliberations and a happy stay in Tanzania.

Lastly, the Chairperson, distinguished guests, ladies and gentlemen, it is now my honour and pleasure to declare the 2012 EAUMP CONFERENCE OFFICIALLY OPENED.

I thank you all for your attention

A Proposed Research Agenda in Mathematics Education in Africa

by

M. E. A. El Tom

Garden City College for Science and Technology, Sudan

Abstract

Efforts at capacity building in mathematics in Africa have not been sufficiently sensitive to the importance of mathematics education. They do not appear to have been informed by the fact that mathematical research and mathematics education are organically linked: a weakness in either will undermine the other as well as the science and & technology base, which is vital for meaningful sustainable development.

The paper attempts to identify the most pressing issues and questions for mathematics education in Africa. The proposed research agenda in mathematics education are based on these issues and questions.

1. Introduction

A major aim of the East African Universities Mathematics Programme (EAUMP) network is to strengthen mathematical research in departments of mathematics participating in it. It is useful to think of this aim as part of the broader goal of promoting mathematics in the continent, which is shared in common by the African mathematical community. The achievement of this goal is far from straightforward and requires considerable effort. For, mathematics in Africa is 'young' (most African countries could not boast a single Ph. D. in mathematics at the time of independence in early 1960s. Moreover, the role of mathematics in society is "subtle and not generally recognised in the needs of people in everyday life and most often it remains totally hidden in scientific and technological advancements" (Brown 2007). I consider in the next section some specific obstacles that seem to stand between African mathematicians and the achievement of the goal of promoting mathematics. Also, the section cites some of the problems facing mathematics in specific African countries. The level of research output in mathematics education in Africa is discussed in section 3. A review of the literature dealing with factors that play a role in mathematics achievement is presented in section 4. A proposed research agenda in mathematics education in Africa are presented in the final section.

2. Some problems of mathematics in Africa

The International Mathematical Union (IMU) observes in a recent study that in most African countries "mathematical development is limited by low numbers of secondary school teachers and mathematicians at the masters and PhD levels." Furthermore, the study observes that "Talented students are dissuaded from careers in mathematics by low salaries, a poor public image, and a shortage of mentors engaged in exciting mathematical challenges" (IMU 2009).

Overall, the IMU study concludes, "the story of mathematical development in Africa is one of potential unfulfilled. Based on the achievements of some outstanding individuals and institutions, it is clear that no African country lacks talented potential mathematicians. But without a stronger educational structure at all levels, few of them are able to reach their potential." The last statement in this quotation is further articulated in the observation that there is an

almost universally held conviction held by mathematicians and mathematics educators, "that each mathematical level of learning is grounded pyramid-like in the previous ones, and that lack of quality or capacity at any level of a country's mathematical infrastructure weakens all the levels above. Conversely, the absence of some kind of pinnacle deprives the lower levels of leadership, training and context" (IMU 2009).

2.1 Cracks in the foundation

There are important indications that educational systems in most African countries exhibit cracks in their respective systems. Awareness of these cracks and devising appropriate measures for dealing with them are prerequisites for effective promotion of mathematics in the continent.

- Image of mathematics

Achievement in mathematics is influenced by, among other factors, beliefs about and attitudes towards mathematics. How do parents, teachers and students themselves view mathematics? Do these groups attribute success in mathematics largely to ability or effort? A questionnaire was designed and distributed to 24 leading mathematicians working in departments of mathematics in universities of different African countries to try and find answers to such questions. The response was highly limited, only 6 questionnaires were completed and returned: Ghana, Mali, Kenya, Nigeria, Sudan and Tunisia. The responses from 4 of these countries, namely Ghana, Mali, Nigeria and Sudan turned out to be similar and they are presented in Figure 1.

Although it is not permissible to generalize on the basis of very limited response to the questionnaire, the Figure suggests that general education students in Ghana, Mali, Nigeria and Sudan have a negative image of mathematics, characterizing it as very difficult, unrelated to reality and only for the clever. Also, society in the four countries seem to share in common with general students the perception that mathematics is both difficult and only for the clever. In contrast, policy-makers seem to have a positive image of mathematics, indicating awareness of its importance for economic development. Indeed, the Nigerian Federal Minister of Education said "there could be no meaningful progress in the country without promoting the study of mathematics and sciences" (AfricaSTI. 4 March 2012)

Response to the questionnaire from Tunisia indicate that both general education students and society at large perceive mathematics as very difficult and unrelated to reality. Also, policy-makers view mathematics as very important for economic development.

In Tanzania, mathematics is characterized as Math characterized as the "(most) difficult subject taught in schools" (Philemon 2010). In their review of the strengthening mathematics and science in secondary education (SMASSE) science project in Kenya, Onderi and Malala (2011) believe that the documented poor performance of students in mathematics could be attributed to students' negative attitude towards the subject. They go on to ascribe this attitude to "low entry behavior, belief that these subjects are hard, peer pressure, lack of proper learning facilities, teacher absenteeism and theoretical approach to teaching mathematics." However, the response to the questionnaire indicate that policy-makers in Kenya attach great value for mathematics.

The data reported above pertain to 7 countries belonging to different regions of the continent (North, East and West Africa), exhibit important differences in their educational systems, and differ in the levels of their respective economic development. Thus, it is not unreasonable to conclude that the image of mathematics in most African countries is similar to that reported for the 7 countries mentioned above.

Figure 1: : General education students', society's and policy-makers' image of mathematics, selected African countries, 2012

	General education students	Society at large	Policy-makers
Very difficult			
Unrelated to reality			
Only for the clever			
Very important for passing examinations			
Very important for economic development			

Source: responses to questionnaire from mathematicians in Ghana, Mali, Nigeria, and Sudan.

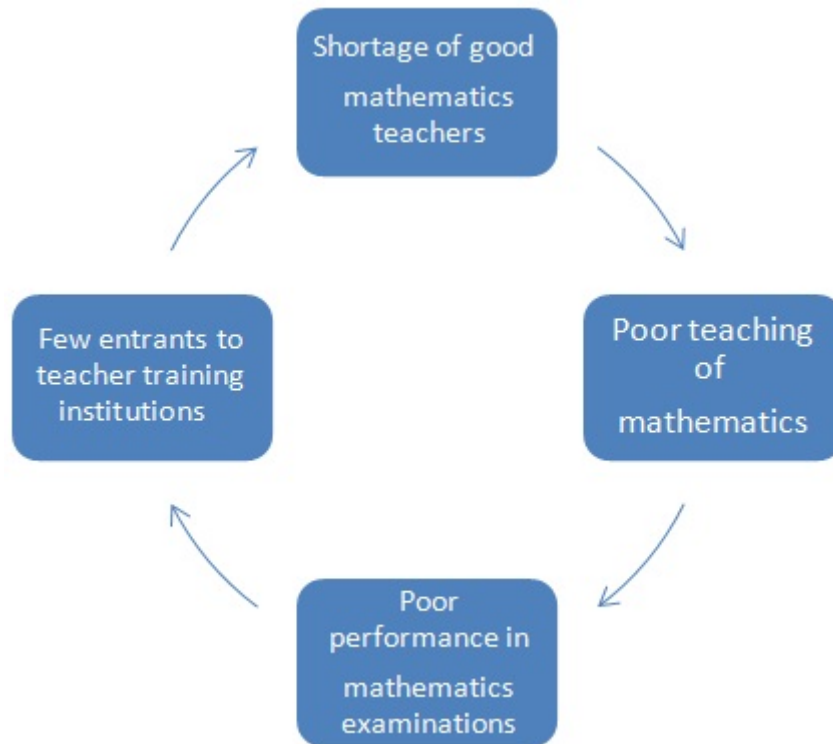
- Teachers of mathematics

Hanushek and Rivkin (2006) make the important observation that "The most consistent finding across a wide range of investigations is that the quality of the teacher in the classroom is one of the most important attributes of schools". Yet the identification of good teachers has been complicated by the fact that the simple measures commonly used - such as teacher experience, teacher education, or even meeting the required standards for certification - are not closely correlated with actual ability in the classroom (Harbison and Hanushek (1992); Hanushek (1995); Hanushek and Luque (2003); Hanushek and Rivkin (2006)). But, however one perceives of good teaching (e.g. Goe (2007)), there is data to suggest strongly that 'good' teachers of mathematics are in short supply in most African countries.

Indeed, in most African countries, mathematical development is limited by low numbers of secondary school teachers and mathematicians at the masters and PhD levels. An important contributing factor to this situation is that talented students are dissuaded from careers in mathematics by low salaries, a poor public image, and a shortage of mentors engaged in exciting mathematical challenges (Developing Countries Strategies Group (DCSG), 2009). South Africa, Tanzania and Uganda provide examples of this problem.

In South Africa, Adler (1994) reported that "72% of mathematics teachers in African schools, are under-qualified" Obviously, these shortages of 18 years ago pose an enormous challenge well into the future. Indeed, Adler noted that projections "for the next ten years indicate that there is a need to produce 135 700 primary and 93400 secondary teachers in order to reach the targeted average teacher-pupil ratio of 1:35. That the immediate areas of attention need to be [mathematics and science] is highlighted in numerous policy proposals" More recently, the South African Department of Education (2004:10)

Figure 2: Vicious cycle of shortage of good teachers of mathematics



expressed concern that the teaching of mathematics in schools was often never a first choice to talented mathematics graduates. Consequently, mathematics was often taught by inadequately qualified teachers and this led to a vicious cycle of poor teaching, poor learner achievement and a constant under-supply of competent teachers.”

In Tanzania, Danielle (2012) observes that “Enrollment rates are low and failure rates are high. Resources and learning materials are limited. But perhaps more than anything, the country suffers from a severe lack of qualified teachers.” In a recent World Bank study (Mulkeen, 2009) it is reported that in Zanzibar, 970 students passed A-level examinations in 2006, but only 53 of these passed mathematics, which leads to shortage of qualified entrants to teacher training colleges. The resulting vicious cycle is shown in Figure 2.

A vicious cycle similar to that in Zanzibar is found in Uganda. For, despite a lowering of admission requirement, Uganda found it difficult to fill places for secondary mathematics and science teacher training in the national training colleges. This reflects the imbalance in examination results. In the 2006 Uganda Advanced Certificate of Education (UACE) examination, 25836 students passed history, but only 5776 passed mathematics. This weakness in mathematics can be seen as a vicious cycle.

Research has shown a positive correlation between teachers’ content knowledge and their students’ learning (Villegas-Reimers 2003, UIS 2006). Despite the importance of adequate content knowledge, there are concerns that some teachers in Africa do not reach the level of knowledge required. SACMEQ data show that in several countries the aver-

Table 1: Percentage of women who hold a doctorate degree of the total doctorate holders in mathematics: selected African countries

Country	Proportion of women holding a PhD in mathematics (%)
Algeria	16
Botswana	31
Burkina Faso*	0
Djibouti	100 (only doctorate holder is female)
Egypt	20
Malawi	25
Mali*	0
Mauritania*	0
Mauritius	17
Somalia	50 (1 out of 2)
South Africa	19
Sudan*	8.3
Swaziland	50 (3 out of 6)
Tanzania	2.6
Tunisia	18

* Author's observations. Source: El Tom (2008); Gerdes (2007).

age teacher did not perform significantly better in reading and mathematics tests than the highest performing sixth-grade students (UNESCO 2006).

- African women and mathematics

It is widely recognized that women are severely underrepresented in the fields of science and engineering worldwide (UNESCO: The World's Women 2010: Trends and Statistics).

A significant feature of mathematics in Africa is that it is male-dominated. Based on first-hand experience of mathematics in several African countries and the data compiled by Gerdes (2007) about African doctorates in mathematics, I estimate the proportion of women mathematicians in Africa to be, on average, less than 10%. Table 2 below shows some relevant data.

The seriousness of this situation led the African Mathematics Millennium Science Initiative (AMMSI) to organize a Symposium in 2008 on African Woman and Mathematics, Maputo, Mozambique. Participants noted that the following factors, among others, influenced the motivation of the girl child towards mathematics and led to lack of self-esteem in the subject:

- Belief that mathematics was a tough subject.
- Lack of role models in the area of maths.
- Early pregnancies.

- Cultural, economic and religious backgrounds that impeded the access of children in general, and the girl child in particular, from accessing quality education.

Unless the representation of African females in mathematics is improved significantly, the pool of potential mathematicians will remain restricted and, consequently, efforts at capacity building in mathematics education and mathematical research will be hampered.

- Performance of students The performance of students in mathematics is described as poor in many African countries. For example, in Kenya, the consistently poor performance in mathematics and science subjects became a matter of serious concern in the late 1990s and the Ministry of Education, Science and Technology felt that it had to intervene in order to improve the situation. Thus, a project entitled 'Strengthening mathematics and science in secondary education' (SMASSE) was introduced in 1998 (Phase I) in cooperation with the Japanese International Cooperation Agency (JICA).

Feeling that they share in common with Kenya the same problem of poor performance, several countries joined SMASSE. In 2011, SMASSE membership included Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Congo, Cote d'Ivoire, Egypt, Ethiopia, Gambia, Ghana, Lesotho, Madagascar, Malawi, Mali, Mauritius, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Seychelles, Sierra Leone, South Africa, Sudan, Swaziland, Tanzania, Uganda, Zambia, and Zimbabwe (Mutahi 2011, cited in Onderi and Malala).

More than a decade since the introduction of SMASSE, Onderi and Malala (January 2011) find that "teaching in schools is examination oriented and rote learning is the order of the day in most schools. Little attention is paid to individual differences, teaching and effective evaluation methods and classroom management. This has been therefore reflected in the declining performance in Mathematics and Sciences in the national examination, with only a few exceptions.

In Tanzania, only 24.3% passed B/Mathematics in Certificate of Secondary Education Examinations (CSEE) in 2008 (compared with a pass rate of 46.3% biology and 53.6% physics. The pass rates in CSEE 2009 for mathematics and science subjects were: Bio 43.2%; B/Maths 17.8%; Physics 55.5%; Chem. 57.1%. Interestingly, boys performed better than girls in B/Mathematics, CSEE 2009: 10.6% girls passed vs. 23.9% boys (Philemon (2010)).

In 1995, 15 ministries of education in southern and east Africa launched a consortium for monitoring education quality, which is popularly known as SACMEQ. South Africa participated in the second study conducted by SACMEQ. "A random sample of 3 416 grade 6 learners from 169 South African public schools was tested in reading (literacy) and mathematics (numeracy). The learners performed particularly poorly in mathematics" (Moloi; undated).

It appears that education authorities in a few African countries have chosen to participate in international student achievement studies as a means of improving teaching and learning in mathematics and science. Two highly regarded such studies are the Trends in International Mathematics and Science Study (TIMSS), and the Programme for International Student Assessment (PISA). While TIMSS is conducted every four years, PISA is conducted every three years.

Few African countries have so far participated in either PISA or TIMSS. Only two African countries have ever participated in PISA during the period 2000-2012, namely Mauritius (2009) and Tunisia (2000 (3) 2012). Participation of African countries in TIMSS was 7 in 1999, 6 in 2003, 5 in 2007 and 7 in 2011 (2011 results will be released in December 2012).

I present in Table 1 below the average scores for eighth-grade students in Singapore and in participating African countries as well as the average international score in mathematics for 1999 (4) 2007.

The data show that students in all African countries scored below the international average and, moreover, the ranking of every African country, except Algeria, has deteriorated over the study years.

If one assumes that participation in international assessments indicate that education authorities in participating countries are seriously concerned about the quality of mathematics education in their respective countries and that they are exerting efforts to improve it, then one might conclude that performance of students in mathematics in other African countries is unlikely to be better than that of their counterparts in participating countries.

Table 2: Average mathematics scores for eighth-grade students in Singapore, participating African countries and for all participating countries: 1999 (4) 2007*.

	1999	2003	2007
Singapore	604 (1)	605 (1)	593 (3)
International average	487	467	500
Algeria			387 (39)
Botswana		366 (42)	364 (43)
Egypt		406 (36)	391 (38)
Ghana		276 (44)	309 (47)
Morocco	337 (37)	387 (40)	
South Africa	275 (38)	264 (45)	
Tunisia	448 (29)	410 (35)	
Number of participating countries	38	48	48

* Ranking of a country is indicated in parentheses

- Language of instruction The language of instruction in many African countries is the colonial language, especially at post-primary levels. However, it is widely believed that the best medium for teaching a child is her/his mother tongue. Yet, as (UNESCO 1953) observes, "it is not always possible to use the mother tongue in school, and, even when possible, some [political, linguistic, educational, socio-cultural, economic, financial, practical] factors may impede or condition its use."

The issue of teaching children mathematics in a language other than their mother-tongue is widely discussed in extant literature due to the perceived gap in academic performance between children with different proficiency level in the language of instruction (for example, Cuevas (1984); Adler (1998); Abedi and Lord (2001); Howie (2003); Zakaria and Abd Aziz (January 2011). The Standards for Educational and Psychological Testing

underscored that for "all test takers, any test that employs language is, in part, a measure of their language skills" (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 1999, p. 91). Thus, if certain students have not yet sufficiently acquired language skills, they may not be able to adequately demonstrate their knowledge in a content-based assessment (Abedi, et al. 2006). Clearly, the language of instruction plays an important role in the performance of school children in mathematics.

3. Research in mathematics education in Africa

The discussion of the previous section demonstrate that mathematics education in Africa faces many challenging problems. Measures and policies for improving the quality of mathematics education in a country must be informed by research. It is of interest to inquire about the level and foci of research in mathematics education in Africa. Resource constraints make it difficult to undertake a comprehensive inquiry and I limit myself in what follows to an inquiry about the level of research in mathematics education in selected African countries. In view of the variations among the selected countries, it is reasonable to assume that the findings apply to most African countries. The level of research output in both mathematics and mathematics education and mathematics in 20 African countries over the period 1980-2010. The regional distribution of selected countries is as follows: 5 (East Africa), 3 (North Africa), 4 (Southern Africa) and 8 (West Africa). The countries show important variations in their level of development, scientific and technological capacity, population size, and the size of their educational systems. As such they may be considered to be representative of the whole continent. The data in the Table show that the annual level of research output in mathematics education during the 31-year period 1980-2010 in most African countries is negligible. For, on average, each country in the Table, excluding South Africa, published about a single paper per a decade. If one considers publications in mathematics, then a contrasting picture emerges. We find, after excluding the four countries with more than 1000 publications during the period of the data (Algeria, Egypt, Nigeria and South Africa), one finds that each of the remaining 16 countries published, on average, about 5 papers every 2 years. What explains this contrasting situation? Significant differences in research capacity or relative neglect of mathematics education, or both? I conclude this section by observing that the low level of research output in both disciplines (each country in the Table averaged just under 19 publications per year during the 31-year period) is perhaps an indication of the fact that mathematical research and mathematics education are organically linked: a weakness in either will undermine the other.

4. Proposed research agenda in mathematics education in Africa

The proposed research agenda in mathematics education in Africa reflect largely the problems presented in section 2 above. While the agenda are not meant to be comprehensive, I claim that they are fundamental to any efforts towards improving the teaching and learning of mathematics in African schools. In view of the vital role of the teacher in formal education, it is natural that our first three proposed items concern the teacher.

4.1 Unqualified teachers

Many African countries face shortages of qualified math teachers, especially in secondary school. While the obvious long-term solution is to increase the supply of trained teachers, there is a considerable delay before such an increase has an impact. Indeed, most countries

Table 3: Level of research output in mathematics education and mathematics in selected African countries: 1980-2010.

Country	Number of publications in Mathematics education	Mathematics
Algeria	2	1174
Benin	0	41
Burkina Faso	1	37
Equatorial Guinea	0	0
Egypt	8	3481
Ethiopia	1	78
Ghana	4	16
Cote d'Ivoire	0	5
Kenya	12	111
Malawi	1	14
Mali	1	7
Nigeria	10	596
Senegal	0	84
South Africa	165	4419
Sudan	4	56
Tanzania	2	47
Tunisia	5	1585
Uganda	4	29
Zambia	2	16
Zimbabwe	4	112
Total (20 countries)	226	
Total (19 countries, excluding South Africa)	61	

Source: Thomson Reuters Web of Science databases.

have little option but to allow recruitment of unqualified teachers. What kind of in-service training is needed to bring them to qualified status?

4.2 Qualified teachers (pre-service programmes)

How are present mathematics teachers being prepared? Is their subject knowledge adequate? Is their pedagogical knowledge adequate? How closely should their mathematics curriculum be aligned to the needs of the classroom (i.e. school mathematics curriculum)?

4.3 Qualified in-service teachers (continuous professional development)

Given the education and experience of qualified practicing teachers, what are appropriate programmes for their continuous professional development? How often should in-service programmes be offered? And where should they be offered? What modes of delivery are effective?

4.4 Mathematics curricula

The need for reform of mathematics curricula is predicated by, among other factors,

- (a) Advances in mathematics (including, how mathematics interacts with other disciplines)
 - (b) Advances in mathematics education (e.g. learning theories)
 - (c) Advances in technology.
- To what extent are mathematics curricula in African education systems influenced by such factors?
 - What is the role of the teacher in curriculum reform?
 - What are the differences between the intended, implemented and achieved curriculum?
 - Does the secondary school mathematics curriculum address the needs of all students adequately?

4.5 Gender

What explains the observation that in many African countries girls are less successful than boys in science-based subjects and are less "keen on" them? How to identify and nurture girls that demonstrate ability in mathematics?

4.6 Language of instruction

I noted in section 2.1 above that the learning of mathematics requires a variety of linguistic skills that second-language learners may not have mastered. Furthermore, special problems of reliability and validity arise in assessing the mathematics achievement of students from a language minority (Cuevas 1984).

- *What is the student's attitude towards the use of an official language as a medium of instruction in learning mathematics?*
- *What is the teacher's attitude towards the use of an official language as a medium of instruction in teaching and learning mathematics?*

- *Are there significant differences in the mathematics performance of official language learners and proficient speakers of the official language?*

It should be obvious from the foregoing that research problems in mathematics education are typically multi-faceted and require an awareness of the complexity of the teaching and learning of mathematics and the surrounding social context. In view of the responsibility of departments of mathematics for the promotion of mathematics in Africa (El Tom, 1984), it cannot be overemphasized that mathematicians should strive to participate actively in this multidisciplinary activity. Indeed, in the context of Africa, mathematics education is too important to be left for non-mathematicians.

References

- [1] Adler, J. (1994). Mathematics teachers in the South African transition. *Mathematics Education Research Journal* 6, 2, 101-112.
- [2] Adler, J. (1998). A language for teaching dilemmas: Unlocking the complex multilingual secondary mathematics classroom. *For the Learning of Mathematics*, 18, 24-33.
- [3] Abedi, J. & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, 14(3), 219-234
- [4] Abedi, J. Courtney, M., Leon, S. Kao, J., and Azzam, T. (2006). *English Language Learners and Math Achievement: A Study of Opportunity to Learn and Language Accommodation* (CSE Report 702, 2006). Los Angeles: University of California, Center for the Study of Evaluation/National Center for Research on Evaluation, Standards, and Student Testing.
- [5] Cuevas, Gilberto J.(1984). Mathematics learning in English as a second language. *Journal for Research in Mathematics*, Vol. 15, No. 2, 134-144
- [6] Developing Countries Strategies Group (2009). *Mathematics in Africa: Challenges and Opportunities. A Report to the John Templeton Foundation. International Mathematical Union.* <http://www.mathunion.org/publications/reports-recommendations>.
- [7] El Tom, M. E. A. (1984). The role of Third World university mathematics institutions in promoting mathematics. Invited paper presented at the 5th International Congress on Mathematical Education, Adelaide, Australia, 24-30 August 1984.
- [8] El Tom, M. E. A. (2008). A new model for building capacity in mathematical research in Africa. Open University, U. K., Working Paper Series No. 342. Gerdes, P. (2007). *African Doctorates in Mathematics: A Catalogue*, Lulu.com
- [9] Goe, L. (2007). The link between teacher quality and student outcomes: A research synthesis. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved June 21, 2012 from [http://www.ncctq.org/publications/Link Between TQ and Student Outcomes.pdf](http://www.ncctq.org/publications/Link%20Between%20TQ%20and%20Student%20Outcomes.pdf)

- [10] Hanushek, E. A. (1995). Interpreting recent research on schooling in developing countries. *The World Bank Research Observer*, 10, 227-46.
- [11] Hanushek, E. A., & Luque, J. A. (2003). Efficiency and equity in schools around the world. *Economics Of Education Review*, 22, 481-502.
- [12] Hanushek, Eric A., and Steven G. Rivkin (2006). "Teacher Quality." In Eric A. Hanushek and Finis Welch, eds., *Handbook of the Economics of Education*. Amsterdam: North Holland.
- [13] Harbison, R. W. & Hanushek, E. A. (1992). *Educational performance of the poor: Lessons from rural northeast Brazil*. New York: Oxford University Press.
- [14] Howie, Sarah J. (2003). Language and other background factors affecting secondary pupils' performance in Mathematics in South Africa. *African Journal of Research in Mathematics Science and Technology Education*, 7:1-20
- [15] Moloi, M. Q. (undated). *Mathematics achievement in South Africa: A comparison of the official curriculum with pupil performance in the SACMEQ II Project*.
- [16] Onderi, H. and Malala, G. (January 2011). A review on extent of sustainability of educational projects: A case of strengthening of mathematics and science in secondary education (SMASSE) projects in Kenya. A review on extent of sustainability of educational projects: A case of strengthening of mathematics and science in secondary education (SMASSE) projects in Kenya. *International Journal of Physical and Social Sciences*, Vol. 2, Issue 1 (Social Sciences <http://www.ijmra.us>).
- [17] Philemon, C. (2010). Development of Science and Mathematics. Paper presented at the Annual Meeting, Mathematical Association of Tanzania, September 2010 (retrieved on 10 May 2012 from <http://www.maths.udsm.ac.tz/mat/DEVELOPMENT%20OF%20SCIENCE%202010%20NEW.pdf>)
- [18] UNESCO (1953). *The use of vernacular languages in education. Monograph on fundamental education VIII*. Paris. UNESCO. (<http://unesdoc.unesco.org/images/0000/000028/002897eb.pdf>)
- [19] UNESCO. 2006. *Teachers and Educational Quality: Monitoring Global Needs for 2015*. Montreal: UNESCO Institute for Statistics.
- [20] Villegas-Reimers, E. (2003). *Teacher professional development: An international review of the literature*. Paris: UNESCO, International Institute for Educational Planning.. Retrieved June 21, 2012 from <http://unesdoc.unesco.org/images/0013/001330/133010e.pdf>
- [21] Zahiah Zakaria and Mohd Sallehudin Abd Aziz (January 2011). *Assessing Students Performance: The Second Language (English Language) Factor*. *International Journal of Educational and Psychological Assessment* January 2011, Vol. 6(2)

Mathematical Competitions for Gifted Students: Organization and Training

by

Paul Vaderlind

Stockholm University, Sweden

What are the competitions?

In addition to regular competitions, problem-solving sessions during a limited time, like national Olympiads or multiple-choice question exams, the World Federation of National Mathematics Competitions has formally defined competitions as including enrichment courses and activities in mathematics, mathematics clubs or "circles", mathematics days, mathematics camps, including live-in programs in which students solve open-ended or research-style problems over a period of days, and other similar activities. These activities all have in common the values of creativity, enrichment beyond the normal syllabus, opportunities for students to experience problem solving situations and provision of challenge for the student. Competitions give students the opportunity to be drawn by their own interest to experience some mathematics beyond their normal classroom experience.

Short history

Among all the methods for identifying gifted students, mathematical competitions probably has the longest and most successful history. The idea of competitions in mathematics goes back to the Hungarian Etvös/Kurschak Contest, 1894. First after forty years later came the St. Petersburg (1934) and Moscow (1935) Mathematical Olympiads. The competitions gained a lot of popularity after the Second World War and resulted among other things in the first International Mathematical Olympiad (1959). The success of the IMO was such that within a few years the number of participating countries grew from 7 to 20. Today more than 100 countries from all continents participate in the IMO and those countries cover more than 85% of the population of our Planet. Many more countries have their national competitions but can't afford sending a team to the IMO. This is a case with many developing countries.

The goals:

There are several goals of competitions in Mathematics.

1. An ultimate method for identifying gifted students,
2. To give students an opportunity to discover a latent talent in mathematics and provide a stimulus for improving learning. Competitions provide opportunity for creativity and independent thinking, as students often solve problems in unexpected and innovative ways.
3. To provide resources for the classroom activities: competitions are an important part of learning mathematics and a fun activity for students of all ages. The success of competitions over the years, particularly the resurgence in the last 50 years, indicates that these are events in which students enjoy mathematics. A long-term objective of the organizing committee of a mathematical contest should definitely be rising of the national education level.

4. To highlight the importance of mathematics: competitions provide a focus on problem solving, sometimes giving students an opportunity to be associated with a cutting edge area of mathematics in which new methods may evolve and old methods be revived.

The practice.

Competitions come in a number of categories:

1. Local competitions on a school level, community level or town.
2. Provincial competitions within a country, which often are a part of more general national Olympiad.
3. National mathematical Olympiads.
4. Regional Olympiads, like Baltic Way Mathematical Contest, Asian-Pacific Mathematical Olympiad, Balkan Olympiad, Pan African Math Olympiad and so on.
5. International contests: IMO, Tournament of Towns, Kangaroo Mathematical Contest.

Other categories:

6. Competitions for girls only: China Girls' Math Olympiad and European Girls' Math Olympiad.
7. Team competitions: Baltic Way Team Competition and even contests involving whole classes, giving a very different feel to the competition.
8. Competitions for Primary schools and competitions for University students.

Competitions today

As we mentioned earlier, most countries have a permanent competitions activities although very often those activities are limited to at most national level. Most of the time the reason is lack of funds for travels and for training camps. However, the recent development shows that more and more private companies (banks, investment corporations, telephone companies and internet providers) discover a need of skilled, well-educated co-workers and are willing to sponsor different elite-search events, one of which is obviously mathematical competitions.

The questions offered at the competitions are most of the time non-standard problems being non-routine, provocative, fascinating, and challenging, often with elegant solutions. The topics assume little prior knowledge beyond school curriculum and covers most of the school mathematics: geometry, trigonometry, algebra, inequalities, number theory and combinatorics.

Organizing a competition:

1. An organizing committee. Preferably consisting of a group of University teachers and a group of Secondary school teachers.

2. Getting acquainted with the "competitional mathematics". This usually goes beyond the secondary schools curriculum; demands some (accessible) knowledge and a good part of creativity. There are hundreds of books and numerous websites with different kind of competitions on different levels. Some help may be received from mathematicians coming from countries with a long tradition in organizing olympiads.
3. Preparing a competition. Could be a competition covering only some schools of only one city (the capital) or a number of cities and slowly, in the following years, extending it to the whole country.
4. Getting in touch with at least one teacher of mathematics in each (if possible) school in the country and prepare him/her for arranging a competition.
5. The first stage could be a multiple-choice questions. This is easily marked by the teacher and the results are then send to the national committee. In smaller countries, like Sweden, the papers are marked by the national committee during a weekend-long working session.

The best students may be then selected for the next stage. For example 20-50 students. It may be a provincial competition or already a national final. It is important however that at this stage the questions demand a full solution, not a multiple-choices alternatives.

6. Training of the most successful and promising students for further, international competitions.
7. Participating in an International regional competition, for example PAMO, or creating smaller events, like East African Mathematical Challenge. It doesn't have to involve travels (the students, up to 10 from each country, can work in their schools, but the papers may be marked by one "hosting country", which may vary from year to year.

This year PAMO will take place September 8-16 this year in Tunisia. The country registered that far are Mali, Tunisia, Burkina Faso, Algeria, Tanzania, Kenya, Gambia, Cte d'Ivoire, Nigeria, Egypt and South Africa.

www.pamo – official.org

8. IMO - the queen of all competitions. In the latest one, in Argentina, July 2012, participated 100 countries from all over the world, but only six from Africa (Uganda, Ivory Coast, South Africa, Nigeria, Tunisia and Morocco). Next IMO will take place in Colombia (2013) and then in Cape Town (2014), for the first time on the African soil.

www.imo – official.org

www.artofproblemsolving.com

Estimation of IBNR Claims Reserves Using Linear Models

by

Patrick G. O. Weke

School of Mathematics, University of Nairobi, Kenya

Abstract

Stochastic models for triangular data are derived and applied to claims reserving data. The standard actuarial technique, the chain ladder technique is given a sound statistical foundation and considered as a linear model. The chain ladder technique and the two-way analysis of variance are employed for purposes of estimating and predicting the IBNR claims reserves.

1. Introduction

If claims runoff triangles are to be analysed statistically, as a data analysis exercise, it is desirable to express them as linear models. If the claims are analysed using a model for each row, then it may be straightforward to write down a linear model. The use of linear models to analyse the data by row can give useful insights into the nature of the data, but it is the linear model which is close to the chain ladder technique that is of greatest interest to actuaries. This linear model, whose connection with the chain ladder technique was first identified by Kremer[4] is described in sections 3 and 4.

The data are assumed to be lognormally distributed and is first logged before a linear model is applied. The transformation from the raw data to the logged data is, obviously, straightforward, but the reverse transformation, once the analysis has been carried out, is not simple. This is dealt with in section 5. The process is represented in Figure 1.

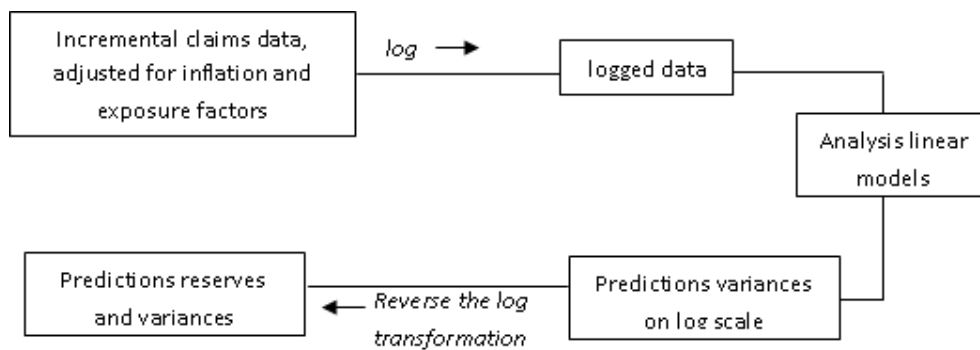


Figure 1:

Prediction from linear models when the data are lognormally distributed was first considered by Finney [3]. Finney considered a sample of independently, identically distributed data, and the theory was generalized to a sample of independently, but not necessarily identically distributed data by Bradu and Mundlak [2]. Subsequent papers by Renshaw [5], Verrall [6], and Weke [8]

have considered the properties of the estimators in more detail. The techniques outlined in this paper have been implemented in GLIM (Baker and Nelder [1] and the results shown.

2. Linear Models

The linear model to be considered is

$$y = X\beta + \epsilon \quad (2.1)$$

where y is a data vector of length n , β is an $n \times p$ design matrix and ϵ is an error vector of length n . The error vector ϵ is assumed to have mean zero and variance-covariance matrix Σ .

The minimum variance linear unbiased estimators of the parameters, β , are the weighted least-squares estimators, $\hat{\beta}$, where

$$\hat{\beta} = (X'\Sigma^{-1}X)^{-1} X'\Sigma^{-1}y \quad (2.2)$$

If the errors, ϵ , are assumed to be jointly normally distributed, then the estimators, $\hat{\beta}$, are also the maximum likelihood estimators. Since a logarithmic transformation will be applied to the data, the reverse transformation to estimate actual claims will depend on the estimation method being used. One estimator can be obtained by simply substituting the estimators into the equations. This is used in the lemmas which show the similarity between the chain ladder technique and a certain linear model. However, these estimators, and indeed the maximum likelihood estimators, are biased, and it may be better to use unbiased estimators. If the errors are assumed to be uncorrelated with equal variance then equation (2.1) simplifies to

$$\hat{\beta} = (X'X)^{-1}X'y \quad (2.3)$$

which is a form which will also be used.

The distributional properties of the maximum likelihood estimators, $\hat{\beta}$, are well-known. Assuming that the errors are independently, identically distributed with variance σ^2 ,

$$\hat{\beta} \sim N(\beta, \sigma^2(X'X)^{-1}) \quad (2.4)$$

3. The Chain Ladder Technique as a Linear Model

Kremer [4] showed that the chain ladder technique is very similar to a two-way analysis of variance and investigated the properties of the estimators. This section describes the connection between the actuarial chain ladder technique and the statistical analysis of variance method. Assuming a triangular data set (without loss of generality) the cumulative claims data, to which the chain ladder technique is applied, are

$$\{C_{ij} : i = 1, \dots, t; j = 1, \dots, t - i + 1\} \quad (3.1)$$

The differenced data, to which the analysis of variance model is applied, are

$$\{Z_{ij} : i = 1, \dots, t; j = 1, \dots, t - i + 1\} \quad (3.2)$$

where

$$\begin{aligned} Z_{ij} &= C_{ij} - C_{i,j-1}, \quad j \geq 2 \\ Z_{i,1} &= C_{i1} \end{aligned}$$

The chain ladder technique is based on the model

$$E[C_{ij}] = \lambda_j C_{i,j-1}; \quad j = 2, \dots, t. \quad (3.3)$$

The parameter λ_j is estimated by $\hat{\lambda}_j$, where

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{i-j+1} C_{ij}}{\sum_{i=1}^{i-j+1} C_{i,j-1}} \quad (3.4)$$

The expected ultimate loss, $E[C_{ij}]$, is estimated by multiplying the latest loss, $C_{i,i-j+1}$, by the appropriate estimated λ -values:

$$\text{estimate of } E[C_{ij}] = \left(\prod_{j=t-i+2}^t \hat{\lambda}_j \right) C_{i,i-j+1} \quad (3.5)$$

The chain ladder technique produces forecasts which have a row effect and a column effect. The column effect is obviously due to the parameters $\{\lambda_j : j = 2, \dots, t\}$. There is also a row effect since the estimates for each row depend not only on the parameters $\{\lambda_j : j = 2, \dots, t\}$, but also on the row being considered. The latest cumulative claims, $C_{i,t-i+1}$, can be considered as the row effect. This leads to consideration of other models which have row and column effects, in particular the two-way analysis of variance model. The connection is first made with a multiplicative model (see [7]). This uses the non-cumulative data, Z_{ij} , and models them according to:

$$E[Z_{ij}] = U_i S_j \quad (3.6)$$

where U_i is a parameter for row i , and S_j is a parameter for row j .

A multiplicative error structure is assumed and also

$$\sum_{j=1}^t S_j = 1 \quad (3.7)$$

In this model, S_j is the expected proportion of ultimate claims which occur in the j th development year; and U_i is the expected total ultimate claim amount for business year i (neglecting any tail factor). The estimates of U_i will be compared with the estimates of $E[C_{it}]$ in equation (3.5) and S_j and λ_j will be related to each other.

The analysis of variance estimators are based on the model (3.6):

$$E[Z_{ij}] = U_i S_j$$

and the chain ladder technique is based on the model (3.3):

$$E[C_{ij}] = \lambda_j C_{i,j-1}; \quad j = 2, \dots, t.$$

In terms of the models, ignoring for the moment the estimation of the parameters, this simply represents a reparameterisation.

Under the chain ladder model, the expected claim total for business year is

$$\prod_{j=t-i+2}^t \lambda_j C_{i,t-i+1} \quad (3.8)$$

and the expected claim amount in development year $t - i + 2$ is

$$\lambda_{t-i+2} C_{i,t-i+1} - C_{i,t-i+1} \quad (3.9)$$

The equivalent quantities under the multiplicative model (3.6) are

$$U_i \quad (3.10)$$

$$\text{and } U_i S_{t-i+2} \quad (3.11)$$

Equating (3.8) and (3.9) with (3.10) and (3.11), respectively, gives

$$S_{t-i+2} = \frac{\lambda_{t-i+2} - 1}{\prod_{j=t-i+2}^t \lambda_j}$$

The expected claim amount for development year $t - i + 3$ under each model is

$$\lambda_{t-i+3} \lambda_{t-i+2} \lambda_{t-i+1} - \lambda_{t-i+2} C_{i,t-i+1} \quad (3.12)$$

and

$$U_i S_{t-i+3} \quad (3.13)$$

which gives

$$S_{t-i+3} = \frac{\lambda_{t-i+3} - 1}{\prod_{j=t-i+3}^t \lambda_j}$$

In general, the expected proportion of ultimate claims can be written in the form

$$S_j = \frac{\lambda_j - 1}{\prod_{l=j}^t \lambda_l} \quad (3.14)$$

Considering year of business t , the expected total claim amount under each model is

$$\left[\prod_{j=2}^t \lambda_j \right] C_{t1}$$

and U_t .

The claim amount in development year 1, C_{t1} , is modeled by $U_t S_1$, and so it can be seen that

$$S_1 = \frac{1}{\prod_{l=1}^t \lambda_l} \quad (3.15)$$

To summarize, the chain ladder model (3.3) is equivalent to the multiplicative model given by equation (3.6) with the following relationships between the parameters:

$$\begin{aligned} S_1 &= \left(\prod_{l=2}^t \lambda_l \right)^{-1} \\ S_j &= \left(\prod_{l=2}^t \lambda_l \right)^{-1} (\lambda_j - 1) \\ U_i &= E(C_{it}). \end{aligned}$$

Equations (3.4) and (3.5) give the estimates of $\{\lambda_j : j = 2, \dots, t\}$ and $E(C_{it})$. Estimators of $\{S_i : i = 2, \dots, t\}$ and $\{U_j : j = 2, \dots, t\}$ can be obtained by applying a linear model to the logged incremental claims data. Taking logs of both sides of equation (3.6), and assuming that the incremental claims are positive, results into

$$E(Y_{ij}) = \mu + \alpha_i + \beta_j \quad (3.16)$$

where $Y_{ij} = \log Z_{ij}$ denotes the cumulative claims in development year j in respect of accident year i , and the errors now have an additive structure and are assumed to have mean zero. The errors will be assumed to be identically distributed with variance σ^2 , although this distributional assumption can be relaxed. Kremer [4] defines as the mean of the $\log U_i$ s and $\log S_j$ s, so that the restriction

$$\sum_{i=1}^t \alpha_i = \sum_{j=1}^t \beta_j = 0$$

is imposed.

An alternative assumption is that $\alpha_1 = \beta_1 = 0$. In this case

$$\alpha_i = \log U_i - \log U_1 \quad (3.17)$$

$$\beta_j = \log S_j - \log S_1 \quad (3.18)$$

$$\mu = \log U_1 + \log S_1 \quad (3.19)$$

The latter set of assumptions are more appropriate for the more sophisticated techniques. However, prediction and estimation of the claims is unaffected by the choice of the assumptions.

The assumption that error terms, ε_{ij} , are independently, identically distributed with variance σ^2 will be used, so that the estimators are given by equation (2.3) Now equation (3.16) can be written in the form of equation (2.1). Suppose, for example, there are three years of data then

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{13} \\ y_{22} \\ y_{31} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_2 \\ \alpha_3 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{21} \\ \varepsilon_{13} \\ \varepsilon_{22} \\ \varepsilon_{31} \end{bmatrix} \quad (3.20)$$

clearly gives the form of the parameter vector and the design matrix.

The following lemma, due to Kremer [4], gives the normal equations for the chain ladder linear model.

Lemma 3.1 For n years data, the best linear unbiased estimators of μ, α_i, β_j are the solutions of

$$\hat{\alpha}_i = \frac{1}{t-i+1} \sum \left(Y_{ij} - \frac{1}{t-j+1} \sum_{l=1}^{t-j+1} (Y_{1l} - \hat{\alpha}_l) \right); i = 2, \dots, t \quad (3.21)$$

$$\hat{\beta}_j = \frac{1}{t-j+1} \sum \left(Y_{ij} - \frac{1}{t-j+1} \sum_{l=1}^{t-j+1} (Y_{1l} - \hat{\beta}_l) \right); j = 2, \dots, t \quad (3.22)$$

$$\hat{\mu} = \frac{1}{t(t+1)} \sum_{i=1}^t \sum_{j=1}^{t-i+1} (Y_{ij} - \hat{\alpha}_i - \hat{\beta}_j). \quad (3.23)$$

Proof:

The normal equations, (2.3), are

$$(t-i+1)\hat{\mu} + (t-i+1)\hat{\alpha}_i + \sum_{j=2}^{t-i+1} \hat{\beta}_j = \sum_{j=1}^{t-i+1} Y_{ij}; i = 2, \dots, t \quad (3.24)$$

$$(t-i+1)\hat{\mu} + \sum_{j=2}^{t-i+1} \hat{\alpha}_i + (t-j+1)\hat{\beta}_j = \sum_{j=1}^{t-i+1} Y_{ij}; i = 2, \dots, t \quad (3.25)$$

$$\frac{t(t+1)}{2}\hat{\mu} + \sum_{i=2}^t (t-i+1)\hat{\alpha}_i + \sum_{j=2}^t (t-j+1)\hat{\beta}_j = \sum_{i=1}^t \sum_{j=1}^{t-i+1} Y_{ij} \quad (3.26)$$

Noting that $\hat{\alpha}_1 = \hat{\beta}_1 = 0$, equations (3.26) and (3.23) are equivalent. Also equations (3.24) and (3.25) can be written as

$$\hat{\alpha}_i = \frac{1}{t-i+1} \sum_{j=1}^{t-i+1} (Y_{ij} - \hat{\beta}_j) - \hat{\mu} \quad (3.27)$$

$$\hat{\beta}_i = \frac{1}{t-i+1} \sum_{j=1}^{t-i+1} (Y_{ij} - \hat{\alpha}_j) - \hat{\mu} \quad (3.28)$$

Substituting equation (3.27) into equation (3.28) and vice versa gives equations (3.21) and (3.22). \square

4. Relationship between the Estimators of the Linear Model and the Chain Ladder Model

The previous section derived the relationship between the parameters of the multiplicative model and the chain ladder technique. The parameters are estimated in different ways according to which method is used, and this section is devoted to examining the relationships between the estimators of the parameters.

This section contains two lemmas. The first deals with the estimation of S_j and U_i - the parameters of the multiplicative model using the chain ladder technique. The second lemma derives the estimators of S_j and U_i using the two-way analysis of variance model. The two sets of estimators are then shown to be similar. Thus, it will be shown that the chain ladder method will

produce results which are similar to those produced by the analysis of variance method. The latter has been studied in great depth in statistical literature and the method has the advantage of a great deal of theoretical background. The theory of analysis of variance will be applied to insurance data, bearing in mind that the main method in use in the industry is the chain ladder method.

Lemma 4.2 *If*

$$S_j = \frac{\lambda_j - 1}{t}; j = 2, \dots, t \quad (4.1)$$

$$\prod_{l=j} \lambda_l$$

and λ_j is estimated by $\tilde{\lambda}_j$, where

$$\tilde{\lambda}_j = \frac{\sum_{i=1}^{t-j+1} C_{ij}}{\sum_{i=1}^{t-j+1} C_{ij-1}} \quad (4.2)$$

then the estimators of S_j, \tilde{S}_j , satisfy the relationship

$$\tilde{S}_j = \frac{\sum_{i=1}^{t-j+1} Y_{ij}}{\sum_{i=1}^{t-j+1} C_{i,t-i+1} / \left(1 - \sum_{l=t-i+2}^t \tilde{S}_l\right)} \quad (4.3)$$

Also, the estimate of U_i is \tilde{U}_i , where

$$\tilde{U}_i = \frac{\sum_{j=1}^{t-i+1} Z_{ij}}{\sum_{j=1}^{t-i+1} \tilde{S}_j} \quad (4.4)$$

Proof:

Equations (4.1) and (4.2) imply that

$$\tilde{S}_j = \frac{\tilde{\lambda}_j - 1}{t} = \frac{\sum_{i=1}^{t-j+1} C_{ij} - \sum_{i=1}^{t-j+1} C_{ij-1}}{\sum_{i=1}^{t-j+1} C_{ij} \prod_{l=1}^t \tilde{\lambda}_l} \quad (4.5)$$

Now, it can be shown by induction that (see [4])

$$\sum_{i=1}^{t-j+1} C_{ij} = \sum_{i=1}^{t-j+1} C_{i,t-i+1} / \prod_{l=j+1}^{t-i+1} \tilde{\lambda}_l \quad (4.6)$$

Substituting equation (4.6) into equation (4.5) gives

$$\tilde{S}_j = \frac{\sum_{i=1}^{t-j+1} Z_{ij}}{\left(\sum_{i=1}^{t-j+1} C_{i,t-i+1} / \prod_{l=j+1}^{t-i+1} \tilde{\lambda}_l \right) \prod_{l=j+1}^t \tilde{\lambda}_l} = \frac{\sum_{i=1}^{t-j+1} Z_{ij}}{\sum_{i=1}^{t-j+1} C_{i,t-i+1} \prod_{l=t-i+2}^t \tilde{\lambda}_l} \quad (4.7)$$

It can also be shown by induction that

$$\left[\prod_{l=k}^t \lambda_l \right]^{-1} = 1 - \sum_{l=k}^t S_l.$$

This is true for $k = 2$ by virtue of (3.15) and the relationship

$$1 - \sum_{l=2}^t S_l = S_1.$$

Suppose it is true for k . Then for $k + 1$:

$$\begin{aligned} 1 - \sum_{l=k+1}^t S_l &= 1 - \sum_{l=k}^t S_l + S_k \\ &= \left[\prod_{l=k}^t \lambda_l \right]^{-1} + \frac{\lambda_k - 1}{\prod_{l=k}^t \lambda_l} = \left[\prod_{l=k+1}^t \lambda_l \right]^{-1} \end{aligned} \quad (4.8)$$

Hence, by induction, the result holds. Substituting this result into (4.7) gives

$$\tilde{S}_j = \frac{\sum_{i=1}^{t-j+1} Z_{ij}}{\sum_{i=1}^{t-j+1} C_{i,t-i+1} / \left(1 - \sum_{l=t-i+2}^t \tilde{S}_l \right)} \quad (4.9)$$

as required.

Now, since $C_{i,t-i+1} = \sum_{j=1}^{t-i+1} Z_{ij}$

and $\prod_{j=t-1+2}^t \tilde{\lambda}_j = \left(1 - \sum_{j=t-i+2}^t \tilde{S}_j \right)^{-1}$ the estimate of total expected outstanding claims for row i ,

$$C_{i,t-i+1} \prod_{j=t-i+2}^t \tilde{\lambda}_j$$

can be written as
$$\frac{\sum_{j=1}^{t-i+1} Z_{ij}}{1 - \sum_{j=t-i+2}^t \tilde{S}_j}.$$

This can be written as

$$\sum_{j=1}^{t-i+1} Z_{ij} \Big/ \sum_{j=1}^{t-i+1} \tilde{S}_j \quad (4.10)$$

since $1 - \sum_{j=t-i+2}^t \tilde{S}_j = 1 - \sum_{j=i}^{t-i+1} \tilde{S}_j.$ □

Lemma 4.3 *Using the estimation method of Lemma 3.1, an estimate of total expected claims for accident year i , \hat{U}_i , is given by*

$$\hat{U}_i = \left[\prod_{j=1}^{t-i+1} \frac{Z_{ij}}{w_j} \right]^{\frac{1}{t-i+1}} \cdot \sum_{j=1}^{t-i+1} w_j \quad (4.11)$$

where

$$w_j = \frac{\left[\prod_{j=1}^{t-i+1} Z_{ij} \right]^{\frac{1}{t-j+1}}}{\left[\prod_{i=1}^{t-j+1} \left(\prod_{l=1}^{t-i+1} Z_{il} \right)^{\frac{1}{t-i+1}} \Big/ \left(\prod_{l=1}^{t-i+1} w_l \right)^{\frac{1}{t-i+1}} \right]^{\frac{1}{t-j+1}}} \quad (4.12)$$

Further,

$$\hat{U}_i = \frac{\left[\prod_{j=1}^{t-i+1} Z_{ij} \right]^{\frac{1}{t-i+1}}}{\left[\prod_{j=1}^{t-i+1} \hat{S}_j \right]^{\frac{1}{t-i+1}}} \quad (4.13)$$

This lemma can be used to show that the estimates of expected total outstanding claims for each row have similar forms using each method, and can be expected to behave in similar ways. The estimate of U_i is obtained by “hatting” the parameters in the identity

$$U_i = e^{\alpha_i} e^{\mu} \sum_{j=1}^t e^{\beta_j}$$

which is derived in the proof of this lemma. The resulting estimate of U_i is not the maximum likelihood estimate, neither is it unbiased, but it does serve the purpose of illustrating the similarity between the chain ladder technique and the two-way analysis of variance.

Proof: The equations (3.17) to (3.19) imply that

$$e^{\alpha_i} = \frac{U_i}{U_1} \quad (4.14)$$

$$e_j^\beta = \frac{S_j}{S_1} \quad (4.15)$$

and

$$e^\mu = U_1 S_1. \quad (4.16)$$

Since $\sum_{j=1}^t S_j = 1$, $S_1 = \left(\sum_{j=1}^t e^{\beta_j} \right)^{-1}$.

This, together with equations (4.14) and (4.16) gives

$$U_i = e^{\alpha_i} e^\mu \sum_{j=1}^t e^{\beta_j} \quad (4.17)$$

Now let $w_j = e^{\beta_j}$; then equation (3.22) is equivalent to equation (4.12).

The best linear unbiased estimate of $\alpha_i + \mu$ is obtained from equation (3.27). Substituting the estimates of $\alpha_i + \mu$ and β_j into equation (4.17) gives the estimate of U_i in equation (4.11).

Now, equation (4.15) implies that $\hat{S}_j = w_j / \sum_{l=1}^t w_l$ and so equations (4.11) and (4.12) can be written as

$$\hat{U}_i = \frac{\left[\prod_{j=1}^{t-i+1} Z_{ij} \right]^{\frac{1}{t-i+1}}}{\left[\prod_{j=1}^{t-i+1} \hat{Z}_j \right]^{\frac{1}{t-i+1}}}$$

and

$$\hat{S}_j = \frac{\left[\prod_{j=1}^{t-i+1} Z_{ij} \right]^{\frac{1}{t-j+1}}}{\left[\prod_{i=1}^{t-j+1} \left(\prod_{l=1}^{t-i+1} Z_{il} \right)^{\frac{1}{t-i+1}} / \left(\prod_{l=1}^{t-i+1} \hat{S}_l \right)^{\frac{1}{t-i+1}} \right]^{\frac{1}{t-j+1}}} \quad (4.18)$$

Now, if all the geometric means are replaced by arithmetic means in equation (4.18), the recurrence relation for the estimators of becomes the same as that in Lemma 4.2. Similarly the estimators of U_i are equivalent if geometric means are replaced by arithmetic means. Thus the two estimation methods, the chain ladder method and the linear model, will produce similar results. The structure of the models is identical and the only difference is the estimation technique. It can be argued that the linear model estimates are best in a statistical sense, but it

should be emphasized that in using the linear model instead of the crude chain ladder technique, there are no radical changes.

5. Unbiased Estimation of Reserves and Variances of Reserves

It has been shown that the chain ladder can be considered as a two-way analysis of variance. This linear model, and other linear models, can be used effectively for analyzing claims data and producing estimates of expected total outstanding claims for each year of business. The methods have in common the assumption that the data is lognormally distributed, and the linear models are therefore applied to the logged incremental claims rather than the raw incremental claims data. The problem therefore arises of reversing the log transformation to produce estimates on the original scale. It is this problem which is addressed in this section; in particular the unbiasedness of the estimates is considered. It is important that estimates should be unbiased in order that they are aiming at the correct target and do not yield values which consistently under- or over-estimate. It is also important to consider unbiased estimation of the standard error of the estimates of expected total outstanding claims, in order that some measure of the order of the errors can be attached to the predictions. The procedure for analyzing claims data using loglinear models is illustrated by Figure 1.

The final stage in this procedure reversing the log transformation is considered here and unbiased estimates of total outstanding claims are derived. Unbiased estimates of the variances of these estimates are derived. The theory is applied to claims data (obtained from [7]) using the analysis of variance linear model and the unbiased estimates compared with some alternatives. In order to make the analysis more easily assimilable, a sample of independently, identically distributed observations is considered first. The theory is then extended to the more general case of independent, but not necessarily identically distributed observations. It is the more general theory which is applicable to claims data.

5.1 Unbiased Estimates of Total Outstanding Claims

The purpose of the analysis of the claims data is to produce estimates of the expected total outstanding claims, R_i , for each year of business, and the total outstanding claims, R , for the whole triangle.

An unbiased estimate of R_i is \hat{R}_i , where

$$\hat{R}_i = \sum_{j=t-i+2}^t \hat{\theta}_{ij}. \quad (5.1)$$

and

$$\hat{\theta}_{ij} = \exp(\underline{X}_{ij}\hat{\beta} + \hat{\sigma}^2) \quad (5.2)$$

is the maximum likelihood estimate of the expected value of the lognormally distributed data, θ_{ij} , which is related to the mean and variance of the normally distributed data by

$$\theta_{ij} = \exp(\underline{X}_{ij}\beta + \sigma^2/2).$$

The variance of \hat{R}_i can be calculated from

$$\text{Var}(\hat{R}_i) = \text{Var} \left[\sum_{j=t-i+2}^t \hat{\theta}_{ij} \right] = \sum_{j=t-i+2}^t \left[\text{Var}(\hat{\theta}_{ij}) + 2 \sum_{k=j+1}^t \text{Cov}(\hat{\theta}_{ij}, \hat{\theta}_{ik}) \right]. \quad (5.3)$$

By extending the limits of the summations, the total outstanding claims for the whole triangle can also be considered.

5.2 Prediction Intervals

Having found an unbiased estimate of total outstanding claims, it is now possible to produce a prediction interval for total outstanding claims. The purpose of the analysis so far has been to produce an estimate of total outstanding claims and an estimate of the variance of this estimate. It is often desirable to find a safe value which is unlikely to be exceeded by the total actual claims.

Let R = total outstanding claims for the whole triangle, and \hat{R} be an unbiased estimate of $E(R)$.

Suppose that a $(1 - \alpha) \times 100\%$ upper confidence bound on total claims, R , is required, then it can be found from

$$\hat{R} + Z_{\alpha/2} \sqrt{\text{Var}(R) + \text{Var}(\hat{R})} \quad (5.4)$$

where $\sqrt{\text{Var}(R) + \text{Var}(\hat{R})}$ is the root mean square error of prediction, and an unbiased estimate is used.

6. Numerical Example

This example illustrates and compares the two methods of claims reserving considered in this paper: the chain ladder method and the two-way analysis of variance. For the analysis of variance model, both the unbiased and maximum likelihood estimates of outstanding claims are given. The data used is that from [7]. The estimates of the parameters in the analysis of variance model and their standard errors are shown in Table 1.

The standard errors are obtained from the estimates of the estimates of the variance-covariance matrix of the parameter estimates:

$$(X'X)^{-1}X\hat{\sigma}^2$$

where $\hat{\sigma}^2$ is the estimate of the residual variance. For example, . Since the data is in the form of a triangle (there are the same number of rows and columns) and the matrix X is based solely on the design matrix, the standard errors are the same for each row and column parameter.

The row parameters are contained within a much smaller range than the column parameters: (0.149, 0.673) compared with (-1.393, 0.965). It is to be expected that the row parameters should be contained within a fairly small range, since the rows are expected to be similar. Any pattern in the row parameters gives an insight into, and depends upon, the particular claims experience. It is thus quite common to observe that the row parameters lie in a small range, but not typical that they follow a trend.

The fitted values for the analysis of variance model are shown in Table 2. These are unbiased estimates and are shown with the actual observations for comparison. In this table, the top entries are the estimates and those underneath are the actual observations.

Of most interest to practitioners are the predicted outstanding claims for each year of business, which are the row totals of predicted values. Table 3 shows the maximum likelihood predictions of the outstanding claims in the lower triangle, and Table 4 shows the unbiased predictions. The method does not produce any predictions for the first row, and each row contains one more predicted value.

Table 3: Maximum likelihood predictions of outstanding claims

								101269
							357398	93599
					217465	319835		83761
				335047	243001	357392		93597
			386433	345088	250283	368102		96402
		617309	418743	373941	271209	398880		104462
	1206369	674243	457364	408430	296223	435668		114097
1026594	1053911	589034	399564	356813	258787	380610		99678
888831	913640	937951	524224	355600	317554	230313	338732	88710

Table 4: Unbiased predictions of outstanding claims

								96238
							350362	88841
					215218	313105		79394
				332848	240075	349268		88564
			384305	342028	246696	358900		91006
		613257	415031	369373	266419	387593		98281
	1193906	666126	450811	401216	289387	421005		106752
1006382	1031734	575643	389575	346716	250077	363813		92248
844677	867203	889047	496032	335695	298762	215487	313486	79483

Table 1: Estimates of the row and column and their standard errors

	Estimate	Standard Error
Overall mean	6.106	0.165
Row parameters	0.194	0.161
	0.149	0.168
	0.153	0.176
	0.299	0.186
	0.412	0.198
	0.508	0.214
	0.673	0.239
	0.495	0.281
	0.602	0.379
Column parameters	0.911	0.161
	0.939	0.168
	0.965	0.176
	0.383	0.186
	-0.005	0.198
	-0.118	0.214
	-0.439	0.239
	-0.054	0.281
	-1.393	0.379

It can be seen that the maximum likelihood estimates are all higher than the unbiased estimates, as was to be expected.

The total predicted outstanding claims for each year of business (the row totals of the predicted outstanding claims) are shown in Table 5. There are three estimates given, the maximum likelihood and unbiased estimates from the analysis of variance model, and the chain ladder estimate.

It can be seen that the maximum likelihood estimates differ most significantly from the unbiased estimates in the early and late rows. The estimates for the middle rows are the closest together, which is where the number of observations are used in the estimation is the greatest. The maximum likelihood estimate is asymptotically unbiased, and the greater the number of observations used to estimate the parameters, the closer are the two. The chain ladder estimates are sometimes higher and sometimes lower than the analysis of variance estimates. There is nothing significant that can be inferred from the differences. This confirms that the crude chain ladder method is a reasonable rough and ready method for calculating outstanding claims, although the more proper method, statistically, is the analysis of variance method (using unbiased estimation).

The total predicted outstanding claims are:

Analysis of Variance	Maximum Likelihood	18186154
	Unbiased	17652064
Chain Ladder		18619916

Table 2: Fitted values and the actual observations

286170	711785	731359	750301	418911	283724	252756	182559	266237	67948
357848	766940	610542	482940	527326	574398	146342	139950	227229	67948
410587	1021245	1049329	1076506	601040	407078	362646	261930	381987	
352118	884021	933894	1183289	445745	320996	527804	266172	425046	
379337	943516	969461	994572	555294	376094	335044	241994		
290507	1001799	926219	1016654	750816	146923	495992	280405		
339233	843767	866971	889425	496588	336334	299624			
310608	1108250	776189	1562400	272482	352053	206286			
378676	941872	967773	992840	554327	375439				
443160	693190	991983	769488	504851	470639				
389421	968599	995234	1021012	570056					
396132	937085	847498	805037	705960					
420963	1047052	1075844	1103710						
440832	847631	1131398	1063269						
457887	1138894	1170213							
359480	1061648	1443370							
396651	986582								
376686	986608								
344014									
344014									

Table 6 below shows the unbiased estimates of the total outstanding claims for each year of business, the standard errors of these estimates and the root mean square error of prediction. This table can be used in setting safe reserves, and gives an idea of the likely variation of outstanding claims.

The unbiased estimate of total outstanding claims is 17652064 and the root mean square error of prediction is 2759258. Thus a 95% upper bound on total outstanding claims is

$$17652064 + 1.645 \times 2759258 = 22191043.$$

This is a safe reserve for this triangle according to the chain ladder linear model using unbiased estimation.

Table 5: Total predicted outstanding claims

Row	Analysis of Variance		
	Maximum Likelihood	Unbiased	Chain Ladder
2	101269	96238	94630
3	450997	439203	464668
4	621061	607717	702101
5	1029037	1010755	965576
6	1446307	1422934	1412202
7	2184544	2149953	2176089
8	3592393	4520202	3897142
9	4164990	4056189	4289473
10	4595556	4339873	4618035

Table 6: Unbiased estimates, standard errors and root MSE for each year

Unbiased Estimate	Standard Error	Mean Square Error of prediction
96238	35105	47202
439203	108804	163217
607717	127616	182847
1010755	195739	269224
1422934	273082	357593
2149953	429669	538533
3529202	775256	942851
4056189	1052049	1197009
4339873	1534943	1631306

7. Conclusion

In conclusion, some practical aspects of claims reserving have been considered. These are the stability of the estimation and predictions, the use of the predictions, their standard errors and the safe reserves in practice. The connection between the linear model and the chain ladder technique has been outlined.

References

- [1] Baker, R. J. & Nelder, J. A. (1978). *The GLIM System: Release 3*, Numerical Algorithms Group, Oxford.
- [2] Bradu, D. & Mundlak, Y. (1970). Estimation in Lognormal Linear Models. *JASA*, Vol. 6, 198-211.
- [3] Finney, D. J. (1942). On the Distribution of a Variate whose Logarithm is Normally Distributed. *JRSS Suppl.*, 7, 155-161.
- [4] Kremer, E. (1982). IBNR Claims and the Two-Way Model of ANOVA. *Scand. Act. J.*, Vol. 1, 47-55.
- [5] Renshaw, A. E. (1989). Chain Ladder and Interactive Modelling: Claims Reserving and GLIM. *Journal of the Institute of Actuaries*, Vol. 116, Part 3, 559-587.
- [6] Verrall, R. J. (1989). A State Space Representation of the Chain Linear Model. *Journal of the Institute of Actuaries*, Vol. 116, Part 3, 589-609.
- [7] Weke, P. G. O. (2003). Estimating IBNR Claims Reserves using Gamma Model and GLIM. *The Nigerian Journal of Risk and Insurance*, Vol. 4, No. 1, 1-11.
- [8] Weke, P. G. O. & Mureithi, A. T. (2006). Deterministic Claims Reserving in Short-Term Insurance Contracts. *The East African Journal of Statistics*, Vol. 1, No. 2, 198-213.

Heat Transfer Analysis of a Convecting and Radiating Two Step Reactive Slab

by

O. D. Makinde

Institute for Advanced Research in Mathematical Modeling and Computation,
Cape Peninsula
University of Technology, P.O. Box 1906, Bellville 7535, South Africa

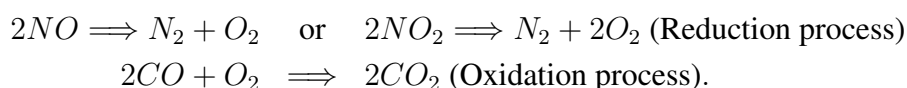
Abstract

This paper examined the heat transfer characteristic of a steady state convecting and radiating two step exothermic reactive slab of combustible materials, taking the diffusion of the reactant into account and assuming a variable (temperature dependent) pre-exponential factor. The nonlinear differential equation governing the reaction-diffusion problem is obtained and tackled numerically using Runge-Kutta- Fehlberg method with shooting technique. The effects of various embedded thermophysical parameters on the temperature field are presented graphically and discussed quantitatively.

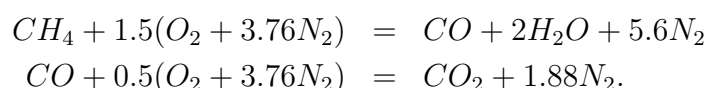
Keywords: Rectangular slab; Two steps exothermic reaction; Convective heat loss, Radiative heat loss; Hermite-Pad approximants

1. Introduction

Heat transfer in a reactive slab of combustible materials due to exothermic chemical reaction plays a significant role in improving the design and operation of many industrial and engineering devices and find applications in power production, jet and rocket propulsion, fire prevention and safety, pollution control, material processing industries and so on [1–3]. For instance, solid propellants used in rocket vehicles are capable of experiencing exothermic reactions without the addition of any other reactants. The theory of heat transfer in reactive materials has long been a fundamental topic in the field of combustion. The chemical reaction may be modelled by considering either a single step or multi step reaction kinetics. For instance catalytic converter used in an automobile's exhaust system provides a platform for a two step exothermic chemical reaction where unburned hydrocarbons completely combust. This helps to reduce the emissions of toxic car pollutant such as carbon monoxide (CO) into the environment. The main chemical reaction schemes in an autocatalytic converter are [4],



Similarly, the combustion taking place within k-fluid is treated as a two step irreversible chemical reaction of methane oxidation as follows [5]:



The vast majority of studies on chemically reactive materials have been concerned with homogeneous boundary conditions ranging from the infinite Biot number case [6] (Frank-Kamenetskii

conditions) to a range of Biot numbers [7] (Semenov conditions). Mathematical models of the problem relating to exothermic reaction in a reactive slab may be extremely stiff owing to the temperature dependence of the chemical reactions. Moreover, the differential equation for the temperature distribution in a convective-radiative reactive slab with temperature dependent pre-exponential factor is highly nonlinear and does not admit an exact analytical solution [8]. Consequently, the equation has been solved either numerically or using a variety of approximate semi-analytical methods [9–11]. The preceding literature clearly shows the work on reacting slab has been confined to convective surface heat loss. No attempt has been made to study the combined effects of convective and radiative heat losses at the slab surface despite its relevance in various technological applications such as aerothermodynamic heating of spaceships and satellites, nuclear reactor thermohydraulics and glass manufacturing. Thermal radiation is characteristic of any material system at temperatures above the absolute zero and becomes an important form of heat transfer in devices that operate at high-temperatures. Radiation is the dominant form of heat transfer in applications such as furnaces, boilers, and other combustion systems.

The present investigation aims to extend the recent work of Makinde [11, 12] to include combined effects of convective and radiative heat losses on a slab of combustible material with internal heat generation due to a two step exothermic reaction. It is hoped that the results obtained will not only provide useful information for applications, but also serve as a complement to the previous studies.

2. Mathematical Model

Let us consider the dynamical thermal behaviour of a rectangular slab of combustible materials with internal heat generation due to a two step exothermic chemical reaction, taking into account the diffusion of the reactant and the temperature dependent variable pre-exponential factor. The geometry of the problem is depicted in Fig. 1. It is assumed that the slab surface is subjected to both convective and radiative heat losses to the environment.

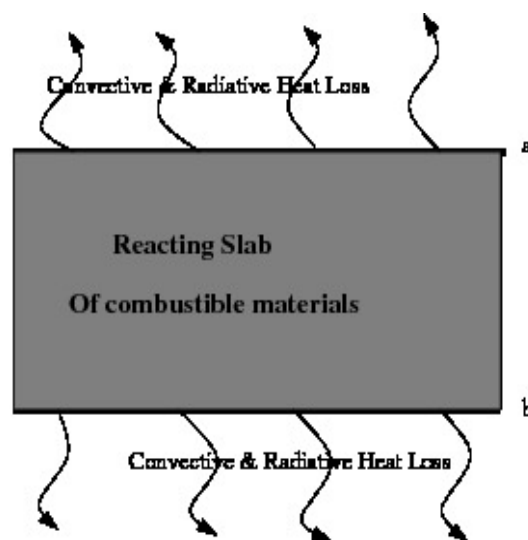


Figure 1: Sketch of the physical model.

The one-dimensional heat balance equation in the original variables together with the boundary

conditions can be written as [1, 3, 10, 11]:

$$k \frac{d^2 T}{d\bar{y}^2} + Q_1 C_1 A_1 \left(\frac{KT}{\nu g} \right)^m e^{\frac{E_1}{RT}} + Q_2 C_2 A_2 \left(\frac{KT}{\nu g} \right)^m e^{\frac{E_2}{RT}} - \varepsilon \sigma [T - T_\infty^4] = 0, \quad (2.1)$$

$$k \frac{dT}{d\bar{y}}(0) = h_1 [T(0) - T_\infty], k \frac{dT}{d\bar{y}} = -h_2 [T(a) - T_\infty], \quad (2.2)$$

where T is the absolute temperature, T_∞ is the ambient temperature, h_1 is convective the heat transfer coefficient at the lower surface, h_2 is convective the heat transfer coefficient at the upper surface, k is the thermal conductivity of the material, ε is the slab surface emissivity, σ is the Stefan Boltzmann constant; Q_1 is the first step heat of reaction, Q_2 is the second step heat of reaction, A_1 is the first step reaction rate constant, A_2 is the second step reaction rate constant, E_1 is the first step reaction activation energy, E_2 is the second step reaction activation energy, ρ is the density, R is the universal gas constant, C_1 is the first step reactant species initial concentration, C_2 is the second step reactant species initial concentration, g is the Plancks number, K is the Boltzmanns constant, ν is vibration frequency, a is the slab half width, \bar{y} is distance measured in the normal direction to the plane c_p is the specific heat at constant pressure and m is the numerical exponent such that $m = \{-2, 0, \frac{1}{2}\}$ represent numerical exponent for Sensitised, Arrhenius and Bimolecular kinetics respectively [1, 3, 9–11]. The following dimensionless variables are introduced into Eqs. (2.1) - (2.3):

$$\theta = \frac{E_1(T - T_\infty)}{RT_\infty^2}, \quad \gamma = \frac{RT_\infty}{E_1}, \quad y = \frac{\bar{y}}{a}, \quad \beta = \frac{Q_2 C_2 A_2}{Q_1 C_1 A_1} e^{\frac{E_1 - E_2}{RT_\infty}}, \quad r = \frac{E_2}{E_1}, \quad (2.3)$$

$$Bi_1 = \frac{ha}{k}, \quad Bi_2 = \frac{ha}{k}, \quad \lambda = \frac{E_1 a^2 Q_1 C_1 A_1}{k R T_\infty^2} \left[\frac{KT_\infty}{\nu g} \right]^m e^{-\frac{E_1}{RT_\infty}}, \quad Nr = \frac{\varepsilon \sigma a E_1 T_\infty^3}{k R},$$

and we obtain the dimensionless governing equation as

$$\frac{d^2 \theta}{dy^2} + \lambda (1 + \gamma \theta)^m \left[e^{\left(\frac{\theta}{1 + \gamma \theta} \right)} + \beta e^{\left(\frac{r \theta}{1 + \gamma \theta} \right)} \right] - Nr [(\gamma \theta + 1)^4 - 1] = 0 \quad (2.4)$$

with

$$\frac{d\theta}{dy}(0) = Bi_1 \theta(0), \quad \frac{d\theta}{dy}(1) = -Bi_2 \theta(1) \quad (2.5)$$

where $\lambda, \gamma, \beta, r, Nr, Bi_1, Bi_2$, represent the Frank-Kamenetskii parameter, activation energy parameter, two step exothermic reaction parameter, activation energy ratio parameter, thermal radiation parameter, the Biot numbers for the slab lower and upper surfaces respectively. Equations (2.4) and (2.5) represent a nonlinear boundary value problem. This nonlinear nature precludes its exact solution, using Runge-Kutta-Fehlberg method with shooting technique, the problem is tackled numerically and the slab surface heat transfer rate $Nu = -\theta'(1)$ is determined.

3. Results and Discussion

We have assigned numerical values to the parameters encountered in the problem in order to get a clear insight into the thermal development in the system. It is very important to note that $\beta = 0$ corresponds to a one step chemical reaction case; an increase in the value $\beta > 0$ signifies an increase in the two step chemical reaction activities in the system.

Table 1: Computation showing the critical values of the reaction rate parameter, $Bi_1 = 1$.

Nr	Bi_2	β	r	m	γ	Nu	λ_c
1	1	0.1	0.1	0.5	0.1	1.0825	0.8179
5	1	0.1	0.1	0.5	0.1	1.1927	1.5468
10	1	0.1	0.1	0.5	0.1	1.2486	2.4632
1	5	0.1	0.1	0.5	0.1	2.3721	1.3129
1	100	0.1	0.1	0.5	0.1	3.0569	1.6614
1	1	0	0.1	0.5	0.1	1.0358	0.8476
1	1	1	0.1	0.5	0.1	1.3910	0.6415
1	1	0.1	1	0.5	0.1	1.0421	0.7706
1	1	0.1	2	0.5	0.1	0.8331	0.6669
1	1	0.1	0.1	0	0.1	1.1662	0.8713
1	1	0.1	0.1	-2	0.1	1.7903	1.1941
1	1	0.1	0.1	0.5	0.2	4.2475	1.3531
1	1	0.1	0.1	0.5	0.4	1.4386	2.6350

Table 1, illustrates the variation in the values of thermal criticality conditions (λ_c) for different combination of embedded parameters. The magnitude of thermal criticality decreases with increasing values of two step reaction rate parameter $\beta > 0$ and the activation energies ratio parameter $r > 0$. This implies that thermal runaway is enhanced by two step exothermic reaction as well as increasing second step activation energy. At very large activation energy ($\gamma = 0$), thermal criticality is independent of the type of reaction as shown in Eq. (2.4). It is interesting to note from the table (1) that thermal runaway will occur faster in bimolecular reaction than in Arrhenius and sensitized reactions. This is reflected in table 1 with lower criticality value for bimolecular reaction. The magnitude of thermal criticality increases with an increase in the Biot number and thermal radiation parameter, thus preventing the early development of thermal runaway and enhancing thermal stability of the system. In figures 2 - 4, we observed that the slab temperature generally increases with increasing values of Frank-Kamenetskii parameter (λ), two step reaction parameter (β) and the activation energy ratio parameter. This can be attributed to an increase in the rate of internal heat generation due to chemical kinetics in the system. Moreover, it is noteworthy that the slab temperature decreases with increasing convective and radiative heat loss as illustrated in figures 5 and 6. Figures 7-9 represent the variation of slab surface heat transfer rate $Nu = -\theta'(1)$ with respect to Frank-Kamenetskii parameter (λ) for different parameter values. In particular, for a given set of parameters value, a thermal critical value λ_c exist such that the thermal system has real solution for $0 \leq \lambda < \lambda_c$. When $\lambda_c < \lambda$ the system has no real solution and displays a classical form indicating thermal runaway. It is interesting to note that the heat transfer rate in the slab is enhanced with increasing radiative and convective heat loss.

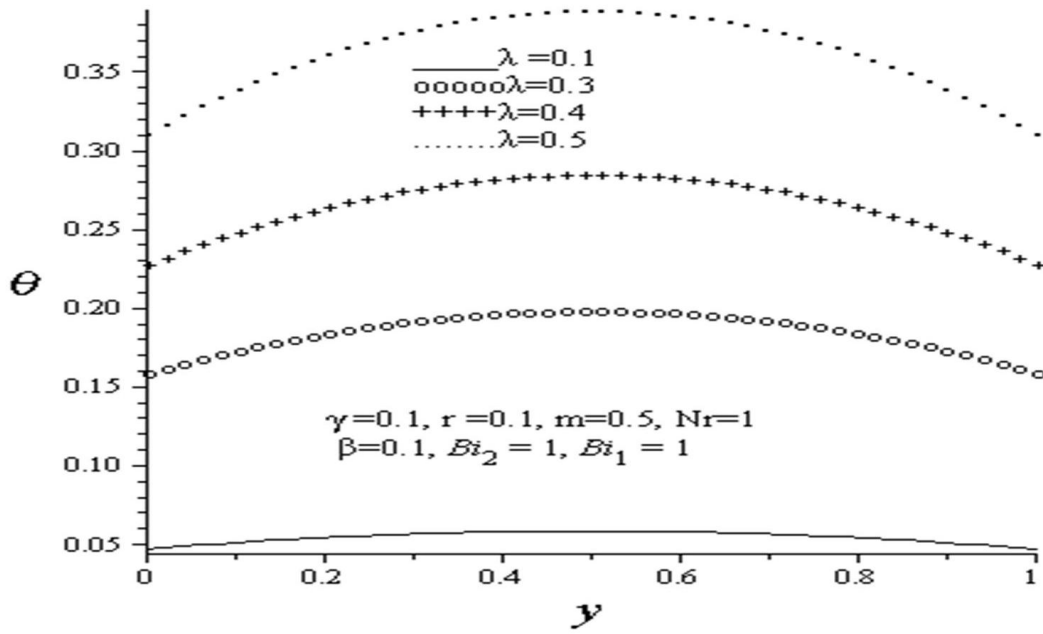


Figure 2: Effects of increasing reaction rate on temperature profiles.

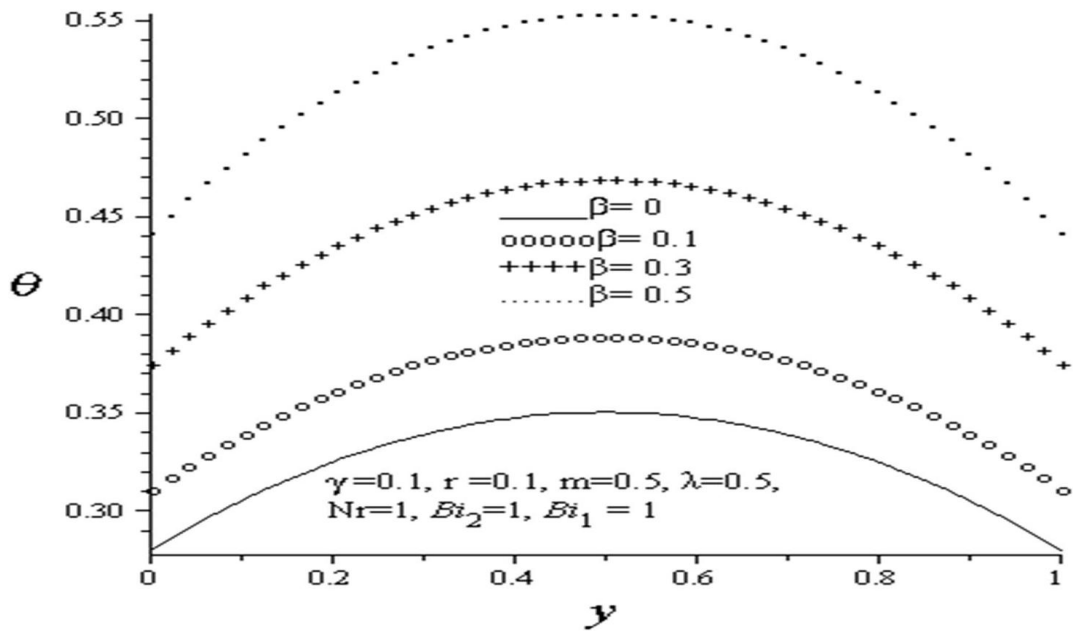


Figure 3: Effects of increasing two stjpg reaction parameter on temperature profiles

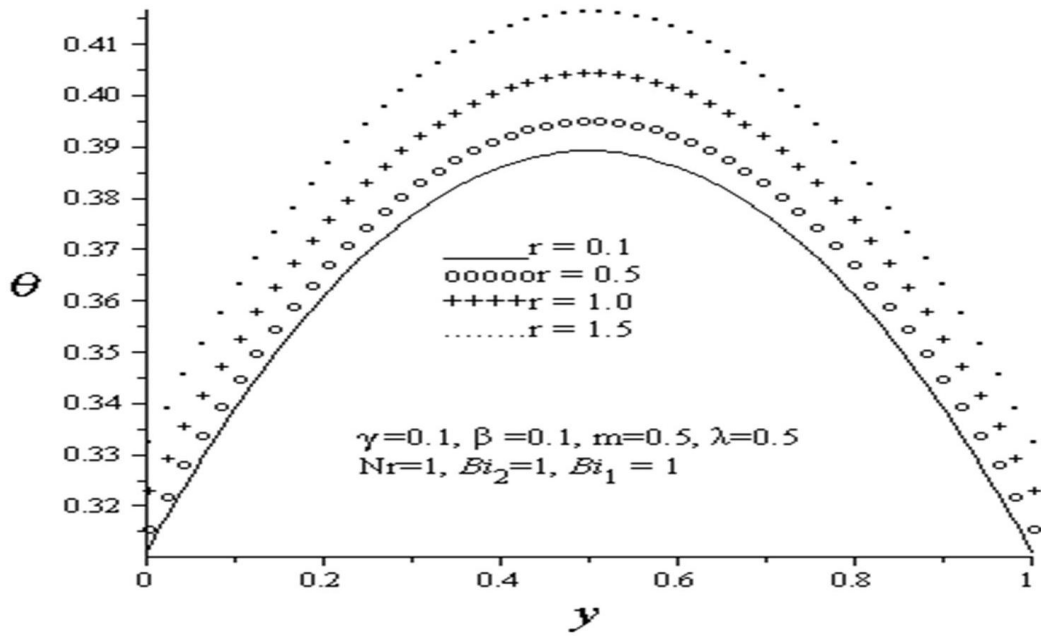


Figure 4: Effects of increasing activation energy ratio parameter on temperature profiles

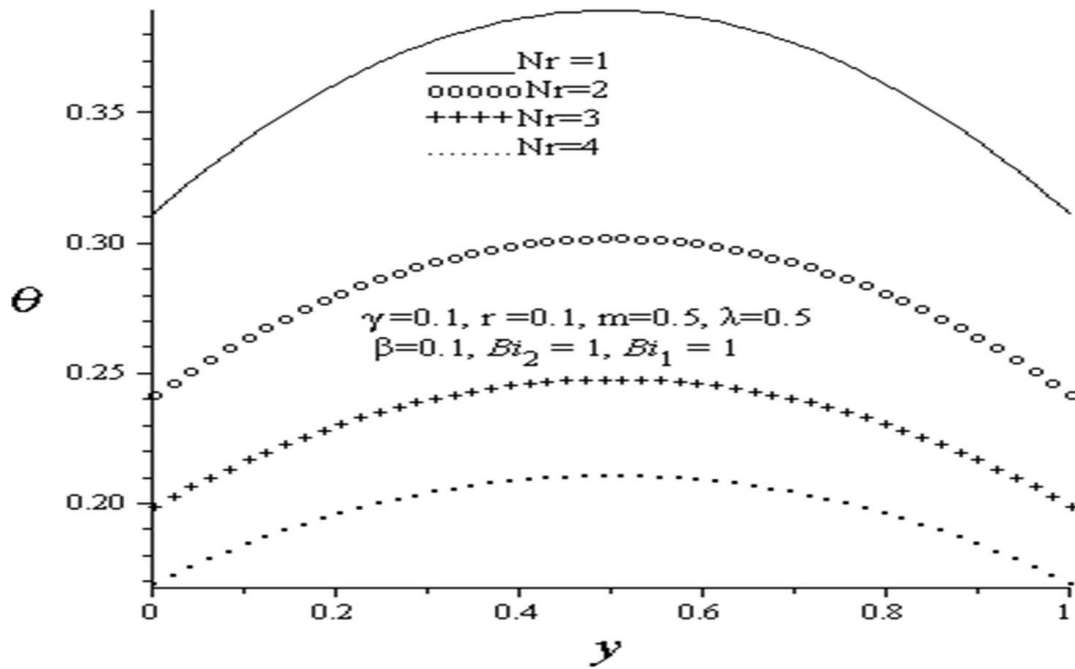


Figure 5: Effects of increasing radiative heat loss on temperature profiles

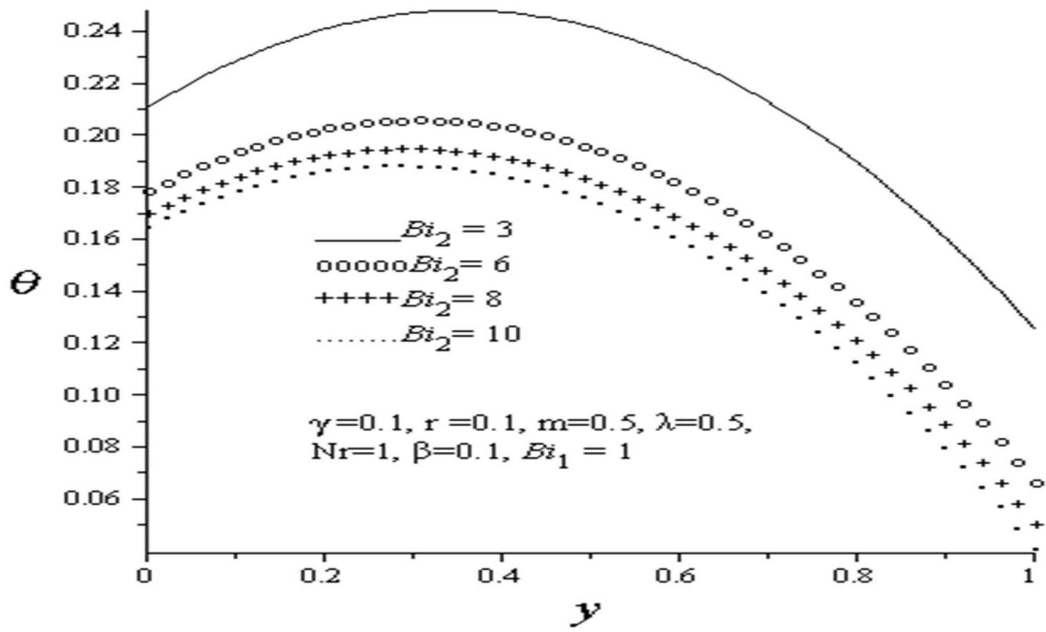


Figure 6: Effects of asymmetrical convective heat loss on temperature profiles

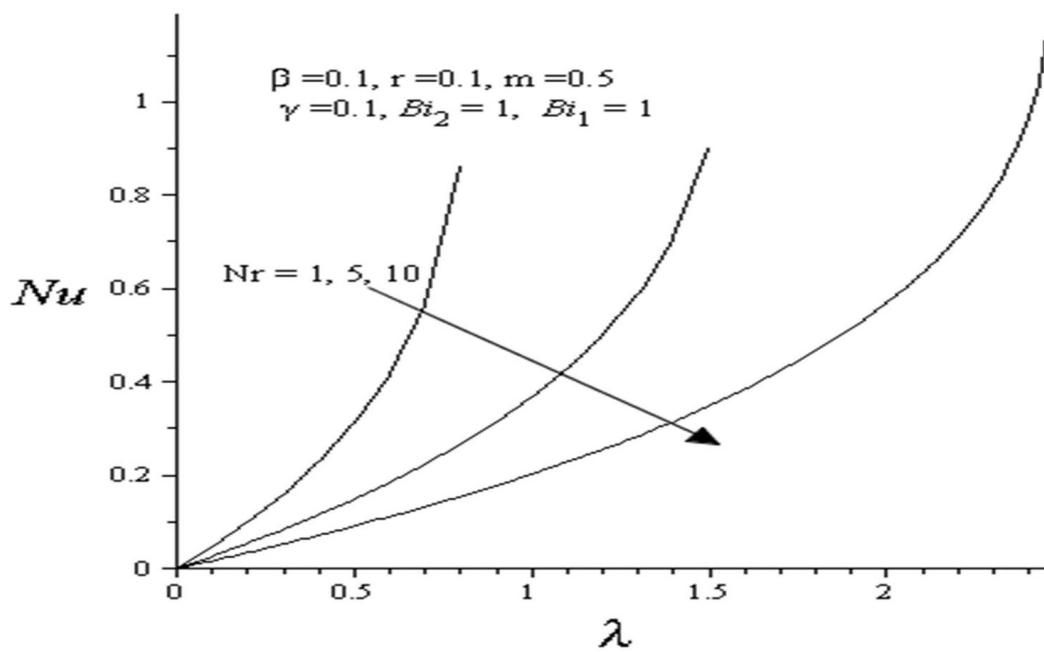


Figure 7: Effects of increasing radiative heat loss on critical value of λ_c

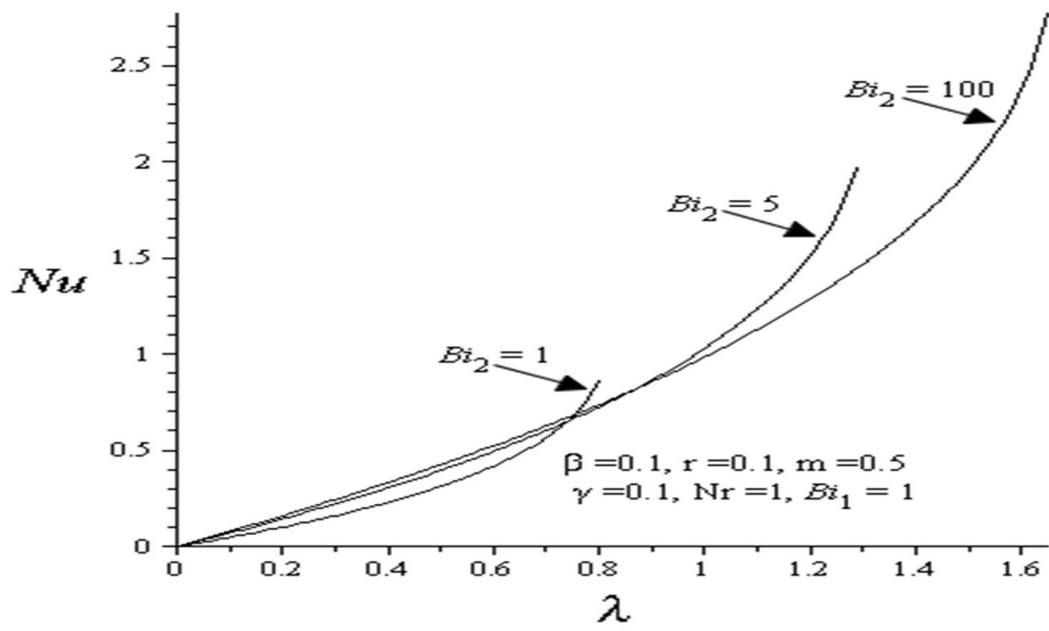


Figure 8: Effects of increasing convective heat loss on critical value of λ_c

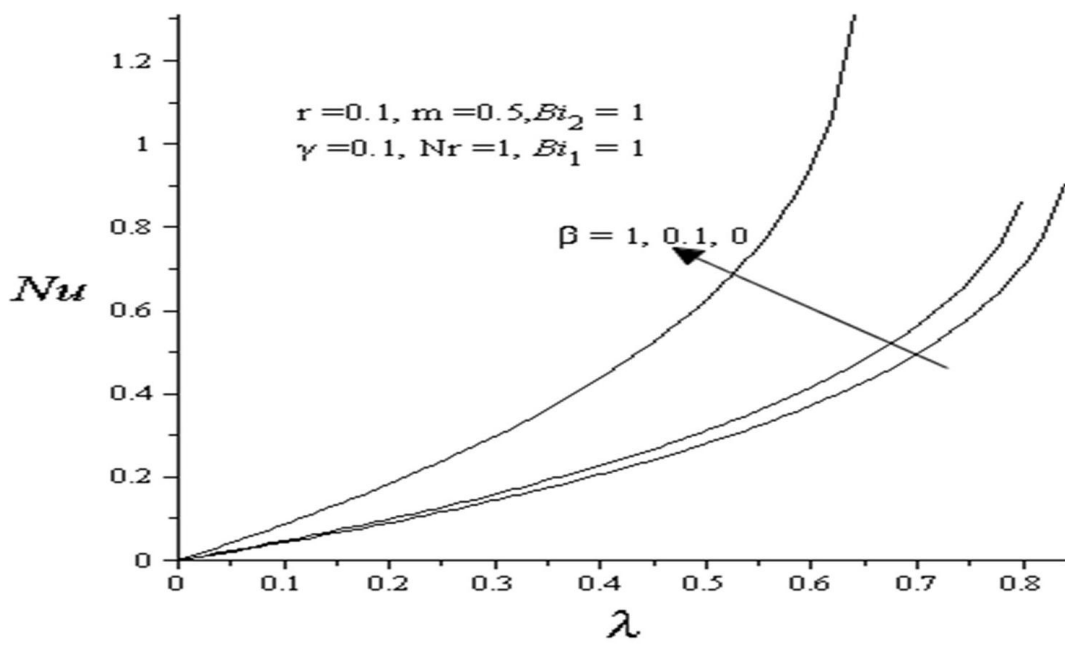


Figure 9: Effects of two step reaction parameter on critical value of λ_c

4. Conclusions

Heat transfer analysis in a convecting and radiating two step reactive slab is presented. The model nonlinear governing differential equation is tackled numerically using Runge-Kutta-Fehlberg method with shooting iteration technique. Our results reveal among others, that the thermal runaway in the system is enhanced by two step exothermic reaction, while an increase in the convective and radiative heat loss stabilizes the system.

References

- [1] Bebernes J. and Eberly D, Mathematical problems from combustion theory. Springer-Verlag, New York, (1989).
- [2] Chambr P. L, On the solution of the Poisson-Boltzmann equation with application to the theory of thermal explosions. *J. Chem. Phys.* 20, 1795 – 1797 (1952).
- [3] Williams F. A, Combustion theory. Second Edition. Benjamin & Cuminy publishing Inc. Menlo Park, California. (1985).
- [4] Makinde O. D, Thermal stability of a reactive viscous flow through a porous-saturated channel with convective boundary conditions. *Applied Thermal Engineering*, 29, 1773 – 1777 (2009).
- [5] Szabo Z. G, Advances in kinetics of homogeneous gas reactions. Methusen and Co. Ltd, Great Britain (1964).
- [6] Frank-Kamenetskii D. A, Diffusion and heat transfer in chemical kinetics. Plenum Press, New York, (1969).
- [7] Barenblatt G.I, Bell J. B, Crutchfield W. Y, The thermal explosion revisited. *Proc. Natl. Acad. Sci. USA*, 95, 13384 – 13386, (1998).
- [8] Sundn B, Transient in a composite slab by a time-varying incident heat-flux combined with convective and radiative cooling. *Int. Commun. Heat Mass Transfer* 13, 515 – 522, (1986).
- [9] Makinde O. D, Exothermic explosions in a slab: A case study of series summation technique. *Int. Comm. Heat Mass Transfer*, 31 (8), 1227 – 1231 (2004).
- [10] Legodi A. M. K, Makinde O. D, A numerical study of steady state exothermic reaction in a slab with convective boundary conditions. *Int. J. Phy. Sci.* 6(10), 2541 – 2549, (2011).
- [11] Makinde O. D, Hermite-Pade approach to thermal stability of reacting masses in a slab with asymmetric convective cooling. *Journal of the Franklin Institute*, 349, 957 – 965, (2012).
- [12] Makinde O. D, On the thermal decomposition of reactive materials of variable thermal conductivity and heat loss characteristics in a long pipe. *Journal of Energetic Materials*, 30, 283 – 298, (2012).

Non-Definite Sturm-Liouville Problems Two Turning Points

by

Mervis Kikonko

University of Zambia,

Abstract

We study the non-definite Sturm-Liouville problem with a weight function having two turning points on a finite closed interval. We find the piecewise smooth solution over the closed interval and give the dispersion relation for the eigenvalues. The dispersion relation is then solved numerically using *Maple* software in order to calculate some eigenvalues. We also find the piecewise smooth eigenfunctions associated with each of the eigenvalues. Moreover, we present graphs of some of the eigenfunctions to check oscillation numbers of the eigenvalues associated with these functions. Finally, we point out a number of interesting open questions for further research

1. Introduction

1.1 General Sturm-Liouville Theory

The Sturm-Liouville equation, named after Jacques Charles François Sturm (1803-1855) and Joseph Liouville (1809-1882) is a real second-order linear differential equation of the form

$$-(p(x)u'(x))' + q(x)u(x) = \lambda w(x)u(x), \quad (1.1)$$

on the bounded or unbounded interval (α, β) . The endpoints α and β can be finite or infinite, and u is a function of the independent variable x . The parameter λ (generally complex) for which the equation (1.1) has a solution u (non-identically zero) in (α, β) is called an **eigenvalue** and the corresponding function u is called an **eigenfunction**. In the case of a regular Sturm-Liouville problem, u is required to satisfy the boundary conditions

$$\alpha_1 u(\alpha) + \alpha_2 p(\alpha) u'(\alpha) = 0, \quad (1.2)$$

$$\beta_1 u(\beta) + \beta_2 p(\beta) u'(\beta) = 0, \quad (1.3)$$

α_1 and α_2 are not both zero, similarly for β_1 and β_2 . The functions $p, q, w : [\alpha, \beta] \rightarrow \mathfrak{R}$, have the following properties:

$$p(x) > 0, \quad q, w, \frac{1}{p} \in L(\alpha, \beta) \text{ and } \int_{\alpha}^{\beta} |w(s)| ds > 0.$$

Suppose that $w(x) > 0$, $p(x)$, $p'(x)$, $q(x)$, and $w(x)$ are continuous functions over the finite interval $[\alpha, \beta]$, then the eigenvalues $\lambda_1, \lambda_2, \lambda_3, \dots$ of problem (1.1),(1.2),(1.3) are real and can be ordered such that $\lambda_1 < \lambda_2 < \lambda_3 < \dots < \lambda_n < \dots < \infty$. Also, corresponding to each eigenvalue λ_n is a unique (up to a normalization constant) eigenfunction $u_n(x)$, which has exactly n zeros in (α, β) . The eigenfunction $u_n(x)$ is called the n th fundamental solution satisfying the regular Sturm-Liouville problem (1.1)-(1.2)-(1.3).

Much has been written on Sturm-Liouville theory since the work of Sturm and Liouville in the 19th century. In the period 1836 – 38, Sturm and Liouville published a remarkable set of papers which initiated the subject, which led to what is now called the *qualitative theory* of differential equations, (see [3]). In the same book, the authors point out that, Sturm was mainly concerned with the qualitative behavior of the eigenfunctions, while Liouville was more concerned with the eigenfunction expansions.

This theory is important in applied mathematics, where SturmLiouville problems occur very commonly. The differential equations considered here arise directly as mathematical models of motion according to Newton’s law, but more often as a result of using the method of separation of variables to solve the classical partial differential equations of physics, such as Laplace’s equation, the heat equation, and the wave equation. Sturm-Liouville problems have been discovered as describing the mathematics underlying a variety of physical phenomena. Thus they have been applied in various fields of study like Engineering and Physics.

1.2 General Non-Definite Sturm-Liouville Problems

Let (1.1) be written as

$$Tu = \lambda wu, \text{ where } T = -\frac{d}{dx}\left(p(x)\frac{d}{dx}\right) + q(x). \quad (1.4)$$

Then the problem (1.4)-(1.2)-(1.3) is called left-definite if the form (Tu, u) is definite on the domain of definition for each $u \neq 0$. The problem is called right-definite if the form (wu, u) is definite. In the case that neither (Tu, u) nor (wu, u) is definite, then the problem is called non-definite. Here, $(,)$ denotes the inner product of the usual Hilbert space $L^2[\alpha, \beta]$.

If we consider problem (1.1)-(1.2)-(1.3), and assume that $w(x)$ changes sign on $[\alpha, \beta]$ and the problem is non-definite, then the spectrum is discrete, always consists of a doubly infinite sequence of real eigenvalues, has no finite limit point, and has at most a finite and even number of non real eigenvalues (necessarily occurring in complex conjugate pairs) along with at most finitely many real non-simple eigenvalues (see e.g, [7]). Furthermore, the eigenfunction corresponding to the smallest positive eigenvalue need not necessarily be of one sign in (α, β) . Let M be the number of pairs of distinct non-real eigenvalues of the problem and N be the number of distinct negative eigenvalues of the same problem, then $M \leq N$.

Moreover, there is an integer n_R , called the *Richardson index* having the property that whenever $n \geq n_R$, there are exactly two eigenfunctions of the problem oscillating n -times in (α, β) (see e.g [7]). There is a positive number λ^+ , called the Richardson number defined as

$\lambda^+ = \inf\{x \in \Re : \forall \lambda > x, \int_{\alpha}^{\beta} |u(x, \lambda)|^2 w(x) dx > 0\}$ (see [4]). Generally speaking, a non definite problem will tend not to have a real ground state (positive eigenfunction) (see e.g [7]). If the positive eigenvalues λ_n^+ of a given non definite problem are labeled in such a way that λ_n^+ has an eigenfunction with precisely n zeros in (α, β) , then

$$\frac{\lambda_n^+}{n^2} \sim \frac{\pi^2}{\left(\int_{\alpha}^{\beta} \sqrt{\left(\frac{w(x)}{p}\right)_+} dx\right)^2}, \quad n \rightarrow \infty,$$

where $\left(\frac{w(x)}{p}\right)_+ = \max\left\{\frac{w(x)}{p}, 0\right\}$ is the *positive part* of $w(x)$.

It is shown in [8] that if $w(x)$ changes sign only once, then the roots of real and imaginary parts u, v of any non real eigenfunction $y = u + iv$ corresponding to a non real eigenvalue, separate

one another. Consequently, any non-real eigenfunction u of the problem cannot have a zero for $x \in (\alpha, \beta)$.

In this paper, we briefly present and discuss some of the results from my Master thesis [6] (see Section 2). Furthermore, a final discussion and conclusion can be found in section 3. In particular, a number of interesting open questions are pointed out.

2. A Non definite Sturm-Liouville problem with weight function having two turning points

We consider the Dirichlet problem

$$u''(x) + (\lambda w(x) + q(x)) u(x) = 0 \quad (2.1)$$

on $[-1, 2]$ given by the boundary conditions

$$u(-1) = 0 = u(2). \quad (2.2)$$

Here, $q(x) = q_0 \in \Re$ for all $x \in [-1, 2]$ and $w(x)$ is a piecewise constant step-function described by the relations

$$w(x) = \begin{cases} a, & \text{if } x \in [-1, 0], \\ b, & \text{if } x \in (0, 1], \\ c, & \text{if } x \in (1, 2], \end{cases}$$

where we assume, without loss of generality that, $a < 0, b > 0, c < 0$. We note that (2.1) is in Sturm-Liouville form with $q(x)$ replaced by $-q(x)$.

Let $H^2[-1, 2]$ be the subspace of $L^2[-1, 2]$ consisting of all continuously differentiable functions $u \in C^1[-1, 2]$ such that u' is absolutely continuous on $[-1, 2]$ and $u'' + q(x)u \in L^2[-1, 2]$. Let T be the linear operator in $L^2[-1, 2]$ defined by

$$\begin{aligned} D(T) &= \{u \in H^2[-1, 2] \mid u(-1) = 0 = u(2)\} \\ Tu &= u'' + q(x)u. \end{aligned} \quad (2.3)$$

We have the following result:

Theorem 2.4 *The forms (Tu, u) and (wu, u) arising from the operator T defined by (2.3) are generally indefinite.*

Proof: We have

$$\begin{aligned}
(Tu, u) &= \int_{-1}^2 (u'' + q(x)u)\bar{u}dx \\
&= \int_{-1}^2 u''\bar{u}dx + \int_{-1}^2 q(x)|u|^2dx \\
&= - \int_{-1}^2 u'\bar{u}'dx + \int_{-1}^2 q(x)|u|^2dx \\
&= \int_{-1}^2 (q(x)|u|^2 - |u'|^2)dx.
\end{aligned}$$

From the above result, we deduce that for $q \leq 0$, $(Tu, u) < 0$ for all $u \in D(T)$ which means that $-T \geq 0$. However, when $q > 0$, we see that the form (Tu, u) may be sign-indefinite. By Sturm-Liouville theory we recall that there are always infinitely many eigenvalues having a fixed sign (positive or negative). Let us choose q so that T has both positive and negative eigenvalues. Then, it is easy to see that if we choose u to be an eigenfunction corresponding to a positive eigenvalue of T (defined by (2.3)) then $(Tu, u) > 0$. On the other hand, by assumption, since T has a negative eigenvalue with eigenfunction v then $(Tv, v) < 0$. So, fixing such a general value of q , there may exist functions u for which $(Tu, u) > 0$ and a possibly different set of u 's for which $(Tu, u) < 0$ and so the form (Tu, u) is generally indefinite.

Similarly,

$$\begin{aligned}
(wu, u) &= \int_{-1}^2 w|u|^2dx \\
&= -|a| \int_{-1}^0 |u|^2dx + b \int_0^1 |u|^2dx - |c| \int_1^2 |u|^2dx.
\end{aligned}$$

It is clear that the sign of $\lambda(wu, u)$ is indefinite since it generally depends on the relative sizes of a, b, c . Thus, both forms are indefinite.

The proof is complete. □

Therefore, in accordance with accepted terminology (see [8]), the problem (2.1),(2.2) is *non-definite*.

2.1 Explicit solution for the problem(2.1),(2.2) and results

For the special case with $a = -1$, $b = 2$, $c = -1$, we found the explicit solution of (2.1),(2.2) given below

$$u(x) = \begin{cases} X(x), & \text{if } x \in [-1,0], \\ Y(x), & \text{if } x \in (0,1], \\ Z(x), & \text{if } x \in (1,2], \end{cases}$$

Table 1: Summary of results of the spectrum of problem (2.1)-(2.2)

q_0	# of complex pairs	# of negative eigenvalues	total # of eigenvalues	Smallest oscillation #
π^2	1	9	18	2
$2\pi^2$	3	6	18	3
$3\pi^2$	3	7	18	3
$5\pi^2$	4	6	20	4
$6\pi^2$	4	7	20	5

where $a < 0, b > 0, c < 0$, and

$$\begin{aligned}
 X(x) &= \frac{\sin(\sqrt{-\lambda|a| + q_0}(x+1))}{\sqrt{-\lambda|a| + q_0}} \\
 Y(x) &= \frac{\sqrt{\lambda b + q_0} \sin(\sqrt{-\lambda|a| + q_0}) \cos(\sqrt{\lambda b + q_0}x)}{\sqrt{-\lambda|a| + q_0} \sqrt{\lambda b + q_0}} \\
 &\quad + \frac{\sqrt{-\lambda|a| + q_0} \cos(\sqrt{-\lambda|a| + q_0}) \sin(\sqrt{\lambda b + q_0}x)}{\sqrt{-\lambda|a| + q_0} \sqrt{\lambda b + q_0}} \\
 Z(x) &= \frac{\sin(\sqrt{-\lambda|a| + q_0}) \cos(\sqrt{\lambda b + q_0}) \cos(\sqrt{-\lambda|c| + q_0}(x-1))}{\sqrt{-\lambda|a| + q_0}} \\
 &\quad + \frac{\cos(\sqrt{-\lambda|a| + q_0}) \sin(\sqrt{\lambda b + q_0}) \cos(\sqrt{-\lambda|c| + q_0}(x-1))}{\sqrt{\lambda b + q_0}} \\
 &\quad + \frac{\cos(\sqrt{-\lambda|a| + q_0}) \cos(\sqrt{\lambda b + q_0}) \sin(\sqrt{-\lambda|c| + q_0}(x-1))}{\sqrt{-\lambda|c| + q_0}} \\
 &\quad - \frac{\sqrt{\lambda|b| - q_0} \sin(\sqrt{-\lambda|a| + q_0}) \sin(\sqrt{\lambda b + q_0}) \sin(\sqrt{-\lambda|c| + q_0}(x-1))}{\sqrt{-\lambda|a| + q_0} \sqrt{-\lambda|c| + q_0}}.
 \end{aligned}$$

The solution is found by piecing together the various solutions on the intervals $(-1,0)$, $(0,1)$ and $(1,2)$ so as to obtain a continuously differentiable function on $(-1,2)$. By solving the dispersion relation

$$\begin{aligned}
 0 &= \sqrt{-\lambda|c| + q_0} \sqrt{\lambda b + q_0} \sin(\sqrt{-\lambda|a| + q_0}) \cos(\sqrt{\lambda b + q_0}) \cos(\sqrt{-\lambda|c| + q_0}) \\
 &\quad + \sqrt{-\lambda|c| + q_0} \sqrt{-\lambda|a| + q_0} \cos(\sqrt{-\lambda|a| + q_0}) \sin(\sqrt{\lambda b + q_0}) \cos(\sqrt{-\lambda|c| + q_0}) \\
 &\quad + \sqrt{-\lambda|a| + q_0} \sqrt{\lambda b + q_0} \cos(\sqrt{-\lambda|a| + q_0}) \cos(\sqrt{\lambda b + q_0}) \sin(\sqrt{-\lambda|c| + q_0}) \\
 &\quad - (\lambda b + q_0) \sin(\sqrt{-\lambda|a| + q_0}) \sin(\sqrt{\lambda b + q_0}) \sin(\sqrt{-\lambda|c| + q_0}).
 \end{aligned}$$

we calculated a few eigenvalues in the cases $q_0 = \pi^2, 2\pi^2, 3\pi^2, 5\pi^2$ and $6\pi^2$ in the rectangle $D = \{\lambda \in \mathcal{C} : |\Re\lambda| < 300 \text{ and } |\Im\lambda| < 300\}$ using the Maple package `RootFinding[Analytic]`. Tables 1 and 2 show summaries of the results on the spectrum.

Table 2: Non-real eigenvalues obtained inside the rectangle D for the cases $q_0 = \pi^2, 2\pi^2, 3\pi^2, 5\pi^2, 6\pi^2$

q_0	Eigenvalues	No. of zeros of	
		$Re u(x, \lambda_i)$	$Im u(x, \lambda_i)$
π^2	$3.2465 \pm 5.6334i$	3	1
$2\pi^2$	$-8.307 \pm 5.5991i$	4	2
	$-4.220 \pm 5.7435i$	3	3
	$12.940 \pm 6.6651i$	4	2
$3\pi^2$	$5.1614 \pm 7.7537i$	4	4
	$-2.452 \pm 10.506i$	5	3
$5\pi^2$	$7.0223 \pm 10.935i$	6	4
	$20.750 \pm 12.134i$	5	5
	$-19.75 \pm 7.2174i$	6	4
	$-16.37 \pm 10.338i$	5	5
$6\pi^2$	$-6.434 \pm 14.431i$	6	6
	$-13.40 \pm 13.525i$	7	5
	$52.026 \pm 7.0997i$	6	4
	$21.552 \pm 15.247i$	7	5
$5\pi^2$	$7.0223 \pm 10.935i$	6	4
	$20.750 \pm 12.134i$	5	5
	$-19.75 \pm 7.2174i$	6	4
	$-16.37 \pm 10.338i$	5	5
$6\pi^2$	$-6.434 \pm 14.431i$	6	6
	$-13.40 \pm 13.525i$	7	5
	$52.026 \pm 7.0997i$	6	4
	$21.552 \pm 15.247i$	7	5

2.2 Example(the case $q_0 = 6\pi^2$)

The asymptotic distribution of the eigenvalues satisfies

$$\frac{\lambda_n^+}{n^2} \sim \frac{\pi^2}{\left(\int_{-1}^2 \sqrt{w_+(x)} dx\right)^2} \approx 4.9348, \quad n \rightarrow \infty.$$

In this case, the Richardson index is 5 and from the data we see that $\lambda^+ \leq 106.4765$ while the oscillation number for $\lambda = 106.4765$ is 5, and so the Richardson index is 5, as expected.

3. Discussion and conclusion

3.1 Discussion

In all the cases considered in this paper, we have both real and non-real eigenvalues. It can be seen from the graphs of the eigenfunctions that generally oscillation numbers decrease as the parameter value increases, but then oscillations will stabilize and then the usual oscillation theorem may be claimed. This leads to the estimation of λ^+ and n_R .

For example in figure 1, we see that the oscillation number of the smallest positive eigenvalue is greater than that of the second one. However, oscillations stabilize from the third onwards,

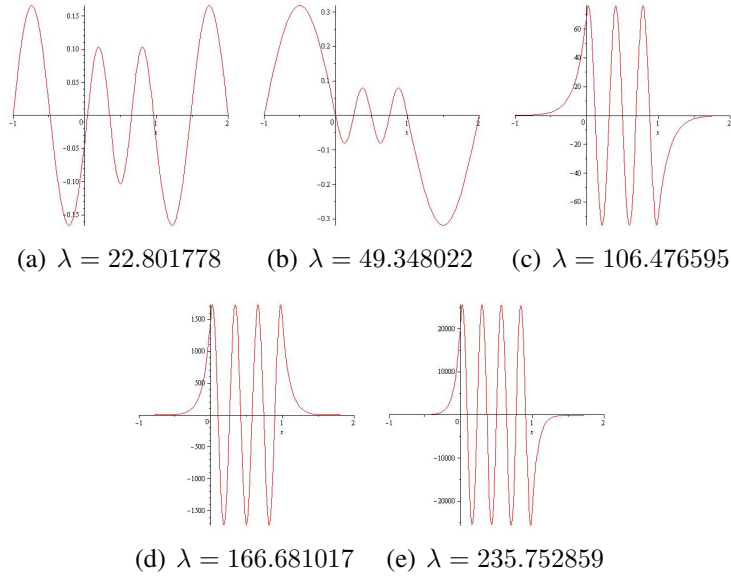


Figure 1: eigenfunctions corresponding to positive eigenvalues for the case $q_0 = 6\pi^2$

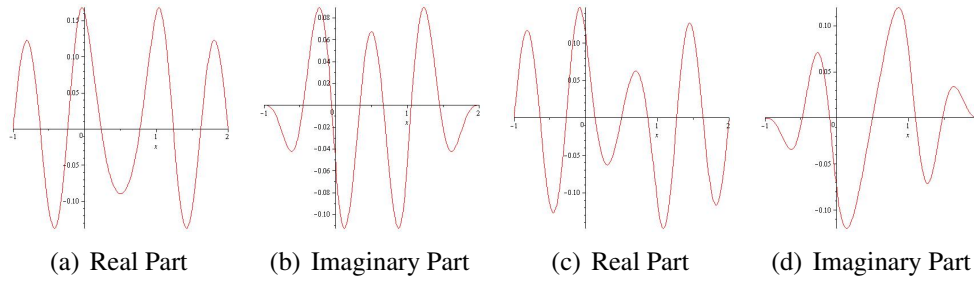


Figure 2: The cases $\lambda = -6.4344 - 14.4314i$, $\lambda = -13.4034 - 13.5248i$

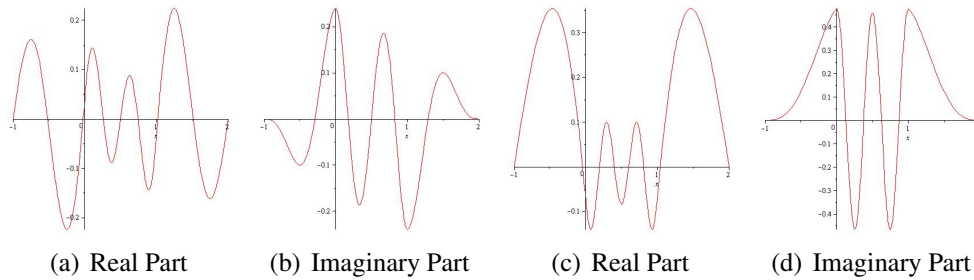


Figure 3: The cases $\lambda = 21.5520 - 15.2468i$, $\lambda = 52.0258 + 7.0997i$

that is, for all positive eigenvalues such that $\lambda \geq 106.476595$ each eigenvalue has a unique oscillation number. This shows that $\lambda^+ \leq 106.476595$ and $n_R = 5$.

Generally speaking, the number of non-real eigenvalues seems to increase with increasing q_0 . The number of pairs of distinct non-real eigenvalues of the problem does not exceed the number of negative eigenvalues in all the cases considered. For all values of q_0 considered (cases

where there are non-real eigenvalues), the smallest oscillation number is 2 and so there is no positive eigenfunction in $(-1, 2)$. Furthermore, the real and imaginary parts of the non-real eigenfunctions do not have interlacing zeros. It is evident from the graphs in figures 2 and 3 that the number of zeros of the imaginary part is less than that of the real part by 2 in some cases, and equal in other cases. However, the non-real eigenfunctions do not vanish in $(-1, 2)$.

3.2 Conclusion

A huge number of papers by mathematicians and others, have been published on Sturm-Liouville problems since their origins in 1836. Yet, remarkably, this subject is still an intensely active field of research today. In this paper, we undertook a numerical study of the non-real eigenfunctions and eigenvalues of a non-definite Sturm-Liouville problem with two turning points, paralleling the study in [4] in the case of one turning point. Ultimately, our aim was to examine the behavior of the eigenfunctions, both real and non-real, of this non-definite Sturm-Liouville problem.

One feature of the non-definite problem is the possible existence of non-real eigenvalues. This may sound paradoxical, as the equation is (formally) self-adjoint and so all eigenvalues should be real. However, this is where the problem lies: the formal self-adjointness of an equation does not necessarily imply the self-adjointness of the corresponding operator. It follows that whenever there is a non-real eigenvalue the corresponding operator cannot be self-adjoint in a Hilbert space.

We have indeed verified that if the problem (3.1)-(3.2) has at least one (complex conjugate) pair of non-real eigenvalues, then there is no real eigenvalue whose corresponding eigenfunction has one zero in the interval $(-1, 2)$ (in conformity with the results in [2], [1]). We also showed that, in the cases considered here, the complex eigenfunctions (corresponding to non-real eigenvalues) are never zero in $(-1, 2)$. Whether this is an accident or the result of a more general yet unproven theorem, is unknown, but we strongly believe it is so and pose this as an open question for future research.

For these same examples, we also showed that the real and imaginary parts of these eigenfunctions do not have interlacing zeros (although they are expected to since the non real eigenfunctions do not vanish). In fact, these zeros interlace in the one-turning-point case as shown by Richardson (see [8]). However, in the case of two turning points we see that there are examples where the zeros do not interlace at all. This fact too, raises an interesting open question for future research.

Furthermore, the number of zeros of the real part of each of the non-real eigenfunctions considered is greater than or equal to the number of zeros of the imaginary part. This may also be a consequence of a more general theorem which we don't know, so then, we have a third interesting open question. In future studies on this subject, there is need for the formulation of general theorems that could explain some or all of these observations. Finally, we observed that even for real eigenvalues the corresponding eigenfunctions do not behave in conformity with the Sturm oscillation theorem as was postulated and proved by Haupt (1911) in [5], and Richardson [8].

ACKNOWLEDGEMENTS: I should like to express my special thanks to my Master of Science thesis supervisor Dr. Angelo B. Mingarelli (Carleton University, Ottawa, Canada) for suggesting the topic. Working under his supervision was a pleasure and a good experience. I feel greatly honored to still have

him supervise my PhD research in which he is co-supervising with Professor Lars-Erik Persson and Professor Peter Wall both from Luleå University of Technology. I thank the three professors for reading through my manuscript and giving valuable suggestions and comments.

References

- [1] W. Allegretto and A. B. Mingarelli. On the non-existence of positive solution for a Schrödinger equation with an indefinite weight-function. *C. R. Math. Rep. Acad. Sci. Canada*, 8:69–72, 1986.
- [2] W. Allegretto and A. B. Mingarelli. Boundary problems of the second order with an indefinite weight-function. *J. Reine Angew. Math.*, 398:1–24, 1989.
- [3] T. G. Anderson, R. C. Brown, and D. B. Hinton. Perturbation theory for a one-term weighted differential operator. In *Spectral theory and computational methods of Sturm-Liouville problems (Knoxville, TN, 1996)*, volume 191 of *Lecture Notes in Pure and Appl. Math.*, pages 149–170. Dekker, New York, 1997.
- [4] F.V Atkinson and D Jabon. Indefinite Sturm-Liouville problems. In Proceeding of the 1984 workshop on Spectral Theory of Sturm-Liouville Differential Operators, pages 31–44. Argon National Laboratory, Argon, Illinois, 1984.
- [5] O. Haupt. *Untersuchungen über Oszillationstheoreme*. Diss. Würzburg. Leipzig: B. Z. Teubner. 50 S. 8° , 1911.
- [6] M. Kikonko. *Non Definite Sturm-Liouville problem with weight function having two turning points*. Thesis (Masters). Carleton University, 2011.
- [7] A.B. Mingarelli. A survey of the regular weighted Sturm-Liouville problem—the nondefinite case. In *International workshop on applied differential equations (Beijing, 1985)*, pages 109–137. World Scientific Publishing, Singapore, 1986.
- [8] R. G. D. Richardson. Contributions to the Study of Oscillation Properties of the Solutions of Linear Differential Equations of the Second Order. *Amer. J. Math.*, 40(3):283–316, 1918.

The Three Layers Maize Crop Optimal Distribution Network in Tanzania

by

S. A. Sima, M. Ali and I. Campbell

University of the Witwatersrand, Johannesburg, South Africa

Abstract

This paper is part of ongoing PhD research work. The three layers, namely plants, distribution centers and customers, are considered in a two level location routing problem (LRP). The flow of maize from plants to customers via distribution centers(DC) is designed at a minimum cost. The application of LRP is studied for maize production, storage and distribution to the customers in Tanzania at minimum cost. The capacity of plants and distribution centers are subject to constraints in an optimization problem where cost for transporting maize in three layers is computed optimally using CPLEX software. The four regions, five DCs and seventeen regions (customers) in Tanzania have been modeled using actual data from respective departments. The results give the optimal allocation of DCs to plants and customers to DCs with a decrease in cost of 8.2% – 10.3%.

Keywords: optimal distribution, Tanzanian maize crop, location routing problem.

1. Introduction

Distribution network design problems consist of determining the best way to transfer goods from supply to demand points by choosing the structure of the network such that the overall cost is minimized [3]. Here, the network is considered from a graph theory point of view. It is a connected graph with sets of vertices and edges. Production centers, warehouses (distribution centers) and customer zones/demand zones are assumed to be vertices while edges are roads and/or railways. Associated with this network, there are two problems: facility location [6, 12, 18, 20] and vehicle routing problem (VRP) [13, 18]. There are a number of papers that deal with these two problems, both individually and combined forms [4, 6, 13, 17, 19, 20].

In the classical facility location problem (FLP), it is required to determine the optimal location of facilities or resources so as to minimize costs, time, distance and risks in relation to supply and demand points. Some examples of such facilities are schools, warehouses, hospitals, markets, industries, post offices and worship places. In FLP, the constraints such as distance between facilities and customers are often imposed. Other typical constraints are the number of customers (people using these facilities), number of facilities and their capacities [20].

On the other hand, VRP can be defined as the problem of designing least-cost delivery routes from a depot to a set of geographically scattered customers, subject to side constraints (capacity, distance, time, etc). In VRP the number of vehicle routes created are such that (i) each route starts and ends at a depot, (ii) each customer is visited exactly once by a single vehicle, (iii) the total demand of a route does not exceed vehicle capacity, and (iv) the total length of a route does not exceed a preset limit [13].

The location routing problem (LRP) integrates FLP and VRP in a single framework. It is an optimization problem that has attracted many academicians and practitioners in recent years. LRP have been studied with different mathematical approaches in the literature [4, 9, 18]. The models, solution procedures, and applications of LRP began to appear in the literature in the 1970's. LRP models can be deterministic or stochastic [4]. The two main solution approaches to LRP are exact and approximate (heuristic). LRP arises in many applications in various forms. Many recent papers on LRP focus on the distribution of

consumer goods [2, 6, 10, 12, 15–18, 22]. This network is a two level LRP as the routes occur between the first layer (plants) and the second layer (distribution centers), and also between the second layer and the third layer (customers). If there is more than one level of routing involved in LRP, then it is called multi-level LRP. The problem we study here is a two-level LRP.

The practical problem at hand is a deterministic model which is part of a PhD work in progress.

The problem is formulated mathematically as a mixed integer linear programming problem which was then solved using actual data. The considered practical example has a number of unique features which make the research worthwhile. In particular, maize crop (single commodity) transportation in Tanzania is considered as an application of the two-level LRP. So our study is three folds: First, we model the optimization problem as a two-level LRP; Second, we design an algorithm to solve such model and third we consider its practical application. This is the first application of LRP to a practical and real problem in Tanzania.

In the next section we present the deterministic mathematical model for a multi-level LRP. Section 3 presents the maize production and distribution network in Tanzania, and the research data and solution approach is in Section 4. The last section, 5, presents conclusion and recommendations.

In this research, we use plant and production center, warehouse and distribution center, and customer and demand zones, interchangeably. The distribution centers or warehouses are also referred to as depots.

2. The deterministic mathematical model for multi-level LRP

The multi-level LRP models explored in the literature, are either single-commodity or multi-commodity. We now present the mathematical model for a multi-level LRP dealing with a single commodity. The model is adapted from Elhedhli et al [7] and the references [8, 12, 14]. The multi-commodity model used by Elhedhli.

Data for the model:

j : index for plants, where $j = 1, 2, \dots, J$. J is the total number of plants.

k : index for possible distribution center sites, where $k = 1, 2, \dots, K$. K is the total number of the candidate distribution center sites.

l : index for customer demand zones, where $l = 1, 2, \dots, L$. L is the total number of customer demand zones.

S_j : supply (production capacity) for a product at plant j .

D_l : demand of product at customer zone l .

V_k : maximum capacity for DC at site k

C_{jk} : average unit cost of shipping (routing) a product from plant j to DC k .

C_{kl} : average unit cost of shipping a product from DC k to customer zone l .

f_k : fixed cost of the annual possession and operating cost for a DC at site k .

Decision variables for the model:

X_{jk} : amount of product shipped from plant j to DC k .

Y_{kl} : 1 if the DC k serves customer zone l , and 0 otherwise.

Z_k : 1 if a DC k is opened at site k , and 0 otherwise.

The formulated mixed integer linear programming model is as follows:

$$\min \sum_{jk} C_{jk} X_{jk} + \sum_{kl} C_{kl} D_l Y_{kl} + \sum_k f_k Z_k \quad (2.1)$$

Subject to:

$$\sum_k X_{jk} \leq S_j, \forall j, \quad (2.2)$$

$$\sum_j X_{jk} \leq V_k Z_k, \forall k, \quad (2.3)$$

$$\sum_j X_{jk} = \sum_l D_l Y_{kl}, \forall k, \quad (2.4)$$

$$\sum_k Y_{kl} = 1, \forall l, \quad (2.5)$$

$$X_{jk} \geq 0, \forall j, k. \quad (2.6)$$

$$Y_{kl} \in \{0, 1\}, \forall k, l. \quad (2.7)$$

$$Z_k \in \{0, 1\}, \forall k. \quad (2.8)$$

The objective function (2.1) minimize the total distribution cost, including transportation costs and fixed costs for DC. Constraints (2.2) are the supply constraints. Constraints (2.3) refer to DCs capacity and allows the use of opened DCs only. Constraints (2.4) ensure that demands are met at customer zone. Constraints (2.5)(single-sourcing constraints) specify that each customer zone must be served by a single DC. Constraints (2.6) are the non-negativity conditions. Constraints (2.7) and (2.8) are binary variables.

The tasks involved are the determination of the number, size and locations of distribution centers, allocation of distribution centers to production centers and customers to the distribution centers.

3. Maize production and distribution network in Tanzania

The real life application that we considered here is maize crop in Tanzania. In Tanzania, physical access to food is affected by inadequate transportation infrastructure. Maize production is concentrated in the southern highland regions (Rukwa, Mbeya, Iringa, Morogoro and Ruvuma) and peripheral areas of the country, while the traditional food deficit areas are located mostly in the central corridor (Singida, Dodoma and Tabora) and northern part (Arusha, Manyara, Kilimanjaro and Tanga), and other parts as shown in the map of Tanzania (Figure 1). Before reaching the customers, maize is stored in DCs which are allocated in different parts of the country. There are seven DCs with a total capacity of 241 thousands tons which are Arusha (39 tons capacities), Dar Es Salaam (52 tons), Dodoma (39 tons), Shinyanga (14.5 tons), Makambako-Iringa (34 tons), Songea (24 tons) and Sumbawanga (38.5 tons). The first five DCs are used in this study as per actual data collected and their existing distribution system.

Due to long distances between food producing centers, DCs and deficit areas, together with inadequate and unreliable transportation network, high transportation costs are unavoidable. This results in high food prices in deficit areas, and therefore affects access to food by both low income, rural and urban populations [21].

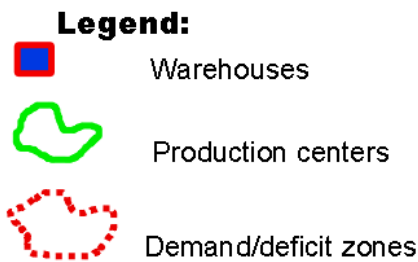
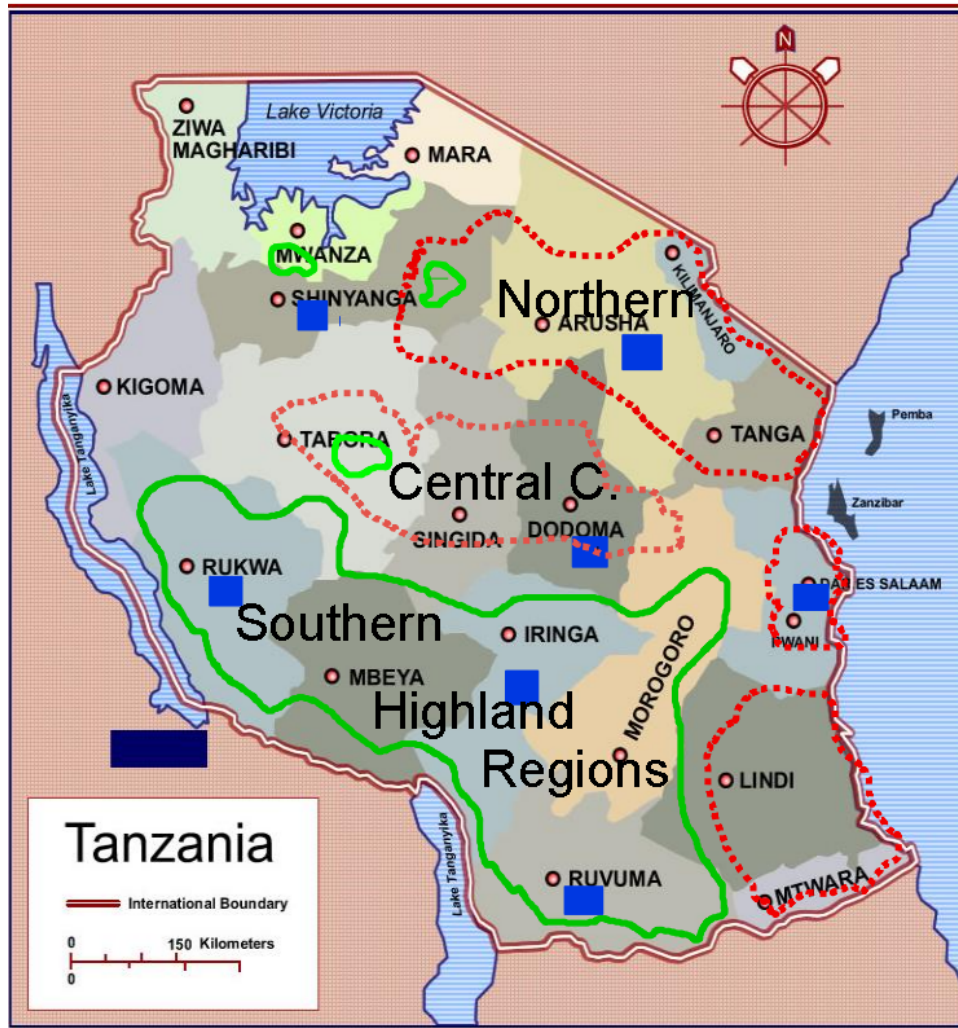


Figure 1: The map of Tanzania showing the food production centers, warehouses and deficit areas

This study is useful as it provides a mechanism for reducing food prices within the country. This will contribute to the June 2009 Tanzania policy involving priority of agriculture also known as ‘Agriculture First’ (Kilimo Kwanza) and as stipulated in ten implementation pillars. For instance, one of the pillars involves identification of priority areas for strategic food commodities to increase the country’s food self sufficiency. The pillars mention a price stabilization mechanism, which includes the expansion of storage capacity and improvement of railway and road systems [1]. In the 2012/2013 Ministry of Agriculture budget speech, the price stabilization for maize flour in cities was addressed, after the price had

decreased by 38% to 40%. This was as a result of about 41,000 tons of maize from warehouses being sold to public markets in regional cities/municipal to cater for a maize deficit [5].

4. Research data and solution approach

The two-level LRP model is a commodity customer delivery model with the following assumptions: (i) demands in each customer zone is known, (ii) number of plants, capacities and their locations are known, (iii) all possible warehouses, their capacities and locations are to be redetermined optimally, (iv) transportation costs from plants to warehouses and from warehouses to customers are known.

Optimization is carried out to determine the values of the decision variables in the model. The optimal decisions to be made are: (i) number of distribution centers, capacities and their locations, (ii) allocation of distribution centers to plants, (iii) allocation of customers to distribution centers, (iv) routes designing from plants to distribution centers and from distribution centers to customers with their associated costs.

4.1 Research data

The food security system in Tanzania is based on maize production, storage and final distribution to the customers. The required data for the model is from several sources; .

All sources of data are based in Tanzania

- Tanzania National Roads Agency (TANROADS): Road distances between regions as per March 2009 and road classification (collected in January, 2011).
- Ministry of Agriculture, Food Security and Cooperatives: Maize production capacity and surplus from "Volume 1: The 2010/11 Final Food Crop Production Forecast for 2011/12 Food Security EXECUTIVE SUMMARY". This is accessible from <http://www.kilimo.go.tz/publications>.
- National Food Reserve Agency (NFRA): Warehouses capacity and transfer quantities from plants to warehouses in 2009/2010 (Obtained in January, 2011).
- Prime Minister Office-Disaster Department: Regional demand quantities of maize between 2004 and 2010 and distances between DCs and demand zones (districts). The maximum annual demand in each region had been considered in this work. The data is as presented in Table 1 and 2.

Table 1: Plants and DCs: Distances ('000km) and Capacities ('000Mt)

Plants	DCs - Road Distances					Capacity	No of DCs
	Arusha	D'Salaam	Dodoma	Makambako	Shinyanga		
Iringa	0.689	0.492	0.264	0.12	0.802	100	3
Mbeya	1.02	0.822	0.594	0.21	0.761	251	7
Rukwa	1.348	1.15	0.922	0.538	0.79	140	4
Ruvuma	1.144	0.947	0.719	0.335	1.257	41	1
DC ca- pacity.	39	52	39	34	14.5		
Total production capacity						532	
Total DCs Capacity						178.5	

Table 2: DCs and Customers: Distances ('000'km) and demands ('000'Mt)

S/N	Cust/DC	Arusha	Dar	Dodoma	Makambako	Shinyanga	Customer Demand
1	Arusha	0.071	0.717	0.496	0.88	0.695	14.378
2	Coast	0.734	0.088	0.539	0.7	1.077	8.626
3	Dodoma	0.495	0.521	0.07	0.454	0.608	15.216
4	Iringa	0.728	0.621	0.403	0.139	0.941	7.008
5	Kagera	1.137	1.502	1.051	1.433	0.513	1.559
6	Kilimanjaro	0.125	0.611	0.55	0.934	0.749	8.940
7	Lindi	1.185	0.539	0.99	1.151	1.528	4.093
8	Manyara	0.229	0.875	0.318	0.809	0.517	15.214
9	Mara	1.133	1.498	1.047	1.431	1.509	11.497
10	Mbeya	0.997	0.8	0.572	0.188	1.11	2.835
11	Morogoro	0.814	0.385	0.452	0.613	0.89	7.689
12	Mtwara	1.256	0.61	1.061	1.17	1.599	3.876
13	Mwanza	0.863	1.228	0.777	1.161	0.289	10.398
14	Shinyanga	0.725	1.19	0.639	1.024	0.101	9.702
15	Singida	0.662	0.688	0.237	0.621	0.493	9.434
16	Tabora	0.807	1.209	0.721	1.105	0.183	5.773
17	Tanga	0.772	0.337	0.788	0.756	1.066	8.906
	DC Customer Capacity	5	6	5	4	2	
		Average demand					8.538
		Total Demand					145.144

In table 1 the values in the last column (No of DCs) were obtained by dividing the relevant plant capacity by average DC capacity. The demands in the last column of table 2, were obtained from 93 districts' demands in Tanzania. The given districts demand were then summed in each region to have a regional customer demand. The last row of this table gives the number of customers (from 17 customers) that can be saved by the particular DC (DC customer capacity). The serial number, 1-17 in table 2 is coded as corresponding customer (customers saved) in the programming results from table 3, 4 and 5.

4.2 Solution approach and computational results

The solution of a mixed integer linear programming deterministic model is obtained by using CPLEX software. The real life data from Tanzania in tables 1 and 2 was used. The problem was solved in two stages. The first stage is solved as a FLP or DC location problem, where DCs and customers are involved. The second stage considered all three layers simultaneously as a two level LRP. The solution obtained is optimal, hence it is an exact method.

Facility Location Problem solution

In the FLP, the target is to find the optimal number and location of warehouses with respect to distance from customers and their capacities while satisfying customers' demands at minimum cost. The model for this case is:

$$\min \sum_{kl} C_{kl} D_l Y_{kl} + \sum_k f_k Z_k \quad (4.1)$$

Subject to constraint sets (2.4), (2.5),(2.7) and (2.8). Using CPLEX software (IBM ILOG CPLEX Optimization Studio), several results were obtained by changing the capacity of warehouses from the existing capacities and observing the influence on the fixed cost. Some important results are summarised below.

Table 3: Results for changing capacities of DC while fixed cost is ZERO

DC/RUN	R1 (6,5,5,4,2: Existing Cap). Customers saved	R2 (4,5,3,1,5: Actual max use). Customers saved	R3 (6,5,5,2,4 = 6,5,4,2,5). Customers saved
Dar	2, 7, 11, 12, 17	2, 7, 12, 17	2, 7, 11, 12, 17
Arusha	1, 6, 8	1, 6, 8, 9	1, 6, 8
Dodoma	3, 5, 9, 13, 15	3, 4, 11	3, 9, 15
Makambako	4, 10	10	4, 10
Shinyanga	14, 16	5, 13, 14,15, 16	5, 13, 14, 16
Objective Value (km)	6.177	5.824	5.151
Run Time (Sec)	8	16	10
NOTE: All warehouses are open and all customers are saved			

From table 3 where fixed cost is zero, R1, R2 and R3 are three runs (computations) with different results from different capacities of DCs. R1 consider the existing capacities as constructed, R2 is maximum use of the DC as per 2004-2010 demands. R3 is optimal location and capacities as from objective value (last but one row). First it shows all DCs are open, and all customers (total 17) are allocated to DCs. The optimal capacities require the Makambako DC to save 2 customers and Shinyanga DC to save 4 or 5 customers with objective value of 5.151km. This result of having 1.026km (1,026km) (savings) for

Table 4: Results for changing capacities of DC while fixed (f) cost is 10 for each DC.

DC/RUN	R1 (6,5,5,4,2: Existing Cap). Customers saved	R2 (4,5,3,1,5: Actual max use). Customers saved	R3 (6,5,5,2,4). Customers saved	R4 (6,5,6,1,4). Customers saved
Dar	2, 7, 11, 12, 17	2, 7, 12, 17	2, 7, 11, 12, 17	2, 4, 7, 11, 12, 17
Arusha	1, 6, 8, 9, 16	1, 4, 6, 8, 9	1, 6, 8	1, 6, 8, 9, 16
Dodoma	3, 4, 10, 13, 15	3, 10, 11	3, 4, 9, 10, 15	3, 5, 10, 13, 14, 15
Makambako	-	-	-	-
Shinyanga	5, 14	5, 13, 14, 15, 16	5, 13, 14, 16	-
Objective Value (km)	46.997	46.533	45.799	38.291
Run Time (Sec)	17	9	17	8
NOTE: At least one warehouse is closed for $f \geq 0.3$ and all customers are saved				

constructed capacity and 673km for maximum use DC capacity. The customers as per each DC can be designed a direct route from allocated DC.

Table 4 results indicates the influence of fixed costs in FLP where only 3 DC are to be opened in the optimal costs. As shown in last column, the objective value is 38.291km, that is less by 8.706km and 8.242km from constructed capacities and maximum use capacities respectively.

The two level LRP solution

This is a second stage solving where plants, warehouses and customers distribution cost are now considered simultaneously. Minimum cost attained by optimal location of warehouses to plants and also customers to open warehouses. In this case, all the five warehouses are open for both capacities and fixed cost consideration contract to first stage.

The whole model (2.1) to (2.8) is used with some modification of decision variable from quantities to a binary. X_{jk} is now 1 if a transfer from plants to open warehouse, and 0 otherwise. So we have an additional single source constraints to link plants and warehouses as:

$$\sum_j X_{jk} = 1, \forall k. \tag{4.2}$$

The most important computational results are summarised in the table 5.

From all runs conducted, the optimal DCs allocation to plants are; Iringa plant supplies to Dar, Arusha and Dodoma, and Makambako and Shinyanga supplied by Mbeya plant. This is from the fact that the plants capacities are more than sufficient as compared to DCs' capacities (See table 1). The Ruvuma and Rukwa plants remain untouched!

Table 5 results as indicated, optimal location and allocation attained with objective value of 7.567km for zero fixed costs and 57.567km for $f = 10$ where capacities of Makambako and Shinyanga in particular are redetermined. The optimal computed objective value is 11.9% and 8.2% less than constructed and maximum use capacities respectively. For both zero and non zero fixed costs, all DCs are to be opened. So the two level LRP is solved with 8.2% - 11.2% saving costs. The direct routes are then designed from optimal location and allocation as resulted from computations.

Table 5: The results for two level LRP with variation of capacities and fixed costs

DC/RUN	R1 (6,5,5,4,2: Existing Cap). Cus- tomers saved	R2 (4,5,3,1,5: Actual max use). Cus- tomers saved	R3 (6,5,5,2,4 = 6,5,4,2,5) Customers saved	R4 (6,5,5,4,2). Customers saved	R5 (6,5,5,2,4 = 6,5,4,2,5). Customers saved
Dar	2, 7, 11, 12, 17	2, 7, 12, 17	2, 7, 11, 12, 17	2, 7, 11, 12, 17	2, 7, 11, 12, 17
Arusha	1, 6, 8	1, 6, 8, 9	1, 6, 8	1, 6, 8	1, 6, 8
Dodoma	3, 5, 9, 13, 15	3, 4, 11	3, 9, 15	3, 5, 9, 13, 15	3, 9, 15
Makambako	4, 10	10	4, 10	4, 10	4, 10
Shinyanga	14, 16	5, 13, 14,15, 16	5, 13, 14, 16	14, 16	5, 13, 14, 16
Objective Value (km)	8.593	8.24	7.567	58.593	57.567
Run Time (Sec)	15	10	8	14	13
Fixed cost	0	0	0	10	10

NOTE: All DCs are open and all customers are saved for capacities and fixed cost variations

5. Conclusion and recommendations

The optimal costs of the two stages (FLP and LRP) are of great importance to distribution network design for food security in Tanzania. As far as food security in Tanzania is concerned, the FLP is more important to Prime Minister's office (Disaster Department) as they work as independent to NFRA that store foods in DCs. In order to meet minimum cost, they might ask NFRA to use only Dar Es Salaam, Arusha and Dodoma DCs in order reduce cost by 11.2%. And from optimal LRP, NFRA can buy only maize from Iringa for that matter and save 8.2% of their costs.

The given saving cost which is in km distance, can be converted to money value. For example the unit cost in 2010 was Tshs 148 per Km per Kg (NFRA source), equivalent to 673km x 1000kg (one ton) x 148 = Tshs 99,604,000 at least. It is an important cost savings to consider. From the fact that only maize has been stored in DC and not other cereals (Rice, Sorghum, Millet and Wheat), it is recommended to store and trade all cereals. This can be done by having enough DCs. For example in 2010/11 cereal surplus was 714,543 tons and only 241,000 can be stored (only 34%). This is only maize from four regions. So re-evaluation of the storage strategies is highly needed. For the current and future DCs expansion as mentioned theoretically to have total capacity of 400,000 (construction of 159,000 capacity DCs), two strategic places or zones as per this study are Shinyanga and Arusha. This is drawn from actual maximum usage of the DCs (each 5 customers to service). The second reason is a frequent food deficit in our neighbour countries which are Somalia, South Sudan and Kenya as from 2012/13 budget speech [5].

Generally, the food storage infrastructure, capacity and location is still alarming to most crops and hence a serious address is highly needed. Further research should be done by considering distances from each demand district to the DCs.

References

- [1] Agriculture first document. <http://www.tzonline.org/pdf/tenpillarsokilimokwanza.pdf> (accessed on 07/05/2010), 2009.
- [2] Ahuja, R., Magnanti, T. L. and Orlin, J. B. *Network flows: Theory, algorithms and applications*. Prentice Hall, Englewood Cliffs, NJ., 1993.
- [3] Ambrosino, D. and Scutella, M.G. *Distribution network design: New problems and related models*. European Journal of Operational Research, **165** , 610-624, 2005.
- [4] Berger, R. T., Coullard, C. R. and Daskin, M. S. *Location-routing problems with distance constraints*. Transportation Science, Vol. 14, **1**, 29-43, 2007.
- [5] The 2012/2013 MAFC budget. <http://www.agriculture.go.tz/speeches/budgetspeeches/budgetspeeches.htm>. (accessed on 04/08/2012), 2012.
- [6] Eiselt, H. A. *Locating landfills-optimization vs. reality*. European Journal of Operational Research **179**, 1040-1049, 2007.
- [7] Elhedhli, S. and Goffin, J. L. *Efficient Production-distribution system design*. Management science, vol. 51, **7** , 1151-1164, 2005.
- [8] Hindi, K. S., Basta, T. and Pienkosz, K. *Efficient solution of a multi-commodity, two stage distribution problem with constraints on assignment of customers to distribution centre's*. Information Systems in Logistics and Transportation, Vol. 5, **6** , 519-527, 1998.
- [9] Jabal-Amelia, M. S., Aryanezhada, M. B. and Ghaffari-Nasaba, N. *A variable neighborhood descent based heuristic to solve the capacitated location-routing problem*. International Journal of Industrial Engineering Computations 2, 141154, 2011.
- [10] Jacobsen, S. K. and Madsen, O. B. G. *A comparative study of heuristics for a two-level location-routing problem*. European Journal of Operational Research, **5**, 378-387, 1980.
- [11] Jiang, W., Tang, L. and Xue, S. *A hybrid algorithm of Tabu search and Benders decomposition for multi-product production distribution network design*. Proceedings of the International Conference on Automation and Logistics, 2009.
- [12] Koksalan, M., Sural, H. and Kirca, O. *A location-distribution application for a beer company*. European Journal of Operational Research **80**, 16-24, 1995.
- [13] Laporte, G. *Fifty Years of Vehicle Routing*. Transportation Science Vol. 43, **4**, pp. 408-416, 2009.
- [14] Lashine, S. H., Fattouh, M. and Issa, A. *Location/allocation and routing decisions in supply chain network design*. Journal of Modelling in Management Vol. 1, **2**, 173-183, 2006.
- [15] Madsen, O.B.G. *Methods for solving combined two level location-routing problems of realistic dimensions*. European Journal of Operational Research **12**, 295-301, 1983.
- [16] Mantel R. J. and Fontein, M. *A practical solution to a newspapers distribution problem*. International Journal of Production Economics, **30-31**, 591-599, 1993.

- [17] Min, H., Jayaraman, V. and Srivastava, R. *Combined location-routing problems: A synthesis and future research directions*. European Journal of Operational Research **108**, 1-15, 1998.
- [18] Nagy, G. and Salhi, S. *Location-routing: Issues, models and methods*. European Journal of Operational Research, **177**, 649-672, 2007.
- [19] Sajjadi, S. R., *Integrated supply chain: Multi products location routing problem integrated with inventory under stochastic demand*. PhD thesis, 2008.
- [20] Skipper, J. B. *An optimization of the hub-and-spoke distribution network in United States European command*. MSc thesis, 2002.
- [21] *United Republic of Tanzania - Food Security report*, 2006.
- [22] Yamada, T., Russ, B. F., Castro, J. and Taniguchi, E. *Designing multi-modal freight transport networks: A heuristic approach and applications*. Transportation Science, Vol. 43, **2**, 129-143, 2009.

A Survey of the Development of Fixed Point Theory

by

Santosh Kumar

University of Dodoma, Tanzania

Abstract

In this survey paper, we collected the developmental history of fixed point theory. Some important results from beginning up to now are incorporated in this paper.

1. Introduction

The fixed point theorem states the existence of fixed points under suitable conditions. Recall that in a case $f : X \rightarrow X$ is a function then y is a fixed point of f if $fy = y$ is satisfied. The topological fixed point theorem started by L. E. J. Brouwer. The famous Brouwer fixed point theorem was given in 1912 [5].

2. Brouwer fixed point theorem:

The theorem states that if $f : B \rightarrow B$ is a continuous function and B is a ball in \mathbb{R}^n , then f has a fixed point. This theorem simply guarantees the existence of a solution, but gives no information about the uniqueness and determination of the solution. For example, if $f : [0, 1] \rightarrow [0, 1]$ is given by $fx = x^2$, then $f0 = 0$ and $f1 = 1$, that is, f has 2 fixed points.

Several proofs of this theorem are given. Most of them are of topological in nature. A classical proof due to **Birkhoff and Kellogg was given in 1922**, Similar classical proof was given in Linear Operators Volume 1, **Dunford and Schwartz 1958**. Brouwer theorem gives no information about the location of fixed points. However, effective methods have been developed to approximate the fixed points. Such tools are useful in calculating zeros of functions.

A polynomial equation $px = 0$ can be written as $Fx = x$ where $Fx - x = Px$. For example, consider $x^2 - 7x + 12$, where $Px = x^2 - 7x + 12$. We can write $Fx - x = Px = x^2 - 7x + 12$, so $x = (x^2 + 12)/7 = Fx$. Here F has two fixed points, $F3 = 3$ and $F4 = 4$.

The following books cover a good deal of fixed point theorems. [1], [2], [4] and [31]. This theorem is not true in infinite dimensional spaces. For example, if B is a unit ball in an infinite dimensional Hilbert space and $f : B \rightarrow B$ is a continuous function, then f need not have a fixed point. This was given by Kakutani in 1941 [18]. The first fixed point theorem in an infinite dimensional Banach space was given by [29]. It is stated below.

3. Schauder fixed point theorem

If B is a compact, convex subset of a Banach space X and $f : B \rightarrow B$ is a continuous function, then f has a fixed point [29]. The Schauder fixed point theorem has applications in approximation theory, game theory and other scientific area like engineering, economics and optimization theory. The compactness condition on B is a very strong one and most of the problems in analysis do not have compact setting. It is natural to prove the theorem by relaxing the condition of compactness. Schauder proved the following theorem [29].

Theorem 3.5 *If B is a closed bounded convex subset of a Banach space X and $f : B \rightarrow B$ is continuous map such that $f(B)$ is compact, then f has a fixed point.*

The above theorem was generalized to locally convex topological vector spaces by Tychonoff in 1935 [32].

Theorem 3.6 *If B is a nonempty compact convex subset of a locally convex topological vector space X and $f : B \rightarrow B$ is a continuous map, then f has a fixed point.*

Further extension of Tychonoff's theorem was given by Ky Fan [10]. A very interesting useful result in fixed point theory is due to Banach known as the Banach contraction principle [3].

Theorem 3.7 *Recall that a map $f : X \rightarrow X$ is said to be a contraction map, if $d(fx, fy) \leq kd(x, y)$ where X is a metric space, $x, y \in X$ and $0 \leq k < 1$. Every contraction map is a continuous map, but a continuous map need not be a contraction map.*

For example, $fx = x$ is a continuous map but it is not a contraction map. The method of successive approximation introduced by Liouville in 1837 and systematically developed by Picard in 1890 culminated in formulation by Banach known as the Banach contraction principle (BCP) is stated as below [3].

Theorem 3.8 *If X is a complete metric space and $f : X \rightarrow X$ is a contraction map, then f has a unique fixed point or $fx = x$ has a unique solution.*

Proof:

The proof of this theorem is constructive. Let $x_{n+1} = fx_n, n = 1, 2, 3, \dots$. Then the sequence $\{x_n\}$ is a Cauchy sequence to y in X . It is easy to show that $y = fy$, that is, y is a fixed point of f . Since f is a contraction map so y is a unique fixed point. \square

The Banach contraction principle is important as a source of existence and uniqueness theorems in different branches of sciences. This theorem provides an illustration of the unifying power of functional analytic methods and usefulness of fixed point theory in analysis.

The important feature of the Banach contraction principle is that it gives the existence, uniqueness and the sequence of the successive approximation converges to a solution of the problem. The important aspect of the result is that existence, uniqueness and determination, all are given by Banach contraction principle.

Definition 3.9 *If $f : X \rightarrow X$ such that $d(fx, fy) \leq d(x, y)$, for all $x, y \in X$, then f is said to be a nonexpansive map.*

A nonexpansive map need not have a fixed point in a complete metric space. For example, if $f : R \rightarrow R$ given by $fx = x + k$ where k is any number, then f has no fixed point.

A translation map has no fixed point. In case we have an identity map $I : R \rightarrow R$, then each point of I is a fixed point. The above examples illustrate that a nonexpansive map, unlike contraction map, need not have a fixed point and if it has a fixed point, then it may not be unique.

The famous fixed point theorem for nonexpansive maps was given by Browder [6], Kirk [19] and Godhe [13] independently in 1965.

Theorem 3.10 *If B is a closed bounded convex subset of a Hilbert space H and is a nonexpansive map, then f has a fixed point.*

The following interesting question was proved by Browder in 1967 [7].

If B is closed convex subset of a Banach space X and $f : B \rightarrow B$ is a nonexpansive map. If for each $r_i \in [0, 1)$ and any $y \in B$, we define $f_{r_i}x = r_ix + (1 - r_i)y$ for all $x \in B$, then $f_{r_i} : B \rightarrow B$, and each f_{r_i} is a contraction map with Lipschitz constant r_i . For r_i sufficiently close to 1, f_{r_i} is a contractive approximation of f .

By Banach contraction principle, each has a unique fixed point say $f_{r_i}x_{r_i} = x_{r_i}$ for each r_i , that is $x_{r_i} = f_{r_i}x_{r_i} = r_ifx_{r_i} + (1 - r_i)y$. It is natural to ask if the sequence $\{x_{r_i}\}$ converges to a fixed point of f .

Since a nonexpansive map need not have a fixed point so in general the result is not affirmative. However, the following is due to Browder [7].

Theorem 3.11 *If C is a closed bounded convex subset of a Hilbert space H and $f : C \rightarrow C$ is a nonexpansive map. Define $f_r x = rfx + (1 - r)y$ for some $y \in C$ and $0 < r < 1$. Let $x_r = f_r x_r$. Then, the sequence $\{x_r\}$ converges to a fixed point of f , closest to y .*

In case C is not bounded and f is not a self map, then a similar result is given in [30]. In the study of fixed point theorems of nonexpansive mappings the following topics are of interest.

- (i) The sequence of iterates $x_{n+1} = fx_n$ need not converge.
For example, if we consider $fx = -x$, for $x \in \mathbb{R}$, then the sequence of iterates is an oscillatory sequence.
- (ii) The nonexpansive map need not have a fixed point. Therefore for the study of nonexpansive map it is important to find that under what conditions the mapping is going to have a fixed point.

Here we give a brief development of the above areas.

The method of successive approximation is useful in determining the solutions of equations. An early result dealing with the convergence of the sequence of iterates was given by Krasnoselskii in 1955. It is stated below [20].

Theorem 3.12 *If C is a closed bounded convex subset of a Banach space X and $f : C \rightarrow C$ a nonexpansive mapping with closure of $f(C)$ compact, then the sequence of iterates given by $(f_{1/2})^n x$ where $f_{1/2}x = \frac{1}{2}fx + \frac{1}{2}x$, converges to a fixed point of f .*

We note here that the fixed point of f and $f_{1/2}$ is the same. For example, if $fy = y$, then $f_{1/2}y = y$. The limit of the sequence $\{(f_{1/2})^n x\}$ converges to a fixed point of f .

More generally, if C is a closed bounded convex subset of a Banach space X , then for $f : C \rightarrow C$, we consider $f_r x = rfx + (1 - r)x$. In this case it is easy to see that $fy = y$ if and only if $f_r y = y$ and the sequence of iterates $(f_{1/2})^n x$ converges to a fixed point of f .

Further extensions of iteration process due to Mann [21], Ishikawa [15], and Rhoades [27] are worth mentioning. Recently several interesting results for sequence of iterates are used to find the solutions of the Variational Inequality Problems (VIP). In most of the cases the basic tool has been the sequence of successive approximation used in the study of fixed point theory. A good deal of work has been associated with the nonexpansive maps. As the sequence of iterates for a nonexpansive map need not always converge therefore several researchers have tried to give techniques for convergence of the sequence of iterates. The following result deals with the contraction maps in the study of variational inequality [23].

Theorem 3.13 Let C be a nonempty closed convex subset of a Hilbert space H and $f : C \rightarrow H$ a continuous map such that $I - rf$ is a contraction map. Then the sequence of iterates

$$u_{n+1} = Po(I - rf)u_n u_0 \in C$$

converges to u where u satisfies the variational inequality $\langle fu, y - u \rangle \geq 0$ for all $y \in C$.

Singh et. al. proved the following result for nonexpansive maps [31].

Theorem 3.14 Let C be a closed convex subset of a Hilbert space H and $f : C \rightarrow H$ a continuous function such that $I - f$ is a nonexpansive map and let $(I - f)C$ be bounded. Then the sequence of iterates $u_{n+1} = Po(I - f)u_n, n = 1, 2, \dots, u_i \in C$ converges to u where u is a solution of the variational inequality $\langle fu, x - u \rangle \geq 0$ for all $x \in C$, provided that $\lim_{n \rightarrow \infty} d(u_n, F) = 0$, where F is the set of fixed points of $Po(I - f) : C \rightarrow C$.

The VIP is also closely associated with the best approximation problem so this technique can be applied to problems in approximation theory.

The following example is worth mentioning [8].

Theorem 3.15 Let C_1 and C_2 be two closed convex sets in Hilbert space H and $g = P_1P_2$ of proximity maps. Convergence of $\{x_n\}$ to a fixed point of g is guaranteed if either

- (i) one set is compact or
- (ii) one set is finite dimensional and the distance between the sets is attained.

The contraction, contractive and nonexpansive maps have been further extended to densifying, and 1- set contraction maps in 1969. Several interesting results of fixed points were proved recently. A few results were proved separately for contraction maps and compact mappings (A continuous map with compact image is called a compact mapping). Both maps are densifying maps. Thus a fixed point theorem for densifying maps includes both for contraction and compact maps.

If $f : B \rightarrow \mathbb{R}^n$, then f is said to be a nonself map. Most of the fixed point theorems have been given for self-maps. In 1937 Rothe [9] gave a fixed point theorem for nonself maps [see also [2], [32].

Theorem 3.16 If $f : B \rightarrow \mathbb{R}^n$ is a continuous map, such that

$$f(\partial B) \subseteq B$$

then f has a fixed point.

The following condition for nonself map is called the Altman's condition (1955)

$$|fx - x|^2 \geq |fx|^2 - |x|^2.$$

. There were a few results in fixed point theory dealing with combination for two maps- say one is contraction and the other one is compact.

Note that if we have f and g both continuous functions, then $f + g$ is also a continuous map and the fixed point theorem for continuous map is applicable for $f + g$. However, if f is a contraction map, then Banach contraction theorem is applied and if g is a compact map, then Schauder fixed point theorem is applicable. However, in such a case when f is contraction and g is a compact map, then for $f + g$ the fixed point theorem of densifying map is applicable.

We record a few definitions [2], [31]:

Definition 3.17 Let C be a bounded subset of a metric space X . Define the measure of noncompactness $\alpha(C) = \inf\{\varepsilon > 0 / C \text{ has a finite covering of subsets of diameter } \leq \varepsilon\}$.

The following properties of α are well known.

Let A be a bounded subset of a metric space X . Then

$\alpha(A) \leq \delta(A)$, $\delta(A)$ is the diameter of A .

If $A \subseteq B$, then $\alpha(A) \leq \alpha(B)$,

$\alpha(\text{closure of } A) = \alpha(A)$

$\alpha(A \cup B) = \max\{\alpha(A), \alpha(B)\}$

$\alpha(A) = 0$ if and only if A is a precompact.

Definition 3.18 A continuous mapping $f : X \rightarrow X$ is called a densifying map if for any bounded set A with $\alpha(A) > 0$, we have $\alpha f(A) < \alpha(A)$.

In case $\alpha f(A) \leq \alpha(A)$, then f is said to be 1-set contraction. Note that a nonexpansive map is an example of 1-set contraction.

A contraction map is densifying and so is the compact mapping, that is, a function mapping closed sets to compact sets.

The following is a well known result [12], [24], [28].

Theorem 3.19 Let $f : C \rightarrow C$ be a densifying map, where C is closed bounded and convex subset of a Banach space X . Then f has at least one fixed point in C .

The contraction, contractive and nonexpansive maps have been further extended to densifying, and 1-set contraction maps in 1969. Several interesting results of fixed points were proved recently [26]. A few results were proved separately for contraction maps and compact mappings (A continuous map with compact image is called a compact mapping). Both maps are densifying maps. Thus a fixed point theorem for densifying maps includes both for contraction and compact maps.

In 1966 Hartman and Stampacchia [14] gave the following interesting result in variational inequalities.

Theorem 3.20 If B is a unit ball in \mathbb{R}^n and $f : B \rightarrow \mathbb{R}^n$ a continuous function, then there is a y such that

$$\langle fy, x - y \rangle \geq 0 \text{ for all } x \in B. \quad (3.1)$$

Note: Let P be a metric projection onto B . Then $P(I - f)$ has a fixed point in B if and only if (1) has a solution.

The variational inequality theory is a very effective tool for handling problems in different branches of mathematics, engineering and theoretical physics. Hartman and Stampacchia [14] theorem yields Brouwer fixed point theorem as an easy corollary.

Let $g : B \rightarrow B$ be a continuous function, where B is a closed ball in \mathbb{R}^n . We have to show that g has a fixed point.

Let $f = I - g$. Then f is continuous and $f : B \rightarrow \mathbb{R}^n$. Hence by using Hartman and Stampacchia theorem we get that there is a $y \in B$ such that $\langle fy, x - y \rangle \geq 0$ for all $x \in B$.

Thus, $\langle (I - g)y, x - y \rangle \geq 0$, that is $\langle y - gy, x - y \rangle \geq 0$. Since $g : B \rightarrow B$, so by taking $x = gy$, we have $\langle y - gy, gy - y \rangle \geq 0$. This is true only when $y = gy$. Hence g has a fixed point.

In 1969 the following result was given by Ky Fan commonly known as the best approximation theorem [11].

Theorem 3.21 *If C is a nonempty compact convex subset of a normed linear space X and $f : C \rightarrow X$ is a continuous function, then there is a $y \in C$ such that*

$$|fy - y| = \inf\{x - fy\} \text{ for all } x \in C. \quad (3.2)$$

If P is a metric projection onto C , then Pof has a fixed point if and only if 3.2 holds.

Recall that $d(x, C) = \inf\{\|x - y\|\}$ for all $y \in C, x \notin C$.

The Ky Fan's theorem has been widely used in approximation theory, fixed point theory, variational inequalities, and other branches of mathematics.

Theorem 3.22 *If $f : B \rightarrow X$ is a continuous function and one of the following boundary conditions are satisfied, then f has a fixed point. Here B is a closed ball of radius r and center 0 (δB stands for the boundary of the ball B).*

(i) $f(\delta B) \subseteq B$ (Rothe condition)

(ii) $|fx - x|^2 \geq |fx|^2 - |x|^2$, (Altman's condition)

(iii) If $fx = kx$ for $x \in \delta B$ then $k \leq 1$, (Leray Schauder condition)

(iv) If $f : B \rightarrow X$ and $fy \neq y$, then the line segment $[y, fy]$ has at least two elements of B . (Fan's condition).

In this survey we have restricted our presentation to single valued maps only. A vast literature is available for the fixed point theorems of multivalued maps. In Kakutani [18] gave the following generalization of the Brouwer fixed point theorem to multivalued maps.

Theorem 3.23 *If F is a multivalued map on a closed bounded convex C subset of \mathbb{R}^n , such that F is upper semicontinuous with nonempty closed convex values, then F has a fixed point.*

Recall that x is a fixed point of F if $x \in Fx$.

The fixed point theory of multivalued maps is useful in economics, game theory and minimax theory.

An important application of Kakutani fixed point theorem was made by Nash [17] in the proof of existence of an equilibrium for a finite game. Other applications of fixed point theorem of multivalued mapping are in mathematical programming, control theory and theory of differential equations.

Popa [33],[34] introduced implicit functions which are proving fruitful due to their unifying power besides admitting new contraction conditions. We also introduce an implicit function to prove our results [22]. The main theorem is listed below:

Theorem 3.24 *Let $\{S_1, S_2, \dots, S_m\}, \{T_1, T_2, \dots, T_n\}, \{I_1, I_2, \dots, I_p\}$ and $\{J_1, J_2, \dots, J_q\}$ be four families of self-mappings of a metric space (X, d) with $S = S_1 S_2 \dots S_m, T = T_1 T_2 \dots T_n, I = I_1 I_2 \dots I_p, J = J_1 J_2 \dots J_q$ satisfying the following conditions:*

- (a) $S(X) \subset J(X), T(X) \subset I(X)$,
- (b) *one of $S(X), T(X), I(X)$ and $J(X)$ is complete subspace of X ,*
- (c) $F(d(Sx, Ty), d(Ix, Jy), d(Ix, Sx), d(Jy, Ty), d(Ix, Ty), d(Jy, Sx)) \leq 0$ for all $x, y \in X$ and $F \in \tau$. Then
- (d) (S, I) have a point of coincidence,
- (e) (T, J) have a point of coincidence.

Moreover if $S_i S_j = S_j S_i, I_k I_l = I_l I_k, T_r T_s = T_s T_r, J_t J_u = J_u J_t, S_i I_k = S_k I_i, I_k T_r = T_r I_k, T_r J_t = J_t T_r, S_i J_t = J_t S_i, S_i T_r = T_r S_i$ and $J_t I_k = I_k J_t$ for all $i, j \in I_1 = \{1, 2, \dots, m\}, k, l \in I_2 = \{1, 2, \dots, p\}, r, s \in I_3 = \{1, 2, \dots, n\}$ and $t, u \in I_4 = \{1, 2, \dots, q\}$.

Then (for all $i \in I_1, k \in I_2, r \in I_3$ and $t \in I_4$), S_i, I_k, T_r and J_t have a common fixed point.

The most recent result for implicit functions is due to Javid Ali and M. Imdad [16]. They introduce an implicit function to prove their results because of their versatility of deducing several contraction conditions in one go.

References

- [1] Agarwal, R. P., Meehan, M. and O' Regan, D., Fixed point theory and applications, Cambridge University Press 2001.
- [2] Andrzej, G. and Dugundji, A., "Fixed Point Theory", Springer Verlag, New York (2003).
- [3] Banach, S., Sur les operations dans les ensembles abstraits et leur applications aux equations integrales, Fund. Math. 3 (1922), 133 - 181.
- [4] Border, K. C., Fixed point theorems with applications to economics and game theory, Cambridge University Press, 1985.
- [5] Brouwer, L. E. J., Uber Abbildung von Mannigfaltigkeiten, Math Ann. 71 (1912), 97 - 115.

- [6] Browder, F. E., Fixed point theorems for noncompact mappings in Hilbert spaces, Proc. Nat. Acad. Sci. USA 53 (1965), 1272 - 1276.
- [7] Browder, F. E., Convergence of approximants to fixed point of nonexpansive nonlinear mappings in Banach spaces, Arch. Rat. Mech. Anal. 21 (1966), 259 - 269.
- [8] Cheney, E. W. and Goldstein, A. A., Proximity maps for convex sets, Proc. Amer. Math. Soc. 10 (1959), 448 - 450.
- [9] Rothe, E. Zur Theorie der topologischen Ordnung und der Vektorfelder in Banachschen Raumen, Compositio Math. 5 (1937), 177-197.
- [10] Fan, Ky, A generalization of Tychonoff's fixed point theorem, Math. Ann. 142 (1961), 305 - 310
- [11] Fan, Ky, Extensions of two fixed point theorems of F. E. Browder, Math. Z. 112 (1969), 234 - 240.
- [12] Furi, M. and Vignoli, A., Fixed point theorems in complete metric spaces, Bull Unione Mat Italiana 2 (1969), 505 - 509.
- [13] Gohde, D., Zum prinzip der kontraktiven abbildung, Math. Nachr. 30(1965), 251 - 258.
- [14] Hartman, P. and Stampacchia, G., On some nonlinear elliptic differential equations, Acta Math. 115 (1966), 271 - 310.
- [15] Ishikawa, S., Fixed points by new iteration method, Proc. Amer Math Soc. 44 (1974), 147 - 150.
- [16] Javid Ali and M. Imdad, An implicit function implies several contraction conditions, Sarajevo Journal of Mathematics 4(17) (2008), 269 - 285.
- [17] John F. Nash Jr., "Equilibrium points in n-person game", Proceedings of the National Academy of Science, USA, 36 (1950), 48 - 49.
- [18] Kakutani, S., A generalization of Brouwer fixed point theorem, Duke Math. Journal, 8 (1941), 457 - 459.
- [19] Kirk, W. A., A fixed point theorem for mappings which do not increase distances, Amer. Math. Monthly, 72 (1965), 1004 - 1006.
- [20] Krasnosel'skii, M. A., Two remarks on the method of successive approximations, Uspekhi Mat. Nauk, 10: 1 (63) (1955), 123 - 127.
- [21] Mann, W. R., Mean value methods in iteration, Proc. Amer. Math. Soc. 4 (1953), 506- 510.
- [22] M. Imdad and S. Kumar, Remarks on some fixed point theorems satisfying implicit relations, Radovi Matematicki, 11(1) (2002), 135-143.
- [23] Noor, M. A., An iterative algorithm for variational inequalities, J. Math. Anal. Appl. 158 (1991), 448-455.
- [24] Nussbaum, R. D., Some fixed point theorems, Bull. Amer. Math. Soc. 77 (1971), 360 - 365.
- [25] Park, S., Ninety years of the Brouwer fixed point theorem, Vietnam J. Math 27 (1999), 187 - 222.
- [26] Petryshyn, W. V., Structure of fixed point sets of the k-set contractions, Arch. Rat. Mech. Anal. 40 (1971), 312 - 328.

- [27] Rhoades, B. E., Comments on two fixed point iteration methods, *J. Math. Anal. Appl.* 56 (1976), 741 - 750.
- [28] Sadovskii, B. N., A fixed point principle, *Funct. Anal. Appl.*, 1 (1967), 151- 153.
- [29] Schauder, J., Der Fixpunktsatz in Funktionalraumen, *Studia Math.* 2 (1930), 171- 180.
- [30] Singh, S. P. and Watson, B., On approximating fixed points, *Proc. Symp. Pure Math. Amer. Math. Soc. Ed. F. E. Browder* 45 (1986), 393- 395.
- [31] Singh S. P., Watson. B. and Srivadtava, P., *Fixed Point Theory and Best Approximation: The KKM map Principle*, Kluwer Academic Publishers (1997) p. 220.
- [32] Tychonoff, A., Ein Fixpunktsatz, *Math. Ann.* 111 (1935), 767-776.
- [33] Popa, V. Fixed point theorems for implicit contractive mappings, *Stud. Cercet. Stiint. Ser. Mat. Univ. Bacau*, 7 (1997), 127 - 133.
- [34] Popa, V. Some fixed point theorems for compatible mappings satisfying an implicit relation, *Demonstratio Math.* 32(1) (1999), 157 - 163.
- [35] Zeidler, E., *Nonlinear functional analysis and applications I*, Springer Verlag, New York (1985).

Epidemiological Modelling at Macro and Micro Levels: The Case of HIV/AIDS

by

Livingstone S. Luboobi

Makerere University, Uganda.

Ecological epidemiology (macro and micro levels)

Ecological models are very important in Epidemiology:

- No disease or epidemic can progress without a population or individual.
- Population dynamics in single/multi-species communities facilitate the epidemiological studies through processes of:
 - Births/reproduction;
 - Deaths;
 - Immigration/immigration.
- Interactions between individuals/species:
 - Prey-predator relationships;
 - Competition;
 - Symbiosis;
 - Obligatory cooperation;
 - Food chain.

Modelling at macro level

- Requires a community/ecosystem
 - Individuals
 - Species
- There are interactions between individuals/species
- Hence ecological considerations are important

Concerned with what happens to/within an individual

- Interplay of different systems of cells within body such as the immune system, the nervous system (the brain), heart, liver, etc.
- Thus an “ecosystem” within an individual/organ.

Models of Interactions of Multi-species communities

Inter-interactions as well as intra-interactions: The rate of growth of i -th species sub-population and n species community through an equation such as:

$$\frac{dN_i}{dt} = g_i(N_1, N_2, \dots, N_n); \quad i = 1, 2, \dots, n$$

The form of g_i depends on the type of interaction.

Stage/Age structured models

- Human populations
 - Immature age-group
 - Mature age-group (i.e. the adults capable of reproduction)
 - Even a third age-group that have stopped giving births
 - Sex-age structured model could be closer to reality
- Application to HIV/AIDS epidemic: 0 – 5, 5 – 12/15 years, adults sub-populations.
- Method of analysis” delay differential equations.

Stochastic models

Why?

- Can derive details
 - Expectation
 - Variance
 - Probability distributions
- E.g. in the birth-death process $\{N(t), t \geq 0\}$
 - Deterministic model indicates exponential growth or decline $N(t) = N_0 e^{(\lambda - \mu)t}$
 - In the corresponding stochastic process $\{N(t), t \geq 0\}$ we can show:
 - * There is a possibility of extinction of the population
 - * Extinction is certain when the birth rate is equal or less than the death rate

Epidemiological

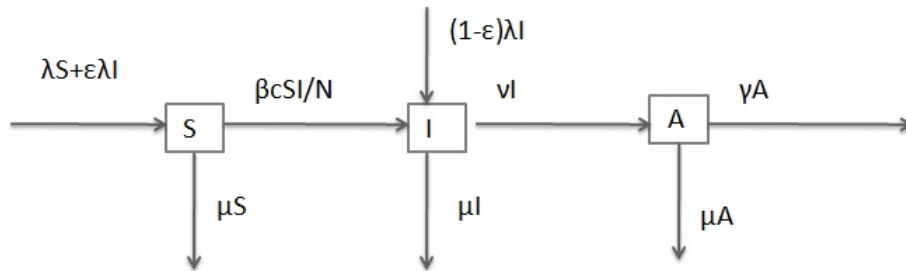
At macro level

- Infectious diseases cannot spread or be transmitted without a population(s)
- Mode of transmission is key in study of the epidemiology of a disease
- Examples

- Compartmentalised/structured populations such as *Susceptibles–Latents–Infectives– Recovered–immunes*
- There may be other stages

HIV/AIDS Macro Level Model

Simple model (early stages)



where

- $S(t)$ = number of susceptibles (i.e., the 'non-infected') at time t
- $I(t)$ = number of infectives (i.e. the infected and are infectious) at time t ;
- $A(t)$ = number of the AIDS cases (bedridden or too weak to interact) at time t .

The equations:

$$\begin{aligned}\frac{dS}{dt} &= \lambda S + \epsilon \lambda I - \beta c S \frac{I}{N} - \mu S \\ \frac{dI}{dt} &= \beta c S \frac{I}{N} + (1 - \epsilon) \lambda I - \nu I - \mu I \\ \frac{dA}{dt} &= \nu I - \mu A - \gamma A\end{aligned}$$

Quick analysis of early stages of HIV in a community:

$S(t) \approx N(t)$ hence

$$\frac{dI}{dt} \approx (\beta c + (1 - \epsilon) \lambda - \nu - \mu) I$$

Thus if

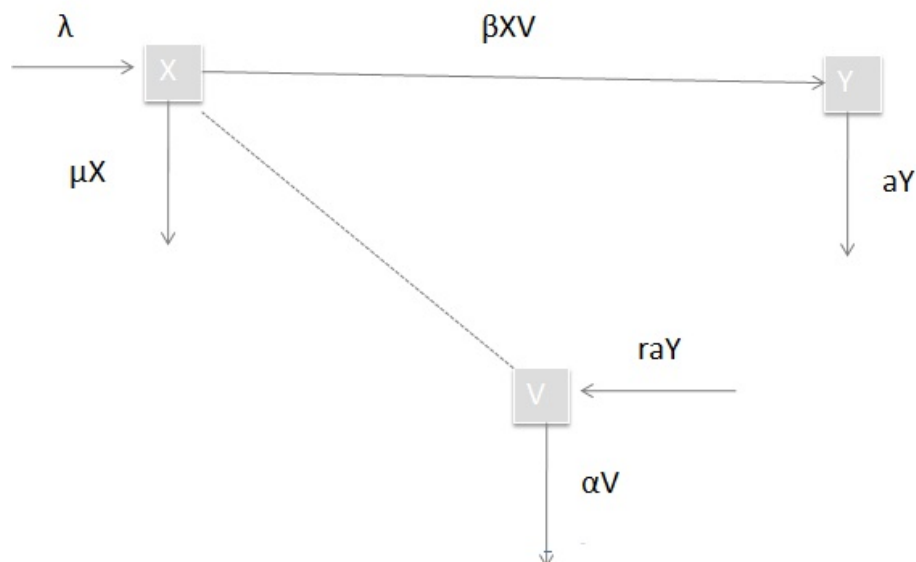
$$R_0 = \frac{\beta c + (1 - \epsilon) \lambda}{\mu + \nu} < 1$$

then HIV/AIDS epidemic would not develop in that community.

HIV/AIDS micro level model

- The development of AIDS is associated with the depletion of the CD4+ helper T lymphocyte.
- HIV relies on a host to assist reproduction.

- Since the CD4+ cells are depleted over time, strengthening cytotoxic responses cannot occur.
- Initially the transformation of immun-sensitivity to resistant genotypes occurs by the generation of mutations primarily due to reverse transcriptase.
- The extreme heterogeneity and diversity of HIV makes the design of effective vaccines extremely difficult.
- The understanding of the dynamics of antigenic escape from immunological response has been that a mutation may enable the virus to have a selection advantage.
- Because there is an asymmetric interaction between immunological specificity and viral diversity, the antigen diversity makes it difficult for the immune system to control the different mutants simultaneously and the virus runs ahead of the immune response.
- While most productively infected cells have a relatively short life span, many cells are latently infected and are very long lived.
- A simple model for the interaction between the human immune system and HIV was developed by Perelson (2002).
- A stochastic model for the HIV pathogenesis under anti-viral drugs has been developed.
- Thus:
 - The immune system offers a natural and the most reliable defense mechanism against HIV:
 - * Interactions of the Virions, CD4+ and CD8+ T-cells of the immune system
 - * Hence the terms viral load and “CD4 cell count”
 - HIV also infects the liver cells: the hepatocytes.



where X is the uninfected CD4 cells, Y the infected CD4 cells and V the free HIV virions.

The parameters are described as follows:

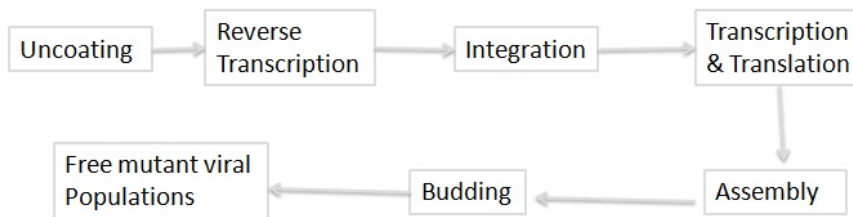
- β = CD4+ T-cell infection rate by HIV.
- a = the death rate of infected CD4+ T-cells.
- α = the rate of removal of free virus from the system.
- r = number of free virus particles from an infected cell as result of bursting.
- λ = constant rate of production of uninfected CD4+ T-cells.
- μ = death rate of uninfected CD4+ T-cells.

The model equations are

$$\begin{aligned} \frac{dX}{dt} &= \lambda - \beta XV - \mu X \\ \frac{dY}{dt} &= \beta XV - aY \\ \frac{dV}{dt} &= arY \end{aligned}$$

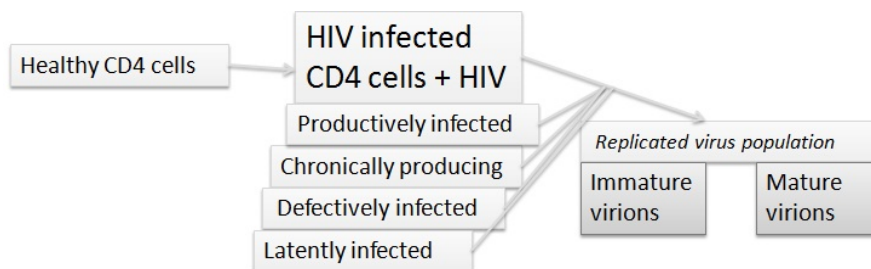
Combined macro-micro epidemiologic dynamics of HIV/AIDS

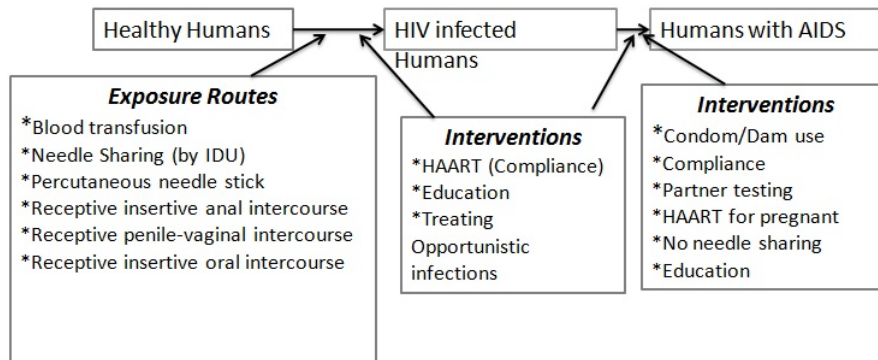
Micro level intracellular level kinetics



Intervention Strategies:

- Inhibition of binding. Blocking of the gp41 conformational changes that permit viral fusion
- Nucleoside/Nucleotide Reverse Transcriptase Inhibitors (NRTIs) & Non-Nucleotide Reverse Transcriptase Inhibitors
- Integrase inhibitors
- Antisense antivirals or transcription Inhibitors (TIs)
- Protease inhibitors (PI) [Tameru et al., 2010: in Ethnicity & Disease, vol.20, pp SI-207-210]





Therapy

- The ARVs are used as drugs to control the effects of HIV
- But they are toxic to the hepatocytes hepatotoxicity
- Hence an optimal therapeutic programme is the concern of the Research team:

Best way of generating models in epidemiological research

There is need to work with: Ecologists, public health officers, physicians, pharmacologists, hematologists, gastrosurgeons etc

Analysis of Cell Phone User's Loyalty in Tanzania Using Markov Chains

by

Christian Baruka Alphonice

University of Dar es salaam, Tanzania

Abstract

Markov chains, applied in marketing problems, are principally used for Brand Loyalty studies. Especially, Markov chains are strong techniques for forecasting long term market shares in oligopolistic markets. The concepts of marketing studies are thought as discrete from the time and place view point. And so finite Markov chains are applicable for this kind of process. The aim of this study was to examine the cell phone user's Loyalty using Markov chains. In this study, the data to examine cell phone user's Loyalty were obtained by interviewing 400 subscribers from ten out of twenty seven wards of Kinondoni District in Dar es salaam, Tanzania.

1. Introduction

The Tanzanian mobile communication market has enjoyed impressive growth in terms of numbers of operators as well as number of subscribers over the past few years. As illustrated in Table 1. Currently there are eight licensed companies, out of which six are currently operational. There are over 10 million voice subscribers [11]. The operational companies are Vodacom, Airtel(zain,celtel) Tigo(Buzz, Mobitel), zantel, TTCL and Benson. The first company to provide mobile phone services in Tanzania was Mobitel. Tritel company, which no longer exists, was the second mobile operator. Four more operators joined later: Vodacom, Celtel(Airtel), TTCLMobile and Benson. These operators and their subscriber bases are shown in Table 2.

Table 1: Voice Telecommunication Operators in Tanzania since 2000

Years	Voice Telecom Operators	Application Service (Internet and other Data)
2000	5	11
2001	6	17
2002	6	20
2003	5	22
2004	5	23
2005	5	23
2006	6	25
2007	8*	34
June – 08	8*	42

Source: Tanzania Communications Regulatory Authority (2008)

8* licensed and 6 operational.

Tanzania has the second largest mobile communications market in East Africa with 11% penetration rate while Uganda and Kenya have 6% and 15% penetration rate respectively [13]. The rate at which Tanzanians are embracing mobile communications technology indicates that there is significant potential for future growth. On the other hand landline telephone growth is insignificant over the past eight years

Table 2: Number of Mobile and Fixed Phone Voice Subscribers

YEAR	BENSON	CELTEL	TIGO	TTCL Fixed	TTCL Mobile	VODACOM	ZANTEL Mobile	ZANTEL Fixed	TOTAL x1000
2000	-	-	56,511	173,591	-	50,000	4,007	-	284
2001	-	-	89,056	177,802	-	180,000	6,501	-	453
2002	-	120,089	160,000	161,590	-	300,000	26,770	-	768
2003	-	320,000	210,000	147,006	-	700,000	68,000	-	1,445
2004	-	504,000	303,000	148,360	-	1,050,000	85,000	-	2,090
2005	-	882,693	422,500	154,420	-	1,562,435	96,109	-	3,118
2006	-	1,516,832	760,874	150,897	6,390	2,975,580	355,246	747	5,767
2007	3,300	2,505,546	1,191,678	157,816	72,729	3,870,843	678,761	5,453	8,486
2008/ june	3,000	2,819,828	1,701,433	153,230	155,251	4,520,120	1,069,035	6,140	10,428

Source: Tanzania Communications Regulatory Authority (2008)

if compared to mobile phone growth. This is due to problems with land line technology; problems such as unreliable fixed lines, common fixed lines faults, frequent connection break downs, frequent wrong bills, lack of innovative ideas and poor maintenance services. In the past it used to take a very long time to get a fixed telephone line installed, while today a walk to a mobile shop is all it takes to get a reliable affordable mobile phone.

The increase of voice subscribers and teledensity (Figures 1 and 2) could be attributed, firstly, to affordability and ease of maintenance of mobile phones, but, secondly, to the introduction of value added services in the mobile phone services, such as caller number display, voice mail, call forwarding, call waiting, conference calls, long distance Internet Protocol (IP) telephony, and short message services (SMS). In an effort to keep up with mobile commerce worldwide, these operators are aiming at launching nation-wide wireless application protocol (WAP) services. WAP is expected to offer mobile banking, stock trading, news, weather reports, and email services to a wide audience of subscribers.

2. Brand Loyalty

Customer loyalty has been a major focus of strategic marketing planning and offers an important basis for developing a sustainable competitive advantage - an advantage that can be realized through marketing efforts [1]. It is reported that academic research on loyalty has largely focused on measurement issues [2] and correlations of loyalty with consumer property in a segmentation context.

Many studies have been conducted on brand loyalty. However, in all of these studies brand loyalty (e.g. repeat purchase) has been measured from the behavioural aspect without considering the cognitive

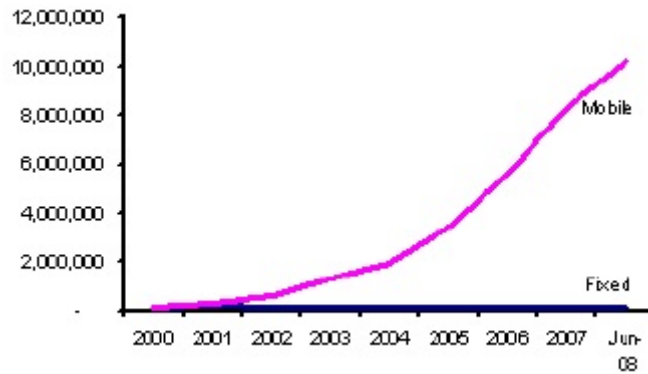


Figure 1: Voice Telecommunication Subscribers.

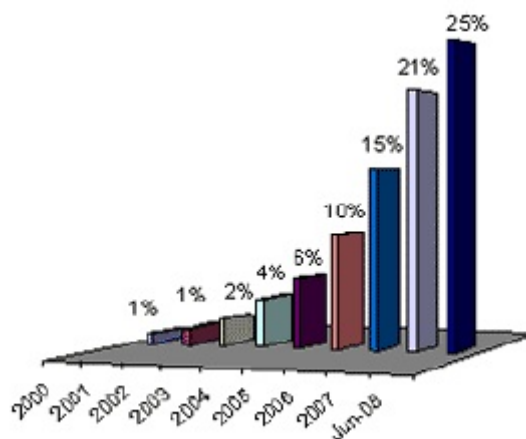


Figure 2: Teledensity in Tanzania.

aspects.

However, brand loyalty is not a simple uni-dimensional concept, but a very complex multi-dimensional concept. Wilkie [3] defines brand loyalty as a "favourable attitude toward, and consistent purchase of a particular brand". But such a definition is too simple for understanding brand loyalty in the context of consumer behaviour. This definition implies that consumers are brand loyal only when both attitudes and behaviours are favourable. However, it does not clarify the intensity of brand loyalty, because it excludes the possibility that a consumer's attitude may be unfavourable, even if he/she is making repeat purchases. In such a case, the consumer's brand loyalty would be superficial and shallow - rooted.

Another definition of brand loyalty that compensates for the incompleteness of Wilkie's definition [3] was offered by Jacobs and Chestnut [4]. They provided a conceptual definition where brand loyalty is (1) biased (i.e. non random), (2) behavioural response (i.e. purchase), (3) expressed over time, (4) by some decision making unit, (5) with respect to one or more brands out of a set of such brands, and is a function of psychological (decision-making, evaluative) processes.

Based on the behavioural element of brand loyalty, Lyong [5] provides an operational definition that "brand loyalty is a function of a brands' relative frequency of purchase in both time-independent and time dependent situations".

Brand loyalty represents a favourable attitude toward a brand resulting in consistent purchase of the

brand over time [6]. Two approaches to the study of brand loyalty have dominated marketing literature. The first is an instrumental conditioning approach, that views consistent purchasing of one brand over time as an indication of brand loyalty. Repeat purchasing behaviour is assumed to reflect reinforcement and a strong stimulus-to-response link. The research that takes this approach uses probabilistic models of consumer learning to estimate the probability of a consumer buying the same brand again, given a number of past purchases of that brand. This is a stochastic model rather than a deterministic model of consumer behaviour, as it does not predict one specific course of action. Rather, the prediction is always in probability terms.

The second approach to the study of brand loyalty is based on cognitive theories. Some researchers believe that behaviour alone does not reflect brand loyalty. Loyalty implies a commitment to a brand that may not be reflected by just measuring continuous behaviour.

Several authors have made distinctions between brand loyalty (in terms of repeat purchasing), and brand commitment (implying some degree of high involvement). The brand loyalty that is defined here is the observed behaviour of repeat purchasing of the same brand.

Behavioural measures have defined loyalty by the sequence of purchases (purchased Brand A give times in a row) and/or the proportion of purchases, in the event that the customer is satisfied with the brand purchase and repeats it in a relatively short period of time [7].

In order for managers to cope with the forces of disloyalty among consumers, there is a need to have an accurate method to measure and predict brand loyalty. However it was impossible to obtain an objective and general measurement of brand loyalty, because brand loyalty has been defined in many different ways and operationalized by a number of scholars. The diverse definition and operationalization of brand loyalty in the past has been due to the various aspects of brand loyalty (e.g. behavioural and attitudinal brand loyalty).

A transition matrix was used as a forecasting instrument for determining the market environment in the future by Stan and Smith in a research conducted in 1964. This paper shows the potential of using Markov Chains in determining the intensive transitional probabilities for a new product. These probabilities may help marketing management by comparing the intensiveness gained in a certain period of time with product life cycle. Thereby it may be possible to take the situation under control by taking corrective action.

Although the Markov Chains Method is quite successful in forecasting (predicting) brand switching, this model still has some limitations:

1. Customers do not always buy products in certain intervals and they do not always buy the same amount of a certain product. This means that in the future, two or more brands may be bought at the same time.
2. Customers always enter and leave markets, and therefore markets are never stable.
3. The transition probabilities of a customer switching from an *i* brand to an *j* brand are not constant for all customers, these probabilities may change from customer to customer and from time to time. These transitional probabilities may change according to the average time between buying situations.
4. The time between different buying situations may be a function of the last brand bought.
5. The other areas of the marketing environment such as sales promotions, advertising, competition etc. were not included in these models.

3. Markov Chains Method

The basic concepts of Markov Chains Method has been introduced by the Russian mathematician, Andrey Andreyevich Markov, in 1970. After this date many mathematicians have conducted research on Markov Matrix and has helped it to develop. Markov Chains Method is used intensively for research conducted on such social topics as the brand selection of customers, income distribution, immigration as a geographic structure, and the occupational mobility (for examples and references please see [8], [9], [10]). In marketing, Markov Chains Model is frequently used for topics such as "brand loyalty" and "brand switching dynamics". Although it is very complicated to transform marketing problems in to mathematical equations, Markov Chains Method comes out as the primary and most powerful technique in predicting the market share a product will achieve in the long term especially in an oligopolistic environment and in finding out the brand loyalty for a product.

The stochastic process is defined as a set of random variables $\{X_t\}$ where the unit time parameter t is taken from a given set T . All the special values the random variables take on are named as a state. Therefore, a state variable name is given to the X_t random variable. The set that accepts each X_t random variable is called an "example space" or a "state space". If the S state space includes whole number discontinuous values then it is called a stochastic process that is separate stated and these separate stated spaces may be countable and finite or countable and infinite. If X_t is defined in the $t \in (-\infty, \infty)$ interval it is classified as a stochastic process that is real valued. Being a special type of stochastic process, the Markov Chain,

$$P(X_{t+1} = x_t | X_0 = x_0, X_1 = x_1, \dots, X_t = x_t) = P(X_{t+1} = x_{t+1} | X_t = x_t); (t = 0, 1, 2, \dots)$$

is a chain that has Markovian property and the Markovian property stresses that given the present (or preceding) state, the conditional probability of the next state is independent of the preceding states. $P(X_{t+1} = x_{t+1} | X_t = x_t)$ are conditional probabilities and are named as transitional probabilities.

If the relationship

$$P(X_{t+1} = x_{t+1} | X_t = x_t) = P(X_1 = X_{t+1} | X_0 = x_t); (t = 0, 1, 2, \dots)$$

exists, the one step transitional probabilities are usually shown as P_{ij} and named as stationary and the transitional probabilities that have this property do not change in time and the relationship

$$P(X_{t+n} = x_{t+1} | X_t = x_t) = P(X_n = x_{t+1} | X_0 = x_t); (n = 0, 1, 2, \dots)$$

becomes valid. These conditional probabilities are named as n-step transitional probabilities and are shown as P_{ij}^n . P_{ij}^n explains that the process that is in the i state, will be in the j state n steps later. This is because P_{ij}^n are conditional probabilities and must be non-negative and also the relationship given below is valid.

$$\sum_{j=1}^m P_{ij}^{(n)} = 1 ; i = 1, 2, \dots, n = 0, 1, 2, \dots$$

At this point n-step transitional probabilities matrix, $S = \{S_0, S_1, \dots, S_m\}$ state space may be shown as

$$P^{(n)} = \begin{matrix} & S_0 & S_1 & \dots & S_m \\ \begin{matrix} S_0 \\ S_1 \\ \vdots \\ S_m \end{matrix} & \begin{pmatrix} P_{00}^{(n)} & P_{01}^{(n)} & \dots & P_{0m}^{(n)} \\ P_{10}^{(n)} & P_{11}^{(n)} & \dots & P_{1m}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ P_{m0}^{(n)} & P_{1m}^{(n)} & \dots & P_{mm}^{(n)} \end{pmatrix} \end{matrix}$$

If $n = 1$ is taken, then a stochastic process is a Markov Chain that has Markovian properties'. In this research, only the Markov Chains that are finite and have stationary transitional probabilities will be considered.

4. Chapman - Kolmogorov Equations

The $P_{ij}^{(n)}$ n-stepped transitional probabilities stress the probability of transition from the i state to the j state at the $n(> 1)$ step. The Chapman -Kolmogorov equations,

$$P_{ij}^{(n)} = \sum_{k=0}^m P_{ik} P_{kj}^{(n-1)} \quad \forall ij \text{ and } 0 \leq m \leq n$$

helps in forming a method for calculating the n-step transitional probabilities. In the special occasions where $m = 1$ and $m = n - 1$, the equations

$$P_{ij}^{(n)} = \sum_{k=0}^m P_{ik} P_{kj}^{(n-1)}; (\forall ij \text{ and } n) \text{ and } P_{ij}^{(n)} = \sum_{k=0}^m P_{ik}^{(n-1)} P_{kj}$$

are obtained. These equations stress the fact that the n-step transitional probabilities may be calculated from the one step transitional probabilities. For example, for $n = 2$

$$P_{ij}^{(2)} = \sum_{k=0}^m P_{ik} P_{kj}$$

is obtained and a P_{ij}^2 are the elements of the P^2 matrix. P^2 is obtained from the multiplication of P by P . Therefore, the n-step probabilities matrix may be calculated from the

$$P^n = P \cdot P^{(n-1)} = P^{(n-1)} \cdot P$$

relationship.

5. The long-term Behaviour of the Markov Chain

The ergodic chain (matrix) is defined as a chain where from one state it is possible to transform into all other states and where it contains no zero element that is at the powers of the P regular chain (matrix). Therefore it can be concluded that a regular matrix is ergodic but the opposite is not true. For the case where a T matrix is obtained by P having sufficiently big powers, if all of the line vectors of this T matrix are the same, it could be said that the P transitional matrix reaches a balance and there exists a balancing vector. A regular Markov Chain contains a single balance vector.

If $v = [v_1, v_2, \dots, v_m]$ is a probability vector, then the relationship $vp = v$ is valid and v is named as a balance vector.

6. Research Methodology

The purpose of this study was to examine the cell phone users loyalty using the Markov Chains Method. For this study data has been collected for cell phone users loyalty for 400 subscribers from ten out of twenty seven wards of Kinondoni District in Dar es salaam, Tanzania.

For the purpose of this study, seven states were considered which stand for the network operators represented in the form of a set as follows:

$$S = \{\text{tiGO, Zantel, Vodacom, Airtel, Sasatel, Benson on Line, TTCLMobile}\}$$

The data from the interviews was used to study cell phone users loyalty in two corresponding periods.

Following the two months interval, the switching behaviour of customers as well as their likelihood to exist in a given state was studied and the results were summarised in tabular form. The obtained results

Table 3: The transition results showing customers behaviour from July - August 2010

NETWORK	Vodacom	tiGO	Airtel	Zantel	TTCLMobile	Total
Vodacom	60	2	0	0	0	62
tiGO	3	288	2	0	0	293
Airtel	0	0	34	0	0	34
Zantel	0	0	0	9	0	9
TTCLMobile	0	0	0	0	2	2
Total	63	290	36	9	2	400

shown in Tables 3 up to 8 were used to verify the hypothesis that cell phone users loyalty can be analysed using Ergodic Transition Probability Matrices.

From Table 3, the sixth and seventh states, which stand for Benson on line and Sasatel companies, were assumed to be redundant states because the results show that there were no chances for customers to exist into those states. As a result a reduced table made up of five states was obtained.

In addition, the non communicating states (also known as dominant states) were abandoned in the search for an Ergodic transition matrix. Table 4 is the transition probability matrix obtained from Table 3 after removing the dominant states.

Table 4: The transition Probability Matrix from July to August, 2010

NETWORK	Vodacom	tiGO	Airtel	Total
Vodacom	0.9677	0.0323	0.0000	1.0000
tiGO	0.0102	0.9829	0.0068	1.0000
Airtel	0.0000	0.0000	1.0000	1.0000

The Transition probability matrix in Table 4 above is not Ergodic as it violates the condition of ergodicity [14], [15] hence it was not used in the analysis.

The transitions from September to October 2010 are summarized in Table 5.

Table 5: The transition results showing customers behaviour from September - October 2010

NETWORK	Vodacom	tiGO	Airtel	Total
Vodacom	58	1	1	60
tiGO	1	290	2	293
Airtel	0	0	36	36
Total	59	291	39	389

From Table 5 the following transition probability matrix was obtain.

The transition probability matrix in Table 6 above could not be used in the analysis of cell phone users loyalty due to the existence of the dominant state Airtel. Customers switching behavior for the period from November to December 2010 were summarized in Table 7.

From Table 7 the following transition probability matrix was obtained. Up to this point TTCLMobile was still acting as a dominant state so it was excluded in the construction of the following transition probability matrix:

The above table indicated that during this period cell phone users were loyal to tiGO, followed by Airtel, then Vodacom and lastly Zantel with probabilities 0.9732, 0.9697, 0.9310 and 0.8750 respectively.

Table 6: The transition Probability Matrix from September to October, 2010

NETWORK	Vodacom	tiGO	Airtel	Total
Vodacom	0.9667	0.0167	0.0167	1.0000
tiGO	0.0034	0.9898	0.0068	1.0000
Airtel	0.0303	0.0000	1.0000	1.0000

Table 7: The transition results showing customers behaviour from November - December 2010

NETWORK	Vodacom	tiGO	Airtel	Zantel	Total
Vodacom	54	3	1	0	58
tiGO	5	291	2	1	299
Airtel	1	0	32	0	33
Zantel	0	1	0	7	8
Total	60	295	35	8	398

Finally the switching behavior of cell phone users for four months of January to April 2011 were obtained as summarized in Table 9 below: This was because the switching behaviors of cell phone users for the months of January to February showed close similarity to that of March to April 2011. Also TTCLMobile was still a dominant state.

Table 9: The transition results showing customers behaviour from January - April 2011

NETWORK	Vodacom	tiGO	Airtel	Zantel	Total
Vodacom	46	12	2	0	60
tiGO	4	284	7	0	295
Airtel	1	6	28	0	35
Zantel	0	2	0	6	8
Total	51	304	37	6	398

The transition probability matrix constructed from Table 9 is as follows:

From Table 10 it was found that the leading company in terms of likelihood of having loyal customers was respectively tiGo, Airtel, Vodacom, and lastly Zantel with probabilities 0.9627, 0.8000, 0.7667 and 0.7500 respectively. The results in Table 10 shows only the existing situation of customer preferences, but the future stand or the long run forecast of the cell phone users distribution was analyzed by applying the Chapman-Kolmogorov equation.

7. Steady State Probability Vector

MATLAB software was used in conjunction with the Chapman-Kolmogorov equation to perform iterations on the transition probability matrix obtained from July, 2010 to April, 2011. The analysis resulted in a steady state probability matrix called the stability situation. This result was obtained after 44 iterations which is equivalent to 7 years and 4 months. This result verified the hypothesis that a stabilized Ergodic Transition Probability Matrix plays a significant role in determining the steady state probability vector shown in table 11 below:

From Table 11 the results show that if things continue as they are now in the long run the most preferred network company will be tiGO with the probability of 0.8297. The next preferred network will be Airtel with the probability of 0.1086. Then Vodacom will follow with probability of 0.0617. Finally Zantel will lose all its customers.

Table 8: The transition Probability Matrix from November to December, 2010

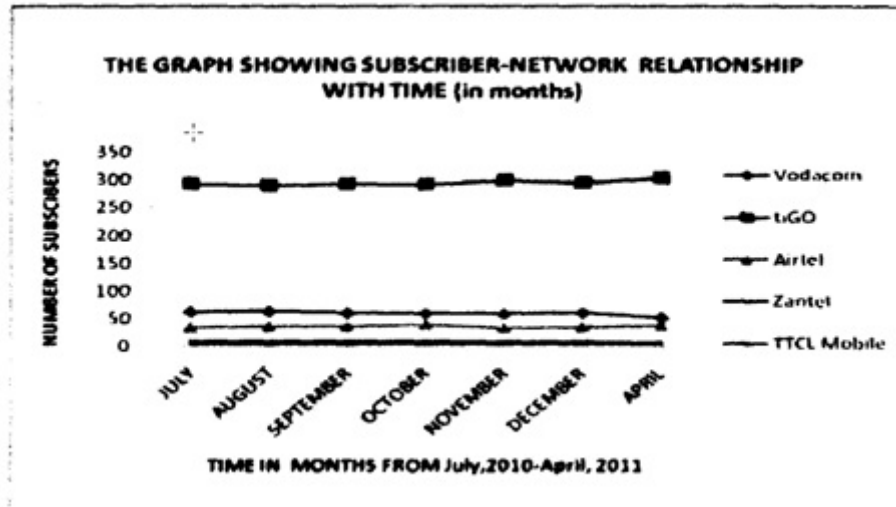
NETWORK	Vodacom	tiGO	Airtel	Zantel	Total
Vodacom	0.9310	0.0517	0.0172	0.0000	1.0000
tiGO	0.0167	0.9732	0.0067	0.0033	1.0000
Airtel	0.0303	0.0000	0.9697	0.0000	1.0000
Zantel	0.0000	0.1250	0.0000	0.8750	1.0000

Table 10: The transition Probability Matrix from December, 2010 to April 2011

NETWORK	Vodacom	tiGO	Airtel	Zantel	Total
Vodacom	0.7667	0.2000	0.0333	0.0000	1.0000
tiGO	0.0136	0.9627	0.0237	0.0000	1.0000
Airtel	0.0286	0.1714	0.8000	0.0000	1.0000
Zantel	0.0000	0.2500	0.0000	0.7500	1.0000

Finally the graphical analysis was done and the general network - customer status show that for the period from July 2010 to April 2011, the leading network was tiGO followed in descending order by Vodacom, Airtel, Zantel, and lastly TTCLMobile as seen in the following figure:

Figure 3: The data for the graph were collected from Kinondoni District.



8. Conclusion

According to the results discussed above, it was observed that Transition Probability Matrix for the periods from July to August 2010 and September to October 2010 shown in Table ?? and 4 were not Ergodic due to the existence of the absorbing state (i.e. the switching from Airtel to Airtel). On contrary, for the periods December 2010 to January 2011 the Transition probability Matrices were Ergodic and the analysis from these matrices showed that cell phone users were loyal to tiGo network followed by Airtel then Vodacom and finally Zantel.

The graphical analysis shown in Figure 1 reveals that the number of cell phone users loyal to tiGO is significantly larger than those loyal to Vodacom, Airtel Zantel and TTCLMobile. However in April 2011 customers loyal to Airtel company seemed to approach those loyal to Vodacom, indicating the existence of customer competition among the two networks.

Table 11: The steady state probability vector

NETWORK OPERATOR	PROBABILITY
tiGO	0.8297
Airtel	0.1086
Vodacom	0.0617
Zantel	0.0000

The steady state vector computed from the Ergodic transition probability matrix using the Chapman-Kolmogorov equation revealed that in the long run tiGo network company will have a large share with about 82.56 percent of cell phone users, followed by Airtel and Vodacom with 11.27 percent and 6.17 percent respectively. On the other hand, the predictions show that Zantel network is going to lose all its customers and may go out of business.

9. Recommendations

The results and conclusions above were based on a small sample of 400 cell phone users in Kinondoni District. The reasons for using such a small sample were lack of time and budgetary constraints. We recommend that a study using a large sample from a large area of Tanzania should be done to get more reliable results. In fact, instead of using a two months interval a six months interval could be used.

Further studies should be done to analyze customers loyalty to other companies in Tanzania such as insurance companies, banks etc. Another potential area where such analysis can play a significant role is investigating the long time voters loyalty to current political parties in Tanzania to see which will survive and which are likely to die out.

References

- [1] Dick, A. S. , Basu, K. (1994), Consumer Loyalty: toward an integrated conceptual framework, *Journal of Academy of Marketing Science*, Vol. 22, No. 2, p. 99-113.
- [2] Kahn, B. E., Kalwani, M. U.; Morrison, D. G. (1986), Measuring Variety Seeking and Reinforcement Behaviours Using Panel Data, *Journal of Marketing Research*, Vol. 23, May, p. 89-100.
- [3] Wilkie, W. L., (1994), *Consumer Behaviour*, John Wiley and Sons, New York.
- [4] Jacoby, J. And Chestnut, R. W.(1978), *Brand Loyalty: Measurement and Management*, John Wiley and Sons, New York.
- [5] Lyong Ha, Choong (1998), The Theory of Research Action Applied Brand Loyalty, *Journal of Product and Brand Management*, Vol. 7, Issue 1.
- [6] Assael, Henry (1992), *Consumer Behaviour and Marketing Action*, Fourth Edition, PWS-KENT Publishing Company.
- [7] Charnatony, Leslie de; McDonald, Malcolm H. B. (1992), *Creating Powerful Brands*, Butte worth-Heinemann Ltd., Linacre House, Jourdan Hill, Oxford.
- [8] Frydman, H. (1984), Maximum Likelihood Estimation in the Moverstayer Model, *Journal of the American Statistical Association*, 79, p. 632-638.

- [9] Geweke, J. Marshall, R., Zarkin, G. (1986), "Mobility Indices in Continuous Time Markov Chains, *Econometrica*, 54, p. 1407-1423.
- [10] Simper, B., Spilerman, S. (1976), Some Methodological Issues in the Analysis of Longitudinal Surveys, *Annals of Economic and Social Measurement*, 5, 447- 474.
- [11] Mtenzi, F. J., Chachage, B. L., Ngumbike F. "The Growth of Tanzania Mobile Phone Sector: Triumph of Quantity, Failure of Quality" Proceedings of M4D, Karlstad University Sweden, p. 55 - 61
- [12] Tanzania Communication Regulatory Authority (2008) Market Statistics: Telecommunication Available at <http://www.tcra.go.tz/market%20info/statsTelecom.html>
- [13] Hisali, E. (2007) Review of Sector Taxation Policies and Determining the Elasticity of Penetration and Price of the Various Telecommunication commissions Services in Uganda Available at <http://www.ucc.co.ug/>
- [14] Karlin S. and Taylor H. M. (1975) *A First Course in Stochastic Processes* 2nd Edition, Academic Press Inc., New York
- [15] Stone C. J., Hoel P. G. and Part S. C. (1972) *Introduction to Stochastic Processes* Houghton Miffling Company, Los Angeles.

Derivatives over Certain Finite Rings

by

Soud K. Mohamed

University of Dodoma, Tanzania

Abstract

We introduce a derivative of a relation over the ring of integers modulo an odd number which is base on the very fundamental concepts which helped in the evolution of derivative of a function over the real number field, namely slope. Then, for a prime field $GF(p)$, we use the derivatives to construct an algorithm that find all the directions, in the sense of [9], of graphs of certain exponential relations over R .

1. Introduction

Derivatives plays a very fundamental role in the analysis of functions over the real and the complex number fields. In these fields, their properties and applications are well-studied, since they reflect well on our every day lives. Over finite rings the notion of a derivative first appeared some 75 year ago in the paper [7] by H. Hasse. This derivative is the so called *Hasse derivative*, and has been successfully used in areas where finite fields play an important role, such as Coding Theory [8].

Suppose that R is a commutative ring and let $f(x) = \sum_{i=0}^n a_i x^i$ be a polynomial over R . Then the r th-Hasse derivative of $f(x)$ is $f^{[r]}(x) = \sum_{i=0}^n \binom{i}{r} a_i x^{i-r}$ with $\binom{i}{r} = 0$ for $i < r$. It is well-known that over a finite field K all functions are polynomial. In fact, if $|K| = m$, then there are m^m functions over K . In addition, there is a 1-1 correspondence between a function $f: K \rightarrow K$ and polynomial of degree less than m . So with Hasse derivatives one has every thing as far as a derivative of a function over K .

If R is a finite commutative ring, then only a fraction of functions on R are polynomials [6]. So for a function f on R which can not be represented by a polynomial over R , its Hasse derivative can not be determined. The aim of this paper is to introduce a derivative on a set of relations on certain rings.

Suppose that R is a finite ring and consider a relation ρ on R in a variable x , which shall be usually denoted by $\rho(x)$. Then the image of ρ may sometimes be an array [10]. For instance, the image of a function on R is an $s \times 1$ -array. In Section 2 we will look into exponential relations and their arrays over the ring $R = \mathbb{Z}_n$ for an integer n , and then we give sufficient conditions for an exponential relation to be a function over R .

Let R be the finite ring \mathbb{Z}_n , where n is an odd integer. Given a relation ρ , and a point $a \in R$, what should be the derivative of ρ at a ? In real analysis we take the slope of a tangent line at a , provided it exists. Moreover for a relation on the real number fields, we have at least one slope at a point: one along “each column” (picture the derivative of $\sqrt{(x)}$). Over a finite field this is not possible, because a point has more than one tangent! In addition, slopes of tangents can be computed along a column of ρ as well as across it. In Section 3 we show that the slope of the “closest” secant to a point $x \in R$ along a column k is the “best candidate” for the derivative of ρ at x , which shall be denoted by $\text{Dr}_k^{(1)}(\rho(x))$ or simply $\rho_k^{(1)}(x)$, the k -th derivative of ρ at x . This derivative has similar properties as the derivative of the real number field, like: linearity; product and quotient rules and how it acts on polynomials and exponential relations. The definition of the derivative requires some ordering of the elements of R . In Section 1, we consider the ring R as cyclically ordered set, which is very natural since R is a finite set.

Let $R = \mathbb{F}_q$ be a finite field with q elements and let $\rho : R \rightarrow R$ be a relation. Define the set of directions of ρ (slopes of secants of the graph of ρ) by:

$$D(\rho) := \left\{ \frac{\rho(a) - \rho(b)}{a - b} \mid a \neq b \in R \right\} \quad (1.1)$$

The problem of determining the bound on the size of $D(\rho)$ has been studied extensively both geometrically and combinatorially. The references [1], [3], [4] are some of the papers where this has been done. However, there has not been any attempt on computing the directions themselves, so far.

Given a graph of a column of a relation ρ over R , the size s of the graph is the number of elements in the domain of ρ . Note that s is less than or equal to q . Now, ideally if one wants to find all directions, one may have to compute up to $s(s-1)/2$ directions. The algorithm is as follows: you start at the first point and then find $s-1$ directions, then move to the second point and compute $s-2$ directions, and so on. Since $D(\rho)$ is a subset of R [3], there is a lot of unnecessary computations in this algorithm. In Section 4 we show that for some relations ρ over prime fields, the derivatives of ρ_k is all one needs to find all the directions of the graph of ρ .

2. Preliminaries

In this section we collect some of the preliminaries that will be needed in this paper. We fix the following notation: If R is a ring with unity, then R^* will denote the group of units of R . Unless otherwise specified, by order of an element $a \in R$ we mean the multiplicative order of a .

2.1 Immediate successor and predecessor

Most “people” are very familiar with linear ordering. However, cyclic order is not a household term. We give a formal definition of cyclic order and use it to define some terminologies.

Let X be a set of at least 3 elements. A *ternary relation* C is a subset of the Cartesian product $X \times X \times X$ which satisfies the following *axioms*:

1. *Cyclicity*: if $[a, b, c]$ is in C , then $[b, c, a]$ is in C
2. *Asymmetry*: if $[a, b, c]$ is in C , then $[c, b, a]$ is not in C
3. *Transitivity*: if $[a, b, c]$ and $[a, c, d]$ are in C , then $[a, b, d]$ is in C .
4. *Totality or Completeness*: if a, b and c are distinct, then either $[a, b, c]$ is in C or $[c, b, a]$ is in C .

If C satisfies the first three axioms, then it is called *partial cyclic ordering* on X , and consequently the pair (X, C) is a *partially cyclically ordered set*. If C satisfies all four axioms, it is called (*total*) *cyclic ordering* on X , as a result we get *cyclically ordered set* (X, C) .

If a cyclically ordered set X is finite of cardinality n , then there is a 1-1 correspondence between X and the cyclically ordered set $\{1, 2, \dots, n, 1\}$. We can use this correspondence to identify positions on X . Now, let X be a finite cyclically ordered set and let $x \in X$ be at position i (using the above correspondence), where i is an integer. Then the element in the position $i+1$ will be called an *immediate successor* of x , and will be denoted by x_+ , while that in the position $i-1$ will be called *immediate predecessor* of x and will be denoted by x_- .

2.2 Unity ordering

Let G be a finite group. The cyclic orderings on G which are of interest to us, are those that depend on generators of the group and the binary operation of the group. For example, for the additive group \mathbb{Z}_5 ,

we have the orderings $\{1, 2, 3, 4, 0\}$, $\{3, 1, 4, 2, 0\}$, $\{2, 4, 1, 3, 0\}$ and $\{4, 3, 2, 1, 0\}$, while for the multiplicative group \mathbb{Z}_5^* we have orderings $\{2, 4, 3, 1\}$ and $\{3, 4, 2, 1\}$. Given a generator g of a finite cyclic group, if the group is additive, then the ordering determined by g will be referred to *g-additive cyclic ordering*, where as if the group is multiplicative, then the ordering will be referred to as *g-multiplicative cyclic ordering*

Suppose that G is a finite cyclically ordered group of order $n \geq 3$ with a binary operation $*$. Then G has at least one cyclic ordering, namely the one determined by each generator of G . For $a, b \in G$, define the *length* from a to b , denoted by $l(a, b)$, to be $b * a^{-1} \in G$, where a^{-1} is the inverse of a . For example, in the cyclically ordered set $\mathbb{Z}_5 = \{0, 4, 3, 2, 1\}$, we have that $l(0, 3) = 3$, $l(2, 2) = 0$, while for the multiplicative group $\mathbb{Z}_5^* = \{1, 3, 4, 2\}$ modulo 5 which is cyclically ordered, we have $l(1, 3) = 3$ and $l(3, 2) = 4$.

Now, for our group G above, we have that every element $a \in G$ has an immediate predecessor and successor. Then the length $l(a, a_+)$ will be referred to as the *least length* of a , and will be denoted by $\delta(a)$. For example, for multiplicative cyclically ordered group $\mathbb{Z}_5^* = \{1, 2, 4, 3\}$, we have that $\delta(x) = 3$ for all $x \in \mathbb{Z}_5^*$. The following fact can be easily proved.

Fact 2.25 *Let R be a ring with unity 1_R and isomorphic to the ring \mathbb{Z}_n .*

- i. *For any additive ordering on R the least length $\delta(a) = \delta$ is constant for all $a \in R$.*
- ii. *There is an additive ordering on R such that $\delta(a) = \delta = 1_R$*

The cyclic ordering of Fact 2.25 (ii.) on a ring R will be called the *unity ordering*.

For the rest of the paper, we impose the following assumption on our ring R :

Assumption 1 *R is the ring \mathbb{Z}_q where q is an odd integer. Moreover, the ordering on R is the associated additive cyclic ordering.*

Remark 1 *Under the above assumption our ring R will have a canonical cyclic ordering, which will be fundamental throughout the paper. Moreover, no matter what additive cyclic ordering one takes on R , the element $x_+ - x_- \in R$ will be a unit, since the ordering is determined by a generator of R .*

3. Exponential and Hyperbolic Relations over R

In real analysis, exponential functions α^x are very fundamental, and they can easily be used to define other functions. Over finite ring, the mapping determined by α^x , for a unit α , is not necessarily a function. We have the following definition, which is motivated from [5].

Definition 3.26 *Let $\alpha \in R$ be a unit of order N . Then, the exponential relation $\rho : R \rightarrow R$ is the relation $\rho(x) = \alpha^x$. The image of ρ is an $N \times t$ -array $[\rho_i(x)]$ with the a -th row, $\rho(a) = \{\alpha^a, \alpha^{a+q}, \dots, \alpha^{a+(t-1)q}\}$, where $t \leq N$. The k -th column of ρ , $\rho_k(x) = \alpha^{x+kq}$, where $x = 0, 1, \dots, N-1$, will be called the k -th exponential relation of ρ .*

Example 3.27 *We consider examples:*

1. *Let $R = \mathbb{Z}_7$ and consider the relation $\rho(x) = 2^x \pmod{7}$. Then the array of ρ is*

$$[\rho_k(x)] = \begin{bmatrix} 2^0 & 2^7 & 2^{14} & \dots \\ 2^1 & 2^8 & 2^{15} & \dots \\ 2^2 & 2^9 & 2^{16} & \dots \end{bmatrix} = \begin{bmatrix} 2^0 & 2^1 & 2^2 \\ 2^1 & 2^2 & 2^3 \\ 2^2 & 2^3 & 2^1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 1 \\ 4 & 1 & 2 \end{bmatrix}$$

2. If $R = \mathbb{Z}_9$ and our relation is $\rho(x) = 2^x$, then the array of ρ is

$$[\rho_k(x)] = \begin{bmatrix} 2^0 & 2^9 & 2^{18} & \dots \\ 2^1 & 2^{10} & 2^{19} & \dots \\ 2^2 & 2^{11} & 2^{20} & \dots \\ 2^3 & 2^{12} & 2^{21} & \dots \\ 2^4 & 2^{13} & 2^{22} & \dots \\ 2^5 & 2^{14} & 2^{23} & \dots \end{bmatrix} = \begin{bmatrix} 1 & 8 \\ 2 & 7 \\ 4 & 5 \\ 8 & 1 \\ 7 & 2 \\ 5 & 4 \end{bmatrix}$$

Computation of the array of the relation in Example 3.27 (1.) looks easy, because along columns and rows one just increases the power by one. This is generally true for prime fields.

Lemma 3.28 *If R is a prime field, and $\rho(x) = \alpha^x$ is a relation on R , then the k -th relation of ρ is $\rho_k(x) = \alpha^{x+k}$.*

Proof

We have that $kq \bmod N = kq - k + k \bmod N = k(q-1) + k \bmod N = k \bmod N$, since $N|q-1$.
 \square Consider the relation $\rho(x) = \alpha^x$, where $\alpha \in R$ has order N . Then $\rho_i(x)$ has N rows. However, as shown in the example above the number of columns may vary. If certain conditions are satisfied, then the number of columns of the array of ρ can easily be obtained.

Theorem 3.29 *Let α be a unit in R of order N , and let $\rho(x) = \alpha^x$ be a relation.*

- (i) *If a perfect square is not a factor of q and $\gcd(N, q) = 1$, then $[\rho_i(x)]$ is an $N \times N$ -array.*
- (ii) *If $q = p^m$ for a prime p , where m is a positive integer, then there are $\gcd(N, p-1)$ columns in $[\rho_i(x)]$.*

Proof

Suppose that $[\rho_i(x)]$ is an $s \times t$ -array, and let the a -th row be $\rho(a) = \{\alpha^a, \alpha^{a+q}, \alpha^{a+2q}, \dots, \alpha^{a+(t-1)q}\}$. Then $\rho(a)$ is a coset of the subgroup $H = \langle \alpha^{iq} \rangle$ of R^* of order t . One can then observe that for both cases, $s = N$ and $t \leq N$.

(i) Now $\alpha^a = \alpha^{a+ tq}$ means that $tq = 0 \bmod N$ which implies that $N|tq$. But since N does not divide q , it must divide t . Hence $t = N$.

(ii) We have that $H \leq \langle \alpha \rangle$, since H contains powers of α . Let K be a subgroup of R^* of order $p-1$. Since $\gcd(p^{n-1}, p-1) = 1$, then $R^* \simeq \mathbb{Z}_{p^{n-1}} \oplus \mathbb{Z}_{p-1}$, and hence K is unique. From this we infer that if $\beta \in R^*$ is such that $\beta^{p-1} = 1$, then $\beta \in K$. Now we have that $(\alpha^{iq})^{p-1} = (\alpha^{i\phi(q)})^{ip} = 1$ for $i = 1, \dots, t$, so that $H \leq K$. Hence $t|N$ and $t|p-1$, which implies that $t \leq \gcd(N, p-1) = d$. If $d < t$, then $\alpha^{dq} = \alpha^{sNq+r(p-1)q} = 1$, so that $|H| < t$, a contradiction. \square The following corollary gives a sufficient condition for an exponential relation on R to be a function.

Corollary 3.30 *Suppose that $\alpha \in R$ is a unit of order N .*

- (i) *If a perfect square is not a factor of q and $N|q$, then the relation $\rho(x) = \alpha^x$ is a function.*
- (ii) *If $q = p^m$ for a prime p and α has order p^i for $i = 1, \dots, m-1$, then the relation $\rho(x) = \alpha^x$ is a function.*

Proof

(i) For all $a \in R$ and some positive integer t , we have that $\rho(a) = \{\alpha^a, \alpha^{a+q}, \dots, \alpha^{a+(t-1)q}\} = \{\alpha^a\}$, since $N|q$.

(ii) Follows from Theorem 3.29, since $\gcd(p^i, p-1) = 1$ for $i = 1, \dots, m-1$. \square

Let $\alpha \in R$ be a unit of order N , and consider the relation $\rho(x) = \alpha^x$. Then ρ is periodic of period N . More precisely, for a positive integer s each subset $S_j = \{jN, jN+1, \dots, (j+1)N-1\}$ of the domain of ρ , where $j = 0, 1, \dots, s-1$, determines the same image $\rho(S_j) = \{\rho(jN), \rho(jN+1), \dots, \rho((j+1)N-1)\}$. The subset S_j is referred to as the j -th steps of ρ . As expected, starting at a point in R , there are only a finite number of steps of ρ before one gets back to the same point.

Proposition 3.31 *Let $\alpha \in R$ be a unit of order N , and consider the relation $\rho(x) = \alpha^x$. If $\gcd(N, q) = d$, then ρ has q/d steps.*

Proof

Let ν be the least positive integer with the property that the set $T = \{0, N, 2N, \dots, (\nu-1)N\}$ taken mod q has distinct elements. Observe that find number of steps of ρ is the same as finding the cardinality ν of T . Also note that $\nu \leq q/d$. If $\nu > q/d$, then $|T| < \nu$, a contradiction. \square

Corollary 3.32 *Let $\alpha \in R$ be a unit of order N , and consider the relation $\rho(x) = \alpha^x$. Then the map $\pi = \rho|_{S_j} : S_j \rightarrow \text{im } \rho$ is a permutation.*

Proof

Only need to show that π is 1-1. Suppose that $\pi(jN) = \pi(jN+k)$ for $k \neq 0 \pmod N$. Then $\alpha^{jN} = \alpha^{jN+k}$ which is equivalent to $k = 0 \pmod N$, a contradiction. \square We now define hyperbolic relations over R .

Definition 3.33 *Let $\alpha \in R$ be a unit of order $N \geq 3$. Then*

1. *The k -th hyperbolic sine and cosine relations to the base α over R , denoted by $\sinh_{\alpha,k}(x)$ and $\cosh_{\alpha,k}(x)$ are respectively*

$$\cosh_{\alpha,k}(x) := \frac{\alpha^{x+kq} + \alpha^{-(x+kq)}}{x_+ - x_-}; \quad \sinh_{\alpha,k}(x) := \frac{\alpha^{x+kq} - \alpha^{-(x+kq)}}{x_+ - x_-}.$$

2. *The hyperbolic sine and cosine relations to the base α , denoted by $\cosh_{\alpha}(x)$ and $\sinh_{\alpha}(x)$ are the relations with images the $N \times t$ -arrays made by the columns $\cosh_{\alpha,k}(x)$ and $\sinh_{\alpha,k}(x)$ respectively, where $t < N$.*

Under certain assumption on R , hyperbolic relations on R behave like those over the real.

Proposition 3.34 *Suppose that R has the unity ordering, and let $\alpha \in R$ be a unit of order $N \geq 3$. Then for $k = 0, 1, \dots, N-1$:*

- (i) *the identity $\cosh_{\alpha,k}^2(x) - \sinh_{\alpha,k}^2(x) = 1_R$ holds,*
- (ii) *$\cosh_{\alpha,k}(-x) = \cosh_{\alpha,k}(x)$ and $\sinh_{\alpha,k}(-x) = -\sinh_{\alpha,k}(x)$*

Proof

Follows from the definition. □

4. Derivatives and their Properties

Denote the set of relations from R to R whose image are $s \times t$ -arrays by $\text{Rel}_{st}(R)$, and by $\text{Fun}(R)$ the set of all functions from R to R .

Let q be the order of the ring R and let $x \in R$. For a positive integer k and a relation $\rho \in \text{Rel}_{st}(R)$, define a map $\text{Dr}_k^{(1)} : \text{Rel}_{st}(R) \rightarrow \text{Rel}_{st}(R)$ on the k -th relation ρ_k of ρ by

$$\text{Dr}_k^{(1)}(\rho)(x) := \frac{\rho_k(x_+ + kq) - \rho_k(x_- + kq)}{x_+ - x_-}. \quad (4.1)$$

If the context is clear, then $\text{Dr}_k^{(1)}(\rho)(x)$ will be just denoted by $\rho_k^{(1)}(x)$.

If one looks closely at the this map one will notice that its value at a point $x \in R$ is the slope of the “closest” secant to the point x along ρ_k . It can be also interpreted as the “average” of the two “closest slopes” to x along ρ_k . The result below show that the transformation has good properties too.

Theorem 4.35 *Let $\rho, \gamma \in \text{Rel}_{st}(R), x \in R$ and let $c \in \mathbb{Z}$. Then*

(i) *the transformation $\text{Dr}_k^{(1)}$ is linear.*

(ii) *Product Rule:*

$$\begin{aligned} (\rho_k \gamma_k)^{(1)}(x) &= \rho_k(x_- + kq) \gamma_k^{(1)}(x) + \gamma_k(x_+ + kq) \rho_k^{(1)}(x) \\ &= \rho_k(x_+ + kq) \gamma_k^{(1)}(x) + \gamma_k(x_- + kq) \rho_k^{(1)}(x). \end{aligned}$$

(iii) *If $\gamma_k(x) \neq 0$ for all $x \in R$, then*

$$\left(\frac{\rho_k}{\gamma_k} \right)^{(1)}(x) = \frac{\gamma_k(x_- + kq) \rho_k^{(1)}(x) - \rho_k(x_- + kq) \gamma_k^{(1)}(x)}{\gamma_k(x_+ + kq) \gamma_k(x_- + kq)}.$$

Proof

(i) Follows easily.

(ii) We use the elementary “+ - tricks”.

$$\begin{aligned} (\rho_k \gamma_k)^{(1)}(x) &= \frac{(\rho_k \gamma_k)(x_+ + kq) - (\rho_k \gamma_k)(x_- + kq)}{x_+ - x_-} \\ &= \frac{\rho_k(x_+ + kq) \gamma_k(x_+ + kq) - \rho_k(x_- + kq) \gamma_k(x_- + kq)}{x_+ - x_-} \\ &= \rho_k(x_- + kq) \gamma_k^{(1)}(x) + \gamma_k(x_+ + kq) \rho_k^{(1)}(x) \end{aligned}$$

(iii) Exercise. □

Remark 2 *If R is a prime field, then (4.1) and its subsequent formulas in Theorem 4.35 become much easier, by the use of Lemma 3.28.*

Example 4.36 *Let us look into examples:*

- (i) Let p be an odd prime number and consider the ring \mathbb{Z}_p with the associated additive cyclic ordering. Let $f(x) = ax \pmod p$. Then for each $x \in \mathbb{Z}_p$, $\delta(x) = 1 \in \mathbb{Z}_p$, so that $x_+ - x_- = 2$ is a unit in \mathbb{Z}_p . So, $f^{(1)}(x) = \frac{a(x+1) - a(x-1)}{2} = a$. For $g(x) = bx^2 \pmod p$, we have that $g^{(1)}(x) = 2bx$.
- (ii) Let $R = \mathbb{Z}_9$ has the unity ordering, and consider the relation $\rho \in \text{Rel}(R)$ given by $\rho(x) = 2^x$. Then the image of ρ is 6×2 -array with columns $\rho_0(x) = \{2^0, 2^1, 2^2, 2^3, 2^4, 2^5\}$ and $\rho_1(x) = \{2^3, 2^4, 2^5, 2^0, 2^1, 2^2\}$. One can verify that $\rho_0^{(1)}(x) = \{3, 6, 3, 6, 3, 6\}$ and $\rho_1^{(1)}(x) = \{6, 3, 6, 3, 6, 3\}$ for $x \in R$.
- (iii) Consider the ring \mathbb{Z}_{35} , and let $R = 5\mathbb{Z}_{35} = \{0, 5, 10, 15, 20, 25, 30\}$. Then R is a ring modulo 35 whose unity $1_R = 15$. If one considers the given cyclic ordering on R , then for each $x \in R$, $\delta(x) = \delta = 5$ and $x_+ - x_- = 2\delta = 10$ which is a unit in R . Consider the relation $\rho(x) = 25^{5x} \pmod{35}$ on R . Then the image of ρ is a 3×3 -array, and we have that $\rho_0^{(1)}(x) = \{25, 15, 30\}$, $\rho_1^{(1)}(x) = \{15, 30, 25\}$ and $\rho_2^{(1)}(x) = \{30, 25, 15\}$ for $x \in R$.

The computation in Example 4.36 (i) and (ii) would not be very clear as far as $\rho_k^{(1)}(0)$ is concerned. But we used the following result.

Lemma 4.37 Suppose $\alpha \in R$ is a unit of order $N \geq 3$, and let $\rho(x) = \alpha^x$ be a relation on R . Then $\rho(k) = \rho(sN + j)$, for all integers j, s . In particular, $\rho(0) = \rho(N)$, $\rho(-1) = \rho(N - 1)$ and $\rho(N + 1) = \rho(1)$

Proof

We have that $\rho(j) = \alpha^j = \alpha^{sN+j} = \rho(sN + j)$. □

From the above theorem we see that the transformation $\rho_k^{(1)}$ looks indeed like a derivative on $\text{Rel}_{st}(R)$. For, when applied to a constant function it vanishes, and when applied to a polynomial of degree two it gives a polynomial of degree one, and so on. Also, when the transformation is applied to an exponential relation over R , it produces the relation times a constant. The above behavior are similar to those seen in the derivative transformation of real functions.

Definition 4.38 Let ρ be a relation on R and let $x \in R$. If $\rho_k^{(1)}(x)$ is defined, then it will be called the k -th derivative of the relation ρ at x , and will be denoted by $\rho_k^{(1)}(x)$.

In the real analysis case we are used to very nice derivative formulas for functions. The situation here is the similar, and we have the following result.

Corollary 4.39 Let R has unity 1_R and let δ be the least length. If $\alpha \in R$ is a unit and n is a nonnegative integer greater than 1, then

$$(i) \quad (\alpha^x)_k^{(1)}(x) = \frac{(\alpha^{2\delta} - 1_R)}{2\delta\alpha^\delta} \alpha^{x+kq}$$

$$(ii) \quad (x^n)^{(1)}(x) = \sum_{i=0}^s \binom{n}{2i+1} x^{n-(2i+1)} \delta^{2i}; \quad s = \begin{cases} \frac{n}{2} - 1 & n \text{ even} \\ \frac{n-1}{2} & n \text{ odd} \end{cases}$$

$$(iii) \quad (\sinh_{\alpha,h}(x))^{(1)}(x) = \frac{(\alpha^{2\delta} - 1_R)}{2\delta\alpha^\delta} \cosh_{\alpha,h}(x)$$

$$(iv) \quad (\cosh_{\alpha,h}(x))^{(1)}(x) = \frac{(\alpha^{2\delta} - 1_R)}{2\delta\alpha^\delta} \sinh_{\alpha,h}(x)$$

Proof

One just uses the definition of the derivative. □

Remark 3 1. If the ordering of R in Corollary 4.39 is the unity ordering, then $\delta = 1_R$ and the formulas become much simpler.

2. Theorem 4.35 and Corollary 4.39 can be used to find k -th derivatives of all relations which are linear combinations or products of the set of k -th relations $\{x^n, \alpha_k^x, \sinh_{\alpha,k}(x), \cosh_{\alpha,k}(x)\}$ for a unit $\alpha \in R$.

3. The derivative of a monomial function whose degree is divisible q is not zero in R .

The derivative is a linear transformation on $\text{Rel}_{st}(R)$. So it can be applied on a relation ρ more than once and still preserve the linearity property. The following result, which is a consequence of Theorem 4.35 and Corollary 4.39, gives formulas for computing k -th derivatives of certain relations ρ , t times, which will be called (t, k) -th derivative of ρ . If the array of ρ has only one column, then the $(t, 0)$ -th derivative will be just called t -th derivative.

Corollary 4.40 Let $\alpha \in R$ be a unit, and let δ is the least length. Then for a nonnegative integer t :

(i)

$$(x^n)^{(t)}(x) = \sum_{i=0}^s \binom{n}{2i+1} (x^{n-(2i+1)})^{(t-1)}(x) \delta^{2i}; \quad s = \begin{cases} \frac{n}{2} - 1 & n \text{ even} \\ \frac{n-1}{2} & n \text{ odd} \end{cases}$$

(ii) $(\alpha^x)_k^{(t)}(x) = \left(\frac{\alpha^{2\delta} - 1_R}{2\delta\alpha^\delta}\right)^t \alpha^{x+kq}$

Proof

(i) Follows from (4.1) and Theorem 4.35.

(ii) Exercise. □

Over the real the derivative transformation is not necessarily periodic for exponential functions. Over prime fields the derivative transformation on an exponential relation is periodic for most of the bases.

Theorem 4.41 Let $R = GF(q)$ for a prime number q , and let $\alpha \in R$ be a unit which is not a square root of unity. If $\rho(x) = \alpha^x$, then

$$\rho_h^{(t)}(x) = \rho_h(x) \text{ for some positive integer } t$$

Proof

By the assumption $\alpha^2 - 1 \neq 0 \in R$, so that $\frac{\alpha^2-1}{2\alpha} \in R^*$, which implies that $\left(\frac{\alpha^2-1}{2\alpha}\right)^t = 1 \in R^*$ some positive integer t . □

4.1 “The Exponential Values e ”

In real analysis the exponential relation $\exp(x) = e^x$ is characterized by its derivative, in the sense that it is the only relation with derivative equals the relation itself. Given a finite ring with additive cyclically ordering, do we have an analog of the exponential relation with respect the derivative we have defined? The answer is not necessarily! But some rings have indeed got a pair of e 's. The result below gives a sufficient condition for that to happen.

Theorem 4.42 Let $R = GF(q)$, where $q \geq 3$ is prime. If 2 is a quadratic residue in R , then the elements $e = 1 \pm \sqrt{2}$ has the property that $\rho_k^{(1)}(x) = \rho_k(x)$, where ρ_k is the k -th relation of the relation $\rho(x) = e^x$.

Proof

Consider the relation $\rho(x) = e^x$, where e is in R . Suppose that $\rho(x)$ has the property that $\rho_k^{(1)}(x) = \rho_k(x)$ for all k . Then one has that $e^2 - 2e - 1 = 0 \in R$, which means $e = 1 \pm \sqrt{2}$. \square

5. Finding Directions

Throughout this section $R = GF(q)$ for an odd prime q , and the ordering of R is the unity ordering. Recall that for a relation $\rho : R \rightarrow R$, the set of directions of ρ is denoted by $D(\rho)$ (see (1.1)). Let us denote the set of all (i, k) -th derivatives of ρ by $\text{Dr}_k^{(i)}(\rho)$ i.e.

$$\text{Dr}_k^{(i)}(\rho) := \{\rho_k^{(i)}(x) \mid x \in R\} \quad (5.1)$$

From the definition of derivative we see that $\text{Dr}_k^{(i)}(\rho) \subseteq D(\rho)$ for all i, k .

For a relation $\rho : R \rightarrow R$, the bound on the size of $D(\rho)$ has been well studied. But there are only few relations ρ over R whereby the exact size of $D(\rho)$ is known. So far these are the known ones: linear functions f [$|D(f)| = 1$] (see [9]); functions $f(x) = x^{(q+1)/2}$ [$|D(f)| = (q+3)/2$] (see [3]). We will try to add to this collection. We need the following lemma.

Lemma 5.43 *Let $\alpha \in R$ be a unit of order $N \geq 3$, and consider the relation $\rho(x) = \alpha^x$. Then for all $x \in R$, $\rho_k^{(i)}(x) \neq 0 \in R$ for all i, k .*

Proof

For $x \in R$ such that $jN + 1 \leq x \leq (j+1)N - 1$, we have that $\rho_k^{(i)}(x) \neq 0$ for all j, k , by Corollary 3.32. One can easily verify that $\rho_k^{(i)}(jN) \neq 0$, for all i, j, k . \square Suppose that $\alpha \in R$ is a unit, and let $\rho(x) = \alpha^x$ be a relation. Then by Theorem 4.41, applying the derivative transformation repeatedly gives back $\rho(x)$. We have two cases for $\beta = \frac{\alpha^2-1}{2\alpha}$:

Case 1: If β is in $\langle \alpha \rangle$, then we get a permutation of $\langle \alpha \rangle$ whose order is the order of subgroup generated by β .

Case 2: If β is not in $\langle \alpha \rangle$, then the collection $\{\text{Dr}_k^{(i)}(\rho)\}$ partitions a bigger subgroup of R^* containing $\langle \alpha \rangle$. If this happens, then we say that α partitions the subgroup.

We have the following result.

Lemma 5.44 *Let α be a unit in R , and let $\rho(x) = \alpha^x$ be a relation. Then the order of $\text{Dr}_k^{(i)}(\rho)$ divides $q-1$ for all i, k . In particular, if α is a generator of R^* , then $|\text{Dr}_k^{(i)}(\rho)| = q-1$ for all i, k .*

Proof

We know that the set $\text{Dr}_k^{(i)}(\rho)$ is a coset of $\langle \alpha \rangle$ in R^* for all i, k . Then the result follows, since all cosets of a subgroup have the same size. \square

Now we have our main result of this section, which establishes a connection between $D(\rho)$ and $\text{Dr}_k^{(i)}(\rho)$ for exponential relations $\rho(x)$.

Theorem 5.45 *Suppose that α is a unit in R of order $N \geq 3$, consider the relation $\rho(x) = \alpha^x$, and let s be the order of $\frac{\alpha^2-1}{2\alpha}$.*

(i) *If α partitions R^* , then for all k*

$$D(\rho) = \text{Dr}_k^{(1)}(\rho) \sqcup \text{Dr}_k^{(2)}(\rho) \sqcup \dots \sqcup \text{Dr}_k^{(s)}(\rho) \sqcup \{0\},$$

(ii) If α is a generator of R^* , then

$$D(\rho) = \text{Dr}_k^{(i)}(\rho) \sqcup \{0\} \quad \text{for all } i, k.$$

Proof

(i) Let $T = \text{Dr}_k^{(1)}(\rho) \sqcup \text{Dr}_k^{(2)}(\rho) \sqcup \cdots \sqcup \text{Dr}_k^{(s)}(\rho)$. Then $|T| = q - 1$, since α partitions R^* . We also have that 0 is not in $\text{Dr}_k^{(i)}(\rho)$ for all i, k , by Lemma 5.43. Since $\text{Dr}_k^{(i)}(\rho) \subseteq D(\rho)$ and $D(\rho) \leq q$ for all i , the result follows.

(ii) By Lemma 5.44 we have that $|\text{Dr}_k^{(i)}(\rho)| = q - 1$ and $0 \notin \text{Dr}_k^{(i)}(\rho)$ by Lemma 5.43, for all i, k . Since $\text{Dr}_k^{(i)}(\rho) \subset D(\rho)$ for all i, k , and $D(\rho) \leq q$, the result follows. \square

References

- [1] S. Ball (2003), *The number of directions determined by a function over finite field* J. Combin. Theory Ser. A, **104** 341-435.
- [2] S. Ball (2007), *Functions over finite fields that determine few directions*, Electr. Notes in Discrete Math., **29**, 185-188.
- [3] S. ball (2011) *Lacunary Polynomials over Finite Fields*, Preprint.
- [4] A. Blockhuis, S. Ball, A. E. Brouwer, L. Storme and T. Szonyi (1999) *On the number of slopes of the graph of a function defined on a finite field*, J. Combin. Theory Ser. A, **86** 187-196.
- [5] R. M. Campello de Souza, H. M. de Oliveira, A. N. Kauffman, A. J. A. Paschoal (1998), *Trigonometry in finite fields and a new Hartley transform*, ISIT 1998, Cambridge, MA, USA.
- [6] S. Frisch (1999), *Polynomial fuctions on finite commutative rings*, Lec. Notes in Pure and Appl. Maths 205, Dekker, 323-336.
- [7] H. Hasse (1936), *Theorie der hoheren Differentiale in einem algebraischen Funktionenkorper mit vollkommenen Konstantenkorper bei beliebiger charakteristik*, J. Reine Ang. Math., vol. 175, 50-54.
- [8] J. L. Massey, N. von Seeman, P. Schoeller (1986), *Hasse derivatives and repeated-root cyclic codes*, IEEE Symposium of Info. Theory, Ann. Arbor, USA.
- [9] L. Redei (1973), *Luchenhafte Polynome uber endkichen Korper*, Birkhauser-Verlag, Basel.
- [10] M.M.C. de Souza, H.M. de Oliveira, R.M.C. de Souza, M.M. Vasconcelos (2004), *The Discrete Cosine Transform over Prime Finite Fields*, LNCS, Vol. 3124/2004, 37-59

Bifurcation results on symplectic manifolds

by

Eleonora Ciriza

Università degli Studi di Roma Tor Vergata, Italy

1. Introduction

Existence and multiplicity of periodic trajectories of Hamiltonian vector fields on symplectic manifolds is a traditional field of research, which found new input from the work on Arnold's conjecture. Fitzpatrick, Pejsachowicz and Recht in [8],[9] studied bifurcation of periodic solutions of one-parameter families of (time dependent) periodic Hamiltonian systems in \mathbf{R}^{2n} relating the spectral flow to the bifurcation of critical points of strongly indefinite functionals.

In [6] we extended their results to families of time dependent Hamiltonian vector fields acting on symplectic manifolds and the related problems of bifurcation of fixed points of one parameter families of symplectomorphisms were discussed. Namely we proved that for a 1-parameter family of time dependent Hamiltonian vector fields, acting on a symplectic manifold M which possesses a known trivial branch u_λ of 1-periodic solutions if the relative Conley Zehnder index of the monodromy path along $u_\lambda(0)$ is defined and does not vanish then any neighborhood of the trivial branch contains 1-periodic solutions not in the branch.

Fixed points of Hamiltonian symplectomorphisms are in one to one correspondence with 1-periodic orbits of the corresponding vector field. Hence as a consequence we obtained, assuming that (M, ω) is a closed symplectic manifold with trivial first De Rham cohomology group, for a path $\phi : [0, 1] \rightarrow \text{Sym}p_0(M)$ of symplectomorphisms with a known smooth path $p : [0, 1] \rightarrow U$ of fixed points, *i.e.*, $p(\lambda)$ is a fixed point of ϕ_λ . If the Conley-Zehnder index $CZ(\phi, p)$ of ϕ along p is defined and does not vanish then there is a bifurcation of fixed points of ϕ from the trivial branch p .

The Arnold conjecture states that a generic Hamiltonian symplectomorphism has more fixed points that could be predicted from the fixed point index. More precisely, by the fixed point theory a diffeomorphism isotopic to the identity with non-degenerate fixed points must have at least as many fixed points as the Euler-Poincaré characteristic of the manifold. But the number of fixed points of a Hamiltonian symplectomorphism verifying the same non-degeneracy assumptions is bounded below by the sum of the Betti numbers. Roughly speaking, this can be explained by the presence of a variational structure in the problem. Fixed points viewed as periodic orbits of the corresponding vector field are critical points of the action functional either if the orbits are contractible or when the symplectic form is exact.

Applied to bifurcation of fixed points of one parameter families of Hamiltonian symplectomorphisms our result shows a similar influence on the presence of a variational structure. In order to see the analogy consider a one parameter family of diffeomorphisms $\psi_\lambda; \lambda \in [0, 1]$ of an oriented manifold M , assuming for simplicity that $\psi_\lambda(p) = p$ and that p is a non degenerate fixed point of $\psi_i; i = 0, 1$. The work of Ize [11] implies that the only homotopy invariant determining the bifurcation of fixed points in terms of the family of linearizations $L \equiv \{T_p\psi_\lambda\}$ at p is given by the parity

$$\pi(L) = \text{sign det}(T_p\psi_0) \cdot \text{sign det}(T_p\psi_1) \in \mathbb{Z}_2 = \{1, -1\}.$$

Here det is the determinant of an endomorphism of the oriented vector space T_pM . In other words bifurcation arise whenever the $\text{det}(T_p\psi_\lambda)$ change sign at the end points of the interval. Moreover, any family of diffeomorphisms close enough to ψ in the C^1 -topology and having p as fixed point undergoes bifurcation as well. On the contrary if both sign coincide one can find a perturbation as above with no bifurcation points at all. The integer valued Conley-Zehnder index provides a stronger bifurcation

invariant for one parameter families of Hamiltonian symplectomorphisms. It forces bifurcation of fixed points whenever the Conley-Zehnder index $\mathcal{CZ}(L)$ is non zero even when $\pi(L) = 1$. The relation between the two invariants is $\pi(L) = (-1)^{\mathcal{CZ}(L)}$.

A natural generalization of the classical Arnold's conjecture estimates the number of intersection points of two Lagrangian submanifolds of a symplectic manifold.

The cause that forces Hamiltonian deformation $L_1 = \phi(L)$ of a compact Lagrangian submanifold L of M to have a huge intersection with L can be explained as follows: by a well known theorem of Weinstein the submanifold L has a neighborhood symplectomorphic to a neighborhood of the zero section in the cotangent bundle $T^*(L)$. If L is simply connected and if L_1 is a Lagrangian submanifold that is C^1 close to L then L_1 is given by the image of the differential $dS : L \rightarrow T^*(L)$ of a smooth function $S : L \rightarrow \mathbb{R}^{2n}$ and therefore will have as many intersection points with L as critical points has the function S on L . The latter is bounded from below by Lusternik-Schnirelmann inequalities or by Morse inequalities if the critical points are non-degenerate. Of course L_1 need not be C^1 -close to L . But when $M = T^*(N)$ using an Hamiltonian isotopy ϕ_λ with $\phi_1 = \phi$ one can still produce a family of generating functions $S : N \times \mathbf{R}^k \rightarrow \mathbf{R}$ with k big enough such that critical points of S correspond to intersections of N with L_1 . This is a Theorem of Sikorav [18]. Using this theorem one can still get estimates on the number of intersection points but weaker than in the previous case. Functions S as before are usually called generating families.

In [7] we showed that intersections of one parameter families of Lagrangian submanifolds with a given one have stronger bifurcation properties than the intersections of general submanifolds of right codimension essentially for the same reason as above. For families L_λ close enough in the C^1 topology to a given Lagrangian submanifold L_0 bifurcation of intersection points of L_λ with L_0 reduces, by the above described process, to bifurcation of critical points of one parameter families of smooth functions. In this setting bifurcation arises whenever the spectral flow, or what is the same, the difference between the Morse indexes of the end points of the trivial branch is non-zero. This gives a stronger invariant than the usual bifurcation index obtained by comparing the sign of the determinant of the Jacobian matrix of the gradient at the end points of the trivial branch. Via generating functions we showed that the assumption of being C^1 close can be substituted with a more general one without modifying the conclusions.

Namely the main result in [7] is as follows. Let N be a closed manifold and let $L = \{L_\lambda\}$ be an exact, compactly supported family of Lagrangian submanifolds of the symplectic manifold $M = T^*(N)$ such that L_0 admits a generating family quadratic at infinity. Let $p : [0, 1] \rightarrow M$ be a path of intersection points of L_λ with N . Assume that L_λ is transversal to N at $p(\lambda)$ for $\lambda = 0, 1$ and that the Maslov intersection index $\mu(L, N; p)$ is different from zero. Then arbitrarily close to the branch p there are intersection points of L_λ with N such that do not belong to p .

The results exposed here were obtained in collaboration with J. Pejsachowicz. Symplectic features needed for our purpose are collected in section §2. In section §3 we extend the definitions of the Maslov and Conley-Zehnder indices to manifolds. This relies on the existence of symplectic trivializations of symplectic vector bundles over an interval. In §4 we outline how the bifurcation results of Fitzpatrick, Pejsachowicz and Recht are applied to the situations described above.

I would like to thank Ramadas Ramakrishnan from ICTP and the EAUMP's coordinators John Mango, Egbert Mujuni, Sylvester Rugeihyamu for inviting me to take part on the EAUMP project. I also wish to thank Velleda Baldoni for financial support.

2. Symplectic features

A symplectic manifold M is a differentiable manifold together with a closed nondegenerate differen-

tiable two form ω , *i.e.*,

$$d\omega = 0 \text{ and } \forall Y \neq 0 \exists X : \omega(X, Y) \neq 0, X, Y \in T_m M.$$

Hence M must have even dimension and because $\omega^n/n!$ gives the canonical volumen form it is oriented.

The non-degeneracy condition induces an isomorphism between the tangent $T(M)$ and the cotangent space of the manifold $T^*(M)$ that assigns to each vector field X a 1-form $\iota_X \omega = \omega(X, \cdot)$.

A diffeomorphism $\phi : (M, \omega) \rightarrow (M, \omega)$ that satisfies $\phi^* \omega = \omega$ is called symplectomorphism. In particular, a symplectomorphism preserves the volumen.

The requirement on the 2-form ω to be closed provides a correspondence between closed 1-forms and conservative vector fields since in this case $\mathcal{L}_X \omega = 0$ if and only if $d(\iota_X \omega) = 0$, such vector fields are called symplectic. The flow generated by a symplectic vector field consist of symplectomorphisms, *i.e.*, $\phi_t^* \omega = \omega \forall t$. A vector field is called Hamiltonian if the 1-form $\iota_X \omega$ is exact.

Because on a manifold there are many 1-forms the dimension of the group of symplectomorphisms of M , $Symp(M, \omega)$ is infinity. To the subset of exact 1-forms $\alpha = dH$ corresponds a normal subgroup $Ham(M, \omega)$ of $Symp(M, \omega)$.

In symplectic geometry there are no local invariants like for instance the curvature in Riemannian geometry. Darboux Theorem states that in some neighborhood of a given point one can choose a coordinate system $(U; x_1, \dots, x_n, y_1, \dots, y_n)$ such that the restriction of the form to the neighborhood U is $\omega|_U = \omega_0 := \sum_{i=1}^n dx_i \wedge dy_i$. Hence the universal local model of a symplectic form is the standard symplectic form ω_0 in \mathbb{R}^{2n} . In this case the isomorphism between the tangent and cotangent space is given explicitly by $X = \partial/\partial x_j \rightarrow \iota_X \omega_0 = dy_j$, $X = \partial/\partial y_j \rightarrow \iota_X \omega_0 = -dx_j$.

An important example of symplectic manifolds is the cotangent bundle of any manifold. Let N be an n -dimensional differentiable manifold. Let $T^*(N)$ be the cotangent bundle of N and $\pi : T^*N \rightarrow N$ the projection on N . There is a canonical 1-form λ_N on T^*N defined as follows: let ξ be a tangent vector to T^*N at the point $p \in T^*N$ ($\xi \in T_p(T^*(N))$). Since the element p is a cotangent vector on $T_x(N)$ where $x = \pi(p)$ and $\pi_*(\xi) \in T_x(N)$ define $\lambda_N(\xi) := p(\pi_*(\xi))$. In local coordinates $\lambda_N(\xi) = pdq$ and the symplectic 2-form is $\Omega = d\lambda_N$. Being exact it is closed and it is non-degenerate because in local coordinates $\Omega = dp \wedge dq$.

Let W be a vector subspace of a symplectic vector space (V, ω) , the symplectic orthogonal to W is the vector subspace $W^\omega := \{v \in V/\omega(v, w) = 0 \forall w, w \in W\}$. W is said to be isotropic if $W \subset W^\omega$. It is said to be coisotropic if $W \supset W^\omega$. If it is both isotropic and coisotropic it is called Lagrangian.

If W is isotropic, then W^ω is coisotropic and the symplectic form ω induces a symplectic form ϖ on the quotient space W^ω/W defined by $\varpi(v + W, w + W) = \omega(v, w) \forall v, w \in W^\omega$. The space $(W^\omega/W, \varpi)$ is called the *isotropic reduction*. Moreover if L is a Lagrangian subspace of (V, ω) then $L_W = (L \cap W^\omega)/(L \cap W)$ is a Lagrangian subspace of W^ω/W .

Lagrangian submanifolds of a symplectic manifold (M, ω) are the submanifolds of maximal dimension where the symplectic form vanishes. They are characterized by $TL = (TL)^\omega$. Examples of Lagrangian submanifolds are the vertical fibers of a cotangent bundle T^*N . As for submanifolds transverse to the fibers, any such submanifold is locally the graph of a 1-form $\alpha : N \rightarrow T^*N$. The graph of a 1-form α is Lagrangian if and only if α is closed. If the 1-form is exact, *i.e.*, if $\alpha = dS$ the function S is called a generating function for the corresponding submanifold.

Any Lagrangian submanifold can be generated locally by a function on the product of N with a parameter space, in which case it is called generating family.

The definition goes as follows (see [21]). Let V be a finite dimensional vector space. Consider a smooth function $S : N \times V \rightarrow \mathbb{R}$ such that the differential dS is transversal to the submanifold

$$N^0 = T^*(N) \times V \times \{0\} \text{ of } T^*(N \times V) \cong T^*(N) \times V \times V$$

Denote by S_n the function $S_n : V \rightarrow \mathbb{R}$ defined by $S_n(v) = S(n, v)$ and by S_v the function $S_v : N \rightarrow \mathbb{R}$ defined by $S_v(n) = S(n, v)$. By the implicit function theorem, the set $C = \{(n, v) / dS_n(v) = 0\}$ of vertical critical points of S is a submanifold of $N \times V$ of the same dimension as N .

Let $e : C \rightarrow T^*(N)$ defined by $e(n, v) = dS_v(n)$. The map e is a Lagrangian immersion (but generally not an embedding) of the manifold C into $T^*(N)$. Given a Lagrangian submanifold L of $T^*(N)$, S is said to be a *generating family* for L if there is a diffeomorphism h from C onto L such that $e = ih$, where $i : L \rightarrow T^*(N)$ denotes the inclusion. The generating family S is said to be *quadratic at infinity* if there is a non-degenerate quadratic form Q on V such that $S(n, v) = Q(v)$ for $\|v\|$ big enough.

Diffeomorphisms of a manifold may be identified with their graphs, that is, with submanifolds of $M \times M$ which are mapped diffeomorphically onto M by the projections π_1, π_2 . If M carries a symplectic structure ω , the form $\pi_1^*\omega - \pi_2^*\omega$ defines a symplectic structure on the product manifold $M \times M$. A diffeomorphism ϕ of a symplectic manifold (M, ω) is a symplectomorphism if and only if its graph is a Lagrangian submanifold of $(M \times M, \pi_1^*\omega - \pi_2^*\omega)$. Fixed points of ϕ correspond to intersections of the graph with the diagonal Δ of $M \times M$.

On a closed symplectic manifold (M^{2n}, ω) every smooth time dependent (Hamiltonian) function $H : \mathbb{R} \times M \rightarrow \mathbb{R}$ gives rise to a family of time dependent Hamiltonian vector fields $X : \mathbb{R} \times M \rightarrow TM$ defined by

$$\omega(X(t, x), \xi) = d_x H(t, x)\xi$$

for $\xi \in T_x M$. If H is periodic in time with period 1, then so is X . By compactness and periodicity the solutions $u(t)$ of the initial value problem for the Hamiltonian differential equation

$$\begin{cases} \frac{d}{dt}u(t) = X(t, u(t)), \\ u(s) = x \end{cases} \quad (2.1)$$

are defined for all times t . The flow (or evolution map) associated to X is the two-parameter family of symplectomorphisms $\psi : \mathbb{R}^2 \rightarrow \text{Symplect}(M)$ defined by

$$\psi_{s,t}(x) = u(t)$$

where u is the unique solution of (2.1).

By the uniqueness and smooth dependence on initial value theorems for solutions of differential equations the map $\psi : \mathbb{R}^2 \times M \rightarrow M$ is smooth. The diffeomorphisms $\psi_{s,t}$ verify the usual cocycle property of an evolution operator *i.e.*, $\psi_{s,r} \circ \psi_{r,t} = \psi_{s,t}$ and $\psi_{t,t} = \text{Id}$. From this property it follows that for each fixed s , the map sending u into $u(s)$ is a bijection between the set of 1-periodic solutions of the time dependent vector field X and the set of all fixed points of $\psi_{s,s+1}$. Hence in order to find periodic trajectories of (2.1) we can restrict our attention to the fixed points of $P = \psi_{0,1}$. The map $P = \psi_{0,1}$ is called the period or Poincaré map of X .

A 1-periodic trajectory is called non degenerate if $p = u(0)$ is a non degenerate fixed point of P , *i.e.*, if the monodromy operator $S_p \equiv T_p P : T_p M \rightarrow T_p M$ has no 1 as eigenvalue. Consistently, the eigenvalues of the monodromy operator will be called Floquet multipliers of the periodic trajectory. The particular choice of $s = 0$ is irrelevant to the property of being non degenerate since the Floquet multipliers do not depend on this choice. (see [1])

Every symplectomorphism that can be represented as a time 1-map of such a time dependent Hamiltonian flow is called a Hamiltonian map. If M is simply connected the connected component of the identity map $Symp_0(M, \omega)$ in the space of symplectic diffeomorphisms $Symp(M, \omega)$ consists of Hamiltonian maps (see [12]).

3. The Maslov index and the Conley-Zehnder index

Before going to the manifold setting let us discuss the case of $\mathbb{R}^{2n} = T^*\mathbb{R}^n$ with the standard symplectic form $\omega_0 = \sum dx_i \wedge dy_i$. The group of real $2n \times 2n$ symplectic matrices will be denoted by $\mathbf{Sp}(2n, \mathbb{R})$.

The relative Conley-Zehnder index is a homotopy invariant associated to any path $\psi: [0, 1] \rightarrow \mathbf{Sp}(2n, \mathbb{R})$ of symplectic matrices with no eigenvectors corresponding to the eigenvalue 1 at the end points. This invariant counts algebraically the number of parameters t in the open interval $(0, 1)$ for which $\psi(t)$ has 1 as an eigenvalue. One of the possible constructions uses the Maslov index for non-closed paths. We shall define it along the lines of Arnold [3] for closed paths. For an alternative construction see Robbin and Salamon [15].

The Lagrangian Grassmannian $\Lambda(n)$ consists of all Lagrangian subspaces of \mathbb{R}^{2n} considered as a topological space with the topology it inherits as a subspace of the ordinary Grassmannian of n -planes. Let J be the selfadjoint endomorphism representing the form ω_0 with respect to the standard scalar product in \mathbb{R}^{2n} . Namely, $\omega_0(u, v) \equiv \langle Ju, v \rangle$. Then J is a complex structure, it is indeed the standard one. It coincides with multiplication by i under the isomorphism sending $(x, y) \in \mathbb{R}^{2n}$ into $x + iy$ in \mathbb{C}^n . In terms of this representation, a Lagrangian subspace is characterized by $JL = L^\perp$.

Using the above description one can identify $\Lambda(n)$ with the homogeneous space $U(n)/O(n)$. This can be done as follows: given any orthonormal basis of a Lagrangian subspace L there exist a unique unitary endomorphism $A \in U(n)$ sending the canonical basis of $L_0 = \mathbb{R}^n \times \{0\}$ into the given one and in particular sending L_0 into L . Moreover the isotropy group of L_0 can be easily identified with $O(n)$. Hence we obtain a diffeomorphism between $U(n)/O(n)$ and $\Lambda(n)$ sending the class $[A]$ into $A(L_0)$. Since the determinant of an element in $O(n)$ is ± 1 , the map sending A into the square of the determinant of A factorizes through $\Lambda(n) \equiv U(n)/O(n)$ and hence induces a one form $\Theta \in \Omega^1(\Lambda(n))$ given by $\Theta = [\det^2]^*\theta$, where $\theta \in \Omega^1(S^1)$ is the standard angular form on the unit circle. This form is called the Keller-Maslov-Arnold form.

The Maslov index of a closed path γ in $\Lambda(n)$ is the integer defined by $\mu(\gamma) = \int_\gamma \Theta$. In other words $\mu(\gamma)$ is the winding number of the closed curve $t \rightarrow \det^2(\gamma(t))$. The Maslov index induces an isomorphism between $\pi_1(\Lambda)$ and \mathbb{Z} .

The construction can be extended to non-closed paths as follows: fix $L \in \Lambda(n)$. If L' is any Lagrangian subspace transverse to L then L' can be identified with the graph of a symmetric transformation from JL into itself. It follows from this that the set Λ_L of all Lagrangian subspaces L' transverse to L is an affine space diffeomorphic to the space of all symmetric forms on \mathbb{R}^n and hence it is contractible.

We shall say that a path in $\Lambda(n)$ is admissible with respect to L if the end points of the path are transverse to L . The Maslov index $\mu(\gamma, L)$ of an admissible path γ with respect to L is defined as follows: take any path δ in Λ_L joining the end points of γ and define

$$\mu(\gamma; L) \equiv \mu(\gamma') = \int_{\gamma'} \Theta.$$

where γ' is the path γ followed by δ . The result is independent of the choice of δ . Moreover, since Λ_L is contractible, $\mu(\gamma; L)$ is invariant under homotopies keeping the end points in Λ_L .

Geometrically, the Maslov index $\mu(\gamma; L)$ can be interpreted as an intersection index of the path γ with the one codimensional analytic set $\Sigma_L = \Lambda(n) - \Lambda_L$ (see [16]). From the definition it follows that the index

is additive under concatenation of paths. Namely, given two admissible paths α and β with $\alpha(1) = \beta(0)$

$$\mu(\alpha \star \beta; L) = \mu(\alpha; L) + \mu(\beta; L).$$

Since $\mathbf{Sp}(2n, \mathbb{R})$ is connected it follows from the homotopy invariance that

$$\mu(S\gamma; SL) = \mu(\gamma; L)$$

for any symplectic isomorphism S . This allows to extend the notion of Maslov Index to paths of Lagrangian subspaces in $\Lambda(V)$, where (V, ω) is any finite dimensional symplectic vector space.

Graphs of symplectic endomorphisms are Lagrangian subspaces of the symplectic vector space $V \times V$ endowed with the symplectic form $\omega \times (-\omega)$. The graph of $P \in \mathbf{Sp}(2n, \mathbb{R})$ is transversal to the diagonal $\Delta \subset V \times V$ if and only if 1 is not an eigenvalue of P . A path $\phi: [0, 1] \rightarrow \mathbf{Sp}(2n, \mathbb{R})$ will be called admissible if 1 is not in the spectrum of its end points. For such a path the *relative Conley-Zehnder index* is defined by

$$\mathcal{CZ}(\phi) = \mu(\text{Graph}\phi, \Delta). \quad (3.1)$$

From the above discussion it follows that $\mathcal{CZ}(\phi)$ is invariant under admissible homotopies and it is additive with respect to concatenation of paths. If the fixed point space of $\phi(\lambda)$ reduces to $\{0\}$ for all λ then $\mathcal{CZ}(\phi) = 0$.

There is one more property of the Conley-Zehnder index that we use in the sequel. Namely, that for any $\alpha: [0, 1] \rightarrow \mathbf{Sp}(2n, \mathbb{R})$ and any admissible path ϕ

$$\mathcal{CZ}(\alpha^{-1}\phi\alpha) = \mathcal{CZ}(\phi). \quad (3.2)$$

This can be seen as follows. Since the spectrum is invariant by conjugation, the homotopy $(t, s) \rightarrow \alpha^{-1}(s)\phi(t)\alpha(s)$ shows that $\mathcal{CZ}(\alpha^{-1}\phi\alpha) = \mathcal{CZ}(\alpha^{-1}(0)\phi\alpha(0))$. Now (3.2) follows by the same argument applied to any path joining $\alpha(0)$ to the identity.

The property (3.2) allows to associate a Conley-Zehnder index to any admissible symplectic automorphism of a symplectic vector-bundle over an interval. Let I be the interval $[0, 1]$, then any symplectic bundle $\pi: E \rightarrow I$ over I has a symplectic trivialization. If $S: E \rightarrow E$ is a symplectic endomorphism of E over I well behaved at the end points, then we can define the Conley-Zehnder index of S as follows: if $T: E \rightarrow I \times \mathbb{R}^{2n}$ is any symplectic trivialization, then $TST^{-1}(\lambda, v)$ has the form $(\lambda, \phi_T(\lambda)v)$ where ϕ_T is an admissible path on $\mathbf{Sp}(2n, \mathbb{R})$. Any change of trivialization induces a change on ϕ_T that has the form of the left hand side in (3.2) and hence $\mathcal{CZ}(\phi_T)$ is independent of the choice of trivialization. Thus the Conley-Zehnder index of S is defined to be $\mathcal{CZ}(S) \equiv \mathcal{CZ}(\phi_T)$.

Now let's define the relative Conley-Zehnder index of a path of symplectomorphisms along a path of fixed points: let M be a closed symplectic manifold and let $Symp(M)$ be the group of all symplectomorphisms endowed with the C^1 topology. Let $\phi: I \rightarrow Symp(M)$ be a smooth path of symplectomorphisms of M . Let $p: I \rightarrow M$ be a path in M such that $p(\lambda)$ is a fixed point of $\phi(\lambda)$. Floquet multipliers of $\phi(\lambda)$ at $p(\lambda)$ are by definition the eigenvalues of $S_\lambda = T_{p(\lambda)}\phi(\lambda): T_{p(\lambda)}(M) \rightarrow T_{p(\lambda)}(M)$. A fixed point will be called non degenerate if none of its Floquet multipliers is one. Consistently, we will call the pair (ϕ, p) admissible whenever $p(i)$ is a non degenerate fixed point of $\phi(i)$ for $i = 0, 1$.

Let $E = p^*[T(M)]$ be the pullback by p of the tangent bundle of M (we use the same notation for the bundle and its total space). The family of tangent maps $S_\lambda = T_{p(\lambda)}\phi(\lambda)$ induces a symplectic automorphism $S: E \rightarrow E$ over I . Define the *relative Conley-Zehnder index of ϕ along p* by

$$\mathcal{CZ}(\phi; p) \equiv \mathcal{CZ}(S). \quad (3.3)$$

From the properties discussed above it follows that the relative Conley-Zehnder index $\mathcal{CZ}(\phi; p)$ of ϕ along p is invariant by smooth pairs of homotopies $(\phi(s, t), p(s, t))$ such that $\phi(s, t)(p(s, t)) = p(s, t)$ and such that for $i = 0, 1$; $p(s, i)$ is a non degenerate fixed point of $\phi(s, i)$.

The index is additive under concatenation. It follows from (3.2) that it has another interesting property, which for simplicity we state in the case of a constant path $p(t) \equiv p$. If $\phi, \psi: I \rightarrow \text{Sym}p(M)$ are two admissible paths in the isotropy subgroup of p then

$$\mathcal{CZ}(\psi \circ \phi, p) = \mathcal{CZ}(\phi \circ \psi, p).$$

In other words \mathcal{CZ} is a trace.

Finally let us define the Maslov intersection index of two families of Lagrangian submanifolds L_λ and N_λ of a symplectic manifold M along a given path $p: I \rightarrow M$ of intersection points.

Since the interval I is contractible, the pullback $p^*(TM)$ by p of the tangent bundle of M is a trivial bundle whose fiber over λ is the tangent space $T_{p(\lambda)}M$. Taking any trivialization $T: p^*(TM) \rightarrow I \times \mathbb{R}^{2n}$ of this bundle the images under the trivialization maps $T_\lambda: T_{p(\lambda)}M \rightarrow \mathbb{R}^{2n}$ of the tangent spaces $T_p L_\lambda$ and $T_p N_\lambda$ determine two paths $l(\lambda)$ and $n(\lambda)$ in the space Λ_n of all Lagrangian subspaces of \mathbb{R}^{2n} . Assuming that the paths l, n have transverse intersection at the end points, the path $l \times n$ has endpoints transversal to the diagonal Δ in $(\mathbb{R}^{2n} \times \mathbb{R}^{2n})$. Since the space of Lagrangian subspaces transversal to a given one is contractible, if we take any path δ joining the endpoints of $l \times n$ the Maslov index of a close path made by $l \times n$ followed by δ , is independent of the choice of δ . The index of this closed path is by definition the relative Maslov index $\mu(l, n)$ (cf. [15]). This index is an integer which counts with appropriate multiplicities the points in $(0, 1)$ where $l(\lambda) \cap n(\lambda) \neq \{0\}$. From the invariance of the Maslov index under the action of the symplectic group it follows that $\mu(l, n)$ is independent of the choice of trivialization. We call it (once more!) the *Maslov intersection index* of the family $L = \{L_\lambda\}$ with $N = \{N_\lambda\}$ along p , and we denote it by $\mu(L, N, p)$.

The last crucial property that we need to mention is the invariance of the Maslov index under isotropic reduction. Consider a Lagrangian subspace $L \subset (V, \omega)$ and a path of Lagrangian subspaces $l: [0, 1] \rightarrow \Lambda(V)$ such that the endpoints $l(0)$ and $l(1)$ are transverse to L . If W is an isotropic subspace such that $W \subset L$ which has transverse intersection with $l(t)$ for all $t \in [0, 1]$ then following the lines of Viterbo (cf. Proposition 3 of [20]) it can be proved that the path $l_W: [0, 1] \rightarrow \Lambda(W^\omega/W)$ defined by $l_W(t) := l(t)/W$ is continuous and that the Maslov index of the path l_W relative to the Lagrangian subspace $L_W := L/W$ of W^ω/W coincide with the Maslov index of the path l relative to the Lagrangian subspace L , that is,

$$\mu_{L_W}(l_W) = \mu_L(l).$$

4. Bifurcations

a) FROM PERIODIC ORBITS OF 1-PARAMETER FAMILIES OF TIME DEPENDENT HAMILTONIAN SYSTEMS

Bifurcation theory deals with the problem of existence of nontrivial solutions arbitrary closed to a known family of solutions. For this purpose one takes into consideration a smooth one parameter family of time dependent Hamiltonian functions $H: I \times \mathbb{R} \times M \rightarrow \mathbb{R}$, where $I = [0, 1]$ is the parameter set and each $H_\lambda: \mathbb{R} \times M \rightarrow \mathbb{R}$ is one periodic in time. Let $X \equiv \{X_\lambda\}_{\lambda \in [0, 1]}$ be the corresponding one parameter family of Hamiltonian vector fields. Then the flows $\psi_{\lambda, s, t}$ associated to each X_λ depend smoothly on the parameter $\lambda \in I$. Suppose also that the 1-parameter family of Hamiltonian vector fields X_λ possesses a known smooth family of 1-periodic solutions u_λ ; $u_\lambda(t) = u_\lambda(t + 1)$. Solutions u_λ in this family are called *trivial* and we seek for sufficient conditions in order to find nontrivial solutions arbitrarily close to the given family. Identifying \mathbb{R}/\mathbb{Z} with the circle S^1 we regard the family of trivial solutions either as a path $\tau: I \rightarrow C^1(S^1; M)$ defined by $\tau(\lambda) = u_\lambda$ or as a smooth map $u: I \times S^1 \rightarrow M$.

A point $\lambda_* \in I$ is called a *bifurcation point* of periodic solutions from the trivial branch u_λ if every neighborhood of $(\lambda_*, u_{\lambda_*})$ in $I \times C^1(S^1; M)$ contains pairs of the form (λ, v_λ) where v_λ is a nontrivial periodic trajectory of X_λ .

A necessary condition for a point λ_* to be of bifurcation is that 1 is a Floquet multiplier of u_{λ_*} . This condition is not sufficient (See for example [2] Proposition 26.1). Thus non degenerate orbits cannot be bifurcation points of the branch. In what follows we will assume that $u(0)$ and $u(1)$ are non degenerate and we will seek for bifurcation points in the open interval $(0, 1)$.

Consider the path $p: I \rightarrow M$ given by $p(\lambda) = u_\lambda(0)$. Each $p(\lambda)$ is a fixed point of the symplectomorphism $P_\lambda = \psi_{\lambda,0,1}$. Under our hypothesis, the pair (P, p) is admissible. The number $\mathcal{CZ}(P, p)$ constructed in the previous section will be called the relative Conley-Zehnder index of $X \equiv \{X_\lambda\}_{\lambda \in [0,1]}$ along the trivial family u . We denote it by $\mathcal{CZ}(X, u)$. If this index is not zero one has the following

THEOREM A: *Let $X \equiv \{X_\lambda\}_{\lambda \in [0,1]}$ be a one parameter family of 1-periodic Hamiltonian vector fields on a closed symplectic manifold (M, ω) . Assume that the family X_λ possesses a known, trivial, branch u_λ of 1-periodic solutions such that $u(0)$ and $u(1)$ are non degenerate. If the relative Conley-Zehnder index $\mathcal{CZ}(X, u) \neq 0$ then the interval I contains at least one bifurcation point for periodic solutions from the trivial branch u .*

For the proof (see [6]) we followed an idea of Salamon and Zehnder [17] (Lemma 9.2.) in the nonparametric case. It consist in using appropriate symplectic trivializations and applying Moser's Method [14] to construct local Darboux coordinates $(V, \psi_{\lambda,t})$ on the manifold M adapted to the λ -parameter family $u_\lambda(t)$ of periodic solutions of the Hamiltonian differential equation

$$\begin{cases} \frac{d}{dt}u_\lambda(t) = X_\lambda(t, u_\lambda(t)), \\ u_\lambda(s) = x \end{cases} \quad (4.1)$$

i.e., we showed the existence of an open neighborhood V of 0 in \mathbb{R}^{2n} and of a family of symplectomorphisms $\psi_{\lambda,t}: V \rightarrow M$ that satisfies $\psi_{\lambda,t}(0) = u_\lambda(t)$ and $\psi_{\lambda,t}^*\omega = \omega_0$ on V . The new coordinates allowed us to reduce our problem to the Fitzpatrick, Pejsachowicz and Recht's bifurcation theorem in [9].

b) FROM INTERSECTION POINTS OF 1-PARAMETER FAMILIES OF LAGRANGIAN SUBMANIFOLDS

Let $T^*(N)$ be the cotangent bundle of a closed manifold N endowed with the standard symplectic structure. We will consider bifurcations of intersections of $N \equiv 0_N$ identified with the zero section of the bundle $T^*(N)$ with an exact one-parameter family of Lagrangian submanifolds $L = \{L_\lambda\}_{\lambda \in [0,1]}$ such that L_λ coincides with L_0 outside of a compact subset of $T^*(N)$. More precisely we consider families $L_\lambda = i_\lambda(L_0)$ where $i_\lambda: L_0 \rightarrow T^*(N)$ is a smooth family of Lagrangian embeddings with $i_\lambda \equiv i_0$ outside of a compact subset of L_0 . Such a family is said to be *compactly supported*. Moreover L is called *exact* if the one-form $i^*\omega(\frac{\partial}{\partial \lambda}, -)$ is exact on $[0, 1] \times L_0$. The natural topology in the space of all Lagrangian submanifolds of a given manifold is discussed in [22]. Remark that a family i_λ as above induces a continuous path in the space $C^\infty(L_0, T^*(N))$ with respect to the fine C^1 topology. Therefore L_r is C^1 close to L_s whenever r is close enough to s .

Let $I = [0, 1]$ and let $p: I \rightarrow T^*(N)$ be a smooth path such that $p(\lambda) \in L_\lambda \cap N$. A point $p(\lambda_*) \in L_{\lambda_*} \cap N$ is called *bifurcation point* from the given path p of intersection points if any neighborhood of $(\lambda_*, p(\lambda_*))$ in $[0, 1] \times T^*(N)$ contains points (λ, q) with $q \in L_\lambda \cap N$, $q \neq p(\lambda)$.

It follows from the implicit function theorem that a necessary condition for $p(\lambda_*)$ to be a bifurcation point of intersection is that the manifold L_{λ_*} fails to be transversal to N at $p(\lambda_*)$. This means that for $p_* = p(\lambda_*)$ one has that $T_{p_*}L_{\lambda_*} + T_{p_*}N$ is a proper subset of the tangent space $T_{p_*}(T^*(N))$. Since

$\dim T_{p_*}L_{\lambda_*} = \dim T_{p_*}N = \frac{1}{2}\dim T^*(N)$ this turns out to be equivalent to

$$T_{p_*}L_{\lambda_*} \cap T_{p_*}N \neq \{0\}.$$

This condition is not sufficient. Assuming that the manifolds L_0, L_1 are transverse to N , under some extra assumption the nonvanishing of $\mu(L, N, p)$ provides a sufficient condition for the existence of at least one bifurcation point.

THEOREM B: *Let N be a closed manifold and let $L = \{L_\lambda\}$ be an exact, compactly supported family of Lagrangian submanifolds of $T^*(N)$ such that L_0 admits a generating family quadratic at infinity. Let $p : [0, 1] \rightarrow T^*(N)$ be a path of intersection points of L_λ with N . Assume L_λ is transverse to N at $p(\lambda)$ for $\lambda = 0, 1$ and that the Maslov intersection index $\mu(L, N, p) \neq 0$, then there exist a $\lambda_* \in (0, 1)$ such that $p(\lambda_*)$ is a point of bifurcation for intersection points of L_λ with N from the trivial branch p .*

If $L_0 = N$ then the first assumption of the theorem holds by taking $S = 0$.

The basic idea of the proof of Theorem B is to convert our problem to that of finding bifurcations of critical points of one parameter families of functionals. We used a result of Sikorav which guarantees the existence of generating families for deformations of Lagrangian submanifolds under Hamiltonian isotopies (see proposition 1.2 and Remark 1.7 in [18]). More precisely, if ϕ_λ is a Hamiltonian isotopy of $T^*(N)$ and if $L_0 \subset T^*(N)$ is generated by a family quadratic at infinity then there exists a smooth family of functions $S_\lambda : N \times \mathbb{R}^k \rightarrow \mathbb{R}$ quadratic at infinity such that $\phi_\lambda(L_0)$ is generated by the family S_λ . On the other hand Chaperon [4] [5] proved that any one parameter exact compactly supported family of Lagrangian embeddings $L_\lambda = i_\lambda(L_0)$ can be extended to a Hamiltonian isotopy of the ambient manifold. Putting both results together we have that for any smooth family L_λ of Lagrangian submanifolds of $T^*(N)$ there exists a smooth family

$$S : [0, 1] \times N \times \mathbb{R}^k \rightarrow \mathbb{R}$$

quadratic at infinity such that S_λ generates L_λ , where $S_\lambda(n, v) = S(\lambda, n, v)$.

Thus each $L_\lambda = e_\lambda(C_\lambda)$ where $C_\lambda = \{(u, v)/v \text{ is critical of } S_{\lambda, n}\}$, the functions $S_{\lambda, n} : \mathbb{R}^k \rightarrow \mathbb{R}$ and $S_{\lambda, v} : N \rightarrow \mathbb{R}$ are given by $S_{\lambda, n}(v) = S_\lambda(n, v)$ and $S_{\lambda, v}(n) = S_\lambda(n, v)$ and $e_\lambda : C_\lambda \rightarrow T^*(N)$ is defined by $e_\lambda(n, v) = dS_{\lambda, v}(n)$.

Since here each e_λ is an embedding it induces a bijection between critical points of $S_\lambda : N \times \mathbb{R}^k \rightarrow \mathbb{R}$ and intersection points in $L_\lambda \cap N$. Therefore the path of intersection points p has a corresponding path $\tau : I \rightarrow N \times \mathbb{R}^k$ of critical points of S_λ . Because L_0, L_1 are transversal to N at $p(0), p(1)$ it follows that $\tau(0)$ and $\tau(1)$ are non-degenerate critical points. This is a direct consequence of the linear algebra of symplectic reductions. Indeed, let $N' = N \times \mathbb{R}^k$ and consider the symplectic manifold $T^*(N') = T^*(N) \times \mathbb{R}^{2k}$. The manifold $\{0\} \times \mathbb{R}^k$ is an isotropic submanifold of $T^*(N')$ and $T^*(N)$ is the symplectic reduction of $T^*(N')$ modulo the isotropic submanifold $\{0\} \times \mathbb{R}^k$. On the other hand N' and dS_λ are lagrangian submanifolds of $T^*(N')$ whose symplectic reductions are N and L_λ respectively. Since L_λ intersects transversally N at $p(\lambda)$, for $\lambda = 0, 1$, then dS_λ intersects transversally N' . But this is equivalent to the non-degeneracy of the critical point $\tau(\lambda)$ for $\lambda = 0, 1$.

At any critical point the Hessian $H(S_\lambda, \tau(\lambda))$ of S_λ at $\tau(\lambda)$ is a well defined symmetric bilinear form. The Morse index $m(S, x)$ of S at a nondegenerate critical point is the dimension of the negative eigenspace of $H(S, x)$. From Morse theory the inequality $m(S_1, \tau(1)) \neq m(S_0, \tau(0))$ guarantees the existence of bifurcation critical points [13].

Since L_λ is the image of $dS_{\lambda, v} : N \rightarrow T^*(N)$, identifying N with the zero section we have that L_λ is transversal to N for $\lambda = 0, 1$ and by the localization properties of the relative Maslov index (Theorem 2.3 in [16]) it equals the difference of the Morse indices at the endpoints of the path, that is,

$$\mu(dS, N', \tau) = m(S_1, \tau(1)) - m(S_0, \tau(0)).$$

But the Maslov index is invariant under isotropic reduction thus

$$\mu(dS, N', \tau) = \mu(L, N, p).$$

Hence the hypothesis of Theorem B implies that it is possible to find a sequence of critical points of S_λ bifurcating from the trivial branch. Via e_λ those critical points correspond to nontrivial intersections of L_λ with N

c) FROM FIXED POINTS OF A ONE PARAMETER FAMILY OF SYMPLECTOMORPHISMS

We discuss now bifurcations of a path of fixed points of a one parameter family of symplectomorphisms. Consider a closed symplectic manifold (M, ω) . We assume here that the first Betti number $\beta_1(M)$ of M vanishes, since in this case any symplectic diffeomorphism belonging to the connected component of the identity $Symp_0(M)$ of the group of all symplectic diffeomorphisms can be realized as the time one map of a 1-periodic Hamiltonian vector field. The following result can be obtained as a consequence either of Theorem A or of Theorem B.

COROLLARY: *Assume that $\beta_1(M) = 0$. Let ϕ_λ be a path in $Symp_0(M)$ such that $\phi_\lambda(p) = p$ for all λ and such that p is a non degenerate fixed point of ϕ_0 and ϕ_1 . Then if $\mathcal{CZ}(\phi, p) \neq 0$, there exist a $\lambda_* \in (0, 1)$ such that any neighborhood of (λ_*, p) in $I \times M$ contains a point (λ, q) such that q is a fixed point of ϕ_λ different from p (i.e. λ_* is a bifurcation point for fixed points of ϕ_λ from the trivial branch p).*

Moreover the same is true for any close enough path in the C^1 -topology lying in the isotropy group of p .

To each symplectomorphism ϕ_λ there corresponds a time dependent family of vector fields X_λ , and to each of this it corresponds a time dependent family of Hamiltonian function H_λ . In [6] we proved that there exist a family of time dependent hamiltonian functions $H' : I \times I \times V \rightarrow \mathbb{R}$ that depends smoothly on the parameter λ such that ϕ_λ is the time-one map of the corresponding time-dependent Hamiltonian vector field $X'_\lambda : I \times M \rightarrow TM$. Then because of the one to one correspondence between 1-periodic orbits of the Hamiltonian vector field with fixed points of the period map we can apply Theorem A.

Let us discuss now the relationship with intersection points of lagrangian submanifolds. Consider $M \times M$ with the symplectic form $\pi_1^*\omega - \pi_2^*\omega$. Given a path of symplectomorphisms ϕ_λ and a path of fixed points $p(\lambda)$ of ϕ_λ having non-degenerate end points (i.e., such that $T_{p(\lambda)}\phi_\lambda$ is nonsingular for $\lambda = 0, 1$), the path of fixed points corresponds to a path of intersection points of the graph of ϕ_λ with the diagonal Δ and the Maslov intersection index $\mu(\text{Graph } \phi, \Delta, p \times p)$ along the intersection path is well defined and coincides with the relative Conley-Zehnder index of ϕ along p .

By Weinstein's theorem [22] any Lagrangian submanifold of a symplectic manifold has a neighborhood that is symplectomorphic to a neighborhood of the zero section of its own cotangent bundle. We apply Weinstein's theorem to the diagonal Δ in $M \times M$ and then modify the Hamiltonian and the flow $\phi_{\lambda,t}$ outside of a neighborhood of p in such a way that the new flow equals the identity outside of a compact neighborhood of p . There the graph of ϕ_λ coincide with Δ and thus it can be viewed as a one parameter family of Lagrangian submanifolds of $T^*\Delta$ with compact support.

Since $H^1(M, \mathbb{R}) = 0$ we get that $L \equiv L_\lambda$ is exact. Moreover by Sikorav's theorem L_0 possesses a generating family being ϕ_0 isotopic by a Hamiltonian isotopy to the identity map of T_*N . Hence we can apply Theorem B to the family L and Δ .

We close this section with a formula that allows to compute the individual contribution of a regular point in the trivial branch to the Conley-Zehnder index and give an example where bifurcation cannot be detected using the parity.

Assume that λ_0 is an isolated point in the set

$$\Sigma = \{\lambda/p(\lambda) \text{ is a degenerate fixed point of } \phi(\lambda)\}.$$

Define $\mathcal{CZ}_{\lambda_0}(\phi) \equiv \lim_{\epsilon \rightarrow 0} \mathcal{CZ}(\phi; p|_{[-\epsilon, \epsilon]})$. The point λ_0 is called regular (cf. [16]) if the quadratic form Q_{λ_0} on the eigenspace $E_1(S_{\lambda_0}) = \text{Ker}(S_{\lambda_0} - Id)$ corresponding to the eigenvalue 1 defined by $Q_{\lambda_0}(v) = \omega(\dot{S}_{\lambda_0}v, v)$ is nondegenerate.

Here $S_\lambda = T_{p(\lambda)}\phi(\lambda)$ as before and \dot{S}_{λ_0} denotes the intrinsic derivative of the vector bundle endomorphism S (See [10] chap 1 sect 5). If t_0 is a regular point then it is an isolated point in Σ and

$$\mathcal{CZ}_{\lambda_0}(\phi) = -\sigma(Q_{\lambda_0}) \quad (4.2)$$

where σ denotes the signature of a quadratic form. This formula follows from the definition of the intrinsic derivative and formula (2.8) in [9].

EXAMPLE: Let M be the symplectic manifold $S^2 = \mathbb{C} \cup \{\infty\}$. Consider the closed path of symplectic maps $\phi_\theta : S^2 \rightarrow S^2; \theta \in [0, 1]$ defined by

$$\phi_\theta(z) = \begin{cases} e^{i2\pi(\theta-1/2)} \cdot z & \text{if } z \in \mathbb{C}, \\ \infty & \text{if } z = \infty \end{cases}$$

ϕ_θ is a rotation of angle $\theta - 1/2$ so it leaves fixed only the points $z = 0$ and $z = \infty$ except for $\theta = 1/2$, in which case the fixed point set is the sphere S^2 . For each θ the tangent map $T_0\phi_\theta$ of ϕ_θ at the fixed point $z = 0$ equals ϕ_θ . The only value of θ for which 1 is an eigenvalue of the tangent map $T_0\phi_\theta$ is $\theta = 1/2$ for which the corresponding eigenspace is \mathbb{C} . Moreover 0 is a regular degenerate fixed point of $\phi_{1/2}$. The relative Conley-Zehnder index $\mathcal{CZ}_0(\phi; 0)$ of the symplectic isotopy ϕ along the constant path of fixed points $p = 0$ coincides with the signature of the quadratic form $Q_{1/2} = \omega(\dot{\phi}_{1/2}, -)$ that is non degenerate on the eigenspace $E_1(\phi_{1/2})$. Then since

$$\dot{\phi}(1/2) = i2\pi Id$$

it follows from (4.2) that

$$\mathcal{CZ}_0(\phi; 0) = -\sigma[v \rightarrow \omega(\dot{\phi}(1/2)v, v)] = \sigma[v \rightarrow 2\pi \langle v, v \rangle] = 2.$$

Therefore any closed path of symplectomorphisms on the sphere keeping 0 fixed and homotopic to ϕ has nontrivial fixed points close to zero.

References

- [1] R. Abraham and J. Marsden. *Foundations of mechanics*. Second edition. Addison-Wesley, Reading, (1978).
- [2] H. Amann. *O.D.E. an introduction to Nonlinear Analysis*. De Gruyter Studies in Math. **13**, (1990)
- [3] V. I. Arnold. *Mathematical methods in classical mechanics*. Springer-Verlag. (1978).
- [4] M. Chaperon. Questions de géométrie symplectique. *Astérisque* **105-106** (1983), 231-249.
- [5] M. Chaperon. On generating families. *The Floer Memorial Volume*. Progress in Mathematics **133**, Birkhäuser, (1995).
- [6] E. Ciriza. Bifurcation of periodic orbits of time dependent Hamiltonian systems on symplectic manifolds. *Rend. Sem. Mat. Univ. Pol. Torino* Vol 57, 3 (1999), 161-176 (2002)

- [7] E. Ciriza and J. Pejsachowicz. A Bifurcation Theorem for Lagrangian Intersections. *Progress in non linear Differential Equations and their Applications*, Vol. 40. 2000 Birkhuaser Verlag
- [8] M. Fitzpatrick, J. Pejsachowicz and L. Recht. Spectral flow and bifurcation of critical points of strongly-indefinite functionals, Part I: General Theory. *J. Funct. Anal.* **162**, (1999), 52-95.
- [9] M. Fitzpatrick, J. Pejsachowicz and L. Recht. Spectral flow and bifurcation of critical points of strongly-indefinite functionals, Part II: Bifurcation of Periodic Orbits of Hamiltonian Systems. *Jour. of ODE.* **163**, (2000) n.1, 18-40.
- [10] V. Guillemin and S. Sternberg. *Geometric asymptotics*. Mathematical Surveys **14** Amer. Math. Soc. (1977).
- [11] J. Ize. Necessary and sufficient conditions for multiparameter bifurcation. *Rocky Mountain J. of Math.* **18** (1988), 305-337.
- [12] D. McDuff and D. Salamon. *Introduction to symplectic topology*. Oxford Math. Monographs, (1995).
- [13] J. Milnor. Morse Theory. *Princeton Univ. Press, Princeton.* (1963).
- [14] J. Moser. On the volume elements of a manifold. *Trans. AMS* **120**, (1965), 286-294.
- [15] J. Robbin and D. Salamon. The Maslov index for paths. *Topology* **32** (1993), 827-844.
- [16] J. Robbin and D. Salamon. The spectral flow and the Maslov index. *Bull. London Math. Soc.* **27** (1993), 1-33.
- [17] D. Salamon and E. Zehnder. Morse theory for periodic solutions of Hamiltonian system and the Maslov index. *Comm. Pure Appl. Math.* **XLV**, (1992), 1303-1360.
- [18] J. C. Sikorav. Problèmes d'intersections et de points fixed en géométrie hamiltonienne. *Comm. Math. Helv.* **62** (1987), 61-72.
- [19] J. C. Sikorav. Sur les immersions lagrangiennes dans une fibré cotangent. *C. R. Acad: Sci. Paris Ser.* **32** (1986), 119-122.
- [20] C. Viterbo. Intersection de sous variétés Lagrangiennes, fonctionnelles d'action et indice des systèmes Hamiltoniens *Bull. Soc. Math. France* **115** (1987), 361-390.
- [21] A. Weinstein. Lectures on symplectic manifolds *AMS Reg. Conf. Ser. Math.* **29** (1977).
- [22] A. Weinstein. Lagrangian submanifolds and Hamiltonian systems. *Ann. of Math.* **98** (1998), 377-410.

Three Systems of Orthogonal Polynomials and Associated Operators

by

John Musonda

Uppsala University, Sweden

Abstract

In this report, three systems of polynomials, that are orthogonal systems for three different but related inner product spaces, are presented. Three basic operators that are related to the systems are described, and boundedness of two other operators on a few Hilbert spaces is proven.

1. Introduction

More than a decade ago, Professor Sten Kaijser happened to discover two remarkable systems of orthogonal polynomials. The most interesting of the systems was in fact not a standard system, but it had some other useful properties. These discoveries led to a dissertation by Tsehaye K. Araaya [4, 5]. In March this year, Professor Lars Holst [7] presented a new way to calculate the Euler sum, $\sum \frac{1}{n^2} = \frac{\pi^2}{6}$. His calculations inspired Professor Kaijser to calculate a third system of polynomials, a system that turned out to fill a gap related to the previous systems. In this report, we present these three systems of orthogonal polynomials, and discuss some operators related to them.

The weight function that is used in one of the first two systems is the function $\omega_1(x) = 1/(2 \cosh \frac{\pi}{2}x)$, while for the third system we use the self convolution of this function, that is, $\omega_2 = \omega_1 * \omega_1$. The function, ω_1 , has three interesting properties that make it useful as a weight function. The first is that it is the density function of a probability measure, and the second is that it is up to a dilation its own Fourier transform, that is, it is the Fourier transform of the function $1/\cosh t$. The third is that it is closely related to the Poisson kernel for a strip of width two. The second property makes it possible to interpret its moments as values at zero of successive derivatives, while the third can be used for direct computations of many integrals.

This report is organised as follows: In section (), we present preliminaries needed to study and understand the work in the subsequent sections. This section has four subsections. In the first, some of the notation used throughout the report is explained. The second reviews those aspects of the theory of Hilbert spaces which are particularly relevant to our study, while the third reviews different aspects of the theory of orthogonal polynomials of one real variable. In the fourth subsection, we introduce the spaces that are of interest to our study. Our first system which we call the σ -system is presented in section (), while our second system which we call the τ -system is presented in section (). As aforementioned, these two systems were studied in Araaya papers [4, 5], and here we just take an overview of the results so that this report can be self contained. Also in section (), we introduce three operators R , J and Q , which are related to the systems. The third system which we call the ρ -system is presented in section (), and we study this system in detail since it is a new addition filling a gap related to the previous systems. This system of orthogonal polynomials is obtained by applying the Gram-Schmidt procedure to the sequence $\{x^n\}_{n=0}^{\infty}$ on the real line with the ω_2 -weighted L^2 inner product. It turns out that the system has a simple recurrence formula, so that the exponential generating function is easily computed. Using this the orthogonality is proven. In section () we discuss some useful connections between the systems, in terms of the operators. Finally in section (), we present two operators, $T = R^{-1}$ and $S = JR^{-1}$, where J and R are the operators introduced in section (). Boundedness of these two operators on five Hilbert spaces (defined in subsection ()) is proven.

2. Preliminaries

2.1 Some Notations

We use the Kronecker's delta: $\delta_{nm} = 0$ or 1 , according as $n \neq m$, or $n = m$. The symbol \mathbb{F} is used to denote the field of either real numbers \mathbb{R} or complex numbers \mathbb{C} . By $\operatorname{Re}(z)$, $\operatorname{Im}(z)$, $|z|$ and \bar{z} , we mean the real part, the imaginary part, the absolute and the conjugate complex value, respectively, of a complex number z . Closed intervals are denoted by $[a, b]$, open intervals by (a, b) and half-open intervals by $(a, b]$ or $[a, b)$.

We use \mathbb{S} to denote the strip $\{z \in \mathbb{C} : -1 \leq \operatorname{Im}(z) \leq 1\}$, $\partial\mathbb{S}$ for the boundary of the strip \mathbb{S} and \mathcal{P} for the Poisson kernel for the strip \mathbb{S} .

More notation will be introduced as we go on.

2.2 Elementary Theory of Hilbert Spaces

In this subsection, we review those aspects of the theory of separable Hilbert spaces which are particularly relevant to our study.

Definition 2.46 A normed linear space is a pair $(V, \|\cdot\|)$ where V is a vector space over \mathbb{F} , and $\|\cdot\|$ is a function $\|\cdot\| : V \rightarrow \mathbb{R}$ called a norm on V that satisfies the following conditions for all $x, y \in V$ and $\alpha \in \mathbb{F}$:

1. $\|x\| \geq 0$ and $\|x\| = 0$ if and only if $x = 0$.
2. $\|\alpha x\| = |\alpha| \|x\|$.
3. $\|x + y\| \leq \|x\| + \|y\|$.

Definition 2.47 A bounded linear operator from a normed linear space $(V_1, \|\cdot\|_1)$ to a normed linear space $(V_2, \|\cdot\|_2)$ is a function L from V_1 to V_2 that satisfies the following for all $x, y \in V_1$ and $\alpha, \beta \in \mathbb{F}$:

1. $L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$.
2. For some $M \geq 0$, $\|Lx\|_2 \leq M \|x\|_1$.

The smallest such M is called the norm of L , written $\|L\|$. Thus,

$$\|L\| = \sup_{\|x\|_1 \leq 1} \|Lx\|_2.$$

If in the second condition equality holds with $M = 1$, then the operator L is called an isometry and the normed linear spaces $(V_1, \|\cdot\|_1)$ and $(V_2, \|\cdot\|_2)$ are said to be isometric. Isometric normed linear spaces can be regarded as the same as far as their normed linear space properties are concerned.

Definition 2.48 An inner product space is a pair $(V, \langle \cdot, \cdot \rangle)$ where V is a vector space over \mathbb{F} , and $\langle \cdot, \cdot \rangle$ is a function $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{F}$ called an inner product on V that satisfies the following four conditions for all $x, y, z \in V$ and $\alpha \in \mathbb{F}$:

1. $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ if and only if $x = 0$.
2. $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$.

3. $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$.
4. $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

Example 2.49 Let $C[a, b]$ denote the set of complex-valued continuous functions on the interval $[a, b]$. For $f, g \in C[a, b]$, define

$$\langle f, g \rangle = \int_a^b f(x) \overline{g(x)} dx.$$

Then $(C[a, b], \langle \cdot, \cdot \rangle)$ is an inner product space.

Given any inner product space V , we can define $\|x\| = \sqrt{\langle x, x \rangle}$. This is, in fact, a norm on V , and to show this, we need what is known as the Schwarz inequality, that is, $|\langle x, y \rangle| \leq \|x\| \|y\|$ for any two vectors $x, y \in V$ [9, lemma 4.2]. We formally present this result in the following proposition.

Proposition 2.50 Every inner product space V is a normed linear space with the norm $\|x\| = \sqrt{\langle x, x \rangle}$.

Proof:

We verify only the triangle inequality since the other properties follow immediately from definition (2.48). Let $x, y \in V$. Then,

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2 |\langle x, y \rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2 \|x\| \|y\| + \|y\|^2, \quad \text{by Schwarz inequality} \\ &= (\|x\| + \|y\|)^2, \end{aligned}$$

which proves the triangle inequality. □

Definition 2.51 A complete inner product space is called a Hilbert space. (Complete here means that every Cauchy sequence converges.)

Example 2.52 Let $L^2[a, b]$ be the set of complex-valued measurable functions on a finite interval $[a, b]$ that satisfy $\int_a^b |f(x)|^2 dx < \infty$. For $f, g \in L^2[a, b]$ define

$$\langle f, g \rangle = \int_a^b f(x) \overline{g(x)} dx.$$

It can be shown that $L^2[a, b]$ equipped with this inner product is complete and therefore is a Hilbert space.

Definition 2.53 Let V be an inner product space. Two vectors $x, y \in V$ are said to be orthogonal if $\langle x, y \rangle = 0$. A sequence of vectors $\{x_n\}_{n=0}^{\infty}$ in V is called an orthogonal system if

$$\langle x_n, x_m \rangle = h_n \delta_{nm}. \tag{2.1}$$

The system is called orthonormal if $h_n = 1$.

Definition 2.54 A sequence of vectors $\{x_n\}_{n=0}^{\infty}$ in a Hilbert space H is complete if $\langle y, x_n \rangle = 0$ for all $n \geq 0$ implies that $y = 0$.

Definition 2.55 An orthonormal basis is a complete orthonormal system.

The following theorem is standard and can be found in many books, for example, in Reed and Simon [11].

Proposition 2.56 Let $\{x_n\}_{n=0}^{\infty}$ be an orthonormal basis in a Hilbert space H . Then for each $y \in H$,

$$y = \sum_{n=0}^{\infty} \langle y, x_n \rangle x_n \quad \text{and} \quad \|y\|^2 = \sum_{n=0}^{\infty} |\langle y, x_n \rangle|^2. \quad (2.2)$$

The equality in the first expression means that the sum on the right-hand side converges, regardless of order, to y . *Proof*

See Reed and Simon [11, thm. II.6]. □

Corollary 2.57 If $\{x_n\}_{n=0}^{\infty}$ is an orthogonal basis in a Hilbert space H then for each $y \in H$,

$$y = \sum_{n=0}^{\infty} \frac{\langle y, x_n \rangle}{\|x_n\|^2} x_n \quad \text{and} \quad \langle y, z \rangle = \sum_{n=0}^{\infty} \frac{\langle y, x_n \rangle \overline{\langle z, x_n \rangle}}{\|x_n\|^2}. \quad (2.3)$$

2.3 Elementary Theory of Orthogonal Polynomials

We review different aspects of the theory of orthogonal polynomials of one real variable. We require that the domain $X \subset \mathbb{R}$ of polynomials be measurable. X is most commonly either the infinite interval $(-\infty, \infty)$, a semi-infinite interval $[a, \infty)$ or a finite interval $[a, b]$. We also need a weight function described in the following definition.

Definition 2.58 Let $X \subset \mathbb{R}$ be a finite or infinite interval. A function w is called a polynomially bounded weight function if it satisfies the following conditions:

1. w is everywhere nonnegative, integrable over X , and non-zero over a subset of X of positive measure, that is,

$$0 < \int_X w(x) dx < \infty.$$

2. For every $n \in \mathbb{N}$,

$$\int_X x^n w(x) dx < \infty.$$

The quantity $\int_X x^n w(x) dx$ is often called the n^{th} moment of $w(x)$, and is symbolized by μ_n .

Now for a given polynomially bounded weight function w , let $L^2(w)$ denote the space of functions $f : X \rightarrow \mathbb{R}$ whose w -weighted squares have finite integral, that is,

$$f \in L^2(w) \iff \int_X f^2(x) w(x) dx < \infty. \quad (2.4)$$

It follows from condition (2) of definition (2.58) that all polynomials are included in the space $L^2(w)$.

Definition 2.59 Let $\{p_n\}_{n=0}^{\infty}$ be a system of polynomials in the space $L^2(w)$ described above, where the n^{th} polynomial p_n has degree n . Then $\{p_n\}_{n=0}^{\infty}$ is called an orthogonal system with respect to w if

$$\int_X p_n(x)p_m(x)w(x)dx = h_n\delta_{nm}. \quad (2.5)$$

The system is called orthonormal if $h_n = 1$.

More generally if μ is a monotonic non-decreasing function (usually called the distribution function), then we can write equation (2.5) in terms of the Stieltjes integral,

$$\int_X p_n(x)p_m(x)d\mu(x) = h_n\delta_{nm}. \quad (2.6)$$

which is reduced back to (2.5) in case μ is absolutely continuous, that is, if $d\mu(x) = w(x)dx$.

Definition 2.60 If p is a polynomial of degree m and

$$p(x) = c_mx^m + c_{m-1}x^{m-1} + \dots + c_2x^2 + c_1x + c_0, \quad (2.7)$$

then c_m is called the leading coefficient of p . If $c_m = 1$, we say that p is a monic polynomial.

A useful property of real orthogonal polynomials is that they obey a three-term recurrence relation as described in the next proposition [14].

Proposition 2.61 For a weight function w described as in definition (), there exists a unique system of monic orthogonal polynomials $\{p_n\}_{n=0}^{\infty}$. In particular, we can construct $\{p_n\}_{n=0}^{\infty}$ as follows:

$$p_0(x) = 1, \quad p_1(x) = x - a_1 \quad \text{with} \quad a_1 = \frac{\int_X xw(x)dx}{\int_X w(x)dx}$$

and

$$p_{n+1}(x) = xp_n(x) - a_{n+1}p_n(x) - b_{n+1}p_{n-1}(x), \quad (2.8)$$

where

$$a_{n+1} = \frac{\int_X xp_n^2(x)w(x)dx}{\int_X p_n^2(x)w(x)dx} \quad \text{and} \quad b_{n+1} = \frac{\int_X xp_n(x)p_{n-1}(x)w(x)dx}{\int_X p_{n-1}^2(x)w(x)dx}.$$

Remark 4 If w is an even measure, then $a_{n+1} = 0$ since then its integrals with odd polynomials are all zero.

Proof

We begin by proving the existence of monic orthogonal polynomials. The first polynomial p_0 should be monic and of degree zero, and so,

$$p_0(x) = 1.$$

The next polynomial p_1 should be monic and of degree one. It should therefore take the form

$$p_1(x) = x - a_1,$$

and this orthogonal to p_0 implies that

$$0 = \langle p_1, p_0 \rangle = \int_X xw(x)dx - a_1 \int_X w(x)dx.$$

Since w is nonzero on X , it follows that

$$a_1 = \frac{\int_X xw(x)dx}{\int_X w(x)dx}.$$

a_{n+1} and b_{n+1} are found following the same procedure. To prove uniqueness of the sequence $\{p_n\}_{n=0}^{\infty}$ of monic orthogonal polynomials of degree n , assume that $\{q_n\}_{n=0}^{\infty}$ is another sequence of monic orthogonal polynomials of degree n . Then

$$\deg(p_{n+1} - q_{n+1}) \leq n,$$

and since p_{n+1} and q_{n+1} are orthogonal to any polynomial of degree n or less, we have

$$\langle p_{n+1}, p_{n+1} - q_{n+1} \rangle = 0 \quad \text{and} \quad \langle q_{n+1}, p_{n+1} - q_{n+1} \rangle = 0.$$

But this implies that

$$\langle p_{n+1} - q_{n+1}, p_{n+1} - q_{n+1} \rangle = 0,$$

and so, $p_{n+1} - q_{n+1} \equiv 0$ for all $n \geq 0$. □

2.4 Spaces of Interest

The following spaces are of particular interest to our study:

$$L^2(\omega_2), L^2(\omega_1), H^2(\mathbb{S}, \mathcal{P}), L^2(\mathbb{R}), H^2(\mathbb{S}). \quad (2.9)$$

Other useful spaces are:

$$A_0(\mathbb{S}), L^2_{\mathbb{R}}(\omega_2), L^2_{\mathbb{R}}(\omega_1), H^2_{\mathbb{R}}(\mathbb{S}, \mathcal{P}), L^2_{\mathbb{R}}(\mathbb{R}), H^2_{\mathbb{R}}(\mathbb{S}). \quad (2.10)$$

In the above, and indeed throughout this paper, ω_1 denotes the weight function $1/(2 \cosh \frac{\pi}{2}x)$ while ω_2 denote the self convolution of ω_1 , that is, $\omega_2 = \omega_1 * \omega_1$. In fact, it can be shown that $\omega_2(x) = x/(2 \sinh \frac{\pi}{2}x)$.

$L^2(\omega_1)$ denotes the Hilbert space of measurable functions on \mathbb{R} that satisfy $\int_{-\infty}^{\infty} |f(x)|^2 \omega_1(x) dx < \infty$ equipped with the inner product

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x) \overline{g(x)} \omega_1(x) dx = \int_{-\infty}^{\infty} f(x) \overline{g(x)} \frac{dx}{2 \cosh \frac{\pi}{2}x}. \quad (2.11)$$

The Hilbert space $L^2(\omega_2)$ is like the space $L^2(\omega_1)$ but with the weight function ω_2 in place of ω_1 .

$L^2(\mathbb{R})$ denotes the Hilbert space of measurable functions on \mathbb{R} that satisfy $\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty$ equipped with the inner product

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x) \overline{g(x)} dx. \quad (2.12)$$

$H^2(\mathbb{S}, \mathcal{P})$ denotes the Hilbert space of analytic functions on \mathbb{S} that satisfy $\int_{\partial\mathbb{S}} |f(z)|^2 d\mathcal{P}(z) < \infty$ equipped with the inner product

$$\begin{aligned} \langle f, g \rangle &= \int_{\partial\mathbb{S}} f(z) \overline{g(z)} d\mathcal{P}(z) \\ &= \int_{-\infty}^{\infty} f(x+i) \overline{g(x+i)} \left(\frac{\omega_1(x)}{2} \right) dx + \int_{-\infty}^{\infty} f(x-i) \overline{g(x-i)} \left(\frac{\omega_1(x)}{2} \right) dx \\ &= \int_{-\infty}^{\infty} \frac{f(x+i) \overline{g(x+i)} + f(x-i) \overline{g(x-i)}}{2} \omega_1(x) dx \\ &= \int_{-\infty}^{\infty} \frac{f(x+i) \overline{g(x+i)} + f(x-i) \overline{g(x-i)}}{2} \frac{dx}{2 \cosh \frac{\pi}{2} x}. \end{aligned} \quad (2.13)$$

The Hilbert space $H^2(\mathbb{S})$ is like the space $H^2(\mathbb{S}, \mathcal{P})$ but without any weight function.

$A_0(\mathbb{S})$ is the space of functions f that are analytic in \mathbb{S} , continuous on $\partial\mathbb{S}$ and $f(x+iy) \rightarrow 0$ as $|x| \rightarrow \infty$.

The spaces $L_{\mathbb{R}}^2(\omega_2)$, $L_{\mathbb{R}}^2(\omega_1)$ and $L_{\mathbb{R}}^2(\mathbb{R})$ are like the corresponding spaces but restricted to real-valued functions. For the spaces, $H_{\mathbb{R}}^2(\mathbb{S}, \mathcal{P})$ and $H_{\mathbb{R}}^2(\mathbb{S})$, we talk of real-valued functions on the real axis.

3. The τ -System

In this section, we present our first system of orthogonal polynomials which we call the τ -system. This system was studied in Araaya's paper [4], and it was found that it has a simple recurrence relation

$$\tau_{-1} = 0, \tau_0 = 1 \text{ and } \tau_{n+1}(x) = x\tau_n(x) - n^2\tau_{n-1}(x).$$

The first few polynomials for this system are shown below.

$$\begin{aligned} \tau_0 &= 1 \\ \tau_1 &= x \\ \tau_2 &= x^2 - 1 \\ \tau_3 &= x^3 - 5x \\ \tau_4 &= x^4 - 14x^2 + 9 \\ &\vdots \end{aligned}$$

The weight function for this system is $\omega_1(x) = 1/(2 \cosh \frac{\pi}{2} x)$, and as such, we start by looking at two interesting properties of this function that make it useful for this purpose.

Proposition 3.62 *The function ω_1 is a probability density function.*

Proof

This follows from the integration,

$$\int_{-\infty}^{\infty} \omega_1(x) dx = \int_{-\infty}^{\infty} \frac{dx}{2 \cosh \frac{\pi}{2} x} = \left[\frac{1}{\pi} \arctan(\sinh \frac{\pi}{2} x) \right]_{-\infty}^{\infty} = 1.$$

□

The following property makes it possible to interpret the moments of ω_1 as values at zero of successive derivatives.

Proposition 3.63 *The function ω_1 is up to a dilation its own Fourier transform. In particular, it is a Fourier transform of $1/\cosh t$, that is,*

$$\omega_1(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ixt} dt}{\cosh t}$$

Proof

Using the Fourier inversion theorem, we can write

$$\hat{\omega}_1(t) = \int_{-\infty}^{\infty} e^{ixt} \omega_1(x) dx = \int_{-\infty}^{\infty} \frac{e^{(it+\frac{\pi}{2})x}}{e^{\pi x} + 1} dx$$

and show that $\hat{\omega}_1(t) = 1/\cosh t$. For the complete proof, see similar calculations in lemma (3.65). □

We now present the main results for this section. The calculations in proving these results are crucial for proving the main results for the other two systems.

Theorem 3.64 *Let the system $\{\tau_n\}_{n=0}^{\infty}$ be given by the recurrence relation*

$$\tau_{-1} = 0, \tau_0 = 1 \text{ and } \tau_{n+1}(x) = x\tau_n(x) - n^2\tau_{n-1}(x). \quad (3.1)$$

Then

1. *The function τ_n is a monic polynomial of degree n for $n \geq 0$.*
2. *The exponential generating function¹, $G_{\tau}(x, s) = \sum_{n=0}^{\infty} \frac{\tau_n(x)}{n!} s^n$, is given by the function*

$$G_{\tau}(x, s) = \frac{e^{x \arctan s}}{\sqrt{1 + s^2}}.$$

3. *The polynomials $\{\frac{\tau_n}{n!}\}_{n=0}^{\infty}$ are an orthonormal basis in the Hilbert space $L^2(\omega_1)$.*

As aforementioned, the calculations for this proof are similar to those for the other two systems, and since we shall provide a complete proof for the system of section (), we omit this proof. Instead, we provide some tools needed to do this proof and these will also be needed in section ().

¹The exponential generating function of a sequence $\{a_n\}$ is defined as $G(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$.

Lemma 3.65 If $\operatorname{Re}(\alpha) < \frac{\pi}{2}$, then

$$\int_{-\infty}^{\infty} e^{\alpha x} \omega_1(x) dx = \frac{1}{\cos \alpha}.$$

Proof

The complex-valued function $\omega_1(z) = 1/(2 \cosh \frac{\pi}{2} z)$ has a simple pole $z = i$, and so we consider a rectangular contour with vertices $(-R, 0)$, $(R, 0)$, $(R, 2i)$ and $(-R, 2i)$, that is, a contour containing the simple pole. Call this contour C . Then, by the residue theorem, we have

$$\oint_C e^{\alpha z} \omega_1(z) dz = 2\pi i \cdot \operatorname{Res}(i) = 2\pi i \left(\frac{e^{\alpha z}}{\pi \sinh \frac{\pi}{2} z} \Big|_{z=i} \right) = 2e^{\alpha i}. \quad (3.2)$$

Now

$$e^{\alpha z} \omega_1(z) = e^{\alpha z} \frac{1}{2 \cosh \frac{\pi}{2} z} = e^{\alpha z} \frac{2e^{\frac{\pi}{2} z}}{2(e^{\pi z} + 1)} = \frac{e^{(\alpha + \frac{\pi}{2})z}}{e^{\pi z} + 1},$$

and so, we can integrate around the contour C as follows:

$$\begin{aligned} \oint_C e^{\alpha z} \omega_1(z) dz &= \int_{I_1} + \dots + \int_{I_4} \\ &= \int_{-R}^R \frac{e^{(\alpha + \frac{\pi}{2})x}}{e^{\pi x} + 1} dx + i \int_0^2 \frac{e^{(\alpha + \frac{\pi}{2})(R+iy)}}{e^{\pi(R+iy)} + 1} dy \\ &\quad - \int_{-R}^R \frac{e^{(\alpha + \frac{\pi}{2})(x+2i)}}{e^{\pi(x+2i)} + 1} dx - i \int_0^2 \frac{e^{(\alpha + \frac{\pi}{2})(-R+iy)}}{e^{\pi(-R+iy)} + 1} dy. \end{aligned}$$

Along the side I_2 , we have

$$|e^{\alpha z} \omega_1(z)| = \left| \frac{e^{(\alpha + \frac{\pi}{2})(R+iy)}}{e^{\pi(R+iy)} + 1} \right| \leq \frac{e^{(\alpha + \frac{\pi}{2})R}}{e^{\pi R} - 1} = \frac{e^{-\frac{\pi}{2}R} e^{\alpha R}}{1 - e^{-\pi R}}$$

so that by Darboux inequality,

$$\left| \int_{I_2} e^{\alpha z} \omega_1(z) dz \right| \leq \frac{2e^{-\frac{\pi}{2}R} e^{\alpha R}}{1 - e^{-\pi R}} \rightarrow 0 \text{ as } R \rightarrow \infty.$$

Similarly, the integral along I_4 vanishes as $R \rightarrow \infty$.

Thus, taking $R \rightarrow \infty$ and combining with (3.2), we have

$$\begin{aligned}
2e^{\alpha i} &= \lim_{R \rightarrow \infty} \oint_C e^{\alpha z} \omega_1(z) dz \\
&= \lim_{R \rightarrow \infty} \int_{-R}^R \frac{e^{(\alpha + \frac{\pi}{2})x}}{e^{\pi x} + 1} dx - \lim_{R \rightarrow \infty} \int_{-R}^R \frac{e^{(\alpha + \frac{\pi}{2})(x+2i)}}{e^{\pi(x+2i)} + 1} dx \\
&= \lim_{R \rightarrow \infty} (1 + e^{i2\alpha}) \int_{-R}^R \frac{e^{(\alpha + \frac{\pi}{2})x}}{e^{\pi x} + 1} dx \\
&= (1 + e^{i2\alpha}) \int_{-\infty}^{\infty} \frac{e^{(\alpha + \frac{\pi}{2})x}}{e^{\pi x} + 1} dx
\end{aligned}$$

which implies that

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{e^{(\alpha + \frac{\pi}{2})x}}{e^{\pi x} + 1} dx &= \frac{2e^{\alpha i}}{1 + e^{i2\alpha}} \\
&= \frac{1}{\cos \alpha}.
\end{aligned}$$

□

Lemma 3.66 *The following identity holds:*

$$\cos(\alpha + \beta) = \frac{1 - \tan \alpha \tan \beta}{\sqrt{1 + \tan^2 \alpha} \sqrt{1 + \tan^2 \beta}}.$$

Proof

It is a well known fact that $\cos^2 x + \sin^2 x = 1$. Dividing through by $\cos^2 x$ gives $1 + \tan^2 x = \sec^2 x$. Thus,

$$\begin{aligned}
\frac{1 - \tan \alpha \tan \beta}{\sqrt{1 + \tan^2 \alpha} \sqrt{1 + \tan^2 \beta}} &= \frac{1 - \tan \alpha \tan \beta}{\sec \alpha \sec \beta} \\
&= \cos \alpha \cos \beta \left(1 - \frac{\sin \alpha \sin \beta}{\cos \alpha \cos \beta}\right) \\
&= \cos \alpha \cos \beta - \sin \alpha \sin \beta \\
&= \cos(\alpha + \beta).
\end{aligned}$$

□

4. The σ -System and Some Useful Operators

Like the τ -system, this system was studied in Araaya's paper [4], and it was found that it has a simple recurrence relation

$$\sigma_{-1} = 0, \sigma_0 = 1 \text{ and } \sigma_{n+1}(x) = x\sigma_n(x) - n(n-1)\sigma_{n-1}(x).$$

The first few polynomials for this system are shown below.

$$\begin{aligned}
\sigma_0 &= 1 \\
\sigma_1 &= x \\
\sigma_2 &= x^2 \\
\sigma_3 &= x^3 - 2x \\
\sigma_4 &= x^4 - 8x^2 \\
&\vdots
\end{aligned}$$

The first two properties of the function $\omega_1(x) = 1/(2 \cosh \frac{\pi}{2}x)$ were discussed in section (). The third useful property is that, it is closely related to the Poisson kernel for a strip of width two in the manner of the following proposition.

Proposition 4.67 *Let the function f be continuous and harmonic in the strip $\mathbb{S} = \{z \in \mathbb{C} : -1 \leq \text{Im}(z) \leq 1\}$, and suppose further that $|f(z)| < Ce^{a|z|}$ for some $a \in [0, \frac{\pi}{2})$. Then*

$$\begin{aligned}
f(0) &= \int_{-\infty}^{\infty} f(x+i) \frac{dx}{4 \cosh \frac{\pi}{2}x} + \int_{-\infty}^{\infty} f(x-i) \frac{dx}{4 \cosh \frac{\pi}{2}x} \\
&= \int_{-\infty}^{\infty} \frac{f(x+i) + f(x-i)}{2} \frac{dx}{2 \cosh \frac{\pi}{2}x} \\
&= \int_{-\infty}^{\infty} \frac{f(x+i) + f(x-i)}{2} \omega_1(x) dx.
\end{aligned}$$

Proof

This is simply the Poisson integral. □

In the preceding proposition, we used the operator,

$$Rf(x) = \frac{1}{2}(f(x+i) + f(x-i)), \quad (4.1)$$

which is densely defined in $L^2(\omega_1)$. For symmetry, we also consider the operator,

$$Jf(x) = \frac{1}{2i}(f(x+i) - f(x-i)). \quad (4.2)$$

It is clear from the definition of these two operators that

$$(R \pm iJ)f(x) = f(x \pm i). \quad (4.3)$$

In the next section, we shall see that multiplying the ρ -system by x gives a relation to this system, and this is the reason to define the third operator,

$$Qf(x) = xf(x). \quad (4.4)$$

The notation for this operator is inspired by analogies with quantum mechanics, an analogy which seems natural in the light of the following easily verified relations between the operators.

Proposition 4.68 *The operators R, J and Q satisfy the following relations:*

$$RQ - QR = -J \quad (4.5)$$

$$JQ - QJ = R \quad (4.6)$$

$$RJ - JR = 0 \quad (4.7)$$

$$R^2 + J^2 = I \quad (4.8)$$

where I is the identity operator.

Proof

Use the definition of the operators involved. □

We now present the main results for this section which describe an orthogonal basis for the space $H^2(\mathbb{S}, \mathcal{P})$ where \mathcal{P} is the Poisson measure for 0.

Theorem 4.69 *Let the system $\{\sigma_n\}_{n=0}^\infty$ be given by the recurrence relation.*

$$\sigma_{-1} = 0, \sigma_0 = 1 \text{ and } \sigma_{n+1}(x) = x\sigma_n(x) - n(n-1)\sigma_{n-1}(x). \quad (4.9)$$

Then

1. The function σ_n is a monic polynomial of degree n for $n \geq 0$.
2. The exponential generating function, $G_\sigma(x, s) = \sum \frac{\sigma_n(x)}{n!} s^n$, is given by the function

$$G_\sigma(x, s) = e^{x \arctan s}.$$

3. The norm of the polynomial $\frac{\sigma_n}{n!}$ is 1 for $n = 0$ and $\sqrt{2}$ for $n \geq 1$.
4. The polynomials $\{\frac{\sigma_n}{n!}\}_{n=0}^\infty$ are an orthogonal basis in the Hilbert space $H^2(\mathbb{S}, \mathcal{P})$.

Proof

See similar calculations in the proof of theorem (5.73) in the next section. □

5. The ρ -System

We study this system in detail since it is a new addition, filling a gap related to the previous systems. In fact, it is the main motivation behind this thesis. Unlike the two previous systems, the weight function for this system is $\omega_2 = \omega_1 * \omega_1$, the self convolution of $\omega_1(x) = 1/(2 \cosh \frac{\pi}{2}x)$. By the convolution theorem and proposition (3.63), the Fourier transform $\hat{\omega}_2$ of ω_2 is given by $\hat{\omega}_2(t) = \hat{\omega}_1(t) \cdot \hat{\omega}_1(t) = 1/\cosh^2 t$. Abramowitz [1] gives the Maclaurin series expansion

$$\begin{aligned} \frac{1}{\cosh^2 t} &= \left(\sum_{n=0}^{\infty} \frac{E_{2n} t^{2n}}{(2n)!} \right)^2 \\ &= \left(1 - \frac{t^2}{2} + \frac{5t^4}{24} - \frac{61t^6}{720} + \frac{1385t^8}{40320} + \dots \right)^2 \\ &= 1 - t^2 + \frac{2t^4}{3} - \frac{17t^6}{45} + \dots \end{aligned} \quad (5.1)$$

where E_n is the n^{th} Euler number².

²The first few Euler numbers are: 1, -1, 5, -61, 1385, -50521 with alternating signs. For the explicit definition and formula, see [1, p. 804].

Now using the Fourier inversion theorem, $\hat{\omega}_2(t) = \int_{-\infty}^{\infty} e^{ixt} \omega_2(x) dx$, we derive the n^{th} derivative of $\hat{\omega}_2$ evaluated at zero as follows:

$$\begin{aligned} \hat{\omega}_2(t) &= \int_{-\infty}^{\infty} e^{ixt} \omega_2(x) dx, & \hat{\omega}_2(0) &= \int_{-\infty}^{\infty} \omega_2(x) dx \\ \hat{\omega}'_2(t) &= \int_{-\infty}^{\infty} ix e^{ixt} \omega_2(x) dx, & \hat{\omega}'_2(0) &= \int_{-\infty}^{\infty} ix \omega_2(x) dx \\ \hat{\omega}''_2(t) &= \int_{-\infty}^{\infty} (ix)^2 e^{ixt} \omega_2(x) dx, & \hat{\omega}''_2(0) &= \int_{-\infty}^{\infty} (ix)^2 \omega_2(x) dx \\ &\vdots & &\vdots \\ \hat{\omega}_2^{(n)}(t) &= \int_{-\infty}^{\infty} (ix)^n e^{ixt} \omega_2(x) dx, & \hat{\omega}_2^{(n)}(0) &= \int_{-\infty}^{\infty} (ix)^n \omega_2(x) dx \end{aligned}$$

Since $\hat{\omega}_2$ is an even function, it is orthogonal to all odd polynomials. Thus all odd derivatives vanish, and we can rewrite the expression for the n^{th} derivative of $\hat{\omega}_2$ evaluated at zero as

$$\int_{-\infty}^{\infty} x^{2n} \omega_2(x) dx = (-i)^{2n} \hat{\omega}_2^{(2n)}(0) = (-i)^{2n} \left(\frac{d}{dt} \right)^{2n} \left(\frac{1}{\cosh^2 t} \right) \Big|_{t=0}, \quad (5.2)$$

which is then used together with equation (5.1) to find the moments as follows:

$$\begin{aligned} n = 0, & \quad \int_{-\infty}^{\infty} \omega_2(x) dx = (-i)^0 \hat{\omega}_2(0) = 1 \\ n = 1, & \quad \int_{-\infty}^{\infty} x^2 \omega_2(x) dx = (-i)^2 \hat{\omega}_2''(0) = (-i)^2 (-2! \times 1) = 2 \\ n = 2, & \quad \int_{-\infty}^{\infty} x^4 \omega_2(x) dx = (-i)^4 \hat{\omega}_2^{(4)}(0) = (-i)^4 \left(4! \times \frac{2}{3} \right) = 16 \\ n = 3, & \quad \int_{-\infty}^{\infty} x^6 \omega_2(x) dx = (-i)^6 \hat{\omega}_2^{(6)}(0) = (-i)^6 \left(-6! \times \frac{17}{45} \right) = 272 \\ & \quad \vdots \end{aligned}$$

We can now use proposition (2.61) to construct a unique system of monic orthogonal polynomials $\{\rho_n\}_{n=0}^{\infty}$. Set $\rho_0(x) = 1$ and since this has an even power of x , it is orthogonal to all odd polynomials and in particular to $\rho_1(x) = x$. To find the third polynomial, set $\rho_2(x) = x^2 + a$ and this orthogonal to ρ_0 implies that

$$0 = \int_{-\infty}^{\infty} (x^2 + a) \omega_2(x) dx = \int_{-\infty}^{\infty} x^2 \omega_2(x) dx + a \int_{-\infty}^{\infty} \omega_2(x) dx = 2 + a.$$

Thus $a = -2$. To find the fourth polynomial, set $\rho_3(x) = x^3 + bx$ and this orthogonal to ρ_1 implies that

$$0 = \int_{-\infty}^{\infty} (x^3 + bx)x\omega_2(x)dx = \int_{-\infty}^{\infty} x^4\omega_2(x)dx + b \int_{-\infty}^{\infty} x^2\omega_2(x)dx = 16 + 2b.$$

Thus $b = -8$. To find the fifth polynomial, set $\rho_4(x) = x^4 + cx^2 + d$ and this orthogonal to ρ_0 implies that

$$\begin{aligned} 0 &= \int_{-\infty}^{\infty} (x^4 + cx^2 + d)\omega_2(x)dx \\ &= \int_{-\infty}^{\infty} x^4\omega_2(x)dx + c \int_{-\infty}^{\infty} x^2\omega_2(x)dx + d \int_{-\infty}^{\infty} \omega_2(x)dx \\ &= 16 + 2c + d. \end{aligned} \tag{5.3}$$

ρ_4 should also be orthogonal to ρ_2 , and so,

$$\begin{aligned} 0 &= \int_{-\infty}^{\infty} (x^4 + cx^2 + d)(x^2 - 2)\omega_2(x)dx \\ &= \int_{-\infty}^{\infty} (x^6 + cx^4 + dx^2)\omega_2(x)dx \\ &= \int_{-\infty}^{\infty} x^6\omega_2(x)dx + c \int_{-\infty}^{\infty} x^4\omega_2(x)dx + d \int_{-\infty}^{\infty} x^2\omega_2(x)dx \\ &= 272 + 16c + 2d \\ &= 272 + 16c + 2(-16 - 2c), \quad \text{by (5.3)} \\ &= 240 + 12c. \end{aligned}$$

Thus, $c = -20$ and $d = 24$. The rest of the ρ -polynomials are obtained following the same procedure, and we have

$$\begin{aligned} \rho_0(x) &= 1 \\ \rho_1(x) &= x \\ \rho_2(x) &= x^2 - 2 \\ \rho_3(x) &= x^3 - 8x \\ \rho_4(x) &= x^4 - 20x^2 + 24 \\ &\vdots \end{aligned}$$

We now establish the relationship between these polynomials. Setting $\rho_{-1} = 0$, we note that

$$\begin{aligned} \rho_1(x) &= x\rho_0(x) - \rho_{-1}(x) & 0 &= 0 \times 1 \\ \rho_2(x) &= x\rho_1(x) - 2\rho_0(x) & 2 &= 1 \times 2 \\ \rho_3(x) &= x\rho_2(x) - 6\rho_1(x) & 6 &= 2 \times 3 \\ \rho_4(x) &= x\rho_3(x) - 12\rho_2(x) & 12 &= 3 \times 4 \\ &\vdots & &\vdots \end{aligned}$$

where the second column shows the pattern of the coefficients of the second terms on the right hand side of the polynomial equations. This pattern of the coefficients motivates us to define the system of polynomials $\{\rho_n\}_{n=0}^{\infty}$ by the recurrence relation

$$\rho_{-1} = 0, \rho_0 = 1, \text{ and } \rho_{n+1}(x) = x\rho_n(x) - n(n+1)\rho_{n-1}(x),$$

which we will later use to compute the exponential generating function for proving orthogonality of our system.

Before proceeding further, we present two lemmas that will be useful in proving the main results of this section.

Lemma 5.70 *If the function f is integrable on $(-\infty, \infty)$ and*

$$\hat{f}(x) = \int_{-\infty}^{\infty} f(t)e^{ixt} dt \equiv 0,$$

then $f = 0$ almost everywhere.

Proof

See Andrews, Askey and Roy [3, thm. 6.5.1]. □

Lemma 5.71 *If $Re(\alpha) < \frac{\pi}{2}$, then*

$$\int_{-\infty}^{\infty} e^{\alpha x} \omega_2(x) dx = \left(\frac{1}{\cos \alpha} \right)^2. \quad (5.4)$$

Proof

Bearing in mind that $\omega_2 = \omega_1 * \omega_1$, a self convolution, we have

$$\begin{aligned} \int_{-\infty}^{\infty} e^{\alpha x} \omega_2(x) dx &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{\alpha x} \omega_1(x-y) \omega_1(y) dy dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{\alpha(t+y)} \omega_1(t) \omega_1(y) dt dy \quad \text{if we let } x-y = t \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{\alpha t} \omega_1(t) dt \right) e^{\alpha y} \omega_1(y) dy \\ &= \int_{-\infty}^{\infty} e^{\alpha t} \omega_1(t) dt \int_{-\infty}^{\infty} e^{\alpha y} \omega_1(y) dy \\ &= \left(\int_{-\infty}^{\infty} e^{\alpha x} \omega_1(x) dx \right)^2 \quad \text{if we let } t = y = x \\ &= \left(\frac{1}{\cos \alpha} \right)^2, \quad \text{by lemma (3.65)}. \end{aligned}$$

□

Lemma 5.72 For $|x| < 1$,

$$\frac{1}{(1-x)^2} = \sum_{n=0}^{\infty} (n+1)x^n.$$

Proof

Differentiate the geometric series, $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$, with respect to x , that is,

$$\frac{1}{(1-x)^2} = \sum_{n=1}^{\infty} nx^{n-1} = \sum_{n=0}^{\infty} (n+1)x^n.$$

□

We now have all the necessary definitions and lemmas needed to present and prove the main results of this section.

Theorem 5.73 Let the system $\{\rho_n\}_{n=0}^{\infty}$ be given by the recurrence relation

$$\rho_{-1} = 0, \rho_0 = 1 \text{ and } \rho_{n+1}(x) = x\rho_n(x) - n(n+1)\rho_{n-1}(x). \quad (5.5)$$

Then

1. The function ρ_n is a monic polynomial of degree n for $n \geq 0$.
2. The exponential generating function, $G_\rho(x, s) = \sum_{n=0}^{\infty} \frac{\rho_n(x)}{n!} s^n$, is given by the function

$$G_\rho(x, s) = \frac{e^{x \arctan s}}{1 + s^2}.$$

3. The sequence of polynomials $\{\frac{\rho_n}{n!}\}_{n=0}^{\infty}$ is an orthogonal basis in the Hilbert space $L^2(\omega_2)$.

Proof

(1) follows immediately from the definition of the recurrence relation. To prove (2), we multiply the recurrence by $s^n/n!$ and sum over n so that

$$\begin{aligned} 0 &= \sum_{n=0}^{\infty} [\rho_{n+1}(x) - x\rho_n(x) + n(n+1)\rho_{n-1}(x)] \frac{s^n}{n!} \\ &= \sum_{n=0}^{\infty} \rho_{n+1}(x) \frac{s^n}{n!} - x \sum_{n=0}^{\infty} \rho_n(x) \frac{s^n}{n!} + \sum_{n=1}^{\infty} n(n+1)\rho_{n-1}(x) \frac{s^n}{n!} \\ &= G'_\rho(x, s) - xG_\rho(x, s) + \sum_{n=0}^{\infty} (n+1)(n+2)\rho_n(x) \frac{s^{n+1}}{(n+1)!} \\ &= G'_\rho(x, s) - xG_\rho(x, s) + 2s \sum_{n=0}^{\infty} \rho_n(x) \frac{s^n}{n!} + \sum_{n=1}^{\infty} n\rho_n(x) \frac{s^{n+1}}{n!} \\ &= G'_\rho(x, s) - xG_\rho(x, s) + 2sG_\rho(x, s) + \sum_{n=0}^{\infty} (n+1)\rho_{n+1}(x) \frac{s^{n+2}}{(n+1)n!} \\ &= G'_\rho(x, s) - xG_\rho(x, s) + 2sG_\rho(x, s) + s^2G'_\rho(x, s) \\ &= (1+s^2)G'_\rho(x, s) + (2s-x)G_\rho(x, s). \end{aligned}$$

Thus,

$$G'_\rho(x, s) + \frac{2s - x}{1 + s^2} G_\rho(x, s) = 0. \quad (5.6)$$

This is a first-order linear differential equation where all derivatives are with respect to s , holding x fixed. The integrating factor is

$$\begin{aligned} \exp\left(\int \frac{2s - x}{1 + s^2} ds\right) &= \exp\left(\int \frac{2s}{1 + s^2} ds - \int \frac{x}{1 + s^2} ds\right) \\ &= \exp(\ln(1 + s^2) - x \arctan s) \\ &= (1 + s^2)e^{-x \arctan s}. \end{aligned}$$

Multiplying both sides of equation (5.6) by this factor gives

$$\frac{d}{ds} \left((1 + s^2)e^{-x \arctan s} G_\rho(x, s) \right) = 0$$

which implies that

$$G_\rho(x, s) = c \frac{e^{x \arctan s}}{1 + s^2}.$$

Now since $G_\rho(x, s) = \sum_{n=0}^{\infty} \frac{\rho_n(x)}{n!} s^n$, it implies that $G_\rho(x, 0) = 1$. Thus $c=1$ and (2) follows.

To prove (3), we first show that

$$\int_{-\infty}^{\infty} G_\rho(x, s) \overline{G_\rho(x, t)} \omega_2(x) dx = \frac{1}{(1 - st)^2}.$$

Now

$$G_\rho(x, s) \overline{G_\rho(x, t)} = \frac{e^{x \arctan s}}{1 + s^2} \frac{e^{x \arctan \bar{t}}}{1 + \bar{t}^2} = \frac{1}{(1 + s^2)(1 + \bar{t}^2)} e^{x(\arctan s + \arctan \bar{t})}.$$

Set $u = \frac{1}{(1+s^2)(1+\bar{t}^2)}$, $\alpha = \arctan s$, $\beta = \arctan \bar{t}$ and assume that $Re(\alpha + \beta) < \frac{\pi}{2}$. Then we have

$$\begin{aligned} \int_{-\infty}^{\infty} G_\rho(x, s) \overline{G_\rho(x, t)} \omega_2(x) dx &= u \int_{-\infty}^{\infty} e^{(\alpha+\beta)x} \omega_2(x) dx \\ &= u \left(\frac{1}{\cos(\alpha + \beta)} \right)^2, \quad \text{by lemma (5.71)} \\ &= u \left(\frac{\sqrt{1 + \tan^2 \alpha} \sqrt{1 + \tan^2 \beta}}{1 - \tan \alpha \tan \beta} \right)^2, \quad \text{by lemma (3.66)} \\ &= u \left(\frac{\sqrt{1 + s^2} \sqrt{1 + \bar{t}^2}}{1 - st} \right)^2 \\ &= u \cdot \frac{(1 + s^2)(1 + \bar{t}^2)}{(1 - st)^2} \\ &= \frac{1}{(1 - st)^2}. \end{aligned} \quad (5.7)$$

Next, by lemma (5.72) we see that this implies that

$$\int_{-\infty}^{\infty} G_{\rho}(x, s) \overline{G_{\rho}(x, t)} \omega_2(x) dx = \sum_{n=0}^{\infty} (n+1) (s\bar{t})^n. \quad (5.8)$$

But using the definition, $G_{\rho}(x, s) = \sum_{n=0}^{\infty} \frac{\rho_n(x)}{n!} s^n$, gives

$$\begin{aligned} \int_{-\infty}^{\infty} G_{\rho}(x, s) \overline{G_{\rho}(x, t)} \omega_2(x) dx &= \int_{-\infty}^{\infty} \left(\sum_{n=0}^{\infty} \frac{\rho_n(x)}{n!} s^n \right) \left(\sum_{k=0}^{\infty} \frac{\rho_k(x)}{k!} \bar{t}^k \right) \omega_2(x) dx \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} s^n \bar{t}^k \int_{-\infty}^{\infty} \frac{\rho_n(x) \rho_k(x)}{n! k!} \omega_2(x) dx. \end{aligned} \quad (5.9)$$

It therefore follows from (5.8) and (5.9) that

$$\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} s^n \bar{t}^k \int_{-\infty}^{\infty} \frac{\rho_n(x) \rho_k(x)}{n! k!} \omega_2(x) dx = \sum_{n=0}^{\infty} (n+1) (s\bar{t})^n.$$

Comparing the coefficients of the powers of s and \bar{t} proves orthogonality, that is,

$$\left\langle \frac{\rho_n(x)}{n!}, \frac{\rho_k(x)}{k!} \right\rangle = (n+1) \delta_{nk} \quad (5.10)$$

To show that this system of polynomials $\{\frac{\rho_n}{n!}\}_{n=0}^{\infty}$ is a basis in the Hilbert space $L^2(\omega_2)$, we need to show that it is complete. But since the span of $\{\frac{\rho_n}{n!}\}_{n=0}^{\infty}$ is the space of all polynomials, it suffices to show density of the system $\{x^n\}_{n=0}^{\infty}$. Let $\langle f, x^n \rangle = 0$ for some $f \in L^2(\omega_2)$ and all $n \geq 0$. Then

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) e^{itx} \omega_2(x) dx &= \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \int_{-\infty}^{\infty} f(x) x^n \omega_2(x) dx \\ &= \lim_{N \rightarrow \infty} \sum_{n=0}^N \frac{(it)^n}{n!} \int_{-\infty}^{\infty} f(x) x^n \omega_2(x) dx \\ &= \lim_{N \rightarrow \infty} \sum_{n=0}^N \frac{(it)^n}{n!} \cdot 0 \\ &= 0. \end{aligned}$$

By Lemma (5.70), $f\omega_2 = 0$ almost everywhere. But $\omega_2 \neq 0$ and so $f = 0$ almost everywhere which by definition (2.54) implies that $\{x^n\}_{n=0}^{\infty}$ is dense in $L^2(\omega_2)$. Therefore, the system $\{\frac{\rho_n}{n!}\}_{n=0}^{\infty}$ is complete, and in particular, it is an orthogonal basis in the Hilbert space $L^2(\omega_2)$. \square

6. Some Connections Between the Systems

Having presented the three systems of polynomials in the previous sections, we can now discuss some useful connections between them, in terms of the operators R , J and Q . To start with, let us write a few terms for each system. By definition, $\sigma_{-1} = \tau_{-1} = \rho_{-1} = 0$, $\sigma_0 = \tau_0 = \rho_0 = 1$, and $\sigma_{n+1}(x) = x\sigma_n(x) - n(n-1)\sigma_{n-1}(x)$, $\tau_{n+1}(x) = x\tau_n(x) - n^2\tau_{n-1}(x)$ and $\rho_{n+1}(x) = x\rho_n(x) - n(n+1)\rho_{n-1}(x)$.

We thus have

σ	τ	ρ
$\sigma_0 = 1$	$\tau_0 = 1$	$\rho_0 = 1$
$\sigma_1 = x$	$\tau_1 = x$	$\rho_1 = x$
$\sigma_2 = x^2$	$\tau_2 = x^2 - 1$	$\rho_2 = x^2 - 2$
$\sigma_3 = x^3 - 2x$	$\tau_3 = x^3 - 5x$	$\rho_3 = x^3 - 8x$
$\sigma_4 = x^4 - 8x^2$	$\tau_4 = x^4 - 14x^2 + 9$	$\rho_4(x) = x^4 - 20x^2 + 24$
\vdots	\vdots	\vdots

Our three operators are defined by $Rf(x) = \frac{1}{2}(f(x+i) + f(x-i))$, $Qf(x) = xf(x)$ and $Jf(x) = \frac{1}{2i}(f(x+i) - f(x-i))$. Comparing columns 1 and 3, we see that $x\rho_n = \sigma_{n+1}$ which by definition of Q implies that $Q\rho_n = \sigma_{n+1}$. In what follows below, we check the operations of R , J and Q on the three systems of polynomials. We start with the operator R . On the first column, we have

$$\begin{aligned}
 R\sigma_0 &= \frac{\sigma_0(x+i) + \sigma_0(x-i)}{2} = \frac{1+1}{2} = 1 \\
 R\sigma_1 &= \frac{(x+i) + (x-i)}{2} = \frac{2x}{2} = x \\
 R\sigma_2 &= \frac{(x+i)^2 + (x-i)^2}{2} = \frac{2x^2 - 2}{2} = x^2 - 1 \\
 R\sigma_3 &= \frac{(x+i)^3 - 2(x+i) + (x-i)^3 - 2(x-i)}{2} = x^3 - 5x \\
 &\vdots
 \end{aligned}$$

This indicates that the operation of R on column 1 gives column 2. We can therefore claim that $R\sigma_n = \tau_n$ which we will prove later. On column 2, we have

$$\begin{aligned}
 R\tau_0 &= 1 \\
 R\tau_1 &= x \\
 R\tau_2 &= \frac{(x+i)^2 - 1 + (x-i)^2 - 1}{2} = \frac{2x^2 - 4}{2} = x^2 - 2 \\
 R\tau_3 &= \frac{(x+i)^3 - 5(x+i) + (x-i)^3 - 5(x-i)}{2} = x^3 - 8x \\
 &\vdots
 \end{aligned}$$

This indicates that the operation of R on column 2 gives column 3. We can therefore claim that $R\tau_n = \rho_n$ which we will prove later. We now turn to the operator J . On column 1, we have

$$\begin{aligned}
 J\sigma_0 &= \frac{\tau_0(x+i) - \tau_0(x-i)}{2i} = \frac{1-1}{2i} = 0 \\
 J\sigma_1 &= \frac{(x+i) - (x-i)}{2i} = \frac{2i}{2i} = 1 \\
 J\sigma_2 &= \frac{(x+i)^2 - (x-i)^2}{2i} = \frac{4xi}{2i} = 2x \\
 J\sigma_3 &= \frac{(x+i)^3 - 2(x+i) - (x-i)^3 + 2(x-i)}{2i} = 3x^2 - 3 = 3(x^2 - 1) \\
 &\vdots
 \end{aligned}$$

From this, we can claim that $J\sigma_n = n\tau_{n-1}$ which will be proved later. On column 2, we have

$$\begin{aligned}
J\tau_0 &= 0 \\
J\tau_1 &= 1 \\
J\tau_2 &= \frac{(x+i)^2 - 1 - (x-i)^2 + 1}{2i} = \frac{4xi}{2i} = 2x \\
J\tau_3 &= \frac{(x+i)^3 - 5(x+i) - (x-i)^3 + 5(x-i)}{2i} = 3x^2 - 6 = 3(x^2 - 3) \\
&\vdots
\end{aligned}$$

From this, we can claim that $J\tau_n = n\rho_{n-1}$ which will be proved later.

We can now state the main results of this section.

Theorem 6.74 *The following connections between the three systems of orthogonal polynomials $\{\sigma_n\}$, $\{\tau_n\}$ and $\{\rho_n\}$ hold:*

$$R\sigma_n = \tau_n \quad (6.1)$$

$$J\sigma_n = n\tau_{n-1} \quad (6.2)$$

$$R\tau_n = \rho_n \quad (6.3)$$

$$J\tau_n = n\rho_{n-1} \quad (6.4)$$

$$Q\rho_n = \sigma_{n+1} \quad (6.5)$$

Proof

We shall prove only (6.3) and (6.4) since the proofs for the rest follow the same procedure. The idea of the proof is that, given (6.3), we prove by induction (6.4), and viceversa.

We start with (6.3). For $n = 0$, the statement is true since we have

$$R\tau_0(x) = \frac{\tau_0(x+i) + \tau_0(x-i)}{2} = \frac{1+1}{2} = 1 = \rho_0(x).$$

Now assume that both (6.3) and (6.4) hold for all τ_k , $k \leq n$, then

$$\begin{aligned}
R\tau_{n+1}(x) &= R[x\tau_n(x) - n^2\tau_{n-1}(x)], \quad \text{by recurrence relation} \\
&= \frac{(x+i)\tau_n(x+i) + (x-i)\tau_n(x-i)}{2} - \frac{n^2\tau_{n-1}(x+i) + \tau_{n-1}(x-i)}{2} \\
&= x\frac{\tau_n(x+i) + \tau_n(x-i)}{2} + i\frac{\tau_n(x+i) - \tau_n(x-i)}{2} - n^2R\tau_{n-1}(x) \\
&= xR\tau_n(x) - \frac{\tau_n(x+i) - \tau_n(x-i)}{2i} - n^2R\tau_{n-1} \\
&= xR\tau_n(x) - J\tau_n(x) - n^2R\tau_{n-1} \\
&= xR\tau_n(x) - n\rho_{n-1}(x) - n^2R\tau_{n-1}, \quad \text{by (6.4) assumption} \\
&= x\rho_n(x) - n\rho_{n-1}(x) - n^2\rho_{n-1}(x), \quad \text{by induction assumption} \\
&= x\rho_n(x) - n(n+1)\rho_{n-1}(x) \\
&= \rho_{n+1}(x).
\end{aligned}$$

Therefore, since the statement is also true for $n + 1$, it follows by induction that it is true for all integers $n \geq 0$.

The proof for (6.4) follows the same procedure, and as such, we omit it. \square

We now introduce some notations related to the three systems of polynomials. Denote the polynomials $\frac{\sigma_n}{n!}, \frac{\tau_n}{n!}, \frac{\rho_n}{n!}$ by $\tilde{\sigma}_n, \tilde{\tau}_n, \tilde{\rho}_n$ respectively. It follows from Theorems (3.64), (4.69) and (5.73) that the systems $\{\tilde{\sigma}_n\}_{n=0}^{\infty}, \{\tilde{\tau}_n\}_{n=0}^{\infty}$ and $\{\tilde{\rho}_n\}_{n=0}^{\infty}$ are orthogonal bases for the Hilbert spaces $H^2(\mathbb{S}, \mathcal{P}), L^2(\omega_1)$ and $L^2(\omega_2)$ respectively. In fact, the system $\{\tilde{\tau}_n\}_{n=0}^{\infty}$ is orthonormal. In what follows, we look at some consequences of the relations in Theorem (6.74).

Corollary 6.75 *The following connections between the three systems of orthogonal polynomials $\{\tilde{\sigma}_n\}, \{\tilde{\tau}_n\}$ and $\{\tilde{\rho}_n\}$ hold:*

$$R\tilde{\sigma}_n = \tilde{\tau}_n \quad (6.6)$$

$$J\tilde{\sigma}_n = \tilde{\tau}_{n-1} \quad (6.7)$$

$$R\tilde{\tau}_n = \tilde{\rho}_n \quad (6.8)$$

$$J\tilde{\tau}_n = \tilde{\rho}_{n-1} \quad (6.9)$$

$$Q\tilde{\rho}_n = (n + 1)\tilde{\sigma}_{n+1} \quad (6.10)$$

Proof

Divide the equations in Theorem (6.74) through by $n!$. For instance, we have for relation (6.9),

$$\begin{aligned} J\tau_n &= n\rho_{n-1} \\ \frac{J\tau_n}{n!} &= \frac{n\rho_{n-1}}{n!} \Rightarrow J\tilde{\tau}_n = \frac{\rho_{n-1}}{(n-1)!} \Rightarrow J\tilde{\tau}_n = \tilde{\rho}_{n-1}. \end{aligned}$$

\square

Corollary 6.76 *Let the operators K, L, M, A, B and C be defined as follows: $K = RRQ, L = QRR, M = RQR, A = RQJ, B = QJR$ and $C = RJQ$. Then the following relations hold:*

$$K^n(\rho_0) = \rho_n \quad (6.11)$$

$$L^n(\sigma_0) = \sigma_n \quad (6.12)$$

$$M^n(\tau_0) = \tau_n \quad (6.13)$$

$$A(\tau_n) = n\tau_n \quad (6.14)$$

$$B(\sigma_n) = n\sigma_n \quad (6.15)$$

$$C(\rho_n) = n\rho_n \quad (6.16)$$

Proof

We prove only (6.11) and (6.16). The proofs for the rest follow the same procedure.

For (6.11), we proceed by induction. For $n = 0$, the statement is trivially true. Now assume that it is true for some integer $n \geq 0$, then

$$\begin{aligned} K^{n+1}(\rho_0) &= KK^n(\rho_0) \\ &= K\rho_n, \quad \text{by induction assumption} \\ &= RRQ\rho_n \\ &= \rho_{n+1}, \quad \text{by Theorem (6.74)}. \end{aligned}$$

Therefore, since the statement is also true for $n + 1$, it follows by induction that it is true for all integers $n \geq 0$.

For (6.16), we use the definition of C and the relations in Theorem (6.74),

$$\begin{aligned} C(\rho_n) &= RJQ(\rho_n) \\ &= RJ\sigma_{n+1} \\ &= Rn\tau_n \\ &= n\rho_n. \end{aligned}$$

□

Corollary 6.77 *The following relations hold:*

$$\tilde{\tau}_n(x \pm i) = \tilde{\rho}_n(x) \pm i\tilde{\rho}_{n-1}(x) \quad (6.17)$$

$$\tilde{\sigma}_n(x \pm i) = \tilde{\tau}_n(x) \pm i\tilde{\tau}_{n-1}(x) \quad (6.18)$$

Proof

We prove only (6.17) since the proof for (6.18) follows the same procedure. From corollary (6.75), $\tilde{\rho}_n = R\tilde{\tau}_n$ and $\tilde{\rho}_{n-1} = J\tilde{\tau}_n$. Thus,

$$\begin{aligned} \tilde{\rho}_n(x) \pm i\tilde{\rho}_{n-1}(x) &= R\tilde{\tau}_n(x) \pm iJ\tilde{\tau}_n(x) \\ &= (R \pm iJ)\tilde{\tau}_n(x) \\ &= \tilde{\tau}_n(x \pm i), \quad \text{by relation (4.3)}. \end{aligned}$$

□

7. Two Bounded Operators

In this section, we study two more operators, namely $T = R^{-1}$ and $S = JR^{-1}$, where J and R are the operators that were defined and presented in section (). It is clear from the connections in corollary (6.75) that

$$T\tilde{\rho}_n = \tilde{\tau}_n \quad (7.1)$$

$$T\tilde{\tau}_n = \tilde{\sigma}_n \quad (7.2)$$

$$S\tilde{\tau}_n = \tilde{\tau}_{n-1} \quad (7.3)$$

$$S\tilde{\rho}_n = \tilde{\rho}_{n-1} \quad (7.4)$$

Also by relation (4.3),

$$\begin{aligned} Tf(x \pm i) &= (R + iJ)Tf(x) \\ &= RTf(x) + iJTf(x) \\ &= f(x) + iSf(x). \end{aligned} \quad (7.5)$$

The integral representations of these two operators, S and J , were developed and presented in Araaya's paper [5]. For the operator T , we have

$$Tf = \frac{1}{2 \cosh \frac{\pi}{2}x} * f, \quad (7.6)$$

and for the operator S , we have

$$Sf = -\frac{1}{2 \sinh \frac{\pi}{2}x} * f, \quad (7.7)$$

where in both cases $*$ denotes convolution. Using the convolution theorem, the Fourier transforms for T and S were shown to be

$$\widehat{Tf}(t) = \operatorname{sech} t \hat{f}(t) \quad (7.8)$$

and

$$\widehat{Sf}(t) = -i \tanh t \hat{f}(t) \quad (7.9)$$

respectively. We shall also make use of what is known as the Plancherel theorem which states that $\|\hat{f}\| = \|f\|$ for any $f \in L^2(\mathbb{R})$. See [13, thm. 9.13].

Proposition 7.78 *For the operator T , we have the following:*

1. T is linear and bounded from $L^2(\omega_2)$ to $L^2(\omega_1)$.
2. If $L_0^2(\omega_1) = \{f \in L^2(\omega_1) : \langle f, 1 \rangle = 0\}$ and $H_0^2(\mathbb{S}, \mathcal{P}) = \{f \in H^2(\mathbb{S}, \mathcal{P}) : f(0) = 0\}$, then $T/\sqrt{2}$ is a unitary operator from $L_0^2(\omega_1)$ onto $H_0^2(\mathbb{S}, \mathcal{P})$.

Remark 5 *Let $f \in L^2(\omega_1)$ and $b_n = \langle f, \tilde{\tau}_n \rangle$. Then the operator $U : L^2(\omega_1) \rightarrow H^2(\mathbb{S}, \mathcal{P})$ defined by $Uf = b_0 + \frac{1}{\sqrt{2}} \sum_{n=1}^{\infty} b_n \tilde{\sigma}_n$ is unitary.*

Proof

Since all other properties are clear, we prove only boundedness.

1. Let $f \in L^2(\omega_2)$ and $a_n = \langle f, \tilde{\rho}_n \rangle$. By Theorem (5.73), the system $\{\tilde{\rho}_n\}_{n=0}^{\infty}$ is an orthogonal basis in $L^2(\omega_2)$ with norm $\sqrt{n+1}$, and so, by proposition (2.56),

$$f = \sum_{n=0}^{\infty} a_n \tilde{\rho}_n \quad \text{and} \quad \|f\|_{L^2(\omega_2)}^2 = \sum_{n=0}^{\infty} (n+1) |a_n|^2.$$

By relation (7.1),

$$Tf = \sum_{n=0}^{\infty} a_n \tilde{\tau}_n.$$

Since by Theorem (3.64) the system $\{\tilde{\tau}_n\}_{n=0}^{\infty}$ is an orthonormal basis in $L^2(\omega_1)$, we have

$$\begin{aligned} \|Tf\|_{L^2(w)}^2 &= \sum_{n=0}^{\infty} |a_n|^2 \\ &\leq \sum_{n=0}^{\infty} (n+1) |a_n|^2 \\ &= \|f\|_{L^2(\omega_2)}^2, \end{aligned}$$

which proves boundedness of T from $L^2(\omega_2)$ to $L^2(w)$ with norm 1.

2. Let $f \in L_0^2(\omega_1)$ and $b_n = \langle f, \tilde{\tau}_n \rangle$. Then $b_0 = \langle f, \tilde{\tau}_0 \rangle = \langle f, 1 \rangle = 0$, and since by Theorem (3.64) the system $\{\tilde{\tau}_n\}_{n=0}^\infty$ is an orthonormal basis in $L^2(\omega_1)$, we have

$$f = \sum_{n=1}^{\infty} b_n \tilde{\tau}_n \quad \text{and} \quad \|f\|_{L_0^2(\omega_1)}^2 = \sum_{n=1}^{\infty} |b_n|^2.$$

By relation (7.2),

$$Tf = \sum_{n=1}^{\infty} b_n \tilde{\sigma}_n.$$

Since by Theorem (4.69) the system $\{\tilde{\sigma}_n\}_{n=0}^\infty$ is an orthogonal basis in $H^2(\mathbb{S}, \mathcal{P})$ with norm 1 for $n = 0$ and $\sqrt{2}$ for $n \geq 1$, we have

$$\left\| \frac{1}{\sqrt{2}} Tf \right\|_{H_0^2(\mathcal{S}, \mathcal{P})}^2 = \sum_{n=1}^{\infty} |b_n|^2 = \|f\|_{L^2(\omega_1)}^2.$$

This proves that $T/\sqrt{2}$ is an isometry. □

Before proceeding further, we present three lemmas that will be useful in proving the main results of this section. The proof of the next lemma depends on Cauchy's theorem [2, thm. 1.4.2] which says that if two different paths connect the same two points, and a function is holomorphic everywhere in between the two paths, then the two path integrals of the function will be the same.

Lemma 7.79 *If $f \in H^2(\mathbb{S})$ then $\hat{f} \in L^2(\mathbb{R}, \cosh 2t dt)$ where \hat{f} is the Fourier transform of an analytic function f . Furthermore, $\|f\|_{H^2(\mathbb{S})}^2 = \|\hat{f}\|_{L^2(\mathbb{R}, \cosh 2t dt)}^2$.*

Proof

Recall from subsection () that analytic functions in the Hilbert space $H^2(\mathbb{S})$ have the norm

$$\|f\|_{H^2(\mathbb{S})} = \int_{-\infty}^{\infty} \frac{|f(x+i)|^2 + |f(x-i)|^2}{2} dx.$$

For $f(x+i)$, we have the Fourier transform

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x+i) e^{-ixt} dx &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x+i) e^{-i(x+i)t} e^t dx \\ &= e^t \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-ixt} dx, \quad \text{by Cauchy's theorem} \\ &= e^t \hat{f}(t). \end{aligned}$$

Similarly for $f(x-i)$,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} f(x-i) e^{-ixt} dx = e^{-t} \hat{f}(t).$$

It therefore follows by the Plancherel theorem that

$$\begin{aligned}
\|f\|_{H^2(\mathbb{S})} &= \int_{-\infty}^{\infty} \frac{|f(x+i)|^2 + |f(x-i)|^2}{2} dx \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|e^t \hat{f}(t)|^2 + |e^{-t} \hat{f}(t)|^2}{2} dt \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(t)|^2 \left(\frac{e^{2t} + e^{-2t}}{2} \right) dt \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(t)|^2 \cosh 2t dt \\
&= \|\hat{f}\|_{L^2(\mathbb{R}, \cosh 2t dt)}.
\end{aligned}$$

□

The proof of the next lemma depends on the Hadamard three-lines theorem [13, thm. 12.8] which for our particular case says that if $f \in A_0(\mathbb{S})$ and if $M(n) = \max |f(x+in)|$ then $M(0) \leq (M(-1)M(1))^{\frac{1}{2}}$.

Lemma 7.80 *If $f \in H^2(\mathbb{S})$ then $\|f\|_0 \leq (\|f\|_{+1} \cdot \|f\|_{-1})^{\frac{1}{2}}$ where we have used the notation $\|f\|_n^2 = \int_{-\infty}^{\infty} |f(x+in)|^2 dx$.*

Proof

Define the convolution $F = f * \tilde{f} \in A_0(\mathbb{S})$ where $\tilde{f}(x) = \overline{f(-x)}$, that is,

$$F(z) = \int_{-\infty}^{\infty} f(z-x) \tilde{f}(x) dx.$$

Then

$$F(0) = \int_{-\infty}^{\infty} |f(-x)|^2 dx = \|f\|_0^2.$$

Recall from subsection () that $A_0(\mathbb{S})$ is the space of functions that are analytic in \mathbb{S} , continuous on $\partial\mathbb{S}$ and $f(x+iy) \rightarrow 0$ as $|x| \rightarrow \infty$. Thus $f(x+in)$ attains a maximum, say, $F(n) = \max_{x \in \mathbb{R}} |f(x+in)|$. Then by the Hadamard three-lines theorem [13, thm. 12.8],

$$F(0) \leq (F(+1)F(-1))^{\frac{1}{2}},$$

and by the Schwarz inequality,

$$F(+1) \leq \|f\|_{+1} \cdot \|f\|_0 \quad \text{and} \quad F(-1) \leq \|f\|_{-1} \cdot \|f\|_0.$$

Therefore,

$$\begin{aligned}
\|f\|_0^2 = F(0) &\leq (F(+1)F(-1))^{\frac{1}{2}} \\
&\leq (\|f\|_{+1} \cdot \|f\|_0 \cdot \|f\|_{-1} \cdot \|f\|_0)^{\frac{1}{2}} \\
&\leq (\|f\|_{+1} \cdot \|f\|_{-1})^{\frac{1}{2}} \|f\|_0
\end{aligned}$$

so that

$$\|f\|_0 \leq (\|f\|_{+1} \cdot \|f\|_{-1})^{\frac{1}{2}}$$

□

Lemma 7.81 For all $x \geq 0$,

$$\frac{\sqrt{x^2 + 1}}{2 \cosh \frac{\pi}{2}x} \leq \frac{\pi}{2} \frac{x}{2 \sinh \frac{\pi}{2}x}$$

Proof

This is equivalent to proving the inequality

$$\tanh \frac{\pi}{2}x \leq \frac{\pi}{2} \frac{x}{\sqrt{1+x^2}}.$$

For all $x \geq 0$, define

$$f(x) = \frac{\pi}{2} \frac{x}{\sqrt{1+x^2}} - \tanh \frac{\pi}{2}x.$$

Then,

$$f(0) = 0 \quad \text{and} \quad f'(x) = \frac{\pi}{2} \left(\frac{1}{(1+x^2)^{3/2}} - \operatorname{sech}^2 \frac{\pi}{2}x \right).$$

We need to show that $f' > 0$ for all $x > 0$. This is equivalent to proving the inequality

$$\begin{aligned} \left(\frac{1}{(1+x^2)^{3/2}} - \operatorname{sech}^2 \frac{\pi}{2}x \right) &> 0 \\ \cosh^2 \frac{\pi}{2}x &> (1+x^2)^{3/2} \\ \cosh^4 \frac{\pi}{2}x &> (1+x^2)^3 \end{aligned} \tag{7.10}$$

Inequality (7.10) can be proved using Maclaurin series expansion, that is,

$$\left(\cosh \frac{\pi x}{2} \right)^4 > \left(1 + \frac{\pi^2 x^2}{8} \right)^4 > (1+x^2)^4 > (1+x^2)^3.$$

We have thus showed that $f'(x) > 0$ for all $x > 0$, and since $f(0) = 0$, it follows that for all $x \geq 0$,

$$\begin{aligned} f(x) &\geq 0 \\ \frac{\pi}{2} \frac{x}{\sqrt{1+x^2}} - \tanh \frac{\pi}{2}x &\geq 0 \\ \tanh \frac{\pi}{2}x &\leq \frac{\pi}{2} \frac{x}{\sqrt{1+x^2}} \\ \frac{\sqrt{x^2 + 1}}{2 \cosh \frac{\pi}{2}x} &\leq \frac{\pi}{2} \frac{x}{2 \sinh \frac{\pi}{2}x}. \end{aligned}$$

□

We now have all the necessary definitions and lemmas needed to present and prove the main results of this section. In fact, they are the final results for this project thesis, and they will be presented in two separate theorems, one for the operator T and the other for the operator S .

Theorem 7.82 *The operator S is linear and bounded on the following Hilbert spaces with norm 1:*

1. $L^2(\omega_2)$
2. $L^2(\omega_1)$
3. $L^2(\mathbb{R})$
4. $H^2(\mathbb{S}, \mathcal{P})$
5. $H^2(\mathbb{S})$

Proof

Since linearity follows immediately from the fact that S is a convolution, we shall prove only boundedness.

1. Let $\tilde{\rho}_n = \frac{\tilde{\rho}_n}{\sqrt{n+1}}$, $f \in L^2(\omega_2)$ and $a_n = \langle f, \tilde{\rho}_n \rangle$. Since by equation (5.10) $\sqrt{n+1}$ is the norm of the polynomial $\tilde{\rho}_n$ in the Hilbert space $L^2(\omega_2)$, it follows from Theorem (5.73) that $\{\tilde{\rho}_n\}_{n=0}^{\infty}$ is an orthonormal basis in $L^2(\omega_2)$. Thus by proposition (2.56),

$$f = \sum_{n=0}^{\infty} a_n \tilde{\rho}_n \quad \text{and} \quad \|f\|^2 = \sum_{n=0}^{\infty} |a_n|^2.$$

By relation (7.4),

$$\begin{aligned} Sf &= \sum_{n=1}^{\infty} a_n \left(\frac{\tilde{\rho}_{n-1}}{\sqrt{n+1}} \right) \\ &= \sum_{n=0}^{\infty} a_{n+1} \left(\frac{\tilde{\rho}_n}{\sqrt{n+2}} \right) \\ &= \sum_{n=0}^{\infty} a_{n+1} \left(\sqrt{\frac{n+1}{n+2}} \right) \left(\frac{\tilde{\rho}_n}{\sqrt{n+1}} \right) \\ &= \sum_{n=0}^{\infty} \left(\sqrt{\frac{n+1}{n+2}} \right) a_{n+1} \tilde{\rho}_n. \end{aligned}$$

Thus,

$$\|Sf\|^2 = \sum_{n=0}^{\infty} \left(\frac{n+1}{n+2} \right) |a_{n+1}|^2 = \sum_{n=1}^{\infty} \left(\frac{n}{n+1} \right) |a_n|^2 \leq \sum_{n=1}^{\infty} |a_n|^2 \leq \|f\|^2,$$

which proves boundedness of S on $L^2(\omega_2)$ with norm 1.

2. Let $f \in L^2(\omega_1)$ and $b_n = \langle f, \tilde{\tau}_n \rangle$. By Theorem (3.64), the system $\{\tilde{\tau}_n\}_{n=0}^{\infty}$ is an orthonormal basis in $L^2(\omega_1)$ so that by proposition (2.56),

$$f = \sum_{n=0}^{\infty} b_n \tilde{\tau}_n \quad \text{and} \quad \|f\|^2 = \sum_{n=0}^{\infty} |b_n|^2.$$

By relation (7.3),

$$Sf = \sum_{n=1}^{\infty} b_n \tilde{\tau}_{n-1} = \sum_{n=0}^{\infty} b_{n+1} \tilde{\tau}_n.$$

Thus,

$$\|Sf\|^2 = \sum_{n=0}^{\infty} |b_{n+1}|^2 = \sum_{n=1}^{\infty} |b_n|^2 \leq \sum_{n=0}^{\infty} |b_n|^2 = \|f\|^2,$$

which boundedness of S on $L^2(\omega_1)$ with norm 1.

3. Let $f \in L^2(\mathbb{R})$. Using the Plancherel theorem, we have

$$\begin{aligned} \|Sf\|^2 &= \|\widehat{Sf}\|^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\widehat{Sf}(t)|^2 dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} |\tanh t \hat{f}(t)|^2 dt, \quad \text{by (7.9)} \\ &\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(t)|^2 dt, \quad \text{since } |\tanh t| \leq 1 \\ &= \|f\|^2 \end{aligned}$$

which proves boundedness of S on $L^2(\mathbb{R})$ with norm 1.

4. Let $\tilde{\sigma}_n = 1$ if $n = 0$, and $\tilde{\sigma}_n = \frac{\tilde{\sigma}_n}{\sqrt{2}}$ for all $n \geq 1$. Then by Theorem (4.69), the system $\{\tilde{\sigma}_n\}_{n=0}^{\infty}$ is an orthonormal basis in $H^2(\mathbb{S}, \mathcal{P})$. Let $f \in H^2(\mathbb{S}, \mathcal{P})$ and $c_n = \langle f, \tilde{\sigma}_n \rangle$. By proposition (2.56),

$$f = \sum_{n=0}^{\infty} c_n \tilde{\sigma}_n \quad \text{and} \quad \|f\|^2 = \sum_{n=0}^{\infty} |c_n|^2.$$

Since $RJ = JR$ by proposition (4.68), it follows that $S = JR^{-1} = R^{-1}J$. Thus $S\tilde{\sigma}_n = \tilde{\sigma}_{n-1}$ by corollary (6.75) so that

$$\begin{aligned} Sf &= \sum_{n=1}^{\infty} c_n \left(\frac{\tilde{\sigma}_{n-1}}{\sqrt{2}} \right) \\ &= \sum_{n=0}^{\infty} c_{n+1} \left(\frac{\tilde{\sigma}_n}{\sqrt{2}} \right) \\ &= \frac{c_1}{\sqrt{2}} + \sum_{n=1}^{\infty} c_{n+1} \tilde{\sigma}_n. \end{aligned}$$

Thus,

$$\|Sf\|^2 = \frac{|c_1|^2}{2} + \sum_{n=1}^{\infty} |c_{n+1}|^2 = \frac{|c_1|^2}{2} + \sum_{n=2}^{\infty} |c_n|^2 \leq \sum_{n=1}^{\infty} |c_n|^2 \leq \|f\|^2,$$

which proves boundedness of S on $H^2(\mathbb{S}, \mathcal{P})$ with norm 1.

5. Let $f \in H^2(\mathbb{S})$. Then by Lemma (7.79),

$$\begin{aligned}
\|Sf\|_{H^2(\mathbb{S})}^2 &= \|\widehat{Sf}\|_{L^2(\mathbb{R}, \cosh 2t dt)}^2 \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} |\widehat{Sf}(t)|^2 \cosh 2t dt \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} |\tanh t \hat{f}(t)|^2 \cosh 2t dt, \quad \text{by (7.9)} \\
&\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(t)|^2 \cosh 2t dt, \quad \text{since } |\tanh t| \leq 1 \\
&= \|\hat{f}\|_{L^2(\mathbb{R}, \cosh 2t dt)}^2 \\
&= \|f\|_{H^2(\mathbb{S})}^2,
\end{aligned}$$

which proves boundedness of S on $H^2(\mathbb{S})$ with norm 1.

□

Theorem 7.83 *The operator T is linear and bounded on the following Hilbert spaces:*

1. $L^2(\omega_2)$ with norm less than or equal to $\sqrt{\pi}$.
2. $L^2(\omega_1)$ with norm less than or equal to 2.
3. $L^2(\mathbb{R})$ with norm 1.
4. $H^2(\mathbb{S})$ with norm 1.

Proof

Since linearity follows immediately from the fact that S is a convolution, we shall prove only boundedness.

1. We first show that if $f \in L^2_{\mathbb{R}}(\omega_2)$ and $\psi = \sqrt{\omega_2}$, then $Tf\psi \in H^2(\mathbb{S})$. In particular, we show that

$$\|Tf\psi\|_{H^2(\mathbb{S})}^2 = \int_{-\infty}^{\infty} \frac{|Tf(x+i)\psi(x+i)|^2 + |Tf(x-i)\psi(x-i)|^2}{2} dx$$

is finite. Now,

$$|\psi(x \pm i)|^2 = \left| \frac{x \pm i}{2 \sinh \frac{\pi}{2}(x \pm i)} \right| = \left| \frac{x \pm i}{\pm i 2 \cosh \frac{\pi}{2}x} \right| = \frac{\sqrt{x^2 + 1}}{2 \cosh \frac{\pi}{2}x}, \quad (7.11)$$

and by relation (7.5),

$$|Tf(x \pm i)|^2 = |f(x) + iSf(x)|^2 = |f(x)|^2 + |Sf(x)|^2. \quad (7.12)$$

Thus by (7.11) and (7.12),

$$|Tf(x+i)\psi(x+i)|^2 = |Tf(x-i)\psi(x-i)|^2 \quad (7.13)$$

$$= (|f(x)|^2 + |Sf(x)|^2) \frac{\sqrt{x^2+1}}{2 \cosh \frac{\pi}{2}x}. \quad (7.14)$$

Therefore,

$$\begin{aligned} \|Tf\psi\|_{H^2(\mathbb{S})}^2 &= \int_{-\infty}^{\infty} \frac{|Tf(x+i)\psi(x+i)|^2 + |Tf(x-i)\psi(x-i)|^2}{2} dx \\ &= \int_{-\infty}^{\infty} |Tf(x+i)\psi(x+i)|^2 dx, \quad \text{by (7.13)} \\ &= \int_{-\infty}^{\infty} (|f(x)|^2 + |Sf(x)|^2) \frac{\sqrt{x^2+1}}{2 \cosh \frac{\pi}{2}x} dx, \quad \text{by (7.14)} \\ &\leq \frac{\pi}{2} \int_{-\infty}^{\infty} (|f(x)|^2 + |Sf(x)|^2) \frac{x}{2 \sinh \frac{\pi}{2}x} dx, \quad \text{by lemma (7.81)} \\ &= \frac{\pi}{2} \int_{-\infty}^{\infty} |f(x)|^2 \omega_2 dx + \frac{\pi}{2} \int_{-\infty}^{\infty} |Sf(x)|^2 \omega_2 dx \\ &= \frac{\pi}{2} \|f\|_{L_{\mathbb{R}}^2(\omega_2)}^2 + \frac{\pi}{2} \|Sf\|_{L_{\mathbb{R}}^2(\omega_2)}^2 \\ &\leq \frac{\pi}{2} \|f\|_{L_{\mathbb{R}}^2(\omega_2)}^2 + \frac{\pi}{2} \|f\|_{L_{\mathbb{R}}^2(\omega_2)}^2, \quad \text{by Theorem (7.82) part(1)} \\ &= \pi \|f\|_{L_{\mathbb{R}}^2(\omega_2)}^2, \end{aligned} \quad (7.15)$$

which proves that $Tf\psi \in H^2(\mathbb{S})$ for any $f \in L_{\mathbb{R}}^2(\omega_2)$.

Next, we show that T is bounded on $L_{\mathbb{R}}^2(\omega_2)$. Let $f \in L_{\mathbb{R}}^2(\omega_2)$, then

$$\begin{aligned} \|Tf\|_{L_{\mathbb{R}}^2(\omega_2)}^2 &= \int_{-\infty}^{\infty} |Tf(x)|^2 \omega_2(x) dx \\ &= \int_{-\infty}^{\infty} |Tf(x)\sqrt{\omega_2(x)}|^2 dx \\ &= \int_{-\infty}^{\infty} |Tf(x)\psi(x)|^2 dx \\ &\leq \left(\int_{-\infty}^{\infty} |Tf(x+i)\psi(x+i)|^2 dx \right)^{\frac{1}{2}} \left(\int_{-\infty}^{\infty} |Tf(x-i)\psi(x-i)|^2 dx \right)^{\frac{1}{2}} \\ &= \int_{-\infty}^{\infty} |Tf(x+i)\psi(x+i)|^2 dx, \quad \text{by (7.13)} \\ &\leq \pi \|f\|_{L_{\mathbb{R}}^2(\omega_2)}^2, \quad \text{by (7.15)} \end{aligned}$$

where the first inequality follows by Lemma (7.80) since $Tf\psi \in H^2(\mathbb{S})$.

We have shown that T is bounded on $L^2_{\mathbb{R}}(\omega_2)$ with norm $\sqrt{\pi}$. Now since every analytic function g can be written as $g = f + ih$ for $f, h \in L^2_{\mathbb{R}}(\omega_2)$, it follows that T is bounded on $L(\omega_2)$ with norm $\sqrt{\pi}$.

Remark 6 *Since T and S map functions that are real on the real line to functions that also have this property and therefore $T(f + ih) = Tf + iT h$, the same bounds hold for complex-valued functions as for real-valued.*

2. We first show that if $f \in L^2_{\mathbb{R}}(\omega_1)$ and $\varphi = 1/(2 \cosh \frac{\pi}{4}x)$, then $Tf\varphi \in H^2(\mathbb{S})$. In particular, we show that

$$\|Tf\varphi\|_{H^2(\mathbb{S})}^2 = \int_{-\infty}^{\infty} \frac{|Tf(x+i)\varphi(x+i)|^2 + |Tf(x-i)\varphi(x-i)|^2}{2} dx$$

is finite. Now,

$$\begin{aligned} \left| 2 \cosh \frac{\pi}{4}(x \pm i) \right|^2 &= \left| \sqrt{2} \cosh \frac{\pi}{4}x \pm i\sqrt{2} \sinh \frac{\pi}{4}x \right|^2 \\ &= \left(\sqrt{2} \cosh \frac{\pi}{4}x \right)^2 + \left(\sqrt{2} \sinh \frac{\pi}{4}x \right)^2 \\ &= 2 \cosh \frac{\pi}{2}x, \end{aligned}$$

and so,

$$|\varphi(x \pm i)|^2 = \frac{1}{2 \cosh \frac{\pi}{2}x}. \quad (7.16)$$

By relation (7.5),

$$|Tf(x \pm i)|^2 = |f(x) + iSf(x)|^2 = |f(x)|^2 + |Sf(x)|^2. \quad (7.17)$$

Thus by (7.16) and (7.17),

$$|Tf(x+i)\varphi(x+i)|^2 = |Tf(x-i)\varphi(x-i)|^2 \quad (7.18)$$

$$= (|f(x)|^2 + |Sf(x)|^2) \frac{1}{2 \cosh \frac{\pi}{2}x}. \quad (7.19)$$

Therefore,

$$\begin{aligned}
\|Tf\varphi\|_{H^2(\mathbb{S})}^2 &= \int_{-\infty}^{\infty} \frac{|Tf(x+i)\varphi(x+i)|^2 + |Tf(x-i)\varphi(x-i)|^2}{2} dx \\
&= \int_{-\infty}^{\infty} |Tf(x+i)\varphi(x+i)|^2 dx, \text{ by (7.18)} \\
&= \int_{-\infty}^{\infty} (|f(x)|^2 + |Sf(x)|^2) \frac{dx}{2 \cosh \frac{\pi}{2}x}, \text{ by (7.19)} \\
&= \int_{-\infty}^{\infty} |f(x)|^2 \omega_1 dx + \int_{-\infty}^{\infty} |Sf(x)|^2 \omega_1 dx \\
&= \|f\|_{L_{\mathbb{R}}^2(\omega_1)}^2 + \|Sf\|_{L_{\mathbb{R}}^2(\omega_1)}^2 \\
&\leq \|f\|_{L_{\mathbb{R}}^2(\omega_1)}^2 + \|f\|_{L_{\mathbb{R}}^2(\omega_1)}^2, \text{ by Theorem (7.82) part(2)} \\
&= 2\|f\|_{L_{\mathbb{R}}^2(\omega_1)}^2, \tag{7.20}
\end{aligned}$$

which proves that $Tf\varphi \in H^2(\mathbb{S})$ for any $f \in L_{\mathbb{R}}^2(\omega_1)$.

Next, we show that T is bounded on $L_{\mathbb{R}}^2(\omega_1)$. Let $f \in L_{\mathbb{R}}^2(\omega_1)$, then

$$\begin{aligned}
\int_{-\infty}^{\infty} |Tf(x)\psi(x)|^2 dx &= \int_{-\infty}^{\infty} \frac{|Tf(x)|^2}{(2 \cosh \frac{\pi}{4}x)^2} dx \\
&= \int_{-\infty}^{\infty} \frac{|Tf(x)|^2}{2 (\cosh \frac{\pi}{2}x + 1)} dx \\
&\geq \int_{-\infty}^{\infty} \frac{|Tf(x)|^2}{2 (\cosh \frac{\pi}{2}x + \cosh \frac{\pi}{2}x)} dx, \text{ since } \cosh \frac{\pi}{2}x \geq 1 \\
&= \int_{-\infty}^{\infty} \frac{|Tf(x)|^2}{4 \cosh \frac{\pi}{2}x} dx \\
&= \frac{1}{2} \int_{-\infty}^{\infty} |Tf(x)|^2 \omega_1(x) dx
\end{aligned}$$

so that

$$\begin{aligned}
\|Tf\|_{L^2_{\mathbb{R}}(\omega_1)}^2 &= \int_{-\infty}^{\infty} |Tf(x)|^2 \omega_1(x) dx \\
&\leq 2 \int_{-\infty}^{\infty} |Tf(x)\varphi(x)|^2 dx \\
&\leq 2 \left(\int_{-\infty}^{\infty} |Tf(x+i)\varphi(x+i)|^2 dx \right)^{\frac{1}{2}} \left(\int_{-\infty}^{\infty} |Tf(x-i)\varphi(x-i)|^2 dx \right)^{\frac{1}{2}} \\
&= 2 \int_{-\infty}^{\infty} |Tf(x+i)\varphi(x+i)|^2 dx, \text{ by (7.18)} \\
&\leq 4\|f\|_{L^2_{\mathbb{R}}(\omega_1)}^2, \text{ by (7.20)}
\end{aligned}$$

where the second inequality follows by the Lemma (7.80) since $Tf\varphi \in H^2(\mathbb{S})$.

We have shown that T is bounded on $L^2_{\mathbb{R}}(\omega_1)$ with norm 2. Now since every analytic function g can be written as $g = f + ih$ for $f, h \in L^2_{\mathbb{R}}(\omega_1)$, it follows by Remark (6) that T is bounded on $L(\omega_1)$ with norm 2.

3. Let $f \in L^2(\mathbb{R})$. Using the Plancherel theorem, we have

$$\begin{aligned}
\|Tf\|^2 &= \|\widehat{Tf}\|^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\widehat{Tf}(t)|^2 dt \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} |\operatorname{sech} t \hat{f}(t)|^2 dt, \text{ by (7.8)} \\
&\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(t)|^2 dt, \text{ since } |\operatorname{sech} t| \leq 1 \\
&= \|f\|^2
\end{aligned}$$

which proves boundedness of T on $L^2(\mathbb{R})$ with norm 1.

4. Let $f \in H^2(\mathbb{S})$. Then by Lemma (7.79),

$$\begin{aligned}
\|Tf\|_{H^2(\mathbb{S})}^2 &= \|\widehat{Tf}\|_{L^2(\mathbb{R}, \cosh 2t \, dt)}^2 \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} |\widehat{Tf}(t)|^2 \cosh 2t \, dt \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} |\operatorname{sech} t \hat{f}(t)|^2 \cosh 2t \, dt, \quad \text{by (7.8)} \\
&\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(t)|^2 \cosh 2t \, dt, \quad \text{since } |\operatorname{sech} t| \leq 1 \\
&= \|\hat{f}\|_{L^2(\mathbb{R}, \cosh 2t \, dt)}^2 \\
&= \|f\|_{H^2(\mathbb{S})}^2,
\end{aligned}$$

which proves boundedness of T on $H^2(\mathbb{S})$ with norm 1.

□

References

- [1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover Publications, Inc., New York, 1965.
- [2] L. V. Ahlfors. *Complex Analysis: An Introduction to the Theory of Analytic Functions of One Complex Variable*. 3rd ed., Mc-Graw-Hill, 1979.
- [3] G. E. Andrews, R. Askey and R. Roy. *Special Functions*. Cambridge University Press, Cambridge, 2000.
- [4] T. K. Araaya. *The Meixner-Pollaczek Polynomials and a System of Orthogonal Polynomials in a Strip*. Uppsala, 1999.
- [5] T. K. Araaya. *Umbral Calculus and the Meixner-Pollaczek Polynomials*. Uppsala, 2002.
- [6] K. N. Chaudhury. *L^p -Boundedness of the Hilbert Transform*. arXiv:0909.1426v8 [cs.IT], 2010.
- [7] L. Holst. *Probabilistically Proving Some Identities of Euler*. Department of Mathematics, KTH, SE-10044 Stockholm, Sweden.
- [8] E. Koelink. *Spectral Theory and Special Functions*. arXiv:math/0107036v1 [math.CA], 2001.
- [9] T. W. Korner. *A Companion to Analysis: A Second First and First Second Course in Analysis*. American Mathematical Society, 2004.
- [10] E. Laeng. *A Simple Real-Variable Proof that the Hilbert Transform is an L^2 -Isometry*. C. R. Acad. Sci. Paris, Ser. I 348 (2010) 977980.
- [11] M. Reed and B. Simon. *Methods of Modern Mathematical Physics. Vol. 1 Functional Analysis*. Academic Press Inc., New York, 1980.

- [12] W. Rudin. *Principles of Mathematical Analysis*. Third edition, McGraw-Hill International Editions, 1976.
- [13] W. Rudin. *Real and Complex Analysis*. International Student Edition, McGraw-Hill, New York, 1970.
- [14] J. Shen. *Orthogonal Polynomials and Polynomial Approximations: Chapter 3*. Department of Mathematics, Purdue University, 2009.
- [15] G. F. Simmons and S. G. Krantz. *Differential Equations: Theory, Technique, and Practice*. McGraw-Hill International Edition, New York, 2007.
- [16] E. M. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press, 1971, pp. 205-209.
- [17] G. Szego. *Orthogonal Polynomials*. Amer. Math. Soc. Colloq. Pubb. 23, Fourth Edition, 1975, pp 395.
- [18] V. Totik. *Orthogonal Polynomials*. Surveys in Approximation Theory, Volume 1, 2005. pp. 70125.

ABSTRACTS

Asymptotic Properties of The Delannoy Numbers and Similar Arrays

by

Christer Kiselman

kiselman@math.uu.se

Uppsala University, Department of Mathematics, Uppsala, Sweden

Abstract: Abstract: The Delannoy numbers were introduced and studied by Henri-Auguste Delannoy (1833-1915). He investigated the possible moves on a chessboard: the numbers under consideration appear when one studies “la marche de la Reine,” i.e., how the queen moves (the binomial coefficients appear similarly for the moves of the rook). The asymptotic behavior of the array of Delannoy numbers is studied. The regularized upper and lower radial indicators of the array are determined, proved to coincide, and to be concave. We also describe the radial indicator as an infimum of linear functions, which amounts to determining its Fenchel transform. Since the methods developed for this study apply to more general convolution equations, we prove results also for these equations.

The (Dis)connectedness of Products in the Box Topology

by

Vitalij A. Chatyrko

Department of Mathematics, Linköping University, Linköping, Sweden

Abstract: In this talk we suggest two independent sufficient conditions on topological connected spaces which imply disconnectedness, and one sufficient condition which implies connectedness, of products of spaces endowed with the box topology. Some applications of that will be also presented.

Identification of Coefficients in Parabolic Equations Using Measurements on the Boundary

by

Frerik Berntsson

Linköping University, Sweden

Abstract: To determine the thermal conductivity in the interior of a body using measurements on the boundary is an important problem. Applications arise, e.g., in methods for non-destructive testing of adhesive bonds, or crack detection in metallic materials. In our application we measure the time-dependent temperature and heat-flux at certain locations on the boundary, or inside the domain, and attempt to reconstruct the thermal conductivity as accurately as possible. The coefficient identification problem is severely ill-posed in the sense that small changes in the measured data can lead to large changes in the computed solution. The ill-posedness is analyzed using the singular value decomposition. Also, the recorded data may not contain any information about the unknown coefficient. We propose to formulate the coefficient identification problem as a non-linear least squares problem. The problem can be solved using the Gauss-Newton method. The dimension of the least squares problem is reduced by

modelling the unknown coefficient using only a small number of parameters. Numerical tests show that the method works well.

Multidisciplinary Research in Mathematical Sciences With Applications to Real World Problems in Biological, Bio-Inspired and Engineering Systems

by

Padmanabhan Seshaiyer

George Mason University, USA Email: pseshaiy@gmu.edu

Abstract: Computational mathematics, which comprises of mathematical modeling, analysis and simulation, is quickly becoming the foundation for solving most complex applications in biological, bio-inspired and engineering systems. Breakthrough research required to solve such complex problems involves transformative and multidisciplinary efforts spanning scientific and engineering disciplines. In this talk, we will describe research methods using mathematical modeling, analysis and computational techniques to explain fundamental mechanisms needed to understand the quantitative behavior underlying real-world applications. Specific examples of interdisciplinary projects involving analytical and numerical solutions to partial differential equations that can help to encourage students to learn by discovery to enhance their understanding of the multidisciplinary role of mathematics in engineering, science and medicine will be presented.

Epidemic Potential for Malaria in Epidemiological Zones in Kenya

by

Wandera Ogana

School of Mathematics, University of Nairobi, Kenya

Abstract: Malaria is a vector-borne disease which annually results in over one million deaths and five hundred million clinical episodes, most of which occur in sub-Saharan Africa. Since the disease is influenced by climate factors, it is important to assess the possible risk posed by climate change on malaria transmission. A number of indices can be used to assess this risk but the most appropriate appears to be the epidemic potential, which is derived from the basic reproduction number, R_0 . We determine the epidemic potential for selected areas within the four epidemiological zones in Kenya, using modeled temperature and rainfall data. For the years 2009 to 2011, for which detailed malaria data is available, we compare the variation in epidemic potential with malaria incidence. Results show that the variation in epidemic potential, from month to month, reflects a pattern similar to the variation in malaria incidence.

How to Manipulate Derangements

by

Fanja Rakotonandrajao,

University of Antananarivo, Madagascar

Abstract: We are living in a period of change or/and of conflict: for example, the climate change, the trend to electronic systems, and so on. All these changes will also change our way of life. In

combinatorial word, we say that there are derangements. But "derangements are like a rose bush, when you touch them; you need to be careful of thorns." We will give different methods how to manipulate derangements combinatorially and how to analyze them.

The Dynamics of Populations in Wetlands

by

Abdou Sene

Gaston Berger University

Abstract: Some wetlands are centers of important migrations of populations of various species. These interdependent migrations, caused by economic and environmental factors, are governed by Mathematical nonlinear complex systems. This work consists in developing and analysing a Mathematical model of interconnection among:

- The quantity and quality of the water available in the wetland
- The size of the population of fish in the waterways
- The dynamics of migrating birds population
- The size of the tourists population
- The dynamics of human populations going to or coming from the surrounding towns or villages.

Adaptive Markov Chain Monte Carlo Using Variational Bayesian Adaptive Kalman Filter

by

Isambi Sailon Mbalawata and Simo Sarkka

Lappeenranta, Finland

Abstract: When we analyze the high dimensional complex models, the computation of the normalized posterior density and its expectations are intractable hence we employ the numerical approximation techniques such as Markov chain Monte Carlo methods. Markov chain Monte Carlo (MCMC) method is a powerful computational tool for analysis of complex statistical problems; it requires proper tuning of proposal distribution for better mixing of chains and suitable acceptance rate. However, in practice the selection of a proper proposal distribution is not a trivial task because manual tuning of proposal distribution is time consuming and laborious. The most used proposal distribution is the Gaussian distribution, due to its attractive computational and theoretical properties. One problem of Gaussian proposal distribution is how to find a suitable covariance matrix. One way to overcome this problem is to use the adaptive Markov chain Monte Carlo algorithm. The algorithm automatically tunes the covariance matrix during MCMC run. In this work, we propose a new adaptive MCMC algorithm where the covariance matrix is adapted using variational Bayesian adaptive Kalman filter. Numerical results for simulated examples are presented and discussed in detail.

Linear Estimation of Location and Scale Parameters for Logistic Distribution Based on Consecutive Order Statistics

by

Patrick G. O. Weke

School of Mathematics University of Nairobi, Kenya

Abstract: Linear estimation of the scale parameter of the logistic population based on the sum of consecutive order statistics when the location parameter is unknown is discussed. A method based on a pair of single spacing and the 'zero-one' weights rather than the optimum weights is presented and used to compute the bias, variance and relative efficiencies with respect to variance Cramer-Rao lower bound and best linear unbiased estimators (BLUE's) for sample size n . Finally, a comparison of these estimators is discussed.

A Stochastic Model for Planning a Compartmental Education System and Supply of Manpower

by

Lydia Musiga

University of Nairobi, Nairobi, Kenya

Abstract: This paper describes a Markov Chain transition model that encompasses the different compartments of an education system. The model clearly shows transition rates within compartments and also between compartments, thus planners in the country will understand better the flow of students from primary school to university level. Also, the theory of Absorbing Markov Chains, specifically the Chapman-Kolmogorov result, assists in predicting future enrolments. Hence, the model will facilitate more effective planning in the country's education system and in the supply of manpower.

Financial Sector Performance Enhancers

by

Emma Anyika and Patrick Weke

University of Nairobi, Nairobi, Kenya

Abstract: In any state or country there are certain sectors that are relied upon to drive its economy. For many of these countries the financial sector is seen as the driving force of the economy. This is witnessed in many World economic crises which commence with the large organisations in the financial sector. This should aid entrepreneurs to be aware of the areas of emphasis and factors for consideration for positive growth of their organisations. Existing organisations will also benefit by improving the said areas and adopting the factors for continued growth and sustainability. A multiple regression model will be used to relate performance to its causes. Tests of hypothesis will then be made to allow for the generalization of the findings to the whole population

Estimating the List Size Using Bipartite Graph for Colouring Problems

by

Mashaka Mkandawile
University of Dar es Salaam

Abstract: In graph theory a bipartite graph is a graph whose vertices can be divided into two disjoint sets U and V such that every edge connects a vertex in U to one in V ; that is, U and V are independent sets. Bipartite graphs show up in many places and are therefore often used tool to model and calculate with. Let $B(n, m)$ be a bipartite graph with n vertices in each side and m edges. For each vertex we draw uniformly at random a list of size k from a base set N of size n . In this paper we estimate the sizes of n and k so that $B(n, m)$ has a perfect matching with high probability.

Mathematical Modeling of Pneumonia Transmission Dynamics

by

Emaline Joseph, Kgosimore and Teresia Marijani

Department of Mathematics, University of Dar-es-Salaam Tanzania . **Abstract:** Pneumonia is one of the leading causes of serious illness and deaths among children and adults around the world. We therefore, formulate and analyze a mathematical model of the transmission dynamics of pneumonia with the aim of understanding its transmission dynamics. The study also evaluates the impact of control and prevention strategies in curtailing or mitigating against the spread of disease. We derive conditions for the clearance or persistence of the pneumonia infection through the stability of the equilibria. We infer the impact of control strategies on the dynamics of the disease through sensitivity analysis of the reproduction number, R_0 . Numerical simulations are carried out to illustrate the analytical results and test the influence of certain parameters. The result of study showed that treatment and vaccination have the effect of reducing the disease provided that the control reproduction number is reduced to a value below one

Hydrodynamics of Shallow Water Equations: A Case Study of Lake Victoria

by

David Ddumba Walakira
Mathematics Department, Makerere University

Abstract: A three dimensional hydrodynamic model has been applied to lake Victoria. We have assumed a vertical coriolis dominance because of the geographical location of lake Victoria at the equator, leading to a non-hydrostatic approximation. To capture not only the discontinuities, but also the physical and numerical shocks in the shallow water flow of the second largest fresh water body in the world, a min-mod flux limiter has been employed as a numerical finite volume high resolution method to solve a 3D model with Boussines and shallow water approximations. An energy method is applied for the well-posedness of this non-linear hyperbolic system of equations.

Boundary Layer Flow Over a Moving at Surface With Temperature Dependent Viscosity

by

J.W. MwaonANJI

Mathematics and Statistics Dept, The Polytechnic, University of Malawi, Malawi

Abstract: Numerical investigation of two-dimensional laminar boundary layer flow over a moving flat surface with temperature dependent viscosity is studied. The flow is divided into two regimes: with viscous dissipation and without viscous dissipation. The flow is also restricted to a region where both the free-stream and the flat surface are moving in the same direction i.e no reverse flow within the boundary layer. Thus the velocity ratio, ξ , has generally been chosen in such a way that $0 < \xi < 1$. The governing boundary layer equations of the flow are transformed to a dimensionless system of equations using a similarity variable $\zeta(x, y)$. The resulting set of coupled non-linear ordinary differential equations are solved numerically by applying shooting iteration technique together with fourth-order Runge-Kutta integration scheme. The effects of the various parameters of the flow e.g velocity variation variable, ξ , viscosity variation parameter, ϵ , etc. on velocity and temperature distribution in the boundary layer, on the local skin friction and local heat transfer coefficients are investigated.

Absolute-Convective Instability of Mixed Forced-Free Convection Boundary Layers

by

Eunice Mureithi

Mathematics Department, University of Dar es salaam

Abstract: A spatio-temporal inviscid instability of a mixed forced-free convection boundary layer is investigated. The base flow considered is the self-similar flow with free-stream velocity $u \sim x^n$. Such a boundary layer flow presents the unusual behaviour of generating a region of velocity overshoot, in which the stream-wise velocity within the boundary layer exceeds the free-stream speed. A linear stability analysis has been carried out. Saddle points have been located and a critical value for the buoyancy parameter, $G_0 = G_{0C} \approx 3.6896$, has been determined below which the flow is convectively unstable and above which the flow becomes absolutely unstable. Two families of spatial modes have been obtained, one family being of convective nature and the other of absolute nature. The convective type spatial mode shows mode crossing behaviour at lower frequencies. Thermal buoyancy has been shown to be destabilizing to absolutely unstable spatial mode.

Optimal Premium Policy of an Insurance Firm With Delay

by

Moses Mwale

University of Dar es Salaam, Tanzania

Abstract: In this work, we study the optimization problem confronted by an insurance firm whose management can control its cash-balance dynamics by adjusting the underlying premium rate. The firm's objective is to minimize the total deviation of its cash-balance process to some pre-set target levels by selecting an appropriate premium policy. We make two inclusions to the problem; Firstly, we introduce the aspect of time delay to the system. Delay systems may occur in several situations, e.g. in finance and biology where the growth of the state depends not only on the current value of the state but also

on previous state values. Stochastic delay differential equations (SDDE's) model systems with delay. The second inclusion is that we replace the standard expected additive utility function with a Stochastic differential utility (SDU). A SDU is an extension of the notion of recursive utility to a continuous-time, stochastic setting. It allows us to disentangle risk aversion and inter temporal substitution. The paper is devoted to the study of optimal control of stochastic differential delay equations in a firstly in general framework. The Martingale Representation Theorem is applied to obtain a Backward Stochastic Differential Equation (BSDE) which represents the utility function. The resulting system is a Forward-Backward Stochastic Differential Equation which is not fully coupled and we also assume that it is quadratic in the cash-balance and control variables. We establish existence and prove for uniqueness of the solution to our FBSDE. We then also establish sufficient and necessary maximum principles for an optimal control of such systems. We end with a case study of two particular works which fit into our general model.

On the Coexistence of Distributional and Rational Solutions for Ordinary Differential Equations With Polynomial Coefficients

by

G.I. Mirumbe, Vincent SSembatya, Rikard Bgvad and Jan Erik Bjork

Abstract: Given an ordinary differential equation with polynomial coefficients, Wiener & Cooke (1990) gave a necessary and sufficient condition for the simultaneous existence of solutions to ordinary differential equations with polynomial coefficients in the form of finite order linear combination of the Dirac-delta function and its derivatives and the rational function solutions using the Laplace transform and functional differential equations techniques. In this paper, we prove a similar result using the theory of boundary values and the Cauchy transform. This method has an advantage as it gives a closed form expression for the polynomial $q(t)$ in case the finite order distributional solution and the rational function solution do not satisfy the same differential equation but in different variables

On Modelling and Pricing Index Linked Catastrophe Derivatives

by

Philip Ngare

School of Mathematics, University of Nairobi, Kenya

Abstract: We consider the problem of indifference pricing of derivatives written on CAT bonds. The industrial loss index is modeled by a compound Poisson process and the number of claims as doubly stochastic process, such that its intensity varies over time. The insurer can adjust her portfolio by choosing the risk loading, which in turn determines the demand. We probably restrict the policies of the insurance company in a way that does not permit changing the risk loading during catastrophe times. We compute the price of a CAT option written on that index using utility indifference pricing.

Optimal Portfolio Management When Stocks are Driven by Mean Reverting Processes

by

Lusungu Mbiliri, Charles Mahera, and Sure Mataramvura

Abstract: In this paper we present and solve the problem of portfolio optimization within the context of continuous-time stochastic model of financial variables. We consider an investment problem where an investor has two assets, namely, risk-free assets (eg bonds) and risky assets (eg stocks) to invest on and tries to maximize the expected utility of the wealth at some future time . The evolution of the risk-free asset is described deterministically while the dynamics of the risky asset is described by the geometric mean reversion (GMR) model. The controlled wealth stochastic differential equations (SDE) as well as the portfolio problem are formulated. Therefore the portfolio optimization problem is then successfully formulated and solved with the help of the theory of stochastic control technique where the Dynamic programming principle (DPP) and the HJB theory are the perfect tools. We obtain the very interesting results including the solution of the HJB equation which is the non-linear second order partial differential equation and the optimal policy which is the optimal control strategy of the investment process. So we have considered utility functions which are members of HARA, called power and exponential utility. In both cases, the optimal control (investment strategy) has explicit forms and is wealth dependant, in the sense that, as the investor becomes richer, the less he invests on the risky assets.

On Hub Number of Hypercube and Grid Graphs

by

Egbert Mujuni

Mathematics Department, University of Dar es Salaam, Tanzania

Abstract: A set $H \subset V$ is a hub set of a graph $G = (V, E)$ if, for every pair of vertices $u, v \in V - H$, there exists a path from u to v such that all intermediate vertices are in H . The hub number of G is the minimum size of a hub set in G . In this talk we derive the hub numbers of hypercube and grid graphs. Meanwhile, new results on the size of maximum leaf spanning tree of grid graph problem are also obtained.

Fixed Points of Homeomorphisms of Knaster Continua

by

Vincent A Ssembatya

Makerere University, Uganda

Abstract: J. Aarts and R. Fokkink proved that every homeomorphism of the standard (dyadic) Knaster continuum has two fixed points. This answered in the affirmative a question asked by W. Mahavier. In this paper we show that for generalized Knaster continua defined by an arbitrary sequence of primes, this result may be false. On the other hand, there are many circumstances where homeomorphisms on generalized Knaster continua do have more than one fixed point. In certain cases, one can give a lower bound to the number of fixed points of a homeomorphism. Very often this lower bound is in fact very large. We also discuss a generalization of Knaster continua defined for dimensions greater than one. We show that some of the properties of Knaster continua hold for these generalized examples

Continuity of Inversion in the Algebra of Locality - Measurable Operators

by

Isaac Daniel Tembo

University of Zambia, Zambia

Abstract: Let M be a semi-finite von Neumann algebra in a Hilbert space H and τ be a faithful normal semi-finite trace on M . Let M^p denote the lattice of self-adjoint projections in M , I denote the identity of M , and $\|\cdot\|$ denote the C^* -norm on M . The set of all measurable operators \tilde{M} with sum and product defined as the respective closures of the algebraic sum and product is $*$ -algebra. Equipped with a metrisable vector topology called the topology of convergence in measure τ_m , \tilde{M} is a complete metrisable topological $*$ -algebra in which M is dense. For M , it has been shown that [1]

Proposition Let Q be the set of invertible elements in \tilde{M} , and (S_n) a sequence in Q such that $S_n \rightarrow_{\tau_m} I$. Then $S_n^{-1} \rightarrow_{\tau_m} I$, that is to say inversion is τ_m -continuous on Q .

In this talk we present a similar result but for the topology of local convergence in measure, whose definition we shall present.

Uniquely Hamiltonian Graphs

by

Herbert Fleischner

Vienna Technical University, Austria

Abstract: To decide whether a graph G has a hamiltonian cycle is an NP-complete problem. However, if G has no vertices of even degree, then by a theorem of Thomason, every edge belongs to an even number of hamiltonian cycle. In fact, J.Sheehan asked whether there exists a 4-regular uniquely hamiltonian graph (i.e., with precisely one hamiltonian cycle), and J.A. Bondy posed the more general question whether there is a uniquely hamiltonian graph of minimum degree 3. In this talk we show how one can construct uniquely hamiltonian graphs of minimum degree 4 and arbitrary large maximum degree.

The Role of Backward Mutations on the Within Host Dynamics Of HIV-1

by

Kitayimbwa M. John, Joseph Y. T. Mugisha and Robert A. Saenz

Abstract: The quality of life for patients infected with human immunodeficiency virus (HIV-1) has been positively impacted by the use of antiretroviral therapy (ART). However, the benefits of ART are usually halted by the emergence of drug resistance. Drug-resistant strains arise from virus mutations, as HIV-1 reverse transcription is prone to errors, with mutations normally carrying fitness costs to the virus. When ART is interrupted, the wild-type drug-sensitive strain rapidly out-competes the resistant strain, as the former strain is fitter than the latter in the absence of ART. One mechanism for sustaining the sensitive strain during ART is given by the virus mutating from resistant to sensitive strains, which is referred to as backward mutation. This is important during periods of treatment interruptions as prior existence of the sensitive strain would lead to replacement of the resistant strain. In order to assess the role of backward mutations in the dynamics of HIV-1 within an infected host, we analyze a mathematical model of two interacting virus strains in either absence or presence of ART. We study

the effect of backward mutations on the definition of the basic reproductive number, and the value and stability of equilibrium points. The analysis of the model shows that, thanks to both forward and backward mutations, sensitive and resistant strains co-exist. In addition, conditions for the dominance of a viral strain with or without ART are provided. For this model, backward mutations are shown to be necessary for the persistence of the sensitive strain during ART.

Comparative Study of the Distributions Used To Model Dispersion

by

Kipchirchir, I. C.

School of Mathematics, University of Nairobi, Nairobi Kenya

Abstract: The negative binomial distribution has been widely used and to a lesser extent the Neyman Type A distribution, whereas the Polya-Aeppli distribution has received no attention in modeling overdispersed (clustered) populations. On the other hand, the Poisson distribution is naturally used to model random populations. The aim of this paper is to carry out a comparative study of the aforementioned distributions based on index of patchiness, correlation, skewness and kurtosis. The study revealed that the negative binomial, the Neyman Type A and the Polya-Aeppli distributions are equivalent in describing dispersion and they have Poisson as a limiting distribution. However, the distributions differ in terms of skewness and kurtosis, though the Polya-Aeppli is closer to the negative binomial than the Neyman Type A. Thus, in order to discriminate probability models for over dispersion, an index which incorporates skewness and kurtosis need to be devised.

A Within Host Model of Blood Stage Malaria

by

Theresia Marijani

University of Dar es Salaam, Tanzania

Abstract: Malaria is a deadly tropical disease caused by protozoa of the genus plasmodium. The malaria parasite life cycle involves three cycles namely the sporogony (mosquito stages), exoerythrocytic schizogony (human liver stages), and the erythrocytic schizogony (human blood stage). We consider a mathematical model for malaria involving, susceptible red blood cells, latent infected red blood cells, active infected red blood cells, intracellular parasites, extracellular parasites and effector cells. The models is analysed mathematically and numerically. One of the question addressed in our study is: what replicative characteristics offer the parasite opportunities to evade the host immune system? The results showed that the longer it takes to produce the parasites, the higher the chance that an infected red blood cell will be identified and apoptosised by the effector cells. Our sensitivity analysis results show that poor parametric estimation has serious implications on the prognosis of the disease. Treatment results suggest that a high drug efficacy can stop the development of the disease. The study has revealed that the parasite replicative characteristics enable the parasite to evade the immune response during the red blood stage malaria. We have found that the parasite has a strategy of infecting older red blood cells as a strategy to evade immune surveillance. We recommend treatment to be used in areas where anti-malarial drugs do not show resistance to the parasites. We also recommend that individuals with malaria or showing some symptoms should be treated for both malaria and chronic infections.

Application of Stochastic Differential Equations to Model Dispersion of Pollutants in Shallow Water

by

Wilson Mahera Charles

University of Dar-es-salaam, Tanzania

Abstract: A two dimensional stochastic differential equations(SDEs) to describe the dispersion of pollutants in shallow water is developed. By deriving the Kolmogorov's forward partial differential equation or commonly called Fokker-Planck equation, the SDEs model is shown to be consistent with the two-dimensional advection-diffusion equation. To improve the behaviour of the model shortly after the deployment of the pollutant, the SDEs called random flight model is developed too. It is shown that over long simulation periods, this model is again consistent with the advection diffusion equation. The simulated results in an ideal two dimensional domain are presented to predict the dispersion of a pollutant in the shallow waters.

Stochastic Model for In-Host HIV Virus Dynamics With Therapeutic Intervention

by

R. W. Mbogo, L. LuboobiJ. W. Odhiambo

Abstract: Mathematical models are used to provide insights into the mechanisms and dynamics of the progression of viral infection in vivo. Untangling the dynamics between HIV and CD4⁺ cellular populations and molecular interactions can be used to investigate the effective points of interventions in the HIV life cycle. With that in mind, we develop and analyze a stochastic model for In-Host HIV dynamics that includes combined therapeutic treatment and intracellular delay between the infection of a cell and the emission of viral particles, which describes HIV infection of CD4⁺ T-cells during therapy. The unique feature is that both therapy and the intracellular delay are incorporated into the model. Models of HIV infection that include intracellular delays are more accurate representations of the biological data. We show the usefulness of our stochastic approach towards modeling combined HIV treatment by obtaining probability distribution, variance and co-variance structures of the healthy CD4⁺ cell, and the virus particles at any time t . Our analysis show that, when it is assumed that the drug is not completely effective, as is the case of HIV in vivo, the predicted rate of decline in plasma HIV virus concentration depends on three factors: the death rate of the virions, the efficacy of therapy and the length of the intracellular delay.

An Alternating Iterative Procedure for the Cauchy Problem for the Helmholtz Equation

by

F. Berntsson, V. Kozlov, L. Mpinganzima, and B.O. Turesson

Department of Mathematics, Linköping University, Sweden

Abstract: Let ω be a bounded domain in \mathbb{R}^2 with a Lipschitz boundary Γ divided into two parts Γ_0 and Γ_1 which do not intersect one another and have a common Lipschitz boundary. We consider the following Cauchy problem for the Helmholtz equation

$$\begin{cases} \Delta u + k^2 u = 0 & \text{in } \omega \\ U = f & \text{on } \Gamma_0 \\ \partial_\nu u = g & \text{on } \Gamma_1 \end{cases}$$

where k is the wave number, ∂_ν denotes the outward normal derivative, and f and g are specified Cauchy data on Γ_0 . This problem is ill-posed.

The alternating iterative algorithms for solving this problem are developed and studied. These algorithms are based on the alternating iterative schemes suggested in [1] and [2]. Since these original alternating iterative algorithms diverge for a large constant k^2 in the Helmholtz equation, we develop a modification of the alternating iterative algorithms which converge for such k^2 . We also perform numerical tests. The numerical experiments confirm that the proposed modification works.

Index

- Abdou Sene, 157
Alphonse, 87
Anyika, 158
- Berntsson, 155, 165
Bjork, 161
Bgvad, 161
- Chatyrko, 155
Ciriza, 108
- El Tom, 11
Evarest, iv
- Fleischner, 163
- Gahirima, v
- Joseph, 159
- Kasozi, v
Kgosimore, 159
Kikonko, 52
Kipchirchir, 164
Kiselman, 155
Kitayimbwa John, 163
Kozlov, 165
Kumar, 72
- Luboobi, 81, 165
- Mahara, v
Mahera, 162, 165
Makame Mnyaa Mbarawa, 9
Makinde, 43
Mango, v, 3
Marijani, iv, 159, 164
Massawe, 1
Mataramvura, 162
Mbalawata, 157
Mbiliri, 162
Mbogo, 165
Mirumbe, 161
Mkandawile, 159
Montaz Ali, 61
Mpinganzima, 165
Mugisha, 163
Mujuni, iv, 162
Mureithi, iv, 160
- Musiga, 158
Musonda, 120
Mwale, 160
Mwaonanji, 160
- Ngare, 161
- Odhiambo, 165
Ogana, 156
- Rakotondrajao, 156
Rugeihyamu, iv
- Saenz, 163
Sarkka, 157
Seshaiyer, 156
Sima, 61
Soud, 98
Ssembatya, 161, 162
- Tembo, v
Tembo, 163
Turesson, 165
- Vaderline, 23
- Walakira, 159
Weke, v, 26, 158