Proceedings of the 30th Southern Africa Mathematical Sciences Association (SAMSA) Conference

26th - 29th November, 2012
Crossroads Hotel, Lilongwe,
Malawi

Editors
Busiso P. Chisala
Khumbo Kumwenda

SAMSA

SAMSA 2012 participants pose for a group photograph soon after Guest of Honour, Malawi's Minister of Finance Hon. Dr Ken Lipenga (first row, 4th from left) officially opened the conference. He is flanked on the right by Her Excellency Madam Liz Haggins, Irish Ambassador to Malawi, and on the left by Dr Emmanuel Fabiano, Vice Chancellor of the University of Ma-

Foreword

SAMSA conferences have over the last 30 years provided a platform for researchers from Southern Africa and beyond to share recent advances in mathematical sciences including ideas on new approaches to problem solving in various fields of research. The conferences have also provided opportunities for networking and learning. Being conscious that we live in a global village, SAMSA has always extended invitation to experts from all over the world to share ideas and experiences. The 2012 conference, held at Crossroads Hotel in Lilongwe, Malawi, from 26th to 29th November, was no different.

Held under the theme "Strengthening Mathematics and its applications in Developing Countries", the conference was officially opened by the Minister of Finance Hon. Dr. Ken Lipenga, MP, whose thought-provoking speech, included in the proceedings, dwelt on the role mathematics plays in addressing challenges in science, engineering, business and industry; and how it is applied to solve practical problems facing humanity. The conference brought together keynote speakers, researchers and students from as far and wide as Belgium, Botswana, Canada, England, Germany, Indonesia, Italy, Ireland, Kenya, Malawi, Namibia, Norway, Nigeria, Russia, South Africa, Serbia, Singapore, Sweden, Tanzania, U.S.A., Uganda, Zambia and Zimbabwe. Among other dignitaries who graced the official opening were Her Excellency Liz Higgins, Ambassador of Ireland and Dr. Emmanuel Fabiano, Vice Chancellor of the University of Malawi.

SAMSA acknowledges the role played by the following in organising the 2012 conference: The former SAMSA executive led by Dr Chiteng'a John Chikunji; the Local Organising Committee led by Dr Peter Y Mhone; the host institutions, namely, University of Malawi's (UNIMA) Chancellor College and the Malawi Polytechnic, Mzuzu University (MZUNI), Catholic University of Malawi (CUNIMA), and the Lilongwe University of Agriculture and Natural Resources (LUANAR). SAMSA would also like to thank presenters who sent their papers for publication in these conference proceedings, and the various reviewers who responded to the LOC's request for review positively.

Probably the greatest challenge in organizing such a conference is having prior funding. This became an even greater challenge in 2012 with the economic downturn in Malawi. The conference would not have been a success that it was, without the generous financial support from Irish Aid, Malawi Savings Bank, Grey Matter Book Distributors, and travel support to keynote speakers and students by Centre International de Mathématiques Pures et Appliquées (CIMPA, France), Centre of Mathematics for Applications (CMA, Norway), Innovations in Stochastic Analysis and Applications (Innostoch, Norway), National Science Foundation (NSF) through Auburn University (USA), Humboldt University of Berlin (Germany), London Mathematical Society - African Mathematics Millennium Science Initiative (LMS-AMMSI) and many other organizations and institutions who supported speakers and delegates. To all these, SAMSA is truly most grateful.

After the conference, from 29th November - 2nd December 2012, SAMSA held the second Masamu US-Africa Advanced Study Institute and Workshops in Math-

ematical Sciences under the MASAMU program, a collaborative effort between African, United Kingdom, and US mathematicians supported by the National Science Foundation (NSF, USA). Thanks to Professor Overtoun Jenda of Auburn University for his initiative and efforts to contribute to the development of mathematics in the Southern Africa. The primary goal of the Program is to enhance research in mathematical sciences within SAMSA institutions, targeting graduate students and early career faculty. For the present Institute, participants came from Botswana, Kenya, Malawi, South Africa, Tanzania, USA, and Zimbabwe; and presented progress of research work they started at the first Institute in Livingstone at the end of SAMSA 2011. It was encouraging to note that there was a lot of progress, with some groups about to submit papers for publication.

We hope you will enjoy reading these proceedings and trust that you will find them useful.

Dr. Levis Keliyasi Eneya
SAMSA President

SAMSA Executive

Levis Eneya	President
Farai Nyabadza	Vice President
Peter Mhone	Secretary
Moatlhodi Kgosimore	Treasurer
Stanford Shateyi	Committee Member
John Mango	Committee Member
Mervis Kikonko	Committee Member

Scientific Committee

Bernt Oksendal	Norway
Wandera Ogana	Kenya
Edward Lungu	Botswana
John Ryan	Malawi
Mihalis Zervos	UK
Busiso Chisala	Malawi
Overtoun Jenda	USA

Local Organizing Committee

Peter Mhone	Unima (Chairperson)
Paul Kubwalo	Mzuni (V. Chairperson
Levis Eneya	Unima (Secretary)
Milliward Maliyoni	Unima (V. Secretary)
Patrick Musicha	Cunima (Treasurer)
Kondwani Magamba	Cunima
John Ryan	Mzuni
Busiso Chisala	Mzuni
Khumbo Kumwenda	Mzuni
Dun Kasoka	Unima
Tasokwa Kakota	Luanar

Contents

Opening Speech at SAMSA 2012

Hon. Dr Ken Lipenga, MP

Minister of Finance, Republic of Malawi

I am delighted to be here to give the Opening Statement at this Conference whose theme is "Strengthening Mathematics and its Applications in Developing Countries". I welcome all participants to Malawi, and as your hosts we are delighted that Malawi is hosting this International Conference for the third time. I congratulate all universities within SAMSA for sustaining this Conference for almost three decades. These conferences have served as an avenue for the exchange of knowledge, information and provoking new ideas in problem solving among experts.

I wish you well in this conference. I am sure that it will be fascinating and provide opportunities for networking and learning. I hope all students take advantage of the opportunities that this Conference presents to continuously push forward the boundaries of knowledge and to advance research and learning within the Southern African region. I commend and applaud the Local Organizing Committee for all the work you have done in advance of this conference. I thank the Embassy of Ireland for providing a grant to finance this Conference and to the hosting Universities in Malawi for your own contributions to this Conference.

The objective of the Conference is to provide a platform for experts to exchange ideas and establish collaboration. I am pleased that the conference has so many participants from across the SADC Region, Nigeria, Cameroon, Kenya, Norway, Ireland, the United Kingdom, Germany, Sweden and the United States of America. This conference will provide an excellent opportunity to present the most recent advances in pure and applied mathematics, advance debate on problem solving and apply mathematics to the pressing issues of our continent in health, finance, the environment, engineering and commerce.

I am also delighted to see so many students here from Malawi and I welcome students from other SADC universities. I am proud of the recent advances in Mathematics in Malawi including the establishment of a Masters and PhD Programme in Mathematics at Mzuzu University and our advances in pure and applied mathematics.

I also consider it a great privilege and honour that you have invited a non-mathematician to deliver the opening statement. I know that there are many debates between you as to whether mathematics is a science and whether mathematics has sovereignty over the natural sciences. It appears that Albert Einstein was sceptical noting that "In so far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality".

It has been pointed out that all creatures seem to be born with brains that have predisposition for mathematics, and the numerical competence is crucial to survival in the wild. A chimpanzee is less likely to go hungry if he can look up a tree and

quantify the amount of ripe fruits he will have for his lunch. To me, mathematics is a language, a language with universal relevance to economic development. Galileo noted that "The universe cannot be read until we have learned the language and become familiar with the characters in which it is written. Without these, we are wandering in a dark labyrinth."

Mathematicians are famous for seeking out "problems" the more complex, the more exciting. And what could be more complex than reality, especially when it involves new developments in modern key technologies? Industries and businesses, in turn, are increasingly employing mathematics to structure and transfer complex problems into technical, quantitative models. Mathematical approaches allow problems to be better understood, to recognize what is essential, to make predictions in difficult situations and to make better decisions. Its research activities make it possible for important biological and medical processes to be mathematically modelled, tested in virtual laboratories, adapted and optimized. And of course we use a lot of mathematical modelling in my own ministry, whether it is designing the budget and making projections, or in the structuring of salary and payroll systems. Systems such as IFMS (Integrated Financial Management System), which depend so much on mathematical modelling, are designed not only to make work easier for our technicians, but also to combat fraud and corruption, thereby contributing to the never-ending quest for the perfect society. So, if you have ever wondered whether mathematics has a role in the growth of democracy, there is your answer.

As a non mathematician, I have for a very long time been fascinated by the way applied mathematics concerns itself with addressing challenges in science, engineering, business and industry, and applies mathematics to practical problems facing humanity. I note from the agenda that you will discuss papers on mathematical biology including "Raising the Bar; Projected Impact, Cost and Cost-Effectiveness of Alternative CD4 Cell Count Threshold for ART Initiation on the HIV and TB Epidemic in South Africa" Other papers in the section on Codes and Cryptology include "Mobile Banking Scheme, Independent of both Banks and Mobile Network" and "Mathematics: A tool for Promoting Consumer Financial Literacy". There is also a paper for politicians like myself; "Low cost and secure hybrid e-voting system using mobile phones.

Ladies and gentlemen, why do so many people find mathematics so outright bewildering? Today, we require understanding of the linear number line to function in modern society. The linear number line is the basis of measuring, and facilitates calculations. Yet in our dependence on linearity, have we perhaps gone too far in stifling our own, arguably more naive, logarithmic institutions? Could this perhaps be the reason why so many people find mathematics difficult? Could this be the reason why we so easily lose our ability to manipulate exact numbers and default to our logarithmic institutions, judging amounts and perceive distances with approximations and ratios?

There is paradox here, in that the seemingly difficult subject that is mathematics has in fact been largely responsible for making life easier for modern humanity, through the applications that I have referred to that you are going to be dis-

cussing. It is for this reason that I believe it will be mathematicians who will find the formula for demystifying mathematics, making it easier for everyone to follow. I am confident that through your work we will find that we are in fact, all of us, mathematicians under the skin, descendants of Pythagoras and Euclid.

Ladies and Gentlemen, I wish you well as you explore, advance and use mathematics to promote innovation and development. I thank you for your kind attention.

Mathematical Modelling of the Prevalence of Malaria: A Case Study of Zomba District, Malawi

Thomas Basikolo and Levis Eneya

Department of Mathematical Sciences, University of Malawi, Chancellor College, Zomba, Malawi
Emails: thomasbasikolo@gmail.com, leneya@cc.ac.mw

Abstract. Malaria is a major public health concern, especially among pregnant women and children under the age of five. It is a leading cause of morbidity and mortality in Malawi, accounting for fourty percent of out-patient consultations in many health facilities in the country. Mathematical models have an important role to play in making public health decisions about the control of infectious diseases such that they are better informed and more objective. In order to understand the prevalence, transmission and control of the Malaria epidemic, an SEIR model has been used in this study. The model was analysed to determine criteria for control of the malaria epidemic, and was used to compute the basic reproduction and effective reproduction numbers necessary for control of the disease. In this paper an expression for the basic reproduction number R_0 is derived through the next generation method. Numerical results indicate the effect of the two controls: protection and treatment in lowering exposed and infected members of each of the populations. The results also highlight the effects of infection rate and removal rate.

Keywords: SEIR model, next generation method, mathematical modelling, Malaria epidemic

1. Introduction

Malaria is one of the greatest health and development challenges worldwide. The disease is endemic in ninety-one countries, accounting for fourty percent of the world's population. But the greatest load of mortality and morbidity due to malaria is borne by the world's poorest economies, most of them in sub Saharan Africa [5]. Nine of every ten cases of malaria and malaria deaths occur in Africa. Malaria is a major public health concern in Malawi, especially among pregnant women and children under the age of five. It is a leading cause of morbidity and mortality in Malawi, accounting for 40% of Out Patient Department (OPD) consultations in many health facilities in the country [10]. The Malawi

government has embarked on reducing maternal and neonatal mortality and morbidity in Malawi through combating Malaria, HIV and Aids, among other diseases which is goal 6 in the Malawi Government Millennium Development Goals Report of 2008 [11].

Mosquitoes transmit malaria; hence malaria distribution largely depends on the following climatic factors that affect the survival and multiplication of the anopheles mosquitoes: temperature, humidity, and rainfall [4]. The extent and geographic distribution of these factors influences the prevalence (occurrence) of malaria in Malawi. Malaria occurs in all parts of Malawi but with variable transmission patterns. While malaria prevalence data for Malawi is scanty, there are estimates of incidence (frequency) rates. For instance, in 1994 the estimated malaria incidence for Malawi was $49,410$ per $100,000$ population [5]. Even though in a disease with a short natural course like malaria, incidence and prevalence are almost the same, the reported estimates of malaria incidence in an endemic country like Malawi would certainly underestimate prevalence rates of the disease. This is because in people living in endemic areas most malaria infections are asymptomatic and resolve without treatment due to malaria-specific acquired immunity. Thus, one of the most serious problems with estimating malaria prevalence from incidence data is the omission of these asymptomatic malaria cases thereby underestimating incidence rates. In addition, since incidence data are predominantly obtained from malaria case records at clinics, health centres and hospitals, malaria cases that are treated at home are missed out resulting in an underestimation of true malaria incidence.

The incidence of malaria has been growing recently due to increasing parasite drug-resistance and mosquito insecticide-resistance. Therefore, it is important to understand the important parameters in the transmission of the disease and develop effective solution strategies for its prevention and control. A mathematical model is developed to better understand the transmission and spread of malaria. The model divides the human population into four classes: susceptible, exposed, infectious, and recovered (immune). Humans enter the susceptible population through birth or immigration. Susceptible humans get infected at a certain probability when they are bitten by infectious mosquitoes. They then progress through the exposed, infectious, and recovered classes, before re-entering the susceptible class.

According to Dzinjalamala [5], malaria is associated with productivity losses and reduction in profits in the private sector particularly tourism and construction industry, stemming from a sick workforce and hospital bills. In addition, the threat of malaria negatively affects visitors to Malawi. Malaria is a frequent cause of absenteeism in school, not only for students but also teachers, resulting in poor scholastic performance on the part of the student and a negative impact on the ability of the teachers to work hard, either because they are themselves sick or because their children are sick. In some children, cerebral malaria may lead to cognitive impairment or neurological sequelae that negatively impact on their educational attainment. Thus, this paper aims to understand the dynamics of malaria transmission and spread, predict the steady states, and calculate the

threshold parameter R_0 of the disease.

2. Mathematical Modelling

2.1 Ross - MacDonald Model

The earliest attempt to provide a quantitative understanding of the dynamics of malaria transmission is the so called Ross-Macdonald model, which still is the basis for much malarial epidemiology [3]. This model, which captures the basic features of the interaction between the infected proportions of the human host and the mosquito vector population, is defined by the following system of Differential Equations:

$$\begin{aligned} \frac{dx}{dt} &= \alpha y(1-x) - rx \\ \frac{dy}{dt} &= \beta x(1-y) - \mu y \end{aligned} \tag{2.1}$$

where,

- x = proportion of the human infected population
- y = proportion of the female mosquito population with malarial parasites
- α = rate of infection of humans by mosquitoes
- β = rate at which mosquitoes gets malarial parasites from humans

In this model, the total population of both humans and mosquitoes is assumed to be constant. Also the quasimonotonicity condition is satisfied. Usually the parameter α is explicated as follows:

$$\alpha = \frac{a\beta M}{N}, \tag{2.2}$$

where β is defined as above, a is the proportion of infected bites on man that produce an infection, N is the size of the human population; and M is the size of the female mosquito population. The threshold parameter usually referred as the basic reproduction ratio is given by

$$R = \frac{M}{N}\frac{\beta^2\alpha}{\mu r}. \tag{2.3}$$

When $R > 1$ a unique nontrivial endemic state appears with components:

$$\begin{aligned} x^* &= \frac{R-1}{R+\frac{\beta}{\mu}} \\ y^* &= \frac{R-1}{R}\frac{\frac{\beta}{\mu}}{1+\frac{\beta}{\mu}} \end{aligned} \tag{2.4}$$

where x^* and y^* are the steady states.

2.2 SIR Model

The so-called SIR model is one of the simplest and most fundamental of all epidemiological models. It is based upon calculating the proportion of the population in each of the three classes (susceptible, infected and recovered) and determining the rates of transition between these classes. The model, which was described by Kermack and McKendrick in 1927, consists of three compartments: susceptible S, infected I, and Removed R hence it is referred to as SIR model, and is given by

$$\begin{aligned} \frac{dS}{dt} &= \wedge - \frac{\beta SI}{N} - \mu S \\ \frac{dI}{dt} &= \frac{\beta SI}{N} - \mu I - \gamma I \\ \frac{dR}{dt} &= -\mu R + \gamma I \end{aligned} \tag{2.5}$$

where $\wedge$ is the recruitment rate, β the contact rate, μ the death or mortality rate and γ is the removal or recovery rate.

2.3 SEIR Model

With respect to the classical SIR model with vital dynamics, the SEIR models include a further class E of latent individuals. Many diseases like malaria have what is termed as latent or exposed phase E during which an individual is said to be infected but not infectious, i.e. they are not yet capable of transmitting the disease [3]. We will use the SEIR framework to describe a disease with temporary immunity on recovery from infection. SEIR model indicates that the passage of individuals is from the susceptible class S, to the exposed class E, then to infective class I, and finally to the recovery class, R. $S(t)$ represents the number of individuals not yet infected by the malaria parasite at time t, or those susceptible to the disease [8]. In this model $I(t)$ denote the number of individuals who have been infected with malaria and are capable of spreading the disease to those in the susceptible category. This is done through infecting the susceptible mosquitoes. The dynamic transmission of the malaria parasite between and amongst individuals in both species is driven by the mosquito biting habit of the humans. R is the compartment for individuals who have recovered from the disease. These humans cannot transmit the infection to mosquitoes as we assume that they have no plasmodium parasites in their bodies.

For some disease, it takes certain time for an infective agent to multiply inside the host up to the critical level so that the disease, actually manifest itself in the body of the host [13]. This is called an incubation period. We have the same assumptions as in the SIR model, which are homogeneous mixing (mass action principle), constant population size and the rates of change from one compartment

to the other, giving the system:

$$\begin{aligned} \frac{dS}{dt} &= \wedge - \frac{\beta SI}{N} - \mu S \\ \frac{dE}{dt} &= \frac{\beta SI}{N} - \mu E - \delta E \\ \frac{dI}{dt} &= -\mu I + \delta E - \gamma I \\ \frac{dR}{dt} &= -\mu R + \gamma I \end{aligned} \tag{2.6}$$

where δ is rate of infection, and the rest pr parameters are as in model (2.5).

The probability to survive the latency period and to enter the infectious period equals to $\frac{\delta}{\delta+\mu}$.

Therefore the basic reproductive number in this case will be

$$R_0 = \frac{\delta}{\delta+\mu}\frac{\beta}{\gamma+\mu}. \tag{2.7}$$

The steady states are found when $\frac{dI}{dt} = 0$.

3. Model analysis

The next-generation method, developed by Diekmann et al. [19] and Heesterbeek [20], and popularised by van den Driessche and Watmough [26], is a generalisation of the Jacobian method. It is significantly easier to use than Jacobian-based methods, since it only requires the infection states (such as the exposed class, the infected class and the asymptomatically infected class) and ignores all other states (such as susceptible and recovered individuals). This keeps the size of the matrices relatively manageable. In order to determine the matrices F and V, where F accounts for the new infections and V accounts for the transfer between infected compartments, biological insight must be used in order to decide which terms count as new infections and which terms are transfer terms.

Using the Next Generation Method, consider G to be the next generation matrix. It is comprised of two parts, F and V^{-1}, where

$$F = \left[\frac{\partial F_i(x_0)}{\partial x_j}\right] \tag{3.1}$$

and

$$V = \left[\frac{\partial V_i(x_0)}{\partial x_j}\right]. \tag{3.2}$$

The F_i are the new infections, while the V_i are transfers of infections from one compartment to another. In equations (3.1) and (3.2), x_0 is the disease-free equilibrium state.

R_0 is the dominant eigenvalue of the matrix $G = FV^{-1}$. Thus the determination of $G = FV^{-1}$ leads to the determination of R_0 for the system in (2.6). The Disease Free Equilibrium (DFE) is $P_0 = (S_0, E_0, I_0, R_0)$ such that $S_0 = \frac{\wedge}{\mu} = N$, $E_0 = 0$, $I_0 = 0$, $R_0 = 0$. The vector of disease state is

$$\vec{X} = (E, I)^T \quad \text{where,}$$

$$\frac{dE}{dt} = \frac{\beta SI}{N} - \mu E - \delta E$$
$$\frac{dI}{dt} = -\mu I + \delta E - \gamma I$$

Linearization of $\vec{X}$ about the DFE yields

$$\frac{d\vec{X}}{dt} = J\vec{X}$$

where $J := F - V$ is given by

$$J = \begin{pmatrix} -\delta - \mu & \beta \\ \delta & -\mu - \gamma \end{pmatrix}$$

From the above equation, we see that

$$F = \begin{pmatrix} 0 & \beta \\ 0 & 0 \end{pmatrix},$$

and

$$V = \begin{pmatrix} \delta + \mu & 0 \\ -\delta & \mu + \gamma \end{pmatrix}.$$

By definition, the next generation matrix is given by $G = FV^{-1}$. From F and V given above, we compute

$$V^{-1} = \begin{pmatrix} \frac{1}{\delta+\mu} & 0 \\ \frac{\delta}{(\delta+\mu)(\gamma+\mu)} & \frac{1}{\mu+\gamma} \end{pmatrix}$$

and

$$G := FV^{-1} = \begin{pmatrix} \frac{\beta\delta}{(\delta+\mu)(\gamma+\mu)} & \frac{\beta}{\mu+\gamma} \\ 0 & 0 \end{pmatrix}$$

The last matrix FV^{-1} has two eigenvalues, namely, 0 and $\frac{\beta\delta}{(\delta+\mu)(\gamma+\mu)}$. From the foregoing, since R_0 is the dorminant eigen value of G, then we have

$$R_0 = \frac{\beta\delta}{(\delta+\mu)(\gamma+\mu)}. \tag{3.3}$$

4. Results and Discussions

Using a set of parameters, we carried out simulations of the model in Matlab. Some of the parameters were obtained from the National Statistical Office (N.S.O) in Zomba, Malawi and other values that have been used before in literature on Malaria transmission. The values from literature are used within a certain range of parameter values according to an area because there are differences in terms of factors that leads to Malaria prevalence. Estimates for some parameters are scarce or their actual values are unknown precisely thus we resorted to parameter values that are in line with literature on Malaria transmission for their values. Some parameters vary from country to country and some are influenced by demographics, for instance, natural death rate, transmission of the disease and recruitment rate. In this model, the dynamics of the vector have been assumed as the contact rate and is incorporated in the mass action term $\frac{\beta SI}{N}$. Thus the model only deals with the dynamics of the human beings.

We simulate the basic malaria model in the absence of any intervention and then a malaria model with data from Zomba D.H.O on population of Zomba, and people who were suffering from malaria in the previous year. Then we find out the differences. In order to find out the dynamics of the disease in the population when there is no intervention, a simulation of basic model has been conducted. In the absence of intervention strategies, the susceptible population decreases as shown in the Figure 4.1.

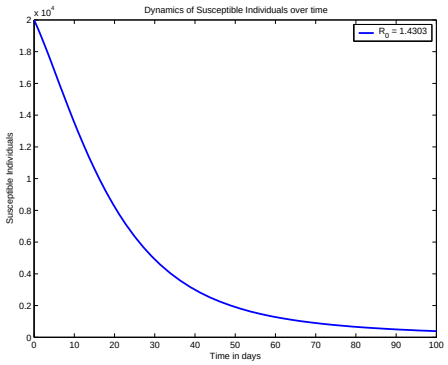


Figure 4.1: Graph of Susceptible Population

Since there are no interventions to eradicate the disease, the susceptible population will continue being exposed to the disease, and as a result of being recruited to another class, the exposed population will slightly increase from the start as shown in Figure 4.2 and then decrease.

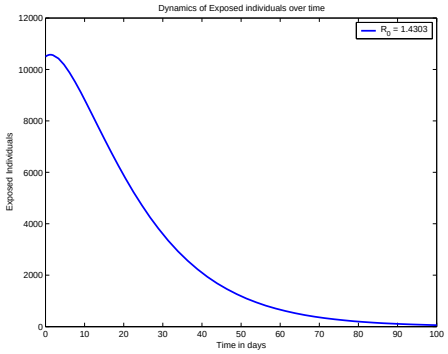


Figure 4.2: Graph of Exposed Population

This means that the plasmodium is continuously multiplying since there are no means of reducing or eradicating it. After some few days the exposed population decreases since they are now becoming infective. Thus the infected population increases due to the increase in the exposure to the disease and the value of R_0 is 1.4303 meaning that a lot more people are getting the disease, see Figure 4.3.

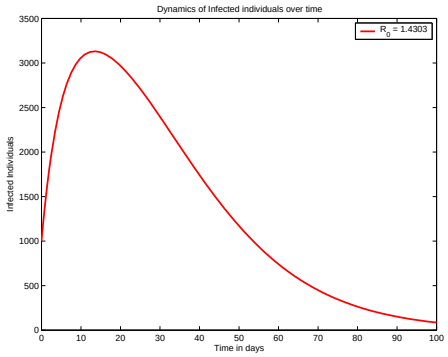


Figure 4.3: Graph of Infected Population

This supports the result that the disease is endemic when $R_0 > 1$. The recovered population decreases as a result of availability of malaria in the community in which no any intervention is being practised. This is illustrated in the Figure 4.4.

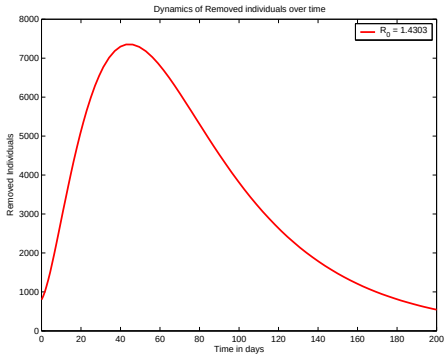


Figure 4.4: Graph of Removed Population

We also looked at the prevalence in the population. Prevalence is defined as the ratio of the number of cases of a disease in a population with the number of individuals in a population at a given time. It is used as an estimate of how common a disease is within a population over a certain period of time. Figure 4.5 shows the prevalence of malaria in the population without interventions.

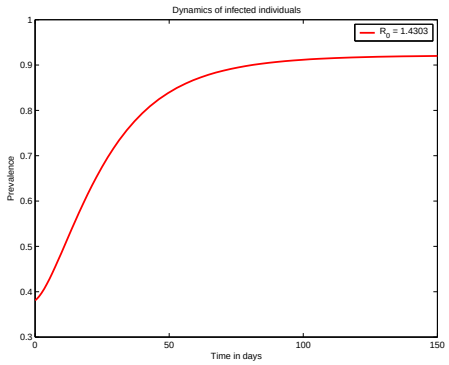


Figure 4.5: Graph of Malaria Prevalence

The prevalence graph, Figure 4.5 with $R_0 = 1.4303$, increases exponentially for a while and then reaches the equilibrium. This happens because of the reduced number of susceptible individuals with time as most of the individuals in the society become affected with the disease. In this we note that the removal rate γ is of paramount significance. When γ is increased from say 10% to 50% the graph of prevalence changes and there is a significant change in terms of the disease dynamics. Figure 4.6 illustrates how γ affects the prevalence and the dynamics of Malaria in general.

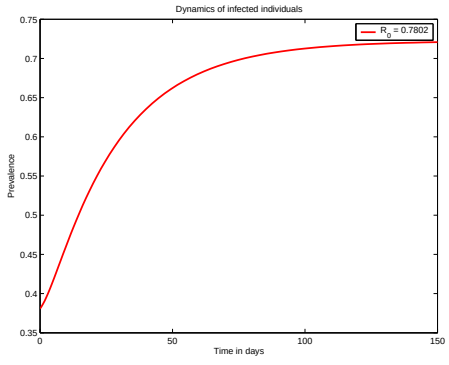


Figure 4.6: Graph of Malaria Prevalence

As the value of γ continues to increase, the prevalence of the disease further decreases and this is also noted by the changing value of R_0. Figure 4.7 illustrates this.

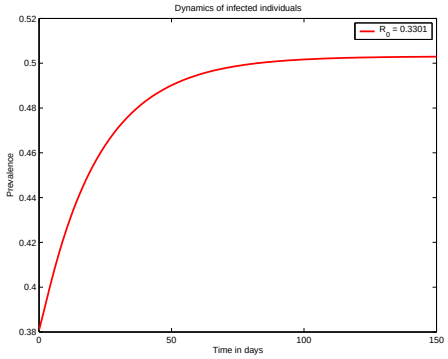


Figure 4.7: Graph of Malaria Prevalence

5. Malaria Dynamics for Zomba

When the interventions are introduced, improved trends of the populations in terms of the proportion of the sick people are observed. This can be seen in Figure 5.1 that the susceptible population decreases though interventions are there to reduce the impact of the disease.

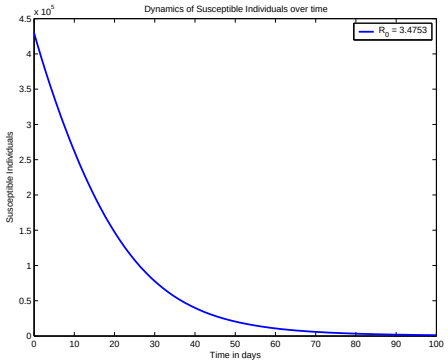


Figure 5.1: Graph of Susceptible Population for Zomba

The dynamics of the infected individuals is such that few people get infected by the disease as compared to the one without interventions.

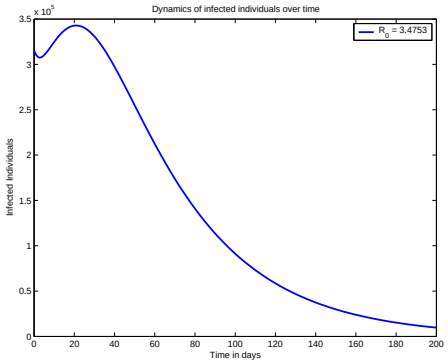


Figure 5.2: Graph of Infected Population for Zomba

In the Figure 5.2, the trend is such that people are recruited to the infected compartment as manifested by the increase in the graph and these are from the susceptible class. After some time they are removed from the population, which entails that people are removed from this class after a certain period of time due to the treatment they get after visiting the hospital.

The prevalence is such that the disease is so prevalent in the population as manifested by the trend in Figure 5.3 and also by the value of R_0 which is so big that one sick person can infect more than four people.

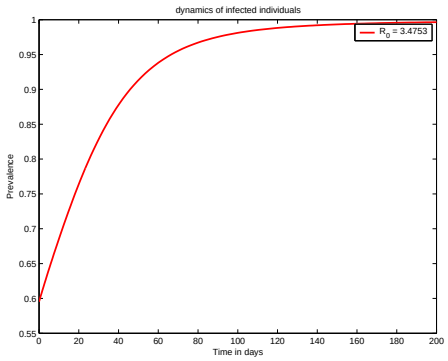


Figure 5.3: Graph of Malaria Prevalence in Zomba

6. CONCLUSION

This study has presented and analysed a basic SEIR model for the transmission of Malaria. It considered a varying total human population since it incorporated recruitment of new individuals into the susceptible class mainly through birth. The model incorporated features that are effective to control the transmission of Malaria disease in the district. The prominent parameter in the model, the basic reproduction number, R_0, as a control intervention measure was computed. In this, when $R_0 < 1$ the disease-free equilibrium is stable and when $R_0 > 1$ the endemic equilibrium is stable.

From the graphs of susceptible individuals in Figure 5.1, we noted that many people were being exposed to the disease in the first days and that the rate of exposure was so high. This entails that there is need for control strategies of the contact rate. Since in our case the vector acts as the contact, there is need for control of the mosquito. This means a lot of people that are susceptible to the disease are supposed to be protected from contacting the disease. Thus, the use of Insecticide Treated Nets (ITNs), Indoor Residual Spraying (IRS), and filling up of stagnant water places.

Another parameter that is of much significance is the removal rate γ. When this parameter is increased by some percentage, it affects a lot the value of R_0. As such, increasing the population of the people that are treated from the infection will mean that the infected class will be reduced and that the transmission of the

disease will also decrease since there will also be a decrease in terms of the contacts, and the secondary infections. Furthermore, prevention measures are important to maintain the reduction or eradication of the disease transmission and lowering the infection. Therefore, the results indicate the effect of control such as protection and treatment in lowering exposed and infected members of each of the populations.

7. Recommendations

Malaria eradication remains a challenge to National Malaria Control Programme in Malawi, hence there is need to strengthen the control strategies at hand as well as looking for some new ones. As the Malaria disease continues to burden individuals and communities in Malawi, policy makers have to be informed about the research results. From the results of this work, we make the following recommendations.

7.1 Protecting individuals from Malaria

With regard to malaria prevention, insecticide-treated nets (ITNs) have a track record of reducing malaria related morbidity and mortality, hence should increasingly be utilized. According to [25], regular use of ITNs can reduce overall under-five mortality rates by about 20 per cent in malaria-endemic areas. Malaria-infected mosquitoes bite at night, and these nets provide to the sleeping individual a physical barrier against the bite of an infected mosquito. In addition, a net treated with insecticide provides much greater protection by repelling or killing mosquitoes that rest on the net. Indoor residual spraying (IRS) is the practice of spraying insecticides on the interior walls of homes in malaria affected areas [6]. After feeding, many mosquito species rest on a nearby surface while digesting the blood meal, so if the walls of dwellings have been coated with insecticides, the resting mosquitoes will be killed before they can bite another victim, transferring the malaria parasite. Most of the reductions in transmission comes from the protection of humans, it is important therefore also to improve the killing effects of insecticide mosquito treated bed-nets (ITNs) and indoor residual spray (IRS). Thus, complete coverage and improved killing effects may be necessary to reach control goals.

7.2 Vector Control

Interventions such as insecticide-treated nets (ITNs) and indoor residual spraying (IRS) are proving effective to combat and prevent the disease in the district and also filling up of stagnant water and other mosquito breeding site. These would reduce the availability of hosts, and kill mosquitoes that are attempting to feed on humans' blood, larva that is growing, and reducing malaria transmission. There are different malaria vector control measures, including [27]:

- Reduction of human mosquito contacting which Insecticide-treated nets, repellents, protective clothing and screening of houses are used.

- Destruction of adult mosquitoes and mosquito larvae where ITNs, IRS, space spraying, and ultra low-volume sprays are used.
- Source reduction small-scale drainage.
- Social participation where there is motivation for personal and family protection, health education and community participation.

7.3 Treatment of individuals

Some of the critical components in the reduction transmission of malaria are: population wide or focused screening and treatment (FSAT) of all infected individuals, active detection of malaria cases (symptomatic) or of infections (asymptomatic). Treatment of all infected persons can substantially reduce the transmission of malaria at community level, since there will be less individuals that are infected thus reducing the number of contacts.

7.4 Awareness Campaigns

There are supposed to be awareness campaigns for people to use nets and other control strategies. Studies show that although people have nets, some do not use them. This is connected to religious beliefs, and problems that some people develop after using the nets. According to [6], education in recognizing the symptoms of malaria has reduced the number of cases in some areas of the developing world by as much as 20%. Recognizing the disease in the early stages can also stop the disease from becoming a killer. Education can also inform people to cover over areas of stagnant, still water e.g. water tanks, which are ideal breeding grounds for the parasite and mosquito, thus cutting down the risk of the transmission between people. Therefore, more awareness campaigns should be conducted in the district to sensitize people on the importance of the ITNs, IRS and getting medical attention in good time.

8. Future Work

As Malaria continues to claim more lives, and affect a lot of people in Malawi, it is imperative to have comprehensive research done in order to explore possible new control strategies of the infection as well as assessing the impact of the existing control strategies. Based on this study, it is proposed that future work should focus on:

1 the dynamics of the vector and interventions in the model;

2 comprehensive study on the compartments of the disease since compartmentalised data is not available; and

3 optimisation of malaria control strategies and interventions.

Bibliography

[1] N.T.J. Bailey, *The Mathematical Theory of Infectious Diseases and its Applications*, Charles Griffin & Company LTD: London, Second Ed., 1975.

[2] J.G. Breman, A. Mills, R.W. Snow, J. Mulligan, C. Lengeler, K. Mendis, B. Sharp, C. Morel, P. Marchesini, N.J. White, R.W. Steketee, and O.K. Doumbo, Conquering Malaria, in *Disease Control Priorities in Developing Countries*, 2nd edition. D.T. Jamison, J.G. Breman, A.R. Measham et al., (Eds.) Washington (DC): World Bank, 2006.

[3] V. Capasso, *Lecture Notes in Biomathematics: Mathematical Structures of Epidemic Systems*, Springer-Verlag Berlin Heidelberg, 2nd Edition., 2008.

[4] E.A. Cahill, *Basic Malaria Epidemiology*, 2010. Available from `academic.reed.edu/epi/papers/malariacahill.doc`

[5] F. Dzinjalamala, *Epidemiology of Malaria in Malawi*, College of Medicine-University of Malawi. Blantyre, 2006. Available from `http://www.medcol.mw/commhealth/publications/epi\%20book/Malaria\%20chapter\%203.pdf`

[6] East African Community Health, *Malaria Prevention and Control*, Technical Report, 2012.

[7] E. Feig, *Fighting malaria in Malawi*, Malaria Griots, 2010. Presentation slides available from `http://www.slideshare.net/ellenfeig/fighting-malaria-in-malawi`

[8] P.M.M Mwantobe, *Modelling the Effects of Multi-Intervention Campaigns for the Malaria Epidemic in Malawi*, MSc Thesis, University of Dar es Salaam, 2010.

[9] Ministry of Health, *Malawi National Malaria Indicator Survey 2010*, Malawi National Malaria Control Programme, Lilongwe: Malawi, 2010.

[10] Government of Malawi, *2008 Malawi Millennium Goals Report*, Ministry of Development Planning and Cooperation, Lilongwe: Malawi, 2008.

[11] Ministry of Health and Population, *To the year 2020: A Vision for the Health Sector in Malawi*, Malawi National Malaria Control Programme, Lilongwe: Malawi, 1999.

[12] J. Ndawala, G. Kalanda, M. Mahy, *Malaria: In National Statistical Office Malawi and ORC Macro. Malawi Demographic and Health Survey 2000*, Zomba, Malawi and Calverton, Maryland, USA, 2001.

[13] L. Bubniakova, *The mathematics of infectious diseases*, MSc Thesis, Comenius University, 2007.

[14] N.R. Chitnis, *Using Mathematical Models in Controlling the Spread of Malaria*, PhD Thesis, University of Arizona, 2005

[15] The malERA Consultative Group on Modeling, A Research Agenda for Malaria Eradication: Modeling, *PLos Med*, **8**(1) (2011). doi:10.1371/journal.pmed.1000403.

[16] Malaria Life Cycle, Available from `www.malariasite.com/malaria/LifeCycle`, 2012.

[17] R. Tumwiine, J.Y.T. Mugisha, and L.S. Luboobi, A Mathematical Model for the Dynamics of Malaria in a Human Host and Mosquito Vector with Temporary Immunity, *Applied Mathematics and Computation*, **189** (2007), 1953-1965.

[18] J. Li, D. Blakeley and R.J. Smith, The Failure of R_0, *Computational and Mathematical Methods in Medicine*, **2011** (2011), 1-26.

[19] O. Diekmann, J.A. Heesterbeek and J.A. Metz, On the Definition and the Computation of the Basic Reproduction Ratio R_0 in Models for Infectious Diseases in Heterogeneous Populations, *J. Math. Biol.*, **28**(1990), 365-382.

[20] O. Diekmann and J. A. P Heesterbeek, *Mathematical Epidemiology of Infectious Diseases*, Chichester, UK: John Wiley & Sons, 2000

[21] H.W, Hethcote, The Basic Epidemiology Models: Models, Expressions for R_0, Parameter Estimation and Applications, *Mathematical Understanding of Infectious Disease Dynamics*, 2008.

[22] H. W. Hethcote and J. A. Yorke, *Gonorrhea Transmission Dynamics and Control:Lecture Notes in Biomathematics*, Springer-Verlag: Berlin, 1984

[23] A. Korobeinikov, Global properties of SIR and SEIR epidemic models with multiple parallel infectious stages. *Bulletin of Mathematical Biology*, **71**(1) (2009), 75-83. Available from `http://www.ncbi.nlm.nih.gov/pubmed/18769976`

[24] M.G. Roberts and J.A.P. Heesterbeek, Model-consistent estimation of the basic reproduction number from the incidence of an emerging infection. *J. Math. Biol.*, **55**(5-6) (2007), 803-816.

[25] Unicef, *Malaria and Children: Progress in Intervention Coverage*, (2007), 5-78.

[26] P. van den Driessche, P. and J. Watmough, Reproduction numbers and subthreshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.* **180** (2002), 29-48.

[27] World Health Organisation, *Malaria Vector Control and Personal Protection*, Technical Report, WHO Technical Report series no. 936, 2006.

[28] M. Y. Li and L. Wang, Global Stability in Some SEIR Epidemic Models. In Castillo-Chavez, Carlos (ed.) et al., *Mathematical Approaches for Emerging and Re-Emerging Infectious Diseases: Models, Methods, and Theory.* Proceedings of a Workshop, Integral part of the IMA Program on Mathematics in Biology. New York, NY: Springer. IMA Vol. Math. Appl. **126** (2002), 295-311.

Maximal Subgroups of Unit Groups of Completely Primary Finite Rings

Chiteng'a J. Chikunji

Department of Basic Sciences, Botswana College of Agriculture, Private Bag 0027, Gaborone, Botswana. Email: jchikunj@bca.bw

Abstract. It is well known that the group of units of a completely primary finite ring R of order p^{nr} is a semi-direct product of a cyclic group of order p^r-1 by a p-group of order $p^{(n-1)r}$. The structure of the p-subgroup is not completely determined. In this paper, we investigate and determine the structure of the $p-$subgroup of the group of units of a commutative completely primary finite ring R of order p^{nr} with unique maximal ideal $\mathcal{J}$ such that $\mathcal{J}^m = (0)$, $\mathcal{J}^{m-1} \neq (0)$, and with characteristic p^k $(1 \leq k \leq m \leq n)$, for any prime number p and positive integers n, $m \geq 2$, k and r. We do this by first determining highest possible orders of elements in the p-subgroup under various conditions on the parameters p^k, n and m, for any r.

Keywords: Finite commutative rings, group of units, direct products of abelian p-groups
Classification (MSC2010): Primary 16P10, 13M05; Secondary 20K01, 20K25.

1. Introduction

Throughout this work we will assume that all rings are finite, commutative (unless otherwise stated) and associative with identities, denoted by $1 \neq 0$, that ring homomorphisms preserve 1, a ring and its subrings have the same 1 and modules are unital. Moreover, we adopt the notation used in [2] and [3], that is, R will denote a finite ring, unless otherwise stated, $\mathcal{J}$ will denote the Jacobson radical of R, and we will denote the Galois ring $GR(p^{kr},\ p^k)$ of characteristic p^k and order p^{kr} by R_o, for some prime integer p, and positive integers $k,\ r$. We denote the group of units of R by G_R; if g is an element of G_R, then $o(g)$ denotes its order, and $< g >$ denotes the cyclic group generated by g. Further, for a subset A of R or G_R, $|A|$ will denote the number of elements in A. The ring of integers modulo the number n will be denoted by $\mathbb{Z}_n$ and the characteristic of R will be denoted by char$R$. We shall use C_r^n or $\binom{n}{r}$, for the binomial coefficient $\frac{n!}{(n-r)!r!}$.

A finite ring R with identity $1 \neq 0$ is *completely primary* if the set $\mathcal{J}$ of all its zero-divisors including the zero element forms an additive group, and hence, its unique maximal ideal.

Completely primary finite rings have been studied in some detail by Raghavendran in [13], and for easy reference, we state without proofs in section 2, most of the results needed for our purpose.

In [8], Fuchs asked for a characterization of abelian groups which could be groups of units in a ring. This question was noted to be too general for a complete answer [14], and a natural course is to restrict the classes of groups or rings to be considered.

Let R be a ring and let G_R be its multiplicative group of unit elements. All local rings R with G_R cyclic were determined by Gilmer [10] and this case was also considered by Ayoub [1] (also proofs are given in [12] and [13]). Pearson and Schneider have found all R where G_R is generated by two elements. Clark [6] has investigated G_R where the ideals form a chain and has shown that if $p \geq 3$, $n \geq 2$ and $r \geq 2$, then the units of the Galois ring $GR(p^{nr},\ p^n)$ are a direct sum of a cyclic group of order $p^r - 1$ and r cyclic groups of order $p^n - 1$ (this was also done independently by Raghavendran [13]). In fact, Raghavendran described the structure of the multiplicative group of every Galois ring. Stewart in [14] considered a related problem to that asked by Fuchs [8] by proving that for a given finite group G (not necessarily abelian), there are, up to isomorphism, only finitely many directly indecomposable finite rings having group of units isomorphic to G.

Ganske and McDonald [9] provided a solution for G_R when the local ring R has Jacobson radical $\mathcal{J}$ such that $\mathcal{J}^2 = (0)$ by showing that

$$G_R = (\oplus \sum_{i=1}^{nt} \epsilon(p) \oplus \epsilon(|K| - 1),$$

where $n = dim_K(\mathcal{J}/\mathcal{J}^2)$, $|K| = p^t$, and $\epsilon(\pi)$ denotes the cyclic group of order π.

In [7], Dolzan found all non-isomorphic rings with a group of units isomorphic to a group G with n elements, where n is a power of a prime or any product of prime powers not divisible by 4; and also found all groups with n elements which can be groups of units of a finite ring, a contribution to Stewart's problem [14]. X.-D. Hou et al. [11] gave an algorithmic method for computing the structure of the group of units of a finite commutative chain ring and further strengthening the known result by listing a set of linearly independent generators for the group of units.

In [2], [3], [4] and [5] (see also Sections 2 and 3 in [13]), the author studied unit groups of commutative completely primary finite rings R with maximal ideals $\mathcal{J}$ such that $\mathcal{J}^3 = (0)$ and $\mathcal{J}^2 \neq (0)$, for various parameters $p,\ n,\ r,\ s,\ ,t$ and λ.

The purpose of the current paper is to extend the above study to a general commutative completely primary finite ring R of order p^{nr} with unique maximal ideal $\mathcal{J}$ such that $\mathcal{J}^m = (0)$, $\mathcal{J}^{m-1} \neq (0)$, characteristic of $R = p^k$, for any prime number p, and any fixed positive integers $m,\ n,\ r,\ k$, where $1 \leq k \leq m \leq n$. The highest possible order of an element in any such group will be determined. Explicit representations are given for the p-subgroup $1 + \mathcal{J}$ of the group of units of the ring R and directions for further research are indicated.

2. PRELIMINARIES

Let R be a completely primary finite ring (not necessarily commutative), $\mathcal{J}$ the set of all zero divisors in R, including the zero element, p a prime, k, n and r be positive integers. The following results will be assumed (see [13]): $|R| = p^{nr}$, $\mathcal{J}$ is the Jacobson radical of R, $\mathcal{J}^n = (0)$, $|\mathcal{J}| = p^{(n-1)r}$, $R/\mathcal{J} \cong GF(p^r)$, the finite field of p^r elements and $charR = p^k$, where $1 \leq k \leq n$; the group of units G_R is a semi-direct product $G_R = (1+\mathcal{J}) \times_\theta < b >$, of its normal subgroup $1+\mathcal{J}$ of order $p^{(n-1)r}$ by a cyclic subgroup $< b >$ of order $p^r - 1$; the multiplicative group G_R is solvable; if G is a subgroup of G_R of order $p^r - 1$, then G is conjugate to $< b >$ in G_R; if G_R contains a normal subgroup of order $p^r - 1$, then the set $K_o =< b > \cup \{0\}$ is contained in the center of the ring R; and $(1+\mathcal{J}^i)/(1+\mathcal{J}^{i+1}) \cong \mathcal{J}^i/\mathcal{J}^{i+1}$ (the left hand side as a multiplicative group and the right hand side as an additive group). If $\mathrm{char}R = p^k$ and $n = k$, it is known that, up to isomorphism, there is precisely one completely primary ring of order p^{kr} having characteristic p^k and residue field $GF(p^r)$. It is called the *Galois ring* $GR(p^{kr},\ p^k)$ and a concrete model is the quotient $\mathbb{Z}_{p^k}[X]/(f)$, where f is a monic polynomial of degree r, irreducible modulo p. Any such polynomial will do: the rings are all isomorphic. Trivial cases are $GR(p^n,\ p^n) = \mathbb{Z}_{p^n}$ and $GR(p^r,\ p) = \mathbb{F}_{p^r}$. In fact, $R = \mathbb{Z}_{p^n}[b]$, where b is an element of R of multiplicative order $p^r - 1$; $\mathcal{J} = pR$ and $Aut(R) \cong Aut(R/pR)$ (see Proposition 2 in [13]). Furthermore, if $\mathrm{char}R = p^k$, then it can be deduced from [13] that R has a coefficient subring R_o of the form $GR(p^{kr},\ p^k)$ which is clearly a maximal Galois subring of R. Moreover, if R_o' is another coefficient subring of R then there exists an invertible element x in R such that $R_o' = xR_ox^{-1}$ (see Theorem 8 in [13]). Furthermore, there exist elements $m_1,\ ...,\ m_h \in \mathcal{J}$ and $\sigma_1,\ ...,\ \sigma_h \in Aut(R_o)$ such that $R = R_o \oplus \sum_{i=1}^{h} R_o m_i$ (as R_o-modules), $m_i r_o = r_o^{\sigma_i} m_i$, for all $r_o \in R_o$ and any $i = 1,\ ...,\ h$ (use the decomposition of $R_o \otimes_{\mathbb{Z}} R_o$ in terms of $Aut(R_o)$ and apply the fact that R is a module over $R_o \otimes_{\mathbb{Z}} R_o$). Moreover, $\sigma_1,\ ...,\ \sigma_h$ are uniquely determined by R and R_o. We call σ_i the automorphism associated with m_i and $\sigma_1,\ ...,\ \sigma_h$ the associated automorphisms of R with respect to R_o.

Now, let $R_o = \mathbb{Z}_{p^k}[b]$ be a coefficient subring of R of order p^{kr} and characteristic p^k and let $K_o =< b > \cup \{0\}$, denote the set of coset representatives of $\mathcal{J}$ in R. Then it is easy to show that every element of R_o can be written uniquely as $\sum_{i=0}^{k-1} \lambda_i p^i$, where $\lambda_i \in K_o$.

Lemma 1. *Let R be a completely primary finite ring of characteristic p^k and order p^{nr}, with maximal ideal $\mathcal{J}$ of index of nilpotence i. Then $1 \leq k \leq i \leq n$.*

Proof. We have only to prove that $k \leq i$. We know that $\mathcal{J} \neq 0$ so that $i \geq 2$. Since $p^k = 0$, it follows that $p \in \mathcal{J}$ and $i \geq k$. ■

It can be seen from the above lemma that taking a fixed index of nilpotence produces limitations on the characteristic of R. Hence, if we take a completely primary finite ring with $\mathcal{J}^2 = (0)$, then $1 \leq k \leq 2$, i.e. R is either of characteristic p or p^2.

3. Constructions

We construct commutative completely primary finite rings R with Jacobson radicals $\mathcal{J}$ such that $\mathcal{J}^m = (0)$, $\mathcal{J}^{m-1} \neq (0)$, characteristic of $R = p^k$, for any prime number p, and any fixed positive integers $m,\ n,\ r,\ k$, where $1 \leq k \leq m \leq n$.

3.1 Rings of characteristic p

Let $R_o = \mathbb{F}_{p^r}$, the finite field with p^r elements for any prime number p and positive integer r. Let $U_1,\ U_2,\ ...,\ U_{m-1}$ be vector spaces over R_o and suppose that U_1 is generated by $u_1,\ \ldots,\ u_s$ over the field R_o; U_2 is generated by the $\frac{s(s+1)}{2}$ distinct products u_iu_j of the elements u_i, over the field R_o; and so on, and U_{m-1} is generated by the $\binom{s+m-2}{m-1} = \frac{s(s+1)\cdots(s+m-2)}{(m-1)!}$ distinct products of $m-1$ of the elements u_i over R_o.

Let R be an additive group direct sum of the R_o vector spaces:

$$R = R_o \oplus U_1 \oplus U_2 \oplus \cdots \oplus U_{m-1}.$$

Define multiplication on R as follows: $u_iu_j \neq 0$ is a linear combination of elements in $U_2 \oplus U_3 \oplus \cdots \oplus U_{m-1}$ over R_o; $u_iu_ju_k \neq 0$ is a linear combination of elements in $U_3 \oplus U_4 \oplus \cdots \oplus U_{m-1}$ over R_o; ...; $u_i \cdots u_{m-1} \neq 0$ is a linear combination of elements in U_{m-1} over R_o; and $u_ia = au_i$, for every $a \in R_o$.

Then it is easy to show that R is a commutative completely primary finite ring with Jacobson radical $\mathcal{J}$ such that $\mathcal{J}^m = (0)$, $\mathcal{J}^{m-1} \neq (0)$, and $\mathrm{char}R = p$.

Clearly, $\mathcal{J} = U_1 \oplus U_2 \oplus \cdots \oplus U_{m-1}$, $\mathcal{J}^{m-1} = U_{m-1}$ and

$$|\mathcal{J}| = \prod_{i=1}^{m-1} p^{\binom{s+i-1}{i}r} = p^{\sum_{i=1}^{m-1}\binom{s+i-1}{i}r},$$

where $\binom{s+i-1}{i}r = n-1$.

3.2 Rings with characteristic p^k, $k > 1$

We assume that $p \in \mathcal{J} - \mathcal{J}^2$ so that $p^2 \in \mathcal{J}^2$ and $p^{k-1} \in \mathcal{J}^{k-1}$. We leave the other cases for future consideration (See Remark 4.5 below).

Let R_o be the Galois ring $GR(p^{kr},\ p^k)$ of order p^{kr} and with characteristic p^k, for a fixed prime p and fixed positive integers $k,\ r$. Let $U_1,\ U_2,\ ...,\ U_{m-1}$ be R_o-modules and suppose that $u_1,\ u_2,\ ...,\ u_s$ generate U_1 as an R_o-module, the $\frac{s(s+1)}{2}$ distinct products u_iu_j $(i,j = 1,\ ...,\ s)$ generate U_2 as an R_o-module, and so on, and U_{m-1} is generated as an R_o-module by the C_{m-1}^{s+m-2} distinct products of $m-1$ of the elements u_i.

Let R be an additive group direct sum of the R_o-modules:

$$R = R_o \oplus U_1 \oplus U_2 \oplus\ \cdots\ \oplus U_{m-1}.$$

Define multiplication on R such that $u_iu_j \neq 0$ is a linear combination of elements in $p^2R_o \oplus pU_1 \oplus U_2 \oplus \cdots \oplus U_{m-1}$ over R_o/pR_o; $u_iu_ju_k \neq 0$ is a linear combination of elements in $p^3R_o \oplus p^2U_1 \oplus pU_3 \oplus U_4 \oplus \cdots \oplus U_{m-1}$ over R_o/pR_o; ...; $u_i \cdots u_{m-1} \neq 0$ is a linear combination of elements in $p^{m-1}R_o \oplus pU_{m-2} \oplus U_{m-1}$ (if $k < m$) or in U_{m-1} (if $k \geq m$), over R_o/pR_o; and $u_ia = au_i$, for every $a \in R_o$.

Then it is routine to verify that R is a commutative completely primary finite ring with Jacobson radical $\mathcal{J}$ such that $\mathcal{J}^m = (0)$, $\mathcal{J}^{m-1} \neq (0)$ and with char$R = p^k$. Moreover,

$$\mathcal{J} = pR_o \oplus U_1 \oplus U_2 \oplus \ \cdots \ \oplus U_{m-1},\ \mathcal{J}^2 = p^2R_o \oplus pU_1 \oplus U_2 \oplus \ \cdots \ \oplus U_{m-1},$$

$\mathcal{J}^{m-1} = p^{m-1}R_o \oplus pU_{m-2} \oplus U_{m-1}$, (if $k < m$) or $\mathcal{J}^{m-1} = U_{m-1}$, (if $k \geq m$).

We note here that the set

$$\{1,\ p,\ p^2,\ ...,\ p^{k-1},\ u_i,\ pu_i,\ u_iu_j,\ pu_iu_j,\ ...,\ u_1 \cdots u_{m-1}\ (m = 1,\ ...,\ s)\}$$

forms a "basis" for R over $R_o/pR_o \cong GF(p^r)$.

4. Structure of the p-subgroup $(1 + \mathcal{J})$

Throughout this section, we assume that R is a commutative completely primary finite ring, $|R| = p^{nr}$, $|\mathcal{J}| = p^{(n-1)r}$, $R/\mathcal{J} \cong GF(p^r)$, $\mathcal{J}^m = (0)$, $\mathcal{J}^{(m-1)} \neq (0)$, and char$R = p^k$, where $1 \leq k \leq m \leq n$.

It is well know that the group of units of a commutative completely primary finite ring R is $G_R \cong < b > \times(1 + \mathcal{J})$, where $< b >$ is cyclic of order $p^r - 1$ and $(1 + \mathcal{J})$ is a $p-$ group of order $p^{(n-1)r}$. In this paper, we investigate and determine the structure of the p-subgroup $1 + \mathcal{J}$ and hence, the structure of G_R. We do this by first determining highest possible orders of elements in $1 + \mathcal{J}$.

4.1 Rings of characteristic p

Lemma 2. *Let R be a ring defined in* 3.1 *with characteristic p and suppose $\mathcal{J}^m = (0)$, $\mathcal{J}^{m-1} \neq (0)$, where $m \leq n$. Let $x \in \mathcal{J} - \mathcal{J}^2$.*
(i) If char$R = p \geq m$, then the highest possible order of $1 + x$ in $1 + \mathcal{J}$ is p.
(ii) If char$R = p < m$, then the highest possible order of $1 + x$ in $1 + \mathcal{J}$ is $p^l \geq m$, for some least positive integer $l > 1$.

Proof. Since the binomial coefficient $\binom{p}{k} \equiv 0 \pmod p$ for every p, where $k = 1,\ ...,\ p - 1$, it follows that $(1 + x)^p = 1 + x^p$.

If $p \geq m$, then $x^p = 0$ and $(1 + x)^p \equiv 1 \pmod p$. This proves the first result.

Now, if $p < m$, then $(1 + x)^p = 1 + x^p \neq 1 \pmod p$, and x^p is a non-zero element of $\mathcal{J}$. Next, $(1 + x)^p(1 + x)^p = 1 + x^{p^2}$. Here, either $x^{p^2} = 0$ if $p^2 \geq m$, in which case $(1 + x)^{p^2} \equiv 1 \pmod p$, or $x^{p^2} \in \mathcal{J}$ is a non-zero element. By the induction process, there exists a least positive integer $l > 1$ such that $m < p^l$ with $m > p^{l-1}$, and $(1 + x)^{p^l} = 1 + x^{p^l} \equiv 1 \pmod p$. ■

The case when char$R = p \geq m$

Proposition 1. *Let R be a commutative completely primary finite ring defined in* 3.1*, of order p^{nr} and with characteristic p. Suppose that $\mathcal{J}^m = (0)$, $\mathcal{J}^{m-1} \neq (0)$ where $m \leq n$. If $p \geq m$ then*

$$1 + \mathcal{J} \cong \underbrace{\mathbb{Z}_p^r \times \ \cdots \ \times \mathbb{Z}_p^r}_{n-1},$$

and hence,

$$G_R \cong \mathbb{Z}_{p^r-1} \times \underbrace{\mathbb{Z}_p^r \times \ \cdots \ \times \mathbb{Z}_p^r}_{n-1}.$$

Proof. Let $\varepsilon_1,\ \varepsilon_2,\ ...,\ \varepsilon_r$ be elements of R_o with $\varepsilon_1 = 1$ so that $\overline{\varepsilon_1},\ \overline{\varepsilon_2},\ ...,\ \overline{\varepsilon_r}$ form a basis of $R_o = GF(p^r)$ over its prime subfield $GF(p)$. First notice that, for $1+\varepsilon_j w_{ik_i} \in 1+\sum_{i=1}^{m-1} \oplus U_i$, and for each $j = 1,\ ...,\ r$; $(1+\varepsilon_j w_{ik_i})^p = 1$ and $g^p = 1$ for all $g \in 1 + \sum_{i=1}^{m-1} \oplus U_i$, where p is a prime integer such that $p =$char$R$, and $k_i = 1,\ 2,\ ...,\ C_{m-1}^{s+m-2}$.

For integers $l_j,\ m_j,\ ...,\ n_j \leq p$, we assert that the equation

$$\prod_{j=1}^{r}\prod_{k_i=1}^{s}\left\{(1+\varepsilon_j w_{1k_i})^{l_j}\right\} \times \prod_{j=1}^{r}\prod_{k_i=s+1}^{\frac{s(s+1)}{2}}\left\{(1+\varepsilon_j w_{2k_i})^{m_j}\right\} \times$$
$$\cdots \times \prod_{j=1}^{r}\prod_{k_i=C_{(m-2)}^{(s+m-3)}}^{C_{(m-1)}^{(s+m-2)}}\left\{(1+\varepsilon_j w_{m-1k_i})^{n_j}\right\} = 1,$$

will imply that $l_j = m_j = \ ... = n_j = p$, for all $j = 1,\ ...,\ r$.

If we set

$$\begin{aligned}
F_j &= \left\{(1+\varepsilon_j w_{1k_i})^l : l = 1,\ ...,\ p;\ k_i = 1,\ ...,\ s\right\},\\
G_j &= \left\{(1+\varepsilon_j w_{2k_i})^m : m = 1,\ ...,\ p;\ k_i = s+1,\ ...,\ s(s+1)/2\right\},\\
&\vdots\\
H_j &= \left\{(1+\varepsilon_j w_{m-1k_i})^n : n = 1,\ ...,\ p;\ k_i = C_{m-2}^{s+m-3},\ ...,\ C_{m-1}^{s+m-2}\right\},
\end{aligned}$$

for all $j = 1,\ ...,\ r$; we see that $F_j,\ G_j,\ ...,\ H_j$ are all cyclic subgroups of $1 + \sum_{i=1}^{m-1} \oplus U_i$ and these are all of order p as indicated in their definition. Intersection of any pair of these subgroups is trivial. The argument above will show that the product of the $\sum_{i=1}^{m-1} \binom{s+i-1}{i} r$ subgroups $F_j,\ G_j,\ ...,$ and H_j is direct. So, their product will exhaust $1 + \sum_{i=1}^{m-1} \oplus U_i$. ■

Corollary 1. *If some of the products in the generating sets for $U_2,\ ...,\ U_{m-1}$ are zeros, and assume that $|\mathcal{J}| = p^{(n-1)r}$. Then*

$$1 + \mathcal{J} \cong \underbrace{\mathbb{Z}_p^r \times \ \cdots \ \times \mathbb{Z}_p^r}_{n-1}.$$

The case when char$R = p < m$

Recall that if char$R = p < m$, then the highest possible order of $a = 1 + x$, with $x \in \mathcal{J} - \mathcal{J}^2$ is $p^l \geq m$, for some least positive integer $l > 1$ [Lemma 4.1]. We give a general description of the structure of $1 + \mathcal{J}$. This can be proved in a similar manner to Proposition 4.2 after noting that $o(1 + x) = p^l$, $l > 1$ since $p < m$.

Proposition 2. *Let R be a commutative completely primary finite ring defined in* 3.1. *Suppose that $\mathcal{J}^m = (0)$ and $\mathcal{J}^{m-1} \neq (0)$, where $m \leq n$. If $p < m$ then,*

$$1 + \mathcal{J} \cong \mathbb{Z}_{p^l}^r \times \mathbb{Z}_{p^{l_1}}^r \times \mathbb{Z}_{p^{l_2}}^r \times \cdots \times \mathbb{Z}_{p^{l_t}}^r,$$

where $l > 1$, $l_i \geq 1$ are positive integers such that $l \geq l_1 \geq l_2 \geq \cdots \geq l_t$ and $l + l_1 + l_2 + \cdots + l_t = n - 1$.

Proof. Similar to Proposition 4.2 with some modifications. ■

Example 1. Suppose $p = 3$, $m = 3$ or $m = 2$ and $n - 1 = 4$, with r arbitrary. Then highest possible order of an element in $1 + \mathcal{J}$ is 3, since $m \leq p = 3$. Now $|1 + \mathcal{J}| = 3^{4r}$ and we have only one partition of 4 in this case in which 1 is the largest number, namely $1 + 1 + 1 + 1$. Therefore,

$$1 + \mathcal{J} \cong \mathbb{Z}_3^r \times \mathbb{Z}_3^r \times \mathbb{Z}_3^r \times \mathbb{Z}_3^r.$$

Example 2. Suppose $p = 3$, $m = 6$ and $n - 1 = 10$, with r arbitrary. Then highest order of an element in $1 + \mathcal{J}$ is 3^3, since $m = 6 > p$. Now $|1 + \mathcal{J}| = 3^{10r}$ and partitions of 10 in this case in which 3 is the largest number are $3+3+3+1$; $3 + 3 + 2 + 2$; $3 + 3 + 2 + 1 + 1$; $3 + 3 + 1 + 1 + 1 + 1$; and so on and lastly $3+1+1+1+1+1+1+1$. Therefore,

$$1 + \mathcal{J} \cong \begin{cases} \mathbb{Z}_{3^3}^r \times \mathbb{Z}_{3^3}^r \times \mathbb{Z}_{3^3}^r \times \mathbb{Z}_3^r; \\ \mathbb{Z}_{3^3}^r \times \mathbb{Z}_{3^3}^r \times \mathbb{Z}_{3^2}^r \times \mathbb{Z}_{3^2}^r; \\ \cdots \qquad \text{or} \\ \mathbb{Z}_{3^3}^r \times \mathbb{Z}_3^r \times \mathbb{Z}_3^r \times \mathbb{Z}_3^r \times \mathbb{Z}_3^r \times \mathbb{Z}_3^r \times \mathbb{Z}_3^r. \end{cases}$$

4.2 Rings with characteristic p^k, $k > 1$

Remark 1. *Let R be a commutative completely primary finite ring of order p^{nr} with Jacobson radical $\mathcal{J}$. If $\mathcal{J}^m = (0)$, $\mathcal{J}^{m-1} \neq (0)$, and char$R = p^k$ $(1 < k \leq m \leq n)$, and since $py = 0$ for every $y \in \mathcal{J}^{m-1}$, it is easy to check that either*
(i) $p \in \mathcal{J}$ and $p^{(k-1)} \in \mathcal{J}^{(k-1)}$ $(k - 1 < m \leq n)$;
(ii) $p \in \mathcal{J}^i$, $i \leq m - 1$ and $p^{k-1} \in \mathcal{J}^{(k-i)}$ $((k - i) < m)$; or
(iii) $p \in \mathcal{J}^i$, $i \leq m - 1$ and $p^l < p^k$, for some positive integer $l < k$.

For example, if char$R = p^2$ and $\mathcal{J}^3 = (0)$, then either $p \in \mathcal{J}$ or $p \in \mathcal{J}^2$. On the other hand, if char$R = p^3$ and $\mathcal{J}^3 = (0)$, then $p \in \mathcal{J}$, $p^2 \in \mathcal{J}^2$. However, $p \notin \mathcal{J}^2$ since in this case, $p^2 \in \mathcal{J}^4 = (0)$, implying that $p^2 \equiv 0 \pmod{p^3}$, which contradicts the definition of the characteristic of R.

In this section, we consider rings satisfying case (i) only, i.e. rings of characteristic p^k in which $p \in \mathcal{J}$ and $p^{(k-1)} \in \mathcal{J}^{(k-1)}$ $(k-1 < m \leq n)$, and we leave the other cases for further work.

Let $x \in \mathcal{J} - \mathcal{J}^2$ and consider the highest possible order of $1+x$ in $1+\mathcal{J}$.

Suppose that $\mathrm{char}R = p^k$, where p is any prime number and k is a fixed positive integer. Then

$$\begin{aligned}(1+x)^{p^k} &= 1 + p^k x + \frac{p^k(p^k-1)}{2}x^2 + \frac{p^k(p^k-1)(p^k-2)}{3!}x^3 + \cdots + x^{p^k} \\ &= 1 + \sum_{i=1}^{p^{(k-1)}} \binom{p^k}{pi} x^{pi} + x^{p^k}\end{aligned}$$

since $\binom{p^k}{pi} \equiv 0 \pmod{p^k}$, for every non-multiple of p.

If $p \geq m$ then $1 + \sum_{i=1}^{p^{(k-1)}} \binom{p^k}{pi} x^{pi} + x^{p^k} \equiv 1 \pmod{p^k}$, and thus, $o(1+x) = p^k$. (If $m = 2$, then $o(1+x) = 2$, since $2x = 0$ in this case.)

If, however, $p < m$ but $p^k \geq m$, then $(1+x)^{p^k} \ncong 1(mod\ p^k)$. But $(1+x)^{p^k} \cdots (1+x)^{p^k} = (1+x)^{p^{(k+1)}} \equiv 1(mod\ p^k)$ since $\binom{p^{(k+1)}}{i} \equiv 0 \pmod{p^k}$, for all $i = 1,\ ...,\ p^{(k+1)} - 1$. Therefore, if $p^k \geq m$, then $1 + x^{p^{(k+1)}} \equiv 1(mod\ p^k)$ so that $o(1+x) = p^{(k+1)}$. (If $m = 2^k$ or $2^{(k-1)}x = 0$, then $o(1+x) = 2^k$, and if $m < 2^k$, then $o(1+x) = 2^{(k+1)}$ as for any prime p.)

Now, if $p^k < m$, then $x^{p^k} \neq 0$, and we proceed by induction on powers of $1+x$, i.e. $(1+x)^{p^k} \cdots (1+x)^{p^k} = 1 + x^{p^{(k+l)}}$ (l times) and this will be $\equiv 1 \pmod{p^k}$) for some least positive integer l such that $p^{(k+l)} \geq m$. (Note that any power p^s with $s < k$ or a non-multiple of p^k will not work since the characteristic of R is p^k.)

Lemma 3. *Let R be a commutative completely primary finite ring defined in* 3.2 *of order p^{nr} with Jacobson radical $\mathcal{J}$. Suppose $\mathcal{J}^m = (0)$, $\mathcal{J}^{m-1} \neq (0)$ and $charR = p^k$ $(1 < k \leq m \leq n)$,. Let $x \in \mathcal{J} - \mathcal{J}^2$.*

If $p \geq m$, then highest possible order of $1+x$ is p^k;

if $p < m$ but $p^k \geq m$, then highest possible order of $1+x$ is $p^{(k+1)}$; and

if $charR = p^k < m$, then highest possible order of $1+x$ is $p^{(k+l)}$, for some least positive integer l such that $p^{(k+l)} \geq m$.

Note that in the above Lemma, 2 cannot be greater than m, so in that case, when $m = 2$, then $o(1+x) = 2$.

Proof. Follows from the above discussion. ■

The case when $\mathrm{char}R = p^k$, with $p \geq m$ and $p^k \geq m > p$

Following the variations in the discussion before Lemma 4.6, from now on, we consider the case when the prime number $p \neq 2$, and leave the case when $p = 2$ for future work. However, in the following results, we do not suppose that $pu_i = 0$ for $i = 1,\ ...,\ s$, since we are only giving a general representation of the group $1+\mathcal{J}$.

Proposition 3. *Let R be a commutative completely primary finite ring defined in* 3.2 *with characteristic p^k. Suppose that $\mathcal{J}^m = (0)$, $\mathcal{J}^{m-1} \neq (0)$ with $1 < k \leq m \leq n$. If $p \geq m$ then,*

$$1 + \mathcal{J} \cong \mathbb{Z}_{p^k}^r \times \mathbb{Z}_{p^{k_1}}^r \times \ \cdots \ \times \mathbb{Z}_{p^{k_t}}^r,$$

where $k > 1$ and $k_i \geq 1$ are positive integers such that $k \geq k_1 \geq \ \cdots \ \geq k_t$ and $k + k_1 + \cdots + k_t = n - 1$.

Proof. Let $p \in \mathcal{J} - \mathcal{J}^2$ so that $p^2 \in \mathcal{J}^2$ and $p^{(k-)} \in \mathcal{J}^{(k-1)}$. Assume that $x \in \mathcal{J} - \mathcal{J}^2$ and let $a = 1 + x$ be an element of $1 + \mathcal{J}$ with the highest possible order. Then $o(a) = p^k$ if $p \geq m$ or $o(a) = p^{(k+1)}$ if $p^k \geq m > p$ (see Lemma 4.6), for any prime number $p \neq 2$.

Notice that $1 + \mathcal{J} = (1 + pR_o) \times (1 + U_1 \oplus U_2 \oplus \ \cdots \ \oplus U_{m-1})$ if $u_i u_j$ and any product from U_i has a zero coefficient in pR_o.

In fact, in this case, the structure of $1 + pR_o$ is well known; it is the direct product of r cyclic groups each of order $p^{(k-1)}$. All that is needed is to determine the structure of $1 + U_1 \oplus \ \cdots \ \oplus U_{m-1}$.

However, we opt to determine the structure of $1 + \mathcal{J}$ for the case where

$$u_i u_j \in p^2 R_o \oplus pU_1 \oplus U_2 \oplus \ \cdots \ \oplus U_{m-1}.$$

In this case $o(1 + p) = p^{(k-1)}$ (see above), and it is not the highest possible.

Let $\varepsilon_1,\ \varepsilon_2,\ ...,\ \varepsilon_r$ be elements of R_o with $\varepsilon_1 = 1$ so that $\overline{\varepsilon_1},\ \overline{\varepsilon_2},\ ...,\ \overline{\varepsilon_r} \in R_o/pR_o \cong GF(p^r)$ form a basis of $GF(p^r)$ over its prime subfield $GF(p)$.

Suppose that $p \geq m$. First notice that, for $1 + \varepsilon_j p \in 1 + pR_o$ and for each $j = 1,\ ...,\ r$; $(1+\varepsilon_j p)^{p^{(k-1)}} = 1$; for $1+\varepsilon_j w_{ik_i} \in 1+U_i$ $(i = 1,\ \ldots,\ k)$, and for each $j = 1,\ ...,\ r$; $(1+\varepsilon_j w_{ik_i})^{p^{(k-i+1)}} = 1$; for $1+\varepsilon_j w_{ik_i} \in 1+U_i$ $(i = k+1,\ \ldots,\ m-1)$, and for each $j = 1,\ ...,\ r$; $(1+\varepsilon_j w_{ik_i})^p = 1$; and $g^{p^k} = 1$ for all $g \in 1+pR_o \oplus \sum_{i=1}^{m-1} \oplus U_i$, where p is a prime integer such that $p^k = \mathrm{char} R$, and $k_i = 1,\ 2,\ ...,\ C_{m-1}^{s+m-2}$.

For integers $l_j \leq p^{(k-1)},\ m_j \leq p^k,\ n_j \leq p^{(k-1)},\ ...\ k_j \leq p$ we assert that

$$\prod_{j=1}^{r}(1+\varepsilon_j p)^{l_j} \prod_{j=1}^{r}\prod_{k_i=1}^{s}\{(1+\varepsilon_j w_{1k_i})^{m_j}\} \times \prod_{j=1}^{r}\prod_{k_i=s+1}^{\frac{s(s+1)}{2}}\{(1+\varepsilon_j w_{2k_i})^{n_j}\} \times$$
$$\cdots \times \prod_{j=1}^{r}\prod_{k_i=C_{(m-2)}^{(s+m-3)}}^{C_{(m-1)}^{(s+m-2)}}\left\{(1+\varepsilon_j w_{m-1k_i})^{k_j}\right\} = 1,$$

will imply that $l_j = p^{k-1},\ m_j = p^k,\ n_j = p^{(k-1)},\ ...,\ k_j = p$ for all $j = 1,\ ...,\ r$.

If we set

$$\begin{aligned}
E_j &= \left\{(1+\varepsilon_j p)^l : l = 1,\ ...,\ p^{(k-1)}\right\},\\
F_j &= \left\{(1+\varepsilon_j w_{1k_i})^m : m = 1,\ ...,\ p^k;\ k_i = 1,\ ...,\ s\right\},\\
G_j &= \left\{(1+\varepsilon_j w_{2k_i})^n : n = 1,\ ...,\ p^{(k-1)};\ k_i = s+1,\ ...,\ \frac{s(s+1)}{2}\right\},\\
&\vdots\\
H_j &= \left\{(1+\varepsilon_j w_{m-1k_i})^k : k = 1,\ ...,\ p;\ k_i = C_{m-2}^{s+m-3},\ ...,\ C_{m-1}^{s+m-2}\right\},
\end{aligned}$$

for all $j = 1,\ ...,\ r$; we see that $E_j,\ F_j,\ G_j,\ ...,\ H_j$ are all cyclic subgroups of $1+pR_o \oplus \sum_{i=1}^{m-1} \oplus U_i$ and these are all of the orders as indicated in their definition. Intersection of any pair of these subgroups is trivial. The argument above will show that the product of the subgroups $E_j,\ F_j,\ G_j,\ ...,$ and H_j is direct. So, their product will exhaust $1 + pR_o \oplus \sum_{i=1}^{m-1} \oplus U_i$. ■

Proposition 4. *Let R be a commutative completely primary finite ring defined in* 3.2 *with characteristic p^k. Suppose that $\mathcal{J}^m = (0)$, $\mathcal{J}^{m-1} \neq (0)$ with $1 < k \leq m \leq n$. If $p^k \geq m > p$ then,*

$$1 + \mathcal{J} \cong \mathbb{Z}^r_{p^{(k+1)}} \times \mathbb{Z}^r_{p^{k_1}} \times \ \cdots \ \times \mathbb{Z}^r_{p^{k_t}},$$

where $k > 1$ and $k_i \geq 1$ are positive integers such that $(k+1) \geq k_1 \geq \ \cdots \ \geq k_t$ and $(k+1) + k_1 + \cdots + k_t = n - 1$.

Proof. This can be proved in a similar manner as 4.7 by first noting that for $1+\varepsilon_j w_{ik_i} \in 1+U_i$ $(i = 1,\ \ldots,\ k-1)$, and for each $j = 1,\ ...,\ r$; $(1+\varepsilon_j w_{ik_i})^{p^{(k-i)}} = 1$. ■

Example 3. Suppose $\mathrm{char}R = p^k = 3^2$, $p \geq m = 3$ and $n - 1 = 6$, with r arbitrary. Then highest order of an element in $1 + \mathcal{J}$ is 3^2, since $m = 3$. Now $|1 + \mathcal{J}| = 3^{6r}$ and partitions of 6 in this case in which 2 is the largest number are $2+2+2$; $2+2+1+1$; and $2+1+1+1+1$. Therefore,

$$1 + \mathcal{J} \cong \begin{cases} \mathbb{Z}^r_{3^2} \times \mathbb{Z}^r_{3^2} \times \mathbb{Z}^r_{3^2}; \\ \mathbb{Z}^r_{3^2} \times \mathbb{Z}^r_{3^2} \times \mathbb{Z}^r_3 \times \mathbb{Z}^r_3; \\ \ldots \qquad\qquad \text{or} \\ \mathbb{Z}^r_{3^2} \times \mathbb{Z}^r_3 \times \mathbb{Z}^r_3 \times \mathbb{Z}^r_3 \times \mathbb{Z}^r_3. \end{cases}$$

Example 4. Suppose $\mathrm{char}R = p^k = 5^2$, $m = 6 > p$ and $n - 1 = 6$, with r arbitrary. Then highest order of an element in $1 + \mathcal{J}$ is 5^3, since $m = 6$. Now $|1 + \mathcal{J}| = 5^{6r}$ and partitions of 6 in this case with 3 the largest number are $3+3$; $3+2+1$; and $3+1+1$. Therefore,

$$1 + \mathcal{J} \cong \begin{cases} \mathbb{Z}^r_{5^3} \times \mathbb{Z}^r_{5^3}; \\ \mathbb{Z}^r_{5^3} \times \mathbb{Z}^r_{5^2} \times \mathbb{Z}^r_5; \quad \text{or} \\ \mathbb{Z}^r_{5^3} \times \mathbb{Z}^r_5 \times \mathbb{Z}^r_5 \times \mathbb{Z}^r_5. \end{cases}$$

The case when char$R = p^k$ and $p^k < m$

Proposition 5. *Let R be a commutative completely primary finite ring defined in* 3.2 *and of characteristic p^k. Suppose that $\mathcal{J}^m = (0)$, $\mathcal{J}^{m-1} \neq (0)$ with $1 < k \leq m$. If $p^k < m$ then,*

$$1 + \mathcal{J} \cong \mathbb{Z}^r_{p^{(k+l)}} \times \mathbb{Z}^r_{p^{k_1}} \times \ \cdots \ \times \mathbb{Z}^r_{p^{k_t}},$$

where $k > 1$ and $k_i \geq 1$ are positive integers such that $(k+l) \geq k_1 \geq \ \cdots \ \geq k_t$ and $(k+l) + k_1 + \cdots + k_t = n - 1$.

Proof. This can be proved in a similar manner to Proposition 4.7 after noticing that if $x \in \mathcal{J} - \mathcal{J}^2$ and $a = 1 + x$ is an element of $1 + \mathcal{J}$ with the highest possible order, then $o(a) = p^{(k+l)}$ since $p^k < m$ (see Lemma 4.6), for any prime number $p \neq 2$. ■

Example 5. Suppose char$R = p^k = 3^2$, $m = 10 > p^k$ and $n - 1 = 11$, with r arbitrary. The order of $1 + \mathcal{J}$ is 3^{11r}. Then highest order of an element in $1 + \mathcal{J}$ is 3^3, since $3x^9 = 0$ and $x^9 \in \mathcal{J}^9$. If, however $m = 11$, then $3x^9 \neq 0$, in which case, the order of $(1 + x)$ is 3^9 so that the least positive integer $l = 7$. The partitions of 11 in the first part with 3 being the largest integer are $3+3+3+2$, $3+3+3+1+1$, $3 + 3 + 2 + 2 + 1$, and so on and lastly $3 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1$; and in the second part with 9 being the largest integer are $9 + 2$ and $9 + 1 + 1$. These correspond with the isomorphism classes of the p-group $1 + \mathcal{J}$.

The case when char$R = p^k$, with $k = m$

Proposition 6. *Let R be a commutative completely primary finite ring of order p^{nr} with characteristic p^k defined in* 3.2, *and suppose that $\mathcal{J}^m = (0)$, $\mathcal{J}^{m-1} \neq (0)$ with $1 < k \leq m \leq n$. If $p^k = p^m$ and $m > 2$, then*

$$1 + \mathcal{J} \cong \mathbb{Z}^r_{p^{(k-1)}} \times \mathbb{Z}^r_{p^{k_1}} \times \ \cdots \ \times \mathbb{Z}^r_{p^{k_t}},$$

where $k > 1$ and $k_i \geq 1$ are positive integers such that $(k-1) \geq k_1 \geq \ \cdots \ \geq k_t$ and $(k-1) + k_1 + \cdots + k_t = n - 1$.

Proof. This can be proved in a similar manner to Proposition 4.7 after noticing that if $x \in \mathcal{J} - \mathcal{J}^2$ and $a = 1 + x$ is an element of $1 + \mathcal{J}$ with the highest possible order, then $o(a) = p^{(k-1)}$ since $p^k = p^m$ (see Lemma 4.6), for any prime number $p \neq 2$ and $m > 2$, and $xp^{(k-1)} = 0$ being an element in $\mathcal{J}^m = (0)$. ■

Remark 2. *It can be seen from the above proposition that the structure of the unit group $1 + pR_o$ of any Galois ring $R_o = GR(p^{kr}, p^k)$ coincides with that determined by Raghavendran in* [13].

Example 6. Suppose char$R = p^k = 3^3$, $m = k = 3$ and $n - 1 = 4$, with r arbitrary. The order of $1 + \mathcal{J}$ is 3^{4r}. Then possible highest order of an element in $1 + \mathcal{J}$ is 3^2, since $3^2 x = 0$ and $3^2 \in \mathcal{J}^2$. The partitions of 4 with 2 the largest

integer in this case are $2+2$ and $2+1+1$. These correspond with the isomorphism classes of the p-group $1+\mathcal{J}$, i.e.

$$1+\mathcal{J} \cong \begin{cases} \mathbb{Z}_{3^2}^r \times \mathbb{Z}_{3^2}^r; & \text{or} \\ \mathbb{Z}_{3^2}^r \times \mathbb{Z}_3^r \times \mathbb{Z}_3^r. \end{cases}$$

Remark 3. *The case where* $1+\mathcal{J} = (1+pR_o) \times (1+U_1 \oplus U_2 \oplus \ \cdots \ \oplus U_{m-1})$ *with* u_iu_j *or any product from* U_i *having a zero coefficient in* pR_o *can be deduced from the above propositions.*

It remains to determine the structure of $1+\mathcal{J}$ in the cases when

(i) $\mathrm{char}R = 2^k$ and $2 \in \mathcal{J}$;

(ii) $\mathrm{char}R = p^k,\ k > 1$, $p \in \mathcal{J}^i$, $i \leq m-1$ and $p^{k-1} \in \mathcal{J}^{(k-i)}$ $(k-i < m)$; and

(iii) $\mathrm{char}R = p^k,\ k > 1$, $p \in \mathcal{J}^i$, $i \leq m-1$ and $p^l \leq p^k$, for some positive integer $l < k$.

Bibliography

[1] C. W. Ayoub, On Finite Primary Rings and their Groups of Units, *Compos. Math.*, **21** (1969), 247-252.

[2] C. J. Chikunji, Unit groups of cube radical zero commutative completely primary finite rings, *Int. J. Math. Math. Sci.*, **4** (2005), 579-592.

[3] C. J. Chikunji, Unit groups of a certain class of completely primary finite rings, *Math. J. Okayama Univ.*, **47** (2005), 39-53.

[4] C. J. Chikunji, On Unit Groups of Completely Primary Finite Rings, *Math. J. Okayama Univ.*, **50** (2008), 149-160.

[5] C. J. Chikunji, Unit Groups of a Certain Class of Finite Rings of Characteristic p^2, *Int. Elect. J. Pure & Appl. Math.*, (to appear).

[6] W. E. Clark, A Coefficient Ring for Finite Non-Commutative Rings, *Proc. Amer. Math. Soc.*, **33** (1972), 25-28.

[7] D. Dolzan, Group of Units in a Finite Ring, *J. Pure Appl. Algebra*, **170** (2002), 175-183.

[8] L. Fuchs, *Abelian Groups*, 3rd ed., International Series of Monographs on Pure and Applied Mathematics, Pergamon Press, New York, 1960.

[9] G. Ganske and B. R. McDonald, Finite Local Rings, *Rocky Mountain J. Math.*, **3**(4) (1973), 521-540.

[10] R. W. Gilmer, Finite Rings having a Cyclic Multiplicative Group of Units, *Amer. J. Math.*, **85** (1963), 447-452.

[11] X.-D. Hou, K. H. Leung & S. L. Ma, On the Group of Units of Finite Commutative Chain Rings, *Finite Fields Appl.*, **9** (2003), 20 - 38.

[12] K. W. Pearson and J. E. Schneider, Rings with a Cyclic Group of Units, *J. Algebra*, **16** (1970), 243-251.

[13] R. Raghavendran, Finite Associative Rings, *Compos. Math.*, **21** (1969), 195-229.

[14] I. Stewart, Finite Rings with a Specified Group of Units, *Math. Z.*, **126** (1972), 51-58.

On the future of HIV-Related Lymphomas (HRLs) in the Western Cape Province of South Africa

Alana De la Harpe & Farai Nyabadza

Department of Mathematical Science,
University of Stellenbosch, Private Bag X1, Matieland, 7602, South Africa.

Abstract. HIV is dramatically changing the demographics of diseases and the resource needs of the entire health care service of South Africa. Several malignancies have an increased risk of developing in Persons Living with HIV (PLWH) and one typical example is HIV related Lymphomas (HRLs). The dynamics and trends of HRLs, thus need to be understood and determined. To track the patterns in the population, there is need to develop mathematical models that aid in making projections of HRLs, on the basis of being informed by the current data being collected. In this paper, a dynamic compartmental model detailing the co-infection dynamics of HRLs and HIV is formulated. The model is fitted to data obtained from the Tygerberg Lymphoma Study Group (TLSG) using the least squared method. Projections are made to depict the likely trends of HRL cases beyond 2010. For specific parameter values the model is found to fit very well to the data that is available. Projections show a continued increase in Lymphoma cases till the year 2020. Infact, the projections show that the number of cases would have almost doubled by the year 2020. Scatter plots are obtained that show the relationships between specific parameters and an outcome variable $A(t)$, representing individuals with AIDS at any time t. The results show that control of HIV-HRLs co-infection is important. We have demonstrated that a very simple model of five compartments can be used as a tool to reproduce the trends of HRLs in the Western Cape province of South Africa. The model helps to improve our understanding of how HIV is transforming the incidence, pattern, prognosis and outcomes of Lymphoproliferative Disorders in the Tygerberg catchments area of the Western Province of South Africa. This has health care implications on the planning and allocation of resources.

Keywords: HIV-related lymphomas, model simulations, sensitivity analysis

1. Introduction

The number of HIV infected individuals in South Africa is higher than in any other country [1]. In 2011, approximately 5.7 million people in South Africa were living

with HIV. The estimated prevalence rate is currently 10.6% and the number of new infections for 2011 was estimated at 380 500 individuals [2]. The impact of this epidemic can be seen in the dramatic change in South Africa's mortality rates [3]. Almost half of all the deaths in the country are due to AIDS [4]. The development of lymphomas in HIV infected individuals has been steadily increasing since the early phases of the HIV/AIDS epidemic [5].

HIV-associated lymphomas are predominantly aggressive B-cell lymphomas [2]. HIV-associated systemic non-Hodgkin's lymphomas (HIV-NHLs) include diffuse large B-cell lymphoma (DLBCL), HIV-associated primary central nervous system lymphoma (HIV-PCNSL), Burkitt's lymphoma (BL), primary effusion lymphoma (PEL) and plasmablastic lymphoma of the oral cavity [3]. A lymphoma is a cancer of the lymphatic cells of the immune system. The lymphoma originates in a single lymphocyte that has undergone mutations that present on it a growth and survival advantage in comparison to its normal cellular counterparts and it presents as an enlarged lymph node (a tumour). It is a cancer that affects the white blood cells of the lymph system, which is part of the body's immune system.

HIV-infected individuals are at a much higher risk of developing lymphomas, especially non-Hodgkin's lymphomas (NHLs) [6, 7, 11]. The development of HIV-related NHLs is related to old age, a low CD4 cell count and the lack of treatment with anti-retroviral therapy (ART) [8]. The risk of developing lymphomas also increases as the immune system weakens. The challenge is that HIV-infected individuals, who are receiving treatment, still have a high risk of developing lymphomas.

There are many mathematical models that have been developed and analysed to model the dynamics of HIV infection, see for instance, [9, 10, 15, 16, 17, 18, 19, 20, 21]. These models differ in terms of the number of variables that represent the compartments and the parameters used. Data is not always available for the models to be fitted, therefore only a few of these models have been fitted to data, see for instance [9, 18, 19, 20, 21]. It is important to note that there has been significant research on HRLs that does not use mathematical models. Abayomi *et al.* [11], estimated the impact of HIV and the South African ART roll-out program on the number of lymphoma cases. They found a dramatic increase in the number of lymphomas cases due to an increase in HIV-related lymphomas. The introduction of ART, did however not appear to have a big impact on the increase in HIV-related lymphoma cases. Their conclusion was that due to the late commencement of ART, its insufficient coverage and ineffective viral suppression of ART, the trend will continue and HIV-related lymphoma cases will continue to increase. One can then ask; by how much will the increase be? Can projections be made based on the available data? These questions can only be answered by the use of mathematical models. The dynamics and trends of HRLs over time can be understood and determined.

Little research has been done on mathematical models for HIV related lymphomas in general. We present a five state compartmental model that is path-breaking in three regards. Firstly, to the best our knowledge, this is the first attempt to model this newly developing phenomenon using compartmental mod-

els. Secondly, mathematical modelling tools are used to improve the understanding of how HIV infection is transforming the incidence and patterns of HRLs in the Western Cape province of South Africa. Thirdly, the model tracks the HRLs patterns in the population. The need to develop mathematical models that aid in making projections of HRLs on the basis of being informed by the current data that is being collected is thus addressed here. This model lays a foundation in shaping the policies that relate to the future trends and the interventions that need to be instituted.

The next section describes the model structure. In section 3, the simulations and least squares curve fitting method are described. The model parameters are also estimated and the sensitivity analysis of parameters presented. Section 4, the simulation results are presented and in section 5 the paper is concluded.

2. The model structure

We assume that we have a closed community. This means that the population stays constant over the modelling period. This assumption holds for populations in the communities from which the data was obtained, see [22, 23]. The population is divided into the following classes: those who are susceptible to infection of HIV and lymphomas, $S(t)$, those infected with HIV only $I_h(t)$, those with lymphomas only $I_l(t)$, those co-infected with HIV and lymphomas $I_{hl}(t)$ and those with AIDS $A(t)$. The total population at any time t is thus given by

$$(t) = S(t) + I_h(t) + I_l(t) + I_{hl}(t) + A(t).$$

Susceptible individuals increase due to a constant recruitment rate Λ. It is assumed that all individuals that enter the population are susceptible. For there to be an infection, there must be interaction between individuals in infected compartments and those that are susceptible. We assume that individuals in the infectious compartments infect the susceptibles at different rates. For instance, individuals with AIDS in compartment A who have a higher viral load, are assumed to be more infections than the individuals in I_h and I_{hl}. We thus assume that η is greater than 1. We assume that individuals in I_h and I_{hl} have the same infectiousness. Of particular importance is the inclusion of the exponential function to model decreased infectivity due to behaviour change that is driven by disease mortality. The parameter q measures how individuals respond to the increase or decrease of mortality due to HIV/AIDS. For instance, in Zimbabwe, fear of AIDS mortality may have influenced behaviour thus leading to a decline in HIV prevalence [18, 24].

The force of infection of HIV is

$$\lambda = \beta_3 e^{-q\delta A}\left(\frac{I_h + I_{hl} + \eta A}{N}\right).$$

In the model, individuals move between compartments as their infection status change. Individuals die at a fixed mortality rate μ, the natural death rate. Those with AIDS will die at an increased rate due to the disease. We thus let the disease induced mortality to be δ. We fully describe the parameters in the table below.

Table 2.1: Definitions of the parameters used in the model.

Parameter	Description
Λ	Recruitment rate
β_1	Rate at which susceptible individuals develop lymphoma
β_2	Rate at which HIV-infected individuals develop lymphoma
β_3	Effective contact rate of HIV
γ_1	Rate of recovery from lymphoma in HIV negative individuals
γ_2	Rate of recovery from lymphoma in HIV positive individuals
ρ_1	Rate at which HIV positive lymphoma individuals develop AIDS
ρ_2	Rate at which HIV positive individuals develop AIDS
δ	Disease related death rate
q	Natural death rate response to the increased or decreased mortality due to HIV/AIDS

Movement between compartments is represented by the transfer diagram below.

Combining the assumptions and transfer diagram, the equations of the model are given by

$$\frac{dS}{dt} = \Lambda - (\beta_1 + \lambda + \mu) S + \gamma_1 I_l, \tag{2.1}$$

$$\frac{dI_h}{dt} = \lambda S - (\beta_2 + \rho_2 + \mu) I_h + \gamma_2 I_{hl}, \tag{2.2}$$

$$\frac{dI_l}{dt} = \beta_1 S - (\gamma_1 + \lambda + \mu) I_l, \tag{2.3}$$

$$\frac{dI_{hl}}{dt} = \lambda I_l - (\gamma_2 + \rho_1 + \mu) I_{hl} + \beta_2 I_h, \tag{2.4}$$

$$\frac{dA}{dt} = \rho_1 I_{hl} + \rho_2 I_h - (\mu + \delta) A, \tag{2.5}$$

with initial conditions

$S(0) = S_0,\ S(0) = S_0,\ \ I_h(0) = I_{h0},\ \ I_l(0) = I_{l0},\ I_{hl}(0) = I_{lh0},\ A(0) = A_0.$

3. Simulations

3.1 The lease squares method

The most common, useful and powerful way to compare data to a theoretical model is to search for a theoretical curve that matches the data as closely as possible. The task here was to find the "best fit" line that goes through the data points.

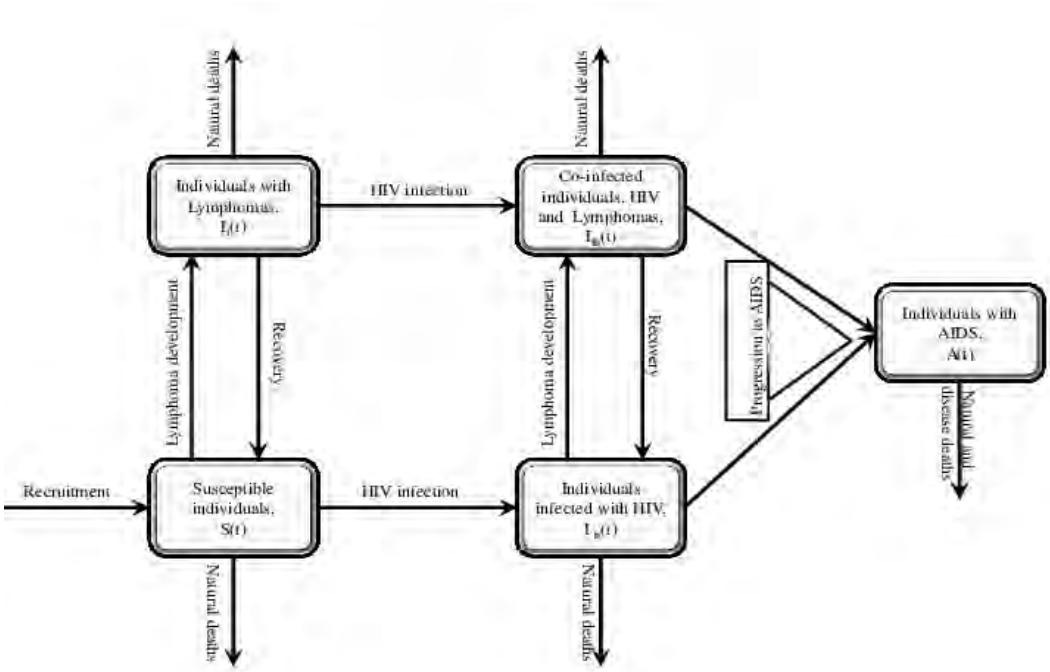


Figure 2.1: Schematic diagram of the model's five compartments and the associated flows.

To evaluate how the "best fit" line agreed with the data, it was important to consider uncertainties i.e. the level of discrepancies between a point and the line. This was done by measuring the vertical distance between the points and the line. We define the function X to be this sum of squares of discrepancies, as we tried different lines, and calculated X for each line. The "best line" was the one with the smallest value of X, i.e. the one with the "least squares." The value of X was obtained through an algorithm that was set in such a way that it always gave the best fit. The algorithm was also set to give the parameter values that produced the "best fit" line.

3.2 Parameter estimation

We now consider the estimation of parameters values. The dynamic model's functionality is subject to a correct estimation of parameter values corresponding to the arrows in Fig 1. Data on progression rates is inevitably limited since no related research has been done for the Western Cape province. We shall therefore make some assumptions to simplify the estimation process. In some cases we quote data from other sources. Because of the difficulties of estimating parameters, the descriptions of our findings are made in a much broader sense. For our numerical simulations, we shall assume that

$$\mu N = \Lambda - \delta A$$

ensuring that our population is always constant. This is because over the modelling time, the population change is small. We begin by estimating the model parameters. The average life expectancy in South Africa is 50 years, thus we consider $\mu = 0.02$ per year. The death rate due to AIDS is assumed to be 0.3, see also [23].

The size of the population considered in this model is estimated to be 50 000. is the recruitment rate, and it is equal to $(\mu + \delta)\, N$.

$$\Lambda = (\mu + \delta)\, N = \ (0.02 + 0.3)\, 50000 = 16000.$$

Table 3.1: The number of lymphoma cases per year for HIV positive (HIV+) and HIV negative (HIV-) individuals, data obtained from [16].

Year	2002	2003	2004	2005	2006	2007	2008	2009
HIV+	2	15	17	16	35	43	43	48
HIV-	33	106	95	62	147	133	150	131
Total	**35**	**111**	**112**	**78**	**182**	**176**	**193**	**179**

The data can best be represented figuratively as in Figure 2.

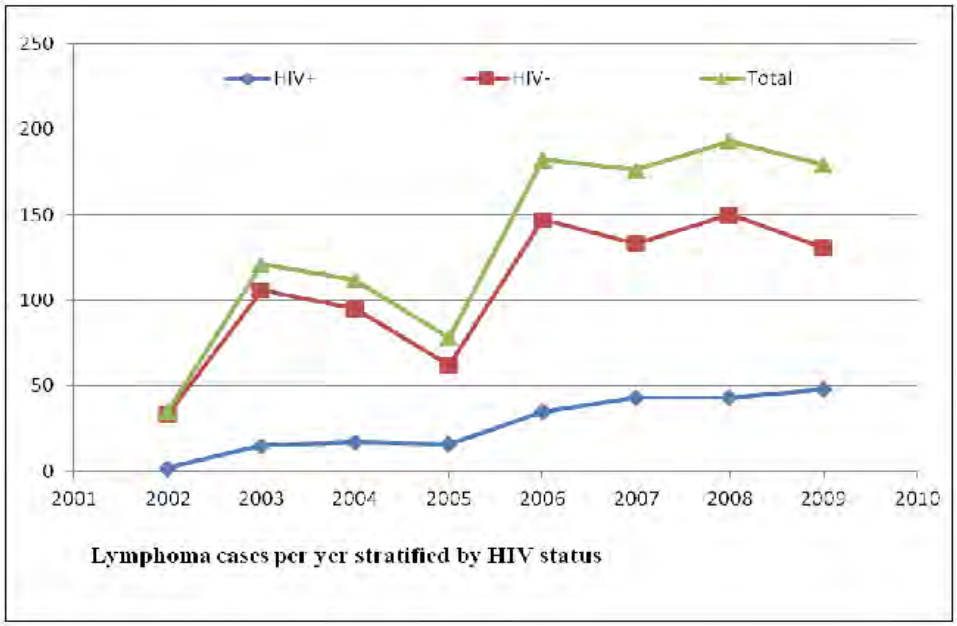


Figure 3.1: Trends of lymphoma incidence by HIV status, 2002–2009.

We fit our model to the total lymphoma cases. The data helps us to set the initial conditions as we endeavor to have the initial values to the model equal to 35 individuals. This value is used to set the initial conditions. The initial conditions are set to

$$(S(0),\ I_h(0),\ I_l(0),\ I_{hl}(0),\ A(0)) = (47275,\ 7500,\ 15,\ 10,\ 2000).$$

The parameter values are set within the following bounds depicted in Table 3.2.

Table 3.2: Upper and lower bounds of the parameters.

Parameter	Bounds
β_1	(0, 1)
β_2	(0.0001, 1)
β_3	(0, 0.5)
γ_1	(0.25, 0.62)
γ_2	(0.008, 0.03)
ρ_1	(0.04, 0.2)
ρ_2	(0, 0.9)
δ	(0.01, 0.5)
q	(1, 3)

3.3 Sensitivity Analysis

To establish what factors affect model outputs, it is important to do a sensitivity analysis as a way to assess the adequacy of a model. In the absence of detailed qualitative mathematical analysis we used Latin Hypercube Sampling and Partial Rank Correlation Coefficients (PRCCs) with 1000 simulations per run. Since the relationship between the model parameters and output variables is not known a priori, using PRCCs is ideal to determine how the parameters affect the model outputs. PRCCs illustrate the degree of the effect that each parameter has on the outcome. Thus PRCCs can be informative on what parameters to target if we want to achieve specific goals (e.g. control or regulatory mechanisms). For example, the most significant set of parameters can be used to determine how to efficiently reduce the infection rate. In other words, our predictions will be strengthened if we can reduce uncertainty and get better estimates on specific parameters of the model.

Latin Hypercube Sampling is a statistical sampling method that allows for an efficient analysis of parameter variations across simultaneous uncertainty ranges in each parameter [12]. To investigate the relationship between our parameters and the variables, we chose the output results for *A(t)* (individuals with AIDS). We

produce scatter plots depicted if Figure 3.2. The figure shows the scatter plots of the 1000 output values at year 20 plotted versus the input parameter.

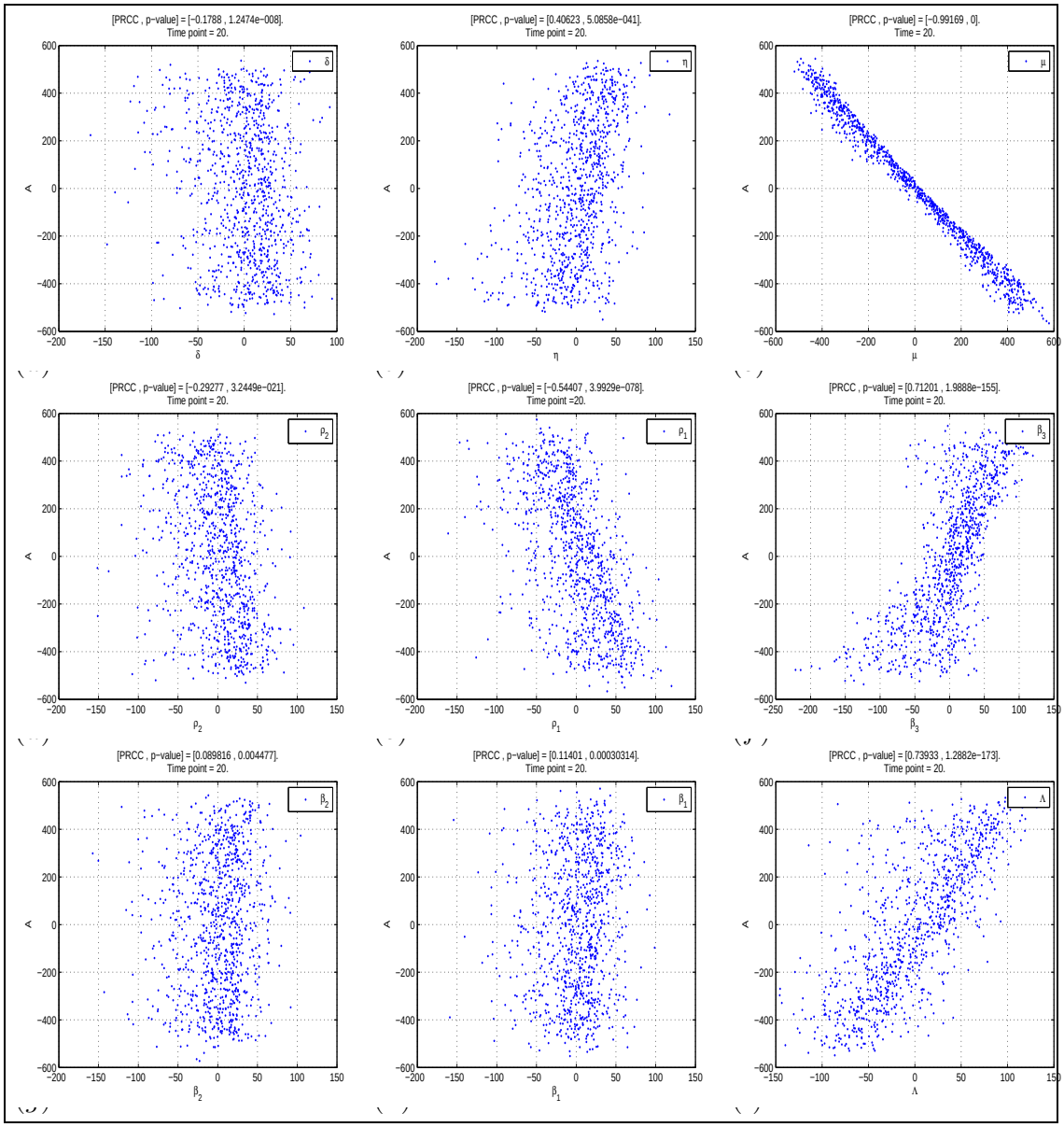


Figure 3.2: PRCC scatter plots of parameters δ, η, μ, ρ_2, ρ_1, β_3,β_2,β_1 and Λ (calculated at year 20, all the parameters are varied at the same time). The abscissa represents the residuals of the linear regression between the rank-transformed values of the parameter under investigation versus the rank-transformed values of all the other parameters. The ordinate represents the residuals of the linear regression between the rank-transformed values of the output versus the rank-transformed values of all the parameter under investigation The title of each plot represents the PRCC value with the corresponding p-value.

Figure 3.2, illustrates the PRCCs using $A(t)$ as an output variable. The parameters with the greatest effect on the outcome are the natural mortality, recruitment

rate and the HIV infection rate. Interestingly recover rates from HRLs are not significant to the outcome variable *A(t)*. This is reasonable since progression of HIV is not linked to the Lymphomas in the model. It is also interesting to note that $\rho_1$is more significant than ρ_2. Progression of individuals dually infected with HIV and Lymphomas negatively affects *A(t)* more than that of those with HIV infection only. This means control of HIV-HRLs co-infection is important.

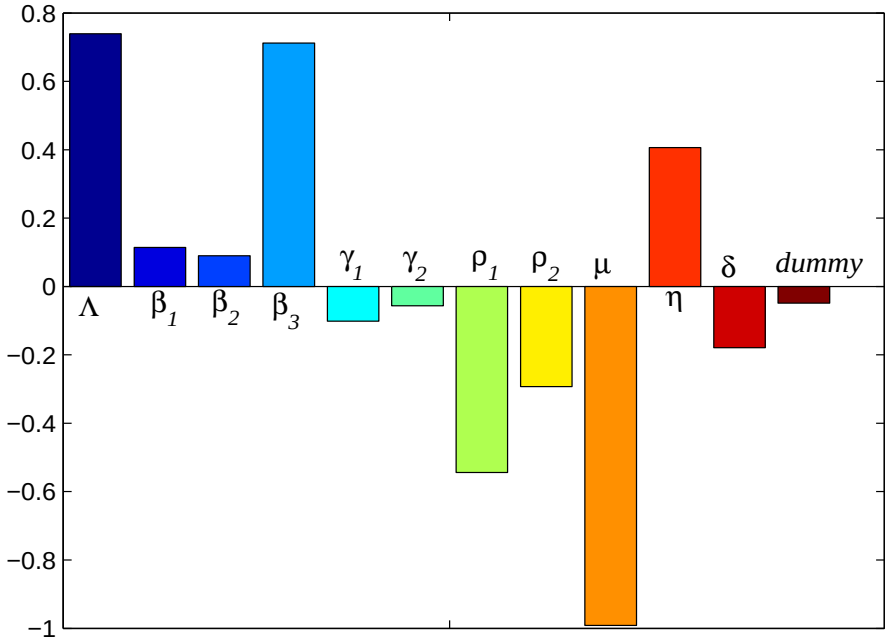


Figure 3.3: PRCC results with sample size 1000. The PRCCs that are clearly significant are η, μ, ρ_2, ρ_1, β_3 and Λ.

4. Results

The model is simulated using Matlab. We specifically use the least squares curve fitting algorithm (LSCFA) to fit the model to the data. The unknown parameters are defined in acceptable ranges, mostly estimated due to the unavailability of data. A set of parameter values are obtained, within the range as set out in Table 3., to produce the best fit. The best fit to the model is shown in Figure 4.1.

We hypothetically project the development of HRLs over 18 years. Such a projection is less likely to be accurate in reality, but such information is vital for public health planning and institution of interventions. Figure 4.2 shows the projected change in the occurrence of lymphoma cases in both HIV positive and HIV negative patients.

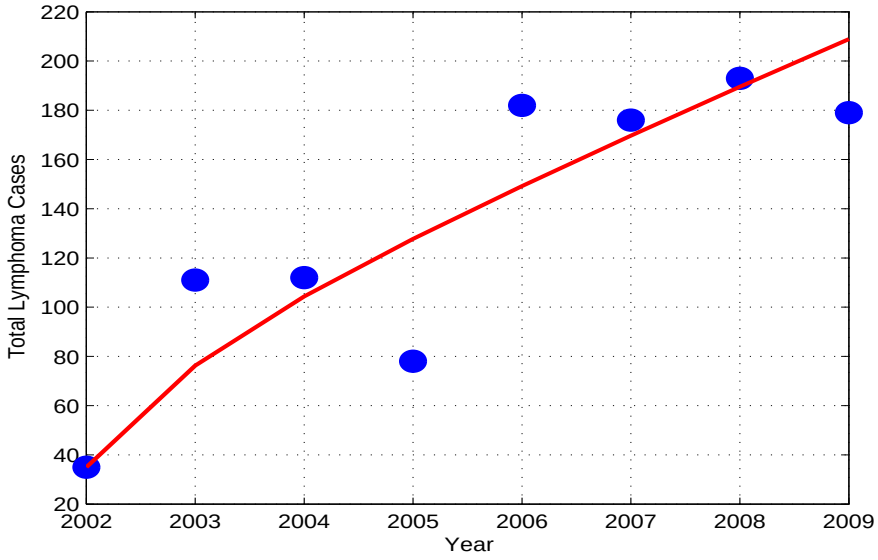


Figure 4.1: shows the model fit to the data. The following parameters were estimated: $\beta_1 = 0.0011$; $\beta_2 = 0.004$; $\beta_3 = 0.1$; $\gamma_1 = 0.718$; $\gamma_2 = 0.4245$; $\rho_1 = 0.8163$; $\rho_2 = 0.0159$; $\delta = 0.03$; $\eta = 1.5$; $\mu = 0.02$; $q = 50$ and $\Lambda = 16000$.

5. Conclusion

HIV is dramatically changing the demographics of disease and the resource needs of the entire health care service of South Africa. Several malignancies have an increased risk of developing in Persons Living with HIV (PLWH), such as the HRLs. PLWH have between 60 and 200 times greater risk when compared to the uninfected population of presenting with one of these serious Lympho-proliferative disorders [13, 14].

In this research work, a model was developed to model the dynamics of HIV-related lymphomas. The research is a component of the Tygerberg Lymphoma Study Group, that looks at using expanded laboratory profiling in conjunction with clinical, imaging and demographic methodologies to improve the understanding of how HIV is transforming the incidence, pattern, prognosis and outcomes of Lymphoproliferative Disorders in the Tygerberg catchments area of the Western Province of South Africa. The model looks at the co-infection of HIV and Lymphomas. Co-infections present immense challenges in disease control. This has been the case with co-infections of HIV and tuberculosis, HIV and malaria, HIV and Hepatitis B virus to mention a few.

We have demonstrated that a very simple model of five compartments can be used a tool to reproduce the trends of HRLs in the Western Cape province of South Africa. Like any modelling exercise, this work presented here is not without shortcomings. The data set we used is small and imperfect, and this can result in poor predictions with the fitting of the models that are deterministic. The epidemiology of HIV-associated lymphomas has changed since antiretroviral therapy

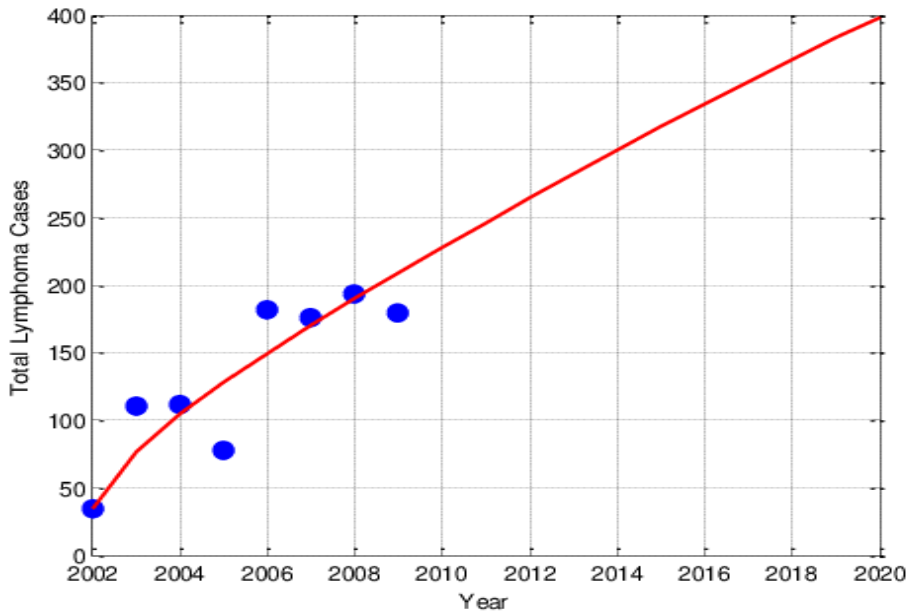


Figure 4.2: shows the total number of lymphoma cases over a period of 18 years till 2020. The parameter values are as given in Figure 4.1.

became available in clinical practice. The treatment of HIV infection, which has evolved rapidly in recent years, has benefited the treatment of lymphomas in HIV infected individuals. It is thus important to include treatment of HIV in such the model and this forms our current research focus. This will help in determining the impact of ART on the prevalence of lymphomas.

Acknowledgements

The authors would like to thank the Department of Mathematical Sciences at the University of Stellenbosch for its support in the production of this manuscript.

Bibliography

[1] UNAIDS, UNAIDS report on the global AIDS epidemic, 2010.

[2] L.D. Kaplan, HIV-associated lymphoma, Best Practice & Research Clinical Haematology, **25** (2012), 101-117.

[3] M. Karunanayake and A. Carol, HIV-associated lymphoma, *Proceedings (Baylor University Medical Center)*, **22** (2009), 74-76.

[4] P.J. Lebohla, Mid-year Population Estimates, *Statistics South Africa*, 2011.

[5] T.M. Rehel, T.B. Hallett, O. Shisana, V.P. Wyk, K. Zuma, H. Carrara, S. Jooste, A decline in new HIV infections in South Africa: Estimating HIV incidence from three National HIV Surveys in 2002, 2005 and 2008, *PLoS ONE* **5**(6) (2010), e11094. doi:10.1371/journal.pone.0011094

[6] M. Mohamed, HIV-associated malignancies have become increasingly important with the spread of HIV infection, *CME*, **26** (2007), 70-74.

[7] U. Tirelli, M. Spina, G. Gaidano, E. Vaccher, S. Franceschi, A. Carbone, Epidemiological, biological and clinical features of HIV-related lymphomas in the era of highly active antiretroviral therapy, *AIDS*, **14** (2000), 1675-1688.

[8] C. Hoffmann, E. Wolf, G. Fatkenheuer, T. Buhk, A. Stoehr, A. Plettenberg, H. Stellbrink, H. Jaeger, U. Siebert, H. Horst, Response to highly active antiretroviral therapy strongly predicts outcome in patients with AIDS-related lymphoma, *AIDS*, **17** (2003), 1524-1529.

[9] F. Nyabadza, Z. Mukandavire, S.D. Hove-Musekwa, Modelling the HIV/AIDS epidemic trends in South Africa: Insights from a simple mathematical model. *Nonlinear analysis*: *Real World Applications*, **12** (2011), 2091-2104.

[10] Z. Mukandavire, K. Bowa, W. Garira, Modelling circumcision and condom use as HIV/AIDS preventative control strategies, *Mathematical and Computer Modeling*, **46** (2007), 1353-1372.

[11] E.A. Abayomi, A. Somers, R. Grewal, G. Sissolak, F. Bassa, D. Maartens, P. Jacobs, C. Stefan, L.W. Ayers, Impact of the HIV epidemic and Antiretroviral treatment policy on lymphoma incidence and subtypes seen in the Western Cape of South Africa, 2002-2009: Preliminary findings of the Tygerberg Lymphoma Study Group, *Transfusion and Apheresis Science*, 2011, **44** (2011), 161-166.

[12] S.M. Blower, H. Dowlatabadi, Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example, *International Statistics Review*, **2** (1994), 229-243.

[13] C. Palmieri et al., AIDS related non Hodgkin's Lymphoma in the first decade of highly active antiretroviral therapy, *QJMed*, **99** (2006), 811-826.

[14] G. Sissolak, A. Abayomi, P. Jacobs, AIDS defining lymphomas in the era of highly active antiretroviral therapy (HAART): An African perspective, *Journal of Transfusion and Apheresis Science*, 37 (2007), 63-70.

[15] J.Y.T. Mugisha, F. Baryarama, L.S. Luboobi, Modelling HIV/AIDS with complacency in a population with declining prevalence, *Computational Mathematical Methods in Medicine*, **7** (2006), 27-35.

[16] F. Nyabadza, On the complexity of modelling HIV/AIDS in Southern Africa. *Mathematical Modelling and Analysis*, **13** (2008), 539-552.

[17] A. Tripathi, R. Naresh, D. Sharma, Modelling the effect of screening of unaware infectives on the spread of HIV infection, *Applied Mathematical Computation*, **184** (2007), 1053-1068.

[18] T.M. Hallet, S. Gregson, O. Mugurungi, E. Gonese, G.P. Garnett, Assessing evidence for behavior change affecting the course of HIV epidemics: A new mathematical modelling approach and application to data from Zimbabwe, *Epidemics*, **1** (2009), 108-117.

[19] T.M. Rehle, O. Shisana, Epidemiological and demographic HIV/AIDS projections: South Africa. *Africa Journal of AIDS Research*, **2** (2003), 1-8.

[20] L.F. Johnson, R.E. Dorrington, Modelling the demographic impact of HIV/AIDS in South Africa and the likely impact of interventions, *Demographic Research*, **14** (2006), 541-574.

[21] N. Bacaër, R. Ouifki, C. Pretorius, R. Wood, B. Williams, Modeling the joint epidemics of TB and HIV in a South African township, *Journal of Mathematical Biology*, **57** (2008), 557-593.

[22] S. Verver et. al., Transmission of tuberculosis in a high incidence urban community in South Africa, *International Journal of Epidemiology*, **33** (2004), 351-357.

[23] Statistical Support and Informatics, *Statistics South Africa: Western Cape*, Census 2001.

[24] S. Gregson, et al., HIV Decline associated with behavior change in Eastern Zimbabwe. *Science*, **311** (2006), 664-666.

Ring of Invariants for Systems with Linear Part $N_{3(n)}$

David M. Malonza[1], Grace W. Gachigua[2] and Johana K. Sigey[3]

[1]*Kenyatta University, Department of Mathematics, Nairobi, Kenya. Email: dmalo2004@gmail.com*

[2]*Corresponding author. Kimathi University College of Technology, Department of Mathematics, Nyeri, Kenya. Email: wambuigachigua@gmail.com*

[3]*Jomo Kenyatta University of Agriculture and Technology, Department of Pure and Applied Mathematics, Nairobi, Kenya. Email: jksigey2002@yahoo.com*

Abstract. The problem of describing the normal form of a system of differential equations at equilibrium with nilpotent linear part is solvable once the ring of invariants associated with the system is known. Our concern in this paper is to describe ring of invariants of differential systems with nilpotent linear part made up of n 3×3 Jordan blocks which is best described by giving the Stanley decomposition of the ring. An algorithm based on the notion of transvectants from classical invariant theory is used to determine the Stanley decomposition for the ring of invariants for the coupled systems when the Stanley decompositions of the Jordan blocks of the linear part are known at each stage.

Keywords: invariants, box product, transvectant, Stanley decomposition, normal form.

1. Introduction

In the study of the qualitative properties of a nonlinear differential equation, $\dot{x} = Ax + \textit{higher order terms}$ near an equilibrium point, one of the most powerful techniques available is to simplify A using nonlinear changes of coordinates which leave the origin fixed. The simplified vector field is called the normal form of A. The theory of normal forms is concerned with finding the simplest form for the system by removing as many terms as possible with the remaining ones having dynamical significance.

There are well-known procedures for putting a system of differential equations

$$\dot{x} = Ax + v(x) \tag{1.1}$$

(where v is a formal power series with quadratic terms) into normal form with respect to its linear part, A. The normal form theory divides into two parts, the

case when $A = S$ is diagonalizable and the case when $A = N$ is nilpotent, that is nilpotent systems. The general case can be solved by combining the results of the two special cases. The goal of this paper is to describe the ring of invariants of a differential system (1.1) when its linear part A is a nilpotent matrix N, where

$$N = \begin{bmatrix} N_3 & & \\ & \ddots & \\ & & N_3 \end{bmatrix} \quad \text{and} \quad N_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Our main result is a procedure that solves the description problem where N, the nilpotent matrix is in Jordan form, with coupled n Jordan blocks, provided that the description problem is already solved for each Jordan block of N taken separately. Our method is based on adding one block at a time.

The problem of finding Stanley decomposition for the equivariants of $N_{22,\ldots,2}$ was first solved by Cushman et al. [2] using a method called "covariants of special equivariants". Their method begins by creating a scalar problem that is larger than the vector problem. Their procedures derived from classical invariant theory. Thus it was necessary to repeat calculations of classical theory at the levels of equivariants. Malonza [5] solved the same problem by "Groebner" basis methods found in Adams et al. [1] rather than borrowing from classical results.

Murdock and Sanders [8] developed an algorithm based on the notion of transvectants to determine the normal form of a vector field with a nilpotent linear part, when the normal form is known for each Jordan block of the linear taken separately. The algorithm is based on the notion of transvectant, from classical invariant theory. Malonza [5] using the algorithm in [7] for transvectants computed the Stanley decomposition for Takens-Bogdanov systems and his results agreed with his previous work in [5].

Namachchivaya et al [8] studied a generalized Hopf bifurcation with non-semisimple 1:1 Resonance. The normal form for such a system contains only terms that belong to both the semisimple part of linear part and the normal form of the nilpotent, which is a couple Takens-Bogdanov system with linear part

$$A = \begin{bmatrix} i\omega & 1 & & \\ & i\omega & & \\ & & i\omega & 1 \\ & & & i\omega \end{bmatrix}.$$

This example illustrates the physical significance of the study of normal forms for systems with nilpotent linear part.

Our results are mainly based on the work found in [7], that is, the application of tranvectant's method (also known as box product) for computing Stanley decompositions for the ring of invariants of nilpotent systems.

2. Invariant and Stanley decompositions

Let $\mathcal{P}_j(\mathbb{R}^n, \mathbb{R}^m)$ denote the vector space of homogeneous polynomials of degree j on $\mathbb{R}^n$ with coefficients in $\mathbb{R}^m$, where $\mathbb{R}$ denotes the set of real numbers. Let $\mathcal{P}(\mathbb{R}^n, \mathbb{R}^m)$ be the vector space of all such polynomials of any degree and let $\mathcal{P}_*(\mathbb{R}^n, \mathbb{R}^m)$ be the vector space of formal power series. If $m = 1$, $\mathcal{P}_*(\mathbb{R}^n, \mathbb{R})$ becomes the ring of formal power series on $\mathbb{R}^n$. For such smooth vectors fields, it is sufficient to work with polynomials. For any nilpotent matrix N, we define the Lie operator

$$L_N : \mathcal{P}_j(\mathbb{R}^n, \mathbb{R}^n) \to \mathcal{P}_j(\mathbb{R}^n, \mathbb{R}^n)$$

by

$$(L_N v)x = v'(x)Nx - Nv(x) \tag{2.1}$$

and the differential operator

$$\mathcal{D}_{Nx} : \mathcal{P}_j(\mathbb{R}^n, \mathbb{R}) \to \mathcal{P}_j(\mathbb{R}^n, \mathbb{R})$$

by

$$(\mathcal{D}_{Nx} f)(x) = f'(x)Nx = (Nx.\bigtriangledown)f(x). \tag{2.2}$$

Then $\mathcal{D}_N$ is a derivation of the ring $\mathcal{P}(\mathbb{R}^n, \mathbb{R})$, meaning that

$$\mathcal{D}(fg) = (\mathcal{D}_N f)g + f\mathcal{D}_N g). \tag{2.3}$$

In addition,

$$L_N(fv) = (\mathcal{D}_N f)v + fL_N v. \tag{2.4}$$

A function f is called an invariant of Ax if $\frac{\partial}{\partial t} f(e^{At}x) \mid_{t=0} = 0$ or equivalently $f \in ker\mathcal{D}_A$. Since

$$\mathcal{D}_N(f + g) = \mathcal{D}_N f + \mathcal{D}_N g$$

$$\mathcal{D}_N fg = f\mathcal{D}_N g + g\mathcal{D}_N f$$

it follows that if f and g are invariants, so are $f + g$ amd fg; that is $ker\ \mathcal{D}_N$ is both a vector space over $\mathbb{R}$ and also a subring of $\mathcal{P}(\mathbb{R}^n, \mathbb{R})$, known as the *ring of invariants*.

Similarly a vector field v is called an *equivariant* of Ax if $\frac{\partial}{\partial t}(e^{-At}v(e^{At}x)) \mid_{t=0} = 0$ that is $v \in ker L_A$.

There are two normal form styles in common use for nilpotent systems, the *inner product normal form* and the *sl(2) normal form.* The inner product normal form is defined by $\mathcal{P}(\mathbb{R}^n, \mathbb{R}^n) = imL_N \oplus ker L_{N^*}$ where N^* is the conjugate transpose of N. To define the sl(2) normal form, one first sets $X = N$ and constructs matrices Y and Z such that

$$[X, Y] = Z, \qquad [Z, X] = 2X, \qquad [Z, Y] = -2Y. \tag{2.5}$$

An example of such an $sl(2)$ triad $\{X, Y, Z\}$ is

$$X = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad Y = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \qquad Z = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Having obtained the triad $\{X, Y, Z\}$ we create two additional triads $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$ and $\{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$ as follows

$$\mathcal{X} = \mathcal{D}_Y, \qquad \mathcal{Y} = \mathcal{D}_X, \qquad \mathcal{Z} = \mathcal{D}_Z \tag{2.6}$$

$$\mathsf{X} = L_Y, \qquad \mathsf{Y} = L_X, \qquad \mathsf{Z} = L_Z \tag{2.7}$$

The first of these is a triad of differential operators and the second is a triad of Lie operators. Both the operators $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$ and $\{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$ inherit the triad properties (2.5). Observe that the operators $\{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$ map each $\mathcal{P}(\mathbb{R}^n, \mathbb{R}^n)$ into itself. It then follows from the representation theory $sl(2)$ that

$$\mathcal{P}(\mathbb{R}^n, \mathbb{R}^n) = im\mathsf{Y} \oplus ker\mathsf{X} = im\mathsf{X} \oplus ker\mathsf{Y}. \tag{2.8}$$

Clearly the *ker* $\mathcal{X}$ ia s subring of $\mathcal{P}(\mathbb{R}^n, \mathbb{R})$, the ring of invariants and it follows from (2.4) that *ker* X is a module over this subring. This is the *sl(2) normal form module.*

The most effective way of describing the invariant ring associated with a nilpotent matrix N is by a device from communicative algebra called a Stanley decomposition, introduced for this purpose in [8]. We write $\mathbb{R}[[x_1, ..., x_n]]$ for the ring of (scalar) power series in variables $x_1, ..., x_n$. A subalgebra $\Re$ of $\mathbb{R}[[x_1, ..., x_n]]$ is a subset that is both a subring and a vector subspace. The subalgebra is graded if

$$\Re = \bigoplus_{d=0}^{\infty} \Re_d,$$

where $\Re_d$ is the vector subspace of $\Re$ consisting of elements of degree d. To define Stanley decomposition of a graded subalgebra, we begin with the definition of a Stanley term. A Stanley term is an expression of the form $\mathbb{R}[[f_1, ..., f_k]]\varphi$, where the elements $f_1, ..., f_k$ and φ are homogeneous polynomials and $f_1, ..., f_k$ not including φ are required to be algebraically independent. The Stanley term $\mathbb{R}[[f_1, ..., f_k]]\varphi$ denotes the set of all expressions of the form $F(f_1, ..., f_k)\varphi$ where F is a formal power series in k variables. When $\varphi = 1$, φ is omitted, and the Stanley term is a subalgebra, otherwise it is only a subspace. A Stanley decomposition is a finite direct sum of Stanley terms. A polynomial f is called *doubly homogeneous of type (d,w)* if every monomial in f has a degree d and weight w. A vector subspace V of *ker* $\mathcal{X}$ is doubly graded if

$$V = \bigoplus_{d=0}^{\infty} \bigoplus_{w=0}^{\infty} V_{dw},$$

where V_{dw}, is the vector subspace of V consisting of doubly homogeneous polynomials of degree d and weight w. A doubly graded Stanley decomposition of a doubly graded subalgebra $\mathfrak{R}$ of $ker\ \mathcal{X}$ is an expression of $\mathfrak{R}$ as a direct sum of vector subspaces of the form $\mathbb{R}[[f_1, ..., f_k]]\varphi$, where $f_1, ..., f_k, \varphi$ are doubly homogeneous polynomials. All Stanley decomposition considered from here on will be of this kind and the words "doubly graded" will be omitted.

A standard monomial associated with a Stanley decomposition is an expression of the form $f_1^{m_1}, ..., f_k^{m_k}\varphi$, where $\mathbb{R}[[f_1, ..., f_n]]\varphi$ is a term in the Stanley decomposition. Notice that monomial here means a monomial in the basic invariants $x_1, ..., x_n$, which are polynomials in the original variables $x_1, ..., x_n$. Given a Stanley decomposition of $ker\ \mathcal{X}$, its standard monomials of a given degree (or of a given type) form a basis for the (finite-dimensional) vector space of invariants of that degree (or type). Next, we give Stanley decompositions for the rings of invariants associated with N_2 and N_{22} using the notion in [5]. The ring of invariants of N_2 in $\mathbb{R}[x_1, y_1]$ is $ker\ \mathcal{X}_2$. This ring clearly contains $\alpha = x_1$ which is of a type (1,1), and in fact every element of $ker\ \mathcal{X}_2$ can be written uniquely as a formal power series $f(x_1)$ in x_1 alone. We express this by the Stanley decomposition

$$ker\ \mathcal{X}_2 = \mathbb{R}[[\alpha]].$$

The ring of N_{22} in $\mathbb{R}[[x_1, x_2, y_1, y_2]]$ is described by the Stanley decomposition

$$ker\ \mathcal{X}_{22} = \mathbb{R}[[\alpha_1, \alpha_2, \beta_{12}]]$$

with

$$\alpha_1 = x_1, \alpha_2 = x_2, \beta_{12} = x_1y_2 - x_2y_1.$$

Here α_1, α_2 are of the type (1,1) and β_{12} is of type (2,0).

To prove that we have obtained all the invariants (transvectants), we need to generate the table function of the Stanley decomposition. We replace each term of the decomposition by a rational function P/Q in d and w (d for degree and w for weight) construct as follows: for each basic invariants appearing inside the square brackets, the denominator will contain a factor $1 - d^p w^q$, where p and q are the degree and weight of the invariants; the numerator will be $d^p w^q$, where p and q are the degree and weight of the standard monomials of that term. When the rational function P/Q from each term of the Stanley decomposition are summed up we obtain the table function T given by $T = \sum_i P_i/Q_i$. Thus, for examples above, the table function is $T_2 = \frac{1}{1-dw}$ and $T_{22} = \frac{1}{(1-dw)^2(1-d^2)}$.

The following lemma found in [7] gves a method to check that enough basic invariants have been found.

Lemma 1. *Let $\{X, Y, Z\}$ be a triad of $n \times n$ matrices, let $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$ be the induced triad and suppose that $I_1, ..., I_n$ is a finite set of polynomials in ker $\mathcal{X}$. Let $\mathcal{R}$ be a subring of $\mathbb{R}[I_1, ..., I_n]$; suppose that the Stanley terms have been found, and that the Stanley decomposition and its associated table function $T(d, w)$ have been determined. Then $\mathcal{R} = ker\ \mathcal{X} \subset \mathcal{P}(\mathbb{R}^n, \mathbb{R}^n)$ if and only if*

$$\frac{\partial}{\partial w} wT \mid_{w=1} = \frac{1}{(1-d)^n}.$$

3. BOX PRODUCTS OF STANLEY DECOMPOSITIONS

Let $V_k, k = 1, 2$ be $sl(2)$ representation spaces with triads $\{X_k, Y_k, Z_k\}$. Then $V_1 \otimes V_2$ is a representation space with triad $\{X, Y, Z\}$, where $X = X_1 \otimes I + I \otimes X_2$ (and similarly for Y and Z). We define the box product of $ker X_1$ and $ker X_2$ by

$$(ker X_1 \boxtimes ker X_2) = ker X \tag{3.1}$$

To begin to put the box product into computationally useful form, we use the notion of *external transvectants* introduced for this purpose in [7]. Consider a system with nilpotent linear part

$$N = \begin{bmatrix} \hat{N} & 0 \\ 0 & \tilde{N} \end{bmatrix}$$

where $\hat{N}$ and $\tilde{N}$ are nilpotent matrices of sizes $\hat{n} \times \hat{n}$ and $\tilde{n} \times \tilde{n}$ respectively $(\hat{n} + \tilde{n} = n)$, in (upper) Jordan form, and each may consist of one or more Jordan blocks. Let $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$, $\{\hat{\mathcal{X}}, \hat{\mathcal{Y}}, \hat{\mathcal{Z}}\}$ and $\{\tilde{\mathcal{X}}, \tilde{\mathcal{Y}}, \tilde{\mathcal{Z}}\}$ be the associated triads of operators acting on $\mathbb{R}[[x_1, ..., x_n]], \mathbb{R}[[x_1, ...x_{\hat{n}}]]$ and $\mathbb{R}[[x_{\hat{n}+1}]]$ respectively. Suppose that $f = f(x_1, ..., x_{\hat{n}}) \in ker\ \hat{\mathcal{X}}$ and $g = g(x_{\hat{n}+1}, ..., x_n) \in ker \tilde{\mathcal{X}}$ are weight invariants of weights w_f and w_g, and i is an integer in the range $0 \leq i \leq min(w_f, w_g)$. Then we define *external transvectant* of f and g of order i to be the polynomial $(f, g)^i \in \mathbb{R}[[x_1, ..., x_n]]$ given by

$$(f, g)^i = \sum_{j=0}^{i} (-1)^j W_{f,g}^{i,j} (\hat{\mathcal{Y}}^j f)(\tilde{\mathcal{Y}}^{j-1} g) \tag{3.2}$$

where

$$W_{f,g}^{i,j} = \binom{i}{j} \frac{(w_f - j)!}{(w_f - i)!} . \frac{(w_g - i + j)!}{(w_g - i)!}$$

We say that a transvectant $(f, g)^i$ is well defined if i is in the proper range for f and g. Notice that the zeroth transvectant is always well-defined and reduces to the product:$(f, g)^0 = fg$. Given Stanley decompositions for $ker\ \hat{\mathcal{X}}$ and $ker\ \tilde{\mathcal{X}}$, the following results found in [7] section 6, provide the first step toward obtaining a Stanley decomposition for $ker\ \mathcal{X}$.

Theorem 1. *Each well-defined transvectant $(f, g)^i$ of $f \in ker \hat{\mathcal{X}}$ and $g \in ker \tilde{\mathcal{X}}$ belongs to ker $\mathcal{X}$. If f and g are doubly homogeneous polynomials of types (df, wf) and (dg, wg) respectively, $(f, g)^i$ is a doubly homogeneous polynomial of type $(df + dg, wf + wg - 2i)$. Suppose that Stanley decompositions for ker $\hat{\mathcal{X}}$ and ker $\tilde{\mathcal{X}}$ are given, then a basis for the (finite-dimensional) subspace $(ker\ \mathcal{X})_d$ of homogeneous polynomials in ker $\mathcal{X}$ with degree d is given by the set of all well-defined transvectants $(f, g)^i$ where f is a standard monomial of the Stanley decomposition for ker $\hat{\mathcal{X}}$ and g is a standard monomial of the Stanley decomposition for ker $\tilde{\mathcal{X}}$ and $df + dg = d$.*

The bases given by Theorem 2 are sufficient to determine $ker\ \mathcal{X}$ one degree at a time, but to find all of $ker\ \mathcal{X}$ in this way would require finding infinitely many transvectants. A Stanley decomposition for $ker\ \mathcal{X}$ must be based on a finite number of basic invariants. To construct such a decomposition, we must find an alternative basis for each $(ker\ \mathcal{X})_d$ that uses only a finite number of transvectants overall. Such alternatives bases can be found by the following replacement theorem found in [7].

Theorem 2. *Any transvectant $(f,g)^i$ in the basis given by Theorem 1 can be replaced by a product $(f_1,g_1)^{i_1}...(f_j,g_j)^{i_j}$ of transvectants, provided that $f_1...f_j = f, g_1...g_j = g$ and $i_1 + ... + i_j = i$.*

The following corollary of the Replacement Theorem 3 will play a crucial role in our calculations.

Corollary 1. *If $w_h = w_k = r$ so that $(h,k)^{(r)}$ has weight zero, then whenever $(fh,gk)^{(i+r)}$ is well defined, it may be replaced by $(f,g)^{(i)}(h,k)^{(r)}$.*

Proof. Clearly $(fh,gk)^{(i+r)}$ and $(f,g)^{(i)}(h,k)^{(r)}$ have the same stripped form and total transvectant order. It is only necessary to observe that $(f,g)^{(i)}$ is well-defined. But $w_{fh} = w_f + w_h = w_f + r \geq i + r$, so $w_f \geq i$ and similarly $w_g \geq i$. □

The next lemma is now trivial, but essential to our method.

Lemma 2. *Box distributes over direct sums of admissible subspaces: If $\hat{V} \subset ker\hat{X}, \tilde{V}_1 \subset ker\tilde{X}$, and $\tilde{V}_2 \subset ker\tilde{X}$ are admissible subspaces, with $\tilde{V}_1 \cap \tilde{V}_2$, then $\tilde{V}_1 \oplus \tilde{V}_2$ is admissible and $\hat{V} \boxtimes (\tilde{V}_1 \oplus \tilde{V}_2) = (\hat{V} \boxtimes \tilde{V}_1) \oplus (\hat{V} \boxtimes \tilde{V}_2$, and similarly for $(\tilde{V}_1 \oplus \tilde{V}_2) \boxtimes \hat{V}$.*

We complete this section by the following theorem which is Theorem (9) in [7], and outlines the procedure for computing $ker\ \mathcal{X}$.

Theorem 3. *[7, Theorem 9] A Stanley decomposition of $ker\ \mathcal{X} = ker\ \hat{\mathcal{X}} \boxtimes ker\ \tilde{\mathcal{X}}$ is computable in a finite number of steps given decomposition of $ker\ \hat{\mathcal{X}}$ and $ker\ \tilde{\mathcal{X}}$.*

The proof of the theorem is given in [7], but we will briefly outline the ideas used in the proof important in our calculations. By Lemma 2, we can compute $ker\ \mathcal{X}$ if we can compute any box product of the form $\mathbb{R}[[f_1,...,f_k]]\varphi \boxtimes \mathbb{R}[[g_1,...,g_l]]]\psi$, where each factor is a Stanley term from the given decompositions of $ker\ \hat{\mathcal{X}}$ and $ker\ \tilde{\mathcal{X}}$.

Let p be the number of elements of weight > 0 in $f_1,...,f_k$ and q the number of such elements in $g_1,...,g_l$. We proceed by double induction on p and q. Suppose $p = q = 0$, Then the box product is spanned by transvectants of the form $(f_1^{m_1},...,f_k^{m_k}\varphi, g_1^{n_1},...,g_l^{n_l}\psi)$, which is well-defined if and only if $0 \leq i \leq r$, where $r = min(w_\varphi, w_\psi)$. The f and g factors add no weight, and cannot support any higher transvectants. By Theorem 3 each transvectant may be replaced by $f_1^{m_1}...f_k^{m_k}, g_1^{n_1}...g_l^{n_l}(\varphi,\psi)^i$ which remains well-defined. Therefore

$$\mathbb{R}[[f_1,...,f_k]]\varphi \boxtimes \mathbb{R}[[g_1,...,g_l]]\psi \cong \bigoplus_{i=0}^{r} \mathbb{R}[[f_1,...,f_k,g_1,...,g_l]](\varphi,\psi)^i.$$

Now we make induction hypothesis that all cases with $p = 0$ are computable up through the case $q - 1$, and we discuss case q. Choose one of the q elements of $g_1, ..., g_l$ having positive weight; we assume the chosen element is g_1. Then we may expand

$$\mathbb{R}[[g_1, ..., g_l]]\psi = \Big(\bigoplus_{v=0}^{t-1} \mathbb{R}[g_2, ..., g_l]]g_1\psi\Big) \oplus \mathbb{R}[[g_1, ..., g_l]]g_1^t\psi,$$

where t is the smallest integer such that $w_{g_1^t\varphi} > w_\psi$. This decomposition corresponds to classifying monomials according to the power of g_1 that occurs, with all powers greater than or equal to t assigned to the last term. Now take the box product of $\mathbb{R}[[f_1^{m_1}, ..., f_k^{m_k}]]\varphi$ times this expression, and distribute the product according to Lemma 5. All of the terms except the last are computable by the induction hypothesis. We claim the last term is computable by the formula

$$\mathbb{R}[[f_1, ..., f_k]]\varphi \boxtimes \mathbb{R}[[g_1, ..., g_l]]g_1^t\psi \cong \bigoplus_{i=0}^{w_\varphi} \mathbb{R}[[f_1, ..., f_k, g_1, ..., g_l]](\varphi, g_1^t\psi)^{(i)}.$$

This is because w_φ is an absolute limit to the order of transvectants in this box product that will be well-defined, and any such transvectant $(f_1^{m_1}...f_k^{m_k}\varphi, g_1^{n_1}...g_l^{n_l}...g_l^t\psi)^i$ can be replaced by $f_1^{m_1}...f_k^{m_k}\varphi, g_1^{n_1}...g_l^{n_l}(g_1^t\varphi, \psi)^{(i)}$.

Now we make the induction hypothesis that cases $(p - 1, q), (p, q - 1)$, and $(p - 1, q - 1)$ can be handled, and we treat the case (p, q). Choose one of the p functions in $f_1, ..., f_k$ having positive weight; we assume the chosen element is f_1. Similarly, choose a function of positive weight from $g_1, ..., g_l$ and suppose it is g_1. Let s and t be the smallest integers such that $s.w_{f_1} = t.w_{g_1}$

Expand

$$\mathbb{R}[[f_1, ..., f_k]]\varphi = (\bigoplus_{u=0}^{s-1} \mathbb{R}[f_2, ..., f_k]]f_1^\mu\varphi) \oplus \mathbb{R}[[f_1, ..., f_k]]f_1^s\varphi$$

and

$$\mathbb{R}[[g_1, ..., g_l]]\psi = (\bigoplus_{v=0}^{t-1} \mathbb{R}[g_2, ..., g_l]]g_1^v\psi) \oplus \mathbb{R}[[g_1, ..., g_l]]g_1^\nu\psi$$

Taking the box product of these last two expansions and distribute the product. There are four kinds of terms. Terms that are missing both f_1 and g_1 in square brackets are of type $(p - 1, q - 1)$. Terms that are missing f_1 in square brackets, but not g_1 are of type $(p-1, q)$ and there are likewise terms of type $(p, q-1)$. All of these can be handled by the induction hypothesis. Finally, there is the term

$$\mathbb{R}[[f_1, ..., f_k]]f_1^s\varphi \boxtimes \mathbb{R}[[g_1, ..., g_l]]g_1^t\psi.$$

There is no upper limit to the transvectant order that can occur here, since in general there remain terms of positive weight in the square brackets. However, setting $r = s.w_{f1} = t.w_{g1}$ it can be shown that this box product is equivalent to

$$\mathbb{R}[[f_1, ..., f_k]]f_1^s\varphi \boxtimes \mathbb{R}[[g_1, ..., g_l]]g_1^t\psi \cong$$

$$\Big(\bigoplus_{i=0}^{r-1}\mathbb{R}[[f_1,...,f_k,g_1,...,g_l]](f_1^s\varphi,g_1^t\psi\Big)^{(i)}\oplus\Big(\mathbb{R}[[f_1,...,f_k]]\varphi\boxtimes\mathbb{R}[[g_1,...g_l]]\psi)(f_1^s,g_1^t\Big)^{(r)}$$

The final term is quite different from any other considered so far, since it involves a box product of subspaces as the coefficient (f_1^s, g_1^t). At this point we have reduced the calculation of $\mathbb{R}[[f_1,...,f_k]]\varphi\boxtimes\mathbb{R}[[g_1,...,g_l]]\psi$ in the case (p,q) to a number of terms computable by the induction hypothesis plus one special term that lead in circles since it involves the very same box product that we are trying to calculate. Thus our result has the form

$$\mathfrak{R}=\delta\oplus\mathfrak{R}\theta$$

where $\theta=(f_1^s,g_1^t)^{(r)}$ has weight zero. But this implies $\mathfrak{R}=\delta\oplus(\delta\oplus\mathfrak{R}\theta)\delta=\delta\oplus\delta\theta\oplus\mathfrak{R}\theta^2$ which reduces to $\mathfrak{R}=\delta[[\theta]]$.

This simply means that we erase the term $(\mathbb{R}[[f_1,...,f_k]]\varphi\boxtimes\mathbb{R}[[g_1,...,g_l]]\psi)(f_1^s,g_1^t)^{(r)}$ from our computation, and instead insert $\theta=(f_1^s,g_1^t)^{(r)}$ into the square brackets in all the coefficient rings that have already been computed. This does not affect the induction, because the new elements added have weight zero, and the induction is on the numbers p and q of elements of positive weight.

4. The ring of invariants for coupled $N_{33...3}$ Systems

The ring of invariants of N_3 in $R[x,y,z]$ is $ker\ N_3$. Let $N=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}$.

The $sl(2)$ triad will be as follows;

$$X=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix},\qquad Y=\begin{bmatrix}0&0&0\\2&0&0\\0&2&0\end{bmatrix},\qquad Z=\begin{bmatrix}2&0&0\\0&0&0\\0&0&-2\end{bmatrix}.$$

Having obtained the triad $\{X,Y,Z\}$, we create additional triad $\{\mathcal{X},\mathcal{Y},\mathcal{Z}\}$ as

$$\mathcal{X}=\mathcal{D}_Y=2x\frac{\partial}{\partial y}+2y\frac{\partial}{\partial z}$$

$$\mathcal{Y}=\mathcal{D}_X=y\frac{\partial}{\partial x}+z\frac{\partial}{\partial y}$$

$$\mathcal{Z}=\mathcal{D}_Z=2x\frac{\partial}{\partial x}-2z\frac{\partial}{\partial z}$$

The differential operators $\{\mathcal{X},\mathcal{Y},\mathcal{Z}\}$ map each vector space of homogeneous scalar polynomials into itself, with $\mathcal{X}$ and $\mathcal{Y}$ being nilpotent and $\mathcal{Z}$ semisimple. The eigenvectors of $\mathcal{Z}$ (called weight vectors) are the monomials x^m and the associated eigenvalues(called weights) are $<m,\mu>$ where $\mu=(\mu_1,...,\mu_n)$ are the

eigenvalues of $\mathcal{Z}$ that is $\mathcal{Z}(x^n) =< m, \mu > x^m$.

The basic invariants can be shown to be $\alpha = x, \beta = y^2 - 2xz$. Here α is of degree 1 weight 2 and β is of degree 2 weight 0. Every element of $ker\ \mathcal{X}_3$ can be written uniquely as a formal series $f[\alpha, \beta]$ in x, y, z. We describe this by the Stanley decomposition $ker\ \mathcal{X}_3 = \mathbb{R}[[\alpha, \beta]]$.

4.1 Linear part N_{33}

From above the Stanley decomposition of $ker\ \mathcal{X}_3 = \mathbb{R}[[\alpha_1, \beta_1]]$, we have by Theorem 6 that $ker\ \mathcal{X}_{33} = ker\ \mathcal{X}_3 \boxtimes ker\ \bar{\mathcal{X}}_3$, where $ker\ \bar{\mathcal{X}}_3 = \mathbb{R}[[\alpha_2, \beta_2]]$ corresponds to the second block in N_{33}. Expanding, we have

$$ker\ \mathcal{X}_3 = \mathbb{R}[[\beta_1]] \oplus \mathbb{R}[[\alpha_1, \beta_1]]\alpha_1$$

$$ker\ \bar{\mathcal{X}}_3 = \mathbb{R}[[\beta_2]] \oplus \mathbb{R}[[\alpha_2, \beta_2]]\alpha_2.$$

Note that β_1 and β_2 are terms of weight zero and we do not expand along terms of weight zero, so they are suppressed and will have them appear in every square brackets of the box product we compute. Therefore

$$ker\ \mathcal{X}_{33} \quad = \quad [\mathbb{R} \oplus \mathbb{R}[[\alpha_1]]\alpha_1] \boxtimes [\mathbb{R} \oplus \mathbb{R}[[\alpha_2]]\alpha_2]$$

Distributing the box product considering well defined transvectants $(f, g)^i$ according to Lemma 5 gives three kinds of terms:

1. Two terms that are immediately computed in final form: $\mathbb{R} \oplus \mathbb{R}[[\alpha_1]]\alpha_1$
2. One box product: $\mathbb{R} \boxtimes \mathbb{R}[[\alpha_2]]\alpha_2 = \mathbb{R}[[\alpha_2]]\alpha_2$
3. One box product $\mathbb{R}[[\alpha_1]]\alpha_1 \boxtimes \mathbb{R}[[\alpha_2]]\alpha_2$. This will recycle to $\mathbb{R}[[\alpha_1]] \boxtimes \mathbb{R}[[\alpha_2]]$.

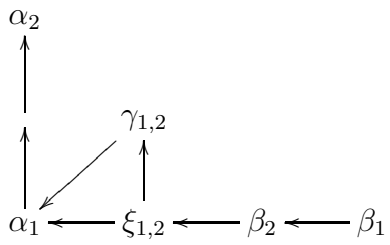
Indeed:

$$\mathbb{R}[[\alpha_1]]\alpha_1 \boxtimes \mathbb{R}[[\alpha_2]]\alpha_2 = \mathbb{R}[[\alpha_1, \alpha_2]]\alpha_1\alpha_2 \oplus \mathbb{R}[[\alpha_1, \alpha_2]](\alpha_1, \alpha_2)^{(1)} \oplus \mathbb{R}[[\alpha_1, \alpha_2]](\alpha_1, \alpha_2)^{(2)}.$$

Let $(\alpha_1, \alpha_2)^{(1)} = \gamma_{1,2}$ and $(\alpha_1, \alpha_2)^{(2)} = \xi_{1,2}$. According to recycling rule the last term will be deleted and $\xi_{1,2}$ which has weight zero will be inserted to all square brackets along side the suppressed invariants. Collecting and recombining all the terms, whenever possible we have:

$$ker\ \mathcal{X}_{33} = \mathbb{R}[[\alpha_1, \alpha_2, \beta_1, \beta_2, \xi_{1,2}]] \oplus \mathbb{R}[[\alpha_1, \alpha_2, \beta_1, \beta_2, \xi_{1,2}]]\gamma_{1,2}.$$

The same Stanley decomposition can also be obtained from the lattice diagram below where the Stanley terms are viewed as a sum of the path from β_1 to α_2 and path from β_1 to α_2 with a corner at $\gamma_{1,2}$.



with the monotone paths:
$(\beta_1 \to \beta_2 \to \xi_{1,2} \to \alpha_1 \to \alpha_2)$
$(\beta_1 \to \beta_2 \to \xi_{1,2} \to \alpha_1 \to \alpha_2)\gamma_{1,2}$.

4.2 Linear part N_{333}.

When $n = 3$ the Stanley decomposition for the ring of invariants is given by

$$\begin{aligned} ker\mathcal{X}_{333} &= ker\mathcal{X}_{33} \boxtimes ker\mathcal{X}_3 \\ &= [\mathbb{R}[[\alpha_1, \alpha_2, \beta_1, \beta_2, \xi_{1,2}]] \oplus \mathbb{R}[[\alpha_1, \alpha_2, \beta_1, \beta_2, \xi_{1,2}]]\gamma_{1,2}] \boxtimes \mathbb{R}[[\alpha_3, \beta_3]]. \end{aligned}$$

There are two cases to consider. Distributing the box products and recombining terms where possible we have:

$$\begin{aligned} ker\mathcal{X}_{333} &= \mathbb{R}[[\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \xi_{1,2}, \xi_{1,3}]] \oplus \\ &\quad \mathbb{R}[[\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \xi_{1,2}, \xi_{1,3}]]\gamma_{1,2} \oplus \\ &\quad \mathbb{R}[[\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \xi_{1,2}, \xi_{1,3}]]\gamma_{1,3} \oplus \\ &\quad \mathbb{R}[[\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \xi_{1,2}, \xi_{1,3}]]\gamma_{1,2}\gamma_{1,3} \oplus \\ &\quad \mathbb{R}[[\alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \xi_{1,2}, \xi_{1,3}, \xi_{2,3}]]\gamma_{2,3} \oplus \\ &\quad \mathbb{R}[[\alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \xi_{1,2}, \xi_{1,3}, \xi_{2,3}]]\xi_{2,3} \oplus \\ &\quad \mathbb{R}[[\alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \xi_{1,2}, \xi_{1,3}, \xi_{2,3}]]\gamma_{1,2}\gamma_{2,3} \oplus \\ &\quad \mathbb{R}[[\alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \xi_{1,2}, \xi_{1,3}, \xi_{2,3}]]\gamma_{1,2}\xi_{2,3} \oplus \\ &\quad \mathbb{R}[[\alpha_3, \beta_1, \beta_2, \beta_3, \xi_{1,2}, \xi_{1,3}, \xi_{2,3}]](\gamma_{1,2}, \alpha_3)^{(1)} \oplus \\ &\quad \mathbb{R}[[\alpha_3, \beta_1, \beta_2, \beta_3, \xi_{1,2}, \xi_{1,3}, \xi_{2,3}]](\gamma_{1,2}, \alpha_3)^{(2)}. \end{aligned}$$

To verify that this is the true Stanley decomposition consider the table function which in this case is given by

$$\begin{aligned} T_9 = &\frac{dw^2}{(1-dw^2)^3(1-d^2)^5} + \frac{d^3w^4}{(1-dw^2)^3(1-d^2)^5} + \frac{d^2w^2}{(1-dw^2)^3(1-d^2)^5} \\ &+\frac{d^4w^4}{(1-dw^2)^3(1-d^2)^5} + \frac{1}{(1-dw^2)^2(1-d^2)^6} + \frac{d^2w^2}{(1-dw^2)^2(1-d^2)^6} \\ &+\frac{d^4w^4}{(1-dw^2)^2(1-d^2)^6} + \frac{d^3w^2}{(1-dw^2)(1-d^2)^6} + \frac{d}{(1-dw^2)(1-d^2)^6} \end{aligned}$$

By multiplying the table function by w, differentiating with respect to w and putting $w = 1$, it can be easily be shown that

$$\frac{\partial}{\partial w} wT_9 \mid_{w=1} = \frac{1}{(1-d)^9}$$

Hence the table function is perfect thus all the tranvectants have been found.

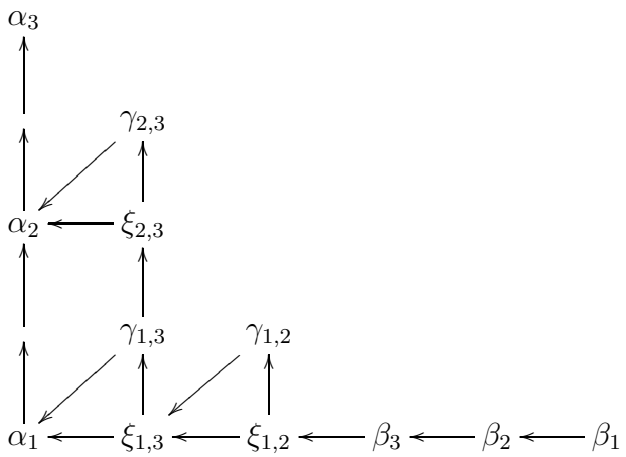
The following observations are made from the Stanley decomposition above:

(a) first term of Stanley decomposition has no product outside the square bracket.

(b) the transvectants $\gamma_{1,2}, \gamma_{1,3}$ and $\gamma_{2,3}$ never appears inside the square brackets.

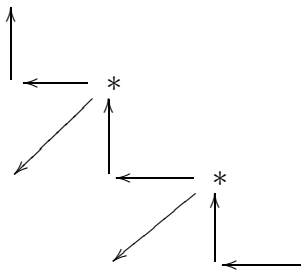
(c) the transvectants $\xi_{1,2}, \xi_{1,3}$ and $\xi_{2,3}$ appears inside as well as outside the square brackets.

It is evident that the same Stanley decomposition of N_{333} can be obtained from sum of the paths in the following lattice diagram:



where

- Every path takes the form:



  We refer to $*$ as a corner.

- Each square brackets of the Stanley decomposition contains all invariants in a path except $\gamma_{k,l}$ and the product of transvectants outside the square bracket is the product of the invariants at the corners. -

- Stanley decomposition of the ring of invariants $ker\ \mathcal{X}_{333}$ is then given by the sum of the terms T_1 and T_2, where

  T_1=$\bigoplus_j \mathbb{R}$[[invariants on the j^{th} path]](product of corners on the j^{th} path) exiting at α_k and ending at α_3 where $k = 1, 2$.

  T_2=$\bigoplus_j \mathbb{R}$[[invariants on the j^{th} path]](product of corners on the j^{th} path, $\alpha_3)^{(i)}$ exiting at α_2 through $\gamma_{2,3}$ and ending at α_3 where $i = 1, 2$.

From the above examples, we conclude that:

- For every additional n, there are new transvectants $(\alpha_k, \alpha_l)^{(i)}$ where $i = 1, 2$ and $1 \leq k < l \leq n$.

- The lattice diagram of $N_{3^{(n)}}$ is obtained by adding these new transvectants together with α_n to the lattice diagram for $N_{3^{(n-1)}}$.

- The first term of the Stanley decomposition has no products of transvectants outside the square brackets.

- The transvectants $(\alpha_k, \alpha_l)^{(1)} = \gamma_{k,l}$ where $1 \leq k < l \leq n$ never appears inside the square brackets.

- The transvectants $(\alpha_k, \alpha_l)^{(2)} = \xi_{k,l}$ where $1 \leq k < l \leq n$ appears inside as well as outside the square brackets.

- The Stanley decomposition of $N_{3^{(n)}}$ is the sum of terms T_1 and T_2 where

  $T_1 = [[$*invariants on the jth path*$]]($*products of corners on the jth path*$)$. *Exit at* α_k *and end atr* α_n *wkere* $k = 1, ..., n-1$.

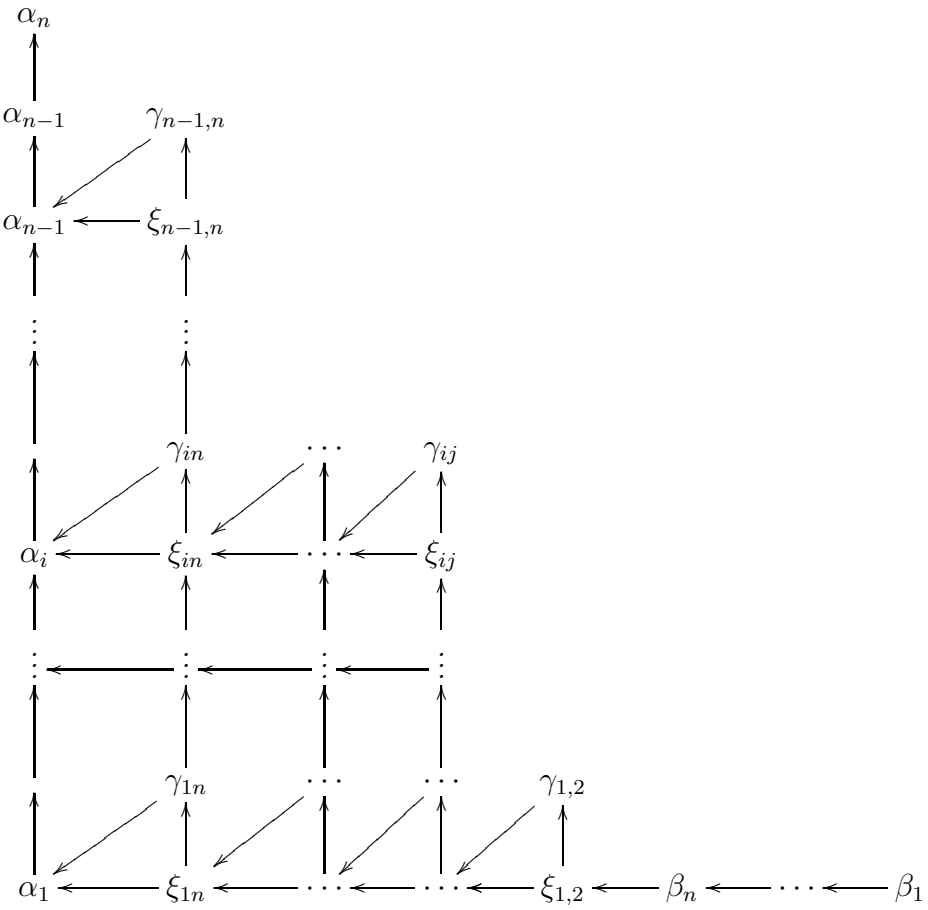
  $T_2 = [[$*invariants on the jth path*$]]($ *product of corners on the jth path,* $\alpha_m)^{(i)}$. *Exit at* α_m *through* $\gamma_{m-1,m}$ *and ending at* α_n, *where* $i = 1, 2$ *and* $m = 3, 4, ..., n$.

In general, we have the following theorem for obtaining the Stanley decomposition of systems with linear part $N_{3^{(n)}}$ as:

Theorem 4. *The Stanley decomposition of the ring of invariant of* $kerN_{(3)^n}$ *is given by the sum of terms* $\bigoplus_j T_1$ *and* $\bigoplus_j T_2$ *where* j *will range over all possible number of paths for* $kerX_{3^{(n)}}$. *Where,*

$T_1 = [[$*invariants on the jth path*$]]($*products of corners on the jth path*$)$ *exit* α_k *and end at* α_n *where* $k = 1, ..., n-1$.

$T_2 =$ $[[$*invariants on the jth path*$]]($ *product of corners on the jth path,* $\alpha_m)^{(i)}$ *exit at* α_m *through* $\gamma_{m-1,m}$ *and ending at* α_n, *where* $i = 1, 2$ *and* $m = 3, 4, ..., n$.



Proof. We prove by induction on n. It is true for $n = 2$ and $n = 3$, by the above examples. We suppose that it is true for $k = n-1$ and show that it hold for $k = n$. Since

$$kerX_{3^{(n)}} = kerX_{3^{(n-1)}} \boxtimes kerX_3$$

Suppressing all transvectants of the form $\beta_1, ...\beta_n$ and $\xi_{k,l}$ for $1 \leq k \leq l \leq n$ since they are of weight zero and noting that they will be added to every square brackets depending on the terms they are found we have: $\mathbb{R}[[\alpha_i, ..., \alpha_{n-1}]]\varphi \boxtimes \mathbb{R}[[\alpha_n]]$

Expanding the box product:

$\mathbb{R}[[\alpha_i, ..., \alpha_{n-1}]]\varphi \boxtimes \mathbb{R}[[\alpha_n]] = \big(\mathbb{R}[[\alpha_{i+1}, ..., \alpha_{n-1}]]\varphi \oplus \mathbb{R}[[\alpha_i, \alpha_{i+1}, ..., \alpha_{n-1}]]\alpha_i\varphi\big) \boxtimes \big(\mathbb{R} \oplus \mathbb{R}[[\alpha_n]]\big)$

Distributing the box product gives three kinds of terms.

1. Two terms that are computed to final form:

   $$\mathbb{R}[[\alpha_{i+1}, ..., \alpha_{n-1}]]\varphi \oplus \mathbb{R}[[\alpha_i, \alpha_{i+1}, ..., \alpha_{n-1}]]\alpha_i\varphi$$

2. One box product: $\mathbb{R}[[\alpha_{i+1}, ..., \alpha_{n-1}]]\varphi \boxtimes \mathbb{R}[[\alpha_n]]\alpha_n$, that must be computed by further expansions.

3. One box product that recycles: $\mathbb{R}[[\alpha_i, \alpha_{i+1}, ..., \alpha_{n-1}]]\alpha_i\varphi \boxtimes \mathbb{R}[[\alpha_n]]\alpha_n$
$= \mathbb{R}[[\alpha_i, \alpha_{i+1}, ..., \alpha_{n-1}, \alpha_n]]\alpha_i\alpha_n\varphi \oplus \mathbb{R}[[\alpha_i, \alpha_{i+1}, ..., \alpha_{n-1}, \alpha_n]](\alpha_i\varphi, \alpha_n)^{(1)}$
$\oplus [\mathbb{R}[[\alpha_i, \alpha_{i+1}, ..., \alpha_{n-1}]]\varphi \boxtimes \mathbb{R}[[\alpha_n]]](\alpha_i, \alpha_n)^{(2)}$

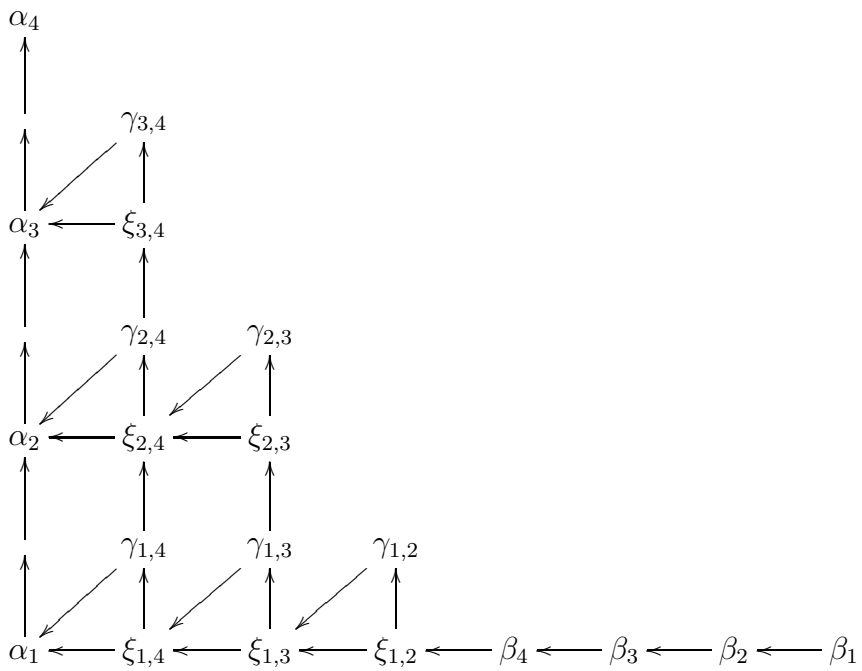
We delete the last term and insert $(\alpha_i, \alpha_n)^{(2)} = \xi_{i,n}$ in all square brackets together with other suppressed tranvectants. Recombining terms whenever possible, we finally find the Stanley decomposition of $\ker X_{3^{(n)}}$.

Equivalently, finding all the additional transvectants for $\ker X_{(3)^n}$ of the form $\gamma_{i,n}$ and $\xi_{i,n}$ where $1 \leq i < n$ and adding these together with α_n to the lattice diagram of $\ker X_{(3)^{n-1}}$, we obtain the lattice diagram for ker $X_{(3)^n}$ as shown above and the sum of the jth paths of the form T_1 and T_2 gives the Stanley decomposition of $\ker X_{3^n}$ as required. □

We summarize our work by applying Theorem 4 in finding the Stanley decomposition for the ring of invariants with linear part N_{3333}.

The new transvectants created are $\gamma_{1,4}, \gamma_{2,4}, \gamma_{3,4}, \xi_{1,4}, \xi_{2,4}, \xi_{3,4}$

By adding these transvectants and α_4 to the lattice diagram for N_{333}, we illustrate how to get the Stanley decomposition from the lattice diagram below:



The Stanley decomposition of the ring of invariants $\mathfrak{X}_{3333}$ is then given by the sum of the terms $\bigoplus_j T_1$ and $\bigoplus_j T_2$. Let $\mathfrak{R} = \mathbb{R}[[\beta_1, \beta_2, \beta_3, \beta_4, \xi_{1,2}, \xi_{1,3}]]$. The final results are:

$$
\begin{aligned}
ker\mathcal{X}_{3333} \;=\;& \mathfrak{R}[[\alpha_1,\alpha_2,\alpha_3,\alpha_4,\xi_{1,4}]] \oplus \mathfrak{R}[[\alpha_1,\alpha_2,\alpha_3,\alpha_4,\xi_{1,4}]]\gamma_{1,2} \oplus \\
& \mathfrak{R}[[\alpha_1,\alpha_2,\alpha_3,\alpha_4,\xi_{1,4}]]\gamma_{1,3} \oplus \mathfrak{R}[[\alpha_1,\alpha_2,\alpha_3,\alpha_4,\xi_{1,4}]]\gamma_{1,4} \oplus \\
& \mathfrak{R}[[\alpha_1,\alpha_2,\alpha_3,\alpha_4,\xi_{1,4}]]\gamma_{1,2}\gamma_{1,3} \oplus \mathfrak{R}[[\alpha_1,\alpha_2,\alpha_3,\alpha_4,\xi_{1,4}]]\gamma_{1,2}\gamma_{1,4} \oplus \\
& \mathfrak{R}[[\alpha_1,\alpha_2,\alpha_3,\alpha_4,\xi_{1,4}]]\gamma_{1,3}\gamma_{1,4} \oplus \mathfrak{R}[[\alpha_1,\alpha_2,\alpha_3,\alpha_4,\xi_{1,4}]]\gamma_{1,2}\gamma_{1,3}\gamma_{1,4} \oplus \\
& \mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4}]]\gamma_{2,4} \oplus \mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4}]]\gamma_{1,2}\gamma_{2,4} \oplus \\
& \mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4}]]\gamma_{1,3}\gamma_{2,4} \oplus \mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4}]]\gamma_{1,2}\gamma_{1,3}\gamma_{2,4} \oplus \\
& \mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4}]]\xi_{2,4} \oplus \mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4}]]\gamma_{1,2}\xi_{2,4} \oplus \\
& \mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4}]]\gamma_{1,3}\xi_{2,4} \oplus \mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4}]]\gamma_{1,2}\gamma_{1,3}\xi_{2,4} \oplus \\
& \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4},\xi_{3,4}]]\gamma_{3,4} \oplus \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4},\xi_{3,4}]]\gamma_{1,2}\gamma_{3,4} \oplus \\
& \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4},\xi_{3,4}]]\gamma_{1,3}\gamma_{3,4} \oplus \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4},\xi_{3,4}]]\gamma_{1,2}\gamma_{1,3}\gamma_{3,4} \\
& \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4},\xi_{3,4}]]\xi_{3,4} \oplus \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4},\xi_{3,4}]]\gamma_{1,2}\xi_{3,4} \oplus \\
& \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4},\xi_{3,4}]]\gamma_{1,3}\xi_{3,4} \oplus \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{1,4},\xi_{2,4},\xi_{3,4}]]\gamma_{1,2}\gamma_{1,3}\xi_{3,4} \oplus \\
& \mathfrak{R}[[\alpha_4,\xi_{1,4},\xi_{2,4},\xi_{3,4}]](\gamma_{1,2},\alpha_4)^{(1)} \oplus \mathfrak{R}[[\alpha_4,\xi_{1,4},\xi_{2,4},\xi_{3,4}]](\gamma_{1,2},\alpha_4)^{(2)} \oplus \\
& \mathfrak{R}[[\alpha_4,\xi_{1,4},\xi_{2,4},\xi_{3,4}]](\gamma_{1,3},\alpha_4)^{(1)} \oplus \mathfrak{R}[[\alpha_4,\xi_{1,4},\xi_{2,4},\xi_{3,4}]](\gamma_{1,3},\alpha_4)^{(2)} \oplus \\
& \mathfrak{R}[[\alpha_4,\xi_{1,4},\xi_{2,4},\xi_{3,4}]](\gamma_{1,2}\gamma_{1,3},\alpha_4)^{(1)} \oplus \mathfrak{R}[[\alpha_4,\xi_{1,4},\xi_{2,4},\xi_{3,4}]](\gamma_{1,2}\gamma_{1,3},\alpha_4)^{(2)} \oplus \\
& \mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4}]]\gamma_{2,3} \oplus \mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4}]]\gamma_{2,3}\gamma_{2,4} \oplus \\
& \mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4}]]\gamma_{1,2}\gamma_{2,3}\mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4}]]\gamma_{1,2}\gamma_{2,3}\gamma_{2,4} \oplus \\
& \mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4}]]\xi_{2,3} \oplus \mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4}]]\gamma_{2,4}\xi_{2,3} \oplus \\
& \mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4}]]\gamma_{1,2}\xi_{2,3}\mathfrak{R}[[\alpha_2,\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4}]]\gamma_{1,2}\xi_{2,3}\gamma_{2,4} \oplus \\
& \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4},\xi_{3,4}]]\gamma_{2,3}\gamma_{3,4}\mathfrak{R}[[\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4},\xi_{3,4}]]\gamma_{1,2}\gamma_{2,3}\gamma_{3,4} \oplus \\
& \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4},\xi_{3,4}]]\gamma_{2,3}\xi_{3,4} \oplus \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4},\xi_{3,4}]]\gamma_{1,2}\gamma_{2,3}\xi_{3,4} \oplus \\
& \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4},\xi_{3,4}]]\gamma_{3,4}\xi_{2,3} \oplus \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4},\xi_{3,4}]]\gamma_{1,2}\gamma_{3,4}\xi_{2,3} \oplus \\
& \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4},\xi_{3,4}]]\xi_{2,3}\xi_{3,4} \oplus \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{2,3},\xi_{2,4},\xi_{3,4}]]\gamma_{1,2}\xi_{2,3}\xi_{3,4} \oplus \\
& \mathfrak{R}[[\alpha_4,\xi_{2,3},\xi_{2,4},\xi_{3,4}]](\gamma_{2,3},\alpha_4)^{(1)} \oplus \mathfrak{R}[[\alpha_4,\xi_{2,3},\xi_{2,4},\xi_{3,4}]](\gamma_{2,3},\alpha_4)^{(2)} \oplus \\
& \mathfrak{R}[[\alpha_4,\xi_{2,3},\xi_{2,4},\xi_{3,4}]](\gamma_{1,2}\gamma_{2,3},\alpha_4)^{(1)} \oplus \mathfrak{R}[[\alpha_4,\xi_{2,3},\xi_{2,4},\xi_{3,4}]](\gamma_{1,2}\gamma_{2,3},\alpha_4)^{(2)} \oplus \\
& \mathfrak{R}[[\alpha_4,\xi_{2,3},\xi_{2,4},\xi_{3,4}]](\gamma_{1,2}\xi_{2,3},\alpha_4)^{(1)} \oplus \mathfrak{R}[[\alpha_4,\xi_{2,3},\xi_{2,4},\xi_{3,4}]](\gamma_{1,2}\xi_{2,3},\alpha_4)^{(2)} \oplus \\
& \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{2,3},\xi_{3,4}]](\gamma_{1,2},\alpha_3)^{(1)} \oplus \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{2,3},\xi_{3,4}]](\gamma_{1,2},\alpha_3)^{(2)} \oplus \\
& \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{2,3},\xi_{3,4}]](\gamma_{1,2}\gamma_{3,4},\alpha_3)^{(1)} \oplus \mathfrak{R}[[\alpha_3,\alpha_4,\xi_{2,3},\xi_{3,4}]](\gamma_{1,2}\gamma_{3,4},\alpha_3)^{(2)} \oplus \\
& \mathfrak{R}[[\alpha_4,\xi_{2,3},\xi_{3,4}]](\gamma_{1,2},\alpha_3)^{(1)},\alpha_4)^{(1)} \oplus \mathfrak{R}[[\alpha_4,\xi_{2,3},\xi_{3,4}]]((\gamma_{1,2},\alpha_3)^{(1)},\alpha_4)^{(2)}
\end{aligned}
$$

.

To verify that this is the true Stanley decomposition consider the table function which in this case is given by

$$
\begin{aligned}
T_{12} = \frac{w+3d^2w^3+3d^4w^5+d^6w^7}{(1-dw^2)^4(1-d^2)^7} &+ \frac{2d^2w^3+4d^4w^5+2d^6w^7+2d^2w+4d^4w^3+2d^6w^5}{(1-dw^2)^3(1-d^2)^8} \\
&+ \frac{d^2w^3+3d^4w^5+2d^6w^7+d^2w+4d^4w^3+3d^6w^5+d^4w+d^6w^3}{(1-dw^2)^2(1-d^2)^9}
\end{aligned}
$$

$$+\frac{3d^3w^3+3d^3w+2d^5w^5+3d^5w^3+d^5w}{(1-dw^2)(1-d^2)^9}+\frac{d^3w^3+d^3w+d^5w^5+d^5w^3}{(1-dw^2)^2(1-d^2)^8}$$
$$+\frac{d^4w^3+d^4w}{(1-dw^2)(1-d^2)^8}$$

By multiplying the table function by w, differentiating with respect to w and putting $w=1$, it can be shown that

$$\frac{\partial}{\partial w}wT_{12}\mid_{w=1}=\frac{1}{(1-d)^{12}}.$$

Bibliography

[1] W. W. Adams and P. Loustaunau, *An Introduction to Groebner Bases*, American Mathematical Society, Providence, 1994.

[2] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties and Algorithms*, Springer, New York, 1997.

[3] R. Cushman, J.A. Sanders, and N. White, Normal Form for the (2;n) Nilpotent Vector Field Using Invariant Theory, *Phys. D*, **30** (1988), 399 - 412.

[4] D. M. Malonza, Normal Forms for Coupled Takens-Bogdanov Systems, *J. Nonlinear Math. Phys.*, **11** (2004), 376 - 398.

[5] D. M. Malonza, Stanley Decomposition for Coupled Takens-Bogdanov Systems, *J. Nonlinear Math. Phys.*, **17**(1) (2010), 69-85.

[6] J. Murdock, *Normal Form and Unfoldings for Local Dynamical Systems*, Springer-Verlag, New York, 2003.

[7] J. Murdock and J.A. Sanders, A New Transvectant Algorithm for Nilpotent Normal Forms. *J. Differential Equations*, **238** (2007), 234 - 256.

[8] N.S. Namachchivaya et al., Normal form for generalized Hopf Bifurcation with Nonsemisimple 1:1 Resonance. *Z. Angew. Math. Phys*, **45**(2) (1994), 312 - 335.

Modelling Distribution of Under-Five Child Diarrhoea Across Malawi

Tsirizani T. Kaombe & Jimmy J. Namangale

Department of Mathematical Sciences, Chancellor College, University of Malawi, P.O. Box 280, Zomba, Malawi

Abstract. Analysis of diarrhoea data in Malawi has been mostly done using classical methods. However, of late new approaches such as Bayesian methods have been introduced in literature. This study aimed at trying out new statistical techniques in comparison with classical ones as well as finding out how each isolates dominant risk factors for childhood diarrhoea. This was done by fitting Logit, Poisson, and Bayesian models to 2006 Malawi Multiple Indicator Cluster Survey data, which was collected with an aim of estimating indicators of women and child health per district. The comparison between Logit and Poisson models was done via chi-square's goodness-of-fit test. Confidence and Credible Intervals were used to compare Bayesian and Logit/ Poisson model estimates. Modelling and inference in Bayesian method was done through MCMC techniques. The results showed agreement in directions and significance of estimates from Bayesian and Poisson/ Logit models, but Poisson provided better fit than Logit model. Further, all models identified child's age, breastfeeding status, region of stay, and toilet-sharing status as significant factors for determining the child's risk. The models also agreed in ruling out effects of mothers education, area of residence (rural or urban), and source of drinking water on the risk. But, Bayesian model proved significant closeness to lake/river factor, which was not the case with Poisson/ Logit model. The findings imply that classical and semi-parametric models are equally helpful, while Poisson is better than Logit model when estimating the child's risk to diarrhoea.

1. Introduction

Diarrhoea is an increase in volume of stool or frequency of defecation. It is the commonest clinical sign of gastrointestinal diseases, but it can also reflect primary disorders outside of the digestive system [27]. The disease can be manifested in different levels of clinical intensity, ranging from acute to chronic or severe stages. Acute diarrhoea appears rapidly and may last from five to ten days. Chronic diarrhoea lasts much longer than this [35]. An estimated 3.5 million deaths each year are attributable to diarrhoea worldwide, 80% of which occur in children under the age of 5 years [57]. In Malawi, the disease accounts for 11% of deaths

in children aged below 5 years [30]. To that effect, Malawi government recently adopted UNICEF/ WHO's seven-point plan for diarrhoea control. This includes fluid replacement to prevent dehydration; zinc treatment; rotavirus and measles vaccinations; promotion of early and exclusive breastfeeding and vitamin A supplementation; promotion of hand washing with soap; improved water supply quantity and quality, including treatment and safe storage of household water; and community-wide sanitation promotion [32]. Based on these guidelines, diarrhoea control policies were formulated between 2010 and 2011 by a committee with representatives from Ministries of Health and Irrigation, University of Malawi-The Polytechnic and other organisations. The committee recommended achieving political support for raising the profile of diarrhoeal disease; ensuring that policies are effectively coordinated and implemented; increasing collaboration and integration through a Technical Working Group (TWG); developing national programs; and information, education and communication to allow one clear message to be disseminated at national level [32].

However, there is still knowledge gap in citizens of sub-Saharan Africa on possible prevention strategies for childhood diarrhoea. Research indicate that morbidity and mortality from childhood diarrhoea are compounded by inappropriate household case management and frequent misuse of antibiotics by citizens of countries such as Gambia, Guinea-Bissau, Kenya, Malawi, Nigeria, Tanzania, DRC and Sudan. In addition, limited knowledge, among health care providers in these countries, of proper treatment of diarrhoea also contributed to poor outcomes [13]. More recent cross-sectional survey study in Temeke Municipality, Dar es Salaam-Tanzania indicate poor Mothers' knowledge on predisposing factors of childhood diarrhoea, which was directly correlated with education level. It was also found out that only about one third of the respondents were aware of risk factors for childhood diarrhoea that cited poor sanitation and water as main factors. Further, diarrhoea episodes were perceived wrongly as normal growth stage and that they were caused by several other "illnesses" [27]. These results agree with those of Munthali [27] in Malawi, who found out that mothers and caregivers in Rumphi wrongly associated child diarrhoea to child's teeth development and breastfeeding by a pregnant mother. A multivariate analysis in Accra, Ghana, on incidence of diarrhoea among children, by Boadi and Kuitunen [5] has attributed risk to morbidity due to diarrhoea to household economic status, mother education, access to water and sanitation facilities, hygiene practices, flies infestation and regular consumption of street food.

Most of these studies employed classical approaches whose restrictions are often violated in circumstances in which the data are collected. This often overshadows true effects of variables on risk to diarrhoea. This can be avoided by employing modern nonparametric Bayesian techniques. Fresh applications of such methods on childhood diarrhoea in sub-Saharan Africa are due to Kandala et al [17] on 1999 and 2003 Nigerian DHSs and Kazembe et al. [19] on 2000 Malawian DHS datasets. Using a Bayesian geo-additive model based on Markov-Chain-Monte-Carlo estimation, Kandala et al [17] observed that overall prevalences of diarrhoea, cough, and fever recorded in 1999 (among children aged 3 years) were similar to

those seen in 2003 (among children aged 5 years). But morbidity attributable to each of the three causes varied differently at state level. In addition, place of birth (hospital versus other), type of feeding (breastfed only versus other), parental education, maternal visits to antenatal clinics, household economic status, marital status of the mother, and place of residence (urban versus rural) were each significantly associated with the childhood morbidity studied. Further, both surveys revealed that children from urban areas were found to have a significantly lower risk of fever than their rural counterparts. It was also found out that most other factors affecting diarrhoea, cough, and fever differed in the two surveys. Besides, the risk of developing each of the three conditions increased in the first 6-8 months after birth, but then gradually declined. Likewise, through a logistic model with spatial random effects that were partitioned into shared and specific effects Kazembe et al [19] observed that shared area-specific effects were persistently high in central and southern regions of the country. On the other hand, fever-specific effects were high along the Lakeshore areas, and diarrhoea-specific effects were excessive in central and south-eastern zones of the country.

While modern approaches offer alternatives to classical methods, most studies that have used these methods in sub-Saharan Africa have applied them in isolation to classical techniques, which may not easily demonstrate their usefulness. It is important to mention that inferences from Bayesian models are usually more reasonable compared to classical estimates. Often times, classical regression provides identical estimates for all levels of a variable. For instance, an estimate may be obtained that predicts amount of change in number of diarrhoea cases corresponding to one unit change in region of stay, which may be particularly inappropriate for Bayesian multilevel application whose goal is to identify the locations in which residents are at high risk. In addition, classical regression model may over fit the data, for example giving an implausibly high estimate of average number of cases. This can happen in areas where only few diarrhoea observations were available. Bayesian multilevel models avoid this by taking into account variations in data at both individual and group levels [12]. Further, multilevel models have ability to separately estimate predictive effects of individual predictor and its group-level mean which are sometimes interpreted as "direct" and "contextual" effects of the predictor [12].

In addition, most parametric models often lack capability of identifying nonlinear relationships between dependent and independent variables. The use of Bayesian semi-parametric approaches avoids these shortcomings [16]. Nevertheless, classical models are essential in a wealth of applications where one needs to compensate for paucity of data [25]. The various approaches to data analysis (frequentist, Bayesian, machine learning, exploratory or other) should be seen as complementary to one another rather than as competitors for outright domination [25]. Unfortunately, parametric formulations become easy targets for criticism when, as occurs rather often, they are constructed with too little thought [25]. The lack of demands on the user made by most statistical packages does not help matters and, despite enthusiasm one may have for Markov chain Monte Carlo (MCMC) methods, their ability to fit very complicated parametric formulations

can be a mixed blessing [25].

This study aimed at investigating variations in the risk of diarrhoea in under-five children in Malawi by applying statistical models to explain the diarrhoea incidence. Specifically the study compared estimates found using classical and modern semi-parametric models; identified appropriate classical model to use when explaining a child's risk to diarrhoea; and evaluated influence of socio-economic and bio-demographic factors on the child's risk.

2. Methodology

2.1 Study design

This study was an applied quantitative research employing Bayesian semi-parametric additive, Logistic, and Poisson regression modelling on 2006 Malawi Multiple Indicator Cluster Survey (MICS) data. The national survey data was used with surety to allow implementation of intended statistical analyses, since estimates of a random variable from a large random sample are believed to possess all optimal properties of an estimator. Perhaps due to the rigorous process of random sampling employed surveys usually give accurate estimates of population parameters [14], a property that is desirable in statistical inference. Further, national survey data had cross-sectional information from all districts which would make it possible to estimate distributions of child diarrhoea and amount of risk posed by various parts of the country.

Further, logistic model was used due to the fact that the outcome variable, two-week total number of diarrhoea cases, was believed to follow binomial distribution. Introduced in the 1940s, Logistic regression is an example of a GLM where the random component is a Bernoulli random variable whose distribution is specified by probabilities $P(Y = 1) = \pi$ of success and $P(Y = 0) = 1 - \pi$ of failure. If the outcome of a trial can only be either a success or a failure, then the trial is called a Bernoulli trial. The total number of successes $\sum(Y = 1)$ in one Bernoulli trial, which can be 1 or 0, is called a Bernoulli random variable, $\sum(Y = 1)\ Ber(\pi)$. When many independent and identical Bernoulli trials n have been carried out, the resulting sequence of identically and independently distributed Bernoulli variables is called a Bernoulli process [1]. For n independent observations on a binary response with parameter π, the total number of successes, $\sum(Y = 1)$ has the Binomial distribution specified by the indices n and π, $\sum(Y = 1)\ Bin(n, \pi)$, and belonging to the exponential family of distributions, that is, probability mass function has form $f(y; \theta) = exp[a(y)b(\theta) + c(\theta) + d(y)]$.

Each diarrhoea observation in MICS had two possible outcomes; either a child suffered from diarrhoea or did not. Thus, each outcome was a Bernoulli process. Further, it is known that total number of children that were observed is fixed, with $n = 15,018$, and from the 2004 DHS, prevalence of under-five child diarrhoea in Malawi was 22%, hence $\pi = 0.22$, which was believed to be constant from one observation to another in the children population. In addition, each observed child was an individual and therefore the outcome in an observed child could not

influence that of the next child. Hence, outcomes were independently distributed in the children's population. Therefore, the total number of cases in the country at any time of observation was a binomial random variable. Its probability mass function is specified as

$$\begin{aligned} F(y;(n,\pi)) &= \binom{n}{y}\pi^y(1-\pi)^{n-y} \\ &= \binom{n}{y}(1-\pi)^n\left[\frac{\pi}{1-\pi}\right]^y \\ &= exp\left[\log\binom{n}{y} + n\log(1-\pi) + y\log\frac{\pi}{1-\pi}\right], \end{aligned}$$

which is an exponential form with $c(\pi) = (1-\pi)^n$, $d(y) = \binom{n}{y}$, $b(\pi) = \log(\pi/(1-\pi))$ and $a(y) = y$. The natural parameter is therefore $\log(\pi/(1-\pi))$, log of odds of response 1, the logit of π, it's canonical link. Because of this link function, the binomial or logistic model is also called logit model.

The actual value of π in the population can vary as the value, x of X varies; hence the notation π may be replaced by $\pi(x)$ to reflect its dependence on that value [2].

The relationship between x and $\pi(x)$ is a non-linear S-shaped curve, called logistic function, given by:

$$\pi(x) = E\left(\sum(Y=1)_{x_1,\cdots,x_p}\right) = \frac{exp(\beta_0+\beta_1x_1+\cdots+\beta_px_p)}{1+exp(\beta_0+\beta_1x_1+\cdots+\beta_px_p)},$$

where β_0, $\cdots$, β_p are parameters to be estimated from the data y. In situation where explanatory variable x_1 is binary exposure of interest, $exp(\beta_1)$ is adjusted ratio of odds of outcome occurring in exposed group versus non-exposed group, adjusting for effects of other explanatory variables x_2, $\cdots$, x_p [1, 2]. As x gets large, (x) approaches 0 if $\beta < 0$ and it approaches 1 if $\beta > 0$.

The transformation given below, logarithm of odds of success, called Logit transform, linearises the logistic function;

$$\hat{\pi} = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0+\beta_1x_1+\cdots+\beta_px_p.$$

The estimation of parameters is usually done through Maximum Likelihood technique. The Maximum Likelihood estimates of parameters β, and consequently of the probabilities $\pi i = g(\mathbf{x_i^T}\beta)$, are obtained by maximizing the log-likelihood function;

$$l(\pi;y) = \sum_{i=1}^{N}\left[y_i\log\pi_i + (n_i-y_i)\log(1-\pi_i) + \log\binom{n_i}{y_i}\right]$$

using iterative weighted least squares procedure (see Dobson, 2002).

Since for n Bernoulli iid observations, total number of diarrhoea cases, $\sum(Y=1)$, at any time was a positive integer, then Poisson distribution was assumed for

total number of cases throughout the four-month period. When the response is a count, one can use a count regression model to explain this response in terms of given predictors. Sometimes, total count is bounded, in which case a binomial response regression should probably be used. In other cases, the counts might be sufficiently large that a normal approximation is justified so that a normal linear model may be used [9]. One of the common distributions for counts is Poisson. If $\sum(Y = 1)$ is a Poisson random variable with mean $\mu = E[\sum(Y = 1)] = Var[\sum(Y = 1)] > 0$, then:

$$P(\sum(Y = 1) = y) = \frac{exp(\mu)\mu^y}{y!} = exp\left[\mu - \log(y!) + y\log\mu\right], y = 0, 1, 2, \cdots$$

From the exponential form, it is clear that the link function relating μ with predictors is log link given by $\log\mu = \alpha + \sum\beta x$, where the parameters are estimated using usual procedure of MLE. According to Faraway (2006), Poisson distribution arises naturally in several ways. For instance, if the count is some number out of some possible total, then the response would be more appropriately modeled as a binomial. However, for small success probabilities and large totals, the Poisson is a good approximation and can be applied. For example, in modelling the incidence of rare forms of cancer, the number of people affected is a small proportion of the population in a given geographical area. A Poisson regression model can be used in preference to a binomial. If $\mu = n\pi$ while $n \to \infty$, then $B(n, \pi)$ is well approximated by Pois(μ). Also, for small π, logit(π) $\approx \log\pi$, so that the use of the Poisson with a log link is comparable to the binomial with a logit link.

It is important to mention that to allow for correlation within households, robust standard error was to be calculated using residuals at the cluster level. An important result concerning Poisson random variables is that their sum is also Poisson. Specifically, suppose that Y_i Pois(μ_i) for $i = 1, 2,$ and are independent, then $\sum_i Y_i$ Pois($\sum_i \mu_i$). This is useful because sometimes one has access only to the aggregated data. If we assume individual-level data is Poisson, then so is the summed data and Poisson regression can still be applied [8, 9]. From 2004 Malawi DHS, $\pi = 0.22$ and $n \approx 10,000$ for under-five child diarrhoea. Therefore, π was considerably small while n being a large number. Hence, approximation of binomial regression with logit link by a Poisson regression with a log link was valid.

For the Bayesian semi-parametric model, the term comes from the fact that both parametric and nonparametric forms of relationship are assumed in one model. In this respect, continuous covariates are treated non-parametrically with the help of smoothing functions whereas categorical variables are related parametrically to the response variable [16, 17]. In general, a Bayesian model is considered to be a regression (linear or generalized linear model) in which the parameters the regression coefficients are given a probability model [11]. The use of semi-parametric model was therefore thought of in order to capture both linear effects of discrete covariates and non-linear effects of continuous covariates on the child's risk to diarrhoea. Further, the data had some categorical explanatory variables

with more than two levels; hence the model was employed in order to show the results in reduced form of covariates. Assuming that total number of observed cases at any time in the four months of MICS study, $\sum(Y = 1)$ is a random variable belonging to an exponential family with parameters n and π, then $\sum(Y = 1)$ satisfies the logistic model

$$\text{logit}(\pi) = \alpha + \sum \beta x,$$

where α and β stand for parameter components, and x for a vector of factors or covariates.

Further, it was assumed that α and β were distributed as gamma random variables with fixed scale and location parameters, $u = v = 0.001$, except for a continuous variable child age whose parameters were assumed to have normal prior distributions with 0 means and inverse gamma distributed variances.

It was assumed that regression parameters in this model are not static, but vary at: (1) child's individual-level; with focus on child's age and breast feeding status; (2) child's family-level; focusing on mother's education, family source of drinking water, and whether or not the family toilet is shared and (3) child's residential location; with focus on region of stay (north, centre, or south), closeness to the lake or river (lake/river shore or highland), and area of residence (rural or urban).

Briefly, the GLMs assume that, given covariates u and unknown parameters, the distribution of the response variable y belongs to an exponential family, with mean $\mu = E(y|y, \gamma)$ linked to a linear predictor η by

$$\mu = h(\eta), \eta = u\gamma.$$

Here, h denotes a known response function, and γ are unknown regression parameters. The following structured additive predictor was used in this study to estimate a flexible Bayesian semi-parametric model that was fitted to MICS data [6]:

$$\eta_r = f_1(x_{r1}) + \cdots + f_p(x_{rp}) + u_1\gamma,$$

where r is a generic observation index, x_{rj} denote generic covariates of different type and dimension, and $f_j(j = 1, 2, , p)$ are (not necessarily smooth) functions of the covariates. The functions f_j may comprise the usual non-linear effects of continuous covariates, time trends and seasonal effects, two-dimensional surfaces, varying coefficient terms, i.i.d. random intercepts and slopes, spatially correlated effects, and geographically weighted regression [6].

Once a model of this type is specified, inferences can be drawn from available data for population means at any level of data. These estimators, which can be regarded from a Bayesian perspective as posterior means or from a Frequentist perspective as Best Linear Unbiased Predictors (BLUPs), often have better properties than simple sample-based estimators using only data from the unit in question. This makes them useful in the problem of "small-area estimation," that is, making estimates for units or domains for which there is a very limited amount of information [33].

2.2 Geographic Location and Population Distribution

As earlier alluded to, MICS was conducted in all districts in Malawi, a country that is located in south-east Africa, landlocked between Mozambique to its eastern and southern sides, Zambia to its western side, and Tanzania to its northern side. It covers a total earth surface area of $118,484\text{km}^2$, of which $94,276\text{km}^2(79.6\%)$ is made of land and $24,208\text{km}^2(20.4\%)$ of water. By 2008, the country had a population of $13,077,160$ people and its land was divided into three major regions: central, $35,592\text{km}^2$ had $5,510,195$ people (42.14% of national population); northern, $26,931\text{km}^2$ had $1,708,930$ people (13.08% of national population); and southern region, $31,753\text{km}^2$ with $5,858,035$ people (44.80% of national population). About 90% of the countrys population lives in rural areas where, among other things, access to health services and poverty are major hardships [31]. The country's population had $7,157,985$ (45.1%) people in the age group of $0-14$ years ($3,586,696$ males; $3,571,298$ females), as of October, 2011 [24].

2.3 Study Population and Sampling Techniques

The MICS study sampled 31,200 occupied households and interviewed 30,553 (97.9%) of them. In addition, 23,238 under-five children were listed from interviewed households, of which questionnaires for 22,994 (98.9%) were completed. Further, 27,073 women (age 15-49 years) were identified from interviewed households, of which 26,259 (97%) were interviewed and 8,556 men (age 15-49 years) were identified in every third household and 7,636 (89.2%) of them were interviewed [28]. With an aim to obtain estimates, at district level, on key indicators related to well-being of children and women, MICS study targeted a sample of size 1,200 households (HHs) per district to obtain statistically valid estimates at 95% CL for majority of indicators. By then, there were 28 districts in Malawi, two of which (Likoma and Neno) were too small to draw 1,200 HHs out of total available HHs. As a result, Likoma was merged with Nkhata Bay and Neno with Mwanza, thereby reducing number of study districts to 26. Weighted estimates for the three regions and Malawi as a whole were obtained based on data from the 26 districts [28].

A two-stage cluster sampling design was used to select the households, where within each district 40 census enumeration areas (identified as clusters) were selected, and within each cluster a systematic sample of 30 households was drawn. A total of 31,200 HHs (26 districts multiplied by 1,200 HHs) were selected in 1,040 clusters (26 districts multiplied by 40 clusters) in that process. The 1,040 selected clusters were all visited during the fieldwork period [28].

The targeted population in this study was children aged at most 5 years. The outcome variable of interest was cases/non-cases of diarrhoea as in 2006 MICS. The explanatory variables included child's age, child's breastfeeding status (weaned or still breastfeeding), child's area of residence (rural or urban), region of stay (northern, central or southern), toilet facility (shared between families or not), mother's education, source of drinking water, and closeness to lake/river.

2.4 Instrumentation and Data Collection

The MICS study, conducted from July to November, 2006, used four questionnaires that were translated into Chichewa and Tumbuka vernacular languages to collect data. One questionnaire, termed household questionnaire, administered to head of household or any person who was able to provide information was used to identify all eligible persons for specific forms. It collected information regarding household listing, education, water and sanitation, household characteristics, insecticide treated nets, orphan-hood, child labour, and salt iodization. The other questionnaire, called under-five children questionnaire, administered to mothers or caretakers of under-five children collected information on Vitamin A, breastfeeding, care of illness, diarrhoea, malaria, immunization, and anthropometry. Another questionnaire, termed women questionnaire, administered to women aged 15-49 years gathered data on child mortality, birth history, tetanus toxoid, maternal and newborn health, marriage/union, contraception, sexual behaviour, HIV/AIDS, and maternal mortality. The fourth questionnaire, called men questionnaire, administered to men aged 15-49 years collected data on marriage/union, contraception, sexual behaviour, and HIV/AIDS.

2.5 Confidentiality and Ethical Clearance on Data Use

The MICS data do not show identities and particulars of respondents. This study has maintained confidentiality of participants in reporting of results. The data was used with permission from National Statistical Office of Malawi.

2.6 Data Analysis Procedures

The sample data were examined in Stata package to check for completeness of values for all variables. The children with incomplete data in some variables were dropped from analysis, with randomness assumption. Further, baseline characteristics of children were analysed in Stata Version 10 package. These included totals and percentages of studied children based on individual, household, cluster location, and regional characteristics. The variable-specific estimates of two-weeks diarrhoea incidences were calculated in Stata Version 10 package. This explored incidences before applying statistical models to data in light of study objectives. The crude odds ratios (ORs) estimating a child's risk to diarrhoea given two levels of a particular factor were calculated in Stata Version 10 package. This aimed at foreshadowing findings to third objective in this study before fitting the models to data. To achieve the objectives of this study, logistic model (with logit link), Poisson model (with log link), and Bayesian semi-parametric model were fitted to the data, using Stata Version 10 package for classical models, and BayesX package for the Bayesian model. The results from logistic model are reported as odds ratios (ORs) of effects of levels of the factors together with their corresponding 95% CIs. The relative risks with their 95% CIs are reported from the Poisson model. On the other hand, results from the Bayesian method are reported as estimates

of posterior mean effects of factors, together with their corresponding 95% CrIs, and contextual non-parametric effects, with CrIs are reported for the non-linear variable age.

The consistency of estimates between the Bayesian semi-parametric model and either Binomial or Poisson model was compared through estimates for sizes of credible and confidence intervals. This answered the first objective of this study. The logistic and Poisson models were compared based on chi-square's goodness-of-fit test results. This answered the second objective of this study. The test was preferred to the usual coefficients of determination (R2) since the two models were non-nested and the data used was enumerative (counts) in nature. A goodness-of-fit test is a statistical test of how well the data at hand support an assumption about the distribution of a population or random variable of interest [1]. The test determines how well an assumed distribution fits the data. If the data are collected in a table of k cells with at least 5 counts per cell, and observed counts in cell i are denoted O_i while expected counts are denoted E_i, then the statistic,

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i},$$

has chi-square distribution with $k-1$ degrees of freedom (that is, $E = np$ for a binomial random variable).

For a 1-tailed test, if the computed $X^2 > \text{chi} - \text{square}(k-1,)$ from distribution tables, then the null hypothesis for particular assumed distribution is rejected at α level, otherwise the null hypothesis is accepted. The closer the value observed in each cell to the expected value in that cell from the assumed distribution the higher the chances of accepting the distributional assumption of the model. Further, model adequacy statistics, such as pseudo-R^2 and parameter p-values, for individual models were studied before each model was compared with another.

For the Bayesian model, adequacy was checked via Deviance Information Criterion (DIC) and posterior predictive checking was done via posterior credible intervals. The DIC is a generalization of the Akaike Information Criterion (AIC) and Bayesian information criterion (BIC), also termed Schwarz criterion. It is most applicable in Bayesian model selection problems where the posterior distributions of the models have been obtained through Markov chain Monte Carlo (MCMC) simulation. The DIC is an asymptotic approximation as sample size gets large, just like the AIC or BIC. It is only valid when the posterior distribution is approximately multivariate normal. Deviance can be defined as $D(\theta) = -2\log(p(y|\theta)) + C$, where y is the data, θ are the unknown parameters of the model and $p(y|\theta)$ is the likelihood function. C is a constant that cancels out in all calculations that compare different models and, which therefore, does not need to be known. The expectation $\overline{D} = E[D(\theta)]$ is a measure of how well the model fits the data; the larger this is, the worse the fit. The effective number of parameters of the model is computed as $p_D = \overline{D} - D(\overline{\theta})$, where $\overline{\theta}$ is the expectation of θ. The larger this is, the better it is for the model to fit the data. The Deviance Information Criterion is calculated as

$$\text{DIC} = p_D + \overline{D}.$$

The idea is that models with smaller DIC should be preferred to models with larger DIC. Models are penalized both by the value of $\overline{D}$, which favours a good fit, but also (in common with AIC and BIC) by the effective number of parameters p_D. Since $\overline{D}$ will decrease as the number of parameters in a model increases, the term compensates for this effect by favouring models with a smaller number of parameters. Hence, DIC is a compromise between model fit and complexity [26]. The advantage of DIC over other criteria, for Bayesian model selection, is that it is easily calculated from the samples generated by the MCMC simulation. AIC and BIC require calculating the likelihood at its maximum over θ, which is not readily available from the MCMC simulation. But to calculate DIC, simply compute $\overline{D}$ as the average of $D(\theta)$ over the samples of θ, and $D(\overline{\theta})$ as the value of D evaluated at the average of the samples of θ. Then the DIC follows directly from these approximations.

2.7 Checking Randomness of Outcome Variable

The models were fitted with assumption that diarrhoea variable as well as the error resulting from fitting each parametric model was a random variable. This assumption had to be proved in the process of fitting the models. A procedure to employ depends on several factors, such as type of outcome variable (discrete or continuous), the way in which the data are observed and recorded (sequentially or not), and the nature of the study design (cluster or not), among others. One simplest method used for a binary variable recorded sequentially and randomized individually is a non-parametric test called Runs Test. A run is a sequence of like elements that are preceded and followed by different elements or no element at all [1]. By arranging the diarrhoea cases and non-cases in the order they were recorded, it was easy to come up with the Runs, and, hence, the probabilities of obtaining any number of runs. The logic behind the Runs Test for randomness is that if one obtains an extreme number of runs (too many or too few), then it can be decided that the elements in the sequence under study were not generated in a random fashion [1]. Thus, it sufficed to prove randomness using the Runs Test in this study.

The test was performed in StataSE 10 package with the assumption that data was recorded sequentially and randomized individually. A two-tailed hypothesis test that was conducted was as follows: H_0: Diarrhoea observations were generated randomly versus H_1: Diarrhoea observations were not randomly generated. The test statistic is R = number of runs. The decision rule is to reject H_0 at level α, if $R \leq C_1$ or $R \geq C_2$. In this case, C_1 and C_2 are critical values obtained from cumulative distribution function $F(r)$ for the total number of runs R in samples of sizes n_1 for cases and n_2 for non-cases, with total tail probability $P(R \leq C_1 + R \leq C_2) = \alpha$.

2.8 Validity and Reliability of Estimates

The investigators in MICS study pre-tested the questionnaires during the month of June 2006 in Chichewa and Tumbuka speaking areas of the country and in both urban and rural settings. Based on the results of the pre-test, modifications were made to the wording and translation of the questionnaires [28]. This ensured internal validity of the findings that can be gotten using MICS data. The fact that random sampling techniques were used to collect MICS data, external validity as well as reliability of results can also be assumed. However, in case the randomness test disapproves of assumption of randomness of diarrhoea variable then use of Bayesian semi-parametric model strengthens external validity and reliability of estimates from classical models where the results from the two types of models tallied.

3. Results and Interpretations

3.1 Baseline and Cross-Classification Results

There were 22, 994 under-five children who were surveyed in 2006 MICS. A total of 15, 018 (65.3%) of these had complete information on all studied variables and hence, their data was analysed in this study. The incomplete data was dropped based on randomness assumption. That is, dropped data points could produce similar results if analysed separately. Further, the large sample that remained ensured that dropping incomplete data points could not seriously distort findings. The results presented in Table 3.1 show that the study involved almost equal numbers of female (50.4%) and male (49.6%) children. Further, it is shown that most of the studied children were in age group $12-23$ months (23%), with mean age of 28 months and standard deviation of 16 months. In addition, a large proportion (56.1%) of the children was weaned, with more children (87.1%) residing in rural areas. Furthermore, most children (85.1%) had mothers whose highest education was primary. It is also indicated that most children (39.5%) were living in southern region of Malawi. Besides, more children (62.2%) were living in families that were not sharing toilets. Similarly, a large proportion of the children (71.7%) were drinking from piped water source. Finally, the results show that more children (53.1%) were residing along Lake Malawi and Shire Valley areas. In this respect, Lake Malawi and Shire Valley districts included Karonga, Rumphi, Nkhata-bay, Nkhotakota, Salima, Dedza, Mangochi, Balaka, Machinga, Zomba, Mwanza, Blantyre, Chikhwawa, Thyolo, and Nsanje. Whereas Chitipa, Mzimba, Kasungu, Ntchisi, Dowa, Lilongwe, Mchinji, Ntcheu, Phalombe, Chiradzulu, and Mulanje were regarded as highland districts.

On incidence rate, the results indicate that $3,282$ (21.85%) children had diarrhoea at some time in two weeks preceding the survey. In addition, the incidence rate was proportionally distributed in males (10.97%) and females (10.88%). Further, the rate was highest in age group $12-23$ months (8.5%). It was also high in the breastfed children (13.4%). Furthermore, the rate was proportional between

children who were living along Lake Malawi and Shire River valley (10.86%) and those from highlands (10.99%). Additionally, the rate was highest in central region (9.36%) compared to the other two regions. Similarly, the rate was higher in children who were living in rural areas of the country (19.38%). Likewise, incidence was higher in children whose families were not sharing toilets (12.45%). Besides, the rate was highest in children whose mothers' highest education was primary (18.87%) compared to other studied education levels. Finally, the rate was higher in children who were drinking from piped water (15.28%) compared to other sources.

The crude odds ratios show that female children were as likely as male children to catch diarrhoea, although gender is not significant. Further, it is shown that weaned children had 59.8% reduced odds of catching diarrhoea than those who were breastfed. Further, children aged $12-23$ months had 84.1% higher odds of catching diarrhoea than those aged $0-11$ months. While children aged $24-35$; $36-47$; and $48-59$ months had respectively 18.8%; 52.8%; 65.4% reduced odds of catching diarrhoea compared to those aged $0-11$ months. Furthermore, children from rural areas had 20.7% higher odds of catching diarrhoea than those from urban areas. Likewise, children who were living along Lake Malawi and Shire River banks had 16.2% reduced odds of catching diarrhoea than those from highlands. On region of stay, the results show that children who were living in central and southern regions had respectively 76.1% and 37.6% higher odds of catching diarrhoea compared to those who were living in the northern region. In addition, a child from secondary educated mother had 11.3% reduced odds of catching diarrhoea than the one from primary educated mother. But the results show no difference between odds of children from primary and tertiary educated mothers catching diarrhoea. Besides, the results show that children whose families were sharing toilets had 32.8% increased odds of catching diarrhoea than those whose families were not. Finally, it is shown that children who were drinking from unprotected well had 15.3% increased odds of catching diarrhoea compared to those who were drinking from piped water. But there was no significant difference in odds of catching diarrhoea between children who were drinking from piped water and those drinking from protected well or surface water.

3.2 Logistic and Poisson Model Results

The results in Table 3.2 show that logistic model as a whole fits the diarrhoea data significantly better than an empty model, that is, a model with no predictors ($LR = 985.24, p < 0.001$). However, chi-square's goodness-of-fit test result leads to rejection, at 5% level, of the binomial distribution assumption of total number of cases at any time of observation ($GoF = 1019, p = 0.0015$). For Poisson model, the output for unconditional mean and variance of diarrhoea cases give mean of 0.2185 and variance of 0.1708. The values, though for unconditional mean and variance, indicate slight under-dispersion. However, the variance is not substantially smaller than the mean, $E(\sum(Y=1)) \approx \text{var}(\sum(Y=1)) \approx 0.2$, and thus predictor variables could be of help. Further, using Microsoft Excel 'rand'

Characteristic	**Total(%)**	**Incidence (%)**	**OR(95% CI, p-value)**
Overall	15, 018 (100)	3,282 (21.85)	
Gender: Male (ref)	7,450 (49.61)	1,648 (10.97)	
Female	7,568 (50.39)	1,634(10.88)	0.969 (0.897-1.047, 0.432)
Age: 0-11(ref)	2,826 (18.82)	682 (4.54)	
12-23	3,458 (23.03)	1,277 (8.5)	1.841 (1.646-2.058, <0.001)
24-35	3,400 (22.64)	698 (4.65)	0.812 (0.72-0.916, <0.001)
36-47	3,054 (20.34)	399 (2.66)	0.472 (0.412-0.542, <0.001)
48-59	2,280 (15.18)	226 (1.5)	0.346 (0.293-0.408, <0.001)
Breastfeeding: Breastfed (ref)	6,585 (43.85)	2,013 (13.40)	
Weaned	8,433 (56.15)	1,269 (8.45)	0.402 (0.371-0.436, <0.001)
Area of residence Urban (ref)	1,936(12.89)	371 (2.47)	
Rural	13,082 (87.11)	2,911 (19.38)	1.207 (1.07-1.362, 0.002)
Altitudinal locale: Highland (ref)	7,037 (46.86)	1,651 (10.99)	
Lakeshore/riverine	7,981 (53.14)	1, 631 (10.86)	0.838 (0.775-0.905, < 0.001)
Region: Northern (ref)	3,650 (24.30)	604 (4.02)	
Central	5,429 (36.15)	1,405 (9.36)	1.761(1.582-1.96, <0.001)
Southern	5,939 (39.55	1,273 (8.48)	1.376 (1.236-1.532, <0.001
Mothers education: Primary (ref)	12,779 (85.09)	2,834 (18.87)	
Secondary	2,165 (14.42)	437 (2.91)	0.887 (0.793-0.994, 0.038)
Higher	74 (0.49)	11 (0.07)	0.613 (0.322-1.164, 0.13)
Family toilet: Not shared (ref)	9,348 (62.25)	1,869 (12.45)	
Shared	5,670 (37.75)	1,413 (9.41)	1.328 (1.228-1.437, < 0.001)
Drinking water source: Piped (ref)	10,766 (71.69)	2,294 (15.28)	
Protected well	818 (5.45)	188 (1.25)	1.102 (0.93-1.305, 0.26)
Unprotected well	2,455 (16.35)	584 (3.89)	1.153 (1.039-1.279, 0.007)
Surface water	979 (6.52)	216 (1.44)	1.045 (0.893-1.224, 0.58)

Table 3.1: Baseline analysis and cross-classification results for child diarrhoea

function, random samples of 100, 1000, 5000, 10000 and 15000 generated from the diarrhoea variable produced prevalence rates of 0.27, 0.251, 0.242, 0.241, and 0.239 respectively, indicating that increasing sample size resulted in reduction of prevalence rate. So, it was reasonable to approximate binomial model with logit link by Poisson model with log link, but with robust standard errors to account for clustering of data. The results for Poisson model with robust (residual-based) standard errors indicate that the model is significantly better than an empty model ($LR = 973, p < 0.001$). Further, the goodness-of-fit test is accepted at 5% level ($GoF = 9225, p = 1.00$), showing that the data give no statistical evidence that the diarrhoea variable does not follow Poisson distribution.

The estimates from Logit and Poisson models show that, adjusting for other factors, a weaned child had respectively 30.5% and 23.2% reduced odds and risk of catching diarrhoea compared to a breastfed child. In addition, the two models show that children who were living in central region had respectively 67.5% and 47.2% higher odds and risk of catching diarrhoea than those who were living in northern region, adjusting for other factors. Likewise, children from southern region had respectively 36.5% and 27.2% adjusted higher odds and risk of catching diarrhoea compared to children from the north. Furthermore, it is indicated that odds and risk of catching diarrhoea increased by 27.3% and 19.2% respectively in children whose families shared toilets compared to those whose families did not, controlling for other factors. The results also show that adjusted odds and risk of catching diarrhoea in children aged $12 - 23$ months were respectively higher by 92.8% and 57% than in children aged $0 - 11$ months, while there was no difference in odds or risk between age group $24 - 35$ and age $0 - 11$ months. However, the adjusted odds and risk were respectively lower by 33.1% and 30% in children aged $36 - 47$ and lower by 50.4% and 46.2% in age $48 - 59$ compared to those aged $0 - 11$ months. Similarly, both models showed that children living in families that shared toilets had 27.3% and 19.2% respectively higher odds and risk of catching diarrhoea.

Finally, the two models showed no evidence of difference in adjusted odds and risk of catching diarrhoea between children living in rural and urban areas, lakeshore/riverside areas and highlands, primary educated and higher than primary educated mothers, and in children drinking from piped and other sources of drinking water. These results agree with crude ORs reported before.

3.3 Runs Test for Randomness Results for Diarrhoea Variable

The results from Runs Test for randomness of diarrhoea variable analysed in Stata for $n = 15,018$, using either continuity or split mean as cut-off points, with or without continuity correction produced number of runs statistic, $r = 4,963$ ($z = -3.99, p < 0.0001$). Hence, the data provide no evidence, at 5% level, that the diarrhoea observations were generated in randomly. This was expected as 2006 MICS sampling was done at cluster level and not individual level. Since, analysis of the data is done list wise, it is very likely to find that observations are not a random sample viewed from case by case situation rather than cluster wise.

Variable	**Logit, OR (95%CI, p)**	**Poisson, RR (95%CI, p.)**
Breastfeeding: ref (Breast-fed)	0.695 (0.6-0.8, p<0.001)	0.768 (0.693 - 0.85, p<0.001)
Age in months ref (0-11)		
12-23	1.928 (1.17-2.16, p < 0.001)	1.57 (1.452-1.699, p<0.001)
24-35	1.09 (0.93-1.29, p = 0.29)	1.055(0.936-1.19, p = 0.378)
36-47	0.669 (0.55-0.81, p < 0.001)	0.7(0.604-0.813, p< 0.001)
48-59	0.496 (0.4-0.61, p < 0.001)	0.538(0.453-0.637, p<0.001)
Area of residence: ref(Urban)	1.122 (0.985, 1.28, p=0.08)	1.09(0.989-1.202, p=0.083)
Attitudinal local: ref(highland)	0.918 (0.84, 1.002, p=0.055)	0.939(0.881-1.001, p=0.052)
Region: ref (Northern)		
Central	1.678(1.495-1.88, p<0.001)	1.472 (1.348-1.608, p<0.001)
Southern	1.365(1.22-1.528, p<0.001)	1.272 (1.167-1.387, p<0.001)
Mothers education: ref(primary)		
Secondary	0.922 (0.818, 1.04, p=0.185)	0.941(0.861-1.029, p=0.182)
Higher	0.783 (0.404, 1.52, p=0.47)	0.822 (0.485-1.395, p=0.468)
Family toilet: ref(not shared)	1.273 (1.17-1.38, p<0.001)	1.192 (1.123-1.266, p<0.001)
Drinking water source: ref(piped)		
Protected well	0.997(0.84, 1.19, p=0.97)	0.995 (0.878-1.129, p=0.942)
Unprotected well	1.036(0.93, 1.16, p=-0.53)	1.025 (0.947-1.109, p=0.539)
Surface water	1.084(0.92, 1.28, p=0.343)	1.063 (0.943-1.2, p=0.321)
Overall model fit	GoF=1019, p=0.002; LR=985, p<0.001	GoF=9225, p=1.00; W=973, p<0.001

Table 3.2: Logit and Poisson model, adjusted OR and RR, results

3.4 Bayesian semi-parametric model results

The results presented in this section were run in BayesX package Version 2.0.1, using the following code:

```
b.regress ca1 = cage_11(psplinerw2) + bf2 + ed3a + ws8 + water3
   + dist3 + hh6 + ufreg, family =binomial iterations=12000 burnin
   =2000 step=10 predict using d.
```

The Markov chain Monte Carlo (MCMC) simulations were run on the set of full conditional posterior distributions in order to derive the full posterior estimates for all parameters of interest (see [10]). The options iterations, burning and step define the total number of iterations, the burn in period, and the thinning parameter of the MCMC simulation run [6]. Specifying step=10 as above forces BayesX to store only every 10th sampled parameter which leads to a random sample of length 1000 for every parameter in this case. Therefore, a sample of 10000 random numbers is obtained with the above specifications. The model presented in Table 3.3 has DIC based on un-standardized deviance results, Deviance $(bar_m u)$ = 14847.347, pD = 11.887, DIC = 14871.121, and DIC based on saturated deviance results, Deviance (bar_mu) = 14847.347, pD = 11.887, DIC = 14871.121.

The results show that posterior mean amount of diarrhoea cases were expected to be low in weaned children, in children whose mothers education was higher than primary, and in children who lived close to Lake Malawi or Shire River. However, the posterior mean amount of cases were expected to be high in children whose families were sharing toilets, in children who were drinking from non-piped water source, in children who were living in rural areas, and in children who were living in other regions than northern region. These results were supported by direct fixed-effects results for each categorical variable that were analysed in BayesX as well. Finally, it is clear that the 80% credible intervals indicate significance of all variables studied. While, the 95% credible intervals show that mothers education, source of drinking water and area of residence were not significant factors for determining a child risk to diarrhoea, but the rest variables are proven significant. These results agree with those from logit and Poisson models.

For non-linear effects of age, it is clear that expected posterior mean cases of diarrhoea was low in age groups $0-11$ months, $36-47$ months, and $48-59$ months, but high in age groups $12-23$ months and $24-35$ months. However, the 95% credible intervals show that age groups $0-11$ and $24-35$ months have no significant effects. But the most vulnerable age group to diarrhoea is $12-23$ months as found in logit and Poisson models.

Variable	Posterior mean	95%CrI	80%CrI
Constant	-1.137	-1	-1
Breastfeeding: ref(breastfed)	-0.376	-1	-1
Mothers education: ref (primary)	-0.111	0	-1
Family toilet: ref(not shared)	0.258	1	1
Drinking water source: ref(piped)	0.029	0	1
Altitudinal locale: ref(highland)	-0.202	-1	-1
Area: ref(urban)	0.110	0	1
Region: ref(northern)	0.123	1	1
Age group in months: 0-11	- 0.016	0	0
12-23	0.624	1	1
24-35	0.078	0	1
36-47	-0.411	-1	-1
48-59	-0.714	-1	-1

Table 3.3: Bayesian model results

4. Discussion of Results

4.1 Consistency of estimates found by Bayesian and Logit/ Poisson models

The presented results have shown that significance and direction of estimates from Bayesian semi-parametric model and Poisson or logit model were generally similar. The exception is in closeness to lake/river variable which was found to be statistically significant using the Bayesian model but insignificant using logit or Poisson model. The three models have coincidentally ruled out usefulness of mother's education, area of residence (rural or urban), and source of drinking water in determining the child's risk of diarrhoea.

4.2 Classical Models Comparison

The chi-square's goodness-of-fit tests' results suggest that Poisson log-linear regression model, with robust standard errors, fits the diarrhoea data set well than the logistic regression model. This was expected as, unlike logistic model, Poisson model with robust standard errors takes into account household correlations due to clustering of data.

4.3 Risk Factors for Child Diarrhoea

The results suggest that gender of a child has little (if any) to do with a child's risk to diarrhoea, as female were as likely as male children to catch diarrhoea. This may imply that biological make-up of a child's body gives no bias to gender in terms of vulnerability to diarrhoea. Further, it has been found out that breastfeeding status of a child is useful for determining a child's risk to diarrhoea. Thus, weaned

children were found to have lower chances of catching diarrhoea than breastfed children. This may reflect low possibilities of gastro transmission in a weaned child who chooses what to put to the mouth independent of the mother. It may also reflect on low hygiene considerations in breastfeeding mothers when giving food items to breastfeeding babies in the country. In addition, age of a child was found to be a useful factor in estimating a child's risk to diarrhoea. To that effect, age group $12-23$ months has been found to be the most risky group to diarrhoea compared to all other age groups studied. The results also suggest that the risk is lower in age $0-11$ months and after 23months of a child's life.

These variations across age groups may reflect breastfeeding stages of a child. For instance, the weaning time, which is reported to pose more threats of diarrhoea attacks to a child [22], is around 17.6 months [19] which is spanned in age $12-23$ months. Further, the low risk in age $0-11$ months may reflect the fact that the data had a mixture of exclusively and predominantly breastfed children who are reported to be at low risk of morbidity and mortality due to diarrhoea [3, 14, 15, 34, 57, 37] and the general breastfed children. The 2006 MICS, whose data was analysed in this study, reported that approximately 56% of children aged less than 6 months were exclusively breastfed. Thus, the observed low risk of catching diarrhoea in age $0-11$ months, which overlaps age $0-6$ months, reflect the high percentage of exclusively breastfed children analysed in the study. Furthermore, the results show a shift of most risky age group upward from age $6-11$ months reported in 2004 Malawi DHS or age $6-8$ months reported by Kandala et al (2008) for the 1999 and 2003 Nigerian DHSs to age 12-23 months reported in this study. The shift also seems to mimic the trend of diarrhoea in breastfeeding children reported recently by researchers in Malawi. Although this study did not intend to explore interactions of studied factors, there seems to be an interaction between child's age and breastfeeding status. Studies in Malawi have shown an increase in diarrhoea during and following weaning time among exclusively breastfed infants reportedly weaned at 6 months [7]. The fact that weaning time in Malawi is around 17.6 months [19] which is within $12-23$months age group then these results are not a surprise as earlier observed. As per tradition, weaning time entails introduction of complementary infant foods which, may in turn spread diarrhoea to the child if not hygienically prepared by the mother. This is why researchers have recommended that greater emphasis should be placed on hygienic preparation of weaning foods and water purification in order to decrease infant diarrhoeal morbidity in resource-limited settings [21].

Furthermore, region of stay has been found to be a significant factor for determining child's risk to diarrhoea. Thus, the results suggest that children from the central and southern regions are at higher risk compared to those from northern region. Compared with southern region, children from central region have higher chances of catching diarrhoea. The causes of such differences can be far from speculation. However, the findings agree with the 2004 Malawi DHS results and a study report by Kazembe et al (2009).

Likewise, the findings have also shown that children whose mothers' highest education qualification is secondary have marginally lower chances of catching di-

arrhoea compared to those with primary educated mothers. But the findings suggest no difference in the risk of diarrhoea between children with primary educated mothers and those with tertiary educated mothers, as well as between children with secondary educated mothers and tertiary educated mothers. However, mother's education was found to be statistically insignificant factor for determining child diarrhoea. This may reflect the way the study was designed, which just sought differences in academic qualification and not in health education of mothers. Although other studies in sub-Saharan Africa have supported influence of mother's academic qualification on a child's chances of catching diarrhoea (see [18]), there cannot be any immediate reason as to why one can think that mere differences in levels of academic or formal education achievements (other than health education) can result in differences in child's risk to diarrhoea. No wonder there was no difference in effects between secondary and tertiary education, the same results could be expected if levels of tertiary education were compared (for instance, diploma and degree). What is felt to have an effect on child's health is the mother's knowledge in health, which is richly provided in the primary education curriculum in Malawi. But also, mere health education literacy of the mother that can be attained through attendance of antenatal or postnatal care services could serve the purpose of controlling a child's health.

Besides, the findings suggest that children from rural areas have high chances of catching diarrhoea compared to those from urban areas, although area of residence was found to be statistically insignificant factor in determining a child's risk. This agrees with results from the 2004 Malawi DHS. The situation may reflect low rates of exclusive breastfeeding practices in rural areas of the country. It is reported that exclusive breastfeeding reduces diarrhoea threats in under-five children [57]. A study report by Kerr et al (2007) has indicated that only 4% of Malawian children are exclusively breastfed for 6 months in rural areas of Ekwendeni, Mzimba district. Thus, a majority of mothers living in rural areas of Ekwendeni do not practice exclusive breastfeeding during the first 6 months of a child's life. If the situation is true in other rural parts of the country, then the high risk to diarrhoea findings for rural children noted in this study may not be a surprise.

Furthermore, the results have found closeness to lake/river as a useful factor in determining a child's risk to diarrhoea. The findings suggest that children living along Lakeshore or river banks have reduced chances of catching diarrhoea compared to those from highlands. The opposite was expected, but these results may reflect high utilization of water sanitation interventions rolled by government and other stakeholders, such as free water guard in drinking water [23] and improved drinking water sources, such as piped water and boreholes, targeted to lakeshore/riverine dwellers in recent years who were previously believed to be at high risk of diarrhoea than highlanders. Thus, high use of safe and clean drinking water by residents of lakeshore or shire valley has reversed the old trend of child diarrhoea cases between highlands and lakeshore areas.

The findings also suggest that there is no difference in tendencies of catching diarrhoea in children who drink from protected well or surface water to those drinking from piped water. But children who drink from unprotected well were

found to have marginally increased chances of catching diarrhoea compared to those drinking from piped water. However, source of drinking water was found to be statistically insignificant factor in determining a child's risk to diarrhoea. But the findings may reflect splash effects of the water sanitation interventions projects, such as free water guard, which were underway in many parts of the country around or during the time of MICS study which could not bring significant differences in diarrhoea cases in children who were drinking from different water sources.

Finally, the findings suggest family toilet facility is a useful factor in estimating a child's risk to diarrhoea. Thus, children from families that shared toilets were found to have increased chances of catching diarrhoea than those whose families did not share toilets. This may reflect high possibilities of diarrhoea germs transmission from other people who use the same toilet as the child or her mother to the child.

5. Conclusions and Recommendations

5.1 Conclusions

The findings suggest that estimating child's risk to diarrhoea using Bayesian semi-parametric model is as good as using logistic or Poisson model. This is the case since the two groups of models have agreed in isolating most significant as well as insignificant factors for determining the child's risk to diarrhoea. But of the two classical models used, the goodness-of-fit of Poisson regression, with robust standard errors, is better than logistic model. It can further be concluded that region from which a child comes (northern, central, or southern), child's age, whether or not a child is still breastfeeding, whether or not a child comes from a family that shares toilet with other families, and closeness to lake/river are statistically significant factors in determining likelihood of the child catching diarrhoea. However, under-five child diarrhoea has little (if any) to do with area of residence (rural or urban), source of drinking water, and mother's education.

5.2 Implications of Findings

The findings suggest that applying Bayesian semi-parametric models together with classical models can help to confirm classical model estimates or this can provide alternative estimates which can be trusted when one is not sure of level of satisfaction of classical model assumptions. Thus, various approaches to data analysis should be seen as complementary to one another rather than as competitors for outright domination [25].

5.3 Limitations of the study

The study sample had $7,976$ (34.69%) children with missing values in at least one variable of interest. This may have influenced the results in this study in one way or another as the nature of variability of the dropped data was not known, but was

just assumed to be random. In addition, the study did not exhaust all possible models for the diarrhoea data since it was just an application study on use of statistical models in explaining under-five diarrhoea incidence. It is important to mention that other models, such as Negative Binomial, Generalized Estimating Equations were possible, especially in situations where serious under-dispersion could be noted when fitting the Poisson model and where inter-cluster correlations were possible.

The study findings on actual epidemiology of under-five child diarrhoea incidence in the Malawian population may not be accurate since the survey data used are from 2006 which is not the most current one in the country. Thus, focus of this study was on whether a statistical model can be used to explain/ predict the child's risk of diarrhoea rather than on whether the findings on diarrhoea situation in the country reflect the true current situation on the ground. Thus, much attention was on formal theory of the applied statistical models and their practical outcomes rather than on diarrhoea findings. Finally, the models applied were not extended to capture seasonality of child diarrhoea in Malawi although it is obvious that findings on seasonality of the disease could add more meaning to the study, as study reports have indicated that there is higher probability of infant diarrhoea in the rainy compared to the dry season in Malawi [7].

5.4 Recommendations

It is recommended that Bayesian semi-parametric models should be employed in parallel with classical models as a checking tool when the researcher is in doubts of meeting classical model assumptions. Further, researchers should consider fitting Poisson regression model with robust standard errors when analysing child diarrhoea data that is randomized at cluster level compared to ordinary logistic regression model. In addition, more interventions in child diarrhoea are needed in central region of the country by government and other stakeholders in health in order to contain the problem in the region. Also, Ministry of Health and other stakeholders should continue mobilising for better hygiene practices in breastfeeding mothers in the country, especially around weaning period. The ministry and other stakeholders may initiate campaign for independent family toilets in the country as child diarrhoea is associated with sharing of toilets. Above all, there is need for another study that may try to find causes of high risk to diarrhoea in children from central region of Malawi.

Bibliography

[1] A.D. Aczel, and J. Sounderpandian, *Complete Business Statistics* (5th Ed.), New York: McGraw-Hill, 2002.

[2] A. Agrest, *An Introduction to Categorical Data Analysis*, New York: John Wiley, 1996.

[3] S. Arifeen, R.E. Black, G. Antelman, A. Baqui, L. Caulfield, and S. Becker, Exclusive Breastfeeding Reduces Acute Respiratory Infection and Diarrhoea Deaths among Infants in Dhaka Slums, *Pediatr*, **108**, 67, 2001.

[4] A.P. Betran, M. Onis, J.A. Lauer, and J. Villar, Ecological Study of Effects of Breastfeeding on Infant Mortality in Latin America, *Br Med J*, **323** (2001), 15.

[5] K.O. Boadi, and M. Kuitunen, Childhood Diarrhoea Morbidity in the Accra Metropolitan Area, Ghana: Socio-economic, Environmental and Behavioral Risk Determinants, *Journal of Health and Population in Developing Countries*, (ISSN 1095-8940), 2005.

[6] A. Brezger, T. Kneib, and S. Lang, BayesX: Analyzing Bayesian Structured Additive Regression Models, *Journal of Statistical Software*, **14**(11), 2005.

[7] P. Clayden, Diarrhoea in Uninfected Infants of HIV-Positive Mothers who stop Breastfeeding at 6 Months, *HIV i-Base*, 2007.

[8] A.J. Dobson, *An Introduction to Generalized Linear Models* (2nd Ed.), New York: Chapman and Hall/CRC, 2002.

[9] J.J. Faraway, *Extending the Linear Model with r Generalized Linear, Mixed Effects and Nonparametric Regression Models*, New York: Taylor and Francis Group, 2006.

[10] A. Ferreira da Silva, cudaBayesreg: Bayesian computation in CUDA. *The R Journal*, **2**(2), 2010.

[11] A. Gelman, and J. Hill, *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge: University Press, 2011.

[12] A. Gelman, *Multilevel (Hierarchical) Modelling: What it can and cant do*, New York: Columbia University, 2005.

[13] H.D. Hamer, J. Simon, D. Thea, and G.T. Keusch, Childhood Diarrhoea in Sub-Saharan Africa. *Child Health Research Project Special Report*, April 1998.

[14] M.H. Hansen, W.N. Hurwitz and W.G. Madow, *Sample Survey Methods and Theory, vol. II*, New York: John Wiley, 1996.

[15] L.A. Hanson, R. Ashraf, S. Zaman et al, Breast Feeding is a Natural Contraceptive and Prevents Disease and Death in Infants, Linking Infant Mortality and Birth Rates, *Acta Paediatr*, **83** (1994), 36.

[16] A. Jerak and S. Wagner, *Estimating Probabilities of EPO Patent Oppositions in a Bayesian semi-parametric Regression Framework*, Department of Statistics, Munich: University of Munich, 2003.

[17] N.B. Kandala et al, semi-parametric Analysis of the Socio-Demographic Determinants of Under Nutrition in Two African Countries, Research in Official Statistics, *EUROSTAT*, **4**(1) (2001), 81-100.

[18] N.B. Kandala, C. Ji, N. Stallard, S. Stranges, and F.P. Cappuccio, Morbidity from Diarrhoea, Cough and Fever among Young Children in Nigeria, *Annals of Tropical Medicine and Parasitology*, **102**(5) (2008), 427-445.

[19] L.N. Kazembe, Spatial Modelling of Initiation and Duration of Breastfeeding: Analysis of Breastfeeding Behaviour in Malawi I, *World Health Population*, **10**(3) (2008), 14-31.

[20] L.N. Kazembe, L.S. Muula, and C. Simoonga, Joint Spatial Modelling of Common Morbidities of Childhood Fever and Diarrhoea in Malawi, *Health and Place*, **15**(1) (2009), 165-172.

[21] R.B. Kerr, P.R. Berti and M. Chirwa, Breastfeeding and Mixed Feeding Practices in Malawi: Timing, Reasons, Decision Makers, and Child Health Consequences, *Food and Nutrition Bulletin*, **28**(1), 2007.

[22] A. Kourtis, D. Fitzgerald, L. Hyde et al, Diarrhoea in Uninfected Infants of HIV- Infected Mothers who stop Breastfeeding at 6 months: The BAN Study Experience. *14th CROI, Los Angeles*, Abstract 772, 2007.

[23] S. Kumwenda, *Assessment of Water Guard use at Household Level in Chikwawa District*, Masters Thesis, University of Malawi, College of Medicine, 2009.

[24] Index Mundi, *Malawi Age Structure-Demographics*, Accessed on February, 10, 2012 from: `http://www.indexmundi.com`, 2011.

[25] P. McCullagh, What is a Statistical Model? *The Annals of Statistics*, 30(5) (2002), 12251310.

[26] T. Mesele, *Bayesian Approach to Identify Predictors of Children Nutritional Status in Ethiopia* Unpublished Masters Thesis, Addis Ababa University, Addis Ababa, 2009.

[27] A.C. Munthali, Change and Continuity in the Management of Diarrhoeal Diseases in Under-Five Children in Rural Malawi, *Malawi Med Journal*, **16**(2) (2005), 43-46.

[28] K.D. Mwambete and R. Joseph, Knowledge and perception of mothers and caregivers on childhood diarrhoea and its management in Temeke Municipality, Tanzania, *Tanzania Journal of Health Research*, **12**(1), 2010.

[29] National Statistical Office of Malawi and UNICEF-Malawi, *Monitoring the Situation of Children and Women: Malawi Multiple Indicator Cluster Survey 2006*, Final Report, Zomba and Lilongwe, Malawi: UNICEF and NSO, 2008.

[30] National Statistical Office of Malawi, *Malawi Demographic and Health Survey, 2004*, Calverton, MD: ORC Macro, 2005.

[31] National Statistical Office of Malawi, *Population and Housing Census 2008*, Zomba: NSO, 2008.

[32] PATH, *Charting the Course for Integrated Diarrhoea Control in Malawi: A Way Forward for Policy Change*, Washington, DC 20001 USA: PATH, 2011.

[33] C.J. Skinner, D. Holt and T.M.F. Smith, *Analysis of Complex Surveys*, Chichester: Wiley, 1989.

[34] C.G. Victora, S.R. Huttly, S.C. Fuchs et al, Deaths Due to Dysentery, Acute and Persistent Diarrhoea Among Brazilian Infants, *Acta Paediatr Suppl*, **381** (1992), 711.

[35] *Medical News Today*, What is diarrhoea? What causes diarrhoea? Accessed on December 30, 2011 from: `http://www.medicalnewstoday.com/articles/158634.php`

[36] WHO, Collaborative Study Team on the Role of Breastfeeding on the Prevention of Infant Mortality, Effect of Breastfeeding on Infant and Child Mortality Due to Infectious Diseases in Less Developed Countries: A Pooled Analysis. *Lancet*, **355** (2000), 451455.

[37] P.W. Yoon, R.E. Black, L.H. Moulton, and S. Becker, Effects of not Breastfeeding on the Risk of Diarrhoea and Respiratory Mortality in Children Under 2 Years of Age in Metro Cebu, The Philippines, *Amer J Epidemiol*, **143** (1996), 11421148.

Normal Form for Systems with Linear Part $N_{(3)^n}$

David M. Malonza[1], Grace W. Gachigua[2] and Johana K. Sigey[3]

[1] *Corresponding author. Kenyatta University, Department of Mathematics, Nairobi, Kenya. Email: dmalo2004@gmail.com*

[2] *Kimathi University College of Technology, Department of Mathematics, Nyeri, Kenya. Email: wambuigachigua@gmail.com*

[3] *Jomo Kenyatta University of Agriculture and Technology, Department of Pure and Applied Mathematics, Nairobi, Kenya. Email: jksigey2002@yahoo.com*

Abstract. The concept of normal form is used to study the dynamics of non-linear systems. The set of systems of differential equations that are in normal form with respect to a particular linear part has the structure of a module of equivariants and is best described by giving a Stanley decomposition of that module. In this work we describe the normal form for vector fields on $\mathbb{R}^{3n}$ with linear nilpotent part made up of coupled n 3×3 Jordan blocks. We use an algorithm based on the notion of transvectants from classical invariant theory known as boosting to equivariants in determining the normal form when the Stanley decomposition for the ring of invariants is already known.

Keywords: transvectant, equivariants, box product, Stanley decomposition.

1. Introduction

There are well-known procedures for putting a system of differential equations $\dot{x} = Ax + v(x)$, where v is a formal power series beginning with quadratic terms, into normal form with respect to its linear part A, which can be found in [3] and [7]. Our concern in this paper is to describe the normal form space of A, that is the set of all v such that $Ax+v(x)$ is in normal form. Our main result is a procedure that solves the description problem when A is a nilpotent matrix $N_{33\ldots3}$ with coupled n 3×3 Jordan blocks, provided that the description problem is already for the ring of invariants of the system. This procedure will be illustrated with examples and then be generalized.

A coupled system with n 3×3 Jordan blocks has the form

$$\begin{bmatrix} \dot{x_1} \\ \dot{y_1} \\ \dot{z_1} \\ \vdots \\ \dot{x_n} \\ \dot{y_n} \\ \dot{z_n} \end{bmatrix} = \begin{bmatrix} N_3 & & & & & & \\ & N_3 & & & & & \\ & & N_3 & & & & \\ & & & \ddots & & & \\ & & & & N_3 & & \\ & & & & & N_3 & \\ & & & & & & N_3 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ \vdots \\ x_n \\ y_n \\ z_n \end{bmatrix} + \cdots$$

That is,

$$\dot{x} = N_{33\ldots3}x + \cdots, \tag{1.1}$$

where, $x \in \mathbb{R}^{3n}$, $N_{33\ldots3} = \begin{bmatrix} N_3 & & & \\ & N_3 & & \\ & & \ddots & \\ & & & N_3 \end{bmatrix}$, $N_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$,

and the dots denote higher order terms starting with quadratic terms.

A Stanley decomposition of a ring or a module is a way of expressing each element of the ring or module uniquely as a linear combination of Stanley basis elements with coefficients that are polynomials in restricted set of variables. For example, in the ring $\mathbb{R}[[x,y]]/\langle xy - x - 1\rangle$ of polynomials in x and y modulo the ideal generated by the syzygy (relation) $xy - x - 1 = 0$, any terms divisible by xy can be eliminated and each polynomial can then be expressed uniquely as $f(x) + g(y)y$. The Stanley decomposition of the ring is therefore:

$$\mathbb{R}[[x,y]]/\langle xy - x - 1\rangle = \mathbb{R}[[x]] \oplus \mathbb{R}[[y]]y.$$

The Stanley basis is $\{1, y\}$ and the coefficients are arbitrary polynomials $f(x) \in \mathbb{R}[[x]]$ and $g(y) \in \mathbb{R}[[y]]$.

The idea of normal form (simplification) near an equilibrium point goes back at least to Poincare (1880), who was among the first to bring forth the theory in a more definite form. Poincare considered the problem of reducing a system of nonlinear differential equations to a system of linear ones. The formal solution of this problem entails finding near-identity coordinate transformations which eliminate the analytic expressions of the nonlinear terms.

Cushman et al. in [2], using a method called "covariant of special equivariant" solved the problem of finding Stanley decomposition of $N_{22,\ldots,2}$ system. Their method begins by creating a scalar problem that is larger than the vector problem and their procedures are derived from classical invariant theory. Their approach made it necessary to repeat calculations of classical invariants theory at the levels of equivariants. Malonza in [5], solved the same problem by "Groebner basis" methods found in [1] rather than borrowing from classical theory.

Murdock and Sanders in [8], developed an algorithm based on the notion of transvectants to determine the form of normal form of a vector field with nilpotent linear part, when the normal form is known for each Jordan block of the linear part taken separately. The algorithm is based on the notion of transvectants from the classical invariant theory known as boosting to module of equivariants when the Stanley decomposition for the ring of invariants is known.

Sri Namachchivaya et al. in [9], have studied a generalized Hopf bifurcation with nonsemisimple 1:1 resonance having equilibrium point with the linear part governed by the matrix

$$\mathrm{A}=\begin{bmatrix} i\omega & 1 & & \\ & i\omega & & \\ & & i\omega & 1 \\ & & & i\omega \end{bmatrix}.$$

The normal form for such a system contains only terms that belong to both the *semisimple* part of A and the normal form of the nilpotent part, which is a coupled Takens-Bogdanov system with N_{22}. This example illustrates the physical significance of the study of normal forms for systems with a nilpotent linear part. In this paper we study coupled Takens-Bogdanov systems with an arbitrary number of 2×2 blocks. We shall use $N_{(3)^n}$ to denote $N_{33\ldots3}$ with n 3×3 Jordan blocks.

Our results are mainly based on the work found in [8], that is, application of transvectant's method for computing normal form for the module of equivariants of nilpotent systems. In Section 2, 3 and 4 we put together background knowledge for understanding the content of this work. In Section 5, which forms the central part of this paper we shall compute the module of equivariants, that is, the normal form of $N_{(3)^n}$.

2. Invariants and Stanley Decompositions

Let $\mathcal{P}_j(\mathbb{R}^n, \mathbb{R}^m)$ denote the vector space of homogeneous polynomials of degree j on $\mathbb{R}^n$ with coefficients in $\mathbb{R}^m$. Let $\mathcal{P}(\mathbb{R}^n, \mathbb{R}^m)$ be the vector space of all such polynomials of any degree and let $\mathcal{P}_*(\mathbb{R}^n, \mathbb{R}^m)$ be the vector space of formal power series. If $m = 1$, $\mathcal{P}_*(\mathbb{R}^n, \mathbb{R})$ becomes the ring of formal power series on $\mathbb{R}^n$, where $\mathbb{R}$ denotes the set of real numbers. From the viewpoint of smooth vectors fields, it is most natural to work with formal power series (Taylor series), but since in practice these must be truncated at some degree, it is sufficient to work with polynomials. Now, for any nilpotent matrix N, we define the Lie operator

$$L_N : \mathcal{P}_j(\mathbb{R}^n, \mathbb{R}^n) \to \mathcal{P}_j(\mathbb{R}^n, \mathbb{R}^n)$$

by

$$(L_N v)x = v'(x)Nx - Nv(x) \tag{2.1}$$

and the differential operator

$$\mathcal{D}_{Nx} : \mathcal{P}_j(\mathbb{R}^n, \mathbb{R}) \to \mathcal{P}_j(\mathbb{R}^n, \mathbb{R})$$

by

$$(\mathcal{D}_{Nx}f)(x) = f'(x)Nx = (Nx.\bigtriangledown)f(x). \tag{2.2}$$

In addition observe that,

$$L_N(fv) = (\mathcal{D}_N f)v + fL_N v. \tag{2.3}$$

A function f is called an invariant of Nx if $\frac{\partial}{\partial t}f(e^{Nt}x) \mid_{t=0}= 0$ or equivalently $f \in ker\mathcal{D}_N$. Since

$$\mathcal{D}_N(f+g) = \mathcal{D}_N f + \mathcal{D}_N g$$
$$\mathcal{D}_N fg = f\mathcal{D}_N g + g\mathcal{D}_N f$$

it follows that if f and g are invariants, so are $f+g$ and fg; that is $ker\ \mathcal{D}_N$ is both a vector space over $\mathbb{R}$ and also a subring of $\mathcal{P}(\mathbb{R}^n, \mathbb{R})$, known as the *ring of invariants.* Similarly a vector field v is called an *equivariants* of Nx, if $\frac{\partial}{\partial t}(e^{-Nt}v(e^{Nt}x)) \mid_{t=0}= 0$ that is $v \in kerL_N$.

There are two normal form styles in common use for nilpotent systems, the *inner product normal form* and the *sl(2) normal form.* The inner product normal form is defined by $\mathcal{P}(\mathbb{R}^n, \mathbb{R}^n) = imL_N \oplus kerL_{N^*}$ where N^* is the conjugate transpose of N. To define the sl(2) normal form, one first sets $X = N$ and constructs matrices Y and Z such that

$$[X, Y] = Z, \qquad [Z, X] = 2X, \qquad [Z, Y] = -2Y. \tag{2.4}$$

An example of such an $sl(2)$ triad $\{X, Y, Z\}$ is

$$X = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad Y = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \qquad Z = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Having obtained the triad $\{X, Y, Z\}$ we create two additional triads $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$ and $\{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$ as follows:

$$\mathcal{X} = \mathcal{D}_Y, \qquad \mathcal{Y} = \mathcal{D}_X, \qquad \mathcal{Z} = \mathcal{D}_Z \tag{2.5}$$

$$\mathsf{X} = L_{\mathsf{Y}}, \qquad \mathsf{Y} = L_{\mathsf{X}}, \qquad \mathsf{Z} = L_{\mathsf{Z}} \tag{2.6}$$

The first of these is a triad of differential operators and the second is a triad of Lie operators. Both the operators $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$ and $\{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$ inherit the triad properties (2.4). Observe that the operators $\{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$ map each $\mathcal{P}(\mathbb{R}^n, \mathbb{R}^n)$ into itself. It follows from the representation theory $sl(2)$ that

$$\mathcal{P}(\mathbb{R}^n, \mathbb{R}^n) = im\mathsf{Y} \oplus ker\mathsf{X} = im\mathsf{X} \oplus ker\mathsf{Y} \tag{2.7}$$

Clearly the $ker\ \mathcal{X}$ ia s subring of $\mathcal{P}(\mathbb{R}^n, \mathbb{R})$, the ring of invariants and it follows from (2.3) that $ker\ \mathsf{X}$ is a module over this subring. This is the *sl(2) normal form module.*

3. Box Products of Stanley Decompositions

Let X_1, Y_1, Z_1 be a triad acting on a vector space V_1, and X_2, Y_2, Z_2 be a triad acting on V_2. There is a natural triad X, Y, Z defined on the tensor product space $V = V_1 \oplus V_2$ by the equation $X = X_1 \boxplus X_2, Y = Y_1 \boxplus Y_2, Z = Z_1 \boxplus Z_2$, where

$$A \boxplus B = A \oplus I + I \oplus B.$$

The tensor product may be replaced with ordinary product of polynomials if there is no overlap between the variables appearing in the polynomials in the two spaces being tensored. (This nonoverlap condition implies that the ordinary product satisfies the algebraic requirements for a tensor product map.) See [6].

Let $V_k, k = 1, 2$ be $sl(2)$ representation spaces with triads $\{X_k, Y_k, Z_k\}$. Then $V_1 \otimes V_2$ is a representation space with triad $\{X, Y, Z\}$, where $X = X_1 \otimes I + I \otimes X_2$ (and similarly for Y and Z).

We define the box product of $ker\ X_1$ and $ker\ X_2$ by

$$(ker\ X_1 \boxtimes ker\ X_2) = kerX.$$

Now, Consider a system with nilpotent linear part

$$N = \begin{bmatrix} \hat{N} & 0 \\ 0 & \tilde{N} \end{bmatrix},$$

where $\hat{N}$ and $\tilde{N}$ are nilpotent matrices of sizes $\hat{n} \times \hat{n}$ and $\tilde{n} \times \tilde{n}$ respectively $(\hat{n} + \tilde{n} = n)$, in (upper) Jordan form. Let $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\}, \{\hat{\mathcal{X}}, \hat{\mathcal{Y}}, \hat{\mathcal{Z}}\}$, and $\{\tilde{\mathcal{X}}, \tilde{\mathcal{Y}}, \tilde{\mathcal{Z}}\}$ be the associated triads of operators acting on $\mathbb{R}[[x_1, ..., x_n]], \mathbb{R}[[x_1, ..., x_{\hat{n}}]]$ and $\mathbb{R}[[x_{\hat{x}+1}, ..., x_n]]$ respectively. Suppose that $f = f(x_1, ..., x_{\hat{n}}) \in \ker \hat{\mathcal{X}}$ and $g = g(x_{\hat{n}+1}, ..., x_n) \in \ker \tilde{\mathcal{X}}$ are weight invariants of weight w_f and w_g, and i is an integer in the range $0 \leq i \leq \min(w_f, w_g)$. We define *external transvectant* of f and g of order i to be the polynomial $(f, g)^{(i)} \in \mathbb{R}[[x_1, ..., x_n]]$ given by

$$(f, g)^{(i)} = \sum_{j=0}^{i} (-1)^j W_{f,g}^{i,j} (\hat{\mathcal{Y}}^j f)(\tilde{\mathcal{Y}}^{i-j} g), \tag{3.1}$$

where

$$W_{f,g}^{i,j} = \begin{pmatrix} i \\ j \end{pmatrix} \frac{(w_f - j)!}{(w_f - i)!} . \frac{(w_g - i + j)!}{(w_g - i)!}.$$

We say that a transvectant $(f, g)^{(i)}$ is *well-defined* if i is in the proper range for f and g. Notice that the zeroth transvectant is always well-defined and reduces to the product: $(f, g)^{(0)} = fg$. Now given Stanley decomposition for $\ker \hat{\mathcal{X}}$ and $\ker \tilde{\mathcal{X}}$, the following results found in [8] provide the first steps towards obtaining a Stanley decomposition for $\ker \mathcal{X}$.

Theorem 1. *Each well-defined transvectant* $(f,g)^i$ *of* $f \in ker\ \hat{\mathfrak{X}}$ *and* $g \in ker\ \tilde{\mathfrak{X}}$ *belongs to ker* $\mathfrak{X}$. *If* f *and* g *are doubly homogeneous polynomials of types* (df, wf) *and* (dg, wg) *respectively,* $(f,g)^i$ *is a doubly homogeneous polynomial of type* $(df + dg, wf + wg - 2i)$. *Suppose that Stanley decompositions for ker* $\hat{\mathfrak{X}}$ *and ker* $\tilde{\mathfrak{X}}$ *are given, then a basis for the (finite-dimensional) subspace* $(ker\ \mathfrak{X})_d$ *of homogeneous polynomials in ker* $\mathfrak{X}$ *with degree* d *is given by the set of all well-defined transvectants* $(f,g)^i$ *where* f *is a standard monomial of the Stanley decomposition for ker* $\hat{\mathfrak{X}}$ *and* g *is a standard monomial of the Stanley decomposition for ker* $\tilde{\mathfrak{X}}$ *and* $df + dg = d$.

Remark: A *standard monomial* associated with a Stanley decomposition is an expression of the form $f_1^{m_1} \cdots f_k^{m_k}\varphi$ of $\mathbb{R}[[f_1, ..., f_k]]\varphi$, where $\mathbb{R}[[f_1, ..., f_k]]\varphi$ is a term in the Stanley decomposition. Notice that "monomial" here means a monomial in the basic invariants, which are polynomials in the original variables $x_1, ..., x_n$. Given a Stanley decomposition of $\ker \mathfrak{X}$, its standard monomials of a given type(or degree) form a basis for the (finite dimensional) vector space of invariants of that type(or degree).

4. Boosting Rings of Invariants to Module of Equivariants

In this section we describe the procedure for obtaining a Stanley decomposition of the module of equivariants (or normal form space) *ker* X of a system, given the Stanley decomposition of the ring of invariants *ker* $\mathfrak{X}$ of the system.

It is well known that the module of all formal power series vector fields on $\mathbb{R}^n$ can be viewed as the tensor product $\mathbb{R}[[x_1, ..., x_n]] \otimes \mathbb{R}^n$, see[8]. In fact the tensor product can be identified with the ordinary product (of a field times a constant vector) since (just as in the case of a tensor product of two polynomial spaces with nonoverlapping variables, as mentioned in section 3) the ordinary product satisfies the same algebraic rules as a tensor product. This also follows fron the following theorem found in [6].

Theorem 2. *The unique map* $\mathcal{P}_{j+1}(\mathbb{R}^n, \mathbb{R}) \otimes \mathbb{R}^n \to \mathcal{P}_j(\mathbb{R}^n, \mathbb{R}^n)$ *sending* $f \otimes v \to fv$ *is an isomorphism of vector spaces, under which the triad* $\{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$ *on* $\mathcal{P}_j(\mathbb{R}^n, \mathbb{R}^n)$ *corresponds with* $\{\mathfrak{X} \boxplus (-Y), \mathcal{Y} \boxplus (-X), \mathcal{Z} \boxplus (-Z)\}$.

Specifically, every formal power series vector field can be written as

$$f_1(x)e_1 + ... + f_n(x)e_n = \begin{bmatrix} f_1(x) \\ . \\ . \\ . \\ f_n(x) \end{bmatrix}$$

where the e_i are the standard basis vectors of $\mathbb{R}^n$.

Next, the Lie derivative $\mathsf{X} = \mathsf{L}_{N^*}$ can be expressed as the tensor product of $\mathcal{X}$ and $-N^*$, that is $\mathsf{X} = \mathcal{X} \otimes I + I \otimes (-N^*)$. Under the identification of $\otimes$ with ordinary product, this means $\mathsf{X}(fv) = (\mathcal{X}f)v + f(-N^*)$, where $f \in \mathbb{R}[[x_1, ..., x_n$ and $v \in \mathbb{R}^n$ in agreement with the following calculation, in which $v' = 0$ because v is constant.

$$\begin{aligned} \mathsf{X}(fv) &= \mathsf{L}_{N^*}(fv) \\ &= (\mathcal{D}_{N^*}f)v + f(\mathsf{L}_{N^*}v) \\ &= (\mathcal{D}_{N^*}f)v + f(v'N^*x - N^*v) \\ &= (\mathcal{D}_{N^*})v + f(-N^*v). \end{aligned}$$

This kind of calculation also shows that $sl(2)$ representation (on vector fields) with triad $\{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$ is the tensor product of the representation (on scalar fields) with triad $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$ and the representation on $\mathbb{R}^N$ with triad $\{-N^*, -M^*, -H\}$ that is

$$ker\ \mathsf{X} = ker\ \mathcal{X} \boxtimes \mathbb{R}e_r.$$

It follows (as in Theorem 1 above) that a basis from the normal form space $ker\ \mathsf{X}$ is given by well defined transvectants $(f, v)^{(i)}$ as f ranges over a basis for $ker\ \mathcal{X} \subset \mathbb{R}[[x_1, ..., x_n]]$ and v ranges over a basis for $ker\ N^* \subset \mathbb{R}^n$. The first of these bases is given by the standard monomials of a Stanley decomposition for $ker\ \mathcal{X}$. The second is given by the standard basis vectors $e_r \in \mathbb{R}$ such that r is the index of the bottom row of a Jordan block in N. It is useful to note that the weight of such an e_r is one less than the size of the block. Then we define the transvectant $(f, v)^i$ as

$$(f, e_r)^i = \sum_{j=0}^{i} (-1)^j W^{i,j}_{f,e_r} (\mathcal{Y}^j f)((-M^*)^{j-1} e_r)$$

$$= (f, g)^i = \sum_{j=0}^{i} (-1)^j W^{i,j}_{f,g} (\mathcal{Y}^j f)((M^*)^{j-1} g).$$

From here, the computational procedures of box products are the same as those used in describing rings of invariants in [8], except that infinite iterations never arise.

5. Normal Form for Systems with Linear Part $N_{3^{(n)}}$

In this section we describe the module of equivariants (or normal form space) $ker\ \mathsf{X}$ of a system, given the Stanley decomposition of the ring of invariants $ker\ \mathcal{X}$ of the system.

First, we shall consider the normal form for nonlinear systems with linear part having two and three blocks, that is N_{33} and N_{333} as examples before generalizing to $N_{333\ldots3}$.

5.1 System with linear part N_{33}.

The Stanley decomposition for the ring of invariants with linear part N_{33} is given by: $ker\ \mathcal{X}_{33} = \mathbb{R}[[\alpha_1, \alpha_2, \beta_1, \beta_2, \xi_{1,2}]] \oplus \mathbb{R}[[\alpha_1, \alpha_2, \beta_1, \beta_2, \xi_{1,2}]]\gamma_{1,2}$ as found in [4]. Since β_1, β_2 and $\xi_{1,2}$ has weight zero, it is convenient to remove them since we do not expand along terms of weight zero by setting $\mathcal{R} = \mathbb{R}[[\beta_1, \beta_2, \xi_{1,2}]]$ and write

$$\begin{aligned} ker\ \mathcal{X}_{33} &= \mathcal{R}[[\alpha_1, \alpha_2]] \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\gamma_{1,2} \\ &= \mathcal{R}[[\alpha_2]] \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\alpha_1 \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\gamma_{1,2} \\ &= \mathcal{R} \oplus \mathcal{R}[[\alpha_2]]\alpha_2 \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\alpha_1 \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\gamma_{1,2} \end{aligned}$$

In this case the basis elements are e_3 and e_6. Therefore we need to compute the box product of the ring $ker\ \mathcal{X}_{33}$ with $\mathbb{R}e_3 \oplus \mathbb{R}e_6$ which are both of weight 2.

Therefore $ker\ \mathsf{X}_{33}$=$(ker\ \mathcal{X}_{33})\boxtimes(\mathbb{R}e_3 \oplus \mathbb{R}e_6)$. Distributing the box product there are two cases to consider.
Case 1: $[\mathcal{R} \oplus \mathcal{R}[[\alpha_2]]\alpha_2 \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\alpha_1 \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\gamma_{1,2}] \boxtimes \mathbb{R}e_3$.
There are four products namely:

a) $\mathcal{R} \boxtimes \mathbb{R}e_3 = \mathcal{R}e_3$

b)$\mathcal{R}[[\alpha_2]]\alpha_2 \boxtimes \mathbb{R}e_3 = \mathcal{R}[[\alpha_2]]\alpha_2 e_3 \oplus \mathcal{R}[[\alpha_2]](\alpha_2, e_3)^{(1)} \oplus \mathcal{R}[[\alpha_2]](\alpha_2, e_3)^{(2)}$

c) $\mathcal{R}[[\alpha_1, \alpha_2]]\alpha_1 \boxtimes \mathbb{R}e_3 = \mathcal{R}[[\alpha_1, \alpha_2]]\alpha_1 e_3 \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\alpha_1, e_3)^{(1)} \oplus \mathcal{R}[\alpha_1, [\alpha_2]](\alpha_1, e_3)^{(2)}$

d)$\mathcal{R}[[\alpha_1, \alpha_2]]\gamma_{1,2} \boxtimes \mathbb{R}e_3 = \mathcal{R}[[\alpha_1, \alpha_2]]\gamma_{1,2} e_3 \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\gamma_{1,2}, e_3)^{(1)} \oplus \mathcal{R}[\alpha_1, [\alpha_2]](\gamma_{1,2}, e_3)^{(2)}$

Recombining terms gives

$[\mathcal{R} \oplus \mathcal{R}[[\alpha_2]]\alpha_2 \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\alpha_1 \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\gamma_{1,2}] \boxtimes \mathbb{R}e_3 =$
$\mathcal{R}[[\alpha_1, \alpha_2]]e_3 \oplus \mathcal{R}[[\alpha_2]](\alpha_2, e_3)^{(1)} \oplus \mathcal{R}[[\alpha_2]](\alpha_2, e_3)^{(2)} \mathcal{R}[[\alpha_1, \alpha_2]](\alpha_1, e_3)^{(1)} \oplus$
$\mathcal{R}[[\alpha_1, \alpha_2]](\alpha_1, e_3)^{(2)} \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\gamma_{1,2} e_3 \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\gamma_{1,2}, e_3)^{(1)} \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\gamma_{1,2}, e_3)^{(2)}$.

Case 2: Similarly we have,

$[\mathcal{R} \oplus \mathcal{R}[[\alpha_2]]\alpha_2 \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\alpha_1 \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\gamma_{1,2}] \boxtimes \mathbb{R}e_6 =$
$\mathcal{R}[[\alpha_1, \alpha_2]]e_6 \oplus \mathcal{R}[[\alpha_2]](\alpha_2, e_6)^{(1)} \oplus \mathcal{R}[[\alpha_2]](\alpha_2, e_6)^{(2)} \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\alpha_1, e_6)^{(1)} \oplus$
$\mathcal{R}[[\alpha_1, \alpha_2]](\alpha_1, e_6)^{(2)} \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\gamma_{1,2} e_6 \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\gamma_{1,2}, e_6)^{(1)} \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\gamma_{1,2}, e_6)^{(2)}$

Adding terms in case 1 and 2 we obtain:

$ker\ \mathsf{X}_{33} =$
$\mathcal{R}[[\alpha_1, \alpha_2]]e_3 \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\alpha_1, e_3)^{(1)} \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\alpha_1, e_3)^{(2)} \oplus \mathcal{R}[[\alpha_2]](\alpha_2, e_3)^{(1)} \oplus$
$\mathcal{R}[[\alpha_2]](\alpha_2, e_3)^{(2)} \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\gamma_{1,2} e_3 \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\gamma_{1,2}, e_3)^{(1)} \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\gamma_{1,2}, e_3)^{(2)} \oplus$
$\mathcal{R}[[\alpha_1, \alpha_2]]e_6 \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\alpha_1, e_6)^{(1)} \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\alpha_1, e_6)^{(2)} \oplus \mathcal{R}[[\alpha_2]](\alpha_2, e_6)^{(1)} \oplus$
$\mathcal{R}[[\alpha_2]](\alpha_2, e_6)^{(2)} \oplus \mathcal{R}[[\alpha_1, \alpha_2]]\gamma_{1,2} e_6 \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\gamma_{1,2}, e_6)^{(1)} \oplus \mathcal{R}[[\alpha_1, \alpha_2]](\gamma_{1,2}, e_6)^{(2)}$.

Finally, to complete the calculation, it is necessary to compute the transvectants that appear. These are of the form $(f, e_3)^{(i)}$ and $(f, e_6)^{(i)}$ for $i = 0, 1, 2$ where

$f = \{\alpha_1, \alpha_2, \gamma_{1,2}\}$.

$$(f, e_3)^{(0)} = \begin{bmatrix} 0 \\ 0 \\ f \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned} (f, e_3)^{(1)} &= w_f f \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 2\mathcal{Y}f \\ 0 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ -1 w_f f \\ -2\mathcal{Y}f \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -\mathcal{X}\mathcal{Y}f \\ -2\mathcal{Y}f \\ 0 \\ 0 \\ 0 \end{bmatrix}. \end{aligned}$$

Observe that the weight of $w_f = \mathcal{X}\mathcal{Y}$.

$$\begin{aligned} (f, e_3)^{(2)} &= w_f(w_f - 1)\mathcal{Y}f(M^*)^2 e_3 - 2(w_f - 1)\mathcal{Y}f M^* e_3 + 2\mathcal{Y}^2 f e_3 \\ &= -2 \begin{bmatrix} \mathcal{X}^2\mathcal{Y}^2 f \\ \mathcal{X}\mathcal{Y}^2 f \\ \mathcal{Y}^2 f \\ 0 \\ 0 \\ 0 \end{bmatrix}. \end{aligned}$$

We ignore the nonzero constants -1 and -2 because we are concerned with computing basis elements. Similarly for the basis e_6 we have:

$$(f, e_6)^{(0)} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ f \end{bmatrix}, \qquad (f, e_6)^{(1)} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \mathcal{X}\mathcal{Y}f \\ \mathcal{Y}f \end{bmatrix} \qquad (f, e_6)^{(2)} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \mathcal{X}^2\mathcal{Y}^2 f \\ \mathcal{X}\mathcal{Y}^2 f \\ \mathcal{Y}^2 f \end{bmatrix}.$$

Therefore the normal form $ker\ \mathsf{X}_{33}$ for system with linear part N_{33} is :

$ker\ \mathsf{X}_{33} =$

$$\mathcal{R}[[\alpha_1,\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]\begin{bmatrix}0\\0\\1\\0\\0\\0\end{bmatrix}\oplus\mathcal{R}[[\alpha_1,\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]\begin{bmatrix}0\\\mathcal{X}\mathcal{Y}\alpha_1\\\mathcal{Y}\alpha_1\\0\\0\\0\end{bmatrix}\oplus\mathcal{R}[[\alpha_1,\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]$$
$$\begin{bmatrix}\mathcal{X}^2\mathcal{Y}^2\alpha_1\\\mathcal{X}\mathcal{Y}^2\alpha_1\\\mathcal{Y}^2\alpha_1\\0\\0\\0\end{bmatrix}\oplus\mathcal{R}[[\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]\begin{bmatrix}0\\\mathcal{X}\mathcal{Y}\alpha_2\\\mathcal{Y}\alpha_2\\0\\0\\0\end{bmatrix}\oplus\mathcal{R}[[\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]\begin{bmatrix}\mathcal{X}^2\mathcal{Y}^2\alpha_2\\\mathcal{X}\mathcal{Y}^2\alpha_2\\\mathcal{Y}^2\alpha_2\\0\\0\\0\end{bmatrix}\oplus$$
$$\mathcal{R}[[\alpha_1,\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]\begin{bmatrix}0\\0\\\gamma_{1,2}\\0\\0\\0\end{bmatrix}\oplus\mathcal{R}[[\alpha_1,\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]\begin{bmatrix}0\\\mathcal{X}\mathcal{Y}\gamma_{1,2}\\\mathcal{Y}\gamma_{1,2}\\0\\0\\0\end{bmatrix}\oplus\mathcal{R}[[\alpha_1,\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]$$
$$\begin{bmatrix}\mathcal{X}^2\mathcal{Y}^2\gamma_{1,2}\\\mathcal{X}\mathcal{Y}^2\gamma_{1,2}\\\mathcal{Y}^2\gamma_{1,2}\\0\\0\\0\end{bmatrix}\oplus\mathcal{R}[[\alpha_1,\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]\begin{bmatrix}0\\0\\0\\0\\0\\1\end{bmatrix}\oplus\mathcal{R}[[\alpha_1,\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]\begin{bmatrix}0\\0\\0\\0\\\mathcal{X}\mathcal{Y}\alpha_1\\\mathcal{Y}\alpha_1\end{bmatrix}\oplus$$
$$\mathcal{R}[[\alpha_1,\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]\begin{bmatrix}0\\0\\0\\\mathcal{X}^2\mathcal{Y}^2\alpha_1\\\mathcal{X}\mathcal{Y}^2\alpha_1\\\mathcal{Y}^2\alpha_1\end{bmatrix}\oplus\mathcal{R}[[\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]\begin{bmatrix}0\\0\\0\\0\\\mathcal{X}\mathcal{Y}\alpha_2\\\mathcal{Y}\alpha_2\end{bmatrix}\oplus\mathcal{R}[[\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]$$
$$\begin{bmatrix}0\\0\\0\\\mathcal{X}^2\mathcal{Y}^2\alpha_2\\\mathcal{X}\mathcal{Y}^2\alpha_2\\\mathcal{Y}^2\alpha_2\end{bmatrix}\oplus\mathcal{R}[[\alpha_1,\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]\begin{bmatrix}0\\0\\0\\0\\0\\\gamma_{1,2}\end{bmatrix}\oplus\mathcal{R}[[\alpha_1,\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]\begin{bmatrix}0\\0\\0\\0\\\mathcal{X}\mathcal{Y}\gamma_{1,2}\\\mathcal{Y}\gamma_{1,2}\end{bmatrix}$$
$$\oplus\mathcal{R}[[\alpha_1,\alpha_2,\beta_1,\beta_2,\xi_{1,2}]]\begin{bmatrix}0\\0\\0\\\mathcal{X}^2\mathcal{Y}^2\gamma_{1,2}\\\mathcal{X}\mathcal{Y}^2\gamma_{1,2}\\\mathcal{Y}^2\gamma_{1,2}\end{bmatrix}.$$

5.2 System with linear part N_{333}

The Stanley decomposition for ring of invariants of a system with linear part N_{333} is given by:

$$
\begin{aligned}
&ker\ \mathcal{X}_{333} = \\
&\mathbb{R}[[\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3,\xi_{1,2},\xi_{1,3}]] \oplus \mathbb{R}[[\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3,\xi_{1,2},\xi_{1,3}]]\gamma_{1,2}\oplus \\
&\mathbb{R}[[\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3,\xi_{1,2},\xi_{1,3}]]\gamma_{1,3} \oplus \mathbb{R}[[\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3,\xi_{1,2},\xi_{1,3}]]\gamma_{1,2}\gamma_{1,3}\oplus \\
&\mathbb{R}[[\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3,\xi_{1,2},\xi_{1,3},\xi_{2,3}]]\gamma_{2,3} \oplus \mathbb{R}[[\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3,\xi_{1,2},\xi_{1,3},\xi_{2,3}]]\xi_{2,3}\oplus \\
&\mathbb{R}[[\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3,\xi_{1,2},\xi_{1,3},\xi_{2,3}]]\gamma_{1,2}\gamma_{2,3} \oplus \mathbb{R}[[\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3,\xi_{1,2},\xi_{1,3},\xi_{2,3}]]\gamma_{1,2}\xi_{2,3}\oplus \\
&\mathbb{R}[[\alpha_3,\beta_1,\beta_2,\beta_3,\xi_{1,2},\xi_{1,3},\xi_{2,3}]](\gamma_{1,2},\alpha_3)^{(1)} \oplus \mathbb{R}[[\alpha_3,\beta_1,\beta_2,\beta_3,\xi_{1,2},\xi_{1,3},\xi_{2,3}]](\gamma_{1,2},\alpha_3)^{(2)}.
\end{aligned}
$$

as found in [4].

The basis elements for $ker\ \mathcal{X}_{333}$ are e_3, e_6 and e_9. Therefore we need to compute the box product of the invariants ring $ker\ \mathcal{X}_{333}$ with $\mathbb{R}e_3 \oplus \mathbb{R}e_6 \oplus \mathbb{R}e_9$. Thus $ker\ \mathsf{X}_{333} = ker\mathcal{X}_{333} \boxtimes [\mathbb{R}e_3 \oplus \mathbb{R}e_6 \oplus \mathbb{R}e_9]$. Let $\mathcal{R} = \mathbb{R}[[\beta_1,\beta_2,\beta_3,\xi_{1,2},\xi_{1,3}]]$, then
$ker\ \mathsf{X}_{333} = [\mathcal{R}[[\alpha_1,\alpha_2,\alpha_3]]\oplus\mathcal{R}[[\alpha_1,\alpha_2,\alpha_3]]\gamma_{1,2}\oplus\mathcal{R}[[\alpha_1,\alpha_2,\alpha_3]]\gamma_{1,3}\oplus\mathcal{R}[[\alpha_1,\alpha_2,\alpha_3]]\gamma_{1,2}\gamma_{1,3}\oplus$
$\mathcal{R}[[\alpha_2,\alpha_3,\xi_{2,3}]]\gamma_{2,3}\oplus\mathcal{R}[[\alpha_2,\alpha_3,\xi_{2,3}]]\xi_{2,3}\oplus\mathcal{R}[[\alpha_2,\alpha_3,\xi_{2,3}]]\gamma_{1,2}\gamma_{2,3}\oplus\mathcal{R}[[\alpha_2,\alpha_3,\xi_{2,3}]]\gamma_{1,2}\xi_{2,3}\oplus$
$\mathcal{R}[[\alpha_3,\xi_{2,3}]](\alpha_3,\gamma_{1,2})^{(1)} \oplus \mathcal{R}[[\alpha_3,\xi_{2,3}]](\alpha_3,\gamma_{1,2})^{(2)}] \boxtimes [\mathbb{R}e_3 \oplus \mathbb{R}e_6 \oplus \mathbb{R}e_9]$.

There are three cases to consider. Computing, simplifying and recombining the three cases we obtain the normal form as:
$ker\ \mathsf{X}_{333} =$

$$
\begin{aligned}
&\mathcal{R}[[\alpha_1,\alpha_2,\alpha_3]]e_{3n} \oplus \mathcal{R}[[\alpha_3]](\alpha_3,e_{3n})^{(i)} \oplus \mathcal{R}[[\alpha_2,\alpha_3]](\alpha_2,e_{3n})^{(i)} \oplus \mathcal{R}[[\alpha_1,\alpha_2,\alpha_3]](\alpha_1,e_{3n})^{(i)} \\
&\oplus\mathcal{R}[[\alpha_1,\alpha_2,\alpha_3]]\gamma_{1,2}e_{3n} \oplus \mathcal{R}[[\alpha_1,\alpha_2,\alpha_3]](\alpha_1\gamma_{1,2},e_{3n})^{(i)} \oplus \mathcal{R}[[\alpha_2,\alpha_3]](\alpha_2\gamma_{1,2},e_{3n})^{(i)} \\
&\oplus\mathcal{R}[[\alpha_3]](\gamma_{1,2},e_{3n})^{(i)} \oplus \mathcal{R}[[\alpha_1,\alpha_2,\alpha_3]]\gamma_{1,3}e_{3n} \oplus \mathcal{R}[[\alpha_1,\alpha_2,\alpha_3]](\alpha_1\gamma_{1,3},e_{3n})^{(i)} \\
&\oplus\mathcal{R}[[\alpha_2,\alpha_3]](\alpha_2\gamma_{1,3},e_{3n})^{(i)} \oplus \mathcal{R}[[\alpha_3]](\gamma_{1,3},e_{3n})^{(i)} \oplus \mathcal{R}[[\alpha_1,\alpha_2,\alpha_3]]\gamma_{1,2}\gamma_{1,3}e_{3n} \\
&\oplus\mathcal{R}[[\alpha_1,\alpha_2,\alpha_3]](\alpha_1\gamma_{1,2}\gamma_{1,3},e_{3n})^{(i)} \oplus \mathcal{R}[[\alpha_2,\alpha_3]](\alpha_2\gamma_{1,2}\gamma_{1,3},e_{3n})^{(i)} \oplus \mathcal{R}[[\alpha_3]](\gamma_{1,2}\gamma_{1,3},e_{3n})^{(i)} \\
&\oplus\mathcal{R}[[\alpha_2,\alpha_3,\xi_{2,3}]]\gamma_{2,3}e_{3n} \oplus \mathcal{R}[[\alpha_3,\xi_{2,3}]](\alpha_3\gamma_{2,3},e_{3n})^{(i)} \oplus \mathcal{R}[[\alpha_2,\alpha_3,\xi_{2,3}]](\alpha_2\gamma_{2,3},e_{3n})^{(i)} \\
&\oplus\mathcal{R}[[\alpha_2,\alpha_3,\xi_{2,3}]]\xi_{2,3}e_{3n} \oplus \mathcal{R}[[\alpha_3,\xi_{2,3}]](\alpha_3\xi_{2,3},e_{3n})^{(i)} \oplus \mathcal{R}[[\alpha_2,\alpha_3,\xi_{2,3}]](\alpha_2\xi_{2,3},e_{3n})^{(i)} \\
&\oplus\mathcal{R}[[\alpha_2,\alpha_3,\xi_{2,3}]]\gamma_{1,2}\gamma_{2,3}e_{3n} \oplus \mathcal{R}[[\alpha_3,\xi_{2,3}]](\alpha_3\gamma_{1,2}\gamma_{2,3},e_{3n})^{(i)} \oplus \mathcal{R}[[\alpha_2,\alpha_3,\xi_{2,3}]](\alpha_2\gamma_{1,2}\gamma_{2,3},e_{3n})^{(i)} \\
&\oplus\mathcal{R}[[\alpha_2,\alpha_3,\xi_{2,3}]]\gamma_{1,2}\xi_{2,3}e_{3n} \oplus \mathcal{R}[[\alpha_3,\xi_{2,3}]](\alpha_3\gamma_{1,2}\xi_{2,3},e_{3n})^{(i)} \oplus \mathcal{R}[[\alpha_2,\alpha_3,\xi_{2,3}]](\alpha_2\gamma_{1,2}\xi_{2,3},e_{3n})^{(i)} \\
&\oplus\mathcal{R}(\gamma_{1,2},\alpha_3)^{(1)}e_{3n} \oplus \mathcal{R}[[\alpha_3,\xi_{2,3}]]\Big((\gamma_{1,2},\alpha_3)^{(1)},e_{3n})\Big)^{(i)} \oplus \mathcal{R}(\gamma_{1,2},\alpha_3)^{(2)}e_{3n} \\
&\oplus\mathcal{R}[[\alpha_3,\xi_{2,3}]]\Big((\gamma_{1,2},\alpha_3)^{(2)},e_{3n})\Big)^{(i)}
\end{aligned}
$$

where $i = 1, 2$ and $n = 1, 2, 3$ such that $e_{3(1)} = e_3$, $e_{3(2)} = e_6$ and $e_{3(3)} = e_9$.

In general, from the above examples we conclude that the normal form of a system with linear part $N_{(3)^n}$ is obtained by computing the box product

$$ker\ \mathsf{X}_{(3)^n} = ker\ \mathcal{X}_{(3)^n} \boxtimes (\oplus_{k=1}^{n}\mathbb{R}e_{3k}).$$

The basis of the normal form of $ker\ \mathsf{X}_{(3)^n}$ are transvectants of the form: $(f, e_{3k})^{(i)}$ where f are the standard monomials of the Stanley decomposition of the ring of invariants, $ker\ \mathcal{X}_{(3)^n}$, $i = 0, 1, 2$ and $k = 1, ..., n$.

Bibliography

[1] W.W. Adams and P. Louatsana, *An Introduction to Groebner Bases*, American Mathematical Society, Providence, 1994.

[2] R. Cushman, J.A. Sanders and N. White, Normal form for the (2;n) Nilpotent Vector Field Using Invariant Theory. *Phys. D*, **30**, 399 - 412, 1988.

[3] R. Cushman and J.A. Sanders, A Survey of Invariant Theory Applied to Normal Forms of Vector Fields with Nilpotent Linear Part. In D. Stanton (Ed.), *Invariant Theory and Tableaux,* Springer - Verlag, Nwe York, 1990.

[4] G. Gachigua and D.M. Malonza, Stanley Decomposition of Coupled N_{333} System. *Proceedings of the first Kenyatta University International Mathematics Conference*, 6 - 10 June 2011, Kenyatta University, Nairobi, Kenya.

[5] D.M. Malonza, Normal Forms for Coupled Takens-Bogdanov Systems, *Journal of Nonlinear Mathematical Physics*, **11**, 376 - 398, 2004.

[6] J. Murdock, On the Structure of Nilpotent Normal Form Modules.*Journal of Differential Equations* **180**, 198 - 237, 2002.

[7] J. Murdock, *Normal Form and Unfoldings for Local Dynamical Systems*, Springer-Verlag, New York, 2003.

[8] J. Murdock and J.A. Sanders, A New Transvectant Algorithm for Nilpotent Normal Forms, *Journal of Differential Equations*, 238, 234 - 256, 2007.

[9] N.S. Namachchivaya et al., Normal Form for Generalized Hopf Bifurcation with Nonsemisimple 1:1 Resonance, *Z. Angew. Math. Phys*, **45**(2),312 - 335, 1994.

Determination of the Term Structure of CDS Spread by an Antithetic Variate Monte Carlo Algorithm

Walter Mudzimbabwe

School of Business and Management Sciences,
Harare Institute of Technology,
P. O. Box BE 277, Belvedere, Harare, Zimbabwe
Email: mudzmudz@gmail.com

Abstract. We determine the Credit Default Swap spread term structure where the firm's asset price process is driven by a jump-diffusion model. In order to get an efficient implementation of a Monte Carlo algorithm, we propose a antithetic variate as a variance reduction procedure. The antithetic variate is derived from the compound Poisson process in the jump-diffusion process.

1. Introduction

Among the various methods that can be used to determine the insurance premium that the protection buyer of a Credit Default Spread (CDS) pays to the protection seller, for protection against the credit risk for an exposure on assets of the reference entity, is the cash flow method (see, e.g., Hull [4], Hull and White [5]). In this article, we consider an extension of the Merton model where a firm has issued both equity and debt. The total value of the firm assets V_t which comprises equity and debt is modelled as a jump-diffusion model (Zhou [7]):

$$\frac{dV_t}{V_t} = \mu dt + \sigma_v dW_t + (\Pi - 1)dN_t, \tag{1.1}$$

where μ represents expected return on firm's assets; W_t is a Weiner process; N_t is a Poisson process with intensity λ; Π is the jump size with expected value of $\nu + 1$; dW_t, dN_t and Π are assumed to be mutually independent. According to [7], modelling firm asset values by a geometric Brownian motion does not produce the term structure of CDS spreads observed from market data. As an extension [7], has incorporated jumps in the asset process.

We characterise default of the firm by a first-passage model. The firm defaults when the value of its assets falls below a certain fixed barrier.

The rest of the article is as follows in Section 2, we consider the valuation method and we present the valuation formula for the CDS spread. We determine the risk neutral risk dynamics in section 3. In section 4, we discuss a basic Monte

Carlo technique and how it can be made efficient by variance reduction techniques. We give numeric results based of implementing the algorithm and conclude in section 6.

2. Valuation of a CDS

We consider the discounted cash flow method to determine the periodic premium payments that the protection buyer of a CDS makes to the protection seller. Following [5], the CDS spread is found by discounting the future payments at the risk free rate r. We assume that there is a risk-neutral probability Q equivalent to the real world probability. Our approach is a risk-neutral pricing methodology.

The first passage-time model entails that a default event happens at τ which is such that

$$\tau = \min\{t : V_t \leq K, t \geq 0\}, \tag{2.1}$$

where K is the fixed barrier. In general, the time to default τ has no tractable distribution (see Giesecke et al. [1]).

The protection buyer makes periodic payments at times t_i for $i = 1, 2, \cdots, d$ until maturity ($t = T$) or default, whichever happens first. There are two components of the buyer payments. At time t_i the buyer pays $Ns(t_i - t_{i-1})\mathbb{1}_{\{\tau > t_i\}}$, where s is the CDS spread of a newly issued CDS and N is the notional value of assets protected by the CDS and $\mathbb{1}_{\{.\}}$ is the indicator function. The other part is the accrual payment of $Ns(\tau - t_i)\mathbb{1}_{\{t_i < \tau < t_{i+1}\}}$. The present value of the total payments to the seller is

$$\mathbb{E}_Q\left[\sum_{i=1}^{d}\left(e^{-rt_i}Ns\Delta t_i\mathbb{1}_{\{\tau>t_i\}} + e^{-r\tau}Ns(\tau - t_i)\mathbb{1}_{\{t_i<\tau<t_{i+1}\}}\right)\right],$$

where $\Delta t_i = t_i - t_{i-1}$. Assuming that the CDS is cash settled, if the firm defaults, the protection seller makes a default payment of $N(1 - RR)$, where RR is there recovery rate which we will assume to be deterministic and is known at inception $t_0 = 0$ of contract. The present value is given by $\mathbb{E}_Q[e^{-r\tau}N(1 - RR)\mathbb{1}_{\{\tau \leq T\}}]$.

To the buyer the value of the CDS is the surplus of payments to the receipts. The CDS spread is the value s that makes the surplus zero and its value is given by:

$$s = \frac{\mathbb{E}_Q[e^{-r\tau}(1 - RR)\mathbb{1}_{\{\tau \leq T\}}]}{\mathbb{E}_Q\left[\sum_{i=1}^{d}\left(e^{-rt_i}\Delta t_i\mathbb{1}_{\{\tau>t_i\}} + e^{-r\tau}(\tau - t_i)\mathbb{1}_{\{t_i<\tau<t_{i+1}\}}.\right)\right]} \tag{2.2}$$

The formula (2.2) cannot be evaluated by analytic means because the random variable τ does not have a closed form distribution. In section 4, we develop a Monte Carlo algorithm to approximate s.

3. Risk-neutral dynamics

Since our approach will be a risk-neutral valuation, we determine the corresponding firm dynamics under Q. Assuming a finite number of jumps and a jump sizes, we can apply a general Ito formula (Hanson [3]) for semi-martingale processes such as (1.1). In this case the solution for (1.1) is

$$d\ln(V_t) = (\mu - \sigma_v^2/2)dt + \sigma_v dW_t + \ln(\Pi)dN_t, \tag{3.1}$$

which upon being integrated yields

$$V_t = V_0 \exp((\mu - \sigma_v^2/2)t + \sigma_v W_t + \ln(\Pi)N_t). \tag{3.2}$$

In a small interval $\Delta t_i = t_i - t_{i-1}$, the Poisson process N_t is such that the number of jumps is 0 or 1. Using a uniform spacing, i.e., $\Delta t_i = \Delta t = T/n$, where n is the number of partitions of [0,T]. Similarly to [7], we assume that the jump size is lognormally distributed, i.e., $\ln(\Pi) \sim N(\mu_\pi, \sigma_\pi^2)$, where $N(,)$ denotes the normal distribution. From the properties of the lognormal distribution, $\nu = \mathbb{E}(\Pi) - 1 = \exp(\mu_\pi + \sigma_\pi^2/2) - 1$. Letting $X_t = \ln(V_t)$, we may write

$$\Delta X_{t_i} = (\mu - \sigma_v^2/2)\Delta t + \sigma_v \Delta W_{t_i} + \sum_{j=1}^{dN_{t_i}} \pi_{ij}. \tag{3.3}$$

From which it follows that

$$X_{t_i} - X_{t_{i-1}} = x_i + y_i \pi_i, \tag{3.4}$$

where

$$x_i \sim N((\mu - \sigma_v^2/2)T/n, \sigma_v^2 T/n), \quad \pi_i \sim N(\mu_\pi, \sigma_\pi^2) \tag{3.5}$$

$$y_i = \begin{cases} 0, & \text{with probability } 1 - \lambda T/n \\ 1, & \text{with probability } \lambda T/n, \end{cases} \tag{3.6}$$

see [7] for this decomposition.

In the risk-neutral world, $\mathbb{E}_Q(V_t) = V_0 e^{rt}$. One can determine the expected value of V_t from (3.2) by conditioning on $N(t)$. The same idea has been used by Merton [6] in option pricing where the underlying price follows a jump diffusion process. In this case $\mathbb{E}(V_t) = V_0 e^{(\mu+\lambda\nu)t}$. From which it follows that $\mu = r - \lambda\nu$. Finally the risk-neutral dynamics of the firm's asset value is given by

$$\begin{aligned} \frac{dV_t}{V_t} &= (r - \lambda\nu)dt + \sigma_v dW_t + (\Pi - 1)dN_t, \\ &= r dt + \sigma_v dW_t + (\Pi - (\nu + 1))dN_t + d\tilde{N}_t, \end{aligned}$$

where $\tilde{N}_t = N_t - \lambda t$ is a compensated Poisson process and the corresponding solution may be written as:

$$d\ln(V_t) = (r - \sigma_v^2/2 - \lambda\nu)dt + \sigma_v dW_t + \ln(\Pi)dN_t, \tag{3.7}$$

4. Monte Carlo Algorithm

In an elementary Monte Carlo algorithm, in order for the standard deviation to be reduced by ten the number of Monte Carlo loops has to be increased by a hundred fold. However Monte Carlo algorithms can be derived with lower variances. We will use an antithetic procedure for variance reduction.

To illustrate the antithetic variate method, from (3.5), may write $x_i = (\mu - \sigma_v^2/2)T/n + \sigma_v\sqrt{T/n}Z_1$ and $\pi_i = \mu_\pi + \sigma_\pi Z_2$, where Z_1 and Z_1 are standard normal variates. The two variables x_i and π_i are usually referred to as the *thetic* variables. To obtain the antithetic variates, we let $x_i^{(a)} = (\mu - \sigma_v^2/2)T/n - \sigma_v\sqrt{T/n}Z_1$ and $\pi_i^{(a)} = \mu_\pi - \sigma_\pi Z_2$. The compound Poisson process $\sum_{j=1}^{dN_{t_i}} \pi_{ij}$ in (3.3), involves a summation normal variables and for this reason derivation of antithetic variates is straight forward. See [3], for instance. In an implementation an algorithm based on as antithetic technique, Z_1 and Z_2 are a prior draw and are reused in leu of being generated afresh. For more details on the antithetic control variate method in finance, see e.g., Glasserman [2].

For each Monte Carlo simulations we determine whether or not there has been default based on the two paths X_t and $X_t^{(a)}$ hitting the fixed barrier $\ln(K)$. Finally, the Monte Carlo algorithm with the antithetic variate to determine the CDS spread is as follows:

1. Subdivide $[0, T]$ into n equal parts. Denote $t_i = T \cdot i/n$.

2. Determine protection payment times t_k such that $t_k = t_{k-1} + \delta t$ for $k = 1, 2, \cdots, N_k$.

3. Perform Monte Carlo simulations for $j = 1, 2, \cdots, M$. For each j perform the following:

   (a) Generate the mutually independent random vector (x_i, π_i, y_i) according to (3.5) and (3.6)

   (b) i. Let $X_{t_0} = \ln(V_0)$ and calculate X_{t_i} according to (3.4).
   ii. Let $X_{t_0}^{(a)} = \ln(V_0)$, $X_{t_i}^{(a)} = x^{(a)} + y_i\pi_i^{(a)}$

   (c) i. Find the smallest integer $i \leq n$: $X_{t_i} \leq \ln(K)$. If such i exists, set $\tau = i$.
   ii. Find the smallest integer $i \leq n$: $X_{t_i}^{(a)} \leq \ln(K)$. If such $i^{(a)}$ exists, set $\tau^{(a)} = i^{(a)}$.

   (d) i. Calculate the discounted protection payment PP_j $\forall k$: $\tau \geq t_k$ when i exists, otherwise $PP_j = 0$.
   ii. Calculate the discounted protection payment $PP_j^{(a)}$ $\forall k$: $\tau^{(a)} \geq t_k$ when $i^{(a)}$ exists, otherwise $PP^{(a)} = 0$.

   (e) i. Calculate the discounted accrual payment $AP_j = (\tau - t_k)e^{-r\tau}$ when i exists, otherwise $AP_j = 0$.

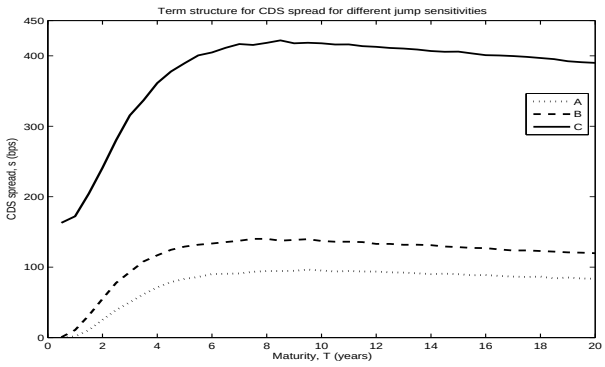


Figure 5.1: Jump parameters: C: $\lambda = 0.1$, $\sigma_\pi = 1.0$, B: $\lambda = 1.0$, $\sigma_\pi = 0.1$ and A: $\lambda = 10.0$, $\sigma_\pi = 0.01$

ii. Calculate the discounted accrual payment $AP_j^{(a)} = (\tau^{(a)} - t_k)e^{-r\tau^{(a)}}$ when $i^{(a)}$ exists, otherwise $AP_j^{(a)} = 0$.

(f) i. Calculate the discounted default payment DP_j when i exists, otherwise $DP_j = 0$.

ii. Calculate the discounted default payment $DP_j^{(a)}$ when $i^{(a)}$ exists, otherwise $DP_j^{(a)} = 0$.

(g) Set $aDP_j = 0.5(DP_j + DP_j^{(a)})$, $aAP_j = 0.5(AP_j + aAP_j^{(a)})$ and $aPP_j = 0.5(PP_j + PP_j^{(a)})$

4. Calculate

$$s = \frac{(1/M)\sum_{j=1}^{M} aDP_j}{(1/M)\left(\sum_{j=1}^{M} aPP_j + \sum_{j=1}^{M} aAP_j\right)}.$$

5. Numerical Results

In this section we present our results based on the Monte Carlo method with antithetic variates to determine CDS spread. Figure 5.1 shows the CDS spread for maturities ranging from 0.5 to 20 years. The graph has been plotted for different jump parameters. The rest of the parameters are as follows: $V_0 = 100.0$, $r = 0.05$, $\sigma_v = 0.2$, $\mu = 0.0$, we also assume a constant recovery rate RR of 0.4 and quarterly insurance buyer payments. The graph shows that the CDS spread decrease as with maturity. This is mainly due the positive drift of the asset price process. An increase in jump intensity λ and σ_π increases the CDS spread since the likelihood of default is higher. However to get term structure spreads consistent with market data we increase λ and decrease σ_π.

T (years)	λ	σ_π	s (bps)
0.5	0.1	1.0	157.10
	1.0	0.1	0.36
1.0	0.1	1.0	177.15
	1.0	0.1	9.38
7.0	0.1	1.0	416.11
	1.0	0.1	137.22
18.0	0.1	1.0	393.51
	1.0	0.1	122.37

Table 5.1: CDS spreads for different maturities and different jump sensitivities

From Table 5.1, we confirm that generally the CDS spread increases with maturity but for longer T it decreases. If a firm survives the first few years it becomes well adapted to its industry and hence the likelihood of default decreases.

6. Conclusion

In this paper we derived a Monte Carlo method which is based on the antithetic technique of variance reduction. The algorithm is used to determine the CDS spread for a firm that has issued debt whose asset price dynamics follow a jump-diffusion process. The results based on the model developed confirm market observations.

Bibliography

[1] K. Giesecke, B. Kim, and S. Zhu, Monte Carlo Algorithms for Default Timing Problems, *Management Science*, 1–15, 2011.

[2] P. Glasserman, *Monte Carlo Methods in Financial Engineering (Stochastic Modelling and Applied Probability)*, Springer, 2003.

[3] F. B. Hanson, *Applied Stochastic Processes and Control for Jump-diffusion: Modeling, Analysis and Computation*, SIAM, 2007.

[4] J. C. Hull, *Options, Futures, and other Derivatives*, Pearson Prentice Hall, 2009.

[5] J. C. Hull and A. White, Valuing Credit Default Swaps 1: No Counterparty Default Risk, *Journal of Derivatives*, **8**, 29 - 40, 2000.

[6] R. C. Merton, Option Pricing when the Underlying Stock Returns are Discontinuous, *J. Financial Economics*, **83**, 125 - 144, 1976.

[7] C. Zhou, The term structure of credit spreads with jump risk, *Journal of Banking & Finance*, **25**, 2015–2040, 2001.

A Transport Optimization Model of Retail Distribution : A Case Study of Zomba Bakery

Elias P. Mwakilama and Levis K. Eneya

Department of Mathematical Sciences, University of Malawi, Chancellor College, Zomba, Malawi.
Emails: mwakilamae@gmail.com, leneya@cc.ac.mw

Abstract. This paper concerns finding optimal routes that minimize the overall cost of transporting confectionery products from the Zomba Bakery by applying mathematical modelling. We used the baseline information collected from the company to estimate variables of interest and to understand the problem. The study was designed as a Multiple Travelling Salesman Problem (m-TSP) with distribution centres as nodes and in-between distances as weighted edges. An optimization problem was then formulated using linear programming and graph theory techniques; and implemented using Prim's algorithm [20]. From the distribution centres, it was possible to find optimal routes or Minimum Spanning Tree (MST) that reduce cost of delivery considerably by 40%. Other than the derived MST, other routes registered higher fuel and operation costs. MST also provided the possibility of using one vehicle to visit all distribution centres thereby reducing the number of vehicles to be used. The study also demonstrated the possibility of obtaining solution to seemingly complex problems using relatively simple mathematical ideas.

Keywords: Transport Optimization, Linear programming, Prim's algorithm, Minimum Spanning Tree

1. Introduction

The optimal route finding problem, e.g. for a fleet of vehicles, has long been studied. The reader is referred to [10, 13, 3, 4, 7] among other references for detailed exposition on the topic. Transport optimisation models or vehicle routing problems (VRP) [11] find their application in diverse fields including food distribution, air plane and train scheduling.

In 1941, Hitchcock [10, p.319] developed the first transportation model. This was later modified by Dantzig [10] who used a simplex method to compute optimal solution of the problem.

Literature to date describes several variants of VRPs. For instance, Muthukannan [14] modelled an urban transit system utility by relating the demand of a node to the aggregate cost of travel, travel time and accessibility. Tavares [24] proposed the application of Geographic Information System (GIS) in 3D route modelling for waste collection taking relief into consideration to minimize fuel consumption. In their findings, the optimization for the lowest fuel consumption yielded 52% savings on fuel, yet travelling a 34% longer distance. The study hence showed the importance of considering simultaneously the relief of the territory and lowest fuel consumption criterion when optimizing vehicle routes.

Several studies on transport models have focused on variants of Travelling Salesman Problem (TSP) [13] with particular constraints, e.g. Multiple Travelling Salesman Problem (m-TSP), Euclidean Travelling Salesman Problem (ETSP) and Multiple Travelling Salesman Problem with Time Window (m-TSPTW) [23, 6, 5, 11]. Most of these studies have utilized comprehensive and special evolutionary algorithms to solve the problems with multi-objective functions, and have hence been prone to uncertainty factors [1, 18, 22, 25]. However they have shown advantages and disadvantages of solution approaches based on each algorithm capabilities. The studies have also suggested that many problems in transportation logistics can be modelled and solved similarly whenever routes can be enumerated. Hence the present study aims at finding the optimal routes for delivering products to markets in a Malawian setup, where application of such relatively simple ideas is scanty.

In this paper we apply a mathematical model that describes distribution of goods and compute an optimal solution using simple algorithms. In [23], Shafie points out that robot (vehicle) counts can be less important when robots (vehicles) and manpower (labour) costs are low. He thus concludes that by considering minimal cost (distance), energy consumption is reduced. As such, in our study, only distance minimization was considered. The study demonstrates the application of mathematics to the industry through use of existing techniques in problem identification and solution method at a locally manufacturing company in Zomba district.

2. Materials and Methods

A sample of data from the company were scrutinized to understand the study characteristics. They included distribution system, travel times, product demands and fuel savings. In addition, information regarding choice of routes and the number of vehicles used when distributing the products were also collected.

At the time of the study, Zomba bakery had 13 active distribution centres with only one source which is the Bakery itself. Since the overall aim was to minimize the total costs of moving the products from the Bakery to all of the 13 centres, we decided that the problem was an m-TSP- a special case of VRP, and hence we designed a Multi-Objective (MO) function (2.1) that depended on five criteria, namely; fuel costs (per litre), time of delivery, distance (in km), labour (number

of personnel involved per trip) and number of vehicles needed for the execution of the entire process. Hence, the following structural MO function was formulated;

$$\sum_{k=1}^{5} Z_k \tag{2.1}$$

where k is the decision criterion. For instance, if k=1, fuel cost criterion, then

$$Z_1 = \sum_{j=1}^{13} C_{1j}^1 X_{1j} = C_{1,1}^1 X_{1,1} + C_{1,2}^1 X_{1,2} \ldots + C_{1,13}^1 X_{1,13}. \tag{2.2}$$

Similarly for $k = 2$ (labour),

$$Z_2 = \sum_{j=1}^{13} C_{1j}^2 X_{1j} = C_{1,1}^2 X_{1,1} + C_{1,2}^2 X_{1,2} \ldots + C_{1,13}^2 X_{1,13} \tag{2.3}$$

$k = 3$ (number of vehicles) yields

$$Z_3 = \sum_{j=1}^{13} C_{1j}^3 X_{1j} = C_{1,1}^3 X_{1,1} + C_{12}^3 X_{1,2} \ldots + C_{1,13}^3 X_{1,13}; \tag{2.4}$$

$k = 4$ (distance),

$$Z_4 = \sum_{j=1}^{13} C_{1j}^4 X_{1j} = C_{1,1}^4 X_{1,1} + C_{1,2}^4 X_{1,2} \ldots + C_{1,13}^4 X_{1,13} \tag{2.5}$$

and $k = 5$ (time of delivery) gives

$$Z_5 = \sum_{j=1}^{13} C_{1j}^5 X_{1j} = C_{1,1}^5 X_{1,1} + C_{1,2}^5 X_{1,2} \ldots + C_{1,13}^5 X_{1,13} \tag{2.6}$$

for all $j = 1, 2, ..., 13$.

Here, $i = 1$ represents the central store, while $j = 1, 2, ..., 13$ represents all distribution centres. C_{1j}^k is the cost incurred to transport one unit of goods from the central store 1, the Bakery, to jth distribution centre with respect to criterion k; and X_{1j} is the amount of goods transported from the Bakery to destination j. Since only distance minimization was considered in this study, our m-TSP involved minimizing Z_4 in equation (2.5):

$$\text{Minimize} \sum_{j=1}^{13} C_{1j}^4 X_{1j} \tag{2.7}$$

subjective to non-negativity constraints $X_{1j} \geq 0$.

m-TSP consists of determining a set of routes for m salesmen who all start from

and turn back to home city (depot). Thus, in general the m-TSP could be defined as follows; Given a set of nodes, let there be m salesmen located at a single depot node. The remaining nodes (cities) that are to be visited are called intermediate nodes [5]. Then, the problem consists of finding tours for all m salesmen, who all start and end at the depot, such that each intermediate node is visited exactly once and the total cost of visiting all nodes is minimised. In our case, the cost metric was defined in terms of distance.

The first data category on the distance minimization procedure was the route data. This was because the selection of freight nodes and route segments, as well as their distances, was a major factor in minimizing the costs of transporting goods from the Bakery to a destination. Data relevant to the master route list and node definitions available in the map were sourced from the National Spatial Data Centre through google earth software. Table 2.1 displays this information, where the abbreviations stand for the centres: Zomba Bakery, Malosa, Changalume, Thondwe, Namadzi, Nasawa, Mayaka, Jali, Turn-off Centre, Kachulu, Mpasakamwa, Migowi and Phalombe, respectively.

Table 2.1: Distribution centres, their codes and minimum distance from Zomba Bakery

Abbreviation	Za	Ma	Ce	Th	Nz	Na	Mk	Ji	Tc	Ku	Mp	Mg	Ph
Code	1	2	3	4	5	6	7	8	9	10	11	12	13
Distance (km)	0	24	15	16	27	37	35	21	19	31	26	61	67

Recognizing that the cost system is inherently complex, sensitive to market changes, that it varies from company to company and that the breakdown of costs is rarely made available, the study only utilised values as in (Table 2.2). Any user of this model is free to modify these costs to suit his or her problem characteristics.

Table 2.2: Estimates of Fuel Usage for some distances. Source: Zomba Bakery

Retail Centre	Chiponde	Kachulu	Phalombe	Zomba City
Fuel usage (litres)	50	30	45	35
Distance (kilometres)	156	31	67	20
Consumption(litres per 100 km)	32.05	96.77	67.16	175
Wear and tear costs	K 84,240	K 16,740	K 36,180	K 10,800

Distance Minimization: The Spanning Tree

Here, we illustrate how the distance-dependencies of the transportation network could be modelled. Then we considered Kruskal's and Prim's algorithms [16, 20] to solve the problem. All routes and distribution points mimic an undirected connected weighted graph with distribution points as vertices while distances between any two points acted as weighted edges of the graph. Because our route system involved a lot of edges, we employed the Prim's algorithm to determine the total minimum distance that vehicles

could travel by visiting each distribution point exactly once.

Using information in Table 2.1, we came up with the graph (route map) depicting the distribution of the products from Zomba Bakery to the distribution centres as shown in Figure 2.1, where numbers over edges indicate distances bewteen the two connecetd nodes identified by their code numbers.

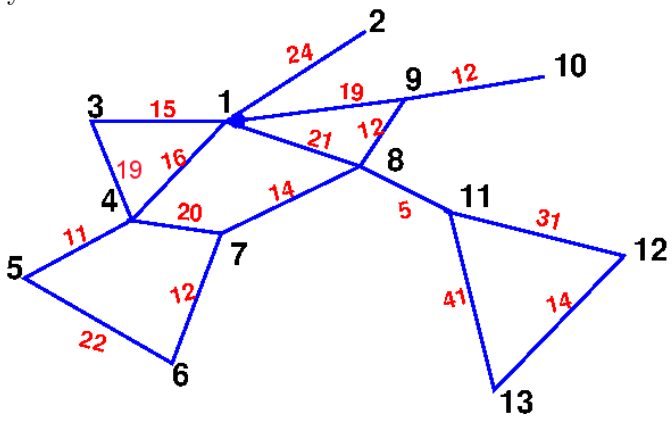


Figure 2.1: Route Map

The undirected graph in Figure 2.1 was then coded in Matlab giving the sparse matrix in Table 2.3 in which the *(i,j)th* entry in the matrix is the weight (distance) of the edge (route) between nodes (distribution centres) i and j, where a weight of zero means that no edge existed between the nodes.

Table 2.3: Distance Matrix

	Za	Ma	Ce	Th	Nz	Na	Mk	Ji	Tc	Ku	Mp	Mg	Ph
Za	0	24	15	16	0	0	0	21	19	0	0	0	0
Ma	24	0	0	0	0	0	0	0	0	0	0	0	0
Ce	15	0	0	19	0	0	0	0	0	0	0	0	0
Th	16	0	19	0	11	0	20	0	0	0	0	0	0
Nz	0	0	0	11	0	22	0	0	0	0	0	0	0
Na	0	0	0	0	22	0	12	0	0	0	0	0	0
Mk	0	0	0	20	0	12	0	14	0	0	0	0	0
Ji	21	0	0	0	0	0	14	0	12	0	5	0	0
Tc	19	0	0	0	0	0	0	12	0	12	0	0	0
Ku	0	0	0	0	0	0	0	0	12	0	0	0	0
Mp	0	0	0	0	0	0	0	5	0	0	0	31	41
Mg	0	0	0	0	0	0	0	0	0	0	31	0	14
Ph	0	0	0	0	0	0	0	0	0	0	41	14	0

3. Results and Discussion

Appendix I gives the MATLAB code implementing Prim's algorithm for the problem; and Table 3.1 gives results (Minimum Spanning Tree (MST) upon executing the program.

Table 3.1: Results of Minimum Spanning Tree Matrix

	ZB	Ma	Ce	Th	Nz	Na	Mk	Ji	Tc	Ku	Mp	Mg	Ph
ZA	0	24	15	16	0	0	0	0	19	0	0	0	0
Ma	0	0	0	0	0	0	0	0	0	0	0	0	0
Ce	0	0	0	0	0	0	0	0	0	0	0	0	0
Th	16	0	0	0	11	0	0	0	0	0	0	0	0
Nz	0	0	0	0	0	0	0	0	0	0	0	0	0
Na	0	0	0	0	0	0	0	0	0	0	0	0	0
Mk	0	0	0	0	0	12	0	0	0	0	0	0	0
Ji	0	0	0	0	0	0	14	0	0	0	5	0	0
Tc	0	0	0	0	0	0	0	12	0	12	0	0	0
Ku	0	0	0	0	0	0	0	0	0	0	0	0	0
Mp	0	0	0	0	0	0	0	0	0	0	0	31	0
Mg	0	0	0	0	0	0	0	0	0	0	0	0	14
Ph	0	0	0	0	0	0	0	0	0	0	0	0	0

From the the matrix in Table 3.1, non-zero entries are edges that connect two adjacent nodes, giving the Minimum spanning Tree (MST) or optimal routes shown in Figure 3.1.

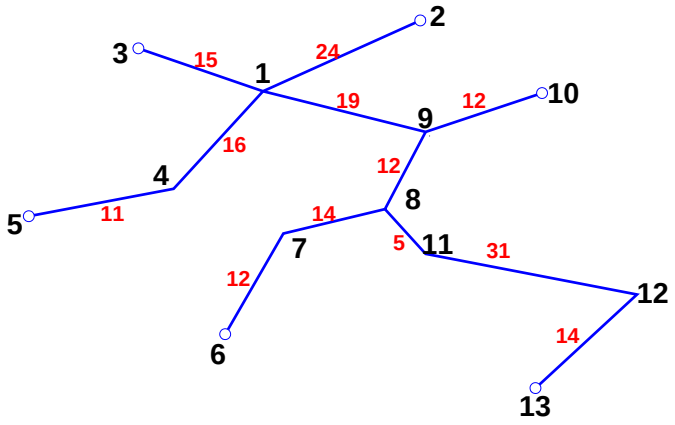


Figure 3.1: Minimum Spanning Tree

It can be seen from the Minimum Spanning Tree which optimal routes vehicles ought to take when delivering products to the outlets. Total distance, time taken and fuel used to visit all the centres using the above computed MST were then compared with similar values that would be obtained if all the centres were visited using the current random routes that the company uses. The results are presented in Table 3.2.

Table 3.2: Comparison Between Optimal and Other Routes

Route	I	II	III	IV	**MST**
Distance(Km)	308	210	195	227	**185**
Fuel usage(litres)	51	35	32	39	**30**
Fuel costs	K 23,100	K 15,750	K 14,550	K 17,025	**K13,875**
Total Time (Hr)	6.16	4.2	3.9	4.54	**3.7**

Using the computed optimal routes to visit each of the 13 centres, one finds that the total fuel cost is K 13,875. Table 3.2 shows that the longer the distance vehicles could take in delivering goods to the centres, the higher the fuel costs, especially when the routes are repeated.

Furthermore, from Table 3.2, one finds that if the proposed optimised routes were used by the company, both distance and time costs would decrease by 40% on average. In addition, the MST indicates that Management of Zomba Bakery would not need more vehicles than they have in order to visit each distribution centre because even one vehicle could visit all the centres at a minimum cost.

Bibliography

[1] M.I. Andreica et al., Algorithmic solutions to some Transportation Optimization Problems with Applications in the Metallurgical Industry, *Metalurgia International Journal*, **14**(2), 46-53,2008.

[2] C.H. Arthur et al., Promoting Undergraduate Research in Mathematics, *Industrial Mathematics and Statistics Research for Undergraduates at WPI*, November,21, 2006.

[3] O. Apaydim et al., Route Optimization for Solid Waste Collection : Trabzon (Turkey) Case Study, *Global NEST journal*, **9**(1), 6-11, 2007.

[4] O. Braysy et al., The Potential of Optimization in Commercial routing problem;case studies from Finland, *Journal of Transport Geography*, 2009.

[5] T. Bektas, The Multiple Travelling Salesman Problem: An Overview of Formulations and Solution Procedures, *The International Journal of Management Science*, 2006.

[6] M. De berg et al., The Travelling Salesman Problem Under Squared Euclidean Distances, *Symposium Theoretical Aspects of Computer Science*, Nancy, France, 239-250, 2010.

[7] T.C Detofeno et al., Optimizing Routes for Collection of Urban Solid Waste: A case Study for the City of Joinville, State of Santa Catarina, *Journal of Industrial Engineering* **2**(1), 124-136, 2010.

[8] Edwards, *A guide to Mathematical Modelling*,McMillan Education LTD, London, 1989.

[9] M. J. Fyer, *An Introduction to Linear Programming and Matrix Game Theory*, Edward Arnold, London, 1978.

[10] S.I. Gass,*Linear Programming-Methods and Applications*, McGraw-Hill Book Company (5th edition), New York, 1985.

[11] Gendreau et al., Vehicle Routing Problem with Time Windows, Part 1: Route Construction and Local Search Algorithm, *Transportation Science Journal*, 2005.

[12] M.M. Hoque et al, Optimization of Retail Distribution Network : A Case Study, *Proceedings of the International Conference on Mechanical Engineering*, Dhaka Bangladesh, 2009.

[13] G. Laporte, The Travelling Salesman Problem: An Overview of Exact and Approximate Algorithms, *European Journal of Operational Research*, 1992.

[14] M. Muthukannan et al., Modelling for Optimization of Urban Transit System Utility: A Case Study, *APRN Journal of Engineering and Applied Sciences*, 2008.

[15] S. Munzir et al., Liner Programming for Parking Slot Optimisation: A Case Study at JI.T. Panglima, *Proceedings of the 6th IMT-GT Conference on Mathematics, Statistics and its Application, ICMSA 2010*, University of Tunku Abdul Rahman, Malaysia, 462-472.

[16] E. Neuman, *Linear Programming with MATLAB*, Department of Mathematics, Southern Illinois University at Carbondale, 2010.

[17] T.C. Ogwueleka et al., Route Optimization for Solid Waste Collection: Onitsha (Nigeria) Case Study, *Journal of Applied Sciences and Environmental Management*, **13**(2), 37-40, 2009.

[18] E. Ozcoylan, A Decision Support System to Compare the Transportation Models in Logistics, *International Journal of Lean Thinking*, 2010.

[19] W.J. Palm, *Introduction to MATLAB for Engineers*, McGraw-Hill Companies, 1998.

[20] Rakha et al., *Introduction to MATLAB-Optimization(Linear Programming) - Computer Applications in Civil Engineering*, Spring 2000, McGraw-Hill Companies, 1998.

[21] T. Roukainen et al., Optimization of Deliveries from Distribution Points, 2007.

[22] Sanchez - Rodrigues et al., Measuring the Time and Distance Impact of Transport Uncertainty: A FMCG Case Study.

[23] A.A. Shafie, Optimal Tour Constructions for Multiple Mobile Robots, *Journal of Engineering Science and Technology*, 2011.

[24] G. Tavares et al., A Case Study of Fuel Savings through Optimization of MSW Transportation Routes,*An International Journal of Management of Environmental Quality*, 2008.

[25] H. Tang et al., Inventory - Transportation Integrated Optimization problem:A model of product Oil logistics, *Internal Journal of Non-linear Science*, 2009.

[26] G.H. Tzeng et al., Fuzzy Bi-criteria Multi - Index Transportation Problems for Coal Allocation Planning of Taipower, *European Journal of Operational Research*, 1996.

[27] Z. Yang et al., Optimizing the Scale and Spatial Location of City Logistics Terminals, *Journal of the Eastern Asian Society for Transportation Studies*, **6**, 2931-2937, 2005.

Orthogonality of Elementary Operators in Normed Spaces and their Applications

N.B. Okello[1] & J.O. Agure[2]

Bondo University College, Box 210-40601, Bondo, Kenya.
School of Mathematics, Applied Statistics and Actuarial Science,
Maseno University, Box 333, Maseno, Kenya.

Abstract. We present new notions and aspects of orthogonality of elementary operators in normed spaces. Characterizations, generalizations and applications of orthogonality are also considered. In particular, results on orthogonality of the range and the kernel of elementary operators and the operators inducing them in norm-attainable classes are established.

Keywords: orthogonality, normed space, elementary operator, norm-attainable classes
Classification: Primary: 47B47; Secondary: 47A30.

1. Introduction

Orthogonality in normed spaces is a concept that has been analyzed for quite a period of time. Benitez [4] described several types of orthogonality which have been studied in real normed spaces namely: Robert's orthogonality, Birkhoff's orthogonality, Orthogonality in the sense of James, Isoceles, Pythagoras, Carlsson, Diminnie, Area among others. Some of these orthogonalities are described as follows. For $x \in \mathcal{M}$ and $y \in \mathcal{N}$ where $\mathcal{M}$ and $\mathcal{N}$ are subspaces of E which is a normed linear space, we have,

(i) Roberts: $\|x - \lambda y\| = \|x + \lambda y\|,\ \forall, \lambda \in \mathbb{R}$.

(ii) Birkhoff: $\|x + y\| \geq \|y\|$.

(iii) Isosceles: $\|x - y\| = \|x + y\|$.

(iv) Pythagorean: $\|x - y\|^2 = \|x\|^2 + \|y\|^2$.

(v) a-Pythagorean: $\|x - ay\|^2 = \|x\|^2 + a^2\|y\|^2,\ a \neq 0$.

(vi) Diminnie: $\sup\{f(x)g(y) - f(y)g(x) : f,\ g \in S'\} = \|x\|\|y\|$ where S' denotes the unit sphere of the topological dual of E.

(vii) Area: $\|x\|\|y\| = 0$ or they are linearly independent and such that $x, -x, y, -y$ divide the unit ball of their own plane (identified by $\mathbb{R}^2$) in four equal areas.

[1]bnyaare@yahoo.com
This author's appreciations go to National Council for Science and Technology (NCST) of Kenya and DAAD- Germany for financial support to this research.
[2]johnagure@yahoo.com

Consider a normed space $\mathcal{A}$ and let $T_{A,B} : \mathcal{A} \to \mathcal{A}$. An elementary operator T has the following representation:

$$T(X) = \sum_{i=1}^{n} A_i X B_i, \ \forall \ X \in \mathcal{A},$$

where $A_i, \ B_i$ are fixed in $\mathcal{A}$. Let $\mathcal{A} = B(H)$. For $A, \ B \in B(H)$ we define the particular elementary operators:

(i) the left multiplication operator $L_A : B(H) \to B(H)$ by:
$L_A(X) = AX, \ \forall \ X \in B(H)$.

(ii) the right multiplication operator $R_B : B(H) \to B(H)$ by :
$R_B(X) = XB, \ \forall \ X \in B(H)$.

(iii) the generalized derivation (implemented by $A, \ B$) by:
$\delta_{A,B} = L_A - R_B$.

(iv) the normal derivation (implemented by A) by:
$\delta_{A,A} = L_A - R_A$.

(v) the basic elementary operator(implemented by $A, \ B$) by:
$M_{A, \, B}(X) = AXB, \ \forall \ X \in B(H)$.

(vi) the Jordan elementary operator(implemented by $A, \ B$) by:
$\mathcal{U}_{A, \, B}(X) = AXB + BXA, \ \forall \ X \in B(H)$.

Regarding orthogonality involving elementary operators, Anderson[1] established the orthogonality of the range and kernel of normal derivations. Others who have also worked on orthogonality include: Kittaneh [22], Mecheri [30] among others. For details see [1-28, 32-34]. We shall investigate the orthogonality of the range and the kernel of several types of important elementary operators.
Some of these orthogonalities are described as follows. For $x \in \mathcal{M}$ and $y \in \mathcal{N}$ where $\mathcal{M}$ and $\mathcal{N}$ are subspaces of E which is a normed linear space, we have,

(i) Roberts: $\|x - \lambda y\| = \|x + \lambda y\|, \ \forall, \lambda \in \mathbb{R}$.

(ii) Birkhoff: $\|x + y\| \geq \|y\|$.

(iii) Isosceles: $\|x - y\| = \|x + y\|$.

(iv) Pythagorean: $\|x - y\|^2 = \|x\|^2 + \|y\|^2$.

(v) a-Pythagorean: $\|x - ay\|^2 = \|x\|^2 + a^2\|y\|^2, \ a \neq 0$.

We extend this study to the $\mathcal{NA}(\mathcal{H})$ and $\mathcal{NA}(\mathcal{H})$-classes.

2. Orthogonality in $\mathcal{NA}(\mathcal{H})$-classes

Definition 1. *Let $T : \mathcal{NA}(\mathcal{H})$ be the class of all norm-attainable operators an a Hilbert space H and let $T : \mathcal{NA}(\mathcal{H}) \to \mathcal{NA}(\mathcal{H})$ be defined by*

$$T(X) = \sum_{i=1}^{n} A_i X B_i, \ \forall \ X \in \mathcal{NA}(\mathcal{H}),$$

where A_i, B_i are fixed in $\mathcal{NA}(\mathcal{H})$. We define the range of T by

$$RanT = \{Y \in \mathcal{NA}(\mathcal{H}) : Y = T(X),\ \forall\ X \in \mathcal{NA}(\mathcal{H})\},$$

and the Kernel of T by

$$KerT = \{X \in \mathcal{NA}(\mathcal{H}) : T(X) = 0,\ \forall\ X \in \mathcal{NA}(\mathcal{H})\}.$$

It is known [30] that for any of the examples of the elementary operators defined in Section 1 (inner derivation, generalized derivation, basic elementary operator, Jordan elementary operator), the following implications hold for a general bounded linear operator T on a normed linear space W , i.e.

$$\begin{aligned} RanT \perp KerT &\Rightarrow \overline{RanT} \cap KerT = \{0\} \\ &\Rightarrow RanT \cap KerT = \{0\}. \end{aligned}$$

Here $\overline{RanT}$ denotes the closure of the range of T and $KerT$ denotes the kernel of T and $RanT \perp KerT$ means $RanT$ is orthogonal to the Kernel of T in the sense of Birkhoff.

Lemma 1. *Let $A, B, C \in \mathcal{NA}(\mathcal{H})$ with $CB = I$ (I is an identity element of $\mathcal{NA}(\mathcal{H})$). Then for a generalized derivation $\delta_{A,B} = AX - XB$ and an elementary operator $\Theta_{A,B}(X) = AXB - X$, $R_B(\overline{Ran\delta_{A,C}} \cap Ker\delta_{A,C}) = \overline{Ran\Theta_{A,B}} \cap Ker\Theta_{A,B}$. Moreover, if $\overline{Ran\delta_{A,C}} \cap Ker\delta_{A,C} = \{0\}$ then $\overline{Ran\Theta_{A,B}} \cap Ker\Theta_{A,B} = \{0\}$.*

Proof. First, we see that if $CB = I$ then $R_B\delta_{A,C} = \Theta_{A,B}$. To see this, $\forall\ X \in \mathcal{NA}(\mathcal{H})$, $R_B\delta_{A,C}(X) = AXB - XCB = AXB - X = \Theta_{A,B}$. Suppose that $P \in R_B(\overline{Ran\delta_{A,C}} \cap Ker\delta_{A,C})$. Now, it is a fact that the uniform norm assigns to real- or complex-valued continuous bounded operator R_B defined on any set $\mathcal{NA}(\mathcal{H})$ the nonnegative number $\|R_B\|_\infty = \sup\{\|R_B(X)\| : X \in \mathcal{NA}(\mathcal{H})\}$. Since $R_B\delta_{A,C} = \Theta_{A,B}$ and R_B is continuous for the uniform norm, then $P \in \overline{Ran\Theta_{A,B}} \cap Ker\Theta_{A,B}$. Conversely, since R_C is continuous for the uniform norm, then by the same argument we prove that if $P \in R_B(\overline{Ran\Theta_{A,B}} \cap Ker\Theta_{A,B})$ then $P \in R_B(\overline{Ran\delta_{A,C}} \cap Ker\delta_{A,C})$. □

Remark 4. *Let $A, B, C \in \mathcal{NA}(\mathcal{H})$ with $CB = I$ (I is an identity element of $\mathcal{NA}(\mathcal{H})$). Then $R_B(\overline{Ran\delta_{A,C}} \cap Ker\delta_{A,C}) = Ran\Theta_{A,B} \cap Ker\Theta_{A,B}$. Indeed, since $Ran\Theta_{A,B} \subseteq \overline{Ran\Theta_{A,B}}$, then by adapting the proof of Lemma above, the equality holds.*

Let $A \in \mathcal{NA}(\mathcal{H})$. The algebraic numerical range $V(A)$ of A is defined by:

$$V(A) = \{f(A) : f \in \mathcal{NA}(\mathcal{H})' \text{ and } \|f\| = f(I) = 1\}$$

where $\mathcal{NA}(\mathcal{H})'$ is the dual space of $\mathcal{NA}(\mathcal{H})$ and I is the identity element in $\mathcal{NA}(\mathcal{H})$.

Definition 2. *If $V(A) \subseteq \mathbb{R}$, then A is called a Hermitian element. Given two Hermitian elements S and R, such that $SR = RS$ then $D = S + Ri$ is called normal.*

Proposition 1. *Let S and R be Hermitian elements. Then $\delta_{S,R}$ is also Hermitian.*

Proof. From [24], it is known that if X is a Banach space then $V(\delta_{S,R}) = V(S) - V(R)$ for all $S, R \in B(X)$. Therefore, $V(\delta_{S,R}) \subseteq V(L_S) - V(L_R) = V(S) - V(R) \subseteq \mathbb{R}$. □

We note that the converse of Proposition 1 is true if $\mathcal{NA}(\mathcal{H}) = B(X)$ where X is a Banach space.

Lemma 2. *If D and E are normal elements in $\mathcal{NA}(\mathcal{H})$ then $\delta_{D,E}$ is also normal.*

Proof. Assume $D = S + Ri$ and $E = T + Ui$ where S, R, T, U are Hermitian elements in $\mathcal{NA}(\mathcal{H})$ such that $SR = RS$ and $TU = UT$. Then $\delta_{D,E} = \delta_{S,T} + i\delta_{R,U}$ with $\delta_{S,T}\delta_{R,U} = \delta_{R,U}\delta_{S,T}$. Since S, R, T, U are Hermitian, then by Proposition 1 $\delta_{R,U}$ and $\delta_{S,T}$ are Hermitian and so is $\delta_{D,E}$. □

Lemma 3. *([24]) Let X be a Banach space and $T \in B(X)$. If T is a normal operator, then $RanT \perp KerT$.*

Theorem 1. *Let $\mathcal{NA}(\mathcal{H})$ be as defined above.If D and E are normal elements in $\mathcal{NA}(\mathcal{H})$ then*

$$Ran\delta_{D,E} \perp Ker\delta_{D,E}.$$

Proof. Assume that D and E are normal elements in $\mathcal{NA}(\mathcal{H})$. Then by Lemma 2, $\delta_{D,E}$ is normal and by Lemma 3 $Ran\delta_{D,E} \perp Ker\delta_{D,E}$.

□

Corollary 1. *If $A, B \in \mathcal{NA}(\mathcal{H})$ are normal and there exists $C \in \mathcal{NA}(\mathcal{H})$ such that $BC = I$ then $\overline{Ran\Theta_{A,C}} \cap Ker\Theta_{A,C} = \{0\}$.*

Proof. If $A, B \in \mathcal{NA}(\mathcal{H})$ are normal elements, then by Theorem 1,
$Ran\delta_{A,B} \perp Ker\delta_{A,B}$. This implies that $\overline{Ran\delta_{A,B}} \cap Ker\delta_{A,B} = \{0\}$. Using Lemma 1, we conclude that $\overline{Ran\Theta_{A,C}} \cap Ker\Theta_{A,C} = \{0\}$. □

The following theorem from Kittaneh [22] gives an orthogonality condition for generally linear operators. The proof is omitted.

Theorem 2. *Let be $\mathcal{NA}(\mathcal{H})$ a normed algebra with the norm $\|.\|$ satisfying $\|XY\| \leq \|X\|\|Y\|$ for all X, $Y \in \mathcal{NA}(\mathcal{H})$ and let $\delta : \mathcal{NA}(\mathcal{H}) \to \mathcal{NA}(\mathcal{H})$ be a linear map with $\|\delta\| \leq 1$. If $\delta(Y) = Y$ for some $Y \in \mathcal{NA}(\mathcal{H})$, then $\|\delta(X) - X + Y\| \geq \|Y\|$, for all $X \in \mathcal{NA}(\mathcal{H})$.*

We utilize the Theorem 2 to prove some results for general elementary operators. Let $T : \mathcal{NA}(\mathcal{H}) \to \mathcal{NA}(\mathcal{H})$ be an elementary operator defined by $T(X) = \sum_{i=1}^{n} A_i X B_i$, $\forall\ X \in \mathcal{NA}(\mathcal{H})$. From the above theorem we get the following lemma.

Lemma 4. *Suppose that $T(Y) = Y$ for some $Y \in \mathcal{NA}(\mathcal{H})$. If $\|T\| \leq 1$, then $\|T(X) - X + Y\| \geq \|Y\|$, for all $X \in \mathcal{NA}(\mathcal{H})$.*

Proof. The proof follows immediately from the proof of Theorem 2. □

Theorem 3. *Suppose that $T(Y) = Y$ for some $Y \in \mathcal{NA}(\mathcal{H})$. If*

$$\left\|\sum_{i=1}^{n} A_i A_i^*\right\|^{\frac{1}{2}} \left\|\sum_{i=1}^{n} B_i^* B_i\right\|^{\frac{1}{2}} \leq 1,$$

then $\|T(X) - X + Y\| \geq \|Y\|$, for all $X \in \mathcal{NA}(\mathcal{H})$.

Proof. We only need to show that $\|T\| \leq 1$. Let $Z_1 = [A_1, ..., A_n]$ and $Z_2 = [B_1, ..., B_n]^T$. Taking $Z_1 Z_1^*$ and $Z_2^* Z_2$ shows that

$$\|Z_1\| = \left\| \sum_{i=1}^{n} A_i A_i^* \right\|^{\frac{1}{2}}$$

and

$$\|Z_2\| = \left\| \sum_{i=1}^{n} B_i^* B_i \right\|^{\frac{1}{2}}.$$

From [14], it is known that $T(X) = Z_1(X \otimes I_n)Z_2$, where I_n is the identity of $M_n(\mathbb{C})$. Therefore it follows that $\|T(X)\| \leq \|Z_1\|\|Z_2\|\|X\|$. Hence $\|T\| \leq 1$. □

3. Orthogonality of Elementary Operators

At this point we consider the orthogonality of Jordan elementary operators. We later consider the necessary and sufficient conditions for their normality. We state the following theorem from [21] on orthogonality.

Theorem 4. *Let $A, B \in B(H)$ be normal operators, such that $AB = BA$, and let $\mathcal{U}(X) = AXB - BXA$. Furthermore, suppose that $A^*A + B^*B > 0$. If $S \in Ker\mathcal{U}$, then $|||\mathcal{U}(X) + S||| \geq |||S|||$.*

Remark 5. *The norm $|||.|||$ is a unitarily invariant norm.*

A unitarily invariant norm is any norm defined on some two-sided ideal of $B(H)$ and $B(H)$ itself which satisfies the following two conditions. For unitary operators $U, V \in B(H)$ the equality $|||UXV||| = |||X|||$ holds, and $|||X||| = s_1(X)$, for all rank one operators X. It is proved that any unitarily invariant norm depends only on the sequence of singular values. Also, it is known that the maximal ideal, on which $|||UXV|||$ has sense, is a Banach space with respect to that unitarily invariant norm. Among all unitarily invariant norms there are few important special cases. The first is the Schatten p-norm ($p \geq 1$) defined by $\|X\|_p = (\Sigma_{j=1}^{+\infty} s_j(X)^p)^{1/p}$ on the set $\mathcal{C}_p = \{X \in B(H) : \|X\|_p < +\infty\}$. For $p = 1, 2$ this norm is known as the nuclear norm (Hilbert-Schmidt norm) and the corresponding ideal is known as the ideal of nuclear (Hilbert-Schmidt) operators. The ideal $\mathcal{C}_2$ is also interesting for another reason. Namely, it is a Hilbert space with respect to the $\|.\|_2$ norm. The other important special case is the set of so-called Ky Fan norms $\|X\|_k = \Sigma_{j=1}^{k} s_j(X)$. The well-known Ky Fan dominance property asserts that the condition $\|X\|_k \leq \|Y\|$ for all $k \geq 1$ is necessary and sufficient for the validity of the inequality $|||X||| \leq |||Y|||$ in all unitarily invariant norms. For further details refer to [21].
We extend Theorem 4 to distinct operators $A, B, C, D \in B(H)$ in the theorem below.

Theorem 5. *Let $A, B, C, D \in B(H)$ be normal operators, such that $AC = CA$, $BD = DB$, $AA^* \leq CC^*$, $B^*B \leq D^*D$. For an elementary operator $\mathcal{U}(X) = AXB - CXD$ and $S \in B(H)$ satisfying $ASB = CSD$, then $\|\mathcal{U}(X) + S\| \geq \|S\|$, for all $X \in B(H)$.*

Proof. From $AA^* \leq CC^*$ and $B^*B \leq D^*D$, let $A = CU$, and $B = VD$, where U, V are contractions. So we have $AXB - CXD = CUXVD - CXD = C(UXV - X)D$. Assume C and D^* are injective, $ASB = CSD$ if and only if $USV = S$. Moreover, C and

U commute. Indeed from $A = CU$ we obtain $AC = CUC$. Therefore, $C(A - UC) = 0$. The result follows since $A = CU$ and C is injective. Similarly, D and V commute. So,

$$\begin{aligned} \|\mathcal{U}(X) + S\| &= \|[AXB - CXD] + S\| \\ &= \|[U(CXD)V - CXD] + S\| \\ &\geq \|S\|, \forall X \in B(H). \end{aligned}$$

Now, under the condition of Theorem 5, A and C have operator matrices $A = \begin{pmatrix} A_0 & 0 \\ 0 & 0 \end{pmatrix}$ and $C = \begin{pmatrix} C_0 & 0 \\ 0 & 0 \end{pmatrix}$ with respect to the space decomposition $H = \overline{\mathcal{R}(C)} \oplus \mathcal{N}(C)$, respectively. Here, A_0 is a normal operator on $\overline{\mathcal{R}(C)}$ and C_0 is an injective and normal operator on $\overline{\mathcal{R}(C)}$. B and D have operator matrices $B = \begin{pmatrix} B_0 & 0 \\ 0 & 0 \end{pmatrix}$ and $D = \begin{pmatrix} D_0 & 0 \\ 0 & 0 \end{pmatrix}$ with respect to the space decomposition $H = \overline{\mathcal{R}(D)} \oplus \mathcal{N}(D)$, respectively. Here, B_0 is a normal operator on $\overline{\mathcal{R}(D)}$ and D_0 is an injective and normal operator on $\overline{\mathcal{R}(D)}$. X and S have operator matrices $X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}$ and $S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$ which are as operator from the space decomposition $H = \overline{\mathcal{R}(D)} \oplus \mathcal{N}(D)$ into the space decomposition $H = \overline{\mathcal{R}(C)} \oplus \mathcal{N}(C)$, respectively.
In this case, $\mathcal{U}(X) = AXB - CXD = \begin{pmatrix} A_0X_{11}B_0 - C_0X_{11}D_0 & 0 \\ 0 & 0 \end{pmatrix}$ and $A_0S_{11}B_0 - C_0S_{11}D_0 = 0$. Therefore, $\|A_0X_{11}B_0 - C_0X_{11}D_0 + S_{11}\| \geq \|S_{11}\|$.
Hence,

$$\begin{aligned} \|\mathcal{U}(X) + S\| &= \| \begin{pmatrix} A_0X_{11}B_0 - C_0X_{11}D_0 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \| \\ &= \| \begin{pmatrix} A_0X_{11}B_0 - C_0X_{11}D_0 + S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \| \\ &\geq \| \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \|. \end{aligned}$$

□

Next consider the case of ideals of compact operators.We recall the following theorem in [16,Theorem III. 12.2].

Theorem 6. *If $\mathcal{J}$ is a separable ideal of compact operators, associated with some unitarily invariant norm, then its dual is isometrically isomorphic with another ideal of compact operators (not necessarily separable) and it admits the representation: $f_Y(X) = tr(XY)$.*

Definition 3. *Let $\mathcal{J}$ be some separable ideal of compact operators, and let $T : \mathcal{J} \to \mathcal{J}$ be some elementary operator given by $T(X) = \sum_{i=1}^{n} A_iXB_i$. Then its conjugate operator $T^* : \mathcal{J}^* \to \mathcal{J}^*$ has the form $T^*(Y) = \sum_{i=1}^{n} B_iYA_i$.*

Indeed from the theorem above we have

$$\begin{aligned} f_Y(T(X)) &= tr(T(X)Y) \\ &= tr(\sum_{i=1}^{n} A_i X B_i Y) \\ &= tr(X \sum_{i=1}^{n} B_i Y A_i) \\ &= tr(XT^*(Y)) \\ &= f_{T^*(Y)}(X). \end{aligned}$$

Definition 4. *The vector ξ in a Hilbert space H is a smooth point of the sphere $\mathcal{S}(0, \|\xi\|)$ if there exists a unique support functional $\phi_\xi \in H^*$, such that $\phi_\xi \xi = \|\xi\|$ and $\|\phi_\xi\| = 1$.*

Definition 5. *If there exists the Gateaux derivative of the norm at the point ξ in a Hilbert space H i.e. if there exists the limit*

$$\lim_{\mathbb{R} \ni t \to 0} \frac{\|\xi + t\zeta\| - \|\xi\|}{t} = 0,$$

then it is equal to $Re\phi_\xi\zeta$,where ϕ_ξ is the functional from the Definition 4. Moreover, in this case ζ is orthogonal to ξ if and only if $\phi_\xi\zeta = 0$.

Remark 6. *It is also well known in general that if Banach space X has a strictly convex dual space then every nonzero point is a smooth point of the corresponding sphere. For details see [3].*

Consider an arbitrary separable ideal of compact operators $\mathcal{J}$, such that $\mathcal{J}^*$ is strictly convex. According to Anderson's result, for all $Z \in \mathcal{J}$ there exists a unique operator $Z_0 \in \mathcal{J}^*$ such that $Z_0(Z) = \|Z\|$ and $\|Z_0\| = 1$. We recall that an ideal $\mathcal{J}$ is reflexive if it is equal to the algebra of all bounded linear operators which leave invariant each subspace left invariant by every operator in $\mathcal{J}$. So if we suppose that $\mathcal{J}$ is reflexive then the mapping $Z \to Z_0$, $Z_0 = \tau(Z)$ is a bijection (and also involution) of the unit spheres of the spaces $\mathcal{J}$ and $\mathcal{J}^*$. Moreover, S is orthogonal to Z in the space $\mathcal{J}$ (see [16]) if and only if $Z_0(Y) = 0$.

Lemma 5. *Let $\mathcal{J}$ be a reflexive ideal in $B(H)$ such that $\mathcal{J}^*$ is strictly convex, and let $T : \mathcal{J} \to \mathcal{J}$ be some elementary operator given by $T(X) = \sum_{i=1}^{n} A_i X B_i$. Then RanT is orthogonal (in the sense of Birkhoff) to the operator R if and only if $\tau(R) = R_0 \in KerT^*$.*

Proof. From Definitions 3 and 5, we see that Ran$T$ is orthogonal (in the sense of Birkhoff) to the operator R implies that for all $Z \in \mathcal{J}$, $R_0(T(Z)) = 0$ or $(T^*(R_0)Z) = 0$ for all $Z \in \mathcal{J}$ and as a consequence $T^*(R_0) = 0$. □

Next, consider the elementary operator acting on the ideal of operators in $B(H)$. We prove that for such elementary operators its range is orthogonal to its kernel as we see in the theorem below.

Theorem 7. *Let $\mathcal{J}$ be a reflexive ideal in $B(H)$ such that $\mathcal{J}^*$ is strictly convex, and let $T : \mathcal{J} \to \mathcal{J}$ be some elementary operator given by $T(Z) = AZB + CZD$, where $A, B, C, D \in B(H)$ are distinct normal operators, such that $AC = CA$, $BD = DB$, $AA^* \leq CC^*$, $B^*B \leq D^*D$. Then $\mathcal{J} = \overline{RanT} \oplus KerT$.*

Proof. From Lemma 5 and [21,Theorem 3], we get $T(R) = 0$ implies that for all $Z \in \mathcal{J}$, $|||T(Z) + R||| \geq |||R|||$, which implies that $T^*(R) = 0$ which further implies that for all $Z \in \mathcal{J}^*$, $|||T^*(Z) + R_0||| \geq |||R_0|||$, which implies that $T^{**}(R_{00}) = 0$ if and only if $T(R) = 0$. We conclude that $T(R) = 0$ if and only if $T(Z)$ is orthogonal to R for all $Z \in \mathcal{J}$. Next consider $\xi_n \in \overline{RanT}$ and $\zeta_n \in KerT$. If $\xi_n + \zeta_n$ tends to η then, from the inequality $\|\zeta_n - \zeta_m\| \leq \|\xi_n + \zeta_n - \xi_m - \zeta_m\|$ we come to a conclusion that ζ_n is Cauchy and hence $\zeta_n \to \zeta \in KerT$. Moreover, $\xi_n \to \eta - \zeta \in \overline{RanT}$. Therefore, $\eta \in \overline{RanT} + KerT$. Hence it follows that $\overline{RanT} + KerT$ is closed. Now assume that $\|T(Z) + Y + X\|_{\mathcal{J}} \geq \|X\|_{\mathcal{J}}$ for all Z, and for all $Y \in KerT$. If we choose $Y = 0$ then $X \in KerT$. Put $Y = -X$ and $Z = 0$, this implies that $X = 0$. It follows that $(\overline{RanT} + KerT)_\perp = \{0\}$. In fact, this last assertion follows from [33, Lemma 3.6]. □

4. Finite Operators and Orthogonality

Definition 6. *Let H be a complex separable and infinite dimensional Hilbert space and $B(H)$ be the algebra of all bounded linear operators on H. An operator $A \in B(H)$ is called finite if $\|I - (AX - XA)\| \geq 1$, $\forall\ X \in H$. Equivalently, we say that A is a finite operator if the distance from the identity to the range of the inner derivation δ_A is greater or equal to 1.We denote the set of all finite operators by $F(H)$.*

Examples of finite operators include: Hyponormal, p-hyponormal, log-hyponormal, dominant, quasihyponormal, k-quasihyponormal, (p,k)-quasihyponormal, paranormal and normaloid operators. We define these operators as follows:

Definition 7. *Let H be a complex separable and infinite dimensional Hilbert space and $B(H)$ be the algebra of all bounded linear operators on H. An operator $A \in B(H)$ is called normal if $AA^* = A^*A$, where A^* is the adjoint of A, hyponormal if $A^*A - AA^* \geq 0$, p-hyponormal if $(A^*A)^p - (AA^*)^p \geq 0$, for $0 < p \leq 1$, quasihyponormal if $A^*(A^*A - AA^*)A \geq 0$, k-quasihyponormal if $A^{*k}(A^*A - AA^*)A^k \geq 0$, for $k \in \mathbb{N}$, (p, k)-quasihyponormal if $A^{*k}((A^*A)^p - (AA^*)^p)A^k \geq 0$, for $k \in \mathbb{N}$,$0 < p \leq 1$, log-hyponormal if A is invertible and satisfies: $log(A^*A) \geq log(AA^*)$, paranormal if $\|Ax\|^2 \leq \|A^2x\|\|x\|$, $\forall\ x \in H$, normaloid if $\|A\| = r(A)$, where $r(A)$ is the spectral radius of A. If A satisfies $\|A^2\| \geq \|A\|^2$, then A is said to be in class A operators and A is called a dominant operator if for all complex β, $Ran(A - \beta) \subset Ran(A - \beta)^*$ or equivalently if there is a real number $M_\beta \geq 1$ such that $\|(A - \beta)^*x\| \leq M_\beta\|(A - \beta)x\|$ for all $x \in H$.*

Remark 7. *We have*

(a) *$Normal \subset Hyponormal \subset p - hyponormal \subset p - quasihyponormal \subset (p,k) - quasihyponormal \subset normaloid$.*

(b) *log-hyponormal $\subset$hyponormal $\subset$ paranormal.*

(c) *$Normal \subset$ hyponormal $\subset$ paranormal $\subset$ normaloid.*

Lemma 6. *An operator $A \in B(H)$ is finite if and only if it's adjoint is finite.*

Proof. The proof is obvious, so it is omitted. □

One way of defining the numerical range is through states. Let $B(H)$ denote a complex Banach algebra with identity I. A state on $B(H)$ is a functional $f \in (B(H))^*$ such that $f(I) = 1 = ||f||$. For $T \in B(H)$, the numerical range of T is defined by $W_0(T) = \{f(T) : f \text{ is a state on } B(H)\}$, [34]. $W_0(T)$ is a compact convex set containing $\overline{convh\sigma(T)}$, the convex hull of the spectrum of T. An element $A \in B(H)$ is finite if $0 \in \overline{W_0(AX - XA)}$ for each $X \in B(H)$
We see that the following conditions are equivalent from [34, Theorem 4].

Proposition 2 (34, Theorem 4)**.** *For $A \in B(H)$ the following are equivalent: (i) $0 \in \overline{W_0(AX - XA)}$ for each $X \in B(H)$; (ii) $\|I - (AX - XA)\| \geq 1$, $\forall\ X \in H$. (iii) there exists a state f on $B(H)$ such that $f(AX) = f(XA)$ for every $X \in B(H)$.*

Proposition 3. *Let $A \in B(H)$ be such that there is a nonzero selfadjoint linear functional ϕ on $B(H)$ such that $\phi(AX) = \phi(XA)$ for all $X \in B(H)$. Then A is a finite operator.*

Proof. Let $\mathcal{C} = \{T \in B(H) : \phi(TX) = \phi(XT),\ for\ all\ X \in B(H)\}$. Since $\mathcal{C}$ is a $\mathcal{C}^*$-subalgebra of $B(H)$ which contains A, $\mathcal{C}^*(A) \subseteq \mathcal{C}$. Let $\phi = \phi^+ + \phi^-$. Then it follows that $\phi^+(UXU^*) = \phi^+(X)$ and $\phi^-(UXU^*) = \phi^-(X)$ for all unitaries $U \in U(\mathcal{C})$ and $X \in B(H)$. Thus $\phi^+(TX) = \phi^+(XT)$ and $\phi^-(TX) = \phi^-(XT)$ for all $X \in B(H)$ and $T \in \mathcal{C}$, hence $\phi^+(AX) = \phi^+(XA)$ and $\phi^-(AX) = \phi^-(XA)$ for all $X \in B(H)$. Since at least one of ϕ^+, ϕ^- must be nonzero, it follows that A must be a finite operator. □

Definition 8. *Let $A \in B(H)$. The reduced approximate spectrum of A, $\sigma_{ra}(A)$, is the set of scalars λ for which there exists a normed sequence $\{x_n\}$ in H satisfying $(A-\lambda I)x_n \to 0$ and $(A - \lambda I)^* x_n \to 0$. We define the set $\overline{R_0} = \{A \in B(H) : \sigma_{ra}(A) \neq \emptyset\}$.*

Theorem 8. *$A \in B(H)$ is a finite operator if $A \in \overline{R_0}$.*

Proof. Let $A \in B(H)$. If $A \in \overline{R_0}$, then there exists $\lambda \in \sigma_{ra}(A)$ and a normed sequence $\{x_n\}$ in H such that $(A - \lambda I)x_n \to 0$ and $(A - \lambda I)^* x_n \to 0$. Now, for all $X \in B(H)$, we have

$$\begin{aligned} \|I - (AX - XA)\| &= \|I - ((A - \lambda I)X - X(A - \lambda I))\| \\ &= |I - (\langle (A - \lambda I)Xx_n, x_n\rangle - \langle X(A - \lambda I))x_n, x_n\rangle)|. \end{aligned}$$

Taking limits as $n \to \infty$, we obtain,

$$\|I - (AX - XA)\| = 1,\ \forall\ X \in H.$$

and hence A is finite. □

Next we consider the Berberian technique involving *-isometric isomorphisms which preserve order. We state the following theorem from [31].

Theorem 9. *Let H be a complex Hilbert space. Then there exists a Hilbert space $H' \supset H$ and $\varphi : B(H) \to B(H)$ $(A \mapsto A')$ satisfying: φ is an *-isometric isomorphism preserving the order such that*
(i) $\varphi(A^*) = \varphi(A)^*, \varphi(I) = I', \varphi(\alpha A + \beta B) = \alpha\varphi(A) + \beta\varphi(B), \varphi(AB) = \varphi(A)\varphi(B), ||\varphi(A)|| = ||A||, \varphi(A) \leq \varphi(B), if\ A \leq B\ for\ all\ A, B \in B(H)$ and for all $\alpha, \beta \in \mathbb{C}$.
(ii) $\sigma(A) = \sigma(A') = \sigma_a(A) = \sigma_a(A') = \sigma_p(A')$, where $\sigma_a(A)$ is the approximate spectrum of A and $\sigma_p(A)$ is the point spectrum of A.

Proof. See proof in [31]. □

We apply the above theorem in the next result as follows:

Theorem 10. *Let $B(H)$ be a C^* -algebra, and let A be a finite operator then I does not belong to $\overline{Ran\delta_A}$.*

Proof. By Theorem 9 it is known that there exists a *-isometric isomorphism $\varphi : B(H) \to B(H)$ which preserves the order such that

$$\begin{aligned} \|I - (AX - XA)\| &= \|\varphi(I - (AX - XA))\| \\ &= |I - (\varphi(A)\varphi(X) - \varphi(X)\varphi(A))|. \end{aligned}$$

with $\varphi(A) \in B(H)$ for a finite operator $A \in B(H)$.
So,

$$\begin{aligned} \|I - (AX - XA)\| &= \|\varphi(I - (AX - XA))\| \\ &= \|I - (\varphi(A)\varphi(X) - \varphi(X)\varphi(A))\| \\ &\geq 1, \end{aligned}$$

which implies that I does not belong to $\overline{Ran\delta_A}$. □

Theorem 11. *Let $A \in B(H)$ be a normal operator. Then A is finite if $\sigma_{ra}(A) \neq \emptyset$.*

Proof. From Remark 7, A being normal implies A is normaloid. But for normaloid operators, it is a fact that $\|A\| = r(A)$. So there exists $\lambda \in \sigma_{ra}(A)$ such that $\|A\| = |\lambda|$. But, λ is in the boundary of $\sigma_{ra}(A)$, therefore, there exists a normed sequence $\{x_n\}$ in H such that $(A - \lambda I)x_n \to 0$ and $(A - \lambda I)^* x_n \to 0$, because $\|A\| = |\lambda|$. From [31], see the assertion for finiteness of A. □

Theorem 12. *Let H be a complex separable and infinite dimensional Hilbert space and $B(H)$ be the algebra of all bounded linear operators on H. For each positive integer k let $\overline{R_k}$ denote the set of operators on H that have an k-dimensional reducing subspace. Then $\overline{R_k}$ is a subspace of $F(H)$.*

Proof. Let H be decomposed as $H = H_1 \oplus H_2$ where H_1 is a k-dimensional reducing subspace of A given by the following matrix representation, $A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$, relative to the decomposition of H. Let $X = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}$ for all $X \in B(H)$. A simple computation shows [31] that

$$\|I - (AX - XA)\| \geq \|I_1 - (A_1X_1 - X_1A_1)\| = \|I_2 - (A_2X_4 - X_4A_2)\|,$$

where $I = \begin{pmatrix} I_1 & 0 \\ 0 & I_2 \end{pmatrix}$ is the identity operator on H. This shows that $\overline{W_0(AX - XA)}$ contains the numerical range of $A_1X_1 - X_1A_1$. Since H_1 is finite dimensional, the latter commutator has trace 0 and thus by [34],

$$0 = \frac{1}{k} tr(A_1X_1 - X_1A_1) \in \overline{W_0(A_1X_1 - X_1A_1)}$$

since the numerical range is a convex set whose closure contains the spectrum. If $A = X \oplus P$ where P is an operator of finite rank, then the vector sum of the ranges of P and P^* is a finite dimensional reducing subspace for A, hence $A \in \bigcup_k R_k$. Therefore, $F(H)$ contains every operator that can be written as a uniform limit of operators each having a summand of finite rank. In particular, $F(H)$ contains every operator with a compact direct summand. □

The following example is a nice finite operator.

Example 1. *Let* $A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & A_{2,4} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \in B(H^{(4)})$. *Then* A *is finite if and only if* $A_{2,4} \in B(H)$ *is finite.*

Indeed, following the proof in [34, Theorem 8], a tedious computation based on the same argument shows that A is finite.

5. Generalized Finite Operators and Orthogonality

Let H be a complex separable and infinite dimensional Hilbert space and $B(H)$ be the algebra of all bounded linear operators on H. Let $A, B \in B(H)$. We define the generalized derivation $\delta_{A,B} : B(H) \to B(H)$ by

$$\delta_{A,B}(X) = AX - XB, \ \forall \ X \in B(H).$$

Definition 9. *Let* H *be a complex separable and infinite dimensional Hilbert space and* $B(H)$ *be the algebra of all bounded linear operators on* H. *Operators* $A, B \in B(H)$ *are called general finite if* $\|I - (AX - XB)\| \geq 1, \ \forall \ X \in B(H)$. *We denote the set of all general finite operators by* $GF(H)$. *So,*

$$GF(H) = \{(A, B) \in B(H) \times B(H) : \|I - (AX - XB)\| \geq 1, \ \forall \ X \in B(H)\}.$$

Theorem 13. *For* $A, B \in B(H)$, *the following statements are equivalent (i)* $\|I-(AX-XB)\| \geq 1, \ \forall \ X \in B(H)$
(ii) There exists a state f *such that* $f(AX) = f(XB)$, *for all* $X \in B(H)$.
(iii) $0 \in W_0(AX - XB), \forall \ X \in B(H)$.

Proof. For proof, see [31]. □

For orthogonality in the sense of Birkhoff. Let E be a complex Banach space, we say that $y \in E$ is orthogonal to $x \in E$ if for all complex α there holds $\|x + \alpha y\| \geq \|x\|$. This definition has a natural geometric interpretation. Namely, $y \perp x$ if and only if the complex line $\{x + \alpha y : \alpha \in \mathbb{C}\}$ is disjoint with the open ball $K(0, \|x\|)$, i.e., if and only if this complex line is a tangent one. Note that if y is orthogonal to x, then x need not be orthogonal to y. If E is a Hilbert space, then $\langle x, y\rangle = 0$, i.e., orthogonality in the usual sense. Concerning elementary operators, the range of $\delta_{A,B}$ is orthogonal to the null space of $\delta_{A,B}$. In particular the inequality $||T-(AX-XB)|| \geq ||T||$ means that the $Ran\delta_{A,B}$ is orthogonal to $Ker\delta_{A,B}$ in the sense of Birkhoff. We see that if $Ran\delta_{A,B}$ is orthogonal to $Ker\delta_{A,B}$, then (A, B) is generalized finite if and only if $T = I \in Ker\delta_{A,B}$.

Bibliography

[1] Anderson J., *On normal derivations,* Proc. Amer. Math. Soc. **38** (1973), 135-140.

[2] Bachir A. and Segres A., *Numerical range and orthogonality in normed spaces*, Filmat, **23** (2009), 21-41.

[3] Bachman G. and Narici L., *Functional Analysis*, Academic press, New York, 2000.

[4] Benitez C., *Orthogonality in normed linear spaces: a classification of the different concepts and some open problems,* Revista Mathematica, **2** (1989), 53-57.

[5] Bhattacharya D. K and Maity A. K., *Semilinear tensor product of* Γ- *Banach algebras,* Ganita Vol. 40. No. 2, (1989), 75-80.

[6] Bonsall F. and Duncan J., *Complete normed algebra,* springer verlag, New York, 1973.

[7] Bonyo J. O and Agure J. O., *Norm of a derivation and hyponormal operators,* Int. Journal of Math. Analysis, Vol 4. No. 14, (2010), 687-693.

[8] Bouali S. and Bouhafsi Y., *On the range of the elementary operator* $X \mapsto AXA-X$, Math. Proc. Royal Irish Academy, **108** (2008), 1-6.

[9] Cabrera M. and Rodriguez A., *Nondegenerately ultraprime Jordan Banach algebras,* Proc. London Math. Soc. **69** (1994), 576-604.

[10] Canavati J. A., Djordjevic S. V. and Duggal B. P., *On the range closure of an elementary operator,* Bull Korean Math. Soc. **43** (2006), 671-677.

[11] Du H. K. and Ji H. X., *Norm attainability of Elementary Operators and derivations*, Northeast. Math. J. **3** (1994), 394-400.

[12] Du H. K. Wang H. K. and Gao G. B., *Norms of Elementary Operators,* Proc. Amer. Math. Soc. **4** (2008), 1337-1348.

[13] Dutta T. K., Nath H. K and Kalita R. C., α*-derivations and their norms in projective tensor products of* Γ*-Banach algebras*, J. London math. Soc. Vol. 2, No. 2 (1998), 359-368.

[14] Franka M. B., *Tensor products of C*-algebras, operator spaces and Hilbert C*-modules*, Math. Comm. **4** (1999), 257-268.

[15] Gajendragadkar P., *Norm of Derivations of Von-Neumann algebra*, Journ. Trans. Amer. Math. Soc. **170** (1972), 165-170.

[16] Gohberg I. C. and Krein M. G., *Introduction to the Theory of Linear Nonselfadjoint Operators,* Transl. Math. Monogr., vol. **18**, Amer. Math. Soc., Providence, RI, 1969.

[17] Helemskii A., *Lectures and exercises on Functional Analysis*, Translation of mathematical monographs, Vol. **233**, American Mathematical Society, New York, 2006.

[18] Iain R. and Dana P. W., *Equivalence and Continuous-Trace C*-algebras,* American Mathematical Society, New York, 1998.

[19] Kadison R., Lance C. and Ringrose J., *Derivations and Automorphisms of Operator Algebra II*, Math. Journ. Functional Analysis **1** (1967), 204-221.

[20] Keckic D. J., *Orthogonality in C_1 and C_∞ spaces and normal derivations*, J. Operator theory **51** (2004), 89-104.

[21] Keckic D. J., *Orthogonality of the range and kernel of some elementary operators*, Proc. Amer. Math. Soc. **11** (2008), 3369-3377.

[22] Kittaneh F., *Normal derivations in normal ideals*, Proc. Amer. Math. Soc. **6** (1995), 1979-1985.

[23] Kreyzig E., *Introductory Functional Analysis with Applications*, John Wiley and sons, New York, 1978.

[24] Kyle J., *Norms of Derivations*, Journ. London Math. Soc. **16** (1997), 297-312.

[25] Magajna B., *The norm of a symmetric elementary operator*, Proc. Amer. Math. Soc. **132** (2004), 1747-1754.

[26] Mathieu M., *Properties of the product of two derivations of a C*- algebra*, Canad. Math. Bull. **42** (1990), 115-120.

[27] Mathieu M., *More properties of the product of two derivations of a C*- algebra*, Canad. Math. Bull. **42** (1990), 115-120.

[28] Mathieu M., *Elementary operators on Calkin algebras*, Irish Math. Soc. Bull. **46** (2001), 33-42.

[29] Masamichi T., *Theory of Operator Algebras I,* Springer-Verlag, New York, Inc., 1979.

[30] Mecheri S., *On the range and kernel of the elementary operators $\sum_{i=1}^{n} A_i X B_i - X$*, Acta Math. Univ. Comnianae, **52** (2003), 119-126.

[31] Mecheri S., *Finite operators*, Demonstratio Math. **35** (2002), 355-366.

[32] Murphy J. G., *C*-algebras and Operator Theory,* Academic Press Inc., Oval Road, London, 1990.

[33] Turnsek A., *Orthogonality in $\mathcal{C}_p$ classes,* Monatsh. Math. **132** (2001), 349 - 354.

[34] Williams J. P., *Finite operators,* Proc. Amer. Math. Soc. **26** (1970), 129- 135.

Mathematical Model for Pneumonia Dynamics among Children

Ong'ala J. Otieno[1], J. Odhiambo and J.Y.T. Mugisha

Center for Applied Research in Mathematical Sciences,
Strathmore University, Box 59857 - 00200 City Square, Nairobi, Kenya.
Department of Mathematics,
Makerere University, Box 7062, Kampala, Uganda

Abstract. Major advances have been made to understand the epidemiology of infectious diseases. However, more than 2 million children in the developing countries still die from pneumonia each year. The efforts to promptly detect, effectively treat and control the spread of pneumonia is possible if its transmission dynamics are understood. In this paper, we develop a mathematical model for pneumonia among children under five years of age. The model is analyzed using the theory of ordinary differential equations and dynamical systems. We derive the basic reproduction number, R_0, analyze the stability of equilibrium points and bifurcation analysis. The results of the analysis show that there exists a locally stable disease free equilibrium point, E^f when $R_0 < 1$ and a unique endemic equilibrium, E^e when $R_0 > 1$.The analysis also shows that there is a possibility of a forward bifurcation.

Keywords: Pneumonia Model, Basic reproduction number, forward bifurcation, Stability, Carriers

1. Introduction

Pneumonia is a high-incidence respiratory disease characterized by an inflammatory condition of the lungs. It is caused by bacteria, fungi, parasites and viruses. Among the four micro-organisms with potential to cause pneumonia, bacteria is the leading cause [35, 32] especially *Streptococcus Pneumoniae* [8, 18, 25]. When bacteria enters the lungs, it settles in the alveoli and passages of the lungs where they rapidly grow and multiply in number. The area of the lung that is invaded then becomes filled with fluid and pus as the body attempts to fight off the infection [22]. This makes breathing difficult, painful and limits the intake of oxygen.

Most cases of pneumonia are as a result of inhaling small droplets of coughs or sneezes containing the bacteria. These droplets get into the air when an infected person coughs or sneezes [22, 35]. The bacteria can also be carried in the mouth or flora of nasopharynx of a healthy person without causing any harm [18, 22, 35]. Such people are referred to as carriers. For carriers, the bacteria invades the lungs and causes infection [16],[22]. This is possible when the immunity of the individual is lowered.

[1]Corresponding author. Email: jotieno@strathmore.edu

There is limited information on the transmission patterns of the pneumococcal disease in the developing world [8]. However, it is pointed out that the risk factors associated with the spread of the disease include malnutrition, lack of exclusive breastfeeding, indoor pollution and antecedent viral infection [8, 25].

Despite the increasing focus on the Millennium Development Goal (MDG) 4 of the United Nations [36], "to reduce child mortality", almost 1.9 million children still die from pneumonia each year in the developing countries, accounting for 20 % of deaths globally [11]. In Kenya, pneumonia contributes up to 16 % of child mortality [7]. It is evident that the management of the disease is challenging due to overlap of its symptoms with that of malaria hence a possibility of mistreatment with antimalarial drugs [40]. Deaths due to pneumonia can occur within three days of illness and any delays in proper treatment may threaten life [28].

Therefore to realize MDG 4, research should be done to promptly diagnose, effectively treat and deduce other prevention strategies for pneumonia. For this to be achieved, accurate projections on possibility of epidemic or endemic and strategies to put up control measures is required. Mathematical models integrated in epidemiological research are powerful tools in studying the dynamics of diseases and to find threshold parameters necessary for controlling the disease. In this paper, we develop and analyze a mathematical model for pneumonia dynamics in children. Our work is based on an initial model by Doura et al. [47].

2. Derivation of the Model

The transmission dynamics of pneumonia in the population under study is considered between four compartments based on the disease status, that is: Susceptible, Carriers, Infectious and Recovered. At time t, the total population size N is divided into: susceptible S, infected I, carriers C and recovered R such that

$$N = S + C + I + R. \tag{2.1}$$

The per capita recruitment rate into the susceptible population is denoted ν. We assume that the infected immigrants are not included because they are not able to travel. New infections can be due to effective contact with either a carrier or a symptomatically infected individual, where the force of infection of susceptibles is denoted by λ. A newly infected individual joins a carrier class with a probability ρ or a symptomatically infected class with a probability of $1 - \rho$. Carriers can change their status to show symptoms (infected) [24] at the rate π. Infected individuals recover at the rate η. A proportion q of the recovered individuals clear all the bacteria from the body and gain temporal immunity while $1 - q$ of them will still carry the bacteria [52, 51]. The carriers can also recover to gain temporal immunity at the rate β. In this model, the temporal immunity is a result of all possible ways that may lead to recovery from the disease. Studies in [34, 23, 29] show that there is a possibility of reinfection at the rate δ. There is a natural death rate of μ and a disease induced death rate of α. We define the force of infection as:

$$\lambda = \psi \left(\frac{I + \varepsilon C}{N} \right) : \psi = \kappa \mathcal{P}. \tag{2.2}$$

Where κ is the rate of contact and $\mathcal{P}$ be the probability that a contact is efficient to cause infection. Combining all the definitions and assumptions, the model for the transmission dynamics of pneumonia is given by the following system of differential equations:

$$\begin{aligned}
\frac{dS}{dt} &= \nu + \delta R - (\lambda + \mu)S, \\
\frac{dI}{dt} &= (1-\rho)\lambda S + \pi C - (\mu + \alpha + \eta)I, \\
\frac{dC}{dt} &= \rho(\lambda)S + (1-q)\eta I - (\mu + \pi + \beta)C, \\
\frac{dR}{dt} &= q\eta I + \beta C - (\mu + \delta)R
\end{aligned} \tag{2.3}$$

3. Positivity and boundedness of solutions

We can show from Model (2.3) that the state variables are non-negative and the solutions remain positive for all time $t \geq 0$. Here the parameters in the model are assumed to be positive. We also show that the feasible solutions are bounded in a region:

$$\Phi = \{(S, I, C, R) \in \mathcal{R}_+^4 : N(t) \leq \frac{\nu}{\mu}\}$$

Lemma 1. *Let the initial values of the parameters be $S(0) \geq 0$, $I(0) \geq 0$, $C(0) \geq 0$, $R(0) = 0$ and $N(0) \geq 0 \in \Phi$. Then the solution set $\{S(t), I(t), C(t), R(t), N(t)\}$ is positive for all $t \geq 0$*

Proof. Consider the first equation in (2.3):

$$\frac{dS}{dt} = \nu + \delta R - \lambda S - \mu S.$$

We have

$$\begin{aligned}
\frac{dS}{dt} &\geq -(\lambda + \mu)S \\
\Rightarrow \int \frac{dS}{S} &\geq \int -(\lambda + \mu)dt \\
\Rightarrow S &\geq S_0 e^{-(\lambda+\mu)t}
\end{aligned}$$

Hence, $S \geq 0$.

Next, we consider the second equation in (2.3)

$$\begin{aligned}
\frac{dI}{dt} &= (1-\rho)\lambda S + \pi C - (\mu + \alpha + \eta)I \\
\frac{dI}{dt} &\geq -(\mu + \alpha + \eta)I \\
\int \frac{dI}{I} &\geq \int -(\mu + \alpha + \eta)dt \\
I &\geq I_0 e^{-(\mu+\alpha+\eta)t}
\end{aligned}$$

Hence, $I \geq 0$.

Similarly, we can prove the positivity of C, R and N. □

Lemma 2. *The solutions for the system (2.1) are contained and remain in the region* Φ *for all time* $t \geq 0$

Proof. Consider Equation (2.1). Taking the derivatives with respect to time t of (2.1) and substituting onto it the set of equations in (2.3), we have,

$$\begin{aligned} \frac{dN(t)}{dt} &= \nu - \alpha I - \mu N \\ \Rightarrow \quad \frac{dN}{dt} &\leq \nu - \mu N \\ \Rightarrow \quad N &\leq \frac{\nu}{\mu} + \left(N_0 - \frac{\nu}{\mu}\right) e^{-\mu t} \end{aligned}$$

Where N_0 is initial population size. Thus,

$$\lim_{t \to \infty} N(t) \leq \frac{\nu}{\mu}.$$

Using this result together with Lemma 1 and equation 2.1, we have that $0 \leq N(t) \leq \frac{\nu}{\mu}$ which implies that N and all other variable (S, I, C and R) are bounded and all the solutions starting in Φ approach, enter or stay in Φ. □

4. Analysis of the Model

We analyze the model for pneumonia transmission in the following sub sections to determine the basic reproduction number and other threshold parameters for pneumonia dynamics.

4.1 Stability analysis of the disease-free equilibrium (DFE)

The DFE of model (2.3) is obtained by equating the right-hand sides of the equations in the model to zero and it describes the model in absence of disease or infection. Here we define carrier and infected classes as diseased classes, DFE denoted by $E^f = (S^f, I^f, C^f, R^f)$ is then given by $E^f = \left(\frac{\nu}{\mu}, 0, 0, 0\right)$.

Theorem 1. *There is a unique DFE* (E^f) *for the model represented by the system of equations in (2.3)*

Proof. This lemma is proven by substituting E^f into the system of Equations (2.3). The results show that all derivatives are equal to zero, hence DFE is an equilibrium point. □

To establish the linear stability of E^f, we use the next-generation operator approach [53] on the system (2.3) to compute the basic reproduction number R_0. Using the notation of the matrices F and V as in [53], we have,

$$\mathbf{F} = \begin{pmatrix} (1-\rho)\psi & (1-\rho)\psi\varepsilon \\ \\ \rho\psi & \rho\psi\varepsilon \end{pmatrix} \text{ and } \mathbf{V} = \begin{pmatrix} h_1 & -\pi \\ \\ -(1-q)\eta & h_2 \end{pmatrix}$$

where $h_1 = \mu + \alpha + \eta$ and $h_2 = \mu + \pi + \beta$ The eigenvalues for the matrix FV^{-1} are

$$0, \psi\left(\frac{\rho[\varepsilon h_1 + \pi] + (1-\rho)[h_2 + (1-q)\varepsilon\eta]}{h_1 h_2 - (1-q)\pi\eta}\right).$$

Thus from Theorem 2 of [53] we have

$$R_0 = \psi\left(\frac{\rho[\varepsilon h_1 + \pi] + (1-\rho)[h_2 + (1-q)\varepsilon\eta]}{h_1 h_2 - (1-q)\pi\eta}\right). \tag{4.1}$$

Lemma 3. *The disease-free equilibrium E^f of (2.3) is locally asymptotically stable whenever $R_0 < 1$ and unstable when $R_0 > 1$.*

Proof. Consider the Jacobian matrix for the model (2.3) at E^f given as

$$\mathcal{J}(E^f) = \begin{pmatrix} -\mu & -\psi & -\psi\varepsilon & \delta \\ 0 & -h_1 & \pi & 0 \\ 0 & (1-q)\eta & -h_2 & 0 \\ 0 & q\eta & \beta & -(\mu+\delta) \end{pmatrix}$$

and

$$\left.\begin{array}{l} Trace\left[\mathcal{J}(E^f)\right] = -(2\mu + \delta + h_1 + h_2) < 0 \\ Det\left[\mathcal{J}(E^f)\right] = \mu(\delta+\mu)[h_1 h_2 - (1-q)\pi\eta] > 0 \end{array}\right\} \tag{4.2}$$

Since the parameters μ, δ, h_1 and h_2 are all positive, $-(2\mu + \delta + h_1 + h_2) < 0$. Therefore $Trace\left[\mathcal{J}(E^f)\right] < 0$. On the other hand, R_0 can never be negative. Since the numerator $\{\rho[\varepsilon h_1 + \pi] + (1-\rho)[h_2 + (1-q)\varepsilon\eta]\}$ is positive, the denominator must also be positive, i.e. $h_1 h_2 - (1-q)\pi\eta > 0$. This implies that $Det\left[\mathcal{J}(E^f)\right] > 0$, since $\mu(\delta+\mu) > 0$ and $[h_1 h_2 - (1-q)\pi\eta] > 0$. Thus

$$R_0 = \psi\left(\frac{\rho[\varepsilon h_1 + \pi] + (1-\rho)[h_2 + (1-q)\varepsilon\eta]}{h_1 h_2 - (1-q)\pi\eta}\right) < 1.$$

The solutions in (4.2) imply that E^0 is locally asymptotically stable whenever $R_0 < 1$ □

4.2 Stability of the Endemic equilibrium (EE) and Bifurcation analysis)

The endemic equilibrium is denoted by E^e and defined as a steady state solution for the model (2.3). This can occur when there is a persistence of the disease. Hence $E^e = \{S^e, I^e, C^e, R^e\}$ can be expressed as shown below.

$$\begin{aligned}
S^e &= \frac{N}{R_0} \\
C^e &= \frac{(\mu+\delta)\left((1-\rho)(1-q)\eta + h_1\rho\right)(R_0-1)\nu}{R_0\left((\mu+\delta)(h_2h_1 - (1-q)\pi\eta) - \delta\left(\rho(\eta\pi q + h_1\beta) + (1-\rho)(\eta q h_2 + (1-q)\eta\beta)\right)\right)} \\
I^e &= \frac{(\delta+\mu)(\pi\rho + (1-\rho)h_2)(R_0-1)\nu}{R_0\left((\delta+\mu)(h_2h_1 - (1-q)\pi\eta) - \delta\left(\rho(\pi q\eta + h_1\beta) + (1-\rho)(\eta q h_2 + (1-q)\eta\beta)\right)\right)} \\
R^e &= \frac{\left(\rho(\eta\pi q + h_1\beta) + (-\rho+1)(h_2 q\eta + (1-q)\eta\beta)\right)(R_0-1)\nu}{R_0\left((\delta+\mu)(h_2h_1 - (1-q)\pi\eta) - \delta\left(\rho(\pi q\eta + h_1\beta) + (1-\rho)(\eta q h_2 + (1-q)\eta\beta)\right)\right)}
\end{aligned}$$

Lemma 4. *For $R_0 > 1$ a unique endemic equilibrium point E^e exists. There is no endemic equilibrium otherwise.*

Proof. For the disease to be endemic, then $\frac{dI}{dt} > 0$ and $\frac{dC}{dt} > 0$, that is,

$$\begin{aligned}(1-\rho)\psi\frac{S}{N}(I+\varepsilon C)+\pi C-h_1 I>0\\ \rho\psi\frac{S}{N}(I+\varepsilon C)+(1-q)\eta I-h_2 C>0\end{aligned} \tag{4.3}$$

From the first inequality of 4.3 we have

$$h_1 I < (1-\rho)\psi\frac{S}{N}(I+\varepsilon C)+\pi C.$$

Using the fact that $\frac{S}{N} \leq 1$,

$$I < \frac{(1-\rho)\psi I+(1-\rho)\psi\varepsilon C+\pi C}{h_1}, \tag{4.4}$$

and from the first inequality of 4.3 we have

$$C < \frac{\rho\psi I+(1-q)\eta I}{h_2-\rho\psi\varepsilon}. \tag{4.5}$$

Substituting 4.5 into 4.4, we have

$$\begin{aligned}I &< \frac{(1-\rho)\psi I+[(1-\rho)\psi\varepsilon+\pi]\left[\frac{\rho\psi I+(1-q)\eta I}{h_2-\rho\psi\varepsilon}\right]}{h_1}\\ 1 &< \frac{(1-\rho)\psi I+[(1-\rho)\psi\varepsilon+\pi][\rho\psi I+(1-q)\eta I]}{h_1h_2-h_1\rho\psi\varepsilon}\\ h_1h_2-h_1\rho\psi\varepsilon &< (1-\rho)\psi h_2+\rho\psi\pi+(1-\rho)\psi\varepsilon(1-q)\eta+(1-q)\eta\pi\\ 1 &< \frac{\psi[\rho(h_1\varepsilon+\pi)+(1-\rho)(h_2+(1-q)\varepsilon\eta)]}{h_1h_2-(1-q)\eta\pi}=R_0\end{aligned} \tag{4.6}$$

□

Thus there exists a unique endemic equilibrium when $R_0 > 1$.

Local Stability analysis of the Endemic Equilibrium

We study the local stability of the endemic equilibrium by applying the Routh-Hurwitz criterion [54].

Theorem 2. *If $R_0 > 1$ then the endemic equilibrium E^e of system 2.3 is locally asymptotically stable in $\mathcal{G}$.*

Proof. Consider the Jacobian matrix at endemic equilibrium denoted by J_{E^e}.

$$J_{E^e}=\begin{bmatrix}-\overline{\lambda}-\mu & 0 & 0 & \delta\\ (1-\rho)\overline{\lambda} & -h_1 & \pi & 0\\ \rho\overline{\lambda} & (1-q)\eta & -h_2 & 0\\ 0 & q\eta & \beta & -\mu-\delta\end{bmatrix}$$

where $\overline{\lambda}$ is defined as the force of infection at endemic equilibrium. We obtain a characteristic equation $P(\lambda) = |\lambda I - J_{E^e}|$ where I is a 4×4 unit matrix. So that the characteristic equation becomes, $P(\lambda) = \lambda^4 + a_1\lambda^3 + a_2\lambda^2 + a_3\lambda + a_4$ Hence from Routh-Hurwitz criterion, we have the matrix

$$\begin{bmatrix} 1 & a_2 & a_4 & \lambda^4 \\ a_1 & a_3 & 0 & \lambda^3 \\ a_2 - \frac{a_3}{a_1} & a_4 & 0 & \lambda^2 \\ a_3 - \frac{a_1 a_4}{\left(a_2 - \frac{a_3}{a_1}\right)} & 0 & 0 & \lambda \\ a_4 & 0 & 0 & 1 \end{bmatrix}$$

where,

$$\begin{aligned} a_1 &= 2\,\mu + \delta + h_2 + h_1 + \overline{\lambda} \\ a_2 &= 2\,h_2\mu + 2\,h_1\mu + \mu\,\overline{\lambda} + \mu^2 + h_2\delta + h_1\delta + \delta\,\overline{\lambda} + \delta\,\mu - \eta\,\pi + \eta\,\pi\,q + h_2h_1 + h_2\overline{\lambda} + h_1\overline{\lambda} \\ a_3 &= (\mu + \delta)\,(h_1 + h_2)\left(\mu + \overline{\lambda}\right) + (h_2h_1 - (1-q)\,\eta\,\pi)\left(\overline{\lambda} + \delta + 2\,\mu\right) - \overline{\lambda}\delta\,\left((1-\rho)\,q\eta + \beta\,\rho\right) \\ a_4 &= (1-\rho)\,(q\pi - h_2q - (1-q)\,\beta)\,\eta\,\overline{\lambda}\delta - \left(\mu + \overline{\lambda}\right)(\mu + \delta)\,(\eta\,\pi - h_2h_1) \\ &\quad -\beta\,\rho\,\overline{\lambda}\delta\,h_1 + \eta\,\pi\,\mu\,q\left(\mu + \overline{\lambda} + \delta\right) \end{aligned}$$

According the the Routh-Hurwitz criterion, For $R_0 > 0$, the endemic equilibrium (E^e) is locally asymptotically stable if $a_1 > 0, a_2 - \frac{a_3}{a_1} > 0, a_3 - \frac{a_1 a_4}{\left(a_2 - \frac{a_3}{a_1}\right)} > 0$ and $a_4 > 0$. □

Bifurcation analysis

A bifurcation is a qualitative change in the nature of the solution trajectories due to a parameter change. The point at which this change take place is called a bifurcation point. At the bifurcation point, a number of equilibrium points, or their stability properties, or both, change. When $R_0 < 1$, the infectious disease will not invade the population unless otherwise. We prove using the Center Manifold theorem the possibility of bifurcation at $R_0 = 1$.

Let $S = x_1$, $I = x_2$, $C = x_3$ and $R = x_4$, so that $N = x_1 + x_2 + x_3 + x_4$. Then (2.3) is re-written in the form:

$$\left.\begin{aligned} \tfrac{dx_1}{dt} &= f_1 = \nu + \delta x_4 - \psi\tfrac{x_2}{x_1+x_2+x_3+x_4}x_1 - \psi\varepsilon\tfrac{x_3}{x_1+x_2+x_3+x_4}x_1 - \mu x_1 \\ \tfrac{dx_2}{dt} &= f_2 = (1-\rho)\psi\tfrac{x_2}{x_1+x_2+x_3+x_4}x_1 + (1-\rho)\psi\varepsilon\tfrac{x_3}{x_1+x_2+x_3+x_4}x_1 + \pi x_3 - h_1x_2 \\ \tfrac{dx_3}{dt} &= f_3 = \rho\psi\tfrac{x_2}{x_1+x_2+x_3+x_4}x_1 + \rho\psi\varepsilon\tfrac{x_3}{x_1+x_2+x_3+x_4}x_1 + (1-q)\eta x_2 - h_2x_3 \\ \tfrac{dx_4}{dt} &= f_4 = q\eta x_2 + \beta x_3 - (\mu+\delta)x_4 \end{aligned}\right\} \quad (4.7)$$

Suppose that we choose ψ_c as a bifurcation parameter. Using (4.1), we solve for ψ_c at $R_0 = 1$ as:

$$\psi_c = \frac{\mu^2 + \mu\,\alpha + \mu\,\eta + \Pi\,\mu + \Pi\,\alpha + \beta\,\mu + \beta\,\alpha + \beta\,\eta + \eta\,q\Pi}{\rho\,\epsilon\,\mu + \rho\,\epsilon\,\alpha + \mu + \Pi + \beta + \epsilon\,\eta - \epsilon\,\eta\,q - \rho\,\mu - \rho\,\beta + \rho\,\epsilon\,\eta\,q} \quad (4.8)$$

The liberalization matrix of (4.7) at a disease free Equilibrium (E^f) corresponding to $\psi = \psi_c$ is given by:

$$\mathcal{J}(E^f)\,|_{\psi=\psi_c} = J_{\psi_c} = \begin{pmatrix} -\mu & -\psi_c & -\psi_c\varepsilon & \delta \\ 0 & -h_1 & \pi & 0 \\ 0 & (1-q)\eta & -h_2 & 0 \\ 0 & q\eta & \beta & -(\mu+\delta) \end{pmatrix}.$$

Zero is a simple eigenvalue of J_{ψ_c} if $h_1 = \frac{\eta\pi(1-q)}{h_2}$. A right eigenvector w of J_{ψ_c} associated with the zero eigenvalues is given by $w = (w_1, w_2, w_3, w_4)^T$ where

$$\begin{aligned} w_1 &= \frac{w_3\left(\eta\,\delta\,\left(qh_2 + (1-q)\,\beta\right) - \psi\,\left(\mu+\delta\right)\left(h_2 + (1-q)\,\epsilon\,\eta\right)\right)}{\eta\,\left(1-q\right)\left(\mu+\delta\right)\mu}, \\ w_2 &= \frac{w_3 h_2}{\eta\,\left(1-q\right)}, \\ w_3 &= w_3, \\ w_4 &= \frac{w_3\left(qh_2 + (1-q)\,\beta\right)}{(1-q)\,(\mu+\delta)}. \end{aligned}$$

and a left eigenvector v of J_{ψ_c} corresponding to the zero eigenvalues is given by $v = (v_1, v_2, v_3, v_4)^T$ where

$$\begin{aligned} v_1 &= 0, \\ v_2 &= \tfrac{v_3 h_2}{\pi}, \\ v_3 &= v_3, \\ v_4 &= 0 \end{aligned}$$

We now reproduce a theorem stated by Castillo-Chavez and Song [55] .

Theorem 3. *[55]. Consider the following general system of ordinary differential equations with parameter ϕ*

$$\tfrac{dx}{dt} = f(x,\phi), f : R^n \times R \text{ and } f \in C^2(R^n \times R)$$

where 0 is an equilibrium point of the system (that is $f(0,\phi) \equiv 0$ for all ϕ) and assume:

1. *$A = D_x f(0,0) = (\frac{\partial f_i}{\partial x_j}(0;0))$ is the linearization matrix of the system around the equilibrium point 0 with ϕ evaluated at 0.*
2. *Zero is a simple eigenvalue of A and all other eigenvalues of A have negative real parts.*
3. *Matrix A has a right eigenvector w and a left eigenvector v corresponding to the zero eigenvalue.*

Let f_k be the k^{th} component of f and

$$a = \sum_{k,i,j=0}^{n} v_k w_i w_j \frac{\partial^2 f_k}{\partial x_i \partial x_j}(0,0), \tag{4.9}$$

$$b = \sum_{k,i=0}^{n} v_k w_i \frac{\partial^2 f_k}{\partial x_i \partial \phi}(0,0). \tag{4.10}$$

Then the local dynamics of system 1 around the $x = 0$ are totally determined by a and b. In particular,

1. *$a > 0, b > 0$, when $\phi < 0$ with $||\phi|| \ll 1$, $(0,0)$ is locally asymptotically stable and there exists a positive unstable equilibrium; when $0 < \phi \ll 1$, $(0;0)$ is unstable and there exists a negative and locally asymptotically stable equilibrium.*

2. *$a < 0, b < 0$, when $\phi < 0$ with $||\phi|| \ll 1$, $(0,0)$ is unstable; when $0 < \phi \ll 1$, $(0;0)$ is locally asymptotically stable and there exists exists a positive unstable equilibrium.*

3. *$a > 0, b < 0$, when $\phi < 0$ with $||\phi|| \ll 1$, $(0,0)$ is unstable and there exists locally asymptotically stable equilibrium; when $0 < \phi \ll 1$, $(0,0)$ stable and positive unstable equilibrium appears.*

4. *$a < 0, b > 0$, when ϕ changes from negative to positive, $x = 0$ changes its stability from stable to unstable. Correspondingly, a negative unstable equilibrium becomes locally asymptotically stable*

Algebraic calculations from Theorem 3 are shown in the working below.

$$\begin{array}{ll}
\frac{\partial^2 f_2}{\partial x_2 \partial x_2} = -2\,\frac{(1-\rho)\psi}{x_1}, & \frac{\partial^2 f_3}{\partial x_2 \partial x_2} = -2\,\frac{\rho\,\psi}{x_1} \\
\frac{\partial^2 f_2}{\partial x_2 \partial x_3} = (1-\rho)\,\psi\,x_1\left(-\frac{1}{{x_1}^2} - \frac{\epsilon}{{x_1}^2}\right), & \frac{\partial^2 f_3}{\partial x_2 \partial x_3} = \rho\psi\,x_1\left(-\frac{1}{{x_1}^2} - \frac{\epsilon}{{x_1}^2}\right) \\
\frac{\partial^2 f_2}{\partial x_2 \partial x_3} = -\frac{(1-\rho)\psi}{x_1}, & \frac{\partial^2 f_3}{\partial x_2 \partial x_3} = -\frac{\rho\psi}{x_1} \\
\frac{\partial^2 f_2}{\partial x_3 \partial x_3} = -2\,\frac{(1-\rho)\psi\,\epsilon}{x_1}, & \frac{\partial^2 f_3}{\partial x_3 \partial x_3} = -2\,\frac{\rho\psi\,\epsilon}{x_1} \\
\frac{\partial^2 f_2}{\partial x_3 \partial x_4} = -\frac{(1-\rho)\psi\,\epsilon}{x_1} & \frac{\partial^2 f_3}{\partial x_3 \partial x_4} = -\frac{\rho\psi\,\epsilon}{x_1} \\
\frac{\partial^2 f_2}{\partial x_2 \psi_c} = 1-\rho, & \frac{\partial^2 f_3}{\partial x_2 \psi_c} = \rho \\
\frac{\partial^2 f_2}{\partial x_3 \psi_c} = (1-\rho)\epsilon, & \frac{\partial^2 f_3}{\partial x_3 \psi_c} = \rho\epsilon
\end{array}$$

Note: $\frac{\partial^2 f_k}{\partial x_i \partial x_j} = \frac{\partial^2 f_k}{\partial x_j \partial x_i}$

The rest of the second derivatives that in (4.9) and (4.10) are zero. Hence,

$$\begin{array}{rcl} a & = & \left(-2\,\frac{v_3 {w_3}^2 \psi}{\pi\,x_1}\right)(\rho\,\pi + (1-p)\,h_2)\left(\frac{(h_2+(1-q)\epsilon\,\eta)(h_2\eta\,q+(1-q)\eta\,\beta+(\mu+\delta)(h_2+(1-q)\eta))}{\eta^2(-1+q)^2(\mu+\delta)}\right), \\ a & < & 0. \end{array} \tag{4.11}$$

$$\begin{array}{rcl} b & = & \frac{v_3 {h_2}^2 w_3(1-\rho)}{\pi\,\eta\,(1-q)} + \frac{v_3 h_2 w_3((1-\rho)\epsilon)}{\pi} + \frac{v_3 w_3 h_2 \rho}{(1-q)\eta} + v_3 w_3\,(\rho\,\epsilon)\,, \\ b & > & 0. \end{array} \tag{4.12}$$

Using the results in Theorem 3, the results in (4.11) and (4.12) indicate that there is a forward bifurcation at $\psi = \psi_c$ and there exist at least one stable endemic equilibrium when $R_0 > 1$.

5. Numerical Simulation

To observe the dynamics of pneumonia model over time, numerical simulations are done using MAPLE 14.0. The parameters in table 5.1 are based on the data of children under

five years of age and are used in the simulation. Some parameters have been derived from epidemiological literature and WHO database while other parameters have been allowed to vary within the possible intervals.

Table 5.1: Parameter Value

Parameter	Value	Source
ν	μN_0	[47]
κ	1-10 per day	Estimated
$\mathcal{P}$	0.89 to 0.99	[47]
ψ	$\kappa\mathcal{P}$	Expressed as in (2.2)
ϵ	0.001124	[47]
ρ	0.338	[59, 60]
π	0.00274 to 0.01096 per day	[47]
η	0.0238 to 0.0476 per day	[58]
q	0.5 to 1	[47]
α	0.33	Estimated
δ	0.2	Estimated
μ	0.0002 per day	[57]
β	0.0115	[56]

Using the parameter values, the numerical simulations show that a transcritical (forward) bifurcation is likely to occur at $\psi = \psi_c = 0.47$, ($R_0 = 1$) where there is only one stable equilibrium point if $R_0 < 1$ (disease-free equilibrium) and a low endemicity when R_0 is slightly above one (See Figure 5.1 (a) and (b)). This is important to conclude that there can only be one stable endemic equilibrium when $R_0 > 1$. In models with multi-group infectious classes, forward bifurcations commonly exist [61]. This could be the reason for the existence of a forward bifurcation for pneumonia transmission dynamic.

The mathematical technique involved in determining the global stability of the endemic equilibrium is quite complicated and therefore in this paper we determine the global stability of the endemic equilibrium using a numerical simulation (See Figure 5.2). We observe from Figure 5.2 that starting with any number of infected individuals with the initial population $N_0 = 100$, the number of the susceptible and the infected will always converge to a stable value ($S^e = 24.41243257, I^e = 4.549013989 \times 10^{-2}$). Assuming that we reduce the transfer rates between the Carriers and the Infected with the aim or reducing R_0, the infected population also reduces.

Using MAPLE 14, data for the infected population in both cases (dotted line and continuous line in Figure 5.3) was generated and checked for any significant difference in the two populations. Table 5.2 shows the results of the statistical analysis. Since the P-value is equal to 2.1344×10^{-15} for the t-test for mean difference is less than 0.01, we conclude that there is a strong significant effect of reducing the rates of transfer between the carriers and the infected on reducing the infected populations. We also simulate the effect of different proportions of carriers on transmission by considering different initial proportions of carriers and different rates of transfer leading to increase in carrier proportion in the population.

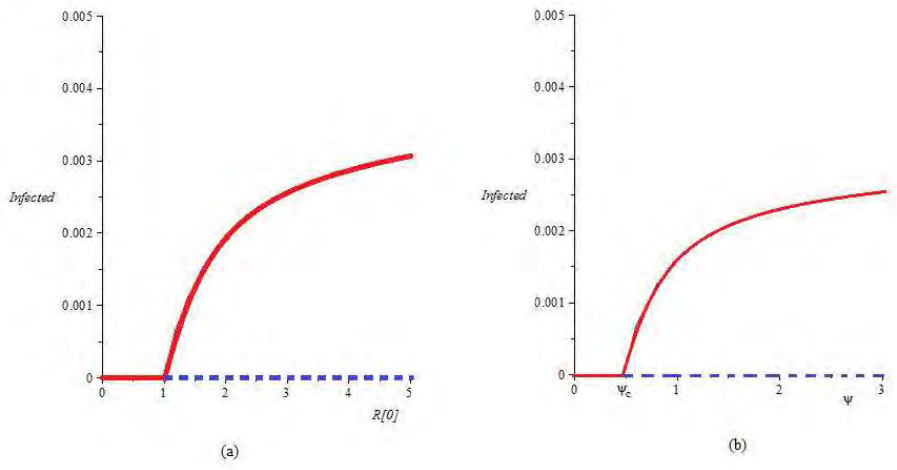


Figure 5.1: Forward bifurcation diagram in (a) plane I, R_0 and in (b) plane I, ψ. The continuous line represents a stable equilibrium. There are two stable equilibriums (disease free equilibrium for $R_0 < 1$ and an endemic equilibrium for $R_0 > 1$). The dotted line represents the unstable disease free equilibrium

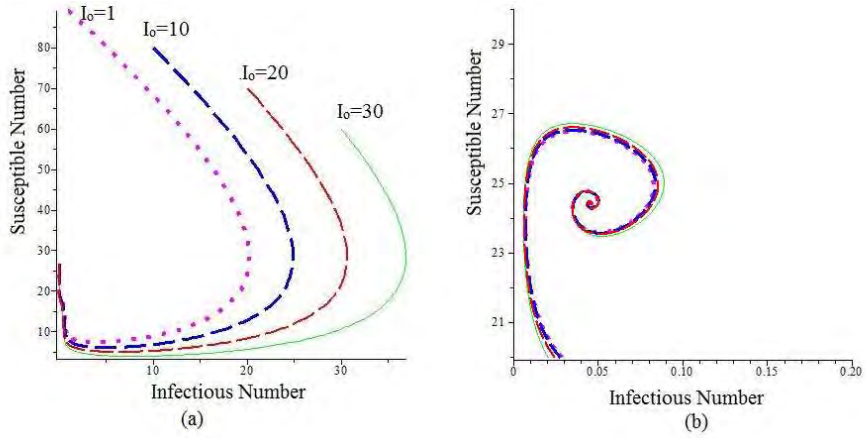


Figure 5.2: The phase plane portrait of S vs I for 2.3. The four curves correspond to the initial conditions I_0= 1, 10, 20 and 30 respectively. They all converge at $I^e = 4.549013989$ and $S^e = 24.41243257$ as in (b) when the plot is magnified. showing global asymptotic stability of the endemic equilibrium

6. Interpretation of the model and Biological Implication

The results from the analysis of the model indicate that a possible disease control strategy would be to reduce the number of new secondary infections i.e. reducing the value of the basic reproduction number, R_0. Rewriting (4.1) into,

$$R_0 = \kappa\mathcal{P}\left(\frac{\rho[\varepsilon(\mu+\alpha+\eta)+\pi]+(1-\rho)[\mu+\beta+\pi+(1-q)\varepsilon\eta]}{(\mu+\alpha+\eta)(\mu+\beta+\pi)-(1-q)\pi\eta}\right),$$

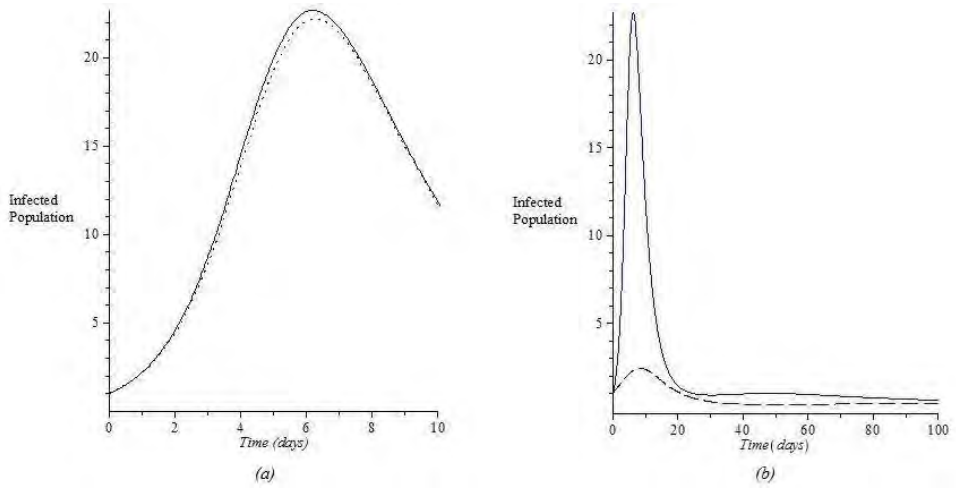


Figure 5.3: (a) Dynamics considering carrier-infected interaction rates. The continuous line is plotted when $\pi = 0.005$, and $q = 0.75$, while the dotted line is plotted when $\pi = 0$ and $q = 0.999$. (b) Dynamics considering recovery rate of the infected and carrier-infected transfer rates. The continuous line is plotted when $\eta = 0.03$,$\pi = 0.005$ and $q = 0.75$ while The dotted line is plotted when $\eta = 0.6$, $\pi = 0$ and $q = 0.999$

Statistics	Infected Population ($\pi = 0.005$, $q = 0.75$)	Infected population ($\pi = 0$, $q = 0.999$)
Mean	2.486300321	2.170599263
Variance	20.90683033	20.72202111
Observations	100	100
Hypothesis	Mean Difference = 0	
df	99	
t Statistics	9.408577558	
P- Value	2.1344×10^{-15}	
t Critical	1.9842169	

Table 5.2: t-Test: Paired Two Sample for Means (testing for the significant different between infected populations when the values of π β and q are varied respectively)

it is evident that R_0 is directly proportional to the contact rate κ and to the mean time spent in the diseased classes, $\frac{1}{(\mu+\alpha+\eta)(\mu+\beta+\pi)-(1-q)\pi\eta}$. The implication of reducing the contact rate $\kappa \to 0$ and mean time spent in the diseased classes ensures that $R_0 \to 0$. It is possible to reduce mean time spent in the diseased classes when the transfer rates between the carrier and the infected classes are reduced (i.e. $\pi \to 0$ and $q \to 1$) and when the transfer rate out from the diseased classes are increased (i.e. $\alpha, \beta \to \infty$). This indicates that quarantine (where possible), prompt and effective diagnosis and treatment of the carriers and the infected individuals may lead to possible reduction of the new infections to zero. A justification of controlling pneumonia by reducing the R_0 is indicated by the forward bifurcation results in the analysis. The presence of a forward

bifurcation implies that a disease can be cleared from the population by just reducing the R_0.

7. Discussion

Mathematical models of infectious diseases have been used to successfully explain the transmission dynamics of many diseases and the use of such models has grown exponentially from mid 20th century [62]. Our main aim in this paper was to provide a mathematical explanation of pneumonia transmission dynamics, taking into consideration the role of carriers and recovery measures in the transmission. We only considered bacteremic pneumonia since it is the most common among children who are under five years of age.

The model that we have discussed here is based on the initial model that was studied by Doura et al. [47]. When studying the transmission dynamics of infectious diseases with an aim of suggesting control measures, it is natural to consider the stability of equilibrium points and possibility of bifurcation. In this paper we have established R_0, existence and stability of the equilibrium points and existence of bifurcation points. Our main results indicates that when $R_0 < 1$ then the disease free equilibrium is stable. It becomes unstable when $R_0 > 1$. The local stability of the endemic equilibrium point E^e changes its nature to unstable when ψ crosses the critical value ψ_c via a forward bifurcation. This is a clear indication that the effective control measure for pneumonia is achieved when R_0 is reduced.

Most of the results in this paper are in agreement with those of [47]. However, we find some interesting results in the numerical simulation: reducing the transfer rates between the carrier and the infected class reduces prevalence of the disease. This is a control strategy that can be employed for pneumonia dynamics.

Bibliography

[1] A. Brueggemann and V. G. Doern, Resistance Among Streptococcus pneumoniae: Patterns, Mechanisms, Interpreting the Breakpoints, *The American Journal of Managed Care*, **6**(23), 1189-1196, 2000.

[2] J. Anthony, G. Scott, W. Abdullah, P. Walik, H. Douglas and M. Kim, Pneumonia research to reduce childhood mortality in the developing world, *Journal of Clinical Investigation*, **118**(4), 1291-1300, 2008.

[3] P. C. Appelbaum, Resistance among Streptococcus Pneumoniae: Implications for Drug Selection, *Oxford Journals of Clinical Infectious Diseases*, **34**(12), 1613-1620, 2003.

[4] B. Ashby and C. Turkington, The Encyclopedia of Infectious Diseases (3rd Ed), `http://books.google.co.uk/books?id=4Xlyaipv3dIC/pg=PA242.`, 2007.

[5] B.M. Gray and G.M. Converse and J. Dillon, Epidemiologic Studies of Streptococcus pneumoniae in Infants: Acquisition, Carriage, and Infection during the First 24 Months of Life, *The Journal of Infectious Diseases*, **142**(6), 923-933, 1980.

[6] R.E. Black and S. Moris and J. Bryce, Where and why are 10 Million Children Dying Every Year, *Lancet*, **361**(9376), 2226-2234, 2003.

[7] S. Cousens, E.R. Black, L. H. Johnson, I. Rudan, D. G. Bassani, R. Cibulskis, Global, Regional, and National Causes of Child Mortality in 2008: A Systematic Analysis, *Lancet*, **375**, 1969-1987, 2010.

[8] B. Greenwood, The Epidemiology of Pneumococcal Infection in Children in the Developing World, The Royal Society, **354**, 777-785, 1999.

[9] A.B. Brueggemann, D.T. Griffiths, E. Meats, T. Peto, D.W. Crook and B.G. Spratt, Clonal Relationships between Invasive and Carriage Streptococcus pneumoniae and Serotype- and Clone-Specific Differences in Invasive Disease Potential, The Journal of Infectious Diseases, **187**(9), 1424-1432, 2003.

[10] B. H. Singer and D. E. Kirschner, Influence of Backward Bifurcation On Interpretation Of R_0 in A Model of Epidemic Tuberculosis With Reinfection, Mathematical Biosciences and Engineering, **1**(1), 81-93, 2004.

[11] , J. Bryce, C. B. Pinto, K. Shibuya and R. Black, Childhood Pneumonia in Developing Countries, *WHO Child Health Epidemiology Reference Group*, **365** (9465), 1147-1152, 2005.

[12] J.C. Butler, R.F. Breiman, J.F. Campbell, H.B. Lipman, C.V. Broome and R.R. Facklam, Pneumococcal Polysaccharide Vaccine Efficiency. An Evaluation of Current Recommendation, *JAMA*, **270**(15), 1826-1831, 1993.

[13] C. Colijn, T. Cohen and M. Murray, Mathematical Model of Tuberculiosis: Accomplishment and Future Challenges, *World Scientific Connecting Great Minds*, 2006.

[14] Centers for Disease Control, Centers for Disease Control and Prevention :Pneumococcus: Questions and AnswersInformation about the disease and vaccines, `www.immunize.org/catg.d/p4213.pdf`, 2010.

[15] Centers for Disease Control, Centers for Disease Control and Prevention :Drug-Resistant Streptococcus Pneumoniae (DRSP) Disease, `http://www.cdc.gov/ncidod/dbmd/diseaseinfo/drugresisstreppneum.t.htm`, 2010.

[16] C. Davis, Pneumoniae, `http://www.medicinenet.com/pneumonia/article.htm2howdo.`, 2010.

[17] P. Coen, T. Heath, A. Barbuor and G. Garnett, Mathematical model of haemophilus influenzae type b, *Epidemiology and infection*, **120** (3), 195-281, 1998.

[18] D. Pessoa, *Modelling the Dynamics of Streptococcus Pneumoniae Transmmission in Children*, University of De Lisboa, Masters Thesis, 2010.

[19] L. Dunn, Pneumonia: Classification, Diagnosis and Nursing Management, Nursing Standard (Royal College of Nursing), **19**(42), 50-54, 2005.

[20] C. Evagelia, A. Gerard, J.Morris, J. Albert and R.C. Stephen, Metabolomic Analysis in Severe Childhood Pneumonia in Teh Gambia, West Africa, Journal of Plos One, **5**(9), 2010.

[21] W. Fonseca, B. Kirkwood, C. Victora and S. Fuchs, Risk factors for childhood pneumonia among the urban poor in Fortaleza, Brazil: a case-control study, *World Health Organization*, 1996.

[22] G. Schiffman and C. Melissa, Pneumonia, `http://www.medicinenet.com/pneumonia/article.htm2howdo`, 2010.

[23] K. Ghosh and G.B. Clements, Surveillance of Mycoplasma Pneumoniae Infections in Scotland 1986-1991, *Journal of Infectious Diseases*, **25**(2), 221-227, 1992.

[24] , G. McKenzie, The Pneumococcus Carrier, *The Mary Imogene Bassett Hospital*, 88-100, 1999.

[25] I. Rudan, C. Boschi-Pinto, Z. Biloglav, K. Mulholland and H. Campbell, Epidemiology and Entomology of Childhood Pneumonia, *Bulletin of the World Health Organization*, 2011.

[26] , M.R. Jacobs, H.J. Koornhof, R.M. Browne, Emergence of Multiply Resistant Pneumococci. N Engl, *J Med*, **299**, 735-740, 1778.

[27] , J. Starr, G. Fox and J. Clayton, Streptococcus Pneumoniae: An Update on Resistance Patterns in the United States, *Journal of Pharmacy Practice*, **21**(5), 363-370, 2008.

[28] L. Niessen, A. Hove, H. Hilderink, M. Weber, K. Mulholland and M. Ezzati, Comparative Impact Assessment of Child Pneumonia Interventions, *Bulletin of the World Health Organization*, **87**, 472-480, 2009.

[29] M. Ekman, T. Grayston, R. Visakorpi, M. Kleemola, C. Kuo and P. Saikku, An Epidemic of Infections Due to Chlamydia pneumoniae in Military Conscripts, *Clinical Infectious Diseases*, **17**(3), 420-425, 1993.

[30] M. Darboe, A. Fulford, O. Secka and A. Prentice, The dynamics of nasopharyngeal streptococcus Pneumoniae carriage among rural Gambian Mother-child pairs, *BioMed Central Infectious Diseases*, **10**(195), 1471-2334, 2010.

[31] J.Y.T. Mugisha, Lecture Notes in Biomathematics: A Course conducted at Nelson Mandela Metropolitan University, Department of Mathematics, Makerere University, 2009.

[32] S. Obaro and R. Adegbola, The Pneumococcal: Carriage, Disease and Conjugate Vaccine, *J Med Microbiol*, **51**(2), 98-104, 2002.

[33] J. Nuorti and C. Whitney, Prevention of Pneumococcal Disease Among Infants and Children - Use of 13-Valent Pneumococcal Conjugate Vaccine and 23-Valent Pneumococcal Polysaccharide Vaccine, Division of Bacterial Diseases, National Center for Immunization and Respiratory Diseases, Recommendations of the Advisory Committee on Immunization Practices (ACIP), no. 59, 2010.

[34] E. C. Rosenow, Studies in Influenza and Pneumonia: Study X. The Immunologic Properties of the Green- Producing Streptococci from Influenza,*The Journal of Infectious Diseases*, **26**(6), 597-613, 1920.

[35] K. Todar, Streptococcus, http://textbookofbacteriology.net/S.pneumoniae.html, 2011.

[36] , United Nations, Milleniam Development Goals, Available from http://www.un.org/millenniumgoals/, 2011.

[37] E. Wiah and A. Adetunde, A mathematical model of Cerebrospinal Meningitis Epidemic: A case study for Jirapa District, Ghana, *KMITL Sci Tech. J.*, **10**(2), 2010.

[38] K. Kallander, J. Nsungwa-Sabiiti, A. Balyeku, G. Pariyo, G. Tomson and S.Peterson, Home and community management of acute respiratory infections in children in eight Ugandan districts, *Annals of Tropical Paediatrics*, **25**, 283-291, 2005.

[39] D. Kalajdzievska and M. Li, Modeling the effect of carriers on transmission daynamicsof infectious diseases, *Mathematical Bioscience and engineering*, **8**(3), 711-722, 2011.

[40] K. Kallander, H. Hildenwall, P. Waiswa, E. Galiwango, S. Peterson and G. Pariyo, Delayed care seeking for fatal pneumonia in children aged under five years in Uganda: a case-series study, *Bulletin of the World Health Organization*, **86** (5), 321-416, 2008.

[41] Statement of Alex Palacios on the Global Alliance for Vaccines and Immunizations (GAVI) Alliance, 2011.

[42] D. T. Jamison, World Bank and Disease Control Priorities Project, Disease control priorities in developing countries, Part 611 Stand Alone Series, *World Bank Publications*, 2006, 2nd Ed., http://books.google.co.ke/books?id=Ds93H98Z6D0C, ISBN: 9780821361795, Stand Alone Series, 485-497, 2006.

[43] Abdul Ghafoor, Nasreen Khalid Nomani, Zahid Ishaq and Sohail Zahoor Zaidi, Diagnosis of Acute Lower Respiratory Tract Infections in Children in Rawalpindi and Islamabad, Pakistan, *Reviews of Infectious Diseases*, **12** (8), S907-S914. Available from http://www.jstor.org/stable/4455727, 1990.

[44] Catchup Study Group, Clinical Efficacy of Co-Trimoxazole versus Amoxicillin Twice Daily for Treatment of Pneumonia: A Randomized Controlled Clinical Trial in Pakistan, *Archives of Disease in Childhood*, **86**, 2002.

[45] R.A. Adegbola, S.O. Usen, M. Weber, N. Lloyd-Evans and K. Jobe, Haemophilus influenzae Type B Meningitis in The Gambia after Introduction of a Conjugate Vaccine, *Lancet*, **354**(9184), 1091-1092, 1999.

[46] N. W. Schluger, Vaccinations Bacterial, for Pneumonia, *Encyclopedia of Respiratory Medicine*, 389-392, 2006.

[47] K. Doura, D. Malendez-Morales, G. Mayer and L.E Perez, An S-I-S Model of Streptococcal Disease with a Class of Beta Hemolytic Carriers, 2000.

[48] , K.E. Lamb, D. Greenhalgh and C. Robertson, A simple mathematical model for genetic effects in pneumococcal carriage and transmission, *Journal of Computational and Applied Mathematics*, **235**, 1812 - 1818, 2011.

[49] D. Greenhalgh, K.E. Lamb and C. Robertson, A mathematical model for the spread of Streptococcus pneumoniae with transmission dependent on serotype, *Journal of Biological Dynamics*, iFirst, 1-16, 2011.

[50] Carlos Castillo-Chavez and Zhilan Feng, To Treat or not to Treat: The Case of Tuberculosis, *Journal of Mathematical Biology*, **35**, 629 - 656, 1995.

[51] H. Tacio, Pneumonia: What you should know about it, Sun.Star Davao newspaper, 2011. Available from `http://www.sunstar.com.ph/davao/lifestyle/pneumonia-what-you-should-know-about-it`

[52] Division of Disease Control, What Do I Need To Know: Streptococcus Pneumoniae, North Dakota: Department of Health, `http://www.ndhealth.gov/Disease/Documents/faqs/Pneumonia.pdf`, 2006.

[53] P. van den Driessche and J. Watmough, Reproduction Numbers and Sub-threshold Endemic Equilibria for Compartmental Models of Disease Transmission, *Mathematical Biosciences*, **180**, 29-48, 2002.

[54] Lina, Tayseer, Yves, Isaac, The Routh-Hurwitz Criterion and Application to the Stability of the Viral Free Steady State, *Biomathematics Lecture Notes*, *unpublished*, 2005.

[55] C Castillo-Chavez and B Song, Dynamical Models of Tuberculosis and their Applications, *Mathematical Biosciences and Engineering*, **1**(2), 361-404, 2004.

[56] P. Hill, Y. Cheung, A. Akisanya, K. Sankareh, G. Lahai, B. Greenwood and R. Adegbola, Nasopharyngeal Carriage of Streptococcus Pneumoniae in Gambian Infants: A Longitudinal Study, Clinical Infectious Disease, **46**(6), 807-814, 2008.

[57] World Health Organization, *Global Health Observatory Data Repository*, `http://apps.who.int/ghodata/?vid=1320`, 2010.

[58] R. Thadani, Pneumonia Recovery Time, `http://www.buzzle.com/articles/pneumonia-recovery-time.html`, 2011.

[59] D. Kateete, H. Kajumbula, K. Deogratias and A. Ssevviri, Nasopharyngeal Carriage Rate of Streptococcus Pneumoniae in Ugandan Children with Sickle Cell Disease, BMC Research Notes, **5**(28), 321-416, 2012.

[60] Y. Kimura, S. Kotami and Y. Siokawa, Recent Advances in Streptococci and Streptoccocal Diseases, *Bracknell*, 1984.

[61] J. Dushoff, W. Huang, C.Carlos, Backwards Birfurcations and Catastrophe in Simple Model of Fatal Diseases, *Journal of Mathematical Biology*, **36**, 227-248, 1998.

[62] W. Hethcote, The Mathematics of Infectious Diseases, *Society for Industrial and Applied Mathematics*, **42**(4), 599-653, 2000.

A New Connection between Irreducible and Extended Irreducible Goppa Codes

John Ryan

Department of Mathematics, Mzuzu University,
P/Bag 201, Mzuzu 2, Malawi.

Abstract. We identify sets of irreducible Goppa codes which become equivalent when extended by a parity check bit.

Keywords: classical Goppa codes, extended codes, irreducible Goppa codes, equivalent codes

1. Introduction

We begin by defining a degree r irreducible Goppa code $\Gamma(\mathbf{L}, g)$ over $\mathbb{F}_q$ of length q^n in terms of a single field element α of degree r over $\mathbb{F}_{q^n}$. We then define the extended code $\overline{\Gamma(\mathbf{L}, g)}$. We give the well known sufficient conditions on two elements α and β of degree r for

1. the corresponding irreducible Goppa codes to be equivalent;
2. the corresponding extended irreducible Goppa codes (the Goppa codes defined by α and β extended by a parity check bit) to be equivalent.

We then give a relationship between a set of all elements related by conditions in 2) to a bunch of sets of elements related by the conditions in 1) above.

2. Preliminaries

Let q be a power of a prime number, let $\mathbb{F}_q$ be the field of order q, $\mathbb{F}_{q^n}$ its extension of order n. The family of Goppa codes was first introduced by V.D. Goppa in 1970 [2]. For our purposes, in this paper we focus on irreducible Goppa codes, and define irreducible Goppa codes as follows.

Definition 10. Let $g(z) \in \mathbb{F}_{q^n}[z]$ be irreducible of degree r and let $L = \mathbb{F}_{q^n} = \{\zeta_i : 0 \leq i \leq q^n - 1\}$. Then the irreducible Goppa code $\Gamma(L, g)$ is defined as the set of all vectors $\underline{c} = (c_0, c_1, \ldots, c_{q^n-1})$ with components in $\mathbb{F}_q$ which satisfy the condition

$$\sum_{i=0}^{q^n-1} \frac{c_i}{z - \zeta_i} \equiv 0 \bmod g(z).$$

The polynomial g is called the Goppa polynomial. Since g is irreducible over $\mathbb{F}_{q^n}$, g does not have any root in L and the code is called an *irreducible Goppa code*. Since $g(z)$ is of degree r the code $\Gamma(L, g)$ is called a *Goppa code of degree* r. In this paper $g(z)$ is always irreducible of degree r over $\mathbb{F}_{q^n}$.

3. Extended Irreducible Goppa codes

Definition 11. Let $\Gamma(\mathbf{L}, g)$ be a code of length q^n over $\mathbb{F}_q$. Then the extended code $\overline{\Gamma(\mathbf{L}, g)}$ is defined by

$$\overline{\Gamma(\mathbf{L}, g)} = \{(c_0, c_1, \cdots, c_{q^n}) : (c_0, c_1, \cdots, c_{q^n-1}) \in \Gamma(\mathbf{L}, g) \text{ and } \sum_{i=0}^{q^n} c_i = 0\}.$$

The extended code $\overline{\Gamma(\mathbf{L}, g)}$ is often described as the code obtained from $\Gamma(\mathbf{L}, g)$ by adding a parity check bit to each codeword of $\Gamma(\mathbf{L}, g)$.

4. Irreducible Goppa codes defined by a field element

It is shown in [3] that if α is any root of the Goppa polynomial g then $\Gamma(\mathbf{L}, g)$ is completely described by any root α of $g(z)$ and a parity check matrix $\mathbf{H}(\alpha)$ is given by

$$\mathbf{H}(\alpha) = \begin{pmatrix} \frac{1}{\alpha - \zeta_0} & \frac{1}{\alpha - \zeta_1} & \cdots & \frac{1}{\alpha - \zeta_{q^n-1}} \end{pmatrix} \tag{4.1}$$

where $L = \mathbb{F}_{q^n} = \{\zeta_i : 0 \leq i \leq q^n - 1\}$. We may denote this code by $\mathbf{C}(\alpha)$.

Note that $\overline{\mathbf{C}}(\alpha)$ denotes the same code as $\overline{\Gamma(\mathbf{L}, g)}$, where $g(\alpha) = 0$. Also, by using the parity check matrix $H(\alpha)$ to define $C(\alpha)$, we are implicitly fixing an order on L and, consequently, an order on the components of the codewords in the code $C(\alpha)$.

Considering that any Goppa code (or extended Goppa code) can be defined by an element of degree r over $\mathbb{F}_{q^n}$, and conversely, any element of degree r defines a Goppa code (or extended Goppa code), we make the following definition.

Definition 12. The set $\mathbb{S}$ is the set of all elements in $\mathbb{F}_{q^{nr}}$ of degree r over $\mathbb{F}_{q^n}$.

5. Equivalent Goppa codes

Consider the map $\pi_{\zeta,\xi}$ defined on $\mathbb{S}$ by

$$\pi_{\zeta,\xi} : \alpha \mapsto \zeta\alpha + \xi$$

where $\zeta, \xi \in \mathbb{F}_{q^n}$
It is well known that if $\pi_{\zeta,\xi}(\alpha) = \beta$ then

$$(\Gamma(\mathbf{L}, g))\,\mathbf{C}(\alpha) \text{ is equivalent to } (\Gamma(\mathbf{L}, f))\,\mathbf{C}(\beta)$$

where g and f are the minimal polynomials of α and β respectively.

6. Equivalent Extended Irreducible Goppa codes

Consider the maps $\pi_{\zeta_1,\zeta_2,\xi_1,\xi_2}$ defined on $\mathbb{S}$ by

$$\pi_{\zeta_1,\zeta_2,\xi_1,\xi_2} : \alpha \mapsto \frac{\zeta_1\alpha + \xi_1}{\zeta_2\alpha + \xi_2}$$

for fixed ζ_j, ξ_j where $\zeta_j, \xi_j \in \mathbb{F}_{q^n}, j = 1, 2$. For simplicity, where there is no confusion, we write π for $\pi_{\zeta_1,\zeta_2,\xi_1,\zeta_2}$.

It is well known that if $\pi(\alpha) = \beta$ then $\overline{\mathbf{C}}(\alpha)$ is equivalent to $\overline{\mathbf{C}}(\beta)$. See [1]

Remark 8. Note that in the definition of π the scalars ζ_j and ξ_j are defined up to scalar multiplication. Hence we may assume that $\zeta_2 = 1$ or $\xi_2 = 1$ if $\zeta_2 = 0$.

In light of the foregoing, we make the following two definitions:

Definition 13. *$F1$ denotes the set of all maps $\{\pi_{\zeta,\xi} : \zeta, \xi \in \mathbb{F}_{q^n}\}$.*

Remark 9. It is not difficult to see that $F1$ together with the operation of composition of maps is a group which acts on the set $\mathbb{S}$.

Definition 14. *$F2$ denotes the set of all maps $\{\pi_{\zeta_1,\zeta_2,\xi_1,\xi_2} : \zeta_j, \xi_j \in \mathbb{F}_{q^n}, j = 1, 2, \zeta_1\xi_2 - \zeta_2\xi_1 \neq 0\}$.*

Remark 10. It is not difficult to see that $F2$ together with the operation of composition of maps is a group which acts on the set $\mathbb{S}$.

7. The New Connection

The new connection, spoken of in the title of this paper, is merely the observation that one orbit in $\mathbb{S}$ under the action of the group $F2$ (all elements of this orbit define equivalent extended Goppa codes) is made up of $q^n + 1$ orbits in $\mathbb{S}$ under the action of the the group $F1$. Specifically,

1. the orbit in $\mathbb{S}$ containing α under $F1$ is
$$A1(\alpha) = \{\zeta\alpha + \xi : \zeta \neq 0, \xi \in \mathbb{F}_{q^n}\};$$
2. the orbit in $\mathbb{S}$ containing α under $F2$ is
$$A2(\alpha) = \left\{ \frac{\zeta_1\alpha + \xi_1}{\zeta_2\alpha + \xi_2} : \zeta_j, \xi_j \in \mathbb{F}_{q^n}, j = 1, 2 \text{ and } \zeta_1\xi_2 - \zeta_2\xi_1 \neq 0 \right\}.$$

Our task now is to prove the following Theorem:

Theorem 1. *For any $\alpha \in \mathbb{S}$,*
$$A2(\alpha) = A1(\alpha) \cup A1\left(\frac{1}{\alpha}\right) \cup A1\left(\frac{1}{\alpha + \xi_1}\right) \cdots \cup A1\left(\frac{1}{\alpha + \xi_{q^n-1}}\right)$$
where $\mathbb{F}_{q^n} = \{0, \xi_1, \cdots \xi_{q^n-1}\}$

Proof. Considering the definitions of the sets $A1$ and $A2$, it is clear that $A1(\alpha) \subset A2(\alpha)$ and $A1(\frac{1}{\alpha+\xi_i}) \subset A2(\alpha)$ for any i. That is the right hand side of the equation in the Theorem is contained in the left hand side. It remains to show that both have the same number of elements. So we use a counting argument.

It is easy to see that there are $q^n + 1$ distinct sets on the right hand side and each of these sets contain $q^n(q^n - 1)$ elements. So the right hand side has a total of $(q^n + 1)(q^n)(q^n - 1)$ elements. On the left hand side
$$A2(\alpha) = \left\{ \frac{\zeta_1\alpha + \xi_1}{\zeta_2\alpha + \xi_2} : \zeta_j, \xi_j \in \mathbb{F}_{q^n}, j = 1, 2 \text{ and } \zeta_1\xi_2 - \zeta_2\xi_1 \neq 0 \right\}.$$

1. If $\zeta_1 = 0$, then w.l.o.g. $\xi_1 = 1$ and there are $q^n(q^n - 1) = q^{2n} - q^n$ possibilities
2. If $\zeta_1 = 1$, then we need to exclude the cases when $\xi_2 = \zeta_2\xi_1$:
   (a) $\xi_2 = 0$, then exclude
      i. the q^n cases when $\zeta_2 = 0$ and $\xi_1 \in \mathbb{F}_{q^n}$
      ii. the $q^n - 1$ cases when $\zeta_2 \neq 0 \in \mathbb{F}_{q^n}$ and $\xi_1 = 0$
   (b) $\xi_2 \neq 0$. There are $q^n - 1$ such cases. In each such case, for each $\xi_1 \neq 0$ (and there are $q^n - 1$ of them) there is a unique solution for ζ_2. Hence there are $(q^n - 1)^2$ possibilities when $\xi_2 \neq 0$.

   So the total number of possibilities under item 2 is $q^{3n} - (2q^n - 1) - (q^n - 1)^2 = q^{3n} - q^{2n}$

Adding the possibilities under 1) and 2) we get $q^{3n} - q^{2n} + q^{2n} - q^n = q^{3n} - q^n = (q^n + 1)(q^n)(q^n - 1)$. The proof is complete. □

Remark 11. The significance of this Theorem is that we have identified $q^n + 1$ sets of Goppa codes, all of which become equivalent when extended by a parity check bit. The author used $A1$ type sets to count the number of Goppa codes in a previous work [4]. This new connection will facilitate a similar count of extended Goppa codes.

Bibliography

[1] P. Berger, Goppa and Related Codes Invariant under a Prescribed Permutation, *IEEE Trans. Inform. Theory*, **46**, 2000.

[2] V.D. Goppa, Rational representation of codes and (L,g) codes, *Prob. Peredach. Inform*, **7**, 41-49, 1971.

[3] C.-L. Chen, Equivalent irreducible Goppa codes, *IEEE Trans. Inform. Theory*, **24**, 766-769, 1978.

[4] J.A. Ryan, *Irreducible Goppa Codes*, PhD Thesis, University College Cork, 2004.

List of SAMSA 2012 Participants

1. Ash Abebe, Department of Mathematics and Statistics, Auburn University, Alabama 36849-5310, U.S.A.

2. Leif Abrahamsson, International Science Programme, Uppsala University. Email: leifab@math.uu.se

3. Komi Afassinou, University of KwaZulu-Natal, School of Mathematics Sciences, Pietermaritzburg, South Africa

4. T. Basikolo, Department of Mathematical Sciences, Chancellor College, University of Malawi, P.O.Box 280, Zomba, Malawi. Email: thomasbasikolo@gmail.com

5. Erik Baurdoux, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom.

6. Ronald Becker, Department of Mathematical Sciences, University of Zululand, Private Bag X1001, Kwa-Dlangezwa, 3886, South Africa

7. B.T. Bekele, South African Centre for Epidemiological Modelling and Analysis, University of Stellenbosch, Stellenbosch, South Africa. Email: bekelebt@sun.ac.za

8. Jocelyne Bion-Nadal, CNRS - CMAP, Ecole Polytechnique, 91128 Palaiseau Cedex, France. Email: jocelyne.bion-nadal@cmap.polytechnique.fr, Web: `http://www.cmapx.polytechnique.fr/~bionnada/`

9. Madalitso Chamba, Consumer Education and Consumer Protection, Reserve Bank of Malawi. Email: mchamba@rbm.mw

10. Chiteng'a J. Chikunji, Department of Basic Sciences, Botswana College of Agriculture, Private Bag 0027, Gaborone, Botswana. Email: jchikunj@bca.bw

11. Thokozani Chimkono, The Malawi Polytechnic

12. Faraimunashe Chirove, University of KwaZulu-Natal, Pietermaritzburg, South Africa. Email: Chirovef@ukzn.ac.za

13. Busiso P. Chisala, Department of Mathematics, Mzuzu University, Malawi

14. Angwenyi Nyachae David, Department at Mathematics, University of Nairobi, Kenya

15. Levis Eneya, Department of Mathematical Sciences, Chancellor College, University of Malawi, P.O.Box 280, Zomba, Malawi. Email: leneya@cc.ac.mw

16. Grace Gachigua, Kimathi University College of Technology, Department of Mathematics, P.0. Box 657,00100, Nyeri, Kenya. Email: wambuigachigua@gmail.com

17. S.D. Hove-Musekwa, Department of Applied Mathematics, National University of Science and Technology, P. O. Box AC 939, Ascot, Bulawayo, Zimbabwe. Emails: senelani@aims.ac.za, sdmusekwa@nust.ac.zw

18. Peter Imkeller, HU Berlin. Email: imkeller@mathematik.hu-berlin.de

19. Overtoun M.G. Jenda, Department of Mathematics and Statistics, Auburn University, Alabama 36849-5310, U.S.A.
Email: jendaov@auburn.edu

20. Peter Johnson, Department of Mathematics and Statistics, Auburn University, Alabama 36849-5310, U.S.A. Email: johnspd@auburn.edu

21. E. Joseph, Department of Mathematics, University of Dar-es-Salaam, P.O.BOX 35062, Dar-es-Salaam, Tanzania

22. Ezekiel Kachisa, Malawi Institute of Education, Curriculum Development Department-Mathematics Section, Malawi. Email: ezekiel@live.ie

23. Tsirizani M. Kaombe, Science Department, University of Livingstonia - Laws Campus, P.O. Box 37, Livingstonia, Malawi. Email: tsirizanikaombe@yahoo.com

24. Elias Kaphesi, Department of Mathematics and Statistics, The Polytechnic, University of Malawi

25. Dun Kasoka, Malawi Polytechnc, University of Malawi

26. Tasokwa Kakota, Lilongwe University of Agriculture and Natural Resources

27. M. Kgosimore, Botswana college of Agriculture, Private Bag 0027, Gaborone, Botswana. Email: mkgosi@bca.bw

28. Chaudry M. Khalique, International Institute for Symmetry Analysis and Mathematical Modelling, Department of Mathematical Sciences, North-West University, Mafikeng Campus, Private Bag X 2046, Mmabatho 2735, Republic of South Africa. Email: Masood.Khalique@nwu.ac.za

29. Mervis Kikonko, University of Zambia, mervis.kikonko@unza.zm

30. Khumbo Kumwenda, Department of Mathematics, Mzuzu University, P/Bag 201, Mzuzu 2, Malawi. Email khumbo@aims.ac.za

31. E.M. Lungu, Department of University of Botswana, Gaborone, Botswana. Emails: lunguliz@gmail.com, lunguem@mopipi.ub.bw

32. A.G. Luwanda, Department of Mathematics, Mzuzu University *and* School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal

33. Mavuta Trocco Mafuta, Department of Mathematics, Mzuzu University

34. Kondwani Magamba, Department of Mathematical Sciences, Catholic University of Malawi. Email: magambakho@yahoo.co.uk

35. David M. Malonza, Kenyatta University, Department of Mathematics, Nairobi, Kenya. E-mail: dmalo2004@gmail.com

36. Vincent Manda, Lilongwe University of Agriculture and Natural Resources.

37. T. Marijani, Department of Mathematics, University of Dar-es-Salaam, P.O.BOX 35062, Dar-es-Salaam, Tanzania

38. A.J. Meir, Department of Mathematics and Statistics, Auburn University, Alabama 36849-5310, U.S.A.

39. Cynthia Mikateko Mazinu, SACEMA, University of Stellenbosch, South Africa

40. M. Mokgolele, Botswana College of Agriculture, Private Bag 0027, Gaborone, Botswana.

41. Eunice Mphako-Banda, School of Mathematics, University of the Witwatersrand. Email: Eunice. Mphako-Banda@wits.ac.za

42. Ben Muatjetjeja, International Institute for Symmetry Analysis and Mathematical Modelling, Department of Mathematical Sciences, North-West University, Mafikeng Campus, Private Bag X 2046, Mmabatho 2735, Republic of South Africa. Email: Ben.Muatjetjeja@nwu.ac.za

43. Walter Mudzimbabwe, School of Business and Management Sciences, Harare Institute of Techonology, P. O. Box BE 277, Belvedere, Harare, Zimbabwe. Email: mudzmudz@gmail.com

44. Rosaliah Mughogho, Department of Mathematics, Mzuzu University, Malawi. Email: rosamughogho@gmail.com

45. Egbert Mujuni, University of Dar-es-Salaam, Tanzania. Email: emujuni@maths.udsm.ac.tz

46. P. Musicha, Department of Mathematics, Catholic University of Malawi

47. John Musonda, Department of Mathematics and Statistics, University of Zambia. Email: j.musonda@yahoo.com

48. E. Mwakilama, Department of Mathematical Sciences, Chancellor College, University of Malawi, P.O.Box 280, Zomba, Malawi. Email: mwakilamae@gmail.com

49. George Mzengo, Kamuzu Academy, Malawi

50. SC Oukouomi Noutchie, North-West University, Mafikeng Campus, South Africa

51. G.P. Nthoiwa, Botswana College of Agriculture, Private Bag 0027, Gaborone, Botswana. Email: gnthoiwa@bca.bw

52. Giulia Di Nunno, Centre of Mathematics for Applications, University of Olso, Norway. Email: g.d.nunno@cma.uio.no

53. F. Nyabadza, Department of Mathematics, University of Stellenbosch, P. Bag X 1, Matieland, 7602, South Africa

54. Elias R. Offen, University of Botswana, Private Bag UB00704, Gaborone, Botswana. E-mail: offen@mopipi.ub.bw

55. Bernt Øksendal, Center of Mathematics for Applications (CMA), Dept. of Mathematics, University of Oslo, P.O. Box 1053, Blindern, N-0316 Oslo, Norway *and* Norwegian School of Economics and Business Administration, Helleveien 30, N-5045 Bergen, Norway. Email: oksendal@math.uio.no

56. Eva Øksendal, Norway

57. Jacob Otieno Ong'ala, Center for Applied Research in Mathematical Sciences, Strathmore University, P.O. Box 59857 - 00200 City Square, Nairobi, Kenya. Email: jotieno@strathmore.edu

58. Nicolas Privault, Nanyang Technological University, Singapore. Email: nprivault@ntu.edu.sg

59. Mangatiana A. Robdera, Department of Mathematics, University of Botswana, Private Bag 0022, Gaborone, Botswana

60. John A. Ryan, Department of Mathematics, Mzuzu University

61. S. Shateyi, Department of Mathematics, University of Venda, P Bag X5050, Thohoyandou 0950, South Africa.
E-mail: stanford.shateyi@univen.ac.za

62. Virginie Konlack Socgnia, Programme in Advanced Mathematics of Finance, School of Computational and Applied Mathematics, University of the Witwatersrand, Wits 2050, Johannesburg, South Africa. E-mail: virginie.konlack@wits.ac.za

63. Mihalis Zervos, London School of Economics, United Kingdom. Email: M.Zervos@lse.ac.uk