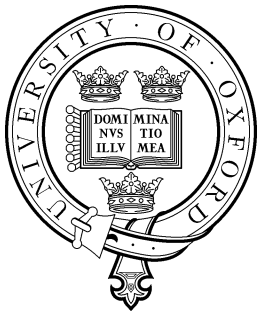
Mathematical Evolutionary Epidemiology: Limited Epitopes, Evolution of Strain Structures and Age-specificity



Alhaji Cherif

St Edmund Hall

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2015

To those who had no opportunity for an education
To Pax Africana
To Hadja Felani Saran Kalon Cherif (Jan. 12, 1929–Nov. 27, 2007)
To Mamadi Cherif (–July 24, 2012)

Acknowledgements

Asiyefunzwa na mamae hufunzwa na ulimwengu, which in Kiswahili, roughly means: it takes a village to raise a child. This adage has been particularly true in both my personal and academic life. I am deeply indebted to the people who have helped and influenced me along my journey, and have taught me both formally and informally. There is no way I will be able to mention all of their names, and therefore I beg your forgiveness if I forget to do so. In particular, I would like to thank my supervisors Professors Philip K. Maini and Sunetra Gupta and Dr. Janet Dyson for their unwavering guidance, limitless enthusiasm, encouragement and constructive advice, and for placing a considerable faith in me. I thank Prof. Gupta for introducing me to the exciting world of evolutionary epidemiology and for suggesting the research problems addressed in this thesis, Prof. Maini for his considerable patience and exceptional mentorship, and Dr. Dyson for lending her expertise and guidance in abstract dynamical systems. I would not have been able to pursue my doctorate study without the financial support from all the organisations who have opened the various worry-free doors leading to my current academic endeavour. Most importantly, I thank the U.S. National Science Foundation (NSF) for funding my doctoral work under the auspices of the Graduate Research Fellowship Program. I would also like to thank Dr. Mario Recker and Dr. Paul Wikramaratna (for debugging and a more efficient implementation of the codes); Kamal Barley and Shaquetta Tatum (for help at various stages, including some editorial works) and Dr. Fengqi Yi for introducing me to bifurcation in infinite-dimensional dynamical systems and his kind hospitality at Harbin Engineering University, Harbin, China. Finally, I would like to thank my family for all their love, faith, unwavering support, and interests throughout my education. This work is dedicated to those who had no opportunity for an education, to those who try to make education accessible and affordable to all, and to teachers and mentors who have impacted the lives of their students and mentees.

Abstract

We investigate the biological constraints determined by the complex relationships between ecological and immunological processes of host-pathogen interactions, with emphasis on influenza viruses in human, which are responsible for a number of pandemics in the last 150 years. We begin by discussing prolegomenous reviews of historical perspectives on the use of theoretical modelling as a complementary tool in public health and epidemiology, current biological background motivating the objective of the thesis, and derivations of mathematical models of multi-locus-allele systems for infectious diseases with co-circulating serotypes. We provide detailed analysis of the multi-locus-allele model and its age-specific extension. In particular, we establish the necessary conditions for the local asymptotic stability of the steady states and the existence of oscillatory behaviours. For the age-structured model, results on the existence of a mild solution and stability conditions are presented. Numerical studies of various strain spaces show that the dynamic features are preserved. Specifically, we demonstrate that discrete antigenic forms of pathogens can exhibit three distinct dynamic features, where antigenic variants (i) fully self-organize and co-exist with no strain structure (NSS), (ii) sort themselves into discrete strain structure (DSS) with non-overlapping or minimally overlapping clusters under the principle of competitive exclusion, or (iii) exhibit cyclical strain structure (CSS) where dominant antigenic types are cyclically replaced with sharp epidemics dominated by (1) a single strain dominance with irregular emergence and re-emergence of certain pathogenic forms, (2) ordered alternating appearance of a single antigenic type in periodic or quasi-periodic form similar to periodic travelling waves, (3) erratic appearance and disappearance of synchrony between discrete antigenic types, and (4) phase-synchronization with uncorrelated amplitudes. These analyses allow us to gain insight into the age-specific immunological profile in order to untangle the effects of strain structures as captured by the clustering behaviours, and to provide public health implications. The age-structured model can be used to investigate the effect of age-specific targeting for public health purposes.

Contents

1 Prolegomenon **1**
- 1.1 Historical Introduction . . . 1
- 1.2 Motivation and Aims . . . 4
- 1.3 Organization of Thesis . . . 8

2 Epidemiology of Influenza Virus **10**
- 2.1 Introduction . . . 10
- 2.2 Epidemiology of Influenza . . . 11
  - 2.2.1 Antigenic Drift and Shift in Influenza A . . . 13
- 2.3 Limited Epitopes in Influenza A . . . 13
- 2.4 Simple Compartmental Models . . . 16
- 2.5 Generalization of Multiple Strain Model . . . 17
- 2.6 Model Reductions . . . 20
  - 2.6.1 Auxiliary Quantities . . . 20
  - 2.6.2 Reduction of Dynamic Eqs. (2.12)-(2.14) . . . 21
- 2.7 Conclusion . . . 24

3 Mathematical Analysis of Multi-locus-allele System **26**
- 3.1 Introduction . . . 26
- 3.2 Model Description: Multi-locus-allele Systems . . . 28
- 3.3 Positivity and Invariant Feasible Region . . . 31
- 3.4 Steady State and Stability Analysis . . . 33
  - 3.4.1 Disease-Free Equilibrium . . . 33
  - 3.4.2 Endemic Equilibria . . . 34
    - 3.4.2.1 Weak Endemicity: Principle of Competitive Exclusion and Discrete Strain Structure . . . 35
    - 3.4.2.2 Strong Endemicity: No Strain Structure (NSS) . . . 38
    - 3.4.2.3 Strong Endemicity: Cyclical Strain Structures and Existence of Oscillatory Behaviours . . . 42
- 3.5 Discussion . . . 43
- 3.6 Conclusion . . . 49

4 Derivation of Age-structured Multi-locus-allele Model **52**
- 4.1 Introduction . . . 52
- 4.2 Antigenic Thrift Model with Age-specificity . . . 53

4.3 Model Reduction via Auxiliary Quantity 55

5 Age-specificity I: Existence of Mild and Steady-State Solutions 59
5.1 Antigenic Thrift Model with Age-specificity 59
5.2 Existence of a Positive Solution: 60
5.3 Equilibria: Steady-State Analysis 67
5.3.1 Existence of Steady States 67
5.3.2 Properties of the Steady State Operator 68
5.3.3 Next-Generation Operator: Threshold Property 73
5.3.4 Existence and Threshold Conditions for Equilibria 76

6 Age-specificity II: Stability Analysis of Equilibria 81
6.1 Introduction 81
6.2 Preliminaries: Networks and Symmetry 84
6.2.1 Preliminaries and Definitions 85
6.2.2 Notion of Synchronous Equilibrium Clusters 87
6.3 Stability of Equilibria 91
6.3.1 Stability of Disease-free Equilibrium 93
6.3.2 Stability of Endemic Equilibria 102
6.3.2.1 Stability of Weakly Endemic Equilibria 105
6.3.2.2 Stability of Strongly Endemic Equilibrium 116
6.4 Summary 119

7 Extended Numerical Investigation of Multi-locus-allele Age-structured Models with Different Allelic Variations 121
7.1 Introduction 121
7.2 Parameter Descriptions 122
7.2.1 Demographic Parameters 122
7.2.2 Epidemiological Parameters 123
7.2.3 Reproductive Numbers, $\mathcal{R}_0$ 124
7.3 Model Validation with Serological Data 125
7.4 Numerical Simulations: The Effect of Epitope Architectures 131
7.4.1 Model Dynamics I: Uniform Age-specific Parameters 132
7.4.2 Model Dynamics II: Differential Age-specific Parameters . . . 134
7.5 Conclusion 140

8 Interaction Between Shifting Demographic and Immunological Landscapes in Influenza Epidemics 146
8.1 Introduction 146
8.2 Antigenic Evolution under Changing Immunological and Demographic Landscapes 148
8.3 Effects of Single Strain Dominance on Age-specifi-city 151
8.4 Antigenic Evolution of Pathogens: Limited Epitopes and Age-specificity 157
8.5 Public Health Policy based on Dynamic Immunological Profiles of the Population 159

8.6 Conclusions 160

9 Closing Remarks **162**
9.1 Summary 162
9.2 Future Directions 163

A Chapter 3 Appendix **164**
A.1 Proof of Lemmas and Theorems 164
A.1.1 Proof of Theorem 3.4.1 164
A.1.2 Proof of Theorem 3.4.2 164
A.1.3 Proof of Theorem 3.4.4 165
A.1.4 Proof of Theorem 3.4.7 165
A.2 Other Calculations 166
A.2.1 Monotonicity of the Infected Steady State, y^* 166
A.2.2 No Strain Structure Characteristic Polynomial $P_k(\zeta)$ 166
A.3 Illustrative Example of a 4-strain System 167
A.3.1 Equilibrium Types of the form a-1 to a-4 in Fig. 3.1 169
A.3.2 Equilibrium Types of the form b-1 to b-4 in Fig. 3.1 169
A.3.3 Equilibrium Types of the form c-1 to c-4 in Fig. 3.1 170
A.3.4 Equilibrium Types of the form d-1 to d-4 in Fig. 3.1 171

B Chapter 4 Appendix **172**
B.1 Numerical Methods 172
B.1.1 Method of Lines *à la* Stable Age-distribution 172
B.1.1.1 Derivation 172
B.1.1.2 Application: Age-structured Multi-locus-allele System 175
B.1.2 Model Comparison under Constant Parameters 176

C Chapter 5 Appendix **178**
C.1 Proofs 178
C.1.1 Proof of Theorem 5.3.26(see Krasnoselskii, 1964) 178

D Chapter 6 Appendix **180**
D.1 Proofs 180
D.1.1 Proof of Theorem 6.3.4 180

Bibliography **181**

List of Figures

1.1 Time evolution of Influenza A virus 6
1.2 Age-specificity in the Strength of Micro-neutralization 7

2.1 Historical Episodes of H1N1 Influenza epidemics and pandemics . . . 12
2.2 Limited Epitope/Antigenic Thrift Hypothesis 14

3.1 Multi-locus System (Possible Equilibria for Four Strain System on a Square Lattice with $\mathbb{D}_4$-Symmetry) 35
3.2 DSS Time Evolution (Eqs. (3.4)-(3.6)) 39
3.3 Functional Properties of $y^*(\gamma, \mathcal{R}_0)$ (Eqs. (3.16)-(3.17)) 40
3.4 Time Evolution of No Strain Structure (NSS) (Eqs. (3.4)-(3.6)) . . . 41
3.5 Time Evolution of Cyclic Strain Structure (Eqs. (3.4)-(3.6)) 44
3.6 Parameter Space $\mathcal{R}_0$-γ for 2-locus-2-allele system 45
3.7 Steady-State Bifurcation Diagram 46
3.8 Strain Space Structures. 48

6.1 Multi-locus System (Possible Equilibria for Four Strain System on a Square Lattice with $\mathbb{D}_4$-Symmetry) and Quotient Network Representation 84
6.2 Network and its Quotient. 87
6.3 Proof Sketch 99
6.4 Proof Sketch of Endemic Equilibrium 105
6.5 Domain $\mathcal{D}_l$. 114

7.1 Age-Specific Demographic Patterns. 122
7.2 Representative Contact Rate. 123
7.3 Age-specificity in Serological Surveys of Influenza A Virus 126
7.4 Estimated Parameters 127
7.5 Comparison of Model (4.18)-(4.20) and Data with Four Age Classes . 129
7.6 Estimated Force of Infection for Four Age Classes 130
7.7 Potential Infectivity 131
7.8 Genetic Relatedness of Hemagglutination-Inhibition Assay 132
7.9 Three Long-term Age-specific Dynamical Patterns 135
7.10 Uniform Age-Specificity in NSS and Average Temporal Evolution . . 136
7.11 Uniform Age-specificity and CSS of the 3-locus-2-allele System 137
7.12 Uniform Age-specificity and CSS of the 2-locus-(2,3)-allele System . . 138

7.13 Age-specificity in DSS Dynamics 139
7.14 Differential Age-Specificity in NSS and Average Temporal Evolution . 141
7.15 Differential Age-specificity and CSS of the 3-locus-2-allele System . . 142
7.16 Differential Age-specificity and CSS of the 2-locus-(2,3)-allele System 143
7.17 Differential Age-specific and CSS of the 2-locus-(2,3)-allele System . . 144
7.18 Differential Age-specificity in DSS Dynamics 145

8.1 Mechanism of Changing Immunological Landscape and Restructuring of Epidemiological Structure 149
8.2 Dominance of a Single Strain Dynamics 152
8.3 Temporal Evolution of an Age-specific 5-Locus-2-Allele System 153
8.4 Age-specificity in Relative Proportions of Immune Individuals for Multi-locus-allele Systems 154
8.5 Age-specificity in Relative Proportions of Infected Individuals for Multi-locus-allele Systems 155
8.6 Overall Peak-statistics for Multi-locus-allele Systems 156
8.7 Serological Survey by Age-cohorts 158

B.1 Model Comparision 177

List of Tables

7.1 Contact, Fertility and Mortality Rates 125

1 Prolegomenon

Contagions have been studied for many years by sociologists, epidemiologists and theoreticians concerned with the diffusion of innovation and cultures, and disease spread. In recent decades, efforts have been devoted to predicting social and biological contagion processes in the economy, public policy, national security and public health. This chapter will motivate the need for the work presented in this thesis by providing the historical precedence of mathematical approaches to the study of infectious diseases, and its critical importance in public health decision-making. We will conclude with the outline of our contributions and summary of the contents of each chapter.

1.1 Historical Introduction

Infectious diseases–including both great epidemics, such as plague and smallpox, which have devastated human populations from ancient times to modern eras, and less dramatic viral and bacterial infections causing high infant morbidity and mortality– have profoundly impacted both human history and biology; and have claimed more lives than all wars and natural disasters combined (Anderson & May, 1979, 2002; Anderson, 1991; Hyams et al., 2001; Azambuja, 2004; Chea, 2005; Becker, 2005; Wong & Yuen, 2006). The Bubonic plague claimed approximately one-third of the total European population in the fourteenth century, while smallpox infected most of the indigenous population in the Americas (Shrewsbury, 1970). For instance, at least 1.75 million Aztecs died from smallpox in 1520. HIV/AIDS, tuberculosis and malaria account for more than 6 million deaths a year (Asamaoh-Odei et al., 2004; Ruxrungtham et al., 2004). In 2005, HIV/AIDS was estimated to have caused 3.1 million deaths; tuberculosis is responsible for 1.7 million deaths a year; malaria is credited with more than a million deaths, and more than 400 million people clinically affected annually (WHO, 2005a-d; Boraschi et al., 2008). Globally, between 250,000 and 500,000 deaths, and up to millions in some pandemic years, are due to influenza every year. In the United States alone, approximately 40,000 people die each year from influenza (Cox & Subbarao, 2000). Thus, understanding and controlling the

spread of infectious diseases is of crucial importance, providing challenges for trans-disciplinary research with contributions to be made from various approaches in which traditional disciplinary boundaries are irrelevant.

Because the relevant research requires the synthesis of theoretical frameworks from different fields, interdisciplinary scholars, public health officials and decision-makers have, in recent decades, turned to mathematical models as complementary tools in dealing with epidemiological events ranging from the outbreak of malaria, measles, HIV/AIDS, cholera and tuberculosis to influenza. As a result, we have seen an increase in the use of theoretical models (e.g., meta-population, multi-group, assortative and non-homogeneous mixing, network, individual-based and spatio-temporal, and models structured by age, gender and size) that have revealed the subtleties in the dynamics of contact processes underlying the transmission mechanisms and the biological constraints of host-pathogen systems (Watts & Strogatz, 1999; Keeling, 1999; Mollison, 1995; Diekmann & Heesterbeek, 2000; Brauer & Castillo-Chavez, 2001).

Mathematical studies in epidemiology attempt to reconstruct epidemiological patterns of disease transmission and evolution in populations of interest (e.g., human and animals). Theoreticians have contributed significantly to the understanding of the prevalence and spread of infectious diseases in human populations. The first quantitative attempts to predict the extent of disease outbreaks took two bifurcating routes. The first attempts were *a posteriori* statistical bio-demographic models in nature, and the earliest of these models was attributed to John Graunt (1620-1674) in his *Natural and Political Observations made upon the Bills of Mortality* published in 1665, in which he illustrated the extent to which plague affected London. The second milestone in the use of statistical methods came in the late 1800s when Farr provided mathematical descriptions of smallpox and cattle-plague epidemics using parametric curve-fitting procedures. Farr's models were restricted to describing the shape of the epidemic curve via the use of suitable functions with satisfactory fits (Farr, 1866), but the models failed to provide interpretations of the parameters in terms of biology (Whitehead, 2000). It was Ransome (1868) who speculated the underlying bio-demographic mechanism, and recognized the importance of susceptible density on the contagion of an epidemic (see Hilts, 1980, for a detailed historical account of statistical contributions in epidemiology).

The second route was the use of dynamic models to quantify the time course of epidemics. The first mathematical contribution to epidemiology in this regard is attributed to Daniel Bernoulli, who showed the effect of *variolation* on smallpox using an age-specific catalytic model at the Royal Academy of Sciences in Paris (Dietz & Heesterbeek, 2002). The need to identify the mechanisms that generate the various

epidemiological patterns, and to provide strategic measures for eradication or elimination of infectious diseases (Anderson, 1991; Kermack & McKendrick, 1927, 1932, 1933; also see Molyneux, 2008 and references therein for review) has played a significant role in the development of mathematical biology. In particular, theoretical epidemiology is the fastest growing interdisciplinary field, with an extensive list of introductory textbooks, monographs, and refereed papers (Bailey, 1975; Mollison, 1995; Diekmann & Heesterbeek, 2000; Brauer & Castillo-Chavez, 2001).

Since the first catalytic compartmental model was introduced by Bernoulli (1760) and its generalization by Ross in a series of papers (Ross, 1910, 1911, 1916; Ross & Hudson, 1917a-b), the acceptance of the germ theory of disease, and independent formulation of the mass action assumption for epidemics by Brownlee (1907), Ross (1916), and Kermack and McKendrick (1927, 1932, 1933), theoretical and mathematical epidemiologists have acknowledged the need to include more biologically realistic properties in order to understand transmission mechanisms and dynamics fully. In addition, researchers have suggested various preventive measures in order to eradicate, control or reduce incidence or prevalence. Models in the past have suggested various control measures of isolation, quarantine, and vaccination, some of which have been effective in reducing and/or eliminating diseases. Adhering to the mathematical traditions of Bernoulli (1769), Ross (1916), McKendrick and Kermack (1927, 1932, 1933), and Anderson and May (1979), various models of different diseases have appeared in the literature with more sophisticated epidemiological assumptions and observational approximations. In particular, models incorporating effects such as functional birth rates (Castillo-Chavez & Velasco-Hernandez, 1998a-b), attainment of immunity, identification of competition mediated by cross-immunity, co-infection, development of symptoms or lack thereof in stages, age-dependence or age-structure and vaccination (Sinko & Streifer, 1967; Hethcote, 1997; Read & Keeling, 2003) have been studied. Other variations incorporating delays (e.g., discrete or distributed) in transmission of disease have been proposed (Fine, 1975, 1977, 1979; Busenberg et al., 1983; Brauer, 1995; Castillo-Chavez & Velasco-Hernandez, 1998a-b; Ingemar & Scott, 2001; Karp & Auwaerter, 2007). While some models have included additional compartments, others have incorporated seasonality and spatio-temporal patterns, and heterogeneity through the use of network, stochastic effects and structurally extended models (Bartlett, 1957, 1960, 1964; Bartholomew, 1982; Klovdahl et al., 1994; Watts & Strogatz, 1998; Keeling, 1999, 2000a-b; Morse, 2004; Keeling & Eames, 2005; Kuske et al., 2007; Rost & Wu, 2008).

The majority of these models have been useful in characterizing and evaluating different biological assumptions and phenomena that are important in determining

causes of outbreaks. In addition, an increasing use of a mathematical and computational modelling framework has provided better insight into data collection, and implementation of optimal strategies for public health measures for diseases such as SARS and rubella, to name but a few. However, despite all these advances and efforts, incorporation of more realistic assumptions has been a significant challenge, and infectious diseases still remain the leading cause of death, accounting for the majority of global mortality (Azambuja, 2004). The persistence of many of the infectious diseases modelled by these approaches has been credited to the emergence and re-emergence of new zoonoses such as gonorrhea, cholera, AIDS/HIV, SARS, malaria, dengue, tuberculosis; and evolution of pathogen genetic traits seen in many influenza subtypes such as H1N1 swine-origin flu and H5N1 avian flu, and the surge in the existence of more drug-resistant strains among diseases such as tuberculosis, pneumonia and gonorrhea.

1.2 Motivation and Aims

Due to advances in epidemiological surveillance, the temporal patterns of some of these diseases are known, providing opportunities to examine the biological assumptions underpinning theoretical models, and to characterize the dynamics of host-parasite systems in order to implement effective public health measures for improvement of preventive protocols (e.g., control management, detection, forecasting, and eradication). Despite a considerable amount of empirical and theoretical investigation, the recurrent patterns in infectious diseases such as Influenza A virus, dengue, and P. falciparum parasites are not well-understood. This lack of understanding has been attributed to the uncertainty associated with the strength of host immune response and other factors relevant to disease ecology, evolutionary dynamics, rapid pathogen evolution, and the availability of limited amount of surveillance on different pathogenic subtypes. As a result, these aspects are the subject of much current epidemiological research. In this thesis, we propose to study the evolution of infectious agents and their observed population patterns, using Influenza A virus as a motivating application. We attempt to provide explanations as to how evolutionary mechanisms and ecological factors regulate the diversity and age-specificity of Influenza A viruses.

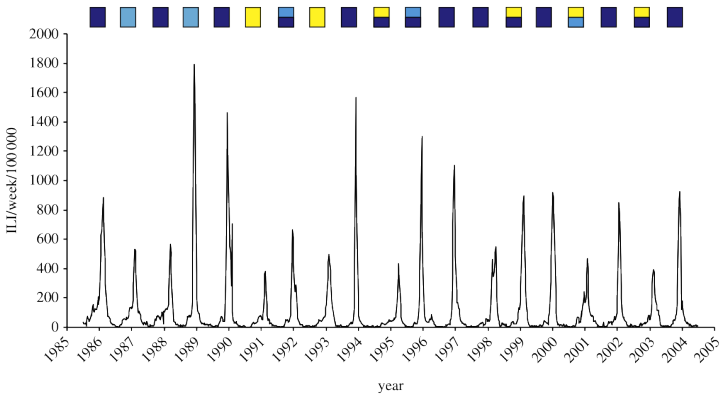
Influenza viruses are known to have caused a number of pandemics in the last 150 years, most notably that of the Spanish flu of 1918. Once an influenza subtype has established itself in a population, it appears to settle into a pattern of epidemic behaviour whereby the seasonalities are characterized by annual epidemics with clear signals of multi-annual cycles in strain abundance as illustrated in mortality data (see Fig 1.1). This process has generally been understood as an *antigenic drift* whereby

the virus incrementally accumulates mutations that render it effective against prevailing herd immunity, eventually allowing it to breach such an immunological barrier and enabling it to elude the antibodies made against the parent virus (Lin et al., 1999; Hampson, 2002; Recker et al., 2007). Formal mathematical, theoretical and empirical investigations of this hypothesis show the link between this process, patterns of influenza dynamics and genetic diversity to be far more tenuous (Andreasen et al., 1997; Lin et al., 1999; Gog & Grenfell, 2002; Ferguson et al., 2003; Grenfell et al., 2004; Andreasen & Sasaki, 2006; Koelle et al., 2006; Gog, 2008).

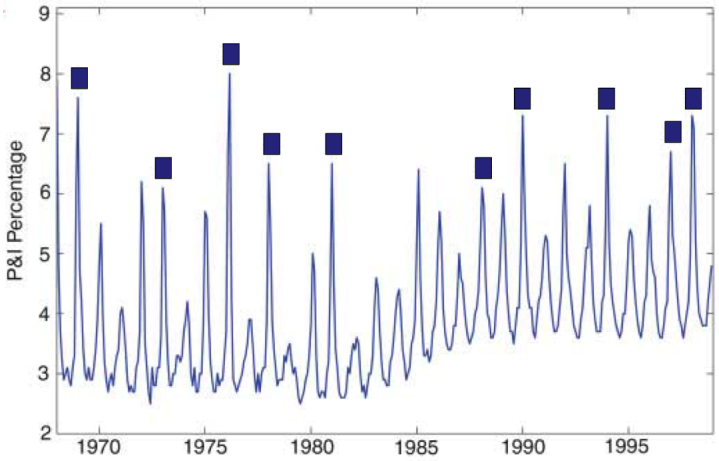
An alternative explanation has recently been formulated where, due to functional constraints on the defining epitopes, the virus population is characterized by a limited set of antigenic types, all of which may be continuously generated by mutation within pre-existing strains and other processes. Under this antigenic *thrift* hypothesis, as highlighted by Gupta et al. (1998) and Recker et al. (2007), influenza outbreaks arise in a different manner to that which is envisaged by the conventional antigenic drift model, in that the successful viral type is one that possesses the correct combination of epitopes to find a hole in the complex landscape of prevailing epitope-specific host immune responses rather than one that has simply accumulated enough mutations to become sufficiently different from the previous strains.

A mathematical model of this hypothesis explains the irregular emergence of certain pathogenic forms and limited boundedness of diversity. In addition, a specific prediction of the results is that the antigenic distance between influenza strains does not necessarily increase with time (Gupta et al., 1998; Recker et al., 2007). Instead, it visualizes the virus population zigzagging in antigenic (higher dimensional hypercube characteristic) space, and over a period returning to where it started. For example, Fig. 1.1b shows a monthly percent of influenza-associated mortality in the United States from 1968 to 1998 for elderly persons in which the squares denote the seasons dominated by recycling of H3N2 cluster.

Recent empirical studies of cross-reactivity of human sera to the current swine-origin H1N1 virus consistently reveal that individuals above a certain age have neutralizing antibodies to this new virus (see Fig. 1.2). These observations are in line with the prediction that the same antigenic type is likely to re-emerge after a certain period. Figure 1.2, for example, shows data on the strength of a microneutralization of H1N1 among subjects categorized by decade of birth (Hancock et al., 2009). While numbers are small, it is clear that the most intense recognition is among those born prior to 1930, but significant responses are observed also among those born afterwards thus suggesting that strains carrying similar epitopes were still in circulation. This corroborates earlier anecdotal work on the sera collected from some elderly persons in 1950, where the sera were found to have contained antibodies to



(a) France from 1984 – 2004



(b) USA from 1968 – 1998

Figure 1.1: **Time evolution of Influenza A virus.** Typical time evolution dynamics of influenza mortality and morbidity. **(a)** Weekly temporal incidence dynamics for medically attended influenza-like illness (ILI) in France from 1984−2004 (adapted from Denoeud et al., 2007, Truscott et al., 2011). **(b)** Monthly percentage of deaths associated with pneumonia and influenza (P&I) in USA from 1968 – 1998 for all persons aged 65 years and older (adapted from Greene et al., 2006, and Koelle et al., 2006). The coloured rectangles mark the first season with dominant types and subtypes of circulating influenza viruses. Dark blue, A/H3N2; light blue, A/H1N1; yellow, B. Bottom: rectangles mark the first season dominated by A/H3N2 after a cluster transition.

the H2N2 subtype before H1N1 viruses had been displaced by H2N2 in the Asian Influenza pandemic of 1957. A similar observation was made in 1968, when sera from elderly populations were found to have possessed antibodies to H3N2 viruses before they too displaced H2N2 viruses in the epidemics of 1968 (Hong Kong flu). In addition, there are numerous empirical findings suggesting that a new mutant cluster can replace a resident cluster as illustrated by the invasion Beijing/1993 cluster, which provoked the epidemics from 1992 to 1996 in New York state before being replaced

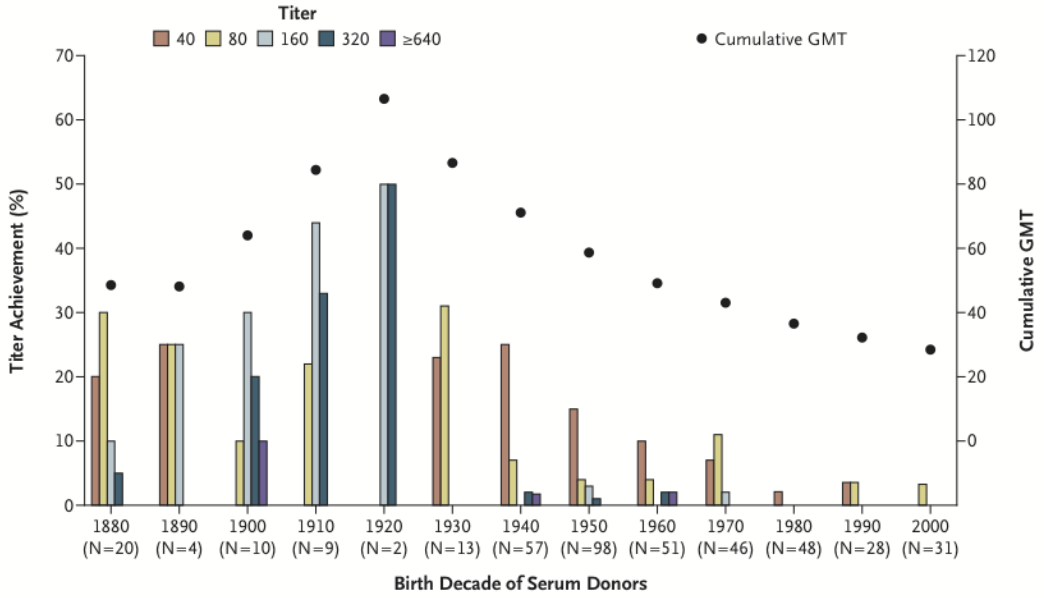


Figure 1.2: **Age-specificity in the Strength of Micro-neutralization.** Strength of micro-neutralization of H1N1 among individuals in different decadal birth group, where titer achievement is proportion with neutralizing antibody titers, and GMT is the geometric mean titer on the right ordinate (Reproduced with permission from Hancock et al., 2009, Copyright Massachusetts Medical Society)

.

by Wuhan/1995-like viruses (Rambaut et al, 2008).

Although the dynamics in Recker et al. (2007) captured the irregular emergence and re-emergence phenomena, it failed to provide explanations of the mechanisms that constrain the patterns of age-specificity illustrated in the above anecdotal dynamics. This re-appearance or recycling phenomenon of long-vanished subtypes of Influenza A virus cannot be viewed without considering the age patterns of Influenza patients. Therefore, a finer stratification by age provides an opportunity further to reassess the hypothesis of antigenic thrift (limited epitope) on these data. By incorporating an age stratification of antigenic evolution of Influenza, we propose to extend the antigenic thrift (or limited epitopes) model characterized by the host immune response with meta-population structure of antigenic types in Recker et al. (2007).

The main thrust or contribution of this thesis is the exploration of whether the serological data exemplified in Fig. 1.2 can be explained by the population dynamics of our extended model and the interrogation of the *Limited Epitope/Antigenic Thrift* hypothesis put forth by Recker et al. (2007); more broadly we investigate whether age-structured patterns of infection, disease and serological reactivity can be used to discriminate between different hypotheses regarding the diversity in the antigenic evolution of diseases such as Influenza A viruses, *Staphylococcus aureus, Enterococcus*

faecium and various species of pathogenic *Neisseria*. Furthermore, our goal is to use our understanding of the dynamic processes in the model to speculate the most effective public health policy. As such, we aim to accomplish the following three objectives:

1. Develop predictive models based on biological and ecological studies and assumptions to understand the observed complex epidemiological dynamic patterns, to propose theories to enhance predictability of experimental and empirical data and observations, and to make testable predictions that can further our understanding of the mechanisms involved in generating the observed patterns;

2. Provide a rigorous mathematical analysis of the models in order to gain generic insights into the dynamic behaviours of the models;

3. Provide public health policy-related implications and recommendations based on the predictions and understanding of the dynamic behaviours of the model.

The thesis extends the work of Gupta et al. (1998), in which the authors incorporate immunological heterogeneity in the population, and assume that cross-immunity is acquired through previous infections and mediates competition among strains. The directions from which we approach our investigation are set out in the next section.

1.3 Organization of Thesis

The thesis consists of 9 chapters. Chapter 2 starts with overviews of influenza epidemiology with a brief discussion of mechanisms that maintain strain diversity and the underlying biological assumptions of antigenic drift and shift. We then provide an alternative hypothesis – *antigenic thrift* – that might explain the limited boundedness of influenza diversity. Later, we derive multi-locus models incorporating multiple strain interactions via cross-immunity where an unstructured model defined by a system of ordinary differential equations is discussed, and Chapter 4 extends the model presented in Chapter 2 where age-specificity is incorporated using a system of coupled hyperbolic partial differential equations.

Furthermore, we present a detailed local analysis of both models discussed in Chapters 2 and 4. In particular, Chapter 3 gives detailed mathematical analyses of the limited epitope model developed and numerically studied by Gupta and colleagues (Gupta et al., 1998; Recker et al., 2007), which has provided a promising explanation as to how strains might compete in the presence of host immunological history and how phenotypes are perceived differently by hosts. In addition, Chapter 3 establishes

a positivity result and provides local stability conditions for the equilibria and conditions for extinction of the disease. Partial coexistence via the principle of competitive exclusion and full coexistence of strains are discussed. The impact of cross-immunity on the temporal evolution and strain community structures is illustrated, where we show the existence of three distinct dynamical regions with some exotic dynamic patterns (i.e., full and partial synchronization, limit cycles, periodic travelling waves, period-doubling, intermittency, chaotic, self-organizing and synchronous solutions). Using results from infinite-dimensional dynamical systems (i.e., semigroup theory, spectral theory for operators, bifurcation theory) and relying on the inherent insights from Chapter 3, Chapters 5–6 study positivity, existence of a mild solution and of stability conditions, among other model characteristics necessary to establish rigorous results needed to understand the model behaviours.

In Chapters 7-8, we perform numerical studies of various dimensions of strain spaces, use four serological data sets to fit the model, and discuss public health implications. Chapter 9 summarizes the contents, highlighting the usefulness as well as the limitations of our models. In addition, we propose future extensions, list open problems and outline future directions.

2 Epidemiology of Influenza Virus

The morphological, genome composition, envelope components, molecular structures, surface glycoproteins and replication functions of influenza virus are the most studied and well-understood of any virus (Earn et al., 2002; Carrat & Flahault, 2007; Fraser et al., 2009; Plotkin, 2006; WHO, 2003, 2010). However, predicting epidemic and pandemic behaviours of influenza viruses is still a challenging problem. Hence, most current research has focused on understanding the interaction between demographic, epidemic and pandemic, and evolutionary patterns of the virus. In this chapter, we provide both biological and epidemiological backgrounds of influenza viruses, and review some basic mathematical models of multiple strain systems and provide a derivation of the multi-locus-allele model.

2.1 Introduction

In April of 2009, the first strain of swine A/H1N1 influenza virus was identified and reported in Mexico City (WHO, 2010). By the end of the month, more than 140 reported cases were confirmed in various US cities and around the world, moving the World Health Organization pandemic alert classification[1] from level 3 to level 5 and later to level 6 despite a lower mortality rate as compared to ordinary seasonal influenza outbreaks. This move in the pandemic classification within the first two-week period created hysteria in the US and around the globe.

The emergence of this recent swine A/H1N1 influenza across the global, as well as the occasional infections of humans attributed to avian flu, with a disconcertingly high number of deaths, has caused a public health concern about a potential outbreak of a new pandemic and epidemic of influenza (WHO, 2010). Influenza is one of the recurrent infectious diseases with the most significantly high annual morbidity and mortality, resulting in the death of between $250,000$ and $500,000$ people with more than $150,000$ hospitalizations (Cox & Subbarao, 2000). In the pandemic year of 2009 A/H1N1, fewer than 0.05% of people developing clinical symptoms died and 1.4% were hospitalized (WHO, 2003, 2010). Demographically, high incident rates

[1] A preparedness plan that defines the stages of a pandemic.

are usually seen among young children, elders and immunologically compromised individuals (CDC, 2009a-b). In addition, Influenza is the cause of most of the modern severe pandemics of the last century, ranging from the 1889 H2N2 Russian Influenza, 1918 H1N1 Spanish Influenza, 1957 Asian H2N2, 1968 Hong Kong to the most recent 2009 Mexico H1N1 Influenza (see Fig. 2.1, and see Cox & Subbarao, 2000; WHO, 2003, 2010; CDC, 2009a-b). Pandemics are usually caused by the emergence of new strains in humans when an existing human strain either picks up new genes from bird and/or pig strains (e.g., 2009 A/H1N1 Swine), or jumps from animals to humans (e.g., A/H5N1).

2.2 Epidemiology of Influenza

Influenza viruses, segmented single-stranded negative-sense RNA viruses of the family *Orthomyxoviridae* contained within enveloped virions, are classified into three phylogenetically distinct types–A, B, and C–related to their genera–Influenzavirus A, Influenzavirus B, and Influenzavirus C, respectively (Smith et al., 1999; Pensaert et al., 1981; Brown, 2000; Choi et al., 2004). Each of these influenza types is understood to have originated from aquatic avian reservoirs, with type C as the most ancestral type, and types A and B as the closest phylogenetic relatives and clinically significant (with significant seasonal morbidity and mortality in humans). Influenza A virus are further subdivided into serotypes[2] according to its surface glycoproteins, of which at least 17 hemagglutinin (HA) and 10 neuraminidase (NA) antigenically distinguishable subtypes have currently been identified. Each influenza virion contains approximately 500 HA and 200 NA glycoproteins on its surface, in which HA facilitates binding and entry into targeted epithelial cells; NA permits the release of the offspring of infected cells, hence allowing the cellular propagation of the infection (Bouvier & Palese, 2008; Cox & Subbarao, 2000).

The ability of influenza to continuously evolve plays a significant role in its transmission dynamics and evolutionary properties. Its long-term epidemiological success as a result of being an adept and rapidly evolvable pathogen has long been accredited to antigenic variations in HA and NA, in which antigenic variation renders an individual susceptible with no clinical protection to new strains despite previous infection, vaccination and/or exposure to antigenically similar strains. This evolution of influenza viruses via antigenic variation has been understood as a process either of point mutation (antigenic drift) or reassortment (antigenic shift) in the HA and NA genes.

[2]Serotypes are distinct variations within a bacterium or virus species.

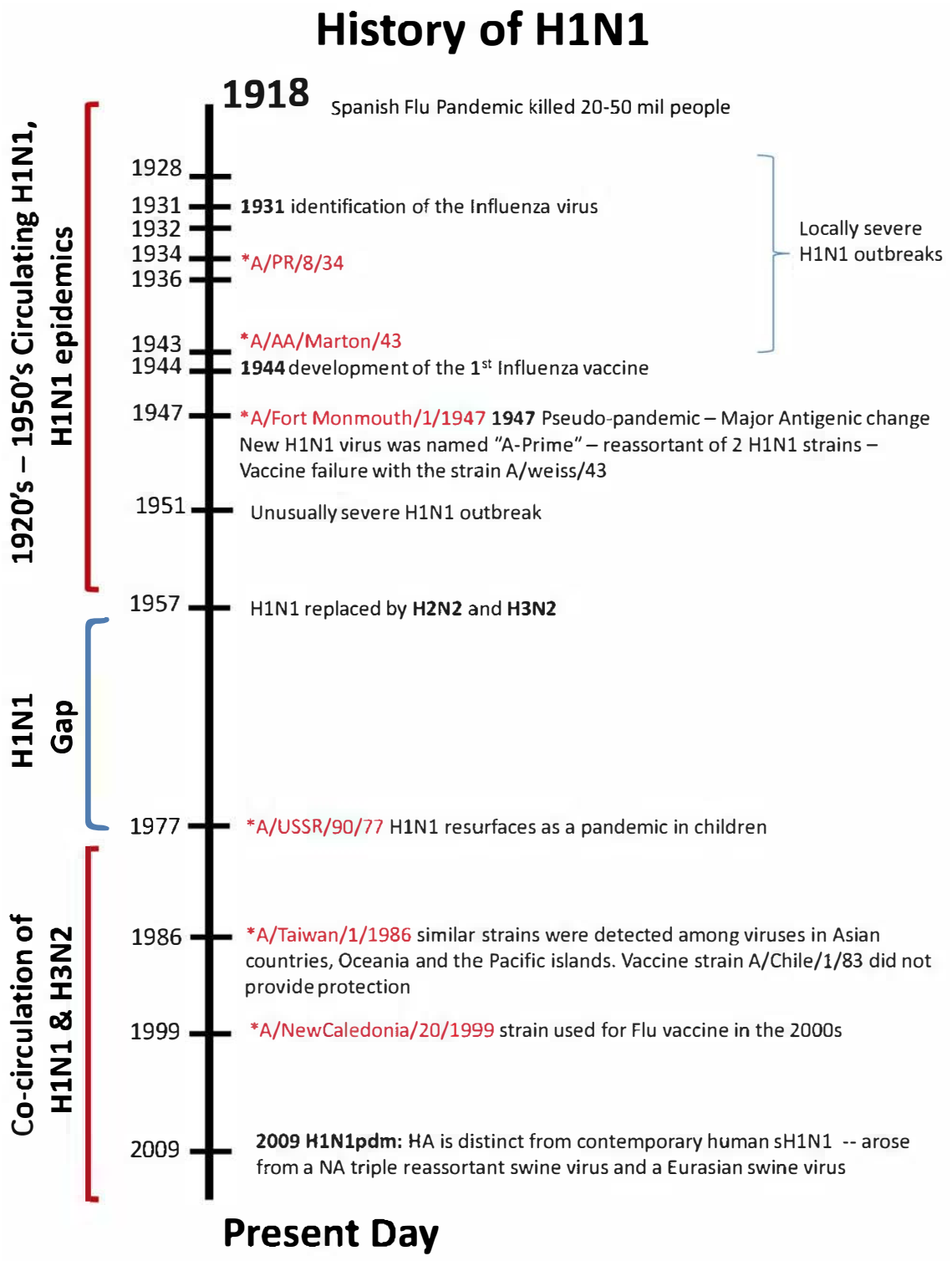


Figure 2.1: **Historical Episodes of H1N1 Influenza epidemics and pandemics** (courtesy of Huang et al., 2013).

2.2.1 Antigenic Drift and Shift in Influenza A

During the process of antigenic drift, a variety of mutations (e.g., substitutions, deletions, insertions) cause accumulation of minor changes in the HA and NA on the surface of the virus in which genetic variations are produced due to the transcription of viral RNA polymerase. This mutational process frequently occurs because of the lack of a proofreading mechanism of the virus' RNA polymerase. Antigenic drift permits the virus to eventually escape neutralization by antibody generated by pre-existing strains, previous infections and/or vaccinations in some hosts (Cox & Subbarao, 2000; Earn et al., 2002; Hampson, 2002; Bouvier & Palese, 2008). Because of minor changes, the new strains are usually antigenically similar, and some individuals will be clinically protected. This has been believed to be responsible for regular flu seasonal epidemics, and the locations and numbers of mutations needed to confer the greatest amount of immune escape has been an important topic (Bush et al., 1999; Fitch et al., 1997; Smith et al., 2004).

Because antigenic drift posits a continuous process where the virus has limited access to an unlimited repertoire of epitopes[3], one will suspect an increase in the number of strains in the population. However, this is not the case since influenza has a limited genetic diversity (Gog & Grenfell, 2002; Gomes et al., 2002; Smith et al., 2004; Ferguson et al., 2003; Minayev & Ferguson, 2009a-b; Koelle et al., 2006; Wikramaratna & Gupta, 2009; for a detailed mini-review).

In contrast, the second mode of variation in Influenza A virus is antigenic shift which is far less frequent and causes genetic reassortment events (Brown et al., 1998; Peiris et al., 2001). The reassortment process occurs during an infection by at least two virions (an entire virus particle), say between human and avian strains, in which the resulting strains acquire new antigens.

2.3 Limited Epitopes in Influenza A

In addition to these processes of variation, a new hypothesis has recently been put forth by Gupta and colleagues (Gupta et al., 1996, 1998; Wikramaratna & Gupta, 2009; Recker et al., 2007). It is called the *antigenic thrift or limited epitope* hypothesis. In antigenic *thrift*, it is hypothesized that new strains are produced by interactions of a fixed set of viral strains within the population with an evolving set of immunities to different viral strains (see Fig. 2.2). The consequence of this hypothesis is that the epitope-specific nature of the immunity restricts (or even prevents) the rise of all antigenic variants with the exception of a particular antigenic strain type consisting

[3]Epitope is antigenic determinant or is the part of an antigen that is recognized by the immune system, specifically by antibodies, B cells, or T cells.

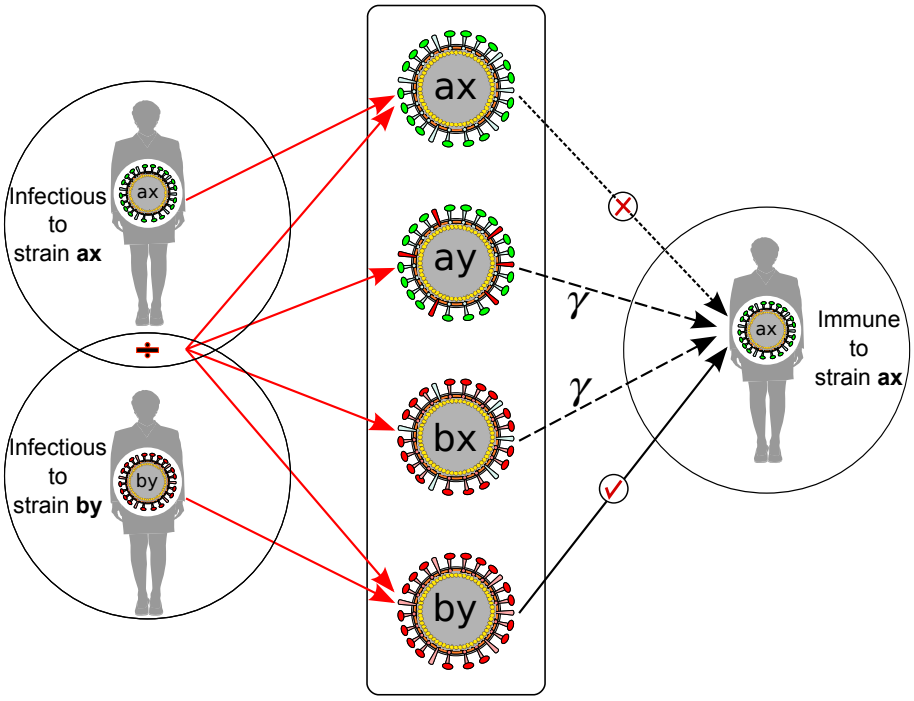


Figure 2.2: **Limited Epitope/Antigenic Thrift Hypothesis.** Schematic representation of a multi-locus-allele system, where loci encode a set of dominant immunogens. For the case where there are two immunologically dominant loci, each with two alleles (e.g., 2-loci-2-allele system), there are four possible strain subtypes **ax, ay, bx** and **by**, where **a** and **b** denote alleles at one locus, and **x** and **y** denote alleles on the other locus. Individuals infected with strains **ax** and **by** are able to generate four genotypes, while those infected with a single subtype can only produce infectious progeny of the same genotype. Those infected with a single strain, say **ax**, will be immune to that same strain in future infection, but will be susceptible to an unrelated strain, say **by**. However, transmission of related strains (e.g. **bx** and/or **ay**) within an individual immune system previously exposed to strain **ax** is determined by a cross-immunity factor γ.

of a set of antigenic variants to which the population is naively susceptible, and future epidemics and pandemics will be composed of different combinations of these limited sets of epitopes instead of emergence of entirely new antigenic types, as in the case of antigenic shift for pandemics.

The epidemiological patterns of this assumption can easily be observed in the appearance and re-appearance phenomena of limited number of strains, and its consequence on the relationship between re-infections, co-existence of subtypes, cluster formation, and viral severity (Smith et al., 2004; Gupta et al., 1998; Recker et al., 2007). When the swine-origin A/H1N1/Mexico was first isolated and reclassified to level 5 pandemic level, most of the associated concerns were due to its similarity to the age-specific structure of the Spanish flu of 1918. In particular, instead of the usual bath-tube structure (or U-shape) of incidence and mortality rates of seasonal strains,

where the very young and old were significantly affected, it exhibited the W-shape that was observed in 1918 H1N1 (Taubenberger & Morens, 2006). However, the incidence and mortality rates of the swine-origin H1N1 were far less severe than those of 1918 H1N1. The consequence of the limited epitope/antigenic thrift hypothesis, in this regard, is the antibody levels associated with both the H1 and N1 may have facilitated the reduction of the swine-origin subtype because of the previous existence of related subtypes. Similarly, if one compares the incidence rates of 1918 H1N1 and H2N2 pandemic subtypes, the susceptibility level was higher in 1918 because there may have been no previous exposure to H1N1, prior to 1918. In contrast, the H2N2 incidence rate was less severe because some exposure to similar strains, such as 1890 H2N2 and 1902 H3N2 subtypes, existed.

In addition to the empirical studies given in support of antigenic thrift or the limited epitopes hypothesis, we return to the recent epidemiological event of the H1N1 pandemic to survey some sero-epidemiological studies. Early findings of swine-origin H1N1 indicate a low level of infection in the older population (65 year of age and above) with the virus and greater susceptibility in people 35 years of age and younger (Allwin et al., 2010; Bandaranayake et al., 2010; Chen et al., 2010; Deng et al., 2010; Gilbert et al., 2010; Miller et al., 2010; Skowronski et al., 2010; Tandale et al., 2010; Wu et al., 2010; Zimmer et al., 2010). In particular, it was demonstrated that the HA gene of 2009 H1N1 belongs to the classical swine lineage, which is antigenically similar to the human H1N1/1918 and to classical swine lineage with greater similarity to the H1N1/Fort Dix/1976. In addition, HA of swine-origin H1N1 of 2009 was antigenically and genetically different from HA of previous seasonal H1N1. This suggests the presence of the cross-reactive immunity in the older population. Hancock et al. (2009) assessed the pre-existing immunity to 2009 H1N1 and showed that children under the age of 10 years have little or no pre-existing immunity to the 2009 H1N1, while individuals over the age of 60 exhibit some level of cross-reactive antibody to the pandemic H1N1 strain of 2009. Similar age-specificity and distribution results were also observed in more extensive studies using serum samples from the general population that spanned 130 years (Hancock et al., 2009), and in experimental studies conducted by Manicassamy et al. (2010) for outbreaks of previous pandemic years. The benefits of these studies are numerous. The use of serological surveillance provides a means of estimating the total number of infected individuals, from which accurate estimates of disease severity, fatality, and the proportion of the population susceptible to future outbreaks can be made. These studies also provide information that can be used to interrogate the biological assumptions about some of these diseases, and to scientifically predict the possibility of future pandemics as the higher frequency of HI antibodies in the population diminishes the likelihood of an outbreak.

2.4 Simple Compartmental Models

In the field of classical epidemic theory, nearly all equations are based on the catalytic compartmental model of Kermack & McKendrick (1927) in which the population is divided into three groups: namely the susceptible $S(t)$, the infected $I(t)$ and recovered $R(t)$. We let N denote the total population, so $N = S + I + R$. The model assumes a constant population with no disease-induced mortality, and death rate μ and birth rate are the same. In addition, all newborns are assumed to be susceptible. Furthermore, individuals have an equal probability of contracting the disease with a rate β (called transmission rate), and the total rate of infection is given by $\mathcal{B}(\beta, S, I)$ with $\mathcal{B}(\beta, 0, I) = \mathcal{B}(\beta, S, 0) = \mathcal{B}(\beta, 0, 0) = 0$, where the functional forms of $\mathcal{B}(\beta, S, I)$ are determined by assumptions on the contact rate and mixing (i.e., mass action incidence, standard incidence). Moreover, infected individuals can recover at a rate σ (recovery rate), where $\frac{1}{\sigma}$ is the period of infection.

With these assumptions, the corresponding equations describing the dynamics are given as follows:

$$\begin{aligned} \frac{dS}{dt} &= \mu N - \mathcal{B}(\beta, S, I) - \mu S, && (2.1)\\ \frac{dI}{dt} &= \mathcal{B}(\beta, S, I) - \mu I - \sigma I, && (2.2)\\ \frac{dR}{dt} &= \sigma I - \mu R, && (2.3) \end{aligned}$$

where the initial conditions are $S(0) = S_0 \geq 0,\ I(0) = I_0 \geq 0,$ and $R(0) = R_0 \geq 0$. Equations (2.1)-(2.3) comprise the standard SIR-type model with demographic terms (i.e., birth-death process). Note that $N = S + I + R$ satisfies $\frac{dN}{dt} = 0$.

Before proceeding, we note that Eqs. (2.1)-(2.3) can also be rewritten in terms of overlapping compartments. That is, we let $\tilde{S}$ still denote the susceptibles, and we define $\tilde{R}$ as the number of hosts who have been infected at some time in their life. Then $\tilde{R}$ is the number of hosts who are immune to the infection and/or are currently infected. Therefore, $\tilde{R} = I + R$, and Eqs. (2.1)-(2.3) can be rewritten as follows:

$$\begin{aligned} \frac{d\tilde{S}}{dt} &= \mu \tilde{N} - \mathcal{B}(\beta, \tilde{S}, I) - \mu \tilde{S}, && (2.4)\\ \frac{d\tilde{R}}{dt} &= \mathcal{B}(\beta, \tilde{S}, I) - \mu \tilde{R}, && (2.5)\\ \frac{dI}{dt} &= \mathcal{B}(\beta, \tilde{S}, I) - \mu I - \sigma I, && (2.6) \end{aligned}$$

where $\tilde{N} = \tilde{S} + \tilde{R}$. The functional structure of $\mathcal{B}(\beta, S, I)$ can be formulated as the mass-action (density dependent) form (e.g., $\mathcal{B}(\beta, S, I) = \beta SI$), or the standard

incidence (frequency-dependent) form ($\mathcal{B}(\beta, S, I) = \beta S(t)\frac{I(t)}{N(t)}$). We will use the mass-action form in the generalization of an SIR model (Andreasen et al., 1997; Ferguson & Andreasen, 2001).

2.5 Generalization of Multiple Strain Model

We now discuss a further extension of the model above to incorporate more additional strains, which was given by Andreasen et al. (1997), and is widely used as the foundational basis for most of the existing models with arbitrary number of strains in the SIR-type structure. In this framework, the population is divided into complex partitioning of host population according to detailed subclasses based on the immunological history of individuals. Let us use an index set notation, where $\mathcal{K} = \{1, 2, \ldots, M\}$ denotes the set of all possible strains. From the SIR-model, $S_\varnothing$ represents the number of individuals who have never been infected, and $S_\mathcal{J}$ denotes the susceptible individuals with immunity to the strains in the set $\mathcal{J}$, where $\mathcal{J} \subseteq \mathcal{K}$. Similarly, $I^i_\mathcal{J}$ represents the number of individuals currently infected by strain i who have previously been exposed to, and have recovered from, infections with all strains in the set $\mathcal{J}$ (i.e., $i \notin \mathcal{J}$).

Here we assume life-long immunity as in the previous section. Furthermore, past infections confer partial protection against future infection with antigenically related strains. Hence, immunity is enhanced with previous infections, and the cross-immunity is determined by the cross-reactivity of the most related previous infections. In other words, rather than developing antibodies to every new antigenically similar epitope that is encountered, the immune system reproduces antibodies against the epitopes of a previously experienced serotype instead of producing new immunity to new antigenic sites. The consequence of this is that either the first experienced strain dominates all subsequent immunological responses rather than developing antibodies to every new epitope that is encountered, provided there are antigenically significant similarities, or the most recently acquired antibodies are important. This assumption is known as *original antigenic sin* (Francis, 1960; Virelizier et al., 1974a-b; Webster, 1968; Webster et al., 1976). Once protective antibodies have been developed prior to the infection, there are two possible outcomes. The first is that partial immunity to the infecting strain reduces the severity and the duration of infection. Secondly, the production of new types of antibodies is reduced. The reduced infectivity assumption neglects the secondary effect of immunity when individuals are completely immune to the infecting strains, whereby individuals with complete immunity are prevented from transmitting the strain while still increasing the repertoire of protective antibodies,

resulting in ultra-immunization of individuals. To circumvent this problem, it is assumed that the cross-reactivity is determined by the most related antigenically strains which might be the first experienced strain or the most recently acquired antibodies. Because immune response is a competition between existing memory B-cells and naive ones for viral antigen, subsequent works have shown that these two assumptions occurs, where the persisting antibodies to the original (i.e., first) strain are numerous to prevent the activation of naive B cells. In addition, it has also been shown that subsequent responses can be activated using the most recently acquired immunity (Kim et al., 2009; Lambertet al., 2005). As a result, we assume in this thesis that, with respect to strain $i \in \mathcal{K}$, there are m levels of increasing protection, the sets $\mathcal{L}_1(i) \subseteq \mathcal{L}_2(i) \subseteq \ldots \subseteq \mathcal{L}_k(i) \subseteq \ldots \subseteq \mathcal{L}_m(i) = \mathcal{K}$, that describe collections of strains eliciting less and less cross-reaction to strain i. Therefore, we assume that additional immunity can only increase cross-protection and hence $\gamma^i_{\mathcal{J}} \leq \gamma^i_{\mathcal{K}}$ where $\gamma^j_{\mathcal{X}}$ is the level of cross-immunity and is defined in reverse sense (i.e., $\gamma^i_{\mathcal{J}} = 1 - \gamma$). Inheriting all the assumptions made in the case of the standard SIR model, the equations for multiple co-circulating strains that extend Eqs. (2.1)-(2.3) under the reduced susceptibility assumption are given as follows:

$$\frac{dS_{\varnothing}}{dt} = \mu N - \sum_{j \in \mathcal{K}} \Lambda^j S_{\varnothing} - \mu S_{\varnothing}, \tag{2.7}$$

$$\frac{dS_{\mathcal{J}}}{dt} = \sum_{j \in \mathcal{J}} \sigma I^j_{\mathcal{J} \setminus j} - \sum_{i \notin \mathcal{J}} \gamma^i_{\mathcal{J}} \Lambda^i S_{\mathcal{J}} - \mu S_{\mathcal{J}}, \tag{2.8}$$

$$\frac{dI^i_{\mathcal{J}}}{dt} = \gamma^i_{\mathcal{J}} \Lambda^i S_{\mathcal{J}} - \mu I^i_{\mathcal{J}} - \sigma I^i_{\mathcal{J}}, \tag{2.9}$$

$$\Lambda^i = \beta_i \sum_{\mathcal{J} \subseteq 2^{\mathcal{K} \setminus i}} I^i_{\mathcal{J}}, \text{ and } N = S_{\varnothing} + \sum_{\mathcal{J}} S_{\mathcal{J}} + \sum_{i,\mathcal{J}} I^i_{\mathcal{J}}, \tag{2.10}$$

where β_i is the transmissibility or infection rate, σ denotes the recovery rate, $\gamma^i_{\mathcal{J}}$ is the cross-immunity level, and Λ^i is the force of infection. Here we use reversed cross-immunity for notational purposes. That is, $\gamma^i_{\mathcal{J}} = 0$ implies complete protection, and $\gamma^i_{\mathcal{J}} = 1$ implies no protection (e.g., $\gamma^i_{\varnothing} = 1$ for all i). In Eq. (2.7), the inflow is due to birth at a rate μN and the outflows are due to death and infections. In Eq. (2.8), there are two outflows and one inflow. The inflow is given by the sum of all recoveries from infections with different strain i for individuals with immunological profile $\mathcal{J}$ after recovery, and the outflows are again due to deaths $\mu S_{\mathcal{J}}$ and due to infections with strains not yet experienced, $i \notin \mathcal{J}$. Similarly, for the third equation (see Eq. (2.9)), the outflow is determined by death and recovery, and the inflow comes from infection. Note that the cross-immunity $\gamma^i_{\mathcal{J}}$ reduces the force of infection for strain i for susceptibles $S_{\mathcal{J}}$ who have previously been infected with strains in $\mathcal{J}$. The growth

rate of infection and the prevalence due to infection in $S_{\mathcal{J}}$ are also reduced. We note that $\dfrac{dN}{dt} = 0$, hence, N is constant. In addition, we observe that Eq. (2.9) can be reformulated in terms of force of infection Λ^i as follows

$$\frac{d\Lambda^i}{dt} = \beta_i \sum_{\mathcal{J}\subseteq 2^{\mathcal{K}\setminus i}} \Lambda^i S_{\mathcal{J}} - \mu\Lambda^i - \sigma\Lambda^i. \tag{2.11}$$

Generalizing Eqs. (2.1)-(2.3) yields an analogous system of equations for multiple co-circulating strains under reduced infectivity. Using Eqs. (2.4)-(2.6), we have:

$$\frac{dS_{\varnothing}}{dt} = \mu N - \sum_{j\in\mathcal{K}} \Lambda^j S_{\varnothing} - \mu S_{\varnothing}, \tag{2.12}$$

$$\frac{dS_{\mathcal{J}}}{dt} = \sum_{j\in\mathcal{J}} \Lambda^j S_{\mathcal{J}\setminus j} - \sum_{i\notin\mathcal{J}} \Lambda^i S_{\mathcal{J}} - \mu S_{\mathcal{J}}, \tag{2.13}$$

$$\frac{d\Lambda^i}{dt} = \sum_{\mathcal{J}\subseteq\mathcal{K}} \tau^i_{\mathcal{J}} \Lambda^i S_{\mathcal{J}} - \mu\Lambda^i - \sigma\Lambda^i, \tag{2.14}$$

where $\tau^i_{\mathcal{J}}$ is also a cross-immunity level which reduces the infectivity of individuals previously infected with similar strains. Combining Eqs. (2.7)-(2.11) with Eqs. (2.12)-(2.14), we can obtain trivially a more general model which incorporates both reduced susceptibility and infectivity (Ferguson & Andreasen, 2001):

$$\frac{d\tilde{S}_{\varnothing}}{dt} = \mu N - \sum_{j\in\mathcal{K}} \Lambda^j \tilde{S}_{\varnothing} - \mu\tilde{S}_{\varnothing}, \tag{2.15}$$

$$\frac{d\tilde{S}_{\mathcal{J}}}{dt} = \sum_{j\in\mathcal{J}} \gamma^j_{\mathcal{J}\setminus j} \Lambda^j \tilde{S}_{\mathcal{J}\setminus j} - \sum_{i\notin\mathcal{J}} \gamma^i_{\mathcal{J}} \Lambda^i \tilde{S}_{\mathcal{J}} - \mu\tilde{S}_{\mathcal{J}}, \tag{2.16}$$

$$\frac{d\Lambda^i}{dt} = \beta_i \sum_{\mathcal{J}\subseteq\mathcal{K}} \gamma^i_{\mathcal{J}} \tau^i_{\mathcal{J}} \Lambda^i \tilde{S}_{\mathcal{J}} - \mu\Lambda^i - \sigma\Lambda^i. \tag{2.17}$$

These equations describe the general multiple strain SIR system and are used as a basis for the majority of the models investigating multiple co-circulating strains. The differences between the models are usually in the assumptions and formulation of the structure of cross-immunity, $\gamma^i_{\mathcal{J}}$. The qualitative dynamics are similar, and it appears that the dynamics are insensitive to the chosen assumption (Ferguson & Andreasen, 2001; Dawes & Gog, 2002). In this chapter and throughout this thesis, we assume reduced infectivity in part because this is the one form of cross-immunity that can be used to decrease the dimensionality of the equations under appropriate auxiliary epidemiological quantities without imposing some form of closure assumptions (see Section 2.4). To discuss the dynamics of Eqs. (2.7)-(2.11), or equivalently Eqs. (2.12)-(2.14), one must impose or assume a particular structure on the host immunity

which mediates strain competition through cross-immunity. Therefore, steady state analysis beyond the disease-free equilibrium is not trivial and may not be feasible (see Elveback et al., 1964; Dietz, 1979; Castillo-Chavez et al., 1989; Lloyd & Wodarz, 2006; Nuno et al., 2009 for discussion of the system under a particular cross-immunity and uniformity of strains (i.e., $\beta_i = \beta_j$ and all strains share the same properties)). Rather than discussing the dynamics of the general system herein, the aim of this section is to show the general relationship between the above approaches (i.e., reduced susceptibility and infectivity) in the context of history-based framework, and to derive the model of Gupta et al. (1998) analyzed in Chapter 3 using an SIR-type model under reduced infectivity (see Gupta et al., 1997a-b for different derivation approach).

2.6 Model Reductions

Understanding the qualitative dynamics of multiple strain systems using a more general abstraction is extremely difficult, and analytically and numerically cumbersome. One of the major problems arising from the extension of the SIR framework is that the number of state variables and parameters increases exponentially, resulting in both computational and conceptual challenges. To circumvent this problem, there have been various attempts that introduce auxiliary quantities to reduce the higher dimensionalities of both state variables and parameter variables.

2.6.1 Auxiliary Quantities

In Section 2.5, we outlined two approaches both using an SIR-type formulation with minimal strain-centric indexing. The first approach uses an SIR model with the reduced susceptibility assumption (i.e., Eqs. (2.7)-(2.10)), while the second approach employs the reduced infectivity assumption (i.e., Eqs. (2.12)-(2.14)). In this section, we will focus only on the reduced infectivity model because of state space dimensional reducibility. To reduce the dimensionality of the systems, we employ auxiliary quantities.

The use of auxiliary quantities allows us to define some epidemiological quantities of interest that can be found by summing over the dynamic equations. One example is using the total number of susceptibles, $S = \sum_{\mathcal{J}} S_{\mathcal{J}}$. For instance, for the dynamic equations in Sec. 2.5, we can define $S = \sum_{\mathcal{J}} S_{\mathcal{J}} = \sum_{\mathcal{J} \in 2^{\mathcal{K}}} S_{\mathcal{J}} = \sum_{\mathcal{J} \subseteq \mathcal{K}} S_{\mathcal{J}}$. Similarly, we can also use force of infection for each strain i as an auxiliary quantity. That is, $\Lambda^i = \beta_i \sum_{J \subseteq \mathcal{K} \setminus i}$. However, the auxiliary quantity S does not provide any insight into the dynamics of interacting strains, and the second auxiliary quantity described by the force of infection does not reduce the number of dynamic equations (see Eqs.

(2.7)-(2.10) or Eqs. (2.12)-(2.14)). There are other auxiliary quantities that can be used, and are listed below:

1. $S_i = \sum_{\mathcal{J}:i\in\mathcal{J}} S_{\mathcal{J}}$: The number of susceptibles who have experienced at least strain i.

2. $W_k^i = \sum_{\mathcal{L}_k(i)\bigcap\mathcal{J}\neq\emptyset} S_{\mathcal{J}}$: The number of individuals who have immunity to one or more strains in the $\mathcal{L}_k(i)$-neighborhood of i, for all $k = 0, 1, 2, \ldots, m$ levels of protection, where we define $\mathcal{L}_k(i)$ as follows: $\mathcal{L}_k(i) = \{j \in \mathcal{K} : d_{\mathcal{H}}(i, j) \leq k, k \in \mathbb{Z}_+\}$ with $d_{\mathcal{H}}(i, j)$ being the Hamming distance (i.e., the number of positions at which the corresponding i and j are different).

3. $S_{\mathcal{M}} = \sum_{\mathcal{J}\subseteq\mathcal{M}} S_{\mathcal{J}}$: Total number of susceptible individuals with immunity only to strain in $\mathcal{M}$.

4. $T_{\mathcal{M}} = \sum_{\mathcal{J}\subseteq\mathcal{M}} m_{\mathcal{J}} S_{\mathcal{J}} + \sum_{\mathcal{J}\subseteq\mathcal{M}} \sum_{i\notin\mathcal{J}} m_{\mathcal{J}} I^i_{\mathcal{J}},\ m_{\mathcal{J}} = \sum_{j\notin\mathcal{J}} \gamma^j_{\mathcal{J}}$: Number of individuals who are immune only to some of the strains in $\mathcal{M}$ weighted by their remaining susceptibility.

5. $N_{\mathcal{M}} = \sum_{\mathcal{J}\subseteq\mathcal{M}} S_{\mathcal{J}} + \sum_{\mathcal{J}\subseteq\mathcal{M}} \sum_{i\notin\mathcal{J}} I^i_{\mathcal{J}}$: Total number of individuals who have been exposed only to some of the strains in $\mathcal{M}$ (i.e. individuals that have not recovered from infections with any of the strains in $\mathcal{K}\backslash\mathcal{M}$).

Along with the above quantities, other auxiliary quantities, such as $R_{\mathcal{M}}$ and R_i denoting respectively the number of individuals who are immune to all strains in $\mathcal{M}$, and the total number of the population that is immune to strain i, can be used as auxiliary variables. In the next section, we will use S_i and W_k^i with the overlapping compartmentalization (i.e., Eqs. (2.12)-(2.14)) of the system to reduce the dimension, and later show that the dynamic equation for S_i is related to that of W_k^i. The definition of W_k^i allows us to biologically-motivated assumption on the relationship between strains as defined by their epitopes using the notion of Hamming distance.

2.6.2 Reduction of Dynamic Eqs. (2.12)-(2.14)

To reduce the dynamic equations obtained in Section 2.5, we use the first auxiliary variable $S_i = \sum_{\mathcal{J}:i\in\mathcal{J}} S_{\mathcal{J}}$ for susceptible individuals who have experienced at least strain i. To derive a dynamic expression for S_i, we use Eq. (2.8) and sum over all $i \in \mathcal{J}$. That is,

$$\begin{aligned}\sum_{\mathcal{J}:i\in\mathcal{J}} \frac{dS_{\mathcal{J}}}{dt} &= \sum_{\mathcal{J}:i\in\mathcal{J}} \sum_{j\in\mathcal{J}} \Lambda^j S_{\mathcal{J}\backslash j} - \sum_{\mathcal{J}:i\in\mathcal{J}} \sum_{j\notin\mathcal{J}} \Lambda^j S_{\mathcal{J}} - \sum_{\mathcal{J}:i\in\mathcal{J}} \mu S_{\mathcal{J}},\\ &= \sum_{\mathcal{Y}:i\notin\mathcal{Y}} \Lambda^i S_{\mathcal{Y}} + \sum_{\mathcal{J}:i\in\mathcal{J}} \sum_{j\in\mathcal{J},j\neq i} \Lambda^j S_{\mathcal{J}\backslash j} - \sum_{\mathcal{J}:i\in\mathcal{J}} \sum_{j\notin\mathcal{J}} \Lambda^j S_{\mathcal{J}} - \mu \sum_{\mathcal{J}:i\in\mathcal{J}} S_{\mathcal{J}}\end{aligned}$$

$$= \sum_{\mathcal{J}\subseteq\mathcal{K}} \Lambda^i S_{\mathcal{J}} - \sum_{\mathcal{J}:i\in\mathcal{J}} \Lambda^i S_{\mathcal{J}} - \mu S_i = (N - S_i)\,\Lambda^i - \mu S_i. \tag{2.18}$$

We now use the second auxiliary variable W_k^i, where $W_k^i = \sum_{\mathcal{L}_k(i)\bigcap\mathcal{J}\neq\emptyset} S_{\mathcal{J}}$ denotes the number of individuals who have immunity to one or more strains in the $\mathcal{L}_k(i)$-neighborhood of i for all k levels of protection (i.e., $k = 0, 1, 2, \ldots, m$). For instance, W_0^i represents the number of individuals who have been exposed, and are now immune, to strain i, W_1^i denotes the number of individuals who have been exposed to strain i or any strain related to i, and so on. So by definition $W_m^i = N$. Then we can reduce the dynamic equations in Sec. 2.5 (i.e., Eqs. (2.12)-(2.14)), which can be rewritten in more compact forms. While observing that the number of individuals who do not have $\mathcal{L}_k(i)$ immunity is $N - W_k^i$, we sum the dynamic equation for $S_{\mathcal{J}}$ over $\mathcal{L}_k(i)\bigcap\mathcal{J} \neq \emptyset$, to obtain the following expressions:

$$\frac{dW_k^i}{dt} = \left(N - W_k^i\right) \sum_{j\in\mathcal{L}_k(i)} \Lambda^j - \mu W_k^i, \tag{2.19}$$

$$\frac{d\Lambda^i}{dt} = \beta_i \Lambda^i \sum_{k=1}^{m} \tau_k \left(W_k^i - W_{k-1}^i\right) - (\mu + \sigma)\,\Lambda^i, \tag{2.20}$$

where $\tau_{\mathcal{J}}^i = \tau_k$ if $\mathcal{J}\bigcap(\mathcal{L}_k(i)\backslash\mathcal{L}_{k-1}(i))$, and $0 = \tau_0 \leq \tau_1 \leq \tau_2 \leq \ldots \leq \tau_m = 1$, $W_m^i = N$. Here τ_k reduces disease transmissibility for individuals whose most related infection is at k-level (Ferguson & Andreasen, 2001). We note that when $m = 2$, Eqs. (2.19)-(2.20) reduce to,

$$\frac{dW_1^i}{dt} = \left(N - W_1^i\right) \sum_{j\in\mathcal{L}_1(i)} \Lambda^j - \mu W_1^i, \tag{2.21}$$

$$\frac{d\Lambda^i}{dt} = \beta_i \Lambda^i \left[\tau_1 \left(W_1^i - W_0^i\right) + \left(N - W_1^i\right)\right] - (\mu + \sigma)\,\Lambda^i, \tag{2.22}$$

$$\frac{dW_0^i}{dt} = \left(N - W_0^i\right)\Lambda^i - \mu W_0^i. \tag{2.23}$$

Note that these equations are closely related to the model studied by Gupta and colleagues (Gupta et al., 1996, 1998), except that τ_k is defined in the opposite sense (i.e., reverse cross-immunity, $\tau_1 = 1 - \tau$) and N was normalized to unity without loss of generality (i.e., $N = 1$). The benefit of using Eqs. (2.19)-(2.20) over Eqs. (2.7)-(2.10) or Eqs. (2.12)-(2.14) is that Eqs. (2.19)-(2.20) reduce the dimension of the dynamical system from $2^M + M2^{M-1}$ or $2^M + M$ to $M\,(1 + m)$, where M is the number of strains and it is related to the number alleles in m epitopes. That is, defining strains by their epitopes, M is the product of the alleles. For example, let pathogen strains be defined by a finite (say up to three) immuno-dominant loci, a, b and c with n_a, n_b and n_c number of alleles, respectively. Then, strains are

characterized by the combination of $a_1 \dots a_{n_a}$, $b_1 \dots b_{n_b}$ and $c_1 \dots c_{n_c}$, and the total number of strains is $n_a \times n_b \times n_c$. In general, there are $M := \prod_{i=1}^{m} n_i$ strains, where m denotes the number of loci and n_i denotes the number of alleles at the ith locus. For instance, for a system with say $m = 3$ with 2 alleles each (i.e., $\{2, 2, 2\}$), then $M = 2 \times 2 \times 2 = 8$ (see Chapter 3). Furthermore, Eqs. (2.19)-(2.20) are easier to analyze and/or work with computationally. As the number of strains increases, the complexity of Eqs. (2.19)-(2.20) increases linearly or algebraically.

Let Ω_{ij} be the similarity and relatedness matrix, where $\Omega_{ij} = 1$ if strain i is similar or related to strain j or $i = j$ and $\Omega_{ij} = 0$ otherwise. Then we can rewrite $\sum_{j \in \mathcal{L}_1(i)} \Lambda^j = \sum_j \Omega_{ij} \Lambda^j$, and Eq. (2.21) becomes

$$\frac{dW_1^i}{dt} = \left(N - W_1^i\right) \sum_j \Omega_{ij} \Lambda^j - \mu W_1^i. \tag{2.24}$$

Furthermore, we observe that one can also reformulate Eqs. (2.21)-(2.23) using another set of auxiliary expressions. Let $\tilde{S}_i = N - W_1^i$ denote the number of individuals who do not have immunity to strain i, $V_i = W_1^i - W_0^i$ represent individuals with partial immunity to strain i, and $Z_i = W_0^i$ denote individuals fully immune to strain i, then the above equations become:

$$\frac{d\tilde{S}_i}{dt} = \mu N - \tilde{S}_i \sum_j \Omega_{ij} \Lambda^j - \mu \tilde{S}_i, \tag{2.25}$$

$$\frac{dV_i}{dt} = \tilde{S}_i \sum_{j \neq i} \Omega_{ij} \Lambda^j - V_i \Lambda^i - \mu V_i, \tag{2.26}$$

$$\frac{d\Lambda^i}{dt} = \beta_i \Lambda^i \left[\tau_1 V_i + \tilde{S}_i\right] - (\mu + \sigma) \Lambda^i, \tag{2.27}$$

$$\frac{dZ_i}{dt} = \left(\tilde{S}_i + V_i\right) \Lambda^i - \mu Z_i. \tag{2.28}$$

However, if we use a locus-allele centric representation, we can rewrite the equations above using two indices where, say for the 2-locus system, the first index represents the first locus and the second index represents the second locus. For the 2-locus system, $\sum_j \Omega_{ij} \Lambda^j$ can be rewritten as $\sum_j \Omega_{ij} \Lambda^j := \sum_l \Lambda^{lj} + \sum_k \Lambda^{ik} - \Lambda^{ij}$. Then Eqs. (2.21)-(2.23) become

$$\frac{dW_1^{ij}}{dt} = \left(N - W_1^{ij}\right) \left(\sum_l \Lambda^{lj} + \sum_k \Lambda^{ik} - \Lambda^{ij}\right) - \mu W_1^{ij}, \tag{2.29}$$

$$\frac{d\Lambda^{ij}}{dt} = \beta_i \Lambda^{ij} \left[\tau_1 \left(W_1^{ij} - W_0^{ij}\right) + \left(N - W_1^{ij}\right)\right] - (\mu + \sigma) \Lambda^{ij}, \tag{2.30}$$

$$\frac{dW_0^{ij}}{dt} = \left(N - W_0^{ij}\right) \Lambda^{ij} - \mu W_0^{ij}. \tag{2.31}$$

Similarly, Eqs. (2.25)-(2.28) can be rewritten as,

$$\frac{d\tilde{S}_{ij}}{dt} = \mu N - \tilde{S}_{ij}\left(\sum_{l}\Lambda^{lj} + \sum_{k}\Lambda^{ik} - \Lambda^{ij}\right) - \mu\tilde{S}_{ij}, \tag{2.32}$$

$$\frac{dV_{ij}}{dt} = \tilde{S}_{ij}\left(\sum_{l\neq i}\Lambda^{lj} + \sum_{k\neq j}\Lambda^{ik}\right) - V_{ij}\Lambda^{ij} - \mu V_{ij}, \tag{2.33}$$

$$\frac{d\Lambda^{ij}}{dt} = \beta_{ij}\Lambda^{ij}\left[\tau_1 V_{ij} + \tilde{S}_{ij}\right] - (\mu+\sigma)\,\Lambda^{ij}, \tag{2.34}$$

$$\frac{dZ_{ij}}{dt} = \left(\tilde{S}_{ij} + V_{ij}\right)\Lambda^{ij} - \mu Z_{ij}. \tag{2.35}$$

The multi-loci-alleles system is very easy to study, especially when we want to study analytically higher strain structures. The convenience of using multiple locus-allele lexicographic notion is that we can take advantage of the *Mendelian* structure and the inherent symmetries to simplify and reduce the analysis into a lower dimensional dynamical system (i.e., 2-locus-2-allele) and then extend these analyses to a more general n-locus-m-allele system by using the phenotypic space. Equations (2.25)-(2.28) are sometimes more useful than Eqs. (2.29)-(2.31). In particular, in order to understand the immunological profile of the population, using Eqs. (2.25)-(2.28) or Eqs. (2.32)-(2.35) will be more appropriate since $S_i + \tau_1 V_i$ or $S_{ij} + \tau_1 V_{ij}$ provides a measure of population immunity. It measures how much infectivity could be created if all hosts were infected. However, the relevance of this measure is in the study of age-specificity where understanding age-specificity in immunological profile is important; and Eqs. (2.32)-(2.35) will be used in Chapters 5-6 instead of Eqs. (2.29)-(2.31), hence their inclusion in this section. Using multi-locus formulation and the inherent symmetries, extending the analysis of the lower dimensional strain spaces to higher strain spaces becomes much more convenient since the symmetries in the *Mendelian* strain spaces persist in all dimensions (hyperoctahedral group). Additionally, these lower dimensional *Mendelian* structures serve as building blocks in the higher dimensional strain spaces.

2.7 Conclusion

We have provided a brief overview of epidemic models ranging from the catalytic SIR model to high dimensional extensions incorporating higher strain numbers. In the last two sections, we focused on a general model based on the SIR framework, and then showed that the model can further be reduced using a particular or appropriate auxiliary quantity which produces a hierarchical structure with overlapping compartments. The reduced model has a similar structure to the model described by Gupta

et al. (1996, 1998) and others (Ferguson & Andreasen, 2001; Recker et al., 2007). In the proceeding chapters, we explore the reduced model described in Sec. 2.6.2 and its age-specificity in Chapter 4. Chapter 3 analyzes Eqs. (2.29)-(2.31) to obtain insight into the dynamics of strain structure and into the mathematical underpinning of the various numerical studies that have been carried out by Gupta and colleagues (Gupta et al., 1997a-b, 1998; Recker et al., 2007). In particular, we focus on local properties of the steady states. In Chapters 5-6, an age-specific extension of the second formulation in Eqs. (2.25)-(2.28) is used, for this formulation facilitates the use of abstract and concrete approaches to the steady-state analysis of the age-structured model. In the next chapter, we provide the analysis of the multi-locus-allele model.

3 Mathematical Analysis of Multi-locus-allele System

In this chapter, we revisit one of the multi-strain models that has been used to investigate the dynamics of infectious diseases with co-circulating serotypes, and provide analytical results underlying the numerical investigations. In particular, we establish the necessary conditions for the local asymptotic stability of the steady states and the existence of oscillatory behaviours.

3.1 Introduction

To understand and explain the mechanisms underlying antigenic variation and diversity maintenance on co-evolution of pathogens, several researchers have formulated mathematical models incorporating the diversifying selection mediated by host immunity. All of these mathematical models use the standard compartmentalized structure of the SIR model in a metapopulation framework incorporating multiple strain groups (Andreasen et al., 1997; Gupta et al., 1998; Gog & Swinton, 2002). These theoretical models link infectious disease dynamics to antigenic diversity using different modelling assumptions in order to show how immunological cross-reactive strains invade a partially protected susceptible. The pioneering work of Elveback et al. (1964) provided one of the first Monte Carlo simulations of multiple strain dynamics; Dietz (1979) provided the mathematical description of Elveback et al. (1964) which was later modified by Castillo-Chavez et al. (1989) to include cross-immunity in both discrete and continuous age-structured models. Similar extensions and generalizations of these models have appeared in the literature (Andreasen et al., 1997; Lin et al., 1999; Plotkin et al., 2002; Gog & Grenfell, 2002; Cliff et al., 1986; Gupta et al., 1994, 1996; Gupta & Anderson, 1999; Gupta & Galvani 1999; Buckee et al., 2004; Buckee & Gupta 2010; Recker & Gupta, 2005; Calvez et al., 2005; Ferguson & Andreasen, 2001; Ferguson et al., 2003; Sasaki & Haraguchi, 2000; Nuno et al., 2009) and are characterized by one of the two dichotomies based on the way in which competition

between strains is incorporated or how the properties of cross-immunity are mediated within the host and the ecology of the pathogens.

The first dichotomy is related to the bookkeeping of the phenotypic structure of the pathogen of interest. In particular, there have been two commonly used approaches to keep track of either the immunological history or status of the hosts. The history-based framework characterizes hosts by their immunological histories, whereby the hosts may have some protection against any antigenically similar strains to which they have previously been exposed (Castillo-Chavez et al., 1989; Elveback et al.,1964; Dietz, 1979; Andreasen et al., 1997; Gomes et al., 2002; Gupta et al., 1996, 1998; Lin et al., 1999; Recker et al., 2007) and the hosts are grouped into classes by their immunological histories or disease histories defined in terms of all the phenotypes to which an individual has ever been exposed. In contrast, the status-based framework highlights the importance of the current immune state and is immunologically path-independent (Gog & Swinton 2002; Kryazhimsky et al., 2007). That is, the status-based approach is only concerned with the current, but not previous, immunological memories defined in terms of all strains against which an individual has gained immunity.

The second dichotomy is based on how the properties of cross-immunity are mediated within the hosts. That is, whether a secondary infection reduces the chance of an immune individual becoming re-infected or the infectivity of an immune individual. These are known as either reduced susceptibility or reduced infectivity, respectively. Mathematically, models constructed under both susceptibility and infectivity reduction assumptions within the history-based framework have been shown to exhibit similar qualitative behaviours (see Ferguson & Andreasen, 2001; Dawes & Gog, 2002; for dynamic similarity and differences in the various underlying modelling assumptions). On the other hand, status-based models require one to impose assumptions on the nature of susceptibility and infectivity reductions. In addition, the reduced susceptibility models may yield qualitative and quantitative behaviours that are different from reduced infectivity models. History-based models require fewer assumptions about the biology, whereas status-based models force one to make a choice between the two extreme assumptions: reduced susceptibility (Kryazhimsky et al., 2007) and reduced infectivity (Gupta et al., 1998; Recker et al., 2007). The reduced susceptibility and infectivity assumptions affect how susceptible hosts are infected and how infectious hosts infect others, respectively. These are usually formulated either via product cross-immunity whereby hosts gain immunity as a result of repeated exposures with antigenically similar strains, or via minimal cross-immunity whereby host susceptibility is determined to be the least that would result from previous exposures.

Whilst there have been some mathematical studies establishing generic stability results for multiple strain systems, generalization of the results usually becomes intractable because the number of possible equilibria increases exponentially. For instance, for the multi-locus-allele system, the number of possible equilibria is 2^N, where N is the number of strains. It should be noted that this is not true in general. Given the difficulty in fully analyzing these models, there have been various approaches designed for studying both their steady state solutions and dynamic features. For instance, Andreasen et al. (1997) and Lin et al. (1999) considered, respectively, four-strain and three-strain epidemic models with partial cross-immunity between strains. For both models, the authors studied the stability of two of the boundary equilibria (equilibria representing discordant or unrelated strains for the former model) and the coexistence equilibrium. Dawes & Gog (2002) also considered a four-strain system and, using an unfolding approach to study the bifurcation structure of the system, found similar behaviours to those in the aforementioned papers of Andreasen et al. (1997) and Lin et al. (1999). Recker & Gupta (2005) have also considered a variant model of Gupta et al. (1998) and studied the stability of the partially synchronous and fully synchronous steady states via an ε-fluctuation method and numerically showed the existence of cyclic dynamics. More recently, Chan & Yu (2013) and Blyuss (2013) used the groupoid formalism of Golubitsky et al. (2005) and centre-manifold reduction, and an equivariant dynamical system approach to study synchronous clustering and symmetric dynamics, respectively. In the proceeding sections, we present a different approach to perform a detailed study of the multi-locus-allele system of Gupta et al. (1996, 1998) with particular emphases on stability analysis of the equilibria and the existence of a Hopf bifurcation.

3.2 Model Description: Multi-locus-allele Systems

To incorporate multiple groups into the standard compartmentalization, strain structures are typically imposed, based on assumptions about both immunological and genetic processes (Gog & Grenfell, 2002; Dawes & Gog, 2002; Gog & Swinton, 2002). In order to circumvent the imposition of strain structures and to simplify the descriptions of various epidemiological groupings and status, Gupta et al. (1996, 1998) and Recker et al. (2007) introduced the multi-locus-allele system, where pathogenic strains are defined by a unique major epitope (locus) and a set of non-unique minor epitopes (allele). The major epitope encodes a set of dominant *immunogens* that elicit the strongest epitope-specific immune responses which, in turn, profoundly affect transmission successes. In this framework of a multi-locus-allele system, the

combination of alleles at different loci constitutes a strain where each strain can incorporate either asymmetric (differential) or symmetric (uniform) numbers of allelic combinations. For example, let pathogen strains be defined by a finite (say up to three) immuno-dominant loci, a, b and c with m_a, m_b and m_c number of alleles, respectively. Then, strains are characterized by the combination of $a_1 \dots a_{m_a}$, $b_1 \dots b_{m_b}$ and $c_1 \dots c_{m_c}$, and the total number of strains is $m_a \times m_b \times m_c$. In general, there are $N := \prod_{i=1}^{n} m_i$ strains, where n denotes the number of loci and m_i denotes the number of alleles at the ith locus. We denote the case when $m_i \neq m_j,\ i \neq j$ as differential allelic (combination) variations, and $m_i = m_j,\ i \neq j$ as a uniform allelic combination. In the case of uniform allelic combination, where each locus has the same number of alleles, $m_i = m$ and the total number of strains is given by $N = m^n$.

Herein, we denote the general multi-locus-allele system by $\mathcal{G}_n(m_1, m_2, \dots, m_n)$. Different alternatives of this model have been used in various scenarios investigating infectious diseases ranging from dengue (Wikramaratna et al., 2010), *Neisseria meningitidis* (Gupta et al., 1996) to Influenza A virus (Gupta et al., 1998; Recker et al., 2007, Buckee et al., 2004, 2008, 2011). This framework automatically inherits some of the internal intrinsic order associated with genetic processes with immunological distance based on *allelomorphic* similarity. For example, in the case of the $\mathcal{G}_2(2, 2)$ (e.g., 2-locus-2-allele) system mentioned above, there are four strains **ax**, **ay**, **bx** and **by**. Using a multi-dimensional graph where the number of loci determines the dimension of the strain space and vertices represent the strain, the structural space of four strains in the case of $\mathcal{G}_2(2, 2)$ is a square, while the structure spaces for eight strains and for sixteen strains organized on the $\mathcal{G}_3(2, 2, 2)$ and $\mathcal{G}_4(2, 2, 2, 2)$ systems, respectively, are three-dimensional and four-dimensional hypercubes. In general, the strain space structure of n loci and m alleles is an n-dimensional hypercube with m^n strains, with each sharing alleles with $m^n - (m-1)^n - 1$ other strains. In addition, when the allelic combination is differential (asymmetric), the structures of the strain spaces are represented as lattice graphs. This multi-locus-allele framework provides a one-to-one correspondence between the pathogen genotypic structure and antigenic phenotype where each locus, as defined above, is assumed to encode antigenically dominant epitopes (e.g., HA and NA in Influenza A virus). This simple model (Gupta et al., 1996, 1998) and other extensions (Recker et al., 2007) have been numerically studied extensively to provide valuable insights into pathogen formation, diversity maintenance and genetic organization of strains.

We focus on the $\mathcal{G}_2(2, 2)$ system for the sake of simplicity and provide a detailed analysis of the model and mathematical results underlying the numerical results of Gupta et al. (1996, 1998). However, we will introduce additional parameters into the model (to become apparent later) in order to make the results easily extendable

to a more general multi-locus-allele system $\mathcal{G}_n(m_1, m_2, \ldots, m_n)$ when the symmetric structures of strain space are considered. To study the maintenance of strain diversity using the aforementioned framework, Gupta et al. (1996, 1998) characterized individuals based on their immunological history and divided the host population into three partially overlapping epidemiological compartments for each strain j: the proportion of the population that has been exposed to pathogenic type j and/or is now either infected or recovered (immune to strain j), z_j; the proportion of the individuals who have been exposed and are immune to any antigenic type sharing alleles with strain j, w_j (note that $w_j = \bigcup_{j\sim i} z_j$, where $j \sim i$ means j shares alleles with i) and the proportion of the population infectious with strain j, y_j. The dynamics of these compartments are given by the following set of differential equations:

$$\frac{dz_j}{dt} = \beta(1 - z_j)y_j - \mu z_j \tag{3.1}$$

$$\frac{dw_j}{dt} = \beta(1 - w_j)\sum_{j\sim i} y_i - \mu w_j \tag{3.2}$$

$$\frac{dy_j}{dt} = \beta\left[(1 - w_j) + (1 - \gamma)(w_j - z_j)\right] y_j - (\mu + \sigma)y_j, \tag{3.3}$$

where the force of infection is given by $\lambda_i = \beta y_i$, and β is the transmission rate. In addition, the model assumes that individuals leave the population at a constant rate μ due to natural death (i.e., $\frac{1}{\mu}$ is the average life-expectancy of the host), and recover from the infectious class at the rate of σ (where $\frac{1}{\sigma}$ is the average period or séjour of infectiousness).

Furthermore, the model assumes that immunity to the infecting strain is obtained immediately after infection and is lifelong; in addition, multiple concurrent infections with *allelomorphically* different strains are possible. However, individuals who have never been infected with strain type j or any strain that shares alleles with j, have no protection against the strain. Those that have previously been exposed to strains sharing alleles with strains j excluding j itself, have probability $(1 - \gamma)$ of becoming infected due to cross-immunity, where $\gamma \in [0, 1]$. In other words, strains that are *allelomorphically* different do not affect each others' transmissibility. However, infection with *allelomorphically* similar strains provides some cross-protection. Here, γ denotes the strength of cross-immunity and it is assumed that sharing one allele has the same protective outcome as sharing more than one. All the epidemiological parameters are considered to be the same for all strains.

Here we have used a 2-indexing notation. In the 2-locus system, we use two indices i, j to indicate each locus, where $i = 1, 2,\ j = 1, 2$. Then we can rewrite $\sum_{j\sim k} \lambda_j$ as $\sum_l \lambda_{lj} + \sum_k \lambda_{ik} - \lambda_{ij}$ where i is one locus and j the other locus. Using the same

overlapping compartmentalization discussed earlier, Eqs. (3.1)-(3.3) become:

$$\frac{dz_{ij}(t)}{dt} = \beta(1-z_{ij})y_{ij} - \mu z_{ij} \tag{3.4}$$

$$\frac{dw_{ij}(t)}{dt} = \beta\left(\sum_{l} y_{lj} + \sum_{k} y_{ik} - y_{ij}\right)(1-w_{ij}) - \mu w_{ij} \tag{3.5}$$

$$\frac{dy_{ij}(t)}{dt} = \beta\left[(1-w_{ij}) + (1-\gamma)(w_{ij}-z_{ij})\right]y_{ij} - (\mu+\sigma)y_{ij}, \tag{3.6}$$

where z_{ij} is the proportion of the population that has been exposed to pathogenic type ij and/or is now either infected or recovered (immune to strain j), w_{ij} denotes the proportion of the individuals who have been exposed and are immune to any antigenic type sharing alleles with strain ij, and y_{ij} represents the proportion of the population infectious with strain ij.

3.3 Positivity and Invariant Feasible Region

The model (3.4)-(3.6) describes population dynamics, and therefore it is important to show the feasible region for which the model is epidemiologically relevant. We need to show that the solutions to Eqs. (3.4)-(3.6) with non-negative initial conditions will remain non-negative for all time $t \in [0,\infty)$, i.e., $1 \geq z_{ij}(t) \geq 0$, $1 \geq w_{ij}(t) \geq 0$, and $1 \geq y_{ij}(t) \geq 0$ for all strains ij. First, we provide the positivity condition for y_{ij} and then argue that w_{ij} and z_{ij} are both positive. This can be done in the standard way. We prove the positivity of y_{ij} by contradiction. Let $y_{ij}(0) \geq 0$ be the initial condition of $y_{ij}(t)$ and $T_0 := \inf\{t \in [0,\infty) \mid y_{ij}(t) = 0\}$ be the first time when $y_{ij}(T_0) = 0$. Noting that $y_{ij}(T_0) = 0$ implies $y_{ij}(t) \geq 0$ for all $t \in [0, T_0]$, we let

$$\psi := \min_{0\leq t\leq T_0}\{\beta\left[(1-w_{ij}) + (1-\gamma)(w_{ij}-z_{ij})\right] - (\sigma+\mu)\}.$$

Then, for $t \in [0, T_0]$, the expression for y_{ij} becomes:

$$\frac{dy_{ij}}{dt} = \beta\left[(1-w_{ij}) + (1-\gamma)(w_{ij}-z_{ij})\right]y_{ij} - (\sigma+\mu)y_{ij} \geq \psi y_{ij}.$$

Therefore, $y_{ij}(T_0) \geq y_{ij}(0)e^{\psi T_0}$. Clearly, if $y_{ij}(0) = 0$, then $y_{ij}(t) = 0$ for all $t \geq 0$. However, if $y_{ij}(0) > 0$, then $y_{ij}(T_0) \geq y_{ij}(0)e^{\psi T_0} > 0$, which implies that $y_{ij}(t) > 0$ for all $t \in [0, T_0]$.

For $z_{ij}(t)$ and $w_{ij}(t)$, we let T_1 and T_2 be defined similarly to T_0 and assume that there exist first times T_1 and T_2, respectively, such that $z_{ij}(t)$ and $w_{ij}(t)$ become zero. Let $z_{ij}(0) \geq 0$ and $w_{ij}(0) \geq 0$. This implies that $z_{ij}(t) \geq 0$ and $w_{ij}(t) \geq 0$ for all time

$t \in [0, T_1)$ and $t \in [0, T_2)$, respectively; in addition,

$$\begin{aligned}
\frac{dz_{ij}(T_1)}{dt} &= \beta y_{ij} > 0, \\
\frac{dw_{ij}(T_2)}{dt} &= \beta \left(\sum_l y_{lj} + \sum_k y_{ik} - y_{ij} \right) = \beta \left(\sum_{l \neq i} y_{lj} + \sum_{k \neq j} y_{ik} + y_{ij} \right) > 0,
\end{aligned}$$

because $y_{ij}(t) > 0$ for time $t \geq 0$ and for all ij. These contradict our assumptions, for they imply that $z_{ij}(t)$ and $w_{ij}(t)$ must be negative for t right before T_1 and T_2, which violate the choices of T_1 and T_2. We have therefore shown the positivity of our solutions for all times. To show the solutions of Eqs. (3.4)-(3.6) are bounded, we unfold the overlap of immuno-epidemiological classes (e.g., $w_{ij}(t) = \bigcup_{<i,j>} z_{ij}(t)$ where $< i, j >$ means strains sharing alleles with ij). That is, we extract the SIR-type dynamics in Eqs. (3.4)–(3.6). For convenience, we let $s_{ij}(t) = 1 - w_{ij}(t)$, $v_{ij}(t) = w_{ij}(t) - z_{ij}(t)$ and $z_{ij}(t) = 1 - s_{ij}(t) - v_{ij}(t)$. Using these variables and substituting them in (3.4)–(3.6), we transform the equations into the following system:

$$\frac{dz_{ij}(t)}{dt} = \beta(s_{ij}(t) + v_{ij}(t))y_{ij}(t) - \mu z_{ij}(t), \tag{3.7}$$

$$\frac{ds_{ij}(t)}{dt} = \mu - \beta s_{ij}(t) \left(\sum_l y_{lj}(t) + \sum_k y_{ik}(t) - y_{ij}(t) \right) - \mu s_{ij}(t), \tag{3.8}$$

$$\frac{dv_{ij}(t)}{dt} = \beta s_{ij}(t) \left(\sum_{l \neq i} y_{lj}(t) + \sum_{k \neq j} y_{ik}(t) \right) - \beta v_{ij}(t) y_{ij}(t) - \mu v_{ij}(t), \tag{3.9}$$

where $y_{ij}(t)$ is the same as in Eq. (3.6) with $1 - w_{ij}(t)$ and $w_{ij}(t) - z_{ij}(t)$ replaced by $s_{ij}(t)$ and $v_{ij}(t)$, respectively. It suffices to show the positivity of Eqs (3.7)-(3.9) in order to determine the boundedness of Eqs. (3.4)-(3.6). From Eq. (3.8), it follows that

$$\begin{aligned}
s_{ij}(t) &= s_{ij}(0) \exp\left(-\beta \int_0^t \Xi^{ij}_{<l,k>}(\tau) d\tau - \mu t \right) + \exp\left(-\beta \int_0^t \Xi^{ij}_{<l,k>}(\tau) d\tau - \mu t \right) \\
&\quad \times \int_0^t \mu \exp\left(\beta \int_0^\tau \Xi^{ij}_{<l,k>}(\nu) d\nu - \mu \tau \right) d\tau \geq 0,
\end{aligned}$$

where $\Xi^{ij}_{<l,k>}(\nu) := \sum_l y_{lj}(\nu) + \sum_k y_{ik}(\nu) - y_{ij}(\nu) \geq 0$. This indicates that $s_{ij}(t) \geq 0$ for $t \in [0, \infty)$. Similarly, using the same method, it can be shown that $v_{ij} \geq 0$ and $z_{ij} \geq 0$ for all $t \in [0, \infty)$. Since $s_{ij} \geq 0$, $v_{ij} \geq 0$, $z_{ij} \geq 0$ and $w_{ij} \geq 0$, it follows, from $w_{ij} = 1 - s_{ij}$ and $z_{ij} = 1 - s_{ij} - v_{ij}$, that $z_{ij} \leq 1$ and $w_{ij} \leq 1$. Furthermore, we observe, from Eqs. (3.4)-(3.6), that $\frac{dz_{ij}}{dt}|_{z_{ij}=w_{ij}=y_{ij}=1} = -\mu < 0$, $\frac{dw_{ij}}{dt}|_{z_{ij}=w_{ij}=y_{ij}=1} = -\mu < 0$, and $\frac{dy_{ij}}{dt}|_{z_{ij}=w_{ij}=y_{ij}=1} = -(\sigma + \mu) < 0$. Clearly, w_{ij}, z_{ij} and y_{ij} are bounded by 1.

We have thus shown that the model described by Eqs. (3.4)-(3.6) is dissipative and all feasible solutions are uniformly bounded and remain positive for all time $t > 0$, for any given positive initial data in a proper subset, say $\Gamma \subset \mathbb{R}_+^{3M}$, where M is the number of strains. Therefore, the bounded region:

$$\Gamma := \{z_{ij}, w_{ij}, y_{ij} \in \mathbb{R}_+^{3M} : 0 \leq z_{ij},\ w_{ij},\ y_{ij} \leq 1;\ i, j = 1, \ldots, m\}$$

is positively invariant. Thus, it is sufficient to consider the flow generated by the system (3.4)-(3.6) in Γ, where the model is both biologically relevant and mathematically well-posed. The above results are summarized in the following theorem:

Theorem 3.3.1. *Suppose the initial data $z_{ij}(0) \geq 0, w_{ij}(0) \geq 0, y_{ij}(0) \geq 0$ and are inside the invariant biologically feasible simplex Γ. Then the solutions $z_{ij}(t), w_{ij}(t), y_{ij}(t)$ of the model are non-negative for all time $t \in \mathbb{R}_+$. In addition, the following hold:*

1. *If $y_{ij}(0) = 0$ for all i, j, then all solutions $y_{ij}(t)$ remain zero,*
2. *If $y_{ij}(0) > 0$ for all i, j, then all solutions remain inside the feasible region Γ.*

Moreover, $\limsup_{t\to\infty} w_{ij} \leq 1$, $\limsup_{t\to\infty} z_{ij} \leq 1$ and $\limsup_{t\to\infty} y_{ij} \leq 1$.

3.4 Steady State and Stability Analysis

In this section, we establish the conditions for both the existence and stability of equilibria in the model given by (3.4)-(3.6).

3.4.1 Disease-Free Equilibrium

We observe that the system of Eqs. (3.4)-(3.6) admits a trivial steady state solution where $y_{ij} = 0$ for all $i, j = 1, 2$ and we call this the disease-free equilibrium, denoted by $\mathcal{E}_0 := \{z_{11}, \ldots, z_{22}; w_{11}, \ldots, w_{22}; y_{11}, \ldots, y_{22}\} = \{0, \ldots, 0; 0, \ldots, 0; 0, \ldots, 0\}$. The linear stability of the disease-free steady state $\mathcal{E}_0$ can easily be investigated and depends on the basic reproductive number $\mathcal{R}_0 = \dfrac{\beta}{\sigma + \mu}$, which can be calculated by the next generation method (see, for example, Diekmann & Heesterbeek, 2000; Castillo-Chavez et al., 2002). Here, $\mathcal{R}_0$ denotes the average number of secondary cases caused by each infected individual at the beginning of infection in the population with no pre-existing immunity, and measures the initial growth rate of the epidemic. The following condition determines the stability of the disease-free equilibrium, $\mathcal{E}_0$.

Lemma 3.4.1. *The disease-free equilibrium $\mathcal{E}_0 := (z_{ij}, w_{ij}, y_{ij})_{i,j=1,2} = (0, 0, 0)$ exists for all $\mathcal{R}_0$. Moreover, it is asymptotically stable if and only if $\mathcal{R}_0 < 1$. It is unstable when $\mathcal{R}_0 > 1$.*

The proof of Theorem 3.4.1 relies on the eigenvalues of the Jacobian matrix generated by Eqs. (3.4)-(3.6) and evaluated at the disease-free equilibrium point $\mathcal{E}_0$. It is trivial and has been ignored herein (see A.1.1). If the disease-free equilibrium point is the only steady state in Γ, then we obtain the following statement.

Theorem 3.4.2. *Suppose $z_{ij}(0) \geq 0$, $w_{ij}(0) \geq 0$ and $y_{ij}(0) \geq 0$ and are inside the biologically invariant feasible simplex Γ. Then there is only one steady state $\mathcal{E}_0$, and is globally asymptotically stable provided $\mathcal{R}_0 < 1$.*

Proof. See Appendix A.1.2 □

Whenever the condition $\mathcal{R}_0 > 1$ holds, $\mathcal{E}_0$ is unstable. From the positivity and boundedness of Γ, and dissipativity, if there are no equilibrium points or attractors in Γ, then all the trajectories inside $\Gamma \backslash \mathcal{E}_0$ leave the feasible region which violates the boundedness of the solutions. Therefore, there must be at least one non-trivial equilibrium or an attractor in Γ.

3.4.2 Endemic Equilibria

To determine the existence of non-trivial equilibrium points or endemic steady states, we use Fig. 3.1 to help us simplify the search for endemic equilibria and stability analysis. Suppose we have a 2-strain system, then there are three possible endemic configurations when the reproductive number $\mathcal{R}_0 > 1$: either strain 1 persists while strain 2 goes to extinction, or strain 2 persists while strain 1 goes to extinction, or both strains 1 and 2 coexist. Hence, we denote the persistent sets as follows: $\{1\}$, $\{2\}$ and $\{1, 2\}$. Figure 3.1 illustrates all the possible endemic configurations of different clustering type for a 4-strain system, where black circles ($\bullet$) denote the persistent strains while white circles ($\circ$) represent extinct strains. Using the representation summarized in Fig. 3.1, we propose the following mathematical notion to provide a natural framework for the analysis of the model. Let $\mathcal{P}$ denote the set of persistent strains (i.e., strains that can survive) and be defined as

$$\mathcal{P} := \begin{cases} \{i \in \{1, 2, \ldots, \mathcal{N}\} : \mathcal{R}_0^{(i)} > 1\}, & \mathcal{R}_0 > 1 \\ \varnothing, & \mathcal{R}_0 < 1, \end{cases} \tag{3.10}$$

where $\mathcal{R}_0^{(i)}$ is the strain-specific reproductive number when the epidemiological and demographic parameters are different for each strain; and $\mathcal{R}_0 = \max_{i \in \mathbb{N}}\{\mathcal{R}_0^{(i)}\}$. We note that when $\mathcal{R}_0^{(i)} = \mathcal{R}_0^{(j)}$, $i \neq j$ for all i (which is the case herein), $\mathcal{R}_0^{(i)} = \mathcal{R}_0$. In addition, we observe that the possible endemic configurations outlined previously can be classified into two notions of endemicity whether there is a principle of competitive

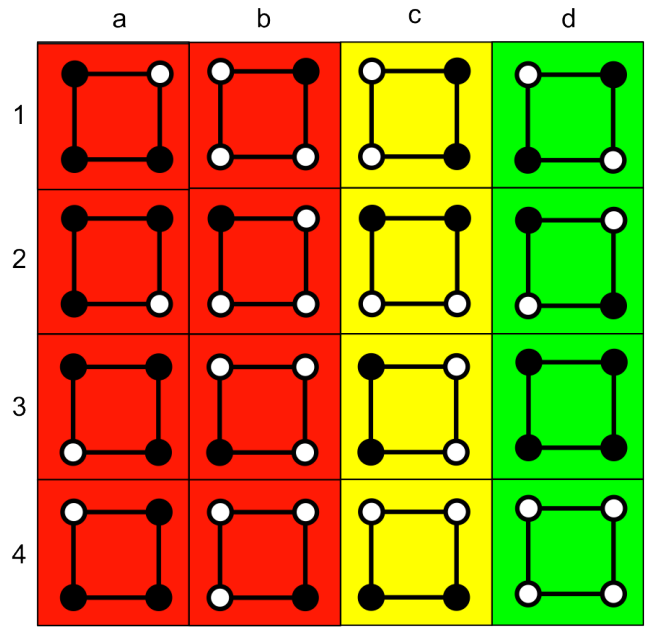


Figure 3.1: **Multi-locus System (Possible Equilibria for Four Strain System on a Square Lattice with $\mathbb{D}_4$-Symmetry).** The sketch shows different configurations of the equilibrium clusters, where black circles (•) denote the persistent strains while white circles (○) represent extinct strains. For labelling purposes, the columns are denoted *a*-*d*, and the rows 1-4.

exclusion or full coexistence. We define the following notion of non-negative solutions (Thieme, 2003):

Definition 3.4.3. *1. A non-trivial and non-negative equilibrium point $(z^*_{ij}, w^*_{ij}, y^*_{ij})$ is called weakly endemic if $y_{ij} = y^*_{ij} \neq 0$ for some pair ij.*

2. A non-trivial solution is called strongly endemic if $y_{ij} = y^*_{ij} \neq 0$ for all ij.

In the definition above, a weakly endemic equilibrium is equivalent to having a principle of competitive exclusion, and a strongly endemic equilibrium represents full coexistence. Throughout the proceeding sections, we will use the multi-locus-allele notation for ease of illustration and will focus on the 2-locus-2-allele system.

3.4.2.1 Weak Endemicity: Principle of Competitive Exclusion and Discrete Strain Structure

In principle, one could calculate all the endemic equilibria. However, this could be very difficult as the dimensionality of the system is increased. Instead, we simplify our study and break up the steady state analysis according to equilibrium clusters (see Fig. 3.1). From the numerical results of Gupta et al. (1998), we know that strains have the tendency to self-organize into discordant sets containing non-overlapping alleles under

high antigenic interference and strong immune response (i.e., for high cross-immunity, γ). Suppose we are interested in studying one of the equilibrium cluster types shown in Fig. 3.1 (say, any of the structures in Fig. 3.1 except d-3 and d-4), then using the definition of weak endemicity, we can place strains into two disjoint sets based on whether they persist or not. We denote the two disjoint sets of equilibrium clusters as $\mathcal{D}_p$ and $\mathcal{D}_\emptyset$, where $\mathcal{D}_p$ contains all the strains that are persistent while $\mathcal{D}_\emptyset$ contains all the strains that go extinct. Observe that the set $\mathcal{D}_p \cup \mathcal{D}_\emptyset$ contains all the strains (i.e., $\mathcal{D}_p \cup \mathcal{D}_\emptyset = \{\mathbf{ax},\ \mathbf{ay},\ \mathbf{bx},\ \mathbf{by}\}$) and $\mathcal{D}_p \cap \mathcal{D}_\emptyset = \emptyset$. For example, if we are interested in deriving the necessary conditions and stability of the steady state in the forms of d-1 and d-2 in Fig. 3.1 (i.e., discordant set), we set $\mathcal{D}_p$ to contain **ax** and **by** and $\mathcal{D}_\emptyset$ to contain **ay** and **bx**. Then $y_{ax} = y^*_{ax} > 0$ and $y_{by} = y^*_{by} > 0$ where $ij \in \mathcal{D}_p := \{\mathbf{ax}, \mathbf{by}\}$; and $y_{ay} = y^*_{ay} = 0$ and $y_{bx} = y^*_{bx} = 0$ for $kl \in \mathcal{D}_\emptyset := \{\mathbf{ay}, \mathbf{bx}\}$. Using this, we separate the coordinates of the steady state subspace for weak endemicity into sub-coordinates that are persistent and those that go to extinction. Let $\rho_p \in \mathbb{Z}_+$ be the number of discordant strains in $\mathcal{D}_p$, and $\rho_\emptyset \in \mathbb{Z}_+$ denote the number of discordant strains in $\mathcal{D}_\emptyset$. We begin by first considering the simple balanced case where $\rho = \rho_p = \rho_\emptyset$. However, it should be noted that it is possible to have $\rho_p \neq \rho_\emptyset$, say for example, when $\mathcal{D}_p = \{\mathbf{ax}\}$ and $\mathcal{D}_\emptyset := \{\mathbf{ay}, \mathbf{bx}, \mathbf{by}\}$ (see of one the cases of b-1, b-2, b-3 or b-4 in Fig. 3.1) or when $\mathcal{D}_\emptyset = \{\mathbf{ax}\}$ and $\mathcal{D}_p = \{\mathbf{ay}, \mathbf{bx}, \mathbf{by}\}$ (see one of the case of a-1, a-2, a-3 or a-4 in Fig. 3.1). However, these cases are unbalanced. We rewrite Eqs. (3.4)-(3.6) according to whether strains belong to $\mathcal{D}_p$ or not. For strains belonging to $\mathcal{D}_p$, $\sum_l y_{lj} + \sum_k y_{ik} - y_{ij}$ reduces to y_{ij}. For those strains which do not belong to $\mathcal{D}_p$ but are members of $\mathcal{D}_\emptyset$, $\sum_i y_{ik} + \sum_m y_{km} + y_{kl}$ reduces to ρy_{ij}. Solving for the weak endemic steady state, we find

$$\begin{aligned} z^*_{ij} &= 1 - \frac{1}{\mathcal{R}_0},\ \ w^*_{ij} = 1 - \frac{1}{\mathcal{R}_0},\ \ y^*_{ij} = \frac{\mu}{\beta}\left(\mathcal{R}_0 - 1\right),\ \text{for } ij \in \mathcal{D}_p, \\ z^*_{kl} &= 0,\ \ w^*_{kl} = \frac{\rho\left(\mathcal{R}_0 - 1\right)}{\rho\left(\mathcal{R}_0 - 1\right) + 1},\ \ y^*_{kl} = 0,\ \text{for } kl \in \mathcal{D}_\emptyset, \end{aligned}$$

where we recall $\mathcal{R}_0 = \dfrac{\beta}{\mu + \sigma}$. Note that these points only exist in biologically relevant regimes if $\mathcal{R}_0 > 1$; when $\mathcal{R}_0 = 1$, we recover the disease-free equilibrium $\mathcal{E}_0$. In addition, by symmetry, interchanging the role of ij and kl, we can obtain the other weakly endemic equilibrium. Whenever the equilibrium clustering type in d-1 exists in Fig. 3.1, equilibrium clustering type d-2 also exists. Similarly, if the equilibrium clustering type in b-1 exists in Fig. 3.1, then b-2, b-3 and b-4 also exist.

Evaluating the Jacobian matrix generated by the model at the above equilibrium point, we obtain the following theorem which establishes the stability of the discrete strain structure endemic steady state:

Theorem 3.4.4. *Suppose $\mathcal{R}_0 > 1$ and $\rho \neq 0$. Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be defined as:*

$$\mathcal{E}_1 = \left(1 - \frac{1}{\mathcal{R}_0}, 1 - \frac{1}{\mathcal{R}_0}, \frac{\mu}{\beta}(\mathcal{R}_0 - 1), 0, \frac{\rho(\mathcal{R}_0 - 1)}{\rho(\mathcal{R}_0 - 1) + 1}, 0\right), \tag{3.11}$$

$$\mathcal{E}_2 = \left(0, \frac{\rho(\mathcal{R}_0 - 1)}{\rho(\mathcal{R}_0 - 1) + 1}, 0, 1 - \frac{1}{\mathcal{R}_0}, 1 - \frac{1}{\mathcal{R}_0}, \frac{\mu}{\beta}(\mathcal{R}_0 - 1)\right), \tag{3.12}$$

be the steady states corresponding to weakly endemic equilibria of discordant type, where ρ is the number of discordant sets (non-overlapping set or minimally overlapping set), and $\mathcal{E}_i$, $i = 1, 2$ condenses the steady state expression according to clusters. The weak endemic equilibria exist provided that $\mathcal{R}_0 > 1$. In addition, the following hold:

1. *The weakly endemic equilibria $\mathcal{E}_1$ and $\mathcal{E}_2$ are locally asymptotically stable provided*

$$\gamma > \overline{\gamma}(\mathcal{R}_0) =: \frac{1 + \rho(\mathcal{R}_0 - 1)}{\rho\mathcal{R}_0}. \tag{3.13}$$

2. *If $\gamma < \overline{\gamma}(\mathcal{R}_0)$, then the weakly endemic equilibria $\mathcal{E}_1$ and $\mathcal{E}_2$ are unstable.*

Proof. See Appendix A.1.3.. □

Corollary 3.4.5. *Suppose $\mathcal{R}_0 > 1$, $\rho_p = \rho_\emptyset$ and let $\rho_p \in \mathbb{Z}_+$ and $\rho_\emptyset \in \mathbb{Z}_+$ be the number of discordant strains in $\mathcal{D}_p$ and in $\mathcal{D}_\emptyset$, respectively, where $\mathcal{D}_p$ is the set containing persistent strains while $\mathcal{D}_\emptyset$ is the set containing non-persistent (extinct) strains. If $\rho = \min\{\rho_p, \rho_\emptyset\} \neq 0$, then all weakly endemic equilibria with the steady states having such discrete strain structures are locally asymptotically stable provided that $\gamma > \overline{\gamma}(\mathcal{R}_0)$.*

Conjecture 3.4.6. *Suppose $\mathcal{R}_0 > 1$ and let $\rho_p \in \mathbb{Z}_+$ and $\rho_\emptyset \in \mathbb{Z}_+$ be the number of discordant strains in $\mathcal{D}_p$ and in $\mathcal{D}_\emptyset$, respectively, where $\mathcal{D}_p$ is the set containing persistent strains while $\mathcal{D}_\emptyset$ is the set containing non-persistent (extinct) strains. Then the following holds:*

1. *Suppose $\rho_p = \rho_\emptyset$. If $\rho = \min\{\rho_p, \rho_\emptyset\} = 0$, then all weakly endemic equilibria with the steady states having such discrete strain structures are unstable.*

2. *Suppose $\rho_p \neq \rho_\emptyset$. If $\rho = \min\{\rho_p, \rho_\emptyset\} = 0$, then all weakly endemic equilibria with the steady states having such discrete strain structures are always unstable.*

It should also be noted that, in Theorem 3.4.4 $\overline{\gamma}(\mathcal{R}_0) \to 1$ as $\mathcal{R}_0 \to +\infty$. The proof of Corollary 3.4.5 follows directly from Theorem 3.4.4. For Conjecture 3.4.6, the symmetry structure of the equilibrium clusters can be used to show that the

results hold (see Appendix A.3 for illustrative examples for the 4-strain system). In particular, we observe that equilibrium cluster types of the form shown in Fig. 3.1 associated with panels ranging a-1 to a-4 and b-1 to b-4, where $\rho_p \neq \rho_\emptyset$ (i.e., $\rho_p = 2$ and $\rho_\emptyset = 0$, and $\rho = \rho_p = 0$ and $\rho = \rho_\emptyset = 2$, respectively), cannot be stable. That is, when $\rho = \rho_p = 2$, the condition in Theorem 3.4.4 can be satisfied but when $\rho = \rho_\emptyset = 0$ the condition in Theorem 3.4.4 is violated. However, from the symmetry of the antigenic space, these two discrete strain structures are conjugate structures where the structures are related by either $\mathbb{Z}_4$-rotations or reflections along the horizontal/vertical axes. Hence, this gives rise to a contradiction. Moreover, if we let $\rho = \min\{\rho_p, \rho_\emptyset\}$, then we can easily see that the stability condition in Theorem 3.4.4 is automatically violated when $\rho = \min\{\rho_p, \rho_\emptyset\} = 0$, for $\overline{\gamma} \to \infty$ as $\rho \to 0$. But $\gamma \in [0, 1]$ and $\gamma > \overline{\gamma} = \infty$ implies instability. Similarly, equilibrium cluster types in Fig. 3.1 in the panels ranging c-1 to c-4 cannot exist because $\rho = 0$, i.e. $\rho_p = 0$ and $\rho_\emptyset = 0$ because they violate the necessary stability condition in Theorem 3.4.4. We note that weakly endemic equilibria consist of strain clusters that are structured discretely (i.e., not all strains are present due to the principle of competitive exclusion), hence they exhibit discrete strain structures (DSS). Figure 3.2 illustrates the temporal evolution of the prevalence z_i where the pathogen population is structured into discrete discordant sets where one discordant pair of strains, strain **ay** and strain **bx**, dominates in equal abundance into perpetuity while the other discordant pair of strains, strain **ax** and **by**, is suppressed.

3.4.2.2 Strong Endemicity: No Strain Structure (NSS)

In order to find the unique strongly endemic equilibrium, we use the second part of Def. 3.4.3(2), that is, $y_{ij} = y^*_{ij} > 0$ for all $i, j = 1, 2$. Since we have assumed that all strains have the same transmissibility success and $\mathcal{R}_0$ is the same for all strains, the abundance of y^*_{ij} becomes $y^*_{ij} = y^*$, $z^*_{ij} = z^*$ and $w^*_{ij} = w^*$.Using this, we note that $y_{ij} = y_{kl}$ for $ij \neq kl$, and $\sum_l y_{lj} + \sum_k y_{ik} - y_{ij}$ in Eqs. (3.4)-(3.6) reduces to αy_{ij} where $\alpha = m^n - (m-1)^n$, $\alpha \in \mathbb{Z}_+$ is the number of strains sharing alleles with strain ij (including ij itself). Then at steady state, the strong endemic equilibrium point $\mathcal{E}_*$ is given by:

$$z^* = \frac{\beta y^*}{\mu + \beta y^*} \tag{3.14}$$

$$w^* = \frac{\alpha\beta y^*}{\mu + \alpha\beta y^*} \tag{3.15}$$

$$y^* = \frac{\mu}{2\alpha\beta}\left(-\Xi + \sqrt{4\alpha\left(\mathcal{R}_0 - 1\right) + \Xi^2}\right) \tag{3.16}$$

$$\Xi = 1 - \gamma\mathcal{R}_0 + \alpha\left(1 - \mathcal{R}_0\left(1 - \gamma\right)\right). \tag{3.17}$$

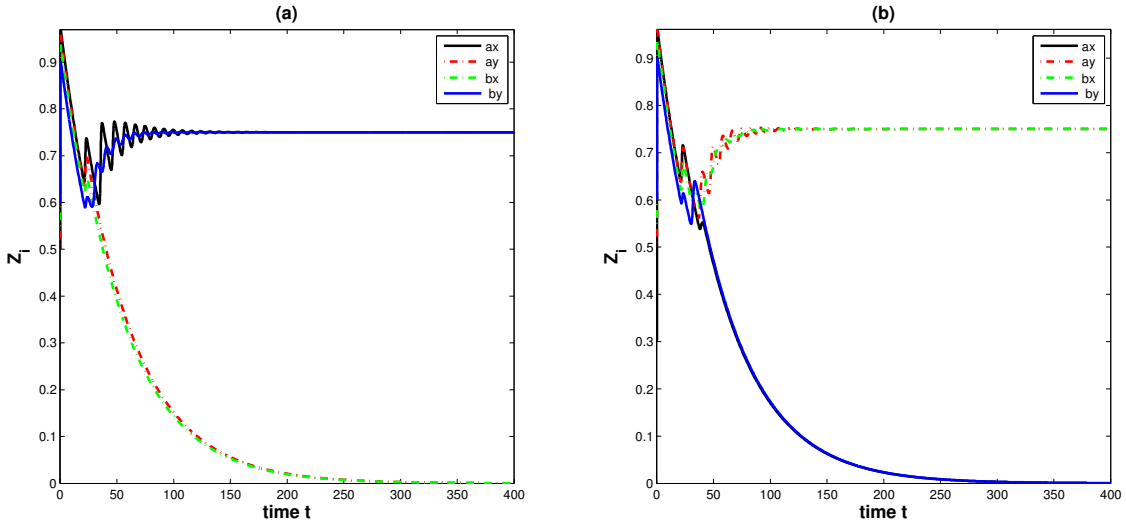


Figure 3.2: **DSS Time Evolution (Eqs.** (3.4)-(3.6)**).** Dynamics of a stable DSS for a 2-locus-2-allele system where two strains coexist in equal abundance, and the other two strains are suppressed and go to extinction. Here z_i denotes the prevalence of strain i, where $i = \{\mathbf{ax}, \mathbf{ay}, \mathbf{bx}, \mathbf{by}\}$ and $\mathcal{R}_0 = 4$, $\sigma = 9.98$, $\mu = 0.02$, $\rho = 2$, and $\gamma = 0.95$ are used to reproduce the dynamics in Gupta et al. (1998). The dynamics in both panels have the same parameters but with different initial conditions. Equations (3.4)-(3.6) were solved using the Runge-Kutta fourth- and fifth-order method solver ODE45 in MATLAB (The Mathworks Inc., Natwick, MA, USA) with the following initial conditions: (**a**) z(0) = [0.5 0.51 0.55 0.6], w(0) = [0.4 0.3 0.3 0.30] and y(0) = [0.5 0.4 0.3 0.2] and (**b**) z(0) = [0.3 0.51 0.55 0.6], w(0) = [0.4 0.3 0.3 0.30] and y(0) = [0.5 0.4 0.3 0.2]; and relative and absolute error tolerance of 10^{-5} and 10^{-6}, respectively.

It should be noted that y^* is a function of both γ and $\mathcal{R}_0$, for $\alpha \geq 1$. When $\alpha = 1$, the steady state for system (3.4)-(3.6) reduces to a single strain system. Henceforth, we assume $\alpha > 1$. In addition, for any given value of $\mathcal{R}_0$, the proportion of infected individuals y^* decreases with respect to the levels of cross-reactivity γ (see Fig. 3.3 and Appendix A.2.1).

Having calculated the strong endemic equilibrium point or NSS endemic steady state, we now establish its linear stability. To do this, we calculate the Jacobian matrix evaluated at the NSS endemic equilibrium point. The Jacobian matrix $\mathcal{J}$ is:

$$\mathcal{J} := \begin{bmatrix} -(\beta y^* + \mu)1_{m^2\times m^2} & 0_{m^2\times m^2} & \beta(1-z^*)1_{m^2\times m^2} \\ 0_{m^2\times m^2} & -(\beta\alpha y^* + \mu)1_{m^2\times m^2} & \beta(1-w^*)\Omega_{m^2\times m^2} \\ -\beta(1-\gamma)y^*1_{m^2\times m^2} & -\gamma\beta y 1_{m^2\times m^2} & 0_{m^2\times m^2} \end{bmatrix}, \quad (3.18)$$

where $\Omega_{m^2\times m^2} = \mathbb{I}_{m^2\times m^2} + \mathbb{A}_{m^2\times m^2}$, $\mathbb{I}_{m^2\times m^2}$ is the identity matrix and $\mathbb{A}_{m^2\times m^2}$ is the adjacency (strain similarity) matrix for strains sharing alleles in common. Note that the characteristic equation of this Jacobian is an 12^{th} order polynomial and its characterization is difficult. However, because of the sparseness of the Jacobian matrix, we were able to obtain the characteristic polynomial function of the

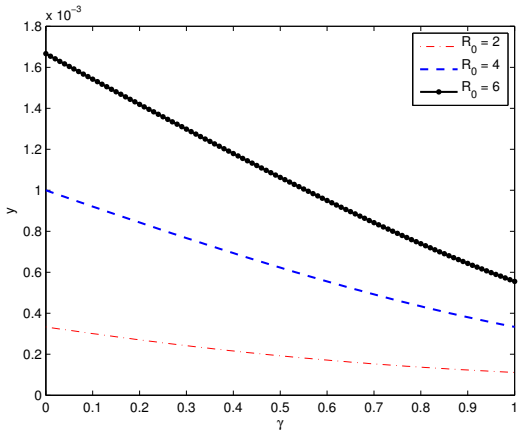


Figure 3.3: **Functional Properties of $y^*(\gamma, \mathcal{R}_0)$ (Eqs. (3.16)-(3.17)).** The plot shows that the proportion of individuals y who are infectious is a decreasing function of cross-immunity level γ, for $\mathcal{R}_0 > 1$ and $\alpha = 3$. Here $\sigma = 9.98$ and $\mu = 0.02$ and $\alpha = 3$ are used.

form of $P(\zeta) = \prod_{k=1}^{4} P_k(\zeta)$, where $P_k(\zeta)$ is a cubic polynomial function of the form $P_k(\zeta) = \sum_{l=0}^{3} \mathcal{A}_{k,3-l}\zeta^l$ and where $\mathcal{A}_{k,0} = 1$ for all $k = 1, 2, 3, 4$; and the coefficients are given in Appendix A.2.2. It should be noted that, using the symmetry property of the 2-locus-2-allele system and isotopic decomposition of the phase space to diagonalize the Jacobian, Blyuss (2013) obtained similar expressions for the fully symmetric steady state. We observe that the characteristic polynomials $P_k(\zeta)$ have positive coefficient $\mathcal{A}_{k,l},\ k = 1, 2, 3, l = 0, 1, \ldots 3$ and their Hurwitz determinants are all positive. That is, $\mathcal{A}_k > 0$ for $k = 1, 2, 3$. Hence, from the Routh-Hurwitz stability criterion, the characteristic polynomials $P_k(\zeta),\ k = 1, 2, 3$ have all their eigenvalues with negative real parts. For the characteristic polynomial $P_4(\zeta)$, we observe that $\mathcal{A}_{4,0} > 0$ and $\mathcal{A}_{4,1} > 0$. It suffices to show that the coefficients $\mathcal{A}_{4,2}$ and $\mathcal{A}_{4,3}$, and Hurwitz determinant $\mathcal{A}_4$ are positive. In order to determine the condition for the stability of the strongly endemic or co-existence equilibrium, we must solve a fixed problem $\gamma^* = \overline{\gamma}_r(\gamma^*, \mathcal{R}_0)$ for any fixed $\mathcal{R}_0 > 1$, where $\overline{\gamma}_r(\gamma, \mathcal{R}_0)\ r = 1, 2, 3$ are given below (see Appendix A.2.2). For the strongly endemic equilibrium to be locally asymptotically stable, we state the following results:

Theorem 3.4.7. *Suppose $\mathcal{R}_0 > 1$ and $\gamma \in [0, 1]$. Let $\overline{\gamma}_r(\gamma, \mathcal{R}_0),\ r = 1, 2, 3$ be defined as follows:*

$$\overline{\gamma}_1(\gamma, \mathcal{R}_0) = \frac{(\beta y^* + \mu)(\beta\alpha y^* + \mu) + \beta^2 y^*(1 - z^*)}{\beta^2 y^*(1 - w^*) + \beta^2 y^*(1 - z^*)},$$

$$\overline{\gamma}_2(\gamma, \mathcal{R}_0) = \frac{(1 - z^*)(\beta\alpha y^* + \mu)}{(1 - w^*)(\beta y^* + \mu) + (1 - z^*)(\beta\alpha y^* + \mu)},$$

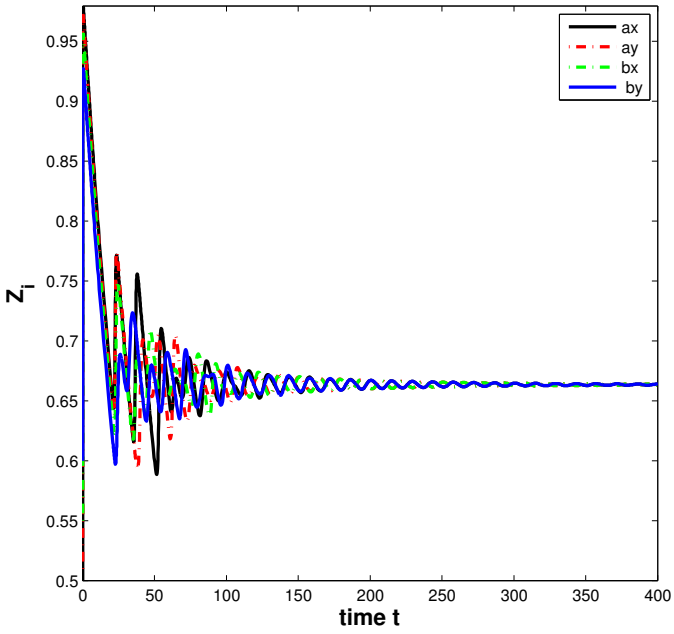


Figure 3.4: **Time Evolution of No Strain Structure (NSS) (Eqs.** (3.4)-(3.6)**).** Dynamics of a stable NSS for the 2-locus-2-allele system where all strains coexist in equal abundance. Here z_i denotes the prevalence of strain i and $\mathcal{R}_0 = 4$, $\sigma = 9.98$, $\mu = 0.02$ and $\gamma = 0.95$ are used to reproduce the dynamics in Gupta et al. (1998). Equations (3.4)-(3.6) were solved using the Runge-Kutta fourth- and fifth-order method solver ODE45 in MATLAB (The Mathworks Inc., Natwick, MA, USA) with the following initial conditions $z(0) = [0.50\ 0.51\ 0.55\ 0.6]$, $w(0) = [0.40\ 0.30\ 0.3\ 0.30]$ and $y(0) = [0.50\ 0.40\ 0.30\ 0.20]$, and relative and absolute error tolerance of 10^{-5} and 10^{-6}, respectively.

$$\overline{\gamma}_3(\gamma, \mathcal{R}_0) = \frac{[\beta(1+\alpha)y^* + 2\mu]\left[(\beta y^* + \mu)(\beta\alpha y^* + \mu)\right]}{\beta^2 y^*(1-w^*)\left[\beta(\alpha+2)y^* + 3\mu\right] + (\beta^2 y^*(1-z^*))\left[\beta(2\alpha+1)y^* + 3\mu\right]} + \frac{\beta^2 y^*(1-z^*)\left[\beta(2\alpha+1)y^* + 3\mu\right]}{\beta^2 y^*(1-w^*)\left[\beta(\alpha+2)y^* + 3\mu\right] + (\beta^2 y^*(1-z^*))\left[\beta(2\alpha+1)y^* + 3\mu\right]}.$$

Then the strongly endemic (no strain structure (NSS)) equilibrium is locally asymptotically stable provided that $\gamma < \overline{\gamma}_1(\gamma, \mathcal{R}_0)$, $\gamma > \overline{\gamma}_2(\gamma, \mathcal{R}_0)$ *and* $\gamma < \overline{\gamma}_3(\gamma, \mathcal{R}_0)$. *The strongly endemic equilibrium is unstable if any one of the above conditions is false. That is, if* $\gamma \geq \overline{\gamma}_1(\gamma, \mathcal{R}_0)$, $\gamma \leq \overline{\gamma}_2(\gamma, \mathcal{R}_0)$, *and/or* $\gamma \geq \overline{\gamma}_3(\gamma, \mathcal{R}_0)$, *then the strongly endemic steady state is unstable.*

Proof. See Appendix A.1.4 □

Whenever the conditions in Theorem 3.4.7 hold, the steady state is fully symmetric and all strains coexist in equal abundance (see Fig. 3.4). It exhibits *no strain structure* (NSS) because there is no distinction between the strains in the pathogen population structure.

3.4.2.3 Strong Endemicity: Cyclical Strain Structures and Existence of Oscillatory Behaviours

To consider the existence of cyclical behaviours, we look for the necessary condition for a Hopf bifurcation in the system. Because of the forms of $P(\zeta)$ and $P_k(\zeta),\ k = 1, 2, 3$, we note that they always have eigenvalues with negative real part. As a result, we focus on $P_4(\zeta)$. If $\mathcal{A}_{4,0} > 0$, $\mathcal{A}_{4,1} > 0$, $\mathcal{A}_{4,2} > 0$ and $\mathcal{A}_{4,3} > 0$, then a Hopf bifurcation can take place at the transit through the surface defined by the Hurwitz determinant $\mathcal{A}_4 = 0$. The strongly endemic equilibrium is unstable if there is pair of complex conjugate roots for $P_4(\zeta)$. Let $\zeta_{1,2} = \pm i\omega$, then $\zeta_1 + \zeta_2 + \zeta_3 = -\mathcal{A}_{4,1}$. Hence $\zeta_3 = -\mathcal{A}_{4,1}$ and $\omega^2 = \mathcal{A}_{4,2}$. Then $P_4(\zeta_3) = -\mathcal{A}_{4,1}\mathcal{A}_{4,2} + \mathcal{A}_{4,3}\mathcal{A}_{4,0} = 0$ and $\gamma = \gamma_3^*$, where $\gamma_3^* \in [0, 1]$ is the unique solution for the fixed point problem $\gamma = \overline{\gamma}_3(\gamma, \mathcal{R}_0)$. Thus, a Hopf bifurcation may occur at the strongly endemic equilibrium. We state the following results:

Lemma 3.4.8. *Suppose $\mathcal{R}_0 > 1$, and let $\mathcal{H}_0$ or $\mathcal{H}_1$ be defined as:*

$$\mathcal{H}_0 = \{(\beta, \mu, \sigma, \gamma) \in \mathbb{R}_+^4 : \mathcal{A}_4 = 0\},\ \textit{or equivalently} \tag{3.19}$$

$$\mathcal{H}_1 = \{(\mathcal{R}_0, \gamma) \in \mathbb{R}_+^2 : \mathcal{A}_4 = 0\}. \tag{3.20}$$

Then the characteristic polynomial $P_4(\zeta)$ has one real negative and two purely imaginary solutions if and only if $(\beta, \mu, \sigma, \gamma) \in \mathcal{H}_0$ (or $(\mathcal{R}_0, \gamma) \in \mathcal{H}_1$). In addition, the roots are given by $\zeta_1 = -\mathcal{A}_{4,1}$ and $\zeta_{2,3} = \pm i\omega$, where $\omega^2 = \mathcal{A}_{4,2}$.

Lemma 3.4.9. *Suppose $\mathcal{R}_0 > 1$. Let $\overline{\gamma}_r(\gamma, \mathcal{R}_0),\ r = 1, 2, 3$ be defined as above for $\gamma \in [0, 1]$ and suppose that they are differentiable and monotone on $\gamma \in (0, 1)$. Let*

1. *$1 > \overline{\gamma}_r(0, \mathcal{R}_0) > \overline{\gamma}_r(1, \mathcal{R}_0) > 0$, or*
2. *$0 < \overline{\gamma}_r(0, \mathcal{R}_0) < \overline{\gamma}_r(1, \mathcal{R}_0) < 1$.*

Then there exist $\gamma_r^*,\ r = 1, 2, 3$ such that $\gamma_r^* = \overline{\gamma}_r(\gamma_r^*, \mathcal{R}_0)$ and $\gamma_r^* \in (\overline{\gamma}_r(0, \mathcal{R}_0), \overline{\gamma}_r(1, \mathcal{R}_0))$. In addition, the following hold:

1. *The strongly endemic equilibrium undergoes a steady state bifurcation when $\gamma = \gamma_1^*$ and/or $\gamma = \gamma_2^*$.*
2. *The strongly endemic equilibrium undergoes a Hopf bifurcation when $\gamma = \gamma_3^*$, assuming the transversality condition holds.*

Theorem 3.4.10. *Assume $\mathcal{R}_0 > 1$ and $\mathcal{A}_{4,2} > 0$ (or equivalently $\gamma < \gamma_1(\gamma, \mathcal{R}_0)$) and $\mathcal{A}_{4,3} > 0$ (or equivalently $\gamma > \gamma_2(\gamma, \mathcal{R}_0)$). When γ passes through a critical value γ_3^*, where γ_3^* is the solution of $\gamma = \overline{\gamma}_3(\gamma, \mathcal{R}_0)$, the system undergoes a Hopf bifurcation, provided the transversality condition holds.*

Due to the complicated form of $\gamma = \overline{\gamma}_3(\gamma, \mathcal{R}_0)$, we will not address the properties of the Hopf bifurcation (i.e., direction of bifurcation, orbital stability of periodic solutions). In particular, we will not discuss the direction, stability of the periodic orbits, uniqueness of the orbit and period of solutions. Instead, we use numerical simulations to show that the system indeed undergoes a Hopf bifurcation at some critical value of γ (see Fig. 3.5). Figure 3.5 shows the dynamics of the state variables as the value of cross-immunity, γ, changes. As γ passes through a subcritical threshold defined in Theorem 3.4.10, the full synchronicity of NSS, where all strains are in equal abundance, is broken, and the temporal dynamics of the strain population denoted by the strongly endemic equilibrium becomes oscillatory via a Hopf bifurcation. The strains exhibit cyclical strain structure.

3.5 Discussion

Using the multi-locus-allele framework, we have shown the exact conditions under which the linear (local) stability of disease-free and endemic equilibria holds. We established that the disease-free equilibrium is globally asymptotically stable whenever the basic reproductive number is less than unity (e.g., $\mathcal{R}_0 < 1$), and is unstable when this threshold condition is violated. When $\mathcal{R}_0 > 1$, at least one endemic equilibrium exists. However, this condition alone is not sufficient for the stability of the endemic equilibria. We found additional conditions that depend on the value of cross-immunity level in the population. We showed that the existence of discrete antigenic forms of pathogens can exhibit three distinct dynamic features, where antigenic variants (i) fully self-organize and co-exist with NSS for low γ, (ii) sort themselves into DSS with non-overlapping or minimally overlapping clusters under the principle of competitive exclusion for high γ, or (iii) exhibit CSS where dominant antigenic types exist cyclically for intermediary level of γ. These observations are consistent with numerical simulations presented herein and in Gupta et. al. (1996, 1998) and others (see Andreasen et al., 1997; Ferguson & Andreasen, 2001; Dawes & Gog, 2002; Recker et. al., 2004; Recker & Gupta, 2005; Calvez et al., 2005; Recker et. al., 2007; Chan & Yu, 2013; Blyuss, 2013). Figures 3.6-3.7 summarize these results.

From the results of the analysis presented herein for the basic 2-locus-2-allele system, we can extend our results to more general cases of the multi-locus-allele model, and use the structure (see Fig. 3.8) envisioned by the pathogen population to postulate the nature of the immune responses against the various serotypes. As we increase the number of loci, we observe that the number of cluster types and structures increase and their dynamics become highly involved. However, the general characterization of the behaviours is conserved. That is, the general properties (e.g.,

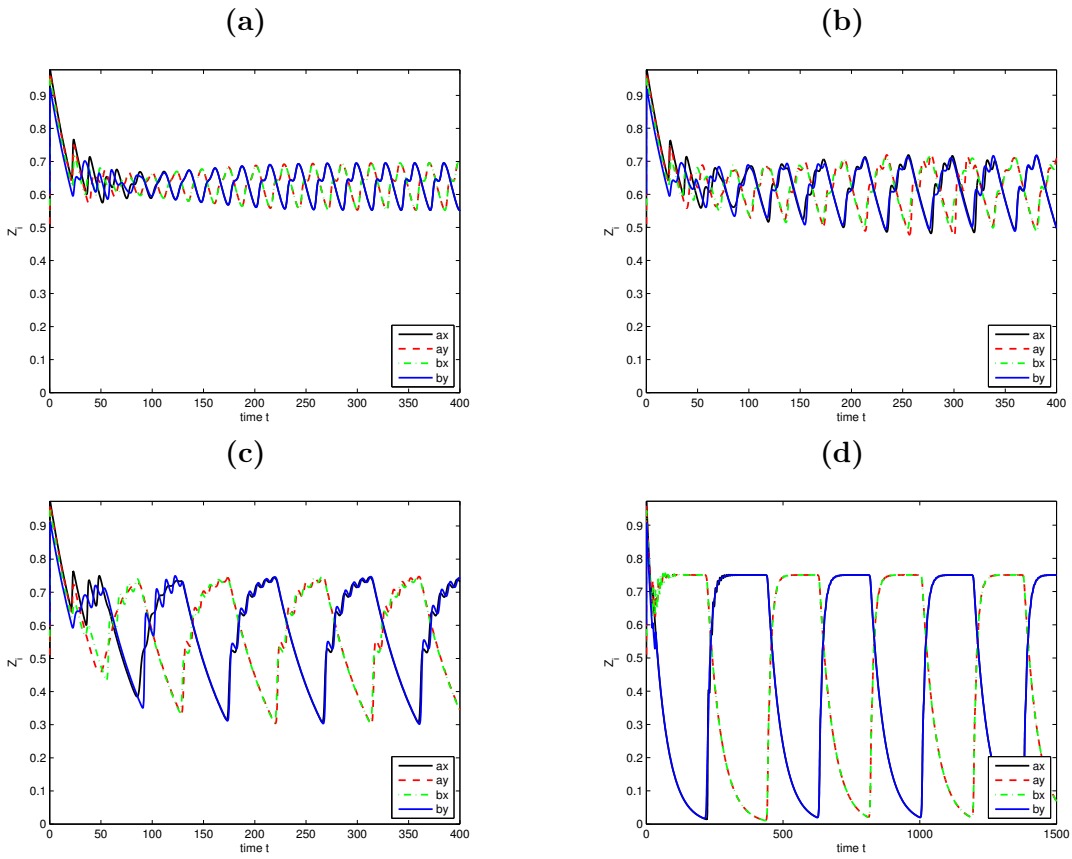


Figure 3.5: **Time Evolution of Cyclic Strain Structure (Eqs. (3.4)-(3.6)).** The graphs illustrate the dependency of the dynamics of the model on γ. **(a)** $\gamma = 0.58$, **(b)** $\gamma = 0.62$, **(c)** $\gamma = 0.72$, and **(d)** $\gamma = 0.85$. As the cross-immunity level γ is varied, in the CSS region, the strains are self-organized into sets of discordant strains alternating in abundance, where each discordant set is half a period out of phase. Equations (3.4)-(3.6) were solved using the Runge-Kutta fourth- and fifth-order method solver ODE45 in MATLAB (The Mathworks Inc., Natwick, MA, USA) with the following initial conditions $z(0) = [0.50\ 0.51\ 0.55\ 0.6]$, $w(0) = [0.40\ 0.30\ 0.3\ 0.30]$ and $y(0) = [0.50\ 0.40\ 0.30\ 0.20]$, and relative and absolute error tolerance of 10^{-5} and 10^{-6}, respectively.

NSS, CSS and DSS) of parameter space shown in Fig. 3.6 remain consistent. To illustrate this concretely, we look at three examples where the number of loci is two, three, and four, respectively, with two alleles each, which are summarized in Fig. 3.8. In the illustration, we focus mainly on weakly endemic equilibria types exhibiting DSS, for there is always a unique strongly endemic steady state, and its stability depends on the conditions in Theorem 3.4.7. When the conditions of Theorem 3.4.7 are violated, the unique strongly endemic equilibrium is unstable and undergoes a Hopf bifurcation, exhibiting cyclical strain structures.

In the case of the 2-locus-2-allele system $\mathcal{G}_2(2,2)$, we observe a natural pairing of

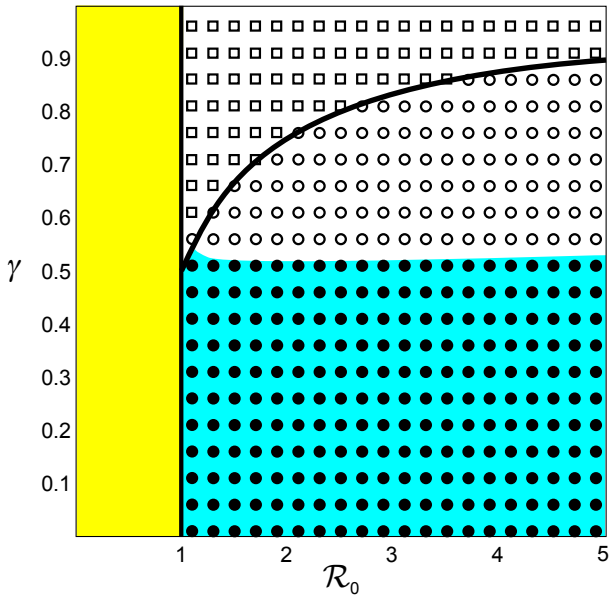


Figure 3.6: **Parameter Space $\mathcal{R}_0$-γ for 2-locus-2-allele system.** Different dynamic behaviours are possible depending on the combination of $\mathcal{R}_0$ and γ. The solid curve is given by inequality (3.13). In the yellow region (i.e., $\mathcal{R}_0 < 1$), the disease-free equilibrium is the only biologically relevant steady state and is globally asymptotically stable. In the white region with black squares, the weakly endemic equilibrium exhibiting discrete strain structures (DSS) are locally asymptotically stable, and they are unstable outside this region. In the region of stability, the maintenance of invading strains related to the discordant persisting set is impossible. In the cyan region with black dots, the unique strongly endemic equilibrium is locally asymptotically stable and exhibits no strain structures (NSS); it is unstable outside this region. In the white region between the black line and cyan boundary, the unique strongly endemic equilibrium is unstable and exhibits cyclical behaviours (hence cyclical strain structures, CSS, see Theorem 3.4.7). The black lines are steady state bifurcation curves (see inequality (3.13)), and the cyan boundary is a Hopf bifurcation curve. The parameter values $\sigma = 9.98$, $\mu = 0.02$, $\rho = 2$ and $\alpha = 3$ are used.

discordant sets resulting in two clusters where strains marked by blue belong to one cluster (e.g., $C_1 = \{00; 11\}$ or **ax** and **by** of the motivating notation) and by black belong to the other (e.g., $C_2 = \{10; 01\}$ or **ay** and **bx**) as shown in Fig. 3.8(**a**). The dynamics corresponding to this system is illustrated in Figs. 3.2, 3.4 and 3.5. In this system only DSS of discordant types, CSS with discordant sets exhibiting anti-phase periodic solution and fully synchronized steady state exhibiting NSS are possible.

For the 3-locus-2-allele system, there are numerous weakly endemic equilibria, but most of them are unstable. Discrete strain structures containing an balanced number of persistent strains are locally asymptotically stable provided $\rho \neq 0$. For example, in the 3-locus-2-allele system, one clustering structure consists of discordant strains, where there are four clusters with two discordant strains corresponding to the main diagonal of the 3-hypercube in Fig. 3.8(b) (e.g., $\mathcal{C}_1 = \{000, 111\}$, $\mathcal{C}_2 = \{001, 110\}$,

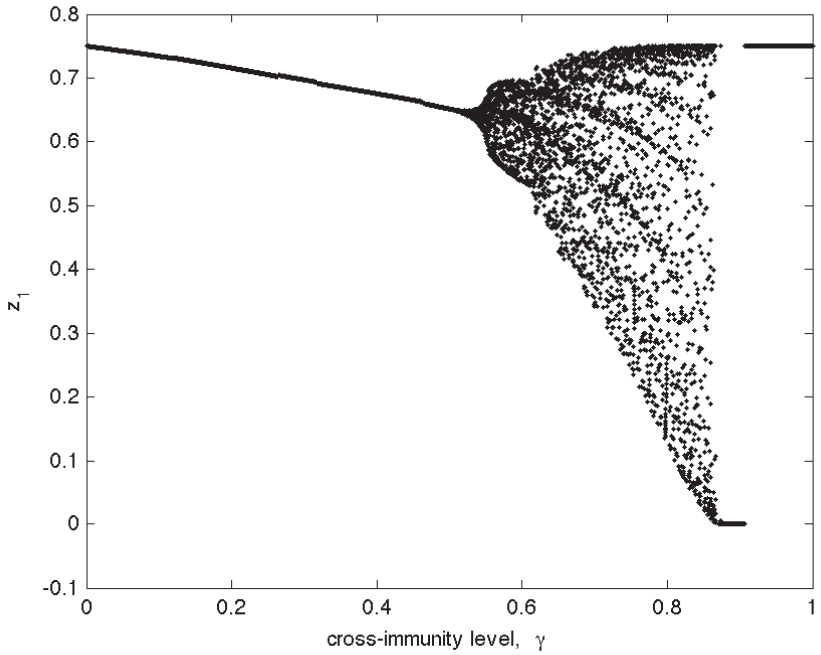


Figure 3.7: **Steady-State Bifurcation Diagram.** For the 2-locus-2-allele system, the stability of the original system (see Eqs. (3.4)-(3.6)) varies with respect to cross-immunity level γ, for $\mathcal{R}_0 = 4$, $\sigma = 9.98$, $\mu = 0.02$. Note the three distinct regions emerging, namely no strain structure, cyclic strain structure, and discrete strain structure (see text for details). Here, we are showing only z_1 (i.e., z_{ax}) and linearly stable states. Equations (3.4)-(3.6) were solved using the Runge-Kutta fourth- and fifth-order method solver ODE45 in MATLAB (The Mathworks Inc., Natwick, MA, USA) with the following initial conditions $z(0) = [0.50\ 0.51\ 0.55\ 0.6]$, $w(0) = [0.40\ 0.30\ 0.3\ 0.30]$ and $y(0) = [0.50\ 0.40\ 0.30\ 0.20]$, and relative and absolute error tolerance of 10^{-5} and 10^{-6}, respectively.

$\mathcal{C}_3 = \{011, 100\}$ and $\mathcal{C}_4 = \{010, 101\}$). Any one of these clusters can persist provided the initial condition is in the appropriate stability region (i.e., region of attraction). In addition, we also observe that the clustering $\mathcal{C}_1 = \{000, 100, 011, 111\}$ and $\mathcal{C}_2 = \{010, 110, 001, 101\}$ can also persist since $\rho \neq 0$. Furthermore, the cluster type of two tetrahedral structure with four strains each (strains marked by a blue colour form one cluster, and those marked by a black colour form another), where members of each cluster share one allele at the same locus location (e.g., $\mathcal{C}_1 = \{000, 011, 101, 110\}$ and $\mathcal{C}_2 = \{111, 100, 010, 001\}$) cannot persist since $\rho = 0$. These observations are consistent with the groupoid approach of Chan & Yu (2013).

For the third example, the system formed by the 4-locus-2-allele results in a 4-hypercube, the number of weakly endemic equilibria increases substantially. As in the previous case, all clusters containing an unbalanced number of persistent strains are unstable. For simplicity, we focus on three cluster types. The first structure, as in all the above cases, consists of eight clusters with two discordant strains that

correspond to the main diagonal of the 4-hypercube in Fig. 3.8(c). The second case has eight clusters with two strains where members of each cluster share one allele that is at the same locus. This has the same structure of the second cluster type in the 3-hypercube, whereas strains in this case belong to the diagonal of the 3-hypercube face of the 4-hypercube. The third type is the formation of two clusters (marked by black colour and blue colour) of eight strains where members of each cluster share either no allele or two alleles as illustrated in Fig. 3.8(c) (e.g., the first cluster is marked by a blue colour and the second cluster is marked by a black colour). Under the results of Corollary 3.4.5, the first and the third cases are possible since $\rho \neq 0$.

From these examples, one observes that as we increase the number of loci, the clusters formed consist of clusters in the lower dimensional $(n-1)$-hypercubes and additional clusters that account for the inherited structure of the n-hypercube. Using the *Mendelian* combinatorial property of strain space and the type of cluster structures possible, we can predict the type of weakly endemic equilibria that are locally asymptotically stable (hence the dynamics of partial synchrony). This is due to the fact the structures of higher dimensional strain spaces are built from the building blocks of the lower dimensional strain space. Observation of different cluster types is possible depending on the relative cross-immunity.

Different aspects of the observations outlined above have been discussed elsewhere (Buckee et al., 2004; Calvez et al., 2005; Chan & Yu, 2013; Blyuss, 2013). In particular, for a 3-locus-2-allele system, Buckee et al. (2004) illustrated the formation of these clusters and their relationship with maintenance of diverse strain populations. Using a modified version of the model studied herein, Calvez et al. (2005) made similar observations on the type and structure of clusters possible as the number of loci is increased beyond two and provided a detailed discussion on the behaviour of force of infection (specifically relative force of infection) of individual clusters, mainly the regularity of force of infection of individual clusters and chaotic dynamics of individual strains. The exposition in Calvez et al. (2005) focused primarily on the dynamics in CSS regions and did not investigate the behaviour of cluster types and structures in DSS and NSS regions. Recently, Chan & Yu (2013) showed, using the groupoid formalism of Golubitsky et. al. (2005) and centre-manifold reduction, that DSS consisting of discordant sets are linearly stable provided $\gamma\rho > 1$, where ρ was given in terms of the quotient network parameters (i.e., difference between the number of each strain in $\mathcal{D}_p$ sharing alleles with strains in $\mathcal{D}_\emptyset$ and number of strains sharing alleles in the same set $\mathcal{D}_p$). But the stability of the unique strongly endemic equilibrium and the nature of its bifurcation structure were not addressed in Chan & Yu (2013). Blyuss (2013) has recently provided some detailed discussion of the type of cyclical behaviours possible based on equivariant bifurcation theory and the H/K Theorem

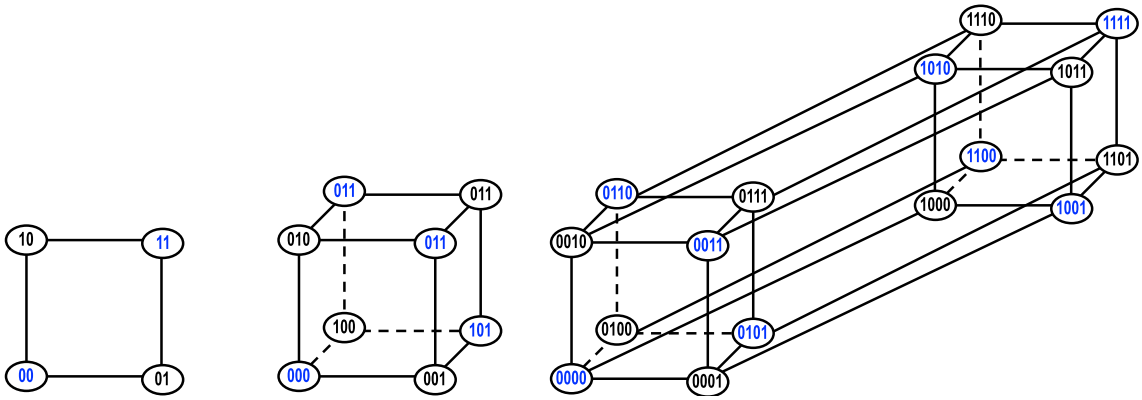


Figure 3.8: **Strain Space Structures.** See text for details.

to explain some of these oscillatory behaviours. However, the results presented herein are more general than those obtained by Chan & Yu (2013) and Blyuss (2013), and their results are consistent with our findings when $\rho = 2$ and $\alpha = 3$.

Our analytical results suggest that, because these clusterings are inherent properties of the Mendelian combinatorial structure of the strain space, the clustering behaviours are not restricted to the CSS region. In addition, one can also observe simple cluster transition with complex or exotic strain dynamics with a varying degree of cluster-specific sequential domination, as illustrated in Calvez et al. (2005), Gupta et al. (1996, 1998), Recker et al. (2007, 2009), Minaev & Ferguson (2009a-b), Chan & Yu (2013), Blyuss (2013) (see Chapter 7 for some examples). To fully understand the parameter regions where these different clustering types and structures, and their respective dynamics, are observed, we must go to a co-dimensional parameter space higher than two or develop a relative cross-immunity measure capable of capturing the diversifying cluster selective properties of multiple cross-immunity levels in order to obtain a parameter space diagram similar to Fig. 3.6. Further analysis of these dynamic patterns and the complex forms of synchronization observed, may be the essential ingredient to understanding and predicting epidemiological episodes, and designing effective and optimal public health schemes

A major question arising from the numerical results of different versions of the model is whether the inherent dynamical behaviours are observed in epidemiological systems. Many pathogens, such as influenza, plasmodium falciparum and neisseria meningitidis, show structured strain diversity which is organized into distinct determinant strain lineages (Gupta et al., 1996; Buckee et al., 2008; Recker et al., 2007). In *Neisseria meningitidis*, for example, the PorA protein (one of the subcapsular outer member proteins) which has two antigenic variable regions VR1 and VR2, is known to exhibit high diversity at population level. These regions are structured into discordant combinations. This is consistent with the results of the aforementioned model

in which strong immune selection organizes the strain population into DSS (Gupta et al., 1996; Buckee et al., 2008). Similarly in Influenza A virus, Gupta and colleagues (Gupta et al., 1998; Recker et al., 2007) have captured the limited diversity, antigenic evolution, cyclical temporal dynamics and population strain structure of influenza using the multi-locus-allele model reviewed above. In addition, it was shown in Recker et al. (2007) that the antigenic distance between epidemics does not increase with time as illustrated in an empirical study (Smith et al., 2004). However, for the model to realistically capture the dynamics observed in influenza, a high number of loci is needed (e.g., five loci or more). The ecology of antigenically variable pathogens is related to their evolution, where these pathogens generate different patterns of antigenic variability. For instance, on the one hand, pathogens such as influenza virus, *S. pneumoniae* and *N. meningitidis* show little variation within a single host. While *S. pneumoniae* and *N. meningitidis* have extensive population-wide variation that is spatiotemporally variable, influenza virus exhibits little variation across the population at any given time but rapid variation over time on a temporal scale of years. On the other hand, infectious diseases such as HIV/AIDS induce extensive replacement of dominant strains with high diversity over time within a single host and across the population. These differences in the patterns of antigenic variability are due to the way in which inter-strain competition provokes or induces cross-reactive immunological response (e.g., immunological interference). For instance, infection with one strain can confer partial cross-immunity or protection to other antigenically similar strains (e.g., influenza virus), or enhance severity of infections by other strains (e.g., dengue). As a result, a multi-locus-allele modelling framework with host immune system mediating competition between strains via cross-immunity can be used to provide insight into various infectious diseases such as dengue, meningitis, malaria and influenza viruses, to name but a few examples of antigenically variable infectious diseases.

3.6 Conclusion

Infectious diseases with antigenically diverse pathogens, such as influenza, remain a major threat to humanity. The recurrent epidemiological patterns, continuous emergence and re-emergence of new subtypes, the ability to evade immunological detection, and public health challenges, have made these diseases more intriguing to scientific inquiries. Herein, we have studied one of the models often used to investigate the interplay between epidemiological, immunological and evolutionary dynamics. In particular, we have provided a detailed analysis of a multi-locus-allele model proposed

previously by Gupta et al. (1996, 1998) and focused our attention on the fundamental model of the 2-locus-2-allele system. Using both biological insights and some mathematical definitions of endemicity and persistent sets, we split the steady state analysis into two parts based on the nature of endemicity, namely weakly endemicity and strongly endemicity where in the former some strains are persistent while others go extinct and in the latter all strains coexist. Reducing the system according to the steady state types allows us to simplify our analysis. Analyzing these steady state types, we found that weakly endemic equilibria can only undergo steady-state bifurcation while the unique strong endemic equilibrium can undergo bifurcation either to a steady-state or to an oscillatory solution (Hopf).

Although the results presented herein have focused on local properties of the equilibria, the assumptions of life-long immunity, uniformity of model parameters, and that a strain sharing one allele has the same cross-reactivity effect as having more than one allele in common, one could pursue further research in a variety of directions. The easiest extensions of the model are to relax the uniformity assumption of the epidemiological parameters and of allelic variation in the general multi-locus-allele system with cross-immunity based on differential allelic distance which will introduce a weighted serotypic cross-immunity network. Furthermore, one can consider the effect of general graph or network topology (e.g., small network topology, random network, real-world strain network motivated by biological evidence) on the dynamical behaviours of the multi-locus-allele systems where the symmetry structure of the network is not readily observable. We can use algebraic graph-theoretical tools to reconcile the approach presented herein, the groupoid formalism of Golubitsky and colleagues (Golubitsky et al., 2005) used by Chan & Yu (2013), and the equivariant dynamical systems approach of Blyuss (2013) and others (see Chossat & Lauterbach, 2000, and references therein) to study the dynamical behaviours and to characterize all generic dynamic patterns of multi-locus-allele systems with both uniform and differential allelic variations. These algebraic approaches can then be used to simplify the model and the analysis and to explore both local and global properties and existence conditions for the evolution of strain serotypic cluster structures. In addition, the assumption of a life-long immunity can be relaxed to include exponential, discrete and/or distributed strain-transcending immunity which can be accomplished by the use of delay and/or functional differential equations. This extension provides a deterministic generalization of the multi-locus-allele framework of Gupta et al. (1996, 1998). Another possibility is incorporating the strain-transcending time-since-infection or the time-since-recovery with or without varying rates of loss of immunity assumption and/or age-specificity into the model. Another mathematical direction one could explore is the effect of intrinsic stochasticity on the transitions between

different cluster types and structures. Recently, Buckee et al. (2011) and Minayev and Ferguson (2009b) have considered the effect of stochasticity on the dynamical patterns of the model proposed in Gupta et al. (1998) and its variation, whereas in Buckee et al. (2011) the authors show that although the self-organization of strains into discordant/non-overlapping or minimally overlapping sets will break down under the differential allelic combination assumption, self-organization is recovered in the stochastic system. It would be interesting to investigate the underlying mechanism of this process mathematically.

In the next chapter, we investigate the age-structured multi-locus-allele system derived in Chapter 4. In particular, we reformulate the multi-locus-allele model as an abstract Cauchy semi-linear evolutionary problem on an infinite-dimensional Banach space and show the existence of a positive mild solution. In addition, we determine a threshold-like condition and examine the existence of both trivial and nontrivial equilibria using this threshold.

4 Derivation of Age-structured Multi-locus-allele Model

This chapter extends the multi-locus-allele model discussed in Chapter 2 by incorporating age-specificity, using hyperbolic partial differential equations.

4.1 Introduction

Structured population models (e.g., spatial structure, temporal variation, multiple groups, differences in activity or risk factors) have been widely studied in the mathematical biology literature for many years. In particular, models written as evolutionary systems of partial differential equations (PDEs) for the density of individuals incorporating a specific structural variable, for example age and/or size for the age-structured and/or size-structured equations (Webb, 1985, 2004; McKendrick, 1926); phenotypic trait for the selection-mutation equations (Burger, 2000); physiologically structured equations (Diekmann & Getto, 2005; Diekmann et al., 2001; Diekmann & Metz, 1986). These structural variables are usually included in the models to enhance features not captured by standard unstructured population models (see Burger, 2000; Perthame, 2007). In particular, structured models have been used to capture differential burdens due to heterogeneity across ages, classes, sizes, physiological traits, and so on, in the presence and/or absence of other prior predisposing risk factors. Furthermore, structural variables such as age and/or size have been incorporated into models in order to induce oscillatory dynamics in the cases where unstructured models do not exhibit such behaviours (Magal & Ruan, 2009a-b, 2010). Therefore, studying epidemiological processes within a structural variable framework is critical to understanding factors that are important in disease transmission and evolution; and to designing effective organized public health measures. In this chapter, we extend the model derived in Chapter 2 (i.e., Section 2.5) in which age is incorporated to account for the age-specificity in influenza dynamics and other pathogenically diverse infectious diseases.

4.2 Antigenic Thrift Model with Age-specificity

We start with the reduced infectivity assumption used in Chapter 2 (i.e., Section 2.5). Let us use the same index set notation, where $\mathcal{K} = \{1, 2, \ldots, M\}$ denotes the set of all possible strains. Let $S_\varnothing$ denote the age density of individuals of age a who, at time t have never been infected and are susceptible, and $S_\mathcal{J}$ denote the age density of susceptible individuals of age a with immunity to the strains in the set $\mathcal{J}$, but no immunity to strains in $\mathcal{K}\backslash\mathcal{J}$ where $\mathcal{J} \subseteq \mathcal{K}$. Similarly, $I^i_\mathcal{J}$ represents the age density of individuals of age a currently infected by strain i who have previously been exposed to, and have recovered from, infections with all strains in the set $\mathcal{J}$ (i.e., $i \notin \mathcal{J}$). Then $I^i = \sum_{\mathcal{J}\in\mathcal{K}\backslash i} I^i_\mathcal{J}$ represents the age density of individuals of age a currently infectious with strain i (see Section 2.5 for details). All the biological assumptions stated in the previous chapters hold. Extending Eqs. (2.7)-(2.10) in Section 2.5, the disease dynamics is given by

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) S_\varnothing(t,a) = -\mu(a)S_\varnothing(t,a) - \sum_{j\in\mathcal{K}} \Lambda^j(t,a)S_\varnothing(t,a) \tag{4.1}$$

$$\begin{aligned}\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) S_\mathcal{J}(t,a) &= \sum_{j\in\mathcal{J}} \Lambda^j(t,a)S_{\mathcal{J}\backslash j}(t,a) - \mu(a)S_\mathcal{J}(t,a) \\ &\quad - \sum_{i\notin\mathcal{J}} \Lambda^i(t,a)S_\mathcal{J}(t,a),\end{aligned} \tag{4.2}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) I^i(t,a) = \sum_{\mathcal{J}\subseteq\mathcal{K}\backslash i} \gamma^i_\mathcal{J}\Lambda^i(t,a)S_\mathcal{J}(t,a) - (\mu(a)+\sigma(a))I^i, \tag{4.3}$$

where the demographic and epidemiological parameters $\mu(a)$, $f(s)$ and $\sigma(a)$ represent the age-specific per capita mortality, per capita fertility rate and per capita recovery rate, respectively, and $\gamma^i_\mathcal{J}$ is the age-specific cross-immunity level. That is, $\gamma^i_\mathcal{J} = 0$ implies complete protection, and $\gamma^i_\mathcal{J} = 1$ implies no protection (e.g., $\gamma^i_\varnothing = 1$ for all i). The force of infection is denoted by Λ^i, which incorporates the infection rate per unit time per capita and the contact rate. The force of infection rate is given as follows:

$$\Lambda^i(t,a) = \frac{\int_0^\infty \beta(a,s)I^i(t,s)ds}{\int_0^\infty P(t,s)ds}, \tag{4.4}$$

where $\beta(a,s)$ represents the combined age-specific transmission and contact rates between individuals of age a and age s, and $N(t) = \int_0^\infty P(t,s)ds$. In addition, we assume that all individuals are born susceptible. Thus, the initial and boundary conditions are given as follows:

$$\begin{aligned}S_\varnothing(t,0) &= \int_0^\infty f(s)P(t,s)ds, \quad S_\mathcal{J}(t,0) = I^i(t,0) = 0 \\ S_\varnothing(0,a) &= S^0_\varnothing(a), \quad S_\mathcal{J}(0,a) = S^0_\mathcal{J}(a), \quad \text{and} \quad I^i(0,a) = I^i_0(a) = 0,\end{aligned}$$

where $S^0_\varnothing(a)$ is the initial age-distribution of susceptibles, and the total population $P(t,a)$ is of the form given by

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)P(t,a) \;=\; -\mu(a)P(t,a), \tag{4.5}$$

with $P(t,0)=\int_0^\infty f(s)P(t,s)ds$ and $P(0,a)=P^0(a)$, and is equivalent to the McKendrick – von Forester equation. The consequence of this equation is that the total population dynamics and demographic rates (e.g., birth and mortality rates) are not affected by the disease.

We now provide a two strain case to illustrate the use of the notation. For the two strain system $\mathcal{K}=\{1,2\}$, and $S_\varnothing$ is susceptible to all strains, S_1 is immune to strain 1 but susceptible to strain 2, S_2 is immune to strain 2 but susceptible to strain 1, S_{12} is immune to both strains 1 and 2; and $I^1=I_1+I_{12}$ and $I^2=I_2+I_{21}$ are infectious with strain 1 and 2 respectively, then the dynamical system is described as follows:

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)S_\varnothing(t,a) \;=\; -\mu(a)S_\varnothing(t,a)-\left(\Lambda^1(t,a)+\Lambda^2(t,a)\right)S_\varnothing(t,a), \tag{4.6}$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)S_1(t,a) \;=\; \Lambda^1(t,a)S_\varnothing(t,a)-\mu(a)S_1(t,a)-\Lambda^2(t,a)S_1(t,a), \tag{4.7}$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)S_2(t,a) \;=\; \Lambda^2(t,a)S_\varnothing(t,a)-\mu(a)S_2(t,a)-\Lambda^1(t,a)S_2(t,a), \tag{4.8}$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)S_{12}(t,a) \;=\; \Lambda^1(t,a)S_2(t,a)+\Lambda^2(t,a)S_1(t,a)-\mu(a)S_{12}(t,a), \tag{4.9}$$

$$\begin{aligned}\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)I^1(t,a) \;&=\; \left(\gamma^1_\varnothing S_\varnothing(t,a)+\gamma^1_2 S_2(t,a)\right)\Lambda^1(t,a)\\ &\quad -(\mu(a)+\sigma(a))I^1,\end{aligned} \tag{4.10}$$

$$\begin{aligned}\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)I^2(t,a) \;&=\; \left(\gamma^2_\varnothing S_\varnothing(t,a)+\gamma^2_1 S_1(t,a)\right)\Lambda^2(t,a)\\ &\quad -(\mu(a)+\sigma(a))I^2,\end{aligned} \tag{4.11}$$

where $\gamma^1_\varnothing=\gamma^2_\varnothing=1$ and $\gamma^1_2=\gamma^2_1=1-\gamma$. The initial and boundary conditions are given by

$$\begin{aligned}S_\varnothing(t,0) \;&=\; \int_0^\infty f(s)P(t,s)ds,\;\; S_1(t,0)=S_2(t,0)=S_{12}(t,0)=I^1(t,0)=I^2(t,0)=0\\ S_\varnothing(0,a) \;&=\; S^0_\varnothing(a),\;\; S_1(0,a)=S^0_1(a),\;\; S_2(0,a)=S^0_2(a),\;\; S_{12}(0,a)=S^0_{12}(a),\\ &\quad I^1(0,a)=I^1_0(a),\;\text{ and }\; I^2(0,a)=I^2_0(a).\end{aligned}$$

Observe that this is the age-structured version of the two strain system with reduced infectivity discussed in Section 2.5 (see Chapter 2).

4.3 Model Reduction via Auxiliary Quantity

In Chapter 2, we use two auxiliary quantities to reduce the model. Here we use the auxiliary variables to reduce the dimension of the model. However, since we are using a non-standard summation, we illustrate the use of the notation. The first auxiliary $\tilde{S}_i(t,a) = \sum_{\mathcal{J}:i\in\mathcal{J}} S_{\mathcal{J}}(t,a)$ denotes susceptible individuals who have experienced at least strain i. For the two strain system, $\tilde{S}_1 = S_1 + S_{12}$ and $\tilde{S}_2 = S_2 + S_{12}$. Then using the auxiliary variable and Eqs. (4.6)-(4.11), we have

$$\begin{aligned}
\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right) S_{\varnothing}(t,a) &= -\mu(a)S_{\varnothing}(t,a) - \left(\Lambda^1(t,a)+\Lambda^2(t,a)\right) S_{\varnothing}(t,a),\\
\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right) \tilde{S}_1(t,a) &= \Lambda^1(t,a)\left(P-\tilde{S}_1(t,a)\right) - \mu(a)\tilde{S}_1(t,a),\\
\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right) \tilde{S}_2(t,a) &= \Lambda^2(t,a)\left(P-\tilde{S}_2(t,a),\right) - \mu(a)\tilde{S}_2(t,a),\\
\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right) I^1(t,a) &= \left(S_{\varnothing}+(1-\gamma)\left(P-S_{\varnothing}-\tilde{S}_1\right)\right)\Lambda^1(t,a) - (\mu(a)+\sigma(a))I^1,\\
\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right) I^2(t,a) &= \left(S_{\varnothing}+(1-\gamma)\left(P-S_{\varnothing}-\tilde{S}_2\right)\right)\Lambda^2(t,a) - (\mu(a)+\sigma(a))I^2.
\end{aligned}$$

where $\gamma^i_{\varnothing} = 1$ for all $i = 1, 2$, $\gamma^i_j = 1-\gamma$, and $P = S_{\varnothing} + S_1 + S_2 + S_{12}$. We note that the auxiliary quantity $\tilde{S}_i$ reduces the original systems of equations by one in the two strain system.

We now turn to the use of the second auxiliary variable $W^i_k(t,a)$, where $W^i_k(t,a) = \sum_{\mathcal{L}_k(i)\bigcap\mathcal{J}\neq\emptyset} S_{\mathcal{J}}(t,a)$ denotes the number of individuals who have immunity to one or more strains in the $\mathcal{L}_k(i)$-neighborhood of i for all k levels of protection (i.e., $k = 0, 1, 2, \ldots, m$). Here $\mathcal{L}_k(i)$ represents $\text{dist}(i,\mathcal{J}) \leq k$, where $\text{dist}(x,y)$ is the distance between elements x and y. Therefore, the summation $W^i_k(t,a) = \sum_{\mathcal{L}_k(i)\bigcap\mathcal{J}\neq\emptyset} S_{\mathcal{J}}(t,a)$ counts all strains satisfying the collection of all sets that are up to k distance (e.g., Hamming distance) away. For example, $W^i_0(t,a)$ represents the age density of individuals who have been exposed to strain i at some point in their life and are now immune to it, $W^i_1(t,a)$ denotes the age density of individuals who have been exposed to, and are immune to, strain i or any strain related to i up strains that differ by 1 epitope, and so on. So by definition $W^i_m(t,a) = P(t,a)$. For the two strain system, $W^1_0 = S_1 + S_{12}$, $W^2_0 = S_2 + S_{12}$, $W^1_1 = S_1 + S_2 + S_{12} = W^2_1$. Note that here $W^1_1 = W^2_1$ because in our early assumption we assume that strain 1 and 2 are similar. Using the second auxiliary, we observe that W^i_0 is identical to $\tilde{S}_i$, and is given by

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right) W^1_0(t,a) = \Lambda^1(t,a)\left(P-W^1_0(t,a)\right) - \mu(a)W^1_0(t,a),$$

and because $W_1^1 = W_1^2$, we can ignore W_1^2. Using the definition of W_1^1, and re-expressing I^1, we obtain

$$\begin{aligned}
\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right) W_1^1(t,a) &= \left(P - W_1^1\right)\left(\Lambda^1+\Lambda^2\right) - \mu(a) W_0^1(t,a) \\
\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right) I^1(t,a) &= \left(P - W_1^1 + (1-\gamma)\left(W_1^1 - W_0^1\right)\right)\Lambda^1(t,a) - (\mu(a)+\sigma(a)) I^1,
\end{aligned}$$

and similarly for W_1^2 and I^2. We note that since $W_1^2 = W_1^1$, the resulting system of equations is reduced to five equations and $S_\varnothing$ is no longer included in the final equations. For a more general expression, we can reduce the dynamic equations in Eqs. (4.1)-(4.3), and the resulting equations can be rewritten in more compact forms. While observing that the number of individuals who do not have $\mathcal{L}_k(i)$ immunity is $P(t,a) - W_k^i(t,a)$, we sum the dynamic equation for $S_{\mathcal{J}}(t,a)$ over $\mathcal{L}_k(i) \bigcap \mathcal{J} \neq \emptyset$, to obtain the following expressions:

$$\frac{\partial W_k^i}{\partial t} + \frac{\partial W_k^i}{\partial a} = \left(P(t,a) - W_k^i(t,a)\right) \sum_{j \in \mathcal{L}_k(i)} \Lambda^j(t,a) - \mu(a) W_k^i(t,a), \quad (4.12)$$

$$\begin{aligned}
\frac{\partial I^i}{\partial t} + \frac{\partial I^i}{\partial a} &= \Lambda^i \sum_{k=1}^{m} \gamma_k \left(W_k^i(t,a) - W_{k-1}^i(t,a)\right) - \mu(a) I^i(t,a) \\
&\quad -\sigma(a) I^i(t,a), \qquad (4.13)
\end{aligned}$$

where $\gamma_{\mathcal{J}}^i = \gamma_k$ if $\mathcal{J} \bigcap (\mathcal{L}_k(i) \backslash \mathcal{L}_{k-1}(i))$, and $0 = \gamma_0 \leq \gamma_1 \leq \gamma_2 \leq \ldots \leq \gamma_m = 1$, $W_m^i(t,a) = P(t,a)$. Here γ_k reduces disease transmissibility for individuals whose most related infection is at the k-level. Hence this is the age-structured extension of the model studied by Ferguson & Andreasen (2001). We note that when $m = 2$, Eqs. (4.12)-(4.13) reduce to an age-structured extension of the equations in Gupta et al. (1997a-b, 1998), except that γ_k is defined in the opposite sense (i.e., reverse cross-immunity, $\gamma_1 = 1-\gamma$). Henceforth, we assume that $m = 2$. Furthermore, it should be noted that if we let Ω_{ij} denote the $ij-$element of the similarity matrix generated by the interaction between strains, then Ω_{ij} based on the lexicographic representation chosen for any given strain space can be used to rewrite $\sum_{j \in \mathcal{L}_1(i)} \Lambda^j = \sum_j \Omega_{ij} \Lambda^j$.

For $m = 2$, let $\tilde{S}_i = P - W_1^i$ denote the number of individuals who do not have immunity to strain i; $V_i = W_1^i - W_0^i$ represent individuals with partial immunity to strain i; $Z_i = W_0^i$ denote individuals fully immune to strain i; and $Y_i = I^i$ denote infectious individuals. Then the above equations become:

$$\frac{\partial \tilde{S}_i}{\partial t} + \frac{\partial \tilde{S}_i}{\partial a} = -\tilde{S}_i(t,a) \sum_j \Omega_{ij} \Lambda^j(t,a) - \mu(a) \tilde{S}_i(t,a), \quad (4.14)$$

$$\frac{\partial V_i}{\partial t} + \frac{\partial V_i}{\partial a} = \tilde{S}_i(t,a) \sum_{j \neq i} \Omega_{ij} \Lambda^j(t,a) - V_i(t,a) \Lambda^i(t,a) - \mu(a) V_i(t,a), \quad (4.15)$$

$$\begin{aligned}
\frac{\partial Y_i}{\partial t} + \frac{\partial Y_i}{\partial a} &= \Lambda^i(t,a)\left[(1-\gamma)\,V_i(t,a) + \tilde{S}_i(t,a)\right] \\
&\quad - (\mu(a)+\sigma(a))\,Y_i(t,a), \qquad (4.16)\\
\frac{\partial Z_i}{\partial t} + \frac{\partial Z_i}{\partial a} &= \left(\tilde{S}_i(t,a) + V_i(t,a)\right)\Lambda^i(t,a) - \mu(a)Z_i(t,a), \qquad (4.17)
\end{aligned}$$

with the following initial conditions:

$$\begin{aligned}
\tilde{S}_i(t,0) &= \int_0^\infty f(s)P(t,s)ds \\
V_i(t,0) &= Z_i(t,0) = Y_i(t,0) = 0 \\
\tilde{S}_i(0,a) &= \tilde{S}_i^0(a),\ \ V_i(0,a) = V_i^0(a),\ \ Z_i(0,a) = Z_i^0(a),\ \text{ and }\ Y_i(0,a) = Y_i^0(a).
\end{aligned}$$

Since the age-specific mortality and fertility (including zero net-migration) rates are constant, we can use the stable age-distribution assumption (see, for example, discussions in Webb (1985), Iannelli (1995) and Thieme (2003) for details) to find that $P(t,a) \to P^\infty(a)$ as $t \to \infty$, where $P^\infty(a)$ is the stable population age-distribution. In particular, we use the notion of population ergodicity where the population eventually forgets its initial age-distribution and takes the features determined by the vital rates to which it is exposed. It should be noted that stable population (in age-distributional sense) is different from stationary population, where stationary population has mortality and birth rates are equal (i.e., zero growth) while stable population has constant growth, mortality and birth rates and age-distribution. Hence, stationary population can be seen as a special form of stable population.

Let $s_i = \dfrac{\tilde{S}_i}{P^\infty}$, $v_i = \dfrac{V_i}{P^\infty}$, $z_i = \dfrac{Z_i}{P^\infty}$, and $y_i = \dfrac{Y_i}{P^\infty}$, where $P(t,a) = P^\infty(a)$. Then, the system can be rewritten as follows:

$$\begin{aligned}
\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) s_i(a,t) &= -s_i \sum_j \Omega_{ij}\lambda_j \qquad (4.18)\\
\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) v_i(a,t) &= s_i \sum_{j\neq i} \Omega_{ij}\lambda_j - v_i\lambda_i \qquad (4.19)\\
\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) y_i(a,t) &= \left[s_i + (1-\gamma(a))v_i\right]\lambda_i - \sigma(a)y_i \qquad (4.20)\\
\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) z_i(a,t) &= (s_i + v_i)\lambda_i \qquad (4.21)
\end{aligned}$$

where $s_i(t,0) = 1,$ and $v_i(t,0) = z_i(t,0) = y_i(t,0) = 0$ and $s_i(0,a) = s_i^0(a),\ v_i(0,a) = v_i^0(a),\ z_i(0,a) = z_i^0(a),$ and $y_i(0,a) = y_i^0(a)$, and the force of infection is given by:

$$\lambda_i(t,a) = \int_0^\infty \beta(a,s)\psi(s)y_i(t,s)ds, \text{ and } \psi(a) = \frac{P^\infty(a)}{\int_0^\infty P^\infty(a)da},$$

where ψ is the normalized stable age distribution and $\int_0^\infty \psi(s)ds = 1$. In Appendix B.1.2, we numerically show that, when all the age-specific parameters $\beta,\ f,\ \sigma,\ \gamma,$ and μ are independent of age, the age-structured model reduces to the model of Gupta et al. (1998) derived in Chapter 2. In Chapters 5-6, we provide some rigorous results on the existence of a mild positive solution and non-trivial steady state solutions, and on the stability of the equilibria using insight gained in Chapter 3.

5 Age-specificity I: Existence of Mild and Steady-State Solutions

In this chapter, we provide a theoretical investigation of the multi-strain system derived in Chapter 4. We focus on an age-structured multi-locus-allele model and provide some analyses. We formulate the multi-locus-allele model as an abstract Cauchy semi-linear evolutionary problem on an infinite-dimensional Banach space and show the existence of mild solutions. In addition, under some appropriate (sometimes restrictive) assumptions, we determine a threshold-like condition and examine the existence of both trivial and nontrivial equilibria using this threshold.

5.1 Antigenic Thrift Model with Age-specificity

In this chapter, we consider a system of coupled initial-boundary value problems that models the evolution of the multiple strain age-structured population derived in the previous chapter. Let s_i, v_i, z_i, and y_i denote the proportion of individuals who do not have immunity to strain i or similar to it (i.e., susceptible to any strain similar to strain i and including i itself), the proportion of individuals with partial immunity to strain i, the proportion of individuals immune to strain i, and the proportion of individuals infectious with strain i, respectively. The epidemiological process is described by the following partial differential equations:

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right) s_i(t,a) = -s_i \sum_j \Omega_{ij}\lambda_j, \tag{5.1}$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right) v_i(t,a) = s_i \sum_{j\neq i} \Omega_{ij}\lambda_j - v_i\lambda_i, \tag{5.2}$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right) y_i(t,a) = \left[s_i + (1-\gamma(a))v_i\right]\lambda_i - \sigma(a)y_i, \tag{5.3}$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right) z_i(t,a) = (s_i+v_i)\lambda_i, \tag{5.4}$$

where Ω_{ij} is the similarity or relatedness matrix, and $\Omega_{ij}=1$ if strain i is similar to strain j or $i=j$, and $\Omega_{ij}=0$ otherwise; in addition, the initial and boundary

conditions are given as follows $s_i(t,0) = 1$, and $v_i(t,0) = z_i(t,0) = y_i(t,0) = 0$ and $s_i(0,a) = s_i^0(a)$, $v_i(0,a) = v_i^0(a)$, $z_i(0,a) = z_i^0(a)$, and $y_i(0,a) = y_i^0(a)$, $i = 1,2,\ldots,\mathcal{N}$ and $\mathcal{N}$ is the number of strains. The force of infection is given by:

$$\begin{aligned}\lambda_i(t,a) &= \int_0^\infty \beta(a,s)\psi(s)y_i(t,s)ds,\\ \psi(a) &= \frac{P^\infty(a)}{\int_0^\infty P^\infty(a)da} = \frac{\mathcal{E}(a)}{\int_0^\infty \mathcal{E}(s)ds},\end{aligned}$$

where we have assumed a stable age distribution $P^\infty(a) = \mathcal{E}(a)$, where $\mathcal{E}(a) := \exp\left(-\int_0^a \mu(s)ds\right)$; $\int_0^\infty \psi(s)ds = 1$ and ψ is the normalized stable age distribution; and the parameters $\beta(a,s)$, $\mu(a)$, $\gamma(a)$ and $\sigma(a)$ denote age-specific contact rate, mortality rate, cross-immunity level and recovery rate, respectively. We also note that $L = \int_0^\infty \mathcal{E}(s)ds$ is the life-expectancy and $D = \int_0^\infty \exp\left(-\int_0^a \sigma(s)ds\right)da$ is infectious period/or mean-séjour. In addition, because the other equations are independent of the variable z_i, we can eliminate z_i from the system. Hence, we will consider only Eqs. 5.1-5.3 to study existence and stability questions. First, we restate the technical and biologically reasonable assumptions made in Chapter 4. That is,

Assumption 5.1.1. *We assume the following properties on demographic patterns:*

1. $\mathbf{H_1}$**:** $\mu(a) \in L^\infty([0,\infty))$, $\mu(a) \geq 0$ *in* $[0,\infty)$.

5.2 Existence of a Positive Solution:

We now consider the well-posedness of the evolution problem described by Eqs. (5.1)–(5.3) on the Banach Space

$$\mathcal{X} := L^1([0,\infty);\mathbb{R}^{3\mathcal{N}}) = L^1([0,\infty))^{3\mathcal{N}},$$

endowed with the usual norm

$$||\phi|| = \sum_{i=1}^{3}\sum_{j=1}^{\mathcal{N}} ||\phi_j^{(i)}||,$$

for $\phi(a) = \{\phi_1^{(1)}(a), \phi_2^{(1)}(a), \ldots, \phi_{\mathcal{N}}^{(1)}(a), \phi_1^{(2)}(a), \phi_2^{(2)}(a), \ldots, \phi_{\mathcal{N}}^{(2)}(a), \phi_1^{(3)}(a), \phi_2^{(3)}(a), \ldots, \phi_{\mathcal{N}}^{(3)}(a)\}^T \in \mathcal{X}$, where $||*||$ is the norm of $L^1([0,\infty))$ and T denotes the transpose operation. In addition, $\phi_k = \{\phi_j^{(k)}\}_{j=1,2,\ldots,\mathcal{N}} \in \mathbb{R}^{\mathcal{N}\times 1}$, $k = 1,2,3$. The state space of the normalized age distribution for the system induced by the semi-flow Eqs. (5.1)–(5.3) is given as follows:

$$\Gamma := \{\phi := (\phi_1,\phi_2,\phi_3) \in \mathcal{X}_+ : 0_{\mathcal{N}\times 1} \leq \phi_1 + \phi_2 + \phi_3 \leq 1_{\mathcal{N}\times 1}\},$$

where $\mathcal{X}_+ = L^1_+([0,\infty))^{3\mathcal{N}}$ and $L^1_+([0,\infty))$ denotes the positive cone of $L^1([0,\infty))$ (i.e., $L^1_+ = \{f \in L^1 : f \geq 0 \text{ a.e.}\}$).

Let $\mathcal{A}$ be the linear operator on $\mathcal{X}$ defined by:

$$\mathcal{A}[\phi](a) := \begin{pmatrix} \mathcal{A}_1 & 0 & 0 \\ 0 & \mathcal{A}_2 & 0 \\ 0 & 0 & \mathcal{A}_3 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix},$$

where $\mathcal{A}_1 := -\mathbb{I}_{\mathcal{N}\times\mathcal{N}}(\frac{\partial}{\partial a})$, $\mathcal{A}_2 := -\mathbb{I}_{\mathcal{N}\times\mathcal{N}}(\frac{\partial}{\partial a})$ and $\mathcal{A}_3 := -\mathbb{I}_{\mathcal{N}\times\mathcal{N}}(\frac{\partial}{\partial a} + \sigma(a))$. In addition, $\phi = \{\phi_1, \phi_2, \phi_3\}^T \in \mathcal{D}(\mathcal{A})$. The domain $\mathcal{D}(\mathcal{A})$ is defined as $\mathcal{D}(\mathcal{A}) := \{\phi \in \mathcal{X} : \phi_1, \phi_2, \phi_3 \text{ are absolutely continuous on } [0,\infty) \text{ and } \phi_1(0) = 1_{\mathcal{N}\times 1}, \phi_2(0) = 0_{\mathcal{N}\times 1}, \phi_3(0) = 0_{\mathcal{N}\times 1}\}$. We also let $\mathcal{F}$ be a nonlinear operator $\mathcal{F} : \mathcal{X} \to \mathcal{X}$ and is defined by:

$$\begin{aligned} \mathcal{F}[\phi](a) &:= \begin{pmatrix} \mathcal{F}_1[\phi](a) \\ \mathcal{F}_2[\phi](a) \\ \mathcal{F}_3[\phi](a) \end{pmatrix}, \\ \mathcal{F}_k[\phi](a) &:= \begin{pmatrix} \mathcal{F}_1^{(k)}[\phi](a) \\ \mathcal{F}_2^{(k)}[\phi](a) \\ \cdots \\ \mathcal{F}_{\mathcal{N}}^{(k)}[\phi](a) \end{pmatrix}, \quad k = 1,2,3, \end{aligned}$$

where

$$\begin{aligned} \mathcal{F}_i^{(1)}[\phi](a) &:= -\phi_i^{(1)}(a) \sum_j \Omega_{ij} \lambda_j[\phi](a) \\ \mathcal{F}_i^{(2)}[\phi](a) &:= \phi_i^{(1)}(a) \sum_{j\neq i} \Omega_{ij} \lambda_j[\phi](a) - \phi_i^{(2)} \lambda_i[\phi](a) \\ \mathcal{F}_i^{(3)}[\phi](a) &:= \left(\phi_i^{(1)}(a) + (1-\gamma(a))\,\phi_i^{(2)}(a)\right) \lambda_i[\phi](a), \quad i = 1,2,\ldots,\mathcal{N}. \end{aligned}$$

The operator $\lambda_j[\phi](a) : \mathcal{X} \to L^1([0,\infty])$ is the linear operator:

$$\lambda_j[\phi](a) = \int_0^\infty \beta(a,s)\psi(s)\phi_j^{(3)}(s)ds.$$

Before going any further, we make the following assumption on the age-specific parameters:

Assumption 5.2.1. *The following properties on the epidemiological parameters hold:*

1. **$\mathbf{H}_2$:** *$\beta(a,s) \in L^\infty_+([0,\infty)\times[0,\infty))$ and $\beta(a,s) \geq 0$. In addition, $\beta(a,s) \leq \tilde{\beta}(a)$ with $\tilde{\beta}(a) \in L^1([0,\infty))$ and $\overline{\beta} = ess\ sup\ \beta(a,s)$.*

2. **$\mathbf{H}_3$:** *$\sigma(a) \in L^\infty_+([0,\infty))$ and $\sigma(a) \geq 0$ with $\overline{\sigma} = ess\ sup\ \sigma(a)$.*

3. **$\mathbf{H}_4$:** *$\gamma(a) \in L^\infty_+([0,\infty))$ and $0 \leq \gamma(a) \leq 1$ with $\overline{\gamma} = ess\ sup\ \gamma(a)$.*

4. **$\mathbf{H}_5$:** *$\psi(a) \in L^\infty_+([0,\infty))$ and $\psi(a) \geq 0$ with $\overline{\psi} = ess\ sup\ \psi(a)$.*

Now let $u = \{u_1, u_2, u_3\} \in L^1([0,\infty), \mathbb{R}^{3\mathcal{N}}_+)$, where $u_1 = \{s_i\}_{i=1,2,...,\mathcal{N}}$, $u_2 = \{v_i\}_{i=1,2,...,\mathcal{N}}$ and $u_3 = \{y_i\}_{i=1,2,...,\mathcal{N}}$, $u_k \in L^1([0,\infty), \mathbb{R}^{\mathcal{N}}_+)$. Then Eqs. (5.1)–(5.3) can be reformulated as the abstract semi-linear Cauchy evolution equation,

$$\frac{du}{dt} \quad = \quad \mathcal{A}u + \mathcal{F}(u), \tag{5.5}$$

where $\mathcal{A}$ and $\mathcal{F}$ satisfy the following properties:

Lemma 5.2.2. *The linear operator $\mathcal{A}$ is the infinitesimal generator of a strongly continuous semigroup (C_0-semigroup) of bounded linear operators $\{T_{\mathcal{A}}(t)\}_{t\geq 0} := \{e^{t\mathcal{A}}\}_{t\geq 0}$ on $\mathcal{X}$. In addition, the linear operator $T_{\mathcal{A}}(t)$ is defined by*

$$T_{\mathcal{A}}(t)\,[\phi] := \begin{pmatrix} T_1(t)\,[\phi] \\ T_2(t)\,[\phi] \\ T_3(t)\,[\phi] \end{pmatrix},$$

where $\{T_i(t)[\phi]\}_{t\geq 0},\ i = 1,2$ is a nilpotent right translation C_0-semigroup of operators given by $T_1(t)\,[\phi] := \phi_1(a-t)\mathcal{H}(a-t) + \mathcal{H}(t-a)$, $T_2(t)\,[\phi] := \phi_2(a-t)\mathcal{H}(a-t)$, and $T_3[\phi]$ is a combination of the nilpotent right translation and bounded multiplication C_0-semigroup of operators given by $T_3(t)\,[\phi] := e^{-\int_{a-t}^{a}\sigma(\tau)d\tau}\phi_3(a-t)\mathcal{H}(a-t)$, where $\mathcal{H}(s)$ is a Heaviside function (see Pazy, 1983, pg. 4; Engel & Nagel, 2000, Ch. I).

Lemma 5.2.3. *If $1 \leq p \leq \infty$ and if $\{f_n\}$ is a Cauchy sequence in L^p, with limit f, then $\{f_n\}$ has a subsequence that converges pointwise almost everywhere to f.*

Proof. See any standard and/or advanced analysis textbook (e.g., Rudin, 1966). □

Lemma 5.2.4. *The state space Γ is closed and convex, and it is invariant with respect to the semi-flow defined by $\{T_{\mathcal{A}}(t)\}_{t\geq 0}$.*

Proof. 1. To show that Γ is closed, it suffices to show that every sequence in Γ that converges in $\mathcal{X}$ also converges in Γ. Suppose $\phi_n \in \Gamma$, $\phi \in \mathcal{X}$ and $\phi_n \to \phi$, then from Lemma 5.2.3 there exists n_k such that for each component $\phi^{(j)}_{i,n_k} \to \phi^{(j)}_i$, $j = 1,2,3$, $i = 1,2,\ldots,\mathcal{N}$ pointwise almost everywhere. If $0 \leq \phi^{(1)}_{i,n} + \phi^{(2)}_{i,n} + \phi^{(3)}_{i,n} \leq 1$, then taking any subsequence $n_k \in \mathbb{N}$ as $k \to \infty$, this converges to $0 \leq \phi^{(1)}_i + \phi^{(2)}_i + \phi^{(3)}_i \leq 1$. Hence $\phi \in \Gamma$. Thus Γ is closed.
2. To show that Γ is convex, it suffices to show that, for all $x_1, x_2 \in \Gamma$ and $\varepsilon \in [0,1]$, $\varepsilon x_1 + (1-\varepsilon)x_2 \in \Gamma$. Let $x_1 = (x_{11}, x_{12}, x_{13})^T \in \Gamma$ and $x_2 = (x_{21}, x_{22}, x_{23})^T \in \Gamma$ such that $0 \leq x_{11} + x_{12} + x_{13} \leq 1$ and $0 \leq x_{21} + x_{22} + x_{23} \leq 1$, respectively. If $x_{ij},\ i =$

$1, 2,\ j = 1, 2, 3$ are non-negative, then $\varepsilon x_{1j} + (1-\varepsilon)x_{2j},\ j = 1, 2, 3$ are non-negative. Hence, we observe that $\varepsilon x_1 + (1-\varepsilon)x_2 = (\varepsilon x_{11} + (1-\varepsilon)x_{21}, \ldots, \varepsilon x_{13} + (1-\varepsilon)x_{23})^T$. The sum of the elements can be re-arranged as $[\varepsilon x_{11} + (1-\varepsilon)x_{21}] + [\varepsilon x_{12} + (1-\varepsilon)x_{22}] + [\varepsilon x_{13} + (1-\varepsilon)x_{23}] = \varepsilon[x_{11} + x_{12} + x_{13}] + (1-\varepsilon)[x_{21} + x_{22} + x_{23}] \leq \varepsilon + 1 - \varepsilon = 1$. Thus $\varepsilon x_1 + (1-\varepsilon)x_2 \in \Gamma$. Therefore, Γ is convex.
3. To show Γ is invariant under the semi-flow, we have, from Lemma 5.2.2,

$$T_1(t)\,[\phi] := \begin{cases} \phi_1(a-t), & a > t \\ 1, & a \leq t \end{cases} \tag{5.6}$$

$$T_2(t)\,[\phi] := \begin{cases} \phi_2(a-t), & a > t \\ 0, & a \leq t \end{cases} \tag{5.7}$$

$$T_3(t)\,[\phi] := \begin{cases} e^{-\int_{a-t}^{a}\sigma(\tau)d\tau}\phi_3(a-t), & a > t \\ 0, & a \leq t \end{cases}. \tag{5.8}$$

Hence, $T_{\mathcal{A}}(t)[\Gamma] \subset \Gamma$. This completes the proof. □

Before we state the next theorems, we state the following definitions and inequalities:

Definition 5.2.5. *Let $\mathcal{Y}$ and $\tilde{\mathcal{Y}}$ be Banach spaces. The Fréchet derivative of an operator $\mathcal{O} : \mathcal{Y} \to \tilde{\mathcal{Y}}$ is the bounded linear operator $\mathcal{D}_a\mathcal{O} : \mathcal{Y} \to \tilde{\mathcal{Y}}$ such that*

$$\lim_{||h||\to 0} \frac{||\mathcal{O}(a+h) - \mathcal{O}(a) - \mathcal{D}_a\mathcal{O}(a)h||}{||h||} = 0.$$

Definition 5.2.6. *Let $\omega_1 = ||\Omega||_{L^1} = \max_{1\leq j\leq \mathcal{N}} \sum_{i=1}^{\mathcal{N}} |\Omega_{ij}|$ and $\omega_\infty = ||\Omega||_{L^\infty} = \max_{1\leq i\leq \mathcal{N}} \sum_{j=1}^{\mathcal{N}} |\Omega_{ij}|$; and Ω_{ij} is an element of Ω. Because Ω is symmetric, $\omega = \omega_1 = \omega_\infty$. Define $\Xi_i^{(2)}[\phi](a) := \sum_{j\neq i} \Omega_{ij}\lambda_j[\phi](a)$.*

Inequality 5.2.7. *Suppose Assumption 5.2.1 holds. Then the following inequalities hold and $\lambda_i : \mathcal{X} \to L^1([0,\infty))$, where*

$$\begin{aligned}
\lambda_i[\phi](a) &= \int_0^\infty \beta(a,s)\psi(s)\phi_i^{(3)}(s)ds \leq \overline{\beta}\int_0^\infty \psi(s)\phi_i^{(3)}(s)ds \leq \overline{\beta\psi}||\phi||, \\
\lambda_i[\phi](a) &= \int_0^\infty \beta(a,s)\psi(s)\phi_i^{(3)}(s)ds \leq \overline{\beta}\int_0^\infty \psi(s)ds = \overline{\beta},\ \ \phi \in \Gamma, \\
||\lambda_i[\phi]|| &= \int_0^\infty \left|\int_0^\infty \beta(a,s)\psi(s)\phi_i^{(3)}(s)ds\right| da \leq \int_0^\infty \left(\int_0^\infty \beta(a,s)da\right)\psi(s)\left|\phi_i^{(3)}(s)\right| ds \\
&\leq ||\tilde{\beta}||\int_0^\infty \psi(s)\left|\phi_i^{(3)}(s)\right| ds \leq ||\tilde{\beta}||\overline{\psi}||\phi||.
\end{aligned}$$

From the above inequalities, it follows that

$$\begin{aligned}
\Xi_i^{(2)}[\phi](a) &= \sum_{j\neq i}\Omega_{ij}\lambda_j[\phi](a) \leq \sum_{j\neq i}\Omega_{ij}\overline{\beta} \leq (\omega-1)\,\overline{\beta},\ \ \phi\in\Gamma,\\
\Xi_i^{(2)}[\phi](a) &= \sum_{j\neq i}\Omega_{ij}\lambda_j[\phi](a) \leq (\omega-1)\,\overline{\beta\psi}||\phi||,\\
||\Xi_i^{(2)}[\phi]|| &\leq (\omega-1)\,||\tilde{\beta}||\overline{\psi}||\phi||.
\end{aligned}$$

Lemma 5.2.8. *The force of infection operator* $\lambda_i[\phi],\ i=1,2,\ldots,\mathcal{N}$ *defined by*

$$\lambda_i[\phi](a) = \int_0^\infty \beta(a,s)\psi(s)\phi_i^{(3)}(t,s)ds$$

is continuously Fréchet differentiable on $\mathcal{X}$*.*

This is true because $\lambda_i[\phi]$ is a linear operator, and the Fréchet derivative of a bounded linear operator is itself (i.e., $\mathcal{D}\lambda_i[h]=\lambda_i[h]$). We now turn to a property of $\mathcal{F}$ in the following lemma:

Lemma 5.2.9. *The operator* $\mathcal{F}$ *is continuously Fréchet differentiable on* $\mathcal{X}$*.*

Proof. We calculate $\mathcal{F}[\phi+h]-\mathcal{F}[\phi]$:

$$\mathcal{F}[\phi+h]-\mathcal{F}[\phi] := \begin{pmatrix} \mathcal{F}_1[\phi+h]-\mathcal{F}_1[\phi]\\ \mathcal{F}_2[\phi+h]-\mathcal{F}_2[\phi]\\ \mathcal{F}_3[\phi+h]-\mathcal{F}_3[\phi]\end{pmatrix}.$$

From the definition of norm in $\mathcal{X}$, it is sufficient to show that each component is continuously Fréchet differentiable. Let $\Xi_i^{(1)}[\phi](a) := \Xi_i^{(2)}[\phi](a)+\lambda_i[\phi](a)$. Then,

$$\begin{aligned}
\mathcal{F}_i^{(1)}[\phi+h]-\mathcal{F}_i^{(1)}[\phi] &= -\left(\phi_i^{(1)}+h_i^{(1)}\right)\Xi_i^{(1)}[\phi+h]+\phi_i^{(1)}\Xi_i^{(1)}[\phi]\\
&= \mathcal{L}_i^{(1)}[\phi,h]+\mathcal{R}_i^{(1)}[\phi,h], \qquad (5.9)
\end{aligned}$$

where $\mathcal{L}_i^{(1)}[\phi,h] := -\phi_i^{(1)}\Xi_i^{(1)}[h]-h_i^{(1)}\Xi_i^{(1)}[\phi]$ and $\mathcal{R}_i^{(1)}[\phi,h] := -h_i^{(1)}\Xi_i^{(1)}[h]$. Rearranging Eq. (5.9), taking the norm, dividing by $||h||$ and taking the limit as $||h||\to 0$, yields

$$\lim_{||h||\to 0}\frac{||\mathcal{F}_i^{(1)}[\phi+h]-\mathcal{F}_i^{(1)}[\phi]-\mathcal{L}_i^{(1)}[\phi,h]||}{||h||} = \lim_{||h||\to 0}\frac{||\mathcal{R}_i^{(1)}[\phi,h]||}{||h||}.$$

By Definition 5.2.5, it suffices to show that $\lim_{||h||\to 0}\dfrac{||\mathcal{R}_i^{(1)}[\phi,h]||}{||h||}=0$. From Definition 5.2.6 and Inequality 5.2.7, we observe that

$$\begin{aligned}
||\mathcal{R}_i^{(1)}[\phi,h]|| &= ||h_i^{(1)}\Xi_i^{(1)}[h]||\\
&\leq \omega||\tilde{\beta}||\overline{\psi}||h||\int_0^\infty \left|h_i^{(1)}(a)\right| da \leq \omega||\tilde{\beta}||\overline{\psi}||h||^2 = \mathcal{M}||h||^2,
\end{aligned}$$

where $\mathcal{M} := \omega||\tilde{\beta}||\overline{\psi}$. Then, $\lim_{||h||\to 0} \frac{||\mathcal{R}_i^{(1)}[\phi,h]||}{||h||} \leq \lim_{||h||\to 0} \mathcal{M}||h|| = 0$. Clearly, $\mathcal{L}_i^{(1)}[\phi,h]$ is continuous. Hence, $\mathcal{F}_i^{(1)}$ is continuously Fréchet differentiable with Fréchet derivative $\mathcal{L}_i^{(1)}[\phi,h]$. Similarly, using the same argument, it can be shown that $\mathcal{F}_i^{(2)}[\phi]$ and $\mathcal{F}_i^{(3)}[\phi]$ are also continuously Fréchet differentiable. Hence, the operator $\mathcal{F}$ is continuously Fréchet differentiable. This completes the proof. □

Corollary 5.2.10. *Suppose Assumption 5.2.1 holds. Then the map $\mathcal{F}|_\Gamma\colon \Gamma \to \mathcal{X}$ is Lipschitz continuous and there exists a number $K > 0$ such that $(Id|_{\mathcal{X}} + K\mathcal{F})(\Gamma) \subset \Gamma$.*

Proof. From the inequalities in 5.2.6–5.2.7, we can show the Lipschitz continuity of $\mathcal{F}(\phi)$. Note that $||\mathcal{F}[\phi] - \mathcal{F}[\eta]|| = \sum_j^3 \sum_k^{\mathcal{N}} ||\mathcal{F}_k^{(j)}[\phi] - \mathcal{F}_k^{(j)}[\eta]||$. We will use a component-wise approach. That is, using inequalities. 5.2.6–5.2.7,

$$\begin{aligned} ||F_i^{(1)}(\phi) - F_i^{(1)}(\eta)|| &= \int_0^\infty \left| -\phi_i^{(1)}(a)\Xi_i^{(1)}[\phi](a) + \eta_i^{(1)}(a)\Xi_{(i}^{(1)}[\eta](a) \right| da \\ &\leq \mathcal{J}_1^{(1)} + \mathcal{J}_2^{(1)}, \end{aligned}$$

where,

$$\begin{aligned} \mathcal{J}_1^{(1)} &= \int_0^\infty \Xi_i^{(1)}[\phi](a) \left| \phi_i^{(1)}(a) - \eta_i^{(1)}(a) \right| da \leq \omega\tilde{\beta}||\phi - \eta||, \\ \mathcal{J}_2^{(1)} &= \int_0^\infty \eta_i^{(1)}(a) \left| \Xi_i^{(1)}[\phi](a) - \Xi_i^{(1)}[\eta](a) \right| da \leq \omega||\tilde{\beta}||\overline{\psi}||\phi - \eta||. \end{aligned}$$

Hence, using the triangle inequality, for $\phi,\ \eta \in \Gamma$, we obtain

$$||F_i^{(1)}(\phi) - F_i^{(1)}(\eta)|| \leq \omega \left(\overline{\beta} + ||\tilde{\beta}||\overline{\psi} \right) ||\phi - \eta||.$$

Similarly, it can be shown that,

$$\begin{aligned} ||F_i^{(2)}(\phi) - F_i^{(2)}(\eta)|| &\leq \omega \left(\overline{\beta} + ||\tilde{\beta}||\overline{\psi} \right) ||\phi - \eta|| \\ ||F_i^{(3)}(\phi) - F_i^{(3)}(\eta)|| &\leq 2 \left(\overline{\beta} + ||\tilde{\beta}||\overline{\psi} \right) ||\phi - \eta||. \end{aligned}$$

Therefore, from the definition of norm in $\mathcal{X}$, $||\mathcal{F}[\phi] - \mathcal{F}[\eta]|| \leq L||\phi - \eta||$, where $L = 2\mathcal{N}(1+\omega)\left(\overline{\beta} + ||\tilde{\beta}||\overline{\psi} \right)$. Hence, $\mathcal{F}[\phi]$ is Lipschitz continuous.

For the second part of the lemma, we let $\phi = (\phi_1, \phi_2, \phi_3)^T \in \mathcal{X}_+$ such that $\phi_i^{(1)} \geq 0$, $\phi_i^{(2)} \geq 0$, $\phi_i^{(3)} \geq 0$, and $\phi_i^{(1)} + \phi_i^{(2)} + \phi_i^{(3)} \leq 1$. In addition, we define $\tilde{\phi} \in \mathcal{X}_+$ such that $\tilde{\phi}_i := \left(\tilde{\phi}_i^{(1)}, \tilde{\phi}_i^{(2)}, \tilde{\phi}_i^{(3)} \right)^T = (Id + K\mathcal{F})(\phi)$. Then, we have

$$\begin{aligned} \tilde{\phi}_i^{(1)} + \tilde{\phi}_i^{(2)} + \tilde{\phi}_i^{(3)} &= \phi_i^{(1)} + \phi_i^{(2)} + \phi_i^{(3)} + K \left(\mathcal{F}_i^{(1)}(\phi) + \mathcal{F}_i^{(2)}(\phi) + \mathcal{F}_i^{(3)}(\phi) \right) \\ &= \phi_i^{(1)} + \phi_i^{(2)} + \phi_i^{(3)} - K \left(\gamma\lambda_i[\phi]\phi_i^{(2)} \right) \\ &\leq \phi_i^{(1)} + \phi_i^{(2)} + \phi_i^{(3)} \leq 1 \end{aligned}$$

by definition of the set Γ. It now suffices to show that $\phi_i^{(j)} + K\mathcal{F}_i^{(j)}(\phi) \geq 0,\ j = 1,2,3,\ i = 1,2,\ldots,\mathcal{N}$. We observe that

$$\begin{aligned}
\phi_i^{(1)} + K\mathcal{F}_i^{(1)}(\phi) &= \phi_i^{(1)}\left(1 - K\sum_j \Omega_{ij}\lambda_j[\phi]\right) \geq 0,\\
\phi_i^{(2)} + K\mathcal{F}_i^{(2)}(\phi) &= \phi_i^{(2)}\left(1 - K\lambda_i[\phi]\right) + K\phi_i^{(1)}\sum_{j\neq i}\Omega_{ij}\lambda_j[\phi] \geq 0,
\end{aligned}$$

provided that $K \leq \dfrac{1}{\omega\overline{\beta}}$, and $K \leq \dfrac{1}{\overline{\beta}}$, respectively. Furthermore, we note that $\phi_i^{(3)} + K\mathcal{F}_i^{(3)}(\phi) = \phi_i^{(3)} + K\left(\phi_i^{(1)} + (1-\gamma)\,\phi_i^{(2)}\right)\lambda_i[\phi] \geq 0$ because $\phi_i^{(3)} \geq 0$ and $\phi_i^{(1)} + (1-\gamma)\,\phi_i^{(2)} \geq 0$ for $\gamma \in [0,1]$ almost everywhere on $a \in [0,\infty)$. Hence, if we choose $K > 0$ such that $K \leq \min\left(\dfrac{1}{\omega\overline{\beta}}, \dfrac{1}{\overline{\beta}}\right)$, then $(Id|_{\mathcal{X}} + K\mathcal{F})(\Gamma) \subset \Gamma$. This completes the proof. □

Theorem 5.2.11. *For each $u^0 \in \mathcal{X}_+$, there is a unique continuous mild solution $u(t,u^o) \in \mathcal{X}_+,\ t \in [0,\infty)$ for Eq. (5.5) such that*

$$u(t) = e^{t\mathcal{A}}u^0 + \int_0^t e^{\mathcal{A}(t-s)}\mathcal{F}(u)ds.$$

In addition, if $u^0 \in \mathcal{D}(\mathcal{A})$, then the solution is a classical one.

Proof. From Lemma 5.2.9 and Corollary 5.2.10, we note that, given a choice of K, Eq. (5.5) can be reformulated as follows:

$$\frac{du}{dt} = \left(\mathcal{A} - K^{-1}\right)u + K^{-1}\left(Id + K\mathcal{F}\right)(u),$$

where K is chosen such that $Id + K\mathcal{F} : \Gamma \to \Gamma$. Using variation of constants, the solution satisfies:

$$u(t) = e^{t\left(\mathcal{A}-K^{-1}\right)}u^0 + K^{-1}\int_0^t e^{\left(\mathcal{A}-K^{-1}\right)(t-s)}\left[u + K\mathcal{F}(u)\right](s)ds.$$

We observe that this can be rewritten as

$$u(t) = e^{-K^{-1}t}T_{\mathcal{A}}(t)u^0 + \left(1 - e^{-K^{-1}t}\right)f(u),$$

where $f(u) = \dfrac{K^{-1}}{1-e^{-K^{-1}t}}\int_0^t e^{-K^{-1}(t-s)}T_{\mathcal{A}}(t-s)\left[u + K\mathcal{F}(u)\right](s)ds$. Our aim is to show that there is a mild solution. That is, we define $\{u^n\}_{n\in\mathbb{N}}$ by

$$u^0(t) = u^0, \tag{5.10}$$

$$u^{n+1}(t) = e^{-K^{-1}t}T_{\mathcal{A}}(t)u^0 + \left(1 - e^{-K^{-1}t}\right)f(u^n), \tag{5.11}$$

$$f(u^n) = \frac{K^{-1}}{1-e^{-K^{-1}t}}\int_0^t e^{-K^{-1}(t-s)}T_{\mathcal{A}}(t-s)\left[u^n + K\mathcal{F}(u^n)\right](s)ds. \tag{5.12}$$

By Lemma 5.2.4, $\mathcal{T}_{\mathcal{A}}(t)\Gamma \subset \Gamma$. So if $u \in \Gamma$, then by Corollary 5.2.10 $u + K\mathcal{F}(u) \in \Gamma$, for suitable K, and thus $\mathcal{T}_{\mathcal{A}}(t)\left[u + K\mathcal{F}(u)\right] \in \Gamma$. We now want to prove that $u^n \in \Gamma$ implies $u^{n+1} \in \Gamma$. Suppose $u^n \in \Gamma$. Then $f^{(j)}(u^n) \geq 0,\ j = 1, 2, 3$ and hence

$$0 \ \leq \ \sum_{j=1}^{3} f^{(j)}(u^n) \leq 1$$

because

$$\frac{K^{-1}}{1 - e^{-K^{-1}t}} \left(\int_0^t e^{-K^{-1}(t-s)} ds \right) \ = \ 1.$$

Hence, $f(u^n) \in \Gamma$. Therefore, $u^{n+1} \in \Gamma$ because the right-hand-side of Eq. (5.11) is a convex combination of elements in Γ. If u^n converges to u, then $u \in \Gamma$, as Γ is closed (from Lemma 5.2.4). From Corollary 5.2.10, u^n converges to the mild solution uniformly in Γ. Hence, there is a mild solution of the abstract Cauchy problem. This mild solution is a classical solution if $u^0 \in \mathcal{D}(\mathcal{A})$ as $\mathcal{F}$ is continuously differentiable (see Pazy, 1983, Ch. 6). This completes the proof. □

5.3 Equilibria: Steady-State Analysis

In this section, we consider the existence of equilibria and their properties.

5.3.1 Existence of Steady States

A steady state solution $s_i^*(a)$, $v_i^*(a)$ and $y_i^*(a)$ of Eqs. (5.1)–(5.3) exists if it satisfies the following time-independent system of differential equations:

$$\frac{ds_i^*(a)}{da} \ = \ -s_i^*(a) \sum_j \Omega_{ij} \lambda_j^*(a), \tag{5.13}$$

$$\frac{dv_i^*(a)}{da} \ = \ s_i^*(a) \sum_{j \neq i} \Omega_{ij} \lambda_j^*(a) - v_i^*(a) \lambda_i^*[y](a), \tag{5.14}$$

$$\frac{dy_i^*(a)}{da} \ = \ \left[s_i^*(a) + (1 - \gamma(a)) v_i^*(a) \right] \lambda_i^*[y](a) - \sigma(a) y_i^*(a), \tag{5.15}$$

with boundary conditions $s_i^*(0) = 1$ and $v_i^*(0) = y_i^*(0) = 0$, where $\lambda^* := (\lambda_1^*, \lambda_2^*, \ldots, \lambda_{\mathcal{N}}^*)^T \in \mathbb{Z}_+ := [L_+^1]^{\mathcal{N}}$. Furthermore,

$$\lambda_i^*\left[y\right](a) \ = \ \int_0^\infty \beta(a, s) \psi(s) y_i^*(s) ds,$$

By integrating Eqs. (5.13)–(5.15) with respect to a, we obtain the following:

$$s_i^*(a) \ = \ e^{-\int_0^a \Xi_i^{(1)}[\lambda^*](\tau) d\tau} \tag{5.16}$$

$$v_i^*(a) = \int_0^a e^{-\int_0^b \overline{\Xi}_i^{(1)}[\lambda^*](\tau)d\tau} \overline{\Xi}_i^{(2)}[\lambda^*](b) e^{-\int_b^a \lambda_i^*(\tau)d\tau} db \tag{5.17}$$

$$y_i^*(a) = \int_0^a e^{-\int_b^a \sigma(\tau)d\tau} \left[s_i^*(b) + (1-\gamma(b))v_i^*(b)\right] \lambda_i^*(b) db. \tag{5.18}$$

where $\overline{\Xi}_i^{(1)}[\lambda^*](a) := \sum_j \Omega_{ij}\lambda_j^*(a)$ and $\overline{\Xi}_i^{(2)}[\lambda^*](a) := \sum_{j\neq i} \Omega_{ij}\lambda_j^*(a)$. Substituting y_i^* into the equation for $\lambda_i^*[y](a)$, we obtain the following nonlinear integral equation:

$$\lambda_i^*\left[y^*\right](a) = \Phi_i[\lambda^*](a) = \Phi_i^{(1)}[\lambda^*](a) + \Phi_i^{(2)}[\lambda^*](a), \tag{5.19}$$

$$\Phi_i^{(1)}[\lambda^*](a) = \int_0^\infty \beta(a,s)\psi(s) \left[\int_0^s e^{-\int_b^s \sigma(\tau)d\tau} e^{-\int_0^b \overline{\Xi}_i^{(1)}[\lambda^*](\tau)d\tau} \lambda_i^*(b) db\right] ds, \tag{5.20}$$

$$\begin{aligned}\Phi_i^{(2)}[\lambda^*](a) = \int_0^\infty \beta(a,s)\psi(s) \Bigg(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau}(1-\gamma(b)) \Bigg[\int_0^b e^{-\int_0^{b^*} \overline{\Xi}_i^{(1)}[\lambda^*](\tau)d\tau} \\ \times \overline{\Xi}_i^{(2)}[\lambda^*](b^*) e^{-\int_{b^*}^b \lambda_i^*(\tau)d\tau} db^*\Bigg] \lambda_i^*(b) db\Bigg) ds.\end{aligned} \tag{5.21}$$

From Eq. (5.19), it is clear that one trivial solution is $\lambda_i^* = 0$ for all $i = 1, 2, \ldots, \mathcal{N}$. This solution corresponds to the disease-free equilibrium, where no disease exists. To investigate the existence of non-trivial positive solutions for Eq. (5.19), we let $\Phi_i[\lambda^*](a) := \Phi_i^{(1)}[\lambda^*](a) + \Phi_i^{(2)}[\lambda^*](a)$, with $i = 1, 2, \ldots, \mathcal{N}$ be the nonlinear integral operator on the Banach space $\mathcal{X}$. In addition, we let $\lambda^* := (\lambda_1^*, \lambda_2^*, \ldots, \lambda_{\mathcal{N}}^*)^T$ and $\Phi[\lambda](a) := (\Phi_1[\lambda](a), \Phi_2[\lambda](a), \ldots, \Phi_{\mathcal{N}}[\lambda](a))^T$. It follows that the solution of Eq. (5.19) corresponds to the fixed points of the nonlinear operator $\Phi[\lambda^*]$, where $\lambda_i^* \in L_+^1$ and $\lambda^* \in [L_+^1]^{\mathcal{N}}$. In the next section, we will provide some properties of $\Phi[\lambda^*]$.

5.3.2 Properties of the Steady State Operator

We investigate the continuity and compactness properties of the nonlinear steady-state operator $\Phi_i\left[\lambda^*\right](a) = \Phi_i^{(1)}\left[\lambda^*\right](a) + \Phi_i^{(2)}\left[\lambda^*\right](a)$. Before continuing, we first provide the following estimates that will be used in the proofs later.

Lemma 5.3.1. *The following estimates hold for $\lambda_i \in L_+^1(0, a_m)$:*

1. $\int_0^s e^{-\int_0^b \overline{\Xi}_i^{(1)}[\lambda](\tau)d\tau} \lambda_i(b) db \leq 1$
2. $\int_0^s e^{-\int_0^{b^*} \overline{\Xi}_i^{(1)}[\lambda](\tau)d\tau} \overline{\Xi}_i^{(2)}[\lambda](b^*) db^* \leq 1$

Proof. We observe that

$$\begin{aligned}\int_0^s e^{-\int_0^b \overline{\Xi}_i^{(1)}[\lambda](\tau)d\tau} \lambda_i(b) db &= \int_0^s e^{-\int_0^b \overline{\Xi}_i^{(2)}[\lambda](\tau)+\lambda_i(\tau)d\tau} \lambda_i(b) db \\ &\leq \int_0^s e^{-\int_0^b \lambda_i(\tau)d\tau} \lambda_i(b) db = \left(1 - e^{-\int_0^s \lambda_i(\tau)d\tau}\right) < 1.\end{aligned}$$

We carry out a similar calculation for the second estimate. That is,

$$\begin{aligned}\int_0^s e^{-\int_0^b \overline{\Xi}_i^{(1)}[\lambda](\tau)d\tau}\overline{\Xi}_i^{(2)}[\lambda](b)db &\leq \int_0^s e^{-\int_0^b \overline{\Xi}_i^{(2)}[\lambda](\tau)d\tau}\overline{\Xi}_i^{(2)}[\lambda](b)db \\ &= \left(1-e^{-\int_0^s \overline{\Xi}_i^{(2)}[\lambda](\tau)d\tau}\right)<1.\end{aligned}$$

This completes the proof. □

In addition to the assumptions made previously, we make the following assumptions:

Assumption 5.3.2. *Let a_m be the maximal age. We assume*

1. $\mathcal{H}_7$*: $\beta(a,s)\in L_+^\infty([0,a_m)\times[0,a_m))$ and $\lim_{h\to 0}\int_0^{a_m}|\beta(a+h,s)-\beta(a,s)|\,da=0$ uniformly for $s\in\mathbb{R}_+$, where β is extended by $\beta(a,s)=0$ for almost all $(a,s)\in(-\infty,0)\bigcup(a_m,\infty)$.*

2. $\mathcal{H}_8$*: There exist a^* with $a_m>a^*>0$ and $\varepsilon>0$ such that $\beta(a,s)\geq\varepsilon$ for almost all $(a,s)\in(0,a_m)\times(a_m-a^*,a_m)$.*

One must be very careful with the behaviour of the rates when $a_m=\infty$. However, the rates at $a_m=\infty$ can be disregarded since they do not add anything biologically important to the model. Hence, the Assumption 5.3.2 assumes that maximum age a_m is finite, which is biologically realistic. Alternatively, we can disregard this assumption provided that the time ranges are comparable with the lifespan of the population, or fertility rate and survival probability are constrained within a finite window. We now state the following properties for $\Phi[\lambda]$. It is easy to see that

Theorem 5.3.3. *Let $\mathcal{Z}_+:=\left(L_+^1([0,a_m))\right)^{\mathcal{N}}$. Then, $\Phi:\mathcal{Z}_+\to\mathcal{Z}_+$ is positive.*

Before establishing further properties of $\Phi(\lambda)$, we need the following inequality.

Inequality 5.3.4. *Recall $\overline{\Xi}_i^{(1)}[\lambda](a)=\overline{\Xi}_i^{(2)}[\lambda](a)+\lambda_i(a)$. Then, the following estimates hold for $\lambda_i\in L_+^1(0,a_m)$:*

$$\begin{aligned}\Xi_i^{(12)}[\lambda](b) &= \int_0^b e^{-\int_0^{b^*}\overline{\Xi}_i^{(2)}[\lambda](\tau)d\tau}\overline{\Xi}_i^{(2)}[\lambda](b^*)e^{-\int_0^b\lambda_i(\tau)d\tau}db^* \\ &= \left(1-e^{-\int_0^b\overline{\Xi}_i^{(2)}[\lambda](\tau)d\tau}\right)e^{-\int_0^b\lambda_i(\tau)d\tau}\leq 1,\\ \int_0^s\eta_i(b)\left|\Xi_i^{(12)}[\lambda](b)-\Xi_i^{(12)}[\eta](b)\right|db &= \int_0^s\eta_i(b)\left|\left(1-e^{-\int_0^b\overline{\Xi}_i^{(2)}[\lambda](\tau)d\tau}\right)e^{-\int_0^b\lambda_i(\tau)d\tau}\right. \\ &\quad\left.-\left(1-e^{-\int_0^b\overline{\Xi}_i^{(2)}[\eta](\tau)d\tau}\right)e^{-\int_0^b\eta_i(\tau)d\tau}\right|db, \\ &\leq ||\eta||||\lambda-\eta||+||\eta||(\omega||\lambda-\eta||)\,, \\ \int_0^s\left|\Xi_i^{(12)}[\lambda](b)\lambda_i(b)-\Xi_i^{(12)}[\eta](b)\eta_i(b)\right|db &\leq \int_0^s\left|\Xi_i^{(12)}[\lambda](b)\lambda_i(b)-\Xi_i^{(12)}[\lambda](b)\eta_i(b)\right|db \\ &+ \int_0^s\left|\Xi_i^{(12)}[\lambda](b)\eta_i(b)-\Xi_i^{(12)}[\eta](b)\eta_i(b)\right|db \\ &\leq ||\lambda-\eta||(1+||\eta||+\omega||\eta||)\,.\end{aligned}$$

Theorem 5.3.5. *Let* $\mathcal{Z}_+ := \left(L^1_+([0,a_m))\right)^{\mathcal{N}}$. *The operator* $\Phi : \mathcal{Z}_+ \to \mathcal{Z}_+$ *is locally Lipschitz continuous and* $\Phi(\mathcal{Z}_+)$ *is bounded.*

Proof. We have $||\Phi_i[\lambda] - \Phi_i[\eta]|| \leq ||\Phi_i^{(1)}[\lambda] - \Phi_i^{(1)}[\eta]|| + ||\Phi_i^{(2)}[\lambda] - \Phi_i^{(2)}[\eta]||$, where

$$\begin{aligned}
||\Phi_i^{(1)}[\lambda] - \Phi_i^{(1)}[\eta]|| &= \int_0^{a_m} \left[\int_0^{a_m} \beta(a,s)\psi(s) \left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau} \left| \lambda_i(b) e^{-\int_0^b \overline{\Xi}_i^{(1)}[\lambda](\tau)d\tau} \right.\right.\right. \\
&\quad \left.\left.\left. - \eta_i(b) e^{-\int_0^b \overline{\Xi}_i^{(i)}[\eta](\tau)d\tau} \right| db \right) ds \right] da \\
&\leq \int_0^{a_m} \left[\int_0^{a_m} \beta(a,s)\psi(s) \left(\int_0^s \left| \lambda_i(b) e^{-\int_0^b \overline{\Xi}_i^{(1)}[\lambda](\tau)d\tau} \right.\right.\right. \\
&\quad \left.\left.\left. - \eta_i(b) e^{-\int_0^b \overline{\Xi}_i^{(i)}[\eta](\tau)d\tau} \right| db \right) ds \right] da.
\end{aligned}$$

We observe that

$$\begin{aligned}
&\int_0^s \left| \lambda_i(b) e^{-\int_0^b \overline{\Xi}_i^{(1)}[\lambda](\tau)d\tau} - \eta_i(b) e^{-\int_0^b \overline{\Xi}_i^{(1)}[\eta](\tau)d\tau} \right| db \\
\leq &\int_0^s |\lambda_i - \eta_i| \, e^{-\int_0^b \overline{\Xi}_i^{(1)}[\lambda](\tau)d\tau} db + \int_0^s \eta_i \left| e^{-\int_0^b \overline{\Xi}_i^{(1)}[\lambda](\tau)d\tau} - e^{-\int_0^b \overline{\Xi}_i^{(1)}[\eta](\tau)d\tau} \right| db \\
\leq &||\lambda_i - \eta_i|| + \int_0^s \eta_i(b) \int_0^s \left| \overline{\Xi}_i^{(1)}[\lambda](\tau) - \overline{\Xi}_i^{(1)}[\eta](\tau) \right| d\tau db \\
\leq &||\lambda - \eta|| (1 + \omega||\eta||) \, .
\end{aligned}$$

Therefore,

$$\begin{aligned}
||\Phi_i^{(1)}[\lambda] - \Phi_i^{(1)}[\eta]|| &\leq ||\lambda - \eta|| (1 + \omega||\eta||) \int_0^{a_m} \left[\int_0^{a_m} \beta(a,s)\psi(s) ds \right] da \\
&\leq ||\lambda - \eta|| (1 + \omega||\eta||) \int_0^{a_m} \left[\int_0^{a_m} \tilde{\beta}(a)\psi(s) ds \right] da \\
&\leq ||\lambda - \eta|| (1 + \omega||\eta||) \, ||\tilde{\beta}||_{L^1}.
\end{aligned}$$

Similarly, we have $||\Phi_i^{(2)}(\lambda) - \Phi_i^{(2)}(\eta)|| \leq ||\tilde{\beta}||||\lambda - \eta||(1 + ||\eta|| + \omega||\eta||)$. Therefore, $||\Phi_i(\lambda) - \Phi_i(\eta)|| \leq ||\tilde{\beta}||(2 + 2\omega||\eta|| + ||\eta||) \times ||\lambda - \eta||$. Hence, $||\Phi(\lambda) - \Phi(\eta)|| \leq \mathcal{N}||\tilde{\beta}|| \times (2 + 2\omega||\eta|| + ||\eta||) \, ||\lambda - \eta||$, and Φ is Lipschitz (local) continuous.

In addition, it follows that $||\Phi_i(\lambda)|| \leq ||\Phi_i^{(1)}(\lambda)|| + ||\Phi_i^{(2)}(\lambda)||$, where by Lemma 5.3.1,

$$\begin{aligned}
||\Phi_i^{(1)}(\lambda)|| &= \int_0^{a_m} \left| \left[\int_0^{a_m} \beta(a,s)\psi(s) \left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau} \lambda_i(b) e^{-\int_0^b \overline{\Xi}_i^{(1)}[\lambda](\tau)d\tau} db \right) ds \right] \right| da \\
&\leq \int_0^{a_m} \left| \left[\int_0^{a_m} \beta(a,s)\psi(s) \left(\int_0^s \lambda_i(b) e^{-\int_0^b \overline{\Xi}_i^{(1)}[\lambda](\tau)d\tau} db \right) ds \right] \right| da \\
&\leq \int_0^{a_m} \left| \left[\int_0^{a_m} \beta(a,s)\psi(s) ds \right] \right| da \leq \int_0^{a_m} \left| \tilde{\beta}(a) \right| \left[\int_0^{a_m} \psi(s) ds \right] da \leq ||\tilde{\beta}||.
\end{aligned}$$

Similarly, it can be shown that $||\Phi_i^{(2)}(\lambda)|| \leq ||\tilde{\beta}||$. Therefore, $||\Phi|| \leq 2||\tilde{\beta}||\mathcal{N} := \mathcal{K}$. Hence $\Phi(\mathcal{Z}_+)$ is bounded. This completes the proof. □

Definition 5.3.6. *Let $\mathcal{Y}$, $\overline{\mathcal{Y}}$ be Banach spaces and $\mathcal{L} : \mathcal{Y} \to \overline{\mathcal{Y}}$.*

1. *The operator $\mathcal{L} : \mathcal{Y} \to \overline{\mathcal{Y}}$ is said to be* **bounded** *if it maps any bounded subset of $\mathcal{Y}$ into a bounded subset of $\overline{\mathcal{Y}}$.*

2. *An operator $\mathcal{L} : \mathcal{Y} \to \overline{\mathcal{Y}}$ is* **compact** *if it maps bounded subsets of $\mathcal{Y}$ into relatively compact (pre-compact) subsets of $\overline{\mathcal{Y}}$.*

3. *The operator $\mathcal{L} : \mathcal{Y} \to \overline{\mathcal{Y}}$ is said to be* **completely continuous** *if it is continuous and compact.*

4. *The operator $\mathcal{L}$ is said to be of* **finite rank** *if $\mathcal{L}(\mathcal{Y})$ lies in a finite-dimensional subspace of $\overline{\mathcal{Y}}$.*

We note that an operator $\mathcal{L}$ is compact if and only if for every bounded sequence $\{u_k\}_{k\in\mathbb{N}_+} \in \mathcal{Y}$ the sequence $\{\mathcal{L}(u_k)\}_{k\in\mathbb{N}_+}$ has a convergent subsequence. We recall the Riesz-Frechet-Kolmogorov theorem:

Theorem 5.3.7. (Riesz-Frechet-Kolmogorov Compactness) *Let $F \subset L^p(\mathbb{R}^n)$ be bounded, $1 \leq p < \infty$, and suppose $\lim_{|h|\to 0}||\rho_h[f] - f||_p = 0$ uniformly for $f \in F$, where $\rho_h[f](x) = f(x+h)$, $h \in \mathbb{R}^n$ (i.e., for every $\varepsilon > 0$, there exists $\delta > 0$ such that $||\rho_h[f] - f||_p < \varepsilon$, for all $h \in \mathbb{R}^n$ with $|h| < \delta, f \in F$). Then, $F|_\Omega$ has compact closure in $L^p(\Omega)$, for each $\Omega \subset \mathbb{R}^n$, Ω measurable and bounded (see Brezis, 2010, pg. 111, Theorem 4.26).*

Theorem 5.3.8. *The operator $\Phi : \mathcal{Z} \to \mathcal{Z}$ is completely continuous.*

Proof. Let $\mathcal{S}$ be a bounded subset of $L^1(0, a_m)$ and $F_i := \Phi_i(\mathcal{S}) \subset L^1(0, a_m)$ for each $i = 1, 2, \ldots, \mathcal{N}$. We want to prove that $\Phi(\mathcal{S})$ has a compact closure. We will use Theorem 5.3.7. Then, if $\lambda_i \in \mathcal{S}$, we have $||\lambda_i|| \leq c_0$ for some positive constant c_0. In addition, with the result of Inequality 5.2.7, we have $||\Phi_i[\lambda]|| \leq ||\Phi_i^{(1)}[\lambda]|| + ||\Phi_i^2[\lambda]||$, where

$$\begin{aligned}
||\Phi_i^{(1)}[\lambda]|| &= \int_0^{a_m} \left[\int_0^{a_m} |\beta(a,s)|\, \psi(s) \left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau} e^{\int_0^b \left|\Xi_i^{(1)}[\lambda](\tau)\right| d\tau} |\lambda_i(b)|\, db \right) ds \right] da \\
&\leq e^{\omega c_o} \int_0^{a_m} \left|\tilde{\beta}(a)\right| \left[\int_0^{a_m} \psi(s) \left(\int_0^s |\lambda_i(b)|\, db \right) ds \right] da \\
&\leq e^{\omega c_o} ||\tilde{\beta}|| ||\lambda_i|| \leq e^{\omega c_o} ||\tilde{\beta}|| c_0.
\end{aligned}$$

Using similar argument, we observe that $||\Phi_i^{(2)}[\lambda]|| \leq e^{\omega c_o}||\tilde{\beta}||(\omega - 1)c_0^2$. Therefore, $||\Phi_i[\lambda]|| \leq e^{\omega c_o}||\tilde{\beta}||c_0 + e^{\omega c_o}||\tilde{\beta}||(\omega-1)c_0^2$, and $||\Phi[\lambda]|| \leq \mathcal{N} e^{\omega c_o}||\tilde{\beta}||c_0(1 + (\omega - 1)c_0)$. Hence, Φ and Φ_i are bounded and so is F_i; the first condition of Theorem 5.3.7 holds. It now suffices to show that $||\rho_h[\Phi[\lambda]](a) - \Phi[\lambda](a)|| \to 0$ uniformly. We observe that,

using Lemma 5.3.1 and the consequence of Assumption 5.3.2 (i.e. for every $\varepsilon > 0$, there exists δ such that $\int_0^{a_m} |\beta(a+h,s) - \beta(a,s)|\, da \leq \varepsilon$, for all $s \in \mathbb{R}_+$ and $|h| \leq \delta$), we have

$$||\rho_h[\Phi_i[\lambda]](a) - \Phi_i[\lambda](a)|| \leq ||\rho_h[\Phi_i^{(1)}[\lambda]](a) - \Phi_i^{(1)}[\lambda](a)|| + ||\rho_h[\Phi_i^{(2)}[\lambda]](a) - \Phi_i^2[\lambda](a)||,$$

where we calculate $||\rho_h[\Phi_i^{(1)}[\lambda]](a) - \Phi_i^{(1)}[\lambda](a)||$ as follows:

$$\begin{aligned}
||\rho_h[\Phi_i^{(1)}[\lambda]](a) - \Phi_i^{(1)}[\lambda](a)|| &\leq \int_0^{a_m} \left[\int_0^{a_m} |\beta(a+h,s) - \beta(a,s)|\, \psi(s) \right. \\
&\quad \left. \times \left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau} e^{\int_0^b \left| \Xi_i^{(1)}[\lambda](\tau) \right| d\tau} |\lambda_i(b)|\, db \right) ds \right] da \\
&\leq e^{\omega c_o} ||\lambda|| \int_0^{a_m} \left[\int_0^{a_m} |\beta(a+h,s) - \beta(a,s)|\, \psi(s) ds \right] da \\
&\leq e^{\omega c_o} c_0 \varepsilon.
\end{aligned}$$

Using similar approach for $\Phi_i^{(2)}$, $||\rho_h \left[\Phi_i^{(2)}[\lambda]\right](a) - \Phi_i^{(2)}[\lambda](a)|| \leq \varepsilon e^{\omega c_o}(\omega - 1)c_0^2$. Then, $||\rho_h[\Phi[\lambda]](a) - \Phi[\lambda](a)|| \leq \mathcal{N}\varepsilon e^{\omega c_o} c_0(1 + (\omega - 1)c_0)$. Therefore, the second condition of Theorem 5.3.7 is satisfied. Hence, $F_i = \{\Phi_i(u) : u \in \mathcal{S}\}$ is precompact for each $i = 1, 2, \ldots, \mathcal{N}$ in $L^1(0, a_m)$. Let $F = \{\Phi(u) : u \in \mathcal{S}\}$, then $F \subset F_1 \times F_2 \times \ldots \times F_{\mathcal{N}}$ implies that $\overline{F} \subset \overline{F_1 \times F_2 \times \ldots \times F_{\mathcal{N}}} = \overline{F}_1 \times \overline{F}_2 \times \ldots \times \overline{F}_{\mathcal{N}}$, where $\overline{F}$ denotes the closure of F. Because $\overline{F}_i$ is compact, $\overline{F}_1 \times \overline{F}_2 \times \ldots \times \overline{F}_{\mathcal{N}}$ is compact, so $\overline{F}$ is compact. Hence, Φ is completely continuous. This completes the proof. □

Theorem 5.3.9. *The operator Φ is Fréchet differentiable at $\lambda = 0$. In addition, let $\mathcal{T} : [L_+^1([0, a_m))]^{\mathcal{N}} \to [L_+^1([0, a_m))]^{\mathcal{N}}$ and $\mathcal{T} = \mathcal{L}_\lambda \Phi(0)$ be the Fréchet derivative of Φ at $\lambda = 0$, then the next-generation operator is given by,*

$$\mathcal{T}[\lambda](a) := \begin{pmatrix} \mathcal{T}_1[\lambda](a) \\ \mathcal{T}_2[\lambda](a) \\ \ldots \\ \mathcal{T}_{\mathcal{N}}[\lambda](a) \end{pmatrix}, \tag{5.22}$$

$$\mathcal{T}_i[\lambda](a) = \int_0^{a_m} \Psi(a,b)\lambda_i(b)db,\ i = 1, 2, \ldots, \mathcal{N}, \tag{5.23}$$

$$\Psi(a,b) = \int_b^{a_m} \beta(a,s)\psi(s)\Theta(b,s)ds, \tag{5.24}$$

$$\Theta(b,s) = e^{-\int_b^s \sigma(\tau)d\tau}. \tag{5.25}$$

Proof. To establish the Fréchet differentiability of $\Phi[\lambda](a)$ with respect to λ at zero, we show that each of the elements of $\Phi[\lambda](a) = (\Phi_1[\lambda](a), \Phi_2[\lambda](a), \ldots, \Phi_{\mathcal{N}}[\lambda](a))$ is Fréchet differentiable. We observe that for $\Phi_i^{(1)}[\lambda](a)$, $\Phi_i^{(1)}[0+h](a) - \Phi_i^{(1)}[0](a) =$

$\Phi_i^{(1)}[h](a) = \mathcal{T}_i[h](a) + \mathcal{R}_i^{(1)}[h](a)$, where $\mathcal{T}_i[h](a)$ is given above and $\mathcal{R}_i^{(1)}[h](a)$ is defined as follows

$$\mathcal{R}_i^{(1)}[h](a) = \int_0^{a_m} \beta(a,s)\psi(s)\left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau} h_i^{(1)}(b)\left[e^{-\int_0^b \Xi_i^{(1)}[h](\tau)d\tau} - 1\right]db\right)ds.$$

We note that $\frac{||\mathcal{R}_i^{(1)}[h](a)||}{||h||} \leq \omega||\tilde{\beta}||||h||$, which goes to zero as $||h|| \to 0$. Thus, $\mathcal{T}_i[h](a)$ is the Fréchet derivative of $\Phi_i^{(1)}[\lambda](a)$ at $\lambda = 0$. Similarly for $\Phi_i^{(2)}[\lambda](a)$, we observe that $\Phi_i^{(2)}[0+h](a) - \Phi_i^{(2)}[0](a) = \Phi_i^{(2)}[h](a)$ and $\frac{||\Phi_i^{(2)}[h](a)||}{||h||} \leq (\omega - 1)||\tilde{\beta}||||h|| \to 0$ as $||h|| \to 0$. Thus, the Fréchet derivative of $\Phi_i^{(2)}[\lambda](a)$ at $\lambda = 0$ is zero. Hence, the Fréchet derivative of $\Phi_i[\lambda](a)$ at $\lambda = 0$ is $\mathcal{T}_i[h](a)$. This completes the proof. □

5.3.3 Next-Generation Operator: Threshold Property

In this section we study the property of the next-generation operator that characterizes the existence of a non-trivial solution. From now on, we will drop $*$ in Eqs. (5.19)-(5.21). It is clear from Eqs. (5.19)-(5.21) that, $\lambda = \Phi[\lambda]$, where $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_{\mathcal{N}})^T$, and $\Phi[\lambda] := (\Phi_1[\lambda], \Phi_2[\lambda], \ldots, \Phi_{\mathcal{N}}[\lambda])^T$ admit a trivial solution $\lambda_i = 0$, for all $i = 1, 2, \ldots, \mathcal{N}$. This solution corresponds to the disease-free equilibrium where no disease exists. Here we are interested in investigating the existence of positive non-trivial solutions for Eqs. (5.19)-(5.21).

Let $\Phi[\lambda]$ be a nonlinear integral operator on the Banach space $\mathcal{Z} := [L^1(0, a_m)]^{\mathcal{N}}$, with $\Phi[0] = 0$. In addition, let $\mathcal{T} = \mathcal{L}_\lambda \Phi[\lambda = 0]$ be the Fréchet derivative of $\Phi[\lambda]$ at $\lambda = 0$, and $\mathcal{T} : \mathcal{Z} \to \mathcal{Z}$. Here $\mathcal{T}$ defines the next-generation operator which transforms the distribution of infected population to the distribution of secondary cases in the initial phase of epidemiological invasion. As a result, the basic reproductive number $\mathcal{R}_0$ is given by the spectral radius of $\mathcal{T}$ (i.e., spectral radius of $\mathcal{T}$ is the maximum eigenvalue of $\mathcal{T}$, see Diekmann & Heesterbeek 2000). In addition, we note that $\mathcal{T}[\lambda] := (\mathcal{T}_1[\lambda], \mathcal{T}_2[\lambda], \ldots, \mathcal{T}_{\mathcal{N}}[\lambda])^T$ is symmetric across strains because we have assumed that all disease-related age-specific parameters are the same. We can focus on just one of the next-generator operators.

To establish the properties of the threshold, or next-generation operator $\mathcal{T}[\lambda]$, we briefly summarize the necessary Perron-Frobenius theory for positive monotone operators on the partially ordered Banach space (Krein & Rutman, 1948; Marek, 1970; Sawashima, 1964). Let $\mathcal{Y}$ be a real Banach space and $\mathcal{Y}^*$ be its dual, the space of all linear functionals on $\mathcal{Y}$. In addition, let $F \in \mathcal{Y}^*$ at $\phi \in \mathcal{Y}$ be denoted by $\langle F, \phi \rangle$. Then, we have the following definitions:

Definition 5.3.10. *A closed subset $\mathcal{Y}_+$ of $\mathcal{Y}$ is called a* **cone** *if the following properties hold:*

1. $\mathcal{Y}_+ + \mathcal{Y}_+ \subset \mathcal{Y}_+$,
2. $c\mathcal{Y}_+ \subset \mathcal{Y}_+$ *for* $c \in \mathbb{R}_+ = [0, \infty)$,
3. $\mathcal{Y}_+ \bigcap (-\mathcal{Y}_+) = \{0\}$, *and*
4. $\mathcal{Y}_+ \neq \{0\}$.

Definition 5.3.10 gives the closedness, convexity, and multiplicative invariance (under multiplication by a positive constant) of $\mathcal{Y}_+$. Furthermore, we write

Definition 5.3.11.
1. $u \leq v$ *if and only if* $v - u \in \mathcal{Y}_+$
2. $u < v$ *if and only if* $v - u \in \mathcal{Y}_+ \backslash \{0\}$.

Definition 5.3.12.
1. *The cone $\mathcal{Y}_+$ is called* **total** *if the set $\{\phi - \psi : \phi, \psi \in \mathcal{Y}_+\}$ is dense in $\mathcal{Y}$.*
2. *The* **dual cone** *$\mathcal{Y}_+^*$ is the subset of $\mathcal{Y}^*$ that consists of all positive linear functionals on $\mathcal{Y}$.*

Definition 5.3.12(2) implies that $F \in \mathcal{Y}_+^*$ if and only if $F \in \mathcal{Y}^*$ and $\langle F, \phi \rangle \geq 0$ for all $\phi \in \mathcal{Y}_+$.

Definition 5.3.13.
1. *$\phi \in \mathcal{Y}_+$ is called a* **non-supporting/quasi-interior point** *if $\langle F, \phi \rangle > 0$ for all $F \in \mathcal{Y}_+^* \backslash \{0\}$.*
2. *A positive functional $F \in \mathcal{Y}_+^*$ is called* **strictly positive** *if $\langle F, \phi \rangle > 0$ for all $\phi \in \mathcal{Y}_+ \backslash \{0\}$.*

Definition 5.3.14. *Let $\mathcal{B}(\mathcal{Y})$ be a set of bounded linear operators that map $\mathcal{Y}$ into itself. The operators $\mathcal{L}_1, \mathcal{L}_2 \in \mathcal{B}(\mathcal{Y})$ are said to be positive with respect to the cone $\mathcal{Y}_+$ if $\mathcal{L}_1(\mathcal{Y}_+) \subset \mathcal{Y}_+$ and $\mathcal{L}_2(\mathcal{Y}_+) \subset \mathcal{Y}_+$. We say $\mathcal{L}_1 \geq \mathcal{L}_2$ if $(\mathcal{L}_1 - \mathcal{L}_2)(\mathcal{Y}_+) \subset \mathcal{Y}_+$.*

Definition 5.3.15. (Sawashima, 1964)

1. *The operator $\mathcal{L} \in \mathcal{B}(\mathcal{Y})$ is called* **semi-non-supporting** *if and only if for every $v \in \mathcal{Y}_+ \backslash \{0\}$, $F \in \mathcal{Y}_+^* \backslash \{0\}$, there exists a positive $p = p(v, F) \in \mathbb{Z}_+$ such that $\langle F, \mathcal{L}^p v \rangle > 0$,*
2. *The operator $\mathcal{L} \in \mathcal{B}(\mathcal{Y})$ is called* **non-supporting** *if and only if for every pair $v \in \mathcal{Y}_+ \backslash \{0\}$, $F \in \mathcal{Y}_+^* \backslash \{0\}$, there exists a positive $p = p(v, F) \in \mathbb{Z}_+$ such that $\langle F, \mathcal{L}^n v \rangle > 0$ for all $n \geq p$.*

We now state the following properties:

Lemma 5.3.16. *The threshold/next-generation operator $\mathcal{T}$ is positive and bounded.*

Proof. By Assumption 5.3.2, and using Eqs. (5.22)-(5.25), we observe that if $\lambda_i \in L^1_+([0,a_m])$,

$$\mathcal{T}_i[\lambda](a) = \int_0^{a_m}\left(\int_b^{a_m}\beta(a,s)\psi(s)\Theta(b,s)ds\right)\lambda_i(b)db \geq 0.$$

In addition, if $\lambda_i[\phi] \in L^1_+([0,a_m))\backslash\{0\}$, then $\mathcal{T}_i[\lambda](a) > 0$ also holds. Hence $\mathcal{T}[\lambda](a) \in [L^1_+(0,a_m)]^{\mathcal{N}}\backslash\{0\}$. Therefore, $\mathcal{T}$ is positive. Moreover,

$$\begin{aligned}
||\mathcal{T}_i[u]|| &= \int_0^{a_m}\left|\left(\int_0^{a_m}\left(\int_b^{a_m}\beta(a,s)\psi(s)\Theta(b,s)ds\right)u_i(b)db\right)\right|da \\
&\leq \int_0^{a_m}\left(\int_0^{a_m}\left(\int_b^{a_m}\beta(a,s)\psi(s)ds\right)|u_i(b)|\,db\right)da \\
&\leq \int_0^{a_m}\tilde{\beta}(a)\left(\int_0^{a_m}\left(\int_b^{a_m}\psi(s)ds\right)|u_i(b)|\,db\right)da \leq ||\tilde{\beta}||||u||.
\end{aligned}$$

Hence, $\mathcal{T}_i$ is bounded for each $i = 1,2,\ldots,\mathcal{N}$. This completes the proof. □

Theorem 5.3.17. *The threshold/next-generation operator $\mathcal{T}$ is non-supporting.*

Proof. It now suffices to show that the Definition 5.3.15 holds. We follow the approach of Inaba (1988, 1990). We define $\tilde{s}(a) := \varepsilon$ if $a \in (a_m - a^*, a_m)$, $a_m \in (0,\infty)$, and $\tilde{s}(a) := 0$ otherwise. Then it follows from Assumption 5.3.2(2) that $\beta(a,s) \geq \tilde{s}(s)$ for all $(a,s) \in (0,a_m)\times(0,a_m)$). Let f be the linear functional on $L^1(0,a_m)$ defined as

$$\langle f, u_i\rangle = \int_0^{a_m}\left[\int_b^{a_m}\tilde{s}(\tau)\psi(\tau)\Theta(b,\tau)d\tau\right]u_i(b)db.$$

So, the definition of $\tilde{s}$ implies that $f \in \left[L^1_+\right]^*$. We observe that $\int_b^{a_m}\psi(s)\Theta(b,s)ds > 0$ for all $b \in (0,a_m)$, and f is strictly positive. Hence, if $u_i \in L^1_+(0,a_m)$, then from $\beta(a,s,) \geq \tilde{s}(s)$, we have $\mathcal{T}_i[u](a) \geq \langle f,u_i\rangle e(a)$, $e(a) = 1$ for all $a \in \mathbb{R}_+$. Therefore, for any $n \in \mathbb{Z}_+$, we have $\mathcal{T}_i^{n+1}u \geq \langle f,u_i\rangle\langle f,e\rangle^n e(a)$, $e(a) = 1$. Moreover, if the functional $F \in \left[L^1_+(0,a_m)\right]^*\backslash\{0\}$, then $\langle F, \mathcal{T}_i^n[u]\rangle \geq \langle f,u_i\rangle\langle f,e\rangle^{n-1}\langle F,e\rangle > 0$, for $n \geq 1$ for every any pair of u_i and $F \in L^1_+\backslash\{0\}$. Hence $\mathcal{T}_i$ is non-supporting for all $i = 1,2,\ldots,\mathcal{N}$ by Definition 5.3.15, and consequently $\mathcal{T}$ is non-supporting. This completes the proof. □

Lemma 5.3.18. *Under Assumption 5.3.2, for all $\varepsilon > 0$, there exists $\delta > 0$ such that if $|h| < \delta$,*

$$\int_0^{a_m}|\Psi(a+h,b) - \Psi(a,b)|\,da < \varepsilon, \quad \textit{for all } b \in [0,a_m),$$

where Ψ is defined in Eq. (5.24).

Proof. We observe that, from Assumption 5.3.2, $\mathcal{H}_7$ and $\mathcal{H}_8$ imply that for every ε, there exists δ such that $\int_0^{a_m} |\beta(a+h,s)-\beta(a,s)|\,da \leq \varepsilon$ holds for all $s \in \mathbb{R}$ and $|h| < \delta$. Hence,

$$\begin{aligned}\int_0^{a_m} |\Psi(a+h,b)-\Psi(a,b)|\,da &= \int_0^{a_m} \left| \left(\int_b^{a_m} \beta(a+h,s)-\beta(a,s)\psi(s)\Theta(b,s)ds \right) \right| da \\ &\leq \int_0^{a_m} \left(\int_0^{a_m} |\beta(a+h,s)-\beta(a,s)|\,\psi(s)ds \right) da \leq \varepsilon.\end{aligned}$$

This completes the proof. □

Theorem 5.3.19. *The next-generation operator $\mathcal{T} : \mathcal{Z} \to \mathcal{Z}$ is compact.*

Proof. Let $\mathcal{S}$ be a bounded subset of $\mathcal{Z}$ and let $F_i := \mathcal{T}_i(\mathcal{S}) \subset L^1(0,a_m)$ for each $i = 1,2,\ldots,\mathcal{N}$. From Lemma 5.3.16, $\mathcal{T}_i$ is bounded, so F is bounded. Hence, the first criterion of Theorem 5.3.7 is satisfied. It now suffices to show that the second criterion also holds. That is, by Assumption 5.3.2 and Lemma 5.3.18, we have

$$\begin{aligned}||\rho_h \mathcal{T}_i[u] - \mathcal{T}_i[u]|| &= \int_0^{a_m} |\mathcal{T}_i[u](a+h) - \mathcal{T}_i[u](a)|\,da \\ &\leq \int_0^{a_m} \left(\int_0^{a_m} |\Psi(a+h,\tau)-\Psi(a,\tau)|\,da \right) |u_i(\tau)|\,d\tau \leq \varepsilon ||u_i|| \leq \varepsilon c_0,\end{aligned}$$

where $||u|| \leq c_0$ for all $u \in \mathcal{S}$. Hence the second criterion of Theorem 5.3.7 is also satisfied. Therefore, $\mathcal{T}_i(\mathcal{S})$ is precompact or relatively compact in $L^1(0,a_m)$. This completes the proof. □

Because $\mathcal{T}$ is linear and compact, it is continuous. Hence, we can conclude that $\mathcal{T}$ is also completely continuous (Kreyszig, 1978).

5.3.4 Existence and Threshold Conditions for Equilibria

To establish the necessary condition for the existence of nontrivial equilibria, we use some concepts from positive operator theory. Here we focus on the extension of Perron-Frobenius theory in the context of infinite-dimensional Banach spaces. In particular, we recall some important results from spectral theory that will allow us to characterize the properties of the operators and to establish a threshold-like condition for the existence of fixed point solutions. Given a linear next-generation operator $\mathcal{T}$ on $\mathcal{Z}$, we make the following definitions (Engel & Nagel, 2000):

Definition 5.3.20. *Let $\mathcal{L} : D(\mathcal{L}) \subset \mathcal{Y} \to \mathcal{Y}$ be a closed linear operator. Then,*

1. *Spectrum of $\mathcal{L}$: $\Sigma(\mathcal{L}) := \{\zeta \in \mathbb{C} : \zeta - \mathcal{L}$ is not bijective$\}$,*

2. *Resolvent set of* $\mathcal{L}$*:* $\rho(\mathcal{L}) := \mathbb{C}\backslash\Sigma(\mathcal{L})$,

3. *Resolvent of* $\mathcal{L}$*:* $R(\zeta, \mathcal{L}) := (\zeta - \mathcal{L})^{-1}$ *at* $\zeta \in \rho(\mathcal{L})$,

4. *Point spectrum of* $\mathcal{L}$*:* $\Sigma_p(\mathcal{L}) := \{\zeta \in \mathbb{C} : \zeta - \mathcal{L}$ *is not injective*$\}$,

5. *Spectral radius of* $\mathcal{L}$*:* $r(\mathcal{L}) := \sup\{|\zeta| : \zeta \in \Sigma(\mathcal{L})\}$,

6. *Spectral bound of* $\mathcal{L}$*:* $b(\mathcal{L}) := \sup\{\mathbb{R}e\zeta : \zeta \in \Sigma(\mathcal{L})\}$,

Note that if $\mathcal{L} : \mathcal{Y} \to \mathcal{Y}$ is a bounded linear operator, then it is closed. Using these definitions, we summarize the following results from linear operator theory (Krein & Rutman, 1948; Sawashima, 1964; Kato, 1984; Marek, 1970):

Theorem 5.3.21. (Krein & Rutman, 1948) *Let* $\mathcal{K}$ *be a cone in a Banach space* $\mathcal{Y}$, $\mathcal{K} \subset \mathcal{Y}$ *which is total. In addition, let* $\mathcal{L} : \mathcal{Y} \to \mathcal{Y}$ *be a completely continuous linear operator such that* $\mathcal{L}(\mathcal{K}) \subset \mathcal{K}$, *and* $r(\mathcal{L}) > 0$. *Then* $r(\mathcal{L})$ *is an eigenvalue of* $\mathcal{L}$ *(of* $\mathcal{L}^*$*) and there exists* $v \in \mathcal{K}\backslash\{0\}$ *(*$\eta \in \mathcal{K}^*\backslash\{0\}$*) such that* $\mathcal{L}v = r(\mathcal{L})v$ *(*$\mathcal{L}^*\eta = r(\mathcal{L})\eta$*). In addition, if* $\mathcal{L}$ *is strictly positive with respect to* $\mathcal{K}$, *then*

1. $\mathcal{L}$ *has one and only one eigenvector* $v \in Int(\mathcal{K})$, $\mathcal{L}v = r(\mathcal{L})v$, *where* $Int(\mathcal{K})$ *is the interior of* $\mathcal{K}$.

2. $\mathcal{L}^*$ *has one and only one eigenvector* $\eta \in \mathcal{K}^*\backslash\{0\}$ *and* η *is strictly positive, and* $\mathcal{L}^*\eta = r(\mathcal{L})\eta$.

3. $|\zeta| < r(\mathcal{L})$, *for all* ζ *eigenvalues of* $\mathcal{L}$ *such that* $\zeta \neq r(\mathcal{L})$.

Corollary 5.3.22. (Krein & Rutman, 1948) *Let* $\mathcal{L}$ *be a compact positive linear operator on* $\mathcal{Y}$ *with positive spectral radius,* $r(\mathcal{L}) > 0$, *then* $r(\mathcal{L})$ *is an eigenvalue of* $\mathcal{L}$ *with corresponding eigenvector* $\nu \in \mathcal{Y}_+\backslash\{0\}$.

The theorems above imply that $r(\mathcal{L})$ is an eigenvalue that possesses a positive eigenvector in $Int(\mathcal{K})$ with geometric multiplicity of one. However, because $L^1_+([0,\infty))$ in our case is a total cone with empty interior, the theorems above are not sufficient. Therefore, we need a different class of positive linear operators. The Sawashima extension seems to be more appropriate. Below, we outline the theorem:

Theorem 5.3.23. (Sawashima, 1964) *If* $r(\mathcal{L}) > 0$, $\mathcal{L} \in \mathcal{B}(\mathcal{Y})$ *is a semi-nonsupporting positive linear operator on* $\mathcal{Y}$, *and* $r(\mathcal{L})$ *is a pole of the resolvent* $R(\zeta, \mathcal{L})$, *then*

1. $r(\mathcal{L}) \in \Sigma_p(\mathcal{L})\backslash\{0\}$ *and* $r(\mathcal{L})$ *is a simple pole*

2. *The eigenspace corresponding to $r(\mathcal{L})$ is a one-dimensional subspace of $\mathcal{Y}$ spanned by a quasi-interior point $\psi \in \mathcal{Y}_+$. If $\phi \in \mathcal{Y}_+,\ c \in \mathbb{R}$ satisfying $\mathcal{L}\phi = c\phi$, then $c = r(\mathcal{L})$, and there exists a positive constant $k \in \mathbb{R}$ such that $\phi = k\psi$*

3. *$r(\mathcal{L}) \in \Sigma_p(\mathcal{L}^*)$ and the eigenspace of $\mathcal{L}^*$ corresponding to $r(\mathcal{L})$ is a one-dimensional subspace of $\mathcal{Y}^*$ spanned by a strictly positive function f.*

Theorem 5.3.24. (Marek, 1970) *Let $\mathcal{Y}$ be a Banach lattice, and let $\mathcal{L}_1,\ \mathcal{L}_2$ be positive bounded linear operators on $\mathcal{Y}$. Then the following comparison properties hold:*

1. *If $\mathcal{L}_1 \leq \mathcal{L}_2$ then $r(\mathcal{L}_1) \leq r(\mathcal{L}_2)$*

2. *If $\mathcal{L}_1,\ \mathcal{L}_2$ are semi-nonsupporting and compact, and $\mathcal{L}_1 \leq \mathcal{L}_2,\ \mathcal{L}_1 \neq \mathcal{L}_2,\ r(\mathcal{L}_2) \neq 0$ then $r(\mathcal{L}_1) < r(\mathcal{L}_2)$.*

By Theorems 5.3.17–5.3.19 and 5.3.21–5.3.24, we summarize the above discussion in the statement below.

Remark 5.3.25. *Since $r(\mathcal{T})$ is one of the eigenvalues of $\mathcal{T}$ (in fact the largest eigenvalue, see Theorem 5.3.21-5.3.21 Corollary 5.3.22), $r(\mathcal{T})$ is a pole of the resolvent. In addition, if $\mathcal{T}$ is non-supporting, then $r(\mathcal{T})$ is a simple pole (see Theorem 5.3.23). Hence, the spectral radius $r(\mathcal{T})$ of the threshold operator $\mathcal{T}$ is the only positive eigenvalue with a positive eigenvector in $L^1(0, a_m)$. Furthermore, $r(\mathcal{T})$ is an eigenvalue of $\mathcal{T}^*$ with a strictly positive eigenfunctional.*

Using the above Theorems 5.3.23–5.3.24, and Krasnoselskii's fixed point theorem (Krasnoselskii, 1964), we obtain the following theorem:

Theorem 5.3.26. (Krasnoselskii, 1964) *Suppose g is a completely continuous positive nonlinear operator on a real Banach space $\mathcal{Y}$ with positive cone $\mathcal{Y}_+$; and $g(\mathcal{Y}_+)$ is bounded with $g(0) = 0$ and with a Fréchet derivative $\mathcal{L} := g'(0)$ at zero. In addition, let $\mathcal{L}$ have a positive characteristic vector (eigenvector) $\nu_0 \in \mathcal{Y}_+$ corresponding to the characteristic value (eigenvalue) $\zeta_0 \in \mathbb{R}$ such that $\mathcal{L}\nu_0 = \zeta_0\nu_0$, where $\zeta_0 > 1$; and $\mathcal{L}$ has no positive eigenvector in $\mathcal{Y}_+$ corresponding to eigenvalue 1. Then, the nonlinear operator g has at least one non-zero positive fixed point in the cone $\mathcal{Y}_+$.*

Proof. See Appendix C.1.1. □

We will use this result to show that when $r(\mathcal{T}) > 1$, Φ has a non-zero positive fixed point. We now establish the following results:

Lemma 5.3.27. *The inequality $\Phi[\lambda](a) \leq \mathcal{T}[\lambda](a)$ holds.*

Proof. We observe that Φ can be written as, $\Phi_i[\lambda](a) = \tilde{\Phi}_i^{(1)}[\lambda](a) + \tilde{\Phi}_i^{(2)}[\lambda](a)$, where $\tilde{\Phi}_i^{(k)}[\lambda](a)$, $k = 1, 2$ and $i = 1, 2, \ldots, \mathcal{N}$ are given as follows:

$$\begin{aligned}
\tilde{\Phi}_i^{(1)}[\lambda](a) &= \int_0^{a_m} \beta(a,s)\psi(s) \left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau} (1-\gamma(b)) \Lambda_i^{(1)}[\lambda](b)db \right) ds, \\
\tilde{\Phi}_i^{(2)}[\lambda](a) &= \int_0^{a_m} \beta(a,s)\psi(s) \left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau} \gamma(b) \Lambda_i^{(2)}[\lambda](b)db \right) ds,
\end{aligned}$$

where $\Lambda_i^{(1)}[\lambda](b) = \lambda_i(b)e^{-\int_0^b \lambda_i(\tau)d\tau}$, and $\Lambda_i^{(2)}[\lambda](b) = \Lambda_i^{(1)}[\lambda](b)e^{-\int_0^b \Xi_i^{(2)}[\lambda](\tau)d\tau}$. Then

$$\begin{aligned}
\Phi_i[\lambda](a) &\leq \int_0^{a_m} \beta(a,s)\psi(s) \left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau} (1-\gamma(b)) \Lambda_i^{(1)}[\lambda](b)db \right) ds \\
&\quad + \int_0^{a_m} \beta(a,s)\psi(s) \left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau} \gamma(b) \Lambda_i^{(1)}[\lambda](b)db \right) ds, \\
&= \int_0^{a_m} \beta(a,s)\psi(s) \left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau} \lambda_i(b) e^{-\int_0^b \lambda_i(\tau)d\tau} db \right) ds \\
&\leq \int_0^{a_m} \beta(a,s)\psi(s) \left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau} \lambda_i(b) db \right) ds = \mathcal{T}_i[\lambda](a).
\end{aligned}$$

This completes the proof. □

Theorem 5.3.28. *The following properties hold:*

1. *If $r(\mathcal{T}) < 1$, then $u = 0$ is the only solution of $u = \Phi(u)$ in $\mathcal{Z}_+$ and it is the disease-free equilibrium*

2. *If $r(\mathcal{T}) > 1$, then $u = \Phi(u)$ has at least one solution in $\mathcal{Z}_+ \backslash \{0\}$ and it is the endemic equilibrium.*

Proof. Suppose $r(\mathcal{T}) < 1$. We show that when $r(\mathcal{T}) < 1$, $u = 0$ is the only fixed point by contradiction. Suppose there exists a solution $\overline{u} \in \mathcal{Z}_+ \backslash \{0\}$ of $\overline{u} = \Phi(\overline{u})$, then $\overline{u} = \Phi(\overline{u}) \leq \mathcal{T}(\overline{u})$ (i.e., Lemma 5.3.27). From Definitions 5.3.12-5.3.13, we let $\mathcal{F}_0^* \in [\mathcal{Z}_+]^* \backslash \{0\}$ denote the adjoint eigenvector (strictly positive) of $\mathcal{T}$ corresponding to an eigenvalue $r(\mathcal{T})$. Then by duality pairing, $\langle \mathcal{F}_0^*, \mathcal{T}(\overline{u}) - \overline{u} \rangle = (r(\mathcal{T}) - 1) \langle \mathcal{F}_0^*, \overline{u} \rangle$. Because $\mathcal{T}(\overline{u}) - \overline{u} \in \mathcal{Z}_+$ and $\mathcal{F}_0^*$ is strictly positive, we note that $\langle \mathcal{F}_0^*, \mathcal{T}(\overline{u}) - \overline{u} \rangle \geq 0$. In addition, we also observe that $\langle \mathcal{F}_0^*, \overline{u} \rangle > 0$ since, by assumption, $\mathcal{F}_0^*, \overline{u} \in \mathcal{Z}_+ \backslash \{0\}$. Then we have $\langle \mathcal{F}_0^*, \mathcal{T}(\overline{u}) - \overline{u} \rangle = (r(\mathcal{T}) - 1) \langle \mathcal{F}_0^*, \overline{u} \rangle < 0$, because $r(\mathcal{T}) - 1 < 0$. This is a contradiction. Hence, $\overline{u} = 0$ must be the only solution of $\overline{u} = \Phi(\overline{u}) \in \mathcal{Z}_+$.

Now suppose $r(\mathcal{T}) > 1$. To begin, we show that the conditions of Theorem 5.3.26 are satisfied. We note that from Theorem 5.3.8, Φ is a completely continuous operator in the Banach space $\mathcal{Z}$. Moreover, from Theorem 5.3.5, $\Phi(\mathcal{Z}_+)$ is bounded. In addition, Φ is Fréchet differentiable, with the Fréchet derivative of Φ at zero given

by the linear operator $\mathcal{T} = \Phi'(0)$ (i.e., Theorem 5.3.9) and $\Phi(0) = 0$. Because $\mathcal{T}$ is positive, bounded, compact and non-supporting (i.e. Lemma 5.3.16, Theorems 5.3.17–5.3.19), from Theorems 5.3.21-5.3.21, Corollary 5.3.22 and Theorem 5.3.23, we have that $\zeta_0 = r(\mathcal{T}) > 0$ is the eigenvalue of $\mathcal{T}$ satisfying $\mathcal{T}\nu_0 = r(\mathcal{T})\nu_0$, where $\nu_0 \in \mathcal{Z}_+\backslash\{0\}$ is the only positive eigenvector corresponding to the eigenvalue $r(\mathcal{T})$ which is a simple pole (due to compactness of $\mathcal{T}$). Furthermore, by assumption, $r(\mathcal{T}) > 1$ and $\mathcal{T}$ does not have a positive eigenvector in $\mathcal{Z}_+$ corresponding to the eigenvalue equal to one. Hence, the conditions of Theorem 5.3.26 are satisfied. Therefore, Φ has a non-zero positive fixed point. This completes the proof. □

The proof above follows the approaches of Krasnoselskii (1964) and Inaba (1988), and establishes the properties of reproductive number which we denote as $\mathcal{R}_0 = r(\mathcal{T})$. It provides a threshold-like property for the existence of steady states which are the solutions corresponding to a fixed point problem $\lambda = \Phi(\lambda)$. In the proof above, we have not discussed the case when $r(\mathcal{T}) = 1$ because this corresponds to the bifurcation point. Note that Theorem 5.3.26(2) does not establish the uniqueness of the non-zero (non-trivial) positive (endemic) equilibrium. To establish the uniqueness we must provide additional properties in a manner similar to that of Krasnoselskii (1964). However, our goal in this chapter is to show the existence of a positive fixed point and the uniqueness problem is not necessary since we know from numerical simulations that there are different concepts of endemicity corresponding to both the semi-trivial and non-trivial steady states. In the next chapter, we consider the stability of both the disease-free and endemic equilibria.

6 Age-specificity II: Stability Analysis of Equilibria

In this chapter, we provide stability conditions for the equilibria whose existence was previously discussed in Chapter 5. We use graph-theoretical notions, as well as numerical and analytical insights gained in Chapter 3 to simplify the analysis carried out herein. In particular, using the characteristic equation approach, we state sufficient conditions to establish the stability of both weakly (due to the principle of competitive exclusion) and strongly (full co-existence) endemic equilibria.

6.1 Introduction

In the previous chapter, we established a threshold condition, defined by the spectral radius of the next-generation operator $r(\mathcal{T})$ (henceforth $\mathcal{R}_0 := r(\mathcal{T})$, where the linear operator $\mathcal{T} : L^1([0,\infty)) \to L^1([0,\infty))$, for the existence of trivial (disease-free) and non-trivial (endemic) equilibria. This threshold parameter is known as the basic reproductive number or ratio and is defined as the expected number of new infections caused by one infected individual during its entire infectious period. In the case of the homogeneous systems studied in previous chapters, an explicit expression for the basic reproductive number was obtained. However, in heterogeneous systems such as the model considered herein, it is not always straight-forward and/or may not be possible to obtain an explicit expression for $\mathcal{R}_0$. This is only possible in a few cases (e.g., proportionate mixing, intra-group mixing, see Diekmann & Heesterbeek, 2000). For ease of analysis, we make a simplification to the mixing property. We make the following assumption,

Assumption 6.1.1. *The contact rate $\beta(a,s)$ is of the form $\beta(a,s) := \beta_1(a)\beta_2(s)$ where $\beta_1(a) \in L^\infty_+(0,a_m)$ and $\beta_2(s) \in L^\infty_+(0,a_m)$.*

This assumption implies that the ages of susceptible and infected individuals are uncorrelated. Although this proportionate mixing assumption is not needed for the analysis of the disease-free equilibrium, it allows us to calculate the basic reproductive

number explicitly, and to simplify the analysis in the case of the endemic equilibrium. We note that, from the previous chapter, $\mathcal{T}$ is compact and non-supporting. Hence, using the results of Klein & Rutman (1948) and Sawashima (1964) (see Theorems 5.3.21–5.3.21 and 5.3.23, and Corollary 5.3.22), $r(\mathcal{T})$ is the only positive eigenvalue of $\mathcal{T}$, with a positive eigenfunction in $L^1_+(0, a_m)$. Therefore, to calculate $\mathcal{R}_0$ under the proportionate mixing or separable mixing, we observe that the next-generation operator satisfies:

$$\mathcal{T}[u](a) = \int_0^{a_m} \left(\int_b^{a_m} \beta(a,s)\psi(s) e^{-\int_b^s \sigma(\tau)d\tau} ds \right) u(b) db. \tag{6.1}$$

Hence, under proportionate mixing, we have

$$\begin{aligned} \mathcal{T}[u](a) &= \int_0^{a_m} \left(\int_b^{a_m} \beta_1(a)\beta_2(s)\psi(s) e^{-\int_b^s \sigma(\tau)d\tau} ds \right) u(b) db \\ &= \beta_1(a) \int_0^{a_m} \left(\int_b^{a_m} \beta_2(s)\psi(s) e^{-\int_b^s \sigma(\tau)d\tau} ds \right) u(b) db. \end{aligned} \tag{6.2}$$

Substituting β_1 for u and changing the order of integration yields

$$\begin{aligned} \mathcal{T}[\beta_1](a) &= \beta_1(a) \int_0^{a_m} \left(\int_b^{a_m} \beta_2(s)\psi(s) e^{-\int_b^s \sigma(\tau)d\tau} ds \right) \beta_1(b) db \\ &= \beta_1(a) \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s \beta_1(b) e^{-\int_b^s \sigma(\tau)d\tau} db \right) ds. \end{aligned} \tag{6.3}$$

So β_1 is an eigenfunction of $\mathcal{T}$ with positive eigenvalue. Therefore, under the proportionate mixing, the spectral radius of the next-generation operator $r(\mathcal{T})$ is given by $\mathcal{R}_0$, where

$$\mathcal{R}_0 = \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s \beta_1(b) e^{-\int_b^s \sigma(\tau)d\tau} db \right) ds. \tag{6.4}$$

Throughout this thesis, we have assumed that strains have the same epidemiological parameters, where $\beta_i(a,s) = \beta(a,s)$, $\sigma_i(a) = \sigma(a)$ and $\psi_i(a) = \psi(a)$. In general, the epidemiological parameters for each strain will be different. That is, $\beta_i(a,s) \neq \beta_j(a,s)$, $\sigma_i(a) \neq \sigma_j(a)$ and $\psi_i(a) \neq \psi_j(a)$, $i \neq j$. In this case, we can define the strain-specific reproductive number $\mathcal{R}_0^{(i)}$ with proportionate mixing as

$$\mathcal{R}_0^{(i)} = \int_0^{a_m} \beta_{i,2}(s)\psi_i(s) \left(\int_0^s \beta_{i,1}(b) e^{-\int_b^s \sigma_i(\tau)d\tau} db \right) ds. \tag{6.5}$$

Then, the reproductive number $\mathcal{R}_0$ is defined as

$$\mathcal{R}_0 = \max_{1 \leq i \leq \mathcal{N}} \{\mathcal{R}_0^{(i)}\}. \tag{6.6}$$

We note that if $\mathcal{R}_0^{(i)} = \mathcal{R}_0^{(j)}$, then $\mathcal{R}_0 = \max_{1\leq i\leq \mathcal{N}}\{\mathcal{R}_0^{(i)}\}$ reduces to the basic reproductive number in Eq. (6.4). In addition, we showed in the previous chapter that when the spectral radius of the next-generator operator $\mathcal{R}_0 < 1$, the dynamics are simple and there is only a trivial solution corresponding to the disease-free equilibrium (or disease extinction); and when $\mathcal{R}_0 > 1$, the system has at least one non-zero positive solution. However, in the previous chapter, the threshold result does not provide uniqueness of the non-trivial endemic equilibrium. In fact, when $\mathcal{R}_0 > 1$, numerical simulations in the previous chapters suggest that the endemic equilibrium is not unique, and the dynamics are much more involved. In particular, the system can exhibit coexistence and the principle of competitive exclusion whereby in the presence of multiple strains only the strongest can survive depending on the strength of parameters modulating the competition. Suppose we have a 2-strain system, then there are three possible endemic configurations. That is, if $\mathcal{R}_0 > 1$, then either strain 1 persists while strain 2 goes to extinction, or strain 2 persists while strain 1 goes to extinction, or both strains 1 and 2 coexist. Similarly, in a 4-strain system, there are more possibilities. Figure 6.1 illustrates all the possible endemic configurations. For instance, there are 15 possible persistent sets, namely $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\{1,2\}$, $\{2,3\}$, $\{3,4\}$, $\{4,1\}$, $\{1,2,3\}$, $\{2,3,4\}$, $\{3,4,1\}$, $\{4,1,2\}$, $\{1,3\}$, $\{2,4\}$ and $\{1,2,3,4\}$. These are represented in Fig. 6.1, where persistent strains are marked by black circles. The 4-strain system illustrated in Fig. 6.1 is equivalent to a 2-locus-2-allele system resulting on a square lattice. Figure 6.1 also illustrates different types of equilibrium clusters where black circles ($\bullet$) denote the persistent strains while white circles ($\circ$) represent extinct strains.

Using the idea summarized in Fig. 6.1, we introduce the following notation. Let $\mathcal{P}$ denote the set of persistent strains (i.e., strains that can survive) and be defined as

$$\mathcal{P} := \begin{cases} \{i \in \{1,2,\ldots,\mathcal{N}\} : \mathcal{R}_0^i > 1\}, & \mathcal{R}_0 > 1 \\ \varnothing, & \mathcal{R}_0 < 1. \end{cases} \tag{6.7}$$

We note that when $\mathcal{R}_0^i = \mathcal{R}_0^j,\ i \neq j$ for all i, then the set simplifies to a simple set. In addition, we observe that the possible endemic configurations outlined previously can be classified into two notions of endemicity whether there is a principle of competitive exclusion (partial coexistence) or full coexistence. Using the above concept, we define the following notion of positive solutions.

Definition 6.1.2. *1. A steady state solution is called weakly endemic if $y_i(a) = y_i^*(a) \neq 0$ (i.e., $y_i^*(a) > 0$) for some $i = 1,2,\ldots,\mathcal{N}$.*

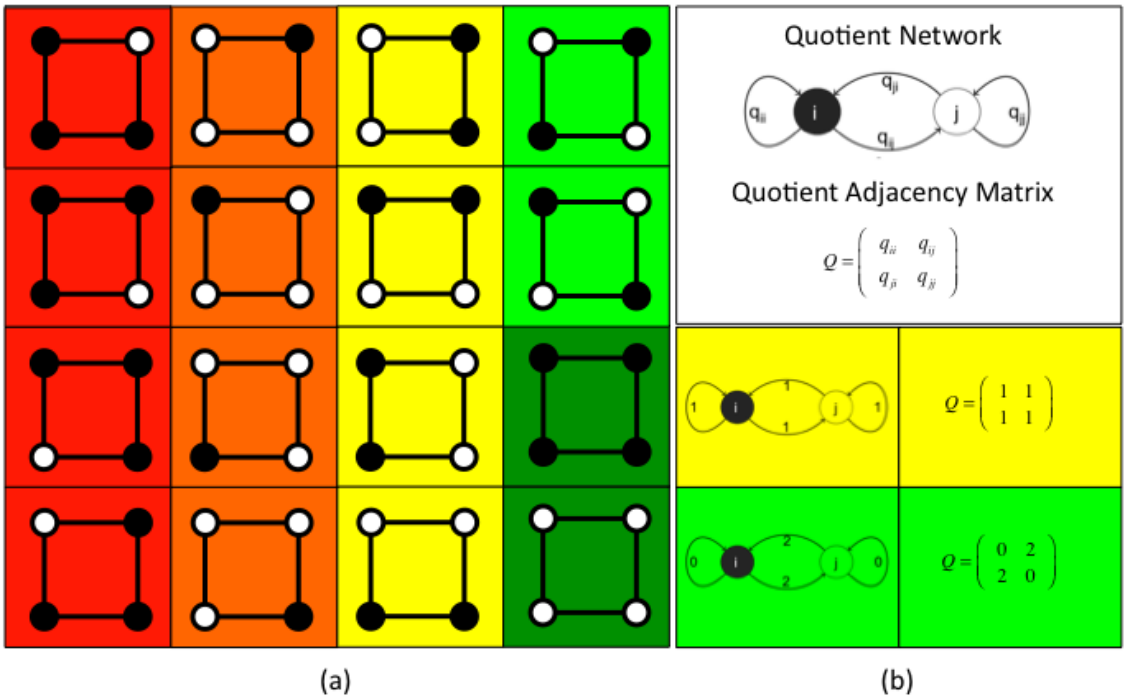


Figure 6.1: **Multi-locus System (Possible Equilibria for Four Strain System on a Square Lattice with $\mathbb{D}_4$-Symmetry) and Quotient Network Representation.** See text for full details (Quotient network sketch is modified from Chan & Yu, 2013).

2. *A steady state solution is called strongly endemic if* $y_i(a) = y_i^*(a) \neq 0$ *(i.e.,* $y_i^*(a) > 0$*) for all* $i = 1, 2, \ldots, \mathcal{N}$.

In the definition above, weakly endemic equilibrium is equivalent to the principle of competitive exclusion and strongly endemic equilibrium represents coexistence. In the proceeding sections, we will use the above notion with structural properties of the strain space to study the asymptotic behaviours of the model and the equilibria. Before we start, we first discuss some preliminary notions from graph theory that will later be used to facilitate our analysis.

6.2 Preliminaries: Networks and Symmetry

Herein, we shall give a brief introduction to symmetry in networks using algebraic graph-theoretical concepts (Biggs, 1974; Godsil & Royle, 2001; Lauri & Scapellato, 2003; Xiao et al., 2008).

A network can be thought of as a graph $G = G(V, E)$ with a set of vertices V and edges $E \subset V \times V$, where vertices are said to be adjacent if there is an edge connecting them. Let G and G' be two graphs, we call G and G' isomorphic and $G \cong G'$ if there exists a bijection $g : V \to V'$ with $xy \in E$ and $g(x)g(y) \in E'$ for

all $x, y, \in V$. Such a map g is called an isomorphism, and if $G \equiv G'$, it is called an automorphism. An automorphism is a set of all permutations of the vertices that can preserve the adjacency of the vertices; and the set of automorphisms under product or composition permutations forms a group $\mathcal{G} = Aut(G)$. This automorphism group describes the symmetry and redundancy structures of the network, where we say a network is symmetric if it possesses a nontrivial automorphism group while it is asymmetric if the network only admits a trivial (identity) permutation. Our goal in this chapter is to leverage the structural properties of the strain space (e.g., strain relatedness structures describing how strains are related and affect each other) and the correspondence between network symmetry and redundancy to reduce the dimensionality of the system; as well as to provide a more rigorously robust result. Instead of analyzing the system directly, we take advantage of the compartmentalization of the system. In the proceeding sections, we will provide a preliminary background and definitions.

6.2.1 Preliminaries and Definitions

The adjacency matrix describes the topology of a simple network G of N nodes and is defined by A with entries given by

$$A_{ij} := \begin{cases} 1, & \text{if there is a link between } i \text{ and } j \\ 0, & \text{otherwise.} \end{cases} \tag{6.8}$$

In general $A_{ij} \neq A_{ji}$ (i.e., directed network). For an undirected network, $A_{ij} = A_{ji}$ and A is symmetric with real spectrum and contains a symmetric structure if we can find a permutation matrix such that the adjacency relations remain unchanged where P is defined by the following entries

$$P_{ij} := \begin{cases} 1, & \text{if permutation of the vertices } v_i \text{ is } v_j \text{ (i.e. } p(v_i) = v_j \text{)} \\ 0, & \text{otherwise,} \end{cases} \tag{6.9}$$

where p is a permutation operator of the vertices of the network G. If $p \in Aut(G)$, then A and P commute. This has a profound implication on the multiplicity and redundancy of the eigenvalues. That is, if v_ζ is an eigenvector of A corresponding to the eigenvalue ζ, then Pv_ζ is also an eigenvector of A corresponding to ζ. Hence, symmetric networks tend to have spectral multiplicity (i.e., eigenvalue multiplicity), and eigenvalues can be decomposed into redundant and non-redundant classes. The formalism adopted herein to remove the redundant portion of the graph and to focus on the non-redundant portion which contains structurally relevant information about

the graph relies on this fact. We now recall some notions from algebraic graph theory. In particular, we focus on automorphism partitions and orbits of graphs. We provide the following definition.

Definition 6.2.1. *Consider p, a permutation of the vertices G. We have*

1. *The support of an automorphism is the set of vertices which p moves, and it is given as*

$$supp(p) \quad = \quad \{v_i \in V : p(v_i) \neq v_i\} \tag{6.10}$$

2. *Automorphisms are support-disjoint if every pair of automorphisms $p \in P$ and $q \in Q$ have disjoint supports.*

From the above definition, if automorphism subgroups are support-disjoint, then the independence of the group action on the network of these subgroups can be used to factorize the automorphism groups of the network into irreducible blocks such that these blocks cannot be further decomposed into support-disjoint blocks (i.e., minimal symmetric structure). For example, if a graph G is an automorphism group $\mathcal{G} = Aut(G)$ generated by a set of $\mathcal{V}$ generators, then $\mathcal{V}$ can be partitioned into a support-disjoint subset of vertices $\mathcal{V} = \mathcal{V}_1 \bigcup \mathcal{V}_2 \bigcup \ldots \bigcup \mathcal{V}_n$ $(\mathcal{V}_i \bigcap \mathcal{V}_j = \varnothing)$, where $\mathcal{V}_i$ cannot be decomposed into a support-disjoint subset and are minimal. Let $\mathcal{G}_i$ be the subgroup generated by $\mathcal{V}_i$. Then we can construct $\mathcal{G} = Aut(G)$ from these subgroups $\mathcal{G}_i$ as $\mathcal{G} = \mathcal{G}_1 \times \mathcal{G}_2 \times \ldots \times \mathcal{G}_k$ since each $\mathcal{G}_i$ may be related to a subgroup of G. These automorphism groups may be used to partition the vertex set into disjoint structural equivalence classes called orbits. That is, for every vertex $v \in V$, v belongs to the orbit $\Delta(v) = \{g \cdot v \in V : g \in \mathcal{G}\}$. Vertices in the same orbits are structurally equivalent and are indistinguishable (see Fig. 6.2 where vertices are coloured by orbits). The vertices can be permuted without changing the adjacency structure of the network. Hence, vertices in the same orbits are vertex invariant with similar degree, eigenvector centrality and clustering coefficient. Therefore, these automorphism partitions can be used as a convenient way of removing structural redundancy and reducing the network into a coarse mesoscale network. This mesoscale network is known as a quotient network, Q.

Figures 6.1(b) and 6.2(b) show the quotient network of the network shown in Figs. 6.1 (a) and 6.2 (a), respectively. We note that quotient networks are generally multigraphs (i.e., graphs with weighted and directed edges). Formally, let $\Delta(v) = \{\Delta(v_1), \Delta(v_2), \ldots, \Delta(v_s)\}$ be the automorphism partition of a network in which the vertices of G are partitioned under the action of $\mathcal{G}$. Then, the adjacency matrix of Q is given by the elements q_{ij}, where the constant q_{ij} is the number of edges

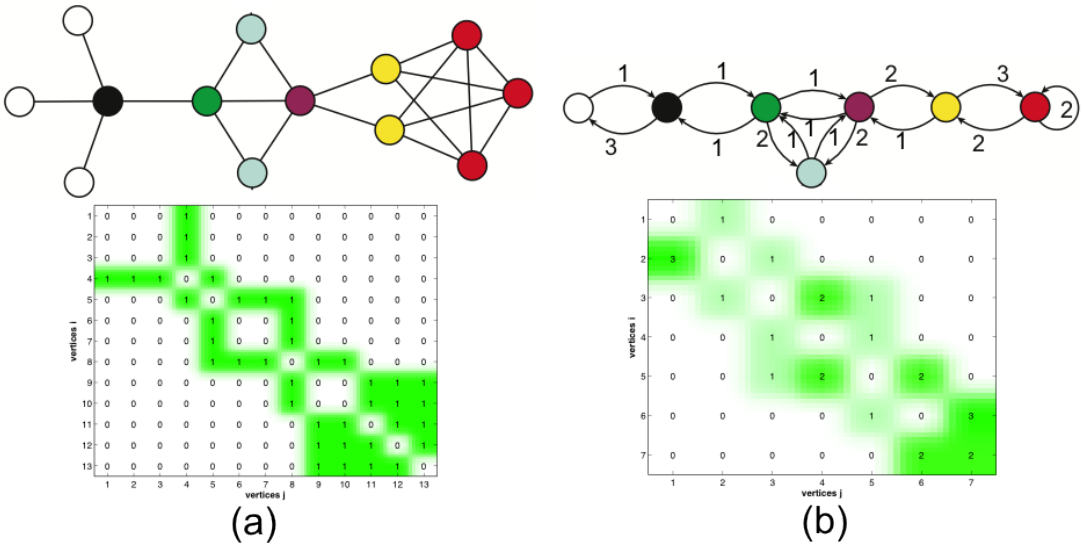


Figure 6.2: **Network and its Quotient.** Modified and adapted with permission from Xiao et al. (2008). Copyrighted by the American Physical Society.

starting from a vertex in Δ_i and ending in vertices in Δ_j; and q_{ij} is constant and independent of the choice of $v \in \Delta_i$ due to disjoint equivalence. Hence, it depends only on i and j. Quotient networks contain all the structural information of the parent graph, excluding redundancy. Therefore, from an algebraic perspective, the eigenvalues of the adjacency matrices of quotient networks Q are a subset of the eigenvalue of the adjacency matrix of the parent network G with fewer singularities or degeneracy. Moreover, quotient networks Q preserve the fundamental properties of G and only contain properties related to the functionality and complexity of the parent network G. In our analysis, we will use this quotient network notion to reduce the dimensionality of the system, for it retains all the structurally relevant information about the parent network and removes all the redundancy. In the next section, we relate the notion of quotient network and the existence of synchronous equilibrium clusters.

6.2.2 Notion of Synchronous Equilibrium Clusters

To tie the above ideas to the existence of synchronous equilibrium clusters, we use the concept of minimal balanced colouring (Golubitsky et al., 2005; Wang & Golubitsky, 2005; Golubitsky & Stewart, 2006). Here colouring is used in a manner different from the definition used in the graph-theoretical context, whereby the graph-theoretical notion of colouring refers to colouring of a graph as an assignment of colours to the vertices (i.e., one colour to each vertex so adjacent vertices have different colours). In our context, two adjacent vertices can have the same colour as long as their dynamics

or functions are given by the same dynamic or functional equations up to a permutation of the relevant variables and/or provided the vertices are isomorphic. This is equivalent to vertices belonging to the same orbits (see Fig. 6.2). We provide the following definitions.

Definition 6.2.2. *1. A colouring of the vertices is balanced if each vertex of colour i gets the same number of inputs from the vertices of colour j, for all i, j.*

2. A minimal balanced colouring is a balanced colouring with minimal number of colours.

Definition 6.2.3. *A quotient network with k-colouring is balanced if the colouring is balanced.*

The existence of synchronous equilibrium clusters yields a decomposition of the network into its support-disjoint subsets of the vertices as defined in the previous section. If this decomposition is invariant under the action of the vector-field, then the corresponding subspace is invariant, and the system can be reduced to synchronous equilibrium clusters. Hence, the dimension of the original system is reduced to a smaller system. Using these concepts, the notion of persistent strain set $\mathcal{P}$ discussed in the introduction section is redefined as

$$\mathcal{P}_Q := \begin{cases} \{i \in \{1, 2, \ldots, n_{\#}(Q)\} : \mathcal{R}_0^i > 1\}, & \mathcal{R}_0 > 1 \\ \varnothing, & \mathcal{R}_0 < 1 \end{cases} \tag{6.11}$$

where $n_{\#}(Q)$ is the number of vertices in the quotient network Q, and $n_{\#}(Q) \leq \mathcal{N}$. We note that $n_{\#}(Q) = \mathcal{N}$ if the parent network G has a trivial automorphism group (completely asymmetric), where the network is completely heterogeneous and all the vertices play a unique structural role. However, if the parent network admits non-trivial automorphism groups (i.e., it is symmetric), $n_{\#}(Q) < \mathcal{N}$. In addition, it should be noted that quotient networks are heterogeneous because they contain only non-equivalent elements, and all vertices play different roles. Using the notion of quotient network, the dynamic equations,

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) s_i(t,a) = -s_i \sum_j \Omega_{ij}\lambda_j, \tag{6.12}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) v_i(t,a) = s_i \sum_{j \neq i} \Omega_{ij}\lambda_j - v_i\lambda_i, \tag{6.13}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) y_i(t,a) = \left[s_i + (1-\gamma(a))v_i\right]\lambda_i - \sigma(a)y_i, \tag{6.14}$$

can be reduced to the quotient network. We note that $\Omega = \mathbb{I} + \overline{\Omega}$, where $\overline{\Omega}$ is the adjacency matrix for interaction between strains. Hence, using the quotient network framework and quotient of the adjacency matrix $\overline{\Omega}$, Ω can be reduced to its quotient $Q_{\overline{\Omega}}$ with entry elements q_{ij}. Therefore, Ω_{ij} can be replaced by $1+q_{ij}$ (see Biggs, 1974; Godsil & Royle, 2001; Lauri & Scapellato, 2003; Golubitsky et al., 2005; Wang & Golubitsky, 2005; Golubitsky & Stewart, 2006). Then, the resulting reduced dynamic equations are given as follows:

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) s_i(t,a) = -s_i\left(\lambda_i + \sum_j q_{ij}\lambda_j\right), \tag{6.15}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) v_i(t,a) = s_i \sum_j q_{ij}\lambda_j - v_i\lambda_i, \tag{6.16}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) y_i(t,a) = \left[s_i + (1-\gamma(a))v_i\right]\lambda_i - \sigma(a)y_i, \tag{6.17}$$

where q_{ij} is the number of neighbours (i.e. external connections) in Δ_j of a vertex $v \in \Delta_i$ and q_{ii} is the number of neighbours (i.e., self-connection) in Δ_i. In Fig. 6.1, we show possible configurations of a persistent set for a 4-strain system on a square lattice with $\mathbb{D}_4$ symmetry, where either 1, 2, 3, or 4 strains can persist and the notion of a persistent set falls into one of the two notions of endemicity defined previously. It should be noted that it is impossible to study all balanced k-colouring for all possible multiple strain systems (even in the case of multi-locus-allele system where the structure of the strain space is well defined and the symmetry properties are classical).

Instead of studying all possible k-colouring, we focus on the case where we can interchange the notion of endemicity and persistent set. That is, if we use for example a 2-colouring framework where one colour presents a persistent synchronous equilibrium cluster and the other colour denotes extinct clusters, then the endemic equilibrium is weakly endemic and the persistent synchronous equilibrium cluster can still synchronize into sub-clusters commensurable with strain-specific reproductive fitness. Recall that a weakly endemic notion is consistent with the principle of competitive exclusion. In addition, in the strongly endemic case, all synchronous equilibrium clusters are persistent and the nature of synchronicity is more complex. In the case where all strains approach the equilibrium point, the synchronous clustering behaviours are simple, in that all strains synchronize into sub-clusters commensurable to their reproductive successes. However, in the case where the dynamic behaviours are cyclical, the synchronous behaviours of the synchronous equilibrium clusters are far more complex. In this chapter, we focus on the stability issues of the equilibrium under 2-colouring framework where we denote the potential persistent strain set with a black vertex ($\bullet$)

while the extinct strain set is represented by a white vertex ($\circ$). Figure 6.1 shows all possible combinations of synchronous equilibrium clusters, where edges are undirected and represent coupling if strains share at least one allele in common in the multi-locus-allele system. In the case of a multi-locus system on hypergraphs, the quotient network is bi-nodal and is shown in Fig. 6.1. While using the definition of a balanced colour, the quotient network of hypercube (Mendelian graph) generated by a multi-locus-allele strain space is balanced if $q_{11} + q_{12} = q_{21} + q_{22}$. Using Fig. 6.1, we note that all the representations in red panels are unbalanced, and only those in yellow and green are balanced, see Fig. 6.1(b) where the q-values are shown. Here we will only focus on one representation of clustering type, since if one of the clustering types is balanced then all the remaining clusterings of the same type are balanced. To relate these topological parameters of the quotient network to the parameters introduced previously (see Chapter 3), we considered a multi-locus-allele system (i.e., n-locus-m-allele) in hypercube strain space. For example, in a 2-locus-2-allele system (see Fig. 6.1(b)), the quotient network is bi-nodal. Let $\rho \in \mathbb{Z}_+$ denote the number of discordant sets in a cluster Δ_i, and $\alpha_0 \in \mathbb{Z}_+$ be the number of strains sharing alleles with strain j and $\alpha = \alpha_0 + 1$ be the number of strains sharing alleles with j (and including j itself). Hence, $\alpha_0 = m^n - (m-1)^n - 1$. For a 2-locus-2-allele system, each strain shares alleles with $\alpha_0 = 2$ other strains. These parameters are related to the topology of the strain space.

Using the notion of a 2-colouring and the bi-nodal structure of the quotient network for the strain space of interest, the dynamic equations further reduce to,

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) s_i(t,a) = -s_i\left[(1+q_{ii})\lambda_i + q_{ij}\lambda_j\right], \tag{6.18}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) v_i(t,a) = s_i\left[q_{ii}\lambda_i + q_{ij}\lambda_j\right] - v_i\lambda_i, \tag{6.19}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) y_i(t,a) = \left[s_i + (1-\gamma(a))v_i\right]\lambda_i - \sigma(a)y_i, \tag{6.20}$$

for $i \neq j$, and $i, j = 1, 2$. We note that all epidemiological parameters $\beta(a,s)$, $\sigma(a)$, $\mu(a)$ and $\gamma(a)$ are the same for all the strains. The notion of strong endemicity is reduced to $y_i(a) = y_i^*(a) = y^*(a)$ while we have $y_i(a) = y_i^*(a) = y^*(a)$ and $y_j(a) = y_j^*(a) = 0$ for all $i \neq j$ for weak endemicity. Using this reduction of the model to a bi-nodal quotient network, we can now proceed with the stability analysis of the equilibria, in particular, synchronous equilibrium clusters.

6.3 Stability of Equilibria

We now consider the stability of both the disease-free and the endemic equilibria. To study the linear stability, we consider small perturbations from the equilibrium solutions and let $s_i(t,a) = s_i^*(a) + \overline{s}_i(t,a)$, $v_i(t,a) = v_i^*(a) + \overline{v}_i(t,a)$, and $y_i(t,a) = y_i^*(a) + \overline{y}_i(t,a)$. Then Eqs. (6.18)-(6.20) become

$$\begin{aligned}
\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right)\overline{s}_i(t,a) &= -\left(\overline{s}_i + s_i^*\right)\left[(1+q_{ii})\overline{\lambda}_i + q_{ij}\overline{\lambda}_j\right] - \overline{s}_i\left[(1+q_{ii})\lambda_i^* + q_{ij}\lambda_j^*\right], \\
\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right)\overline{v}_i(t,a) &= \left(\overline{s}_i + s_i^*\right)\left[q_{ii}\overline{\lambda}_i + q_{ij}\overline{\lambda}_j\right] - \left(\overline{v}_i + v_i^*\right)\overline{\lambda}_i + \overline{s}_i\left[q_{ii}\lambda_i^* + q_{ij}\lambda_j^*\right] \\
&\quad -\overline{v}_i\lambda_i^*, \\
\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right)\overline{y}_i(t,a) &= \left[\left(\overline{s}_i + s_i^*\right) + (1-\gamma(a))\left(\overline{v}_i + v_i^*\right)\right]\overline{\lambda}_i - \sigma(a)\overline{y}_i \\
&\quad + \left[\overline{s}_i + (1-\gamma(a))\overline{v}_i\right]\lambda_i^*,
\end{aligned}$$

where $\overline{s}_i(t,0) = 0$, $\overline{v}_i(t,0) = 0$, $\overline{y}_i(t,0) = 0$, and

$$\begin{aligned}
\overline{\lambda}_i(t,a) &= \int_0^{a_m} \beta(a,s)\psi(s)\overline{y}_i(t,s)ds, \\
\lambda_i^*(a) &= \int_0^{a_m} \beta(a,s)\psi(s)y_i^*(s)ds.
\end{aligned}$$

Define $\mathcal{X}_n := L^1([0,a_m))^n$, where n is an integer. Now, let the linear operator $\mathcal{A}$ on $\mathcal{X}_6 := L^1([0,a_m))^6$ be defined as before by:

$$\mathcal{A}\left[u\right](a) := \begin{pmatrix} \mathcal{A}_1 & 0 & 0 \\ 0 & \mathcal{A}_2 & 0 \\ 0 & 0 & \mathcal{A}_3 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$$

where $\mathcal{A}_1 := -\mathbb{I}_{2\times 2}(\frac{\partial}{\partial a})$, $\mathcal{A}_2 := -\mathbb{I}_{2\times 2}(\frac{\partial}{\partial a})$ and $\mathcal{A}_3 := -\mathbb{I}_{2\times 2}(\frac{\partial}{\partial a} + \sigma(a))$, and in addition $u = \{u_1, u_2, u_3\}^T \in \mathcal{D}(\mathcal{A})$, and $u_1 = \{u_j^{(1)}\}_{j=1,2}$, $u_j^{(1)} = \overline{s}_j$, $u_2 = \{u_j^{(2)}\}_{j=1,2}$, $u_j^{(2)} = \overline{v}_j$, $u_3 = \{u_j^{(3)}\}_{j=1,2}$, $u_j^{(3)} = \overline{y}_j$. The domain $\mathcal{D}(\mathcal{A})$ is defined as $\mathcal{D}(\mathcal{A}) := \{u \in \mathcal{X}_6 : u_1, u_2, u_3$ are absolutely continuous on $[0,\infty)$ and $u_1(0) = 0_{2\times 1}, u_2(0) = 0_{2\times 1}, u_3(0) = 0_{2\times 1}\}$. By reformulating the above dynamic equations into an abstract semilinear Cauchy evolution equation, we have

$$\frac{du}{dt} = \mathcal{A}u + \mathcal{G}(u), \tag{6.21}$$

where $\mathcal{G}$ is a nonlinear operator, $\mathcal{G} : \mathcal{X}_6 \to \mathcal{X}_6$ and is defined by:

$$\begin{aligned}
\mathcal{G}\left[u\right](a) &:= \begin{pmatrix} \mathcal{G}_1\left[u\right](a) \\ \mathcal{G}_2\left[u\right](a) \\ \mathcal{G}_3\left[u\right](a) \end{pmatrix}, \\
\mathcal{G}_k\left[u\right](a) &:= \begin{pmatrix} \mathcal{G}_1^{(k)}\left[u\right](a) \\ \mathcal{G}_2^{(k)}\left[u\right](a) \end{pmatrix}, \quad k = 1,2,3,
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{G}_i^{(1)}[u](a) &:= -\left(u_i^{(1)}(a)+s_i^*(a)\right)\left[(1+q_{ii})\overline{\lambda}_i[u](a)+q_{ij}\overline{\lambda}_j[u](a)\right] \\
&\quad -u_i^{(1)}(a)\left[(1+q_{ii})\lambda_i^*(a)+q_{ij}\lambda_j^*(a)\right], \\
\mathcal{G}_i^{(2)}[u](a) &:= \left(u_i^{(1)}(a)+s_i^*(a)\right)\left[q_{ii}\overline{\lambda}_i[u](a)+q_{ij}\overline{\lambda}_j[u](a)\right] \\
&\quad -\left(u_i^{(2)}(a)+v_i^*(a)\right)\overline{\lambda}_i[u](a)-u_i^{(2)}(a)\lambda_i^*(a) \\
&\quad +u_i^{(1)}(a)\left[q_{ii}\lambda_i^*(a)+q_{ij}\lambda_j^*(a)\right], \\
\mathcal{G}_i^{(3)}[u](a) &:= \left[\left(u_i^{(1)}(a)+s_i^*(a)\right)+(1-\gamma(a))\left(u_i^{(2)}(a)+v_i^*(a)\right)\right]\overline{\lambda}_i[u](a) \\
&\quad +\left[u_i^{(1)}(a)+(1-\gamma(a))u_i^{(2)}(a)\right]\lambda_i^*(a),\ \ i\neq j\ \ i=1,2
\end{aligned}$$

are components of the column vectors $\mathcal{G}_k,\ k=1,2,3$. The operator $\overline{\lambda}_j[\phi](a):\mathcal{X}_6\to L^1([0,a_m])$ is a linear operator:

$$\overline{\lambda}_j[u](a)=\int_0^{a_m}\beta(a,s)\psi(s)u_j^{(3)}(s)ds,\ j=1,2,$$

and $\overline{\lambda}[u]=\left(\overline{\lambda}_1[u],\overline{\lambda}_2[u]\right)^T$. Then the following properties hold:

Lemma 6.3.1. *The linear operator $\overline{\lambda}[u]:\mathcal{X}_6\to\mathcal{X}_2$ is compact.*

Proof. Let $\mathcal{S}$ be a bounded subset of $\mathcal{X}_6$, and let $F_i:=\overline{\lambda}_i[\mathcal{S}]\subset L^1(0,a_m)$ for each i. From Lemma 5.2.8, the linear operator $\overline{\lambda}_i$ is continuous and bounded, so F is bounded. Hence, the first condition of Riesz-Frechet-Kolmogorov Compactness (i.e., Theorem 5.3.7) is satisfied. It is sufficent to show that the second condition holds. That is, using Assumption 5.3.2 (i.e., $\mathcal{H}_7$ and $\mathcal{H}_8$), we have

$$\begin{aligned}
||\rho_h[\overline{\lambda}_i[u]]-\overline{\lambda}_i[u]|| &= \int_0^{a_m}\left|\overline{\lambda}_i[u](a+h)-\overline{\lambda}_i[u](a)\right|da \\
&= \int_0^{a_m}\left(\int_0^{a_m}|\beta(a+h,s)-\beta(a,s)|\,\psi(s)\left|u_i^{(3)}(s)\right|ds\right)da \\
&< \varepsilon\int_0^{a_m}\left|u_i^{(3)}(s)\right|ds\leq\varepsilon||u||\leq\varepsilon c_0,\ \text{for}\ \ h\ \text{small enough},
\end{aligned}$$

and because $||u||\leq c_0$ for some positive constant c_0 if $u\in\mathcal{S}$. Hence, the second criterion of Theorem 5.3.7 also holds. Therefore, $\lambda_i(\mathcal{S})$ is precompact in $L^1(0,a_m)$. Hence, $\overline{\lambda}[u]$ is a compact operator. This completes the proof. □

Lemma 6.3.2. *The nonlinear operator $\mathcal{G}$ is Fréchet differentiable at 0, and the Fréchet derivative $\mathcal{B}$ is given by*

$$\begin{aligned}\mathcal{B}[u](a) &:= \begin{pmatrix}\mathcal{B}_1[u](a)\\ \mathcal{B}_2[u](a)\\ \mathcal{B}_3[u](a)\end{pmatrix},\\ \mathcal{B}_k[u](a) &:= \begin{pmatrix}\mathcal{B}_1^{(k)}[u](a)\\ \mathcal{B}_2^{(k)}[u](a)\end{pmatrix},\ k=1,2,3\end{aligned}$$

where

$$\begin{aligned}\mathcal{B}_i^{(1)}[u](a) &:= -s_i^*(a)\left[(1+q_{ii})\overline{\lambda}_i[u](a)+q_{ij}\overline{\lambda}_j[u](a)\right]\\ &\quad -u_i^{(1)}(a)\left[(1+q_{ii})\lambda_i^*(a)+q_{ij}\lambda_j^*(a)\right],\\ \mathcal{B}_i^{(2)}[u](a) &:= s_i^*(a)\left[q_{ii}\overline{\lambda}_i[u](a)+q_{ij}\overline{\lambda}_j[u](a)\right]-v_i^*(a)\overline{\lambda}_i[u](a)\\ &\quad +u_i^{(1)}(a)\left[q_{ii}\lambda_i^*(a)+q_{ij}\lambda_j^*(a)\right]-u_i^{(2)}(a)\lambda_i^*(a),\\ \mathcal{B}_i^{(3)}[u](a) &:= \left[s_i^*(a)+(1-\gamma(a))v_i^*(a)\right]\overline{\lambda}_i[u](a)+\left[u_i^{(1)}(a)+(1-\gamma(a))u_i^{(2)}(a)\right]\lambda_i^*(a),\end{aligned}$$

for $i=1,2$, $i\neq j$ and if $i=1$, then $j=2$ and vice versa.

Proof. The proof is similar to Lemma 5.2.9. □

In the remainder of the chapter, we follow the standard approach of Inaba (1990), Iannelli (1995) and Webb (1985).

6.3.1 Stability of Disease-free Equilibrium

Linearizing Eq. (6.22), we have

$$\frac{du}{dt} = \mathcal{A}u+\mathcal{B}u, \tag{6.22}$$

where the linear operator $\mathcal{B}$ is the Fréchet derivative of $\mathcal{G}$ at $u=0$, and is given in Lemma 6.3.2. To study the stability of the disease-free equilibrium, we set $y_i^*(a)=0$ and $\lambda_i^*(a)=0$ for all $i=1,2$ and for all $a\in[0,a_m)$. If $\lambda_i^*(a)=0$, then $v_i^*(a)=0$ and $s_i^*(a)=1$ for all i and for all $a\in[0,a_m)$. Then, $\mathcal{B}$ in Eqs. (6.22) above reduces to

$$\begin{aligned}\mathcal{B}_i^{(1)}[u](a) &:= \left[(1+q_{ii})\overline{\lambda}_i[u](a)+q_{ij}\overline{\lambda}_j[u](a)\right],\\ \mathcal{B}_i^{(2)}[u](a) &:= \left[q_{ii}\overline{\lambda}_i[u](a)+q_{ij}\overline{\lambda}_j[u](a)\right],\\ \mathcal{B}_i^{(3)}[u](a) &:= \overline{\lambda}_i[u](a).\end{aligned}$$

Let $\omega_0(T_{\mathcal{A}+\mathcal{B}}(t))$ denote the growth bound of a semigroup $T_{\mathcal{A}+\mathcal{B}}(t)$, which is defined as $\omega_0=\inf_{t>0}\frac{1}{t}\log\|T_{\mathcal{A}+\mathcal{B}}(t)\|$. We note that the growth bound $\omega_0(T_{\mathcal{A}+\mathcal{B}}(t))$ of $\mathcal{A}+\mathcal{B}$ is related to the spectral bound $b(\mathcal{A}+\mathcal{B})$ of $\mathcal{A}+\mathcal{B}$ by $b(\mathcal{A}+\mathcal{B})\leq\omega_0(\mathcal{A}+\mathcal{B})$, where $b(\mathcal{A}+\mathcal{B})$ is defined in Def. 5.3.20(6). We now state the following lemma:

Lemma 6.3.3. *Let $T_{\mathcal{A}+\mathcal{B}}(t)$ be the C_0-semigroup generated by the perturbed operator $\mathcal{A}+\mathcal{B}$. Then $T_{\mathcal{A}+\mathcal{B}}(t)$ is eventually norm continuous; and hence, the growth bound $\omega_0(\mathcal{A}+\mathcal{B})$ of the semigroup $\{T_{\mathcal{A}+\mathcal{B}}(t)\}_{t\geq 0}$ and the spectral bound $b(\mathcal{A}+\mathcal{B})$ of the generator $\mathcal{A}+\mathcal{B}$ are related by*

$$\omega_0(\mathcal{A}+\mathcal{B}) = b(\mathcal{A}+\mathcal{B}) := \sup\{\mathbb{R}e\zeta : \zeta \in \Sigma(\mathcal{A}+\mathcal{B})\}. \tag{6.23}$$

Proof. From the previous chapter, we note that $\mathcal{A}$ generates a nilpotent translation C_0-semigroup $\{T_{\mathcal{A}}\}_{t\geq 0}$. Hence, it is eventually norm continuous (see Engel & Nagel, 2000, pg. 112). That is, there exists a $t_0 \geq 0$ such that the map $t \to T_{\mathcal{A}}(t)$ is continuous from (t_0, ∞) to the space of bounded linear operators on Banach space $\mathcal{Y} = \mathcal{X}_6$. In addition, we know that from Lemma 6.3.1, the linear, continuous and bounded operator $\lambda[u](a)$ is compact, so that $\mathcal{B}[u]$ is compact. Hence, $T_{\mathcal{A}+\mathcal{B}}(t)$ is eventually norm continuous (see Engel & Nagel, 2000, pg. 166, Proposition 3.1.14). Then, the spectral mapping theorem of eventual norm continuity implies that the operator $\mathcal{A}+\mathcal{B}$ satisfies the spectral determined growth condition $\omega_0(\mathcal{A}+\mathcal{B}) = b(\mathcal{A}+\mathcal{B})$ (see Engel & Nagel, 2000; Theorem 4.3.10 and Corollaries 4.3.11-4.3.12). This completes the proof. □

To determine the stability of the equilibrium solutions, we consider the resolvent operator $R(\zeta, \mathcal{A}+\mathcal{B})$ of $\mathcal{A}+\mathcal{B}$. That is, let $(\zeta - [\mathcal{A}+\mathcal{B}])\, w = x,\ w \in \mathcal{D}(\mathcal{A}),\ x \in \mathcal{X}_{6,+},\ \zeta \in \mathbb{C}$. Then, we have

$$\begin{aligned}
\frac{\partial w_i^{(1)}(a)}{\partial a} &= x_i^{(1)}(a) - \zeta w_i^{(1)}(a) - \left[(1+q_{ii})\overline{\lambda}_i[w](a) + q_{ij}\overline{\lambda}_j[w](a)\right], \\
\frac{\partial w_i^{(2)}(a)}{\partial a} &= x_i^{(2)}(a) - \zeta w_i^{(2)}(a) + \left[q_{ii}\overline{\lambda}_i[w](a) + q_{ij}\overline{\lambda}_j[w](a)\right], \\
\frac{\partial w_i^{(3)}(a)}{\partial a} &= x_i^{(3)}(a) - \zeta w_i^{(3)}(a) + \overline{\lambda}_i[w](a) - \sigma(a) w_i^{(3)}(a).
\end{aligned}$$

Integrating the above equations and making the relevant substitutions (i.e., $s_i^*(a) = 1$ for all $i = 1, 2$ and for all $a \in [0, a_m)$), we obtain

$$\begin{aligned}
w_i^{(1)}(a) &= \int_0^a e^{-\int_b^a \zeta d\tau} \left(x_i^{(1)}(b) - \left[(1+q_{ii})\overline{\lambda}_i[w](b) + q_{ij}\overline{\lambda}_j[w](b)\right]\right) db \\
w_i^{(2)}(a) &= \int_0^a e^{-\int_b^a \zeta d\tau} \left(x_i^{(2)}(b) + \left[q_{ii}\overline{\lambda}_i[w](b) + q_{ij}\overline{\lambda}_j[w](b)\right]\right) db \\
w_i^{(3)}(a) &= \int_0^a e^{-\int_b^a \zeta + \sigma(\tau) d\tau} \left(x_i^{(3)}(b) + \overline{\lambda}_i[w](b)\right) db.
\end{aligned}$$

Substituting $w_i^{(3)}$ into $\overline{\lambda}_i[w](a)$, we have

$$\overline{\lambda}_i[w](a) = \int_0^{a_m} \beta(a,s)\psi(s) w_i^{(3)}(s) ds$$

$$
\begin{aligned}
&= \int_0^{a_m} \beta(a,s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \left(x_i^{(3)}(b) + \overline{\lambda}_i[w](b) \right) db \right) ds \\
&= \mathcal{T}_\zeta[x_i^{(3)}](a) + \mathcal{T}_\zeta[\overline{\lambda}_i](a),
\end{aligned}
$$

where the linear operator $\mathcal{T}_\zeta$ on $L^1_+([0,a_m))$ is defined as:

$$
\begin{aligned}
\mathcal{T}_\zeta[\phi_i](a) &= \int_0^{a_m} R_\zeta(a,s)\phi_i(s)ds \\
R_\zeta(a,s) &= \int_s^{a_m} \beta(a,h)\psi(h)e^{-\int_s^h \zeta+\sigma(\tau)d\tau}dh = \int_s^{a_m} \beta(a,h)\psi(h)\Theta(s,h)e^{-\zeta(h-s)}dh,
\end{aligned}
$$

and $\Theta(s,h) = e^{-\int_s^h \sigma(\tau)d\tau}$. Then,

$$
\overline{\lambda}_i[w](a) = \left(\mathbb{I} - \mathcal{T}_\zeta\right)^{-1} \mathcal{T}_\zeta[x_i^{(3)}](a), \tag{6.24}
$$

provided that the inverse exists. In addition, we note that substituting $\overline{\lambda}_i[w](b)$ into $w_i^{(2)}(a)$ and $w_i^{(3)}(a)$, the equations can be expressed in terms of only $x_i^{(k)}(a)$, $k = 1,2,3$. Let

$$
\begin{aligned}
\mathcal{V}_\sigma^\zeta[\phi_i](a) &= \int_0^a e^{-\int_b^a \zeta+\sigma(\tau)d\tau}\phi_i(b)db, \\
\mathcal{J}_\sigma^\zeta[\phi_i](a) &= \int_0^a e^{-\int_b^a \zeta+\sigma(\tau)d\tau} \left(\mathbb{I} - \mathcal{T}_\zeta\right)^{-1} \mathcal{T}_\zeta[\phi_i](b)db.
\end{aligned}
$$

Therefore, $w_i^{(1)}(a)$, $w_i^{(2)}(a)$ and $w_i^{(3)}(a)$ can be rewritten as

$$
w_i^{(1)}(a) = \mathcal{V}_0^\zeta[x_i^{(1)}](a) - \left[(1+q_{ii})\mathcal{J}_0^\zeta[x_i^{(3)}](a) + q_{ij}\mathcal{J}_0^\zeta[x_j^{(3)}](a) \right], \tag{6.25}
$$

$$
w_i^{(2)}(a) = \mathcal{V}_0^\zeta[x_i^{(2)}](a) + \left[q_{ii}\mathcal{J}_0^\zeta[x_i^{(3)}](a) + q_{ij}\mathcal{J}_0^\zeta[x_j^{(3)}](a) \right], \tag{6.26}
$$

$$
w_i^{(3)}(a) = \mathcal{V}_\sigma^\zeta[x_i^{(3)}](a) + \mathcal{J}_\sigma^\zeta[x_i^{(3)}](a). \tag{6.27}
$$

We shall return to the above equations later.

Theorem 6.3.4. *Let $\Omega_c := [0,a_m]$, $\Omega_c \subset \mathbb{R}$, and consider a continuous (real- or complex-valued) function $K \in C(\Omega_c \times \Omega_c)$. We define the following integral operator,*

$$
A[f](x) := \int_0^{a_m} K(x,y)f(y)dy, \text{ for all } x \in \Omega_c. \tag{6.28}
$$

Then, the integral operator $A : L^1(0,a_m) \to L^1(0,a_m)$ is compact.

Proof. See Appendix D.1.1 □

Lemma 6.3.5. *The linear operator $\mathcal{T}_\zeta$ is bounded for all ζ, positive and non-supporting for real ζ, and compact for all ζ, while the operators $\mathcal{V}_\sigma^\zeta$ and $\mathcal{J}_\sigma^\zeta$ (for $\zeta \in \mathbb{C}\backslash\mathbb{E}$) are compact, where $\mathbb{E} := \{\zeta \in \mathbb{C},\ 1 \in \Sigma(\mathcal{T}_\zeta)\}$.*

Proof. From $\mathcal{H}_7$ and $\mathcal{H}_8$ in Assumption 5.3.2, the proof of boundedness, positivity, non-supporting and compactness of $\mathcal{T}_\zeta$ follows directly from similar arguments in Lemma 5.3.16, Theorem 5.3.17 and Theorem 5.3.19, respectively, with the application of the Riesz-Frechet-Kolmogorov Compactness Theorem 5.3.7. Similarly, we note that $\mathcal{V}_\sigma^\zeta$ is a Volterra operator with complex-valued (or real-valued) continuous exponential kernel. Hence, from Theorem 6.3.4, the Volterra operator $\mathcal{V}_\sigma^\zeta$ is compact. It follows that if $\zeta \in \mathbb{C}\backslash\mathbb{E}$, then $\left(\mathbb{I}-\mathcal{T}_\zeta\right)^{-1}$ exists and is bounded. So $\left(\mathbb{I}-\mathcal{T}_\zeta\right)^{-1}\mathcal{T}_\zeta$ is compact. We also observe that $\mathcal{J}_\sigma^\zeta$ is a composition of a compact operator $\mathcal{V}_\sigma^\zeta$ with a bounded operator $\left(\mathbb{I}-\mathcal{T}_\zeta\right)^{-1}\mathcal{T}_\zeta$. Therefore, $\mathcal{J}_\sigma^\zeta$ is compact. This completes the proof. □

Lemma 6.3.6. *For $\zeta \in \rho(\mathcal{A}+\mathcal{B}) = \mathbb{C}\backslash\Sigma(\mathcal{A}+\mathcal{B})$, the resolvent $R(\zeta,\mathcal{A}+\mathcal{B}) = (\zeta-(\mathcal{A}+\mathcal{B}))^{-1}$ is of the form*

$$R\left(\zeta,\mathcal{A}+\mathcal{B}\right)x \;=\; \begin{pmatrix} \mathbb{I}_{2\times2}\mathcal{V}_0^\zeta & 0_{2\times2} & -(\mathbb{I}_{2\times2}+Q)\mathcal{J}_0^\zeta \\ 0_{2\times2} & \mathbb{I}_{2\times2}\mathcal{V}_0^\zeta & Q\mathcal{J}_0^\zeta \\ 0_{2\times2} & 0_{2\times2} & \mathbb{I}_{2\times2}(\mathcal{V}_\sigma^\zeta+\mathcal{J}_\sigma^\zeta) \end{pmatrix} x,$$

where

$$Q = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix},$$

$\mathbb{I}_{2\times2}$ is a 2×2 identity matrix, $0_{2\times2}$ is a 2×2 zero matrix. Also,

$$\mathcal{V}_\sigma^\zeta[\phi_i]+\mathcal{J}_\sigma^\zeta[\phi_i] \;=\; \int_0^a e^{-\int_b^a \zeta+\sigma(\tau)d\tau}\left(\mathbb{I}-\mathcal{T}_\zeta\right)^{-1}\phi_i(b)db. \tag{6.29}$$

In addition, the resolvent is compact, and

$$\Sigma(\mathcal{A}+\mathcal{B}) \;=\; \Sigma_p(\mathcal{A}+\mathcal{B}) = \{\zeta\in\mathbb{C} : 1\in\Sigma_p(\mathcal{T}_\zeta)\} = \mathbb{E}, \tag{6.30}$$

where $\Sigma(\mathcal{A}+\mathcal{B})$ and $\Sigma_p(\mathcal{A}+\mathcal{B})$ are spectrum and point spectrum of $\mathcal{A}+\mathcal{B}$, respectively.

Proof. Let $R(\zeta,\mathcal{A}+\mathcal{B}) : \mathcal{X}_6 \to \mathcal{X}_6$ be the resolvent of $\mathcal{A}+\mathcal{B}$, and let $(\zeta-[\mathcal{A}+\mathcal{B}])\,w = x,\ w\in\mathcal{D}(\mathcal{A}),\ x\in\mathcal{X}_{6,+},\ \zeta\in\mathbb{C}$. From our previous calculations, we obtain the expression for $w_i^{(k)}(a)$ in Eqs. (6.25)-(6.27), which are uniquely determined if $1\in\rho(\mathcal{T}_\zeta)$. We note that the expression for $w_i^{(3)}$ can be simplified. That is,

$$\begin{aligned} w_i^{(3)}(a) &= \int_0^a e^{-\int_b^a \zeta+\sigma(\tau)d\tau}x_i^{(3)}(b)db + \int_0^a e^{-\int_b^a \zeta+\sigma(\tau)d\tau}\left(\mathbb{I}-\mathcal{T}_\zeta\right)^{-1}\mathcal{T}_\zeta[x_i^{(3)}](b)db \\ &= \int_0^a e^{-\int_b^a \zeta+\sigma(\tau)d\tau}\left(\mathbb{I}-\mathcal{T}_\zeta\right)^{-1}\left(\mathbb{I}-\mathcal{T}_\zeta\right)x_i^{(3)}(b) + \left(\mathbb{I}-\mathcal{T}_\zeta\right)^{-1}\mathcal{T}_\zeta[x_i^{(3)}](b)db \\ &= \int_0^a e^{-\int_b^a \zeta+\sigma(\tau)d\tau}\left(\mathbb{I}-\mathcal{T}_\zeta\right)^{-1}x_i^{(3)}(b)db = \mathcal{V}_\sigma^\zeta[x_i^{(3)}](a)+\mathcal{J}_\sigma^\zeta[x_i^{(3)}](a). \end{aligned}$$

Hence, putting the above expressions into matrix notation, we obtain the expression for the resolvent $R(\zeta, \mathcal{A}+\mathcal{B})x$. Since $\mathcal{J}_\sigma^\zeta$ and $\mathcal{V}_\sigma^\zeta$ are compact from Lemma 6.3.5 and $R(\zeta, \mathcal{A}+\mathcal{B})x$ consists of components with linear combinations of an integral operator with continuous kernel and a compact operator, then $R(\zeta, \mathcal{A}+\mathcal{B})x$ is compact on $\mathcal{X}_6$. Hence, $\mathcal{A}+\mathcal{B}$ has a compact resolvent. In addition, $\mathcal{A}+\mathcal{B}$ generates a semigroup (see Lemma 6.3.3) and is closed. Because the spectra of operators with compact resolvent consist of isolated eigenvalues with finite multiplicities, $\Sigma(\mathcal{A}+\mathcal{B}) = \Sigma_p(\mathcal{A}+\mathcal{B})$, and it follows that $\mathbb{C}\backslash\mathbb{E} \subset \rho(\mathcal{A}+\mathcal{B})$ (see Ch. 7-8 of Kreyszig, 1978; Theorem 6.29 of Kato, 1984, pg. 187). That is, $\Sigma(\mathcal{A}+\mathcal{B}) = \Sigma_p(\mathcal{A}+\mathcal{B}) \subset \mathbb{E}$ (see Kato, 1984, Theorem 6.29). Furthermore, from the compactness of $\mathcal{T}_\zeta$, we know that $\Sigma(\mathcal{T}_\zeta)\backslash\{0\} = \Sigma_p(\mathcal{T}_\zeta)\backslash\{0\}$, and for $\zeta \in \mathbb{E}$, there exists an eigenfunction η_ζ such that $\mathcal{T}_\zeta\eta_\zeta = \eta_\zeta$. Then it follows that,

$$\begin{aligned}
\phi_i^{(1)}(a) &= -\int_0^a e^{-\int_b^a \zeta d\tau} \left[(1+q_{ii})\eta_{i,\zeta} + q_{ij}\eta_{j,\zeta}\right](b)db \\
\phi_i^{(2)}(a) &= \int_0^a e^{-\int_b^a \zeta d\tau} \left[q_{ii}\eta_{i,\zeta} + q_{ij}\eta_{j,\zeta}\right](b)db \\
\phi_i^{(3)}(a) &= \int_0^a e^{-\int_b^a \zeta+\sigma(\tau)d\tau}\eta_{i,\zeta}(b)db,\ i,j=1,2,\ i\neq j,
\end{aligned}$$

are the components of an eigenvector of $\mathcal{A}+\mathcal{B}$ corresponding to the eigenvalue ζ. Then, $\mathbb{E} \subset \Sigma(\mathcal{A}+\mathcal{B}) = \Sigma_p(\mathcal{A}+\mathcal{B})$. Hence, $\Sigma(\mathcal{A}+\mathcal{B}) = \Sigma_p(\mathcal{A}+\mathcal{B}) = \mathbb{E}$. This completes the proof. □

We now consider the asymptotic properties of the spectral radius $r(\mathcal{T}_\zeta)$. However, we first state the following properties.

Lemma 6.3.7. *The following limiting properties hold:*

1. $\lim_{\zeta\to-\infty} r(\mathcal{T}_\zeta) = \infty$,
2. $\lim_{\zeta\to\infty} r(\mathcal{T}_\zeta) = 0$.

Proof. Let $f_\zeta \in \left[L_+^1\right]^*$ be the strictly positive linear functional on $L^1([0, a_m))$ defined as

$$\langle f_\zeta, u\rangle = \int_0^{a_m}\left(\int_b^{a_m} \tilde{s}(\tau)\psi(\tau)\Theta(b,\tau)e^{-\zeta(\tau-b)}d\tau\right)u(b)db, \tag{6.31}$$

where $\int_b^{a_m}\psi(s)\Theta(b,s)ds > 0$ for all $b \in (0, a_m)$, $\tilde{s}(a) := \varepsilon$ if $a \in (a_m - a^*, a_m)$, $a_m \in (0,\infty)$ and $\tilde{s}(a) := 0$ otherwise, and $\beta(a,s) \geq \tilde{s}(s)$ for all $(a,s) \in (0,a_m)\times(0,a_m)$ (see $\mathcal{H}_8$ in Assumption 5.3.2). In addition, we observe that $\Theta(b,\tau) = e^{-\int_b^\tau \sigma(s)ds} \geq$

$e^{-\tilde{\sigma}a_m} := c$, for $a_m \geq \tau \geq b \geq 0$. We also observe that when $\zeta = 0$, $f_0 = f$ and $\mathcal{T}_0 = \mathcal{T}$ in Theorem 5.3.17. Because $\beta \geq \tilde{s}$, we note that

$$\mathcal{T}_\zeta[u_i](a) \;\geq\; \langle f_\zeta, u_i\rangle e(a),\ e(a) = 1 \text{ for all } a \in \mathbb{R}_+.$$

Since $\mathcal{T}_\zeta$ is compact (see Lemma 6.3.5) and non-supporting (see Lemma 6.3.5), from Remark 6.3.1, $r(\mathcal{T}_\zeta)$ is an eigenvalue of $\mathcal{T}_\zeta^*$ with strictly positive eigenfunctional $F_\zeta \in \left[L^1_+(0, a_m)\right]^* \backslash \{0\}$ satisfying

$$\langle F_\zeta, \mathcal{T}_\zeta[u_i]\rangle \;\geq\; \langle f_\zeta, u_i\rangle\langle F_\zeta, e\rangle > 0.$$

Hence, $r(\mathcal{T}_\zeta)\langle F_\zeta, u_i\rangle \geq \langle f_\zeta, u_i\rangle\langle F_\zeta, e\rangle$. Suppose $u_i(a) = e(a) = 1$ for all $a \in [0, a_m)$, then $r(\mathcal{T}_\zeta) \geq \langle f_\zeta, e\rangle$. Therefore,

$$\begin{aligned}
\lim_{\zeta\to-\infty} r(\mathcal{T}_\zeta) &\geq \lim_{\zeta\to-\infty}\langle f_\zeta, e\rangle = \lim_{\zeta\to-\infty}\int_0^{a_m}\left(\int_b^{a_m}\tilde{s}(\tau)\psi(\tau)\Theta(b,\tau)e^{-\zeta(\tau-b)}d\tau\right)db\\
&\geq \varepsilon c\lim_{\zeta\to-\infty}\int_{a_m-a^*}^{a_m}\psi(\tau)\left(\int_0^{\tau}e^{-\zeta(\tau-b)}db\right)d\tau\\
&\geq \varepsilon c\lim_{\zeta\to-\infty}\int_{a_m-a^*}^{a_m}\psi(\tau)\left(\frac{1-e^{-\zeta\tau}}{\zeta}\right)d\tau\\
&\geq \varepsilon c\lim_{\zeta\to-\infty}\left(\frac{1-e^{-\zeta(a_m-a^*)}}{\zeta}\right)\int_{a_m-a^*}^{a_m}\psi(\tau)d\tau = \infty.
\end{aligned}$$

For the second part of the proof, let

$$\langle g_\zeta, u_i\rangle \;=\; \overline{\beta}\int_0^{a_m}\left(\int_b^{a_m}\psi(\tau)e^{-\zeta(\tau-b)}d\tau\right)u_i(b)db. \tag{6.32}$$

We note that

$$\begin{aligned}
\mathcal{T}_\zeta[u_i](a) &= \int_0^{a_m}\left(\int_b^{a_m}\beta(a,s)\psi(s)\Theta(b,s)e^{-\zeta(s-b)}ds\right)u_i(b)db\\
&\leq \int_0^{a_m}\left(\int_b^{a_m}\beta(a,s)\psi(s)e^{-\zeta(s-b)}ds\right)u_i(b)db\\
&\leq \overline{\beta}\int_0^{a_m}\left(\int_b^{a_m}\psi(s)e^{-\zeta(s-b)}ds\right)u_i(b)db.
\end{aligned}$$

Hence, $\mathcal{T}_\zeta[u_i](a) \leq \langle g_\zeta, u_i\rangle e(a)$, $e(a) = 1$, $a \in \mathbb{R}_+$, $\zeta \in \mathbb{R}$. So from the above calculation, we have

$$\begin{aligned}
\lim_{\zeta\to\infty} r(\mathcal{T}_\zeta) &\leq \lim_{\zeta\to\infty}\langle g_\zeta, e\rangle = \overline{\beta}\lim_{\zeta\to\infty}\int_0^{a_m}\left(\int_b^{a_m}\psi(s)e^{-\zeta(s-b)}ds\right)db\\
&= \overline{\beta}\lim_{\zeta\to\infty}\int_0^{a_m}\psi(s)\left(\int_0^{s}e^{-\zeta(s-b)}db\right)ds\\
&= \lim_{\zeta\to\infty}\overline{\beta}\int_0^{a_m}\psi(s)\left(\frac{1-e^{-\zeta s}}{\zeta}\right)ds \leq \lim_{\zeta\to\infty}\frac{\overline{\beta}}{\zeta}\int_0^{a_m}\psi(s)ds = 0.
\end{aligned}$$

This completes the proof. □

Lemma 6.3.8. *The spectral radius of the next-generation operator $r(\mathcal{T}_\zeta)$ is continuous and strictly decreasing as a function of $\zeta \in \mathbb{R}$. In addition, there exists $\zeta_0 \in \mathbb{R}$ such that $r(\mathcal{T}_{\zeta_0}) = 1$, and $\zeta_0 \in \mathbb{E} = \{\zeta \in \mathbb{C} : 1 \in \Sigma_p(\mathcal{T}_\zeta)\}$.*

Proof. It is a simple matter to show that $\mathcal{T}_\zeta$ is continuous. Because $r(\mathcal{T}_\zeta)$ is in the point spectrum of $\mathcal{T}_\zeta$ and not an accumulation point of $\Sigma(\mathcal{T}_\zeta)$, $r(\mathcal{T}_\zeta)$ is also continuous (see Section 4.3.5, Kato, 1984). From Lemma 6.3.5, we know that $\mathcal{T}_\zeta$ is nonsupporting and compact for $\zeta \in \mathbb{R}$. If $\zeta < \zeta'$, then we have $\mathcal{T}_\zeta \geq \mathcal{T}_{\zeta'}$, and $\mathcal{T}_\zeta \neq \mathcal{T}_{\zeta'}$ and $r(\mathcal{T}_\zeta) > 0$. Hence, from Theorem 5.3.24, $r(\mathcal{T}_\zeta) > r(\mathcal{T}_{\zeta'})$. Therefore, $r(\mathcal{T}_\zeta)$ is strictly decreasing for $\zeta \in \mathbb{R}$. Moreover, from Lemma 6.3.7 and applying the intermediate value theorem, we see that $r(\mathcal{T}_\zeta) = 1$ has a root $\zeta_0 \in \mathbb{R}$. From the compactness and nonsupporting properties of $\mathcal{T}_\zeta$, we can conclude that $1 = r(\mathcal{T}_{\zeta_0}) \in \Sigma_p(\mathcal{T}_{\zeta_0})$. Hence, $\zeta_0 \in \{\zeta \in \mathbb{C} : 1 \in \Sigma_p(\mathcal{T}_\zeta)\} = \mathbb{E}$. This completes the proof. □

Corollary 6.3.9. *There exists a unique $\zeta_0 \in \mathbb{R} \bigcap \{\zeta \in \mathbb{C} : 1 \in \Sigma_p(\mathcal{T}_\zeta)\} := \mathbb{R} \bigcap \mathbb{E}$ such that $r(\mathcal{T}_{\zeta_0}) = 1$ and the following implications hold:*

1. *$\zeta_0 > 0$, if and only if $r(\mathcal{T}_0) = r(\mathcal{T}) > 1$,*
2. *$\zeta_0 = 0$, if and only if $r(\mathcal{T}_0) = r(\mathcal{T}) = 1$,*
3. *$\zeta_0 < 0$, if and only if $r(\mathcal{T}_0) = r(\mathcal{T}) < 1$.*

Proof. The proof follows directly from Lemma 6.3.8 and monotonicity of $r(\mathcal{T}_\zeta)$. The concept of proof is illustrated in Fig. 6.3. We note that when $\zeta = 0$, $\mathcal{T}_0 = \mathcal{T}$ and $\mathcal{R}_0 := r(\mathcal{T}_0) = r(\mathcal{T})$. This completes the proof. □

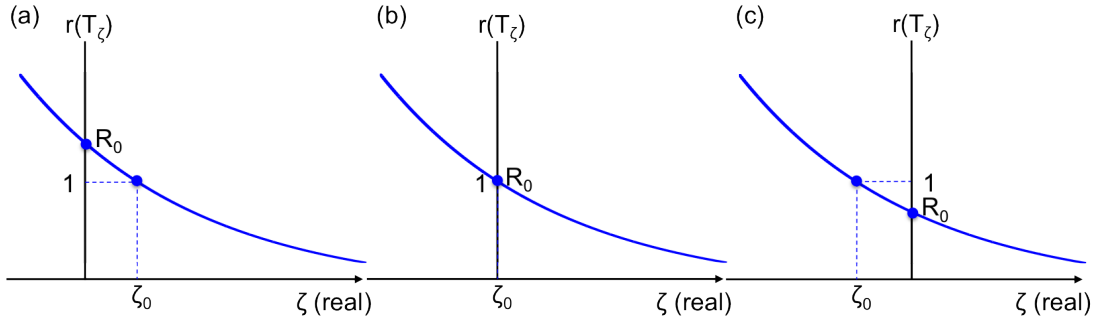


Figure 6.3: **Proof Sketch.** Illustrates the threshold property of $r(\mathcal{T}_\zeta)$ as a function of ζ. For **(a)** $\mathcal{R}_0 > 1$, $\zeta_0 > 0$, **(b)** $\mathcal{R}_0 = 1$, $\zeta_0 = 0$, **(c)** $\mathcal{R}_0 < 1$, $\zeta_0 < 0$.

Lemma 6.3.10. *If there exists a $\zeta \in \{\zeta \in \mathbb{C} : 1 \in \Sigma_p(\mathcal{T}_\zeta)\} = \mathbb{E}$ and $\zeta_0 \neq \zeta$, then $\mathbb{R}e\zeta < \zeta_0$.*

Proof. Suppose that $\zeta_0 \in \mathbb{E},\ 1 \in \Sigma_p(\mathcal{T}_\zeta)$, and $\mathcal{T}_\zeta[\phi](a) = \phi(a)$ for some $\phi \in L^1([0, a_m))$. Let $|\phi|(a) = |\phi(a)|$, then

$$\begin{aligned} |\phi|(a) &= |\mathcal{T}_\zeta[\phi](a)| \leq \int_0^{a_m} \left(\int_b^{a_m} \beta(a,s)\psi(s)\Theta(b,s)|e^{-\zeta(s-b)}|ds \right) \phi(b)db \\ &\leq \int_0^{a_m} \left(\int_b^{a_m} \beta(a,s)\psi(s)\Theta(b,s)e^{-\mathbb{R}e\zeta(s-b)}ds \right) \phi(b)db \\ &= \mathcal{T}_{\mathbb{R}e\zeta}[\phi](a), \text{ for all } a \in [0, a_m). \end{aligned} \tag{6.33}$$

That is, $|\phi|(a) = |\mathcal{T}_\zeta\phi|(a) \leq \mathcal{T}_{\mathbb{R}e\zeta}|\phi|(a)$. Let $F_{\mathbb{R}e\zeta} \in [L^1(0, a_m)]^*$ be a strictly positive eigenfunctional of $\mathcal{T}^*_{\mathbb{R}e\zeta}$. By duality pairing, we have $r(\mathcal{T}_{\mathbb{R}e\zeta})\langle F_{\mathbb{R}e\zeta}, |\phi|\rangle \geq \langle F_{\mathbb{R}e\zeta}, |\phi|\rangle$. Since $< F_{\mathbb{R}e\zeta}, |\phi| >> 0$, we have $r(\mathcal{T}_{\mathbb{R}e\zeta}) \geq 1$. From Lemma 6.3.8, $\mathbb{R}e\zeta \leq \zeta_0$. Suppose $\mathbb{R}e\zeta = \zeta_0$, then $|\phi| \leq \mathcal{T}_{\zeta_0}|\phi|$. If we assume that $|\phi| < \mathcal{T}_{\zeta_0}|\phi|$, then taking duality pairing with the eigenfunctional F_0 of $\mathcal{T}_{\zeta_0}$ corresponding to $r(\mathcal{T}_{\zeta_0}) = 1$ yields

$$\begin{aligned} \langle F_{\zeta_0}, |\phi|\rangle &< \langle F_{\zeta_0}, \mathcal{T}_{\zeta_0}|\phi|\rangle = \langle \mathcal{T}^*_{\zeta_0} F_{\zeta_0}, |\phi|\rangle \\ &= r(\mathcal{T}_{\zeta_0}) < F_{\zeta_0}, |\phi| >= \langle F_{\zeta_0}, |\phi|\rangle, \end{aligned} \tag{6.34}$$

which is a contradiction. Hence, $\mathcal{T}_{\zeta_0}|\phi|(a) = |\phi|(a)$. Now let $\phi_0 \in L^1_+([0, a_m))$ be a nonsupporting eigenvector of $\mathcal{T}_{\zeta_0}$ corresponding to the eigenvalue $r(\mathcal{T}_{\zeta_0}) = 1$, then $|\phi| = c\phi_0$ for some constant $c > 0$. So $\phi(a) = c\phi_0(a)e^{i\kappa(a)}$ for some real-valued function $\kappa : [0, a_m) \to \mathbb{R}$. We observe that $\mathcal{T}_{\zeta_0}\phi_0 = |\mathcal{T}_\zeta\phi|$; and substituting $\phi(a) = c\phi_0(a)e^{i\kappa(a)}$ into $\mathcal{T}_\zeta|\phi| = |\mathcal{T}_\zeta\phi|$, we have $|\mathcal{T}_\zeta\phi(a)| = |\phi|(a) = c\phi_0(a) = c\mathcal{T}_{\zeta_0}\phi_0$. That is, without loss of generality c = 1,

$$\begin{aligned} \mathcal{T}_{\zeta_0}\phi_0(a) &= \int_0^{a_m} \left(\int_b^{a_m} \beta(a,s)\psi(s)\Theta(b,s)e^{-\zeta_0(s-b)}ds \right) \phi_0(b)db \\ &= \left| \int_0^{a_m} \left(\int_b^{a_m} \beta(a,s)\psi(s)\Theta(b,s)e^{-(\zeta_0+i\mathbb{I}m\zeta)(s-b)}ds \right) \phi_0(b)e^{i\kappa(b)}db \right|. \end{aligned}$$

It follows that $f(s,b) = \beta(a,s)\psi(s)\Theta(b,s)e^{-(\zeta_0+i\mathbb{I}m\zeta)(s-b)}\phi_0(b)e^{i\kappa(b)} = f_0(s,b)e^{i\overline{\kappa}}$, for some constant real number $\overline{\kappa} = \kappa(b) - \mathbb{I}m\zeta(s-b)$ and eigenfunction f_0. Then,

$$\begin{aligned} \mathcal{T}_\zeta\phi(a) &= \int_0^{a_m} \left(\int_b^{a_m} \beta(a,s)\psi(s)\Theta(b,s)e^{-\zeta(s-b)}ds \right) \phi_0(b)e^{i\kappa(b)}db \\ &= \int_0^{a_m} \left(\int_b^{a_m} \beta(a,s)\psi(s)\Theta(b,s)e^{-\mathbb{R}e\zeta(s-b)+i\overline{\kappa}}ds \right) \phi_0(b)db \\ &= e^{i\overline{\kappa}}\mathcal{T}_{\mathbb{R}e\zeta}\phi_0(a) = e^{i\overline{\kappa}}\phi_0(a). \end{aligned} \tag{6.35}$$

Hence, $e^{i\kappa}\phi_0(a) = \mathcal{T}_\zeta\phi(a) = e^{i\overline{\kappa}}\phi_0(a)$. Therefore, from $\mathcal{T}_\zeta\phi = \phi$, $e^{\mathbb{I}m\zeta(s-b)} = 1$ for $0 \leq b \leq s \leq a_m$. So $\overline{\kappa} = \kappa(a)$, which implies that $\mathbb{I}m\zeta = 0$. This completes the proof. □

Theorem 6.3.11. *Let $b(\mathcal{A}+\mathcal{B})$ be the spectral bound of $\mathcal{A}+\mathcal{B}$, then $\zeta_0 = b(\mathcal{A}+\mathcal{B})$ so $r(\mathcal{T}_{b(\mathcal{A}+\mathcal{B})}) = 1$.*

Proof. The proof of this theorem is a consequence of the above lemmas and corollaries. That is, from Lemma 6.3.6 and Lemma 6.3.3, $b(\mathcal{A}+\mathcal{B}) = \sup\{\mathbb{R}e\zeta : \zeta \in \mathbb{E}\}$, and from Lemma 6.3.10, we have $\zeta_0 = b(\mathcal{A}+\mathcal{B})$. From Lemma 6.3.8, $r(\mathcal{T}_{\zeta_0}) = 1$. Hence, the assertion holds. This completes the proof. □

Here we define exponential stability of $\mathcal{A}+\mathcal{B}$ in the following equivalent sense: for any $\omega > \omega_0$, there exists M_ω such that for all $t \geq 0$, $\|T_{\mathcal{A}+\mathcal{B}}\| \leq M_\omega e^{\omega t}$. So the semigroup is exponentially stable if $\omega_0 < 0$. Throughout this thesis, we will use locally asymptotically stable to mean that the linearized system is stable. Using the linear and global stability notions of Webb (1985) and Inaba (1990), we state the following stability condition.

Theorem 6.3.12. *If $\mathcal{R}_0 := r(\mathcal{T}) < 1$, then the disease-free equilibrium is locally asymptotically stable; and it is unstable when $\mathcal{R}_0 > 1$.*

Proof. We note that, from Lemma 6.3.3 and Theorem 6.3.11, $\zeta_0 = b(\mathcal{A}+\mathcal{B}) = \omega_0(\mathcal{A}+\mathcal{B})$. Then, by Corollary 6.3.9, $r(\mathcal{T}) < 1$ implies that $\zeta_0 < 0$. So $\omega_0(\mathcal{A}+\mathcal{B}) < 0$ and $\|T_{\mathcal{A}+\mathcal{B}}(t)\| \to 0$ as $t \to \infty$. Hence, the disease-free equilibrium is locally asymptotically stable. However, if $r(\mathcal{T}) > 1$, then $\zeta_0 > 0$ and $\omega_0(\mathcal{A}+\mathcal{B}) > 0$, so the disease-free equilibrium is unstable. This completes the proof. □

Corollary 6.3.13. *If $r(\mathcal{T}) < 1$, then the disease-free equilibrium for initial data in Γ is globally asymptotically stable.*

Proof. We note that the original nonlinear Cauchy evolution equation is,

$$\frac{du}{dt} = \mathcal{A}u + \mathcal{F}(u). \tag{6.36}$$

We focus on the y component of u (i.e., $u_3 = u_j^{(3)},\ j = 1, 2$), and observe that, from $u_i^{(1)} + u_i^{(2)} + u_i^{(3)} \in \Gamma,\ u_i^{(1)} + (1-\gamma)u_i^{(2)} \leq u_i^{(1)} + u_i^{(2)} \leq 1$,

$$\frac{du_3}{dt} = \mathcal{A}_3 u + \mathcal{F}_3(u) \tag{6.37}$$

can be integrated to obtain

$$\begin{aligned} u_3(t) &= T_3(t)u_0 + \int_0^t T_3(t-s)\mathcal{F}_3(u(s))ds \\ &\leq T_3(t)u_0 + \int_0^t T_3(t-s)\mathcal{B}_3(u(s))ds = T_{\mathcal{A}_3+\mathcal{B}_3}u_o, \quad u_0 \in \Gamma, \end{aligned} \tag{6.38}$$

where $\mathcal{F}_3[u](a) = (\mathcal{F}_1^3[u](a), \mathcal{F}_2^3[u](a))^T$, $\mathcal{F}_i^3[u](a) = (u_i^{(1)} + (1-\gamma)u_i^{(2)})\lambda_i[u] \leq \lambda_i[u]$ and $\mathcal{B}_3[u](a) = (\mathcal{B}_1^{(3)}[u](a), \mathcal{B}_2^{(3)}[u](a))^T$, $\mathcal{B}_i^{(3)}[u](a) = s_i^*(a)\lambda_i[u] = \lambda_i[u]$, $s_i^*(a) = 1$ for all i. Hence, $0 \leq u_3(t) \leq T_{\mathcal{A}_3+\mathcal{B}_3}(t)u_0$. From Theorem 6.3.12, the semigroup $T_{\mathcal{A}+\mathcal{B}}$ is exponentially stable. So $T_{\mathcal{A}+\mathcal{B}}(t) \to 0$ as $t \to \infty$. Therefore, $T_{\mathcal{A}_3+\mathcal{B}_3}(t) \to 0$ as $t \to \infty$ and $u_3(t) \leq T_{\mathcal{A}_3+\mathcal{B}_3}(t)u_0 \to 0$, where $T_{\mathcal{A}+\mathcal{B}}(t) = (T_{\mathcal{A}_1+\mathcal{B}_1}(t), T_{\mathcal{A}_2+\mathcal{B}_2}(t), T_{\mathcal{A}_3+\mathcal{B}_3}(t))^T \to 0$. So $y_i(t) := u_i^{(3)}(t) \to 0$ as $t \to \infty$, and $\lambda_i \to 0$, $v_i(t) := u_i^{(2)}(t) \to 0$ and $u_i^{(1)} = s_i \to 1$. This completes the proof. □

6.3.2 Stability of Endemic Equilibria

So far, we have established that, when $\mathcal{R}_0 < 1$, the trivial (disease-free) equilibrium is the only biological relevant steady state and is globally asymptotically stable. Moreover, when $\mathcal{R}_0 > 1$, the disease-free equilibrium is unstable, and there is at least one non-trivial solution. Here we establish the stability of the non-trivial solution(s) using the different notions of endemicity. From Eqs. (6.18)-(6.20), the endemic steady state solution is given as:

$$
\begin{aligned}
s_1^*(a) &= e^{-\int_0^a (1+q_{11})\lambda_1^*(\tau)+q_{12}\lambda_2^*(\tau)d\tau} && (6.39)\\
s_2^*(a) &= e^{-\int_0^a (1+q_{22})\lambda_2^*(\tau)+q_{21}\lambda_1^*(\tau)d\tau} && (6.40)\\
v_1^*(a) &= e^{-\int_0^a \lambda_1^*(\tau)d\tau}\left(1 - e^{-\int_0^a q_{11}\lambda_1^*(\tau)+q_{12}\lambda_2^*(\tau)d\tau}\right) && (6.41)\\
v_2^*(a) &= e^{-\int_0^a \lambda_2^*(\tau)d\tau}\left(1 - e^{-\int_0^a q_{22}\lambda_2^*(\tau)+q_{21}\lambda_1^*(\tau)d\tau}\right) && (6.42)\\
y_1^*(a) &= \int_0^a e^{-\int_b^a \sigma(\tau)d\tau}\lambda_1^*(b)\{s_1^*(b) + (1-\gamma(b))v_1^*(b)\}db && (6.43)\\
y_2^*(a) &= \int_0^a e^{-\int_b^a \sigma(\tau)d\tau}\lambda_2^*(b)\{s_2^*(b) + (1-\gamma(b))v_2^*(b)\}db. && (6.44)
\end{aligned}
$$

Let $c_i = \int_0^{a_m} \beta_2(s)\psi(s)y_i^*(s)ds$, $i = 1, 2$, then $\lambda_i^*(a) = c_i\beta_1(a)$ and let $B(a) = \int_0^a \beta_1(\tau)d\tau$. Substituting y_i^* into c_i, we have

$$
\begin{aligned}
c_1 &= \int_0^{a_m} \beta_2(s)\psi(s)\left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau}c_1\beta_1(b)q_1(b)db\right)ds = c_1H_1(c_1, c_2, \mathcal{R}_0),\\
c_2 &= \int_0^{a_m} \beta_2(s)\psi(s)\left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau}c_2\beta_1(b)q_2(b)db\right)ds = c_2H_2(c_1, c_2, \mathcal{R}_0),
\end{aligned}
$$

where

$$
\begin{aligned}
q_1(b) &= e^{-B(b)[(1+q_{11})c_1+q_{12}c_2]} + (1-\gamma(b))\left(e^{-B(b)c_1} - e^{-B(b)[(1+q_{11})c_1+q_{12}c_2]}\right)\\
q_2(b) &= e^{-B(b)[(1+q_{22})c_2+q_{21}c_1]} + (1-\gamma(b))\left(e^{-B(b)c_2} - e^{-B(b)[(1+q_{22})c_2+q_{21}c_1]}\right).
\end{aligned}
$$

We note that $c_i = 0$ for $i = 1, 2$ corresponds to the disease-free equilibrim, and $c_i \neq 0$ for some i corresponds to a weakly endemic equilibrium while $c_i \neq 0$ for all i corresponds to a strongly endemic equilibrium. Moreover, we observe that $H_1(c_1, c_2, \mathcal{R}_0)$

and $H_2(c_1, c_2, \mathcal{R}_0)$ are analytic in $c = (c_1, c_2)^T$ (see Rudin, 1966; Steinberg, 1968). In addition, we note that $H_1(0, 0, \mathcal{R}_0) = H_2(0, 0, \mathcal{R}_0) = \mathcal{R}_0$. Let $M(c_1, c_2, \mathcal{R}_0) = (M_1, M_2)^T(c_1, c_2, \mathcal{R}_0)$, where $M_i = 1 - H_i,\ i = 1, 2$; and let $c = (c_1, c_2)^T$. Note that $M(0, 0, 1) = 0$.

In addition, we observe that,

$$\begin{aligned}
\frac{\partial H_1(0,0,\mathcal{R}_0)}{\partial c_1} &= -\int_0^{a_m} \beta_2(s)\psi(s)\left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau}\beta_1(b)B(b)\left(1+q_{11}\gamma(b)\right)db\right)ds \\
\frac{\partial H_1(0,0,\mathcal{R}_0)}{\partial c_2} &= -\int_0^{a_m} \beta_2(s)\psi(s)\left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau}\beta_1(b)B(b)q_{12}\gamma(b)db\right)ds \\
\frac{\partial H_2(0,0,\mathcal{R}_0)}{\partial c_1} &= -\int_0^{a_m} \beta_2(s)\psi(s)\left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau}\beta_1(b)B(b)q_{21}\gamma(b)db\right)ds \\
\frac{\partial H_2(0,0,\mathcal{R}_0)}{\partial c_2} &= -\int_0^{a_m} \beta_2(s)\psi(s)\left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau}\beta_1(b)B(b)\left(1+q_{22}\gamma(b)\right)db\right)ds.
\end{aligned}$$

Hence, $\dfrac{\partial M_i(0,0,\mathcal{R}_0)}{\partial c_j} = -\dfrac{\partial H_i(0,0,\mathcal{R}_0)}{\partial c_j} \geq 0,\ i, j = 1, 2$. Define the following,

$$m_0(\mathcal{R}_0) = \int_0^{a_m} \beta_2(s)\psi(s)\left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau}\beta_1(b)B(b)db\right)ds, \tag{6.45}$$

$$m_1(\mathcal{R}_0) = \int_0^{a_m} \beta_2(s)\psi(s)\left(\int_0^s e^{-\int_b^s \sigma(\tau)d\tau}\beta_1(b)B(b)\gamma(b)db\right)ds, \tag{6.46}$$

and let $\mathcal{M}(c_1, c_2, \mathcal{R}_0)$ be the Jacobian matrix of $M = (M_1, M_2)^T$. Then,

$$\mathcal{M}(0,0,\mathcal{R}_0) := \begin{pmatrix} m_0(\mathcal{R}_0) + q_{11}m_1(\mathcal{R}_0) & q_{12}m_1(\mathcal{R}_0) \\ q_{21}m_1(\mathcal{R}_0) & m_0(\mathcal{R}_0) + q_{22}m_1(\mathcal{R}_0) \end{pmatrix}. \tag{6.47}$$

Suppose that γ is chosen such that the determinant $\mathcal{DM}(0, 0, 1) = m_0(1)^2 + m_0(1)$ $m_1(1)(q_{11} + q_{22}) + m_1(1)^2(q_{11}q_{22} - q_{12}q_{21})$ is nonzero. So $\mathcal{M}^{-1}(0, 0, 1)$ exists and is bounded. We can now apply the Implicit Function Theorem of Hildebrandt & Graves (see Hildebrandt & Graves, 1927; Zeidler, 1986, Theorem 4.B). However, to deduce that c_1 and c_2 are continuous functions of $\mathcal{R}_0$, it remains to show the continuity of $H_i(c_1, c_2, \mathcal{R}_0)$ for $i = 1, 2$ with respect to all its arguments $(c_1, c_2, \mathcal{R}_0)$. It is an easy matter to show that $H_i(c_1, c_2, \mathcal{R}_0)$ for $i = 1, 2$ is continuous with respect to (c_1, c_2). However, we are not able to determine the continuity of $H_i(c_1, c_2, \mathcal{R}_0)$ with respect to $\mathcal{R}_0$ because $H_i(c_1, c_2, \mathcal{R}_0)$ is not explicitly expressed in terms of $\mathcal{R}_0$ even though we know that $H_i(0, 0, \mathcal{R}_0) = \mathcal{R}_0$. Because the proceeding discussion in this chapter relies on the continuity of $H_i(c_1, c_2, \mathcal{R}_0)$, we state the following result as a conjecture.

Conjecture 6.3.14. *Let* $\overline{H}(c_1, c_2, \mathcal{R}_0) = (H_1(c_1, c_2, \mathcal{R}_0), H_2(c_1, c_2, \mathcal{R}_0))^T : \mathbb{R}^2_+ \times \mathbb{R}_+ \to \mathbb{R}^2_+$. $\overline{H}(c_1, c_2, \mathcal{R}_0)$ *is continuous near* $(c_1, c_2, \mathcal{R}_0) = (0, 0, 1)$.

If the above conjecture is valid, then, by a direct application of the Implicit Function Theorem of Hildebrandt & Graves (see Hildebrandt & Graves, 1927; Zeidler, 1986), we can show that there exists a branching solution when $\mathcal{R}_0 > 1$ for $\mathcal{R}_0$ small or equivalently, we can find a solution $c(\mathcal{R}_0) = (c_1(\mathcal{R}_0), c_2(\mathcal{R}_0))^T$ bifurcating from the trivial solution at $\mathcal{R}_0 = 1$. We now state the following theorem,

Theorem 6.3.15. *If $M(c_1, c_2, \mathcal{R}_0) : \mathbb{R}^2_+ \times \mathbb{R}_+ \to \mathbb{R}^2_+$ satisfies*

1. *$M(0, 0, 1) = 0$ and*
2. *$\mathcal{DM}(0, 0, 1)$ is nonzero (i.e., $\mathcal{M}(0, 0, 1)$ is nonsingular),*

then there is a solution family $c(\mathcal{R}_0)$ such that $M(c_1(\mathcal{R}_0), c_2(\mathcal{R}_0), \mathcal{R}_0) = 0$ for all $\mathcal{R}_0$ near $\mathcal{R}_0 \approx 1$. Furthermore, if Conjecture 6.3.14 holds, $c_1(\mathcal{R}_0)$ and $c_2(\mathcal{R}_0)$ are continuous.

We note that near $c_i = 0,\ i = 1, 2$ and $\mathcal{R}_0 = 1$, the level set of $M(c_1, c_2, \mathcal{R}_0)$ at $(c_1, c_2, \mathcal{R}_0) = (0, 0, 1)$ given by $\{(c_1, c_2, \mathcal{R}_0) \in \mathbb{R}^2_+ \times \mathbb{R}_+ : M(c_1, c_2, \mathcal{R}_0) = 0\}$ is one-dimensional. Moreover, we observe that from the notion of endemicity discussed previously, there are two possible cases:

1. Weakly Endemic: $c_1 > 0$ and $c_2 = 0$ or $c_1 = 0$ and $c_2 > 0$, for balanced quotient.
2. Strongly Endemic: $c_1 > 0$ and $c_2 > 0$.

We note that, for the first case, whenever one type exists the other also exists. However, which equilibrium is selected will depend on the initial condition, and if one of the weakly endemic equilibria is (un)stable, the other weakly endemic state will be (un)stable due to the inherent symmetry. We now state the following:

Lemma 6.3.16. *Let $H_i(c_1, c_2, \mathcal{R}_0),\ i = 1, 2$ be defined as above. If $\mathcal{R}_0 > 1$, then the following holds:*

1. *In the case of balanced quotient (i.e., $q_{11} + q_{12} = q_{22} + q_{21} = \alpha_0$), there exists a unique strongly endemic equilibrium, $c_1 = c_2 = c > 0$ such that $H(c, c, \mathcal{R}_0) = H_1(c, c, \mathcal{R}_0) = H_2(c, c, \mathcal{R}_0) = 1$.*
2. *There exist weakly endemic equilibria, $c_1 > 0$ and $c_2 = 0$, and $c_1 = 0$ and $c_2 > 0$, such $H_1(c_1, 0, \mathcal{R}_0) = 1$ and $H_2(0, c_2, \mathcal{R}_0) = 1$, respectively.*

However, if $\mathcal{R}_0 < 1$, then there is no non-negative non-trivial steady state.

Proof. For the first result, we note that $c_i = c_j = c,\ i \neq j$ for balanced quotient since $c_i,\ i = 1, 2$ is commensurable to the reproductive fitness and all strains have the same epidemiological parameters. To find c such that $H(c, c, \mathcal{R}_0) = H_1(c, c, \mathcal{R}_0) = H_2(c, c, \mathcal{R}_0) = 1$, we observe that $H(c, c, \mathcal{R}_0)$ is differentiable with respect to c and $\dfrac{\partial H(c, c, \mathcal{R}_0)}{\partial c} < 0$ for all $c > 0$. Hence, $H(c, c, \mathcal{R}_0)$ is strictly monotonically decreasing and satisfies the following limiting behaviours,

1. $\lim_{c \to +\infty} H(c, c, \mathcal{R}_0) \to 0$,
2. $\lim_{c \to -\infty} H(c, c, \mathcal{R}_0) \to +\infty$.

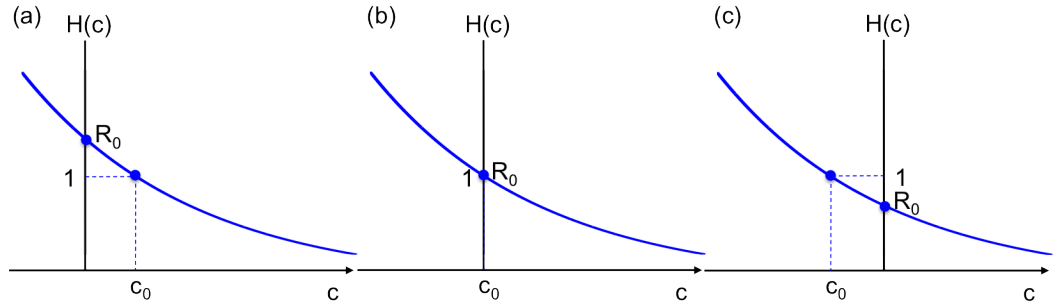


Figure 6.4: **Proof Sketch of Endemic Equilibrium.** Illustrates threshold property of $\mathcal{R}_0$ and $H(c, c, \mathcal{R}_0)$. For **(a)** $\mathcal{R}_0 > 1$, $c_0 > 0$, **(b)** $\mathcal{R}_0 = 1$, $c_0 = 0$, **(c)** $\mathcal{R}_0 < 1$, $c_0 < 0$.

Using similar arguments as in Corollary 6.3.9 and the Intermediate Value Theorem, the result follows (see Fig. 6.4). Hence, when $\mathcal{R}_0 > 1$, there is a unique solution for $H(c, c, \mathcal{R}_0) = 1$, and c is positive. When $\mathcal{R}_0 < 1$, there is no solution on the right-hand half plane. The characteristic equation $H(c, c, \mathcal{R}_0) = 1$ has a negative solution which is biologically irrelevant. For the second part, we observe that $\dfrac{\partial H_1(c_1, 0, \mathcal{R}_0)}{\partial c_1} < 0$ and $\dfrac{\partial H_2(0, c_2, \mathcal{R}_0)}{\partial c_2} < 0$ hold for all c_1 and c_2, respectively, and $H_1(c_1, 0, \mathcal{R}_0)$ and $H_2(0, c_2, \mathcal{R}_0)$ are both monotonically decreasing functions. Hence, the results follow from the Intermediate Value Theorem. This completes the proof. □

In the next sections, we investigate the asymptotic behaviours of these equilibria. We will assume $\mathcal{R}_0 > 1$ but sufficiently close to 1 in the remainder of this chapter.

6.3.2.1 Stability of Weakly Endemic Equilibria

To study the stability of an endemic equilibrium, we start with the linearized equation as in the previous section, and consider the resolvent of $\mathcal{A} + \mathcal{B}$ (i.e. consider the

equation $(\zeta - (\mathcal{A}+\mathcal{B}))\, w = x$). Then, we have for $i,j = 1,2,\ i \neq j$,

$$\begin{aligned}
\frac{\partial w_i^{(1)}(a)}{\partial a} &= x_i^{(1)}(a) - \zeta w_i^{(1)}(a) + \mathcal{B}_i^{(1)}[w](a), \\
\frac{\partial w_i^{(2)}(a)}{\partial a} &= x_i^{(2)}(a) - \zeta w_i^{(2)}(a) + \mathcal{B}_i^{(2)}[w](a), \\
\frac{\partial w_i^{(3)}(a)}{\partial a} &= x_i^{(3)}(a) - \zeta w_i^{(3)}(a) + \mathcal{B}_i^{(3)}[w](a) - \sigma(a) w_i^{(3)}(a),\ i = 1,2
\end{aligned}$$

where $\mathcal{B}_i^{(k)}[w](a) = \overline{\mathcal{B}}_i^{(k)}[w](a) + \tilde{\mathcal{B}}_i^{(k)}[w](a),\ k = 1,2,3$, and

$$\begin{aligned}
\overline{\mathcal{B}}_i^{(1)}[w](a) &:= -s_i^*(a)\left[(1+q_{ii})\overline{\lambda}_i[w](a) + q_{ij}\overline{\lambda}_j[w](a)\right], \\
\tilde{\mathcal{B}}_i^{(1)}[w](a) &:= -w_i^{(1)}(a)\left[(1+q_{ii})\lambda_i^* + q_{ij}\lambda_j^*\right](a), \\
\overline{\mathcal{B}}_i^{(2)}[w](a) &:= s_i^*(a)\left[q_{ii}\overline{\lambda}_i[w](a) + q_{ij}\overline{\lambda}_j[w](a)\right] - v_i^*(a)\overline{\lambda}_i[w](a), \\
\tilde{\mathcal{B}}_i^{(2)}[w](a) &:= w_i^{(1)}\left[q_{ii}\lambda_i^* + q_{ij}\lambda_j^*\right](a) - w_i^{(2)}(a)\lambda_i^*(a), \\
\overline{\mathcal{B}}_i^{(3)}[w](a) &:= \left[s_i^*(a) + (1-\gamma(a))v_i^*(a)\right]\overline{\lambda}_i[w](a), \\
\tilde{\mathcal{B}}_i^{(3)}[w](a) &:= \left[w_i^{(1)}(a) + (1-\gamma(a))w_i^{(2)}(a)\right]\lambda_i^*(a),
\end{aligned}$$

for $i,j = 1,2,\ i \neq j$ and $w_i^{(k)}(0) = 0,\ k = 1,2,3$. Here $w = \{w_1, w_2\}^T$, $w_k = \{w^{(k)},\}_{j=1,2}$, for $k = 1,2,3$, and $w_j^{(1)} = \overline{s}_j$, $w_j^{(2)} = \overline{v}_j$ and $w_j^{(3)} = \overline{y}_j$. Let $c_i = \int_0^{a_m} \beta_2(s)\psi(s)y_i^*(s)ds$, $\overline{c}_i = \int_0^{a_m} \beta_2(s)\psi(s)w_i^{(3)}(s)ds$ and $B(a) = \int_0^a \beta_1(\tau)d\tau$. Then $\lambda_i^*(a) = c_i\beta_1(a)$ and $\overline{\lambda}_i[w](a) = \overline{c}_i\beta_1(a)$. To study the stability of the weakly endemic equilibrium, we use the definition of weak endemicity and either set (i) $c_1 \neq 0$ and $c_2 = 0$ or (ii) $c_1 = 0$ and $c_2 \neq 0$. Here we focus on (i), for the stability of (ii) follows directly from (i) and the analyses of the two cases are identical. If $c_1 \neq 0$ and $c_2 = 0$, substituting the relevant terms in the above expression, we obtain the following,

$$\begin{aligned}
\frac{\partial w_1^{(1)}(a)}{\partial a} &= x_1^{(1)}(a) - \zeta w_1^{(1)}(a) - s_1^*(a)\left[(1+q_{11})\overline{\lambda}_1[w](a) + q_{12}\overline{\lambda}_2[w](a)\right] \\
&\quad - w_1^{(1)}(a)(1+q_{11})\lambda_1^*(a), \\
\frac{\partial w_2^{(1)}(a)}{\partial a} &= x_2^{(1)}(a) - \zeta w_2^{(1)}(a) - s_2^*(a)\left[(1+q_{22})\overline{\lambda}_2[w](a) + q_{21}\overline{\lambda}_1[w](a)\right] \\
&\quad - w_2^{(1)}(a)q_{21}\lambda_1^*(a), \\
\frac{\partial w_1^{(2)}(a)}{\partial a} &= x_1^{(2)}(a) - \zeta w_1^{(2)}(a) + s_1^*(a)\left[q_{11}\overline{\lambda}_1[w](a) + q_{12}\overline{\lambda}_2[w](a)\right] - v_1^*(a)\overline{\lambda}_1[w](a) \\
&\quad + w_1^{(1)}(a)\left[q_{11}\lambda_1^*\right](a) - w_1^{(2)}(a)\lambda_1^*(a), \\
\frac{\partial w_2^{(2)}(a)}{\partial a} &= x_2^{(2)}(a) - \zeta w_2^{(2)}(a) + s_2^*(a)\left[q_{22}\overline{\lambda}_2[w](a) + q_{21}\overline{\lambda}_1[w](a)\right] - v_2^*(a)\overline{\lambda}_2[w](a) \\
&\quad + w_2^{(1)}(a)q_{21}\lambda_1^*(a),
\end{aligned}$$

$$\begin{aligned}
\frac{\partial w_1^{(3)}(a)}{\partial a} &= x_1^{(3)}(a) - \zeta w_1^{(3)}(a) + [s_1^*(a) + (1-\gamma(a))v_1^*(a)]\,\overline{\lambda}_1[w](a) \\
&\quad + \left[w_1^{(1)}(a) + (1-\gamma(a))w_1^{(2)}(a)\right]\lambda_1^*(a) - \sigma(a)w_1^{(3)}(a), \\
\frac{\partial w_2^{(3)}(a)}{\partial a} &= x_2^{(3)}(a) - \zeta w_2^{(3)}(a) + [s_2^*(a) + (1-\gamma(a))v_2^*(a)]\,\overline{\lambda}_2[w](a) - \sigma(a)w_2^{(3)}(a).
\end{aligned}$$

In addition,

$$\begin{aligned}
\overline{\lambda}_j[w](a) &= \int_0^{a_m} \beta(a,s)\psi(s)w_j^{(3)}(s)ds \\
\lambda_1^*(a) &= \int_0^{a_m} \beta(a,s)\psi(s)y_1^*(s)ds.
\end{aligned}$$

Integrating the above equations, we obtain

$$\begin{aligned}
w_1^{(1)}(a) &= \int_0^a e^{-\int_b^a \zeta+(1+q_{11})\lambda_1^*(\tau)d\tau}\Big(x_1^{(1)}(b) - s_1^*(b)\left[(1+q_{11})\overline{\lambda}_1[w](b)\right. \\
&\quad \left.+q_{12}\overline{\lambda}_2[w](b)\right]\Big)\,db \qquad (6.48) \\
w_2^{(1)}(a) &= \int_0^a e^{-\int_b^a \zeta+q_{21}\lambda_1^*(\tau)d\tau}\Big(x_2^{(1)}(b) - s_2^*(b)\left[(1+q_{22})\overline{\lambda}_2[w](b)\right. \\
&\quad \left.+q_{21}\overline{\lambda}_1[w](b)\right]\Big)\,db \qquad (6.49) \\
w_1^{(2)}(a) &= \int_0^a e^{-\int_b^a \zeta+\lambda_1^*(\tau)d\tau}\Big(x_1^{(2)}(b) + s_1^*(b)\left[q_{11}\overline{\lambda}_1[w](b) + q_{12}\overline{\lambda}_2[w](b)\right] \\
&\quad -v_1^*(b)\overline{\lambda}_1[w](b) + w_1^{(1)}(b)q_{11}\lambda_1^*(b)\Big)\,db \qquad (6.50) \\
w_2^{(2)}(a) &= \int_0^a e^{-\int_b^a \zeta d\tau}\Big(x_2^{(2)}(b) + s_2^*(b)\left[q_{22}\overline{\lambda}_2[w](b) + q_{21}\overline{\lambda}_1[w](b)\right] \\
&\quad -v_2^*(b)\overline{\lambda}_2[w](b) + w_2^{(1)}(b)q_{21}\lambda_1^*(b)\Big)\,db \qquad (6.51) \\
w_1^{(3)}(a) &= \int_0^a e^{-\int_b^a \zeta+\sigma(\tau)d\tau}\Big(x_1^{(3)}(b) + [s_1^*(b) + (1-\gamma(a))v_1^*(b)]\,\overline{\lambda}_1[w](b) \\
&\quad + \left[w_1^{(1)}(b) + (1-\gamma(a))w_1^{(2)}(b)\right]\lambda_1^*(b)\Big)\,db \qquad (6.52) \\
w_2^{(3)}(a) &= \int_0^a e^{-\int_b^a \zeta+\sigma(\tau)d\tau}\Big(x_2^{(3)}(b) + [s_2^*(b) + (1-\gamma(a))v_2^*(b)]\overline{\lambda}_2[w](b)\Big)\,db. \qquad (6.53)
\end{aligned}$$

Recall that $\overline{\lambda}_i[u](a) = \overline{c}_i\beta_1(a)$, so substituting the above expressions into

$$\begin{aligned}
\overline{c}_1 &= \int_0^{a_m} \beta_2(s)\psi(s)w_1^{(3)}(s)ds, \\
\overline{c}_2 &= \int_0^{a_m} \beta_2(s)\psi(s)w_2^{(3)}(s)ds,
\end{aligned}$$

we obtain a two-dimensional matrix, $(\mathcal{X}_1, \mathcal{X}_2)^T = (\mathbb{I} - \mathcal{H}(\zeta, c_1(\mathcal{R}_0), 0))\,(\overline{c}_1, \overline{c}_2)^T$, where $\mathcal{H}(\zeta, c_1(\mathcal{R}_0), 0)$ is a 2×2 matrix with element $\mathcal{H}_{ij}(\zeta, c_1(\mathcal{R}_0), 0)_{1\le i,j\le 2}$. The elements

of $\mathcal{H}_{ij}$ of $\mathcal{H}$ and $\mathcal{X}_i$ $(1 \leq i, j \leq 2)$ are defined as follows

$$
\begin{aligned}
\mathcal{H}_{11}(\zeta, c_1(\mathcal{R}_0), 0) &= \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} Q_1^*(b)\beta_1(b)db \right) ds - \int_0^{a_m} \beta_2(s) \\
&\times \psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \lambda_1^*(b) \left[\int_0^b \beta_1(b')(1+q_{11}) s_1^*(b') \right. \right. \\
&\left. \left. \times e^{-\int_{b'}^b \zeta+(1+q_{11})\lambda_1^*(\tau)d\tau} db' \right] db \right) ds + q_{11} \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s \lambda_1^*(b) \right. \\
&\left. \times e^{-\int_b^s \zeta+\sigma(\tau)d\tau} (1-\gamma(b)) \left[\int_0^b s_1^*(b')\beta_1(b') e^{-\int_{b'}^b \zeta+\lambda_1^*(\tau)d\tau} db' \right] db \right) ds \\
&- \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s \lambda_1^*(b) e^{-\int_b^s \zeta+\sigma(\tau)d\tau} (1-\gamma(b)) \left[\int_0^b v_1^*(b') \right. \right. \\
&\left. \left. \times \beta_1(b') e^{-\int_{b'}^b \zeta+\lambda_1^*(\tau)d\tau} db' \right] db \right) ds - \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s \lambda_1^*(b) \right. \\
&\left. \times e^{-\int_b^s \zeta+\sigma(\tau)d\tau} (1-\gamma(b)) \left[\int_0^b e^{-\int_{b'}^b \zeta+\lambda_1^*(\tau)d\tau} q_{11}\lambda_1^*(b') \right. \right. \\
&\left. \left. \times \left(\int_0^{b'} s_1^*(r)(1+q_{11})\beta_1(r) e^{-\int_r^{b'} \zeta+(1+q_{11})\lambda_1^*(\tau)d\tau} dr \right) db' \right] db \right) ds, \\
\mathcal{H}_{12}(\zeta, c_1(\mathcal{R}_0), 0) &= \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \lambda_1^*(b) \left[\int_0^b s_1^*(b') q_{12} \beta_1(b') \right. \right. \\
&\left. \left. \times e^{-\int_{b'}^b \zeta+(1+q_{11})\lambda_1^*(\tau)d\tau} db' \right] db \right) ds + q_{12} \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s \lambda_1^*(b) \right. \\
&\left. \times e^{-\int_b^s \zeta+\sigma(\tau)d\tau} (1-\gamma(b)) \left[\int_0^b s_1^*(b')\beta_1(b') e^{-\int_{b'}^b \zeta+\lambda_1^*(\tau)d\tau} db' \right] db \right) ds \\
&+ q_{11}q_{12} \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \lambda_1^*(b)(1-\gamma(b)) \left[\int_0^b \lambda_1^*(b') \right. \right. \\
&\left. \left. \times e^{-\int_{b'}^b \zeta+\lambda_1^*(\tau)d\tau} \left(\int_0^{b'} e^{-\int_r^{b'} \zeta+(1+q_{11})\lambda_1^*(\tau)d\tau} s_1^*(r)\beta_1(r)dr \right) db' \right] db \right) ds \\
\mathcal{H}_{21}(\zeta, c_1(\mathcal{R}_0), 0) &= 0 \\
\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0) &= \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} Q_2^*(b)\beta_1(b)db \right) ds \\
\mathcal{X}_1(\zeta, c_1(\mathcal{R}_0), 0) &= \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} x_1^{(3)}(b)db \right) ds + \int_0^{a_m} \beta_2(s)\psi(s) \\
&\times \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \lambda_1^*(b) \left[\int_0^b e^{-\int_{b'}^b \zeta+(1+q_{11})\lambda_1^*(\tau)d\tau} x_1^{(1)}(b')db' \right] db \right) ds \\
&+ \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \lambda_1^*(b)(1-\gamma(b)) \left[\int_0^b x_1^{(2)}(b') \right. \right. \\
&\left. \left. \times e^{-\int_{b'}^b \zeta+\lambda_1^*(\tau)d\tau} db' \right] db \right) ds + \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \right.
\end{aligned}
$$

$$
\begin{aligned}
&\times\lambda_1^*(b)(1-\gamma(b))\left[\int_0^b e^{-\int_{b'}^b \zeta+\lambda_1^*(\tau)d\tau} q_{11}\lambda_1^*(b')\left(\int_0^{b'} x_1^{(1)}(r)\right.\right.\\
&\left.\left.\times e^{-\int_r^{b'}\zeta+(1+q_{11})\lambda_1^*(\tau)d\tau}dr\right)db'\right]db\Big)ds\\
\mathcal{X}_2(\zeta,c_1(\mathcal{R}_0),0) &= \int_0^{a_m}\beta_2(s)\psi(s)\left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau}x_2^{(3)}(b)db\right)ds\\
Q_i^*(b) &= s_i^*(b)+(1-\gamma(b))\,v_i^*(b),\ i=1,2\\
s_1^*(b) &= e^{-(1+q_{11})c_1B(b)}\\
s_2^*(b) &= e^{-q_{21}c_1B(b)}\\
v_1^*(b) &= e^{-c_1B(b)}-e^{-(1+q_{11})c_1B(b)}\\
v_2^*(b) &= 1-e^{-q_{21}c_1B(b)}.
\end{aligned}
$$

Let $\mathcal{M}_1(\zeta,\mathcal{R}_0)=(\mathbb{I}-\mathcal{H}(\zeta,c(\mathcal{R}_0)))$ and the determinant of $\mathcal{M}_1(\zeta,c(\mathcal{R}_0))$ be given by $\mathcal{D}_1\mathcal{M}(\zeta,\mathcal{R}_0)$, where for a weakly endemic equilibrium $c(\mathcal{R}_0)=(c_1(\mathcal{R}_0),0)^T$ (or $c(\mathcal{R}_0)=(0,c_2(\mathcal{R}_0))$). From the structure (i.e. upper block matrix) of $\mathcal{H}(\zeta,c(\mathcal{R}_0))$ (and hence $\mathcal{M}_1(\zeta,\mathcal{R}_0)$), we can calculate the determinant as $\mathcal{D}_1\mathcal{M}(\zeta,\mathcal{R}_0)=(1-\mathcal{H}_{11}(\zeta,c(\mathcal{R}_0)))(1-\mathcal{H}_{22}(\zeta,c(\mathcal{R}_0))$. We note that when $\mathcal{R}_0=1$ and $c_i=0,\ i=1,2$, the determinant $\mathcal{D}_1\mathcal{M}(\zeta,c(\mathcal{R}_0))=0$ determines the eigenvalues of the linearized equations for the disease-free equilibrium. Moreover, we observe that at $\zeta=0$ and $\mathcal{R}_0=1$, $\mathcal{H}_{11}(\zeta=0,c_1(\mathcal{R}_0)=0,0)=\mathcal{H}_{22}(\zeta=0,c_1(\mathcal{R}_0)=0,0)=\mathcal{R}_0$ and $\mathcal{D}_1\mathcal{M}(\zeta,\mathcal{R}_0)=0$. From the previous discussion, the resolvent equation is only solvable when $\mathcal{M}_1(\zeta,\mathcal{R}_0)$ is invertible. That is, if the determinant $\mathcal{D}_1\mathcal{M}(\zeta,\mathcal{R}_0)$ of $\mathcal{M}_1(\zeta,\mathcal{R}_0)$ is nonzero. Therefore, to determine the stability of the weakly endemic equilibrium, we state the following results:

Lemma 6.3.17. *Let $\Sigma^{wee}(\mathcal{A}+\mathcal{B})$ and $\Sigma_p^{wee}(\mathcal{A}+\mathcal{B})$ be the spectrum and point spectrum of $\mathcal{A}+\mathcal{B}$, respectively, for the weakly endemic system. Then, $\Sigma^{wee}(\mathcal{A}+\mathcal{B})=\Sigma_p^{wee}(\mathcal{A}+\mathcal{B}):=\{\zeta\in\mathbb{C}:\mathcal{D}_1\mathcal{M}(\zeta,\mathcal{R}_0)=0\}$.*

Proof. We observe that the resolvent equation is solvable if and only if $\mathcal{M}_1(\zeta,\mathcal{R}_0)$is invertible, i.e., $\mathcal{D}_1\mathcal{M}(\zeta,\mathcal{R}_0)\neq 0$. So $\Sigma^{wee}(\mathcal{A}+\mathcal{B})=\{\zeta\in\mathbb{C}:\mathcal{D}_1\mathcal{M}(\zeta,\mathcal{R}_0)=0\}$. Hence $\rho(\mathcal{A}+\mathcal{B})=\{\zeta\in\mathbb{C}:\mathcal{D}_1\mathcal{M}(\zeta,\mathcal{R}_0)\neq 0\}$. Because $\mathcal{M}_1(\zeta,\mathcal{R}_0)$ is invertible for $\zeta\in\rho(\mathcal{A}+\mathcal{B})$, it can be shown that the resolvent $R^{wee}(\zeta,\mathcal{A}+\mathcal{B})$ is compact for any $\zeta\in\rho(\mathcal{A}+\mathcal{B})$ using similar arguments as in Lemma 6.3.6. That is, we first get w in terms of $\overline{c}_i,\ i=1,2$ and x from Eqs. (6.48)-(6.53) by substituting for $\overline{\lambda}_i(a)=\beta_1(a)\overline{c}_i$. Then, we find $\overline{c}_i$ in terms of $\mathcal{X}_i,\ i=1,2$ from $(\overline{c}_1,\overline{c}_2)^T=\mathcal{M}_1^{-1}(\zeta,\mathcal{R}_0)\,(\mathcal{X}_1,\mathcal{X}_2)^T$ and $\mathcal{X}_i$ in terms of x. We note that $\mathcal{M}_1^{-1}(\zeta,\mathcal{R}_0)$ exists and is bounded. Thus, we observe that, after the relevant substitutions, the right-hand sides of w in Eqs. (6.48)-(6.53) are compositions of compact and bounded operators. Hence, the resolvent is

compact. Therefore, $\Sigma^{wee}(\mathcal{A}+\mathcal{B}) = \Sigma_p^{wee}(\mathcal{A}+\mathcal{B})$ (see Kato, 1984, pg. 187; Webb, 1985, Proposition 4.11). This completes the proof. □

Corollary 6.3.18. *$\mathcal{H}_{11}(\zeta, c_1(\mathcal{R}_0), 0) = 1$ or $\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0) = 1$, if and only if $\zeta \in \Sigma_p^{wee}(\mathcal{A}+\mathcal{B})$.*

Theorem 6.3.19. *Suppose $\mathbb{R}e\zeta < 0$ for all $\zeta \in \Sigma_p^{wee}(\mathcal{A}+\mathcal{B})$. Then the weakly endemic equilibrium point is locally asymptotically stable.*

Proof. We first note that the operator $\mathcal{A}$ generates a nilpotent translation C_0-semigroup $\{T_{\mathcal{A}}(t)\}_{t>0}$, and that the operator $\mathcal{B}$ can be broken up into two operators $\mathcal{B} = \overline{\mathcal{B}} + \tilde{\mathcal{B}}$. Then, we observe that $T_{\mathcal{A}+\tilde{\mathcal{B}}}$ is a semigroup of a stable population with a finite age interval, and hence it is eventually compact (see Inaba, 1988, Proposition 2.5). From Assumption 5.3.2 and Lemma 6.3.1, $\overline{\mathcal{B}}$ is compact. Therefore, the C_0-semigroup $\{T_{\mathcal{A}+\mathcal{B}}(t)\}_{t>0}$ is eventually norm continuous (see Engel & Nagel, 2000, pg. 166, Proposition 3.1.14). Hence, $\omega_0(\mathcal{A}+\mathcal{B}) = b(\mathcal{A}+\mathcal{B}) := \sup\{\mathbb{R}e\zeta : \zeta \in \Sigma^{wee}(\mathcal{A}+\mathcal{B})\}$ (see Engel & Nagel, 2000, pg. 281, Corollaries 4.3.11-4.3.12). So, from Lemma 6.3.17, we have $b(\mathcal{A}+\mathcal{B}) = \sup\{\mathbb{R}e\zeta : \zeta \in \Sigma_p^{wee}(\mathcal{A}+\mathcal{B})\} = \zeta_0$. In addition, by Theorem 2.4.18 in Engel & Nagel (2000, pg. 113), the set $\mathcal{S} = \{\zeta \in \Sigma^{wee}(\mathcal{A}+\mathcal{B}) : \mathbb{R}e\zeta \geq -1\}$ is bounded and consists of isolated eigenvalues with finite multiplicities. Therefore, under the given assumption, it follows that the set $\mathcal{S}$ is finite and $\zeta_0 < 0$. Thus, $\lim_{t\to\infty}||T_{\mathcal{A}+\mathcal{B}}(t)|| \to 0$. Hence, the weakly endemic equilibrium is locally asymptotically stable. This completes the proof. □

We can now consider the concrete conditions under which the weakly endemic equilibrium is stable, i.e., there is no ζ with non-negative real part. We note that the condition $\mathcal{D}_1\mathcal{M}(\zeta, \mathcal{R}_0) = 0$ is equivalent to the following: $\mathcal{H}_{11}(\zeta, c_1(\mathcal{R}_0), 0) = 1$ or $\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0) = 1$. In the proceeding, we characterize $\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0) = 1$ and $\mathcal{H}_{11}(\zeta, c_1(\mathcal{R}_0), 0) = 1$. We state the following properties.

Lemma 6.3.20. *The function $\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0)$ is continuous in ζ and strictly decreasing with respect to $\zeta \in \mathbb{R}$, and satisfies the following limiting behaviours:*

1. $\lim_{\zeta\to\infty} \mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0) \to 0$,
2. $\lim_{\zeta\to-\infty} \mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0) \to +\infty$.

Thus, there exists a unique solution $\zeta_{2,0} \in \mathbb{R}$ of $\mathcal{H}_{22}(\zeta_{2,0}, c_1(\mathcal{R}_0), 0) = 1$ and $\zeta_{2,0}$ is a simple root. Moreover, each of the following implications holds:

1. $\zeta_{2,0} < 0$, *if* $\mathcal{H}_{22}(0, c_1(\mathcal{R}_0), 0) < 1$,
2. $\zeta_{2,0} = 0$, *if* $\mathcal{H}_{22}(0, c_1(\mathcal{R}_0), 0) = 1$,

3. $\zeta_{2,0} > 0$*, if* $\mathcal{H}_{22}(0, c_1(\mathcal{R}_0), 0) > 1$.

Furthermore, any solution $\zeta \neq \zeta_{2,0}$ *of* $\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0) = 1$ *is such that* $\mathbb{R}e\zeta < \zeta_{2,0}$.

Proof. We observe that $\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0)$ is holomorphic in ζ. Hence, it is also continuous. In addition, for real ζ, we observe that $\dfrac{d\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0)}{d\zeta} < 0$, so $\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0)$ is strictly monotonically decreasing. Furthermore, using $\mathcal{H}_2$ in Assumption 5.2.1, we note that

$$\begin{aligned}\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0) &= \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta + \sigma(\tau)d\tau} Q_2^*(b)\beta_1(b)db \right) ds \\ &\leq \overline{\beta} \int_0^{a_m} \left(\int_0^s \psi(s) e^{-\int_b^s \zeta d\tau} db \right) ds \leq \frac{\overline{\beta}}{\zeta} \to 0 \text{ as } \zeta \to \infty\end{aligned}$$

as in Lemma 6.3.7(2). For the second part, we observe that $\Theta(b, \tau) = e^{-\int_b^\tau \sigma(s)ds} \geq e^{-\tilde{\sigma} a_m} := c$ for $a_m \geq s \geq b \geq 0$. Let $\overline{Q}_2 = \sup_{b \in [0, a_m]} Q_2^*(b)$. Then, from $\mathcal{H}_8$ in Assumption 5.3.2,

$$\begin{aligned}\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0) &= \int_0^{a_m} \left(\int_b^{a_m} \beta_2(s)\psi(s) e^{-\int_b^s \zeta + \sigma(\tau)d\tau} Q_2^*(b)\beta_1(b)ds \right) db \\ &\geq \varepsilon c \overline{Q}_2 \int_{a_m - a^*}^{a_m} \psi(s) \left(\int_0^s e^{-\int_b^s \zeta d\tau} db \right) ds \to \infty \text{ as } \zeta \to -\infty\end{aligned}$$

as in Lemma 6.3.7(1).

Let $\zeta_{2,0}$ be a critical eigenvalue. Using similar arguments as in Corollary 6.3.8 (see Fig. 6.3), from the monotonicity of $\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0)$ with respect to ζ, the limiting behaviours and Intermediate Value Theorem, it follows that there is a unique real value $\zeta_{2,0} \in \mathbb{R}$ such that $\mathcal{H}_{22}(\zeta_{2,0}, c_1(\mathcal{R}_0), 0) = 1$. Since, $\dfrac{d\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0)}{d\zeta} < 0$ for all $\zeta \in \mathbb{R}$, it is a simple root. Therefore, the assertions follow.

Moreover, let ζ be a solution different from $\zeta_{2,0}$, then

$$\begin{aligned}&\mathcal{H}_{22}(\zeta_{2,0}, c_1(\mathcal{R}_0), 0) = 1 = \mathbb{R}e \left(\int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta + \sigma(\tau)d\tau} Q_2^*(b)\beta_1(b)db \right) ds \right) \\ &= \left(\int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \mathbb{R}e\zeta + \sigma(\tau)d\tau} \cos(\mathbb{I}m\zeta(s-b)) Q_2^*(b)\beta_1(b)db \right) ds \right) \\ &< \left(\int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \mathbb{R}e\zeta + \sigma(\tau)d\tau} Q_2^*(b)\beta_1(b)db \right) ds \right).\end{aligned}$$

Thus, since $\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0)$ is strictly decreasing function, it follows that $\mathbb{R}e\zeta < \zeta_{2,0}$. This completes the proof. □

We observe that $\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0) > 0$ and $\mathcal{H}_{22}(0,0,0) = \mathcal{R}_0$. Similarly, $\mathcal{H}_{11}(\zeta, c_1(\mathcal{R}_0), 0) > 0$ and $\mathcal{H}_{11}(0,0,0) = \mathcal{R}_0$. In addition, $\mathcal{H}_{11}(\zeta, c_1(\mathcal{R}_0), 0)$ can be expressed as a difference of two functions, namely

$$\begin{aligned}
\mathcal{H}_{11}(\zeta, c_1(\mathcal{R}_0), 0) &= \mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0) - \mathcal{H}_{11}^{(2)}(\zeta, c_1(\mathcal{R}_0), 0), \qquad (6.54)\\
\mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0) &= \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta + \sigma(\tau)d\tau} Q_1^*(b)\beta_1(b)db \right) ds \\
&\quad + \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta + \sigma(\tau)d\tau} \lambda_1^*(b)(1-\gamma(b)) \right. \\
&\quad \left. \times \left[\int_0^b e^{-\int_{b'}^b \zeta + \lambda_1^*(\tau)d\tau} s_1^*(b') q_{11} \beta_1(b') db' \right] db \right) ds, \qquad (6.55)\\
\mathcal{H}_{11}^{(2)}(\zeta, c_1(\mathcal{R}_0), 0) &= \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta + \sigma(\tau)d\tau} \lambda_1^*(b) \left[\int_0^b s_1^*(b')(1+q_{11})\beta_1(b') \right. \right. \\
&\quad \left. \left. \times e^{-\int_{b'}^b \zeta + (1+q_{11})\lambda_1^*(\tau)d\tau} db' \right] db \right) ds + \int_0^{a_m} \beta_2(s)\psi(s) \\
&\quad \times \left(\int_0^s e^{-\int_b^s \zeta + \sigma(\tau)d\tau} \lambda_1^*(b)(1-\gamma(b)) \left[\int_0^b v_1^*(b')\beta_1(b') \right. \right. \\
&\quad \left. \left. \times e^{-\int_{b'}^b \zeta + \lambda_1^*(\tau)d\tau} db' \right] db \right) ds + \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s \lambda_1^*(b) \right. \\
&\quad \times (1-\gamma(b)) e^{-\int_b^s \zeta + \sigma(\tau)d\tau} \left[\int_0^b e^{-\int_{b'}^b \zeta + \lambda_1^*(\tau)d\tau} q_{11} \lambda_1^*(b') \left(\int_0^{b'} s_1^*(r) \right. \right. \\
&\quad \left. \left. \left. \times (1+q_{11})\beta_1(r) e^{-\int_r^{b'} \zeta + (1+q_{11})\lambda_1^*(\tau)d\tau} dr \right) db' \right] db \right) ds. \qquad (6.56)
\end{aligned}$$

Lemma 6.3.21. *Let $\mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0)$ be defined as above. Then $\mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0)$ is continuous in ζ and monotonically decreasing with respect to ζ and satisfies,*

1. $\lim_{\zeta \to \infty} \mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0) \to 0$,
2. $\lim_{\zeta \to -\infty} \mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0) \to \infty$.

Proof. The continuity follows directly from the fact that $\mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0)$ is holomorphic with respect to ζ (see Rudin, 1966; Steinberg, 1968). Moreover, we observe that $\dfrac{\partial \mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0)}{\partial \zeta} < 0$, so $\mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0)$ is strictly monotonically decreasing. In addition, let $\overline{B}_1 = \sup_{a \in [0, a_m]} \int_0^a \beta_1(\tau)d\tau$. Then,

$$\begin{aligned}
\mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0) &= \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta + \sigma(\tau)d\tau} Q_1^*(b)\beta_1(b)db \right) ds \\
&\quad + \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta + \sigma(\tau)d\tau} \lambda_1^*(b)(1-\gamma(b)) \right. \\
&\quad \left. \times \left[\int_0^b e^{-\int_{b'}^b \zeta + \lambda_1^*(\tau)d\tau} s_1^*(b') q_{11} \beta_1(b') db' \right] db \right) ds
\end{aligned}$$

$$\leq \ \overline{\beta}\left(1+q_{11}c_1\overline{B}_1\right)\int_0^{a_m}\left(\int_0^s \psi(s)e^{-\int_b^s \zeta d\tau}db\right)ds \to 0 \ \text{ as } \ \zeta\to\infty$$

as in Lemma 6.3.20(1). Moreover, we observe that

$$\begin{aligned}\mathcal{H}_{11}^{(1)}(\zeta,c_1(\mathcal{R}_0),0) &= \int_0^{a_m}\beta_2(s)\psi(s)\left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau}Q_1^*(b)\beta_1(b)db\right)ds\\ &\quad+\int_0^{a_m}\beta_2(s)\psi(s)\left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau}\lambda_1^*(b)(1-\gamma(b))\right.\\ &\quad\left.\times\left[\int_0^b e^{-\int_{b'}^b \zeta+\lambda_1^*(\tau)d\tau}s_1^*(b')q_{11}\beta_1(b')db'\right]db\right)ds,\\ &\geq \int_0^{a_m}\beta_2(s)\psi(s)\left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau}Q_1^*(b)\beta_1(b)db\right)ds\to\infty\end{aligned}$$

as $\zeta\to-\infty$ as in Lemma 6.3.7(1) and Lemma 6.3.20(2). This completes the proof. □

Corollary 6.3.22. *$\mathcal{H}_{11}(\zeta,c_1(\mathcal{R}_0),0)$ is continuous with respect to ζ, and for ζ real, satisfies the following*

1. *$\mathcal{H}_{11}(\zeta,c_1(\mathcal{R}_0),0)\leq\mathcal{H}_{11}^{(1)}(\zeta,c_1(\mathcal{R}_0),0)$, and*
2. *$\lim_{\zeta\to\infty}\mathcal{H}_{11}(\zeta,c_1(\mathcal{R}_0),0)\to 0$.*

Proof. As in Lemma 6.3.20, we observe that $\mathcal{H}_{11}\left(\zeta,c_1\left(\mathcal{R}_0\right),0\right)$ is holomorphic. Therefore, $\mathcal{H}_{11}(\zeta,c_1(\mathcal{R}_0),0)$ is continuous. Moreover, from Eqs. (6.54)-(6.56), we observe that $\mathcal{H}_{11}(\zeta,c_1(\mathcal{R}_0),0)=\mathcal{H}_{11}^{(1)}(\zeta,c_1(\mathcal{R}_0),0)-\mathcal{H}_{11}^{(2)}(\zeta,c_1(\mathcal{R}_0),0)$. Hence, the inequality holds (i.e. $\mathcal{H}_{11}(\zeta,c_1(\mathcal{R}_0),0)\leq\mathcal{H}_{11}^{(1)}(\zeta,c_1(\mathcal{R}_0),0)$). Thus, from Lemma 6.3.21(1), the result of the second part follows. This completes the proof. □

Lemma 6.3.23. *Let $\mathcal{C}_1=1-\mathcal{H}_{11}^{(1)}(\zeta,c_1(\mathcal{R}_0),0)$. Then, there exists a unique solution $\zeta_1\in\mathbb{R}$ such that $\mathcal{C}_1=0$. In addition,*

1. *If $\mathcal{H}_{11}^{(1)}(0,c_1(\mathcal{R}_0),0)<1$, then $\zeta_1<0$,*
2. *If $\mathcal{H}_{11}^{(1)}(0,c_1(\mathcal{R}_0),0)=1$, then $\zeta_1=0$,*
3. *If $\mathcal{H}_{11}^{(1)}(0,c_1(\mathcal{R}_0),0)>1$, then $\zeta_1>0$.*

Moreover, if $\zeta\in\{\zeta\in\mathbb{C}:\mathcal{C}_1=0\}$ then if $\zeta\neq\zeta_1$, $\mathbb{R}e\zeta<\zeta_1$.

Proof. The proof is similar to Lemma 6.3.20. □

We note that, from the properties stated in Lemma 6.3.21 and Corollary 6.3.22, when $\mathcal{H}_{11}(0,c_1(\mathcal{R}_0),0)>1$, there is a solution $\zeta_0>0$ such that $\mathcal{H}_{11}(\zeta_0,c_1(\mathcal{R}_0),0)=1$. That is, let $\Delta(\zeta)=\mathcal{H}_{11}(\zeta,c_1(\mathcal{R}_0),0)-1=0$. We observe that $\Delta(0)>0$ and $\Delta(\infty)<0$. Then, by the Intermediate Value Theorem, there exists a point $\zeta_0\in(0,\infty)$ such that $\Delta(\zeta_0)=0$ or $\mathcal{H}_{11}(\zeta_0,c_1(\mathcal{R}_0),0)=1$. This corresponds to instability. We summarize the above discussion:

Lemma 6.3.24. *Let $\mathcal{R}_0 > 1$ but sufficiently close to 1. If $\mathcal{H}_{11}(0, c_1(\mathcal{R}_0), 0) > 1$, then the characteristic equation admits a positive root.*

In the proceeding Lemma, we will show that under certain condition, the system does not admit eigenvalues with real positive part. We can use compactness arguments to show that, when $\mathcal{H}_{11}^{(1)}(0, c_1(\mathcal{R}_0), 0) < 1$, the system does not admit positive real eigenvalues. However, it does not exclude the case where the system admits complex eigenvalue whose real part is positive. As a result, the compactness argument is not sufficient to exclude all the cases. Hence, in the proceeding lemma, we apply Rouché's Theorem to show the non-existence of eigenvalues with positive real part when $\mathcal{H}_{11}^{(1)}(0, c_1(\mathcal{R}_0), 0) < 1$, for Rouché's Theorem allows us to determine the number of zeros of a function, including complex roots.

Lemma 6.3.25. *Take $\mathcal{R}_0 > 1$ and close to 1. Suppose $\mathcal{H}_{11}^{(1)}(0, c_1(\mathcal{R}_0), 0) < 1$, so that, for c_1 sufficiently small and positive, $\mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0) = 1$ has a unique real negative solution $\zeta_1(c_1)$. Then $\mathcal{H}_{11}(\zeta, c_1(\mathcal{R}_0), 0) = 1$ has no roots with $\mathbb{R}e\zeta > \zeta^*$, where $\zeta^* = \dfrac{\zeta_1(0)}{2} < 0$.*

Proof. Take $\mathcal{R}_0 > 1$ and close to 1 such that c_1 sufficiently small and positive. Assume $\mathcal{H}_{11}^{(1)}(0, c_1(\mathcal{R}_0), 0) < 1$. From Lemma 6.3.23, $\mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0) = 1$ has a unique real negative solution ζ_1, and all roots are such that $\mathbb{R}e\zeta \leq \zeta_1$. From the Implicit Function Theorem, $\zeta_1 := \zeta_1(c_1)$ is a continuous function of c_1 for sufficiently small c_1. So we take δ such that for $|c_1| < \delta$, we have $\zeta_1(c_1) < \dfrac{\zeta_1(0)}{2}$. Let $\zeta^* = \dfrac{\zeta_1(0)}{2} < 0$. Henceforth, we focus on the existence of roots in the half plane $\mathbb{R}e\zeta \geq \zeta^*$. For $l > 0$, we construct a domain $\mathcal{D}_l$ as shown in Fig. 6.5, where $\mathcal{D}_l = \{\zeta : |\zeta| < l \text{ and } \mathbb{R}e\zeta > \zeta^*\}$.

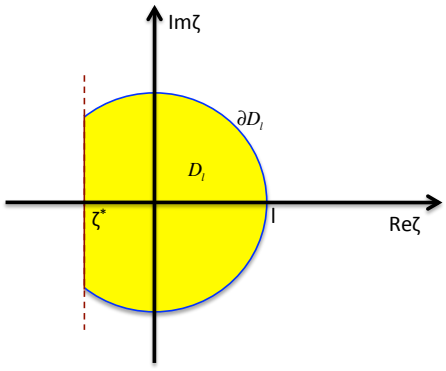


Figure 6.5: **Domain $\mathcal{D}_l$.**

Therefore, in $\mathcal{D}_l$, $\mathbb{R}e\zeta \geq \dfrac{\zeta_1(0)}{2} > \zeta_1(c_1)$. Now let $\Delta(\zeta, c_1(\mathcal{R}_0))$ be defined as follows:

$$\Delta(\zeta, c_1(\mathcal{R}_0)) = \Delta_1(\zeta, c_1(\mathcal{R}_0)) + \Delta_2(\zeta, c_1(\mathcal{R}_0)),$$

where

$$\Delta_1(\zeta, c_1(\mathcal{R}_0)) = 1 - \mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0),$$
$$\Delta_2(\zeta, c_1(\mathcal{R}_0)) = \mathcal{H}_{11}^{(2)}(\zeta, c_1(\mathcal{R}_0), 0).$$

Then, $\Delta_1(\zeta, c_1(\mathcal{R}_0))$ and $\Delta_2(\zeta, c_1(\mathcal{R}_0))$ are entire functions of $\zeta \in \mathbb{C}$. Moreover, we observe that

$$|\Delta_1(\zeta, c_1(\mathcal{R}_0))| = \left|1 - \mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0)\right| \geq 1 - \left|\mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0)\right|.$$

On $\mathcal{D}_l$, we have

$$\left|\mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0)\right| \leq \mathcal{H}_{11}^{(1)}\left(\frac{\zeta_1(0)}{2}, 0, 0\right) + N_0(\zeta_1(0))\,|c_1|,$$

where $N_0(\zeta_1(0))$ is given as follows:

$$N_0(\zeta) \;=\; q_{11}\int_0^{a_m} \beta_2(s)\psi(s)\left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau}\beta_1(b)\left[\int_0^b e^{-\int_{b'}^b \zeta d\tau}\beta_1(b')db'\right]db\right)ds,$$

for real ζ. We note that

$$N_0(\zeta) \leq q_{11}m_0(\mathcal{R}_0)e^{-\zeta_1(0)a_m} \leq q_{11}\overline{B}_1\mathcal{R}_0 e^{-\zeta_1(0)a_m},$$

where $\overline{B}_1 := \sup_{a\in[0,a_m]}\left(\int_0^a \beta_1(b)db\right)$.

If we set $\eta = \mathcal{H}_{11}^{(1)}(\frac{\zeta_1(0)}{2}, 0, 0)$, then $\eta < 1$ because $\mathcal{H}_{11}^{(1)}(\zeta_1(0), 0, 0) = 1$ and $\mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0)$ is decreasing. Therefore, if we choose $|c_1| < \dfrac{1-\eta}{2N_0(\zeta_1(0))}$, then on $\partial\mathcal{D}_l$, we have

$$\left|\mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0)\right| \leq \mathcal{H}_{11}^{(1)}\left(\frac{\zeta_1(0)}{2}, 0, 0\right) + N_0(\zeta_1(0))\,|c_1| \leq \eta + \frac{1-\eta}{2} = \frac{1+\eta}{2},$$

which is less than 1. So, $|\Delta_1(\zeta, c_1(\mathcal{R}_0))| \geq 1 - \dfrac{1+\eta}{2} = \dfrac{1-\eta}{2}$. In addition, we observe that, $\Delta_2(\zeta, 0) = \mathcal{H}_{11}^{(2)}(\zeta, 0, 0) = 0$ and $\Delta_2(\zeta, c_1(\mathcal{R}_0)) = \mathcal{H}_{11}^{(2)}(\zeta, c_1(\mathcal{R}_0), 0)$ satisfies the condition in Lemma 6.3.21 and Lemma 6.3.23. Furthermore, we can show that there exists $M_0(\mathcal{R}_0) = \left(1 + (1+q_{11})(1+\overline{B}_1)\right)m_0$, where $m_0 \leq \overline{B}_1\mathcal{R}_0$, such that $|\Delta_2(\zeta, c_1(\mathcal{R}_0))| = \left|\mathcal{H}_{11}^{(2)}(\zeta, c_1(\mathcal{R}_0), 0)\right| < M_0\,|c_1|$ for c_1 sufficiently small. Hence, if we also take $|c_1| < \dfrac{1-\eta}{2M_0(\mathcal{R}_0)}$, then $|\Delta_2(\zeta, c_1(\mathcal{R}_0))| < \dfrac{1-\eta}{2} \leq |\Delta_1(\zeta, c_1(\mathcal{R}_0))|$ on $\partial\mathcal{D}_l$ for $\mathcal{R}_0 > 1$ and close to 1.

Now we have $|\Delta_1(\zeta, c_1(\mathcal{R}_0))| > |\Delta_2(\zeta, c_1(\mathcal{R}_0))|$ on the boundary $\partial\mathcal{D}_l$ of the contour of any domain $\mathcal{D}_l$ (see Fig. 6.5). Then, by Rouché's Theorem, $\Delta_1(\zeta, c_1(\mathcal{R}_0))$ and

$\Delta(\zeta, c_1(\mathcal{R}_0))$ have the same number of zeros inside the domain $\mathcal{D}_l$ since we know that $|\Delta_1(\zeta, c_1(\mathcal{R}_0))| > |\Delta_2(\zeta, c_1(\mathcal{R}_0))|$ on $\partial\mathcal{D}_l$, and $\Delta_1(\zeta, c_1(\mathcal{R}_0))$ and $\Delta_2(\zeta, c_1(\mathcal{R}_0))$ are holomorphic on $\partial\mathcal{D}_l$ and inside the domain $\mathcal{D}_l$ because they are entire functions of $\zeta \in \mathbb{C}$. Therefore, to show that $\Delta(\zeta, c_1(\mathcal{R}_0)) = 0$ or equivalently $\mathcal{H}_{11}(\zeta, c_1(\mathcal{R}_0), 0) = 1$ has no solution in the right half plane $\mathbb{R}e\zeta \geq \zeta^*$, it suffices to consider the roots of the equation $\Delta_1(\zeta, c_1(\mathcal{R}_0)) = 0$ or equivalently $\mathcal{H}_{11}^{(1)}(\zeta, c_1(\mathcal{R}_0), 0) = 1$. But, from the construction of the domain $\mathcal{D}_l$, $\Delta_1(\zeta, c_1(\mathcal{R}_0)) = 0$ has no solution within $\mathcal{D}_l$. Hence, there are no roots of $\Delta(\zeta, c(\mathcal{R}_0)) = 0$ in $\mathcal{D}_l$. Thus, for $\mathcal{H}_{11}^{(1)}(0, c_1(\mathcal{R}_0), 0) < 1$, there are no zeros of $\Delta(\zeta, c(\mathcal{R}_0)) = 0$ in $\mathcal{D}_l$. But this was for any $l > 0$. Hence, there are no zeros of $\Delta(\zeta, c(\mathcal{R}_0)) = 0$ for $\mathbb{R}e\zeta > \zeta^*$. This completes the proof. □

In the above discussion, we have shown the existence and non-existence of roots (positive real and complex with positive real parts) depending on the conditions imposed on $\mathcal{H}_{11}(0, c_1(\mathcal{R}_0), 0)$ and $\mathcal{H}_{11}^{(1)}(0, c_1(\mathcal{R}_0), 0)$. In particular, we note that the existence and exclusion of positive real roots depends only on $\mathcal{H}_{11}(0, c_1(\mathcal{R}_0), 0)$, while exclusion of complex roots (with positive real part) depends on $\mathcal{H}_{11}(0, c_1(\mathcal{R}_0), 0)$ and $\mathcal{H}_{11}^{(1)}(0, c_1(\mathcal{R}_0), 0)$. Given the above results in Lemma 6.3.24 and Lemma 6.3.25, we establish the condition for the stability of the weakly endemic steady state.

Theorem 6.3.26. *Take $\mathcal{R}_0 > 1$ and close to 1.*

1. *If $\mathcal{H}_{11}^{(1)}(0, c_1(\mathcal{R}_0), 0) < 1$ and $\mathcal{H}_{22}(0, c_1(\mathcal{R}_0), 0) < 1$, then the weakly endemic equilibrium (i.e., discrete strain structure) is locally asymptotically stable.*

2. *If $\mathcal{H}_{11}(0, c_1(\mathcal{R}_0), 0) > 1$ or $\mathcal{H}_{22}(0, c_1(\mathcal{R}_0), 0) > 1$ or both, then the weakly endemic equilibrium is unstable.*

Proof. On one hand, we note that, from Lemma 6.3.25, if $\mathcal{H}_{11}^{(1)}(0, c_1(\mathcal{R}_0), 0) < 1$, the characteristic equation $\mathcal{H}_{11}(\zeta, c_1(\mathcal{R}_0), 0) = 1$ does not admit (real or complex) non-negative characteristic roots. In addition, from Lemma 6.3.20, we observe that the dominant root of $\mathcal{H}_{22}(\zeta, c_1(\mathcal{R}_0), 0) = 1$ is negative when $\mathcal{H}_{22}(0, c_1(\mathcal{R}_0), 0) < 1$. Thus, from Corollary 6.3.18 and Theorem 6.3.19, the weakly endemic equilibrium point is locally asymptotically stable. On the other hand, if one of the conditions is violated, then there is a non-negative eigenvalue $\zeta_0 > 0$, so $\omega_0(\mathcal{A} + \mathcal{B}) > 0$. This completes the proof. □

6.3.2.2 Stability of Strongly Endemic Equilibrium

To study the stability of a strongly endemic equilibrium, we consider the linearized resolvent equation as done in the previous section. We note that, from the definition of strong endemicity, $y_i^*(a) > 0$ for all i. In addition, because the epidemiological

parameters are the same for all strains, $\lambda_i = \lambda_j$ and $y_i = y_j$ for $i \neq j$. Therefore, it is sufficient to look at the asymptotic behaviours of the system as a single strain system provided we take into consideration the balanced condition for the quotient network. That is, $q_{ii} + q_{ij} = q_{jj} + q_{ji}$. Now let $\alpha_0 = q_{ii} + q_{ij} = q_{jj} + q_{ji}$ and $\alpha = \alpha_0 + 1$. Then the system above reduces to

$$\begin{aligned}
\frac{\partial w^{(1)}(a)}{\partial a} &= x^{(1)}(a) - \zeta w^{(1)}(a) - \alpha s^*(a)\overline{\lambda}[w](a) - \alpha w^{(1)}(a)\lambda^*(a), \\
\frac{\partial w^{(2)}(a)}{\partial a} &= x^{(2)}(a) - \zeta w^{(2)}(a) + (\alpha - 1)\, s^*(a)\overline{\lambda}[w](a) - v^*(a)\overline{\lambda}[w](a) \\
&\quad + w^{(1)}(a)\,(\alpha - 1)\,\lambda^*(a) - w^{(2)}(a)\lambda^*(a), \\
\frac{\partial w^{(3)}(a)}{\partial a} &= x^{(3)}(a) - \zeta w^{(3)}(a) + [s^*(a) + (1 - \gamma(a))v^*(a)]\,\overline{\lambda}[w](a) \\
&\quad + \left[w^{(1)}(a) + (1 - \gamma(a))w^{(2)}(a)\right]\lambda^*(a) - \sigma(a)w^{(3)}(a),
\end{aligned}$$

where the steady state solutions $s^*(a)$, $v^*(a)$ and $y^*(a)$ reduce to

$$\begin{aligned}
s^*(a) &= e^{-\int_0^a \alpha\lambda^*(\tau)d\tau}, \\
v^*(a) &= e^{-\int_0^a \lambda^*(\tau)d\tau} - e^{-\int_0^a \alpha\lambda^*(\tau)d\tau}, \\
Q^*(a) &= s^*(a) + (1 - \gamma(a))v^*(a), \\
y^*(a) &= \int_0^a e^{-\int_b^a \sigma(\tau)d\tau}\,(1 - \gamma(b))\,\lambda^*(b)\left[e^{-\int_0^b \lambda^*(\tau)d\tau} - e^{-\int_0^b \alpha\lambda^*(\tau)d\tau}\right] db \\
&\quad + \int_0^a e^{-\int_b^a \sigma(\tau)d\tau}\lambda^*(b)e^{-\int_0^b \lambda^*(\tau)d\tau} db.
\end{aligned}$$

Integrating the resolvent equation above, we have

$$\begin{aligned}
w^{(1)}(a) &= \int_0^a e^{-\int_b^a \zeta + \alpha\lambda^*(\tau)d\tau}\left(x^{(1)} - \alpha s^*(b)\overline{\lambda}[w](b)\right) db, \\
w^{(2)}(a) &= \int_0^a e^{-\int_b^a \zeta + \lambda^*(\tau)d\tau}\left(x^{(2)} + (\alpha - 1)s^*(b)\overline{\lambda}[w](b) - v^*(b)\overline{\lambda}[w](b)\right. \\
&\quad \left. + (\alpha - 1)w^{(1)}(b)\lambda^*(b)\right) db, \\
w^{(3)}(a) &= \int_0^a e^{-\int_b^a \sigma(\tau) + \zeta d\tau}\left(x^{(3)}(b) + q^*(b)\overline{\lambda}[w](b)\right) \\
&\quad + \left[w^{(1)}(b) + (1 - \gamma(b))w^{(2)}(b)\right]\lambda^*(b)\Big) db.
\end{aligned}$$

Let $c = \int_0^{a_m} \beta_2(s)\psi(s)y^*(s)ds$, $\overline{c} = \int_0^{a_m} \beta_2(s)\psi(s)w^{(3)}(s)ds$, and $B(a) = \int_0^a \beta_1(\tau)d\tau$. Then $\lambda^*(a) = c\beta_1(a)$ and $\overline{\lambda}(a) = \overline{c}\beta_1(a)$.

The strongly endemic steady state only exists when $\mathcal{R}_0 > 1$ (see Lemma 6.3.16). To study the stability of the strongly endemic equilibrium, we substitute the above expression for $w_i^{(k)}$, $i = 1, 2$ using strong endemicity into $\overline{c}$. We obtain an equation

of the form $\mathcal{X} = (1 - \mathcal{H}(\zeta, c(\mathcal{R}_0)))\,\overline{c}$, where

$$
\begin{aligned}
\mathcal{H}(\zeta, c(\mathcal{R}_0)) &= \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} Q^*(b)\beta_1(b)db \right) ds \\
&\quad - \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \lambda^*(b) \left[\int_0^b e^{-\int_{b'}^b \zeta+\alpha\lambda^*(\tau)d\tau} \alpha s_1^*(b') \right.\right. \\
&\quad \left.\left. \times \beta_1(b')db' \right] db \right) ds + \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \lambda^*(b) \right. \\
&\quad \left. \times (1-\gamma(b)) \left[\int_0^b e^{-\int_{b'}^b \zeta+\lambda^*(\tau)d\tau} (\alpha-1) s^*(b')\beta_1(b')db' \right] db \right) ds \\
&\quad - \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \lambda^*(b)(1-\gamma(b)) \right. \\
&\quad \left. \times \left[\int_0^b e^{-\int_{b'}^b \zeta+\lambda^*(\tau)d\tau} v^*(b')\beta_1(b')db' \right] db \right) ds - \int_0^{a_m} \beta_2(s)\psi(s) \\
&\quad \times \left[\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \lambda^*(b)(1-\gamma(b)) \left(\int_0^b e^{-\int_{b'}^b \zeta+\lambda^*(\tau)d\tau} (\alpha-1) \right.\right. \\
&\quad \left.\left. \times \lambda^*(b') \left[\int_0^{b'} e^{-\int_\eta^{b'} \zeta+\alpha\lambda^*(\tau)d\tau} \alpha s^*(\eta)\beta_1(\eta)d\eta \right] db' \right) db \right] ds, \\
\mathcal{X} &= \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} x^{(3)}(b)db \right) ds + \int_0^{a_m} \beta_2(s)\psi(s) \\
&\quad \times \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \lambda^*(b) \left[\int_0^b e^{-\int_{b'}^b \zeta+\alpha\lambda^*(\tau)d\tau} x^{(1)}(b')db' \right] db \right) ds \\
&\quad + \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \lambda^*(b)(1-\gamma(b)) \left[\int_0^b x^{(2)}(b') \right.\right. \\
&\quad \left.\left. \times e^{-\int_{b'}^b \zeta+\lambda^*(\tau)d\tau} db' \right] db \right) ds + \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \right. \\
&\quad \left. \times (1-\gamma(b))\lambda^*(b) \left[\int_0^b e^{-\int_{b'}^b \zeta+\lambda^*(\tau)d\tau} (\alpha-1)\lambda^*(b') \left(\int_0^{b'} x^{(1)}(\eta) \right.\right.\right. \\
&\quad \left.\left.\left. \times e^{-\int_\eta^{b'} \zeta+\alpha\lambda^*(\tau)d\tau} d\eta \right) db' \right] db \right) ds.
\end{aligned}
$$

We state the following conditions,

Lemma 6.3.27. *Let $\Sigma^{see}(\mathcal{A}+\mathcal{B})$ and $\Sigma_p^{see}(\mathcal{A}+\mathcal{B})$ be the spectrum and point spectrum of $\mathcal{A}+\mathcal{B}$, respectively, for strongly endemic system, then $\Sigma^{see}(\mathcal{A}+\mathcal{B}) = \Sigma_p^{see}(\mathcal{A}+\mathcal{B}) = \{\zeta \in \mathbb{C} : \mathcal{H}(\zeta, c(\mathcal{R}_0)) = 1\}$.*

Proof. The resolvent equation is solvable if and only if $1 - \mathcal{H}(\zeta, c(\mathcal{R}_0)) \neq 0$. So $\Sigma^{see}(\mathcal{A}+\mathcal{B}) = \{\zeta \in \mathbb{C} : \mathcal{H}(\zeta, c(\mathcal{R}_0)) = 1\}$. Hence, $\rho(\mathcal{A}+\mathcal{B}) = \{\zeta \in \mathbb{C} : \mathcal{H}(\zeta, c(\mathcal{R}_0)) \neq 1\}$. Using similar arguments in Lemma 6.3.17 , if $\zeta \in \rho(\mathcal{A}+\mathcal{B})$, then $\mathcal{A}+\mathcal{B}$ has a

compact resolvent (Kato, 1984, pg. 187; Webb, 1985, Proposition 4.11). Hence, the assertion $\Sigma^{see}(\mathcal{A}+\mathcal{B}) = \Sigma_p^{see}(\mathcal{A}+\mathcal{B})$ holds. This completes the proof. □

Corollary 6.3.28. *$\mathcal{H}(\zeta, c(\mathcal{R}_0)) = 1$, if and only if $\zeta \in \Sigma_p^{see}(\mathcal{A}+\mathcal{B})$.*

Let $\mathcal{H}(\zeta, c(\mathcal{R}_0)) = \mathcal{H}^{(1)}(\zeta, c(\mathcal{R}_0)) - \mathcal{H}^{(2)}(\zeta, c(\mathcal{R}_0))$, where

$$\begin{aligned}
\mathcal{H}^{(1)}(\zeta, c(\mathcal{R}_0)) &= \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} Q^*(b)\beta_1(b)db \right) ds \\
&\quad + \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \lambda^*(b)(1-\gamma(b)) \right. \\
&\quad \left. \times \left[\int_0^b e^{-\int_{b'}^b \zeta+\lambda^*(\tau)d\tau} (\alpha-1)s^*(b')\beta_1(b')db' \right] db \right) ds, \\
\mathcal{H}^{(2)}(\zeta, c(\mathcal{R}_0)) &= \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \lambda^*(b) \left[\int_0^b e^{-\int_{b'}^b \zeta+\alpha\lambda^*(\tau)d\tau} \alpha s_1^*(b') \right. \right. \\
&\quad \left. \left. \times \beta_1(b')db' \right] db \right) ds + \int_0^{a_m} \beta_2(s)\psi(s) \left(\int_0^s \lambda^*(b)(1-\gamma(b)) \right. \\
&\quad \left. \times e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \left[\int_0^b e^{-\int_{b'}^b \zeta+\lambda^*(\tau)d\tau} v^*(b')\beta_1(b')db' \right] db \right) ds + (\alpha-1) \\
&\quad \times \int_0^{a_m} \beta_2(s)\psi(s) \left[\int_0^s e^{-\int_b^s \zeta+\sigma(\tau)d\tau} \lambda^*(b)(1-\gamma(b)) \left(\int_0^b \lambda^*(b') \right. \right. \\
&\quad \left. \left. \times e^{-\int_{b'}^b \zeta+\lambda^*(\tau)d\tau} \left[\int_0^{b'} e^{-\int_\eta^{b'} \zeta+\alpha\lambda^*(\tau)d\tau} \alpha s^*(\eta)\beta_1(\eta)d\eta \right] db' \right) db \right] ds.
\end{aligned}$$

We observe that the functional form of $\mathcal{H}(\zeta, c(\mathcal{R}_0))$ is similar to the functional form of $\mathcal{H}_{11}(\zeta, c_1(\mathcal{R}_0), 0)$. In fact, $\mathcal{H}(\zeta, c(\mathcal{R}_0)) = \mathcal{H}_{11}(\zeta, c(\mathcal{R}_0), 0)$, when $\lambda^*(a) = \lambda_1^*(a)$, and q_{11} is replaced by $\alpha - 1$. Thus, the results and proofs of Lemma 6.3.21, Corollary 6.3.22, Lemma 6.3.23-6.3.25, all carry over to this case and we have:

Theorem 6.3.29. *If $\mathcal{H}^{(1)}(0, c(\mathcal{R}_0)) < 1$, then the strongly endemic equilibrium (i.e., co-existence, fully synchronized steady-state, or no strain structure) is locally asymptotically stable. If $\mathcal{H}(0, c(\mathcal{R}_0)) > 1$, then the strongly endemic equilibrium is unstable.*

6.4 Summary

In this and the previous chapter, we have examined a multi-locus-allele epidemic model describing the dynamics of antigenically variable infectious diseases. We reformulated the model into an abstract semilinear Cauchy evolution equation and investigated the existence of positive solutions. In addition, we employed a fixed point approach to study the steady-states, and the spectral theory of operators to

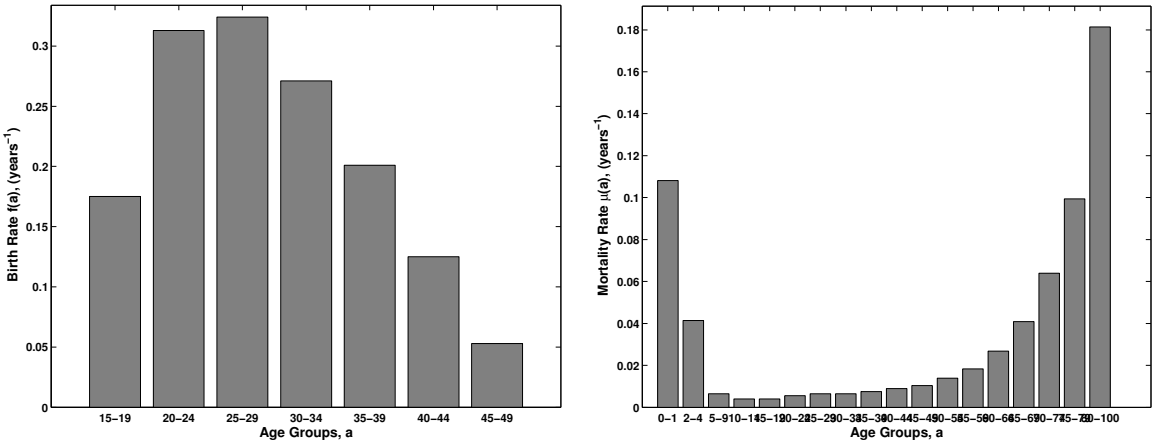
prove the existence of a threshold condition on the next-generation operator. In particular, we showed that if the spectral radius of the next-generation operator $\mathcal{T}$, $r(\mathcal{T})$, is less than 1, then the disease-free equilibrium is the only equilibrium and is globally asymptotically stable. If $r(\mathcal{T}) > 1$, then the disease-free equilibrium is unstable. Moreover, we showed that there is at least one nontrivial equilibrium. However, using the spectral radius $r(\mathcal{T})$ of the next generation operator $\mathcal{T}$, the condition for existence of endemic equilibrium is necessary but not sufficient to establish the stability of the endemic equilibrium. To establish the stability of endemic equilibria, we rely on insights gained from the mesoscale property captured by the quotient network of the strain space and from the analysis presented in Chapter 3. We defined two notions of endemicity (e.g., weakly and strongly endemic equilibria), and used these ideas along with the symmetric structure of the strain space and the uniformity of parameters across strains to simplify our analysis and to provide additional conditions necessary for the stability of the system. It should be noted that the analysis presented herein is based on some assumptions that might be more restrictive for epidemic models. On a fundamental level, we assume a proportionate mixing pattern for the contact rate, a demographically stable population and the uniformity of parameters across strains. The structure of the strain interaction network and its 2-coloured balanced quotient is also restrictive. In particular, we have focused only on a general strain space structure with Mendelian structure or hyperoctahedral symmetry, where the 2-coloured quotient networks are bi-nodal, and the dimension of the system can be reduced. For a more general strain space with no obvious well known symmetry property, we suspect that our analysis can be carried out in a similar manner as long as all the quotient networks can be obtained using tools from algebraic graph theory and all the different combination of synchronous equilibrium clusters can be found. Although, the numerical results of the model described in Chapter 3 show the existence of cyclical behaviours, such dynamics were not investigated herein. This is left as an open problem. The age-specific model described herein can be extended to incorporate different level of cross-immunity including temporary immunity and vaccine-acquired immunity, to name but a few. In Chapters 7-8, we provide some numerical results under different allelic variations.

7 Extended Numerical Investigation of Multi-locus-allele Age-structured Models with Different Allelic Variations

In this chapter, we provide some simulations of the model (Eqs. (4.18)-(4.21)) described in Chapters 4-6 under different allelic variations (e.g., different community structures or strain spaces).

7.1 Introduction

In this section, we discuss the dynamical behaviours of the model under different cross-immunity and allelic assumptions. To investigate the role of different cross-immunity profiles on the dynamics and strain structures of the pathogen, we make some parametric simplifications. Because we do not know *a priori* the exact functional structure of cross-immunity between strains and across the population, we assume that all strains have the same cross-immunity structure, i.e., $\gamma_i(a) = \gamma_j(a) := \gamma(a)$ for $i \neq j$. In addition, we also assume that all epidemiological and demographic parameters are the same for all strains, i.e., $\beta_i(a,s) = \beta_j(a,s) = \beta(a,s)$, $\sigma_i(a) = \sigma_j(a) = \sigma(a)$ and $\mu_i(a) = \mu_j(a) = \mu(a)$ and $f_i(a) = f_j(a) = f(a)$, for $i \neq j$. As a result, the reproductive fitnesses of the serotypes are identical, and are determined by the reproductive number, $\mathcal{R}_0$. Although there is strong empirical evidence that viral populations exhibit variability in their cross-reactivity against other strains due to antigenic variation, we have however assumed that the cross-immunity is identical for all strains and age-independent (i.e., $\gamma_i(a) = \gamma_j(a) := \gamma(a) = \gamma$ for $i \neq j$) for the sake of simplicity. In the proceeding sections, we investigate the effects of cross-immunity structures on the community and strain structures.



(a) (b)

Figure 7.1: **Age-Specific Demographic Patterns.** (**a**) The age-specific fertility and (**b**) mortality rates employed in the numerical simulation studies of the dynamical behaviours of the model. We use the geometric mean of male and female mortality rates. These age-specific distributions represent the demographic patterns of developing countries (e.g., sub-Saharan African countries; data adapted from Anderson et al., 1992).

7.2 Parameter Descriptions

We consider the age-specific model with the following 15 discretized age-cohorts [0-4 years], [5-9 years], [10-14 years], [15-19 years], [20-24 years], [24-29 years], [30-34 years], [35-39 years], [40, 44 years], [45, 49 years], [50, 54 years], [55, 59 years], [60, 64 years], [65, 69 years] and (70+ years], except in Section 7.3. In the proceeding sections, we discuss both demographic and epidemiological parameters.

7.2.1 Demographic Parameters

Most of the model parameters, both demographic and epidemiological, can readily be found in census reports and literature (see Anderson et al., 1992; WHO, 2003; Anderson & May, 1979, 2002; Dushoff et al., 2004). Mortality and fertility moduli are available in census reports and their functional shape varies widely depending on the socio-economic status of the country. That is, in general, developing countries tend to have mortality rates with a *bathtub* shape due to high infant mortality, and high unimodal fertility rates with wider age range. More advanced and developed countries exhibit monotonically increasing mortality and narrowed unimodal fertility rates as a function of age. Figure 7.1 shows representative demographic patterns of mortality and fertility rates for developing countries (e.g., sub-Saharan African countries). The life-expectancy is given as $L = \int_0^\infty \mathcal{E}(a)da$, where $\mathcal{E}(a) = \exp\left(-\int_0^a \mu(\tau)d\tau\right)$ and $L = \dfrac{1}{\mu}$ for constant age-independent μ (e.g., $L = 78$-84 for developed countries).

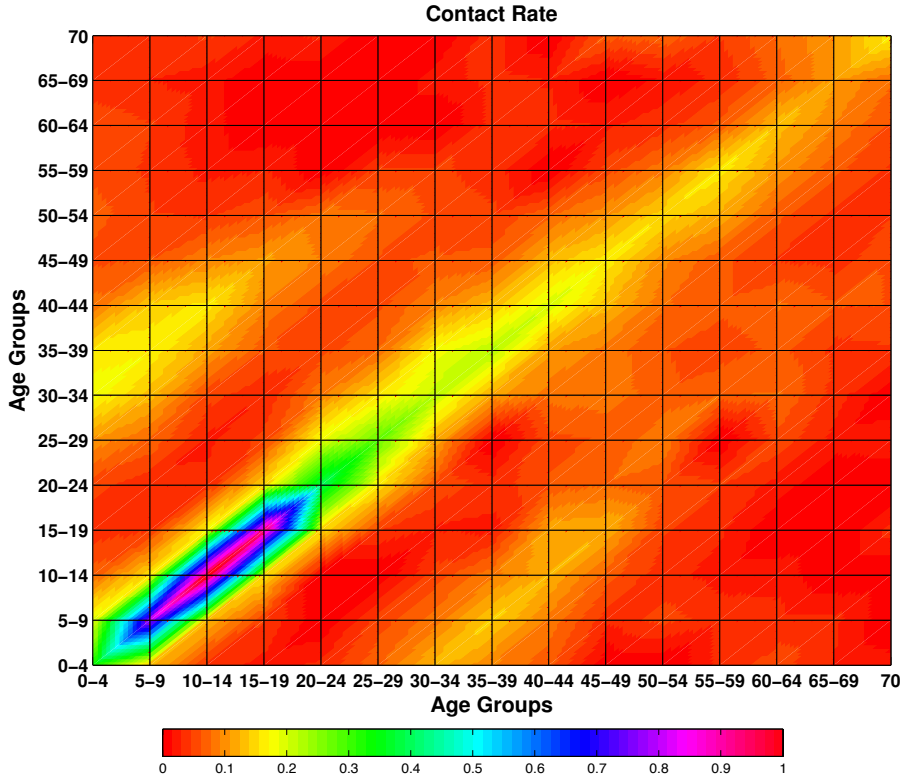


Figure 7.2: **Representative Contact Rate.** Average contact rate of the eight European countries Belgium, Germany, Finland, Great Britain, Italy, Luxembourg, Netherlands and Poland (data taken from Mossong et al., 2008).

7.2.2 Epidemiological Parameters

For epidemiological parameters, we can only provide some rough estimates whenever it is possible. The séjour time (or the mean duration of infection/period of infectiousness) is determined by $D = \int_0^\infty \exp\left(-\int_0^\infty \sigma(\tau)d\tau\right) da$. In order to estimate the séjour time D, we assume that recovery rate $\sigma(a)$ is constant and independent of age due to lack of information (i.e., $D = \frac{1}{\sigma}$). However, this assumption may not hold for immuno-compromised and/or elderly individuals. For a disease-specific value, say influenza, an exposed individual can have a period of infectiousness that ranges from 2 to 7 days (WHO, 2003). Therefore, $D \in (2, 7)$ days, and the recovery rate is between 0.143 days^{-1} and 0.50 days^{-1}.

The infection rate incorporating effective contact rates $\beta(a, s)$, and cross-immunity level are not readily available in scientific reports. These quantities are often estimated from the prevalence or serological data. Since the exact value of the cross-immunity coefficient $\gamma(a)$ is hard to attain, we use γ as an exploratory parameter for our numerical simulations; and assume uniform/homogeneous or constant cross-immunity (i,.e., $\gamma(a) = \gamma$ for all $a \in [0, a_m]$). For the infection rate $\beta(a, s)$, there have been

two main approaches to determining the contact rate. The earliest approach was based on an *ad hoc* assumption on the contact structure that imposes a particular matrix structure (i.e., Who Acquires Infection From Whom, WAIFW, see Anderson & May, 1991). Current approaches emphasize either the use of empirically derived contact structure or the use of the contact surface functional model informed by contact surveys (Farrington & Whitaker, 2005). Here, we parametrize the infection rate by $\beta(a,s) = \beta_0 c(a,s)$ where $c(a,s)$ captures the contact structure and β_0 denotes the disease-specific transmission rate.

We use known disease-specific values of recovery rate along with population-wide unstructured reproductive number $\mathcal{R}_0$ to estimate β. For instance, for influenza, the basic reproductive number ranges from $\mathcal{R}_0 = 1.2$ to $\mathcal{R}_0 = 4$ (Dushoff et al., 2004). However, there have been other estimates that vary widely ($\mathcal{R}_0 = 4 - 20$) (Dushoff et al., 2004; Gog et al., 2003; Gani et al., 2005). For our purpose, we restrict ourselves to $\mathcal{R}_0 = 1.2 - 4$ because the values of recent epidemics of the past decade fall within this range. We note that from the previous chapter, $\mathcal{R}_0 = 7\beta_0$ in the unstructured model because $\mu << \sigma$. Our aim in this section is to investigate the impact of heterogeneity due to age-specificity. In particular, rather than estimating the contact matrix $c(a,s)$, we will use the average of self-reported contact data from a recent large-scale European POLYMOD study (Mossong et al., 2008, and see Farrington & Whitaker, 2005 for a similar approach) that reveals the underlying empirical age-specific contact patterns relevant for infectious diseases in eight European nations as shown in Fig. 7.2. Here, $\beta_0 = 0.49$ days^{-1} is used. The basic reproductive number for the age-structured model can be calculated using the age-discretized next generation method (see the parametrized assumption on the separability of $\beta_{ij} = \beta_0 c_{ij}$ into age-independent disease-relevant (e.g., probability of infection, β_0) and age-specific population contact structure, e.g., c_{ij} below).

7.2.3 Reproductive Numbers, $\mathcal{R}_0$

Using the above parameters, we can calculate the basic reproductive number, $\mathcal{R}_0$, which represents the average number of secondary cases produced by one typical infected person during his/her entire period of infectiousness when introduced into an entirely susceptible population (see Diekmann and Heesterbeek, 2000). In the previous chapter, we showed that $\mathcal{R}_0$ exhibits a threshold property. For $\mathcal{R}_0 < 1$, the infection can invade the population and will eventually go to extinction while for $\mathcal{R}_0 > 1$, the infection can invade the population. Therefore, $\mathcal{R}_0$ captures both the interaction between pathogen and its host, and the epidemic potential because it is influenced by the duration of infection, the contact rate and the probability of infection. It should be noted that, in practice, the reproductive number $\mathcal{R}_0$ is not

observed. In the previous chapter, we showed that the basic reproductive number $\mathcal{R}_0$ can only be calculated explicitly under certain conditions. To approximate the reproductive number, we make an additional assumption. In particular, $\frac{D}{L} << 1$. That is, the infectious period is shorter compared to the timescale of the mortality rate. Hence, discretizing according to the discrete age-cohort given in the previous section, the next-generation matrix is given by

$$G_{ij} = \frac{D}{L}\left[\int_{a_i}^{a_{i+1}} e^{-\int_0^a \mu(\tau)d\tau} da\right]\beta_{ij},$$

where each element of G_{ij} denotes the average number of secondary transmissions caused by a single primary case for age-cohort i to j. The reproductive number $\mathcal{R}_0$ for the whole population is given by the largest eigenvalue of the next-generation matrix. Hence, $\mathcal{R}_0 = r(G_{ij})$, where r is the spectral radius. We let $L = 84$ for a typical developed country (i.e., U.S or U.K), $D = 5$ days (similar to the infection period of influenza), $\beta_{ij} = \beta_0 c_{ij}$ where c_{ij} is the contact rate and β_0 is the probability of infection, and $\beta_0 = 0.49$. We calculate the basic reproductive number to be $\mathcal{R}_0 = 1.97$. Before providing numerical simulations of different allelic variations, we will show how the model performs against serological data.

Table 7.1: **Contact, Fertility and Mortality Rates.** The infection rate with effective contact structure was obtained by averaging the contact rates within each age-group from Figs. 7.1-7.2. The units are per years.

$\beta(a,s)$	0-19	20-39	40-59	60+	$f(a)$	$\mu(a)$
0-19	6.0211	1.4562	1.2836	0.4523	0.175	0.0328
20-39	2.0593	3.3250	1.5573	0.7697	0.2772	0.0065
40-59	1.6843	1.4740	2.4104	1.3689	0.0890	0.0129
60+	0.7003	0.3273	0.8630	2.0037	0.00	0.0718

7.3 Model Validation with Serological Data

Before we apply any estimation scheme, we must first make some assumptions. We assume that the system described by Eqs. (4.18)-(4.20) has reached a steady state. Furthermore, since the data do not yield any information about the similarity matrix Ω_{ij} between strains, and given the nature of the serological data which capture previous exposures or lack of it, we look at the percent of seropositive individuals with strain i and those sharing similar antigens to it. As a result of this assumption, the model can further be reduced to a single strain system with state variable capturing this auxiliary state with constant correction.

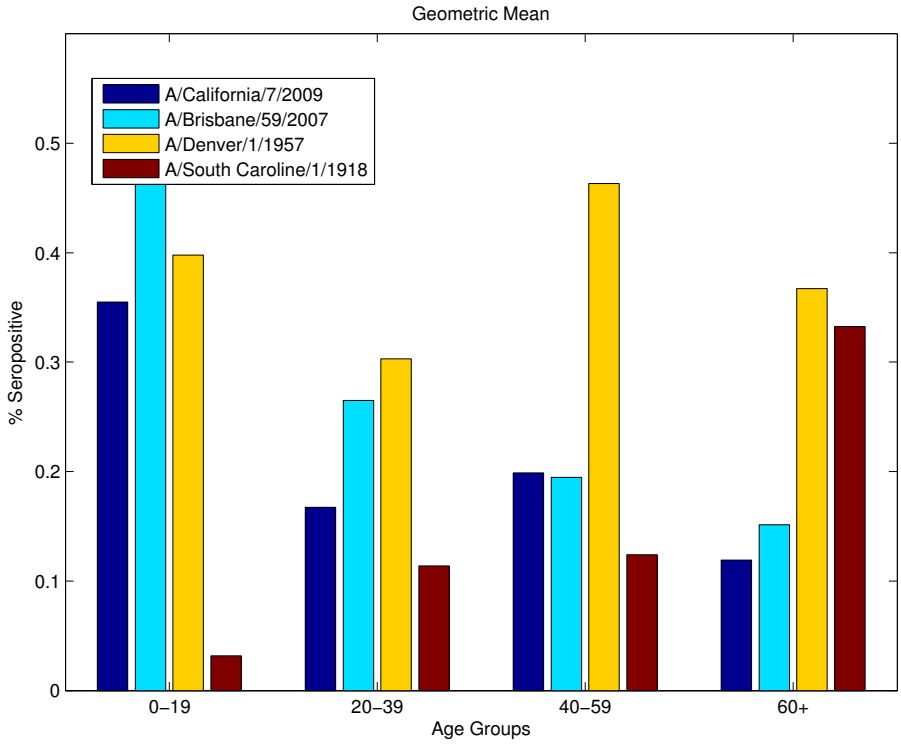


Figure 7.3: **Age-specificity in Serological Surveys of Influenza A Virus.** The proportion of people having antibodies to the Influenza A virus stratified by age is shown (data adapted from Zimmer et al., 2010).

Integrating Eqs. (4.18)-(4.20), we obtain age distributions of the epidemiological variables. Moreover, we assume that the duration of the disease is short compared to the timescale on which the mortality and transmission rates vary. This assumption reduces the force of infection into an expression containing only a single integral. However, estimating the rates can not be done analytically using serological data because the integral equations may not have a closed form solution. To circumvent this problem, we rely on discretization of the age structure model and the integrals, where the force of infection becomes (piece-wise) constant in each age cohort and the contact rate $\beta(a,s) = qc(a,s)$ where $c(a,s)$ is given in Table 7.1 as $\beta_{ij} = qc_{ij}$ and the mortality rate $\mu(a)$ becomes μ_j, for four age cohorts: $[0-19]$, $[20-39]$, $[40-59]$, $[60-100]$.

We now turn to parameter estimation. An inverse problem via combinations of likelihood and sum error square estimators is used to relate the model robustly to serological data for the influenza epidemiological data summarized in Fig. 7.3. We use a two-tier estimation scheme to estimate relevant epidemiological parameters, some of which can later be used to obtain the basic reproductive number. We employ a combination of a generalized estimation scheme based on modifications made to

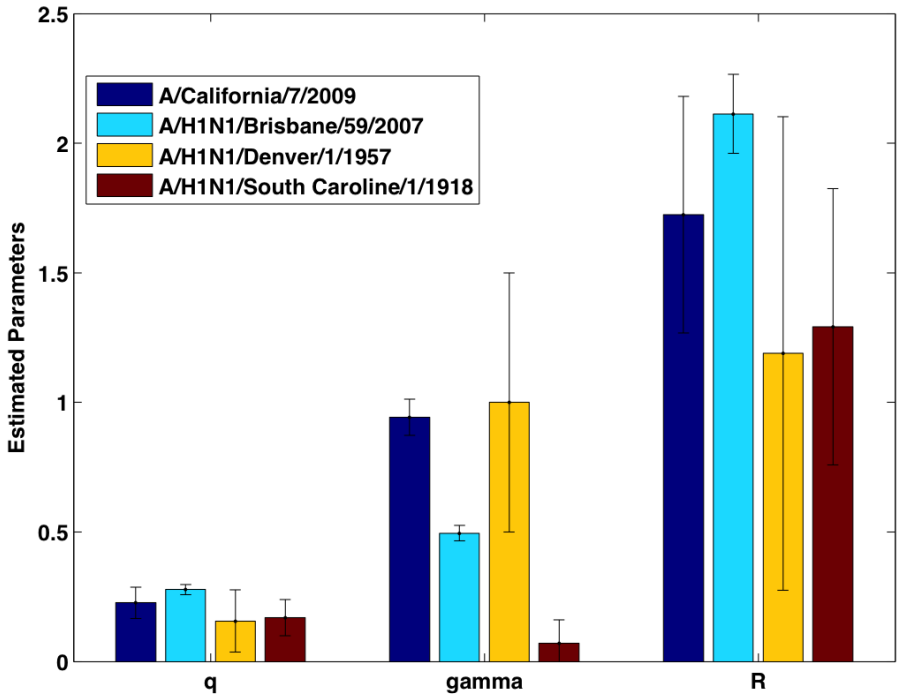


Figure 7.4: **Estimated Parameters.** The parameters were estimated using the scheme discussed in the text. The standard error was calculated via a bootstrapping method.

the approach outlined in Varah (1982), Ramsay & Silverman (2002), Poyton et al. (2006), Cao & Ramsay (2007), and Banks and Tran (2009). In the first tier, we employ a weighted likelihood (specifically, log-likelihood estimators for simplification) to estimate the force of infection. Here it is assumed that the force of infection is constant for each age cohort, resulting in piecewise constant force of infection. A weighted likelihood estimator is used because the sample size of the serological data may vary across each age group. In addition, the serological data are assumed to be binomially distributed. In the second tier, we use the estimated force of infection to estimate the remainder of the parameters, in particular, the cross-immunity level, γ, and the disease-specific transmission rate, q. At this stage, we employ weighted (same weight as in the first tier) sum square error while constraining the estimates to satisfy the fixed point condition on the force of infection. Alternatively, we can combine both the first and the second tiers, and in addition constrain the fixed point criterion in a penalized estimation scheme.

Assuming a 2-locus-2-allele system and using the data given in Table 7.1 for $\beta_{ij} = qc_{ij}$ and the mortality rate μ_j, and in Fig. 7.3 for the serological data for four A/H1N1 subtypes, we estimate q and constant $\gamma(a) = \gamma$. The results are summarized in Figs. 7.4-7.6. We note that, since the estimates in Fig. 7.4 are based on retrospective epidemiological events, γ can be seen as a measure of pre-existing

levels of cross-immunity in the population. Figures 7.5-7.6 show that our model fits most of the serological data for the four Influenza A H1N1 subtypes. Based on the coefficient of determination r^2 values, we observe that A/H1N1/Brisbane/59/2007 ($r^2 = 0.9958$), A/H1N1/California/7/2009 ($r^2 = 0.8557$) and A/H1N1/South Caroline/1/1918 ($r^2 = 0.6441$) have r^2 values above 0.5 (see Figs. 7.5**(a)-(b)** and Fig. 7.5**(d)**). The standard errors for most of the estimated parameters are very small (see Fig. 7.4). Only the subtype A/H1N1/Denver/1/1957 ($r^2 = 0.4301$) has the coefficient of determination r^2 less than 0.5, and larger standard error. For all the subtypes, the basic reproductive number is greater than one and the values are consistent with the values observed for Influenza A H1N1 in the literature (Dushoff et al., 2004; Gog et al., 2003; Gani et al., 2005). From the estimates of the cross-immunity level, we note that the subtype A/H1N1/South Caroline/1/1918 has the lowest cross-immunity level γ ($\gamma = 0.0709$). This is not surprising because retrospective serological studies have shown that individuals in 1918 did not have any previous exposure to a strain similar to the subtype strain of A/H1N1/1918. As a result, the pandemic of 1918 had devastating impact on society. The cross-immunity levels for A/H1N1/California/7/2009 ($\gamma = 0.9420$) and A/H1N1/Denver/1/1957 ($\gamma = 1$) are also very high while A/H1N1/Brisbane/59/2007 is about $\gamma = 0.4955$. The results are consistent with the serological and Hemagglutination-Inhibition Assay (HAI) studies.

Figure 7.7 shows the potential infectivity of the population (i.e., $\mathcal{P}_i(a) = s(a)+(1-\gamma)v(a)$, where $s(a)$ and $v(a)$ represent age-distributions of susceptible and partially immune, see Chapters 5-6), which measures how much infectivity could be created if all hosts aged a were infected, and it reflects the population immunity. High value of $\mathcal{P}_i(a)$ implies that there is least population immunity whereas low $\mathcal{P}_i(a)$ implies high population immunity. We see that the potential infectivity decreases with age, and in addition, certain strains have high $\mathcal{P}_i(a)$. For example, for A/South Caroline/1/1918, individuals in the age-cohort of $[0-19]$ have over 90% potential infectivity. As compared to those aged 60 and over, there is 50% future potential infectivity. Similarly, for the A/H1N1/California/7/2009 strain, the potential infectivity measure is small across all ages, this might explain why the infection rate for the 2009 pandemic was low, due to prior high immunity to the strain. Again, individuals aged 60+ have the lowest potential infectivity for the 2009 and 2007 strains.

Figure 7.8 shows the genetic relatedness of HA from the four historical strains of influenza H1N1 (A/H1N1/California/7/2009, A/H1N1/Brisbane/59/2007, A/H1N1/Denver/1/1957 and A/H1N1/South Caroline/1/1918, see Zimmer et al., 2010). The phylogenetic tree inferred from hemagglutinin amino acid sequences suggests that the antigenic pair of A/H1N1/California/7/2009 and A/H1N1/South Carolina/1/1918

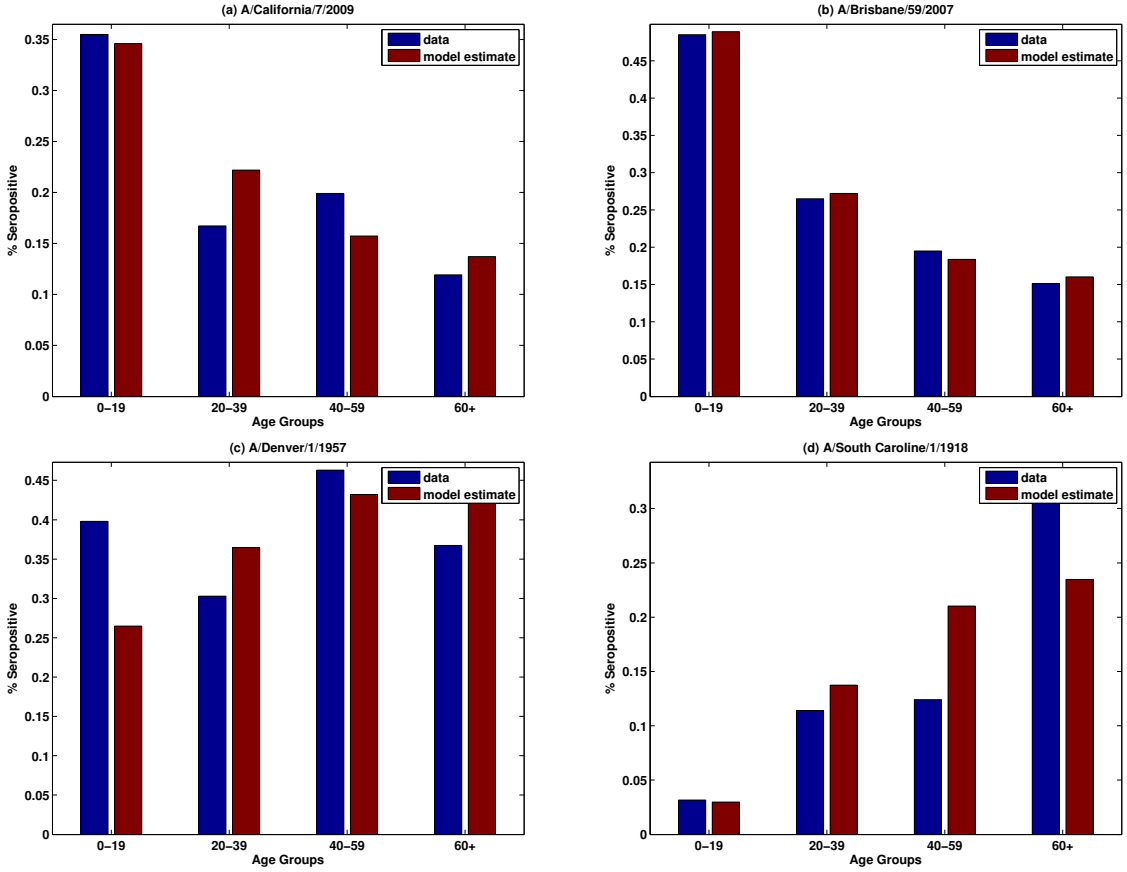


Figure 7.5: **Comparison of Model** (4.18)-(4.20) **and Data with Four Age Classes.** The estimates for the four different influenza subtype strains are **(a)** $\overline{\gamma} = 1 - \gamma = 0.0580$ ($\gamma = 0.9420$) and $q = 0.2265$, $\mathcal{R}_0 = 1.7243$, $r^2 = 0.8557$ for A/H1N1/California/7/2009; (b) $\overline{\gamma} = 1 - \gamma = 0.5045$ ($\gamma = 0.4955$) and $q = 0.2776$, $\mathcal{R}_0 = 2.1130$, $r^2 = 0.9958$ for A/H1N1/Brisbane/59/2007; (c) $\overline{\gamma} = 1 - \gamma = 0.0000$ ($\gamma = 1$) and $q = 0.1562$, $\mathcal{R}_0 = 1.1892$, $r^2 = 0.4301$ for A/H1N1/Denver/1/1957; (d) $\overline{\gamma} = 1 - \gamma = 0.9291$ ($\gamma = 0.0709$) and $q = 0.1696$, $\mathcal{R}_0 = 1.2915$, $r^2 = 0.6441$ for A/H1N1/South Caroline/1/1918, respectively.

has stronger correlations than the antigenic pair of A/H1N1/California/7/2009 and A/H1N1/Brisbane/59/2007 and the pair of A/H1N1/Brisbane/59/2007 and A/H1N1/ South Carolina/1/1918. These observations are in line with the results of Zimmer et al., 2010. The cross-immunity values for A/H1N1/California/7/2009 and A/H1N1/ Denver/1/1957 show that these subtypes have some similarity with previously circulating strains. The value of $\gamma = 1$ for A/H1N1/Denver/1/1957 suggests complete protection resulting in that strain. Historical observation of H1N1 strain A/H1N1/Denver/1/1957 correspond to the last H1N1 circulating prior to the 1957 pandemic of H2N2. This is again consistent with the biology and the epidemiology dynamics discussed previously. Based on historical knowledge of known circulation of H1N1 in the past century, the patterns of serological surveys show expected age

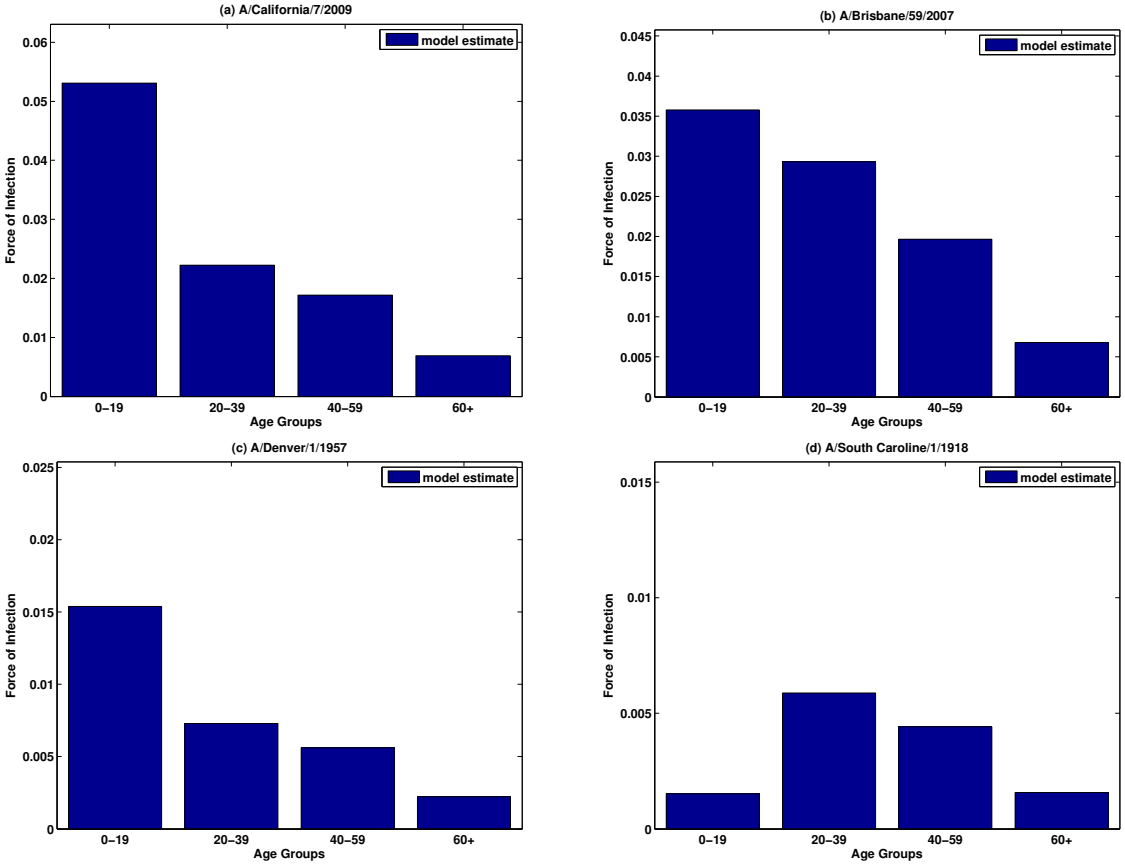


Figure 7.6: **Estimated Force of Infection for Four Age Classes.** The force of infection was estimated from the seroprevalence data for four influenza subtypes for **(a)** A/H1N1/California/7/2009; **(b)** A/H1N1/Brisbane/59/2007; **(c)** A/H1N1/Denver/1/1957; and **(d)** A/H1N1/South Caroline/1/1918).

specificity. The combination of decreased disease pandemic influenza incidence, the high levels of pre-existing antibodies against novel H1N1, and the high seroprevalence of antibodies to the 1918 influenza strain among the elderly, support the role for the long lasting immunologic memory of the initial influenza infection and the estimated cross-immunity levels calculated above. We note that, in Figs. 7.5**(d)** and 7.6**(d)**, the proportion of seropositive and the force of infection for individuals in 0-19 cohort are very small for A/H1N1/South Caroline/1/1918. This is consistent due to the fact that this cohort represents individuals born between 1990 and 2009.

With the predictive success of the age-structured model, in the next section, we will perform various numerical simulations of higher dimensional strain spaces of the model with different allelic combination in order to gain insight into the age-specificity, to characterize the age-specific immunological profile, and to untangle the effects of community structures as captured by the clustering behaviours.

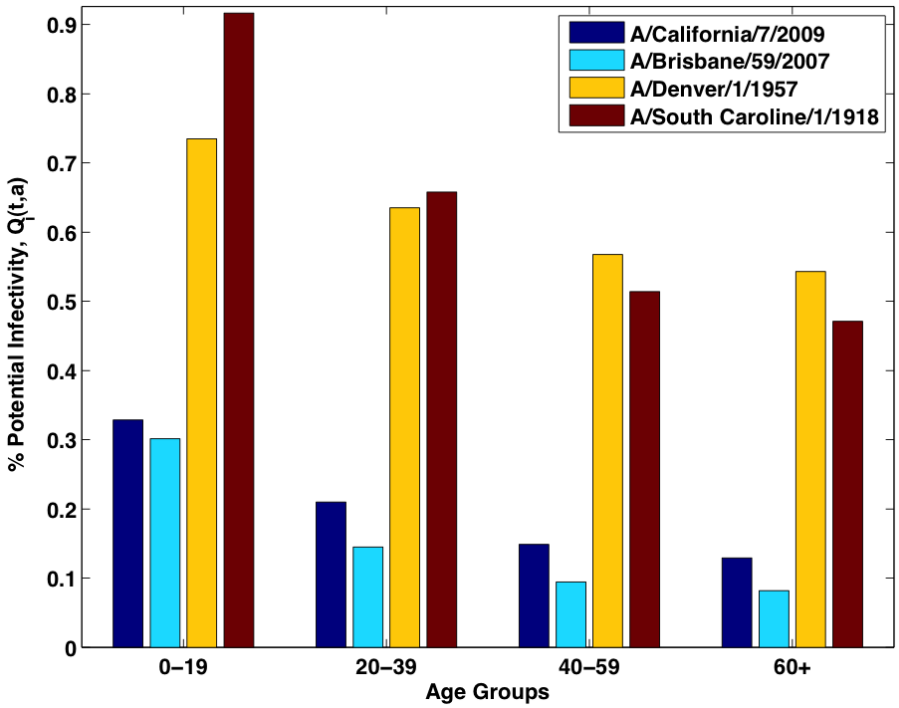


Figure 7.7: **Potential Infectivity.** Potential infectivity measures the population immunity. Given a specific strain, the potential infectivity shows how much infectivity could be created if all hosts in each age-cohort were infected with the specific strain under the cross-immunity values estimated.

7.4 Numerical Simulations: The Effect of Epitope Architectures

In this section, we investigate the effect of the serotypic architectures and the various dynamic features of the model. In particular, we look at the dynamic behaviours of the model under both uniform and differential allelic combinations as we vary the cross-immunity level γ. In the proceeding sections, we study both uniform parameters $\beta(a,s) = \beta$, $\mu(a) = \mu$, $\sigma(a) = \sigma$ and age-specific parameters $\beta(a,s) = qc(a,s)$, $\mu(a)$, $\sigma(a)$. The cross-immunity parameter γ is assumed to be constant and is used as an exploratory parameter. In the case of age-specific parameters, we will scale q such that the basic reproductive number is $\mathcal{R}_0$. Using uniform/homogeneous or constant cross-immunity (i,.e., $\gamma(a) = \gamma$ for all $a \in [0, a_m]$), we observe that dynamic behaviours of the strains naturally tend towards the dynamics observed in the unstructured model discussed in Chapter 3. That is, the system either exhibits co-existence where all strains persist or the principle of competitive exclusion where a subpopulation of strains organize themselves into ordered sets in which at least one cluster goes to extinction.

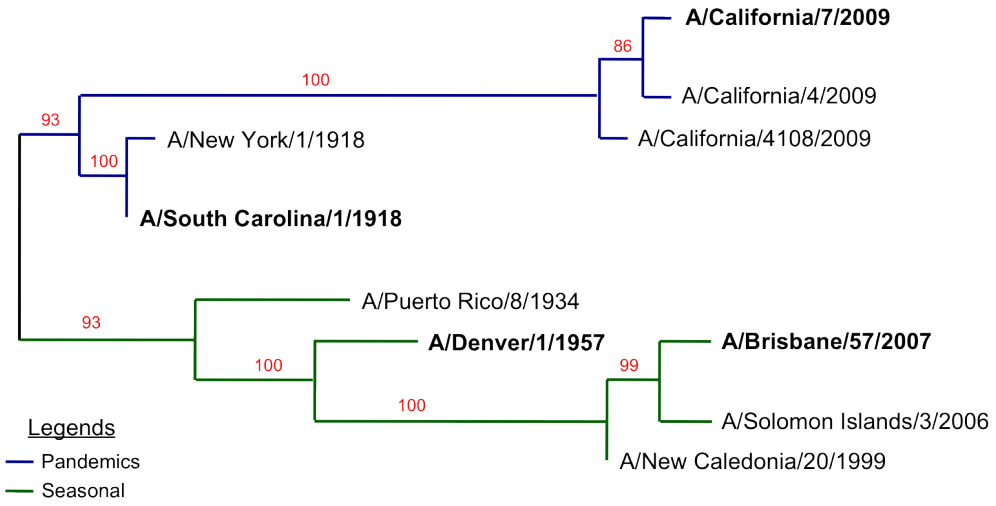


Figure 7.8: **Genetic Relatedness of Hemagglutination-Inhibition Assay (HAI).** The phylogenetic tree inferred from hemagglutinin amino acid sequences via the maximum likelihood method shows a genetic relatedness of HA from historical strains of influenza H1N1 (A/H1N1/California/7/2009, A/H1N1/Brisbane/59/2007, A/H1N1/Denver/1/1957 and A/H1N1/South Caroline/1/1918). The H1N1 influenza isolates are from the classic swine and human influenza lineages (reproduced from Zimmer et al., 2010).

7.4.1 Model Dynamics I: Uniform Age-specific Parameters

When all the epidemiological parameters are independent of age (i.e., $\beta(a,s) = \beta$, $\sigma(a) = \sigma$, $\mu(a) = \mu$), the dynamics are identical to the unstructured homogeneous model described by a system of ordinary differential equations (see Chapter 3). As before, we observe that the system exhibits three distinct dynamical behaviours with a very rich temporal and clustering dynamics and self-organization/synchrony that depends on the relative degrees of cross-immunity, whereby low degree of cross-immunity results in the coexistence of all strains and the maintenance of maximum diversity and low discordance (see Figs. 7.9a, 7.9d, 7.10, and 7.14) while intermediary and high level of immune response result in the formation of a varying degree of clustering with different dynamic patterns (see Figs. 7.9b-7.9c, 7.9e-7.9f, 7.11-7.13, 7.15-7.18).

In Fig. 7.10, we observe that the co-existence of all strains for low cross-immunity: all strains tend to a stable equilibrium which are usually commensurable to their respective reproductive fitnesses. However, because we have assumed that all the strains have the same epidemiological parameters and hence reproductive successes, the strains approach the same abundance level regardless of the allelic architectures. That is, all strains approach the same steady states whether we have uniform (e.g., 2-locus-2-allele, 3-locus-2-allele systems, see Fig. 7.10a) or differential allelic variations (e.g., 2-locus system with 2 alleles at one locus and 3 alleles at the other locus; see Fig.

7.10b). These dynamic patterns correspond to fully synchronized no structure structures (NSS). For high cross-immunity, the system exhibits the principle of competitive exclusion where certain clusters containing discordant serotypes or containing sets of serotypes (i.e., sub-population of strains) with minimal or non-overlapping structures persist while others go to extinction (see Fig. 7.13). In particular, for uniform allelic variations, only discrete strain structures with a single discordant pair with non-overlapping structures are possible (see Figs. 7.13a and 7.13d), whereas in the case of differential allelic combinations we see discrete strain structures with both non-overlapping (see Figs. 7.13c and 7.13f, where persistent strains are discordant to each other and share no alleles in common) and minimally overlapping serotypic structures (see Figs. 7.13b and 7.13e where the dominant persistent strain is discordant to the other two strains that shares alleles in common). In general, the dominant persistent sets share the least number of alleles while the other persistent strains share the most in the case of discrete strain structures with minimally overlapping serotypic structures. These dynamic patterns correspond to partially synchronized discrete strain structures (DSS).

At the intermediary cross-immunity level, the model exhibits cyclic (both periodic, intermittency chaotic) temporal dynamics with exotic cluster-specific sequential domination (see Figs. 7.11-7.12). However, the cyclical dynamical behaviours and their cluster-specific structures at the intermediary cross-immunity level are subtle, as the model exhibits various complex forms of synchronization and dynamic patterns. In particular, we observe that dominant types are cyclically replaced, where we can detect sharp epidemics dominated by, among a number of other behaviours: **(i)** anti-phase periodic solution with alternating dominance of discordant pairs (including large amplitude periodic cycles), **(ii)** a single strain with erratic multi-year abundance similar to the dynamic feature of influenza epidemics with irregular emergence and re-emergence of certain pathogenic forms, **(iii)** ordered alternating appearance of a single antigenic type in periodic or quasi-periodic form similar to discrete periodic travelling waves in strain/antigenic space with spatio-temporal symmetry, **(iv)** chaotic solutions with erratic appearance and disappearance of synchrony between discordant antigenic types or discrete sets of minimally overlapping or non-overlapping structures (periodic solutions with intermittent chaotic episodes, chaotic short-period cycles with decoherence between discordant sets, large-amplitude chaotic cycles with spatio-temporal symmetry), and **(v)** phase-synchronization with (periodic or chaotic) uncorrelated amplitudes. These dynamic behaviours represent cyclical strain structures (CSS). Some of these dynamic patterns described above have been discussed previously in Gupta et al. (1998), Recker et al. (2007), Buckee et al. (2011) and

Chapter 3. The only difference between these behaviours and those discussed previously is that each cohort follows similar dynamics, and there might be age-specific synchronizations, which may not hold in the case of a non-uniform age-specific scenario. In addition, regardless of allelic combinations, for high dimensional systems, we observe that, in the discrete strain structure regions, only differential allelic variations can exhibit discrete strain structures with both non-overlapping discordant and minimally overlapping serotypic structures, whereas uniform allelic combinations show discrete strain structures with discordant pairs.

7.4.2 Model Dynamics II: Differential Age-specific Parameters

Making the epidemiological parameters age-specific, we again investigate the role of cross-immunity level γ in determining the dynamic patterns of both strain structures and population profile. Our results show that the dynamic features observed both in the previous section in the case of uniform/homogeneous age-specific parameters, and in Chapter 3 in the case of the unstructured model given by a set of ordinary differential equations, are preserved in the differential age-specific model. That is, the dynamical behaviours outlined above persist. However, differential age-specificity introduces an additional layer of complexity. In particular, we observe that the abundance level of each strain for individual age-cohorts is commensurable with the age-specific reproductive number. Figures 7.14-7.18 illustrate the main results of our investigation, namely that the organization and dynamical behaviours of strain structures and abundance levels are influenced by the combination of age-specific transmissibility/reproductive number and contact structures, and the cross-immunity level. Figures 7.14-7.18 show the relative proportion of the steady-state age-distribution, the corresponding temporal dynamics, differences in age-specificity between strains and average-specific temporal dynamics. We observe that, for low cross-immunity as in the previous section, the system approaches a stable endemic equilibrium where all strains coexist at equal abundance level commensurable with the age-specific reproductive successes (see Figs. 7.14, where the dynamical behaviours are representative of NSS discussed previously). As we increase the level of cross-immunity, the age-distribution starts to change as the combination of age-specific transmissibility fitness and the cross-immunity level places different age groups into separate dynamic regions. That is, certain age-cohorts can fall into cyclical strain structure (CSS) while some age-cohorts exhibit no strain structure (NSS) as seen in low intermediary cross-immunity level. Similarly, for high intermediary cross-immunity level, we can observe a case where some age-cohorts exhibit CSS while other age-cohorts show DSS. However, we have not presented these age-specific behavioural patterns in the present simulations. Instead, we observe that as we increase the level of cross-immunity, the

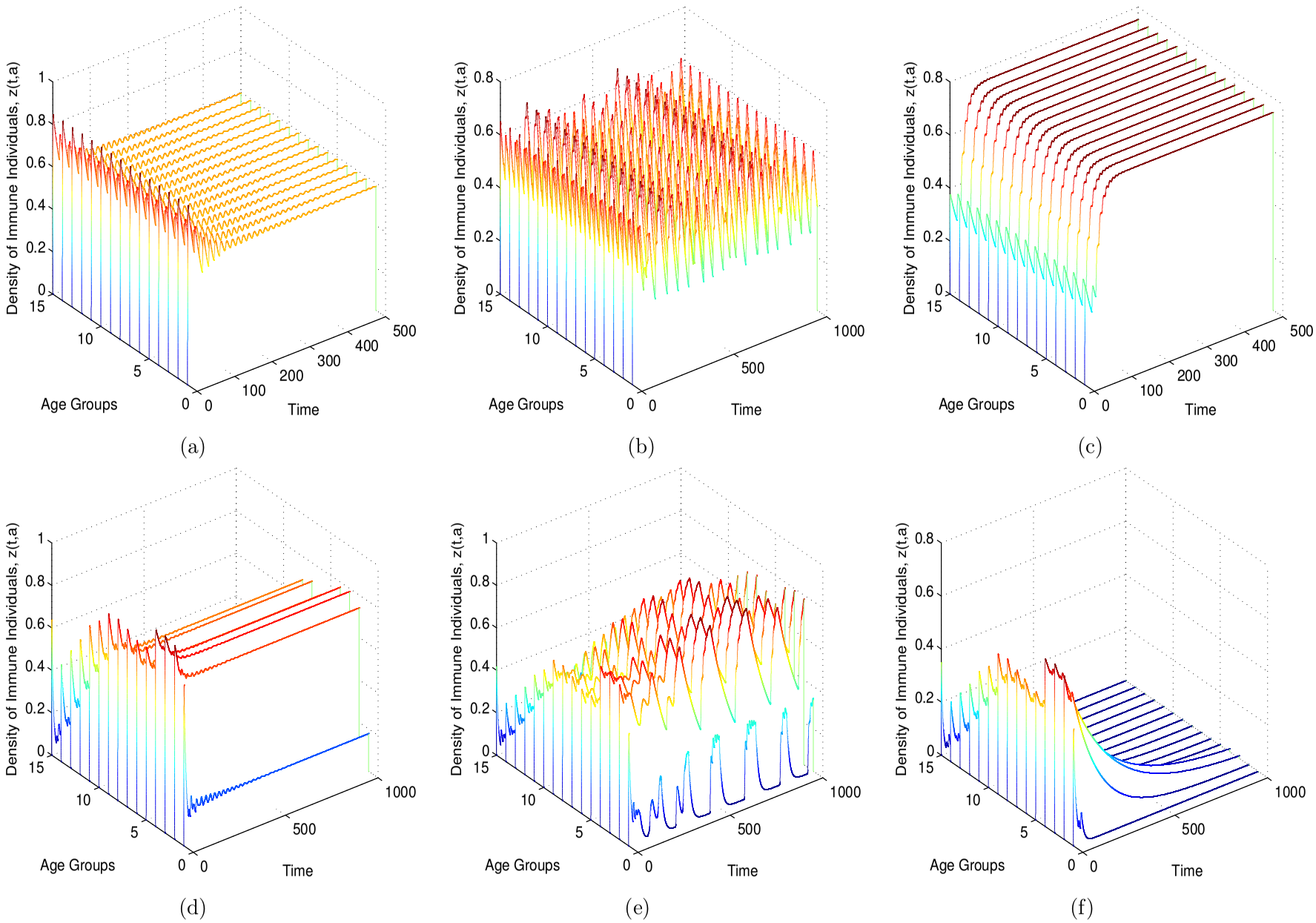


Figure 7.9: **Three Long-term Age-specific Dynamical Patterns.** Dynamics of the model for **(a)-(c)** uniform parameters with the following $\beta(a,s) = \beta = 40$, $\mu(a) = \mu = 0.02$, $\sigma(a) = \sigma = 9.98$, and for **(d)-(f)** age-specific parameters given in Section 7.2.2. Only the first strain is shown for each dynamical pattern. In **(a)** and **(d)**, the system exhibits NSS, with $\gamma = 0.50$, in **(b)** and **(e)**, $\gamma = 0.75$ and we observe CSS, while **(c)** and **(f)** $\gamma = 0.95$ and we observe DSS. The dynamic patterns are for a 3-locus-2-allele system.

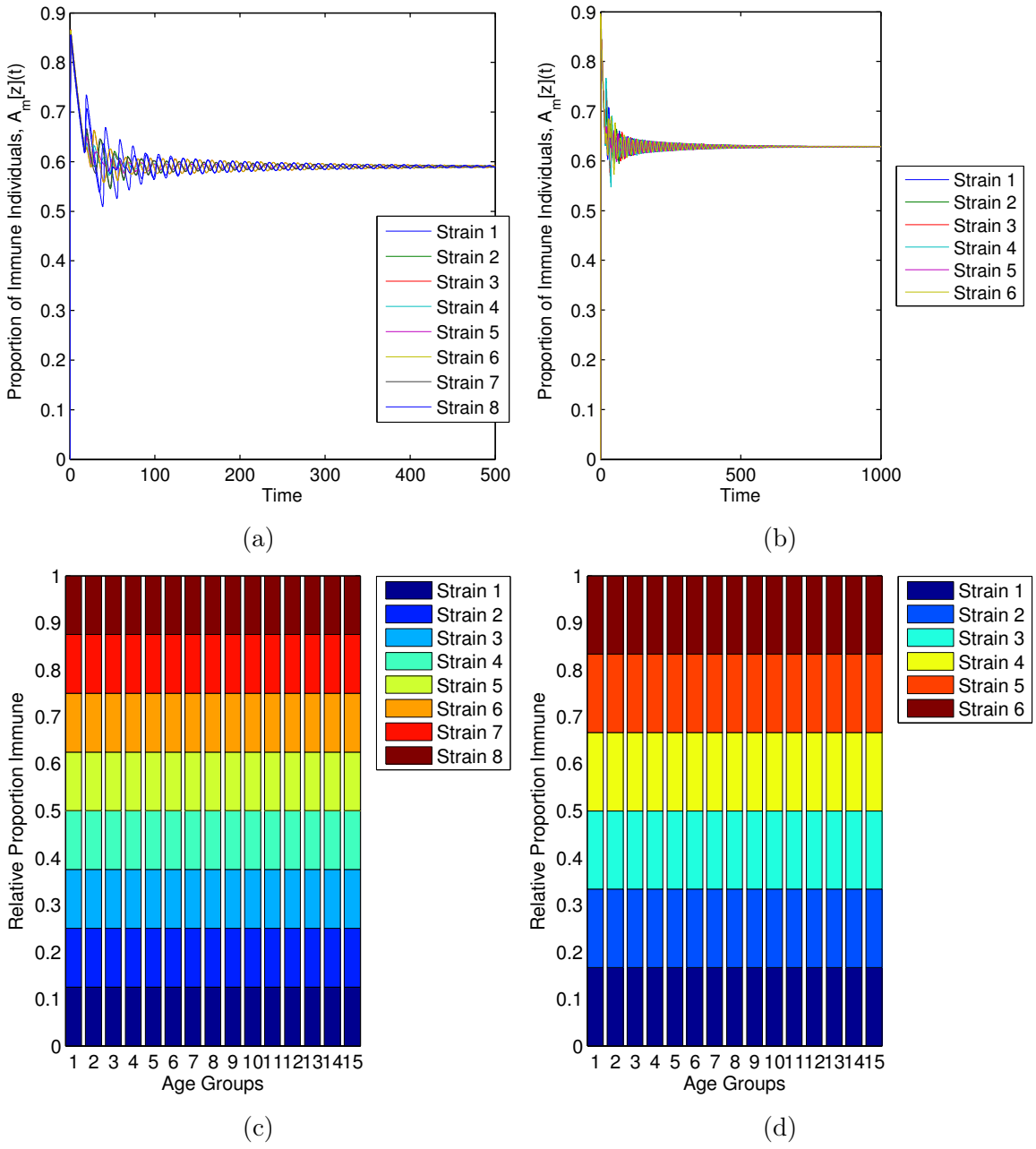


Figure 7.10: **Age-Specificity in NSS and Average Temporal Evolution.** Temporal dynamics show the coexistence of strains approaching the same endemic steady-state for **(a)** a 3-locus-2-allele system and for **(b)** a 2-locus-(2,3)-allele system, where $\mathcal{A}_m[\phi](t) = \frac{1}{|\Omega_m|}\int_{\Omega_m}\phi(t,a)da,\ \Omega_m = [0, a_m]$. The strain-specific scaled relative proportion of immune individuals across different age-cohorts where **(c)** and **(d)** correspond to those shown in **(a)** and **(b)**, respectively. Here, the parameters are as follows: $\beta(a,s) = \beta = 40,\ \mu(a) = \mu = 0.02,\ \sigma(a) = \sigma = 9.98$ and $\gamma = 0.5$, and initial conditions are randomly generated.

dynamics are purely cyclical and consistent with the CSS dynamic patterns discussed above and in Chapter 3. Figures 7.15-7.17 summarize these characteristic patterns

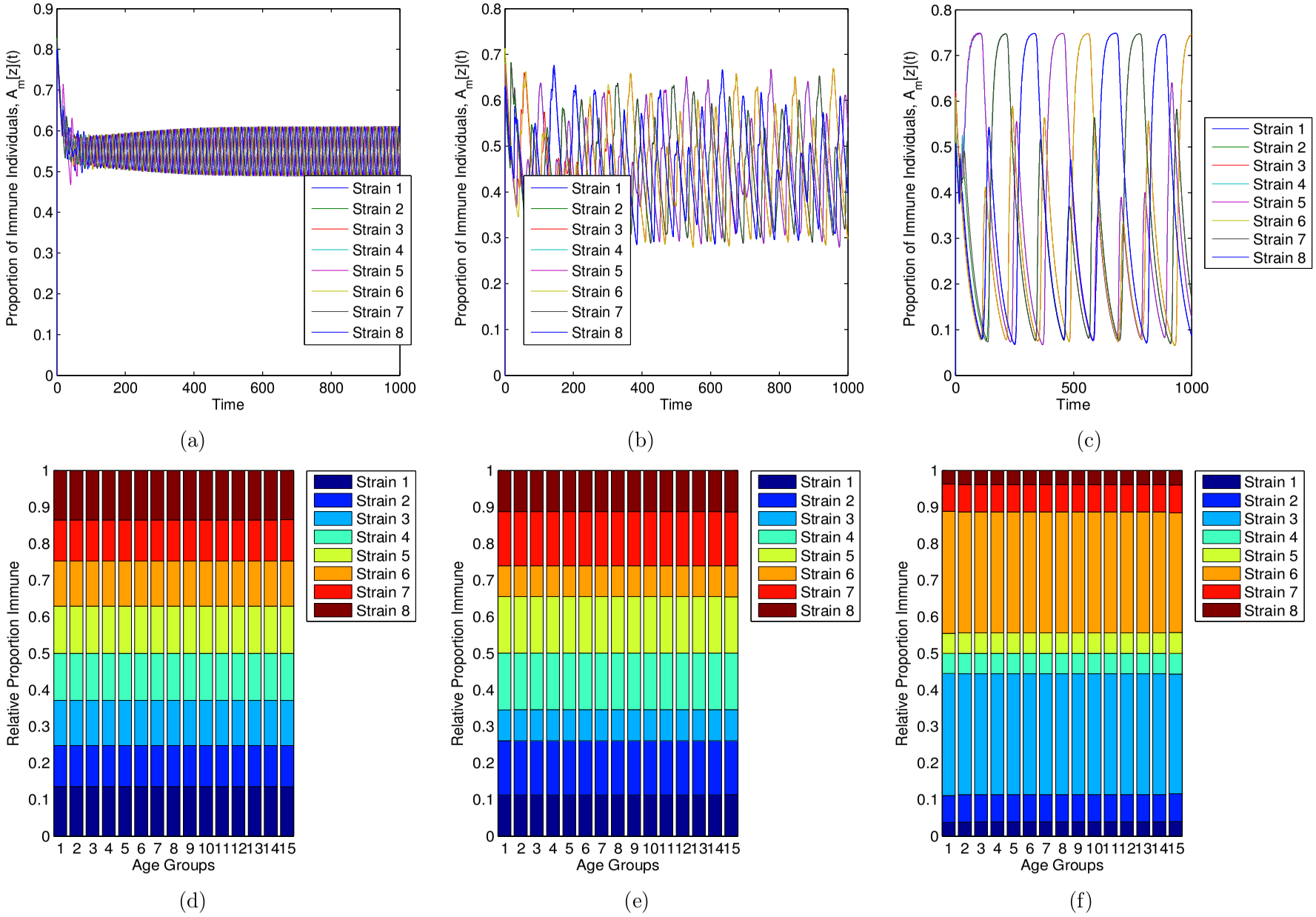


Figure 7.11: **Uniform Age-specificity and CSS of the 3-locus-2-allele System.** Dynamics show the coexistence of eight strains with **(a)** a discrete periodic travelling wave solution, **(b)** chaotic dynamics and **(c)** discrete large-amplitude travelling wave cyclical solutions, with corresponding strain-specific scaled relative proportions of immune individuals **(d)**-**(f)**, respectively, where $\mathcal{A}_m[\phi](t) = \frac{1}{|\Omega_m|}\int_{\Omega_m}\phi(t,a)da$, $\Omega_m = [0, a_m]$. Here, the parameters are as follows: $\beta(a,s) = \beta = 40$, $\mu(a) = \mu = 0.02$, $\sigma(a) = \sigma = 9.98$ and **(a)** $\gamma = 0.58$, **(b)** $\gamma = 0.75$ and **(c)** $\gamma = 0.85$; initial conditions are randomly generated.

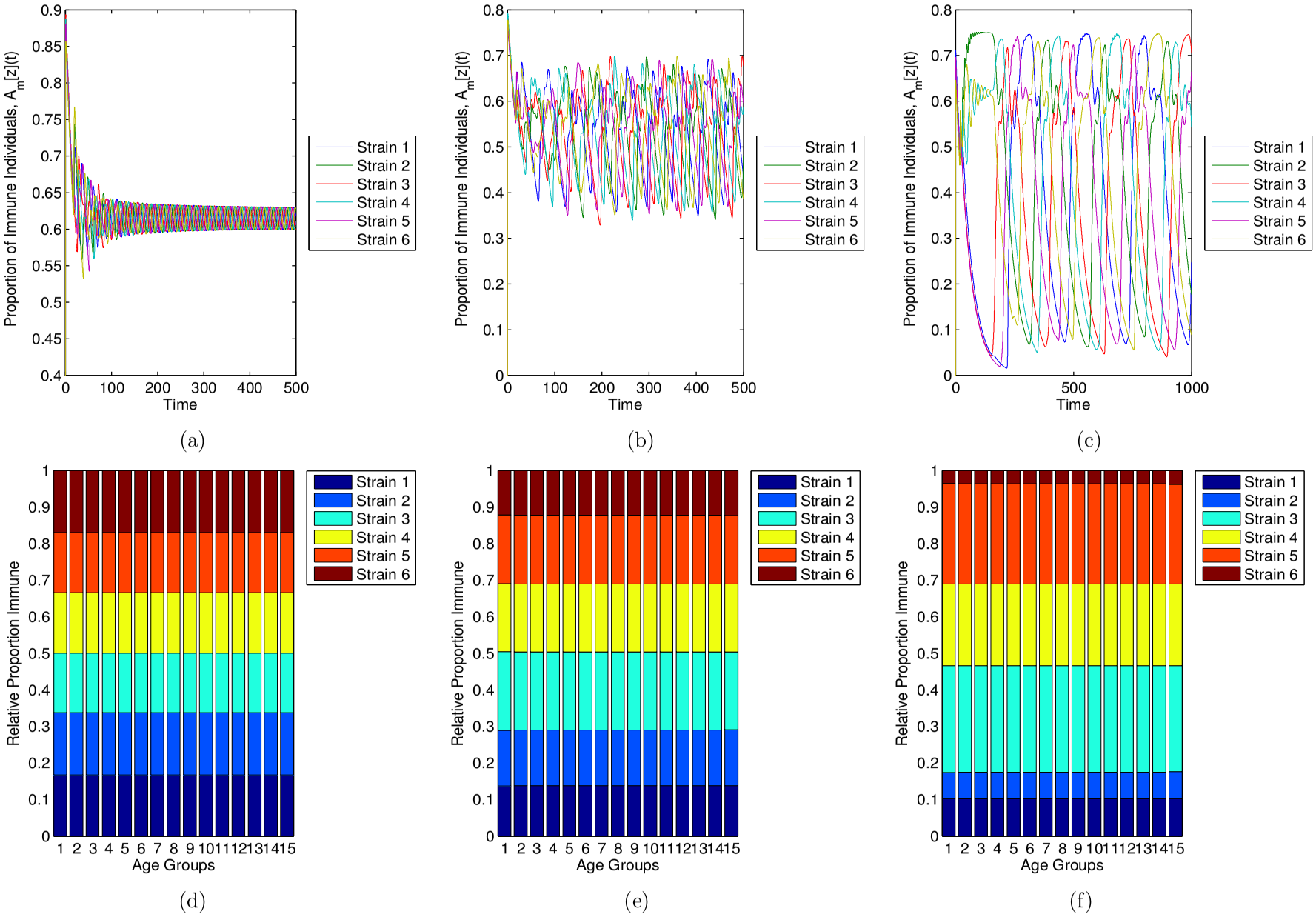


Figure 7.12: **Uniform Age-specificity and CSS of the 2-locus-(2,3)-allele System.** Dynamics show the coexistence of six strains with **(a)** discrete periodic travelling wave solutions and spatio-temporal symmetry across all age-cohorts, **(b)** chaotic dynamics and **(c)** large-amplitude chaotic solutions, with corresponding strain-specific scaled relative proportions of immune individuals **(d)-(f)**, respectively, where $\mathcal{A}_m[\phi](t) = \frac{1}{|\Omega_m|}\int_{\Omega_m}\phi(t,a)da$, $\Omega_m = [0, a_m]$. Here, the parameters are as follows: $\beta(a,s) = \beta = 40$, $\mu(a) = \mu = 0.02$, $\sigma(a) = \sigma = 9.98$ and **(a)** $\gamma = 0.54$, **(b)** $\gamma = 0.75$ and **(c)** $\gamma = 0.90$; initial conditions are randomly generated.

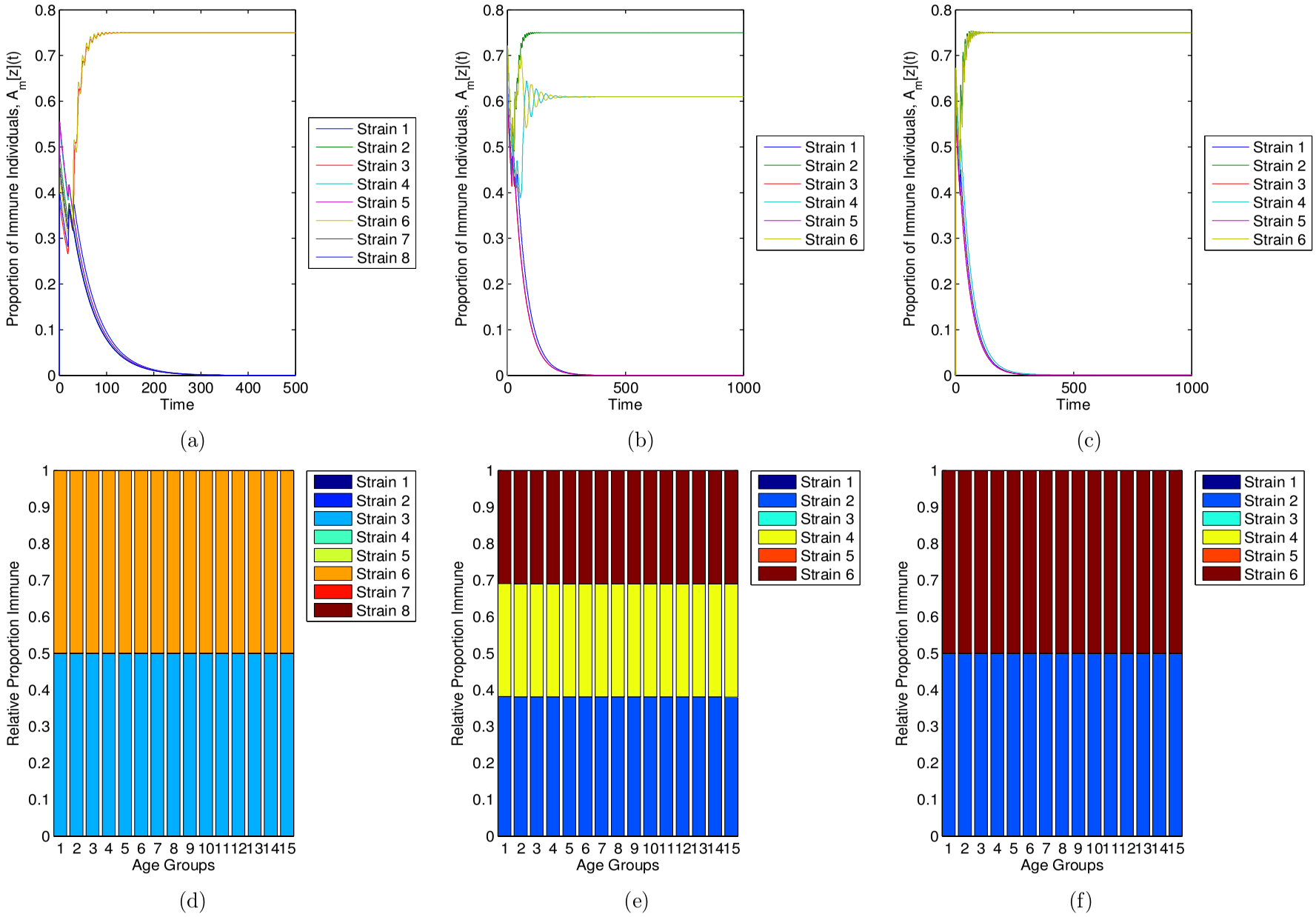


Figure 7.13: **Age-specificity in DSS Dynamics.** Dynamics show the discrete strain structure where one discordant pair persists while the other pairs go to extinction (principle of competitive exclusion) for **(a)** a 3-locus-2-allele system, **(b)**-**(c)** a 2-locus-(2,3)-allele system, with corresponding strain-specific scaled relative proportions of immune individuals across all age-cohorts **(d)**-**(f)**, respectively, where $\mathcal{A}_m[\phi](t) = \frac{1}{|\Omega_m|}\int_{\Omega_m}\phi(t,a)da$, $\Omega_m = [0, a_m]$. Here, the parameters are as follows: $\beta(a,s) = \beta = 40$, $\mu(a) = \mu = 0.02$, $\sigma(a) = \sigma = 9.98$, and **(a)**-**(b)** $\gamma = 0.95$ and **(c)** $\gamma = 1.0$; initial conditions are randomly generated.

ranging from periodic to chaotic behaviours of the model. Furthermore, if we increase the cross-immunity to be very high, the model can exhibit the principle of competitive exclusion resulting in DSS with discordant non-overlapping or minimally overlapping serotypic structures or both depending on the uniformity or differentiability of allelic variations, and on the value of the reproductive number (see Fig. 7.18). If the reproductive number is very high, the region in which DSS is observed might be very narrow, and we may not observe it. This is consistent with the results we obtained in the unstructured model in Chapters 3 and 6. For example, for uniform age-specific parameters and $\gamma = 0.54$, as shown in Figs. 7.12a and 7.12d, we observe that the system exhibits discrete periodic travelling wave solutions with spatio-temporal symmetry, while for differential age-specific parameters with the same cross-immunity level γ, we see that the system approaches a steady-state with NSS pattern (results not shown). In addition, we find that specific clusters of persistent strains in DSS and age-distribution depend on the initial conditions. This is due to the inherent symmetry in the strain space and is also consistent with our theoretical results in Chapters 3 and 6. For differential allelic variations (e.g., 2-locus with two alleles at one locus and three alleles at the other locus), we observe additional features in the DSS region. In particular, we see that for $\gamma = 1$, depending on the initial conditions, we can find DSS with minimally overlapping or discordant serotypic structures. For uniform age-specific parameters, when $\gamma = 1$, only DSS with discordant serotypic structures occur. However, for differential age-specific parameters, both DSS with minimal overlapping and discordant (non-overlapping) serotypic structures (see Figs. 7.18a, 7.18c, 7.18d and 7.18f, and 7.18b and 7.18e, respectively) occur. Dynamic patterns of single strain dominance are usually seen in chaotic regimes and are difficult to detect in the proportion of immune individuals, z_i. However, they are observable in the prevalence of infected individuals, y_i (not shown herein).

7.5 Conclusion

In this chapter, we have examined the dynamic patterns of the multi-locus-allele model stratified/structured by age. In order to investigate the serological consequences of the antigenic thrift model in Section 7.3, we compared the model with four seroprevalence data (A/H1N1/California/7/2009, A/H1N1/Brisbane/59/2007, A/H1N1/ Denver/1/1957 and A/H1N1/South Caroline/1/1918); and found that the estimated parameters, including the reproductive number, are consistent with other studies in the literature. In Section 7.4 we found that all the dynamic features observed in the unstructured model described by the ordinary differential equation system are preserved in the age-structured model with both uniform and differential

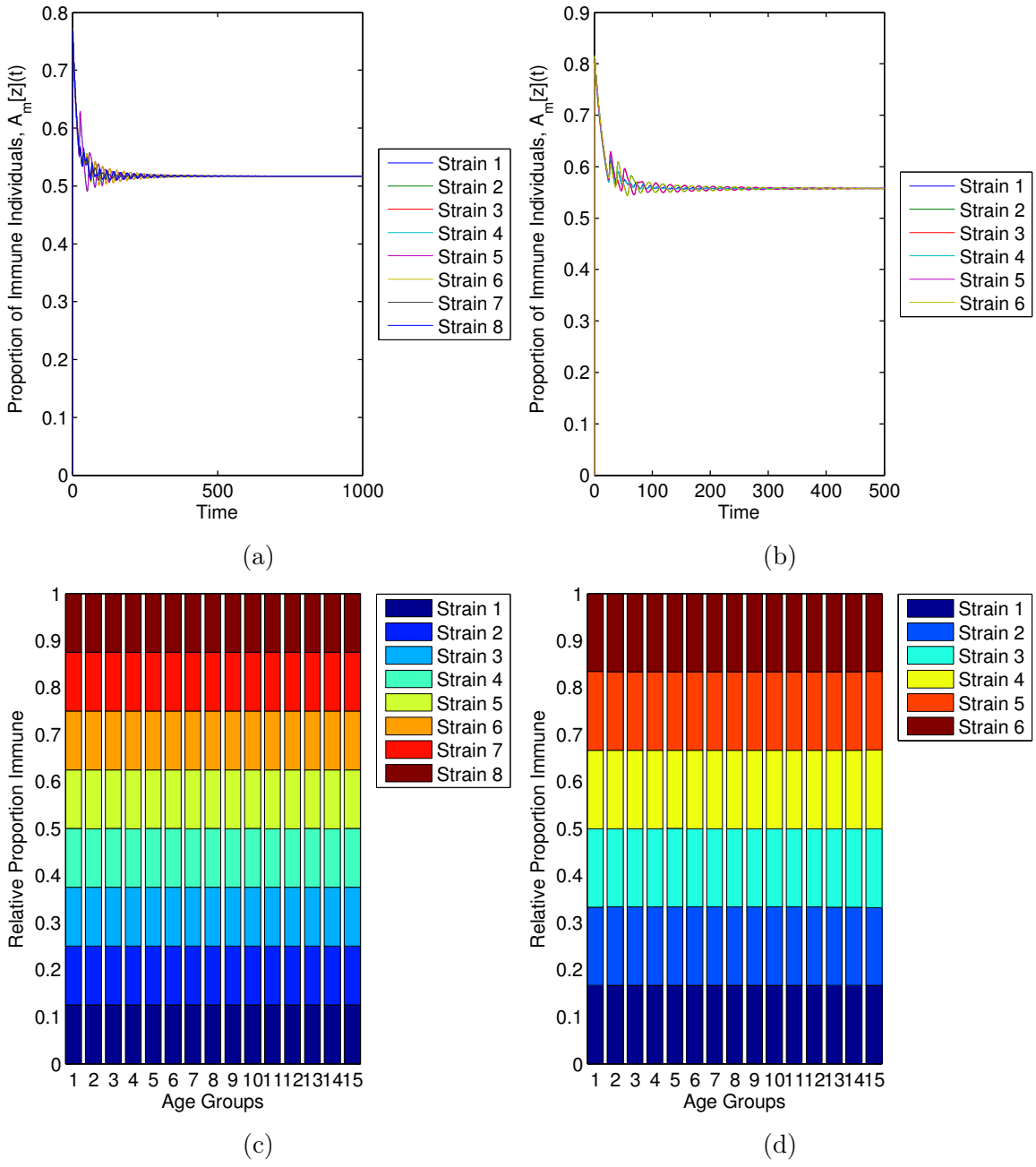


Figure 7.14: **Differential Age-Specificity in NSS and Average Temporal Evolution.** Temporal dynamics show the coexistence of strains approaching the same endemic steady-state for **(a)** a 3-locus-2-allele and **(b)** a 2-locus-(2,3)-allele system, where $\mathcal{A}_m[\phi](t) = \frac{1}{|\Omega_m|}\int_{\Omega_m}\phi(t,a)da,\ \Omega_m = [0, a_m]$. The strain-specific scaled relative proportion of immune individuals across different age-cohorts are shown in **(c)** and **(d)** corresponding **(a)** and **(b)**, respectively. Here, the parameters are given in Section 7.2.2 and $\gamma = 0.5$; initial conditions are randomly generated.

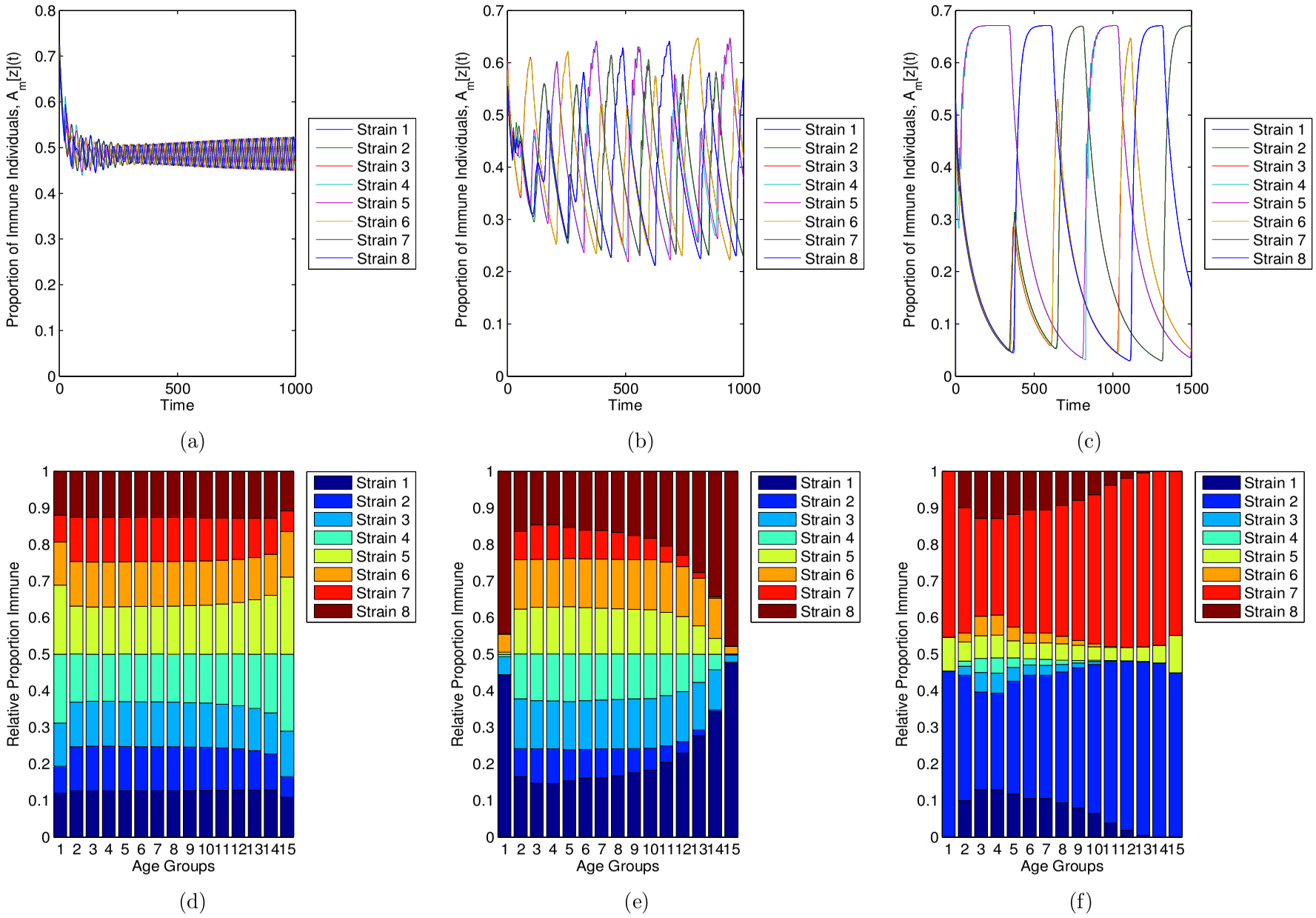


Figure 7.15: **Differential Age-specificity and CSS of the 3-locus-2-allele System.** Dynamics show the coexistence of eight strains with **(a)** discrete periodic travelling wave solutions, **(b)** chaotic dynamics and **(c)** discrete large-amplitude travelling wave cyclical solutions, with corresponding strain-specific scaled relative proportions of immune individuals **(d)**-**(f)**, respectively, where $\mathcal{A}_m[\phi](t) = \frac{1}{|\Omega_m|}\int_{\Omega_m}\phi(t,a)da,\ \Omega_m = [0, a_m]$. Here, the parameters are given in Section 7.2.2, and **(a)** $\gamma = 0.58$, **(b)** $\gamma = 0.75$ and **(c)** $\gamma = 0.85$; initial conditions are randomly generated.

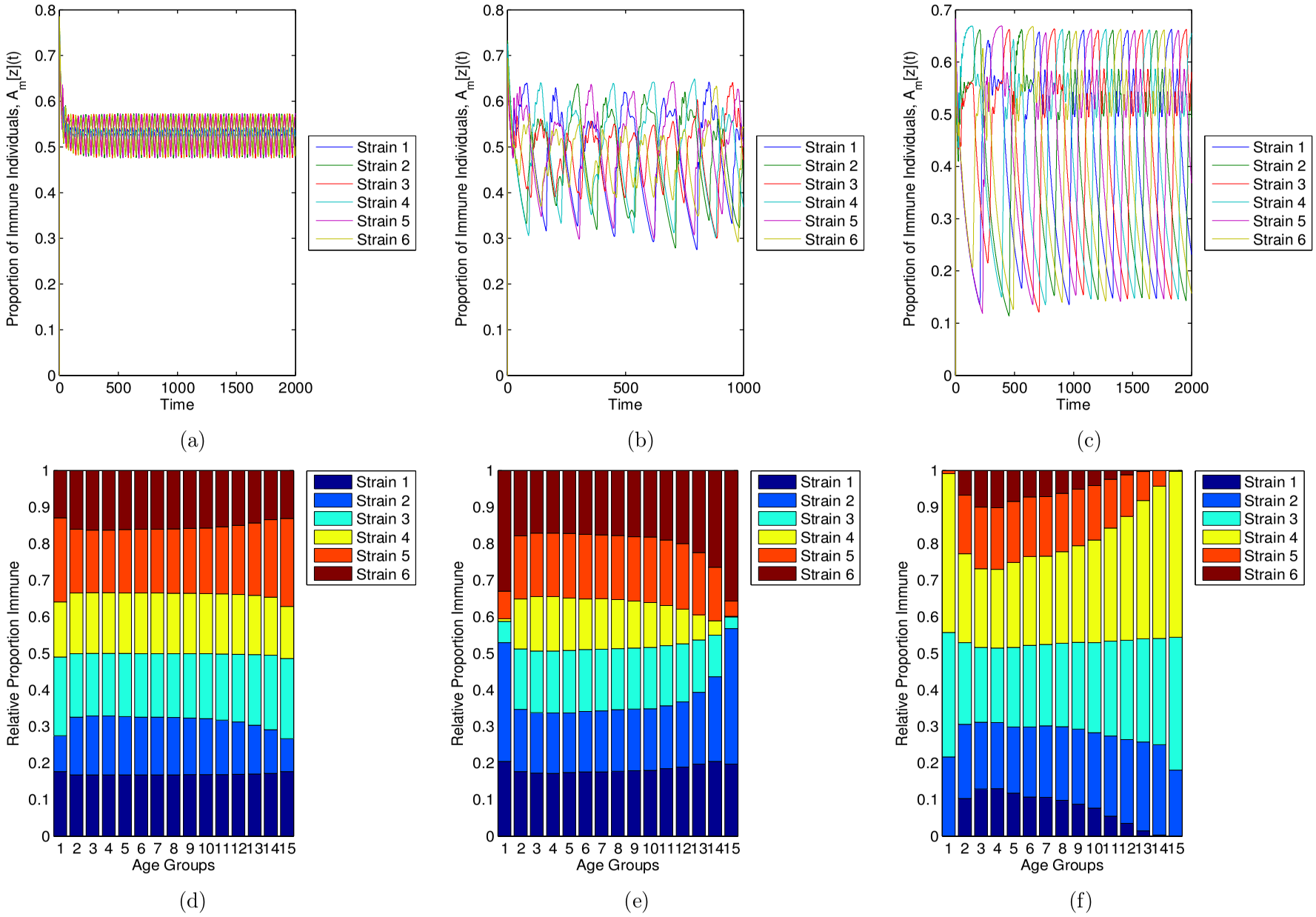


Figure 7.16: **Differential Age-specificity and CSS of the 2-locus-(2,3)-allele System.** Dynamics show the coexistence of six strains with **(a)** discrete periodic travelling wave solutions and spatio-temporal symmetry across all age-cohorts, **(b)** chaotic dynamics and **(c)** discrete large-amplitude periodic travelling wave solutions, with corresponding strain-specific scaled relative proportions of immune individuals **(d)**-**(f)**, respectively, where $\mathcal{A}_m[\phi](t) = \frac{1}{|\Omega_m|}\int_{\Omega_m}\phi(t,a)da,\ \Omega_m = [0, a_m]$. Here, the parameters are given in Section 7.2.2 and **(a)** $\gamma = 0.60$, **(b)** $\gamma = 0.75$ and **(c)** $\gamma = 0.85$; initial conditions are randomly generated.

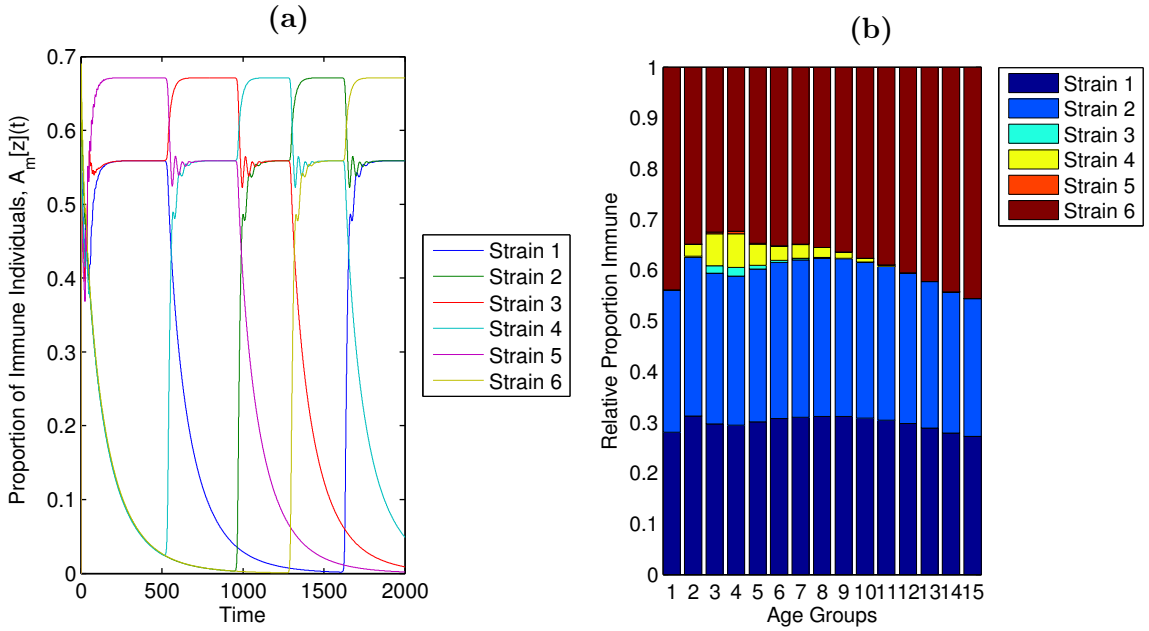


Figure 7.17: **Differential Age-specific and CSS of the 2-locus-(2,3)-allele System.** Dynamics show the coexistence of six strains with **(a)** discrete large-amplitude periodic travelling wave solutions synchronization, and **(b)** corresponding strain-specific scaled relative proportions of immune individuals, where $\mathcal{A}_m[\phi](t) = \frac{1}{|\Omega_m|}\int_{\Omega_m}\phi(t,a)da$, $\Omega_m = [0, a_m]$. Here, the parameters are given in Section 7.2.2, and $\gamma = 0.90$; initial conditions are randomly generated.

age-specific parameters. The only difference is that the abundance levels are commensurable to the age-specific reproductive success. Further analysis of these dynamic patterns at the intermediary cross-immunity level, and the complex forms of synchronization observed, may be an essential ingredient in understanding and predicting epidemiological episodes and designing effective and optimal public health schemes. In addition, for future work, we can investigate the impact of different functional forms of $\gamma(a)$ with other functional forms of $\beta(a,s)$, $\mu(a)$ and $\sigma(a)$ as compared to its age-independent analogue on the dynamic behaviours of the model described in Chapters 5-6. In Chapter 8, we will focus on the cyclical strain structure region, and discuss the role of shifting demographic and immunological landscapes with the antigenic thrift or limited epitopes hypotheses on the various epidemiological features and the serological and molecular profiles of influenza. We will present a mechanism that we believe is responsible for the various dynamic features of influenza, in particular, the emergence and re-emergence phenomena of strains bearing similar epitopes and their age-specific patterns in the host population.

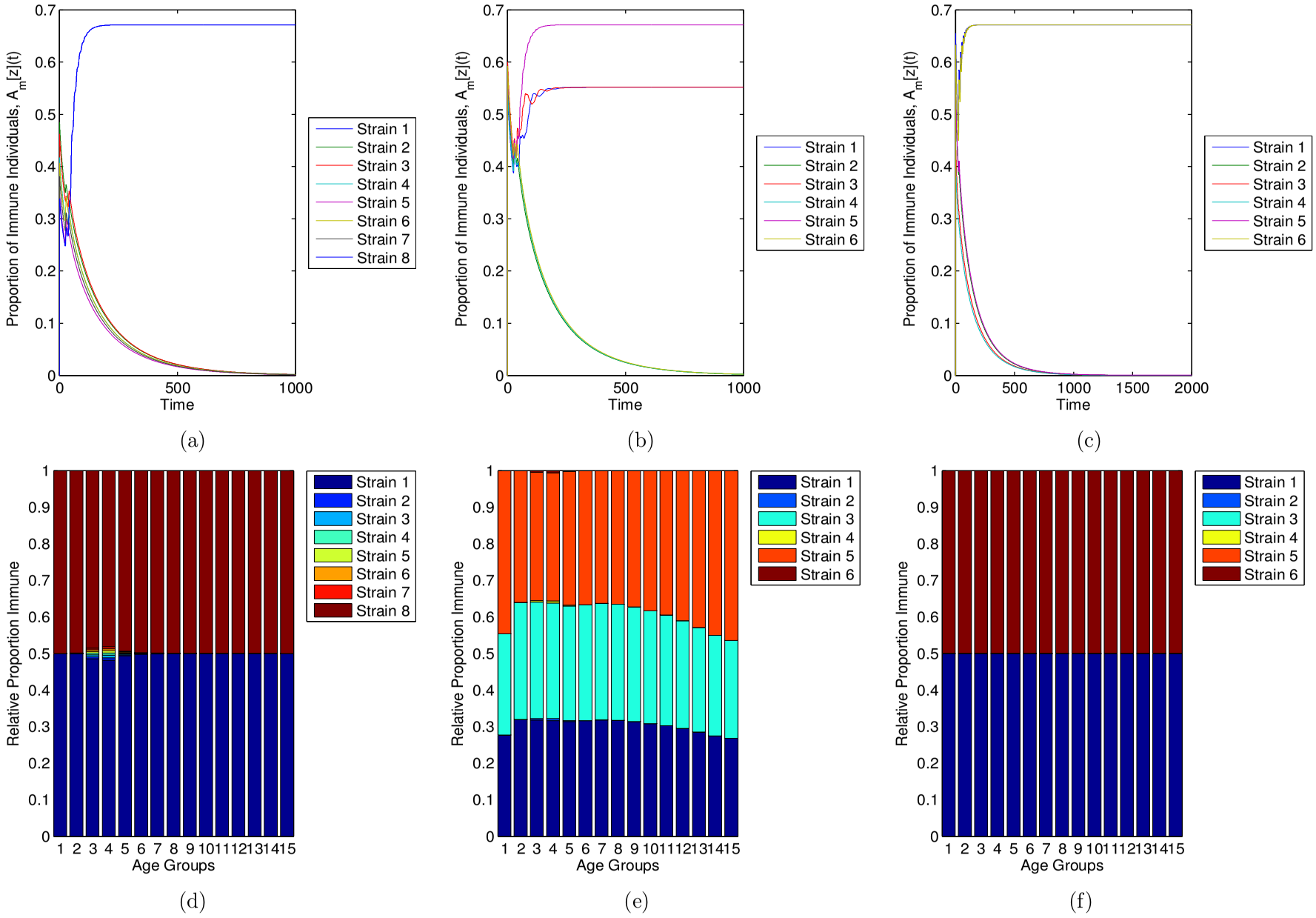


Figure 7.18: **Differential Age-specificity in DSS Dynamics.** Dynamics show the discrete strain structure, where one discordant pair persists while the other pairs go to extinction (principle of competitive exclusion) for **(a)** a 3-locus-2-allele system, **(b)**-**(c)** a 2-locus-(2,3)-allele system, with corresponding strain-specific scaled relative proportions of immune individuals across all age-cohorts **(d)**-**(f)**, respectively, where $\mathcal{A}_m[\phi](t) = \frac{1}{|\Omega_m|}\int_{\Omega_m}\phi(t,a)da$, $\Omega_m = [0, a_m]$. Here, the parameters are given in Section 7.2.2, and **(a)**-**(b)** $\gamma = 0.95$ and **(c)** $\gamma = 1.0$; initial conditions are randomly generated.

8 Interaction Between Shifting Demographic and Immunological Landscapes in Influenza Epidemics

In this chapter we discuss the role of shifting demographic and immunological landscapes under the purview of the antigenic thrift or limited epitopes hypotheses on the various epidemiological features and serological profiles of influenza. We present a mechanism that we believe is responsible for the different dynamic characteristics of the disease. The age-structured model discussed herein and in Chapters 4-7 can be used to investigate the effect of age-specific targeting for public health purposes.

8.1 Introduction

Infectious diseases, as evolutionary agents of selection, exert a substantial selective pressure on the host population, resulting in different mechanisms needed to mitigate and/or reduce the costs and/or threats posed by chronic (e.g., HIV/AIDS) and acute (e.g., influenza, measles) infectious diseases. While mathematical models have successfully been useful in providing some insights into the dynamics and control mechanisms of most diseases, we are still faced with challenges in understanding, predicting and preventing newly emerging and re-emerging pathogens, especially those with diverse antigenic forms, such as HIV and influenza, that are responsible for high mortality and morbidity and are of utmost importance to public health interests. In the past decades, almost half of the newly recognized pathogens impacting human health have been zoonotic viruses; particularly RNA viruses characterized by high mutation rates. Antigenically diverse pathogens are naturally and hierarchically more complex, and their various evolutionary patterns are more puzzling than single strain organisms. One of the major challenges in studying these antigenic variable infectious diseases is untangling the complex biological subtleties of the within-host ecological, epidemiological, immunological and evolutionary constraints and processes. Disentangling these biological complexities requires simultaneous investigation of the multiplex

relationships between the viral kinetics (i.e., the biochemistry of how viruses infect host cells), the complex molecular and cellular system biology of viral interactions with the immune systems, the evolutionary ecology of viral competitions within a single host, and the dynamics of the virus abundance and community structures in the entire population. In particular, how pathogenic genetic variation at the primary level of viruses, modulated and regulated by host immunity and transmission structures, determines the wide variety of emergent phenomena observed at different scales ranging from individuals to population. Here we focus on Influenza A viruses as a motivating organism, for there is no shortage of empirical (temporal, genetic and serological) data on the disease, which provides future opportunity in evaluating predictions made by the insights gained from the model.

Using an immuno-centric approach which does not rely on the process of mutations, Recker et al. (2007) and Wikramaratna et al. (2013) have recently put forward a new alternative mechanism challenging the notion of complete reliance on antigenic drift assumptions as the central explanatory mechanism of dynamic patterns of influenza epidemics. In Recker et al. (2007), the authors posit that, due to functional constraints on the defining epitopes (e.g., antigenic sites on HA and NA) ranging from high to low variability, the virus population is characterized by a limited set of antigenic types and is able to exploit the physio-immunological diversity and population-wide immunological profiles which constrain the fate of newly emerging and/or remerging strains. The successful viral type is one that possesses the correct combination of epitopes to find a hole in the complex landscape of prevailing epitope-specific host immune responses rather than one that has just accumulated enough mutations to become sufficiently different from the previous strains.

Unlike antigenic drift, which is dependent on the idea of a very limited access to a unlimited repertoire of antigenic types, antigenic thrift is predicated on the idea that the virus has unrestricted access to a limited repertoire of antigenic types. The consequence of this antigenic thrift assumption or limited epitopes assumption is that epitope-specific nature of immune responses prevent and/or constrain the emergence and/or re-emergence of certain antigenic variants, hence limiting the nature of future epidemics and pandemics. In this chapter, we attempt to develop a deeper understanding of the complex and puzzling relationship between evolutionary, ecological, epidemiological and immunological processes, and the patterns revealed in genomic (phylogenetic) architectures, population dynamics, as well as community structure of influenza serotypes. Motivated by the empirical sera of cross-reactivity showing significant differential age-specificity in which certain age groups have neutralizing antibodies to 2009/H1N1pdm, we address how functional constraints on defining epitopes characterized by a limited set of antigenic types determine the processes and

dynamic patterns of influenza. In particular, age-specific data and understanding the distributional pattern of immunity (individual physio-immunological diversity) to epidemic strains provide the link needed to relate the role of transmission contact structure, evolutionary (and phylogenetic) dynamics, temporal population dynamics, and serotypic community structures.

8.2 Antigenic Evolution under Changing Immunological and Demographic Landscapes

Although there have been numerous studies providing insights into the impact of antigenic variation and immunity on epidemic dynamics (Gog & Swinton, 2002; Gupta et al., 1998; Recker et al., 2007; Buckee et al., 2011; Wikramaratna et al., 2013) and showing the relationship between population connectivity structure and pathogen evolution (Read & Keeling, 2006; Sasaki & Haraguchi, 2000), we still do not understand how the effect of host immunity network (i.e., immunity acquired through prior outbreak exposures) impacts the future disease dynamics and the evolvability and variability of the pathogens. Here we discuss one mechanism under which prior exposure history interacts with contact structures to constrain the emergence and re-emergence of certain serotypes under the purview of the antigenic thrift conjecture. Figure 8.1 shows the mechanism of changing immunological and demographic profiles. In particular, we posit that infection-acquired immunity and pathogenic evolution and characteristics shape the population structures and future dynamics. That is, during an outbreak, infected individuals may acquire immunity to the same circulating strain thereby restructuring the epidemiological landscape of the population which, in turn, constrains future epidemic outbreaks and drives the pathogen's antigenic evolution allowing spread to otherwise previously immunized individuals.

Three scenarios of the effect of changing immunological landscape are shown in Fig. 8.1, namely (i) $\gamma = 0$, (ii) $\gamma = 1$ and (iii) $\gamma \in (0, 1)$. We note that scenario (i) models the absence or lack of immunity and thus is relevant when the population experiences antigenically distinct serotypes in subsequence epidemic seasons or the disease is non-immunizing, while (ii) is appropriate when the population has full immunity to subsequent serotypes after having prior exposure, or models diseases that are fully immunizing, such as measles. The last scenario (iii) represents partial immunity and models diseases that provide partial immunity to subsequent antigenically similar serotypes after prior exposures. We posit that when a particular strain appears in a naively susceptible population (i.e., stage T_0), the disease spreads through the population affecting high risk (e.g., elderly and infants) and high transmission (e.g., school age children, and/or the most connected portion of the population) age cohorts

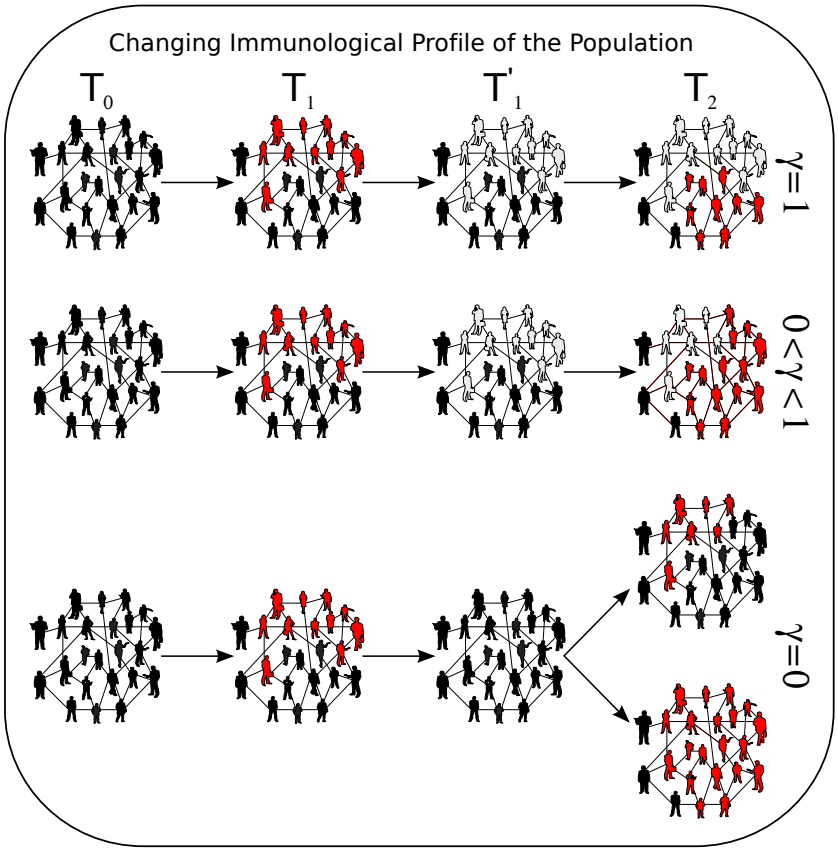


Figure 8.1: **Mechanism of Changing Immunological Landscape and Restructuring of Epidemiological Structure.** The sketch illustrates the changing structure of immunological and epidemiological profiles, where $\gamma = 1$ denotes complete immunity, $\gamma \in (0, 1)$ represents partial immunity, and $\gamma = 0$ denotes lack of immunity. Each column represents a different scenario.

(see stage T_1). After the disease passes through, if the disease is non-immunizing (i.e., $\gamma = 0$), the population, including individuals who have been exposed to the strain previously, become susceptible again (see stage $T_1^{'}$) to a subsequent appearance of previous serotype or antigenically similar strains. However, if the disease is fully (i.e., $\gamma = 1$) or partially immunizing (i.e., $\gamma \in (0, 1)$), the epidemic leaves individuals immune to that particular strain after the disease passes through (see stage $T_1^{'}$), and the remaining susceptible subpopulation will consist of less connected individuals.

For fully immunizing (i.e., $\gamma = 1$), when the homotypic pathogen re-emerges at $T_1^{'}$, it spreads through the remaining susceptible population until it no longer has a susceptible pool. Hence, the potential of a subsequent outbreak is decreased because the pathogen can only spread through these remaining susceptible groups, which are more sparse and homogeneous (see stage T_2). Thus, in order for the pathogen to spread within such a high cross-immunity structure, it must either evolve to become more transmissible and more infectious, or the population contact structure changes

through recruitment of new immunologically naive hosts, or the susceptible pool is replenished through birth. Diseases such as measles rely on this mechanism, and its age-specific seroprevalence is monotonically increasing where younger individuals have less immunity and little immunological history while elderly individuals have previous exposures and are less susceptible. As a result, measles tends to affect mostly school-age individuals. The mechanism is also manifested in the static phylodynamics of measles.

For partially immunizing diseases (i.e., $\gamma \in (0, 1)$), when a new serotype of the same pathogen enters the population that has experienced a prior epidemiological event, its transmissibility success will depend on the extent and pattern of immunological profile of the host population; and the new strain may not be able to invade unless it is antigenically different from the original strain making prior immunity irrelevant or more highly transmissible than the original strain allowing it to potentially reach previously unexposed individuals. High transmissibility can occur when the pathogen evolves to become more infectious or the contact structure changes. The lower the cross-immunity level and infectiousness of the original serotype, the more vulnerable the host population is to a subsequent antigenically similar serotype. At T_1', not only are the remaining susceptible individuals vulnerable but some of the immune individuals will be vulnerable depending on the extent of immunology history and diversity, and antigenicity of the serotypes. Heterotypic diseases such as influenza, *S. pneumoniae*, and *N. meningitides* follow this mechanism. For instance, during the initial outbreaks, age-cohorts with high contact rates have a greater risk of infection than those with smaller contact rates. In the subsequence appearance (say the second season), the moderately connected age-cohorts will have a higher risk. We observe this age-specific gradual shift during most epidemiological events of influenza until a cluster change occurs (see Bansal et al., 2010; and Dávila et al., 2014). For example, during the recrudescent wave of pandemic A/H1N1 influenza in 2009, we observe that during the first season, younger individuals have higher morbidity rate than older individuals as they possess some form of cross-immunity from prior exposures. However, in subsequence seasons, there are gradual age-specific morbidity shifts where older individuals now have a higher risk as a result of an immunological gap being formed in certain age-groups. Bansal et al. (2010) have shown this age-specific gradual shift in the attack rates during and after four influenza pandemics; and Dávila et al. (2014) have found similar gradual age-specific shifts in hospitalization and death rates in three seasons of A/H1N1/2009. As the burden of the disease cascades, immunity is accumulated across different age-cohorts creating networks of cross-protection. This network of cross-reactivity and cross-protection constrains the

emergence and re-emergence of certain antigenic types. The effect of the aforementioned mechanism on age-specificity under the purview of antigenic thrift is discussed below.

8.3 Effects of Single Strain Dominance on Age-specificity

Because influenza exhibits epidemics dominated by a single strain with erratic multi-annual cycles with irregular emergence and re-emergence (or replacement) of certain antigenic types, Recker et al. (2007) developed a measure capable of quantifying a single strain dominance by comparing the average of the relative prevalence of the two most common antigenic variants within a single epidemic. In Recker et al. (2007) and Wikramaratna et al. (2013), it was found that some allelic combinations have a broad range of cross-immunity, γ, where the the metric exceed 0.50 for the model in Chapter 3. In Chapter 7, we showed the dynamic behaviours of the model under both uniform and differential allelic variation. Here, we focus our discussion on the 32 strain-system with different allelic combinations and examine the interactions between shifting immunological and demographic landscapes. We restrict the discussion to the regions of cross-immunity that exhibit some degree of a single strain dominance. Figure 8.2 summarizes the ranges of cross-immunity over which the model shows a single strain dominance for different multi-locus-allele systems. In particular, we observe that different allelic combinations have differential domains and ranges where there is a dominance of single antigenic serotypes. Figures 8.2(c)-(d) illustrate the aforementioned single strain dominance in terms of prevalence (proportion of infected individuals), which usually occurs in chaotic regimes. However, it is hard to detect these behaviours in the proportion of the immune individuals (see Chapter 7). The dynamic features of a single strain dominance can lead to sequences of antigenically related epidemics until eventually a new discordant allelic type is favoured, resulting in antigenic cluster jumps. Furthermore, we also observe that occasionally there is coexistence of two antigenic distinct strains (see Fig. 8.2). There is empirical evidence for all these observed patterns in influenza epidemics/pandemics. For example, subsequent to the 1918 H1N1 pandemic, there were several H1N1 outbreaks occurring from the 1920s to the late 1950s, followed by a 20 year disappearance and a re-emergence in the 1970s. During the H1N1 hiatus, H2N2 and H3N2 emerged. These are consistent with the dynamic features of a single strain dominance.

Figures 8.3-8.6 show the simulation results of six different allelic variations at three different cross-immunity levels including the critical γ_m, where each allelic structure achieves maximum ε with a strong single strain dominance. Unlike in Chapter 7,

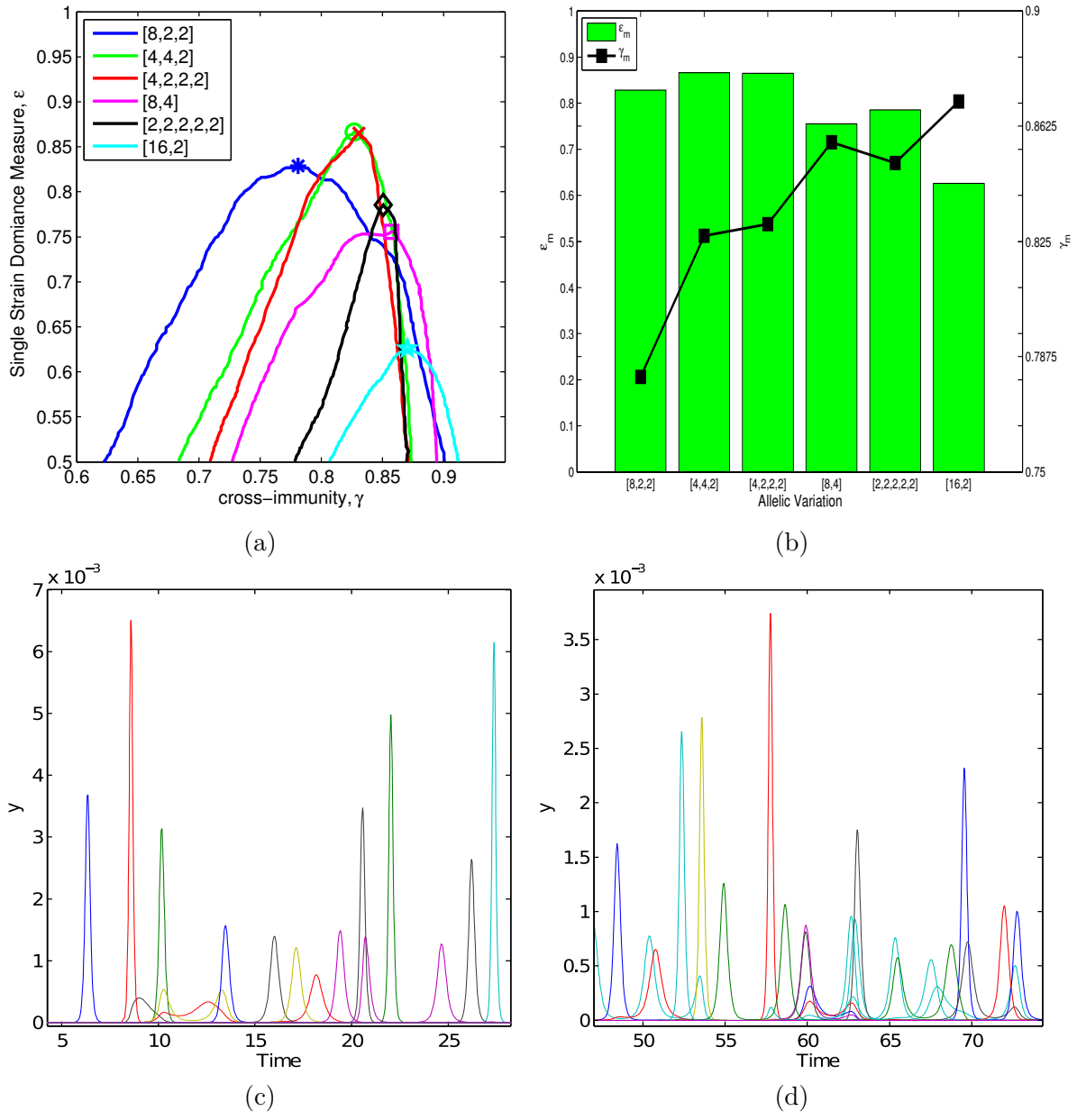
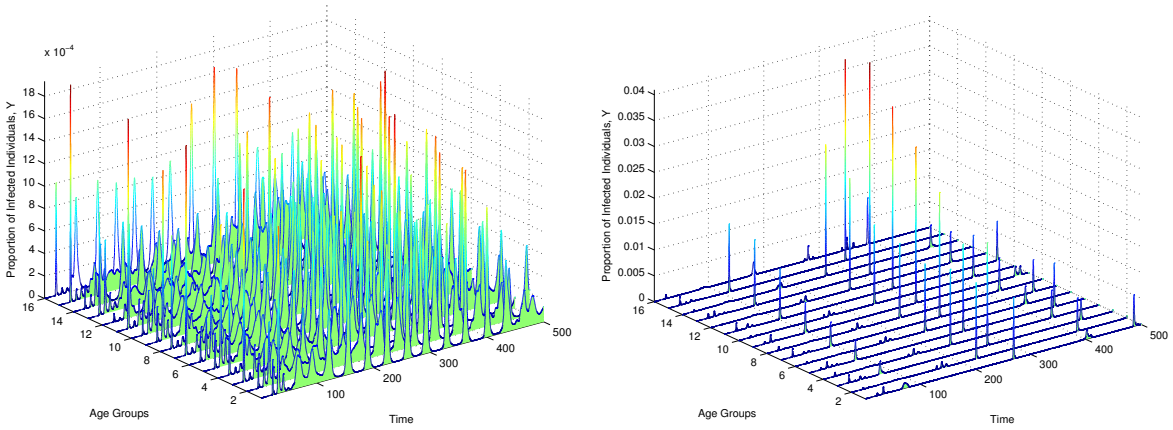


Figure 8.2: **Dominance of a Single Strain Dynamics.** **(a)** The impact of antigenic space on strong single strain dominance, ε, among different allelic combinations for a 32-strain system, where $\varepsilon(y) = \frac{1}{P}\sum_{j=1}^{N_p} \frac{y_m^j - y_s^j}{y_m^j}$ and y_m^j, y_s^j are the prevalence of the two most common antigenic variants within a single epidemic averaged over N_p epidemics (reproduced from Wikramaratna et al., 2013 using Eqs. (3.1)-(3.3) model). **(b)** The value and location where ε attains its maximum for different allelic variations. **(c)-(d)** The dynamic behaviours of single strains in terms of the proportion of infected individuals, where the parameters are given as follows: $\beta(a,s) = \beta = 292$, $\sigma = 73$, $\mu = 0.02$ and $\gamma = 0.78$ with antigenic space of $[8, 2, 2]$. The panels **(c)** and **(d)** show the same realization of prevalence where two time frames are chosen.



(a) (b)

Figure 8.3: **Temporal Evolution of an Age-specific 5-Locus-2-Allele System.** Dynamical behaviours of the model (5.1)-(5.3) in the CSS region. Here we have only included one strain and the other strains have similar dynamic behaviours. All parameter values are the same as for the age-specific cases discussed in Chapter 7, and the cross-immunity level $\gamma = 0.75$ for panel **(a)** and $\gamma = 0.85$ for panel **(b)**.

where we showed only the results of immune individuals, here we look at infected individuals. In Fig. 8.3, we show typical results of a 32-strain system in the CSS region with respect to the proportion of infected individuals. We observe that $\gamma = 0.75$ for the 5-locus-2-allele system (i.e., $[2, 2, 2, 2, 2]$ system) is outside the region where the system exhibits strong single strain dominance, while $\gamma = 0.85$ is within the region of strong single strain dominance emergence. This distinction is also reflected in the temporal evolution of the system shown in Fig. 8.3. In particular, Fig. 8.3(a) shows only chaotic behaviours for strains 1 and 2, respectively. Figure 8.3(b) exhibits single strain dominance, which is a particular type of CSS. In addition, we also observe that, for $\gamma = 0.75$, all the age-cohorts exist in the infectious compartment, although some cohorts will have very low relative proportions in relation to other cohorts. But this is not true for $\gamma = 0.85$ (see Figs. 8.3(b) . In this regime, we can observe a localized epidemic among some individual age-cohorts where some cohorts will have high relative proportions. We also observe travelling wave behaviours across age-cohorts with increasing relative proportion. There is empirical evidence supporting both behavioural patterns observed above. For instance, during the Russian flu of 1977 (H1N1) there was a localized pandemic among children with a limited infectivity to the immunologically naive younger (persons less than 25 years of age) population. The dynamic patterns discussed above are observed in all the six allelic variations concerned herein, and in the proceeding discussion we will not show these temporal evolutions. Figures 8.4-8.6 emphasize peak statistics and evolution of the relative distribution of infected and immune individuals.

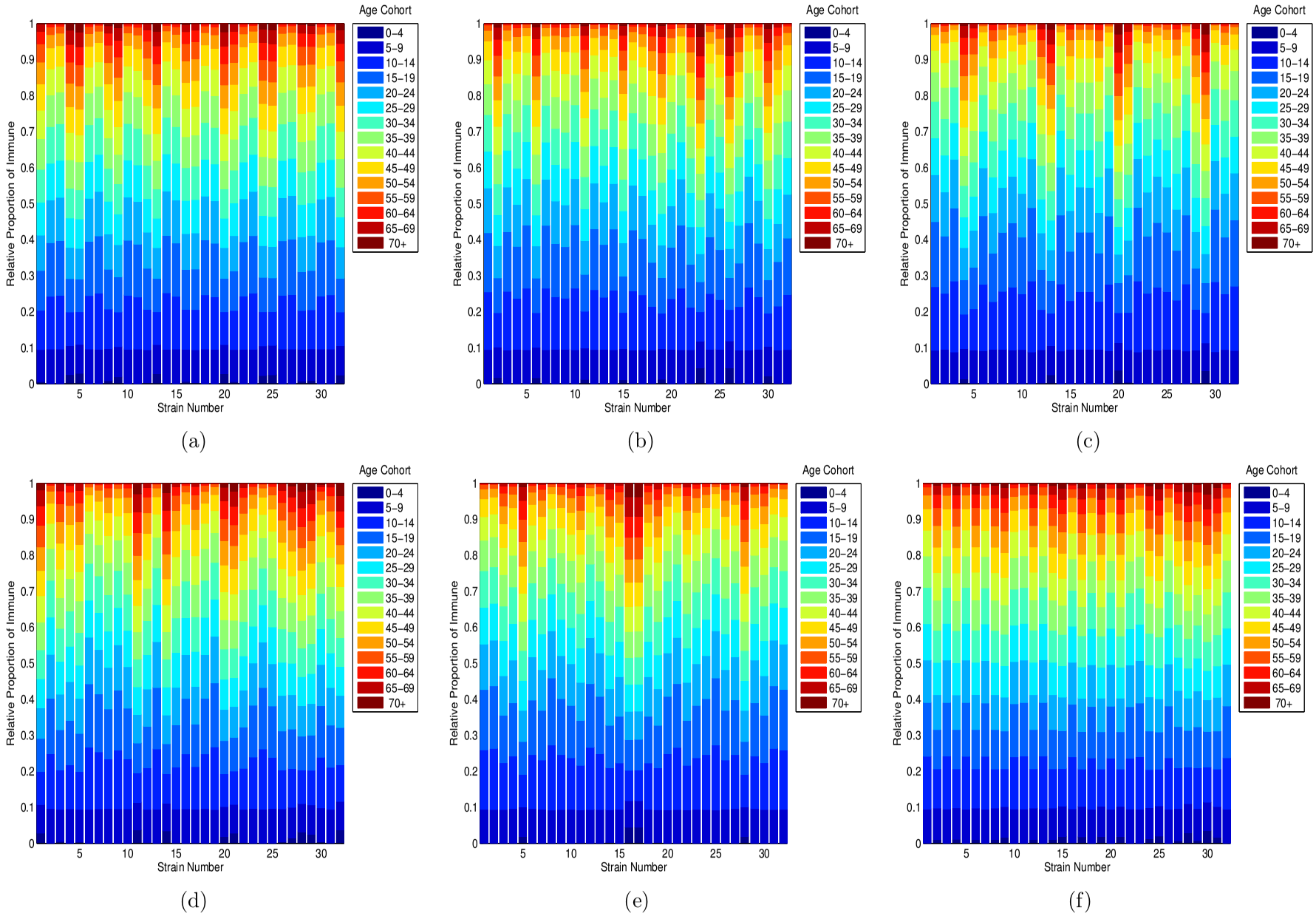


Figure 8.4: **Age-specificity in Relative Proportions of Immune Individuals for Multi-locus-allele Systems.** Illustrations of the age-specific relative proportion of immune individuals are shown for the multi-locus-allele system **(a)** $\mathcal{G}_3(8,2,2)$, $\gamma = \gamma_m = 0.78$, **(b)** $\mathcal{G}_3(4,4,2)$, $\gamma = \gamma_m = 0.83$, **(c)** $\mathcal{G}_4(4,2,2,2)$, $\gamma = \gamma_m = 0.83$, **(d)** $\mathcal{G}_2(8,4)$, $\gamma = \gamma_m = 0.86$, **(e)** $\mathcal{G}_5(2,2,2,2,2)$, $\gamma = \gamma_m = 0.85$ and **(f)** $\mathcal{G}_2(16,2)$, $\gamma = \gamma_m = 0.87$; all other age-specific parameters are discussed in Chapter 7.

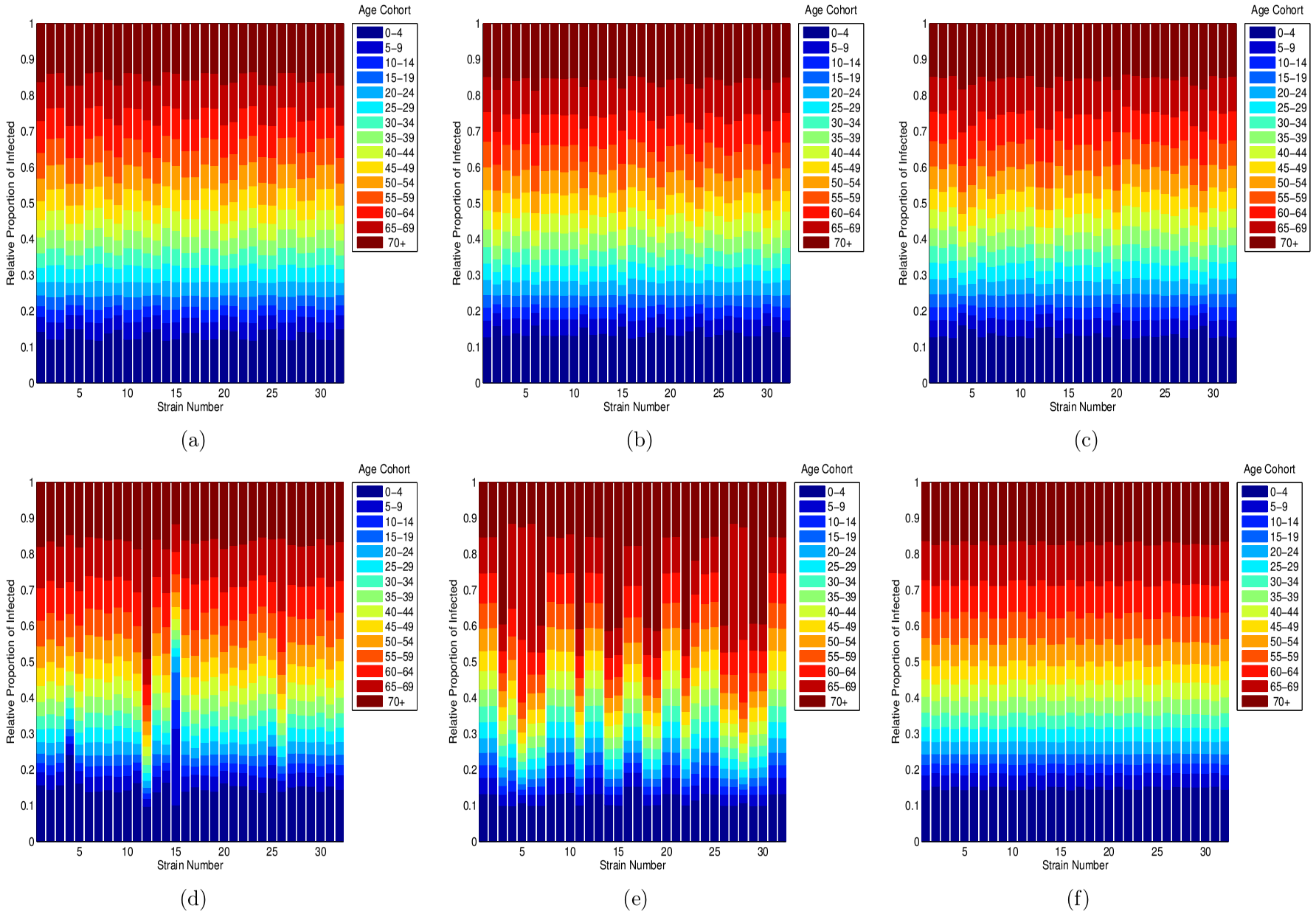


Figure 8.5: **Age-specificity in Relative Proportions of Infected Individuals for Multi-locus-allele Systems.** Illustrations of the age-specific relative proportion of infected individuals are shown for multi-locus-allele system **(a)** $\mathcal{G}_3(8,2,2)$, $\gamma = \gamma_m = 0.78$, **(b)** $\mathcal{G}_3(4,4,2)$, $\gamma = \gamma_m = 0.83$, **(c)** $\mathcal{G}_4(4,2,2,2)$, $\gamma = \gamma_m = 0.83$, **(d)** $\mathcal{G}_2(8,4)$, $\gamma = \gamma_m = 0.86$, **(e)** $\mathcal{G}_5(2,2,2,2,2)$, $\gamma = \gamma_m = 0.85$ and **(f)** $\mathcal{G}_2(16,2)$, $\gamma = \gamma_m = 0.87$; all other age-specific parameters are discussed in Chapter 7.

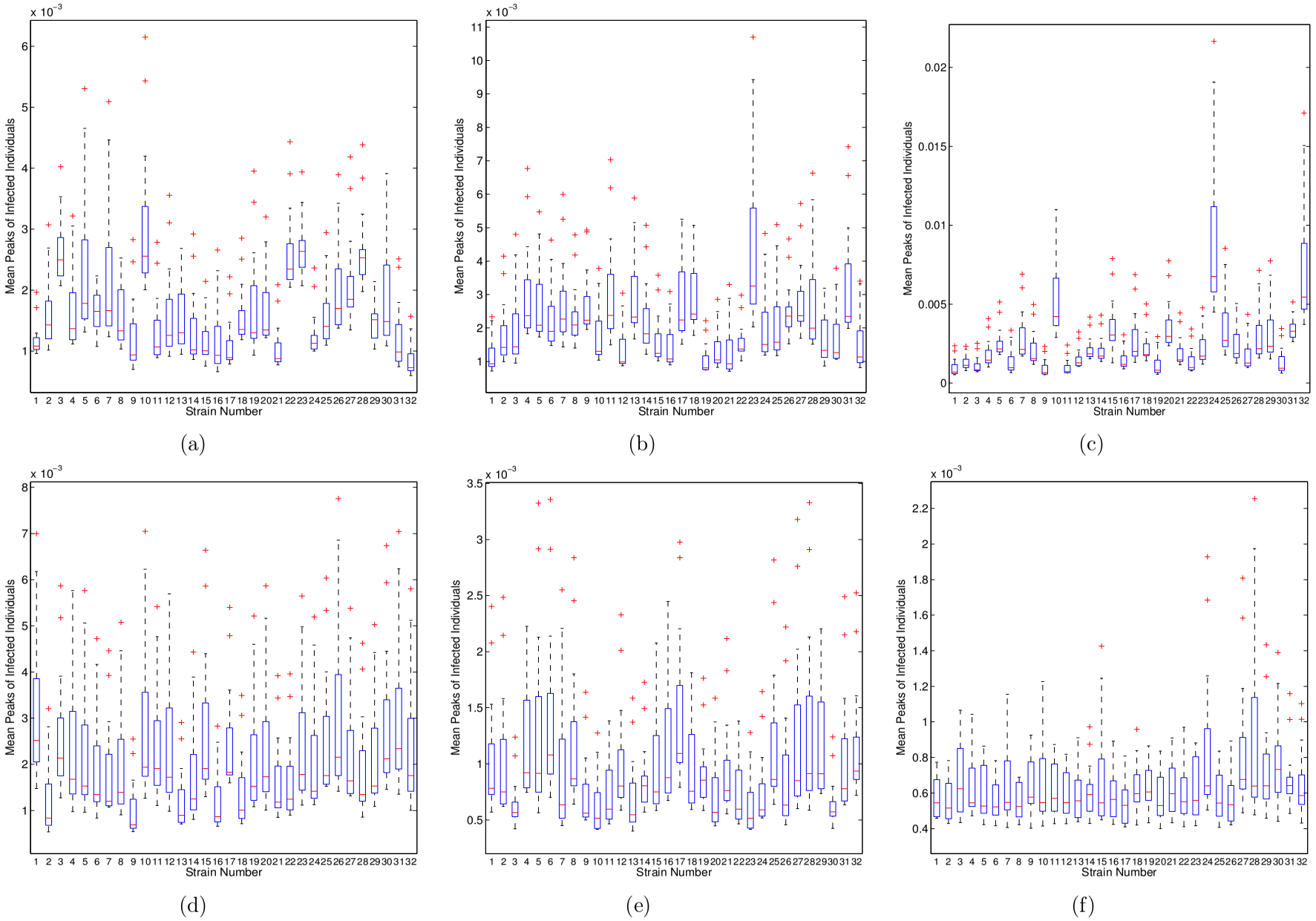


Figure 8.6: **Overall Peak-statistics for Multi-locus-allele Systems.** Illustrations of the average peaks of relative proportion of infected individuals are shown for multi-locus-allele system **(a)** $\mathcal{G}_3(8,2,2)$, $\gamma = \gamma_m = 0.78$, **(b)** $\mathcal{G}_3(4,4,2)$, $\gamma = \gamma_m = 0.83$, **(c)** $\mathcal{G}_4(4,2,2,2)$, $\gamma = \gamma_m = 0.83$, **(d)** $\mathcal{G}_2(8,4)$, $\gamma = \gamma_m = 0.86$, **(e)** $\mathcal{G}_5(2,2,2,2,2)$, $\gamma = \gamma_m = 0.85$ and **(f)** $\mathcal{G}_2(16,2)$, $\gamma = \gamma_m = 0.87$; all other age-specific parameters are discussed in Chapter 7. The blue bottom and top box represent the 1st and 3rd quartile, and the red line in the box denotes the median, end whiskers (black dash vertical lines) are the minimal and maximum of the peaks, and the pluses denotes the outliers.

We note that, in Fig. 8.6, the mean peaks are not constant across different strains. If the systems exhibit periodic patterns (e.g., anti-phase periodic solutions; discrete periodic travelling wave; large-amplitude periodic solutions), then the average peaks will be constant across all the strains (figures not shown herein). Figure 8.6 shows the full range of variation, the likely range of variation (lQR) and a typical median, where high variations suggest that the dynamic features may not be periodic (e.g., chaotic or quasi-periodic) and may present a period 2 or greater. These observations are consistent with dynamic patterns observed in the region of a strong single strain dominance. The whiskers showing low variation may suggest some features of periodicity, for dynamic patterns with small standard deviations tend to be more periodic than those exhibiting high variance. Furthermore, we observe that, in Figs. 8.4-8.5, age-cohorts with low relative proportion of immune individuals (see Fig. 8.4) have high number of infected individuals across all strains (see Fig. 8.5). However, the relationship is not proportional. These patterns are also observed in serological data. For instance, Fig. 8.7 shows the number of individuals in England who were immune before the arrival of the 2009 swine A/H1N1, where substantial numbers of people aged over 45 years had antibodies similar to the H1N1.2009 strain. In addition, we observe higher seroconversions in the young and very few in those over individuals aged 45 in early September. It should be noted that there are some infected individuals in age-cohorts 45-64 and over 65+, but the numbers were small relative to the rest. In addition, we also note that the contact structure alone is not sufficient to explain and generate the patterns shown in Figs. 8.4-8.6. In the proceeding sections, we will discuss one of the possible mechanisms, and provide more serological evidence capturing the dynamic patterns of the limited epitopes conjecture of the antigenic thrift model.

8.4 Antigenic Evolution of Pathogens: Limited Epitopes and Age-specificity

Recently, there have been various sero-epidemiological studies investigating the role of protective immunity from past infection in the population and the cross-immunity nature of influenza viruses (Allwin et al., 2010; Bandaranayake et al., 2010; Chen et al., 2010; Deng et al., 2010; Gilbert et al., 2010; Miller et al., 2010; Skowronski et al., 2010; Tandale et al., 2010; Wu et al., 2010; Zimmer et al., 2010). These serological surveys have recently been used to unravel the subtleties between the biological constraints and the epidemiological patterns of influenza viruses in humans, using particular subpopulations of the community. Titres of hemagglutinin antibodies have been used as a standard measure of protection level acquired previously, either

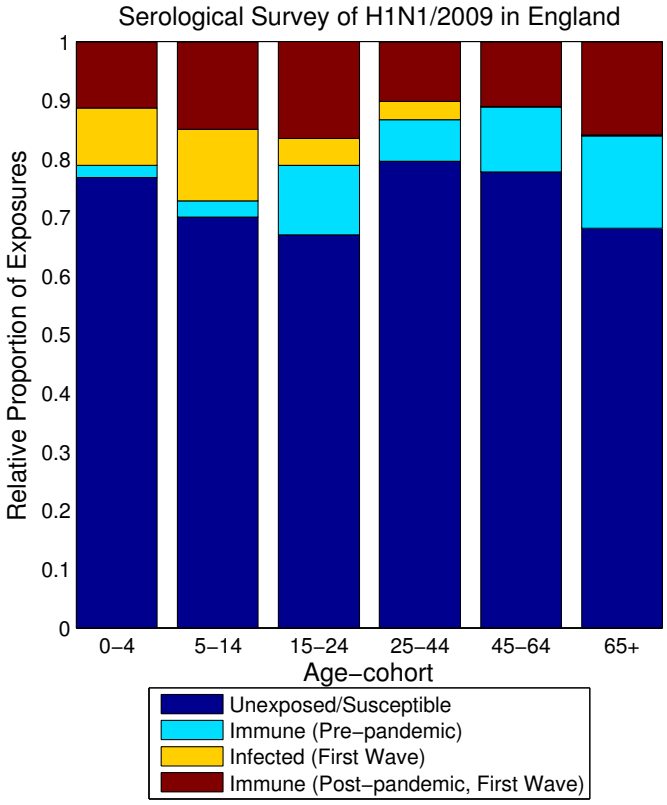


Figure 8.7: **Serological Survey by Age-cohorts.** The relative proportion of the population with different serological states is shown. The number of susceptible or unexposed individuals is calculated using serological and demographic data in each age-cohort (from England). The blue region is for the relative proportion of susceptible individuals, the cyan region for the relative proportion of individuals immune before the pandemics, the yellow region for infected individuals seroconverted between April and September of 2009, and the red region for immune individuals after the first wave (data adapted from Kubiak & McLean, 2012).

through prior infections and/or vaccinations. These studies suggest a striking characteristic profile of immunity against age, which is valid for different subtypes and across different regions. In particular, some of these studies show differential age-specificity in seroprevalence, where some individuals above certain age cohort have neutralizing antibodies to certain serotypes (see Chapter 2 for more details and for review of differential age-specificity in seroprevalence and cross-reactivity in human sera). This seems to explain the low infection rates in the elderly during the recent pandemic of H1N1/2009.

8.5 Public Health Policy based on Dynamic Immunological Profiles of the Population

Mathematical studies in epidemiology attempt to reconstruct epidemiological patterns of diseases and pathogen evolution in order to gain insights needed to provide effective public health measures. Because there are demographic and strain-specific differential mortality, morbidity and attack rates, conventional insights from epidemiological and theoretical studies suggest that school-age and elderly individuals have the highest attack rates and fuel transmission, and highest mortality rates, respectively. Vaccination of school-age individuals (e.g., high transmission age group) has been shown to be an effective control strategy (CDC 2009a-b), for immunizing them can provide herd immunity and break potential transmissibility before the rest of the population is affected. As a result, contact-based, morbidity-based and mortality-based vaccination policies and other intervention strategies are often used in conjunction with prioritization of critical personnel. However, most modern public health intervention strategies ignore the immunological history of the population. Under the purview of the antigenic thrift conjecture of limited epitopes, we suggest differential/dynamic vaccination prioritization and intervention strategies depending on the strain-specific immunological profile of the population. From a public health perspective, individuals with high mortality risks and critical personnel should always have primary prioritization. However, in the context of this thesis, we focus on the secondary prioritization, especially in the face of limited resources.

Under the purview of the antigenic thrift conjecture, vaccination policies should be differential depending on the distributional structure of the population immunity. In the face of a new infection targeting a population with no prior infectious history, the conventional public health measures based on contact, mortality and/or morbidity will suffice, and risk groups (e.g., elderly and infants, school-age children) with the highest potential for becoming infected and transmitting should be vaccinated. However, for seasonal epidemic with antigenically similar serotypes, the highest risk age groups may shift. For example, during the Spanish flu of 1918 (H1N1), the highest risk age group were young adults (20-40years of age), for Asian flu of 1957 (H2N2) and Hong Kong flu of 1968 (H3N2), elderly and infants were affected, for Russian flu of 1977 (H1N1) there was a localized pandemic among children with limited infectivity to the immunologically naive younger (persons less than 25 years of age) population, and for Mexican swine flu of 2009 (H1N1) there was also age-specific disparities in prevalence and severity of infection where older age groups were less susceptible to the disease due to prior exposures to strains with antigenically similar epitopes. From these aforementioned examples, conventional public health

intervention strategies based on contacts may not be effective and/or optimal, for previous infection-acquired immunity may restructure the immunological profile of the population in such a way that highly connected individuals may no longer be the most vulnerable or most likely to transmit infection. Thus, differential immunological profile-based strategy may depend on the recent immunological diversity and history of the population, and knowing and exploring the changing immunological structure (against a particular antigenic variant) may provide both direct and indirect protection by achieving the most optimal herd immunity, for the prioritization of high risk (elderly and infants) and high transmission age cohorts may sometimes overlap. For instance, from the serological survey shown in Fig. 1.2 (see Hancock et al., 2009), we observe that individuals above a certain age have neutralizing antibodies to the Mexican swine flu of 2009 (H1N1) in which the most intense recognition is among those born prior to 1930. This suggests that strains carrying similar epitopes were in circulation. As a result, older were less susceptible to the disease. Consequently, prioritizing the vaccination of older age cohort will not be effective. In this case, younger individuals should receive the highest secondary prioritization. However, in the subsequent years, this prioritization of younger individuals could change depending on the immunological profile and the level of cross-immunity in the population. It has been observed in several studies that during subsequent epidemics there are gradual age-specific morbidity shift where older individuals now have higher risk as a result of immunological gap being formed in certain age-groups (see previous sections).

8.6 Conclusions

In this chapter, we have outlined an additional mechanism involved in shaping both the immunological diversity and population-wide immunological profiles under the assumption of antigenic thrift, and discussed how the interaction between the changing immunological and demographic profiles and network of cross-protection impact the evolution and constrain the fate of emergence and re-emergence of particular antigenic variants. Using various age-specific empirical serological studies, we show that there is mounting evidence supporting both the antigenic thrift conjecture, and the preferential effects of changing immunological and demographic profiles on the observed dynamic of influenza. In addition, we discussed the differential public health intervention strategy in preventing and controlling the spread of the influenza virus. Besides the preliminary goal of preventing and controlling the proliferation of the disease, we note that the differential public health intervention strategy presented above provides a secondary benefit. Because it requires the collection of serological

data on each seasonal influenza strain, having these detailed data might help us improve our understanding of seasonal epidemics, and further allow us to disentangle how individuals build immunity to an antigenically variable infectious disease like influenza, and find correlation between the immunological structure of the population and emergence and re-emergence of certain antigenic serotypes. The model can be used to investigate the effect of age-targeting. It will be worth investigating the effect of cross-immunity and differential immunological profile on the optimality of a public health intervention. This is left as a further research direction to pursue. In the next chapter, we conclude with closing remarks.

9 Closing Remarks

In this chapter, we summarize previous chapters, discuss possible extensions to models discussed in earlier chapters, and discuss some possible broadening future directions.

9.1 Summary

In this study, we have used mathematical modelling as a tool to investigate the mechanisms underlying certain aspects of antigenic evolution of pathogens and their impacts on the evolution and maintenance of strain community structures and the various dynamic behaviours, such as temporal dynamics, age-specificity as well as phylodynamic patterns, observed in antigenically variable pathogens with particular emphasis on influenza viruses. With recent serological evidence, previous mathematical studies had increasingly become unable to explain the mechanisms involved in serological signatures of antigenic evolution or to explain the patterns of antigenic evolution without imposing stringent conditions on various factors (e.g., mode of mutation, temporal strain-transcending immunity).

In Chapters 1-2, we began our study with a brief description of the historical use of mathematical and theoretical modelling (e.g., dynamical and statistical approaches) as complementary tools in public health and epidemiology in order to understand various biological and ecological processes, where in Chapter 1 we motivate the aim of the thesis. In particular, we emphasized the alternative hypothesis of limited epitopes to explain the epidemiological, evolutionary and demographic patterns observed in pathogens such as influenza with diverse antigenic form. Using influenza virus as a motivating application in Chapter 2, with recent new empirical evidence (Hancock et al., 2009) showing differential age-specificity, we introduced the background on various underlying biological assumptions that have been put forward to explain the dynamics of influenza, both temporally and evolutionarily. We focused on the hypothesis of antigenic variation (e.g., antigenic drift and shift) where a range of biological evidence was presented, and we then discussed an alternative process of the limited epitopes hypothesis or antigenic thrift in Influenza A viruses that has recently been put forth to explain the epidemiological, evolutionary and demographic patterns observed in

pathogens with diverse antigenic forms such as flu. Moreover, we supported the antigenic thrift hypothesis with recent serological studies and biological constraints on the immune systems of the host.

In Chapters 3-6, we characterized the behaviours. In particular, we investigated the local stability of the equilibria (e.g., disease-free and endemic equilibria), and showed that the system exhibits two notions of endemicity, namely strong and weak endemicity. In the case of strong endemicity, we showed that the steady state can either approach an equilibrium point where all strains are in equal abundance and fully synchronized, or exhibit cyclical behaviours where a diverse dynamics with subtle cluster-specific dynamics and complex forms of synchronization are observed. In Chapters 7-8, we ended with with numerical studies of various strain spaces where we observed that the dynamic features discussed in 2-6 are preserved. In addition, we fitted four serological data sets to the model, and discussed public health implications

9.2 Future Directions

There are myriad of avenues open to further (theoretical and biological) explorations. Throughout this thesis, we have outlined some possible extensions to the models presented in Chapters 3–8, and we have also pointed to future challenges in each of these chapters. As a result, we will not repeat those challenges herein. Other venues of future investigation include developing accurate and reliable inference-based forecasts of the influenza outbreak features (e.g., temporal and antigenic evolutions) and future epidemics (short-range and long-range scientific prediction) using historical data (e.g., time series, serological, genetic).

A Chapter 3 Appendix

A.1 Proof of Lemmas and Theorems

A.1.1 Proof of Theorem 3.4.1

Proof of Lemma 3.4.1. To determine the linear stability of the disease-free equilibrium $\mathcal{E}_0$, we calculate the Jacobian matrix evaluated at the disease-free equilibrium point. Here the Jacobian matrix is given as:

$$\mathcal{J} := \begin{bmatrix} -\mu\mathbf{Id}_{m^2\times m^2} & \mathbf{0}_{m^2\times m^2} & \beta\mathbf{Id}_{m^2\times m^2} \\ \mathbf{0}_{m^2\times m^2} & -\mu\mathbf{Id}_{m^2\times m^2} & \beta\mathbf{Id}_{m^2\times m^2} \\ \mathbf{0}_{m^2\times m^2} & \mathbf{0}_{m^2\times m^2} & (\beta-\sigma-\mu)\,\mathbf{Id}_{m^2\times m^2} \end{bmatrix},$$

where $m = 2$. We note the upper right triangular structure of the matrix. It suffices to show that the Jacobian matrix above has all its eigenvalues with negative real part, which in this case are determined by its diagonal entries. The eigenvalues for the first $2m^2$ entries are given by $\zeta_{1,2,\ldots,2m^2} = -\mu < 0$, while the remaining eigenvalues are $\zeta_{2m^2+1,2m^2+2,\ldots,3m^2} = \beta - \sigma - \mu$. Therefore, $\mathcal{E}_0$ is locally asymptotically stable if $\zeta_{2m^2+1,2m^2+2,\ldots,3m^2} < 0$. That is, $\mathcal{E}_0$ is linearly stable provided $\mathcal{R}_0 = \frac{\beta}{\sigma+\mu} < 1$. This completes the proof. □

A.1.2 Proof of Theorem 3.4.2

Proof. To show $\mathcal{E}_0$ is globally stable, we employ a differential inequality approach. It suffices to show that $y_{ij}(t) \to 0$ as $t \to \infty$ provided $\mathcal{R}_0 < 1$. Note that,

$$\begin{aligned} \frac{dy_{ij}}{dt} &= \beta\left[(1-w_{ij}) + (1-\gamma)(w_{ij}-z_{ij})\right] y_{ij} - (\sigma+\mu)y_{ij} \\ &= \beta\left[(1-z_{ij}) - \gamma(w_{ij}-z_{ij})\right] y_{ij} - (\sigma+\mu)y_{ij} \\ &\leq \beta(1-z_{ij})y_{ij} - (\sigma+\mu)y_{ij} \leq (\beta-\sigma-\mu)\, y_{ij}, \end{aligned}$$

where $w_{ij} > z_{ij}$ as a consequence of the positivity of Eqs. (3.7)-(3.9) and from the fact that $w_{ij}(t) = \bigcup_{<i,j>} z_{ij}(t)$ where $< i, j >$ means strains sharing alleles with ij. Hence, $y_{ij}(t) \leq y_{ij}(0)e^{(\beta-\sigma-\mu)t}$. Noting that $\beta - \sigma - \mu < 0$ implies that $\mathcal{R}_0 < 1$ and $y_{ij}(t)$ is decreasing exponentially. However, from Theorem 3.3.1, $y_{ij} \in [0, 1]$. Therefore

$\lim_{t\to\infty} y_{ij}(t) \to 0$ for all i, j. Therefore, y_{ij} converges to zero. So the limiting system under $\mathcal{R}_0 < 1$ is $\frac{dz_{ij}}{dt} \leq -\mu z_{ij} < 0$ and $\frac{dw_{ij}}{dt} \leq -\mu w_{ij} < 0$, where the solutions are given as $z_{ij}(t) \leq z_{ij}(0)e^{-\mu t}$ and $w_{ij}(t) \leq w_{ij}(0)e^{-\mu t}$. Hence, as $t \to \infty$, we have $\lim_{t\to\infty} z_{ij} \to 0$, and $\lim_{t\to\infty} w_{ij} \to 0$. Hence $(z_{ij}, w_{ij}, y_{ij})_{i,j=1,2} \to (0, 0, 0)$ as $t \to \infty$. This completes the proof. □

A.1.3 Proof of Theorem 3.4.4

Proof of Theorem 3.4.4. We calculate the Jacobian matrix and observe the following triangular structure $\mathcal{J} := \begin{bmatrix} \mathcal{J}_u & * \\ 0 & \mathcal{J}_l \end{bmatrix}$, where the entries are given by the following matrices:

$$\mathcal{J}_u := \begin{bmatrix} -\mu\mathcal{R}_0 & 0 & \beta\mathcal{R}_0^{-1} \\ 0 & -\mu\mathcal{R}_0 & \beta\mathcal{R}_0^{-1} \\ -(1-\gamma)(\mathcal{R}_0-1) & -\gamma\mu(\mathcal{R}_0-1) & 0 \end{bmatrix},$$

$$\mathcal{J}_l := \begin{bmatrix} -\mu & 0 & \beta \\ 0 & -\rho\mu(\mathcal{R}_0-1)-\mu & \frac{\rho\beta}{\rho(\mathcal{R}_0-1)+1} \\ 0 & 0 & \beta\left(1-\gamma\left[\frac{\rho(\mathcal{R}_0-1)}{\rho(\mathcal{R}_0-1)+1}\right]\right)-\sigma-\mu \end{bmatrix}.$$

Since $\mathcal{J}$ is a triangular block 6×6 matrix, it suffices to consider each block matrix individually. We calculate the characteristic polynomial $P(\zeta) = \zeta^3 + \mathcal{A}_1\zeta^2 + \mathcal{A}_2\zeta + \mathcal{A}_3$ for the upper block matrix $\mathcal{J}_u$, where $\mathcal{A}_1 = 2\mu\mathcal{R}_0 \geq 0$, $\mathcal{A}_2 = (\mu\mathcal{R}_0)^2 + (\mu\mathcal{R}_0)^{-1}\mathcal{A}_3 \geq 0$ and $\mathcal{A}_3 = \mu\beta(\mathcal{R}_0-1)[\mu\gamma + (1-\gamma)] \geq 0$. All the coeffcients $\mathcal{A}_j,\ j = 1, 2, 3$, are non-negative. By the Routh-Hurwitz criterion, it suffices to show that $\mathcal{A}_1\mathcal{A}_2 - \mathcal{A}_3 > 0$. That is, $\mathcal{A}_1\mathcal{A}_2 - \mathcal{A}_3 = 2(\mu\mathcal{R}_0)^3 + \beta\mu(\mathcal{R}_0-1)[\mu\gamma + (1-\gamma)] > 0$. Hence, $\mathcal{J}_u$ has all its eigenvalues with negative real part for $\gamma \in [0, 1]$ and $\mathcal{R}_0 > 1$. Therefore, it suffices to show that the DSS weak endemic equilibrium is linearly stable if the lower block matrix $\mathcal{J}_l$ has all its eigenvalues with negative real part. To show this, we observe that $\mathcal{J}_l$ is a triangular matrix with its eigenvalues determined by its diagonal entries. For the lower left block, notice that the eigenvalues are $\zeta_1 = -\mu < 0$, $\zeta_2 = -\rho\mu(\mathcal{R}_0-1) - \mu < 0$ and $\zeta_3 = \beta\left(1-\gamma\left[\frac{\rho(\mathcal{R}_0-1)}{\rho(\mathcal{R}_0-1)+1}\right]\right) - \sigma - \mu$. It suffices to show that $\mathcal{E}_1$ is locally asymptotically stable if $\zeta_3 < 0$. That is, $\mathcal{E}_1$ is locally stable provided $\gamma > \overline{\gamma} := \frac{1+\rho(\mathcal{R}_0-1)}{\rho\mathcal{R}_0}$. This completes the proof. □

A.1.4 Proof of Theorem 3.4.7

Proof. The proof follows directly from the previous calculations. That is, we note that the characteristic polynomials $P_k(\zeta)$ have positive coefficient $\mathcal{A}_{k,l},\ k = 1, 2, 3, l = 0, 1, \ldots 3$ and their Hurwitz determinants $\mathcal{A}_k$ are positive for $k = 1, 2, 3$. Therefore, the characteristic polynomials $P_k(\zeta),\ k = 1, 2, 3$ have all their eigenvalues with negative

real parts. It suffices to show that the coefficients $\mathcal{A}_{4,2}$ and $\mathcal{A}_{4,3}$, and Hurwitz determinant, $\mathcal{A}_4$ are positive for $P_4(\zeta)$. We observe that $\mathcal{A}_{4,0} > 0$ and $\mathcal{A}_{4,1} > 0$. Hence, the coefficients $\mathcal{A}_{4,2}$, $\mathcal{A}_{4,3}$, and the Hurwitz determinant $\mathcal{A}_4$ are positive provided that $\gamma < \overline{\gamma}_1(\gamma, \mathcal{R}_0)$, $\gamma > \overline{\gamma}_2(\gamma, \mathcal{R}_0)$ and $\gamma < \overline{\gamma}_3(\gamma, \mathcal{R}_0)$, respectively. The instability follows when the conditions do not hold. This completes the proof. □

A.2 Other Calculations

A.2.1 Monotonicity of the Infected Steady State, y^*

It is obvious that as γ increases, Γ increases. Hence y^* decreases as illustrated in Fig. 3.3. That is:

$$\begin{aligned}\frac{\partial y^*}{\partial \gamma} &= \frac{\mu}{2\alpha\beta}\Gamma_\gamma\left(\frac{\Gamma}{\sqrt{4\alpha\left(\mathcal{R}_0-1\right)+\Gamma^2}}-1\right) = \frac{\mu}{2\alpha\beta}\Gamma_\gamma\left(\frac{\Gamma-\sqrt{4\alpha\left(\mathcal{R}_0-1\right)+\Gamma^2}}{\sqrt{4\alpha\left(\mathcal{R}_0-1\right)+\Gamma^2}}\right)\\ &= -\frac{\mu}{2\alpha\beta}\frac{\Gamma_\gamma}{\sqrt{4\alpha\left(\mathcal{R}_0-1\right)+\Gamma^2}}y^* < 0, \quad \text{where } \Gamma_\gamma = \mathcal{R}_0\left(\alpha-1\right).\end{aligned}$$

A.2.2 No Strain Structure Characteristic Polynomial $P_k(\zeta)$

The characteristic polynomial function for the strongly endemic or coexistence equilibrium is of the form of $P(\zeta) = \prod_{k=1}^{4} P_k(\zeta)$, where $P_k(\zeta)$ is a cubic polynomial function of the form $P_k(\zeta) = \sum_{l=0}^{3} \mathcal{A}_{k,3-l}\zeta^l$ and $\mathcal{A}_{k,0} = 1$ for all $k = 1, 2, 3, 4$. Moreover, the coefficients are given as follows:

$$\begin{aligned}
\mathcal{A}_{1,0} &= \mathcal{A}_{2,0} = \mathcal{A}_{3,0} = \mathcal{A}_{4,0} = 1 > 0, && \text{(A.1)}\\
\mathcal{A}_{1,1} &= \mathcal{A}_{2,1} = \mathcal{A}_{3,1} = \mathcal{A}_{4,1} = \beta\left(1+\alpha\right)y^* + 2\mu > 0, && \text{(A.2)}\\
\mathcal{A}_{1,2} &= \mathcal{A}_{2,2} = \left(\beta y^* + \mu\right)\left(\beta\alpha y^* + \mu\right) + \beta^2\gamma y^*\left(1-w^*\right)\\
&\quad +\beta^2 y^*\left(1-\gamma\right)\left(1-z^*\right) > 0, && \text{(A.3)}\\
\mathcal{A}_{1,3} &= \mathcal{A}_{2,3} = \beta^2\gamma y^*\left(1-w^*\right)\left(\beta y^* + \mu\right)\\
&\quad +\beta^2\left(1-\gamma\right)y^*\left(1-z^*\right)\left(\beta\alpha y^* + \mu\right) > 0, && \text{(A.4)}\\
\mathcal{A}_{3,2} &= \left(\beta y^* + \mu\right)\left(\beta\alpha y^* + \mu\right) + \beta^2\gamma\alpha y^*\left(1-w^*\right)\\
&\quad +\beta^2 y^*\left(1-\gamma\right)\left(1-z^*\right) > 0, && \text{(A.5)}\\
\mathcal{A}_{3,3} &= \beta^2\gamma\alpha y^*\left(1-w^*\right)\left(\beta y^* + \mu\right)\\
&\quad +\beta^2\left(1-\gamma\right)y^*\left(1-z^*\right)\left(\beta\alpha y^* + \mu\right) > 0 && \text{(A.6)}\\
\mathcal{A}_{4,2} &= \left(\beta y^* + \mu\right)\left(\beta\alpha y^* + \mu\right) - \beta^2\gamma y^*\left(1-w^*\right)\\
&\quad +\beta^2 y^*\left(1-\gamma\right)\left(1-z^*\right) && \text{(A.7)}\\
\mathcal{A}_{4,3} &= \beta^2\gamma y^*\left(1-w^*\right)\left(\beta y^* + \mu\right)
\end{aligned}$$

$$-\beta^2\left(1-\gamma\right)y^*\left(1-z^*\right)\left(\beta\alpha y^*+\mu\right). \tag{A.8}$$

In addition, the Hurwitz determinants, $\mathcal{A}_k,\ k=1,2,3,4$ for a cubic characteristic polynomial are given by

$$\begin{aligned}
\mathcal{A}_1 &= \mathcal{A}_{1,1}\mathcal{A}_{1,2}-\mathcal{A}_{1,0}\mathcal{A}_{1,3}=\left(\beta\left(1+\alpha\right)y^*+2\mu\right)\left[\left(\beta y^*+\mu\right)\right.\\
&\quad\left.\times\left(\beta\alpha y^*+\mu\right)\right]+\beta^2\gamma y^*\left(1-w^*\right)\left(\beta\alpha y^*+\mu\right)\\
&\quad+\left(\beta^2 y^*\left(1-\gamma\right)\left(1-z^*\right)\right)\left(\beta y^*+\mu\right)>0, \qquad \text{(A.9)}\\
\mathcal{A}_2 &= \mathcal{A}_{2,1}\mathcal{A}_{2,2}-\mathcal{A}_{2,0}\mathcal{A}_{2,3}=\left(\beta\left(1+\alpha\right)y^*+2\mu\right)\left[\left(\beta y^*+\mu\right)\right.\\
&\quad\left.\times\left(\beta\alpha y^*+\mu\right)\right]+\beta^2\gamma y^*\left(1-w^*\right)\left(\beta\alpha y^*+\mu\right)\\
&\quad+\left(\beta^2 y^*\left(1-\gamma\right)\left(1-z^*\right)\right)\left(\beta y^*+\mu\right)>0, \qquad \text{(A.10)}\\
\mathcal{A}_3 &= \mathcal{A}_{3,1}\mathcal{A}_{3,2}-\mathcal{A}_{3,0}\mathcal{A}_{3,3}=\left(\beta\left(1+\alpha\right)y^*+2\mu\right)\left[\left(\beta y^*+\mu\right)\right.\\
&\quad\left.\times\left(\beta\alpha y^*+\mu\right)\right]+\beta^2\alpha\gamma y^*\left(1-w^*\right)\left(\beta\alpha y^*+\mu\right)\\
&\quad+\left(\beta^2 y^*\left(1-\gamma\right)\left(1-z^*\right)\right)\left(\beta y^*+\mu\right)>0, \qquad \text{(A.11)}\\
\mathcal{A}_4 &= \mathcal{A}_{4,1}\mathcal{A}_{4,2}-\mathcal{A}_{4,0}\mathcal{A}_{4,3}=\left(\beta\left(1+\alpha\right)y^*+2\mu\right)\left[\left(\beta y^*+\mu\right)\right.\\
&\quad\left.\times\left(\beta\alpha y^*+\mu\right)\right]-\beta^2\gamma y^*\left(1-w^*\right)\left[\beta(\alpha+1)y^*+2\mu+\beta y^*+\mu\right]\\
&\quad+\left(\beta^2 y^*\left(1-\gamma\right)\left(1-z^*\right)\right)\left[\beta(\alpha+1)y^*+2\mu+\beta\alpha y^*+\mu\right]. \qquad \text{(A.12)}
\end{aligned}$$

It should be noted that w^*, z^* and y^* are functions of both γ and $\mathcal{R}_0$. In order to find the critical γ^* satisfying conditions for $\overline{\gamma}_r(\gamma,\mathcal{R}_0)\ r=1,2,3$, where

$$\begin{aligned}
\overline{\gamma}_1(\gamma,\mathcal{R}_0) &= \frac{\left(\beta y^*+\mu\right)\left(\beta\alpha y^*+\mu\right)+\beta^2 y^*\left(1-z^*\right)}{\beta^2 y^*\left(1-w^*\right)+\beta^2 y^*\left(1-z^*\right)},\\
\overline{\gamma}_2(\gamma,\mathcal{R}_0) &= \frac{\left(1-z^*\right)\left(\beta\alpha y^*+\mu\right)}{\left(1-w^*\right)\left(\beta y^*+\mu\right)+\left(1-z^*\right)\left(\beta\alpha y^*+\mu\right)},\\
\overline{\gamma}_3(\gamma,\mathcal{R}_0) &= \frac{\left[\beta\left(1+\alpha\right)y^*+2\mu\right]\left[\left(\beta y^*+\mu\right)\left(\beta\alpha y^*+\mu\right)\right]+\beta^2 y^*\left(1-z^*\right)\left[\beta(2\alpha+1)y^*+3\mu\right]}{\beta^2 y^*\left(1-w^*\right)\left[\beta(\alpha+2)y^*+3\mu\right]+\left(\beta^2 y^*\left(1-z^*\right)\right)\left[\beta(2\alpha+1)y^*+3\mu\right]},
\end{aligned}$$

we must solve the fixed-point problem $\gamma^*=\overline{\gamma}_r(\gamma^*,\mathcal{R}_0)$ for any fixed $\mathcal{R}_0>1$. The conditions for the stability of the strongly endemic or co-existence equilibrium is stated in Theorem 3.4.7.

A.3 Illustrative Example of a 4-strain System

In Figure 3.1, we listed the different equilibrium types observed in a four strain system in the multi-locus system (3.4)-(3.6) on a square lattice with $\mathbb{D}_4$-symmetry. In Chapter 3, we stated that when $\rho=\min\{\rho_p,\rho_\emptyset\}=0$, the steady states with such equilibrium clusters are unstable. In this section, we sketch the proof of conjecture 3.4.6 for the four-strain system (see Blyuss, 2013, for similar calculations for the 2-locus-2-allele

system). Calculating the Jacobian at the steady-state, we have the following general form:

$$\mathcal{J} := \begin{bmatrix} \mathcal{J}_{11} & \mathbf{0} & \mathcal{J}_{13} \\ \mathbf{0} & \mathcal{J}_{22} & \mathcal{J}_{23} \\ \mathcal{J}_{31} & \mathcal{J}_{32} & \mathcal{J}_{33} \end{bmatrix}, \tag{A.13}$$

where $\mathbf{0}$ is a zero block matrix, and the $\mathcal{J}_{ij}$'s are of the following form:

$$\mathcal{J}_{11} = \begin{bmatrix} -(\beta y_1 + \mu) & 0 & 0 & 0 \\ 0 & -(\beta y_2 + \mu) & 0 & 0 \\ 0 & 0 & -(\beta y_3 + \mu) & 0 \\ 0 & 0 & 0 & -(\beta y_4 + \mu) \end{bmatrix},$$

$$\mathcal{J}_{13} = \begin{bmatrix} \beta(1 - z_1) & 0 & 0 & 0 \\ 0 & \beta(1 - z_2) & 0 & 0 \\ 0 & 0 & \beta(1 - z_3) & 0 \\ 0 & 0 & 0 & \beta(1 - z_4) \end{bmatrix},$$

$$\mathcal{J}_{22} = \begin{bmatrix} -\beta y_{124} - \mu & 0 & 0 & 0 \\ 0 & -\beta y_{123} - \mu & 0 & 0 \\ 0 & 0 & -\beta y_{234} - \mu & 0 \\ 0 & 0 & 0 & -\beta y_{134} - \mu \end{bmatrix},$$

$$\mathcal{J}_{23} = \begin{bmatrix} \beta(1 - w_1) & \beta(1 - w_1) & 0 & \beta(1 - w_1) \\ \beta(1 - w_2) & \beta(1 - w_2) & \beta(1 - w_2) & 0 \\ 0 & \beta(1 - w_3) & \beta(1 - w_3) & \beta(1 - w_3) \\ \beta(1 - w_4) & 0 & \beta(1 - w_4) & \beta(1 - w_4) \end{bmatrix},$$

$$\mathcal{J}_{31} = \begin{bmatrix} -\beta(1 - \gamma)y_1 & 0 & 0 & 0 \\ 0 & -\beta(1 - \gamma)y_2 & 0 & 0 \\ 0 & 0 & -\beta(1 - \gamma)y_3 & 0 \\ 0 & 0 & 0 & -\beta(1 - \gamma)y_4 \end{bmatrix},$$

$$\mathcal{J}_{32} = \begin{bmatrix} -\beta\gamma y_1 & 0 & 0 & 0 \\ 0 & -\beta\gamma y_2 & 0 & 0 \\ 0 & 0 & -\beta\gamma y_3 & 0 \\ 0 & 0 & 0 & -\beta\gamma y_4 \end{bmatrix},$$

$$\mathcal{J}_{33} = \begin{bmatrix} \beta\Xi_1 - (\sigma + \mu) & 0 & 0 & 0 \\ 0 & \beta\Xi_2 - (\sigma + \mu) & 0 & 0 \\ 0 & 0 & \beta\Xi_3 - (\sigma + \mu) & 0 \\ 0 & 0 & 0 & \beta\Xi_4 - (\sigma + \mu) \end{bmatrix},$$

where $y_{ijk} = y_i + y_j + y_k$, $i, j, k = 1, 2, 3, 4$, $i \neq j \neq k$, and $\Xi_l = [(1 - w_l) + (1 - \gamma)(w_l - z_l)]$, $l = 1, 2, 3, 4$. In the proceeding, we use the above Jacobian matrix to obtain the characteristic equations.

A.3.1 Equilibrium Types of the form a-1 to a-4 in Fig. 3.1

There are four equilibria of the form a-1 to a-4 shown in Fig. 3.1, (i.e., equilibrium with three non-zero variants; see a-1 to a-4), and the equilibria are of the following form

$$\mathcal{E}_1^{(a)} = (z_{a,2}, 0, z_{a,2}, z_{a,1}, w_{a,2}, w_{a,3}, w_{a,2}, w_{a,1}, y_{a,2}, 0, y_{a,2}, y_{a,1}), \quad (A.14)$$

$$\mathcal{E}_2^{(a)} = (z_{a,1}, z_{a,2}, 0, z_{a,2}, w_{a,1}, w_{a,2}, w_{a,3}, w_{a,2}, y_{a,1}, y_{a,2}, 0, y_{a,2}), \quad (A.15)$$

$$\mathcal{E}_3^{(a)} = (z_{a,2}, z_{a,1}, z_{a,2}, 0, w_{a,2}, w_{a,1}, w_{a,2}, w_{a,3}, y_{a,2}, y_{a,1}, y_{a,2}, 0), \quad (A.16)$$

$$\mathcal{E}_4^{(a)} = (0, z_{a,2}, z_{a,1}, z_{a,2}, w_{a,3}, w_{a,2}, w_{a,1}, w_{a,2}, 0, y_{a,2}, y_{a,1}, y_{a,2}), \quad (A.17)$$

where $\mathcal{E}_k^{(a)},\ k = 1, 2, 3, 4$ are related by $\mathbb{Z}_4$-rotations. Hence, we will focus on only one of the equilibria. We note that here, $\rho_p \neq \rho_\emptyset$, in particular $\rho_p = 2,\ \rho_\emptyset = 0$ and $\rho = \min\{\rho_p, \rho_\emptyset\} = 0$. We will show that this equilibrium type is unstable.

Substituting the form of the steady state $\mathcal{E}_k^{(a)},\ k = 1, 2, 3, 4$ into the system (3.4)-(3.6) yields, $z_{a,1} = \dfrac{\beta y_{a,1}}{\beta y_{a,1} + \mu}$, $z_{a,2} = \dfrac{\beta y_{a,2}}{\beta y_{a,2} + \mu}$, $w_{a,1} = \dfrac{\beta(y_{a,1} + 2y_{a,2})}{\beta(y_{a,1} + 2y_{a,2}) + \mu}$, $w_{a,2} = \dfrac{\beta(y_{a,1} + y_{a,2})}{\beta(y_{a,1} + y_{a,2}) + \mu}$, $w_{a,3} = \dfrac{2\beta y_{a,2}}{2\beta y_{a,2} + \mu}$, $y_{a,1} = \dfrac{\mu[1 - \mathcal{R}_0 + \mathcal{R}_0 y_{a,2}\,(2 + \mathcal{R}_0(y_{a,2} - 1))]}{\mathcal{R}_0(\mu + \sigma)[\mathcal{R}_0(1 - \gamma) - 1 - \mathcal{R}_0 y_{a,2}]}$, and $y_{a,2} = [x]_+$ is a positive root of the following quartic equation $F(x) = ax^4 + bx^3 + cx^2 + dx + e$, where $a = \mathcal{R}_0^2$, $b = \mathcal{R}_0(2 - \mathcal{R}_0)$, $c = [\mathcal{R}_0^2(4\gamma - 1)(1 - \gamma) + 1 - 2\gamma\mathcal{R}_0]$, $d = [1 + \mathcal{R}_0^2 + 8\mathcal{R}_0\gamma - 2(\mathcal{R}_0 + \gamma) - \mathcal{R}_0\gamma\,(6\gamma(1 - \mathcal{R}_0) + \mathcal{R}_0(5 + 2\gamma^2))]$ and $e = \gamma(\mathcal{R}_0 - \gamma)(\gamma - 1)$. It should be noted that it is difficult to obtain an analytical solution to $F(x)$. As a result, we cannot find a closed form expression for the eigenvalues of linearization near $\mathcal{E}_k^{(a)},\ k = 1, 2, 3, 4$. However, our numerical solutions in Chapter 3 suggest that, for biologically realistic values of the parameters, at least one of these eigenvalues is always positive, suggesting that a steady state $\mathcal{E}_k^{(a)},\ k = 1, 2, 3, 4$ is unstable.

A.3.2 Equilibrium Types of the form b-1 to b-4 in Fig. 3.1

There are four distinct steady states with a single non-zero strain y_i, and the equilibria are of the following forms:

$$\mathcal{E}_1^{(b)} = (0, z_{b,1}, 0, 0, w_{b,1}, w_{b,1}, 0, w_{b,1}, 0, y_{b,1}, 0, 0) \quad (A.18)$$

$$\mathcal{E}_2^{(b)} = (z_{b,1}, 0, 0, 0, w_{b,1}, w_{b,1}, 0, w_{b,1}, y_{b,1}, 0, 0, 0) \quad (A.19)$$

$$\mathcal{E}_3^{(b)} = (0, 0, 0, z_{b,1}, w_{b,1}, 0, w_{b,1}, w_{b,1}, 0, 0, 0, y_{b,1}) \quad (A.20)$$

$$\mathcal{E}_4^{(b)} = (0, 0, z_{b,1}, 0, w_{b,1}, 0, w_{b,1}, w_{b,1}, 0, 0, y_{b,1}, 0), \quad (A.21)$$

with the steady states $\mathcal{E}_k^{(b)},\ k = 1, 2, 3, 4$ being related to each through elements of $\mathbb{Z}_4$- rotations. In addition, we note that here, $\rho_p \neq \rho_\emptyset$, in particular $\rho_p = 0,\ \rho_\emptyset = 2$

and $\rho = \min\{\rho_p, \rho_\emptyset\} = 0$. Substituting into Eqs. (3.4)-(3.6) yields, $y_{b,1} = \dfrac{\mu}{\beta}(\mathcal{R}_0 - 1)$ and $z_{b,1} = w_{b,1} = \dfrac{\beta y_{b,1}}{\beta y_{b,1} + \mu} = 1 - \dfrac{1}{\mathcal{R}_0}$. Substituting these equations into Eqs. (A.13)-(A.14), the system has the following eigenvalues: $\zeta_1 = \zeta_2 = \zeta_3 = \zeta_4 = -\mu$, $\zeta_5 = (\beta - \sigma - \mu)$, $\zeta_6 = \zeta_7 = \zeta_8 = -\mu\mathcal{R}_0$, $\zeta_9 = \zeta_{10} = -\dfrac{(\mathcal{R}_0 - 1)(2 + \mathcal{R}_0(1-\gamma))}{\mathcal{R}_0}$, and solutions to the following sub-characteristic equation $P_2(\zeta) = \zeta^2 + \mu\mathcal{R}_0\zeta + \mu(\mathcal{R}_0 - 1)$. We note that $\mu\mathcal{R}_0 > 0$, and $\mu(\mathcal{R}_0 - 1) > 0$ for $\mathcal{R}_0 > 1$. Hence the trace of the sub-characteristic equation is negative and the determinant is positive. Therefore, $P_2(\zeta) = 0$ has negative characteristic values. Thus, all the eigenvalues of the steady-state $\mathcal{E}_k^{(b)},\ k = 1, 2, 3, 4$ are negative except $\zeta_5 > 0$ since $\mathcal{E}_k^{(b)},\ k = 1, 2, 3, 4$ are only feasible when $\mathcal{R}_0 > 1$. Hence, $\mathcal{E}_k^{(b)},\ k = 1, 2, 3, 4$ are all unstable.

A.3.3 Equilibrium Types of the form c-1 to c-4 in Fig. 3.1

We consider the case for equilibrium clusters of the form c-1 to c-4 in Fig. 3.1, where there are four steady-states with two non-zero strains that are nearest neighbours forming a cluster. Here we note that $\rho_p = \rho_\emptyset = 0$ and $\rho = \min\{\rho_p, \rho_\emptyset\} = 0$. The steady-states are of the following forms:

$$\mathcal{E}_1^{(c)} = (0, z_{c,1}, z_{c,1}, 0, w_{c,1}, w_{c,2}, w_{c,2}, w_{c,1}, 0, y_{c,1}, y_{c,1}, 0), \tag{A.22}$$

$$\mathcal{E}_2^{(c)} = (z_{c,1}, z_{c,1}, 0, 0, w_{22}, w_{c,2}, w_{c,1}, w_{c,1}, y_{c,1}, y_{c,1}, 0, 0), \tag{A.23}$$

$$\mathcal{E}_3^{(c)} = (0, 0, z_{c,1}, z_{c,1}, w_{c,1}, w_{c,1}, w_{c,2}, w_{c,2}, 0, 0, y_{c,1}, y_{c,1}), \tag{A.24}$$

$$\mathcal{E}_4^{(c)} = (z_{c,1}, 0, 0, z_{c,1}, w_{c,2}, w_{c,1}, w_{c,1}, w_{c,2}, y_{c,1}, 0, 0, y_{c,1}), \tag{A.25}$$

where the steady-states $\mathcal{E}_k^{(c)},\ k = 1, 2, 3, 4$ are related by reflections along the horizontal/vertical axes. Substituting these expressions into Eqs. (3.4)-(3.6), we find that the system has the following steady-state solutions: $z_{c,1} = \dfrac{\beta y_{c,1}}{\beta y_{c,1} + \mu}$, $w_{c,1} = \dfrac{\beta y_{c,1}}{\beta y_{c,1} + \mu}$, $y_{c,1} = \dfrac{\mu}{4\beta}\left[\mathcal{R}_0(2-\gamma) - 3 + \sqrt{8(\mathcal{R}_0 - 1) + [\mathcal{R}_0(2-\gamma) - 3]^2}\right]$, and $w_{c,2} = \dfrac{2\beta y_{c,1}}{2\beta y_{c,1} + \mu}$. Substituting these expression into the Jacobian matrix and submatrices Eqs. (A.13)-(A.14), we find the following eigenvalues: $\zeta_1 = \zeta_2 = -\mu$, $\zeta_3 = \zeta_4 = -(\mu + \beta y_{c,1})$, $\zeta_5 = -(\mu + 2\beta y_{c,1})$, $\zeta_6 = \zeta_7 = \dfrac{\mathcal{R}_0\beta y_{c,1}(1-\gamma) + \mu(\mathcal{R}_0 - 1) - \beta y_{c,1}}{\beta y_{c,1} + \mu}$, and solutions to the following sub-characteristic polynomials $P_2(\zeta) = \zeta^2 + a_{21}\zeta + a_{22}$ and $P_3(\zeta) = \zeta^3 + a_{31}\zeta^2 + a_{32}\zeta + a_{33}$, where $a_{21} = \mu + \beta y_{c,1} > 0$, $a_{22} = \dfrac{\mathcal{R}_0^2\mu y_{c,1}(1-\gamma)}{\mu + \beta y_{c,1}} > 0$, $a_{31} = 3\beta y_{c,1} + 2\mu > 0$, $a_{32} = \dfrac{2y_{c,1}^2\mu\beta^2\mathcal{R}_0(\sigma+\mu)(1+6y_{c,1}) + y_{c,1}\mu^2\beta[\beta(1+13y_{c,1}) + \beta\gamma + 6\mu] + 4\beta^4 y_{c,1}^4}{(\mu+\sigma)^2(\beta y_{c,1} + \mu)(2\beta y_{c,1} + \mu)} > 0$, $a_{33} = \dfrac{\mathcal{R}_0\mu\beta y_{c,1}[2y_{c,1}^2\beta^2(2-\gamma) + \mu^2(1+\gamma) + 4\gamma\mu\beta]}{(\mu+\sigma)^2(\beta y_{c,1} + \mu)(2\beta y_{c,1} + \mu)} > 0$, with Hurwitz determinant

$a_{31}a_{32} - a_{33} = \beta\mu y_{c,1}[\beta(1+\gamma) + 13\beta y_{c,1} + 9\mu] + 2(\mu^3 + 3\beta^3 y_{c,1}^3) > 0$. Hence the Routh-Hurwitz criterion implies that the real part of all the roots of $P_2(\zeta)$ and $P_3(\zeta)$ are negative. Now it suffices to show that $\zeta_6 = \zeta_7$ is negative. We note that, substituting the expression for $y_{c,1}$ into $\zeta_6 = \zeta_7$, the eigenvalue can only cross zero when $\mathcal{R}_0 = (1-2\gamma)/(1-\gamma)$ and $\mathcal{R}_0 = 1$. However, since $0 \leq \gamma \leq 1$ and the steady states $\mathcal{E}_k^{(c)},\ k = 1,2,3,4$ are only biologically valid for $\mathcal{R}_0 > 1$, $\mathcal{E}_k^{(c)},\ k = 1,2,3,4$ are always unstable irrespective of the value of γ.

A.3.4 Equilibrium Types of the form d-1 to d-4 in Fig. 3.1

There are three types of equilibrium clusters of the form d-1 to d-4 in Fig. 3.1, namely one disease-free equilibrium, a fully symmetric (co-existence) and two partially symmetric equilibria with a pair of discordant persistent strain set. The equilibrium forms are as follows:

$$\mathcal{E}_1^{(d)} = (0, z_{d,1}, 0, z_{d,1}, w_{d,2}, w_{d,1}, w_{d,2}, w_{d,1}, 0, y_{d,1}, 0, y_{d,1}), \tag{A.26}$$

$$\mathcal{E}_2^{(d)} = (z_{d,1}, 0, z_{d,1}, 0, w_{d,1}, w_{d,2}, w_{d,1}, w_{d,2}, y_{d,1}, 0, y_{d,1}, 0), \tag{A.27}$$

$$\mathcal{E}_3^{(d)} = (z, z, z, z, w, w, w, w, y, y, y, y), \tag{A.28}$$

$$\mathcal{E}_4^{(d)} = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), \tag{A.29}$$

where $y_{d,1} = \dfrac{\mu}{\beta}(\mathcal{R}_0 - 1)$, $z_{d,1} = 1 - \dfrac{1}{\mathcal{R}_0}$, $w_{d,1} = 1 - \dfrac{1}{\mathcal{R}_0}$, $w_{d,2} = \dfrac{2(\mathcal{R}_0 - 1)}{2(\mathcal{R}_0 - 1) + 1}$, $y = \dfrac{\mu}{6\beta}\left[\mathcal{R}_0(3-2\gamma) - 4 + \sqrt{12(\mathcal{R}_0 - 1) + [\mathcal{R}_0(3-2\gamma) - 4]^2}\right]$, $z = \dfrac{\beta y}{\beta y + \mu}$, and $w = \dfrac{3\beta y}{3\beta y + \mu}$. Here, we note that $\mathcal{E}_1^{(d)}$ and $\mathcal{E}_2^{(d)}$ are related by reflection along the diagonal or $\mathbb{Z}_4$-rotations, $\mathcal{E}_3^{(d)}$ is the fully symmetric steady-state, and $\mathcal{E}_4^{(d)}$ is the trivial steady-state. In addition, we observe that $\mathcal{E}_1^{(d)}$ and $\mathcal{E}_2^{(d)}$ are special cases of the cluster set described in the text, where $\rho_p = \rho_\emptyset = 2$ and $\rho = \min\{\rho_p, \rho_\emptyset\} = 2$, and the results of Section 3.4.2.1, Appendix A.1.3, Theorem 3.4.4 and Corollary 3.4.5 hold. For steady-state $\mathcal{E}_3^{(d)}$, the discussion in Section 3.4.2.2, Theorem 3.4.7 and Appendix A.1.4 and Appendix A.2.2 follows where $\alpha = 3$. The results for $\mathcal{E}_4^{(d)}$ follows directly from Section 3.4.1, Lemma 3.4.1, Theorem 3.4.2 Appendix A.1.1, and Appendix A.1.2.

B Chapter 4 Appendix

B.1 Numerical Methods

Here we discuss a method of lines scheme used to approximate and simulate the PDEs in this thesis.

B.1.1 Method of Lines *à la* Stable Age-distribution

Herein, we follow the methodology of Hethcote (1997) under the assumption of stable age distributions in a very short time period which may not be realistic and/or valid over longer time periods, and in developing countries where the age-distribution may not hold. The idea is based on the use of the characteristic equation and relies on the properties of the renewal equation in which it can be shown that the age or size distribution approaches a steady state and the population growth is exponential. The methodology described herein is then applied to the age-structured multi-locus-allele model discussion in Chapter 4 to obtain a set of differential equations describing the age-specificity and to simulate the model dynamics.

B.1.1.1 Derivation

We start with the following hyperbolic PDE

$$\frac{\partial \eta(a,t)}{\partial t} + \frac{\partial \eta(a,t)}{\partial a} = -D(a)\eta(a,t) \tag{B.1}$$

$$\eta(0,t) = \int_0^\infty f(\tau)\eta(\tau,t)d\tau, \tag{B.2}$$

$$\eta(a,0) = \eta_0(a), \tag{B.3}$$

where $D(a)$ is the age-specific death rate for individuals of age a, $\eta(a,0)$ and $\eta(0,t)$ are the boundary condition and initial age distribution, respectively; $f(a)$ is the fecundity function and is well defined; Eqs. (B.1)-(B.3) yield a well-posed problem. Here we adapt the methodology used in Hethcote (1997). Our task is to convert the hyperbolic PDE in (B.1)-(B.3) into a system of n ODEs. For this approach, we assume that n is the total number of size groups, $a \in [0, \infty)$ and that the age interval

can be partitioned such that $0 \leq a_0 < a_1 < \cdots < a_{n-1} < a_n = \infty$ so that $D(a)$, $f(a)$ are constants in each interval $[0, a_1], [a_2, a_3], \ldots, [a_{n-1}, a_n]$, as follows:

$$f(a) = f_i \text{ and } D(a) = d_i \qquad a \in [a_{i-1}, a_i], \tag{B.4}$$

We assume that the population has reached a stationary age-distribution and that the solution is separable with exponential growth, that is, $\eta(a,t) = e^{qt}M(a)$, where q is the growth rate, and $M(a)$ is the population age density. Then the number of individuals with age in $[a_{i-1}, a_i]$ is given by

$$N_i(t) = \int_{a_{i-1}}^{a_i} \eta(a,t)da = \int_{a_{i-1}}^{a_i} e^{qt}M(a)da = e^{qt}\int_{a_{i-1}}^{a_i} M(a)da = e^{qt}P_i, \tag{B.5}$$

where P_i is the population size for each group of age interval $[a_{i-1}, a_i]$. Notice that the boundary condition Eq. (B.2) then becomes,

$$\eta(0,t) = e^{qt}M(0) = \int_0^\infty e^{qt}f(\tau)M(\tau)d\tau \tag{B.6}$$

$$M(0) = \int_0^\infty f(\tau)M(\tau)d\tau, \tag{B.7}$$

which in $a_0 < a_1 < \cdots < a_{n-1} < a_n$, is equivalent to having

$$\begin{aligned} M(0) &= \int_{a_0}^{a_1} f_1(\tau)M(\tau)d\tau + \cdots + \int_{a_{n-1}}^{a_n} f_n(\tau)M(\tau)d\tau \\ &= f_1\int_{a_0}^{a_1} M(\tau)d\tau + f_2\int_{a_1}^{a_2} M(\tau)d\tau + \cdots + f_i\int_{a_{i-1}}^{a_i} M(\tau)d\tau \\ &= f_1P_1 + f_2P_2 + ... + f_nP_n, \qquad i = 1, 2, 3, ..., n. \end{aligned} \tag{B.8}$$

Therefore, $M(0) = \sum_{i=1}^{n} f_iP_i$ is the initial age distribution. Then the change in density with respect to time and with respect to age are given by the equations

$$\frac{\partial \eta}{\partial t} = qe^{qt}M(a), \tag{B.9}$$

$$\frac{\partial \eta}{\partial a} = e^{qt}M'(a). \tag{B.10}$$

Substituting Eqs. (B.9)-(B.10) into Eq. (B.1) we obtain

$$\frac{M'(a)}{M(a)} = -(D(a) + q). \tag{B.11}$$

Fixing $a \in [a_{i-1}, a_i]$, with constant $D(a,t) = d_i$, and integrating Eq. (B.11) yields

$$M(a_i) = M(a_{i-1})\exp[-(d_i + q)(a_i - a_{i-1})]. \tag{B.12}$$

To obtain an expression for P_i, we integrate Eq. (B.12) over $[a_{i-1}, a_i]$ and obtain

$$P_i = \int_{a_{i-1}}^{a_i} M(a)da = \int_{a_{i-1}}^{a_i} M(a_{i-1}) \exp[-(d_i + q)(a - a_{i-1})]da \tag{B.13}$$

which gives $P_i = \frac{M(a_{i-1})}{d_i+q}(1 - \exp\{-(d_i + q)(a_i - a_{i-1})\})$. For $i = 1, 2, ..., n-1$, it is convenient to define constants c_i such that

$$M(a_{i-1}) = c_i P_i. \tag{B.14}$$

Hence $c_i = \dfrac{M(a_{i-1})}{P_i} = \dfrac{d_i + q}{1 - \exp\{-(d_i + q)(a_i - a_{i-1})\}}$. Observe that for a small $[a_{i_1}, a_i]$ or small $(d_i + q)(a_i - a_{i-1})$, the constants c_i can be approximated by $c_i \approx \dfrac{1}{(a_i - a_{i-1})}$, so that $P_i \approx M(a_{i-1})(a_i - a_{i-1})$.

With all the appropriate assumptions we are ready to derive the age-structured system of ODEs. The first equation can be obtained by integrating Eq. (B.1) over $[a_0, a_1]$ as follows:

$$\int_{a_0}^{a_1} \frac{\partial \eta}{\partial t}da + \int_{a_0}^{a_1} \frac{\partial \eta}{\partial a}da = -d_1 \int_{a_0}^{a_1} \eta da, \tag{B.15}$$

which implies that

$$\frac{\partial}{\partial t}\int_{a_0}^{a_1} \eta da + [\eta(a_1) - \eta(a_0)] = -d_1 \int_{a_0}^{a_1} \eta da. \tag{B.16}$$

Hence

$$\frac{dN_1}{dt} + [\eta(a_1) - \eta(a_0)] = -d_1 N_1. \tag{B.17}$$

Since $\eta = e^{qt}M(a)$, $M(a_{i-1}) = c_i P_i$, and $N_i = P_i e^{qt}$, then $\eta = N_i c_i$. Then Eq. (B.17) becomes

$$\frac{dN_1}{dt} + c_1 N_1 - e^{qt} M(a_0) = -d_1 N_1. \tag{B.18}$$

For $i = 1, 2, ..., n$, $\eta(a_0) = e^{qt}M(a_0) = e^{qt}M(0) = \sum_{i=1}^{n} f_i e^{qt} P_i = \sum_{i=1}^{n} f_i N_i$ gives the first age-structured differential equation in $[a_0, a_1]$:

$$\frac{dN_1}{dt} = \sum_{i=1}^{n} f_i N_i - (d_1 + c_1)N_1. \tag{B.19}$$

For groups in $2 \leq i \leq n-1$, we integrate Eq. (B.1) over $[a_{i-1}, a_i]$ to obtain

$$\frac{dN_i}{dt} = c_{i-1}N_{i-1} - (d_i + c_i)N_i. \tag{B.20}$$

Lastly, integrating Eq. (B.1) over $[a_{n-1}, a_n]$ gives us the last group

$$\frac{dN_n}{dt} = c_{n-1}N_{n-1} - (d_n + c_n)N_n, \tag{B.21}$$

where f_i are fertility constants, and c_i are transfer rate constants between successive age groups. Putting Eqs. (B.19)-(B.21) together with the following initial conditions $N_i(0) = N_i^0,\ i = 1, 2, \ldots, n$, yield the discrete age-structured population model. We note that one can use any appropriate general-purpose ODE solver for the time discretization, and can recover various finite difference approximation schemes. For example using an explicit Euler method (with the upwind method) in space yields the explicit (upwind) scheme and when we use an implicit Euler method we get the implicit upwind scheme.

B.1.1.2 Application: Age-structured Multi-locus-allele System

Following the procedures outlined in the previous section and reformulating the age-structured multi-locus-allele system using single indexing representing strains where s_{il} incorporate the strain structure of the multi-locus-allele system, we obtain the following systems of ordinary differential equations where i and j denote strain number and age class, respectively, where for the first age group, $j = 1$, we have

$$\begin{aligned}
\frac{d}{dt}S_i^1 &= \sum_k^{a^*} f_k N_i^k - c(1)S_i^1 - S_i^1 \sum_l s_{il}\lambda_l^1 - \mu^1 S_i^1 && \text{(B.22)}\\
\frac{d}{dt}V_i^1 &= -c(1)V_i^1 + S_i^1 \sum_{l\neq i} s_{il}\lambda_l^1 - V_i^1\lambda_i^1 - \mu^j V_i^{1j} && \text{(B.23)}\\
\frac{d}{dt}Z_i^1 &= -c(1)Z_i^1 + (S_i^1 + V_i^1)\lambda_i^1 - \mu^1 Z_i^1 && \text{(B.24)}\\
\frac{d}{dt}Y_i^1 &= -c(1)Y_i^1 + \left[S_i^1 + (1-\gamma)V_i^1\right]\lambda_i^1 - \sigma^1 Y_i^1, && \text{(B.25)}
\end{aligned}$$

and a^* is the max age group, $c(j) = \dfrac{1}{a_j - a_{j-1}}$, f_k is the discrete fertility modulus and $N_i^j = S_i^j + Z_i^j + V_i^j$. For age groups $j = 2, \ldots, a^* - 1$,

$$\begin{aligned}
\frac{d}{dt}S_i^j &= c(j-1)S_i^{j-1} - c(j)S_i^j - S_i^j \sum_l s_{il}\lambda_l^j - \mu^j S_i^j && \text{(B.26)}\\
\frac{d}{dt}V_i^j &= c(j-1)V_i^{j-1} - c(j)V_i^j + S_i^j \sum_{l\neq i} s_{il}\lambda_l^j - V_i\lambda_i^j - \mu^j V_i^j && \text{(B.27)}\\
\frac{d}{dt}Z_i^j &= c(j-1)Z_i^{j-1} - c(j)Z_i^j + (S_i^j + V_i^j)\lambda_i^j - \mu^j Z_i^j && \text{(B.28)}\\
\frac{d}{dt}Y_i^j &= c(j-1)Y_i^{j-1} - c(j)Y_i^j + \left[S_i^j + (1-\gamma)V_i^j\right]\lambda_i^j - \sigma^j Y_i^j. && \text{(B.29)}
\end{aligned}$$

For $j = a^*$

$$\frac{d}{dt}S_i^j = c(j-1)S_i^{j-1} - S_i^j \sum_l s_{il}\lambda_l^j - \mu^j S_i^j \tag{B.30}$$

$$\frac{d}{dt}V_i^j = c(j-1)V_i^{j-1} + S_i^j \sum_{l\neq i} s_{il}\lambda_l^j - V_i\lambda_i^j - \mu^j V_i^j \tag{B.31}$$

$$\frac{d}{dt}Z_i^j = c(j-1)Z_i^{j-1} + (S_i^j + V_i^j)\lambda_i^j - \mu^j Z_i^j \tag{B.32}$$

$$\frac{d}{dt}Y_i^j = c(j-1)Y_i^{j-1} + \left[S_i^j + (1-\gamma)V_i^j\right]\lambda_i^j - \sigma^j Y_i^j, \tag{B.33}$$

with the following initial conditions $S_i^j(0) = S_i^{j,0}, V_i^j(0) = V_i^{j,0}, Y_i^j(0) = Y_i^{j,0}$, $i = 1, 2, \ldots, a^*, i = 1, 2, \ldots, \mathcal{N}$. In addition, the force of infection becomes $\lambda_l^j := \sum_k \beta_{k,j} Y_l^k$ with $\beta_{k,j}$ incorporating both the contact rate and infectivity.

B.1.2 Model Comparison under Constant Parameters

In this section, we provide numerical results in which we simulate the age-structured multiple strain system. In particular, we will focus mostly on 2-locus-2-allele systems resulting in a four-strain system. Here we will employ the numerical scheme that has extensively been used in structured population models, which relies on the assumption of stable age-distribution (i.e., the age-distribution of the population does not change significantly in the long-term). This formalism reduces the evolution equation describing the structured population dynamics into a system of ordinary differential equations and is equivalent to the method of lines with weighted age-discretization. The method is employed in the numerical solution of Eqs. (4.14)-(4.17) (see Appendix B.1.1 for description of the approach, where the hyperbolic system of partial differential equations is transformed into systems of ordinary differential equations stratified by a finite set of age classes or groups).

Here we will show that the age-structured model of the multiple strain system with *Mendelian* phenotypic strain structure reduces to Gupta *et al.* (1998), Recher *et al.* (2007), Ferguson & Andreasen (2001), and model derived in Chapter 2 and studied in Chapter 3. We use uniform epidemiological and demographic parameter values (i.e., $\mu(a) = \mu$, $\sigma(a) = \sigma$, $\beta(a, s) = \beta$ and $\gamma(a) = \gamma$ across all ages). We consider a coarse age-discretization with 4 age classes (0-19 years], (20-39 years], (40-59 years] and (60+years]. In addition, we chose the same parameter values in Chapter 7 and in Gupta *et al.* (1998), where $\mathcal{R}_0 = 4$, $\mu = 0.02$, $\sigma = 9.98$ and $\beta = 40$. Here we use γ as a free variable to illustrate that the age-structured model under these constant parameters reduces to the model of Gupta *et al.* (1998). With these parameter values, we use the standard general purpose method such as the Runge-Kutta fourth-

and fifth-order solver ODE45 in Matlab (The Mathworks Inc., Natwick, MA, USA) to simulate the resulting initial value ODE of the model in order to compare the two models for 2-locus-allele system. Figure B.1 shows the comparison of the two models. We observe that the models are very closed (almost identical); and all the dynamic features are preserved as γ varies from zero to one. The age-structured model under constant demographic and epidemiological parameter values reduces to the unstructured multi-locus-allele model. Figure B.1 shows that the dynamics are consistent across all age groups. Moreover, the two models are essentially identical.

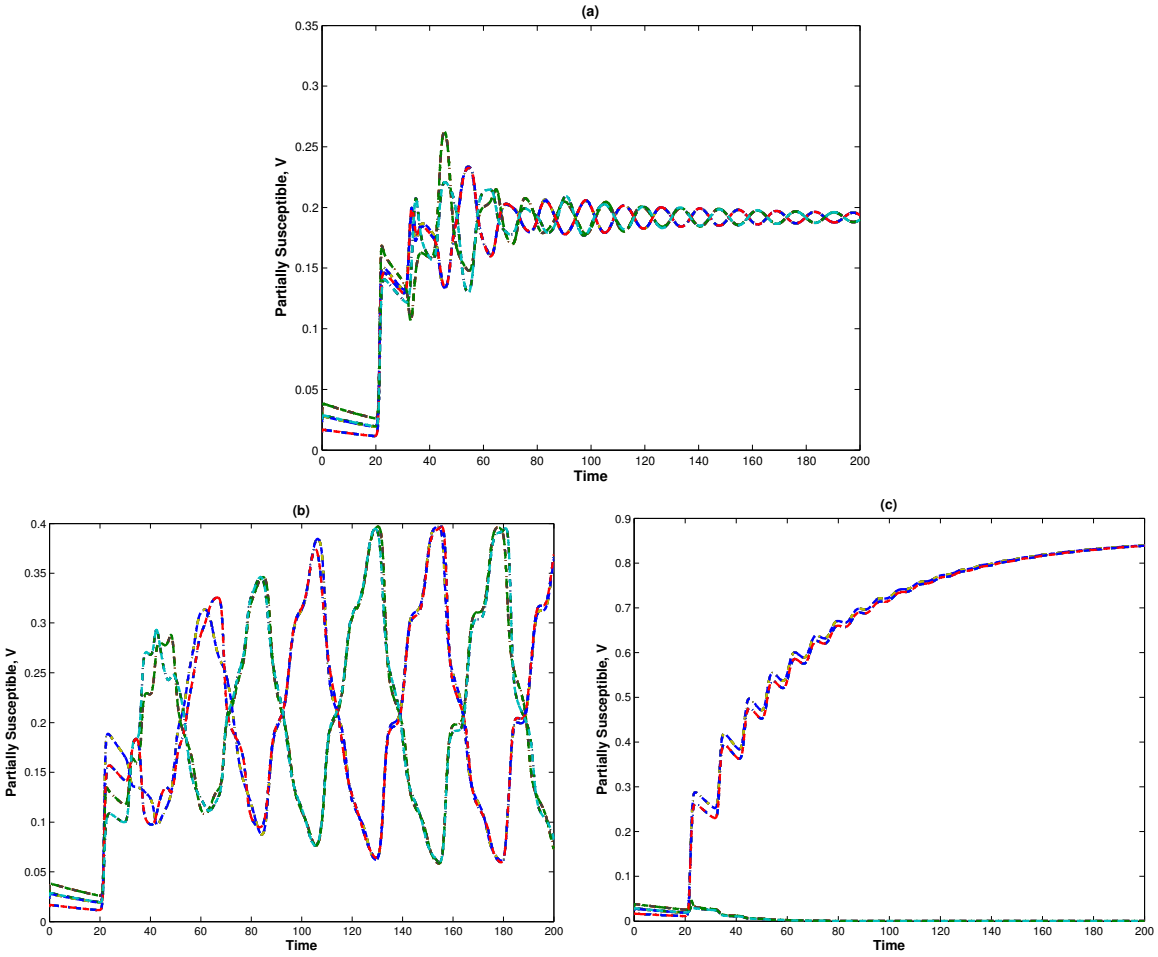


Figure B.1: **Model Comparision.** The dynamics for the 2-locus-2-alleles system where $\mathcal{R}_0 = 4$, $\sigma = 9.98$, $\mu = 0.02$, $\gamma = 0.45(a)$, $\gamma = 0.65(b)$, $\gamma = 0.95(c)$; and $-$. for Gupta et al. (1998) model and $--$ for age-structured model with constant parameters in Eqs. (4.14)-(4.17). All equations were solved using the Runge-Kutta fourth- and fifth-order method solver ODE45 in MATLAB (The Mathworks Inc., Natwick, MA, USA) with relative and absolute error tolerance of 10^{-5} and 10^{-6}, respectively.

C Chapter 5 Appendix

C.1 Proofs

We relegate some of the proofs and properties herein.

C.1.1 Proof of Theorem 5.3.26(see Krasnoselskii, 1964)

Proof. This is a fundamental result needed in establishing the existence of a positive solution in Chapter 5. To prove the above theorem, it suffices to show that the operator g satisfies the complete continuity condition and $g(\mathcal{Y}_+) \subset \mathcal{Y}_+$, and has a Fréchet derivative at zero. This can be done using similar approaches employed in the main text. Once these properties have been established, we can then define an operator $g_r(u)$ as

$$g_r(u) := \begin{cases} g(u), & u \in \mathcal{Z}_+,\ ||u|| \geq r \\ g(u) + (r - ||u||)\nu_0 & u \in \mathcal{Z}_+,\ ||u|| \leq r \end{cases} \tag{C.1}$$

where ν_0 is the positive eigenvector of $\mathcal{L}$ corresponding to the eigenvalue $\zeta_0 > 1$ as defined above. If g is completely continuous, then the complete continuity of g_r is obvious. Applying Schauder's fixed point theorem (in particular Leray extension, i.e. Leray-Schauder alternative; see Zeidler, 1986), we can show that g_r has a fixed point in $\mathcal{Y}_+$. However, this does not tell us whether the fixed point is trivial or non-trivial. To show that the fixed point is positive, it suffices to show that the norms of the fixed points of the operator g_r are greater than r if r is sufficiently small. Then the fixed points of the operator g_r will be non-zero fixed points of the operator g. That is, suppose every operator g_r has a fixed point whose norm is not greater than r. We construct a sequence $u_n \in \mathcal{Y}$ such that $||u_n|| \in (0, \frac{1}{n}]$ and $\lim_{n\to\infty} ||u_n|| \to 0$. That is

$$||u_n|| \leq r_n = \frac{1}{n}, \tag{C.2}$$

$$u_n = g(u_n) + \left(\frac{1}{n} - ||u_n||\right)\nu_0, \tag{C.3}$$

for $n \in \mathbb{Z}_+$. Rewriting Eq. (C.3), we have

$$\left(\frac{1}{n||u_n||} - 1\right)\nu_0 \;=\; \frac{u_n}{||u_n||} - \frac{g(u_n) - g'(0)u_n}{||u_n||} - g'(0)\frac{u_n}{||u_n||}, \tag{C.4}$$

where we note that $\mathcal{L} = g'(0)$ and, suppose $\mathcal{L}$ is compact such that $g'(0)$ transforms the sequence $\frac{u_n}{||u_n||}$ into a compact sequence (or equivalently assume that $g'(0)\frac{u_n}{||u_n||}$ converges). Therefore, we can assume that $g'(0)\frac{u_n}{||u_n||}$ converges in norm to some element $\kappa \in \mathcal{Y}$ (in the contrary case, we can take the subsequence that converges). Then it follows, from Eq. (C.4), that $\lim_{n\to\infty}\left(\frac{1}{n||u_n||} - 1\right)||\nu_0|| \leq 1 + ||\kappa||$, where the second term in Eq. (C.4) vanishes because it is the definition of Fréchet differentiation and $g'(0)$ is the Fréchet derivative of g at zero. Hence $\left(\frac{1}{n||u_n||} - 1\right)$ converges to some non-negative number κ_0 (because of Eq. (C.2)). Therefore, from Eq. (C.4), $\frac{u_n}{||u_n||}$ converges in norm to some vector $u_0 \in \mathcal{Y}$, with $||u_0|| = 1$. Then it follows from Eq. (C.4) that $u_0 = \mathcal{L}u_0 + \kappa_0\nu_0$. Because $\mathcal{L}$ does not have an eigenvector in $\mathcal{Y}$ with eigenvalue equal to one (e.g, $r(\mathcal{L})$ is the only eigenvalue with positive positive eigenvector, $r(\mathcal{L}) > 1$), κ_0 is positive. Hence, there exists a number $c_0 > 0$ such that $u_0 \geq c_0\nu_0$. Whenever $u_0 \geq c\nu_0$, we have $c \leq c_0$, and the following inequality

$$u_0 \;=\; \mathcal{L}u_0 + \kappa_0\nu_0 \geq \mathcal{L}c_0\nu_0 + \kappa_0\nu_0 = (\zeta_0 c_0 + \kappa_0)\,\nu_0 \tag{C.5}$$

holds (i.e., $u_0 \geq c\nu_0$), where $c = \zeta_0 c_0 + \kappa_0 > c_0$ and $\zeta_0 = r(\mathcal{L}) > 1$, which is a contradiction. This completes the proof. □

D Chapter 6 Appendix

D.1 Proofs

We relegate some of the proofs and properties herein.

D.1.1 Proof of Theorem 6.3.4

Proof. Define M_0 as

$$M_0 := \sup_{a\in[0,a_m]} \left(\int_0^{a_m} |K(a,b)|\, db \right). \tag{D.1}$$

In addition, let $\mathcal{S} \subset L^1(0,a_m)$ be a bounded subset of $L^1(0,a_m)$ (say $||f|| \leq c$ for some positive constant c and for all $f \in \mathcal{S}$) and $F := A(\mathcal{S}) \subset L^1(0,a_m)$. Then,

$$||A[f](a)|| = \int_0^{a_m} \left| \left(\int_0^{a_m} K(a,b) f(b) db \right) \right| da \leq M_0 ||f|| \leq M_0 c. \tag{D.2}$$

Hence, $A(\mathcal{S}) \subset L^1(0,a_m)$ is bounded, so F is bounded. The first criterion of Theorem 5.3.7 (i.e., Riesz-Fréchet-Kolmogorov Compactness Theorem) is satisfied. Since K is a continuous function on $\Omega_c \times \Omega_c$, from the Heine-Cantor Theorem, K is also uniformly continuous (i.e., every continuous function on a compact set is uniformly continuous). This implies that, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|K(a+h,b) - K(a,b)| \leq \varepsilon$ for all $h \in \mathbb{R}$ with $|h| \leq \delta$ and for all $a, b \in [0, a_m]$. Therefore,

$$\begin{aligned} ||\rho_h[A[\phi]](a) - A[\phi](a)|| &\leq \int_0^{a_m} \left(\int_0^{a_m} |K(a+h,b) - K(a,b)|\, |f(b)|\, db \right) da \\ &= \int_0^{a_m} \left(\int_0^{a_m} |K(a+h,b) - K(a,b)|\, da \right) |f(b)|\, db \\ &\leq \varepsilon a_m \int_0^{a_m} |f(b)|\, db = \varepsilon a_m ||f|| \leq \varepsilon a_m c. \end{aligned} \tag{D.3}$$

Hence, the second condition of Theorem 5.3.7 also holds. Thus, from Theorem 5.3.7, the operator $A : L^1(0,a_m) \to L^1(0,a_m)$ is compact and the assertion holds. This completes the proof. □

Bibliography

[1] R. Allwinn, J. Geiler, A. Berger, J. Cinatl & H.W. Doerr. (2010). Determination of serum antibodies against swine-origin influenza A virus H1N1/09 by immunofluorescence, haemagglutination inhibition, and by neutralization tests: how is the prevalence rate of protecting antibodies in humans? *Med. Microbiol. Immunol.,* 199:117-121.

[2] R.M. Anderson. (1991). Discussion: the Kermack-McKendrick epidemic threshold theorem. *Bull. Math. Biol.* 53: 3-32.

[3] R.M. Anderson & R.M. May. (1979). Population biology of infectious diseases. *Nature*, 280:361-367.

[4] R.M. Anderson & R.M. May. (2002). *Infectious Diseases of Humans: Dynamics and Control.* Oxford University Press, Oxford.

[5] R.M. Anderson, R.M. May, T.W. Ng & J.T. Rowley. (1992). Age-dependent choice of sexual partners and the transmission dynamics of HIV in Sub-Saharan Africa. *Phil. Trans. R. Soc. B,* 336:135-155.

[6] V. Andreasen, J. Lin & S.A. Levin. (1997). The dynamics of cocirculating influenza strains conferring partial cross-immunity. *J. Math. Biol.,* 35:825-842.

[7] V. Andreasen & A. Sasaki. (2006). Shaping the phylogenetic tree of influenza by cross-immunity. *Theor. Pop. Biol.,* 70:164-173.

[8] F. Ansaldi, P. DÁgaro, D. de Florentiis, S. Puzelli, Y.P. Lin, V. Gregory, M. Bennett, I. Donatelli, R. Gasparini, P. Crovari, A. Hay & C. Campello. (2003). Molecular characterization of influenza B viruses circulating in northern Italy during the 2001-2002 epidemic season. *J. Med. Virol., 70:463-469.*

[9] E. Asamoah-Odei, J.M.G. Calleja & J.T. Boerma. (2004). HIV prevalence and trends in sub-Saharan Africa: no decline and large subregional differences. *The Lancet,* 364(9428):35-40.

[10] M.I.R. Azambuja. (2004). Spanish flu and early 20th-century expansion of a coronary heart disease-prone subpopulation. *J. Texas Heart Institute*, 31:14-21.

[11] P. Chea. (2005). Executive summary: Global antimicrobial resistance alerts and implications. *Clinical Infectious Diseases,* 4:S221-S223.

[12] N.T.J. Bailey. (1975). *The Mathematical Theory of Infectious Diseases and Its Applications*. Hafner Press, 2nd Ed.

[13] D. Bandaranayake, Q.S. Huang, A. Bissielo, T. Wood, G. Mackereth, M.G. Baker, R. Beasley, S. Reid, S. Roberts & V. Hope. (2010). Risk factors and immunity in a nationally representative population following the 2009 influenza A(H1N1) pandemic. *PLoS One* 5:e13211.

[14] H.T. Banks & H.T. Tran. (2009). *Mathematical and experimental modeling of physical and biological processes.* CRC Press.

[15] S. Bansal, B. Pourbohloul, N. Hupert, B. Grenfell & L.A. Meyers. (2010). The Shifting Demographic Landscape of Pandemic Influenza. *PLoS One* 5(2):e9360.

[16] A. Barragan, P.G. Kremsner, W. Weiss, M. Wahlgren & J. Carlson. (1998). Age-related buildup of humoral immunity against epitopes for rosette formation and agglutination in African areas of malaria endemicity. *Infection & Immunity* 66:4783-4787.

[17] D.J. Bartholomew. (1982). *Stochastic Models for Social Processes.* Wiley, London 3rd Ed.

[18] M.S. Bartlett. (1957). Measles periodicity and community size. *J. Royal Stat. Soc. Series A,* 120:48-70.

[19] M.S. Bartlett. (1960). The critical community size for measles in the United States.*J. Royal Stat. Soc. Series A,* 123:37-44.

[20] M.S. Bartlett. (1964). The relevance of stochastic models for large-scale epidemiological phenomena. *Applied Statistics,* 13:2-8.

[21] B.J. Becker. (2005). Plagues and people: infectious and epidemic disease in history. `http://eee.uci.edu/clients/bjbecker/PlaguesandPeople/index.html`.

[22] D. Bernoulli. (1760). Essai d'une nouvelle analyse de la mortalité causée par la petite vérole et des avantages de l'inoculation pour la prévenir. *Mém. Math. Phys. Acad. Roy. Sci., Paris,* 1-45.

[23] H.F. Berg, J. van Gendt, G.F. Rimmelzwaan, M.F. Peeters & P. van Keulen. (2003). Nosocomial influenza infection among post-influenza-vaccinate patients with severe pulmonary diseases. *J. Infec.*, 42:129-132.

[24] N. Biggs. (1974). *Algebraic Graph Theory.* Cambridge University Press, New York.

[25] O.N. Bjornstad, B.F. Finkenstadt & B.T. Grenfell. (2002). Dynamics of measles epidemics: Estimating scaling of transmission rates using a time series SIR model. *Ecological Monographs,* 72:169-184.

[26] B.P. Blackburne, A.J. Hay & R.A. Goldstein. (2008). Changing selective pressure during antigenic changes in human influenza H3. *PLoS Pathog.,* 4:e1000058.

[27] K.B. Blyuss. (2013). The effects of symmetry on the dynamics of antigenic variation. J. Math. Biol., 66:115-137.

[28] D. Boraschi, M.A. Alemayehu, A. Aseffa, F. Chiodi, J. Chisi, G.D. Prete, T.M. Doherty, I. Elhassan, H. Engers, B. Gyan, A.M. Harandi, T. Kariuki, F. Kironde, B. Kouriba, J. Langhorne, T. Laskay, D. Medaglini, O. Olesen, P. Onyebujoh, C. Palma, R. Sauerwein, E. Sibanda, U. Steinhoff, A. Tagliabue, A. Thiel, M. Vahedi & M. Troye-Blomberg. (2008). Immunity against HIV/AIDS, malaria, and tuberculosis during co-infections with neglected infectious diseases: Recommendations for the European Union Research Priorities. *PLoS Negl. Trop. Dis.,* 2:e255.

[29] N.M. Bouvier & P. Palese. (2008). The biology of influenza viruses. *Vaccine*, 26(4):D49-53.

[30] F. Brauer. (1995). Models for diseases with vertical transmission and nonlinear population dynamics. *Math. Biosci.,* 128:13-24.

[31] F. Brauer & C. Castillo-Chavez. (2001). *Mathematical models in population biology and epidemiology.* Springer, New York.

[32] H. Brezis. (2010). *Functional analysis, Sobolev spaces, and partial differential equations.* Universitext, Springer, New York.

[33] I.H. Brown. (2000). The epidemiology and evolution of influenza viruses in pigs. *Vet. Microbiol.,* 74:29-46.

[34] I.H. Brown, P.A. Harris, J.W. McCauley & D.J. Alexander. (1998). Multiple genetic reassortment of avian and human influenza A viruses in European pigs, resulting in the emergence of an H1N2 virus of novel genotype. *J. Gen. Virol.,* 79:2947-2955.

[35] J. Brownlee. (1907). Statistical studies in immunity. The theory of an epidemic. *Proc. Roy. Soc. Edinburgh*, 26:484-521.

[36] C.O. Buckee, K. Koelle, M.J. Mustard & S. Gupta. (2004). The effects of host contact structure on pathogen diversity and strain structure. *PNAS,* 101(29):10839-10844.

[37] C.O. Buckee, K.A. Jolley, M. Recker, B. Penman, P. Kriz, S. Gupta & M.C. Maiden. (2008). Role of selection in the emergence of lineages and the evolution of virulence in Neisseria meningitidis. *PNAS,* 105:15082-15087.

[38] C.O. Buckee & S. Gupta. (2010). A Network Approach to Understanding Pathogen Population Structure, in V. Sintchenko (Ed.), Infectious disease informatics. Springer Verlag, New York

[39] C.O. Buckee, M. Recker, E.R. Watkins & S. Gupta. (2011). Role of stochastic processes in maintaining discrete strain structure in antigenically diverse pathogen populations. *PNAS* 108:15504-15509.

[40] R. Burger. (2000). *The Mathematical Theory of Selection, Recombination, and Mutation.* John Wiley & Sons Ltd., Chichester.

[41] S. Busenberg, K.L. Cooke & M.A. Pozio. (1983). Analysis of a model of a vertically transmitted disease. *J. Math. Bio.,* 17:305-329.

[42] R.M. Bush, W.M. Fitch, C.A. Bender & N.J. Cox. (1999). Positive selection on the H3 hemagglutinin gene of human influenza virus. *Molecular Biology & Evolution,* 16(11):1457-1465.

[43] V. Calvez, A. Korobeinikov & P.K. Maini. (2005). Cluster formation for multi-strain infections with cross-immunity. *J. Theor. Biol.,* 233:75-83.

[44] J. Cao & J.O. Ramsay. (2007). Parameter cascades and profiling in functional data analysis. *Computational Statistics*, 22(3):335-351.

[45] F. Carrat & A. Flahault. (2007). Influenza vaccine: the challenge of antigenic drift. *Vaccine,* 25(39-40):6852-6862.

[46] C. Castillo-Chavez, H.W. Hethcote, V. Andreasen, S.A. Levin & W.M. Liu. (1989). Epidemiological models with age structure, proportionate mixing and cross-immunity. *J. Math. Biol.,* 27:233-258.

[47] C. Castillo-Chavez & J.X. Velasco-Hernandez. (1998a). On the relationship between evolution of virulence and host demography. *J. Theor. Bio.,* 192:437-444.

[48] C. Castillo-Chavez & J.X. Velasco-Hernandez. (1998b). On the relationship between evolution of virulence and host demography. *J. Theor. Bio.,* 192:437-444.

[49] C. Castillo-Chavez, Z. Feng & W. Huang. (2002). On the computation of R_0 and its role on global stability. In: C. Castillo-Chavez, S. Blower, P. van den Driessche and D. Kirschner, (eds). *Mathematical Approaches for Emerging and Reemerging Infectious Diseases Part I: An Introduction to Models, Methods and Theory.* Springer-Verlag, Berlin.

[50] Centers for Disease Control and Prevention. Swine influenza A (H1N1) infection in two children-Southern California, March-April 2009. (2009a) *Morb. Mortal. Wkly Rep.*, 58:400-402.

[51] Centers for Disease Control and Prevention. (2009b). Novel influenza A virus infections-2007 case definition. `http://www.cdc.gov/ncphi/disss/nndss/casedef/novel_influenzaA.htm`.

[52] B.S. Chan & P. Yu. (2013). Bifurcation, stability, and cluster formation of multi-strain infection models. *J Math Biol.* 67(6-7):1507-32.

[53] B.A. Chapman, M.J. Burt, C.M. Frampton, J.A. Collett, K.H. Yeo, I.D. Wilkinson, H.D. Cook, M.J. Barclay, A.G. Ross & P.M. George. (2000). The prevalence of viral hepatitis (HAV, HBC and HCV) in the Christchurch community. *New Zealand Medical Journal,* 113:394-396.

[54] M.I. Chen, V.J. Lee, W.Y. Lim, I.G. Barr, R.T. Lin, G.C.H. Koh, J. Yap, L. Cui, A.R. Cook, K. Laurie, L.W.L. Tan, B.H. Tan, J. Loh, R. Shaw, C. Durrant, V.T.K. Chow, A. Kelso, K.S. Chia & Y.S. Leo. (2010). 2009 influenza A(H1N1) seroconversion rates and risk factors among distinct adult cohorts in Singapore. *JAMA* 303:1381-1391.

[55] A. Cherif, V. Garcia-Horton, G. Meléndez-Rosario, W. Feliciano, B. Crawford, J. Vega-Guzmán, F. Sánchez, & C. Kribs-Zaleta. (2008). *A Tale of Two Regions: A Mathematical Model for Chagas' Disease.* Technical Report, MTBI-05-06M. Arizona State University.

[56] Y.K. Choi, J.H. Lee, G. Erickson, S.M. Goyal, H.S. Joo, R.G. Webster & R.J. Webby. (2004). H3N2 influenza virus transmission from swine to turkeys, United States. *Emerg. Infect. Dis.,* 10:2156-2160.

[57] P. Chossat and R. Lauterbach (2000) Methods in Equivariant Bifurcations and Dynamical Systems. World Scientific, Singapore.

[58] G. Chowell, S. EchevarrŠa-Zuno, C. Viboud, L. Simonsen, J. Tamerius, M.A. Miller & V.H. Borja-Aburto. (2011). Characterizing the Epidemiology of the 2009 Influenza A/H1N1 Pandemic in Mexico. *PLoS Med.,* 8(5):e1000436.

[59] A.D. Cliff, P. Haggett & J.K. Ord. (1986). *Spatial Aspects of Influenza Epidemics.* Pion, London.

[60] N.J. Cox & K. Subbarao. (2000). Global Epidemiology of Influenza: Past and Present. *Annu. Rev. Med.,* 51:407-421.

[61] R.J. Cox & K.A. Brokstad. (1999). The post-vaccination antibody response to influenza virus proteins. *Apmis,* 107:289-296.

[62] J. Dávila, G. Chowell, V.H. Borja-Aburto, C.Viboud, C.G. Muiz & M. Miller. (2014). Substantial Morbidity and Mortality Associated with Pandemic A/H1N1 Influenza in Mexico, Winter 2013-2014: Gradual Age Shift and Severity. *PLoS Curr. 6.*

[63] J.H.P. Dawes & J.R. Gog. (2002). The onset of oscillatory dynamics in models of multiple disease strains. *J. Math. Biol.,* 45:471-510.

[64] G.F. Deboer, W. Back & A.D.M.E. Osterhaus. (1990). An Elisa for detection of antibodies against influenza-A nucleoprotein in humans and various animal species. *Archives of Virology,* 115:47-61.

[65] Y. Deng, X.H. Pang, P. Yang, W.X. Shi, L.L. Tian, B.W. Liu, S. Li, S. J. Cui, Y. Li, G.L. Lu, L. Zhang, X. Zhang, B. Liu, H. Seale, F. Huang & Q.Y. Wang. (2010). Serological survey of 2009 H1N1 influenza in residents of Beijing, China. *Epidemiology & Infection,* 139:1145-1150.

[66] L. Denoeud, C. Turbelin, S. Ansart, A.-J. Valleron, A. Flahault & F. Carrat. (2007). Predicting pneumonia and influenza mortality from morbidity data. *PLoS One,* 2(5):e464.

[67] O. Diekmann & J.A.P. Heesterbeek. (2000). *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation.* Wiley, NY.

[68] O. Diekmann & P. Getto. (2005). Boundedness, global existence and continuous dependence for nonlinear dynamical systems describing physiologically structured populations. *J. Differential Equations,* 215(2):268-319.

[69] O. Diekmann, M. Gyllenberg, H. Huang, M. Kirkilionis, J. A. J. Metz & H. R. Thieme. (2001). On the formulation and analysis of general deterministic structured population models. II. Nonlinear theory. *J. Math. Biol.,* 43(2):157-189.

[70] O. Diekmann & J. A. J. Metz, eds. (1986). The dynamics of physiologically structured populations, volume 68 of Lecture Notes in Biomathematics. Springer-Verlag, Berlin.

[71] U. Dieckmann, J. Metz, M. Sabelis & K. Sigmund. (2002). *Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management.* Cambridge University Press, UK.

[72] K. Dietz. (1979). Epidemiological interference of virus populations. *J. Math. Biol.* 8:291-300.

[73] K. Dietz & J.A.P. Heesterbeek. (2002). Daniel Bernoulli's epidemiological model revisited. *Math. Biosci.,* 180:1-21.

[74] J. Dushoff, J.B. Plotkin, S.A. Levin & D.J. Earn. (2004). Dynamical resonance can account for seasonality of influenza epidemics. *PNAS,* 101:16915-16916.

[75] W.R. Dowdle. (1999). Influenza A virus recycling revisited. *Bull. World Health Organization,* 77:820-828.

[76] D.J.D. Earn, J. Dushoff & S.A. Levin. (2002). Ecology and evolution of the flu. *Trends in Ecology & Evolution,* 17(7):334-340.

[77] L.R. Elveback, J.P. Fox & A. Varma. (1964). An extension of the Reed-Frost epidemic model for the study of competition between viral agents in the presence of interference. *Am. J. Hyg.* 80:356-364.

[78] W. Farr. (1866). On the cattle plague. Letter to the Editor of the Daily News (London).

[79] C.P. Farrington, M.N Kanaan & N.J. Gay. (2001). Estimation of the basic reproductive number for infectious diseases from age-stratified serological survey data. *Appl. Statist.,* 50(3):251-292.

[80] C.P. Farrington & H.J. Whitaker. (2005). Contact surface models for infectious diseases: Estimation from serologic survey data. *J. Am. Stat. Assoc.,* 100(470):370-379.

[81] N.M. Ferguson & V. Andreasen. (2001). The influence of different forms of cross-protective immunity on the population dynamics of antigenically diverse pathogens. In: S. Blower, C. Castillo-Chavez, K.L. Cooke, D. Kirschner, P. van der Driessche (eds). *Mathematical Approaches for Emerging and Re-emerging Infections*, IMAVolumes in Mathematics and its Applications. Springer, NewYork.

[82] N.M. Ferguson, A.P. Galvani & R.M. Bush. (2003). Ecological and immunological determinants of influenza evolution. *Nature,* 422:428-433.

[83] P.E.M. Fine. (1975). Vectors and vertical transmission: an epidemiologic perspective. *Annals NY Acad. Sci.,* 266:173-194.

[84] P.E.M. Fine. (1977). A commentary on the mechanical analogue to the Reed-Frost epidemic model. *Am. J. Epidemiol.,* 106(2): 87-100.

[85] P.E.M. Fine. (1979). John Brownlee and the Measurement of Infectiousness: An Historical Study in Epidemic Theory. *J. Royal Statistical Soc., Series A*, 142(3):347-362.

[86] B.F. Finkenstadt, A. Morton & D.A. Rand. (2005). Modelling antigenic drift in weekly flu incidence. *Stat. Med.,* 24: 3447-3461.

[87] W.M. Fitch, R.M. Bush, C.A. Bender & N.J. Cox. (1997). Long term trends in the evolution of H(3) HA1 human influenza type A. *PNAS,* 94(15):7712-7718.

[88] T. Francis Jr. (1960). On the doctrine of original antigenic sin. *Proc. Am. Phil. Soc.* 104(6):572-578.

[89] C. Fraser, C.A. Donnelly, S. Cauchemez, W.P. Hanage, M.D. Van Kerkhove, T.D. Hollingsworth, J. Griffin, R.F. Baggaley, H.E. Jenkins, E.J. Lyons, T. Jombart, W.R. Hinsley, N.C. Grassly, F. Balloux, A.C. Ghani, N.M. Ferguson, A. Rambaut, O.G. Pybus, H. Lopez-Gatell, C.M. Alpuche-Aranda, I.B. Chapela, E.P. Zavala, D.M.E. Guevara, F. Checchi, E. Garcia, S. Hugonnet, C. Roth, & The WHO Rapid Pandemic Assessment Collaboration. (2009). Pandemic potential of a strain of influenza A (H1N1): Early findings. *Science* 324(5934):1557-1561.

[90] R. Gani, H. Hughes, D. Fleming, T. Griffin, J. Medlock & S. Leach. (2005). Potential impact of antiviral drug use during influenza pandemic. *Emerg. Infect. Dis.,* 11:1355-1362.

[91] J.C. Gaydos, F.H. Top, R.A. Hodder & P.K. Russell. (2006). Swine influenza a outbreak, Fort Dix, New Jersey, 1976. *Emerging Infect. Dis.,* 12(1):23-28.

[92] E. Ghedin, N.A. Sengamalay, M. Shumway, J. Zaborsky, T. Feldblyum, V. Subbu, D.J. Spiro, J. Sitz, H. Koo, P. Bolotov, D. Dernovoy, T. Tatusova, Y. Bao, K. St George, J. Taylor, D.J. Lipman, C.M. Fraser, J.K. Taubenberger & S.L. Salzberg. (2005). Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. *Nature,* 437:1162-1166.

[93] G.L. Gilbert, M.A. Cretikos, L. Hueston, G. Doukas, B. O'Toole, & D.E. Dwyer. (2010). Influenza A (H1N1) 2009 antibodies in residents of New South Wales, Australia, after the first pandemic wave in the 2009 southern hemisphere winter. *PLoS One* 5:e12562.

[94] C. Godsil & G. Royle. (2001). Algebraic graph theory, volume 207 of Graduate Texts in Mathematics. Springer, New York

[95] J.R. Gog. (2008). The impact of evolutionary constraints on influenza dynamics. *Vaccine,* 26(3):C15-C24.

[96] J.R. Gog & B.T. Grenfell. (2002). Dynamics and selection of many-strain pathogens. *PNAS,* 99:17209-17214.

[97] J.R. Gog, G.F. Rimmelzwaan, A.D. Osterhaus & B.T. Grenfell. (2003). Population dynamics of rapid fixation in cytotoxic T lymphocyte escape mutants of influenza A. *Proc. Natl Acad. Sci.,* 100:11143-11147.

[98] J.R. Gog & J. Swinton. (2002). A status-based approach to multiple strain dynamics. *J. Math. Biol.,* 44:169-184.

[99] M. Golubitsky, I. Stewart and A. Torok. (2005). Patterns of synchrony in coupled cell networks with multiple arrows. SIAM J. Appl. Dynam. Sys. 4(1):78-100.

[100] M. Golubitsky & I. Stewart. (2006). Nonlinear dynamics of networks: the groupoid formalism. *Bull. Amer. Math. Soc.* 43:305-364.

[101] M.G.M. Gomes, G. Medley & D.J. Nokes. (2002). On the determinants of population structure in antigenically diverse pathogens. *Proc. R. Soc. Lond. B,* 269:227-233.

[102] S.K. Greene, E.L. Ionides & M.L. Wilson. (2006). Patterns of Influenza-associated Mortality among US Elderly by Geographic Region and Virus Subtype, 1968-1998. *Am. J. Epid.,* 163(4):316-326.

[103] B.T. Grenfell, O.G. Pybus, J.R. Gog, J.L.N Wood, J.M. Daly, J.A. Mumford & E.C. Holmes. (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* 303: 327-332.

[104] S. Gupta, K. Trenholme, R.M. Anderson & K.P. Day. (1994). Antigenic diversity and the transmission dynamics of Plasmodium falciparum. *Science,* 263:961-963.

[105] S. Gupta & K.P. Day. (1994). A strain theory of malaria transmission. *Parasitology Today,* 10:476-481.

[106] S. Gupta, M.C.J. Maiden, I.M. Feavers, S. Nee, R.M. May & R.M. Anderson. (1996). The maintenance of strain structure in populations of recombining infectious agents. *Nat. Med.,* 2:437-442.

[107] S. Gupta, N.M. Ferguson & R.M. Anderson. (1997a). A strain theory of malaria transmission. *Parasitology Today,* 10:476-481.

[108] S. Gupta, N.M. Ferguson & R.M. Anderson. (1997b). Vaccination and the population structure of antigenically diverse pathogens that exchange genetic matierals. *Proc. Royal Soc. Lond. B,* 264:1435-1443.

[109] S. Gupta, N.M. Ferguson & R.M. Anderson. (1998). Chaos, persistence, and evolution of strain structure in antigenically diverse infectious agents. *Science,* 280: 912-915.

[110] S. Gupta & R.M. Anderson. (1999). Population structure of pathogens: the role of immune selection. *Parasitol. Today,* 15:497-501.

[111] S. Gupta & A. Galvani. (1999). The effects of host heterogeneity on pathogen population structure. Proc Roy. Soc. B 354: 711-719

[112] A.W. Hampson. (2002). Influenza virus antigens and antigenic drift. In C. W. Potter. *Influenza*. Elsevier Science B.V., pp. 49-86.

[113] K. Hancock, V. Veguilla, X. Lu, W. Zhong, E.N. Butler, H. Sun, F. Liu, L. Dong, J.R. DeVos, P.M. Gargiullo, T.L. Brammer, N.J. Cox, T.M. Tumpey & J.M. Katz. (2009). Cross-reactive antibody responses to the 2009 pandemic H1N1 influenza virus. *New Engl. J. Med.,* 361:1945-1952.

[114] H.W. Hethcote. (1997). An age-structured model for pertussis transmission. *J. Math. Biosci.,* 145:89-136.

[115] T. Hildebrandt & L. Graves. (1927). Implicit functions and their differentials in general analysis. *Trans. Amer. Math. Soc.* 29:127-153.

[116] V.L. Hilts. (1980). Epidemiology and the statistical movement. In *Aspects of the History of Epidemiology, Times, Places and Persons*, A.M. Lilienfeld (ed). The Johns Hopkins Univ. Press, Baltimore.

[117] S.S.H. Huang, Z. Lin, D. Banner, A.J. León, S.G. Paquette, B. Rubin, S. Rubino, Y. Guan, D.J. Kelvin & A.A. Kelvin. (2013). Immunity toward H1N1 influenza hemagglutinin of historical and contemporary strains suggests protection and vaccine failure. *Scientific Reports,* 3, 1698.

[118] K.C. Hyams, J. Riddle, D.H. Trump & J.T. Graham. (2001). Endemic infectious diseases and biological warfare during the Gulf War: A decade of analysis and final concerns. *Am. J. Trop. Med. Hygiene,* 65:664-670.

[119] M. Iannelli. (1995). *Mathematical Theory of Age-Structured Population Dynamics.* Applied Mathematics Monograph C.N.R. Vol. 7, Giardini editorie stampatori, Pisa.

[120] H. Inaba. (1988). A semigroup approach to the strong ergodic theorem of the multistate stable population process. *Math. Population Stud.,* 1(1):49-77.

[121] H. Inaba. (1990). Threshold and stability results for an age-structured epidemic model. *J. Math. Biol.,* 28:411-434.

[122] F. Ingemar & C. Scott. (2001). Implications of horizontal and vertical pathogen transmission for honey bee epidemiology. *Apidologie,* 32:199-214.

[123] P. H. Karmel. (1947). The relations between male and female reproduction rates. *Pop. Studies*, 1:249-274.

[124] P. H. Karmel. (1948a). The relations between male and female nuptiality in a stable population. *Pop. Studies*, 1:353-387.

[125] P. H. Karmel. (1948b). An analysis of the sources and magnitudes of inconsistencies between male and female net reproductive rates in actual populations. *Pop. Studies*, 2:240-273.

[126] P. H. Karmel. (1948c). A rejoinder to Mr. Hajnal's comments. *Pop. Studies*, 2:361-372.

[127] C.L. Karp & P.G. Auwaerter. (2007). Co-infection with HIV and tropical infectious diseases. Part II: Helminthic, fungal, bacterial, and viral pathogens. *Clinical Infectious Diseases,* 45:1214-1220.

[128] T. Kato. (1984). *Perturbation Theory for Linear Operators*, 2nd Ed. Springer, Berlin.

[129] M.J. Keeling. (1999). The effects of local spatial structure on epidemiological invasions. *Proc. R. Soc. B,* 266:859-867.

[130] M.J. Keeling. (2000a). Evolutionary trade-offs at two time-scales: competition versus persistence. *Proc. Biol. Sci.,* 267:385-391.

[131] M.J. Keeling. (2000b). Evolutionary dynamics in spatial host-parasite systems. In *The Geometry of Ecological Interactions,* U. Dieckmann, R. Law and J.A.J. Metz (Eds.) Cambridge University Press, Cambridge (UK).

[132] M.J. Keeling & K.T.D. Eames. (2005). Networks and epidemic models. *J. R. Soc. Interface,* 2:295-307.

[133] W.O. Kermack & A.G. McKendrick. (1927). Contributions to the mathematical theory of epidemics. *Proc. R. Soc. A,* 115:700-721.

[134] W.O. Kermack & A.G. McKendrick. (1932). A Contributions to the mathematical theory of epidemics, Part 2: The problem of endemicity. *Proc. R. Soc. A,* 138:55-83.

[135] W.O. Kermack & A.G. McKendrick. (1933). Contributions to the mathematical theory of epidemics, Part 3: Further studies of the problem of endemicity. *Proc. R. Soc. A,* 141:94-122.

[136] J.K. Kim, I. Skountzou, R. Compans & J. Jacob. (2009). Original antigenic sin responses to influenza viruses. *J. Immunol.* 183(5):3294-3301.

[137] A.S. Klovdahl, J.J. Potterat, D.E. Woodhouse, J.B. Muth, S.Q. Muth & W.W. Darrow. (1994). Social networks and infectious disease: the Colorado Spring study. *Soc. Sci. Med.,* 38:79-88.

[138] K. Koelle, S. Cobey, B.T. Grenfell & M. Pascual. (2006). Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. *Science,* 314:1898-1903.

[139] M. Koopmans, B. Wilbrink, M. Conyn, G. Natrop, H. van der Nata, H. Vennema, A. Meijer, J. van Steenbergen, R. Fouchier, A. Osterhaus & A. Bosman. (2004). Transmission of H7N7 avian influenza A virus to human beings during a large outbreak in commercial poultry farms in The Netherlands. *Lancet,* 363:587-593.

[140] A. Korobeinikov & G.C. Wake. (2002). Lyapunov functions and global stability for SIR, SIRS, and SIS Epidemiological Models. *Appl. Math. Letters,* 15:955-960.

[141] A. Korobeinikov & P.K. Maini. (2004). A Lyapunov function and global properties for SIR and SEIR epidemiological models with nonlinear incidence. *Math. Biosci. and Eng.* 1(1):57-60.

[142] A. Korobeinikov (2006). Lyapunov functions and global stability for SIR and SIRS epidemiological models with nonlinear transmission. *Bull. Math. Biol.* 30:615-626

[143] M.A. Krasnoselskii. (1964). *Positive Solutions of Operator Equations.* Noordhoff, Groningen.

[144] M.G. Krein & M.A. Rutman. (1948). Linear operators leaving invariant a cone in a Banach space. *Uspehi. Mat. Nauk.,* 3:3-95. (in Russian); English translation: Amer. Math. Soc. Transl. 1(10):199-325, 1950.

[145] E. Kreyszig. (1978). *Introductory Functional Analysis.* John Wiley & Sons, New York.

[146] S. Kryazhimskiy, U. Dieckmann, S.A. Levin & J. Dushoff. (2007). On state-space reduction in multi-strain pathogen models, with an application to antigenic drift in influenza A. *PLoS Comput. Biol.,* 3:e159.

[147] R.J. Kubiak & A.R. McLean. (2012). Why Was the 2009 Influenza Pandemic in England So Small? *PLoS ONE* 7(2):e30223.

[148] R. Kuske, L.F. Gordillo & P. Greenwood. (2007). Sustained oscillations via coherence resonance in SIR. *J. Theor. Bio.,* 245:459-469.

[149] P.-H. Lambert, M. Liu & C.-A. Siegrist. (2005). Can successful vaccines teach us how to induce efficient protective immune responses? *Nat. Med.* 11:S54-S6.

[150] O. Larsen, S. Andersson, Z. da Silva, K. Hedegaard, A. Sandström, A. Nauclér, F. Dias, M. Melbye & P. Aaby. (2000). Prevalences of HTLV-1 infection and associated risk determinants in an urban population in Guinea-Bissau, West Africa. *J. Acquired Immune Deficiency Syndrome,* 25:157-163.

[151] J. Lauri & R. Scapellato. (2003). Topics in graph automorphisms and reconstruction. Cambridge University Press, New York.

[152] J. Lin, V. Andreasen & S.A. Levin. (1999). Dynamics of influenza A drift: the linear three-strain model. *Math. Biosci.* 162:33-51.

[153] A.L. Lloyd & D. Wodarz. (2006). Drug Resistance in Acute Viral Infections: Rhinovirus as a Case Study. In: *Disease evolution: Models, Concepts, and Data Analyses*, Z. Feng, U. Diekmann and S. Levin (eds.), AMS/DIMACS.

[154] P. Magal & S. Ruan. (2007). On integrated semigroups and age structured models in Lp spaces. *Diff. Integr. Equations,* 20:197-239.

[155] P. Magal & S. Ruan. (2009a). Center manifolds for semilinear equations with non-dense domain and applications on Hopf bifurcation in age structured models. *Mem. Amer. Math. Soc.,* 202, no. 951.

[156] P. Magal & S. Ruan. (2009b). On semilinear Cauchy problems with non-densely defined domain. *Adv. Diff. Equations,* 14:1041-1084.

[157] P. Magal & S. Ruan. (2010). Sustained oscillations in an evolutionary epidemiological model of influenza A drift. *Proc. R. Soc. A,* 466:965-992.

[158] B. Manicassamy, & & et al. (2010). Protection of mice against lethal challenge with 2009 H1N1 influenza A virus by 1918-like and classical swine H1N1 based vaccines. *PLoS Pathogens,* 6(1):e1000745.

[159] I. Marek. (1970). Frobenius theory of positive operators: comparison theorems and applications. *SIAM J. Appl. Math.,* 19:607-628.

[160] A.G. McKendrick. (1926). Application of mathematics to medical problems. Proc. Edinb. Math. Soc., 54:98-130

[161] E. Miller, K. Hoschler, P. Hardelid, E. Stanford, N Andrews & M. Zambon. (2010). Incidence of 2009 pandemic influenza A H1N1 infection in England: a cross-sectional serological study, *Lancet,* 375:1100-1108.

[162] P. Minayev & N.M. Ferguson. (2009a). Improving the realism of deterministic multistrain models: implications for modelling influenza A. *J. Royal Soc. Interface,* 6: 509-518.

[163] P. Minayev & N.M. Ferguson. (2009b). Incorporating demographic stochasticity into multi-strain epidemic models: application to influenza A. *J. Royal Soc. Interface,* 6(40):989-996.

[164] D. Mollison. (1995). *Epidemic models their structure and relation to data.* Cambridge University Press, New York.

[165] D.H. Molyneux. (2008). Combating the other diseases of MDG 6: changing the paradigm to achieve equity and poverty reduction? *Trans. R. Soc. Trop. Med. Hygiene,* 102:509-519.

[166] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G.S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska & W.J. Edmunds. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine,* 5(3):e74.

[167] A.S. Monto & A.P. Kendal. (1973). Effect of neuraminidase antibody on Hong-Kong influenza. *Lancet,* 1:623-625.

[168] S.S. Morse. (2004). Factors and determinants of disease emergence. *Revue Scientifique et Technique del Office International des Epizooties,* 23:443-451.

[169] R. Nagel (ed.) (1986). *One-Parameter Semigroups of Positive Operators.* Springer, Berlin.

[170] K-J. Engel & R. Nagel. (2000). *One-parameter semigroups for linear evolution equations.* Springer, Berlin.

[171] M. Nuno, M. Martcheva & C. Castillo-Chavez. (2009). Immune level approach For multiple strain pathogens. *J. Biol. Sys.* 17(4): 713-737.

[172] P. Parham. (2004). *The Immune System,* 2nd Ed. Garland Science, New York.

[173] A. Pazy. (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Springer, Berlin.

[174] J.S.M. Peiris, Y. Guan, D. Markwell, P. Ghose, R. G. Webster & K.F. Shortridge. (2001). Cocirculation of avian H9N2 and contemporary human H3N2 influenza A viruses in pigs in southeastern China: Potential for genetic reassortment? *J. of Virology,* 75:9679-9686.

[175] M. Pensaert, K. Ottis, J. Vanderputte, M.M. Kaplan & P.A. Buchmann. (1981). Evidence for the natural transmission of influenza A virus from wild ducks to swine and its potential for man. *Bull. World Health Organ.,* 59:75-78.

[176] B. Perthame. (2007). *Transport Equations in Biology*. Birkhauser Verlag, Basel.

[177] S.A. Plotkin. (2006). *Mass Vaccination Global Aspects-Progress and Obstacles*. Springer, Berlin.

[178] J.B. Plotkin, J. Dushoff & S.A. Levin. (2002). Hemagglutinin sequence clusters and the antigenic evolution of influenza A virus. *Proc. Natl. Acad. Sci.,* 99:6263-6268.

[179] C.W. Potter. (1998). Chronicle of influenza pandemic. In *Textbook of Influenza*, K. G. Nicholson, R. G. Webster and A. J. Hay (Eds.) Blackwell Science, Oxford, pp. 3-18.

[180] A.A. Poyton, M.S. Varziri, K.B. McAuley, P.J. McLellan, & J.O. Ramsay. (2006). Parameter estimation in continuous-time dynamic models using principal differential analysis. *Computers & chemical engineering*, 30(4):698-708.

[181] A. Rambaut, O.G. Pybus, M.I. Nelson, C. Viboud, J.K. Taubenberger & E.C. Holmes. (2008). The genomic and epidemiological dynamics of human influenza A virus. *Nature,* 453:615-619.

[182] J.O. Ramsay & B.W. Silverman. (2002). *Applied functional data analysis: methods and case studies* (Vol. 77). New York: Springer.

[183] A. Ransome. (1868). On epidemics, studied by means of statistics of disease. *The British Medical Journal,* 2(406):386-388.

[184] J.M. Read & M.J. Keeling. (2003). Disease evolution on networks: the role of contact structure. *Proc. R. Soc. B.,* 270:699-708.

[185] J.M. Read & M.J. Keeling. (2006). Disease evolution across a range of spatio-temporal scales. *Theo. Pop. Biol.* 70:201-213.

[186] M. Recker, et al. (2004). Transient cross-reactive immune responses can orchestrate antigenic variation in malaria. Nature 429:555-558

[187] M. Recker and S. Gupta. (2005). A model for pathogen population structure with cross-protection depending on the extent of overlap in antigenic variant repertoires. J Theor Bio. 232:363-373

[188] M. Recker, O.G. Pybus, S. Nee & S. Gupta. (2007). The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types. *Proc. Natl. Acad. Sci.,* 104:7711-7716.

[189] M. Recker, et al. (2009). Immunological serotype interactions and their effect on the epidemiological pattern of dengue. Proc Roy. Soc. London B 276: 2541-2548

[190] R. Ross. (1910). *The Prevention of Malaria*. Dutton, New York.

[191] R. Ross. (1911). *The Prevention of Malaria*, 2nd Ed., with an addendum on the theory of happenings. Murray, London.

[192] R. Ross. (1916). An application of the theory of probabilities to the study of a priori pathometry, Part I. *Proc. Royal Soc. Series A,* 92(638):204-230.

[193] R. Ross & H.P. Hudson. (1917a). An application of the theory of probabilities to the study of a priori pathometry, Part II. *Proc. Royal Soc. Series A,* 93(650):212-225.

[194] R. Ross & H.P. Hudson. (1917b). An application of the theory of probabilities to the study of a priori pathometry, Part III. *Proc. Royal Soc. Series A,* 93(650):225-240.

[195] G. Rost & J. Wu. (2008). SEIR epidemiological model with varying infectivity and infinite delay. *Math. Biosci. Eng.,* 5:389-402.

[196] W. Rudin. (1966). *Real and complex analysis.* McGraw-Hill, New York.

[197] K. Ruxrungtham, T. Brown & P. Phanuphak. (2004). HIV/AIDS in Asia. *Lancet,* 364:69-82.

[198] A. Sasaki & Y. Haraguchi. (2000). Antigenic drift of viruses within a host: a finite site model with demographic stochasticity. J. Mol. Evol. 51: 245-255.

[199] K. Sato, T. Morishita, E. Nobusawa, Y. Suzuki, Y. Miyazaki, Y. Fukui, S. Suzuki & K. Nakajima. (2000). Surveillance of influenza viruses isolated from travellers at Nagoya International Airport. *Epidemiology & Infection,* 124:507-514.

[200] K. Sato, T. Morishita, E. Nobusawa, K. Tonegawa, K. Sakae, S. Nakajima & K. Nakajima. (2004). Amino-acid change on the antigenic region B1 of H3 haemagglutinin may be a trigger for the emergence of drift strain of influenza A virus. *Epidemiology & Infection,* 132:399-406.

[201] I. Sawashima. (1964). On spectral properties of some positive operators. *Nat. Sci. Rep. Ochanomizu Univ.,* 15(1964):53-64.

[202] J.F.D. Shrewsbury. (1970). *A history of Bubonic Plague in the British Isles.* Cambridge University Press, London.

[203] J.W. Sinko & W. Streifer. (1967). A New model for age-size structure of a population. *Ecological Society of America,* 48(6):910-918.

[204] D.M. Skowronski, T.S. Hottes, N.Z. Janjua, D. Purych, S. Sabaiduc, T. Chan, G. De Serres, J. Gardy, J.E. McElhaney, D.M. Patrick & M. Petric. (2010). Prevalence of seroprotection against the pandemic (H1N1) virus after the 2009 pandemic. *CMAJ,* 182:1851-1856.

[205] D.J. Smith, A.S. Lapedes, J.C. de Jong, T.M. Bestebroer, G.F. Rimmelzwaan, A.D.M.E. Osterhaus & R.A.M. Fouchier. (2004). Mapping the antigenic and genetic evolution of influenza virus. *Science,* 305(5682):371-376.

[206] D.J. Smith, S. Forrest, D.H. Ackley, & A.S. Perelson. (1999). Variable efficacy of repeated annual influenza vaccination. *Proc. Nat. Acad. Sci.*, 96(24):14001-14006.

[207] G.J.D. Smith, D. Vijaykrishna, J. Bahl, S.J. Lycett, M. Worobey, O.G. Pybus, S.K. Ma, C.L. Cheung, J. Raghwani, S. Bhatt, J.S.M. Peiris, Y. Guan & A. Rambaut. (2009). Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature,* 459(7250):1122-1125.

[208] S. Steinberg. (1968). Meromorphic families of compact operators. *Arch. Rational Mech. Anal.,* 31:372-379.

[209] B.V. Tandale, S.D. Pawar, Y.K. Gurav, M.S. Chadha, S.S. Koratkar, V.N. Shelke & A.C. Mishra. (2010). Seroepidemiology of pandemic influenza A (H1N1) 2009 virus infections in Pune, India. *BMC Infect. Dis.,* 10(1):255.

[210] J. Taubenberger & D. Morens. (2006). 1918 Influenza: the mother of all pandemics. *Emerg. Infect. Dis.,* 12(1):15-22.

[211] H.R. Thieme. (2003). *Mathematics in Population Biology.* Princeton University, Princeton, Oxford.

[212] J. Truscott, C. Fraser, S. Cauchemez, A. Meeyai, W. Hinsley, C.A. Donnelly, A. Ghani & N. Ferguson (2011). Essential epidemiological mechanisms underpinning the transmission dynamics of seasonal influenza. *J. Royal Society Interface*, rsif20110309.

[213] J.M. Varah. (1982). A spline least squares method for numerical parameter estimation in differential equations. *SIAM Journal on Scientific and Statistical Computing*, 3(1):28-46.

[214] J.L. Virelizier, R. Postlethwaite, G.C. Schild & A.C. Allison. (1974a). Antibody responses to antigenic determinants of influenza virus hemagglutinin I: Thymus Dependence of Antibody Formation and Thymus Independence of Immunological Memory.. *J. Exp. Med.*, 140(6):1559-1570.

[215] J.L. Virelizier, A.C. Allison & G.C. Schild. (1974b). Antibody responses to antigenic determinants of influenza virus hemagglutinin II. Original antigenic sin: a bone marrow-derived lymphocyte memory phenomenon modulated by thymus-derived lymphocytes. *J. Exper. Med.*, 140(6):1571-1578.

[216] Y. Wang & M. Golubitsky. (2005).Two-color patterns of synchrony in lattice dynamical systems. *Nonlinearity.* 18:631-657

[217] D.J. Watts & S.H. Strogatz. (1998). Collective dynamics of small-world networks. *Nature,* 393:440-4423.

[218] G.F. Webb. (1985). *Theory of Nonlinear Age-Dependent Population Dynamics.* Marcel Dekker, New York.

[219] G. F. Webb. (2004). Structured population dynamics. Polish Acad. Sci. Banach Center Publ. 63:177-186.

[220] R.J. Webby & R.G. Webster. (2003). Are we ready for pandemic influenza? *Science,* 302:1519-1522.

[221] R.G. Webster. (1968). The immune response to influenza virus: III. Changes in the avidity and specificity of early IgM and IgG antibodies. *Immunology* 14(1):39-52.

[222] R.G. Webster, W.J. Bean, O.T. Gorman, T.M. Chambers and Y. Kawaoka. (1992). Evolution and ecology of influenza A viruses. *Microbiol. Rev.* 56:152-179.

[223] R.G. Webster, J.A. Kasel, R.B. Couch, W.G. Laver. (1976). Influenza virus subunit vaccines II. Immunogenicity and original antigenic sin in humans. *J. Infecct. Dis.* 134(1):48-58.

[224] M. Whitehead. (2000). William Farr's legacy to the study of inequalities in health. *Bull. World Health Organ.*, 78(1):86-87.

[225] P.S. Wikramaratna & S. Gupta. (2009). Influenza outbreaks. *Cellular Microbiology,* 11(7):1016-1024.

[226] P.S. Wikramaratna, C.P. Simmons, S. Gupta & M. Recker. (2010). The effects of tertiary and quaternary infections on the epidemiology of dengue. *PLoS One,* 5(8):e12347.

[227] P.S. Wikramaratna, M. Sandeman, M. Recker & S. Gupta. (2013). The antigenic evolution of influenza: drift or thrift? *Phil. Trans. R. Soc. B* 368(1614):20120200.

[228] S.S.Y. Wong & K. Yuen. (2006). Avian influenza virus infections in humans. *Chest,* 129:156-168.

[229] World Health Organization. (2003). Fact sheet number 211: Influenza. *WHO Fact sheets.*

[230] World Health Organization. (2005a). *Global Tuberculosis Control: Surveillance, Planning, Financing.* WHO/HTM/TB/2005.349, Geneva.

[231] World Health Organization. (2005b). *Avian Influenza: Assessing the Pandemic Threat.* WHO/CDS/2005.29, Geneva.

[232] World Health Organization & UNAIDS. (2005c). *AIDS Epidemic Update.* UNAIDS/05.19E.

[233] World Health Organization & UNICEF. (2005d). *World Malaria Report 2005.* WHO/HTM/MAL/2005.1102.

[234] World Health Organization. (2010). Pandemic (H1N1) 2009-update 100. *Disease Outbreak News*, Retrieved 14 May 2010.

[235] J.T. Wu, E.S.K. Ma, C.K. Lee, D.K.W. Chu, P.-L. Ho, A.L. Shen, A. Ho, I.F.N. Hung, S. Riley, L.M. Ho, C.K. Lin, T. Tsang, S.-V. Lo, Y.-L. Lau, G.M. Leung, B.J. Cowling & J.S.M. Peiris. (2010). The infection attack rate and severity of 2009 pandemic H1N1 influenza in Hong Kong. *Clin. Infect. Dis.,* 51:1184-1191.

[236] Y. Xiao, B.D. MacArthur, H.Wang, M. Xiong & W. Wang. (2008). Network quotients: Structural Skeletons of complex systems. *Phys. Rev. E,* (78) 046102

[237] X. Xu, C. B. Smith, B. A. Mungall, S. E. Lindstrom, H. E. Hall, K. Subbarao, N. J. Cox, & A. Klimov. (2002). Intercontinental circulation of human influenza A(H1N2) reassortant viruses during the 2001-2002 influenza season. *J. Infect. Dis.,* 1861490-1493.

[238] X.C. Yu, T. Tsibane, P.A. McGraw, F.S. House, C.J. Keefer, M.D. Hicar, T.M. Tumpey, C. Pappas, L.A. Perrone, O. Martinez, J. Stevens, I.A. Wilson, P.V. Aguilar, E.L. Altschuler, C.F. Basler & J.E. Crowe Jr. (2008). Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors. *Nature,* 455:532-536.

[239] E. Zeidler. (1986). *Nonlinear functional analysis and its applications I, Fixed Point theorems*. Springer, New York.

[240] S.M. Zimmer, C.J. Crevar, D.M. Carter, J.H. Stark, B.M. Giles, R.K. Zimmerman, S.M. Ostroff, B.Y. Lee, D.S. Burke & T.M. Ross . (2010). Seroprevalence following the second wave of Pandemic 2009 H1N1 influenza in Pittsburgh, PA, USA. *PLoS One,* 5:e11601.

[241] Y. Zhou, D.M.W. Ng, W.-H. Seto, D.K.M. Ip, H.K.H. Kwok, E.S.K. Ma, S. Ng, L.L.H. Lau, J.T. Wu, J.S.M. Peiris & B.J. Cowling. (2011). Seroprevalence of antibody to pandemic influenza A (H1N1) 2009 among healthcare workers after the first wave in Hong Kong. *J. Hospital Infection,* 78(4):308-11.