
Electronic Thesis and Dissertation Repository

8-18-2014 12:00 AM


Understanding Recurrent Disease: A Dynamical Systems Approach

Wenjing Zhang
The University of Western Ontario

Supervisor
Pei Yu
The University of Western Ontario Joint Supervisor
Lindi Wahl
The University of Western Ontario

Graduate Program in Applied Mathematics
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy
© Wenjing Zhang 2014

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>

 Part of the [Dynamic Systems Commons](#), [Immunity Commons](#), [Non-linear Dynamics Commons](#), [Ordinary Differential Equations and Applied Dynamics Commons](#), and the [Other Immunology and Infectious Disease Commons](#)

Recommended Citation

Zhang, Wenjing, "Understanding Recurrent Disease: A Dynamical Systems Approach" (2014). *Electronic Thesis and Dissertation Repository*. 2265.
<https://ir.lib.uwo.ca/etd/2265>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

UNDERSTANDING RECURRENT DISEASE: A DYNAMICAL SYSTEMS
APPROACH

(Thesis format: Integrated Article)

by

Wenjing Zhang

Graduate Program in Applied Mathematics

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Wenjing Zhang 2014

Abstract

Recurrent disease, characterized by repeated alternations between acute relapse and long remission, can be a feature of both common diseases, like ear infections, and serious chronic diseases, such as HIV infection or multiple sclerosis. Due to their poorly understood etiology and the resultant challenge for medical treatment and patient management, recurrent diseases attract much attention in clinical research and biomathematics. Previous studies of recurrence by biomathematicians mainly focus on in-host models and generate recurrent patterns by incorporating forcing functions or stochastic elements. In this study, we investigate deterministic in-host models through the qualitative analysis of dynamical systems, to reveal the possible intrinsic mechanisms underlying disease recurrence.

Recurrence in HIV infection is referred to as “viral blips”, that is, transient periods of high viral replication separated by long periods of quiescence. A 4-dimensional HIV antioxidant-therapy model exhibiting viral blips is studied using bifurcation theory. Four conditions for the existence of viral blips in a deterministic in-host model are proposed. Guided by the four conditions, the simplest 2-dimensional infection model which shows recurrence is obtained. One key point for recurrence is identified, that is an increasing and saturating infectivity function. Furthermore, Hopf and generalized Hopf bifurcations, Bogdanov-Takens bifurcation, and homoclinic bifurcation are proved to exist in this 2-dimensional model. Bogdanov-Takens bifurcation and homoclinic bifurcation provide a new mechanism for generating recurrence. From the viewpoint of modelling, the increasing and saturating infectivity function gives rise to a convex incidence rate, which further induces backward bifurcation and Hopf bifurcation, and allows the infection model to exhibit rich dynamical behavior, such as bistability, recurrence, and regular oscillation.

The relapse-remission cycle in autoimmune disease is investigated based on a regulatory T cell model. By introducing a newly discovered class of regulatory T cells, Hopf bifurcation occurs in the autoimmune model with negative backward bifurcation, and gives rise to a recurrent pattern.

The main insight of this thesis is that recurrent disease can arise naturally from the deterministic dynamics of populations. It will provide a starting point for further research in dynamical systems theory, and recurrence in other physical systems.

Co-Authorship Statement

The work in Chapter 2 was published in the following article:

Wenjing Zhang, Lindi M. Wahl and Pei Yu, 2013 Conditions for transient viremia in deterministic in-host models: Viral blips need no exogenous trigger. *SIAM Journal on Applied Mathematics*, 73(2):853–881, 2013. This paper was selected by *SIAM Review* as a SIGEST paper, published in 56(1), 127-155, 2014.

The work in Chapter 3 has been submitted for publication:

Wenjing Zhang, Lindi M. Wahl and Pei Yu, 2014 Modelling and analysis of recurrent autoimmune disease. *Under review for SIAM Journal on Applied Mathematics*.

The work in Chapters 4 and 5 are in preparation to be submitted.

Wenjing Zhang, Lindi M. Wahl and Pei Yu, 2014 Backward bifurcation underlies rich dynamics in simple disease models. *To be submitted to Journal of Mathematical Biology*.

Pei Yu, Wenjing Zhang and Lindi M. Wahl, 2014 Dynamical analysis of a 2-dimensional disease model with convex incidence. *To be submitted to SIAM Journal on Applied Dynamical Systems*.

Acknowledgements

Many people helped and supported me over the four years of graduate study. I would like to express the deepest appreciation to the following:

Dr. Pei Yu gave me this wonderful opportunity to study at Western University. His demand for academic rigour has made me a better researcher. I thank him for being so patient with me throughout this entire process. Without his guidance and persistent help this project would not have been possible.

Dr. Lindi Wahl inspires me by her knowledge, encouragement and patience. Her cheerful attitude and enthusiasm have always made a positive working environment. Also, her instruction has been invaluable in the research and the writing of this thesis.

The administration at the Department of Applied Mathematics has provided terrific support through out these years. I am very grateful for all the professors for their continuous encouragement and their valuable academic discussions. Audrey Kager and Cinthia MacLean have provided tremendous support and assistance. They have provided all the graduate students in the department with an efficient and comfortable learning environment. Financial support received from the Natural Sciences and Engineering Research Council of Canada and Western University are much appreciated.

I would like to thank Dr. Beverly Ulak for guiding me through the adjustment of oversea graduate study life for these years, the Social Committee in the department for organizing many events, such as Morning tea, potluck lunch and Christmas Party, Johannes Middeke for his computer expertise and inspiration, and my fellow students and office mates.

To my parents, I thank them for the unconditional love and support. To myself, I am very glad to experience graduate life and obtain a PhD degree. I am looking forward to a great future.

Keywords: Recurrent diseases, HIV viral blips, recurrent autoimmune diseases, dynamical system theory, bifurcation theory

Contents

Abstract	ii
List of Figures	viii
List of Tables	x
List of Abbreviations, Symbols, and Nomenclature	xi
1 Introduction	1
1.1 Mathematical models for studying recurrence	1
1.1.1 Immunological models	1
1.1.2 Infection models	3
1.2 Mathematical theories and methodologies used to study recurrence in biological models	4
1.2.1 Stability analysis for equilibrium solutions	4
1.2.2 Bifurcation analysis	5
1.2.2.1 Hopf bifurcation	5
1.2.2.2 Bogdanov-Takens bifurcation and homoclinic orbits	6
1.3 Thesis contribution and structure	6
1.4 References	7
2 Conditions for Transient Viremia in Deterministic In-host Models: Viral Blips Need no Exogenous Trigger	9
2.1 Introduction	9
2.2 A 4-dimensional model which exhibits viral blips	11
2.2.1 Well-posedness of the solutions of system (2.1)	13
2.2.2 Equilibrium solutions of (2.1) and their stability	13
2.2.3 Bifurcation analysis	15
2.2.3.1 Transcritical bifurcation	15
2.2.3.2 Hopf bifurcation and limit cycles	17
2.2.4 Conditions for generating viral blips	19
2.3 A simple 3-dimensional in-host infection model producing blips	21
2.3.1 Generalizing ROS to other physical variables	21
2.3.2 Identifying the region of parameter space exhibiting viral blips	22
2.3.2.1 State variable scaling and parameter rescaling	24
2.3.2.2 Equilibrium solutions and their stability	24

2.3.2.3	Parameter A fixed	25
2.3.2.4	Parameter C fixed	25
2.3.3	3-dimensional immunological model	28
2.4	A 2-dimensional in-host infection model	29
2.4.1	2-dimensional in-host model with constant and linear infection rates	29
2.4.2	A 2-dimensional in-host model with saturating infection rate	30
2.4.2.1	Scaling	30
2.4.2.2	Equilibrium solutions and their stability	30
2.4.2.3	Bifurcation analysis	30
2.5	Recurrency in a 5-dimensional model	31
2.5.1	Well-posedness of model (2.22)	32
2.5.2	Equilibrium solutions and their stability	34
2.5.3	Bifurcation analysis for $h \neq 0$	34
2.5.4	Bifurcation analysis for $h \rightarrow 0^+$	35
2.6	Conclusion and discussion	36
2.7	References	38
3	Modelling and Analysis of Recurrent Autoimmune Disease	42
3.1	Introduction	42
3.2	Model development	45
3.2.1	No recurrence in models (3.1) and (3.2)	46
3.2.2	Developing new models	48
3.3	Well-posedness, equilibrium solutions and stability of model (3.6)	49
3.3.1	Well-posedness	49
3.3.2	Equilibrium solutions	51
3.3.3	Stability of the disease-free equilibrium, E_0	51
3.3.4	Stability of the autoimmune disease equilibrium, E_1	53
3.4	Numerical simulation	55
3.5	Model reduction and parameter identification for autoimmune recurrence	57
3.5.1	Model reduction	58
3.5.2	Rescaling on system (3.26)	59
3.6	Conclusion and discussion	65
3.7	References	67
4	Backward Bifurcation Underlies Rich Dynamics in Simple Disease Models	72
4.1	Introduction	72
4.2	Backward bifurcation	76
4.2.1	Backward bifurcation in the infection model with concave incidence	76
4.2.2	Backward bifurcation in the infection model with convex incidence	77
4.3	Hopf bifurcation	80
4.3.1	Hopf bifurcation in the infection model with concave incidence	80
4.3.2	Hopf bifurcation in the infection model with convex incidence	82
4.4	Negative backward bifurcation in an autoimmune disease model	91
4.5	Conclusion	93
4.6	References	94

5	Dynamical Analysis of a 2-dimensional Disease Model with Convex Incidence	99
5.1	Introduction	99
5.2	Dynamics of the 2-D disease model	101
5.2.1	Well-posedness of solutions	101
5.2.2	Construction of generic trapping region	102
5.2.3	Dynamical behavior of (5.1) when $B < D$	104
5.2.3.1	$A = 0.364, C = 0.823, D = 0.057, B = 0.054$	108
5.2.3.2	$A = 0.464, C = 0.523, D = 0.057, B = 0.054$	109
5.2.3.3	$A = 0.264, C = 0.823, D = 0.087, B = 0.054$	111
5.2.3.4	$A = 5.200, C = 0.223, D = 0.087, B = 0.054$	111
5.2.3.5	$A = 0.26302225, C = 0.823, D = 0.087, B = 0.054 (H_1 = 0)$	112
5.2.3.6	$A = 0.260, C = 0.823, D = 0.087, B = 0.054 (H_1 < 0)$	112
5.2.4	Dynamical behavior of (5.1) when $B > D$	113
5.2.5	Dynamical behavior of (5.1) when $B = D$	116
5.3	Hopf and generalized Hopf bifurcations	117
5.3.1	Hopf bifurcation	118
5.3.2	Generalized Hopf bifurcation	121
5.4	Bogdanov-Takens bifurcation	125
5.4.1	Case $B = 0.054, D = 0.057$	125
5.4.2	Case $B = 0.054, D = 0.087$	132
5.4.3	A new mechanism for generating blips	135
5.5	Conclusion and discussion	138
5.6	References	138
6	Conclusion	141
	Curriculum Vitae	143

List of Figures

1.1	Illustration of HIV viral blips	2
2.1	Bifurcation diagram for the 4-dimensional HIV antioxidant-therapy model (2.1)	15
2.2	The saddle-node bifurcation on the center manifold	16
2.3	Simulated limit cycles of system (2.1)	19
2.4	Schematic diagram for explaining the occurrence of blips.	20
2.5	Dynamics and bifurcation of system (2.14)	23
2.6	Parameter space for identifying the oscillating region	24
2.7	Bifurcation diagrams with three values of parameter C , and the corresponding simulated time histories	26
2.8	Bifurcation diagrams with three values of parameter A , and the corresponding simulated time histories	27
2.9	(a) Bifurcation diagram of system (2.19), and the simulated viral blips	28
2.10	(a) Bifurcation diagram of system (2.21), and the simulated viral blips	31
2.11	Bifurcation diagram and simulated viral blips for system (2.22)	35
2.12	Simulated viral blips of system (2.1) with varying amplitude and frequency	37
3.1	Bifurcation diagram for model (3.6)	52
3.2	Simulated time history for system (3.6)	56
3.3	Comparison between the simulated time history and analytical prediction for system (3.6)	56
3.4	Simulated time history for system (3.6)	57
3.5	Simulated time history for system (3.23)	59
3.6	Bifurcation diagrams for the scaled system (3.31)	64
3.7	Numerical simulations for equation (3.31) with nine pairs of parameter values: (\mathcal{E}, C)	65
3.8	Numerical simulations for equation (3.31) with nine pairs of parameter values: $(\mathcal{E}, \mathcal{D})$	66
4.1	Graphs of the incidence function f_3 in system (4.6, 4.7) and f_4 in system (4.8)	78
4.2	Graphs of the incidence functions $f_7(\tilde{X}, Y)$ and $f_7(Y)$	79
4.3	Bifurcation diagrams and simulations associated with the five cases given in Table 4.1	83
4.4	Dynamical behaviors of system (4.4) corresponding to eight cases listed in Table 4.2 and 4.4	90
4.5	Dynamics of system (4.22): bifurcation diagram and simulated time history	93

5.1	Plot of the three curves $H_1 = 0$, $H_2 = 0$ and $H_3 = 0$, on the A - C plane for $B = 0.054$, $D = 0.057$	108
5.2	Plot of the three curves $H_1 = 0$, $H_2 = 0$ and $H_3 = 0$, on the A - C plane for $B = 0.054$, $D = 0.087$	109
5.3	Simulated phase portrait of system (5.1) showing the global stability of E_0	109
5.4	Simulated phase portrait of system (5.1) showing the bistable equilibria E_0 and E_{1-} and an unstable limit cycle	110
5.5	Bifurcation diagram for the bistable equilibrium solutions	110
5.6	Simulated phase portrait of system (5.1) showing the global stability of E_0	111
5.7	Simulated phase portrait of system (5.1) showing the bistable equilibria E_0 and E_{1-}	111
5.8	Simulated phase portrait of system (5.1) showing the global stability of E_0	112
5.9	Plot of two curves $H_2 = 0$ and $H_3 = 0$ on the A - C plane for $B = 0.060$, $D = 0.057$	114
5.10	Simulated blips of system (5.1) for $A = 0.364$, $C = 0.823$, $D = 0.057$, $B = 0.060$	115
5.11	Simulated phase portrait of system (5.1) when $(B, D) = (0.060, 0.057)$	116
5.12	Plot of the three curves $H_1^* = 0$, $H_2^* = 0$ and $H_3^* = 0$, on the A - C plane for $B = D = 0.057$	118
5.13	Simulations of system (5.1) when $B = D = 0.057$, showing stable limit cycles	120
5.14	Simulations of system (5.1) when $B = D = 0.057$	120
5.15	Simulation of two limit cycles for system (5.1) when $B = D = 0.057$, $A = 0.01846287$, $C = 0.11969000$	123
5.16	The BT bifurcation diagram around two critical points	125
5.17	Bifurcation set and phase portraits of system (5.32).	128
5.18	The phase portrait of (5.42) with $\nu_1 = 1$, showing a homoclinic loop Γ	130
5.19	Simulations of system (5.1) when $B = 0.054$, $D = 0.057$, $A = 0.01487968$	132
5.20	Bifurcation set and phase portraits of system (5.46).	134
5.21	Simulations of system (5.1) when $B = 0.054$, $D = 0.087$	135
5.22	Bifurcation diagrams illustrating three Hypotheses	137

List of Tables

2.1	Parameter values used in model (2.1) [39]	13
2.2	Parameter values used in model (2.22) [41].	32
3.1	Parameter definitions and values used in Chapter 3.	71
4.1	Dynamics of system (4.8) for different values of k	82
4.2	Parameter values taken to illustrate various dynamics of system (4.4).	85
4.3	Classification of Hopf bifurcations based on the normal form (4.20).	85
4.4	Classification of Hopf bifurcations appearing in Table 4.2.	86
5.1	Classification of E_{1-} ($H_1 \geq 0$).	107
5.2	Classification of E_{1-} for given parameter values ($D > B = 0.054$).	108

List of Abbreviations, Symbols, and Nomenclature

LIST OF SYMBOLS

HIV – Human Immunodeficiency Virus

HAART – Highly Active Antiretroviral Therapy

ODE – ordinary differential equation

SIRS – Susceptible-Infected-Recovered-Susceptible model

SIR – Susceptible-Infected-Recovered model

SI – Susceptible-Infected model

ROS – reactive oxygen species

CTLs – cytotoxic T lymphocytes

pAPCs – professional antigen presenting cells

IL-2 – interleukin-2

DCs – dendritic cells

T_{Reg} – regulatory T (cells)

nT_{Reg} – natural T_{Reg} (cells)

TCR – T Cell Receptor

R_0 – The basic reproduction number

IC – initial condition

SF – stable focus

UF – unstable focus

SN – stable node

UN – unstable node

DSN – degenerate stable node

DUN – degenerate unstable node

Turning – Turning point

SD – saddle-node bifurcation

Transcritical – transcritical bifurction

HF, Hopf – Hopf bifurcation

subH – subcritical Hopf bifurcation

supH – supercritical Hopf bifurcation

gH – generalized Hopf bifurcation

BT – Bogdanov-Takens bifurcation

Homo – Homoclinic bifurcation curve

Tr – trace

Det – determinant

μ, ρ, θ – the perturbation and the amplitude and phase of motion in the normal form associated with Hopf bifurcation

Chapter 2

E_0 – the uninfected equilibrium

E_1 – the infected equilibrium

Γ – a positively invariant set and attracts all non-negative solutions of (2.1)

$J_{0,1}$ – the Jacobian matrix at $E_{0,1}$

$P_{0,1}$ – the characteristic polynomial of $E_{0,1}$

Λ_s – the Jordan canonical form at saddle-node bifurcation point

T_s – the transformation matrix for Λ_s to diagonal form

Δ_i – the i th Hurwitz arrangement

Model (2.1), (2.14) and (2.20)

x, y – the population densities of uninfected and infected $CD4^+$ T cells

r, a – the density of ROS and antioxidants

$d_{x,y}$ – the death rates for uninfected and infected $CD4^+$ T cells

$d_{r,a}$ – the decay rates for ROS and antioxidants

λ_x – the production rate of $CD4^+$ T cells

ϵ – the effectiveness of drug therapy

λ_r – the natural generation rate of ROS

k – the generation rate of ROS from infected $CD4^+$ T cells

λ_a – antioxidants intaken through diet

α – antioxidants intaken through supplementation

m, p – the decay rates induced by the reaction between ROS and antioxidants, respectively

b_0 – the infection rate in the ROS-absent case

b_{\max} – the maximum infection rate

r_{half} – the ROS concentration at half maximum infection rate

$\beta(r)$ – the infectivity function

$a = b_{\max} - b_0$

$b = b_0$

$c = r_{\text{half}}$

Model (2.21) and (5.1)

X, Y – rescaled state variables x and y in model 2.20

τ – rescaled time variable t

A, B, C, D – rescaled parameter values in model 2.20

Model (2.22)

x, y, z, u, v – the population densities of uninfected target cells, infected target cells, CTLs, antibodies and virions

d, a, b, η, q – the death or clearance rate of uninfected and infected cells, CTLs, antibody and viruses.

β – constant infectivity

p – the clearance rate of infected cells killed by CTLs

c – the natural proliferation rate of CTLs

h – the proliferation rate of CTLs from memory T cells

ξ – the antibody growth rate

k – the binding rate of one antibody with one antigen

e – the virus releasing rate from infected cells

γ – the rate of viruses absorbed by uninfected cells

Chapter 3

A, R_n, R_d, E, G the population of mature pAPCs, active nT_{Reg} cells, terminally differentiated T_{Reg} cells, active auto-reactive effector T cells, and the particular self-antigen of interest

\tilde{v} – per capita rate at which free antigen (G) is taken up by immature pAPCs

f – proportion of antigen molecules that, upon uptake, lead to maturation of the pAPC to enter population A

π_1 – rate (per A , per E) at which active nT_{Reg} cells are generated from the pool of ‘naive’ T_{Reg} cells, due to encounter with mature pAPCs (A) and influence of IL-2 from specific effector T cells

π_3 – rate (per A , per E) at which active nT_{Reg} cells are generated from the pool of ‘naive’ T_{Reg} cells, due to encounter with mature pAPCs (A) and influence of IL-2 from specific effector T cells

β – rate (per A) at which active nT_{Reg} cells are generated from the resting pool, due to encounter with mature pAPCs (A) and influence of IL-2 from other sources

λ_E – rate (per A) at which effector T cells (E) are generated from the resting pool, due to encounter with mature pAPCs (A)

γ – rate (per E) at which self antigen (G) is released due to the actions of effector T cells (E)

$\sigma_{1,3}$ – rate (per capita, R_n or R_d) at which mature pAPCs (A) and effective T cells are effectively eliminated due to suppression by specific active nT_{Reg} cells (R_n) or terminal T_{Reg} cells (R_d)

b_1 – rate (per capita) at which mature pAPCs (A) are effectively eliminated due to suppression by T_{Reg} cells of other specificities or by therapy

b_3 – rate (per capita) at which effective T cells (E) are effectively eliminated due to suppression by T_{Reg} cells of other specificities or by therapy

μ_A – per capita death rate of mature pAPCs

μ_E – per capita death rate of effector T cells (E)

μ_G – per capita rate at which free antigen (G) is cleared, for example due to degradation

μ_n – per capita death rate of active nT_{Reg} cells (R_n)

μ_d – per capita death rate of terminal T_{Reg} cells (R_d)

ξ – proportion of activated nT_{Reg} cells

α – rate (per E) at which immature pAPCs become mature

d – the ratio of suppress effectiveness of nT_{Reg} cells to terminal T_{Reg} cells

c – the fold of matured nT_{Reg} cells expansion and proliferation to terminal T_{Reg} cells

Chapter 4

S, I, R – the population size of susceptible, infective, and recovered individuals

μ, δ, α – the birth/death rate, the recovery rate and the loss of immunity rate in model (4.1)

N – the total population size

Λ – the constant recruitment rate of susceptibles

d, γ, ϵ – the rates of natural death, recovery, and the disease-induced mortality

β the infection rate

k the inhibition effect

α the maximal medical resources per unit time in model (4.8)

ω the half-saturation constant

Chapter 1

Introduction

Recurrent disease, such as several episodes of ear infections or bacterial sinusitis in one year, can be very common and disagreeable. Recurrence can also pose serious health issues, and even fatality [16], and it is often associated with chronic diseases for which there is no known cure, such as human immunodeficiency virus (HIV) infection [3, 13], or lupus [10]. The pattern of recurrent disease is an alternation between acute relapse and long remission [5, 16, 6, 8, 10]. In HIV infection for example, “viral blips” are commonly measured in patients under highly active antiretroviral therapy (HAART), whose blood viral load is controlled for long periods at an undetectable level, but is still punctuated periodically by short episodes of high viral reproduction [14], as shown in Figure 1.1. Although the etiology is not well understood, HIV infected patients chronically suffer from these episodes of acute viral relapse [7]. In addition, important issues in recurrent disease, such as medical treatment and patient management, cry out for new insight. In this study, we apply approaches characteristic of mathematical biology to better understand the intrinsic mechanisms driving recurrent diseases.

1.1 Mathematical models for studying recurrence

Mathematical models using differential equations track changes in biological systems over time, and provide new research tools to investigate and explain clinical and laboratory observations [1, 11, 12]. By translating verbal mechanisms into scientific prediction, mathematical models play a fast-growing and well-recognized role in understanding, predicting, and controlling diseases [11]. In this study, based on traditional epidemic models at the population level, we develop and analyse *in-host* models at the cell-to-cell level to describe the interaction between pathogenic agents and cells.

1.1.1 Immunological models

The body’s defence against foreign pathogen invasion is the immune system. Immunology is the study of the immune system, including its function and possible malfunctions, such as autoimmune disease, hypersensitivities, immune deficiency, and transplant rejection. The immune system is built mainly at the cellular level. Mathematical models in immunology therefore attempt to describe the dynamical world of cells and molecules inside body.

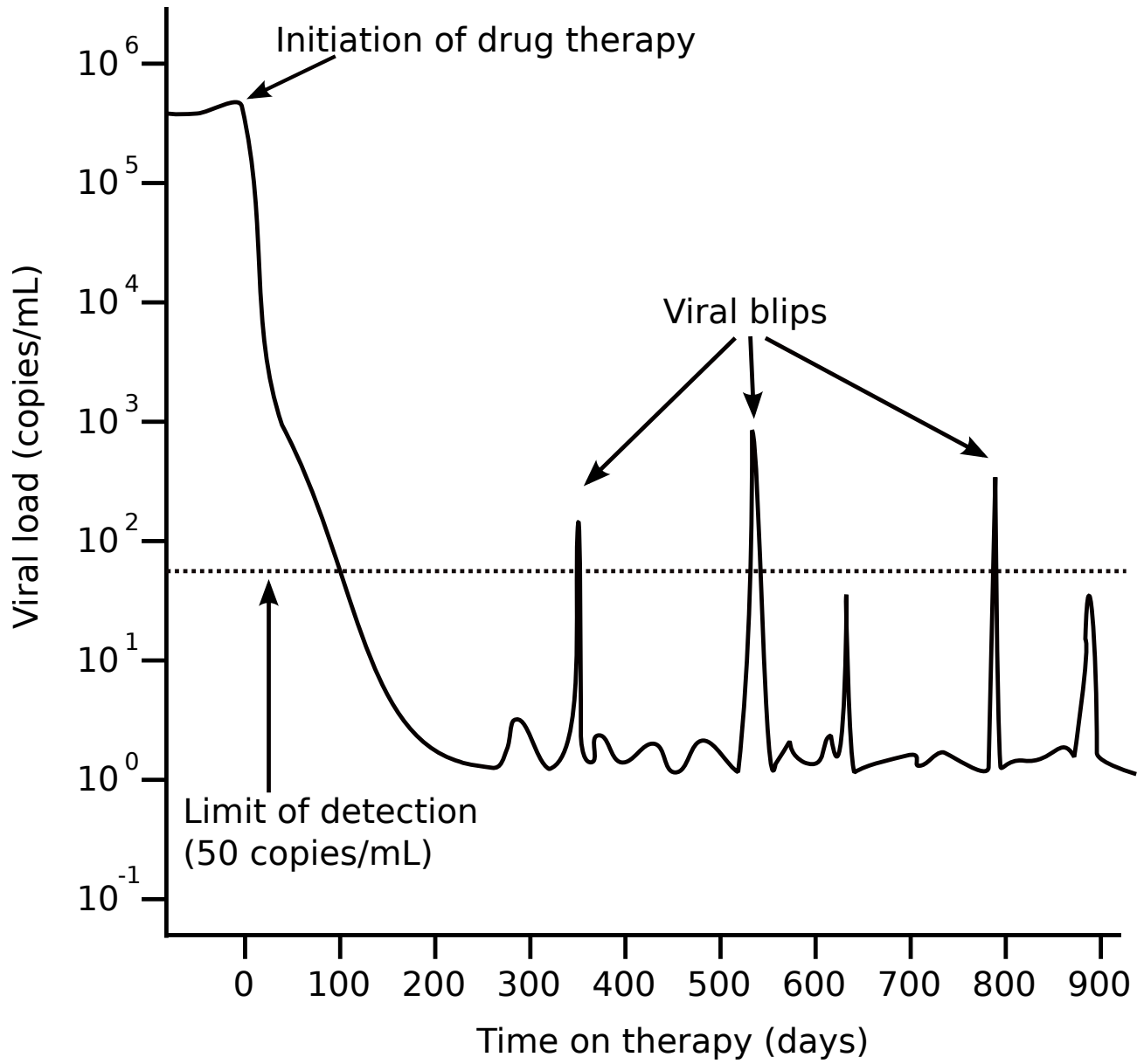


Figure 1.1: Illustration of HIV viral blips

Mathematical models in immunology, typically systems of ordinary differential equations, are well recognized and widely used to describe immune processes, understand the underlying dynamical processes, reveal intrinsic mechanisms, and predict the fate of the disease. In addition, mathematical models can provide a persuasive way to verify verbal assumptions in immunology. Additionally, model simplification can help to identify and emphasize the determining factors in disease. Simplifications, such as quasi-steady state assumptions, are a well-recognized way to reduce model dimension, while retaining the model's main properties. Although cellular processes are key to immune function, the immune response also incorporates processes of a chemical nature, such as the antigen-antibody interaction and enzyme-catalysed reactions, for example the cytokine molecule IL-2 signaling process. Mathematical modeling can incorporate these biochemical factors into immunological models. For example, the influence of reactive oxygen species on HIV infection rate can be modeled according to Michaelis-Menten kinetics, and gives rise to an increasing, saturating HIV infectivity function in Chapter 2. This function further determines the simplest 2-dimensional HIV infection model which shows recurrent behavior, providing a new mechanism for HIV viral blips and a fresh insight into the elusive world of HIV infection.

1.1.2 Infection models

A basic epidemic SIR model divides the population into susceptible, infected and recovered groups, and denotes the numbers in each group as S , I , and R , respectively. An SIR model with no disease-related death is written as

$$\frac{dS}{dt} = bN - \beta IS - dS, \quad \frac{dI}{dt} = \beta IS - \gamma I - dI, \quad \frac{dR}{dt} = \gamma I - dR, \quad (1.1)$$

where the total population size is N , the birth rate is b , the common death rate for each group is d , the infectivity is β , and the recovery rate is γ [1]. The recovery group can be reduced under the assumption that the total population size is constant. Subsequently, the 3-dimensional SIR model is reduced to a 2-dimensional SI model. Similarly, an in-host model tracks the transmission of an infectious agent, for example a virus, from cell to cell within the body of a single infected individual [12]. The basic model in this case also has three variables: uninfected cells, X ; infected cells, Y , and free virus particles, V . These variables can either denote the total population size in an infected individual or population density in blood or tissue [12]. Compared with the host cell, the infectious agent, such as a virus and bacterium, is characterized by a short lifespan and extremely high reproduction rate. Due to these high production and clearance (birth and death) rates, virus particles can be assumed to be in a quasi-steady state with the population of infected cells, and eliminated from the system [12]. This step results in a 2-dimensional within-host model, which is proved to be equivalent to the 2-dimensional epidemic model in Chapter 4.

The spread of disease is a key point in modelling, and the rate at which new individuals are added to the population of infectives is referred to as the incidence rate [4]. The functional form of this term varies according to the properties of the disease and the hypothesis considered. Based on the law of mass action, the spread of disease is usually written as the infection force, multiplied the number of susceptibles [2]. The infection force describes the transition rate from the susceptibles to infectives, and is usually a function of the number of infectives. The most

common form of the infection force is linear, that is βI , where β is the per capita contact rate, with the assumption of homogeneous mixing of both susceptible and infective populations. By considering heterogeneous mixing and saturation effects due to fewer susceptibles being available with the growth of the infective population, the infection force can be modified to be an increasing and concave function in terms of the number of infectives [9]. In contrast, the infection force may take the form of an increasing and convex function, if cooperative effects are considered, for example if infected cells make other host cells more vulnerable to infection [15, 17, 18].

1.2 Mathematical theories and methodologies used to study recurrence in biological models

Biological models are characterized by changes, and differential equations are laws that rule changes. For biological models described by differential equations, the description of the dynamical behavior of the differential equations is the description of the time evolution of the biological system. The differential equations are also referred to as a dynamical system. The solution determines how the dynamical system develops in time. For most differential equations, describing real-world problems, their solution formula or analytic solutions are difficult or even impossible to obtain. Therefore, we apply dynamical systems theory, in particular, qualitative methodologies including stability and bifurcation analysis to extract important information and show the fundamental, long-term qualitative behavior of the system.

In this study, we concentrate on continuous differential equation models, which is a reasonable approximation to describe the continuous overlap of cells' and infectious agents' generations. Nonlinear systems theory and methodologies are applied to investigate the complexity of the biological systems. To reveal intrinsic mechanisms underlying complex phenomena in disease models, we use simple deterministic models to predict the long-term behavior of the disease. In particular, asymptotic behavior is examined, such as local and global stability of equilibrium solutions, and bifurcations from the equilibrium solutions, leading to Hopf bifurcation and even more complex bifurcation such as homoclinic orbits.

1.2.1 Stability analysis for equilibrium solutions

Mathematical analysis of population dynamics usually first proves well-posedness of the solutions, that is, the solutions of the system should be positive and bounded due to their biological meaning. Equilibrium solutions expose the steady-state features of the system, which can be either stable or unstable depending upon whether the solution trajectories of the system converge towards the equilibrium or diverge away from it. The stability of an equilibrium solution can be characterized, in the sense of Lyapunov stability theory, as local or global depending on whether the final state depends on the initial condition. In other words, global asymptotic stability means that any solution trajectory of the system will return to the equilibrium from any initial point in the state space; while for local asymptotic stability this only occurs for initial points near the equilibrium solution. The path of convergence may be either direct, i.e. without oscillating, or with oscillating behavior. Besides equilibrium solutions, many biologi-

cal systems may exhibit complex behavior such as limit cycles, for which the trajectories may approach or diverge from a periodic solution. We can also define the stability of limit cycles, as stable or unstable, depending on whether they attract or repel nearby trajectories. Local stability of the equilibrium solution can be obtained by examining the corresponding characteristic equation, and usually (especially for higher-dimensional dynamical systems) applying the Routh-Hurwitz stability criterion. This process often involves solving multivariate polynomials. The Lyapunov function method (or Lyapunov direct method) is usually applied to prove the well-posedness of the solutions and the global stability of equilibrium solutions. For limit cycles, however, finding their stability is more involved, and requires more sophisticated mathematical methods to be discussed next.

1.2.2 Bifurcation analysis

Bifurcation theory is fundamental for the qualitative study of dynamical systems, and can be used to reveal complex dynamical behaviors of the biological systems under study, such as bistability, recurrence, and regular oscillation. Characterized by a controllable parameter, called the bifurcation parameter, bifurcation occurs at a critical value of this parameter where the properties of equilibria change significantly. These qualitative changes can be illustrated in a bifurcation diagram. Bifurcations can be divided into two principle classes: local bifurcations and global bifurcations. Local bifurcations occur when the local stability of an equilibrium changes, leading to the birth of another equilibrium solution or a limit cycle, as the bifurcation parameter passes through a critical value. Therefore, the characteristic equation and Routh-Hurwitz stability criterion can be applied to study local bifurcations. More precisely, the local bifurcations can be classified as saddle-node, transcritical, and pitch-fork bifurcations, which characterize the “jump” from one equilibrium solution to another equilibrium solution. In this thesis, for the convenience of use in Applied Science and Engineering Society, we call the saddle-node bifurcation point, the “turning point”. Hopf bifurcation, which characterizes the “birth of motion” from an equilibrium solution to periodic motion. Global bifurcations, on the other hand, occur when periodic orbits collide with each other, or with equilibria, and cause changes in the topology of the trajectories out of a small neighborhood. The terminology “unfolding” determines the codimension of a bifurcation, that is, how many bifurcation parameters are required to characterize the fundamental dynamical behavior of the system. In this study, we mainly focus on local bifurcations including saddle-node, transcritical and Hopf bifurcations, which are all codimension-one bifurcations. We will also investigate the well-studied codimension-two bifurcation: Bogdanov-Takens bifurcation, since it can lead to the global bifurcation: homoclinic bifurcation. We will pay more attention to Hopf bifurcation and homoclinic bifurcation, since they provide two mechanisms for generating recurrence.

1.2.2.1 Hopf bifurcation

Hopf bifurcation is perhaps the most typical way to generate limit cycles and recurrent phenomenon. It occurs when the Jacobian matrix of a dynamical system, evaluated at an equilibrium, contains a simple pair of purely imaginary eigenvalues, giving rise to a nonhyperbolic critical point: the Hopf bifurcation point. The stability of the limit cycle is determined by the behavior of the solution trajectories of the system on the center manifold near the Hopf bifurca-

tion point. Center manifold theory provides a means for systematically reducing the dimension of the state space, resulting in a center manifold with reduced dimension. Further simplifying the differential equations, describing the dynamical behavior on the reduced center manifold, by additional coordinate transformations yields the normal form for the Hopf bifurcation. The qualitative picture of the flows of Hopf bifurcation can be revealed by analyzing the stability and bifurcations based on the normal form, and the stability of the bifurcating limit cycle is determined by the coefficients of the normal form. Hopf bifurcations can be classified as supercritical or subcritical, indicating whether the bifurcating limit cycle is stable or unstable. Center manifold theory and normal form theory are the two most powerful and useful mathematical tools in the study of stability and bifurcations for nonlinear dynamical systems.

1.2.2.2 Bogdanov-Takens bifurcation and homoclinic orbits

A homoclinic or saddle-connection bifurcation occurs when a limit cycle collides with a saddle point. It is a global bifurcation and may arise from Bogdanov-Takens bifurcation. Bogdanov-Takens bifurcation is characterized by a double-zero eigenvalue of the linearized system around an equilibrium solution. The existence of homoclinic bifurcation, associated with Bogdanov-Takens bifurcation, may provide a global mechanism for the existence of limit cycles and recurrence. By applying a rescaling or blow-up approach on the normal form obtained associated with Bogdanov-Takens bifurcation, we may obtain a Hamiltonian system and thus properly define a Melnikov function used to determine the homoclinic bifurcation curve, leading to bifurcation of homoclinic orbits. Further, this approach can be employed to identify the parameter region where limit cycles exist between the Hopf bifurcation curve and the homoclinic bifurcation curve.

1.3 Thesis contribution and structure

In this thesis, we study recurrent phenomena in infectious diseases and autoimmune diseases, which are described by deterministic, ordinary differential equations. Local and global mechanisms generating recurrence are provided in explicit mathematical formulae, associated with Hopf bifurcation, Bogdanov-Takens bifurcation and homoclinic bifurcation. Biologically, we find that recurrent behavior can be an intrinsic property in disease dynamics. For infectious disease, an increasing and saturating infectivity function can be the determining component for recurrence. While, for autoimmune disease, recurrence can be attributed to the newly discovered terminally differentiated regulatory T cells. From the viewpoint of modeling, we believe that the investigation of the relation between backward bifurcation and Hopf bifurcation reveals an important finding: a convex incidence function is the key player in determining the bistable, recurrent, and regular oscillating behaviors for a simple 2-dimensional infection model.

In Chapter 2, the dynamics of HIV viral blips are studied by investigating an established 4-dimensional HIV antioxidant therapy model. A new blips-generating mechanism is proposed, that is, infection makes the host more vulnerable to be infected, and is modeled by an increasing, saturating infectivity function. Four conditions are proposed for proving the existence of recurrence in deterministic in-host models.

Chapter 3 is devoted to considering recurrent behavior in an autoimmune model. By introducing a newly discovered regulatory T cell subtype, the autoimmune disease model can exhibit Hopf bifurcation and further generate recurrent behavior.

In Chapter 4, the relation between backward bifurcation and Hopf bifurcation is examined for exploring recurrence, by investigating the infectious disease model established in Chapter 2 as well as the autoimmune model studied in Chapter 3. We identify the parameter region where bistability, recurrence, and regular oscillation can occur.

Chapter 5 provides a further study on the simplest 2-dimensional HIV model (established in Chapter 2) to generate recurrence. More bifurcation parameters are involved in the study to demonstrate complex dynamical behavior. A new mechanism for generating recurrence is obtained from Bogdanov-Takens bifurcation and homoclinic bifurcation.

Finally, the conclusion of the thesis and discussion of future work are given in Chapter 6.

1.4 References

- [1] R. M. Anderson and R. M. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, 1992.
- [2] N. F. Britton. *Essential Mathematical Biology*. Springer Undergraduate Mathematics Series. Springer London, 2003.
- [3] G. Dornadula, H. Zhang, B. VanUitert, J. Stern, L. Livornese Jr., M. J. Ingerman, J. Witek, R. J. Kedanis, J. Natkin, J. DeSimone, and R. J. Pomerantz. Residual HIV-1 RNA in blood plasma of patients taking suppressive highly active antiretroviral therapy. *Journal of the American Medical Association*, 282(17):1627–1632, 1999.
- [4] M. Feinleib. A dictionary of epidemiology, fourth edition - edited by John M. Last, Robert A. Spasoff, and Susan S. Harris. *American Journal of Epidemiology*, 154(1):93–94, 2001.
- [5] I. C. H. Fung, M. Gambhir, A. van Sighem, F. de Wolf, and G. P. Garnett. The clinical interpretation of viral blips in HIV patients receiving antiviral treatment: Are we ready to infer poor adherence? *Journal of Acquired Immune Deficiency Syndromes*, 60(1):5–11, 2012.
- [6] H. J. Girschick, C. Zimmer, G. Klaus, K. Darge, A. Dick, and H. Morbach. Chronic recurrent multifocal osteomyelitis: what is it and how should it be treated? *Nature Clinical Practice Rheumatology*, 3, 2007.
- [7] J. T. Grennan, M. R. Loutfy, D. Su, P. R. Harrigan, C. Cooper, M. Klein, N. Machouf, J. S. G. Montaner, S. Rourke, C. Tsoukas, B. Hogg, J. Raboud, and the CANOC Collaboration. Magnitude of virologic blips is associated with a higher risk for virologic rebound in HIV-infected individuals: A recurrent events analysis. *Journal of Infectious Diseases*, 205(8):1230–1238, 2012.
- [8] R. S. Iyer, M. M. Thapa, and F. S. Chew. Chronic recurrent multifocal osteomyelitis: review. *American Journal of Roentgenology*, 196(6 Suppl):S87S91, 2011.

- [9] A. Korobeinikov and P. K. Maini. Non-linear incidence and stability of infectious disease models. *Mathematical Medicine and Biology*, 22(2):113–128, 2005.
- [10] D. D. Munro. Recurrent subacute discoid lupus erythematosus. *Proceedings of the Royal Society of Medicine*, 56(2):78–79, 1963.
- [11] J. D. Murray. *Mathematical Biology: I. An Introduction*. Interdisciplinary Applied Mathematics. Springer, 2002.
- [12] M. A. Nowak and R. M. May. *Virus Dynamics*. Oxford University Press, New York, 2000.
- [13] S. Palmer, F. Maldarelli, A. Wiegand, B. Bernstein, G. J. Hanna, S. C. Brun, D. J. Kempf, J. W. Mellors, J. M. Coffin, and M. S. King. Low-level viremia persists for at least 7 years in patients on suppressive antiretroviral therapy. *Proceedings of the National Academy of Sciences*, 105(10):3879–3884, 2008.
- [14] L. Rong and A. S. Perelson. Modeling latently infected cell activation: viral and latent reservoir persistence, and viral blips in HIV-infected patients on potent therapy. *PLoS Computational Biology*, 5(10):e1000533, 2009.
- [15] R. D. van Gaalen and L. M. Wahl. Reconciling conflicting clinical studies of antioxidant supplementation as HIV therapy: a mathematical approach. *BMC Public Health*, 9(Suppl. 1):1–18, 2009.
- [16] N. Velez de Mendizabal, J. Carneiro, R. Sole, J. Goni, J. Bragard, I. Martinez-Forero, S. Martinez-Pasamar, J. Sepulcre, J. Torrealdea, F. Bagnato, J. Garcia-Ojalvo, and P. Villoslada. Modeling the effector - regulatory T cell cross-regulation reveals the intrinsic character of relapses in Multiple Sclerosis. *BMC Systems Biology*, 5(1):114, 2011.
- [17] W. Zhang, L. Wahl, and P. Yu. Conditions for transient viremia in deterministic in-host models: Viral blips need no exogenous trigger. *SIAM Journal on Applied Mathematics*, 73(2):853–881, 2013.
- [18] W. Zhang, L. Wahl, and P. Yu. Viral blips may not need a trigger: How transient viremia can arise in deterministic in-host models. *SIAM Review*, 56(1):127–155, 2014.

Chapter 2

Conditions for Transient Viremia in Deterministic In-host Models: Viral Blips Need no Exogenous Trigger

2.1 Introduction

Viruses are infectious intracellular parasites: they can reproduce only inside the living cells of host organisms, and must spread from host to host for continued existence. Animal viruses tend to exhibit either an acute or persistent mode of host infection to ensure this continuity [40]. An acute viral infection is characterised by a relatively short period of symptoms, and resolution within days or weeks. It usually triggers the host immune response to clear the infection, and a memory response can then prevent the same virus from infecting the same host. Pathogens such as influenza virus and rhinovirus typically cause acute viral infections. In contrast, persistent infections [2] establish long-lasting infections in which the virus is not fully eliminated but remains in infected cells. Persistent infections involve both silent and productive infection stages without rapid killing or excessive damage to infected cells. Latent infection is a type of persistent infection.

In latent infection, no clinical signs nor detectable infectious cells can be observed during the silent or quiescent stage of low-level viral replication. However, the virus has not been completely cleared, and recurrent episodes of rapid viral production and release can periodically punctuate relatively long periods in the silent stage. These episodes of recurrent infection are a clinical phenomenon observed in many latent infections [41]. Recurrent infection can also occur in the context of drug treatment for persistent infections. Human immunodeficiency virus (HIV), for instance, can be suppressed by highly active antiretroviral therapy (HAART) to below the limit of detection for months or years [4, 8], nonetheless supersensitive assays can still detect low levels of viremia during this stage [8, 31, 30]. Moreover, these long periods of relative quiescence are typically interrupted by unexplained intermittent episodes of viremia above the detectable limit, termed viral blips [35, 34]. Although these blips have been the focus of much recent research [12, 17, 14, 5], their etiology is still not well-understood [17, 34].

To date, many possible explanations for viral blips during HIV infection have been explored mathematically. An early model of the long-term pathogenesis of HIV [11] incorporates the

activation of T cells in response to antigen, as suggested earlier by [9]. In [11], both HIV and non-HIV antigen exposure are considered in a coupled deterministic-stochastic model. The probability of antigenic exposure evolves continuously in time, and Poisson-distributed exposure events are generated, by simulation, at the appropriate probabilities. This approach captures a number of features of long-term HIV dynamics, including episodic 'bursts' of residual viral replication. Further work [10] considers the number of distinct antigens which activate the $CD4^+$ T cell pool as a random variable, coupled to an ordinary differential equation (ODE) model. Stochastic changes to this number drive fluctuation in the basic reproductive number and viral load. This model is also able to capture the episodic burst-like nature of residual HIV viral replication during long-term infection.

More recent models are based on the recurrent activation of latently-infected lymphocytes, a class of T cells introduced in immunological models by Perelson *et al.* [32] and Rong *et al.* [33], in order to explain the slower second-phase decay of plasma viremia. By introducing antigen concentrations as an explicit variable, Jones and Perelson [23] developed a system of ODEs which exhibits viral blips. The model describes programmed proliferation and contraction of the $CD8^+$ T cell population, and exhibits low viral loads under HAART as expected. Opportunistic or concurrent infection, modelled as an initial concentration of antigen, activates the immune system and is shown by numerical simulation to elicit a transient viral blip. The same authors further showed that occasional intercurrent infections can generate viral blips by the activation of target cells or latently-infected cells, predicting a power law relationship between blip amplitude and viral load [24].

In further work, by considering the asymmetric division of latently-infected cells, Rong and Perelson [34] developed a 4-dimensional ODE model based on the basic model of latent cell activation [32]. This new model not only generated viral blips but also maintained a stable latent reservoir in patients on HAART. In this model, latently-infected cells can divide to produce latently-infected daughter cells, or differentiate into activated, productively-infected cells, depending on antigen concentrations. In a further 5-dimensional ODE model [35], these two types of daughter cells were distinguished as dependent variables, and a contraction phase was added to the activated daughter cells. Numerical simulation showed that both cases gave rise to viral blips and a stable latent reservoir, which were generated from the activated and the latently-infected daughter cells, respectively. In both papers [34, 35], the antigenic stimulation of latently infected cells was modeled as an "on-off" forcing function, and viral blips were initiated during brief pulses in which this activation function was "on".

Most recently, a stochastic model developed by Conway and Coombs [5] presented another possible treatment of latent cell activation. In this model [5], the authors derive the probability generating function for a multi-type branching process describing the populations of productively and latently infected cells, and free virus. A numerical approach is then used to estimate the probability distribution for viral load, which is then used to predict blip amplitudes and frequencies; blip durations are studied by simulation. The authors are able to conclude that with effective drug treatment and perfect adherence to drug therapy, viral blips can *not* be explained by stochastic activation of latently-infected cells, and other factors such as transient secondary infections, or imperfect adherence, must be involved.

In order to elicit transient episodes of high viral replication, the models described above either incorporate transient immune stimulation, for example as a forcing function, or stochastic approaches. In contrast, recent studies have shown that simple deterministic systems can ex-

hibit viral blips. Based on the close relation between recurrent infections and antibody (B-cell) immunodeficiency, Yao *et al.* [41] investigated a 5-dimensional ODE model which included antibody concentrations as an explicit variable, and exhibited transient periods of high viral replication. By numerical simulation at specific, meaningful parameter values, the authors explored factors affecting the interval between recurrent episodes, and their severity. Later, an even simpler 4-dimensional antioxidant-therapy model [39] was explored for HIV, and was similarly used to simulate viral blips with appropriate parameter values. These examples indicate that deterministic systems can produce blips as part of the natural, rich behaviour of the non-linear system. Although to date numerical simulation has been invaluable in describing and delineating the behaviour of these models, there is yet very little analytical work exploring the mathematical underpinnings of recurrent infection. It should be noted that data from clinical studies indicates that HIV viral blips appear to be random biological events, with varying magnitude, frequency and duration. This suggests that stochastic tractable, and their analysis may reveal a global picture or key underlying characteristics of the system. Moreover, non-linear deterministic systems can indeed exhibit varying amplitudes and frequencies of motion, particularly when the underlying parameters are functions of time. We shall return to a discussion of this point in the last section of the paper.

In this paper, we take advantage of dynamical systems theory to reinvestigate deterministic in-host infection models that exhibit viral blips. By examining the bifurcation behaviour in parameter spaces “close” to the region where blips occur, we propose an understanding for the features of the dynamical system which underlie this complex model behaviour. We then propose four conditions which, when satisfied, guarantee that an in-host infection model will exhibit long periods of quiescence, punctuated by brief periods of rapid replication: viral blips. Based on these conditions, we develop very simple 2- and 3-dimensional models that produce blips. Further, we apply stability criteria to determine parameter ranges which may yield blips. Most of the models discussed in this paper share a similar infectivity function, describing the rate at which new infected cells are created. In a final section, we examine a related 5-dimensional immunological model and demonstrate that viral blips are possible in this system even when infectivity is constant.

The rest of the paper is organized as follows. In Section 2, the previously proposed 4-dimensional HIV antioxidant-therapy model is reinvestigated analytically. Based on the insights of our bifurcation analysis, conditions for generating viral blips are proposed. In Section 3, we use these conditions to propose a simpler 3-dimensional in-host infection model, and parameter ranges which will exhibit blips in the simpler model are determined. In Section 4, we develop a 2-dimensional model, characterized by an increasing and saturating infectivity function, which can also generate viral blips. Finally, we demonstrate that a 5-dimensional immunological model [41] can exhibit viral blips with constant infectivity.

2.2 A 4-dimensional model which exhibits viral blips

In this section, we reconsider a 4-dimensional HIV antioxidant-supplementation therapy model which was developed and studied numerically in [39]. This model novelly introduced reactive oxygen species (ROS) and antioxidants to an in-host model of HIV infection. In uninfected individuals, ROS play a positive physiological role at moderate levels [16, 25, 7, 20, 18], but

are harmful at high levels [39].

HIV infection may lead to chronic and acute inflammatory diseases, which may cause high levels of ROS [26] as well as lowered antioxidant levels; this phenomenon has been observed clinically and experimentally [26, 15, 22, 36, 38]. In addition, high levels of ROS may cause damage to CD4⁺ T cells, impair the immune response to HIV [37], and exacerbate infected cell apoptosis, releasing more HIV virions. Thus, infected cells produce high levels of ROS, which in turn increase the viral production by infected cells. To control this cycle, antioxidant supplementation (vitamin therapy) has been suggested as a potential complement to HIV therapy [15, 13], with the aim of counteracting and reducing ROS concentrations [16].

The equations of the 4-dimensional model are described by [39]:

$$\begin{aligned}\dot{x} &= \lambda_x - d_x x - (1 - \epsilon)\beta(r)xy, \\ \dot{y} &= (1 - \epsilon)\beta(r)xy - d_y y, \\ \dot{r} &= \lambda_r + ky - mar - d_r r, \\ \dot{a} &= \lambda_a + \alpha - par - d_a a,\end{aligned}\tag{2.1}$$

where x , y , r and a represent respectively the population densities of the uninfected CD4⁺ T cells, infected CD4⁺ T cells, reactive oxygen species (ROS), and antioxidants. The constant λ_x denotes the production rate of CD4⁺ T cells, and $d_x x$ is the death rate. Uninfected cells become infected at rate $(1 - \epsilon)\beta(r)xy$, where ϵ is the effectiveness of drug therapy, and d_y is the per-capita death rate of infected CD4⁺ T cells. ROS are generated naturally at rate λ_r , and by the infected cells at rate ky ; the concentration of ROS decays at rate $d_r r$, and is eliminated by interaction with antioxidants at rate mar . Antioxidants are introduced into the model through natural dietary intake at a constant rate λ_a , and through antioxidant supplementation at rate α , which is treated as a bifurcation parameter. Antioxidants are eliminated from the system by natural decay at rate $d_a a$, and by reacting with the ROS at rate par , where p is much smaller than m .

An important novel feature of this model is that the infectivity $\beta(r)$ is a positive, increasing and saturating function of r (ROS),

$$\beta(r) = b_0 + \frac{r(b_{\max} - b_0)}{r + r_{\text{half}}},\tag{2.2}$$

where b_0 represents the infection rate in the ROS-absent case, while b_{\max} denotes the maximum infection rate, and r_{half} is the ROS concentration at half maximum. It is obvious that $\beta(r) > 0$, and it is also assumed that $0 < \epsilon < 1$. Therefore, all the parameters in equations (2.1) and (2.2) are positive. The experimental values used for studying model (2.1) are given in Table 2.1. Importantly, these parameters were chosen with careful reference to clinical studies, such that the predicted equilibrium densities are clinically reasonable. Also note that the densities of antioxidants and ROS are of order 10^{13} per μL , while cell densities are of the order 10^2 or 10^3 per μL .

In [39], this model was explored numerically to assess the potential of antioxidant therapy as a complement to HIV drug therapy. In that study, regions of oscillatory behaviour, reminiscent of viral blips, were observed. In the following subsections we perform a thorough equilibrium and stability analysis of the model, in order to shed further light on the factors underlying these rich behaviours.

Table 2.1: Parameter values used in model (2.1) [39]

Parameter	Value
λ_x	60.76 cells μL^{-1} day $^{-1}$
d_x	0.0570 day $^{-1}$
d_y	1.0 day $^{-1}$
λ_a	2.74×10^{13} molecules μL^{-1} day $^{-1}$
d_a	0.0347 day $^{-1}$
ε	$\frac{1}{3}$
b_0	2.11×10^{-4} cell $^{-1}$ μL day $^{-1}$
b_{\max}	0.00621 cell $^{-1}$ μL day $^{-1}$
r_{half}	3.57×10^{13} molecules μL^{-1}
d_r	1.66×10^7 day $^{-1}$
λ_r	1.86×10^{21} molecules μL^{-1} day $^{-1}$
k	1.49×10^{19} molecules cell $^{-1}$ day $^{-1}$
m	1.27×10^{-6} molecule $^{-1}$ μL day $^{-1}$
p	5.04×10^{-14} molecule $^{-1}$ μL day $^{-1}$

2.2.1 Well-posedness of the solutions of system (2.1)

By using the method of variation of constants, we can easily obtain the solutions of (2.1) to show that $x(t) > 0$, $y(t) > 0$, $r(t) > 0$, $a(t) > 0$, $\forall t > 0$, if $x(0) > 0$, $y(0) > 0$, $r(0) > 0$, $a(0) > 0$. To consider the boundedness of the solutions, suppose in general we have the differential inequality: $\dot{T} \leq \lambda - dT$ ($\lambda, d > 0$, $T(0) > 0$). Then if $\dot{T} = \lambda - dT$, we have $\dot{T} + dT = \lambda$. Thus, $T(t) = T(0)e^{-\int_0^t d ds} + \int_0^t \lambda e^{-\int_s^t d du} ds = T(0)e^{-dt} + \frac{\lambda}{d}(1 - e^{-dt})$, which implies that $\lim_{t \rightarrow +\infty} \sup T(t) = \frac{\lambda}{d}$. From the first equation of (2.1), we have $\dot{x} \leq \lambda_x - d_x x$, which yields $\lim_{t \rightarrow +\infty} \sup x(t) = \frac{\lambda_x}{d_x}$. It is also easy to see from the first equation of (2.1) that $x(t) > 0$, $\forall t > 0$. Then, by adding the first two equations of (2.1) we obtain $\frac{d[x(t)+y(t)]}{dt} = \lambda_x - d_x x - d_y y \leq \lambda_x - \tilde{d}(x+y)$, where $\tilde{d} = \min(d_x, d_y)$. Hence, $\lim_{t \rightarrow +\infty} \sup(x(t) + y(t)) = \frac{\lambda_x}{\tilde{d}}$. Therefore, for any given $\varepsilon > 0$, there exists $t^* > 0$, such that $x + y \leq \frac{\lambda_x}{\tilde{d}} + \varepsilon$, for all $t \geq t^*$. For the third equation of (2.1), we similarly have $\frac{dr}{dt} \leq (\lambda_r + k\frac{\lambda_x}{d}) - d_r r$, which results in $\lim_{t \rightarrow +\infty} \sup r(t) = \frac{\lambda_r \tilde{d} + k\lambda_x}{d_r \tilde{d}}$. Finally, for the fourth equation of (2.1), we get $\frac{da}{dt} \leq (\lambda_a + \alpha) - d_a a$, and thus $\lim_{t \rightarrow +\infty} \sup a(t) = \frac{\lambda_a + \alpha}{d_a}$. We define Clearly, Γ is a positively invariant set and attracts all non-negative solutions of (2.1).

2.2.2 Equilibrium solutions of (2.1) and their stability

To find the equilibrium solutions of (2.1), simply setting $\dot{x} = \dot{y} = \dot{r} = \dot{a} = 0$ yields two solutions: the uninfected equilibrium solution E_0 , and the infected equilibrium solution E_1 , given respectively by

$$E_0 : (x_{e0}, y_{e0}, r_{e0}, a_{e0}) = \left(\frac{\lambda_x}{d_x}, 0, r_{e0}, \frac{\lambda_r - d_r r_{e0}}{m r_{e0}} \right), \quad (2.3)$$

where the r_{e0} is determined by the equation

$$F_0(r, \alpha) \equiv \alpha + \lambda_a + \frac{1}{m} \left(p d_r r - \frac{d_a \lambda_r}{r} \right) + \frac{d_a d_r - p \lambda_r}{m} = 0; \quad (2.4)$$

and

$$\begin{aligned} E_1 : (x_{e1}, y_{e1}, r_{e1}, a_{e1}), \quad x_{e1} &= \frac{d_y}{(1 - \epsilon)\beta_r(r_{e1})}, \\ y_{e1} &= \frac{\lambda_x - d_x x_{e1}}{(1 - \epsilon)\beta_r(r_{e1})x_{e1}}, \quad a_{e1} = \frac{\lambda_a + \alpha}{d_a + p r_{e1}}, \end{aligned} \quad (2.5)$$

where r_{e1} is a function in the system parameters, particularly α (see the function F_1 in equation (2.8)). Both E_0 and E_1 are expressed in terms of r (r_{e0} or r_{e1}) for convenience.

We first consider the uninfected equilibrium E_0 . The solution of r_{e0} is determined by (2.4), which is a quadratic equation in r . To simplify the analysis, we use r to express the parameter α since (2.4) is linear in α , and α is treated as a bifurcation parameter. Thus, solving $F_0(r, \alpha) = 0$ for α we obtain

$$\alpha_0(r_{e0}) = -\lambda_a - \frac{1}{m} \left(p d_r r_{e0} - \frac{d_a \lambda_r}{r_{e0}} \right) - \frac{d_a d_r - p \lambda_r}{m}. \quad (2.6)$$

To find the stability of the equilibrium solution E_0 , we first evaluate the Jacobian of system (2.1) at E_0 to get $J_0(r_{e0})$, where (2.6) has been used, and then use $\det(\xi I - J_0)$ to obtain the 4th-degree characteristic polynomial, given by $P_0(\xi, r_{e0}) = (\xi + d_x)[\xi^2 + (p r_{e0} + d_a + \frac{\lambda_r}{r_{e0}})\xi + (\frac{d_a \lambda_r}{r_{e0}} + p d_r r_{e0})](\xi + P_{0r})$, where

$$P_{0r} = d_y - \frac{(1 - \epsilon)\lambda_x(b_0 r_{\text{half}} + r_{e0} b_{\text{max}})}{d_x(r_{e0} + r_{\text{half}})}. \quad (2.7)$$

$P_0(\xi, r_{e0})$ contains three factors: the first one is a linear polynomial of ξ and the second one is a quadratic polynomial of ξ , and both are stable polynomials (i.e., their roots (eigenvalues) have negative real part); and thus the stability of E_0 only depends upon the third factor, a linear polynomial of ξ . Therefore, when $P_{0r} > 0$ ($P_{0r} < 0$), the equilibrium solution E_0 is asymptotically stable (unstable).

The graph for the equation $F_0(r, \alpha) = 0$ given in (2.4) is shown as the red line in Figure 2.1(a), which clearly shows a hyperbola. It is seen from this red line that the relation (2.4) also defines a single-valued function r in α , if only the positive (biologically meaningful) value of r is considered, (i.e., the positive branch of the red line in Figure 2.1(a)). More precisely, it can be shown that the biologically meaningful solution must be located on the first quadrant and above, including the top branch of red line (see Figure 2.1(a)), since E_0 has the component $y_{e0} = 0$.

Next, consider the infected equilibrium solution E_1 . The solution for r_{e1} can be similarly obtained by solving the following equation,

$$F_1(r, \alpha) = \lambda_r + \frac{k\lambda_x}{d_y} - \frac{k d_x (r + r_{\text{half}})}{(1 - \epsilon)(b_0 r_{\text{half}} + b_{\text{max}} r)} - \frac{m r (\lambda_a + \alpha)}{p r + d_a} - d_r r = 0, \quad (2.8)$$

which is again a linear function of α , and we can use r_{e1} to express α as

$$\begin{aligned} \alpha_1(r_{e1}) &= -\lambda_a + \frac{\lambda_r (p r_{e1} + d_a)}{m r_{e1}} + \frac{k \lambda_x (p r_{e1} + d_a)}{m r_{e1} d_y} \\ &\quad - \frac{k d_x (r_{e1} + r_{\text{half}})(p r_{e1} + d_a)}{m r_{e1} (1 - \epsilon)(b_0 r_{\text{half}} + b_{\text{max}} r_{e1})} - \frac{(p r_{e1} + d_a) d_r}{m}. \end{aligned} \quad (2.9)$$

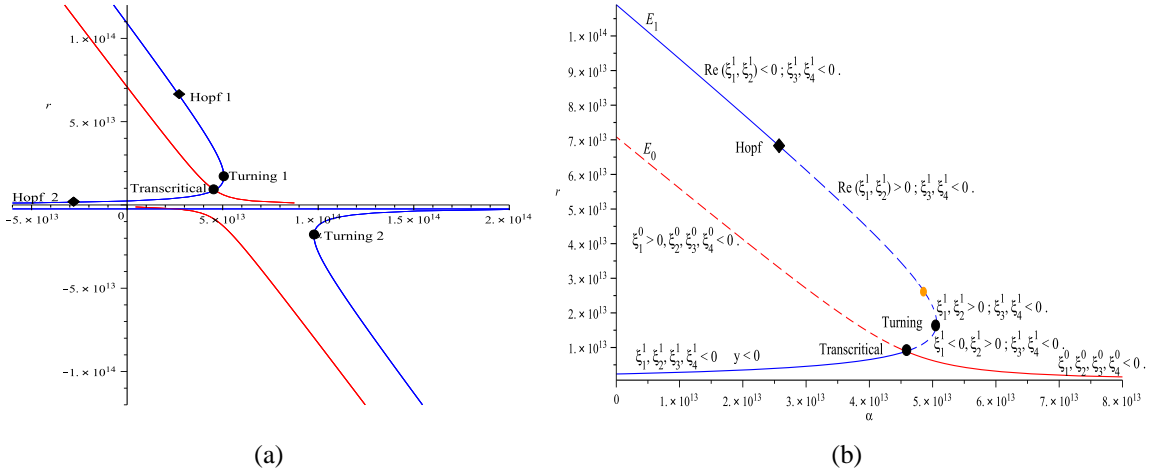


Figure 2.1: (a): Complete bifurcation diagram for the 4-dimensional HIV antioxidant-therapy model (2.1) projected on the r - α plane, with the red and blue lines denoting E_0 and E_1 , respectively; and (b): Bifurcation diagram in (a), restricted in the first quadrant, with the dotted and solid lines indicating unstable and stable, respectively.

The graph of the equations $F_0(r, \alpha) = 0$ given in (2.4) and $F_1(r, \alpha) = 0$ given in (2.8) is shown in Figure 2.1(a). To find the stability of E_1 , in a similar way, we evaluate the Jacobian of (2.1) at E_1 to obtain the 4th-degree characteristic polynomial, $P_1(\xi, r_{e1}) = \xi^4 + a_1(r_{e1})\xi^3 + a_2(r_{e1})\xi^2 + a_3(r_{e1})\xi + a_4(r_{e1})$, where the lengthy expressions for the coefficients $a_1(r_{e1})$, $a_2(r_{e1})$, $a_3(r_{e1})$, and $a_4(r_{e1})$ are omitted here for brevity.

2.2.3 Bifurcation analysis

To understand the conditions underlying oscillatory behaviour and viral blips in this model, we now consider possible bifurcations which may occur from the equilibrium solutions E_0 and E_1 .

2.2.3.1 Transcritical bifurcation

First, for the uninfected equilibrium E_0 , it follows from $P_0(\xi, r_{e0})$ and (2.7) that in general E_0 is stable for $P_{0r} > 0$, and the only possible singularity occurs at the critical point, determined by $P_{0r} = 0$ (see (2.7)). At this point, one eigenvalue of the characteristic polynomial becomes zero (and the other three eigenvalues still have negative real part), leading to a static bifurcation, and E_0 becomes unstable. More precisely, when the parameter values in Table 2.1 are used, the two equilibrium solutions E_0 and E_1 intersect and exchange their stability at the point $(r_t, \alpha_t) \approx (8.89 \times 10^{12}, 4.58 \times 10^{13})$, indicating that a *transcritical bifurcation* occurs at this critical point (see Figure 2.1(b)). Here, the subscript ‘t’ stands for transcritical bifurcation. The value of α_t is obtained by substituting r_t into either $\alpha_0(r_t)$ in (2.6) or $\alpha_1(r_t)$ in (2.9). In fact, $\alpha_0(r_t) = \alpha_1(r_t)$.

As discussed above, the biologically meaningful solutions should be above or on the uninfected equilibrium solution E_0 (the red line shown in Figure 2.1(b)), since solutions below the red line contain the component $y < 0$. It is obvious that there is no Hopf bifurcation from E_0 .

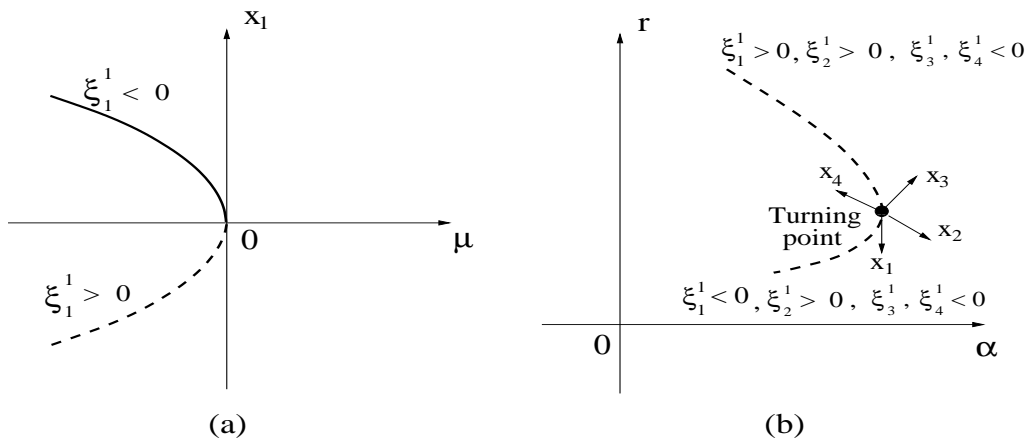


Figure 2.2: The saddle-node bifurcation on the center manifold with the dotted line indicating unstable and the solid line stable: (a) in the transformed x_1 - μ coordinates; and (b) in the original coordinates.

So, the uninfected equilibrium E_0 is asymptotically stable (unstable) when $r < r_t$ ($r > r_t$) or $\alpha > \alpha_t$ ($\alpha < \alpha_t$) (see Figure 2.1(b)).

It should also be noted from Figure 2.1(b) that besides a transcritical bifurcation point, E_1 has a *saddle-node* bifurcation which occurs at the so-called *turning point*. To determine this turning point, using (2.9) and $\frac{d\alpha_1(r)}{dr} = 0$, yields $(r_s, \alpha_s) \approx (1.72 \times 10^{13}, 5.06 \times 10^{13})$, where the subscript 's' denotes saddle-node bifurcation, and $\alpha_s = \alpha_1(r_s)$ by using (2.9). Note that this bifurcation does not change the stability of E_1 , since the characteristic polynomial $P_1(\xi, r_{e1})$ still has an eigenvalue with positive real part when r_{e1} (or α) is varied along E_1 to pass through the turning point (see Figure 2.1(b)).

The saddle-node bifurcation can be seen more clearly if we examine the local dynamics close to the turning point; this analysis will also be useful later for analysing viral blips. At the turning point, the system contains a 1-dimensional center manifold (whose linear part is characterised by the eigenvalue $\xi_1^1 = 0$), a 1-dimensional unstable manifold (whose linear part is characterised by the eigenvalue $\xi_2^1 \approx 0.142$), and a 2-dimensional stable manifold (whose linear part is characterised by the eigenvalues $\xi_3^1 \approx -0.290$ and $\xi_4^1 \approx -1.26 \times 10^8$), as shown in Figure 2.2. It is noted that the eigenvalues ξ_2^1 and ξ_1^1 , which are both positive at the saddle-node point, become a pair of complex conjugates with positive real part at the orange-color point above the saddle-node point (see Figure 2.1(b)), moving towards the Hopf point. So the sub-manifold that is the complement to the centre manifold is still expelling till meeting the Hopf bifurcation point.

In order to find the differential equation described on the center manifold, we first apply the transformation $(x, y, r, a)^T = (x_{e1}, y_{e1}, r_{e1}, a_{e1})^T + T_s(x_1, x_2, x_3, x_4)^T$, where $(x_{e1}, y_{e1}, r_{e1}, a_{e1})$ is the infected equilibrium solution E_1 , and T_s is a constant, non-singular matrix. Under this transformation, the Jacobian of system (2.1) becomes the Jordan canonical form: $\Lambda_s \approx \text{Diag}\{0, 0.142, -0.290, -1.26 \times 10^8\}$. Then, by using center manifold theory [19] on the transformed system of (2.1), we get the differential equation describing dynamics of the system, restricted to the center manifold: $\dot{x}_1 \approx -2.66 \times 10^{-12} \mu - 1.93 \times 10^{-4} x_1^2$, for which the perturbation value of μ near the saddle-node point is roughly $\mu \approx 10^{12}$, about 2% of α (see Figure 2.1(b)), as expected. The bifurcation diagram restricted on the center manifold is depicted in Figure 2.2(a), with

the corresponding bifurcation diagram in the original system, projected in the α - r plane as shown in Figure 2.2(b). It should be noted that the scaling between the graphs in Figures 2.2(a) and 2.2(b) depends upon the transformation matrix T_s . Also, note that the upper half branch in Figure 2.2(a) (denoted by the solid line) indicates that it is stable, but is only restricted to the 1-dimensional center manifold. For the whole system, this branch is still unstable since the system contains an unstable manifold (as shown in Figure 2.2(b)).

2.2.3.2 Hopf bifurcation and limit cycles

To find any possible Hopf bifurcation which may occur from the infected equilibrium E_1 , we first need to determine the critical points at which Hopf bifurcation occurs. The necessary and sufficient conditions for general n -dimensional systems to have a Hopf bifurcation are obtained in [43]. To state the theorem, consider the following general nonlinear differential system:

$$\dot{x} = f(x, \alpha), \quad x \in \mathbb{R}^n, \quad \alpha \in \mathbb{R}^m. \quad (2.10)$$

with an equilibrium determined from $f(x, \alpha) = 0$, as, say, $x_e = x_e(\alpha)$. To find the stability of x_e , evaluating the Jacobian of system (2.10) at $x = x_e(\alpha)$ yields $J(\alpha) = D_x f|_{x=x_e(\alpha)} = \left[\frac{\partial f_i(x_e(\alpha), \alpha)}{\partial x_j} \right]$. The eigenvalues of the Jacobian $J(\alpha)$ are determined by the following characteristic polynomial:

$$\begin{aligned} P_n(\lambda) &= \det[\lambda I - J(\alpha)] \\ &= \lambda^n + a_1(\alpha) \lambda^{n-1} + a_2(\alpha) \lambda^{n-2} + \cdots + a_{n-2}(\alpha) \lambda^2 + a_{n-1}(\alpha) \lambda + a_n(\alpha). \end{aligned} \quad (2.11)$$

Then, by the Hurwitz Criterion [21], we know that the equilibrium solution $x_e(\alpha)$ is asymptotically stable if and only if all the roots of the polynomial $P_n(\lambda)$ have negative real part, or equivalently, if and only if all the following Hurwitz arrangements $\Delta_i(\alpha)$, ($i = 1, 2, \dots, n$) are positive:

$$\Delta_1 = a_1, \quad \Delta_2 = \det \begin{bmatrix} a_1 & 1 \\ a_3 & a_2 \end{bmatrix}, \quad \Delta_3 = \det \begin{bmatrix} a_1 & 1 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{bmatrix}, \quad \cdots \quad \Delta_n = a_n \cdot \Delta_{n-1}.$$

Having defined the Hurwitz arrangements as above, we have the following theorem.

Theorem 2.2.1 [43] *The necessary and sufficient condition for a Hopf bifurcation to occur from the equilibrium solution $x_e(\alpha)$ of system (2.10) is $\Delta_{n-1} = 0$, with $a_n > 0$ and $\Delta_i > 0$, for $1 \leq i \leq n - 2$.*

In order to further consider the post-critical dynamical behaviour of the system and to determine the stability of bifurcating limit cycles, we may apply normal form theory to system (2.10). Assume that at a critical point $\alpha = \alpha_c$, the Jacobian of (2.10) evaluated at the equilibrium x_e contains a pair of purely imaginary eigenvalues $\pm i\omega_c$, and all other eigenvalues have negative real part. Then, the normal form of system (2.10) associated with Hopf bifurcation can be written in polar coordinates as (e.g., see [42])

$$\frac{d\rho}{dt} = \rho \left(v_0 \mu + v_1 \rho^2 + \cdots \right), \quad \frac{d\theta}{dt} = \omega_c + t_0 \mu + t_1 \rho^2 + \cdots, \quad (2.12)$$

where $\mu = \alpha - \alpha_c$, ρ and θ denote the amplitude and phase of motion, respectively. Then, the first equation of (2.12) can be used to approximate the amplitude of bifurcating limit cycles and to determine their stability. The second equation of (2.12) can determine the frequency of periodic motion. The coefficient v_1 , usually called the first-order focus value, plays an important role in determining the stability of limit cycles. When $v_1 < 0$ ($v_1 > 0$, respectively), the Hopf bifurcation is called supercritical (subcritical) and the bifurcating limit cycles are stable (unstable). The Maple program developed in [42] can be easily applied to system (2.10) to obtain the normal form (2.12). The coefficients v_0 and t_0 for the linear part of system (2.10) can be found from a linear analysis, given by [44], $v_0 = \frac{1}{2}(a_{11} + a_{22})$, $t_0 = \frac{1}{2}(a_{12} - a_{21})$, where $a_{ij} = \frac{\partial f_i}{\partial x_j \partial \mu}$, evaluated at the critical point.

We now apply the above formula to consider the infected equilibrium E_1 of system (2.10). To check if there exists Hopf bifurcation from E_1 , based on the fourth-degree characteristic polynomial $P_1(\xi, r_{e1})$, we apply the formula $\Delta_3 = a_1 a_2 a_3 - a_3^2 - a_1^2 a_4 = 0$ and solve this equation for r to obtain a unique value, $r_H > 0$, such that (by using (2.9)) $\alpha_H = \alpha_1(r_H) > 0$. When the parameter values in Table 2.1 are used, these critical values are given by: $(r_H, \alpha_H) \approx (6.72 \times 10^{13}, 2.64 \times 10^{13})$, at which the Jacobian of system (2.1) contains a purely imaginary pair and two negative real eigenvalues: $\pm 0.308i$, -1.66 , and -3.66×10^7 . Thus, as α is varied across α_H , a Hopf bifurcation occurs from E_1 , leading to a family of limit cycles.

To find the approximate solutions of the limit cycles and to determine their stability, we apply normal form theory to this model associated with this singularity. First, we apply a transformation $(x, y, r, a)^T = (x_{e1}, y_{e1}, r_{e1}, a_{e1})^T + T_H(x_1, x_2, x_3, x_4)^T$, where $(x_{e1}, y_{e1}, r_{e1}, a_{e1})$ is the infected equilibrium solution E_1 , and T_H is a constant, non-singular matrix. We obtain a transformed system of (2.1), which is omitted here due to its lengthy expression. Then, applying the formulas $v_0 = \frac{1}{2}(a_{11} + a_{22})$, $t_0 = \frac{1}{2}(a_{12} - a_{21})$ to the transformed system, we obtain $v_0 \approx 3.15 \times 10^{-15}$ and $t_0 \approx 3.33 \times 10^{-15}$. Further, we apply the Maple program [42] to the transformed system to obtain $v_1 \approx -4.18 \times 10^{-7}$, and $t_1 \approx -3.38 \times 10^{-6}$. Thus, the normal form up to third order is given by

$$\begin{aligned} \frac{d\rho}{dt} &\approx \rho(3.15 \times 10^{-15} \mu - 4.18 \times 10^{-7} \rho^2 + \dots), \\ \frac{d\theta}{dt} &\approx 0.308 + 3.33 \times 10^{-15} \mu - 3.38 \times 10^{-6} \rho^2 + \dots \end{aligned} \quad (2.13)$$

The first equation of (2.13) can be used to analyze the bifurcation and stability of bifurcating limit cycles. Setting $\frac{d\rho}{dt} = 0$ results in two solutions: $\rho = 0$, which represents the infected equilibrium solution E_1 ; and $\rho \approx 8.68 \times 10^{-5} \sqrt{\mu}$ ($\mu > 0$), which is an approximation of the amplitude of bifurcating limit cycles. Since $v_1 < 0$, this is a supercritical Hopf bifurcation, and bifurcating limit cycles are stable. For example, choose $\mu = 10^{12}$. Then, the approximate amplitude of the limit cycle is $\rho \approx 86.8$, and the frequency of the limit cycle approximately equals $\omega \approx 0.283$, slightly less than $\omega_c \approx 0.308$. The phase portrait of the simulated limit cycle, projected on the x - y plane, is shown in Figure 2.3(d). It can be seen from Figure 2.3(a) and (d) that the analytical prediction from the normal form, $\rho \approx 86.8$, agrees well with the simulated result.

The above analysis based on normal form theory is for local dynamical behaviour, that is, the limit cycles must be near the Hopf critical point (r_H, α_H) . It can be seen from Figure 2.1(b) that values of α taken from the interval $\alpha \in (\alpha_H, \alpha_t)$ lead to unstable equilibrium solutions

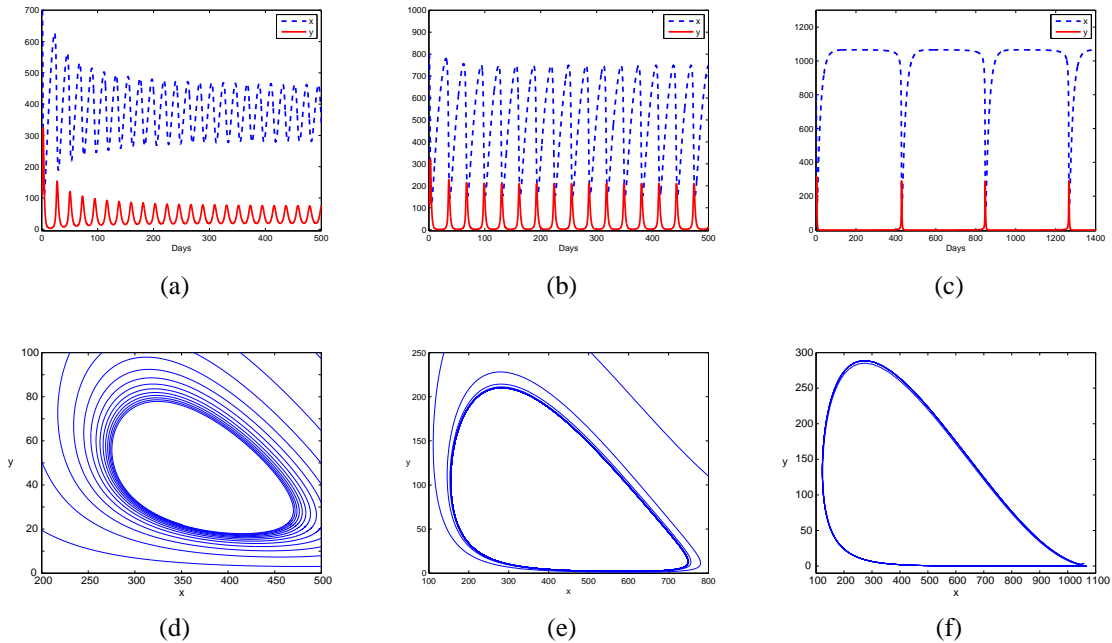


Figure 2.3: Simulated limit cycles of system (2.1) for the parameter values taken from Table 2.1, with the time course of x and y on the top row, and the corresponding phase portraits projected on the x - y plane on the bottom row. For (a) and (d) $\alpha = 2.74 \times 10^{13}$, (b) and (e) $\alpha = 3.50 \times 10^{13}$, and (c) and (f) $\alpha = 4.55 \times 10^{13}$.

(since both E_0 and E_1 are unstable for this interval). However, due to the solutions being non-negative and bounded, we expect that there should exist certain persistent motion such as oscillating solutions for the values of α taken from this interval, and the amplitudes of these oscillations can be large. For example, for $\alpha = 3.50 \times 10^{13}$, the phase portrait of the simulated solution, projected on the x - y plane is shown in Figure 2.3(e), corresponding to the oscillations in time shown in Figure 2.3(b), which have much greater amplitude than the oscillations in Figure 2.3(a).

Now, we take a particular value of α from the interval $\alpha \in (\alpha_H, \alpha_t)$, which is close to α_t , to simulate the system. For example, taking $\alpha = 4.55 \times 10^{13} < \alpha_t \approx 4.58 \times 10^{13}$, we obtain the phase portrait of the simulated oscillating solution, projected on the x - y plane, shown in Figure 2.3(f) with corresponding time history of x and y shown in Figure 2.3(c). This clearly shows viral blips.

Next, we will discuss what conditions are needed for creating the phenomenon of viral blips.

2.2.4 Conditions for generating viral blips

In the previous subsection, we carefully analysed the occurrence of viral blips in a 4-dimensional HIV model (2.1). System (2.1) is an example of *in-host infection model*, an ODE system describing the dynamics of infection within a single infected individual. In-host infection models, based on classical Susceptible-Infected-Recovered (SIR) models in epidemiology [1], typically include populations of uninfected target cells, infected target cells, and the infection dynamics

between the two classes [28]. More complex models also include populations of free virus, latently-infected cells, and various relevant components of the immune response, depending on the infection under study. Although there are many exceptional cases, in-host models typically admit an uninfected equilibrium and at least one infected equilibrium, analogous to the disease-free and endemic equilibria of an SIR model.

Since in-host infection models share many similar features, much of our understanding regarding the behaviour of system (2.1) can be generalized to other models. Based on insights obtained in analysing system (2.1), we propose in the following hypothesis four conditions for an in-host infection model to generate viral blips:

Hypothesis 1: The following conditions are needed for an in-host infection model to generate viral blips:

- (i) there exist at least two equilibrium solutions;
- (ii) there exists a transcritical bifurcation at an intersection of the two equilibrium solutions;
- (iii) there is a Hopf bifurcation which occurs from one of the equilibrium solutions; and
- (iv) large oscillations (or more generally, global, persistent motions) can occur near the transcritical critical point.

The reasons for conditions (i) and (ii) are simple, because when a parameter that reflects infection severity is chosen as a bifurcation parameter, an in-host infection model typically starts at the uninfected equilibrium and then bifurcates to the infected equilibrium as the parameter is increased. Thus, these two equilibrium solutions must exchange their stability, yielding a transcritical bifurcation. For the 4-dimensional model considered in the previous subsection, the uninfected equilibrium E_0 and the infected equilibrium E_1 intersect at the critical point (α_t, r_t) , where they exchange their stability. In fact, E_0 is stable (unstable) for $\alpha > \alpha_t$ ($\alpha < \alpha_t$), while the lower branch of E_1 is stable (unstable) for $\alpha < \alpha_t$ ($\alpha > \alpha_t$), as shown in Figure 2.1(b).

Condition (iii), the existence of a Hopf bifurcation, is necessary to obtain oscillations. It can be seen from Figure 2.1(b) that limit cycles bifurcate from E_1 at the Hopf critical point (α_H, r_H) , and the limit cycles become larger if $\mu = \alpha - \alpha_H > 0$ increases.

The reasoning behind the last condition (iv) is not so obvious. Large oscillations (or global, persistent motions) are necessary, near the transcritical point, for viral blips to emerge. As shown in Figure 2.1(b), both E_0 and E_1 are unstable for $\alpha \in (\alpha_H, \alpha_t)$ (though a part of the lower branch of E_1 is stable but it is biologically meaningless due to $y < 0$). Thus, there exist large oscillations near the transcritical critical point α_t . Moreover, it is noted from Figure 2.1(b) that at the left side

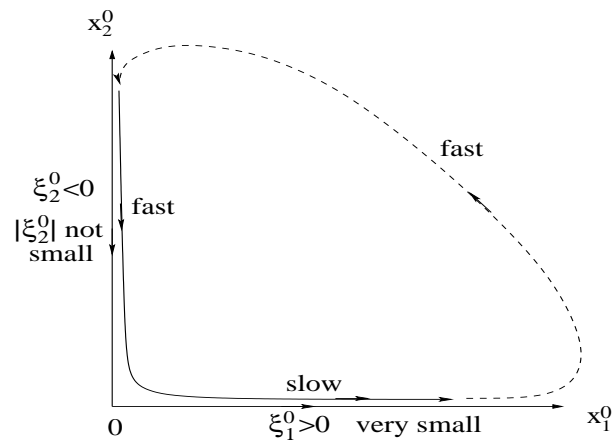


Figure 2.4: Schematic diagram for explaining the occurrence of blips.

of the transcritical point α_t , the eigenvalues evaluated at E_0 are all real, containing one positive eigenvalue ($\xi_1^0 > 0$) and three negative eigenvalues ($\xi_i^0 < 0, i = 2, 3, 4$). In other words, any point on the uninfected equilibrium E_0 for $\alpha < \alpha_t$ is a saddle point. Since ξ_1^0 crosses zero at the critical point $\alpha = \alpha_t$, ξ_1^0 is very small near the critical point for $\alpha < \alpha_t$.

Now suppose we consider a value of $\alpha < \alpha_t$, but near the critical point $\alpha = \alpha_t$ (e.g., $\alpha = 0.455 \times 10^{14}$, as shown in Figure 2.3(c) and (f)). For simplicity, we may consider a submanifold whose linear part is characterized by the eigenvalues ξ_1^0 and ξ_2^0 , and the corresponding coordinates are x_1^0 and x_2^0 , respectively. A solution trajectory of system (2.1) for such a value of α , projected on this submanifold, is depicted in Figure 2.4. Due to $0 < \xi_1^0 \ll 1$, the trajectory moves away from the critical point very slowly near the x_1^0 -axis, while it moves rapidly toward the critical point near the x_2^0 -axis since $|\xi_2^0|$ is not small. Further, due to the global boundedness of solutions, the part of the trajectory which is not close to the saddle point moves rapidly, as shown in Figure 2.4. This fast-slow motion yields the blips phenomenon, with slow changes corresponding to the near-flat section in the time history, and rapid changes occurring during the viral blips, as shown in Figure 2.3(c) and (f). In other words, the trajectory spends relatively long periods in regions of state space which lie very close to the uninfected equilibrium, then transiently visits regions of state space which are close to the infected equilibrium.

2.3 A simple 3-dimensional in-host infection model producing blips

Having established the conditions in Hypothesis 1 for generating viral blips, we are ready to turn to some basic questions such as: what types of in-host infection model can generate blips? and, what is the minimum dimension of such models?

2.3.1 Generalizing ROS to other physical variables

In model (2.1), the variable r represents ROS, which are produced naturally in the body. In HIV infection, extra ROS are generated by infected cells, and these in turn directly accelerate HIV progression [29, 36]. Therefore, infectivity β is an increasing and saturating function of ROS concentrations. However, we note that the form of the infection term is not specific to HIV nor to ROS, and models of a similar form could in fact apply to other infections. To generalize the physical meaning of the variable r , we can for example let r denote any damage caused by the infection, for example to the humoral immune response, to infected organs, or to the infected individual aspecifically. The model assumes that “damage” increases with the extent of the infection at rate ky , and is repaired or cleared at rate $d_r r$. This yields the 3-dimensional system:

$$\dot{x} = \lambda_x - d_x x - \beta(r)xy, \quad \dot{y} = \beta(r)xy - d_y y, \quad \dot{r} = ky - d_r r. \quad (2.14)$$

To achieve an infection term similar to that in model (2.1), we further assume that accrued damage makes target cells more vulnerable to infection, that is, accrued damage increases the infection rate. We thus take $\beta(r)$ to be an increasing, saturating function of r .

In the original model (2.1), r represents ROS, for example H_2O_2 , whose production and decay rates are both extremely fast. For the more general model (2.14), we would like to assess whether viral blips are still possible at more moderate production and repair rates, k and d_r . For ROS the decay rate $d_r = 1.66 \times 10^7 \text{ day}^{-1}$ implies a half life of only 4ms. We decreased d_r by several orders of magnitude; in particular, at $d_r = 1.0 \times 10^3 \text{ day}^{-1}$, a half life of 60s, we find that viral blips are still possible. For this value of d_r , we can take $k = 1.49 \times 10^{15} \text{ molecules cell}^{-1} \text{ day}^{-1}$. Note that λ_r has been set to zero in (2.14) to make the model more general.

For simplicity, let $a = b_{\max} - b_0$, $b = b_0$ and $c = r_{\text{half}}$. Then, the function $\beta(r)$ is rewritten as $\beta(r) = b + \frac{ar}{r+c}$, and a , b , and c are treated as bifurcation parameters. Parameter values λ_x , d_x , d_y , k , d_r , b_0 , b_{\max} , and r_{half} are given in Table 2.1. For practically meaningful solutions, the values of the bifurcation parameters will be chosen close to the values in Table 2.1.

To analyze (2.14), we can follow the same procedure used in the previous section and treat b as a bifurcation parameter. First of all, it is easy to prove the well-posedness of system (2.14). Next, we get the infection-free equilibrium $E_0 : (x_{e0}, y_{e0}, r_{e0}) = (\lambda_x/d_x, 0, 0)$ and the infected equilibrium $E_1 := (x_{e1}, y_{e1}, r_{e1})$, where $x_{e1} = \frac{d_y(r_{e1}+c)}{(a+b)r_{e1}+bc}$, $y_{e1} = \frac{1}{d_y}(\lambda_x - d_x x_{e1})$, and r_{e1} is determined by $F_1(r, c) = d_r d_y (a+b)r^2 + [d_y(d)rbc - kd_x] - k\lambda_x(a+b) r + kc(d_x d_y - b\lambda_x) = 0$. Again, it is easy to show that E_0 and E_1 intersect at the transcritical bifurcation point $(b_t, r_t) \approx (9.38 \times 10^{-4}, 0)$. On the infected equilibrium E_1 , there are two saddle-node bifurcation points (turning points), $(b_{s_1}, r_{s_1}) \approx (-1.49 \times 10^{-3}, 4.18 \times 10^{13})$, and $(b_{s_2}, r_{s_2}) \approx (-5.77 \times 10^{-3}, 3.05 \times 10^{14})$, and a Hopf bifurcation point $(b_H, r_H) \approx (6.56 \times 10^{-4}, 7.24 \times 10^{13})$.

The bifurcation diagram and simulated results are shown in Figure 2.5. All the conditions (i)-(iv) in Hypothesis 1 are satisfied. Blips do appear since the Hopf critical point is close to the transcritical point. However, because E_0 is not globally stable, depending on the initial conditions, the oscillation may converge to the stable equilibrium E_1 (see Figure 2.5(c)), or converge to a limit cycle with large amplitude (blips), as shown in Figure 2.5(d). Convergence to a smaller, regular oscillation due to the Hopf bifurcation is also possible (not shown in Figure 2.5).

2.3.2 Identifying the region of parameter space exhibiting viral blips

Having found viral blip behaviour in the simple 3-dimensional infection model (2.14), we are now further interested in identifying the region of parameter space in which viral blips may occur. This is particularly useful in applications since in reality, all parameters are roughly measured. Thus, we need to study the robustness of the phenomenon to variations in the system parameters. If blips only appear for a very small region in the parameter space, then the results are not practically useful. The main idea of identifying the region where blips may occur is to study the instability of the solutions of the system. Once the unstable region is identified, blips can be found by using the other conditions in Hypothesis 1. In order to simplify the analysis, we first introduce state variable scaling and parameter rescaling into system (2.14).

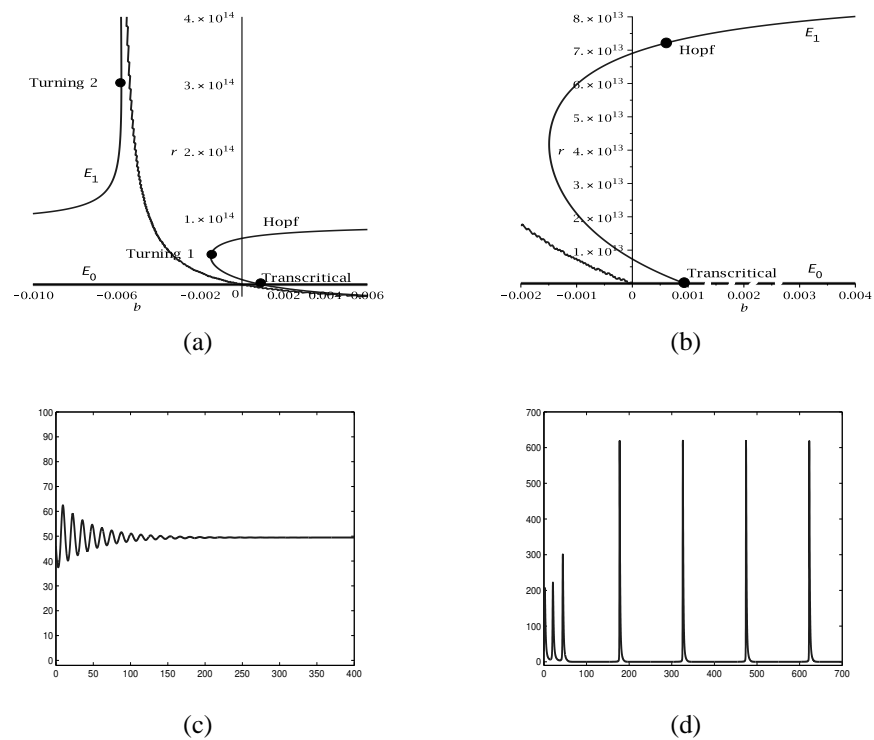


Figure 2.5: Dynamics and bifurcation of system (2.14) for $d_r = 1.0 \times 10^3$, $k = 1.49 \times 10^{15}$: (a) bifurcation diagram projected on the b - r plane; (b) a close-up of part (a); (c) simulated time history $y(t)$ converging to E_1 for $b = 0.001$ with the initial condition $(x, y, r) = (178, 46, 73)$ close to E_1 ; and (d) simulated time history $y(t)$ converging to a stable limit cycle (blips) for $b = 0.001$ with the initial condition $(x, y, r) = (1005, 3, 3)$ close to E_0 .

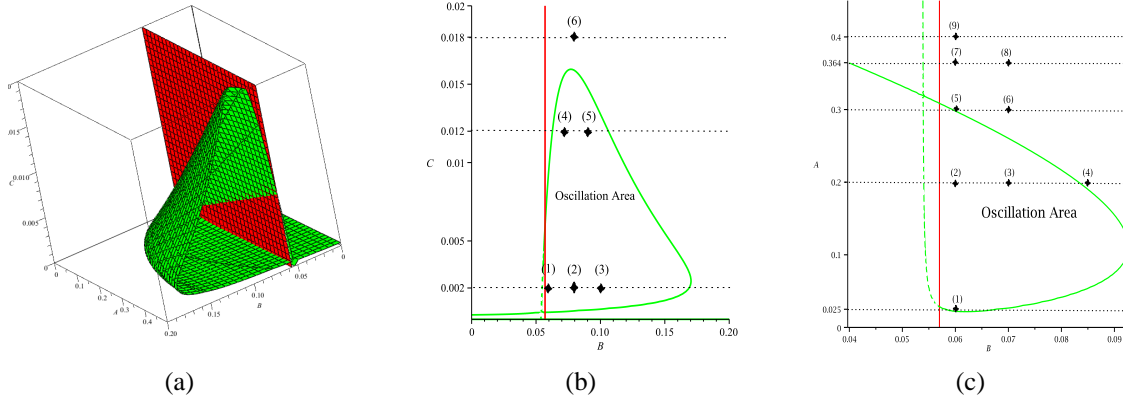


Figure 2.6: (a) Graph of $\Delta_2 = 0$ in the A - B - C parameter space, identifying the region yielding oscillations; (b) cross section of panel (a) where $A = 0.364$; and (c) cross section of panel (a) where $C = 3.94 \times 10^{-4}$.

2.3.2.1 State variable scaling and parameter rescaling

Introducing the following scaling $x = c_1 X$, $y = c_2 Y$, $r = c_3 R$, $t = c_4 \tau$, where $c_1 = \frac{\lambda_x}{d_y}$, $c_2 = \frac{\lambda_x}{d_y}$, $c_3 = \frac{\lambda_x k}{10^{13} d_y^2}$, $c_4 = \frac{1}{d_y}$, to (2.14) and letting $A = \frac{a \lambda_x}{d_y^2}$, $B = \frac{b \lambda_x}{d_y^2}$, $C = \frac{c d_y^2}{10^{12} \lambda_x k}$, $D_x = \frac{d_x}{d_y}$, $D_r = \frac{d_r}{d_y}$, yields the following scaled system

$$\dot{X} = 1 - D_x X - \left(B + \frac{AR}{R+C} \right) XY, \quad \dot{Y} = \left(B + \frac{AR}{R+C} \right) XY - Y, \quad \dot{R} = Y - D_r R, \quad (2.15)$$

which will be used in the following analysis, with the scaled parameter values given by

$$A = 0.364, \quad C = 3.94 \times 10^{-4}, \quad D_x = 0.057, \quad D_r = 1000, \quad (2.16)$$

and B is treated as a bifurcation parameter.

2.3.2.2 Equilibrium solutions and their stability

The bifurcation patterns of the scaled system (2.15) are the same as that of system (2.14). Two equilibrium solutions are $E_0 : (X_{e0}, Y_{e0}, R_{e0}) = (1/D_x, 0, 0)$, and $E_1 : (X_{e1}, Y_{e1}, R_{e1})$, where $X_{e1} = \frac{R_{e1} + C}{(A+B)R_{e1} + BC}$, $Y_{e1} = 1 - \frac{D_x(R_{e1} + C)}{(A+B)R_{e1} + BC}$, and R_{e1} is determined from the equation $F_3(R) = D_r(A+B)R^2 + [D_r(BC+1) - (A+B)]R + D_r - B)C = 0$.

The characteristic polynomial for E_0 is $P_0(\xi) = (\xi + D_x)(\xi + D_r)(\xi - \frac{B}{D_x})$. It is easy to show that E_0 and E_1 exchange stability at the transcritical bifurcation point $B = D_x$. The characteristic polynomial for E_1 is $P_1(\xi) = \xi^3 + a_1(r)\xi^2 + a_2(r)\xi + a_3(r)$, and Hopf critical point is determined by $\Delta_2 = a_1(r)a_2(r) - a_3(r) = 0$. We fix parameters D_r and D_x , and choose A , B and C as bifurcation parameters. Then, we want to find the parameter region where blips may occur. First of all, a Hopf bifurcation is necessary, requiring the condition $\Delta_2(A, B, C) = 0$. The graph of $\Delta_2(A, B, C) = 0$ is plotted in the 3-dimensional A - B - C parameter space, as shown in Figure 2.6(a), where the green hypersurface defines a set of points which are Hopf critical points; and the region bounded by the green surface is unstable for E_1 , leading to oscillations. Thus blips may occur within this region and near the boundary as well, depending on the relative position of the Hopf critical point with respect to the transcritical point.

In the following, we fix either parameter A or C to obtain two-dimensional graphs, which illustrate more clearly the bifurcations necessary for blips.

2.3.2.3 Parameter A fixed

Fix $A = 0.364$, which cuts the surface in Figure 2.6(a) to yield curves, as shown in Figure 2.6(b). The transcritical bifurcation occurs at $B = 0.057$, which is denoted by a red line in Figure 2.6(b). A Hopf bifurcation occurs on the green curve, and the region bounded by the green and red curves indicates where oscillations can happen. It should be noted that the above results are based on local dynamical analysis, thus blips may also appear outside this bounded region, but close to the green curve.

We take three typical values of C (as the three dotted lines shown in Figure 2.6(b)), and obtain the Hopf critical points as follows.

$$\begin{aligned}
 C = 0.002 : & \quad (B_H, R_H) \approx (1.69 \times 10^{-1}, 7.90 \times 10^{-4}), \\
 C = 0.012 : & \quad (B_{H_1}, R_{H_1}) \approx (6.27 \times 10^{-2}, 1.53 \times 10^{-4}), \\
 C = 0.012 : & \quad (B_{H_2}, R_{H_2}) \approx (1.06 \times 10^{-1}, 5.31 \times 10^{-4}), \\
 C = 0.018 : & \quad \text{No Hopf critical point.}
 \end{aligned} \tag{2.17}$$

The bifurcation diagrams corresponding to the three lines, $C = 0.002$, $C = 0.012$ and $C = 0.018$, are shown in the top three graphs in Figure 2.7. Six simulated results are also presented in this figure, corresponding to the six points marked on the three dotted lines in Figure 2.6(b). It is seen that the values taken from the points (1)–(4) generate blips; point (5) leads to a regular oscillation, while point (6) gives a simple stable equilibrium solution, as expected. For this case when parameter A is fixed, no blips have been found for the values outside the region bounded by the red and green curves. It should be noted in the top middle figure of Figure 2.7 that there are two Hopf bifurcation points on the equilibrium E_1 . One of them is supercritical while the other is subcritical, but the two families of the limit cycles bifurcating from these two critical points are both stable, since the stability change is reversed at the two points. In fact, the three eigenvalues along the unstable part of E_1 between the two Hopf bifurcation points contain one negative eigenvalue and a pair of complex conjugates with positive real part. On the two stable parts, the real part of the complex conjugate eigenvalues changes sign to become negative. As the parameter C is increasing from 0.002 to 0.018, the two Hopf bifurcation points merge to a single point on E_1 (corresponding to the turning point on the green curve, see Figure 2.6(b), at which the horizontal line is tangent to the green curve); the corresponding eigenvalues contain a negative eigenvalue and a purely imaginary pair. This indeed characterizes a degenerate Hopf bifurcation (e.g. see [44]), different from the Hopf bifurcation defined by (2.12). A similar discussion applies to the other two Hopf bifurcation points shown in the top left figure in Figure 2.8.

2.3.2.4 Parameter C fixed

Now we fix parameter $C = 3.94 \times 10^{-4}$, which results in curves in the A - B plane by cutting the surface in Figure 2.6(a), as shown in Figure 2.6(c). The transcritical point is kept the same:

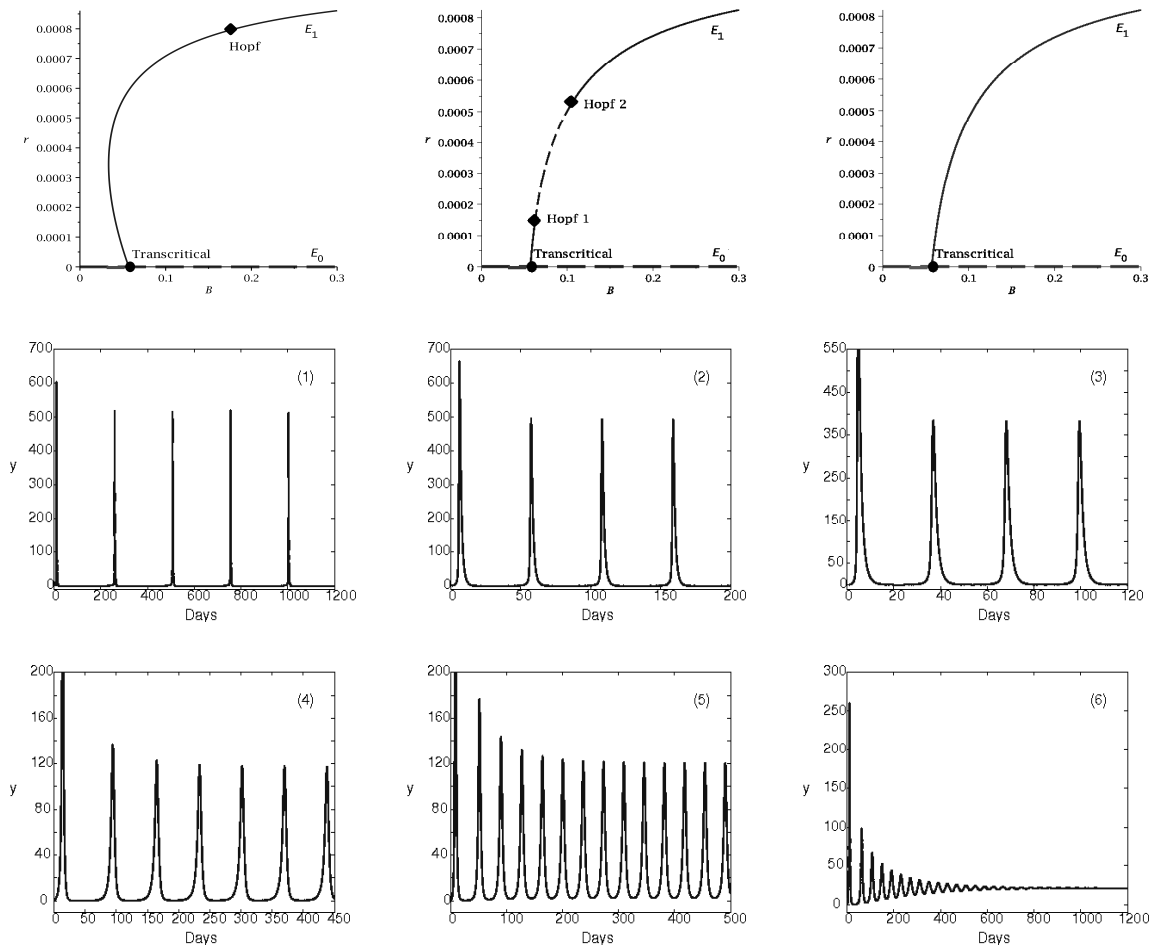


Figure 2.7: Bifurcation diagrams corresponding to $C = 0.002, 0.012$ and 0.018 , respectively, and numerical simulation results for the parameter values $(B, C) = (0.06, 0.002)^{(1)}$, $(0.08, 0.002)^{(2)}$, $(0.10, 0.002)^{(3)}$, $(0.07, 0.012)^{(4)}$, $(0.09, 0.012)^{(5)}$, $(0.08, 0.018)^{(6)}$.

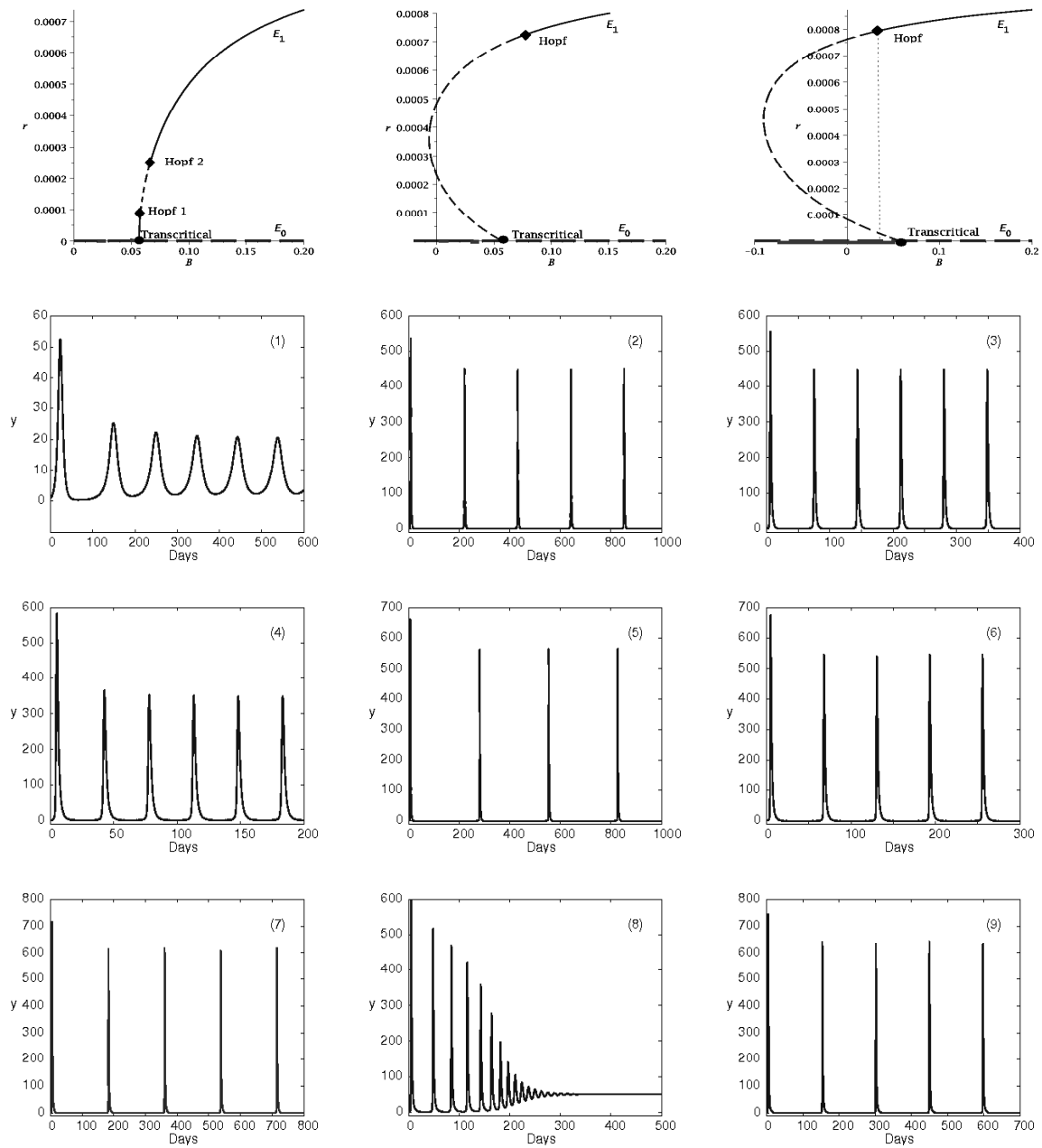


Figure 2.8: Bifurcation diagrams corresponding to $A = 0.025, 0.200, 0.364$, and numerical simulation results for the parameter values $(A, B) = (0.025, 0.060)^{(1)}, (0.200, 0.060)^{(2)}, (0.200, 0.070)^{(3)}, (0.200, 0.085)^{(4)}, (0.300, 0.059)^{(5)}, (0.300, 0.070)^{(6)}, (0.364, 0.060)^{(7)}, (0.364, 0.070)^{(8)}, (0.400, 0.060)^{(9)}$.

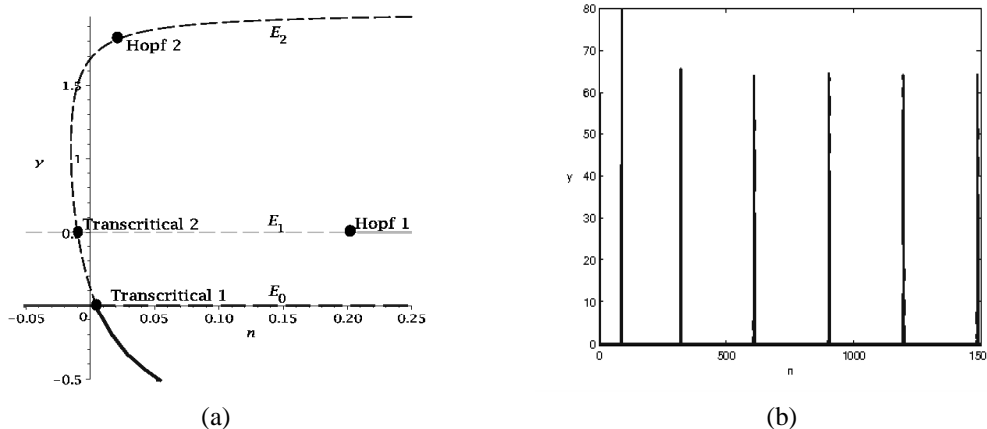


Figure 2.9: (a) Bifurcation diagram of system (2.19), showing the equilibrium solutions E_0 , E_1 , and E_2 with dashed and solid lines denoting unstable and stable, respectively; (b) Simulated viral blips in system (2.19) for $n = 0.007$. Other parameter values used here are: $\lambda = k = p = 1$, $d = 0.01$, $m = b = 0.05$, $a = 0.5$, $c = 0.1$.

$B = 0.057$. We choose three typical values of A , and find the Hopf bifurcation points as follows.

$$\begin{aligned}
 A = 0.025 & : (B_{H_1}, R_{H_1}) \approx (5.82 \times 10^{-2}, 9.84 \times 10^{-5}), \\
 A = 0.025 & : (B_{H_2}, R_{H_2}) \approx (6.75 \times 10^{-2}, 2.65 \times 10^{-4}), \\
 A = 0.200 & : (B_H, R_H) \approx (8.32 \times 10^{-2}, 7.33 \times 10^{-4}), \\
 A = 0.364 & : (B_H, R_H) \approx (3.99 \times 10^{-2}, 7.99 \times 10^{-4}).
 \end{aligned} \tag{2.18}$$

The bifurcation diagrams corresponding to the three lines $A = 0.025$, $A = 0.200$ and $A = 0.364$ are shown in the top three graphs in Figure 2.8. Nine simulated results are also presented in this figure, corresponding to the nine points marked on the five dotted lines in Figure 2.6(c). It is observed from these graphs that among the nine chosen parameter values, seven cases exhibit blips (see the points (2)–(7) and (9) in Figure 2.6(c) with the corresponding simulated results shown in Figure 2.8). It is noted that some of these points are not even close to the red line, nor in the region bounded by the red and green curves, suggesting that a simple 3-dimensional HIV model can generate rich blips.

2.3.3 3-dimensional immunological model

In this subsection, we briefly consider an immunological model [28], and apply Hypothesis 1 to show that the model can have blips. For simplicity, the original 4-d model is reduced (by a quasi-steady state assumption on the virus particles) to a 3-d model, described by

$$\begin{aligned}
 \dot{x} &= \lambda - dx - \beta(y)xy, \\
 \dot{y} &= \beta(y)xy - ay - pyz, \\
 \dot{z} &= cyz - bz,
 \end{aligned} \tag{2.19}$$

where x , y and z represent the densities of the infected cells, uninfected cells, and CTL, respectively. The system (2.19) with constant $\beta(y)$ is well-known [6, 27], which does not exhibit blips. In order to generate viral blips, here we choose $\beta(y) = n + \frac{my}{y+k}$, where n and m are minimum

and maximum infectivity, and k represents the density of infected cells when the infectivity takes its median value. Since the analysis is similar to previous models, we omit the details and only give the results as follows. The system (2.19) has three equilibrium solutions: the infection-free equilibrium, E_0 , the infected equilibrium with CTL, E_1 , and the infected equilibrium without CTL, E_2 . There are two transcritical bifurcation points, one of them, named “transcritical 1” in Figure 2.9(a), is at the intersection of E_0 and E_2 : $(n_{t1}, y_{t1}) \approx (0.005, 0)$, at which E_0 and E_2 exchange their stability. The second one occurs at the intersection of E_1 and E_2 : $(n_{t2}, y_{t2}) \approx (-0.01, 0.5)$, called “transcritical 2” in Figure 2.9(a). However, note that they only exchange their stability if restricted to a one-dimensional manifold, and both of them are unstable in the whole space since one of the eigenvalues keep positive when crossing this transcritical point. E_1 becomes stable until n is increased to cross a Hopf critical point (called “Hopf 1” in Figure 2.9(a)): $(n_{1H}, y_{1H}) \approx (0.206, 0.5)$. Another Hopf bifurcation point (called “Hopf 2” in Figure 2.9(a)) happens on E_2 at $(n_{2H}, y_{2H}) \approx (0.0213, 1.81)$. The limit cycles bifurcating from Hopf 1 are stable, while those from Hopf 2 are unstable, leading to large oscillating motions when the values of n are chosen from the interval (n_t, n_{2H}) . The above results show that all the four conditions in Hypothesis 1 are satisfied, and blips indeed appear. The simulated blips for $n = 0.007$ are depicted in Figure 2.9(b).

2.4 A 2-dimensional in-host infection model

For the generalized 3-dimensional model discussed in Section 2.3, we assume that r is some form of damage to the host or to the host immune system, which increases with the extent of the infection, that is, in proportion to the infected cell density. Here, we further assume that there is a quasi-steady state (as used in (2.14)) between the damage, r , and the infected cell density y . Thus, the 3-dimensional HIV model can be further reduced to a 2-dimensional model, given by

$$\dot{x} = \lambda_x - d_x x - \beta(y)xy, \quad \dot{y} = \beta(y)xy - d_y y, \quad (2.20)$$

Note that system (2.20) is now in the form of an in-host infection model, which includes only uninfected and infected target cell populations, and the most basic “birth” and death rates. However, we now think of the infectivity $\beta(y)$ as a possible function of y ; other parameters have the same meaning as in (2.19). We will show that this simplified 2-dimensional infection model may also be able to generate blips.

2.4.1 2-dimensional in-host model with constant and linear infection rates

First, we consider the case when the infection rate, $\beta(y)$ is simply a constant function, that is $\beta(y) = \beta$. Taking β itself as a bifurcation parameter, it is easy to show that there exist two equilibrium solutions and a transcritical bifurcation point, but no Hopf bifurcation exists. This violates Hypothesis 1, and therefore no blips can appear in this case.

Next, suppose the infection rate $\beta(y)$ is a linear function of the infected cell density, y , that is $\beta(y) = b + ay$, where the parameters a and b represent the same constants as before, and a is treated as a bifurcation parameter. In this case, we have two equilibrium solutions E_0 and E_1 . But E_0 is always stable for all values of a though there exists a Hopf bifurcation on E_1 .

Therefore, no transcritical bifurcation point exists for this case, which violates Hypothesis 1, implying that blips are not possible when $\beta(y)$ is a linear function.

2.4.2 A 2-dimensional in-host model with saturating infection rate

Motivated by our previous results for the 3- and 4-dimensional models, we next assume that infectivity is an increasing saturating function of the infected cell density, y , namely, $\beta(y) = b + \frac{ay}{y+c}$. For our numerical work, we take the same values of a and b , as used in Section 3.1, while c is taken to be $c = 50$, obtained by numerical simulation based on the experimental data given in [39]. Other parameter values are as described for model (2.14).

2.4.2.1 Scaling

For convenience in the following analysis, we first simplify system (2.20) by the following scaling to reduce the number of parameters. Let $x = e_1 X$, $y = e_2 Y$, $t = e_3 \tau$, where $e_1 = \frac{\lambda_x}{d_y}$, $e_2 = \frac{\lambda_x}{d_y}$, $e_3 = \frac{1}{d_y}$, and set $A = \frac{a\lambda_x}{d_y^2}$, $B = \frac{\lambda_x b}{d_y^2}$, $C = \frac{cd_y}{\lambda_x}$, $D = \frac{d_x}{d_y}$. Then, the rescaled system is given by

$$\begin{aligned}\frac{dX}{d\tau} &= 1 - DX - \left(B + \frac{AY}{Y+C}\right)XY, \\ \frac{dY}{d\tau} &= \left(B + \frac{AY}{Y+C}\right)XY - Y,\end{aligned}\tag{2.21}$$

with B treated as a bifurcation parameter. Taking the parameter values from [28], we have the scaled parameter values $A = 0.364$, $C = 0.823$, and $D = 0.057$ for system (2.21).

2.4.2.2 Equilibrium solutions and their stability

By setting $\dot{X} = \dot{Y} = 0$ in (2.21), we get two biologically meaningful equilibrium solutions: the uninfected equilibrium solution $E_0 : (X_0, Y_0) = (1/D, 0)$, and the infected equilibrium solution $E_1 = (X_1, Y_1)$, where $X_1 = \frac{Y_1+C}{(A+B)Y_1+BC}$, and Y_1 is determined by the equation $F_1 = (A+B)Y^2 + (D+BC-A-B)Y + (D-B)C = 0$. This indicates that the condition (i) in Hypothesis 1 is satisfied. Similarly, it is easy to find that E_0 is stable (unstable) if $B < D$ ($B > D$).

2.4.2.3 Bifurcation analysis

By using the characteristic polynomials at E_0 and E_1 , we can show that a transcritical bifurcation occurs at the critical point, $(Y_t, B_t) = (0, 0.057)$, which satisfies the condition (ii) in Hypothesis 1. E_0 and E_1 intersect at this critical point and exchange their stability. Further, a Hopf bifurcation happens at the critical point $(B_H, Y_H) \approx (0.121, 0.811)$. E_1 is stable (unstable) on the right (left) side of the Hopf bifurcation point. Therefore, the condition (iii) in Hypothesis 1 holds for this case. If we take a value of B near B_t on the side where both E_0 and E_1 are unstable, then the condition (iv) in Hypothesis 1, is also satisfied and so blips occur. The bifurcation diagram is shown in Figure 2.10(a), and the simulated viral blips for $B = 0.060$ are depicted in Figure 2.10(b).

Summarizing the results of this section, we conclude that the simple 2-dimensional in-host model is sufficiently complex to exhibit viral blips, provided the infectivity function is an

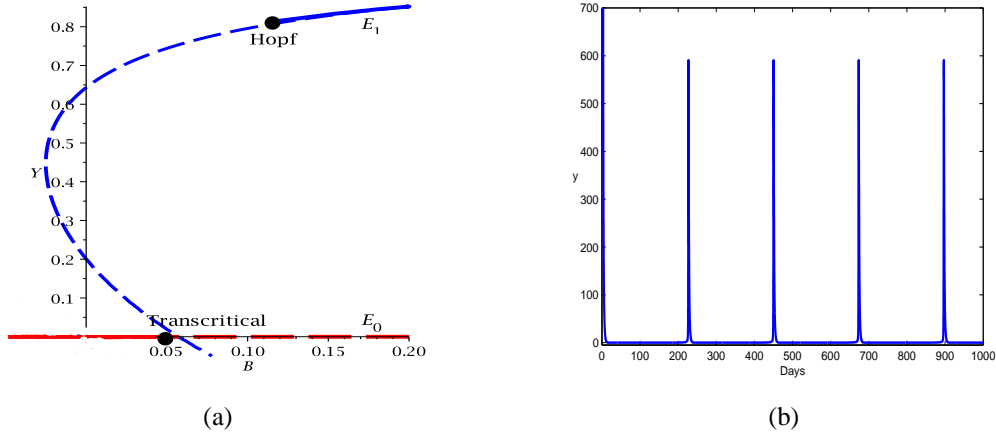


Figure 2.10: (a) Bifurcation diagram of system (2.21) projected on the B - Y plane, with the red and blue lines denoting the E_0 and E_1 , respectively, dotted and solid lines indicating unstable and stable, respectively; and (b) simulated time history of $y(t)$ for $B = 0.060$.

increasing, saturating function of infected cell density. However, for this model, the range of parameter space in which blips occur is relatively restricted, compared with the 3-dimensional model which is established in the previous section.

An interesting question is naturally raised here: does there exist a more general function $\beta(y)$ such that the existence of blips depends upon the general properties of the function like its maximal values and/or its derivatives. In fact, it has been found that by choosing the parameter c large enough in the function β , a threshold is reached beyond which the Hopf bifurcation, and hence also the viral blips, disappear.

2.5 Recurrency in a 5-dimensional model

So far, we have considered 2-, 3- and 4-dimensional in-host infection models with increasing, saturating infectivity functions, and shown that all these models exhibit blips. Moreover, it has been shown for the 2-dimensional model (and can be shown for the 3- and 4-dimensional models, but omitted here) that replacing the infectivity function with a constant or linear function of y will cause blips to disappear. However, in this section we will show that higher-dimensional systems may have blips even with a constant infectivity function.

We consider a previously proposed 5-dimensional immunological model, in which recurrent phenomena or viral blips have been observed via numerical simulation [41]. The model describes antibody concentrations and cytotoxic T lymphocytes (CTLs) explicitly, and is described as follows:

$$\dot{x} = \lambda - dx - \beta xv, \quad (2.22a)$$

$$\dot{y} = \beta xv - ay - pyz, \quad (2.22b)$$

$$\dot{z} = cyz - bz + hy, \quad (2.22c)$$

$$\dot{u} = \xi z - \eta u - kuv, \quad (2.22d)$$

$$\dot{v} = ey - kuv - \gamma xv - qv. \quad (2.22e)$$

Table 2.2: Parameter values used in model (2.22) [41].

Parameter	Value
λ	10^4 cells μL^{-1} day $^{-1}$
d	0.100 day $^{-1}$
β	1.25×10^{-5} virion $^{-1}$ μL day $^{-1}$
p	10^{-4} cells $^{-1}$ μL day $^{-1}$
c	10^{-4} cells $^{-1}$ μL day $^{-1}$
b	0.200 day $^{-1}$
h	$[0, 10^{-4}]$ day $^{-1}$
ξ	10.0 molecules cell $^{-1}$ day $^{-1}$
η	0.040 day $^{-1}$
k	2.50×10^{-5} particle $^{-1}$ μL day $^{-1}$
e	2.50 virions cell $^{-1}$ day $^{-1}$
γ	5.00×10^{-5} cell $^{-1}$ μL day $^{-1}$

Here x , y , z , u and v are respectively the population densities of uninfected target cells, infected target cells, CTLs, antibodies and virions. The parameters λ and d represent uninfected cells' constant growth rate and death rate, respectively. Target cells are infected by virus at rate, βxv . Infected cells die at rate ay , being killed by CTLs at rate pyz . It is assumed that CTLs proliferate at rate cyz , and decrease with the natural death rate bz . The fourth equation describes the antibody growth rate, ξz , which is proportional to the number of CTLs, the natural death rate of antibody, ηu , and the binding rate of one antibody with one antigen, kuv . In the last equation, viruses are released from infected cells at rate ey , and are bound by antibody, absorbed by uninfected cells, or cleared at rates kuv , γxv , and qv , respectively. The term, hy corresponds to the CTL differentiated from memory T cells [41], and should be expressed as $h_M y z_M$, where z_M is the population density of virus-specific memory T cells, which produce activated CTLs with rate $h_M y$. In [41], z_M is assumed to be a constant, and so we have $h = h_M z_M$. We will consider two cases: $h = 0$ and $h \neq 0$; $h = 0$ is due to the absence of memory T cells (that is $z_M = 0$) during the primary effector stage. We will show the relation between the two cases. For simplicity, without loss of the properties of antibodies, we assume $q = 0$ according to [41]. Other experimental parameter values used for studying model (2.22) are given in Table 2.2.

2.5.1 Well-posedness of model (2.22)

Due to physical meaning, negative values of the state variables of system (2.22) are not allowed. Only non-negative initial conditions are considered and the solutions of (2.22) must not be negative. The parameters in (2.22) are all positive due to their biological meaning. Expressing

the solutions of the system (2.22) by variation of constants yields

$$x(t) = x(0) \exp \left[- \int_0^t (d + \beta v(s)) ds \right] + \lambda \int_0^t \exp \left[- \int_s^t (d + \beta v(w)) dw \right] ds, \quad (2.23a)$$

$$y(t) = y(0) \exp \left[- \int_0^t (a + pz(s)) ds \right] + \beta \int_0^t x(s) v(s) \exp \left[- \int_s^t (a + pz(w)) dw \right] ds, \quad (2.23b)$$

$$z(t) = z(0) \exp \left[\int_0^t (cy(s) - b) ds \right] + h \int_0^t y(s) \exp \left[\int_s^t (cy(w) - b) dw \right] ds, \quad (2.23c)$$

$$u(t) = u(0) \exp \left[- \int_0^t (\eta + kv(s)) ds \right] + \xi \int_0^t z(s) \exp \left[- \int_s^t (\eta + kv(w)) dw \right] ds, \quad (2.23d)$$

$$v(t) = v(0) \exp \left[- \int_0^t (ku(s) + \gamma x(s) + q) ds \right] + e \int_0^t y(s) \exp \left[- \int_s^t (ku(w) + \gamma x(w) + q) dw \right] ds. \quad (2.23e)$$

Theorem 2.5.1 *When the initial conditions are taken positive, the solutions of system (2.22) remain positive for $t > 0$. Moreover, they are bounded.*

Proof By the initial condition $x(0) > 0$, it is easy to see from (2.23a) that $x(t) > 0 \forall t > 0$. Next, we show that $y(t) > 0 \forall t > 0$ by an argument of contradiction. Suppose, otherwise, $y(t) < 0$ for some interval $t \in (t_1, t_2)$, $t_1 > 0$. Since $y(0) > 0$, without loss of generality, we may assume t_1 is the first time for y to cross zero, i.e., $y(t) > 0 \forall t \in [0, t_1)$, $y(t_1) = 0$, and $y(t) < 0 \forall t \in (t_1, t_2)$. Thus, from (2.23e) we have $v(t_1) > 0$ due to $v(0) > 0$. On the other hand, it is seen from (2.23b) that $v(t)$ must cross zero to become negative at some $t > t_1$ since $y(t) < 0 \forall t \in (t_1, t_2)$. So let $t = t_3$ be the first time for $v(t)$ to cross zero, i.e., $v(t_3) = 0$ and $v(t) > 0 \forall t \in [t_1, t_3)$. Now, take $t^* = \min(t_2 - \epsilon, t_3)$, satisfying $t^* > t_1$, where $0 < \epsilon \ll 1$. So from the assumption we have $y(t^*) < 0$. However, on the other hand, it follows from (2.23b) that

$$\begin{aligned} y(t^*) &= y(t_1) \exp \left[- \int_{t_1}^{t^*} (a + pz(s)) ds \right] + \beta \int_{t_1}^{t^*} x(s) v(s) \exp \left[- \int_s^{t^*} (a + pz(w)) dw \right] ds \\ &= \beta \int_{t_1}^{t^*} x(s) v(s) \exp \left[- \int_s^{t^*} (a + pz(w)) dw \right] ds > 0, \quad \text{since } v(s) > 0 \forall t \in (t_1, t^*). \end{aligned}$$

leading to a contradiction. Hence $y(t) > 0 \forall t > 0$, and it then follows from (2.23c) and (2.23e) that $z(t) > 0$ and $v(t) > 0 \forall t > 0$. Finally, by the positivity of $z(t)$, (2.23d) gives $u(t) > 0 \forall t > 0$.

It remains to prove that positive solutions of system (2.22) are all bounded. First, consider equation (2.22a), which yields $\dot{x} \leq \lambda - dx$. Given that the exponential functions have negative exponents, we show that $x(t)$ for $t > 0$ is bounded since as $t \rightarrow +\infty$,

$$x(t) \leq \exp \left(- \int_0^t d ds \right) \left[x(0) + \lambda \int_0^t \exp \left(\int_0^s d du \right) ds \right] = x(0) e^{-dt} + \frac{\lambda}{d} (1 - e^{-dt}) \leq \frac{\lambda}{d}.$$

Thus, denote $x_{\max} = \lim_{t \rightarrow +\infty} \sup x(t) = \frac{\lambda}{d}$. It is easy to see $x_{\min} > 0$. Next, we add (2.22a) and (2.22b) together, to obtain $\dot{x} + \dot{y} = \lambda - dx - ay - pyz \leq \lambda - \min(d, a)(x + y)$. Using the same boundedness argument for $x(t)$, we get $x(t) + y(t) \leq \frac{\lambda}{\min(d, a)}$, as $t \rightarrow +\infty$, and thus $y_{\max} = \lim_{t \rightarrow +\infty} \sup y(t) \leq \frac{\lambda}{\min(d, a)}$. Now consider (2.22e), yielding $\dot{v} \leq ey_{\max} - (\gamma x_{\min} + q)v$. Similarly using the same boundedness argument for $x(t)$, we have $\lim_{t \rightarrow +\infty} v(t) \leq \frac{ey_{\max}}{\gamma x_{\min} + q}$. To prove boundedness of $z(t) \forall t > 0$, we use proof by contradiction. Assume $z(t)$ is unbounded, i.e. $\lim_{t \rightarrow +\infty} z(t) \rightarrow +\infty$. Due to positivity of x, y, z, v and boundedness of x, y and v , it follows from (2.22b) that $\dot{y} < 0$ for $z > z^*$, or for $t > t^* > 0$ (z^* and t^* are finite), which implies $\lim_{t \rightarrow +\infty} y(t) \rightarrow 0$. Then, from (2.22c) we have $\dot{z} = (cy - b)z + hy$, so for sufficiently large t , $cy - b < 0$, and so \dot{z}

becomes negative (for some $z > z^*$), implying that z can not increase unboundedly, which is a contradiction. Thus, we denote $z_{\max} = \max\{z(t), t \geq 0\}$. Finally, from equation (2.22d), we have $\dot{u} \leq \xi z_{\max} - \eta u$, which yields $u(t) \leq \frac{\xi z_{\max}}{\eta}$ as $t \rightarrow +\infty$. Hence, we have shown that the solutions of system (2.22) are positive and bounded.

If the initial conditions have some zero elements, it is easy to see from (2.23) that solutions are nonnegative. Hence, system (2.22) is proved to be a well-posed biological model, with nonnegative and bounded solutions.

2.5.2 Equilibrium solutions and their stability

The following results are obtained based on the assumption $q = 0$ [41]. The equilibrium solutions of (2.22) are obtained by simply setting the vector field of (2.22) to zero. There are two equilibrium solutions: the infection-free equilibrium: $E_0 : (x_{e0}, y_{e0}, z_{e0}, u_{e0}, v_{e0}) = (\frac{\lambda}{d}, 0, 0, 0, 0)$, and the infected equilibrium: $E_1 : (x_{e1}, y_{e1}, z_{e1}, u_{e1}, v_{e1})$, where $v_{e1} = \frac{\lambda - dx_{e1}}{\beta x_{e1}}$, $z_{e1} = \frac{u_{e1}(\eta + kv_{e1})}{\xi}$, and $y_{e1} = \frac{v_{e1}(ku_{e1} + \gamma x_{e1})}{e}$. Further, with $h = 10^{-4}$ and other parameter values taken from Table 2.2, u_{e1} can be expressed in terms of x_{e1} , and an equation $F_4(x_{e1}, a) = 0$ is obtained to determine x_{e1} .

The stability analysis for Equilibria E_0 and E_1 is based on the Jacobian matrix of (2.22). Evaluating the Jacobian at the infection-free equilibrium E_0 yields the characteristic polynomial $P_{E_0}(\Psi) = \det[\Psi I - J_0(E_0)] = (\Psi + d)(\Psi + b)(\Psi + \eta)P_{E_{0a}}$, where $P_{E_{0a}} = \Psi^2 + (\frac{\gamma\lambda}{d} + a)\Psi + \frac{(a\gamma - e\beta)\lambda}{d}$. It is easy to see that the stability of E_0 is simply determined by the sign of $(a\gamma - e\beta)$, i.e., E_0 is stable (unstable) if $(a\gamma - e\beta) > 0$ (< 0). In a similar way, we evaluate the Jacobian at E_1 to obtain the 5th-degree characteristic polynomial, from which the fourth Hurwitz determinant Δ_4 can be determined.

2.5.3 Bifurcation analysis for $h \neq 0$

Now we consider possible bifurcations which may occur from the equilibrium solutions E_0 and E_1 . First, for the infection-free equilibrium E_0 , as discussed in the previous subsection, E_0 is stable (unstable) if $(a\gamma - e\beta) > 0$ (< 0). The only possible singularity occurs at the critical point, determined by $a\gamma - e\beta = 0$, at which one eigenvalue of the characteristic polynomial becomes zero (and other four eigenvalues are negative), leading to a static bifurcation. The critical point a_{c0} is solved from $a\gamma - e\beta = 0$ as $a_{c0} = \frac{e\beta}{\gamma}$. Thus, E_0 is stable (unstable) when $a > a_{c0}$ ($a < a_{c0}$), and $x_{c0} = \frac{\lambda}{d}$. With the parameter values in Table 2.2 (with $h = 10^{-4}$), we have $(x_{c0}, a_{c0}) = (0.625, 1.00 \times 10^5)$ which actually holds for both cases $h \neq 0$ and $h = 0$.

As for the infected equilibrium E_1 , one singularity happens when $a_5(x_{e1}, a)$ becomes zero. Thus, the critical point is determined by the equations $a_5(x_{e1}, a) = F_4(x_{e1}, a) = 0$, at which, the characteristic polynomial of E_1 has a zero root. As a result, we obtain one biological meaningful solution, $(x_{c1}, a_{c1}) = (0.625, 1.00 \times 10^5)$. Comparing this critical point with (x_{c0}, a_{c0}) shows that these two critical points are identical, implying that E_0 and E_1 intersect and exchange their stability at this point. Denote this point as $(x_t, a_t) = (0.625, 1.00 \times 10^5)$, which is actually identical for all $h \neq 0$. The bifurcation diagram projected on the a - x plane is shown in Figure 2.11(a). It clearly shows a stability exchange between E_0 and E_1 at the transcritical point.

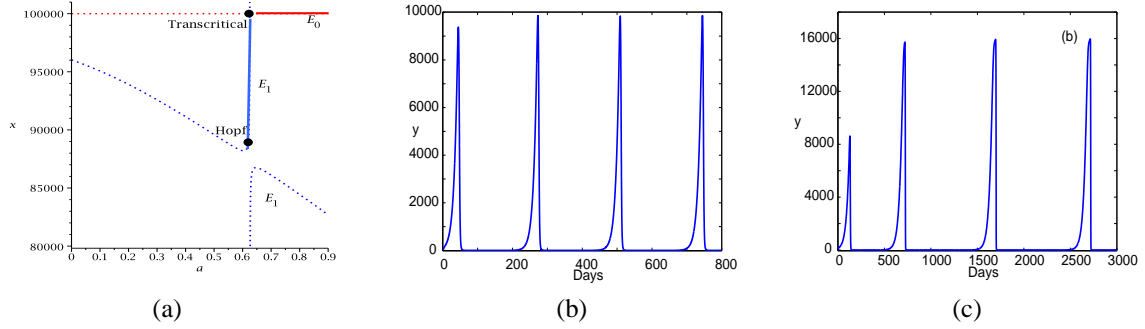


Figure 2.11: Bifurcation diagram and simulated viral blips for system (2.22) with the parameter values taken from Table 2.2 when $a = 0.500$: (a) Bifurcation diagram for $h = 10^{-4}$, with the red and blue lines denoting E_0 and E_1 , respectively, and the dotted and solid lines indicating unstable and stable, respectively (the lower branch of E_1 is biological meaningless, due to negative values in the solution); (b) simulated time history of $y(t)$ for $h = 10^{-4}$; and (c) simulated time history of $y(t)$ for $h = 0$.

Now we turn to possible Hopf bifurcation from E_1 . Since the characteristic polynomial P_{E_1} for E_1 cannot be factorized into polynomials of lesser degree, we will use the Routh-Hurwitz criterion to analyze its stability. The criterion states that the corresponding equilibrium is asymptotically stable if and only if all the Hurwitz determinants are positive [3]. According to [43], the necessary condition for a Hopf bifurcation to occur from the infected equilibrium E_1 is $\Delta_4 = 0$, combined with the equation $F_4(x_{e1}, a) = 0$, since this Hopf bifurcation point is located on the infected equilibrium. Solving these two equations yields a biological meaningful Hopf bifurcation point $(x_H, a_H) \approx (8.85 \times 10^4, 0.617)$. Note that the Hopf bifurcation point is above the turning point $(x_{\text{Turning}}, a_{\text{Turning}}) \approx (8.82 \times 10^4, 0.604)$ in the upper branch of E_1 (see Figure 2.11).

Summarizing the above results shows that the case $h \neq 0$ satisfies all the four conditions in Hypothesis 1 to generate recurrent infection, and indeed recurrence occurs for $a \in (0, a^*)$, where $a^* < a_H$. Moreover, a^* should not be too close to a_H , otherwise the period of limit cycles bifurcating from the Hopf critical point (x_H, a_H) is relatively small. The bifurcation diagram, shown in Figure 2.11(a), indicates that the Hopf critical point a_H is located on the left side of $a = a_t$, where the E_0 is unstable. A simulated time course exhibiting recurrent infection is depicted in Figure 2.11(b).

2.5.4 Bifurcation analysis for $h \rightarrow 0^+$

Now we turn to consider the special case, $h = 0$. It is easy to observe from equation (2.22c) that the solutions of system (2.22) are discontinuous at $h = 0$. Therefore, to have continuity, we should regard the special case $h = 0$, as the limiting case: $h \rightarrow 0^+$. In calculation, we choose a small enough value of h (e.g., $h = 10^{-8}$) and then do the same analysis as done for the case $h \neq 0$. We also get two equilibrium solutions – the infection-free equilibrium E_0 and the infected equilibrium E_1 – a transcritical bifurcation which occurs at the intersection of the two equilibria, a Hopf bifurcation emerging from the infected equilibrium E_1 , and large oscillations occurring near the transcritical critical point on the unstable side of the Hopf critical point, given by $(x_H, a_H) \approx (8.7511 \times 10^4, 0.6249)$. The bifurcation diagram for this case ($h = 10^{-8}$)

is similar to that shown in Figure 2.11(a), except that the two branches of E_1 are much closer, indicating that the Hopf bifurcation point moves down towards the turning point in the upper branch of E_1 , which is also moving down. This implies that one branch of solution E_1 becomes almost a vertical line as $h \rightarrow 0^+$, and the Hopf critical point coincides with the turning point.

For $h = 0$, we treat it as the limit: $h \rightarrow 0^+$. The seemingly vertical line in the bifurcation diagram for $h = 0$ disappears, clearly showing the discontinuity of E_1 at $h = 0$. This causes difficulty in bifurcation analysis. However, if we treat the case $h = 0$ as the limiting case $h \rightarrow 0^+$, the solution E_1 continuously depends on h , and the bifurcation diagram becomes smooth. Therefore, we can still use our theory to explain the occurrence of blips for the case $h = 0$, as shown in Figure 2.11(c). In fact, more precisely, when $h = 0$, a Bogdanov-Takens bifurcation (double-zero singularity) occurs at the point where the Hopf and turning points are merged. This is a codimension-2 bifurcation point, which in general needs two unfolding (bifurcation) parameters to give a complete local dynamical analysis. In our case, the variation of the single parameter α can be considered as a line (ray) in the two-parameter plane. It is well known that in the vicinity of a Bogdanov-Takens bifurcation point, there exists Hopf bifurcation and homoclinic bifurcation. Therefore, the motion generated near the codimension-2 bifurcation point may be due to either the Hopf or homoclinic bifurcation. With respect to the blips phenomenon, the motion is large (not the small motions bifurcating from Hopf or homoclinic bifurcations) and is a globally persistent motion, and so it is not directly related to the Hopf or homoclinic bifurcations. In other words, we are more interested in possible large motions near the transcritical point.

2.6 Conclusion and discussion

In this paper, the problem of recurrent infection (viral blips) in in-host infection models is studied via the qualitative analysis of dynamical systems. A 4-dimensional HIV antioxidant-therapy model [39], which produces viral blips, is investigated in detail using bifurcation theory. A hypothesis consisting of four conditions for the emergence of viral blips is proposed. These conditions describe two equilibrium solutions which intersect at a transcritical bifurcation point, with a Hopf bifurcation which originates from the equilibrium solution. Under these conditions, blips appear for values of the bifurcation parameter near the transcritical point, where equilibrium solutions are unstable.

Guided by the proposed hypothesis, we propose several simpler in-host infection models that can also generate viral blips. We develop a 3-dimensional in-host model with an increasing, saturating infection rate similar to the HIV antioxidant-therapy model, and show that all four conditions in the hypothesis are satisfied, leading to blips. Further, stability and bifurcation analyses determine all possible regions in parameter space where blips may occur. We then investigate an even simpler 2-dimensional in-host model. This very simple model can also exhibit blips, as long as the infection rate is an increasing, saturating function of infected cell density. We also apply the hypothesis to study a standard HIV model with CTL response [28] and find blips by using an increasing, saturating infection rate function.

Overall, our results suggest that simple ODE models of in-host infection dynamics are sufficient to describe transient periods of high viral replication, separated by long periods of quiescence. Rather than needing an exogenous trigger such as stochastic stimulation of the

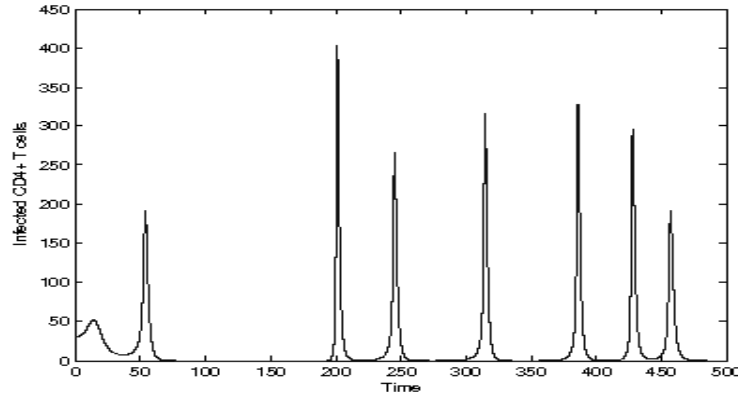


Figure 2.12: Simulated viral blips of system (2.1) with varying amplitude and frequency when using a time-varying function $\alpha(t) = \alpha_T + [-0.31 + 0.3e^{-3\cos(t/50)} \cos(t/100)] \times 10^{13}$, where $\alpha_T = 4.58 \times 10^{13}$ is the transcritical bifurcation value.

immune system, the natural dynamics of such systems may be sufficiently rich, in many cases, to exhibit viral blips. One key to obtaining this rich behaviour is to propose an infection rate which increases, but saturates, with the extent of the infection. This is a natural assumption if the infection itself (high density of infected target cells) makes the host more vulnerable to further infection. Such an assumption is certainly natural for HIV, where the primary target cells are T lymphocytes.

All the simulated oscillating motions and blips presented in this paper show constant amplitudes and frequencies. This is because all parameter values are fixed in these simulations. We note, however, that nonlinear, deterministic systems can indeed generate oscillations with varying amplitudes and phases, called “amplitude modulation” and “frequency modulation” due to nonlinearity. This can be seen from the equation (2.13), where both amplitude and phase are functions of the parameter μ . Since in reality parameters are not constant, time-varying parameters can be seen as analogous to the variation due to random perturbations in stochastic models. Although deterministic models with fixed parameter values cannot generate varying amplitude and phase, deterministic models can generate such variation if the system is nonlinear and some parameters vary with time. For example, Figure 2.12 shows the result of changing the fixed α used in Figure 2.3(c) to a time-varying deterministic function, clearly demonstrating that a deterministic model can generate blips of varying magnitude, frequency and duration.

We note that mathematically, a system of delay differential equations (DDEs) could also generate oscillatory behaviours similar to viral blips. However in this case, the inherent delay would need to be of the same order as the interval between blips, that is, on the order of several months. Since it is difficult to suggest a physiological or immunological process that would impose a delay of this magnitude, it seems unlikely that DDEs are the most natural approach for modeling viral blips.

While we are able to show that linear or constant infection rates do not lead to blips in the 2-, 3- or 4-dimensional models we have studied, further study of a 5-dimensional immunological model reveals that a system with a constant infection rate can also generate blips. This

suggests that the use of an increasing, saturating infection rate function is not necessary, but is effective in low-dimensional models. The results presented here provide a useful tool for the mathematical study of viral blips or other examples of recurrent infection. The conditions in our hypothesis may also be used or generalized to study recurrent phenomena in other physical systems.

2.7 References

- [1] R. M. Anderson and R. M. May. Population biology of infectious diseases: Part I. *Nature*, 280(5721):361–367, 1979.
- [2] I. Boldogh, T. Albrecht, and D. D. Porter. *Medical Microbiology*, chapter 46. University of Texas Medical Branch at Galveston, 4th edition, 1996.
- [3] B. Chan and P. Yu. Bifurcation analysis in a model of cytotoxic T-lymphocyte response to viral infection. *Nonlinear Analysis: Real World Applications*, 13(1):64–77, 2012.
- [4] A. C. Collier, R. W. Coombs, D. A. Schoenfeld, R. L. Bassett, J. Timpone, A. Baruch, M. Jones, K. Facey, C. Whitacre, V. J. McAuliffe, H. M. Friedman, T. C. Merigan, R. C. Reichman, C. Hooper, and L. Corey. Treatment of human immunodeficiency virus infection with Saquinavir, Zidovudine, and Zalcitabine. *New England Journal of Medicine*, 334(16):1011–1018, 1996.
- [5] J. M. Conway and D. Coombs. A stochastic model of latently infected cell reactivation and viral blip generation in treated HIV patients. *PLoS Computational Biology*, 7(4), 2011.
- [6] R. J. De Boer and A. S. Perelson. Target cell limited and immune control models of HIV infection: A comparison. *Journal of Theoretical Biology*, 190(3):201–214, 1998.
- [7] S. Devadas, L. Zaritskaya, S. G. Rhee, L. Oberley, and M. S. Williams. Discrete generation of superoxide and hydrogen peroxide by T cell receptor stimulation: selective regulation of mitogen-activated protein kinase activation and fas ligand expression. *Journal of Experimental Medicine*, 195(1):59–70, 2002.
- [8] G. Dornadula, H. Zhang, B. VanUitert, J. Stern, L. Livornese Jr., M. J. Ingerman, J. Witek, R. J. Kedanis, J. Natkin, J. DeSimone, and R. J. Pomerantz. Residual HIV-1 RNA in blood plasma of patients taking suppressive highly active antiretroviral therapy. *Journal of the American Medical Association*, 282(17):1627–1632, 1999.
- [9] N. M. Ferguson, F. de Wolf, A. C. Ghani, C. Fraser, C. A. Donnelly, P. Reiss, J. M. A. Lange, S. A. Danner, G. P. Garnett, J. Goudsmit, and R. M. Anderson. Antigen-driven CD4⁺ T cell and HIV-1 dynamics: Residual viral replication under highly active antiretroviral therapy. *Proceedings of the National Academy of Sciences the United States of America*, 96(26):15167–15172, 1999.

- [10] C. Fraser, N. M. Ferguson, and R. M. Anderson. Quantification of intrinsic residual viral replication in treated HIV-infected patients. *PNAS*, 98(26):15167–15172, 2001.
- [11] C. Fraser, N. M. Ferguson, F. de Wolf, and R. M. Anderson. The role of antigenic stimulation and cytotoxic T cell activity in regulating the long-term immunopathogenesis of HIV: mechanisms and clinical implications. *Proceedings of the Royal Society B: Biological Sciences*, 268(1481):2085–2095, 2001.
- [12] I. C. H. Fung, M. Gambhir, A. van Sighem, F. de Wolf, and G. P. Garnett. The clinical interpretation of viral blips in HIV patients receiving antiviral treatment: Are we ready to infer poor adherence? *Journal of Acquired Immune Deficiency Syndromes*, 60(1):5–11, 2012.
- [13] M. Garland and W. W. Fawzi. Antioxidants and progression of human immunodeficiency virus (HIV) disease. *Nutrition Research*, 19(8):1259–1276, 1999.
- [14] N. J. Garretta, V. Apeaa, A. Noria, I. Ushiro-Lumbb, A. R. Oliverb, G. Bailya, and D. A. Clarkb. Comparison of the rate and size of HIV-1 viral load blips with Roche COBAS TaqMan HIV-1 versions 1.0 and 2.0 and implications for patient management. *Journal of Clinical Virology*, 53(4):354–355, 2012.
- [15] L. Gil, G. Martinez, I. Gonzalez, A. Tarinas, A. Alvarez, A. Giuliani, R. Molina, R. Tapanes, J. Perez, and O. D. Leon. Contribution to characterization of oxidative stress in HIV/AIDS patients. *Pharmacological Research*, 47(3):217–224, 2003.
- [16] G. Gloire, S. Legrand-Poels, and J. Piette. NF- κ B activation by reactive oxygen species: fifteen years later. *Biochemical Pharmacology*, 72(11):1493–1505, 2006.
- [17] J. T. Grennan, M. R. Loutfy, D. Su, P. R. Harrigan, C. Cooper, M. Klein, N. Machouf, J. S. G. Montaner, S. Rourke, C. Tsoukas, B. Hogg, J. Raboud, and the CANOC Collaboration. Magnitude of virologic blips is associated with a higher risk for virologic rebound in HIV-infected individuals: A recurrent events analysis. *Journal of Infectious Diseases*, 205(8):1230–1238, 2012.
- [18] S. Gropper, J. Smith, and J. Groff. *Advanced Nutrition and Human Metabolism*. Cengage Learning, 2008, 5th edition, 2009.
- [19] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Applied Mathematical Sciences. Springer-Verlag, New York, 4th edition, 1993.
- [20] D. A. Hildeman. Regulation of T-cell apoptosis by reactive oxygen species. *Free Radical Biology and Medicine*, 36(12):1496–1504, 2004.
- [21] D. Hinrichsen and A. J. Pritchard. *Mathematical Systems Theory I: Modelling, State Space Analysis, Stability and Robustness*, volume 48 of *Texts in Applied Mathematics*. Springer, 2nd edition, 2005.

- [22] N. Israel and M. A. Gougerot-Pocidalò. Oxidative stress in human immunodeficiency virus infection. *Cell Molec Life Sciences*, 53(11-12):864–870, 1997.
- [23] L. E. Jones and A. S. Perelson. Opportunistic infection as a cause of transient viremia in chronically infected HIV patients under treatment with HAART. *Bulletin of Mathematical Biology*, 67(6):1227–1251, 2005.
- [24] L. E. Jones and A. S. Perelson. Transient viremia, plasma viral load, and reservoir replenishment in HIV infected patients on antiretroviral therapy. *Journal of Acquired Immune Deficiency Syndromes*, 45(5):483–493, 2007. [PubMed:17496565].
- [25] H. M. Lander. An essential role for free radicals and derived species in signal transduction. *The FASEB Journal*, 11(2):118–124, 1997.
- [26] N. Li and M. Karin. Is NF- κ B the sensor of oxidative stress? *The FASEB Journal*, 13(10):1137–1143, 1999.
- [27] M. A. Nowak and C. R. M. Bangham. Population dynamics of immune responses to persistent viruses. *Science*, 272(5258):74–79, 1996.
- [28] M. A. Nowak and R. M. May. *Virus Dynamics*. Oxford University Press, New York, 2000.
- [29] G. W. Pace and C. D. Leaf. The role of oxidative stress in HIV disease. *Free Radical Biology and Medicine*, 19(4):523–528, 1995.
- [30] S. Palmer, F. Maldarelli, A. Wiegand, B. Bernstein, G. J. Hanna, S. C. Brun, D. J. Kempf, J. W. Mellors, J. M. Coffin, and M. S. King. Low-level viremia persists for at least 7 years in patients on suppressive antiretroviral therapy. *Proceedings of the National Academy of Sciences*, 105(10):3879–3884, 2008.
- [31] S. Palmer, A. P. Wiegand, F. Maldarelli, H. Bazmi, J. M. Mican, M. Polis, R. L. Dewar, A. Planta, S. Liu, J. A. Metcalf, J. W. Mellors, and J. M. Coffin. New real-time reverse transcriptase-initiated PCR assay with single-copy sensitivity for human immunodeficiency virus type 1 RNA in plasma. *Journal of Clinical Microbiology*, 41(10):4531–4536, 2003.
- [32] A. S. Perelson, P. Essunger, Y. Cao, M. Vesanen, A. Hurley, K. Saksela, M. Markowitz, and D. D. Ho. Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature*, 387(6629):188–191, 1997.
- [33] L. Rong, Z. Feng, and A. Perelson. *Mathematical Modeling of Biosystems*, (Eds) Mondaini, R. P. and Pardalos, P. M., chapter Mathematical modeling of HIV-1 infection and drug therapy, pages 77–87. Springer-Verlag, 2008.
- [34] L. Rong and A. S. Perelson. Asymmetric division of activated latently infected cells may explain the decay kinetics of the HIV-1 latent reservoir and intermittent viral blips. *Mathematical Biosciences*, 217(1):77–87, 2009.

- [35] L. Rong and A. S. Perelson. Modeling latently infected cell activation: viral and latent reservoir persistence, and viral blips in HIV-infected patients on potent therapy. *PLoS Computational Biology*, 5(10):e1000533, 2009.
- [36] K. B. Schwarz. Oxidative stress during viral infection: a review. *Free Radical Biology and Medicine*, 21(5):641–649, 1996.
- [37] C. B. Stephenson, G. S. Marquis, S. D. Douglas, and C. M. Wilson. Immune activation and oxidative damage in HIV-positive and HIV negative adolescents. *Journal of Acquired Immune Deficiency Syndromes*, 38(2):180–190, 2005.
- [38] C. B. Stephenson, G. S. Marquis, R. A. Jacob, L. A. Kruzich, S. D. Douglas, and C. M. Wilson. Vitamins C and E in adolescents and young adults with HIV infection. *American Journal of Clinical Nutrition*, 83(4):870–879, 2006.
- [39] R. D. van Gaalen and L. M. Wahl. Reconciling conflicting clinical studies of antioxidant supplementation as HIV therapy: a mathematical approach. *BMC Public Health*, 9(Suppl. 1):1–18, 2009.
- [40] L. P. Villarreal, V. R. Defilippis, and K. A. Gottlieb. Acute and persistent viral life strategies and their relationship to emerging diseases. *Virology*, 272(1):1–6, 2000.
- [41] W. Yao, L. Hertel, and L. M. Wahl. Dynamics of recurrent viral infection. *Proceedings of the Royal Society-Biological Sciences*, 273(1598):2193–2199, 2006.
- [42] P. Yu. Computation of normal forms via a perturbation technique. *Journal of Sound and Vibration*, 211(1):19–38, 1998.
- [43] P. Yu. Closed-form conditions of bifurcation points for general differential equations. *International Journal of Bifurcation and Chaos*, 15(4):1467–1483, 2005.
- [44] P. Yu and K. Huseyin. A perturbation analysis of interactive static and dynamic bifurcations. *IEEE Transactions on Automatic Control*, 33(1):28–41, 1988.

Chapter 3

Modelling and Analysis of Recurrent Autoimmune Disease

3.1 Introduction

The adaptive immune system consists of a set of highly specialized cells and processes that can limit or eradicate the growth of foreign pathogens. Normally, the immune system must be able to mount responses against pathogens that invade the host, but avoid attacking the organism's own tissues; when this discrimination fails, the result is autoimmunity. Autoimmune diseases are often chronic and debilitating. They affect 50 million (or one in five) Americans, but are more common in women (75 percent of cases), according to the American Autoimmune Related Diseases Association [2]. In fact, autoimmune diseases are among the main causes of death of young and middle-aged women in developed countries [9]. Evidence is also mounting that the prevalence of autoimmune disease is increasing: for example, a 3% global increase in type 1 diabetes per year has been reported [26]. Although health care costs related to autoimmune diseases amount to over billion dollars each year in the the U.S.A. alone, patients are still suffering from misdiagnosis and delayed diagnosis due to a lack of understanding of autoimmune disease. These facts illustrate the vital need to focus further research on all autoimmune diseases.

To address autoimmune disease in a mathematical model, we first outline in brief the normal function of the immune system. The cells of the adaptive immune system are T and B lymphocytes: B cells are involved in 'humoral immune responses', while T cells play a large role in the cell-mediated immune responses. Here, we focus on the latter response. Initiation of an adaptive immune response starts when immature dendritic cells (DCs), which are the most important professional antigen presenting cells (pAPCs), settle at a site of infection or inflammation, become activated and undergo maturation. Simultaneously, naive conventional T cells, each bearing a specific antigen receptor, constantly circulate through the peripheral lymphoid tissues, browsing many DCs as they carry out brief contacts, and receiving two signals: discrimination of the antigen presented by DCs and interplay with co-stimulatory molecules on the same DCs. After making a stable interaction with DCs presenting their cognate antigen, naive T cells can be activated and proliferate into effector T cells. The proliferation phase is significant and driven by cytokine interleukin-2 (IL-2), which can be produced by active

conventional T cells themselves and from other sources as well.

Central tolerance is the main mechanism which allows the immune system to avoid mounting a response against the organism's own tissues. In this process, auto-reactive T cells, which have antigen receptors specific to self antigens, are deleted during lymphocyte development in the thymus. Nevertheless, the T cells that leave the thymus are relatively but not absolutely safe. A large body of research has demonstrated that some auto-reactive T cells are present in the periphery of under normal conditions [39]. In this case, peripheral tolerance is established after T cells mature and migrate into the periphery to prevent auto-reactive T cells from directing an immune response toward self-antigens. One mechanism of peripheral tolerance is the population of regulatory T (T_{Reg}) cells.

Regulatory T cells are a subpopulation of $CD4^+$ T cells, that modulate the immune system, preventing the expansion of auto-reactive T cells, and subsequent autoimmune disease [30]. Evidence [29, 4, 18] has shown that human T_{Reg} cells are phenotypically heterogeneous. Most thymus-derived T_{Reg} cells found in the periphery are naive T_{Reg} cells [18, 22, 15], which have not experienced T cell receptor (TCR) stimulation-mediated maturation, and are in a quiescent stage, resistant to apoptosis. Like naive conventional T cells, in order to participate in an immune response, naive T_{Reg} cells require activation by antigen on pAPCs and possible co-stimulation [1, 21]. IL-2 seems to be a necessary factor [10, 33, 31] for T_{Reg} cell proliferation. Activated conventional T cells are believed to be the main source of IL-2 [12, 41], although there also exist other IL-2 sources, such as DCs. Following activation, naive T_{Reg} cells become 'effector' natural T_{Reg} (nT_{Reg}) cells, which have potent suppressive activity.

Recently, a new subset of effector nT_{Reg} cells has been discovered experimentally [5, 27]. This subset of cells have further matured to become terminally differentiated suppressors, which show more efficient suppression, but have a shorter lifespan, than nT_{Reg} cells. Phenotypic analysis has demonstrated that the expression of the cell surface receptor HLA-DR in nT_{Reg} cells is heterogeneous [28], and distinguishes this terminally differentiated subpopulation; in particular $HLA-DR^+$ T_{Reg} cells suppress proliferation of conventional T cells more rapidly than $HLA-DR^-$ T_{Reg} cells. It is believed that activation and expansion of $HLA-DR^-$ effector nT_{Reg} cells provoke the generation of this subset of $HLA-DR^+$ T_{Reg} cells [5].

Despite these multi-layer barriers, self-tolerance mechanisms fail occasionally. Although the activity of auto-reactive T cells in humans is not understood completely, research in non-human primates has indicated that these cells in the periphery can be activated and may provoke a T-cell-mediated attack against self-determinants [37], causing autoimmune disorders. For example, when auto-reactive T cells attack the central nervous system [37], acute focal inflammation may cause a relapse of symptoms in multiple sclerosis [38]. T_{Reg} cells are capable of limiting these attacks, and their deficiency can lead to fatal autoimmune disease which affects multiple organs in mice [6, 14], and human beings [34, 25].

Autoimmune diseases are often chronic, requiring lifelong care and monitoring, despite the fact that symptoms may disappear occasionally. Many autoimmune diseases are characterized by recurrence, that is, disease relapses (return of symptoms) followed by remittance (absence of symptoms, possibly for a long period). In several autoimmune diseases, this relapse-remission behaviour occurs even in the absence of treatment, for example in multifocal osteomyelitis [16, 19], eczema [13], subacute discoid lupus erythematosus [23], and psoriasis [11]. In fact, the subtypes of some diseases are clinically classified based on the patterns of this recurrent behaviour [38]. Therefore, an improved understanding of recurrent dynamics in autoimmune

disease is crucial to promote correct diagnosis, patient management and treatment decisions.

Recently, the relapse-remission behavior of multiple sclerosis was studied using a stochastic differential equation model developed by Mendizabal *et al.* [36]. The authors investigated cross-regulation interactions, modeled as Hill functions, between regulatory and auto-reactive effector T cells. A predator-prey system is adopted in this paper, in where auto-reactive effector T cells act as prey and T_{Reg} cells as predators. The resting auto-reactive effector T cell and resting T_{Reg} cell populations are introduced to the deterministic predator-prey model using stochastic pulse trains [40], which model the probabilistic influx of resting cells. This predator-prey model with stochasticity generates the characteristic relapse-remission behavior of multiple sclerosis. The paper concludes that weakness in the negative feedback between effector and regulatory T cells may allow the immune system to generate the typical recurrent dynamics of autoimmune disease without the need for exogenous triggers.

Recent models introduced by Alexander and Wahl [1] capture the intrinsic feedback cycle of autoimmunity, in which professional antigen presenting cells (pAPCs) present self-antigen, eliciting self-reactive effector T cells, which in turn attack host tissues. The damage to host tissue results in increased concentrations of self-antigen, activating further pAPCs. This cycle is kept in check by the actions of T_{Reg} cells, which limit the self-reactive immune response via several putative mechanisms. These models exhibit equilibria corresponding to tolerance and autoimmunity, but bistability is not observed. Instead, a branching process was used to demonstrate that from identical starting conditions, states of immune tolerance or intolerance could be reached probabilistically. Although this set of related models offers a general approach to autoimmunity and the role of T_{Reg} cells, they do not capture the recurrent behavior which characterizes many autoimmune diseases.

Following the recent experimental discovery of HLA-DR⁺ T_{Reg} cells described above, we chose to expand the model of Wahl and Alexander to include this new class of potentially suppressive cells. In some parameter regimes, we observed numerically that the expanded model exhibits long periods of self tolerance, punctuated by brief episodes of disease recurrence, deterministic dynamics reminiscent of our recent investigations of viral blips in in-host infection models [45]. In these infection models, relapse-remission behaviour may in some cases arise simply from the nonlinear dynamics of the underlying dynamical system, in the absence of stochasticity, therapy, or other trigger mechanisms [42, 35]. By taking advantage of dynamical systems theory, we recently proposed four conditions which guarantee recurrent behavior in deterministic viral infection models [45]. Given the importance of recurrence to autoimmune disease, here we apply a similar approach to gain an analytical understanding of the dynamical features underlying recurrence in the autoimmune model.

The rest of the paper is organized as follows. In Section 2, we introduce two established models [1] describing autoimmune disease. We demonstrate that these two models do not have Hopf bifurcations, and so cannot exhibit the oscillatory behavior which underlies recurrence. Based on recent experimental findings, we introduce the new T_{Reg} subtype and establish a new model. In Section 3, we first prove that the new model is well-posed, and then perform mathematical analysis to find equilibrium solutions and determine their local and global stability. By choosing proper bifurcation parameters, we also identify the transcritical and Hopf bifurcation critical points, showing that the new model should display the recurrent dynamics characteristic of many autoimmune diseases. Further, by applying center manifold theory and normal form theory, we find approximate solutions of the limit cycles and determine their stability. Then, in

Section 4, we use numerical simulation to verify the analytical predictions obtained in Section 4. Moreover, a comparison between the analytical and numerical results for the Hopf bifurcation are given in this section. In order to identify key factors in the mechanism of recurrence, in Section 5 we perform model reduction under a quasi-steady state assumption. This is achieved by reducing the number of state variables and parameters, and also by a rescaling of the time variable. Then, we prove that the original and reduced models exhibit the same dynamical behavior as long as the parameter values are chosen properly. Based on the reduced model, three bifurcation parameters are used to classify the parameter ranges for which recurrence exists. Furthermore, we show that there do not exist homoclinic orbits in either the original or the reduced models, and so the recurrence phenomenon either comes from Hopf bifurcation or is due to persistent oscillations. We conclude with a brief discussion of these results in Section 6.

3.2 Model development

Following recent experimental findings, we sought to introduce terminally differentiated regulatory T cells as an explicit variable into models established by Alexander and Wahl [1]. Their models consider two suppressive mechanisms enacted by T_{Reg} cells. The first of these is the direct suppression of pAPCs by T_{Reg} cells, effectively removing pAPCs from the system. The corresponding model is given by

$$\begin{aligned}\dot{A} &= f\tilde{\nu}G - (\sigma_1 R_n + b_1)A - \mu_A A, \\ \dot{R}_n &= (\pi_1 E + \beta)A - \mu_n R_n, \\ \dot{E} &= \lambda_E A - \mu_E E, \\ \dot{G} &= \gamma E - \tilde{\nu}G - \mu_G G,\end{aligned}\tag{3.1}$$

where the variables A , R_n , E , G represent the populations of mature pAPCs, active nT_{Reg} cells, active auto-reactive effector T cells, and the particular self-antigen of interest. All cell populations are specific for a given self-antigen. Parameter definitions and their numerical values are listed in Table 3.1; meaningful numerical values were carefully chosen in [1] with extended reference to the primary literature. Model (3.1) assumes that pAPCs undergo maturation at a rate of $f\tilde{\nu}G$, while during this process the antigen uptake rate is $\tilde{\nu}G$. The activated auto-reactive effector T cells (E) are produced at a rate of $\lambda_E A$ by resting T cells through an interaction with mature pAPCs (A). After activation, auto-reactive effector T cells (E) can produce IL-2, which is required for T_{Reg} cell proliferation, while other IL-2 sources also exist. Thus nT_{Reg} cells are activated at a rate of $(\pi_1 E + \beta)A$, where $\pi_1 E$ represents IL-2 produced by active auto-reactive effector T cells (E), and β represents background sources of IL-2. nT_{Reg} cells (R_n) then suppress pAPCs (A) at a rate of $\sigma_1 R_n A$, while b_1 represents a level of non-specific background suppression. Auto-reactive effector T cells (E) attack the host tissues, causing the release of self-antigen at a rate of γE , which in turn triggers the maturation of pAPCs, and thus initiates a new cycle of autoimmunity. Here, the death/clearance rates of the populations A , R_n , E , and G are μ_A , μ_n , μ_E , and μ_G respectively.

Another suppressive mechanism is considered in isolation in [1], that is, nT_{Reg} cells may

directly reduce the auto-reactive effector T cell population. This model is described by

$$\begin{aligned}\dot{A} &= f\tilde{v}G - \mu_A A, \\ \dot{R}_n &= (\pi_3 E + \beta)A - \mu_n R_n, \\ \dot{E} &= \lambda_E A - (\sigma_3 R_n + b_3)E - \mu_E E, \\ \dot{G} &= \gamma E - \tilde{v}G - \mu_G G,\end{aligned}\tag{3.2}$$

where the auto-reactive effector T cells are suppressed by nT_{Reg} cells and background suppression at a rate of $(\sigma_3 R_n + b_3)E$. Other terms have the same meaning as the counterparts in model (3.1).

3.2.1 No recurrence in models (3.1) and (3.2)

Since the main purpose of this paper is to study recurrence in autoimmune models, we first want to ask if the above two models (3.1) and (3.2) can exhibit this behavior. According to the Hypothesis given in [45], a Hopf bifurcation is a necessary condition for recurrence. In this section, we will show that the two models (3.1) and (3.2) indeed do not have a Hopf bifurcation. For simplicity, we only briefly outline the proof for model (3.1). Similarly, one can prove this for model (3.2).

First, as usual, we can show that the solutions of model (3.1) are non-negative if the initial conditions are non-negative, and all solutions are bounded. Further, we show that model (3.1) has two equilibrium solutions: one of them is the trivial equilibrium,

$$E_0 : A_0 = R_{n0} = E_0 = G_0 = 0;$$

and the other is the non-trivial equilibrium,

$$E_1 : R_{n1} = \frac{f\tilde{v}\gamma\lambda_E - \mu_E(b_1 + \mu_A)(\tilde{v} + \mu_G)}{\sigma_1\mu_E(\tilde{v} + \mu_G)}, \quad E_1 = \frac{\lambda_E}{\mu_E} A_1, \quad G_1 = \frac{\gamma\lambda_E}{\mu_E(\tilde{v} + \mu_G)} A_1,$$

where

$$A_1 = -\frac{\beta\mu_E}{2\pi_1\lambda_E} + \sqrt{\left(\frac{\beta\mu_E}{2\pi_1\lambda_E}\right)^2 + \frac{f\tilde{v}\gamma\lambda_E - \mu_E(b_1 + \mu_A)(\tilde{v} + \mu_G)}{\sigma_1\pi_1\lambda_E}},\tag{3.3}$$

for $f\tilde{v}\gamma\lambda_E - (b_1 + \mu_A)\mu_E(\tilde{v} + \mu_G) > 0$, and thus $A_1 > 0$.

Then, the stability of E_0 and E_1 can be determined from the linearized system of (3.1) and its characteristic polynomials, associated with these two equilibria. The characteristic polynomial for E_0 is obtained as $P_0(L) = (L + \mu_n)(L^3 + a_{01}L^2 + a_{02}L + a_{03})$, where

$$\begin{aligned}a_{01} &= b_1 + \mu_A + \mu_E + \tilde{v} + \mu_G, \\ a_{02} &= \mu_E(b_1 + \mu_A + \tilde{v} + \mu_G) + (b_1 + \mu_A + \tilde{v} + \mu_G), \\ a_{03} &= \mu_E(b_1 + \mu_A)(\tilde{v} + \mu_G) - f\tilde{v}\gamma\lambda_E.\end{aligned}$$

Further, it is easy to show that

$$\Delta_{02} = a_{01}a_{02} - a_{03} = (b_1 + \mu_A)(\mu_E + \tilde{v} + \mu_G)^2 + (\mu_E + \tilde{v} + \mu_G)[\mu_E(\tilde{v} + \mu_G) + (b_1 + \mu_A)^2] + f\tilde{v}\gamma\lambda_E > 0.$$

Thus, according to the Routh-Hurwitz criterion we can conclude that the equilibrium E_0 is stable (unstable) if $\mu_E(b_1 + \mu_A)(\tilde{v} + \mu_G) - f\tilde{v}\gamma\lambda_E$ is > 0 (< 0). The only possible bifurcation

from E_0 is a static bifurcation which occurs at the critical point, determined by $f\tilde{v}\gamma\lambda_E = \mu_E(b_1 + \mu_A)(\tilde{v} + \mu_G)$. Note that when $\mu_E(b_1 + \mu_A)(\tilde{v} + \mu_G) - f\tilde{v}\gamma\lambda_E > 0$, the equilibrium E_1 does not exist.

Next, similarly we can discuss the stability of E_1 . The characteristic polynomial, associated with E_1 is given by $P_1(L) = L^4 + a_{11}L^3 + a_{12}L^2 + a_{13}L + a_{14}$, where

$$\begin{aligned} a_{11} &= \frac{1}{\mu_E(\tilde{v} + \mu_G)} \left[f\tilde{v}\gamma\lambda_E + \mu_E(\tilde{v} + \mu_G)(\mu_u + \mu_E + \tilde{v} + \mu_G) \right], \\ a_{12} &= \frac{1}{\mu_E(\tilde{v} + \mu_G)} \left\{ \sigma_1(\tilde{v} + \mu_G)(\pi_1\lambda_E A_1 + \beta\mu_E) A_1 + f\tilde{v}\gamma\lambda_E(\mu_n + \mu_E + \tilde{v} + \mu_G) \right. \\ &\quad \left. + \mu_E(\tilde{v} + \mu_G) \left[\mu_E\mu_n + (\mu_n + \mu_E)(\tilde{v} + \mu_G) \right] \right\}, \\ a_{13} &= \frac{1}{\mu_E(\tilde{v} + \mu_G)} \left\{ \sigma_1(\tilde{v} + \mu_G) \left[\lambda_E\pi_1(\tilde{v} + \mu_G + 2\mu_E) A_1 + \beta\mu_E(\mu_E + \tilde{v} + \mu_G) \right] A_1 \right. \\ &\quad \left. + \mu_n \left[\mu_E^2(\tilde{v} + \mu_G)^2 + (\mu_E + \tilde{v} + \mu_G)f\tilde{v}\gamma\lambda_E \right] \right\}, \\ a_{14} &= \sigma_1(\tilde{v} + \mu_G)(2\pi_1\lambda_E A_1 + \beta\mu_E) A_1, \end{aligned}$$

where A_1 is given in (3.3). It is easy to see that $a_{1i} > 0$, $i = 1, 2, 3, 4$. Moreover, we can show that

$$\begin{aligned} \Delta_{12} &= a_{11}a_{12} - a_{13} \\ &= \frac{1}{\mu_E^2(\tilde{v} + \mu_G)^2} \left\{ \mu_E^3(\tilde{v} + \mu_G)^2 \sigma_1 \beta A_1 \right. \\ &\quad \left. + \mu_E^2(\tilde{v} + \mu_G)^2 \left[\mu_n(\mu_E + \tilde{v} + \mu_G)^2 + (\tilde{v} + \mu_G)(\mu_E^2 + \mu_n^2 + \mu_E(\tilde{v} + \mu_G)) \right. \right. \\ &\quad \left. \left. + \mu_E\mu_n(\mu_n + b_1 + \mu_A) \right] \right. \\ &\quad \left. + f\tilde{v}\gamma\mu_E(\tilde{v} + \mu_G)\lambda_E \left[(\mu_n + \tilde{v} + \mu_G)^2 + \mu_E(3(\tilde{v} + \mu_G) + \mu_E + \mu_n) \right] \right. \\ &\quad \left. + (f\tilde{v}\gamma\lambda_E)^2(\mu_n + \mu_E + \tilde{v} + \mu_G) + \mu_n \left[f\tilde{v}\gamma\lambda_E + \mu_n\mu_E(\tilde{v} + \mu_G) \right] \left[f\tilde{v}\gamma\lambda_E - \mu_E(b_1 + \mu_A)(\tilde{v} + \mu_G) \right] \right\} \\ &> 0, \end{aligned}$$

due to $f\tilde{v}\gamma\lambda_E > \mu_E(b_1 + \mu_A)(\tilde{v} + \mu_G)$, as well as $\Delta_{13} = (a_{11}a_{12} - a_{13})a_{13} - a_{14}a_{11}^2 > 0$. Here the lengthy expression of Δ_{13} is omitted for brevity. Therefore, E_1 is stable (unstable) if $f\tilde{v}\gamma\lambda_E - \mu_E(b_1 + \mu_A)(\tilde{v} + \mu_G) > 0 (< 0)$. Noticing the stability condition for E_0 we can see that E_0 and E_1 exchange their stability at the critical point, determined by $f\tilde{v}\gamma\lambda_E = \mu_E(b_1 + \mu_A)(\tilde{v} + \mu_G)$, and only a transcritical bifurcation exists at this critical point. This implies that there is also no Hopf bifurcation which can occur from the equilibrium E_1 .

Further, we can show that the trivial equilibrium of model (3.1) is globally asymptotically stable. This can be achieved by first considering the first, third and the last equations of (3.1), and ignoring the nonlinear term $-\sigma_1 R_n A$ in the first equation, yielding a linear system, which has the characteristic polynomial $P_0(L)$. Thus, by using comparison theory and this linear system (obtained by ignoring the nonlinear term), we can easily prove that the equilibrium E_0 is globally asymptotically stable. Although we have not proved the global stability of the non-trivial equilibrium, we have tried a number of numerical simulations, which show that all solutions converge to E_1 regardless the initial conditions as long as the condition $f\tilde{v}\gamma\lambda_E > \mu_E(b_1 + \mu_A)(\tilde{v} + \mu_G)$ is satisfied. Hence, we conjecture that the two models (3.1) and (3.2) do not have any persistent solutions, except the two equilibrium solutions E_0 and E_1 . This motivates the development of new models for studying relapse-remission dynamics in autoimmune disease.

3.2.2 Developing new models

Now, based on the two models (3.1) and (3.2), we develop new models. First, instead of considering the two immunosuppressive mechanisms in isolation, we combine them to obtain the following 4-dimensional ODE model:

$$\begin{aligned}
\dot{A} &= f\tilde{v}G - (\sigma_1 R_n + b_1)A - \mu_A A, \\
\dot{R}_n &= (\pi_3 E + \beta)A - \mu_n R_n, \\
\dot{E} &= \lambda_E A - (\sigma_3 R_n + b_3)E - \mu_E E, \\
\dot{G} &= \gamma E - \tilde{v}G - \mu_G G,
\end{aligned} \tag{3.4}$$

where the π_1 is replaced by π_3 . Note that the numerical value of either π_1 or π_3 from [1] could be used; the difference is immaterial to our analysis.

As mentioned in the introduction, phenotypic analysis indicates that the effector T_{Reg} cell subset is heterogeneous in the expression of HLA-DR [29], which identifies a terminally differentiated subpopulation of effector T_{Reg} cells, the HLA-DR+ T_{Reg} s. Therefore, we introduce these short-lived but potently suppressive T_{Reg} cells into our model (3.4), denoted by R_d . Then, we get a 5-dimensional model as follows:

$$\begin{aligned}
\dot{A} &= f\tilde{v}G - \sigma_1(R_n + dR_d)A - b_1A - \mu_A A, \\
\dot{R}_n &= (\pi_3 E + \beta)A - \mu_n R_n - \xi R_n, \\
\dot{R}_d &= c \xi R_n - \mu_d R_d, \\
\dot{E} &= \lambda_E A - \sigma_3(R_n + dR_d)E - b_3E - \mu_E E. \\
\dot{G} &= \gamma E - \tilde{v}G - \mu_G G.
\end{aligned} \tag{3.5}$$

For the above model, the possibility remains that HLA-DR⁻ nT_{Reg} cells may be activated to become terminal HLA-DR⁺ T_{Reg} cells [29]. Therefore, we indicate the part of HLA-DR⁻ nT_{Reg} cells which undergo activation as an output term from R_n population, with the activation rate, ' ξR_n '. The activated HLA-DR⁻ nT_{Reg} cells may further experience expansion and proliferation, say three divisions, thus $c = 2^3 = 8$, which contribute an input source of HLA-DR⁺ T_{Reg} cells, denoted by ' $c \xi R_n$ '. From the functional point of view, compared to HLA-DR⁻ T_{Reg} cells, HLA-DR⁺ T_{Reg} cells show more effective suppression of effector conventional T cells and pAPCs, and secrete cytokines more rapidly [5]. Therefore, we assume the suppression rate to pAPCs and effector T cells as ' $\sigma_1 dR_d A$ ' and ' $\sigma_3 dR_d E$ ', respectively, and set $d = 2$. In healthy adults, HLA-DR is expressed by approximately one third of effector T_{Reg} cells in peripheral blood [3], so here we assume in autoimmune patients the ratio is one half, implying that the ratio $\frac{R_d}{R_n}$ is one. We can use this fact to approximate ξ in the quasi-steady state of the R_d population, that is, $c \xi R_n - \mu_d R_d = 0$, yielding $\xi = 0.025$. The death and clearance rates μ_E and μ_A are based on the references given in [1], and are much the same here. Effector T cell lifetimes are approximately 4-5 days [24], so we set $\mu_E = 0.2 \text{ day}^{-1}$. The death rate of mature pAPCs is less certain [20]; we assume the lifetime of a mature pAPC is of the same order as that of a mature effector T cell and take $\mu_A = 0.2 \text{ day}^{-1}$ as well [1]. We likewise assume a similar death rate between the effector T cells and T_{Reg} cells, and set terminal T_{Reg} 's death rate as $\mu_d = 0.2 \text{ day}^{-1}$, and set $\mu_n = 0.1 \text{ day}^{-1}$, due to the rapid death rate of terminally differentiated effector HLA-DR⁺ T_{Reg} cells.

To simplify this model, for which the parameter values are shown in Table 3.1, we impose a quasi-steady state assumption on the free antigen concentration. In particular, we know that the

decay rate of the free antigen molecules (μ_G) is much faster than the dynamics of the effector T cells (E), and we can thus assume that the free antigen is in quasi-steady state with (and proportional to) the effector T cell population. Therefore, in the following, we shall eliminate G from system (3.5) by setting $\gamma E - \tilde{\nu}G - \mu_G G = 0$ to obtain $G = \frac{\gamma}{\mu_G + \tilde{\nu}} E$, to reduce system (3.5) by one dimension. Further, letting $\alpha = \frac{f\tilde{\nu}\gamma}{\mu_G + \tilde{\nu}}$, we obtain a new model, given by

$$\dot{A} = \alpha E - \sigma_1(R_n + dR_d)A - b_1A - \mu_A A, \quad (3.6a)$$

$$\dot{R}_n = (\pi_3 E + \beta)A - \mu_n R_n - \xi R_n, \quad (3.6b)$$

$$\dot{R}_d = c \xi R_n - \mu_d R_d, \quad (3.6c)$$

$$\dot{E} = \lambda_E A - \sigma_3(R_n + dR_d)E - b_3 E - \mu_E E. \quad (3.6d)$$

The parameter definitions and their values are given in Table 3.1. The state variables in (3.6) are defined as follows [1].

A : Mature pAPCs (professional antigen presenting cells), primarily mature dendritic cells, which present a particular self-antigen of interest and express sufficiently high levels of co-stimulatory molecules so as to be capable of activating T cells.

R_n : Activated natural T_{Reg} cells, HLA-DR⁻, specific for the antigen of interest, capable of exerting their suppressor function.

R_d : Terminally differentiated T_{Reg} cells, HLA-DR⁺, with hyper-suppressive ability.

E : Active auto-reactive effector T cells that are specific for the antigen of interest. These may be either CD4⁺ or CD8⁺ T cells, or even a combination of these two; the distinction is not important given the other simplifications we employ.

In the following sections, we study the new model (3.6) in detail, with particular interest in stability and bifurcation behaviors, and show that the model can exhibit cycles of relapse, inter-vened by relatively long periods of remission, which are characteristic of several autoimmune diseases.

3.3 Well-posedness, equilibrium solutions and stability of model (3.6)

First, we investigate the well-posedness of the solutions of model (3.6).

3.3.1 Well-posedness

Due to physical meaning of this autoimmune disease model, only non-negative initial conditions are considered and negative solutions are not allowed. Likewise the parameters in (3.6) are all positive due to their biological meaning. More precisely, we have the following result.

Theorem 3.3.1 *All solutions of system (3.6) are non-negative, if the initial conditions are non-negative. Furthermore, they are bounded.*

Proof Write the equations (3.6a) and (3.6d) as a non-autonomous system:

$$\begin{aligned}\dot{A} &= -[\sigma_1(R_n(t) + dR_d(t)) + b_1 + \mu_A]A + \alpha E, \\ \dot{E} &= -[\sigma_3(R_n(t) + dR_d(t)) + b_3 + \mu_E]E + \lambda_E A.\end{aligned}$$

Thus, according to Theorem 2.1 (P. 81) in [32] we know that $A(t) \geq 0$, and $E(t) \geq 0$ for $t > 0$, provided that $A(0) \geq 0$ and $E(0) \geq 0$. Then, $R_n(t) = R_n(0) \exp[-(\mu_n + \xi)t] + \int_0^t [\pi_3 E(\tau) + \beta] A(\tau) \exp[-(\mu_n + \xi)(t - \tau)] d\tau \geq 0$ for $A(t) \geq 0$, $E(t) \geq 0$ and $R_n(0) \geq 0$. Further $R_d(t) = R_d(0) \exp(-\mu_d t) + \int_0^t c \xi R_n(\tau) \exp[-\mu_d(t - \tau)] d\tau \geq 0$ for $R_n(t) \geq 0$ and $R_d(0) \geq 0$.

Next, we prove that all solutions of system (3.6) are bounded. We first consider two equations (3.6a) and (3.6d). Let

$$\begin{aligned}w_1(t) &= \sigma_1[R_n(t) + dR_d(t)] + (b_1 + \mu_A), \\ w_2(t) &= \sigma_3[R_n(t) + dR_d(t)] + (b_3 + \mu_E).\end{aligned}\tag{3.7}$$

With non-negative initial conditions, we have $w_1(t) > 0$ and $w_2(t) > 0 \forall t > 0$. We construct a Lyapunov-candidate-function of the form $V_1(A, E) = \frac{1}{2}(A^2 + E^2)$, $\forall A, E \geq 0$. It is easy to see that $V_1(A, E) > 0$, $\forall A, E > 0$, and $V_1(0, 0) = 0$. Taking the time derivative of V_1 along the trajectory governed by the differential equation (3.6a) and (3.6d) yields

$$\begin{aligned}\left. \frac{dV_1}{dt} \right|_{(3.6a), (3.6d)} &= A \dot{A} + E \dot{E} = A(-w_1 A + \alpha E) + E(\lambda_E A - w_2 E) \\ &= -(A, E) Q(t) \begin{pmatrix} A \\ E \end{pmatrix},\end{aligned}\tag{3.8}$$

where

$$Q(t) = \begin{bmatrix} w_1(t) & -\frac{1}{2}(\alpha + \lambda_E) \\ -\frac{1}{2}(\alpha + \lambda_E) & w_2(t) \end{bmatrix}.\tag{3.9}$$

To consider the positive definiteness of $Q(t)$, first note that $w_1 > 0 \forall t \geq 0$. For the sign of $\det(Q)$, if we assume $R_n(t)$ is unbounded, i.e., $\lim_{t \rightarrow +\infty} R_n(t) = +\infty$, then it will lead to a contradiction. Due to positivity of $R_d(t)$, σ_1 , σ_3 , d , b_1 , b_3 , μ_A , and μ_E , it follows from (3.7) that $\lim_{t \rightarrow +\infty} w_1(t) = \lim_{t \rightarrow +\infty} w_2(t) = +\infty$, which implies that there exists finite a time $t_1 > 0$, such that $\det[Q(t)] > 0 \forall t > t_1$. That means $Q(t)$ is positive definite, for $t > t_1$. Therefore, it follows from (3.8) that $\dot{V}_1 < 0 \forall t > t_1$. Thus, the equilibrium $(A, E) = (0, 0)$ is proven to be globally asymptotically stable, which implies $\lim_{t \rightarrow +\infty} A(t) = \lim_{t \rightarrow +\infty} E(t) = 0$. However, from (3.6b) we have

$$\lim_{t \rightarrow +\infty} \dot{R}_n(t) = [\pi_3 \lim_{t \rightarrow +\infty} E(t) + \beta] \lim_{t \rightarrow +\infty} A(t) - (\mu_E + \xi) \lim_{t \rightarrow +\infty} R_n(t) = -\infty,\tag{3.10}$$

which indicates that there exists a finite time $t_2 > t_1 > 0$, such that $\dot{R}_n(t) < 0 \forall t > t_2$, leading to $\lim_{t \rightarrow +\infty} R_n(t) = 0$, which is a contradiction with our assumption. Thus, $R_n(t)$ is bounded and we denote $M_{R_n} = \max\{R_n(t), t \geq 0\}$. This also means that equation (3.10) does not hold. Then, there exists an $N > 0$, such that $\lim_{t \rightarrow +\infty} [\pi_3 E(t) + \beta] A(t) = N$. Since the positivity of $E(t)$, π_3 , and β , we have $\pi_3 E(t) + \beta > \beta \forall t > 0$. Thus, there exists $M_A^* > 0$, such that $\lim_{t \rightarrow +\infty} A(t) = M_A^*$, implying that $A(t)$ is bounded, and we denote $M_A = \max\{A(t), t \geq 0\}$.

For the remaining part of the proof, we give a general claim first. Suppose we have the differential inequality : $\dot{T} \leq \lambda - dT$ ($\lambda, d > 0, T(0) > 0$). Then, for $\dot{T} = \lambda - dT$, we have

solution $T(t) = T(0)e^{-dt} + \frac{\lambda}{d}(1 - e^{-dt})$, which implies that $\lim_{t \rightarrow +\infty} \sup T(t) = \frac{\lambda}{d}$. Thus, from the equation (3.6c), we have $\dot{R}_d \leq c\xi M_{R_n} - \mu_d R_d$, which yields $\lim_{t \rightarrow +\infty} \sup R_d(t) = \frac{c\xi M_{R_n}}{\mu_d}$, and so R_d is bounded. Recalling that $A(t)$ is bounded, so for the equation (3.6d), we similarly have $\dot{E} \leq \lambda_E M_A - (b_3 + \mu_E) E$, which yields $\lim_{t \rightarrow +\infty} \sup E(t) = \frac{\lambda_E M_A}{b_3 + \mu_E}$, implying that $E(t)$ is bounded. Hence, the solution of system (3.6) is bounded.

The proof is complete.

Next, we will consider the equilibrium solutions of system (3.6) and determine their stability by using the Routh-Hurwitz criterion [17]. When we consider a Hopf bifurcation, we will use the result given in [44] to determine the Hopf critical condition.

3.3.2 Equilibrium solutions

By setting $\dot{A} = \dot{R}_n = \dot{R}_d = \dot{E} = 0$ in model (3.6), we get two equilibrium solutions: the tolerance equilibrium $E_0 : (\bar{A}_0, \bar{R}_{n0}, \bar{R}_{d0}, \bar{E}_0) = (0, 0, 0, 0)$, and the autoimmune disease equilibrium $E_1 : (\bar{A}, \bar{R}_n, \bar{R}_d, \bar{E})$, where

$$\begin{aligned}\bar{R}_n &= \frac{[\pi_3(b_1 + \mu_A)\bar{A} + \beta\alpha]\mu_d \bar{A}}{\mu_d\alpha(\mu_n + \xi) - \pi_3\sigma_1(\mu_d + d c \xi)\bar{A}^2}, \\ \bar{R}_d &= \frac{c\xi\bar{R}_n}{\mu_d}, \\ \bar{E} &= \frac{[\sigma_1\bar{R}_n(\mu_d + d c \xi) + \mu_d(b_1 + \mu_A)]\bar{A}}{\mu_d\alpha}.\end{aligned}\tag{3.11}$$

and \bar{A} is a function in terms of the system parameters, particularly α , and determined by the following 4th-degree equation, in which the parameter values given in Table 3.1 have been used. Note that the rational numbers given below are obtained using symbolic computation in which all the parameter values given in digital format (see Table 3.1) have been transformed to rational numbers for convenience in computation.

$$F_1(A, \alpha) = \frac{81}{38146972656250} A^4 - \frac{1521\alpha}{625000000} A^2 - \frac{81\alpha}{10000000} A + \frac{5}{8}\alpha^2 - \frac{81\alpha}{640000} = 0.\tag{3.12}$$

The graphs of $A = 0$ and $F_1(A, \alpha) = 0$ as given in (3.12) are shown in Figure 3.1, where Figure 3.1(a) shows the complete bifurcation diagram, while Figure 3.1(b) only depicts the part which is biologically meaningful. Figure 3.1(c) shows a 3-D plot, indicating why the branch in Figure 3.1(a) is biologically meaningless.

3.3.3 Stability of the disease-free equilibrium, E_0

For the stability of E_0 , we have the following result.

Theorem 3.3.2 *When $\alpha < \alpha_t = \frac{1}{\lambda_E}(b_1 + \mu_A)(b_3 + \mu_E)$, the disease-free equilibrium E_0 of the model (3.6) is globally asymptotically stable.*

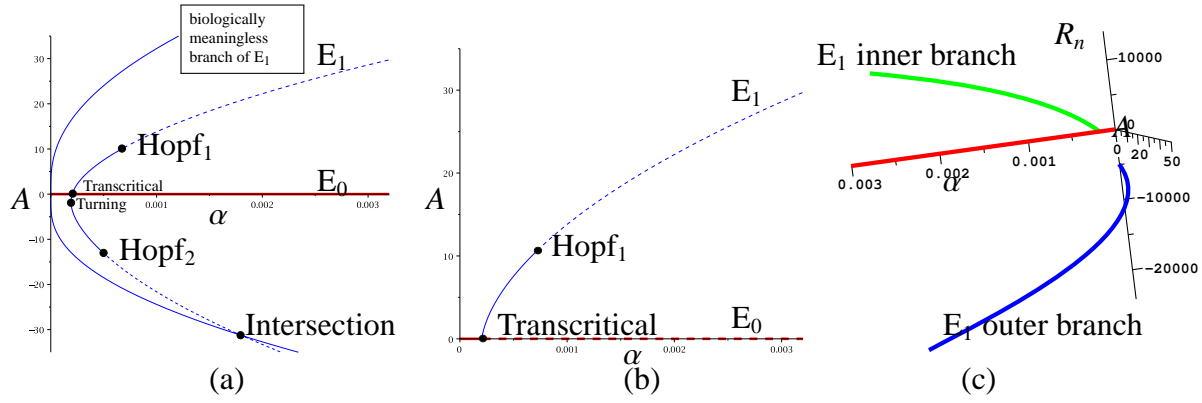


Figure 3.1: (a): Complete bifurcation diagram for model (3.6) projected on the α - A plane, with the red and blue lines denoting E_0 and E_1 , respectively; (b): Bifurcation diagram in (a), restricted to the first quadrant; (c): Bifurcation diagram for model (3.6) projected on the α - A - R_n space, with the red, green and blue lines denoting E_0 , the inner branch of E_1 , and the outer branch of E_1 which is biologically meaningless since R_n takes negative values. Here, the dotted and solid lines indicate unstable and stable equilibria, respectively.

Proof In order to examine the stability of equilibria for system (3.6), we compute the Jacobian matrix of system (3.6), given by

$$J = \begin{bmatrix} -\sigma_1(R_n + dR_d) - (b_1 + \mu_A) & -\sigma_1 A & -\sigma_1 dA & \alpha \\ \pi_3 E + \beta & -(\mu_n + \xi) & 0 & \pi_3 A \\ 0 & c\xi & -\mu_d & 0 \\ \lambda_E & -\sigma_3 E & -\sigma_3 dE & -\sigma_3(R_n + dR_d) - (b_3 + \mu_E) \end{bmatrix}. \quad (3.13)$$

Evaluating the Jacobian (3.13) at $E_0 : (\bar{A}_0, \bar{R}_{n0}, \bar{R}_{d0}, \bar{E}_0) = (0, 0, 0, 0)$, yields $J|_{E_0}$, and then setting $\det(LI - J|_{E_0})$ zero results in a 4th-degree characteristic equation, which can be factorized as

$$P_0(L, \alpha) = (L + \mu_d)(L + \mu_n + \xi) \left[L^2 + (b_3 + \mu_E + b_1 + \mu_A)L + (b_1 + \mu_A)(b_3 + \mu_E) - \lambda_E \alpha \right] = 0. \quad (3.14)$$

The asymptotic stability of E_0 is determined by the sign of real part of the roots of Equation (3.14): if all roots of Equation (3.14) have negative real part, then E_0 is asymptotically stable; if there is at least one root has positive real part, then E_0 is unstable. In fact, $P_0(L, \alpha)$ contains three factors: the first two are linear polynomials in L , with positive parameter values from Table 3.1, both of them are stable (i.e. their roots (eigenvalues) have negative real part); and thus the stability of E_0 only depends upon the third factor, which gives a quadratic equation,

$$L^2 + (b_3 + \mu_E + b_1 + \mu_A)L + (b_1 + \mu_A)(b_3 + \mu_E) - \lambda_E \alpha = 0. \quad (3.15)$$

Using the general formula for solutions of the quadratic equation, we know that whether the two roots of Equation (3.15) have negative real part is determined by the sign of $(b_3 + \mu_E)(b_1 + \mu_A) - \lambda_E \alpha$: the negativity (positivity) of the real part of the two roots of Equation (3.15) is equivalent to $(b_3 + \mu_E)(b_1 + \mu_A) - \lambda_E \alpha > 0 (< 0)$, that is, Equation (3.15) has stable (unstable) roots if $(b_3 + \mu_E)(b_1 + \mu_A) - \lambda_E \alpha > 0 (< 0)$, and a zero eigenvalue root comes out at

$$\alpha_t = \frac{(b_1 + \mu_A)(b_3 + \mu_E)}{\lambda_E}. \quad (3.16)$$

Here, the subscript ‘t’ stands for *transcritical bifurcation*. Using the parameter values from Table 3.1, the transcritical bifurcation point is obtained as $(\alpha_t, A_t) = (2.025 \times 10^{-4}, 0)$. The equilibrium solution E_0 is locally asymptotically stable (unstable), when $\alpha < \alpha_t$ ($\alpha > \alpha_t$).

Next, we want to prove that E_0 is also globally asymptotically stable for $\alpha < \alpha_t$. To achieve this, we construct a Lyapunov function of the form

$$V_2(A, E) = \frac{1}{2} (\lambda_E A^2 + \alpha E^2), \quad (3.17)$$

which is positive-definite and continuously differentiable for all positive bounded values of A and E , i.e., $V_2(0, 0) = 0$ and $V_2(A, E) > 0 \forall A, E > 0$. Moreover, the time derivative of the Lyapunov function V_2 satisfies

$$\begin{aligned} \dot{V}_2 &= \lambda_E A \dot{A} + \alpha E \dot{E} \\ &= \lambda_E A [\alpha E - \sigma_1(R_n + dR_d)A - (b_1 + \mu_A)A] \\ &\quad + \alpha E [\lambda_E A - \sigma_3(R_n + dR_d)E - (b_3 + \mu_E)E] \\ &= -\lambda_E (b_1 + \mu_A) A^2 - \alpha (b_3 + \mu_E) E^2 + 2 \alpha \lambda_E A E \\ &\quad - (\lambda_E \sigma_1 A^2 + \alpha \sigma_3 E^2)(R_n + dR_d) \\ &\leq -\lambda_E (b_1 + \mu_A) A^2 - \alpha (b_3 + \mu_E) E^2 + 2 \alpha \lambda_E A E \\ &= -(A \ E) Q (A \ E)^T, \end{aligned}$$

which is a quadratic form, with

$$Q = \begin{bmatrix} \lambda_E(b_1 + \mu_A) & -\alpha \lambda_E \\ -\alpha \lambda_E & \alpha(b_3 + \mu_E) \end{bmatrix} \quad (3.18)$$

being positive definite for $(b_1 + \mu_A)(b_3 + \mu_E) > \alpha \lambda_E$. Hence, $\dot{V}_2 \leq 0$ and $\dot{V}_2 = 0$ if and only if $(A, E) = (0, 0)$. This yields $A(t), E(t) \rightarrow 0$ as $t \rightarrow +\infty$, for any positive initial conditions. It follows that equation (3.6b) becomes an asymptotically autonomous equation with the limiting equation, $\dot{R}_n = -(\mu_n + \xi)R_n$. By the theory of asymptotically autonomous systems [7], we know that the solution $R_n(t) \rightarrow 0$ as $t \rightarrow +\infty$. Finally, using the same theory on equation (3.6c), we get $R_d(t) \rightarrow 0$ as $t \rightarrow +\infty$. Therefore, under the condition $\alpha < \alpha_t$, the local stability and the global attractivity of E_0 established above give the global asymptotic stability of E_0 .

3.3.4 Stability of the autoimmune disease equilibrium, E_1

In order to examine the stability of E_1 , we evaluate the Jacobian matrix (3.13) of system (3.6) at E_1 , to obtain the characteristic equation $\det(LI - J|_{E_1}) = 0$. By straightforward but tedious computations, the characteristic polynomial of J at E_1 is obtained as the following 4th-degree polynomial:

$$P_1(L, A, \alpha) = L^4 + a_1(A, \alpha)L^3 + a_2(A, \alpha)L^2 + a_3(A, \alpha)L + a_4(A, \alpha) = 0, \quad (3.19)$$

where the coefficients, $a_i(A, \alpha)$, $i = 1, 2, 3, 4$, are expressed in terms of A and α , with other parameter values taken from Table 3.1, and A satisfies $F_1(A, \alpha) = 0$ (see equation (3.12)).

The static bifurcation happens at equilibrium E_1 , when the characteristic polynomial $P_1(L, A, \alpha) = 0$ in (3.19) has zero root (zero eigenvalue). That means $a_4(A, \alpha) = 0$, and A should satisfy $F_1(A, \alpha) = 0$. Thus, we obtain

$$A_s(\alpha_s) = -\frac{21333593750000000\alpha_s^3 + 26617447265625\alpha_s^2 - 49464843750\alpha_s + 8748000}{3525388312500\alpha_s^2 - 4572342000\alpha_s + 979776}, \quad (3.20)$$

where α_s is the root of the following equation $F_2(\alpha_s) = \alpha_s(13530125\alpha_s - 2592) \times (400000\alpha_s - 81) = 0$. Solving $F_2(\alpha_s) = 0$, and then substituting the solutions into $A_s(\alpha_s)$ using Equation (3.20), we get three points. The first one is a transcritical bifurcation point $(\alpha_t, A_t) = (2.025 \times 10^{-4}, 0)$, which is exactly the same as that we obtained from the tolerance equilibrium E_0 . Moreover, at this point, all other Hurwitz arrangements are positive, that is, $\Delta_1 = \frac{49}{40}$, $\Delta_2 = \frac{5863}{16000}$, and $\Delta_3 = \frac{52767}{6400000}$. The two equilibrium solutions E_0 and E_1 intersect and exchange their stability at this critical point. E_1 is stable when $\alpha > \alpha_t$, (E_1 does not exist for $\alpha < \alpha_t$), as shown in Figure 3.1. Here, the subscript ‘t’ stands for transcritical bifurcation. The second point is a turning point $(\alpha_{\text{Turning}}, A_{\text{Turning}}) = (1.9157 \times 10^{-4}, -1.7097)$, which has a negative value for A and so is not biologically interesting (see Figure 3.1(a)). The third one is $(\alpha_s, A_s) = (0, 0)$, which is not allowed since the parameter α cannot take zero.

To check if a Hopf bifurcation exists from the infected equilibrium E_1 of system (3.6), we apply the theorem given in [44] to E_1 defined by (3.11), where A satisfies equation $F_1(A, \alpha) = 0$ in (3.12). Based on the fourth-degree characteristic polynomial $P_1(L, A, \alpha)$ in equation (3.19), we apply the formula in [44], that is, $\Delta_3(A, \alpha) = a_1 a_2 a_3 - a_3^2 - a_1^2 a_4 = 0$. Solving $\Delta_3(A, \alpha) = 0$ and $F_1(A, \alpha) = 0$, together with the parameter values given in Table 3.1, we get two Hopf bifurcation points: $(\alpha_{H1}, A_{H1}) = (7.8666 \times 10^{-4}, 11.4436)$, and $(\alpha_{H2}, A_{H2}) = (5.0387 \times 10^{-4}, -13.1534)$, as shown in Figure 3.1(a). We only consider the biologically meaningful point with two positive entries to obtain a unique Hopf bifurcation point: $(\alpha_H, A_H) = (7.8666 \times 10^{-4}, 11.4436)$. Here, the subscript ‘H’ stands for Hopf bifurcation. At the critical point (α_H, A_H) , other conditions are satisfied: $a_1 = 2.0989$, $a_2 = 0.6311$, $a_3 = 0.1145$, $a_4 = 0.0314$, $\Delta_2 = 1.2100$, $\Delta_3 = -0.1 \times 10^{-18} \approx 0$. Indeed, with these given parameter values, one can numerically calculate the Jacobian matrix of system (3.6) at E_1 , which contains a purely imaginary pair and two negative real eigenvalues: $\pm 0.2335i$, -1.7739 , and -0.325 . Thus, as α is varied across the point $\alpha = \alpha_H$, the equilibrium solution E_1 becomes unstable and a Hopf bifurcation occurs, leading to a family of limit cycles. Summarizing the above results gives the following theorem.

Theorem 3.3.3 *When $\alpha_t < \alpha < \alpha_H$, the disease equilibrium E_1 of model (3.6) is asymptotically stable.*

Now we apply normal form theory and the Maple program developed in [43] to system (3.6) to analyze the Hopf bifurcation which occurs at the critical point $(\alpha_H, A_H) = (7.8666 \times 10^{-4}, 11.4436)$ (with other parameters given in Table 3.1). Using a series of linear and nonlinear transformations and the Maple program [43], we obtain the normal form associated with this Hopf bifurcation up to third order, given by

$$\dot{r} = r(v_0\mu + v_1r^2), \quad \dot{\theta} = \omega_c + \tau_0\mu + \tau_1r^2, \quad (3.21)$$

where $v_0 = 34.2048$, $v_1 = -2.0161 \times 10^{-12}$, $\omega_c = 0.2335$, $\tau_0 = 132.8998$, $\tau_1 = -1.3186 \times 10^{-11}$. The steady-state solutions of equation (3.21) are determined by $\dot{r} = \dot{\theta} = 0$, resulting in $\bar{r}_1 = 0$

and $\bar{r}_2^2 = 0.1697 \times 10^{14} \mu$. The equilibrium solution $\bar{r}_1 = 0$ actually represents the autoimmune equilibrium E_1 of model (3.6). A linear analysis on the first differential equation of (3.21) shows that $\left. \frac{d}{dr}(\dot{r}) \right|_{\bar{r}=\bar{r}_1} = \nu_0 \mu$, and thus $\bar{r}_1 = 0$ (E_1) is stable (unstable) for $\mu < 0$ (> 0), as expected. When μ is increasing from negative to cross zero, a Hopf bifurcation occurs and the amplitude of the bifurcating limit cycles is given by the non-zero steady state solution,

$$\bar{r}(\mu) = 0.4119 \times 10^7 \sqrt{\mu} \quad (\mu > 0). \quad (3.22)$$

Since $\left. \frac{d}{dr}(\dot{r}) \right|_{(3.22)} = 2\nu_1 r^2$, it indicates that the bifurcating limit cycles are stable for $\mu > 0$. We can get the same stability conclusion from $\nu_1 < 0$, implying that the Hopf bifurcation is supercritical and so the bifurcating limit cycles are stable. Equation (3.22) gives the approximate amplitude of the bifurcating limit cycles, while the phase of the motion is determined by $\theta = \omega t$, where ω is given by $\omega = \left. \dot{\theta} \right|_{(3.22)} = 0.2335 - 90.8185 \mu$. We summarize the above results, yielding the following theorem.

Theorem 3.3.4 *At the critical point $\alpha = \alpha_H$, a supercritical Hopf bifurcation occurs, leading to a family of stable limit cycles.*

3.4 Numerical simulation

In this section, we present some simulation results to verify the analytical predictions obtained in the previous section. In particular, we will show the comparison between the analytical and numerical results obtained for the Hopf bifurcation. For convenience in the simulation, we will fix all parameter values, except for α (or μ). We will vary α to demonstrate the stable equilibrium solutions E_0 and E_1 , and the stable limit cycles. Finally, we will also choose a large positive value of μ , which means that this value is far away from the Hopf critical point α_H , to show the relapse-remission phenomenon. Note that the mechanism of generating recurrence in this paper is slightly different from that defined by the conditions in Hypothesis 1 of paper [45] in which recurrence is guaranteed to appear near a transcritical point. In this paper, recurrent oscillations are generated far from the transcritical point $\alpha_t = 2.025 \times 10^{-4}$. In other words, the oscillations described in this paper are determined by more global properties of the system.

Suppose that all parameter values, except for α , are taken from Table 3.1. Then, it follows from formula (3.16) that the equilibrium solution E_0 is asymptotically stable for $0 < \alpha < \alpha_t = 2.025 \times 10^{-4}$. E_0 becomes unstable when α is increased to pass through α_t , and bifurcates into the equilibrium solution E_1 , which is asymptotically stable for $\alpha_t < \alpha < \alpha_H = 7.8666 \times 10^{-4}$. E_1 becomes unstable at $\alpha = \alpha_H$, and a family of limit cycles bifurcates from this Hopf critical point. The normal form for the Hopf bifurcation is given by (3.21). Since $\nu_1 = -2.0161 \times 10^{-12}$, the Hopf bifurcation is supercritical, and the bifurcating limit cycles are stable.

Now, we first take $\alpha = 1.50 \times 10^{-4} < \alpha_t$. The simulation result is shown in Figure 3.2(a), which clearly indicates that E_0 is asymptotically stable, in agreement with the analytical prediction. Next, choose $\alpha_t < \alpha = 4.0 \times 10^{-4} < \alpha_H$, for which the simulation result is depicted in Figure 3.2(b), showing that E_1 is asymptotically stable, which again agrees with the analytical prediction. Further, we select a value of $\mu = 3.0 \times 10^{-12}$ which implies that we take a post-critical value of α near the Hopf critical point. This is a perfect Hopf bifurcation, as shown in Figure 3.3.

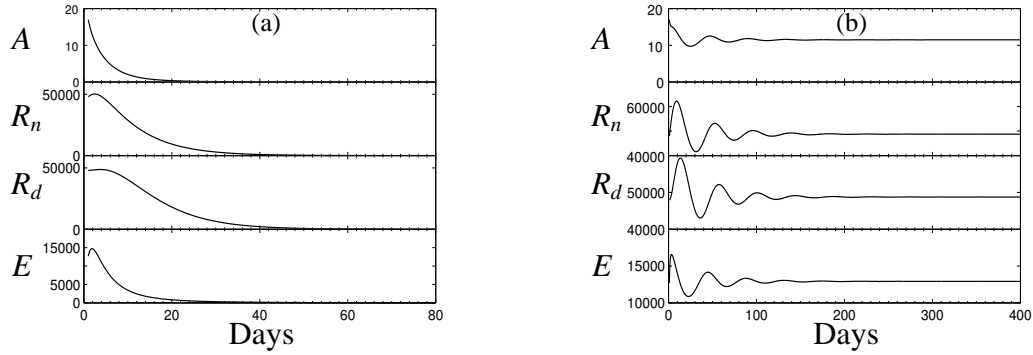


Figure 3.2: Simulated time history for system (3.6) with the initial condition $A(0) = 17$, $R_n(0) = R_d(0) = 48000$, $E(0) = 12700$ for (a) $\alpha = 1.50 \times 10^{-4} < \alpha_t$, converging to E_0 ; and (b) $\alpha = 4.0 \times 10^{-4}$, converging to E_1 .

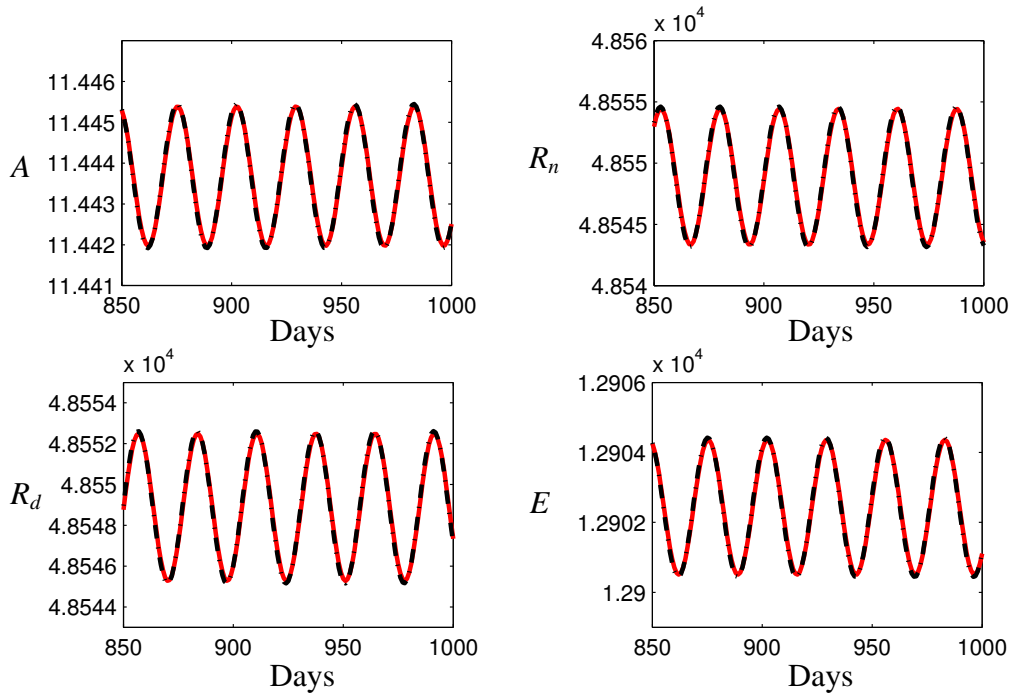


Figure 3.3: Comparison between the simulated time history and analytical prediction for system (3.6) with $\mu = 3 \times 10^{-12}$, the red solid line denoting the simulation results, while the black dash-dot line indicating the analytical predictions, showing stable limit cycles.

The simulations compared with the analytical predictions are depicted in Figure 3.3, showing excellent agreement between simulation results and analytical predictions, particularly for the smaller values of μ , as expected. Note that the analytical predictions are obtained through a series of linear and nonlinear transformations, available from the output of the Maple programs [43]. The details are omitted here for brevity. Finally, we take $\alpha = 3.0 \times 10^{-3} > \alpha_H$, which is not close to α_H . For this case, normal form theory is not applicable since this value of α is not near α_H . In other words, if we apply the above procedure to obtain an approximation, it would have a very large error. The simulation result is given in Figure 3.4, indeed showing the recurrence phenomenon. It should be noted that the vertical axis in Figure 3.4 (c) and (d)

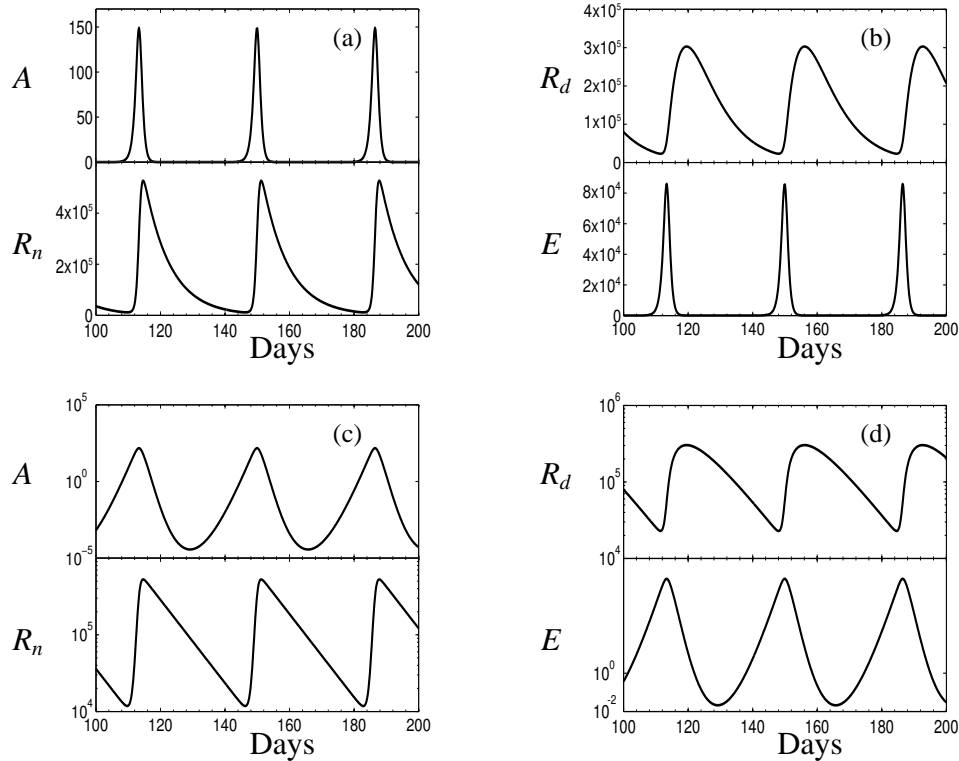


Figure 3.4: Simulated time history for system (3.6) when $\alpha = 3 \times 10^{-3}$, showing recurrence.

have a logarithmic scale so that the minimum level of effector T cells (E) can be clearly seen. The reason for this behavior can be seen from Figure 3.4 (a) and (b) to be: the E population grows very quickly in the absence of R_n and R_d , and then R_n responds very quickly (EA term) and suppresses E , but R_n does not last long. This pattern is of course how the adaptive and innate immune responses work against pathogens, as well. But why is E not eliminated like a pathogen would be? We speculate that the system is now ‘torn between two equilibria’, as described later in the Discussion.

3.5 Model reduction and parameter identification for autoimmune recurrence

In the previous sections, we have studied the 4-dimensional model (3.6) in detail and found recurrence. Now, we are interested in finding the key factors which play the most important roles in generating this phenomenon. To achieve this, a common approach is first to reduce the dimension of the system under a quasi-steady state assumption, and then identify the main system parameters (usually treated as bifurcation parameters) which may effectively influence recurrence so that we may find the mechanism of generating relapse and remission. For model reduction (in particular, the reduction from the 5-d model (3.5) to (3.6) and a further reduction from the 4-d model (3.6) to a 3-d model, which will be considered below in detail), we need to answer a fundamental question: does model reduction alter the dynamical behavior of the system? We have carefully studied this problem and have shown that when proper parameter values are chosen, both the original 5-d model and 4-d model, as well as the 4-d model and

the 3-d model exhibit the same dynamical behavior: recurrence. (Details will be given in a forthcoming paper.) Therefore, in the following, we will not consider the 5-d model (3.5), but the 4-d model and its reduction.

3.5.1 Model reduction

For the model described by (3.6), we assume that at the site of the autoimmune reaction, the influence of IL-2 from other sources, such as dendritic cells [1], is negligible compared to the IL-2 generated by activated effector T cells. Therefore, we can set $\beta = 0$, and the model becomes

$$\begin{aligned}\dot{A} &= \alpha E - \sigma_1(R_n + dR_d)A - (b_1 + \mu_A)A, \\ \dot{R}_n &= \pi_3 EA - (\mu_n + \xi)R_n, \\ \dot{R}_d &= c \xi R_n - \mu_d R_d, \\ \dot{E} &= \lambda_E A - \sigma_3(R_n + dR_d)E - (b_3 + \mu_E)E.\end{aligned}\tag{3.23}$$

It can be shown that model (3.23) still has two equilibrium solutions. One is the tolerance equilibrium $E_0 : (A, R_n, R_d, E) = (0, 0, 0, 0)$, and the other is the autoimmune equilibrium, $E_1 = (\bar{A}, \bar{R}_n(\bar{A}), \bar{R}_d(\bar{A}), \bar{E}(\bar{A}))$. We again choose α as the bifurcation parameter, and find that the two equilibrium solutions exchange their stability at the transcritical bifurcation point $(\alpha_s, A_s) = (2.025 \times 10^{-4}, 0)$. That is, as α increases from $\alpha < \alpha_s$ to cross the critical point $\alpha = \alpha_s$, the stable E_0 becomes unstable, while E_1 emerges from this critical point and is stable. As α continues to increase, a Hopf bifurcation occurs from E_1 at the critical point $(\alpha_H, A_H) = (6.4729 \times 10^{-4}, 12.4401)$. The simulated time history for $\alpha = 3 \times 10^{-3}$ shown in Figure 3.4 displays recurrent autoimmunity, as expected.

In order to further simplify the analysis on model (3.23), here we will adopt a quasi-steady state assumption, which is often used in the study of biochemical and biological systems. The basic idea of the quasi-steady state assumption can be described using the following system [8]:

$$\begin{aligned}\dot{x} &= \epsilon^{-1} f(x, y), & x &\in \mathbb{R}^m, \\ \dot{y} &= g(x, y), & y &\in \mathbb{R}^n,\end{aligned}\tag{3.24}$$

where $0 < \epsilon \ll 1$, f and g are nonlinear functions, and x and y represent ‘fast’ and ‘slow’ variables, respectively. We consider the evolution of the system from an arbitrary initial condition, including a transient period. For the fast variable x , we may rewrite the first equation of (3.24) as $\epsilon \dot{x} = f(x, y)$. Thus, for small ϵ , setting $\epsilon = 0$ results in $f(x, y) = 0$, from which we obtain an algebraic expression for x in terms of the slow variables $x = x(y)$; $\dot{x} \neq 0$ (see [8] for more details on this topic). This leads to a differential equation for the slow variable y in the form $\dot{y} = g(x(y), y)$. Intuitively, although the slow variable y is changing, the fast variable ‘catches up’ so quickly that $f(x, y)$ remains close to zero at all times.

Now, we return to consider system (3.23) and carefully compare the coefficients in the system, finding that the parameter $\lambda_E = 1000$ is greater than all other parameters, which are on the order of $10^{-6} \sim 1$. Thus, we may write the fourth equation of (3.23) as

$$\dot{E} = \lambda_E \left(A - \frac{\sigma_3}{\lambda_E} (R_n + dR_d) E - \frac{b_3 + \mu_E}{\lambda_E} E \right) = \epsilon^{-1} \left(A - \frac{\sigma_3}{\epsilon} (R_n + dR_d) E - \frac{b_3 + \mu_E}{\epsilon} E \right),$$

where $\epsilon = 10^{-3}$. Then, according to the general formula (3.24), we observe that E is a fast variable, while A , R_n , and R_d are slow variables, all of the same order. This is also reflected

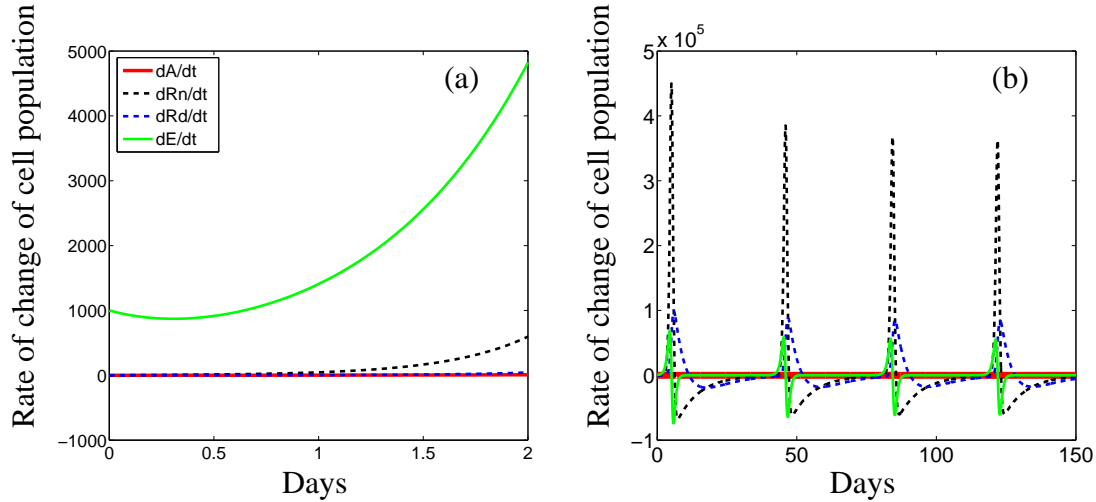


Figure 3.5: Simulated time history for system (3.23) with the initial conditions $A(0) = E(0) = 1$, $R_n(0) = R_d(0) = 0$, for the bifurcation parameter $\alpha = 3 \times 10^{-3}$: (a) for the transient period; and (b) over a longer interval showing periodic behavior. The rates of change of cell populations, \dot{A} , \dot{R}_n , \dot{R}_d , and \dot{E} , are represented by the red solid, black dotted, blue dotted, and green solid curves, respectively.

in the simulated time history for the transient period shown in Figure 3.5(a), which shows the rapid rate of change in E relative to the other populations. Therefore, we can make a quasi-steady state assumption on the fast variable E , yielding

$$E = \frac{\lambda_E A}{\sigma_3(R_n + dR_d) + b_3 + \mu_E}, \quad (3.25)$$

and so the reduced system is given by

$$\begin{aligned} \dot{A} &= \frac{\alpha \lambda_E A}{\sigma_3(R_n + dR_d) + b_3 + \mu_E} - \sigma_1(R_n + dR_d)A - (b_1 + \mu_A)A, \\ \dot{R}_n &= \frac{\pi_3 \lambda_E A^2}{\sigma_3(R_n + dR_d) + b_3 + \mu_E} - (\mu_n + \xi)R_n, \\ \dot{R}_d &= c \xi R_n - \mu_d R_d. \end{aligned} \quad (3.26)$$

3.5.2 Rescaling on system (3.26)

In order to reduce the number of parameters for convenience in analysis, we further attempt to rescale system (3.26) by scaling the state and time variables as

$$R_n = e_1 x, \quad R_d = e_2 y, \quad A = e_3 z, \quad t = e_4 \tau. \quad (3.27)$$

Then, with $\frac{d\tau}{dt} = \frac{1}{e_4}$, the left hand side of system (3.26) becomes

$$\frac{dR_n}{dt} = \frac{e_1}{e_4} \frac{dx}{d\tau}, \quad \frac{dR_d}{dt} = \frac{e_2}{e_4} \frac{dy}{d\tau}, \quad \frac{dA}{dt} = \frac{e_3}{e_4} \frac{dz}{d\tau}. \quad (3.28)$$

Next, we substitute (3.27) and (3.28) into system (3.26) to yield

$$\begin{aligned}\frac{dx}{d\tau} &= \frac{e_3^2 e_4 \lambda_E \pi_3}{e_1^2 \sigma_3 x + e_1 e_2 \sigma_3 d y + e_1 (b_3 + \mu_E)} z^2 - e_4 (\mu_n + \xi) x, \\ \frac{dy}{d\tau} &= \frac{e_1 e_4 c \xi}{e_2} x - e_4 \mu_d y, \\ \frac{dz}{d\tau} &= \frac{e_4 \alpha \lambda_E}{e_1 \sigma_3 x + e_2 \sigma_3 d y + (b_3 + \mu_E)} z - e_1 e_4 \sigma_1 x z - e_2 e_4 \sigma_1 d y z - e_4 (b_1 + \mu_A) z.\end{aligned}\tag{3.29}$$

Further, we set $e_1 e_4 \sigma_1 = 1$, $e_2 e_4 \sigma_1 d = 1$, $e_3^2 e_4 \lambda_E \pi_3 = 1$ and $e_4 \mu_d = 1$ to obtain

$$e_1 = \frac{\mu_d}{\sigma_1}, \quad e_2 = \frac{\mu_d}{\sigma_1 d}, \quad e_3 = \left(\frac{\mu_d}{\lambda_E \pi_3} \right)^{\frac{1}{2}}, \quad e_4 = \frac{1}{\mu_d}.\tag{3.30}$$

Finally, system (3.26) becomes

$$\begin{aligned}\frac{dx}{d\tau} &= \frac{z^2}{\mathcal{A}(x+y) + \mathcal{B}} - C x, \\ \frac{dy}{d\tau} &= \mathcal{D} x - y, \\ \frac{dz}{d\tau} &= \frac{\mathcal{E}}{\mathcal{F}(x+y) + \mathcal{G}} z - x z - y z - \mathcal{H} z,\end{aligned}\tag{3.31}$$

where the new parameters are defined as $\mathcal{A} = \frac{\sigma_3 \mu_d^2}{\sigma_1^2}$, $\mathcal{B} = \frac{\mu_d}{\sigma_1} (b_3 + \mu_E)$, $C = \frac{\mu_n + \xi}{\mu_d}$, $\mathcal{D} = \frac{c \xi d}{\mu_d}$, $\mathcal{E} = \frac{\alpha \lambda_E}{\mu_d}$, $\mathcal{F} = \frac{\sigma_3 \mu_d}{\sigma_1}$, $\mathcal{G} = b_3 + \mu_E$, $\mathcal{H} = \frac{b_1 + \mu_A}{\mu_d}$. Here, we set \mathcal{E} as the bifurcation parameter, since α is used as the bifurcation parameter for the original system (3.6). We then use the parameter values from Table 3.1 to obtain new parameter values for system (3.31) as $\mathcal{A} = \frac{40000}{3}$, $\mathcal{B} = 30000$, $C = \frac{5}{8}$, $\mathcal{D} = 2$, $\mathcal{F} = \frac{1}{5}$, $\mathcal{G} = \frac{9}{20}$, $\mathcal{H} = \frac{9}{4}$. Moreover, it follows from (3.30) that $e_1 = \frac{200000}{3}$, $e_2 = \frac{100000}{3}$, $e_3 = \frac{\sqrt{2}}{16}$, $e_4 = 5$.

The bifurcation patterns of the scaled system (3.31) are the same as that of the original system (3.6), namely, there exist two equilibrium solutions: $\bar{E}_0 : (x_0, y_0, z_0) = (0, 0, 0)$, and $\bar{E}_1 : (x_1, y_1, z_1)$, where $y_1 = \mathcal{D} x_1$, $z_1 = \sqrt{C x_1 [\mathcal{A}(1 + \mathcal{D}) x_1 + \mathcal{B}]}$, and x_1 is determined from the equation: $(1 + \mathcal{D})^2 \mathcal{F} x^2 + [(\mathcal{G} + \mathcal{H} \mathcal{F})(1 + \mathcal{D})] x - \mathcal{E} + \mathcal{H} \mathcal{G} = 0$

Theorem 3.5.1 *The solutions of system (3.31) are non-negative and bounded, provided that the initial conditions are non-negative.*

Proof For the non-negativeness, we write the solutions for z and y of system (3.31) by using the method of constant variations as

$$z(\tau) = z(0) \exp \left[\int_0^\tau \left(\frac{\mathcal{E}}{\mathcal{F}[x(s) + y(s)] + \mathcal{G}} - x(s) - y(s) - \mathcal{H} \right) ds \right],\tag{3.32}$$

and

$$y(\tau) = y(0) e^{-\tau} + \mathcal{D} \int_0^\tau e^{-(\tau-s)} x(s) ds.\tag{3.33}$$

There are two cases.

Case 1. $z(0) = 0$. Then, it follows from (3.32) that $z(\tau) \equiv 0$, $\forall \tau \geq 0$. Thus, the first equation of system (3.31) is reduced to $\frac{dx}{d\tau} = -Cx$, which yields the solution $x(\tau) = x(0)e^{-C\tau}$. Therefore, $x(\tau) \geq 0$, $\forall \tau \geq 0$ if $x(0) \geq 0$. Then, we use (3.33) to obtain $y(\tau) \geq 0$, $\forall \tau \geq 0$ if $y(0) \geq 0$.

Case 2. $z(0) > 0$. Then, it is easy to see from (3.32) that $z(\tau) > 0$, $\forall \tau \geq 0$. We need to discuss four subcases.

Case 2.1. $x(0) > 0$ and $y(0) > 0$. To prove $y(\tau) > 0$, $\forall \tau > 0$, we adopt the argument of contradiction. Since $y(0) > 0$, we assume the first time at which $y(\tau)$ becomes negative is τ_1 , i.e., $y(\tau) > 0$, $\forall \tau \in [0, \tau_1)$, $y(\tau_1) = 0$ and $y(\tau) < 0$, $\forall \tau \in (\tau_1, \tau_2)$. Then, since $y(0)e^{-\tau} > 0$, (3.33) implies that there should exist an interval $(\tau_3, \tau_4) \subset [0, \tau_1)$, such that $x(\tau) < 0$, $\forall \tau \in (\tau_3, \tau_4)$ (τ_1 may equal τ_4). With $x(0) > 0$, we may, without loss of generality, assume τ_3 is the first time $x(\tau)$ become zero, that is, $x(\tau_3) = 0$ and $x(\tau), y(\tau) > 0 \forall \tau \in (0, \tau_3)$. On the other hand,

$$\frac{dx}{d\tau} = \frac{z^2}{\mathcal{A}(x+y) + \mathcal{B}} - Cx > -Cx, \quad \text{for } \tau \in [0, \tau_3]. \quad (3.34)$$

By the comparison principle, we have $x(\tau_3) > x(0)e^{-C\tau_3} > 0$ for $x(0) > 0$, which contradicts that $x(\tau_3) = 0$. Therefore, there is no time for $y(\tau)$ to be zero and then become negative, that is, $y(\tau) > 0$, $\forall \tau \geq 0$. Then, using a similar argument on (3.34), we can prove that $x(\tau) > 0$, $\forall \tau \geq 0$.

Case 2.2 $x(0) = y(0) = 0$. Due to the continuity of the solutions and the conditions $\mathcal{A} > 0$ and $\mathcal{B} > 0$, for the term $\frac{z^2}{\mathcal{A}(x+y) + \mathcal{B}}$, there exists $\tau_5 > 0$, such that, for $\tau \in [0, \tau_5]$, $\frac{z(\tau)^2}{\mathcal{A}[x(\tau)+y(\tau)] + \mathcal{B}} > 0$. Then, $\frac{dx}{d\tau} = \frac{z^2}{\mathcal{A}(x+y) + \mathcal{B}} - Cx > -Cx$, $\forall \tau \in (0, \tau_5]$. Therefore, $x(\tau) > x(0)e^{-C\tau} = 0$ for $\tau \in [0, \tau_5]$. Moreover, the solution of $y(\tau) = \mathcal{D} \int_0^\tau e^{-(\tau-s)} x(s) ds$ indicates $y(\tau) > 0$ for $\tau \in [0, \tau_5]$. Hence, we obtain $x(\tau_5) > 0$ and $y(\tau_5) > 0$. So we can take τ_5 as the initial point and use the conclusion obtained in Case 2.1 to show that $x(\tau) > 0$ and $y(\tau) > 0$ for $\tau \geq \tau_5$. Combining the above two steps proves that $x(\tau) > 0$ and $y(\tau) > 0$ for $\tau > 0$.

Case 2.3 $x(0) = 0$ and $y(0) > 0$.

Case 2.4 $x(0) > 0$ and $y(0) = 0$.

For Cases 2.3 and 2.4, we can apply similar arguments used for proving Cases 2.1 and 2.2 to prove that the solutions of system (3.31) with these initial conditions are non-negative.

The remainder of the proof is devoted to the boundedness of solutions. Suppose that $y(\tau)$ is unbounded, that is, as $\tau \rightarrow +\infty$, $y(\tau) \rightarrow +\infty$. Then, according to the second equation in (3.31), we have $\lim_{\tau \rightarrow +\infty} x(\tau) = +\infty$, and further obtain $\lim_{\tau \rightarrow +\infty} z(\tau) = 0$ by using the third equation in (3.31), and then obtain $\lim_{\tau \rightarrow +\infty} x(\tau) = 0$ from the first equation in (3.31). This leads to a contradiction, and so $y(\tau)$ is bounded. Now applying the boundedness of $y(\tau)$ to the second equation in (3.31) yields the boundedness of $x(\tau)$. Finally, with bounded $x(\tau)$ and $y(\tau)$, the first equation in (3.31) shows that $z(\tau)$ must be bounded as well. Hence, all the solutions of system (3.31) are bounded. The proof is complete.

The characteristic polynomial for \bar{E}_0 is $P_0(L) = (L+1)(L+C)(L\mathcal{G}-\mathcal{E}+\mathcal{H}\mathcal{G})/\mathcal{G}$, from which it is easy to show that \bar{E}_0 is asymptotically stable for $\mathcal{E} < \mathcal{E}_s = \mathcal{H}\mathcal{G}$ and becomes unstable at the critical point $\mathcal{E}_s = \mathcal{H}\mathcal{G}$, from which \bar{E}_1 appears. Further, we can use the characteristic polynomial for \bar{E}_1 to show that the two equilibrium solutions exchange their stability at the transcritical bifurcation point $\mathcal{E}_s = \mathcal{H}\mathcal{G}$. Further, we have the following result for \bar{E}_0 .

Theorem 3.5.2 *The trivial equilibrium $E_0 : (x_0, y_0, z_0) = (0, 0, 0)$ is globally asymptotically stable, for $\mathcal{E} < \mathcal{E}_s = \mathcal{H}\mathcal{G}$.*

Proof We construct the Lyapunov function, $V(x, y, z) = \frac{1}{2}(x^2 + \rho_1 y^2 + \rho_2 z^2)$ for system (3.31), where $\rho_1 = \frac{3C}{\mathcal{D}^2}$, and $\rho_2 = \frac{1}{\mathcal{B}}$. V is continuously differentiable for all positive bounded values of each variable, and positive definite with positive parameter values, i.e., $V(0, 0, 0) = 0$ and $V(x, y, z) > 0, \forall x, y, z > 0$. Then, the derivative of the Lyapunov function V with respect to time, along the solution trajectory of system (3.31), yields

$$\begin{aligned} \left. \frac{dV}{d\tau} \right|_{(3.31)} &= x \frac{dx}{d\tau} + \rho_1 y \frac{dy}{d\tau} + \rho_2 z \frac{dz}{d\tau} \\ &= x \left[\frac{z^2}{\mathcal{A}(x+y) + \mathcal{B}} - Cx \right] + \rho_1 y [\mathcal{D}x - y] + \rho_2 z^2 \left[\frac{\mathcal{E}}{\mathcal{F}(x+y) + \mathcal{G}} - x - y - \mathcal{H} \right] \\ &= \left[\frac{1}{\mathcal{A}(x+y) + \mathcal{B}} - \rho_2 \right] xz^2 - C \left(x - \frac{\rho_1 \mathcal{D}}{2C} y \right)^2 - \rho_1 \left(1 - \frac{\rho_1 \mathcal{D}^2}{4C} \right) y^2 \\ &\quad + \left[\frac{\mathcal{E}}{\mathcal{F}(x+y) + \mathcal{G}} - \mathcal{H} \right] \rho_2 z^2 - \rho_2 xz^2 - \rho_2 yz^2, \end{aligned} \tag{3.35}$$

which implies that $\frac{dV}{d\tau} < 0, \forall x, y, z > 0$ due to $\mathcal{E} < \mathcal{H}\mathcal{G}$. The proof is complete.

The characteristic polynomial for \bar{E}_1 is $P_1(L) = L^3 + a_1(x_1)L^2 + a_2(x_1)L + a_3(x_1)$. $a_3(x_1) = 0$ defines the transcritical point $\mathcal{E} = \mathcal{E}_s$. The Hopf bifurcation point can be determined from the Hurwitz arrangement $\Delta_2 = a_1(x_1)a_2(x_1) - a_3(x_1) = 0$. In general, we may take three parameters, say, C , \mathcal{D} , and \mathcal{E} , as the bifurcation parameters. Therefore, the stability boundary, based in particular on the Hopf critical condition, can be displayed in the 3-dimensional parameter space as a surface. We then try to identify the region in the 3-dimensional parameter space where recurrence may occur. For a clear view of the stability boundary, we use $C = \text{constant}$ or $\mathcal{D} = \text{constant}$ to intersect the surface to obtain planes, as shown in Figure 3.6. The curves shown in Figure 3.6 are the stability boundary determined by the Hopf critical condition. The graphs of $\Delta_2(C, \mathcal{E}) = 0$ and $\Delta_2(\mathcal{D}, \mathcal{E}) = 0$ are plotted in the 2-dimensional $C - \mathcal{E}$ and $\mathcal{D} - \mathcal{E}$ parameter planes, as shown in Figure 3.6. Recurrence may occur on the right side (stable side for bifurcating limit cycles) of the Hopf critical curves. Moreover, in these planes, we select several fixed values for C or \mathcal{D} to obtain the horizontal lines, as shown in Figure 3.6. Then, we choose the points (according to the values of \mathcal{E}) on these lines to perform simulation. Two sets of nine simulated results are presented in Figures 3.7 and 3.8, corresponding to the nine points marked on the five solid lines in each figure of Figure 3.6. It is seen from Figure 3.7 that recurrence becomes more visible when the notation number of the points increases. That is, as

\mathcal{D} is fixed, reducing the value of C (see Figure. 3.6(a)) causes more dramatic recurrence, while changing \mathcal{E} in this case does not change the pattern. Figure. 3.8, on the other hand, shows that when C is fixed at an appropriate value, the changes of \mathcal{D} and \mathcal{E} (see Figure. 3.6(b)) do not play a significant role in determining recurrence. These parameter studies provide us with information regarding which parameters play an important role in generating recurrence: while some parameters mainly change the frequency of the motion, others only affect amplitude.

Finally, we would like to ask a question: since the recurrent pattern (or periodic solution) occurs at the parameter values which are far away from the Hopf critical point, is there any factor other than the Hopf bifurcation contributing to the oscillation. More specifically, do there exist homoclinic orbits? The answer is negative, given in the following theorem.

Theorem 3.5.3 *There exist no homoclinic orbits in the 3-dimensional scaled system (3.31) or the 4-dimensional system (3.6). Thus, the stable limit cycles either come from Hopf bifurcation or are due to persistent oscillations.*

Proof First, for the 3-dimensional scaled system (3.31), note that existence of homoclinic orbits needs a saddle or a saddle-focus point, which requires $\mathcal{E} > \mathcal{H}\mathcal{G}$. Evaluating the characteristic polynomial at $\bar{E}_0 : (0, 0, 0)$ yields three eigenvalues: $\lambda_1 = -C$, $\lambda_2 = -1$, and $\lambda_3 = \frac{\mathcal{E} - \mathcal{H}\mathcal{G}}{\mathcal{G}}$. Their corresponding eigenvectors are $V_1 = (\frac{1-C}{\mathcal{D}}, 1, 0)^\top$, $V_2 = (0, 1, 0)^\top$, and $V_3 = (0, 0, 1)^\top$, starting from \bar{E}_0 . Then, since for \bar{E}_0 the eigenvalue λ_3 is positive, while the other two eigenvalues λ_1 and λ_2 are negative, \bar{E}_0 is a saddle point. If a homoclinic orbit exists, it must connect the saddle point to itself, leaving in the direction tangent to V_3 at \bar{E}_0 , and coming back along a convergent trajectory to \bar{E}_0 , which is located in the stable manifold of system (3.31). It is easy to show that the two eigenvectors V_1 and V_2 actually construct the stable manifold, which is the first quadrant of the x - y plane, denoted by S_1 . The solution on the stable manifold can be expressed as $v = T_1 v_1 + T_2 v_2$, for $T_1, T_2 \in \mathbb{R}^+$, where $v_1 = (\frac{1-C}{\mathcal{D}} e^{-C\tau}, e^{-C\tau}, 0)^\top$ and $v_2 = (0, e^{-\tau}, 0)^\top$. Then it is obvious that S_1 is invariant by verifying the solution v to satisfy system (3.31). The complementary space of S_1 is the z -axis, which is tangent to the unstable manifold. Thus if a homoclinic orbit exists, it must connect the unstable and stable manifolds. However, this is impossible since there is no singular point on S_1 (except for \bar{E}_0), and so it cannot intersect S_1 due to the uniqueness of solutions. Therefore, no homoclinic orbits exist in system (3.31), and thus the stable limit cycles in system (3.31) either come from Hopf bifurcation or are due to persistent oscillations.

Next, we consider the 4-dimensional system (3.6). Note that system (3.6) also has two equilibrium solutions $E_0 : (\bar{A}_0, \bar{R}_{n0}, \bar{R}_{d0}, \bar{E}_0) = (0, 0, 0, 0)$ and $E_1 : (\bar{A}, \bar{R}_n, \bar{R}_d, \bar{E})$, where \bar{A} is determined by equation (3.12), and the other three components are given in equation (3.11). E_0 and E_1 exchange their stability at a transcritical bifurcation point $\alpha = \alpha_t$ defined in (3.16). When $0 < \alpha < \alpha_t$, E_0 is globally asymptotically stable, and E_1 does not exist; when $\alpha_t < \alpha < \alpha_H$, E_0 becomes unstable, while E_1 is asymptotically stable, where α_H is a Hopf bifurcation point at which limit cycles bifurcate from E_1 . When $\alpha > \alpha_H$, E_1 also becomes unstable.

The existence of homoclinic orbits, requires the existence of a saddle or a saddle-focus point, yielding the condition $\alpha > \alpha_t = \frac{1}{\lambda_E}(b_1 + \mu_A)(b_3 + \mu_E)$. The characteristic polynomial for

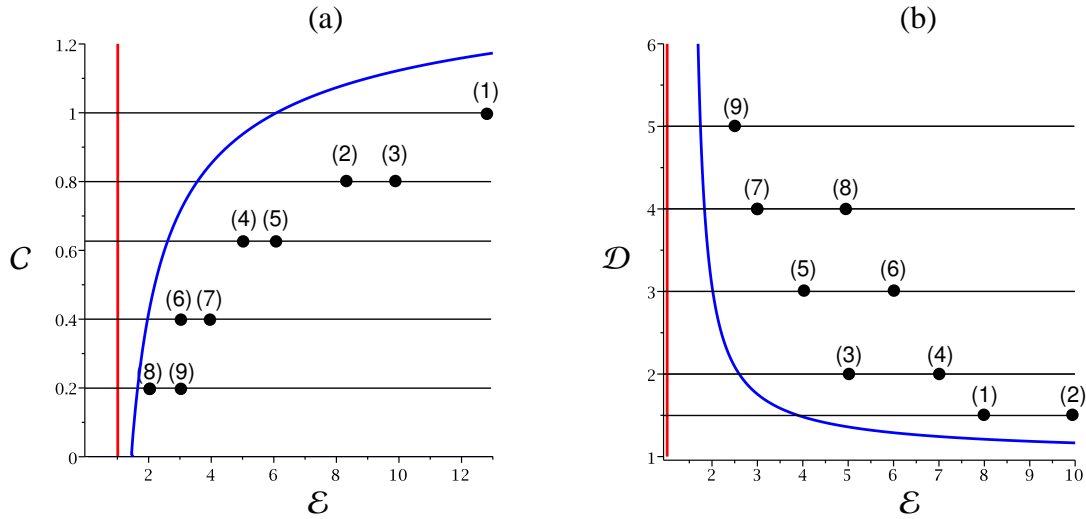


Figure 3.6: Bifurcation diagrams for the scaled system (3.31) in 2-dimensional $C-E$ and $D-E$ parameter spaces, where (a) $D = 2$; and (b) $C = \frac{5}{8}$; with the red and blue lines denoting the transcritical bifurcation and Hopf bifurcation, respectively.

E_0 is given by equation (3.14), from which we obtain four eigenvalues:

$$\begin{aligned}
 L_1 &= -(\mu_n + \xi), \\
 L_2 &= -\mu_d, \\
 L_3 &= \frac{1}{2} \left\{ -(b_1 + b_3 + \mu_A + \mu_E) - \sqrt{(b_1 + b_3 + \mu_A + \mu_E)^2 + 4\lambda_E(\alpha - \alpha_t)} \right\}, \\
 L_4 &= \frac{1}{2} \left\{ -(b_1 + b_3 + \mu_A + \mu_E) + \sqrt{(b_1 + b_3 + \mu_A + \mu_E)^2 + 4\lambda_E(\alpha - \alpha_t)} \right\},
 \end{aligned} \tag{3.36}$$

Since, $\alpha > \alpha_t$ and we have $L_3 < 0$ and $L_4 > 0$, indicating that E_0 is a saddle point when $\alpha > \alpha_t$. The eigenvectors corresponding to the two negative eigenvalues L_1 and L_2 are $V_1 = \left(0, \frac{\mu_d - \mu_n - \xi}{\alpha \xi}, 1, 0\right)^\top$ and $V_2 = (0, 0, 1, 0)^\top$, respectively. It is easy to verify that the solutions: $v_1 = V_1 e^{-(\mu_n + \xi)t}$ and $v_2 = V_2 e^{-\mu_d t}$ satisfy system (3.6). Further, it can be shown that the general solution, $(A, R_n, R_d, E)^\top = T_1 v_1 + T_2 v_2$ also satisfies system (3.6), where $T_1, T_2 \in \mathbb{R}^+$. This implies that the subspace determined by $A = E = 0$, i.e., the first quadrant of the R_n - R_d plane is a two-dimensional invariant stable submanifold, denoted by S_2 . Hence, if a homoclinic orbit exists in system (3.6), it cannot return to E_0 via S_2 , otherwise, it contradicts the uniqueness of solutions. So, the remaining possibility for a homoclinic orbit to appear is in the complementary space of S_2 , which is the first quadrant of the A - E plane, denoted by $C : \{(A, R_n, R_d, E) | A, E \geq 0, R_n = R_d = 0\}$ on which the dynamics are described by $\dot{A} = \alpha E - (b_1 + \mu_A)A$, $\dot{E} = \lambda_E A - (b_3 + \mu_E)E$. However, this system is linear. So no homoclinic orbits can exist in system (3.6), and thus the stable limit cycles in system (3.6) either come from Hopf bifurcation, or are due to persistent oscillations. The proof is complete.

In this section, we have made two reductions, one based on a quasi-steady state assumption and the other based on rescaling. It should be noted that these two reductions have a fundamental difference. The latter one actually generates an equivalent system, i.e., system (3.31) is equivalent to system (3.26), while the former yields system (3.26) which is different from system (3.23). However, system (3.26) still keeps the basic interesting dynamic behaviour (recurrency) as that of the original system (3.23) under the quasi-steady state assumption.

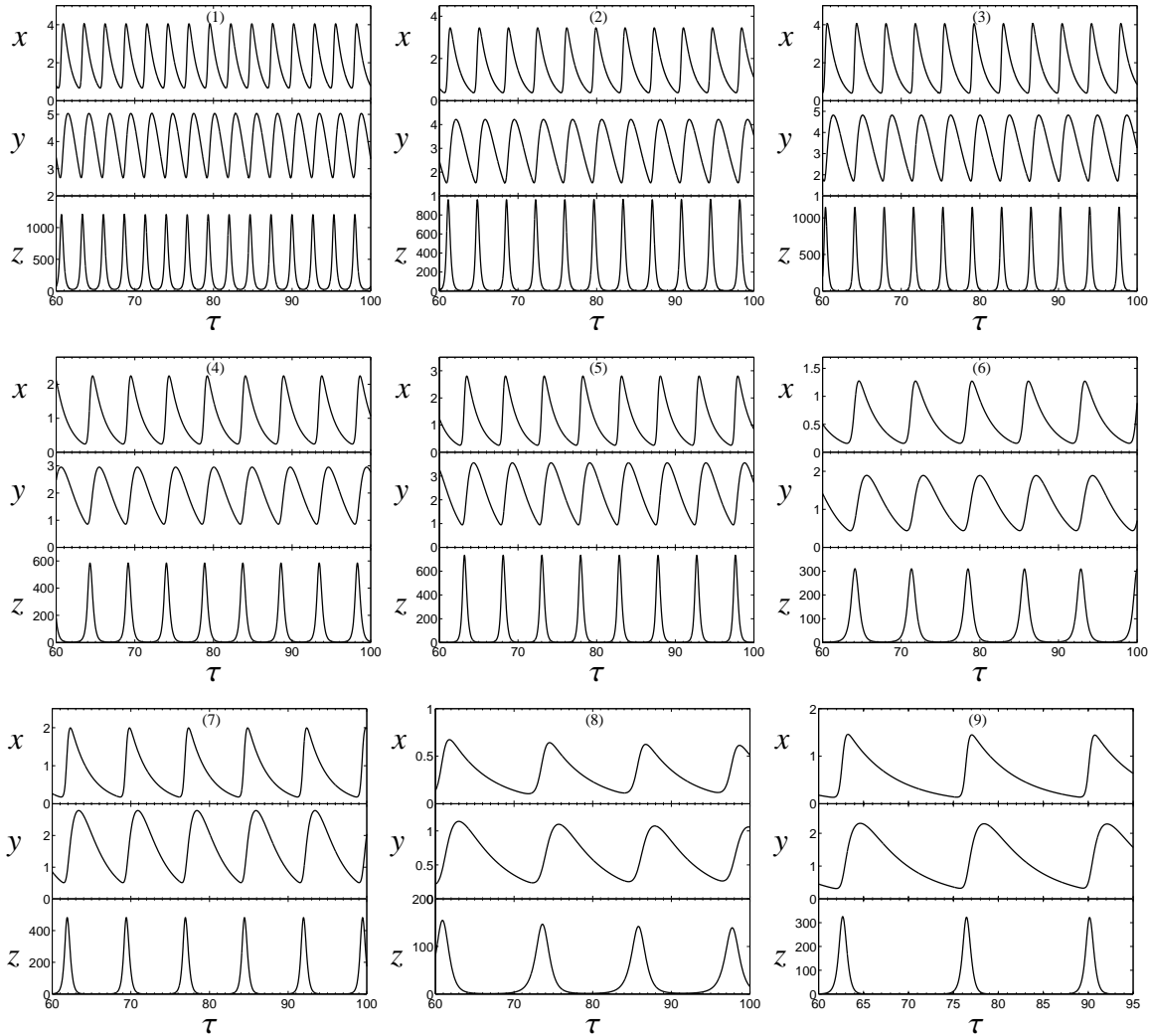


Figure 3.7: Numerical simulations for equation (3.31) with the parameter values (\mathcal{E}, C) taken as: (1) (13, 1.0), (2) (8.5, 0.8), (3) (10, 0.8), (4) (5, 0.625), (5) (6, 0.625), (6) (3, 0.4), (7) (4, 0.4), (8) (2, 0.2), (9) (3, 0.2).

3.6 Conclusion and discussion

Adaptive immunity in vertebrates comprises an extremely complex dynamical system, and much remains to be elucidated, particularly with respect to the role and action of regulatory T cells. In this contribution, we demonstrate that the addition of a newly discovered subclass of T_{Reg} cells, the terminally differentiated HLA-DR⁺ class [5, 27], alters the dynamical behavior of a general model of autoimmune disease [1]. In particular, rather than being restricted to stable equilibria corresponding to self-tolerance and autoimmunity, the system now displays long periods of quiescence, punctuated by brief bursts of autoimmune activity. These cycles of relapse and remission, characteristic of many autoimmune diseases, arise naturally from the dynamical behavior of the system, without the need for stochastic input or exogenous environmental triggers.

As an intuitive explanation for this phenomenon, we argue that the dynamical system is

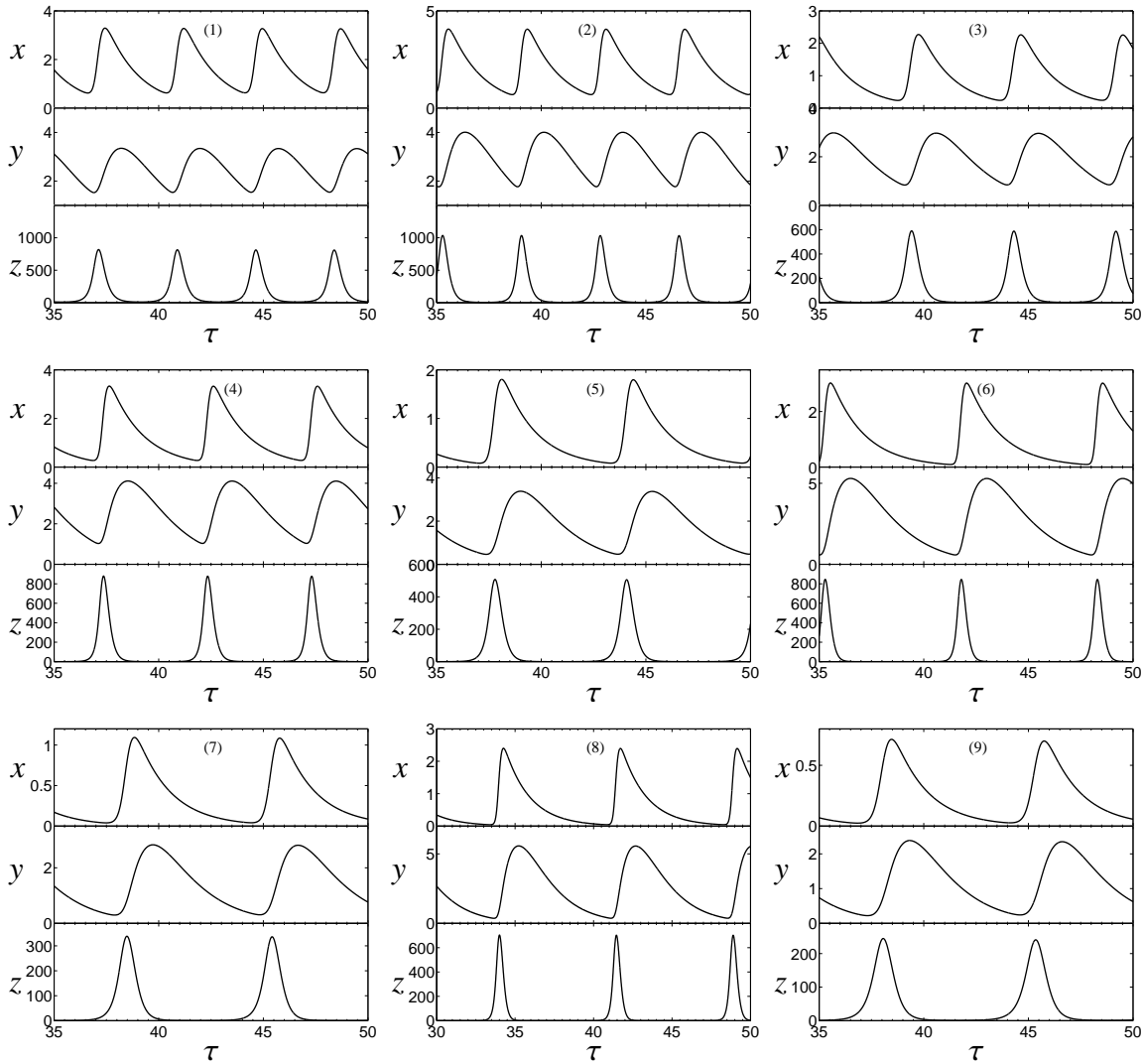


Figure 3.8: Numerical simulation for equation (3.31) with the parameter values $(\mathcal{E}, \mathcal{D})$ taken as: (1) (8, 1.5), (2) (10, 1.5), (3) (5, 2), (4) (7, 2), (5) (4, 3), (6) (6, 3), (7) (3, 4), (8) (5, 4), (9) (2.5, 5).

‘torn between two equilibria’, one of which is the trivial equilibrium corresponding to immune tolerance (self-reactive populations at zero), the other corresponding to a full-blown autoimmune reaction. As a result, after the Hopf bifurcation the model populations remain close to the tolerance equilibrium for long intervals, during which immune regulation (the T_{Reg} population) gradually wanes. When regulatory populations are sufficiently small, the autoreactive effector population escapes immune regulation and a brief episode of autoimmune disease, a relapse, occurs.

Although the cycles of relapse and remission observed in this system occur at regular intervals, we note that even slight fluctuations in the parameter values, or deterministic changes in parameters over time, can result in highly variable intervals between relapse episodes, as demonstrated in [45]. We also note that in any organism, self-antigen is likely to be continually present at low levels. Thus, even if the relevant populations reach extremely low frequencies

during the cycles of remission predicted here, pAPCs specific for self-antigen are likely to be periodically generated, renewing the relapse-remission cycle if they are activated when the T_{Reg} populations have waned. This could be a further factor contributing to variable intervals between relapse episodes.

Clearly, the models we analyse are extreme simplifications of the mechanisms of immune regulation. As the precise mechanisms of action of regulatory T cells are further elucidated, more accurate and predictive models should be possible. Nonetheless we hope that the main insight of this paper, that recurrence in autoimmune diseases can arise naturally from the complex interplay of dynamic populations, will serve as a starting point for further research both in dynamical systems theory, and in theoretical immunology.

3.7 References

- [1] H. K. Alexander and L. M. Wahl. Self-tolerance and autoimmunity in a regulatory T cell model. *Bulletin of mathematical biology*, 73(1):33–71, 2011.
- [2] American Autoimmune Related Diseases Association. Autoimmune disease in women. <http://www.aarda.org/autoimmune-information/autoimmune-disease-in-women/>, Accessed: 2014-01-21.
- [3] C. Baecher-Allan, J. A. Brown, G. J. Freeman, and D. A. Hafler. $CD4^+$ $CD25$ high regulatory cells in human peripheral blood. *The Journal of Immunology*, 167:1245–1253, 2001.
- [4] C. Baecher-Allan and D. A. Hafler. Human regulatory T cells and their role in autoimmune disease. *Immunological Reviews*, 212:203–216, 2006.
- [5] C. Baecher-Allan, E. Wolf, and D. A. Hafler. MHC class II expression identifies functionally distinct human regulatory T cells. *The Journal of Immunology*, 176:4622–4631, 2006.
- [6] D. Buljevac, W. C. Hop, W. Reedeker, A. C. Janssens, F. G. van der Meche, P. A. van Doorn, and R. Q. Hintzen. Self reported stressful life events and exacerbations in multiple sclerosis: prospective study. *BMJ*, 327(7416):646, 2003.
- [7] C. Castillo-Chevez and H. R. Thieme. *Asymptotically autonomous epidemic models*. Mathematical Sciences Institute, Cornell University, 1994.
- [8] A. Ciliberto, F. Capuani, and J. J. Tyson. Modeling networks of coupled enzymatic reactions using the total quasi-steady state approximation. *PLoS Comput Biol*, 3(3):0463–0473, 2007.
- [9] G. S. Cooper and B. C. Stroehla. The epidemiology of autoimmune diseases. *Autoimmunity Reviews*, 2(3):119–125, 2003.
- [10] A. DeFranco, R. Locksley, and M. Robertson. *Immunity: The Immune Response to Infectious and Inflammatory Disease*. London: New Science Press Ltd., 2007.

- [11] E. M. Farber, R. H. Mullen, A. H. Jacobs, and L. Nall. Infantile psoriasis: a follow-up study. *Pediatric Dermatology*, 3:237–243, 1986.
- [12] Z. Fehervari and S. Sakaguchi. Control of FOXP3⁺ CD25⁺CD4⁺ regulatory T cell activation and function by dendritic cells. *International Immunology*, 16:1769–1780, 2004.
- [13] D. M. Fergusson, L. J. Horwood, and F. T. Shannon. Early solid feeding and recurrent childhood eczema: A 10-year longitudinal study. *Pediatrics*, 86(4):541–546, 1990.
- [14] J. D. Fontenot and A. Y. Rudensky. A well adapted regulatory contrivance: regulatory T cell development and the forkhead family transcription factor FOXP3. *Nature Immunology*, 6(4):331–337, 2005.
- [15] B. Fritzsching. Naive regulatory T cells: a novel subpopulation defined by resistance toward CD95L-mediated cell death. *Blood*, 108:3371–3378, 2006.
- [16] H. J. Girschick, C. Zimmer, G. Klaus, K. Darge, A. Dick, and H. Morbach. Chronic recurrent multifocal osteomyelitis: what is it and how should it be treated? *Nature Clinical Practice Rheumatology*, 3, 2007.
- [17] D. Hinrichsen and A. J. Pritchard. *Mathematical Systems Theory I: Modelling, State Space Analysis, Stability and Robustness*, volume 48 of *Texts in Applied Mathematics*. Springer, 2nd edition, 2005.
- [18] T. Ito. Two functional subsets of FOXP3⁺ regulatory T cells in human thymus and periphery. *Immunity*, 28:870–880, 2008.
- [19] R. S. Iyer, M. M. Thapa, and F. S. Chew. Chronic recurrent multifocal osteomyelitis: review. *American Journal of Roentgenology*, 196(6 Suppl):S87S91, 2011.
- [20] P. S. Kim, P. P. Lee, and D. Levy. Modeling regulation mechanisms in the immune system. *Journal of Theoretical Biology*, 246(1):33–69, May 2007.
- [21] J. J. Lafaille, K. Nagashima, M. Katsuki, and S. Tonegawa. High incidence of spontaneous autoimmune encephalomyelitis in immunodeficient anti-myelin basic protein T cell receptor transgenic mice. *Cell*, 78(3):399–408, 1994.
- [22] M. Miyara. Functional delineation and differentiation dynamics of human CD4⁺ T cells expressing the FOXP3 transcription factor. *Immunity*, 30:899–911, 2009.
- [23] D. D. Munro. Recurrent subacute discoid lupus erythematosus. *Proceedings of the Royal Society of Medicine*, 56(2):78–79, 1963.
- [24] K. Murphy, P. Travers, and M. Walport. *Janeway's Immunobiology*. New York: Garland, 7th edition, 2007.
- [25] H. D. Ochs, E. Gambineri, and T. R. Torgerson. IPEX, FOXP3 and regulatory T-cells: a model for autoimmunity. *Immunologic Research*, 38(1-3):112–121, 2007.

- [26] P. Onkamo, S. Vninen, M. Karvonen, and J. Tuomilehto. Worldwide increase in incidence of Type I diabetes – the analysis of the data on published incidence trends. *Diabetologia*, 42(12):1395–1403, 1999.
- [27] A. Putnam. Expansion of human regulatory T-cells from patients with type 1 diabetes. *Diabetes*, 58:652–662, 2008.
- [28] G. Roncador. Analysis of FOXP3 protein expression in human CD4⁺CD25⁺ regulatory T cells at the single-cell level. *European Journal of Immunology*, 35:1681–1691, 2005.
- [29] S. Sakaguchi, M. Miyara, C. M. Costantino, and D. A. Hafler. FOXP3⁺ regulatory T cells in the human immune system. *Nature Reviews Immunology*, 10(7):490–500, 2010.
- [30] S. Sakaguchi, T. Yamaguchi, T. Nomura, and M. Ono. Regulatory T cells and immune tolerance. *Cell*, 133:775–787, 2008.
- [31] E. M. Shevach, R. S. McHugh, C. A. Piccirillo, and A. M. Thornton. Control of T-cell activation by CD4⁺ CD25⁺ suppressor T cells. *Immunological Reviews*, 182:58–67, 2001.
- [32] H. L. Smith. *Monotone dynamical systems: an introduction to the theory of competitive and cooperative systems*, volume 41. American Mathematical Soc., 2008.
- [33] A. M. Thornton and E. M. Shevach. CD4⁺CD25⁺ immunoregulatory T cells suppress polyclonal T cell activation in vitro by inhibiting interleukin 2 production. *The Journal of Experimental Medicine*, 188(2):287–296, 1998.
- [34] H. J. van der Vliet and E. E. Nieuwenhuis. IPEX as a result of mutations in FOXP3. *Clinical and Development Immunology*, 2007.
- [35] R. D. van Gaalen and L. M. Wahl. Reconciling conflicting clinical studies of antioxidant supplementation as HIV therapy: a mathematical approach. *BMC Public Health*, 9(Suppl. 1):1–18, 2009.
- [36] N. Velez de Mendizabal, J. Carneiro, R. Sole, J. Goni, J. Bragard, I. Martinez-Forero, S. Martinez-Pasamar, J. Sepulcre, J. Torrealdea, F. Bagnato, J. Garcia-Ojalvo, and P. Villoslada. Modeling the effector - regulatory T cell cross-regulation reveals the intrinsic character of relapses in Multiple Sclerosis. *BMC Systems Biology*, 5(1):114, 2011.
- [37] P. Villoslada, K. Abel, N. Heald, R. Goertsches, S. L. Hauser, and C. P. Genain. Frequency, heterogeneity and encephalitogenicity of T cells specific for myelin oligodendrocyte glycoprotein in naive outbred primates. *European Journal of Immunology*, 31(10):2942–2950, 2001.
- [38] T. Vollmer. The natural history of relapses in multiple sclerosis. *Journal of the Neurological Sciences*, 256, Supplement 1(0):S5–S13, 2007.
- [39] K. Wing, Z. Fehrvri, and S. Sakaguchi. Emerging possibilities in the development and function of regulatory T cells. *International Immunology*, 18(7):991–1000, 2006.

- [40] Y Wu and W. Q. Zhu. Stochastic analysis of a pulse-type prey-predator model. *Physical Review E statistical, nonlinear, and soft matter physics*, 77(4 Pt 1):041911, 2008.
- [41] S. Yamazaki, K. Inaba, K. V. Tarbell, and R. M. Steinman. Dendritic cells expand antigen-specific FOXP3⁺CD25⁺CD4⁺ regulatory T cells including suppressors of alloreactivity. *Immunological Reviews*, 212(1):314–329, 2006.
- [42] W. Yao, L. Hertel, and L. M. Wahl. Dynamics of recurrent viral infection. *Proceedings of the Royal Society-Biological Sciences*, 273(1598):2193–2199, 2006.
- [43] P. Yu. Computation of normal forms via a perturbation technique. *Journal of Sound and Vibration*, 211(1):19–38, 1998.
- [44] P. Yu. Closed-form conditions of bifurcation points for general differential equations. *International Journal of Bifurcation and Chaos*, 15(4):1467–1483, 2005.
- [45] W. Zhang, L. Wahl, and P. Yu. Conditions for transient viremia in deterministic in-host models: Viral blips need no exogenous trigger. *SIAM Journal on Applied Mathematics*, 73(2):853–881, 2013.

Table 3.1: Parameter definitions and values used in Chapter 3.

Para.	Definition	Values
\tilde{v}	per capita rate at which free antigen (G) is taken up by immature pAPCs	0.0025 day^{-1} per molecule of G
f	proportion of antigen molecules that, upon uptake, lead to maturation of the pAPC to enter population A	1×10^{-4}
π_1	rate (per A , per E) at which active nT_{Reg} cells are generated from the pool of ‘naive’ T_{Reg} cells, due to encounter with mature pAPCs (A) and influence of IL-2 from specific effector T cells	$0.0160 \text{ day}^{-1}/E$ per A (π_1)
π_3	rate (per A , per E) at which active nT_{Reg} cells are generated from the pool of ‘naive’ T_{Reg} cells, due to encounter with mature pAPCs (A) and influence of IL-2 from specific effector T cells	$0.0256 \text{ day}^{-1}/E$ per A (π_3)
β	rate (per A) at which active nT_{Reg} cells are generated from the resting pool, due to encounter with mature pAPCs (A) and influence of IL-2 from other sources	$200 \text{ day}^{-1}/A$
λ_E	rate (per A) at which effector T cells (E) are generated from the resting pool, due to encounter with mature pAPCs (A)	$1000 \text{ day}^{-1}/A$
γ	rate (per E) at which self antigen (G) is released due to the actions of effector T cells (E)	$2000 \text{ day}^{-1}/E$
$\sigma_{1,3}$	rate (per capita, R_n or R_d) at which mature pAPCs (A) and effective T cells are effectively eliminated due to suppression by specific active nT_{Reg} cells (R_n) or terminal T_{Reg} cells (R_d)	$3 \times 10^{-6} \text{ day}^{-1}$ per R_n or R_d per A
b_1	rate (per capita) at which mature pAPCs (A) are effectively eliminated due to suppression by T_{Reg} cells of other specificities or by therapy	$0.25 \text{ day}^{-1}/E$
b_3	rate (per capita) at which effective T cells (E) are effectively eliminated due to suppression by T_{Reg} cells of other specificities or by therapy	$0.25 \text{ day}^{-1}/E$
μ_A	per capita death rate of mature pAPCs	$0.2 \text{ day}^{-1}/A$
μ_E	per capita death rate of effector T cells (E)	$0.2 \text{ day}^{-1}/E$
μ_G	per capita rate at which free antigen (G) is cleared, for example due to degradation	$5 \text{ day}^{-1}/G$
μ_n	per capita death rate of active nT_{Reg} cells (R_n)	$0.1 \text{ day}^{-1}/R_n$
μ_d	per capita death rate of terminal T_{Reg} cells (R_d)	$0.2 \text{ day}^{-1}/R_d$
ξ	proportion of activated nT_{Reg} cells	$0.025R_n$
α	rate (per E) at which immature pAPCs become mature	Bifurcation parameter
d	the ratio of suppress effectiveness of nT_{Reg} cells to terminal T_{Reg} cells	2
c	the fold of matured nT_{Reg} cells expansion and proliferation to terminal T_{Reg} cells	$2^3 = 8$

Chapter 4

Backward Bifurcation Underlies Rich Dynamics in Simple Disease Models

4.1 Introduction

In the mathematical modelling of epidemic diseases, the fate of the disease can be predicted through the uninfected and infected equilibria and their stability. The basic reproduction number, R_0 , represents the average number of new infectives introduced into an otherwise disease-free system by a single infective, and is usually chosen as the bifurcation parameter. If the model involves a forward bifurcation, the uninfected equilibrium is in general globally asymptotically stable [28], characterized by $R_0 < 1$, and infection fails to invade in this parameter regime. The threshold $R_0 = 1$ defines a bifurcation (or critical) point, and when $R_0 > 1$, a stable infected equilibrium emerges. This simple exchange of stability implies that complex dynamics will not typically occur in forward bifurcation.

In contrast, backward bifurcation describes a scenario in which a turning point of the infected equilibrium exists in a region where all state variables are positive, and $R_0 < 1$. This induces multiple infected equilibria, disrupting the global stability of the uninfected equilibrium. Multiple stable states (e.g., bistability) may likewise appear [15, 47, 4, 2]. Instead of converging globally to the uninfected equilibrium when $R_0 < 1$, the solution may approach an infected equilibrium, depending on initial conditions.

In practice, the phenomenon of backward bifurcation gives rise to new challenges in disease control, since reducing R_0 such that $R_0 < 1$ is not sufficient to eliminate the disease [22, 5]. Instead, R_0 needs to be reduced past the critical value given by the turning point [22], since the result in [47] shows that the uninfected equilibrium in backward bifurcation is globally stable if R_0 is smaller than the turning point. Furthermore, an infective outbreak or catastrophe may occur if R_0 increases and crosses unity, while the upper branch of the infected equilibrium remains stable [15, 21, 48, 49]. In addition, oscillation or even recurrent phenomena may occur if uninfected and infected equilibria coexist in a parameter range, and both are unstable [48, 49]. Haderl [22] predicted oscillations arising from backward bifurcation, and Brauer [5] pointed out that the unstable infected equilibrium “commonly arises from Hopf bifurcation”, but did not demonstrate oscillations.

Several mechanisms leading to backward bifurcation have been proposed, such as partially

effective vaccination programs [5, 2], educational influence on infectives' behavior [22], the interaction among multi-group models [10, 9, 25] and multiple stages of infection [40]. In this study, we will investigate the emergence of backward bifurcation in three simple disease models which have arisen in the study of epidemiology, in-host disease and autoimmunity. In each case, we find that backward bifurcation facilitates the emergence of Hopf bifurcation(s), and Hopf bifurcation in turn underlies a range of complex and clinically relevant dynamical behaviors.

A central theme in our investigation is the role of the incidence rate in the epidemiological and in-host disease models. The incidence rate describes the speed at which an infection spreads; it denotes the rate at which susceptibles become infectives. Under the assumptions of mass action, incidence is written as the product of the infection force and the number of susceptibles. For example, if S and I denote the susceptible and infective population size respectively, a bilinear incidence rate, $f(S, I) = \beta SI$ (where β is a positive constant), is linear in each of the state variables: S and I .

The possibility of saturation effects [8, 7] has motivated the modification of the incidence rate from bilinear to nonlinear. Saturation occurs when the number of susceptible contacts per infective drops off as the proportion of infectives increases. A nonlinear incidence rate, therefore, typically increases sublinearly with respect to the growth of the infective population, and may finally reach an upper bound. The development of nonlinear incidence was first investigated in the form $\beta I^p S^q$ (where β , p , and q are positive constants), see [32, 31, 23, 24, 13, 29]. Other forms of nonlinear incidence have also been analysed, such as $kI^p S/(1 + \alpha I^l)$ [32], and $kS \ln(1 + vP/k)$ [6].

Since the nonlinear incidence functions described above were often developed to incorporate saturation effects, these functions are typically concave at realistic parameter values. Korobeinikov and Maini [28] used this feature to derive general results for disease models with concave incidence. They proved that standard epidemiological models with concave incidence functions will have globally asymptotically stable uninfected and infected equilibria for $R_0 < 1$ and $R_0 > 1$, respectively.

More specifically, denoting the incidence rate function as $f(S, I, N)$, where N is the population size, the classical SIRS model considered in [28] takes the form

$$\frac{dS}{dt} = \mu N - f(S, I, N) - \mu S + \alpha R, \quad \frac{dI}{dt} = f(S, I, N) - (\delta + \mu)I, \quad \frac{dR}{dt} = \delta I - \alpha R - \mu R, \quad (4.1)$$

where μ , δ , and α represent the birth/death rate, the recovery rate and the loss of immunity rate, respectively. When $\alpha = 0$, system (4.1) becomes an SIR model. Assuming that the total population size is constant, that is, $N = S + I + R$, the above system can be reduced to a 2-dimensional model:

$$\frac{dS}{dt} = \mu N - f(S, I, N) - \mu S, \quad \frac{dI}{dt} = f(S, I, N) - (\delta + \mu)I. \quad (4.2)$$

Moreover, it is assumed in [28] that the function $f(S, I, N)$, denoting the incidence rate, satis-

fies the following three conditions [28]:

$$f(S, 0, N) = f(0, I, N) = 0, \quad (4.3a)$$

$$\frac{\partial f(S, I, N)}{\partial I} > 0, \quad \frac{\partial f(S, I, N)}{\partial S} > 0, \quad \forall S, I > 0 \quad (4.3b)$$

$$\frac{\partial^2 f(S, I, N)}{\partial I^2} \leq 0, \quad \forall S, I > 0. \quad (4.3c)$$

The first two conditions (4.3a) and (4.3b) are necessary to ensure that the model is biologically meaningful. The third condition (4.3c) implies that the incidence rate $f(S, I, N)$, is concave with respect to the number of infectives. It is also assumed that $\frac{\partial f(S, I, N)}{\partial I}$ evaluated at the uninfected equilibrium is proportional to the basic reproduction number R_0 [42], and thus should be a positive finite number [28]. Korobeinikov and Maini first considered $\dot{I} = 0$, or $f(S, I, N) - (\delta + \mu)I = 0$, and showed that forward bifurcation occurs in model (4.2) with a concave incidence function. They further proved that the uninfected equilibrium $Q_0 = (S_0, I_0) = (N, 0)$ and the infected equilibrium $\bar{Q} = (\bar{S}, \bar{I})$ are globally asymptotically stable, when $R_0 = \frac{1}{\delta + \mu} \frac{\partial f(S_0, I_0, N)}{\partial I} < 1$ and $R_0 > 1$, respectively.

In the sections to follow, for an incidence rate function $f(S, I)$, satisfying (4.3a) and (4.3b), we define $f(S, I)$ as concave, if it satisfies (4.3c); as convex, if $\frac{\partial^2 f(S, I)}{\partial I^2} > 0, \forall I > 0$; and as convex-concave, if there exist $0 < I_1 < I_2 \leq +\infty$, such that $\frac{\partial f(S, I)}{\partial I} > 0, \forall I \in (0, I_2)$, and $\frac{\partial^2 f(S, I)}{\partial I^2} > 0, \forall I \in (0, I_1)$, $\frac{\partial^2 f(S, I)}{\partial I^2} = 0$, for $I = I_1$, $\frac{\partial^2 f(S, I)}{\partial I^2} < 0, \forall I \in (I_1, I_2)$.

Several models closely related to (4.2) have been previously studied. For example, by adding a saturating treatment term to model (4.2) with a concave incidence rate, Zhou and Fan [51] showed that this model may yield backward bifurcation and Hopf bifurcation. With an even more sophisticated nonlinear incidence rate function: $kI^p S / (1 + \alpha I^l)$ [38], where $p = l = 2$, Ruan and Wang [38] proved that a reduced 2-dimensional SIRS model could exhibit backward bifurcation, Hopf bifurcation, and even Bogdanov-Takens bifurcation and homoclinic bifurcation. Although the choice of $p = l = 2$ was not motivated by a specific physical process, this important result demonstrates that a nonlinear incidence rate can induce backward bifurcation, and further generate complex dynamics in a simple disease model.

One of the focal points of our study will be a convex incidence function which arose in a 4-dimensional HIV antioxidant therapy model [43]. In this model, the infectivity of infected cells was proposed to be an increasing function of the density of reactive oxygen species, which themselves increase as the infection progresses. In [43], meaningful parameter values were carefully chosen by data fitting to both experimental and clinical results. In this parameter regime, the model was observed to capture the phenomenon of viral blips, that is, long periods of undetectable viral load punctuated by brief episodes of high viral load. Viral blips have been observed clinically in HIV patients under highly active antiretroviral therapy [11, 14, 35, 34], and have received much attention in the research literature, both by experimentalists [17, 18, 20] and mathematicians [16, 27, 12, 37, 36]. Nonetheless, the mechanisms underlying this phenomenon are still not thoroughly understood [20, 36].

We recently re-examined the model developed in [43], with the aim of providing new insight into the mechanism of HIV viral blips [48, 49]. Focusing on the dynamics of the slow manifold of this model, we reduced the dimension of the 4-dimensional model by using quasi-steady state assumptions. After a further generalization and parameter rescaling process, a

2-dimensional in-host HIV model [48, 49] was obtained, given by

$$\frac{dX}{d\tau} = 1 - DX - \left(B + \frac{AY}{Y+C}\right)XY, \quad \frac{dY}{d\tau} = \left(B + \frac{AY}{Y+C}\right)XY - Y, \quad (4.4)$$

where X and Y denote the concentrations of the uninfected and infected cells respectively. The constant influx rate and the death rate of Y have been scaled to 1. The death rate of X is D . The 2-dimensional infection model above (4.4), reduced from the 4-dimensional HIV model [43], preserves the viral blips observed in the HIV model.

Importantly, system (4.4) is equivalent to the SIR model (4.2), except that the incidence function is convex, as we will show in section 4.2.2. This equivalence can be demonstrated if we set $S = e_1x$, $I = e_2y$, and $t = e_3\tau$ with $e_1 = e_2 = \frac{\mu N}{\delta + \mu}$ and $e_3 = \frac{1}{\delta + \mu}$. In this case, system (4.2) is rescaled to

$$\frac{dx}{d\tau} = 1 - \frac{\mu}{\delta + \mu}x - \frac{1}{\mu N}f(x, y), \quad \frac{dy}{d\tau} = \frac{1}{\mu N}f(x, y) - y, \quad (4.5)$$

which takes the same form as system (4.4). Therefore, although system (4.2) arises in epidemiology and system (4.4) was derived as an in-host model, they are mathematically equivalent in this sense. We will refer to both systems (4.2) and (4.4) as infection models.

In previous work [48, 49], we analyze the recurrent behavior which emerges in system (4.4) in some detail. Recurrence is a particular form of oscillatory behavior characterized by long periods of time close to the uninfected equilibrium, punctuated by brief episodes of high infection [45]. Thus HIV viral blips are an example of recurrent behavior, but recurrence is a more general feature of many diseases [45, 49]. We have demonstrated that the increasing and saturating infectivity function of system (4.4) is critical to the emergence of recurrent behaviour. This form of an infectivity function corresponds to a convex incidence rate function in the associated 2-dimensional infection model (4.4), and can likewise induce recurrence in this model. Convex incidence has been previously suggested to model ‘cooperation effects’ in epidemiology [28], or cooperative phenomena in reactions between enzyme and substrate, as proposed by Murray [33].

The rest of this paper is organized as follows. In Section 2, we study two 2-dimensional infection models, both closely related to system (4.2). We show that system (4.2) with either (a) a concave incidence rate and saturating treatment term or (b) a convex incidence rate as shown in system (4.4), can exhibit backward bifurcation; we then identify the necessary terms in the system equations which cause this phenomenon. In Section 3, we demonstrate that in both models, backward bifurcation increases the likelihood of a Hopf bifurcation on the upper branch of the infected equilibrium. Studying system (4.4) in greater detail, we illustrate how the location of the Hopf bifurcations and their directions (supercritical or subcritical), determine the possible dynamical behaviors, concluding that backward bifurcation facilitates Hopf bifurcation(s), which then underly the rich behaviours observed in these models. In Section 4, we explore backward bifurcation further, presenting an autoimmune disease model which exhibits negative backward bifurcation, that is, a bifurcation for which the turning point when $R_0 < 1$ is located in a region where one or more state variables is negative. Although this bifurcation introduces two branches of the infected equilibrium, we demonstrate that, in the biologically feasible area, only forward bifurcation exists in this model. We then present a modification to this autoimmune model, motivated by the recent discovery of a new cell type,

which generates a negative backward bifurcation and Hopf bifurcation, and allows recurrent behavior to emerge. A conclusion is drawn in Section 5.

4.2 Backward bifurcation

In this section, we study backward bifurcation in two 2-dimensional infection models. In particular, we explore the essential terms and parameter relations which are needed to generate backward bifurcation. Furthermore, we examine the convex incidence rate, and reveal its underlying role in determining the emergence of backward bifurcation.

4.2.1 Backward bifurcation in the infection model with concave incidence

First, we consider the SIR model with concave incidence, described by the following equations [51]:

$$\frac{dS}{dt} = \Lambda - \frac{\beta SI}{1+kI} - dS, \quad \frac{dI}{dt} = \frac{\beta SI}{1+kI} - (d + \gamma + \epsilon)I, \quad \frac{dR}{dt} = \gamma I - dR, \quad (4.6)$$

where S , I and R denote the number of susceptible, infective, and recovered individuals, respectively; Λ is the constant recruitment rate of susceptibles; d , γ , and ϵ represent the rates of natural death, recovery, and the disease-induced mortality, respectively. Note that the function $\frac{\beta SI}{1+kI}$ is an incidence rate of the form $\frac{k^l S}{1+\alpha I^h}$ [32], when $l = h = 1$. Here, β is the infection rate, and k measures the inhibition effect. Since the variable R is not involved in the first two equations, system (4.6) can be reduced to a 2-dimensional model as

$$\frac{dS}{dt} = \Lambda - \frac{\beta SI}{1+kI} - dS, \quad \frac{dI}{dt} = \frac{\beta SI}{1+kI} - (d + \gamma + \epsilon)I. \quad (4.7)$$

In [51] an additional assumption regarding limited medical treatment resources is introduced to the above model, leading to a model with a saturating treatment term, given by

$$\frac{dS}{dt} = f_1(S, I) = \Lambda - \frac{\beta SI}{1+kI} - dS, \quad \frac{dI}{dt} = f_2(S, I) = \frac{\beta SI}{1+kI} - (d + \gamma + \epsilon)I - \frac{\alpha I}{\omega + I}, \quad (4.8)$$

where the real, positive parameter α represents the maximal medical resources per unit time, and the real, positive parameter ω is the half-saturation constant. For simplicity, let the functions on the right-hand side of the equations in (4.8) be f_1 and f_2 , respectively. Then, the equilibrium solutions of system (4.8) are obtained by solving the following algebraic equations: $f_1(S, I) = 0$ and $f_2(S, I) = 0$. From which the disease-free equilibrium can be easily obtained as $\bar{E}_0 = (\Lambda/d, 0)$. For the infected equilibrium $\bar{E} = (\bar{S}, \bar{I})$, \bar{S} is solved from $f_1 = 0$ as $\bar{S}(I) = \frac{\Lambda(1+kI)}{(dk + \beta)I + d}$. Then, substituting $S = \bar{S}(I)$ into $f_2 = 0$ yields a quadratic equation of the form

$$\mathcal{F}(I) = \mathcal{A}I^2 + \mathcal{B}I + \mathcal{C} = 0, \quad (4.9)$$

which in turn gives two roots: $\bar{I}_{1,2} = \frac{-\mathcal{B} \pm \sqrt{\mathcal{B}^2 - 4\mathcal{A}\mathcal{C}}}{2\mathcal{A}}$, where, $\mathcal{A} = (d + \gamma + \epsilon)(dk + \beta)$, $\mathcal{B} = [(dk + \beta)\omega + d](d + \gamma + \epsilon) + (dk + \beta)\alpha - \beta\Lambda$, $\mathcal{C} = [(d + \gamma + \epsilon)\omega + \alpha]d - \beta\Lambda\omega$ for system (4.8).

Since all parameters take positive values, we have $\mathcal{A} > 0$. To get the two positive roots essential for backward bifurcation, it is required that $\mathcal{B} < 0$ and $\mathcal{C} > 0$. Noticing that $\beta, \Lambda, \omega > 0$, we can see that the infection force, β , the constant influx of the susceptibles, Λ , and the effect of medical treatment $\frac{\alpha I}{\omega + I}$ are indispensable terms for backward bifurcation. The number of positive infected equilibrium solutions changes from two to one when the value of C passes from negative to positive, which gives a critical point at $C = 0$, that is, $[(d + \gamma + \epsilon)\omega + \alpha]d = \beta\Lambda\omega$, which is equivalent to $R_0 = \frac{\beta\Lambda}{(d + \gamma + \epsilon + \alpha/\omega)d} = 1$.

On the other hand, we may infer the emergence of backward bifurcation without solving the equilibrium conditions. If we do not consider the medical treatment term $\frac{\alpha I}{\omega + I}$ and remove it from system (4.8), that leads to system (4.7), which is a typical example of an SIR model studied by (4.2). By setting the incidence function as $f_3(S, I) = \frac{\beta SI}{1 + kI}$, we have $f_3(0, I) = f_3(S, 0) = 0$; $\frac{\partial f_3(S, I)}{\partial S} = \frac{\beta I}{1 + kI} > 0$ and $\frac{\partial f_3(S, I)}{\partial I} = \frac{\beta S}{(1 + kI)^2} > 0$ for all $S, I > 0$; and $\frac{\partial^2 f_3(S, I)}{\partial I^2} = -2\beta k S (1 + kI)^{-3} < 0$ for all $S, I > 0$. Therefore, the incidence function $f_3(S, I)$, satisfies the conditions given in (4.3). In particular, the function is concave, and can only have one intersection point with the line $(d + \gamma + \epsilon)I$ in the I - S plane, as shown in Figure 4.1(a). Thus, the uniqueness of the positive infected equilibrium implies that backward bifurcation cannot occur in this case. Moreover, according to the result in [28], the uninfected and infected equilibria are globally asymptotically stable for $R_0 = \frac{\beta\Lambda}{d(d + \gamma + \epsilon)} < 1$ and $R_0 > 1$, respectively. No complex dynamical behavior happens in system (4.7).

In contrast, when we introduce the loss of the infectives due to medical treatment, the dynamics of system (4.8) differ greatly from system (4.7). In particular, backward bifurcation emerges and complex dynamical behaviors may occur. To clarify this effect, we denote the function induced by $\dot{I} = 0$ from (4.8) as $f_4(S, I) = \frac{\beta SI}{1 + kI} - \frac{\alpha I}{\omega + I}$. Note that $f_4(S, I)$ is not an incidence rate. But, if we fix $S = \tilde{S} > 0$, there exist $0 < I_1 < I_2 < +\infty$, such that $\frac{\partial f_4(\tilde{S}, I)}{\partial I} = \frac{1}{(1 + kI)^2(\omega + I)^2} [\beta\tilde{S}(\omega + I)^2 - \alpha\omega(1 + kI)^2] > 0, \forall I \in (0, I_2)$; and $\frac{\partial^2 f_4(\tilde{S}, I)}{\partial I^2} = -2k\beta\tilde{S}(1 + kI)^{-3} + 2\alpha\omega(\omega + I)^{-3} > 0, \forall I \in (0, I_1)$, $\frac{\partial^2 f_4(\tilde{S}, I)}{\partial I^2} = 0$, for $I = I_1$, $\frac{\partial^2 f_4(\tilde{S}, I)}{\partial I^2} < 0, \forall I \in (I_1, I_2)$. Thus, $f_4(\tilde{S}, I)$ actually has a convex-concave 'S' shape, and may have two positive intersection points with the ray line, $g_1(I) = (d + \gamma + \epsilon)I$, in the first quadrant; see Figure 4.1(b). These intersections contribute the two positive equilibrium solutions that are a necessary feature of backward bifurcation.

In summary we may conclude that the necessary terms which should be contained in system (4.8) in order to have backward bifurcation are the constant influx Λ , the infection force β , and the saturating medical treatment $\frac{\alpha I}{\omega + I}$.

4.2.2 Backward bifurcation in the infection model with convex incidence

Now we consider the 2-dimensional infection model (4.4) which exhibits viral blips, studied in [48, 49]. The motivation for this model was a series of clinical discoveries indicating that viral infection can increase the density of a harmful chemical substance [19, 30, 39, 26], thereby amplifying an associated biochemical reaction [41], and thus accelerating the infection rate [19]. This cooperative phenomenon in viral infection is expressed by an increasing, saturating infectivity function: $(B + \frac{AY}{Y+C})$. According to the principle of mass action, the incidence function is then denoted as $(B + \frac{AY}{Y+C})XY$, which is a convex function with respect to the infectives' density Y .

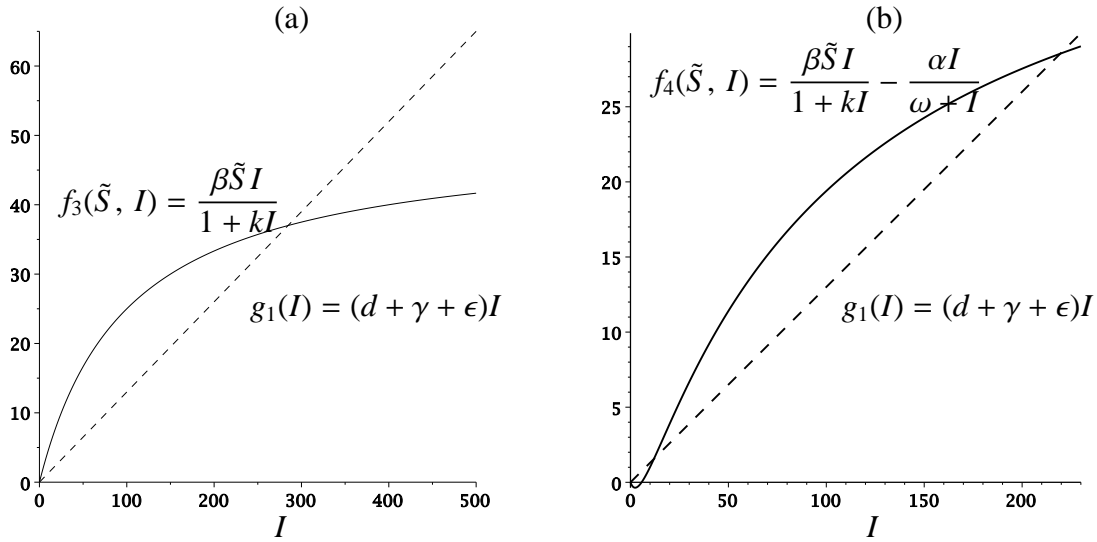


Figure 4.1: Graphs of the incidence function f_3 in system (4.6, 4.7) and function f_4 in system (4.8) with respect to I , for which $\tilde{S} = 50$ has been used. The parameter values are chosen as $\beta = 0.01$, $k = 0.01$, $\alpha = 6$, $\omega = 7$, $d = 0.1$, $\gamma = 0.01$, $\epsilon = 0.02$, according to [51]. The solid lines denote f_3 in (a) and f_4 in (b), while the dashed ray lines in both graphs represent $g_1(I) = (d + \gamma + \epsilon)I$. (a) the incidence function $f_3(S, I) = \frac{\beta SI}{1+kI}$, showing one intersection point with g_1 ; and (b) the function $f_4(S, I) = \frac{\beta SI}{1+kI} - \frac{\alpha I}{\omega+I}$, showing two intersection points with g_1 .

To analyze the occurrence of possible backward bifurcation, we first examine the two equilibrium solutions from the following equations:

$$f_5(X, Y) = 1 - DX - \left(B + \frac{AY}{Y+C}\right)XY = 0, \quad f_6(X, Y) = \left(B + \frac{AY}{Y+C}\right)XY - Y = 0, \quad (4.10)$$

where all parameters A , B , C and D are positive constants. It is easy to find the uninfected equilibrium $\bar{E}_0 = (\bar{X}_0, \bar{Y}_0) = (\frac{1}{D}, 0)$, whose characteristic polynomial has two roots: $\lambda_1 = -D < 0$, and $\lambda_2 = \frac{B}{D} - 1$, which gives $R_0 = \frac{B}{D}$. Consequently, \bar{E}_0 is stable (unstable) for $R_0 < 1$ (> 1). To find the infected equilibrium solution, setting $f_6(X, Y) = 0$ yields $\bar{X}_1(Y) = \frac{Y+C}{(A+B)Y+BC}$, which is then substituted into $f_5(X, Y) = 0$ to give the following quadratic equation:

$$\mathcal{F}_5(Y) = (A+B)Y^2 + (BC + D - A - B)Y + C(D - B) = 0. \quad (4.11)$$

In order to have two real, positive roots, two conditions must be satisfied, that is, $BC + D - A - B < 0$ and $D - B > 0$, or in compact form, $0 < D - B < A - BC$. The condition $D - B > 0$ is equivalent to $0 < R_0 = \frac{B}{D} < 1$, which is a necessary condition for backward bifurcation. Moreover, the positive influx constant, having been scaled to 1, is a necessary term for the positive equilibrium of Y . Therefore, the positive influx rate term and the increasing and saturating infectivity function are necessary for backward bifurcation.

In the rest of the subsection, we further examine the incidence function,

$$f_7(X, Y) = \left(B + \frac{AY}{Y+C}\right)XY, \quad (4.12)$$

without solving the equilibrium solutions. The incidence function f_7 obviously satisfies the condition (4.3a), as well as the condition (4.3b) since $\frac{\partial}{\partial X} f_7(X, Y) = [B + AY(Y+C)^{-1}]Y > 0$

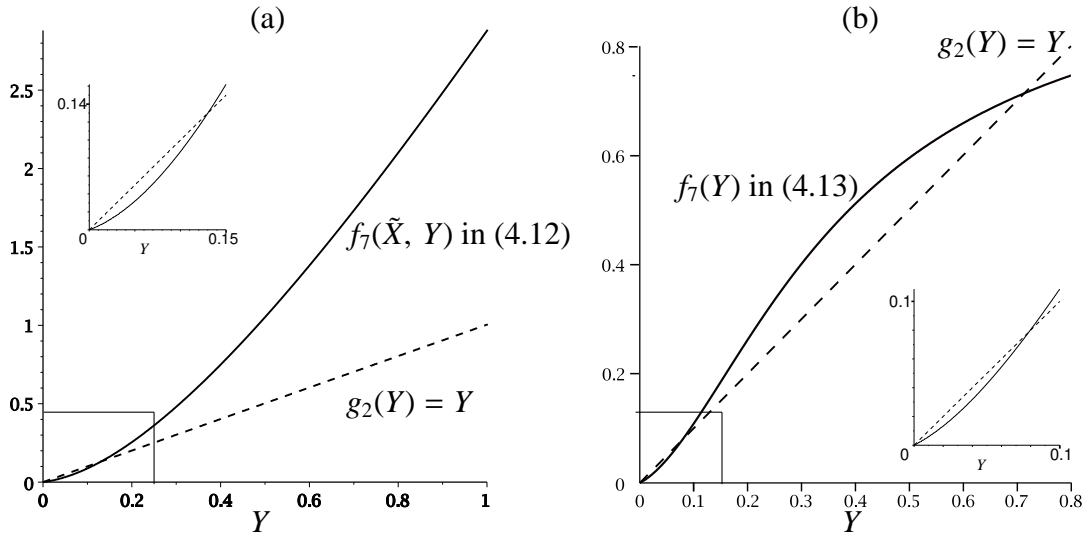


Figure 4.2: Graphs of the incidence functions $f_7(\tilde{X}, Y)$ and $f_7(Y)$ for the parameter values $A = 0.364$, $B = 0.03$, $C = 0.823$, and $D = 0.057$. The incidence functions are denoted by the solid lines, while the ray lines, determined by $g_2(Y) = Y$, are denoted by dotted lines: (a) the incidence function $f_7(\tilde{X}, Y)$, showing one intersection point with g_2 with an inset, with a fixed value $\tilde{X} = 12.54$; and (b) the incidence function $f_7(Y)$, showing two intersection points with an inset.

and $\frac{\partial}{\partial Y} f_7(X, Y) = ACXY(Y + C)^{-2} + [B + AY(Y + C)^{-1}]X > 0$ for all $X, Y > 0$. However, the second partial derivative of $f_7(X, Y)$ with respect to Y , $\frac{\partial^2}{\partial Y^2} f_7(X, Y) = 2AC^2X(X + C)^{-3} > 0$ for all $X, Y > 0$, showing that $f_7(X, Y)$ is a convex function with respect to the variable Y . Consequently, $f_7(X, Y)$ can only have one intersection with $g_2(Y) = Y$, implying that only one equilibrium solution would exist if we only consider the second equation in (4.10), as shown Figure 4.2 (a). However, when considering both conditions given in (4.10) for equilibrium solutions, we will have two intersection points between f_7 and g_2 . According to the first equation in (4.10), that is $f_5(X, Y) = 0$, we can use Y to express X in the equilibrium state as $\tilde{X}(Y) = (Y + C)[(A + B)Y^2 + (BC + D)Y + DC]^{-1}$. Substituting $\tilde{X}(Y)$ into $f_7(X, Y)$ in (4.12), we obtain

$$f_7(Y) = Y[(A + B)Y + BC][(A + B)Y^2 + (BC + D)Y + CD]^{-1}, \quad (4.13)$$

and $\frac{\partial}{\partial Y} f_7(Y) = D[(A + B)Y^2 + 2(A + B)CY + BC^2][(A + B)Y^2 + (BC + D)Y + CD]^{-2} > 0$ for all $X, Y > 0$. However, the sign of $\frac{\partial^2}{\partial Y^2} f_7(Y) = -2D[(A + B)^2Y^3 + 3C(A + B)^2Y^2 + 3(A + B)BC^2Y + (B^2C - AD)C^2][(A + B)Y^2 + (BC + D)Y + CD]^{-3}$, could alter at the inflection point from positive to negative as Y increases. Therefore, with appropriate parameter values, $f_7(Y)$ can have a convex-concave ‘S’ shape, yielding two intersection points with the ray line, $g_2(y)$, in the first quadrant of the X - Y plane, as shown in Figure 4.2 (b). The above discussion, as illustrated in Figure 4.2, implies that system (4.4) can have two positive equilibrium solutions when $R_0 < 1$, and thus backward bifurcation may occur.

Remark 1. Summarizing the discussions and results given in this section indicates that a disease model with a convex-concave incidence function may lead to backward bifurcation, which in turn implies: (a) the system has at least two equilibrium solutions, and the two equilibrium solutions intersect at a transcritical bifurcation point; and (b) at least one of the equilibrium

solutions is determined by a nonlinear equation.

4.3 Hopf bifurcation

In the previous section, we studied backward bifurcation and established the necessary conditions for the occurrence of backward bifurcation in two models. In this section, we turn to Hopf bifurcation, since it typically underlies the change of stability in the upper branch of the infected equilibrium, the key condition in determining whether a model can exhibit oscillation or even recurrence. Again, we will present detailed studies for the two models.

4.3.1 Hopf bifurcation in the infection model with concave incidence

In this subsection, we study two cases of an infection model with concave incidence: system (4.7) and (4.8). First, we discuss the equilibrium solutions and their stability by using the Jacobian matrix, denoted by J , and examining the corresponding characteristic polynomial,

$$P|_J(L) = L^2 + \text{Tr}(J)L + \text{Det}(J). \quad (4.14)$$

Bifurcation analysis is conducted by choosing Λ as the bifurcation parameter.

First, we consider the case without saturating medical treatment, system (4.7). This system satisfies the three conditions in (4.3), and consequently, its uninfected equilibrium $\bar{E}_0 = (\frac{\Lambda}{d}, 0)$ is globally asymptotically stable if $R_0 = \frac{\beta\Lambda}{(d+\gamma+\epsilon)d} \leq 1$, while the infected equilibrium $\bar{E}_1 = (\frac{k\Lambda+d+\gamma+\epsilon}{dk+\beta}, \frac{\beta\Lambda-(d+\gamma+\epsilon)d}{(dk+\beta)(d+\gamma+\epsilon)})$ emerges and is globally asymptotically stable if $R_0 > 1$. Therefore, for this case the system has only one transcritical bifurcation point at $R_0 = 1$ and no complex dynamics can occur.

Next, with the saturating treatment term, system (4.8) violates the conditions established for model (4.3), but leads to the possibility of complex dynamical behaviors. In fact, evaluating the Jacobian matrix $J_1 = J|_{(4.8)}(\bar{E}_0)$ at the uninfected equilibrium, $\bar{E}_0 = (\frac{\Lambda}{d}, 0)$, yields the characteristic polynomial in the form of (4.14), denoted by $P|_{J_1}(L)$, with $\text{Tr}(J_1) = (-\frac{\beta\Lambda}{d} + \epsilon + \frac{\alpha}{\omega} + 2d)$, and $\text{Det}(J_1) = (-\beta\Lambda + d^2 + d\epsilon + \frac{\alpha d}{\omega}) = \text{Tr}(J_1)d - d^2$. This indicates that $\text{Det}(J_1) < 0$ when $\text{Tr}(J_1) = 0$, and thus Hopf bifurcation cannot occur from \bar{E}_0 . On the other hand, a static bifurcation can occur when $\text{Det}(J_1) = 0$, that is, $\Lambda_S = \frac{1}{\beta}(d^2 + d\epsilon + \frac{\alpha d}{\omega})$, where the subscript ‘S’ refers to *static bifurcation*. Therefore, \bar{E}_0 is stable (unstable) for $\Lambda < \Lambda_S$ ($> \Lambda_S$), or $R_0 < 1$ (> 1), with $R_0 = \beta\Lambda d^{-1}(d + \gamma + \epsilon + \frac{\alpha}{\omega})^{-1}$ [51].

We will show that complex dynamical behaviors can emerge in system (4.8) from the infected equilibrium $\bar{E}_1 = (\bar{S}, \bar{I})$, where \bar{I} is determined from the equation $\mathcal{F}(I) = 0$ in (4.9). In the Λ - I plane, the bifurcation diagram as shown in Figure 4.3 (1)-(4), indicates a turning point on the curve with appropriate parameter values, determined by both the quadratic equation (4.9) and the relation $\frac{d\Lambda}{dI} = -\frac{\partial\mathcal{F}}{\partial I} / \frac{\partial\mathcal{F}}{\partial\Lambda} = 0$, which is equivalent to $\frac{\partial\mathcal{F}}{\partial I} = 0$. Solving $\frac{\partial\mathcal{F}}{\partial I} = 0$ yields the turning point of I , denoted by I_T (‘T’ means *turning*), taking the form

$$I_T = \frac{1}{2} \left[\frac{\beta\Lambda_T}{(dk + \beta)(d + \epsilon)} - \omega - \frac{d}{dk + \beta} - \frac{\alpha}{d + \epsilon} \right], \quad (4.15)$$

where Λ_T is obtained from $\mathcal{F}(I_T) = 0$, see (4.9). Thus, when $I_T > 0$ (< 0), the turning point of the quadratic curve appears above (below) the I -axis, meaning that backward bifurcation occurs for $I > 0$ (< 0). Evaluating the Jacobian matrix at the infected equilibrium \bar{E}_1 , and further denoting it as $J_2 = J_{|(4.8)}(\bar{E}_1)$, we obtain the characteristic polynomial in the form of (4.14), with $\text{Tr}(J_2) = a_{11}/[(\omega+I)^2(kI+1)(dkI+\beta I+d)]$ and $\text{Det}(J_2) = a_{21}/[(\omega+I)^2(kI+1)(dkI+\beta I+d)]$, where $a_{11} = a_{1a} - a_{1b}$ and $a_{21} = a_{2a} - a_{2b}$, with $a_{1b} = \beta\Lambda(\omega+I)^2$ and $a_{2b} = da_{1b}$, and a_{1a} and a_{2a} only contain positive terms (their expressions are omitted here for brevity). Therefore, we can rewrite $\text{Det}(J_2) = \frac{a_{21}}{d}/[(\omega+I)^2(kI+1)(dkI+\beta I+d)]$. Determining whether a Hopf bifurcation can occur from \bar{E} is equivalent to finding whether $\text{Det}(J_2)$ remains positive when $\text{Tr}(J_2) = 0$. Ignoring the positive factors in the following subtraction yields

$$h_1(I) = \frac{\text{Tr}(J_2) - \text{Det}(J_2)/d}{(\omega+I)^2(kI+1)(dkI+\beta I+d)} = a_{11} - \frac{1}{d}a_{21} = a_{1a} - \frac{1}{d}a_{2a}, \quad (4.16)$$

where $h_1(I) = \frac{1}{d}(dkI + \beta I + d)[(kI + 1)d^2(\omega + I)^2 - \beta\epsilon I(\omega + I)^2 - \alpha\beta\omega I]$. Thus, when $a_{1a} = 0$, $\frac{1}{d}a_{2a}$ and $h_1(I)$ have opposite signs, implying that when $\text{Tr}(J_2) = 0$, $\text{Det}(J_2)$ could be positive only if $h_1(I)$ is negative. Therefore, the necessary condition for system (4.8) to have a Hopf bifurcation from the infected equilibrium \bar{E}_1 is that $h_1(I)$ is negative.

In the remaining part of this subsection, we demonstrate various dynamics which may happen in system (4.8) with different parameter values of k , as shown in Table 4.1. Taking other parameter values as $\alpha = 6$, $\omega = 7$, $\epsilon = 0.02$, $\gamma = 0.01$, $\beta = 0.01$, and $d = 0.1$, and solving the two equations $\text{Tr}(J_2) = 0$ and $\mathcal{F}(I) = 0$ in (4.9) gives the Hopf bifurcation point candidates, (Λ_H, I_H) , for which $h_1(I_H) < 0$. Since the formula for the transcritical bifurcation point Λ_S has no relation with k , $(\Lambda_S, I_S) = (9.87, 0)$ is a fixed value pair in Table 4.1. Bifurcation diagrams and associated numerical simulations are shown in Figure 4.3 corresponding to the five cases given in Table 4.1. The blue lines and red curves represent the uninfected equilibrium \bar{E}_0 and infected equilibrium \bar{E}_1 , respectively. The stable and unstable equilibrium solutions are shown by solid and dashed lines/curves, respectively. Backward bifurcation occurs in Cases 1, 2, and 3 (see Table 4.1), which are illustrated by the corresponding bifurcation diagrams in Figures 4.3(1), (2), and (3), respectively. For Cases 1 and 2, only one Hopf bifurcation occurs on the upper branch of the infected equilibrium \bar{E}_1 , and this bifurcation point exists at the critical point $\Lambda_H < \Lambda_S$ for Case 1 and $\Lambda_H > \Lambda_S$, for Case 2. For Case 1 with $\Lambda = 9.78$, the simulated time history converges to \bar{E}_0 with initial condition IC = [93.6, 0.44], shown in Figure 4.3(1a), but converges to \bar{E}_1 with initial condition IC = [46.8, 10], shown in Figure 4.3(1b). This clearly indicates the bistable behavior when $\Lambda_H < \Lambda_S$, and an overlapping stable region for both \bar{E}_0 and \bar{E}_1 exists (see Figure 4.3(1)). The recurrent behavior for Case 2 is simulated at $\Lambda = 9.87$ with IC = [50, 5], shown in Figure 4.3(2a). For Case 2, $\Lambda_H > \Lambda_S$, and an overlapping unstable parameter region for both \bar{E}_0 and \bar{E}_1 occurs between Λ_S and Λ_H (see Figure 4.3(2)). For Case 3, two Hopf bifurcations occur on the left side of Λ_S , and a stable part in the upper branch of \bar{E}_1 exists when Λ passes through the critical value $\Lambda = \Lambda_S$. In this case, although backward bifurcation still exists and the turning point is also located above the Λ -axis, giving two branches of biologically feasible \bar{E}_1 , only regular oscillating behavior is observed. The simulated time history is conducted at $\Lambda = 10$, with initial condition IC = [50, 2], shown in Figure 4.3(3a). For Case 4, only forward bifurcation occurs in the biologically feasible region, and the turning point for backward bifurcation moves down to the fourth quadrant, that is, negative backward bifurcation occurs in this case. The whole upper branch of \bar{E}_1 in the first quadrant is stable,

Table 4.1: Dynamics of system (4.8) for different values of k ,
with $\alpha = 6$, $\omega = 7$, $\epsilon = 0.02$, $\gamma = 0.01$, $\beta = 0.01$, $d = 0.1$,
and a fixed transcritical bifurcation point $(\Lambda_S, I_S) = (9.87, 0)$.

Case	k	(Λ_T, I_T)	$h_1(I) < 0$	(Λ_H, I_H)	Dynamics	Notes
1	0.001	(9.48, 4.57)	$I \in [1.72, \infty]$	(9.73, 10.28)	Bistability	$\Lambda_H < \Lambda_S$
2	0.01	(9.71, 2.82)	$I \in [1.76, \infty]$	(9.96, 8.00)	Recurrence	$\Lambda_H > \Lambda_S$
3	0.02	(9.85, 0.84)	$I \in [1.82, \infty]$	(9.88, 2.09), (10.14, 5.62)	Oscillation	Two Hopf critical points
4	0.027	(9.86, -0.65)	$I \in [1.85, 30.65]$	No Hopf	No oscillation	Negative backward bifurcation
5	0.05	No Turning	$I \in [2.01, 15.03]$	(6.18, -22.15)	No oscillation	No backward bifurcation

therefore, no oscillations (or recurrence) can happen. Finally, further increases to the value of k change the shape of the red curves, as shown in Figure 4.3(5), which again indicates that no biologically meaningful backward bifurcation or oscillations can occur. Note that in Figure 4.3(5) a Hopf bifurcation point exists on the lower branch of the equilibrium solution, which is biologically unfeasible since it is entirely below the horizontal axis. In conclusion, interesting dynamical behaviors can emerge in system (4.8) if backward bifurcation occurs.

4.3.2 Hopf bifurcation in the infection model with convex incidence

In this subsection, we return to system (4.4), that is, the 2-dimensional HIV model with convex incidence derived in [48, 49], and analyze the various dynamical phenomena which system (4.4) could possibly exhibit. To achieve this, we set B as the bifurcation parameter, and A as a control parameter; the bifurcation analysis will be carried out for various values of A . Also, simulated time histories are provided to illustrate the dynamical behavior predicted in the analysis.

We first consider the uninfected equilibrium $\bar{E}_0 = (\frac{1}{D}, 0)$, which has two eigenvalues. One of them, given by $\lambda_1|_{\bar{E}_0} = -D$, is always negative. The other one is $\lambda_2|_{\bar{E}_0} = \frac{B}{D} - 1$. Thus, depending upon the relation between B and D , $\lambda_2|_{\bar{E}_0} = 0$ gives a static bifurcation at $B_S = D$ (or $R_0 = \frac{B}{D} = 1$), which is further proved to be a transcritical bifurcation. Here the ‘S’ in subscript stands for *static bifurcation*. Therefore, \bar{E}_0 is stable when $B < D$ (or $R_0 < 1$), loses its stability and becomes unstable when B increases to pass through $B_S = D$, that is $B > D$ (or $R_0 > 1$), and no other bifurcations can happen.

Next, we examine the infected equilibrium $\bar{E}_1 = (\bar{X}, \bar{Y})$. Since $\bar{X}(Y) = \frac{Y+C}{(A+B)Y+BC}$, \bar{Y} is determined by the quadratic equation (4.11), which gives the turning point (B_T, Y_T) as

$$B_T = \frac{-A + D + 2\sqrt{ACD}}{C + 1}, \quad Y_T = \frac{A + B - BC - D}{A + B}, \quad (4.17)$$

where ‘T’ in the subscript stands for *turning bifurcation*. We perform a further bifurcation

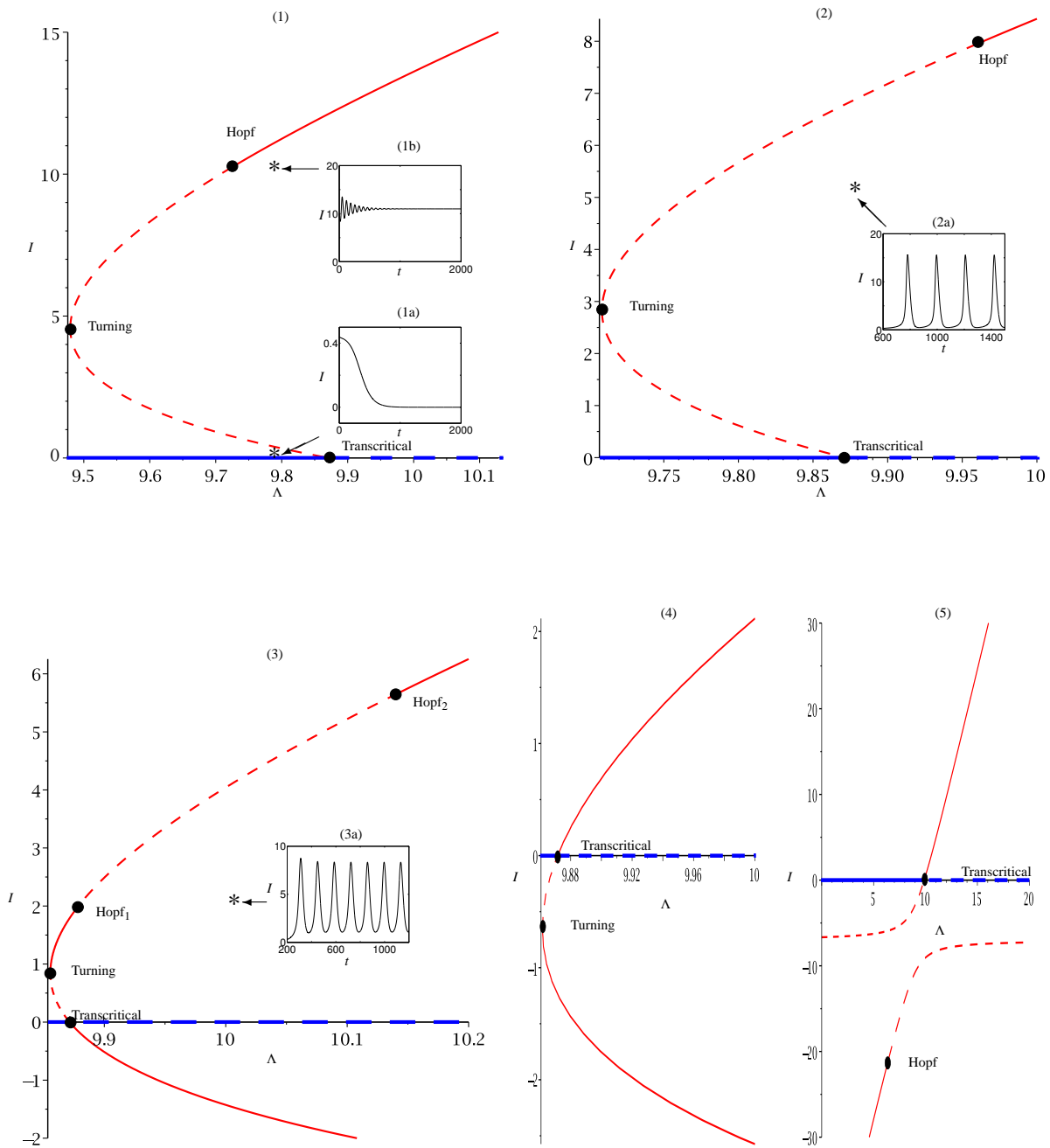


Figure 4.3: Bifurcation diagrams and simulations associated with the five cases given in Table 4.1, demonstrating various dynamical behaviors.

analysis on its corresponding characteristic polynomial (4.14), which takes the form

$$P|_{\bar{E}_1}(\lambda, Y) = \lambda^2 + \frac{a_{1a}}{[(A+B)Y+BC](Y+C)}\lambda + \frac{a_{2a}}{[(A+B)Y+BC](Y+C)}, \quad \text{where}$$

$$a_{1a} = (A+B)^2Y^3 + (2BC+D)(A+B)Y^2 + (B^2C^2+ACD+2BCD-AC)Y + BC^2D,$$

$$a_{2a} = (A+B)^2Y^3 + 2(A+B)BCY^2 + (B^2C-AD)CY. \quad (4.18)$$

Therefore, the sign of the subtraction between the trace and determinant is determined by $h_2(Y) = a_{1a} - a_{2a} = D(A+B)Y^2 + [2CD(A+B) - AC]Y + BC^2D$. Here the equilibrium solution of Y and other parameters satisfy the quadratic equation (4.11), which leads to an explicit expression, given by $\bar{B} = -\frac{AY^2+(D-A)Y+CD}{Y^2+(C-1)Y-C}$. Substituting $B = \bar{B}$ into $h_2(Y)$, we obtain

$$h_2(Y)|_{B=\bar{B}} = a_{1a} - a_{2a} = \frac{[AC(D-1) - D^2]Y^2 - [AC(D-1) + 2CD^2]Y - C^2D^2}{Y-1}. \quad (4.19)$$

Hopf bifurcation may occur when the trace is zero, while the determinant is still positive. This implies $h_2(Y) < 0$, which is possible with appropriately chosen parameter values. Hence, by solving $a_{1a} = 0$ in (4.18) together with the quadratic equation (4.11), we get two pairs of points denoted by (B_{h1}, Y_{h1}) and (B_{h2}, Y_{h2}) , which are candidates for Hopf bifurcation. Then validating the above two points by substituting them back into the characteristic polynomial (4.18), respectively, we denote the Hopf bifurcation point as (B_H, Y_H) if this validation confirms their existence. According to [47], Hopf bifurcation can happen only from the upper branch of the infected equilibrium \bar{E}_1 .

The various dynamical behaviors which may appear in system (4.4) have been classified in Table 4.2 for different values of the parameter A , with fixed values of $C = 0.823$ and $D = 0.057$. Thus, the transcritical bifurcation point is fixed for all cases: $B_S = D = 0.057$ and $Y_S = 0$. The two solutions B_{h1} and B_{h2} are solved from the two equations (4.18) $P|_{\bar{E}_1}(\lambda, Y) = 0$ and (4.11) $\mathcal{F}_5(Y) = 0$, respectively. They become a Hopf bifurcation point only if their corresponding Y values (Y_{h1} and Y_{h2} , respectively) are in the range such that $h_2(Y) < 0$. Otherwise, system (4.4) has a pair of real eigenvalues with opposite signs at (B_{h1}, Y_{h1}) or (B_{h2}, Y_{h2}) , which is denoted by the superscript ‘*’ (which is actually a saddle point) in Table 4.2, while the Hopf bifurcation point is denoted by the superscript ‘H’ in Table 4.2.

Next, we further examine the direction of the Hopf bifurcation, that is, check whether it is a supercritical or subcritical Hopf bifurcation. Since the Jacobian matrix of the system evaluated at the Hopf bifurcation point has a pair of purely imaginary eigenvalues, the linearized system (4.4) does not determine the nonlinear behavior of the system. Therefore, we take advantage of normal form theory to study the existence of the limit cycles bifurcating from the Hopf bifurcation point as well as their stability. As mentioned earlier, Hopf bifurcation can only occur from the upper branch of the infected equilibrium \bar{E}_1 , therefore we first transform the fixed point \bar{E}_1 to the origin by a shifting transformation, and, in addition, make the parameter transformation $B = B_H + \mu$; the Hopf bifurcation point is thus defined as $\mu = \mu_H = 0$. Then the normal form of system (4.4) near the critical point, $\mu = \mu_H = 0$, takes the form up to third-order approximation:

$$\dot{r} = d\mu r + ar^3 + O(r^5), \quad \dot{\theta} = \omega_c + c\mu + br^2 + O(r^4), \quad (4.20)$$

Table 4.2: Parameter values taken to illustrate various dynamics of system (4.4).

The fixed transcritical bifurcation point: $(B_S, Y_S) = (0.057, 0)$

Case	A	(B_T, Y_T)	$h_2(Y) < 0, Y \in$	(B_{h1}, Y_{h1})	(B_{h2}, Y_{h2})	Dynamics	Notes
1	0.80	(-0.1950, 0.5850)	(0.0036, 0.9830)	(0.0355, 0.8725) ^H	(0.054, 0.0034)*	Unstable limit cycle, Bistable	$B_{h1} < B_S$
2	0.71	(-0.1580, 0.5660)	(0.0040, 0.9800)	(0.0539, 0.0038)*	(0.0574, 0.8650) ^H	Recurrence	$B_{h2} > B_S$
3	0.60	(-0.1140, 0.5380)	(0.0048, 0.9769)	(0.0540, 0.0045)*	(0.0819, 0.8530) ^H	Recurrence	$B_{h2} > B_S$
4	0.07	(0.0557, 0.0909)	(0.0476, 0.8030)	(0.0560, 0.0470)*	(0.1015, 0.5612) ^H	Recurrence	$B_{h2} > B_S$
5	0.06	(0.056558, 0.05581)	(0.0574, 0.7700)	(0.056559, 0.0574) ^H	(0.0961, 0.5225) ^H	Recurrence	$B_{h1} < B_S$ < B_{h2}
6	0.05	(0.05697, 0.01442)	(0.0724, 0.7232)	(0.0574, 0.0741) ^H	(0.0894, 0.4701) ^H	Recurrence	$B_{h1} < B_S$ < B_{h2}
7	0.04	(0.0569, -0.0358)	(0.0986, 0.6507)	(0.0592, 0.1071) ^H	(0.0806, 0.3897) ^H	Oscillation	$B_{h1} < B_S$ < B_{h2} $Y_T < 0$
8	0.03	(0.0559, -0.0994)	(0.1611, 0.5149)	—	—	\bar{E}_1 stable	$Y_T < 0$

Table 4.3: Classification of Hopf bifurcations based on the normal form (4.20).

Class	Stability of $\bar{r} = 0$		Stability of $\bar{r}^2 = -\frac{d\mu}{a}$		Hopf bifurcation
	$\mu < 0$	$\mu > 0$	$\mu < 0$	$\mu > 0$	
(a): $d > 0, a > 0$	stable	unstable	unstable	—	subcritical
(b): $d > 0, a < 0$	stable	unstable	—	stable	supercritical
(c): $d < 0, a > 0$	unstable	stable	—	unstable	subcritical
(d): $d < 0, a < 0$	unstable	stable	stable	—	supercritical

where r and θ represent the amplitude and phase of the motion, respectively. The first equation of (4.20) can be used for bifurcation and stability analysis, while the second equation of (4.20) can be used to determine the frequency of the bifurcating periodic motions. The positive ω_c in the second equation of (4.20) is the imaginary part of the eigenvalues at the Hopf bifurcation point. The parameters d and c can be easily obtained from a linear analysis, while a and b must be derived using a nonlinear analysis, with the Maple program available in, say, [46].

Note that the infected equilibrium \bar{E}_1 is represented by the fixed point $\bar{r} = 0$ of system (4.20), while the nonzero fixed point $\bar{r} > 0$ (satisfying $\bar{r}^2 = -\frac{d\mu}{a}$) is an approximate solution for a limit cycle or periodic orbit. The periodic orbit is asymptotically stable (unstable) if $a < 0$ ($a > 0$), and the corresponding Hopf bifurcation is called supercritical (subcritical). According to the Poincare-Andronov Hopf Bifurcation theorem [44], for μ sufficiently small, there are four possibilities for the existence of periodic orbits and their stability, which are classified in Table 4.3, based on the four sets of the parameter values in the normal form (4.20). Then we use the results presented in Table 4.3 with a nonlinear analysis based on normal form theory to classify the Hopf bifurcations appearing in Table 4.2, and the results are shown in Table 4.4.

To illustrate the analytical results given in Tables 4.2 and 4.4, we provide the bifurcation diagrams in Figures 4.4 (1)-(8). These figures depict the uninfected equilibrium \bar{E}_0 and the infected equilibrium \bar{E}_1 in blue and red, respectively. The solid and dashed lines differentiate stable and unstable states of the equilibrium solutions. The bifurcation points on the

Table 4.4: Classification of Hopf bifurcations appearing in Table 4.2.

Case	A	Hopf bifurcation point (B_H, Y_H)	d	a	Stability of limit cycles	Table 4.3 class
1	0.8	(0.0355, 0.8725)	-1.0722	0.2114×10^{-2}	Unstable	(c)
2	0.71	(0.0574, 0.8650)	-1.0726	0.1424×10^{-2}	Unstable	(c)
3	0.6	(0.0819, 0.8530)	-1.0733	0.6755×10^{-3}	Unstable	(c)
4	0.07	(0.1015, 0.5612)	-1.0307	-0.8791×10^{-3}	stable	(d)
5	0.06	(0.056559, 0.0574)	884.27	-0.1019	Stable	(b)
		(0.0961, 0.5225)	-1.0079	-0.8613×10^{-3}	Stable	(d)
6	0.05	(0.0574, 0.0741)	18.232	-0.3145×10^{-2}	Stable	(b)
		(0.0894, 0.4701)	-0.9629	-0.8457×10^{-3}	Stable	(d)
7	0.04	(0.0592, 0.1071)	4.7242	-0.1577×10^{-2}	Stable	(b)
		(0.0805, 0.3897)	-0.8437	-0.8438×10^{-3}	Stable	(d)

equilibrium solutions are highlighted by solid black dots. Moreover, ‘Transcritical’, ‘Turning’, ‘Hopf_{sub}’, and ‘Hopf_{super}’, are used to denote *Transcritical bifurcation*, *Turning point*, *subcritical Hopf bifurcation*, and *supercritical Hopf bifurcation*, respectively. Simulated time histories are used to validate the analytical results, and to show different dynamical behaviors in each case listed in Tables 4.2 and 4.4. Subcritical Hopf bifurcation occurs in Cases 1-3, shown in Figures 4.4 (1)-(3). $A = 0.8$ is used in Figure 4.4 (1) for Case 1. Choosing $B = 0.036$, we have $E_0 = [17.1282566, 0.023689]$ and $E_1 = [2.233533, 0.8726886]$. The simulated solution converges to E_0 or E_1 , with initial condition taken as $IC_d = [17.13, 0.024]$ or $IC_c = [2.233, 0.873]$, shown in Figures 4.4 (1d) and (1c), respectively. Figures 4.4 (1a) and (1b), on the other hand, show the unstable limit cycle bifurcating from the subcritical Hopf bifurcation with $IC_c = [2.233, 0.873]$.

Figure 4.4 (2) corresponds to Case 2 with $A = 0.71$. Choosing $B = 0.0572 \in [B_S, B_H]$ yields recurrence, independent of the initial conditions, see, for example, the result given in Figure 4.4 (2b) with $IC_b = [2.4, 0.5]$. However, for $B = 0.06 > B_H$, the simulated time history converges to E_1 , with an initial condition close to E_1 , such as $IC_a = [2.4, 0.6]$ as shown in Figure 4.4 (2a); or shows recurrence with an initial condition far away from E_1 , such as $IC_c = [2.4, 0.4]$, as shown in Figure 4.4 (2c).

Figure 4.4 (3) plots the result for Case 3 with $A = 0.6$, and shows a broader region between the transcritical and Hopf bifurcation points, associated with a larger recurrent region. Recurrence occurs independent of the initial conditions for $B = 0.083 \in [B_S, B_H]$, giving $E_0 = [12.048, 0]$ and $E_1 = [2.576, 0.852]$, as shown in Figures 4.4 (3a) and (3b), with $IC_a = [2.7, 0.84]$ and $IC_b = [14, 0.1]$, respectively. But if we choose $B = 0.07 > B_H$, we have $E_0 = [14.286, 0]$ and $E_1 = [2.67, 0.8478]$. The time history converges to E_1 with $IC_c = [2.6, 0.8]$, or shows recurrence with $IC_d = [2.6, 0.1]$, as shown in Figure 4.4 (3c) and (3d), respectively.

Supercritical Hopf bifurcations occur in Cases 4-7, as shown in Figures 4.4 (4)-(7). Figure 4.4 (4) depicts the result for Case 4 with $A = 0.07$. Only one supercritical Hopf bifurcation happens in this case, and gives a large recurrent parameter region between the transcritical and Hopf bifurcation points. Although the simulated recurrent behavior does not depend on initial conditions, the recurrent pattern will fade out with the growth of the value of B from the transcritical point to the Hopf bifurcation point, see Figures 4.4 (4a) and (4b) with the same $IC_{a,b} = [8, 0.1]$, but different values of B : $B = 0.06$ and $B = 0.09$, respectively.

Figure 4.4 (5) shows the result for Case 5 with $A = 0.06$. A transcritical bifurcation happens between two supercritical Hopf bifurcations. The recurrent region still starts from the transcritical point and independent of the initial conditions, but is narrower than that shown in Figure 4.4 (4). The simulated recurrent behavior for this case is conducted at $IC = [12, 0.1]$ and $B = 0.06$. Figure 4.4 (6) corresponds to Case 6 with $A = 0.05$, and two supercritical Hopf bifurcations occur on the right side of the transcritical bifurcation point, which makes the recurrent region even narrower and the recurrent pattern less obvious, as shown in the simulated time history with $IC = [10, 0.1]$ and $B = 0.06$. Negative backward bifurcations occur in Cases 7 and 8, as shown in Figure 4.4 (7) and (8). Although two Hopf bifurcations are still present in Case 7, see Figure 4.4 (7), only a regular oscillating pattern exists. For Case 8, no Hopf bifurcation happens in the biologically feasible part of E_1 , and therefore no more interesting dynamics occur.

In general, backward bifurcation, which occurs above the horizontal axis, is much more likely to induce Hopf bifurcation. A Hopf bifurcation can only occur along the upper branch of \bar{E}_1 , since \bar{E}_0 only changes its stability at a transcritical bifurcation point, and any point on the lower branch of \bar{E}_1 is a saddle node [47]. Moreover, Hopf bifurcation can lead to a change in the stability of the upper branch of the infected equilibrium \bar{E}_1 . Thus the system further develops bistable, recurrent, or regular oscillating behavior, corresponding to Cases 1 – 7 in Tables 4.2 and 4.4, and in Figures 4.4 (1)-(7). In particular, bistability happens when both equilibria \bar{E}_0 and \bar{E}_1 share a stable parameter region, see Case 1 in Table 4.2 and Figure 4.4 (1).

As for recurrent behavior, we observe that recurrence is more likely to happen if the following two conditions are satisfied for the upper branch of \bar{E}_1 : (1) the equilibrium remains unstable as the bifurcation parameter increases and crosses the transcritical point, where \bar{E}_0 and \bar{E}_1 intersect, such that the two equilibria share an unstable parameter range; and (2) at least one Hopf bifurcation occurs from \bar{E}_1 . As shown in Cases 2-6 in Table 4.2, and the corresponding Figures 4.4 (2)-(5), the common recurrent parameter region for both subcritical and supercritical Hopf bifurcations starts beside the transcritical point, and is located entirely in the unstable parameter region of \bar{E}_0 and \bar{E}_1 . The simulated recurrent pattern becomes more pronounced if the value of the bifurcation parameter is close to the transcritical point, but approaches an oscillatory pattern as the parameter diverges from the transcritical point, as shown in Figure 4.4 (4a) and (4b). In this common recurrent parameter region, recurrence occurs independent of initial conditions; see Figures 4.4 (3a) and (3b). In addition to the common recurrent region, for subcritical bifurcation, seen in Table 4.2 for Cases (2) and (3) and Figures 4.4 (2) and (3), recurrence may also appear on the stable side of the subcritical Hopf bifurcation point with an initial condition close to \bar{E}_1 . Moreover, the subcritical Hopf bifurcation and the transcritical point should be close to each other for a clear recurrent pattern. When this is not the case, the periodic solutions show a more regular oscillating pattern, as compared in Figures 4.4 (2c) and

(3d). Although two Hopf bifurcation points occur in Table 4.2 for Case 5, see Figure 4.4 (5), the transcritical point is located inside the unstable range of the upper branch of \bar{E}_1 , between the two Hopf bifurcation points. A recurrent pattern still characterizes the dynamical behavior in this case. However, if the unstable range of \bar{E}_1 , between the two Hopf bifurcation points, is located entirely in the unstable range of \bar{E}_0 , and moves further away from the transcritical point, the recurrent motion gradually becomes a regular oscillation, as shown in Figures 4.4 (6) and (7).

Summarizing the results and discussions presented in the previous two sections, we have the following observations.

1. Due to the fact that \bar{E}_0 only changes its stability at the transcritical bifurcation point, and the fact that any point on the lower branch of \bar{E}_1 is a saddle node, Hopf bifurcation can only occur from the upper branch of \bar{E}_1 . A Hopf bifurcation may result in convergent, recurrent, bistable, or regular oscillating behaviors.
2. Backward bifurcation gives rise to two branches in the infected equilibrium \bar{E}_1 . Hopf bifurcation is more likely to happen when the turning point of the backward bifurcation is located on the positive part of the equilibrium solution in the bifurcation diagram, as shown in Figures 4.4 (2)-(6). This means that we have two biologically feasible infected equilibria, which is essential to observe bistability, as shown in Figure 4.4 (1).
3. However, if the turning point on the infected equilibrium \bar{E}_1 , or the backward bifurcation moves down to the negative part of a state variable in the bifurcation diagram, that is, negative backward bifurcation occurs, then Hopf bifurcation is very unlikely to happen. Although Figure 4.4 (7) shows an exceptional case, the parameter range for such a Hopf bifurcation is very narrow.
4. The bifurcation diagram for system (4.4) with $A = 0.03$, shown in Figure 4.4 (8), is a typical model with negative backward bifurcation. Such negative backward bifurcation may occur in higher-dimensional systems. However, by considering more state variables, which make the system more complicated, Hopf bifurcation can happen in the upper branch of the negative backward bifurcation. We will discuss this possibility in more detail in the next section by examining an autoimmune disease model.

The results obtained in this section suggest the following summary.

Remark 2. *If a disease model contains a backward bifurcation on an equilibrium solution, then as the system parameters are varied, there may exist none, one or two Hopf bifurcations from the equilibrium solution, which may be supercritical or subcritical. If further this equilibrium has a transcritical bifurcation point at which it exchanges its stability with another equilibrium, then recurrence can occur between the transcritical and Hopf bifurcation points and near the transcritical point, where both equilibrium solutions are unstable, and bistability happens when Hopf bifurcation makes a shared stable parameter region for both equilibria.*

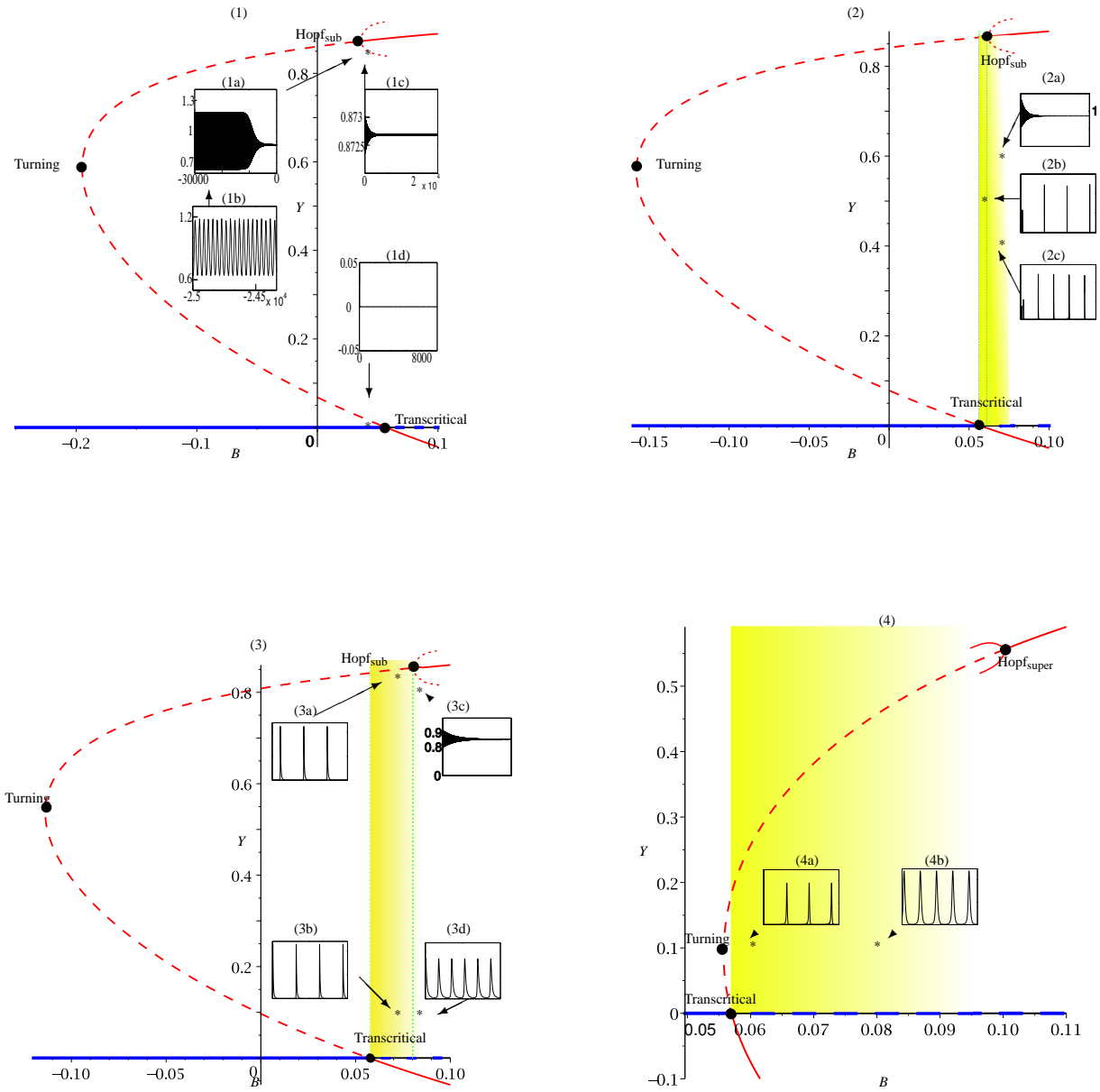


Figure 4.4: Dynamical behaviors of system (4.4) corresponding to eight cases listed in Table 4.2 and 4.4. All insets are simulated time histories of Y vs. τ . The yellow areas fading to white show regions in which recurrent behavior occurs and fades to regular oscillations.

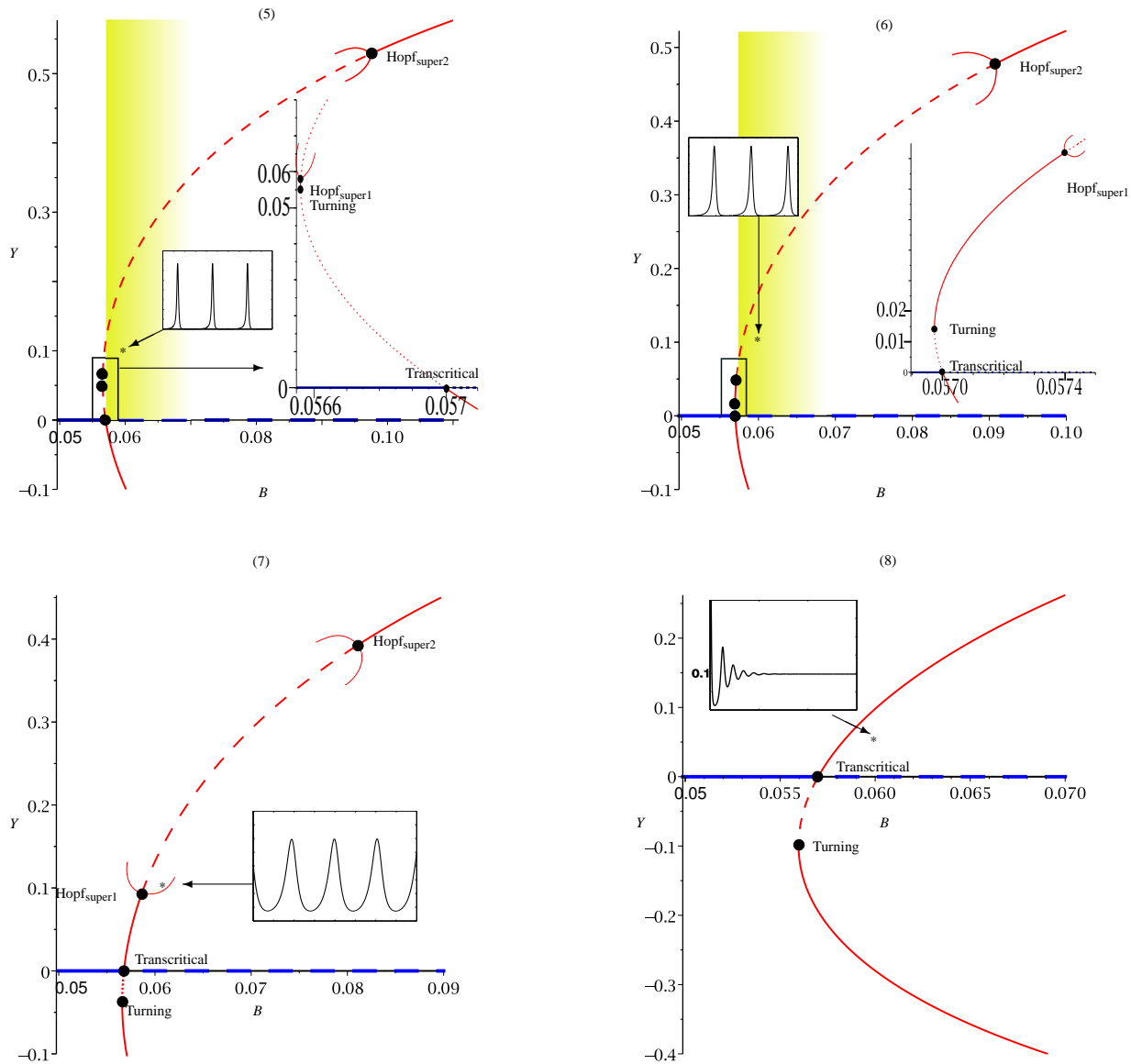


Figure 4.4: Dynamical behaviors of system (4.4) corresponding to eight cases listed in Table 4.2 and 4.4. All insets are simulated time histories of Y vs. τ . The yellow areas fading to white show regions in which recurrent behavior occurs and fades to regular oscillations.

4.4 Negative backward bifurcation in an autoimmune disease model

In the previous section, we examined three cases of negative backward bifurcation: Table 4.1 Case 4 for system (4.8) and Table 4.2 Case (7) and (8) for system 4.4. The analytical and numerical results showed that solutions typically converge to the infected equilibrium in these cases, and the parameter range for Hopf bifurcation is very limited. As a result, negative backward bifurcation tends to give no interesting behavior. In this section, however, we shall explore an established autoimmune model [1] in which negative backward bifurcation occurs. We demonstrate that after modification, the autoimmune model can also exhibit recurrence.

The autoimmune model [1] takes the form

$$\begin{aligned}\frac{dA}{dt} &= f\tilde{v}G - (\sigma_1 R_n + b_1)A - \mu_A A \\ \frac{dR_n}{dt} &= (\pi_1 E + \beta)A - \mu_n R_n \\ \frac{dE}{dt} &= \lambda_E A - \mu_E E \\ \frac{dG}{dt} &= \gamma E - \tilde{v}G - \mu_G G,\end{aligned}\tag{4.21}$$

where mature pAPCs (A) undergo maturation by intaking self-antigen (G), at rate $f\tilde{v}$, and are suppressed by specific regulatory T cells, T_{Reg} cells (R_n), at rate σ_1 ; b_1 represents additional non-specific background suppression. The T_{Reg} cells are activated by mature pAPCs at a rate proportional to the number of auto-reactive effector T cells (E) at rate π_1 , and by other sources at rate β . Active auto-reactive effector T cells (E) come from the activation process initiated by mature pAPCs, at rate λ_E , then attack healthy body tissue and release free self-antigen (G) at rate γ , which is ready for mature pAPCs to engulf; the antigen engulfing rate is \tilde{v} . The death rates of the populations A , R_n , E , and G are denoted by μ_A , μ_n , μ_E , and μ_G , respectively.

Following the steps described in detail in [50], system (4.21), can be reduced via quasi-steady state analysis to a 2-dimensional system:

$$\begin{aligned}\frac{dA}{dt} &= \left[\frac{f\tilde{v}\gamma\lambda_E}{\mu_E(\tilde{v}+\mu_G)} - b_1 - \mu_A \right] A - \sigma_1 R_n A, \\ \frac{dR_n}{dt} &= \left(\frac{\pi_1\lambda_E}{\mu_E} A + \beta \right) A - \mu_n R_n.\end{aligned}\tag{4.22}$$

For simplicity, we set $a = \frac{f\tilde{v}\gamma\lambda_E}{\mu_E(\tilde{v}+\mu_G)} - b_1 - \mu_A$ and $b = \frac{\pi_1\lambda_E}{\mu_E}$. For the stability and bifurcation analysis, we choose λ_E as the bifurcation parameter. System (4.22) has a disease-free equilibrium $\bar{E}_0 = (0, 0)$, which is stable if $a > 0$ or $\lambda_E > \frac{(b_1+\mu_A)(\tilde{v}+\mu_G)\mu_E}{f\tilde{v}\gamma}$; and unstable if $a < 0$ or $\lambda_E < \frac{(b_1+\mu_A)(\tilde{v}+\mu_G)\mu_E}{f\tilde{v}\gamma}$. Thus a static bifurcation occurs on \bar{E}_0 when $a = 0$ or $\lambda_E = \frac{(b_1+\mu_A)(\tilde{v}+\mu_G)\mu_E}{f\tilde{v}\gamma}$. The disease equilibrium is given by $\bar{E}_1 = (\bar{A}, \bar{R}_n)$, in which $\bar{R}_n = \frac{(b\bar{A}+\beta)\bar{A}}{\mu_n}$, and \bar{A} is given by the roots of the following equation,

$$f_8(A) = b\sigma_1 A^2 + \beta\sigma_1 A - \mu_n a.\tag{4.23}$$

Equation (4.23) has two roots with negative signs if $a < 0$, with opposite signs if $a > 0$, and only one zero root if $a = 0$. This means that a negative backward bifurcation is possible in system (4.22) with proper parameter values. We further examine the characteristic equation at \bar{E}_1 , which shares the same form as equation (4.14), with $\text{Tr}(J|_{\bar{E}_1}) = \frac{1}{\mu_n}(b\sigma_1 A^2 + \beta\sigma_1 A + \mu_n^2 - a\mu_n) := a_{11}$ and $\text{Det}(J|_{\bar{E}_1}) = 3b\sigma_1 A^2 + 2\beta\sigma_1 A - a\mu_n := a_{12}$. Solving $f_8(A) = 0$ and

$a_{12} = \text{Det}(J|_{\bar{E}_1}) = 0$, gives the static bifurcation point of \bar{E}_1 at $(\bar{A}, a) = (0, 0)$ or $(\bar{A}, \lambda_E) = (0, \frac{(b_1 + \mu_A)(\tilde{v} + \mu_G)\mu_E}{f\tilde{v}\gamma})$, which is a transcritical bifurcation point between \bar{E}_0 and \bar{E}_1 . Moreover, Hopf bifurcation can happen if and only if $f_8(A) = 0$ and $a_{11} = \text{Tr}(J|_{\bar{E}}) = 0$, which can be satisfied only if $\mu_n = 0$. This implies that the positive branch of \bar{E}_1 is stable for any positive values of μ_n . Thus, this model cannot exhibit recurrence, bistability, or even regular oscillation. The same conclusion was obtained in [50] for the original 4-dimensional model (4.21).

However, a recent experimental discovery [3] has revealed a new class of terminally differentiated T_{Reg} cells. As described in detail in [50], introducing this cell population, denoted R_d , into the model yields the full system

$$\begin{aligned} \frac{dA}{dt} &= f\tilde{v}G - \sigma_1(R_n + dR_d)A - (b_1 + \mu_A)A \\ \frac{dR_n}{dt} &= (\pi_1 E + \beta)A - \mu_n R_n - \xi R_n \\ \frac{dR_d}{dt} &= c\xi R_n - \mu_d R_d \\ \frac{dE}{dt} &= \lambda_E A - \mu_E E \\ \frac{dG}{dt} &= \gamma E - \tilde{v}G - \mu_G G \end{aligned} \quad (4.24)$$

and quasi-steady state analysis then yields a reduced 3-dimensional model in the form

$$\begin{aligned} \frac{dA}{dt} &= \left[\frac{f\tilde{v}\gamma\lambda_E}{(\tilde{v} + \mu_G)\mu_E} - (b_1 + \mu_A) \right] A - \sigma_1(R_n + dR_d)A, \\ \frac{dR_n}{dt} &= \left(\frac{\pi_1\lambda_E}{\mu_E} A + \beta \right) A - \mu_n R_n - \xi R_n, \\ \frac{dR_d}{dt} &= c\xi R_n - \mu_d R_d. \end{aligned} \quad (4.25)$$

Again, here λ_E is chosen as the bifurcation parameter for stability and bifurcation analysis. It is easy to show that system (4.25) still has a disease-free equilibrium \bar{E}_0 as $(A, R_n, R_d) = (0, 0, 0)$, and a disease equilibrium \bar{E}_1 as $(\bar{A}, \bar{R}_n, \bar{R}_d)$, where $\bar{R}_d = \frac{c\xi\bar{R}_n}{\mu_d}$, $\bar{R}_n = \frac{\beta\mu_E + \pi_1\lambda_E\bar{A}}{\mu_E(\mu_n + \xi)}\bar{A}$, and \bar{A} is determined from the following quadratic equation:

$$f_9(A) = \pi_1\lambda_E A^2 + \beta\mu_E A + \frac{\mu_d(\mu_n + \xi)}{(\tilde{v} + \mu_G)(c\xi + \mu_d)\sigma_1} [-f\gamma\tilde{v}\lambda_E + (b_1 + \mu_A)(\mu_G + \tilde{v})\mu_E], \quad (4.26)$$

which gives two negative roots if $\lambda_E < \lambda_{ES} = \frac{(b_1 + \mu_A)(\mu_G + \tilde{v})\mu_E}{f\gamma\tilde{v}}$, and two roots with opposite signs when $\lambda_E > \lambda_{ES}$. The critical point is determined by $\lambda_E = \lambda_{ES}$, which is actually the intersection point of \bar{E}_0 and \bar{E}_1 . The two equilibrium solutions exchange their stability at λ_{ES} , leading to a transcritical bifurcation at $(\bar{A}, \lambda_E) = (0, \lambda_{ES})$. Note that the negative backward bifurcation still happens in system 4.25. Moreover, a Hopf bifurcation occurs from the upper branch of \bar{E}_1 , giving rise to oscillation and recurrence.

Realistic parameter values have been obtained in [50], and are given as follows:

$$\begin{aligned} f &= 1 \times 10^{-4}, \quad \tilde{v} = 0.25 \times 10^{-2}, \quad \sigma_1 = 3 \times 10^{-6}, \quad b_1 = 0.25, \quad \mu_A = 0.2, \quad \pi_1 = 0.016, \\ \beta &= 200, \quad \mu_n = 0.1, \quad \mu_E = 0.2, \quad \gamma = 2000, \quad \mu_G = 5, \quad \mu_d = 0.2, \quad c = 8, \quad d = 2, \quad \xi = 0.025. \end{aligned}$$

For the above parameter values, the Hopf critical point is obtained at $(A_H, \lambda_{EH}) = (5.6739, 1691.6414)$, while the turning point is at $(A_T, \lambda_{ET}) = (-1.4205, 879.9848)$, and the transcritical bifurcation point is at $(A_S, \lambda_{ES}) = (0, 900.45)$. These three bifurcation points and the stability of equilibrium solutions are shown in the bifurcation diagram given in Figure 4.5(a), and the simulated recurrent time history is plotted in Figure 4.5(b) for $\lambda_E = \lambda_{EH} + 1000$.

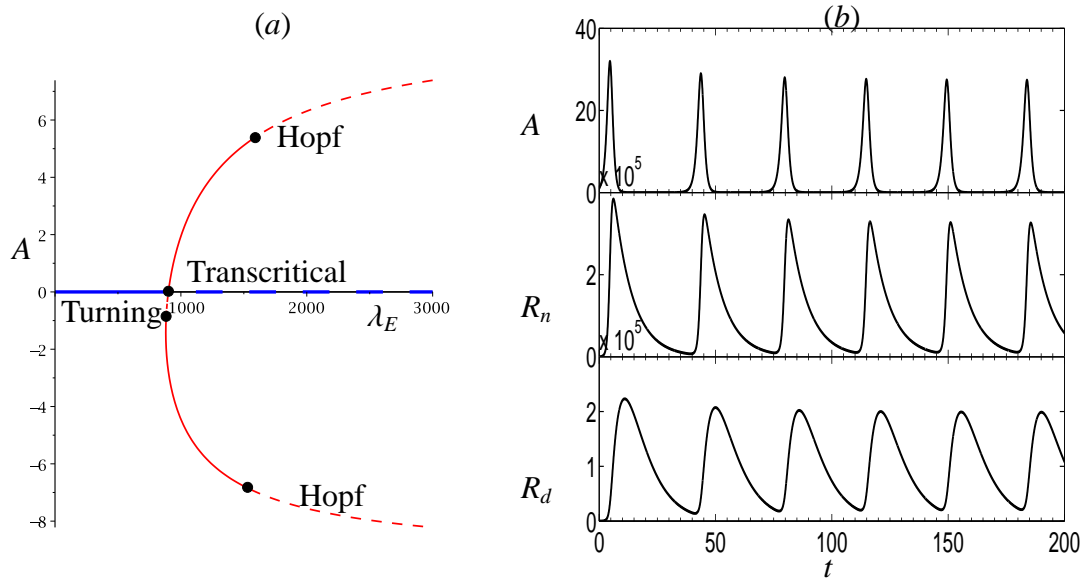


Figure 4.5: Dynamics of system (4.22): (a) bifurcation diagram; and (b) simulated time history for $\lambda_E = \lambda_{EH} + 1000$.

In summary, when negative backward bifurcation occurs, that is, the turning point is located in the negative state variable space, less complex dynamical behavior will be present. Hopf bifurcation in a biologically feasible area does not happen in the reduced 2-dimensional system (4.22), nor in the original system (4.21) [50]. However, if we increase the dimension of the system, Hopf bifurcation and complex dynamical phenomena can emerge, as shown in our results for system (4.25).

4.5 Conclusion

In this paper, we first review the previous work on a reduced 2-dimensional infection model with a concave incidence rate [28]. The authors proved that the disease equilibrium will emerge and be globally stable when the basic reproduction number R_0 is greater than 1. This means that no complex dynamical phenomenon can occur in such models. However, by adding an extra saturating treatment term to this simple 2-dimensional infection model, the resulting system (4.7) considered in [51] can exhibit backward bifurcation, which increases the parameter range for Hopf bifurcation, which in turn leads to recurrent, bistable and regular oscillating behaviors.

Instead of adding an extra term, a 2-dimensional infection model with a convex incidence function can likewise show rich dynamics due to the occurrence of backward bifurcation, giving rise to two types of Hopf bifurcation. Biologically, a convex incidence rate implies that existing infection makes the host more vulnerable to further infection, showing a cooperative effect in disease progression. From the view point of mathematics, the convex incidence function enables backward bifurcation to occur on the positive branch of the disease equilibrium solution, which further generates Hopf bifurcation. The location and direction of Hopf bifurcation(s), determined by parameter values, can further give rise to bistable, recurrent, and regular

oscillating behaviors.

Cooperative effects also occur during the progression of autoimmune disease. However, for an autoimmune model with negative backward bifurcation, in which the turning point is located on the negative state variable space, the biologically feasible parameter range in which Hopf bifurcation may occur is limited. By introducing an additional state variable to the autoimmune model, recurrent phenomenon are once again observed.

4.6 References

- [1] H. K. Alexander and L. M. Wahl. Self-tolerance and autoimmunity in a regulatory T cell model. *Bulletin of mathematical biology*, 73(1):33–71, 2011.
- [2] J. Arino, C. C. McCluskey, and P. van den Driessche. Global results for an epidemic model with vaccination that exhibits backward bifurcation. *SIAM Journal on Applied Mathematics*, 64(1):260–276, 2003.
- [3] C. Baecher-Allan, E. Wolf, and D. A. Hafler. MHC class II expression identifies functionally distinct human regulatory T cells. *The Journal of Immunology*, 176:4622–4631, 2006.
- [4] K. W. Blayneh, A. B. Gumel, S. Lenhart, and T. Clayton. Backward bifurcation and optimal control in transmission dynamics of west nile virus. *Bulletin of Mathematical Biology*, 72(4):1006–1028, 2010.
- [5] F. Brauer. Backward bifurcations in simple vaccination models. *Journal of Mathematical Analysis and Applications*, 298(2):418 – 431, 2004.
- [6] C. J. Briggs and H. C. J. Godfray. The dynamics of insect-pathogen interactions in stage-structured populations. *American Naturalist*, pages 855–887, 1995.
- [7] G. C. Brown and R. Hasibuan. Conidial discharge and transmission efficiency of *Neozygites floridana*, an entomopathogenic fungus infecting two-spotted spider mites under laboratory conditions. *Journal of Invertebrate Pathology*, 65(1):10 – 16, 1995.
- [8] V. Capasso and G. Serio. A generalization of the Kermack-McKendrick deterministic epidemic model. *Mathematical Biosciences*, 42(1):43–61, 1978.
- [9] C. Castillo-Chavez, K. Cooke, W. Huang, and S. Levin. On the role of long incubation periods in the dynamics of acquired immunodeficiency syndrome (AIDS). Part 2: Multiple group models. In Carlos Castillo-Chavez, editor, *Mathematical and Statistical Approaches to AIDS Epidemiology*, volume 83 of *Lecture Notes in Biomathematics*, pages 200–217. Springer Berlin Heidelberg, 1989.
- [10] C. Castillo-Chavez, K. Cooke, W. Huang, and S. A. Levin. Results on the dynamics for models for the sexual transmission of the human immunodeficiency virus. *Applied Mathematics Letters*, 2(4):327 – 331, 1989.

- [11] A. C. Collier, R. W. Coombs, D. A. Schoenfeld, R. L. Bassett, J. Timpone, A. Baruch, M. Jones, K. Facey, C. Whitacre, V. J. McAuliffe, H. M. Friedman, T. C. Merigan, R. C. Reichman, C. Hooper, and L. Corey. Treatment of human immunodeficiency virus infection with Saquinavir, Zidovudine, and Zalcitabine. *New England Journal of Medicine*, 334(16):1011–1018, 1996.
- [12] J. M. Conway and D. Coombs. A stochastic model of latently infected cell reactivation and viral blip generation in treated HIV patients. *PLoS Computational Biology*, 7(4), 2011.
- [13] W. R. Derrick and P. Driessche. A disease transmission model in a nonconstant population. *Journal of Mathematical Biology*, 31(5):495–512, 1993.
- [14] G. Dornadula, H. Zhang, B. VanUitert, J. Stern, L. Livornese Jr., M. J. Ingerman, J. Witek, R. J. Kedanis, J. Natkin, J. DeSimone, and R. J. Pomerantz. Residual HIV-1 RNA in blood plasma of patients taking suppressive highly active antiretroviral therapy. *Journal of the American Medical Association*, 282(17):1627–1632, 1999.
- [15] J. Dushoff, W. Huang, and C. Castillo-Chavez. Backwards bifurcations and catastrophe in simple models of fatal diseases. *Journal of mathematical biology*, 36(3):227–248, 1998.
- [16] C. Fraser, N. M. Ferguson, F. de Wolf, and R. M. Anderson. The role of antigenic stimulation and cytotoxic T cell activity in regulating the long-term immunopathogenesis of HIV: mechanisms and clinical implications. *Proceedings of the Royal Society B: Biological Sciences*, 268(1481):2085–2095, 2001.
- [17] I. C. H. Fung, M. Gambhir, A. van Sighem, F. de Wolf, and G. P. Garnett. The clinical interpretation of viral blips in HIV patients receiving antiviral treatment: Are we ready to infer poor adherence? *Journal of Acquired Immune Deficiency Syndromes*, 60(1):5–11, 2012.
- [18] N. J. Garretta, V. Apeaa, A. Noria, I. Ushiro-Lumbb, A. R. Oliverb, G. Bailya, and D. A. Clarkb. Comparison of the rate and size of HIV-1 viral load blips with Roche COBAS TaqMan HIV-1 versions 1.0 and 2.0 and implications for patient management. *Journal of Clinical Virology*, 53(4):354–355, 2012.
- [19] L. Gil and et al. Contribution to characterization of oxidative stress in HIV/AIDS patients. *Pharmacological research*, 47(3):217–224, 2003.
- [20] J. T. Grennan, M. R. Loutfy, D. Su, P. R. Harrigan, C. Cooper, M. Klein, N. Machouf, J. S. G. Montaner, S. Rourke, C. Tsoukas, B. Hogg, J. Raboud, and the CANOC Collaboration. Magnitude of virologic blips is associated with a higher risk for virologic rebound in HIV-infected individuals: A recurrent events analysis. *Journal of Infectious Diseases*, 205(8):1230–1238, 2012.
- [21] H. Gmez-Acevedo and M. Y. Li. Backward bifurcation in a model for HTLV-I infection of CD4⁺ T cells. *Bulletin of Mathematical Biology*, 67(1):101 – 114, 2005.

- [22] K. P. Hadeler and P. Van den Driessche. Backward bifurcation in epidemic control. *Mathematical Biosciences*, 146(1):15–35, 1997.
- [23] H. W. Hethcote, M. A. Lewis, and P. Van Den Driessche. An epidemiological model with a delay and a nonlinear incidence rate. *Journal of mathematical biology*, 27(1):49–64, 1989.
- [24] H. W. Hethcote and P. Van den Driessche. Some epidemiological models with nonlinear incidence. *Journal of Mathematical Biology*, 29(3):271–287, 1991.
- [25] W. Huang, K. L. Cooke, and C. Castillo-Chavez. Stability and bifurcation for a multiple-group model for the dynamics of HIV/AIDS transmission. *SIAM Journal on Applied Mathematics*, 52(3):835–854, 1992.
- [26] N. Israel and M. A. Gougerot-Pocidalò. Oxidative stress in human immunodeficiency virus infection. *Cell Molec Life Sciences*, 53(11-12):864–870, 1997.
- [27] L. E. Jones and A. S. Perelson. Opportunistic infection as a cause of transient viremia in chronically infected HIV patients under treatment with HAART. *Bulletin of Mathematical Biology*, 67(6):1227–1251, 2005.
- [28] A. Korobeinikov and P. K. Maini. Non-linear incidence and stability of infectious disease models. *Mathematical Medicine and Biology*, 22(2):113–128, 2005.
- [29] M. Y. Li and J. S. Muldowney. Global stability for the SEIR model in epidemiology. *Mathematical Biosciences*, 125(2):155–164, 1995.
- [30] N. Li and M. Karin. Is NF- κ B the sensor of oxidative stress? *The FASEB Journal*, 13(10):1137–1143, 1999.
- [31] W. Liu, H. W. Hethcote, and S. A. Levin. Dynamical behavior of epidemiological models with nonlinear incidence rates. *Journal of Mathematical Biology*, 25(4):359–380, 1987.
- [32] W. Liu, S. Levin, and Y. Iwasa. Influence of nonlinear incidence rates upon the behavior of SIRS epidemiological models. *Journal of Mathematical Biology*, 23(2):187–204, 1986.
- [33] J. D. Murray. *Mathematical Biology: I. An Introduction*. Interdisciplinary Applied Mathematics. Springer, 2002.
- [34] S. Palmer, F. Maldarelli, A. Wiegand, B. Bernstein, G. J. Hanna, S. C. Brun, D. J. Kempf, J. W. Mellors, J. M. Coffin, and M. S. King. Low-level viremia persists for at least 7 years in patients on suppressive antiretroviral therapy. *Proceedings of the National Academy of Sciences*, 105(10):3879–3884, 2008.
- [35] S. Palmer, A. P. Wiegand, F. Maldarelli, H. Bazmi, J. M. Mican, M. Polis, R. L. Dewar, A. Planta, S. Liu, J. A. Metcalf, J. W. Mellors, and J. M. Coffin. New real-time reverse transcriptase-initiated PCR assay with single-copy sensitivity for human immunodeficiency virus type 1 RNA in plasma. *Journal of Clinical Microbiology*, 41(10):4531–4536, 2003.

- [36] L. Rong and A. S. Perelson. Asymmetric division of activated latently infected cells may explain the decay kinetics of the HIV-1 latent reservoir and intermittent viral blips. *Mathematical Biosciences*, 217(1):77–87, 2009.
- [37] L. Rong and A. S. Perelson. Modeling latently infected cell activation: viral and latent reservoir persistence, and viral blips in HIV-infected patients on potent therapy. *PLoS Computational Biology*, 5(10):e1000533, 2009.
- [38] S. Ruan and W. Wang. Dynamical behavior of an epidemic model with a nonlinear incidence rate. *Journal of Differential Equations*, 188(1):135 – 163, 2003.
- [39] K. B. Schwarz. Oxidative stress during viral infection: a review. *Free Radical Biology and Medicine*, 21(5):641–649, 1996.
- [40] C. P. Simon and J. A. Jacquez. Reproduction numbers and the stability of equilibria of SI models for heterogeneous populations. *SIAM journal on Applied Mathematics*, 52(2):541–576, 1992.
- [41] C. B. Stephenson, G. S. Marquis, S. D. Douglas, and C. M. Wilson. Immune activation and oxidative damage in HIV-positive and HIV negative adolescents. *Journal of Acquired Immune Deficiency Syndromes*, 38(2):180–190, 2005.
- [42] P. van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180(12):29 – 48, 2002.
- [43] R. D. van Gaalen and L. M. Wahl. Reconciling conflicting clinical studies of antioxidant supplementation as HIV therapy: a mathematical approach. *BMC Public Health*, 9(Suppl. 1):1–18, 2009.
- [44] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Berlin etc., Springer-Verlag, 1990.
- [45] W. Yao, L. Hertel, and L. M. Wahl. Dynamics of recurrent viral infection. *Proceedings of the Royal Society-Biological Sciences*, 273(1598):2193–2199, 2006.
- [46] P. Yu. Computation of normal forms via a perturbation technique. *Journal of Sound and Vibration*, 211(1):19–38, 1998.
- [47] P. Yu, W. Zhang, and L. Wahl. Dynamical analysis of a 2-dimensional disease model with convex incidence. (*Submitted for publication*), xx(xx):xx–xx, xxxx.
- [48] W. Zhang, L. Wahl, and P. Yu. Conditions for transient viremia in deterministic in-host models: Viral blips need no exogenous trigger. *SIAM Journal on Applied Mathematics*, 73(2):853–881, 2013.
- [49] W. Zhang, L. Wahl, and P. Yu. Viral blips may not need a trigger: How transient viremia can arise in deterministic in-host models. *SIAM Review*, 56(1):127–155, 2014.

- [50] W. Zhang, L. Wahl, and P. Yu. Modelling and analysis of recurrent autoimmune disease. *SIAM Journal on Applied Mathematics (Submitted for publication)*, xx(xx):xx–xx, xxxx.
- [51] L. Zhou and M. Fan. Dynamics of an SIR epidemic model with limited medical resources revisited. *Nonlinear Analysis: Real World Applications*, 13(1):312 – 324, 2012.

Chapter 5

Dynamical Analysis of a 2-dimensional Disease Model with Convex Incidence

5.1 Introduction

Mathematical models in epidemiology and in-host disease share common features, dividing a population of individuals (epidemiology) or cells (in-host) into discrete classes relevant to the disease dynamics, and typically describing their dynamics with a system of ordinary differential equations (ODEs). A key feature of such systems is the incidence function, which defines the spread of the infection to susceptibles.

For example, in classical epidemiological models, the incidence rate is often assumed to take the form $\frac{\beta SI}{N}$, where $S(t)$ is the number of susceptible individuals, $I(t)$ is the number of infectives and β is a constant, the transmission rate [3]. When N , the population size, is constant, this incidence function is also simply written as βSI . Similarly, for in-host models, the rate at which uninfected cells become infected is often described as βxy , where $x(t)$ reflects the uninfected cell density and $y(t)$ denotes the density of infected cells [23].

Bilinear incidence functions of this form have been used extensively and are well-studied in the mathematical literature. As described in greater detail elsewhere [17], a number of possibilities for non-linear incidence functions have also been studied in some detail, including the general form $\beta I^p S^q$, where p and q are positive constants [21, 20, 14, 15, 8, 19], and several more complex forms [21, 6].

Because of physical limitations on the number of new infections possible as disease prevalence increases, a common feature of many incidence functions is their concavity with respect to the number of infectives. In particular, the incidence rate $f(S, I, N)$ typically satisfies the condition

$$\frac{\partial^2 f(S, I, N)}{\partial I^2} \leq 0.$$

Taking advantage of this common feature, Korobeinikov and Maini [17] derived elegant results for all concave incidence functions, showing the global asymptotic stability of the disease-free equilibrium when the basic reproduction number $R_0 \leq 1$, and global asymptotic stability of the endemic equilibrium when $R_0 > 1$, for the standard SIRS model [3] with a constant population size. In other words, the concavity of the incidence rate guarantees the uniqueness and stability

of the endemic equilibrium in these models, and these powerful results apply to any concave incidence function.

In contrast, we have recently analyzed a number of ODE models with convex incidence functions. If incidence is convex, or “synergistic”, the rate at which new infections occur can increase supralinearly with disease prevalence. This situation can arise in a number of realistic scenarios. For example, in in-host models of the human immunodeficiency virus (HIV), increasing the extent of the infection involves greater damage to the immune system, and can thus increase the incidence rate [25]. Similarly, in autoimmune disease, increases to the autoimmune response against self tissue can cause a positive feedback loop which will further increase the incidence rate [1]. While these two examples both arise in in-host disease modelling, catastrophic outbreak or pandemic conditions could also result in convex epidemiological incidence. In particular, an outbreak that is severe enough to compromise health care infrastructure (increasing hospital crowding and front-line worker exposure rates, for example) could involve a supralinear increase in incidence rates with disease prevalence.

In this contribution, we analyze in detail the possible dynamical behaviors of a simple 2-dimensional disease model with a convex, or synergistic, incidence function. The system we analyze is a standard non-dimensionalized SI model which arises in both epidemiology and in-host modelling: it assumes a birth rate into the susceptible population, death rates for both populations, and an incidence rate between the two. The incidence function we study has an analytical form which has arisen in a number of models previously analyzed [25, 1, 27, 28, 29]. Its behavior is such that when the infective population I is small, incidence increases linearly with I ; when I is large, incidence also increases linearly, but with a steeper slope. A convex region of the function connects these limiting behaviors.

In marked contrast to the powerful general conclusions obtained for concave incidence functions [17], we find that a wide range of dynamical behaviors are possible when incidence is synergistic. In particular, as previously analyzed in related higher-dimensional models [27, 28, 29], we note the appearance of *recurrent infection*, that is, cycles consisting of long periods close to the disease free equilibrium, punctuated by brief bursts of disease. This pattern of recurrence occurs in many diseases, including the intriguing pattern of “viral blips” in HIV, as well as the recurrent episodes characteristic of autoimmune diseases, such as multiple sclerosis [7], multifocal osteomyelitis [12, 16], lupus [22], eczema [11], and psoriasis [10]. In this contribution, we explore several mechanisms which can underly these physiologically relevant patterns of infection, finding that when the incidence function is convex, bistable equilibrium solutions, Hopf and generalized Hopf bifurcations and, in particular, homoclinic bifurcations may all contribute to disease recurrence.

In related work, Ruan and Wang [24] analyzed a reduced SI model, which has a zero disease-free equilibrium and a positive endemic equilibrium. In this model, $R_0 = 0$, although it can be shown that the disease can still persist. In [24], the authors also considered Hopf bifurcation, Bogdanov-Takens bifurcation and homoclinic orbits. The structure of the model in [24] is mathematically appealing, such that the authors could transform the model to a Liénard system and then prove the uniqueness of the limit cycle from Hopf bifurcation. Moreover, their analysis of the homoclinic orbit takes the standard form (e.g., see [13]). In contrast, the model we study in this contribution has been derived from physical considerations and has known realistic parameter ranges, however this model cannot be transformed to a Liénard system, and the analysis of homoclinic orbits does not follow the standard form.

The rest of the paper is organized as follows. In next section, we give a detailed dynamical analysis of the simple 2-dimensional disease model. In Section 3, Hopf and generalized Hopf bifurcations are studied in detail, which may be the main features underlying complex dynamical behaviors. Then, in Section 4, Bogdanov-Takens bifurcation and homclinic bifurcation are investigated, giving rise to another scenario/mechanism for generating blips. Finally, conclusions and discussion are given in Section 5.

5.2 Dynamics of the 2-D disease model

Consider the 2-dimensional system:

$$\begin{aligned}\frac{dX}{d\tau} &= 1 - DX - \left(B + \frac{AY}{Y+C}\right)XY, \\ \frac{dY}{d\tau} &= \left(B + \frac{AY}{Y+C}\right)XY - Y,\end{aligned}\tag{5.1}$$

where all parameters, A , B , C and D take positive real values. This system was originally derived as an in-host model of HIV dynamics [25], but has been reduced in dimension and non-dimensionalized using quasi-steady state assumptions as described in [27, 28]. Although arising from in-host disease modeling, the reduced 2-D system is also equivalent to the SIRS model studied in [17], taking the recovery rate, α of [17], to be zero. At appropriate parameter values, system (1) thus represents either an in-host infection (susceptible and infected cells), or an SIR epidemiological model (susceptible and infected individuals). The key difference between system (1) and the class of models studied in [17] is that the incidence function in system (1), $XY(B + AY/(Y + C))$, is convex. Our goal is to understand the dynamical behaviors made possible by this convexity.

In [27, 28], this 2-dimensional model is not analyzed in detail. For example, well-posedness of solutions of this system and the global stability of the disease-free equilibrium were not considered; and a trapping region was proved only for fixed parameter values when $B > D$. In the following subsections, we will provide general proofs for the above mentioned problems with no additional restriction on the positive parameter values.

5.2.1 Well-posedness of solutions

We first prove the positiveness and boundedness of solutions of system (5.1). We have the following result.

Theorem 5.2.1 *Solutions of system (5.1) are non-negative provided the initial conditions are non-negative, and further these solutions are eventually attracted to a bounded region.*

Proof Using the first equation of system (5.1), with the formulae of variation of parameters, we obtain

$$X(\tau) = X(0)e^{-\int_0^\tau [D + (B + \frac{AY(s)}{Y(s)+C})Y(s)]ds} + \int_0^\tau e^{-\int_s^\tau [D + (B + \frac{AY(u)}{Y(u)+C})Y(u)]du} ds, \tag{5.2}$$

which clearly indicates that $X(\tau) > 0$ for $\tau > 0$ if $X(0) \geq 0$. Next, we rewrite the second equation of (5.1) as

$$\frac{dY}{d\tau} = (BX - 1)Y + \frac{A XY^2}{Y + C}.$$

We have shown $X = X(\tau) > 0$ for $\tau > 0$. Suppose $Y(0) \geq 0$. Then, by continuity of solutions, there exists $\tau_1 > 0$ such that

$$\frac{A X(\tau) Y^2(\tau)}{Y(\tau) + C} \geq 0 \quad \text{for } \tau \in [0, \tau_1].$$

Hence, the solution of Y must be non-negative for $\tau \in (0, \tau_1]$, and so $Y(\tau_1) \geq 0$. Now, starting from $\tau = \tau_1$, we apply the above argument to ensure that there exists $\tau_2 > \tau_1$ such that

$$\frac{A X(\tau) Y^2(\tau)}{Y(\tau) + C} \geq 0 \quad \text{for } \tau \in [\tau_1, \tau_2].$$

Repeating the process, we have shown that $Y(\tau) \geq 0$ for $\tau > 0$ as long as $X(\tau) > 0$ ($\tau > 0$) and $Y(0) \geq 0$.

To prove the boundedness of the solutions, we choose (Lyapunov) function

$$L(X, Y) = X + Y, \tag{5.3}$$

which is positive definite for positive solutions. Differentiating L with respect to time τ , along the trajectory of system (5.1) yields

$$\left. \frac{dL}{d\tau} \right|_{(5.1)} = \frac{dX}{d\tau} + \frac{dY}{d\tau} = 1 - DX - Y \begin{cases} < 0 & \text{if } DX + Y > 1, \\ > 0 & \text{if } DX + Y < 1, \end{cases}$$

which indicates that X and Y are bounded.

5.2.2 Construction of generic trapping region

More precisely, we can construct a trapping region for all possible positive parameter values, to which all solutions are attracted. Before stating the theorem, we first note that system (5.1) has two equilibrium solutions obtained by setting $\frac{dX}{d\tau} = \frac{dY}{d\tau} = 0$: one is the disease-free equilibrium, $E_0 = (\frac{1}{D}, 0)$, which is a boundary equilibrium, and other is the endemic equilibrium, $E_1 = (X_1, Y_1)$, which is an interior equilibrium, where

$$Y_1 = 1 - DX_1, \tag{5.4}$$

and X_1 is determined from the quadratic polynomial equation:

$$\begin{aligned} Q(X) &= D(A + B)X^2 - (A + B + D + BC)X + C + 1 \\ &= \frac{1}{D}[(A + B)(1 - DX)^2 \\ &\quad - (A + B - D - BC)(1 - DX) - C(B - D)] \\ &= 0. \end{aligned} \tag{5.5}$$

The existence of E_1 depends on the values of the parameters.

Theorem 5.2.2 *There exists a trapping region G , in the shape of a right triangle, bounded by the X -axis, the Y -axis and the line $X + Y = \max\{1, \frac{1}{D}\} + \varepsilon$ ($0 < \varepsilon \ll 1$).*

Proof First, consider the X -axis. Note that $E_0 = (\frac{1}{D}, 0)$ is located on the X -axis, with two eigenvalues, $\xi_1 = -D$ and $\xi_2 = \frac{B}{D} - 1$, and their corresponding eigenvectors are $v_1 = (1, 0)$ and $v_2 = (1, \frac{D}{B}(1 - D) - 1)$, respectively. Moreover, v_1 is in the direction of the X -axis, which can be shown to be a solution trajectory of the system. With a negative eigenvalue, the trajectory along the X -axis converges to the point E_0 . Thus, the X -axis is a separator (invariant manifold) of the dynamical system, and so no trajectory can cross it due to the uniqueness of solutions. Hence, every trajectory entering the region G cannot escape from this boundary – the X -axis.

On the Y -axis, it is easy to obtain $\frac{dX}{d\tau} = 1$ and $\frac{dY}{d\tau} = -Y$, showing that all trajectories cross the Y -axis from left to right.

Next, we want to prove that all trajectories which cross the line L actually move into the region G . To achieve this, note that the direction of the line L is $(1, -1)$, and so the normal direction of the line in its gradient direction is $(1, 1)$. Define

$$S(Y) = (1, 1) \bullet \left(\frac{dX}{d\tau}, \frac{dY}{d\tau} \right) = \frac{dX}{d\tau} + \frac{dY}{d\tau},$$

where the dot denotes inner product (or dot product). We need to show $S(Y) < 0$ for $0 < Y < \max\{1, \frac{1}{D}\} + \varepsilon$. Simplifying $S(Y)$ yields

$$\begin{aligned} S(Y) &= \frac{dX}{d\tau} + \frac{dY}{d\tau} = \left[1 - DX - \left(B + \frac{AY}{Y+C} \right) XY \right] + \left(B + \frac{AY}{Y+C} \right) XY - Y \\ &= 1 - DX - Y = 1 - DX - \left[\max\left\{1, \frac{1}{D}\right\} + \varepsilon - X \right] \\ &= -\varepsilon + 1 - \max\left\{1, \frac{1}{D}\right\} + (1 - D)X \\ &= \begin{cases} -\varepsilon + (1 - \frac{1}{D}) + (1 - D)X, & D < 1 \\ -\varepsilon, & D = 1 \\ -\varepsilon - (D - 1)X, & D > 1 \end{cases} \text{ for } 0 < X < \max\left\{1, \frac{1}{D}\right\} + \varepsilon \\ &\leq -\varepsilon \min\{1, D\} < 0. \end{aligned}$$

Note that one may set $\varepsilon = 0$ for $D \neq 1$. Hence, for all positive parameter values, there always exists a trapping region G , bounded by the X -axis, the Y -axis, and the line L , and all trajectories move into G when crossing the Y -axis and the line L , and once they enter G , they cannot escape from the X -axis.

In the following, we consider the dynamical behaviour of system (5.1) according to the conditions: $B < D$, $B > D$ and $B = D$. Note that system (5.1) is actually equivalent to the model studied in [17] when Y is small so that $Y^2 \approx 0$. In this case, system (5.1) has bilinear incidence, which is concave, and the local $R_0 = \frac{B}{D}$. Thus, we expect that the disease-free equilibrium, E_0 , when Y is in fact small, is locally stable when $B < D$, and becomes a saddle point when $B > D$.

5.2.3 Dynamical behavior of (5.1) when $B < D$

First, we study the dynamical behavior of system (5.1) when $B < D$. In particular, we want to investigate the global stability of the disease-free equilibrium E_0 . For convenience, define

$$H_1 \triangleq A + B - D - BC - 2\sqrt{C(A+B)(D-B)}, \quad (B < D). \quad (5.6)$$

We have the following result.

Theorem 5.2.3 *When $B < D$, the disease-free equilibrium E_0 of system (5.1) is globally asymptotically stable if $H_1 < 0$, under which the endemic equilibrium E_1 does not exist. Otherwise, there exist two disease equilibria – one of them is a saddle point while the other may be a stable (or an unstable) node or focus – and no definite conclusion can be made regarding the global stability of E_0 .*

Proof First, it is easy to see that when $B < D$, the disease-free equilibrium E_0 is a stable node since both eigenvalues are negative. In order to prove this theorem, we also need the information about the disease equilibrium E_1 . Solving equation (5.5) yields two roots:

$$X_{\pm} = \frac{(A + B + D + BC) \pm \sqrt{\Delta}}{2D(A + B)}, \quad (5.7)$$

where

$$\begin{aligned} \Delta &= (A + B + D + BC)^2 - 4(C + 1)D(A + B) \\ &= (A + B - D - BC)^2 - 4C(A + B)(D - B), \end{aligned} \quad (5.8)$$

which implies that the existence condition for X_{\pm} when $B < D$ is given by

$$\begin{aligned} \Delta &= (A + B - D - BC)^2 - 4C(A + B)(D - B) \\ &= [A + B - D - BC + 2\sqrt{C(A + B)(D - B)}] H_1 > 0. \end{aligned} \quad (5.9)$$

Now, based on H_1 , we discuss the existence condition of biologically meaningful solutions X_{\pm} .

- (i) When $H_1 \geq 0$, it yields $\Delta \geq 0$, for which $0 < X_- \leq X_+ < \frac{1}{D}$, implying that the disease equilibrium E_1 has two solutions E_{1+} : (X_+, Y_+) and E_{1-} : (X_-, Y_-) . In particular, when $H_1 = 0$, $0 < X_- = X_+ < \frac{1}{D}$, indicating a saddle-node bifurcation to occur from the equilibrium E_1 .
- (ii) When $H_1 < 0$, there are two cases.
 - (iia) If $-2\sqrt{C(A+B)(D-B)} < A+B-D-BC < 2\sqrt{C(A+B)(D-B)}$, then $\Delta < 0$, and so there is no real solution for X_{\pm} . Thus, equilibrium E_1 does not exist.
 - (iib) If $A + B - D - BC \leq -2\sqrt{C(A+B)(D-B)}$ under which $\Delta \geq 0$, we then have $X_+ \geq X_- \geq \frac{1}{D}$, showing that there do not exist biologically meaningful equilibria E_1 .

The above discussions show that a biologically meaningful equilibrium E_1 does not exist if $H_1 < 0$ (with $B < D$), and in this case, there exists only one stable equilibrium E_0 on the boundary of the trapping region G . By index theory, this means that all trajectories of system (5.1) must converge to the stable node E_0 , and so the disease-free equilibrium E_0 is globally asymptotically stable if $H_1 < 0$ when $B < D$.

Remark The condition $B \geq D$ guarantees the existence of unique disease equilibrium E_1 , for which the disease-free equilibrium E_0 is a saddle point. (When $B = D$, E_0 is a degenerate saddle point, which will be proved later in Section 5.2.5.) When $B < D$, the disease equilibrium E_1 may or may not exist. The additional condition $H_1 \geq 0$ (with $B < D$) guarantees the existence of two disease equilibria $E_{1\pm}$ ($E_{1-} = E_{1+}$ when $B = D$). It can be easily seen from (5.6) that when $B < D$, $H_1 \geq 0$ implies $A + B - D - BC > 0$, i.e., $A > (D - B) + BC$, indicating that A must pass through a threshold value to generate the disease equilibrium solution E_1 . This is clear from the second equation of (5.1), which can be rewritten as $\frac{dY}{d\tau} = [(BX - 1) + \frac{AY}{Y+C}]Y$, that the first term $BX - 1 < 0$ for $X < \frac{1}{D}$ and $B < D$. Thus, $\frac{dY}{d\tau} < 0$ with small values of A for all values of X , implying that Y will die out. When the value of A exceeds its threshold, $\frac{dY}{d\tau}$ becomes positive at least for some values of X , which makes Y gain a steady state and thus the disease equilibrium E_1 exists. Biologically, the threshold value of the contact rate, A , means that the interaction between X and Y produces sufficient infection such that Y persists.

In the remainder of the proof, we assume that $B < D$ and $H_1 \geq 0$. If $H_1 > 0$, then $0 < X_- < X_+ < \frac{1}{D}$, which implies that two biologically meaningful equilibrium solutions exist for E_1 . When $H_1 = 0$, we have $0 < X_- = X_+ < \frac{1}{D} = X_0$, which means that there is only one solution for equilibrium E_1 . To find the stability of the equilibrium E_1 , evaluating the Jacobian matrix of system (5.1) at E_1 results in

$$\begin{aligned} J(E_1) &= \begin{bmatrix} -D - (B + \frac{AY}{Y+C})Y & -(B + \frac{AY}{Y+C})X - \frac{ACXY}{(Y+C)^2} \\ (B + \frac{AY}{Y+C})Y & (B + \frac{AY}{Y+C})X + \frac{ACXY}{(Y+C)^2} - 1 \end{bmatrix}_{(X,Y)=(X_1,Y_1)} \\ &= \begin{bmatrix} -\frac{1}{X_1} & -1 - \frac{ACX_1Y_1}{(Y_1+C)^2} \\ \frac{1}{X_1} - D & \frac{ACX_1Y_1}{(Y_1+C)^2} \end{bmatrix}. \end{aligned} \quad (5.10)$$

Then, the characteristic equation of E_1 is given by

$$\xi^2 - \text{Tr}(J)\xi + \det(J) = 0, \quad (5.11)$$

where

$$\begin{aligned} \det(J) &= -\frac{ACY_1}{(Y_1+C)^2} + \left(\frac{1}{X_1} - D\right)\left(1 + \frac{ACX_1Y_1}{(Y_1+C)^2}\right) \\ &= \frac{1}{X_1} - D - \frac{CD}{Y_1+C} \frac{AY_1}{Y_1+C} X_1 \\ &= \frac{1}{X_1} - D - \frac{CD}{Y_1+C} \left(\frac{1}{X_1} - B\right) X_1 \\ &= \frac{1}{(1-DX_1+C)X_1} \frac{1}{D} [(D+BC)(1-DX_1)^2 \\ &\quad + 2C(D-B)(1-DX_1) + C(B-D)] \\ &= \frac{-(1-DX_1)}{(1-DX_1+C)X_1} [(A+B+D+BC)X_1 - 2(1+C)] \quad (\text{by using (5.5)}) \\ &= \frac{-(1-DX_1)}{2D(A+B)(1-DX_1+C)X_1} [\sqrt{\Delta}(\sqrt{\Delta} \pm (A+B+D+BC))], \end{aligned} \quad (5.12)$$

in which $Y_1 = 1 - DX_1$ and (5.7) have been used. Since it is assumed that $H_1 > 0$, i.e. $\Delta = (A + B + D + BC)^2 - 4(C + 1)D(A + B) > 0$, we have

$$\det(J) < 0 \quad \text{for } X_1 = X_+, \quad \text{and} \quad \det(J) > 0 \quad \text{for } X_1 = X_-. \quad (5.13)$$

When $\det(J) < 0$, the two eigenvalues of the characteristic polynomial (5.11) are real with opposite signs, and thus the equilibrium point $E_{1+} = (X_+, Y_+)$ is a saddle point.

To consider the property of another equilibrium point $E_{1-} = (X_-, Y_-)$, we need to calculate $\text{Tr}(J)$ as follows:

$$\begin{aligned} \text{Tr}(J) &= -\frac{1}{X_1} + \frac{ACX_1Y_1}{(Y_1+C)^2} \\ &= -\frac{1}{X_1} + \frac{CX_1}{Y_1+C} \left(\frac{1}{X_1} - B \right) \\ &= \frac{-(1-DX_1)-C+CX_1(1-BX_1)}{X_1(Y_1+C)} \\ &= -\frac{1}{X_1(Y_1+C)} [BCX_1^2 - (C + D)X_1 + C + 1] \\ &= -\frac{1}{(Y_-+C)} [(BC - DA - DB)X_- + (A + B + BC - C)] \\ &= -\frac{1}{2D(A+B)(Y_-+C)} [AD(A + B - C) + BC(DA + DB + BC) \\ &\quad -(D - B)(A + B)(C + D) + (DA + DB - BC)\sqrt{\Delta}], \end{aligned} \quad (5.14)$$

which can be positive or negative, depending upon the values of parameters. Therefore, the equilibrium point E_{1-} may be a stable (or an unstable) node or focus.

Summarizing the above results, we have shown that when $B < D$, the boundary equilibrium E_0 is a stable node. Moreover, when $H_1 < 0$, a biologically meaningful disease equilibrium E_1 does not exist and E_0 is the unique equilibrium solution, so it is globally asymptotically stable by applying Theorem 5.2.2. When $H_1 \geq 0$, there exist two disease equilibria, E_{1+} and E_{1-} (E_{1-} coincides E_{1+} if $H_1 = 0$, giving rise to a saddle-node bifurcation), and E_{1+} is a saddle point, while E_{1-} may be a stable (or an unstable) node or focus. In this case, no conclusion can be made regarding for the global stability of the disease-free equilibrium E_0 .

When $\det(J) > 0$, we may use $\text{Tr}(J)$ and $\det(J)$ to further classify the equilibrium point E_{1-} . For convenience, let

$$\begin{aligned} H_2 \triangleq & (D-B)(A+B)(C+D) - AD(A+B-C) - BC(DA+DB+BC) \\ & -(DA + DB - BC)\sqrt{\Delta}, \quad (A + B + D + C - \sqrt{\Delta} > 0), \end{aligned} \quad (5.15)$$

and

$$H_3 \triangleq \text{Tr}^2(J) - 4 \det(J), \quad (A + B + D + C - \sqrt{\Delta} > 0). \quad (5.16)$$

Thus, $\text{Tr}(J)$ has the same sign of H_2 , and $\det(J)$ has the same sign of $A + B + D + C - \sqrt{\Delta}$, but H_2 and $A + B + D + C - \sqrt{\Delta}$ only depends upon the parameters A , B , C and D .

Then, E_{1-} can be classified according to the signs of H_2 and H_3 , as shown in Table 5.1, where SF, UF, SN, UN, DSN and DUN stand for Stable Focus, Unstable Focus, Stable Node, Unstable Node, Degenerate Stable Node and Degenerate Unstable Node, respectively.

Now, we consider the numerical values of the parameters used in [27, 28] to demonstrate different dynamical behaviors of system (5.1) for $B < D$. The typical values used [27, 28] are

$$A = 0.364, \quad C = 0.823, \quad D = 0.057, \quad B = 0.060, \quad (5.17)$$

Table 5.1: Classification of E_{1-} ($H_1 \geq 0$).

	$H_2 < 0$	$H_2 > 0$	$H_2 = 0$
$H_3 < 0$	SF	UF	Center
$H_3 > 0$	SN	UN	—
$H_3 = 0$	DSN	DUN	Double-zero

for which viral blips occur. Note that $B = D = 0.057$ is the transcritical point between the equilibrium solutions E_0 and E_1 , and the oscillating behavior (blips) shown in [27, 28] is for $B > D$. Here, we want to change the parameter values near the above set of values for $B < D$ to demonstrate more interesting dynamical behaviors, in particular, the bistable equilibrium solutions, Hopf and generalized Hopf bifurcations, and Bogdanov-Takens (BT) bifurcation. It is not easy to see the relation between H_1 , H_2 and H_3 in a 4-dimensional parameter space. Thus, we fix $B = 0.054$, and choose two values for $D = 0.057, 0.087$, and then plot the three curves $H_1 = H_2 = H_3 = 0$ on the A - C plane, as shown in Figures 5.1 and 5.2, where the red curve, blue curve and green curve correspond to $H_1 = 0$, $H_2 = 0$ and $H_3 = 0$, respectively. We should point out that in [27, 28] the parameter B ($B > D$) is treated as a bifurcation parameter to explore the blips phenomenon. In this paper, we want to take the parameters A and C as bifurcation parameters and investigate their effects on dynamical behavior, since these two parameters involved in the $\beta(X, Y)$ function play a very important role in the modelling. Figures 5.1 and 5.2 clearly indicate the regions corresponding to the classification shown in Table 5.1. If we vary the parameters B and D , we will obtain more such figures, showing rich patterns of dynamical behaviors. It should be noted from Figure 5.1(a) that the very narrow region bounded by the red curve and green curve corresponds to $H_1 > 0$, $H_2 > 0$ and $H_3 > 0$, and thus taking parameter values from this region generate an unstable node E_{1-} . Each point on the curve $\text{Tr}(J) = 0$ yields a Hopf critical point, leading to bifurcation of limit cycles. At the intersection point of the blue curve ($H_2 = 0$) and the green curve ($H_3 = 0$), as shown in Figures 5.1(b) and Figure 5.2(b), $\text{Tr}(J) = \det(J) = 0$, giving rise to a BT bifurcation, characterized by a double-zero eigenvalue. Thus, by using Figures 5.1 and 5.2, we can easily find different values of A and C to get different types of the equilibrium E_{1-} . Also note from these two figures that the BT bifurcation point, marked by a circle, is actually the intersection point of all three curves $H_1 = H_2 = H_3 = 0$. A number of sets of these parameter values and their corresponding classification of E_{1-} are given in Table 5.2. In this section, we present the results for the non-degenerate cases ($H_2 H_3 \neq 0$), and leave the degenerate cases, leading to Hopf and generalized Hopf bifurcations, and BT bifurcation, to be considered later in Sections 5.3 and 5.4.

Hence, when $B < D$ and $H_1 \geq 0$, for positive parameter values, there may exist bistable equilibrium solutions E_0 and E_1 , and bifurcation of limit cycles or even homoclinic orbits from the BT bifurcation.

In the following, we will further investigate the bistable equilibrium solutions in more details using simulation, and then try to provide some biological explanation. For completeness, we also show the results for the cases $H_1 < 0$ and $H_1 = 0$, see Table 5.2, where $A^{(1)} = 0.09559649$, $A^{(2)} = 0.26302225$. Note that the results for the two sets of values in rows three and eight (see Table 5.2) are obtained by taking a point from the narrow region of Figure 5.1(a) and a point from the narrow region of Figure 5.1(b), respectively. We shall

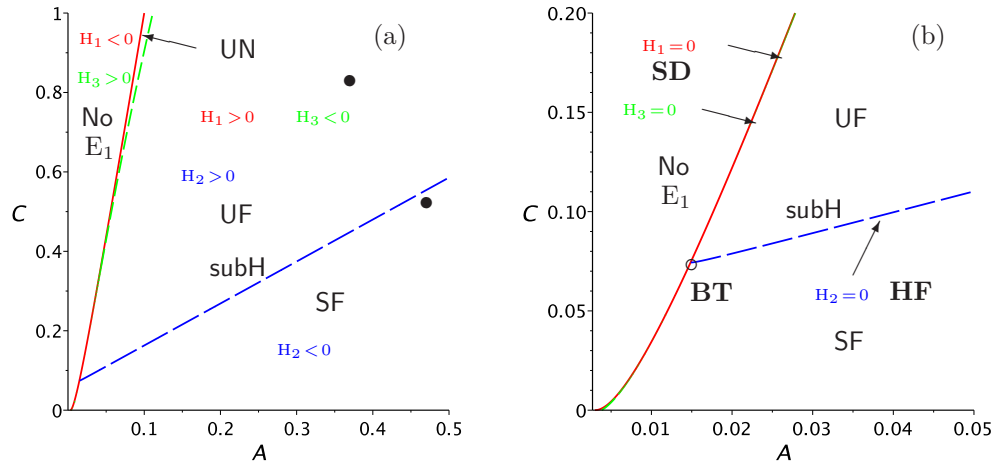


Figure 5.1: (a) Plot of the three curves $H_1 = 0$ (in red), $H_2 = 0$ (in blue) and $H_3 = 0$ (in green), on the A - C plane for $B = 0.054$, $D = 0.057$, with signs of H_1 , H_2 and H_3 indicated; and (b) a zoomed in region near the origin.

Table 5.2: Classification of E_{1-} for given parameter values ($D > B = 0.054$).

A	C	D	E_{1-}	Eigenvalues	H_1	Class
0.100	1.050	0.057	No E_1, E_0 exist	$-0.0570, -0.0526$	< 0	SN
$A^{(1)}$	0.950	0.057	(15.122, 0.1380)	0.0940, 0	$= 0$	DUN
0.100	0.950	0.057	(13.901, 0.2076)	0.0999, 0.0327	> 0	UN
0.364	0.823	0.057	(4.3959, 0.7494)	$0.0858 \pm 0.3747i$	> 0	UF
0.464	0.523	0.057	(2.9509, 0.8318)	$-0.0072 \pm 0.5132i$	> 0	SF
0.260	0.823	0.087	No E_1, E_0 exist	$-0.0870, -0.3793$	< 0	SN
$A^{(2)}$	0.823	0.087	(8.1300, 0.2927)	0.2908, 0	$= 0$	DUN
0.264	0.823	0.087	(7.8326, 0.3186)	0.2719, 0.0165	> 0	UN
0.364	0.823	0.087	(4.9202, 0.5719)	$0.1150 \pm 0.2556i$	> 0	UF
0.364	0.250	0.087	(3.0732, 0.7326)	$-0.0566 \pm 0.4655i$	> 0	SF
5.200	0.223	0.087	(0.2331, 0.9797)	$-1.8817, -2.2251$	> 0	SN

present the simulations for the sets of values in Table 5.2 in the rows 4, 5, 6, 7, 8 and 11, and the corresponding points in the (A, C) parameter space are marked by the black points in Figures 5.1 and 5.2. Also, in Figures 5.1(b) and 5.2(b), the saddle-node (SD) bifurcation, determined by $H_1 = 0$, and the Hopf (HF) bifurcation, determined by $H_2 = 0$, are indicated, and the BT bifurcation is marked by a circle.

5.2.3.1 $A = 0.364, C = 0.823, D = 0.057, B = 0.054$

For this set of parameter values, system (5.1) has three equilibrium solutions: $E_0 = (X_0, Y_0) = (17.5439, 0)$, $E_{1+} = (X_{1+}, Y_{1+}) = (17.4056, 0.0079)$ and $E_{1-} = (X_{1-}, Y_{1-}) = (4.3959, 0.7494)$.

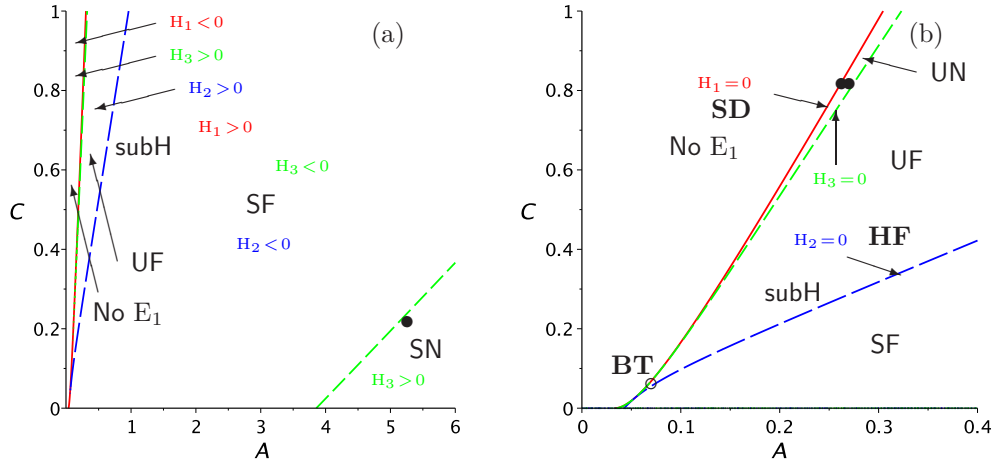


Figure 5.2: (a) Plot of the three curves $H_1 = 0$ (in red), $H_2 = 0$ (in blue) and $H_3 = 0$ (in green), on the A - C plane for $B = 0.054$, $D = 0.087$, with signs of H_1 , H_2 and H_3 indicated; and (b) a zoomed in region near the origin.

It can be shown that E_0 is a stable node, E_{1+} is a saddle point, while E_{1-} is an unstable focus. The phase portrait is shown in Figure 5.3, indicating that there do not exist limit cycles, and the disease-free equilibrium E_0 is actually globally asymptotically stable.

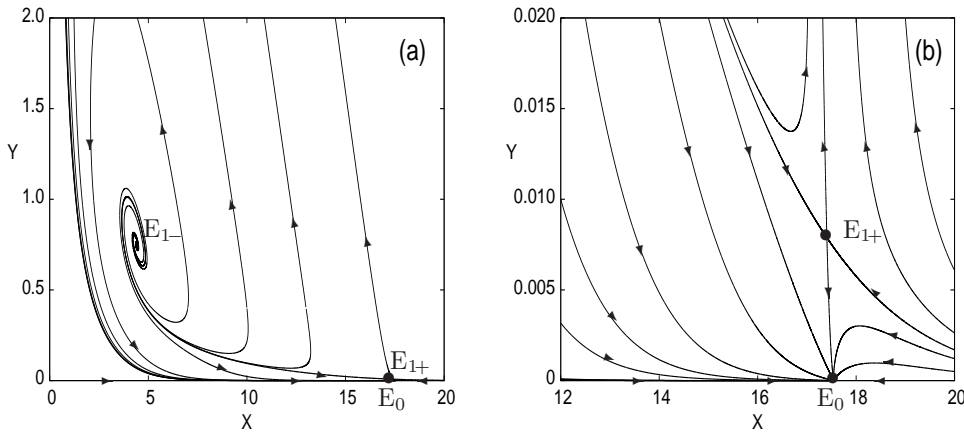


Figure 5.3: Simulated phase portrait of system (5.1) for $A = 0.364$, $C = 0.823$, $D = 0.057$, $B = 0.054$, showing the global stability of E_0 : (a) depicting three equilibrium points E_0 , E_{1+} , E_{1-} ; and (b) showing E_0 , E_{1+} in a zoomed in region.

5.2.3.2 $A = 0.464$, $C = 0.523$, $D = 0.057$, $B = 0.054$

For this set of parameter values, system (5.1) still has three equilibrium solutions: $E_0 = (X_0, Y_0) = (17.5439, 0)$, remains unchanged from the previous case since D is not changed, and is a stable node; $E_{1+} = (X_{1+}, Y_{1+}) = (17.4800, 0.0036)$ is still a saddle point, but now $E_{1-} = (X_{1-}, Y_{1-}) = (2.9509, 0.8318)$ becomes a stable focus. The phase portrait for this case is depicted in Figure 5.4, which shows an unstable limit cycle enclosing the stable focus E_{1-} .

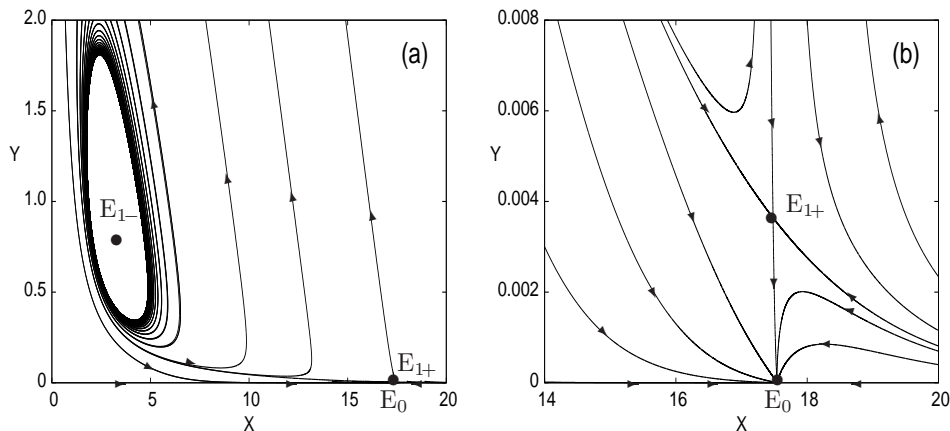


Figure 5.4: Simulated phase portrait of system (5.1) for $A = 0.464$, $C = 0.523$, $D = 0.057$, $B = 0.054$, showing the bistable equilibria E_0 and E_{1-} and an unstable limit cycle: (a) depicting three equilibrium points E_0 , E_{1+} , E_{1-} ; and (b) showing E_0 , E_{1+} in a zoomed in region.

Thus, for this set of parameter values, there exist bistable equilibrium solutions E_0 and E_{1-} . The attracting region for E_{1-} is the region inside the limit cycle, while the area outside the limit cycle is the attracting region for E_0 .

To view the bistable equilibrium solutions, we plot the bifurcation diagram in the A - X plane for fixed values: $C = 0.523$, $D = 0.057$, $B = 0.054$, as shown in Figure 5.5, where the solid red line and blue curve denote the stable equilibria E_0 and E_{1-} , respectively, while the dashed blue line represents the unstable equilibrium E_{1+} . A saddle-node bifurcation point is seen between E_{1-} and E_{1+} , which is actually the underlying cause for the existence of bistable equilibrium solutions. In fact, the saddle-node bifurcation point is the turning point on the solution curve E_1 .

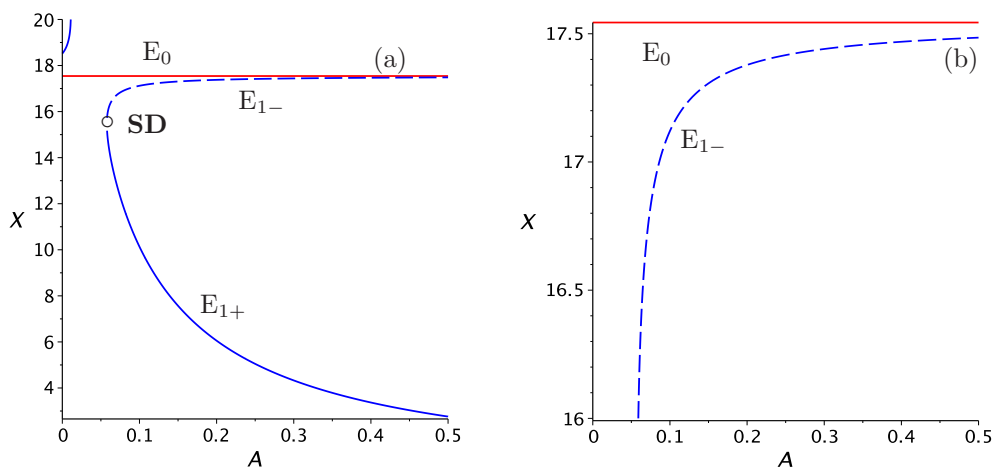


Figure 5.5: (a) Bifurcation diagram for the bistable equilibrium solutions for $B = 0.054$, $C = 0.823$, $D = 0.057$; and (b) a zoomed in region near the equilibrium E_0 .

5.2.3.3 $A = 0.264, C = 0.823, D = 0.087, B = 0.054$

For this set of parameter values, system (5.1) has three equilibrium solutions: $E_0 = (X_0, Y_0) = (11.4943, 0)$, a stable node; $E_{1+} = (X_{1+}, Y_{1+}) = (8.4127, 0.2681)$, a saddle point; and $E_{1-} = (X_{1-}, Y_{1-}) = (7.8326, 0.3186)$, an unstable node. The phase portrait for this case is given in Figure 5.6, showing that there do not exist limit cycles, and the disease-free equilibrium E_0 is actually globally asymptotically stable.

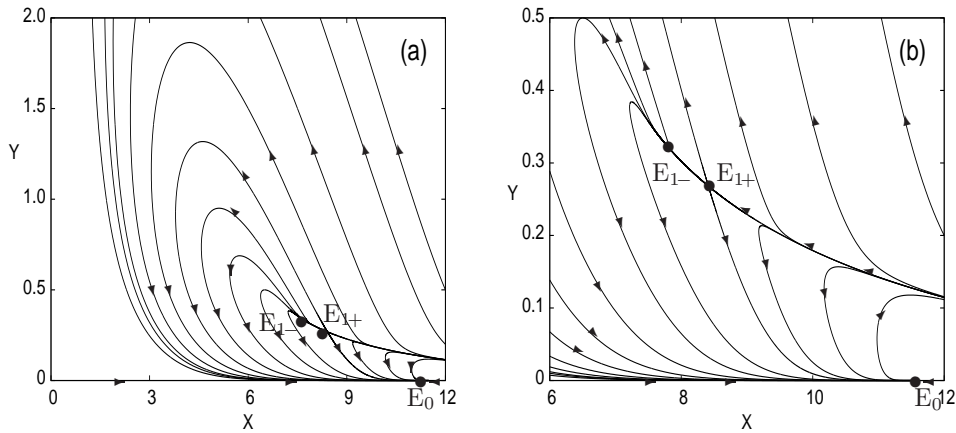


Figure 5.6: Simulated phase portrait of system (5.1) for $A = 0.264, C = 0.823, D = 0.087, B = 0.054$, showing the global stability of E_0 : (a) depicting three equilibrium points E_0, E_{1+}, E_{1-} ; and (b) showing E_0, E_{1+} in a zoomed in region.

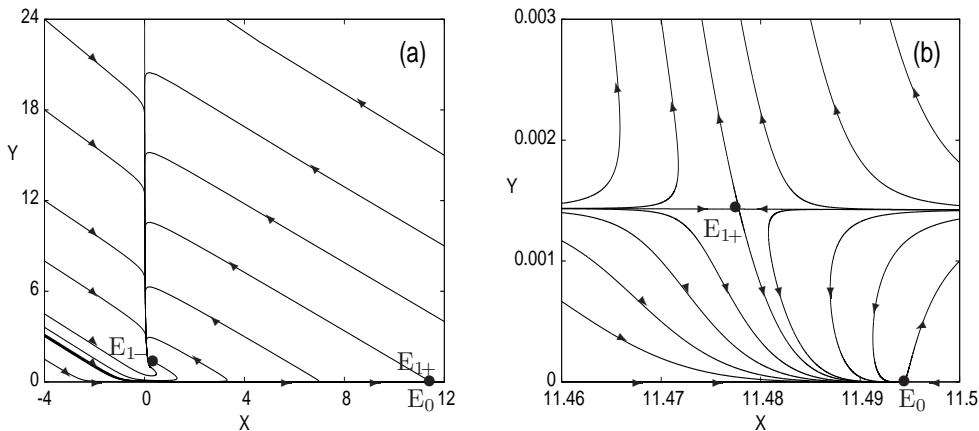


Figure 5.7: Simulated phase portrait of system (5.1) for $A = 5.2, C = 0.223, D = 0.097, B = 0.064$, showing the bistable equilibria E_0 and E_{1-} : (a) depicting three equilibrium points E_0, E_{1+}, E_{1-} ; and (b) showing E_0, E_{1+} in a zoomed in region.

5.2.3.4 $A = 5.200, C = 0.223, D = 0.087, B = 0.054$

For this set of parameter values, system (5.1) still has three equilibrium solutions: $E_0 = (X_0, Y_0) = (11.4943, 0)$, a stable node; $E_{1+} = (X_{1+}, Y_{1+}) = (11.4778, 0.0014)$, a saddle point;

and $E_{1-} = (X_{1-}, Y_{1-}) = (0.2331, 0.9797)$, a stable node. The phase portrait for this case is depicted in Figure 5.7, which shows no limit cycles to exist, but there still exist bistable equilibrium solutions E_0 and E_{1-} . The attracting regions for E_0 and E_{1-} are separated by the two trajectories passing through the saddle point E_{1+} . A similar bifurcation diagram like that given in Figure 5.5 can be obtained.

5.2.3.5 $A = 0.26302225$, $C = 0.823$, $D = 0.087$, $B = 0.054$ ($H_1 = 0$)

For this set of parameter values, $H_1 = 0$ under which $E_{1+} = E_{1-} = E_1 = (8.1300, 0.2927)$, which is an unstable node, and thus the disease-free equilibrium $E_0 = (11.4943, 0)$ is globally asymptotically stable, as shown in Figure 5.8.

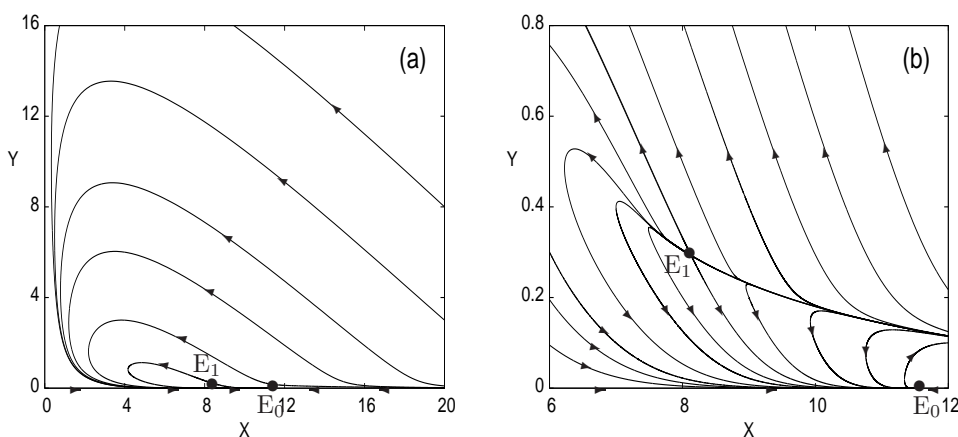


Figure 5.8: Simulated phase portrait of system (5.1) for $A = 0.26302225$, $C = 0.823$, $D = 0.087$, $B = 0.054$, showing the global stability of E_0 : (a) depicting two equilibrium points E_0 , E_1 ; and (b) showing E_0 , E_1 in a zoomed in region.

5.2.3.6 $A = 0.260$, $C = 0.823$, $D = 0.087$, $B = 0.054$ ($H_1 < 0$)

For this set of parameter values, $H_1 = -0.002136 < 0$ under which E_1 does not exist, and so the disease-free equilibrium $E_0 = (11.4943, 0)$ is globally asymptotically stable. The simulated phase portrait is similar to Figure 5.8(a) but without the existence of E_1 .

The most interesting phenomenon found in this section for $B < D$ is the bistable equilibrium solutions E_0 and E_1 . E_0 is always a stable node, while E_{1-} may be a stable focus (see Figure 5.4) or a stable node (see Figure 5.7). The separator between the two attracting regions of the two stable equilibria is either an unstable limit cycle (see Figure 5.4) or the saddle trajectories. Dynamically, this bistable phenomenon is due to the existence of a saddle-node bifurcation on the equilibrium solution E_1 , which has two branches, one of them is stable and the other is unstable. Biologically, this phenomenon is not fully understood. System (5.1) was developed from an in-host model of HIV infection, and there has been evidence of possible bistability in this disease. In particular, the equilibrium viral load, or “viral set point” can differ by orders of magnitude among patients. Several authors have previously suggested bistable equilibrium solutions as an explanation for the phenomenon [2, 18].

It is also noted from Figures 5.3, 5.6 and 5.8 that when E_{1-} is unstable (either focus or node), the equilibrium E_0 seems globally asymptotically stable. This may be explained as follows: first, it can be seen from Figure 5.1 that when the parameters A and C are varied to cross the blue curve, defined by $H_2 = 0$, from the bottom-right to the top-left (e.g., in the negative direction of the A -axis), the equilibrium E_{1-} changes from a stable focus (SF) to an unstable focus (UF). Hopf bifurcation occurs when the parameters are varied to cross the blue curve. The simulations shown in Figures 5.4 and 5.3 correspond to the two points chosen from the SF region and UF region, respectively, implying that the Hopf bifurcation is subcritical. This is why an unstable limit cycle is shown in Figure 5.4, while there is no limit cycle in Figure 5.3 and so all trajectories converge to the stable node E_0 . Similarly, the simulations shown in Figures 5.6 and 5.8 imply that when the parameters A and C are varied to cross the blue curve ($H_2 = 0$ in Figure 5.2) from the bottom-right to the top-left, a subcritical Hopf bifurcation occurs. The proof for the two subcritical Hopf bifurcations will be given in Section 5.3.

5.2.4 Dynamical behavior of (5.1) when $B > D$

Now, we discuss the dynamical behavior of system (5.1) for $B > D$. In this case, E_0 becomes a saddle point, while E_1 always exists, since equation (5.5) always has two roots for

$$\begin{aligned}\Delta &= (A + B + D + BC)^2 - 4(C + 1)D(A + B) \\ &= (A + B - D - BC)^2 + 4C(A + B)(B - D) \\ &> (A + B - D - BC)^2 \quad (\text{due to } B > D) \\ &\geq 0,\end{aligned}$$

and thus $0 < X_- < X_+$. Further, noticing from (5.5) that $Q(0) = C + 1 > 0$ and $Q(\frac{1}{D}) = -C(\frac{B}{D} - 1) < 0$ ($B > D$), we have

$$0 < X_- < \frac{1}{D} < X_+.$$

Thus,

$$X_1 = X_- = \frac{(A + B + D + BC) - \sqrt{\Delta}}{2D(A + B)}, \quad \text{since } X_1 \in \left[0, \frac{1}{D}\right],$$

which guarantees that $0 \leq Y_1 = 1 - DX_1 \leq 1$.

Since E_0 is a saddle point (unstable), and $X_{1+} > \frac{1}{D}$ (which yields $Y_{1+} < 0$) is not biologically meaningful, bistable equilibria cannot exist for this case $B > D$. To find the stability of E_1 (i.e., E_{1-}) when $B > D$, we first show that $\det(J) > 0$. This can be obtained using (5.12) as follows:

$$\begin{aligned}\det(J) &= \frac{1}{X_1} - D + \frac{CD}{Y_1 + C} (BX_1 - 1) \\ &> \frac{1}{X_1} - D + \frac{CD}{Y_1 + C} (DX_1 - 1) \quad (B > D) \\ &= \left(\frac{1}{X_1} - D\right) \left(1 - \frac{CDX_1}{Y_1 + C}\right) \quad (0 < DX_1 < 1, 0 < Y_1 < 1) \\ &> \left(\frac{1}{X_1} - D\right) (1 - DX_1) \\ &= \frac{1}{X_1} (1 - DX_1)^2 > 0.\end{aligned}$$

Therefore, all the formulae derived in the previous section for E_{1-} (when $B < D$) and the results shown in Table 5.1 can be applied here to classify the type of the equilibrium E_{1-} (when $B > D$). Similarly, we may fix B and D and then plot the two curves $H_2 = H_3 = 0$ on the A - C plane to identify the possible parameter values which yield different qualitative behavior of system (5.1). Note that now for $B > D$ we do not need the condition $H_1 > 0$ since $\Delta > 0$ is guaranteed when $B > D$. Two sets of values for $(B, D) = (0.057, 0.060), (0.087, 0.090)$ are chosen to plot the figures. However, it is found that these two figures are quite similar, implying that, unlike the case $B < D$, here slightly varying B and D does not change the behavior of the system. Hence, we only present the result for $(B, D) = (0.057, 0.060)$, as shown in Figure 5.9. It can be seen that for this case, there is no saddle-node bifurcation, nor BT bifurcation, since $H_1 > 0$ for all parameter values.

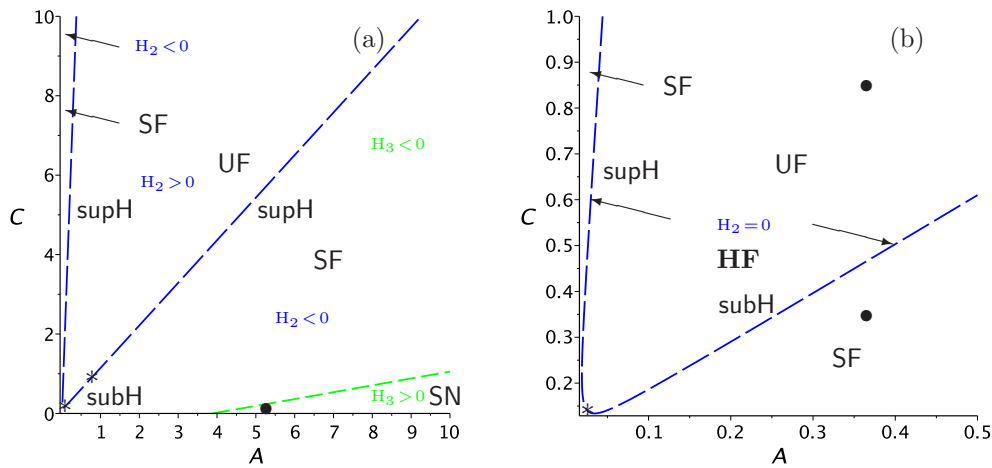


Figure 5.9: (a) Plot of two curves $H_2 = 0$ (in blue) and $H_3 = 0$ (in green), on the A - C plane for $B = 0.060, D = 0.057$, with signs of H_2 and H_3 indicated; and (b) a zoomed in region near the origin.

It is also seen from Figure 5.9 that for most of the parameter values, $H_3 < 0$, in particular for not very large values of A . This means that for most of parameter values, E_1 is a focus. Further, it can be shown that for the points bounded by the blue curve ($H_2 = 0$, i.e. $\text{Tr}(J) = 0$) the equilibrium E_1 is an unstable focus. Therefore, for these parameter values, by Theorem 5.2.2, we can conclude that there exists at least one stable limit cycle inside the trapping region G . When the parameter values are taken from the region outside the region bounded by the blue curve, the equilibrium E_1 is the unique equilibrium inside the trapping region G , and thus the equilibrium E_1 is globally asymptotically stable.

Now, we are ready to prove the following theorem.

Theorem 5.2.4 *When $B > D$ and $H_2 > 0$, system (5.1) has at least one stable limit cycle, and the limit cycle must not bifurcate from a homoclinic orbit.*

Proof First, we show that the positive equilibrium $E_1 = E_{1-} = (X_1, Y_1)$ is inside the trapping region G , defined in Theorem 5.2.2. That is, the point E_1 should be below the line $L: X + Y =$

$\max(1, \frac{1}{D}) + \varepsilon$. Note that $Y_1 = 1 - DX_1$ for $0 < X_1 < \frac{1}{D}$, implying that the point (X_1, Y_1) is on the line, defined by $DX + Y = 1$, which is obviously below the line L .

To prove that limit cycles do not bifurcate from a homoclinic orbit, first note that the only possible homoclinic orbit comes from the saddle point E_0 when $B > D$. Thus, it suffices to show that there do not exist homoclinic orbits passing through this singular point. Otherwise, suppose there exists a homoclinic orbit passing through this point, then the homoclinic orbit must leave this point along the direction of the eigenvector $v_2 = (1, \frac{D}{B}(1 - D) - 1)$ and return to this point along the direction of the eigenvector $v_1 = (1, 0)$, that is, the direction of the X -axis. In other words, the homoclinic orbit must return to the saddle point along the X -axis. But we have already shown that the X -axis itself is a solution trajectory, and thus other trajectories, in particular, the one leaving the saddle point along the v_2 direction, cannot connect to the X -axis due to the uniqueness of solutions.

The proof of Theorem 5.2.4 is complete.

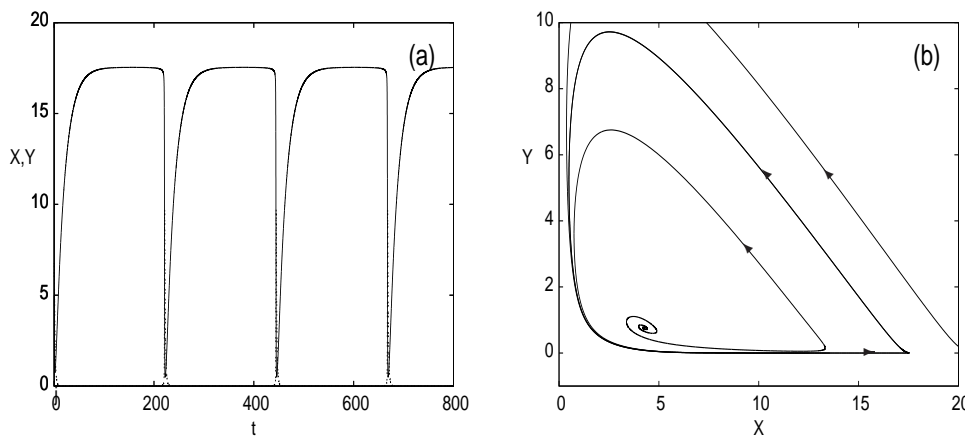


Figure 5.10: Simulated blips of system (5.1) for $A = 0.364$, $C = 0.823$, $D = 0.057$, $B = 0.060$: (a) time history showing blips; and (b) phase portrait showing a limit cycle.

To end this section, we present three simulations for the common parameter values: $D = 0.057$, $B = 0.060$; but for $(A, C) = (0.364, 0.823)$, $(0.364, 0.350)$ and $(5.2, 0.2)$, respectively. The first simulation is shown in Figure 5.10, which yields a blip-like oscillation, as has been discussed in [27, 28]. The simulations for the second and third cases are depicted in Figures 5.11(a) and (b), respectively. Figure 5.11(a) shows that E_1 is asymptotically stable and all trajectories starting from the initial points inside an unstable limit cycle converge to this equilibrium E_1 ; while trajectories outside the unstable limit cycle converge to a separator of the saddle point E_0 . Figure 5.11(b) indicates that E_1 is globally asymptotically stable without the existence of limit cycles.

The results shown in Figures 5.10 and 5.11 clearly indicate that the Hopf bifurcation which occurs on the left branch of the blue curve in Figure 5.9 is supercritical (when, say, A is increasing to cross the blue curve), generating the stable limit cycle (blips) shown in Figure 5.10, and the bifurcation which occurs on the right branch of the blue curve (see Figure 5.9) is subcritical (when, say, A is decreasing to cross the blue curve), leading to the unstable limit cycle shown in figure 5.11(a). The proof for the supercritical and subcritical Hopf bifurcations will be given in Section 5.3.

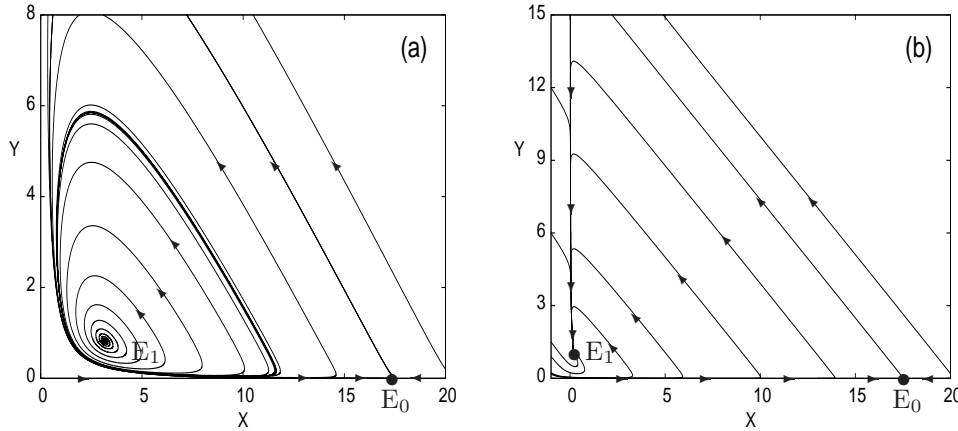


Figure 5.11: Simulated phase portrait of system (5.1) when $(B, D) = (0.060, 0.057)$: (a) for $(A, C) = (0.364, 0.352)$ showing the trajectories inside an unstable limit cycle convergent to the stable equilibrium E_1 ; and (b) for $(A, C) = (5.2, 0.2)$ showing the global stability of E_1 without existence of limit cycles.

5.2.5 Dynamical behavior of (5.1) when $B = D$

We now turn to the case $B = D$. First note that when $B = D$, the equilibrium $E_{1+} = (X_+, Y_+)$ coincides with the disease-free equilibrium E_0 , while the other equilibrium $E_{1-} = (X_-, Y_-) = (\frac{1+C}{A+D}, \frac{A-DC}{A+D})$. In order to have $X_- < \frac{1}{D}$, we require $A + D > D + DC$, or $A > DC$. Note that when $A < DC$, the equilibrium E_{1-} does not exist; and when $A = DC$, the equilibrium E_{1-} also coincides with E_0 . So for the generic case, we assume $A > DC$ in this subsection.

To find the stability of E_0 for this case, we note that the two eigenvalues associated with this equilibrium now become $-D$ and 0 , which is a critical case and the application of center manifold theory is required to determine its stability. To achieve this, we introduce an affine transformation, given by

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \frac{1}{D} \\ 0 \end{pmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & -D \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad (5.18)$$

into (5.1) to obtain a system, expanded around $(u_1, u_2) = (0, 0)$, as

$$\begin{aligned} \frac{du_1}{d\tau} &= -Du_1 + D(D-1)u_1u_2 + \frac{1}{C}(A-DC)(1-D)u_2^2 \\ &\quad + \frac{AD}{C}(1-D)u_1u_2^2 + \dots, \\ \frac{du_2}{d\tau} &= Du_1u_2 - \frac{1}{C}(A-DC)u_2^2 - \frac{AD}{C}u_1u_2^2 + \dots, \end{aligned} \quad (5.19)$$

whose linear part is now in the Jordan canonical form with eigenvalues $-D$ and 0 . To find the center manifold, let $u_1 = h(u_2) = a_2u_2^2 + O(u_2^3)$ and then use (5.19) to find $a_2 = \frac{(1-D)(A-DC)}{DC}$. Therefore, the center manifold up to second order is given by

$$W^C = \{(u_1, u_2) \mid u_1 = \frac{(1-D)(A-DC)}{DC}u_2^2 + O(u_2^3)\},$$

and the differential equation describing the dynamics on the center manifold is

$$\frac{du_2}{d\tau} = -\frac{1}{C}(A - DC)u_2^2 + \frac{(1 - D)(A - DC)}{C}u_2^3 + O(u_2^4). \quad (5.20)$$

Since $Y = -Du_2 > 0$, we only consider $u_2 < 0$. Note that the leading term in (5.20) is $-\frac{1}{C}(A - DC)u_2^2$ with a negative coefficient, implying that u_2 is decreasing from a negative initial value (and so Y is increasing from a positive initial value). Hence, the equilibrium E_0 is a degenerate saddle point, similar to the case when $B > D$.

Next, we consider the stability of E_{1-} . Evaluating the Jacobian (5.10) at this equilibrium yields

$$J(E_{1-}) = \begin{bmatrix} -\frac{A+D}{1+C} & -\frac{A+2AC-DC^2}{A(1+C)} \\ \frac{A-DC}{1+C} & \frac{C(A-DC)}{A(1+C)} \end{bmatrix},$$

which in turn results in two eigenvalues, given by

$$\xi_{\pm} = \frac{-[C(A-DC)-A(A+D)] \pm \sqrt{[C(A-DC)-A(A+D)]^2 - 4A(1+C)(A-DC)^2}}{2A(1+C)}. \quad (5.21)$$

Hence, under the condition $A > DC$, the equilibrium E_{1-} is asymptotically stable (unstable) if $C(A - DC) - A(A + D) < 0$ (> 0), which is a node (focus) when $[C(A - DC) - A(A + D)]^2 - 4A(1 + C)(A - DC)^2 > 0$ (< 0). In order to find parameter values for these four categories, let

$$\begin{aligned} H_1^* &\triangleq H_1^{B=D} = A - DC, \\ H_2^* &\triangleq H_2^{B=D} = C(A - DC) - A(A + D), \\ H_3^* &\triangleq H_3^{B=D} = [C(A - DC) - A(A + D)]^2 - 4A(1 + C)(A - DC)^2. \end{aligned}$$

Then choosing $B = 0.057$, we plot the three curves $H_1^* = H_2^* = H_3^* = 0$ on the A - C plane, as shown in Figure 5.12, from which it is easy to find the parameter values which correspond to different classifications of the equilibrium E_{1-} . Since the equilibrium E_0 is a degenerate saddle node and only one solution exists for E_1 , this case $B = D$ is similar to the case $B > D$. Thus, in general, if E_1 is unstable (either a focus or a node), then there must exist stable limit cycles; if E_1 is stable, then it is globally asymptotically stable. When the parameter values of A and C are chosen from the blue curve (see Figure 5.12) defined by $H_2^* = 0$, Hopf bifurcation occurs, leading to limit cycles. This will be further discussed in the next section.

5.3 Hopf and generalized Hopf bifurcations

In this section, we consider bifurcation of limit cycles due to Hopf and generalized Hopf bifurcations. There are three types of Hopf bifurcations, which occur from the critical blue line $H_2 = 0$ for $B < D$ (see Figures 5.1 and 5.2) and $B > D$ (see Figures 5.9), and from the critical blue line $H_2^* = 0$ for $B = D$ (see Figure 5.12). First we give a detailed analysis for the case $B = D$, and then summarize the results for other cases with representative simulations.

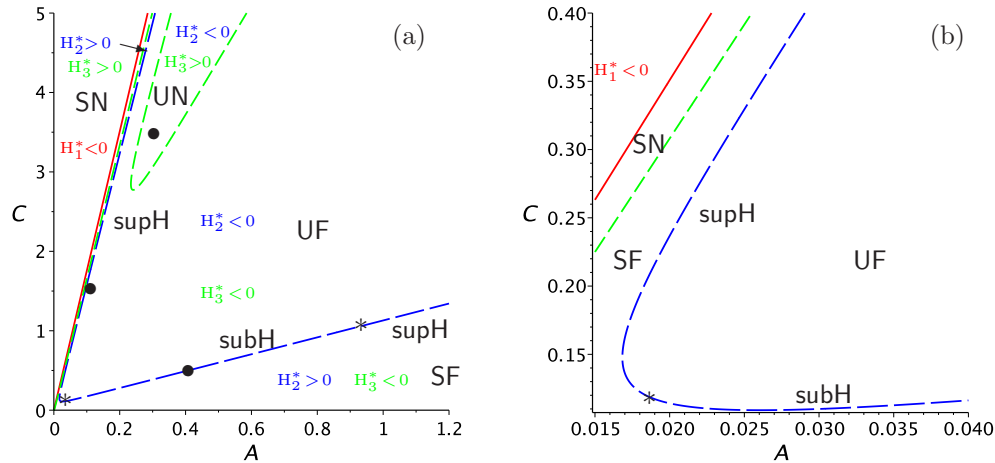


Figure 5.12: (a) Plot of the three curves $H_1^* = 0$ (in red), $H_2^* = 0$ (in blue) and $H_3^* = 0$ (in green), on the A - C plane for $B = D = 0.057$, with the regions indicated for the classifications SF, UF, SN and UN; and (b) a zoomed in region near the turning point.

5.3.1 Hopf bifurcation

We first consider Hopf bifurcation, starting from the case: $B = D = 0.057$, for which the Hopf critical points are located on the blue curve defined by $H_2^* = 0$ (see Figure 5.12) is determined from the equation $A(A + D) - C(A - DC) = 0$, from which we solve for C to obtain

$$C_{\pm} = \frac{500 \pm \sqrt{A(19300A - 3249)}}{57A}, \quad \left(A > \frac{3249}{193000}\right), \quad (5.22)$$

where we use $B = D = \frac{57}{1000}$ to facilitate symbolic computation. Note that the leftmost point on the blue curve is given by $(A, C) = \left(\frac{3249}{193000}, \frac{57}{386}\right)$. The solutions C_- and C_+ correspond to the points (see Figure 5.12(b)) on the upper and lower branches of the $H_2^* = 0$ curve, respectively. In order to apply normal form theory to calculate the first-order focus value (or the first Lyapunov constant), we introduce an affine transformation, given by

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \frac{1000(1+C)}{1000A+57} \\ \frac{1000A-57C}{1000A+57} \end{pmatrix} + \begin{bmatrix} 1 & 0 \\ \frac{-A(1000A+57)}{1000+2000AC-57C^2} & \frac{-1000A(1+C)\omega_c}{1000+2000AC-57C^2} \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix},$$

where $\omega_c = \frac{1000A-57C}{1000\sqrt{A(1+C)}} > 0$ (since $1000A-57C > 0$ due to $Y > 0$), into (5.1) to yield a system to be expanded around $(u_1, u_2) = (0, 0)$ up to third-order terms, and then apply the Maple program for computing the normal forms associated with Hopf and generalized Hop bifurcations [26] to this system to obtain the normal form in polar coordinates up to third-order terms as follows:

$$\frac{dr}{d\tau} = r[v_0\mu + v_1r^2 + o(r^4)], \quad \frac{d\theta}{d\tau} = \omega_c + t_0\mu + t_1r^2 + o(r^4), \quad (5.23)$$

where μ is a perturbation parameter to measure the distance from a critical point on the blue curve $H_2^* = 0$ along the positive direction of the A -axis. v_0 and v_1 are the zero-order and the first-order focus values. The first equation of (5.23) can be used to perform bifurcation analysis and the sign of v_1 determines whether the Hopf bifurcation is supercritical or subcritical. The

values v_0 and t_0 can be found from a linear analysis, while v_1 and t_1 are obtained by applying the Maple program. The calculation shows that

$$v_0 = \frac{57C^2 - 1000A^2}{2000A^2(1+C)}, \quad t_0 = \frac{1000A+57C}{4000A\sqrt{A(1+C)}}, \quad (5.24)$$

and the output of the Maple program gives v_{1-} and v_{1+} , corresponding to C_- and C_+ , respectively, as

$$\begin{aligned} v_{1\pm} = & \frac{-3249(1000A+57)^3}{8000000000A(500A+A_m)(500A+57-A_m)^3(557000A+60249-1000A_m)^3} \\ & \times [(3864999285035000000000000A^5 + 861408257780985000000000A^4 \\ & + 7051942944965614500000A^3 + 223356947766097675500A^2 \\ & - 3214238968494000000A + 38317671392498001) \\ & \pm (8796766369990000000000A^4 + 2033715969208290000000A^3 \\ & + 17848597867145253000A^2 + 759905488695261807A \\ & + 24859340130996000)A_m]. \end{aligned}$$

where $A_m = \sqrt{A(193000A - 3249)}$. It can be shown that $v_{1+} < 0$ for $A > \frac{3249}{193000} \approx 0.0168$. For v_{1-} , it has two real roots: $A = 0.0184$ and $A = 0.9210$ such that $v_{1-} > 0 \forall A \in (0.0184, 0.9210)$ and $v_{1-} < 0 \forall A \in (0.0168, 0.0184) \cup (0.9210, \infty)$. Moreover, it can be shown that $v_0 > 0$ when $C = C_+$ for any values of $A > 0.0168$, and there is a critical point on C_- , defined by $A = 0.0260$, such that when $C = C_-$, $v_0 > 0$ for $A \in (0.0168, 0.0260)$ but $v_0 < 0$ for $A > 0.0260$. Therefore, we can combine the information on the signs of v_0 and v_1 to precisely determine whether a Hopf bifurcation is supercritical or subcritical. In fact, on the upper branch C_+ of the blue curve $H_2^* = 0$, all Hopf bifurcations are supercritical, while on the lower branch C_- , the Hopf bifurcation is supercritical for $A \in (0.0168, 0.0184) \cup (0.9210, \infty)$, and subcritical for $A \in (0.0184, 0.9210)$, as shown in Figure 5.12, where the two points on the blue curves, at $A = 0.0184$ and $A = 0.9210$ are marked by *, where ‘supH’ and ‘subH’ represent supercritical and subcritical Hopf bifurcations, respectively.

It should be pointed out that since E_0 is a degenerate saddle point, for any point inside the region bounded by the blue curve, there must exist stable limit cycles due to Poincaré-Bendixson theory no matter whether E_1 is an unstable focus or node. This seems to imply a contradiction for the subcritical Hopf bifurcation from the lower branch of the blue curve for $A \in (0.0184, 0.9210)$, giving rise to unstable limit cycles below the curve. But on the other side, there exist stable limit cycles. This is because the unstable limit cycle is from a local (Hopf) bifurcation, while the stable limit cycle comes from a global bifurcation. Several representative parameter sets (A, C) are chosen for this case when $B = D = 0.057$ as follows:

$$(A, C) = (0.1, 1.55), \quad (0.3, 3.5), \quad (0.42, 0.50), \quad (0.39, 0.50),$$

which are marked on Figure 5.12 by black points (the last two are at the same place), and the corresponding simulations are shown in Figures 5.13 and 5.14. Note that all of them show the existence of limit cycles. The first two cases confirm that the Hopf bifurcations emerging from the upper branch of the blue curve are indeed supercritical (with the focus value $v_{1+} < 0$), and so the bifurcating limit cycles are stable (see Figure 5.13). The last two points are very close, with one below the curve and one above the curve. The third one yields a typical subcritical

Hopf bifurcation and the bifurcating limit cycle is unstable (see Figure 5.14(a)). The last one is not generated by Hopf bifurcation though the critical point is near the blue curve. It is a big limit cycle, generated due to Poincaré-Bendixon theory, and it is stable since it encloses an unstable focus (see Figure 5.14(b)).

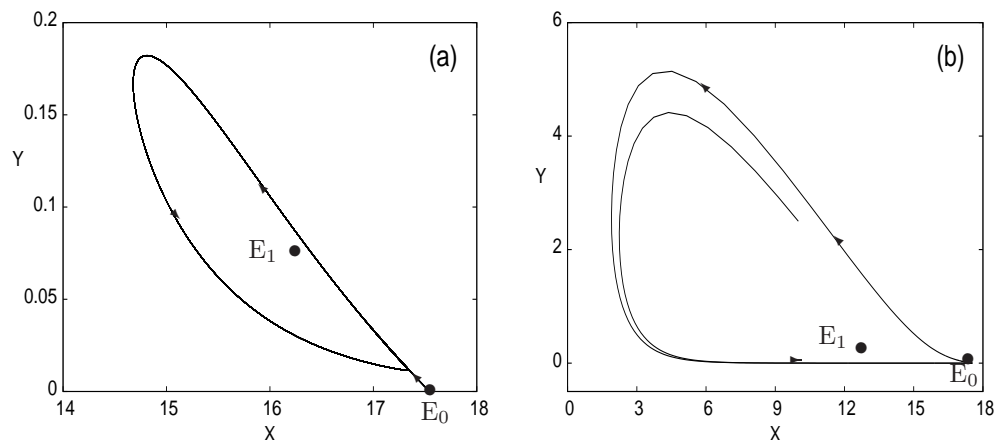


Figure 5.13: Simulations of system (5.1) when $B = D = 0.057$, showing stable limit cycles: (a) $(A, C) = (0.1, 1.55)$ with E_1 being an unstable focus; and (b) $(A, C) = (0.3, 3.5)$ with E_1 being an unstable node.

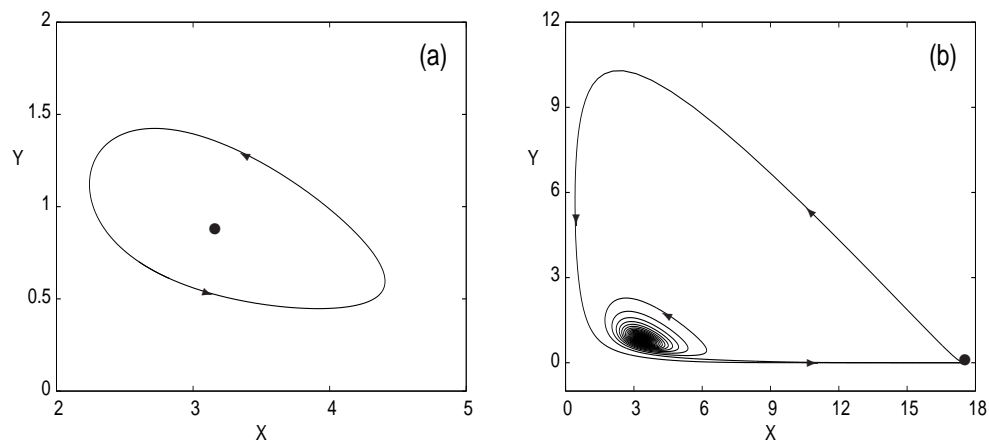


Figure 5.14: Simulations of system (5.1) when $B = D = 0.057$, showing (a) an unstable limit cycle for $(A, C) = (0.42, 0.50)$ with E_1 being a stable focus; and (b) a stable limit cycle for $(A, C) = (0.39, 0.50)$ with E_1 being an unstable focus.

Similarly, we can consider the cases $B < D$ and $B > D$ and determine whether the Hopf bifurcations are supercritical or subcritical. Without giving detailed calculations, we summarize the results as follows. For the case with $B = 0.054 < D = 0.057$, the blue curve actually has a turning point at $A = \frac{729}{49000} \approx 0.014878$ while the BT bifurcation point is above this point at $A = 0.014881$, as shown in Figure 5.16(a) in the next section. On the lower branch of the blue curve, the focus value for the Hopf bifurcation (see the blue curve in Figure 5.1) is shown to have the property that $\nu_1 > 0$ for $A \in (0.014981, 0.9455)$ and $\nu_1 < 0$

for $A \in (0.014878, 0.014981) \cup (0.9455, \infty)$. On the upper branch of the blue curve, $\nu_1 < 0$ for $A \in (0.014878, 0.014881)$. Hence, when $(A, C) = (0.364, 0.823)$, the Hopf bifurcation is subcritical, and the bifurcating limit cycle is unstable, as the example shown in Figure 5.4. We expect that a Hopf bifurcation is supercritical when choosing a point with $A > 0.9455$. For the case with $B = 0.054 < D = 0.087$ (see Figure 5.2), only the upper branch of the blue curve is the solution, which does not contain the turning point, as shown in Figure 5.16(b) (in the next section). It is found that the focus value $\nu_1 > 0$ for $A \in (0.0393, 1.1708)$ and $\nu_1 < 0$ for $A > 1.1708$. But for this case, the BT bifurcation point is at $A = 0.0529$, and the portion for $A < 0.0529$ yields $H_1 < 0$. Therefore, for this case, $\nu_1 > 0$ for $A \in (0.0529, 1.1708)$. Several typical simulations can be seen in Figures 5.3, 5.4, 5.6–5.8.

Finally, we consider the case $B = 0.060 > D = 0.057$ and confirm the conclusion that we made at the end of Section 5.2.4. Note that for this case $H_1 > 0$ for all positive parameter values. Compared to the case $B < D$, now there are two branches on the blue curve (see Figure 5.9). For the upper branch, it can be shown that $\nu_{1+} < 0$ for $A > 0.0189$, and the Hopf bifurcation emerging from the upper branch of the blue curve is supercritical and so the bifurcating limit cycles are stable (see the blips example in Figure 5.10). For the lower branch of the blue curve, it can be shown that the focus value $\nu_{1-} > 0$ for $A \in (0.0214, 0.8964)$ and $\nu_{1-} < 0$ for $A \in (0.0189, 0.0214) \cup (0.8964, \infty)$. Hence, the Hopf bifurcation from the lower branch of the blue curve is subcritical for $A \in (0.0214, 0.8964)$, giving rise to unstable limit cycles (an example is shown in Figure 5.11). When $A \in (0.0189, 0.0214) \cup (0.8964, \infty)$, the Hopf bifurcation becomes supercritical and so the bifurcating limit cycles are stable. This is similar to the case $B = D$ (see Figure 5.12 where supercritical and subcritical Hopf bifurcations are indicated), and thus we omit the details.

By comparing the Figures 5.1, 5.2, 5.9 and 5.12, we have observed an important difference between the different cases: although all the blue curves are defined by a quadratic polynomial in A and C , the case $B < D$ shows no turning point on the blue curve, while the cases $B \geq D$ do have a turning point on the blue curve. As a matter of fact, if we zoomed in the area around the BT point in Figures 5.1 and 5.2 (see Figure 5.16 in the next section), we will see the turning point for the case $B = 0.054, D = 0.057$ since the blue curve contains the turning point, while the blue curve for the case $B = 0.054, D = 0.087$ does not include the turning point. Summarizing the above results, we have the following theorem.

Theorem 5.3.1 *For system (5.1), there always exists Hopf bifurcation which occurs from the disease equilibrium E_1 , for suitable positive parameter values. The bifurcations may be supercritical or subcritical, and a limit cycle bifurcating from a supercritical (subcritical) Hopf critical point is stable (unstable), which encloses an unstable (a stable) focus point – the equilibrium E_1 .*

5.3.2 Generalized Hopf bifurcation

Now we consider possible generalized Hopf bifurcations which may occur from system (5.1), leading to bifurcation of multiple (two) limit cycles from a Hopf bifurcation point. The condition for generalized Hopf bifurcation is that the first-order focus value vanishes, i.e., $\nu_1 = 0$. In other words, on the Hopf bifurcation curve (the blue curves in Figures 5.1, 5.2, 5.9 and

5.12), such a critical point is identified when the Hopf bifurcation changes from supercritical to subcritical, or vice versa.

Again, we first consider the case $B = D = \frac{57}{1000}$, for which there are two generalized Hopf critical points located on the lower branch of the blue curve (see Figure 5.12): $A_{\text{gH}}^{(1)} = 0.0184$ and $A_{\text{gH}}^{(2)} = 0.9210$, where the subscript ‘gH’ denotes ‘generalized Hopf’. Note that in computation we take the accuracy up to 30 decimal points $A_{\text{gH}}^{(1)} = 0.0184128746264075106899349611404$, $A_{\text{gH}}^{(2)} = 0.921012043225272084984762668632$, but only present the results up to 4 decimal points for brevity. The corresponding critical values of C are given by $C_-(A)$ in Eq. (5.22) as $C_{\text{gH}}^{(1)} = C_-(A_{\text{gH}}^{(1)}) = 0.1199$ and $C_{\text{gH}}^{(2)} = C_-(A_{\text{gH}}^{(2)}) = 1.0456$. Then the first equation of the normal form (5.23) associated with the critical point $(A_{\text{gH}}^{(1)}, C_{\text{gH}}^{(1)})$ is given by

$$\frac{dr}{d\tau} = r [v_0 \mu + v_1 r^2 + v_2 r^4 + o(r^6)], \quad (5.25)$$

where $v_1 = 0$ and $v_2 = -0.1076 \times 10^{-3}$, called the second-order focus value, is obtained by using the Maple program [26]. Note that we now take the unfolding term from perturbing the parameter C as $C = C_-(A) + \mu$. Thus, we can perturb A from $A_{\text{gH}}^{(1)}$ to get $v_1 > 0$ such that $v_1 \ll |v_2|$, and then find $v_0 \mu < 0$ satisfying $|v_0 \mu| \ll v_1$. This gives two limit cycles bifurcating from the critical point $(A_{\text{gH}}^{(1)}, C_{\text{gH}}^{(1)})$. For this case, by perturbing C we have

$$v_0 = \frac{A(1057+1000A)-57C(2+C)}{2000A(1+C)^2}.$$

To obtain $v_1 > 0$, we perturb $A = A_{\text{gH}}^{(1)}$ to $A^* = A_{\text{gH}}^{(1)} + 0.00005 = 0.01846287$, for which $C^* = C_-(A^*) = 0.11969100$ and so $v_0 = 0.11653286$. Now for the Hopf bifurcation associated with the critical values (A^*, C^*) , we obtain $v_1 \approx 0.13257095 \times 10^{-4}$ and $v_2 \approx -0.10838198 \times 10^{-3}$. Further, we choose $\mu = -10^{-6} < 0$, i.e. C is decreased to pass through the critical point (A^*, C^*) , yielding $v_0 \mu \approx -0.11653286 \times 10^{-6}$. Finally, we obtain the normal form for this generalized Hopf bifurcation, up to 5th-order terms, in the form of

$$\frac{dr}{d\tau} = r [0.11653286 \times (-10^{-6}) + 0.13257096 \times 10^{-4} r^2 - 0.10838198 \times 10^{-3} r^4],$$

giving two real positive roots, $r_1 \approx 0.09763824$ and $r_2 \approx 0.33583483$, which approximate the amplitudes of the two limit cycles. Since $v_2 < 0$, the larger limit cycle is stable while the smaller limit cycle is unstable, and the equilibrium solution at this critical point is a stable focus.

In order to show the existence of the two limit cycles predicted above, first note that at the parameter values $B = D = 0.057$, $A = A^*$, $C = C^* - 10^{-6}$, the Jacobin matrix evaluated at the fixed point $E_- = (14.8376281, 0.1542552)$ has eigenvalues $-0.11918442 \times 10^{-6} \pm 0.08096077 i$, confirming that this fixed point is a stable focus. But the convergence speed of nearby trajectories to this stable focus is very very slow. Next, we only need to show that there exists a stable limit cycle around this point since $v_2 < 0$, and expect that the convergence speed is also very slow. Therefore, there exists one unstable limit cycle between the stable focus and the stable limit cycle, as shown in Figure 5.15. It can be seen from this figure that the analytical predictions, $r_1 \approx 0.10$ and $r_2 \approx 0.34$, give very good approximations for the amplitudes of the two simulated limit cycles, see Figure 5.15(b).

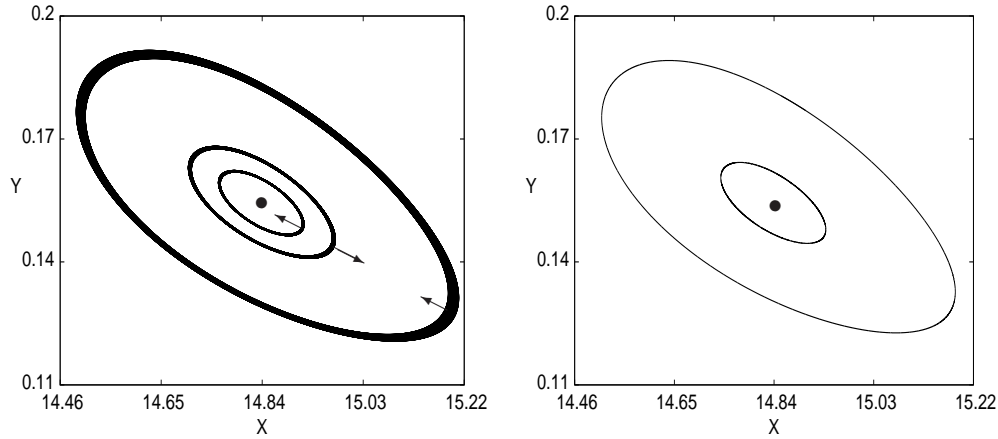


Figure 5.15: Simulation of two limit cycles for system (5.1) when $B = D = 0.057$, $A = 0.01846287$, $C = 0.11969000$: (a) three trajectories with moving directions indicated; and (b) two limit cycles with the inner unstable and outer stable.

Following the above procedure, we can also obtain two limit cycles bifurcating from the other critical point $A_{\text{gH}}^{(2)}$. We give the normal form for this case below without giving details for brevity. Taking $A = A^* = A_{\text{gH}}^{(2)} - 10^{-9}$, $C = C^* = C_-(A^*)$ yields

$$\frac{dr}{d\tau} = r [0.21278281 \times (-10^{-9}) + 0.93716102 \times 10^{-7} r^2 - 0.87730535 \times 10^{-5} r^4],$$

which has two real positive roots, $r_1 \approx 0.05721775$ and $r_2 \approx 0.08607204$, approximating the amplitudes of the two limit cycles bifurcating from this critical point ($A = A^*$, $C = C^*$). Again, since $\nu_2 < 0$, the larger limit cycle is stable and the smaller limit cycle is unstable, and the equilibrium point is a stable focus. For simulation, we should take $A = A^* = 0.9210120422$, $C = C^* - 10^{-9} = 1.0456736673$, which yields the eigenvalues at the equilibrium point $E_- = (2.09166511, 0.88077509)$ as $-0.22645814 \times 10^{-5} \pm 0.62756454i$, and a similar figure to Figure 5.15.

Similarly, we can obtain the five normal forms corresponding to the two critical points for the case of $B = 0.054 < D = 0.057$, one critical point for the case of $B = 0.054 < D = 0.087$, and two critical points for the case of $B = 0.060 > D = 0.057$. We first define the five cases followed by the corresponding five normal forms.

- (a) $B = 0.054 < D = 0.057$ with $A = A_{\text{gH}}^{(1)} = 0.0149805591$
- (b) $B = 0.054 < D = 0.057$ with $A = A_{\text{gH}}^{(2)} = 0.9454739030$
- (c) $B = 0.054 < D = 0.087$ with $A = A_{\text{gH}} = 1.1708464105$
- (d) $B = 0.060 > D = 0.057$ with $A = A_{\text{gH}}^{(1)} = 0.0213860900$

(e) $B = 0.060 > D = 0.057$ with $A = A_{gH}^{(2)} = 0.8963921091$

- (a) $\frac{dr}{dt} = r[0.20278804 \times (-10^{-6}) + 0.89169329 \times 10^{-4} - 0.14900851 \times 10^{-2}]$
 $= 0 \implies r_1 = 0.04866092$ (US), $r_2 = 0.23973712$ (S);
- (b) $\frac{dr}{dt} = r[0.21515679 \times (-10^{-10}) + 0.94588780 \times 10^{-8} - 0.22142107 \times 10^{-6}]$
 $= 0 \implies r_1 = 0.04909881$ (US), $r_2 = 0.20076919$ (S);
- (c) $\frac{dr}{dt} = r[0.21521113 \times 10^{-9} - 0.93765555 \times 10^{-6} + 0.12177368 \times 10^{-3}]$
 $= 0 \implies r_1 = 0.01538841$ (S), $r_2 = 0.08638971$ (US);
- (d) $\frac{dr}{dt} = r[-0.04825749 \times 10^{-9} + 0.15893286 \times 10^{-6} - 0.58166912 \times 10^{-4}]$
 $= 0 \implies r_1 = 0.01865320$ (US), $r_2 = 0.04883049$ (S);
- (e) $\frac{dr}{dt} = r[-0.00236277 \times 10^{-8} + 0.92766615 \times 10^{-7} - 0.16897622 \times 10^{-4}]$
 $= 0 \implies r_1 = 0.01636337$ (US), $r_2 = 0.07226452$ (S),

where US and S denote unstable limit cycle and stable limit cycle, respectively.

Summarizing the above results we have the following result.

Theorem 5.3.2 *For system (5.1), there always exists generalized Hopf bifurcation leading to two limit cycles bifurcating from the disease equilibrium E_1 , for suitable positive parameter values. One of the two limit cycles is stable while the other is unstable.*

This theorem indicates that regardless whether $B < D$ or $B = D$ or $B > D$, the system can always exhibit complex dynamics including different types of bistability or even tristability. More precisely, for Cases (a) and (b) (for which $B < D$), the disease-free equilibrium E_0 is a stable node, the disease equilibrium E_{1-} is a stable focus (another disease equilibrium E_{1+} is a saddle point), and there exist a stable limit cycle, as well as an unstable limit cycle between the stable limit cycle and the stable focus. This indeed shows tristability involving two stable equilibrium solutions and one stable periodic solution. Therefore, the first quadrant of the X - Y plane can be divided into three trapping regions, each corresponding to one of the three stable solutions. Case (c) (again $B < D$) shows a bistable situation, since for this case the disease equilibrium E_{1-} is an unstable focus, and there exist two limit cycles enclosing this unstable focus, with the inner one stable. The disease-free equilibrium E_0 is still a stable node. For Cases (d) and (e) (for which $B > D$) and the two cases when $B = D$, we can see that the disease-free equilibrium E_0 now becomes a saddle point (a degenerate saddle point for $B = D$) and there is only one disease equilibrium E_1 which is a stable focus. There are two limit cycles enclosing the stable focus and the outer one is stable. So this again shows a bistability but it involves one stable equilibrium solution and one stable periodic solution, different from the Hopf bifurcation case.

The above discussion implies that the real situation could be very complex, showing the co-existence of a stable disease-free equilibrium, stable disease equilibria, and even stable oscillating motion, all of which are possible depending upon the initial conditions. Moreover, note that the above seven cases (five cases plus two cases for $B = D$) are obtained for fixed parameter values of B and D . Hence, such phenomena are not uncommon, but quite rich if the parameters B and D are also allowed to be varied.

5.4 Bogdanov-Takens bifurcation

Finally, we consider possible Bogdanov-Takens (BT) bifurcations in system (5.1), characterized by a critical point with a double-zero eigenvalue. First, we have noticed that it is not possible to have a double-zero singularity at $E_0 = (\frac{1}{D}, 0)$ since it has eigenvalues $\xi_1 = -D$ and $\xi_2 = \frac{B}{D} - 1$, implying that it can have at most one zero eigenvalue when $B = D$. Secondly, for the case $B > D$, on the equilibrium solution E_{1-} , $\det(J) > 0$ which cannot have a double-zero eigenvalue. Thirdly, for the case $B = D$, again the equilibrium solution E_{1-} cannot have a double-zero critical point since when $\text{Tr}(J) = 0$, $\det(J) = \frac{A^3(1+C)(A+D)^2}{C^2} > 0$. Thus, the only possibility comes from the case $B < D$ on the equilibrium E_{1-} , which is observed from Figures 5.1 and 5.2. In fact, it can be seen from (5.12) that $\det(J) = 0$ requires $\Delta = 0$, together with (5.15) to solve A and C to obtain the solutions for $B = \frac{27}{500}$, $D = \frac{57}{1000}$ as

$$\text{BT}_1 = (B_1, D_1, A_1, C_1) = \left(\frac{27}{500}, \frac{57}{1000}, \frac{3078507}{206879500}, \frac{61731}{827518} \right) \quad (5.26)$$

which is marked as a circle on Figure 5.1, and for $B = \frac{27}{500}$, $D = \frac{87}{1000}$ as

$$\text{BT}_2 = (B_2, D_2, A_2, C_2) = \left(\frac{27}{500}, \frac{87}{1000}, \frac{118428267}{2237439500}, \frac{219501}{8949758} \right). \quad (5.27)$$

which is marked by a circle on Figure 5.2. For a clear view, the zoomed areas around the two BT bifurcation points in Figures 5.1 and 5.2 are shown in Figures 5.16(a) and (b), respectively. As has been discussed in Section 5.3 that near the BT bifurcation points, the Hopf bifurcation is supercritical when $B = 0.054$, $D = 0.057$; while it is subcritical when $B = 0.054$, $D = 0.087$, which result in stable and unstable limit cycles, respectively. Thus, we will present the results for both cases.

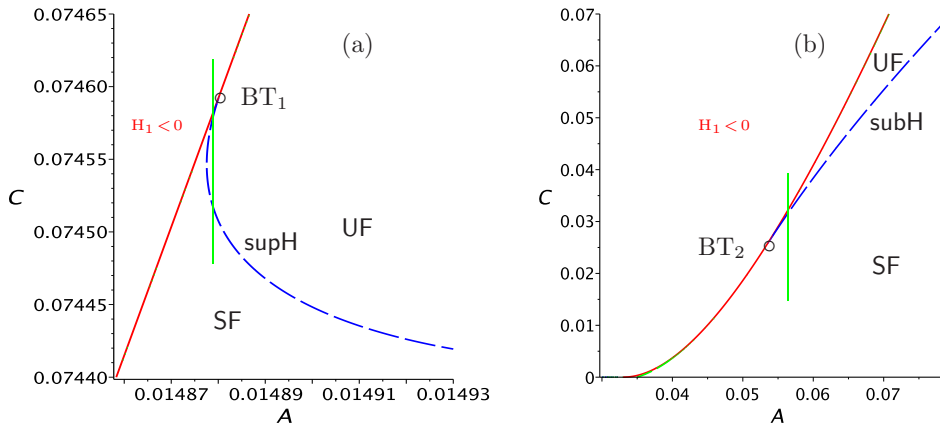


Figure 5.16: The BT bifurcation diagram around the critical points: (a) $(B, D, A, C) = \left(\frac{27}{500}, \frac{57}{1000}, \frac{3078507}{206879500}, \frac{61731}{827518} \right)$, and (b) $(B, D, A, C) = \left(\frac{27}{500}, \frac{87}{1000}, \frac{118428267}{2237439500}, \frac{219501}{8949758} \right)$.

5.4.1 Case $B = 0.054$, $D = 0.057$

We first consider the case $B = 0.054$, $D = 0.057$. We will derive the normal form associated with the BT_1 bifurcation, and then use the normal form to carry out bifurcation analysis. To

achieve this, we introduce the following transformations:

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \frac{943}{57} \\ \frac{57}{1000} \end{pmatrix} + \begin{bmatrix} -\frac{1000}{943} & 0 \\ \frac{57}{943} & 1 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad \begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} \frac{3078507}{206879500} \\ \frac{61731}{827518} \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad (5.28)$$

into (5.1) and expanding the resulting system around the point $(u_1, u_2, \mu_1, \mu_2) = (0, 0, 0, 0)$ up to second order terms yields the system:

$$\begin{aligned} \frac{du_1}{d\tau} &= u_2 + f(u_1, u_2, \mu_1, \mu_2), \\ \frac{du_2}{d\tau} &= f(u_1, u_2, \mu_1, \mu_2) \\ &\equiv \frac{390174737}{1013000000}\mu_1 - \frac{413759}{9500000}\mu_2 - \frac{171196510081}{58491633000}\mu_1\mu_2 + \frac{171196510081}{517272748500}\mu_2^2 \\ &\quad + \left(\frac{315698117}{513084500}\mu_1 - \frac{22326849399}{450404618500}\mu_2\right)u_1 + \left(\frac{619297507081}{58491633000}\mu_1 - \frac{827518}{955259}\mu_2\right)u_2 \\ &\quad - \frac{1624500}{900809237}u_1^2 + \frac{3990000}{900809237}u_1u_2 + \frac{541500}{955259}u_2^2. \end{aligned} \quad (5.29)$$

Next, we apply the near-identity nonlinear transformation (up to second order), given by

$$\begin{aligned} u_1 &= y_1 + \frac{2685896921}{19000000}\bar{\beta}_1 + \left(\frac{2774136507391115169073729}{22529079489117000000000}\bar{\beta}_1 - \frac{5624557927591883}{4137590000000000}\bar{\beta}_2\right)y_1 \\ &\quad + \frac{257312250}{900809237}y_1^2 + \frac{19238822633}{2865777000}y_1y_2 + \frac{6868613670961379723}{72599684000000000}y_2^2 \\ u_2 &= y_2 - \bar{\beta}_1 - \frac{739146260609762232437}{25148272020000000000}\bar{\beta}_1^2 + \frac{6014732664591883}{4137590000000000}\bar{\beta}_1\bar{\beta}_2 \\ &\quad + \left(\frac{17614322633}{2865777000}\bar{\beta}_1 - \frac{57}{1000}\bar{\beta}_2\right)y_1 + \left(\frac{739146260609762232437}{25148272020000000000}\bar{\beta}_1 - \frac{6014732664591883}{4137590000000000}\bar{\beta}_2\right)y_2 \\ &\quad + \frac{1624500}{900809237}y_1^2 + \frac{541500}{955259}y_1y_2 + \frac{17614322633}{2865777000}y_2^2, \end{aligned} \quad (5.30)$$

and the parametrization,

$$\mu_1 = -\frac{55157609919}{171196510081}\bar{\beta}_1 + \frac{55157609919}{171196510081} \quad \mu_2 = -\frac{8836997403671}{342393020162}\bar{\beta}_1 + \frac{975576403671}{342393020162}\bar{\beta}_2,$$

to (5.29) to obtain the normal form:

$$\begin{aligned} \frac{dy_1}{d\tau} &= y_2, \\ \frac{dy_2}{d\tau} &= \bar{\beta}_1 + \bar{\alpha}\bar{\beta}_2 y_1 + \bar{\beta}_2 y_2 - a_1 y_1^2 + a_2 y_1 y_2, \end{aligned} \quad (5.31)$$

where

$$\bar{\alpha} = \frac{57}{1000}, \quad a_1 = \frac{1624500}{900809237}, \quad a_2 = \frac{741000}{900809237}.$$

In order to further simplify system (5.31), we introduce the following scalings:

$$y_1 = m_1 x_1, \quad y_2 = m_2 x_2, \quad \tau_1 = m_3 \tau,$$

into (5.31) to obtain

$$\begin{aligned} \frac{dx_1}{d\tau_1} &= x_2, \\ \frac{dx_2}{d\tau_1} &= \beta_1 + \alpha\beta_2 x_1 + \beta_2 x_2 - x_1^2 + x_1 x_2. \end{aligned} \quad (5.32)$$

Here,

$$\begin{aligned} m_1 &= \frac{a_1}{a_2} = \frac{900809237}{338000}, & m_2 &= \frac{a_1^2}{a_2^3} = \frac{51346126509}{8788000}, & m_3 &= \frac{a_1}{a_2} = \frac{57}{26}, \\ \bar{\beta}_1 &= \frac{a_1^3}{a_2^4} \beta_1 = \frac{2926729211013}{228488000} \beta_1, & \bar{\beta}_2 &= \frac{a_1}{a_2} \beta_2 = \frac{57}{26} \beta_2, & \alpha &= \frac{a_2}{a_1} \bar{\alpha} = \frac{13}{500}. \end{aligned} \quad (5.33)$$

Thus, the relation between the original perturbation parameters (μ_1, μ_2) and the new perturbation parameters (β_1, β_2) is given by

$$\begin{aligned} \mu_1 &= -\frac{161431388159597692837947}{39116348195387528000} \beta_1 + \frac{3143983765383}{4451109262106} \beta_2, \\ \mu_2 &= -\frac{25863498438969955299828723}{78232696390775056000} \beta_1 + \frac{55607855009247}{8902218524212} \beta_2, \end{aligned} \quad (5.34)$$

It should be noted that due to the large values of m_1 and m_2 , very small values of (x_1, x_2) can result in very large values of (y_1, y_2) and so (u_1, u_2) , which are perturbations from the BT critical point (A_T, C_T) . Therefore, we should take small values of x_1 and x_2 when solving system (5.31). Also note in (5.34) that the coefficients of β_1 are large, so we should choose very small values for the perturbation parameter β_1 . Moreover, since in general μ_2 should take negative values (see Figure 5.16(a)), we will show in the following that β_1 must take positive values.

Now, we use the normal form (5.32) to analyze the BT bifurcation. First, we note that in almost all existing articles or books, the unfolding terms (i.e. the terms with the coefficient β_1 or β_2) are usually taken as in a generic form with no direct relation to the original physical system parameters, which may cause difficulty in bifurcation analysis when solving practical problems. Here, we involve perturbation parameters in the nonlinear transformation to obtain the explicit unfolding terms (in terms of β_1 and β_2), which have a direct relation to the original system parameters A and C , and thus facilitate a realistic dynamical study. It should be noted that our system (5.32) is not in the standard normal form for BT bifurcations, given by (e.g. see [13])

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= \beta_1 + \beta_2 x_2 + x_1^2 \pm x_1 x_2. \end{aligned} \quad (5.35)$$

However, we will show in the following that our system (5.32) (or the original system (5.1)) does exhibit interesting dynamics that system (5.35) does, for realistic parameter values, including Hopf bifurcation and homoclinic loops.

The two equilibrium solutions of (5.32) are given by

$$E_{\pm} = (x_{1\pm}, 0), \quad \text{where} \quad x_{1\pm} = \frac{1}{2}[\alpha\beta_2 \pm \sqrt{(\alpha^2\beta_2^2 + 4\beta_1)}]. \quad (5.36)$$

Since we require $\alpha^2\beta_2^2 + 4\beta_1 \geq 0$, we have $\beta_1 > 0$ or $\beta_1 < -\frac{4}{\alpha^2} = -\frac{1000000}{169}$. Thus, for $|\beta_1| \ll 1$, we only consider $\beta_1 \geq 0$. In fact, with (5.5) and (5.34), we obtain

$$H_1 \approx \frac{3063807}{206879500} \mu_1 - \frac{173223}{103439750} \mu_2 = \frac{9632559468266793081}{19558174097693764} \beta_1,$$

and thus the condition $H_1 \geq 0$ yields $\beta_1 \geq 0$. Therefore, in the following analysis we assume $\beta_1 \geq 0$. It is easy to see that when $\beta_1 \geq 0$, $x_{1+} > 0$ and $x_{1-} \leq 0$.

To find the stability of the two equilibrium solutions, we use the Jacobian of (5.32) to obtain the characteristic polynomial $\lambda^2 - \text{Tr} \lambda + \det$, where

$$\text{Tr} = \beta_2 + x_1 \quad \text{and} \quad \det = -\alpha\beta_2 + 2x_1.$$

Defining $\Delta = \text{Tr}^2 - 4 \det$, we have

$$\begin{aligned} \text{Tr}^+ &= \beta_2 + x_{1+} = (1 + \frac{1}{2}\alpha)\beta_2 + \frac{1}{2}\sqrt{\alpha^2\beta_2^2 + 4\beta_1}, \\ \det^+ &= -\alpha\beta_2 + 2x_{1+} = \sqrt{\alpha^2\beta_2^2 + 4\beta_1} > 0, \end{aligned} \quad (5.37)$$

implying that the equilibrium E_+ : $(x_{1+}, 0)$ is either a focus or node, which is stable (unstable) when $\Delta^+ = (\text{Tr}^+)^2 - 4 \det^+ < 0 (> 0)$. Similarly, for the equilibrium E_- : $(x_{1-}, 0)$ we have

$$\begin{aligned} \text{Tr}^- &= \beta_2 + x_{1-} = (1 + \frac{1}{2}\alpha)\beta_2 - \frac{1}{2}\sqrt{\alpha^2\beta_2^2 + 4\beta_1}, \\ \det^- &= -\alpha\beta_2 + 2x_{1-} = -\sqrt{\alpha^2\beta_2^2 + 4\beta_1} < 0, \end{aligned} \quad (5.38)$$

indicating that E_- is always a saddle point. The bifurcation set (only for $\beta_1 \geq 0$) and corresponding phase portraits are shown in Figure 5.17. Note that the Hopf bifurcation near the critical point (denoted by the dashed blue curve in Figure 5.17) is obtained from $\text{Tr}^+ = 0$ as

$$\beta_1 = (1 + \alpha)\beta_2^2 = \frac{513}{500}\beta_2^2. \quad (5.39)$$

There is another curve in Figure 5.17, shown in red, which denotes the bifurcation of homoclinic loop (see [13]).

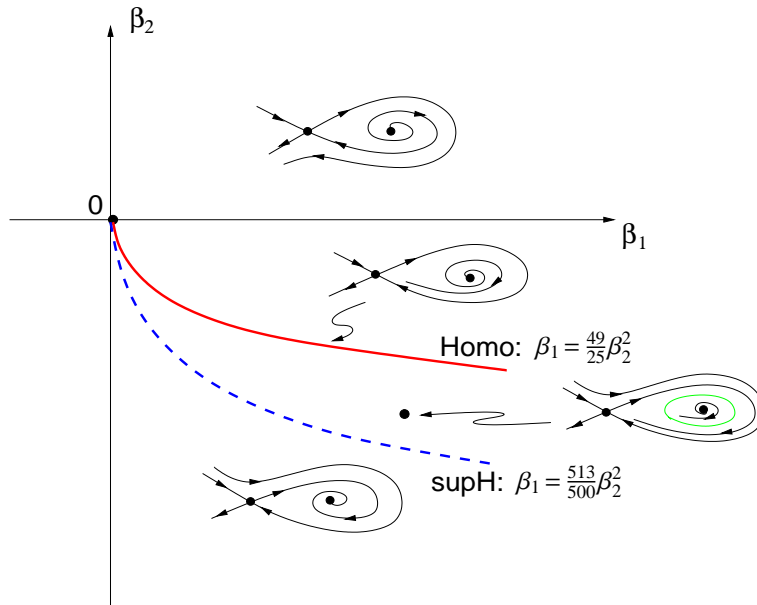


Figure 5.17: Bifurcation set and phase portraits of system (5.32).

Before we derive the equation for the bifurcation of the homoclinic loop, we consider the Hopf bifurcation which occurs from the dashed blue curve. The Hopf critical point on this curve can be defined as

$$\beta_{2H} = -10 \sqrt{\frac{5\beta_1}{513}},$$

and then introducing the transformation: $x_1 = \tilde{x}_1$, $x_2 = \omega_c \tilde{x}_2$ into (5.32) results in the system:

$$\begin{aligned}\frac{d\tilde{x}_1}{d\tau_1} &= \omega_c \tilde{x}_2 \equiv \tilde{f}(\tilde{x}_1, \tilde{x}_2), \\ \frac{d\tilde{x}_2}{d\tau_1} &= -\omega_c \tilde{x}_1 - \frac{1}{\omega_c} \tilde{x}_1^2 + \tilde{x}_1 \tilde{x}_2 \equiv \tilde{g}(\tilde{x}_1, \tilde{x}_2),\end{aligned}$$

where

$$\omega_c = \sqrt{\frac{1013}{8550}} \sqrt{285 \beta_1}.$$

Thus, the first focus value ν_1 is given by

$$\nu_1 = -\frac{1}{16\omega_c^2} (-\tilde{g}_{\tilde{x}_1 \tilde{x}_1} \tilde{g}_{\tilde{x}_1 \tilde{x}_2}) = -\frac{1}{16\omega_c} \times \frac{2}{\omega_c} = -\frac{1}{8\omega_c^2} < 0,$$

indicating that the Hopf bifurcation is supercritical, and bifurcating limit cycles are stable, as shown in Figure 5.17 (see the ellipse in green). The Hopf bifurcation near the BT critical point is not surprising since the original system does have Hopf bifurcations which occur from the blue curve, as shown in Figure 5.16. In fact, as discussed in Section 5.3, we can similarly use the original system to show that the Hopf bifurcations from the blue curve (see Figure 5.16) are indeed supercritical, which agrees with the conclusion obtained above, and so the bifurcating limit cycles are stable.

Next, we consider homoclinic loops which may bifurcate near the BT critical point. Here, we apply the technique of rescaling, as used in [13] to find the equation for the homoclinic curve. Set

$$x_1 = \varepsilon^2 w_1, \quad x_2 = \varepsilon^3 w_2, \quad \beta_1 = \varepsilon^4 \nu_1, \quad \beta_2 = \varepsilon^2 \nu_2, \quad \alpha = \varepsilon \tilde{\alpha}, \quad (0 \leq \varepsilon \ll 1), \quad (5.40)$$

and rescale time $t = \varepsilon \tau_1$, so that (5.32) can be rewritten (up to ε order) as

$$\begin{aligned}\frac{dw_1}{dt} &= w_2, \\ \frac{dw_2}{dt} &= \nu_1 + \varepsilon \tilde{\alpha} \nu_2 w_1 + \varepsilon \nu_2 w_2 + \varepsilon w_1 w_2 - w_1^2.\end{aligned} \quad (5.41)$$

Now, letting $\varepsilon = 0$ in (5.41) yields an integrable Hamiltonian system:

$$\begin{aligned}\frac{dw_1}{dt} &= w_2, \\ \frac{dw_2}{dt} &= \nu_1 - w_1^2.\end{aligned} \quad (5.42)$$

with Hamiltonian

$$H(w_1, w_2) = -\nu_1 w_1 + \frac{1}{3} w_1^3 + \frac{1}{2} w_2^2. \quad (5.43)$$

Taking $\nu_1 = 1$, which corresponds to $\beta_1 \geq 0$, we have two fixed points: $(w_1, w_2) = (\pm 1, 0)$, with $(1, 0)$ being a center and $(-1, 0)$ a saddle point, as shown in Figure 5.18.

The solution on the saddle loop Γ based at the point $(w_1, w_2) = (2, 0)$ is given by

$$(w_1(t), w_2(t)) = \left(3 \operatorname{sech}^2\left(\frac{t}{\sqrt{2}}\right) - 1, 3 \sqrt{2} \operatorname{sech}^2\left(\frac{t}{\sqrt{2}}\right) \left(\tanh\left(\frac{t}{\sqrt{2}}\right)\right) \right). \quad (5.44)$$

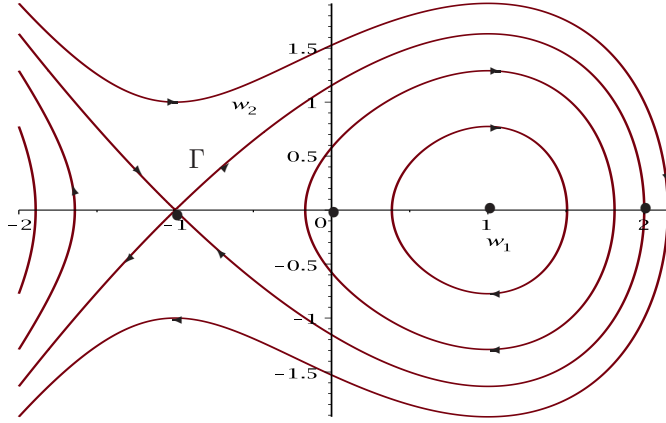


Figure 5.18: The phase portrait of (5.42) with $\nu_1 = 1$, showing a homoclinic loop Γ .

Thus, the first-order Melnikov function $M(t_0)$ on the vector field $\varepsilon(\tilde{\alpha}\nu_2 w_1 + \nu_2 w_2 + w_1 w_2) \frac{\partial}{\partial w_2}$ is independent of time, and can be calculated as

$$\begin{aligned} M(\nu_2) &= \int_{-\infty}^{\infty} w_2(t) [\tilde{\alpha}\nu_2 w_1(t) + \nu_2 w_2(t) + w_1(t)w_2(t)] dt \\ &= \frac{1}{\sqrt{2}} \left[\nu_2 \int_{-\infty}^{\infty} 18 \operatorname{sech}^4 t' \tanh^2 t' dt' \right. \\ &\quad \left. - \int_{-\infty}^{\infty} (3 \operatorname{sech}^2 \tau - 1) 18 \operatorname{sech}^4 t' \tanh^2 t' dt' \right], \end{aligned}$$

where $t' = t/\sqrt{2}$. Note that the first term $\tilde{\alpha}\nu_2 w_1(t)w_2(t)$ yields zero after integration due to the Hamiltonian being symmetric with respect to the w_1 axis, and the negative sign for the integration of the third term $w_1(t)w_2^2(t)$ comes from the definition of the homoclinic loop Γ , which takes t' from $+\infty$ to 0 along the positive w_1 direction while from 0 to $-\infty$ along the negative w_1 direction. Then, solving $M \equiv 0$ for the saddle connection yields

$$\begin{aligned} \nu_2 &\approx \frac{\int_{-\infty}^{\infty} (3 \operatorname{sech}^2 t' - 1) \operatorname{sech}^4 t' \tanh^2 t' dt'}{\alpha \int_{-\infty}^{\infty} \operatorname{sech}^4 t' \tanh^2 t' dt'} \\ &= \frac{\int_{-\infty}^{\infty} \operatorname{sech}^2 t' (2 \tanh^2 t' - 5 \tanh^4 t' + 3 \tanh^6 t') dt'}{\int_{-\infty}^{\infty} \operatorname{sech}^2 t' (\tanh^2 t' - \tanh^4 t') dt'} \\ &= \frac{5}{7}, \end{aligned}$$

where the formula:

$$\int_{-\infty}^{\infty} \operatorname{sech}^2 t' \tanh^k t' dt' = \frac{\tanh^{k+1}(t')}{k+1} \Big|_{-\infty}^{\infty} = \frac{3}{k+1}$$

has been used. Finally, noticing $\nu_1 = 1$, and $\beta_1 = \varepsilon^4$, $\beta_2 = \varepsilon^2 \nu_2$, we obtain the approximate bifurcation curve for the homoclinic loop as

$$\text{Homo : } \beta_1 = \frac{49}{25} \beta_2^2, \quad \beta_2 \leq 0. \quad (5.45)$$

The true bifurcation curve is tangent to the semi-parabola at $\beta_1 = \beta_2 = 0$. Combining this with equation (5.39) for Hopf bifurcation, we indeed see that a second bifurcation curve, denoted as ‘Homo’, is located above the Hopf bifurcation curve and tangent to it (and to $\beta_1 = 0$) at $(\beta_1, \beta_2) = (0, 0)$, and the phase portrait on this bifurcation set has a saddle loop, as shown in Figure 5.17. The sign taken by the Melnikov function M for $\beta_1 < \frac{49}{25}\beta_2^2$ (or $> \frac{49}{25}\beta_2^2$, respectively) gives the relative position of the stable and unstable manifolds (separators of the saddle). Moreover, note that the trace of the “saddle quantity”, given by (5.38),

$$\text{Tr}_{\text{Homo}}^- = (1 + \frac{1}{2}\alpha)\beta_2 - \frac{1}{2}\sqrt{\alpha^2 + \frac{196}{25}|\beta_2|} \quad (\beta_2 < 0),$$

is negative on the ‘Homo’ curve (5.45). Hence, the homoclinic orbit is stable (an ω -limit set) attracting the nearby points. Further, it can be shown (see [13]) that in the region between the Hopf bifurcation curve ‘supH’ and the Homoclinic bifurcation curve ‘Homo’ (see Figure 5.17) the system has a unique attracting limit cycle for each pair of parameter values (β_1, β_2) .

To demonstrate the bifurcation phenomena discussed above, we show simulations using the original system (5.1), rather than the normal form equation (5.32), which gives a more realistic observation. We take seven sets of perturbations on the parameters A and C near the BT_1 critical point (see Figure 5.16(a)) as $A = A_1 + \mu_1$, $C = C_1 + \mu_2$, where A_1 and C_1 are given in (5.26). These seven sets of perturbations denote seven points in the bifurcation diagram (see Figure 5.16) on a same vertical line (see the green line in Figure 5.16(a)) with the same coordinate $A = A_1 - 0.000001$, and different coordinates $C = C_1 + \mu_2$ with μ_2 given from top to the bottom as follows:

$$\begin{aligned} \mu_2 = & -0.0000094, \quad -0.000098, \quad -0.0000106, \quad -0.00003, \\ & -0.0000414239, \quad -0.0000875, \quad -0.0001. \end{aligned}$$

It is noted that the equilibrium E_{1-} is a stable focus at the top and the bottom points, but is an unstable focus at the other five points. Here, we have found an interesting phenomenon that since the Hopf bifurcation curve has a turning point and all nearby points can lead to stable limit cycles in the region where the equilibrium E_{1-} is an unstable focus, there exist two homoclinic loops when one goes through the five points along the vertical line starting with the second point from the top. However, the above normal form theory for the BT bifurcation and the result given in Figure 5.17 only show one homoclinic loop. This is not surprising since the normal form for the BT_1 bifurcation is only applicable for the study of dynamics around the BT_1 point and thus it only predicts the top homoclinic loop. Due to the perturbations being very small, the convergence of the simulating trajectories is very slow. Moreover, the direction of the trajectories near the saddle point is hard to distinguish. Therefore, in order to give a clear view, we, based on the simulating phase portraits which have been rotated by a angle of $\frac{\pi}{55}$, present seven schematic diagrams with exaggerated convergence speed and the part near the saddle point. Since the simulations for the top and bottom points are similar, we will only present one figure for these two points (see Figure 5.19(a)). Of course, they are different quantitatively and the simulation for the bottom point is much clearer than that of the top one. The Figures 5.19(b) to 5.19(f) correspond to the other five points from top to the bottom. The relation between the original coordinates (X, Y) and the new coordinates (\bar{X}, \bar{Y}) shown in Figure 5.19 is given by

$$\bar{X} = \cos\left(\frac{\pi}{55}\right)X - \sin\left(\frac{\pi}{55}\right)Y, \quad \bar{Y} = \sin\left(\frac{\pi}{55}\right)X + \cos\left(\frac{\pi}{55}\right)Y.$$

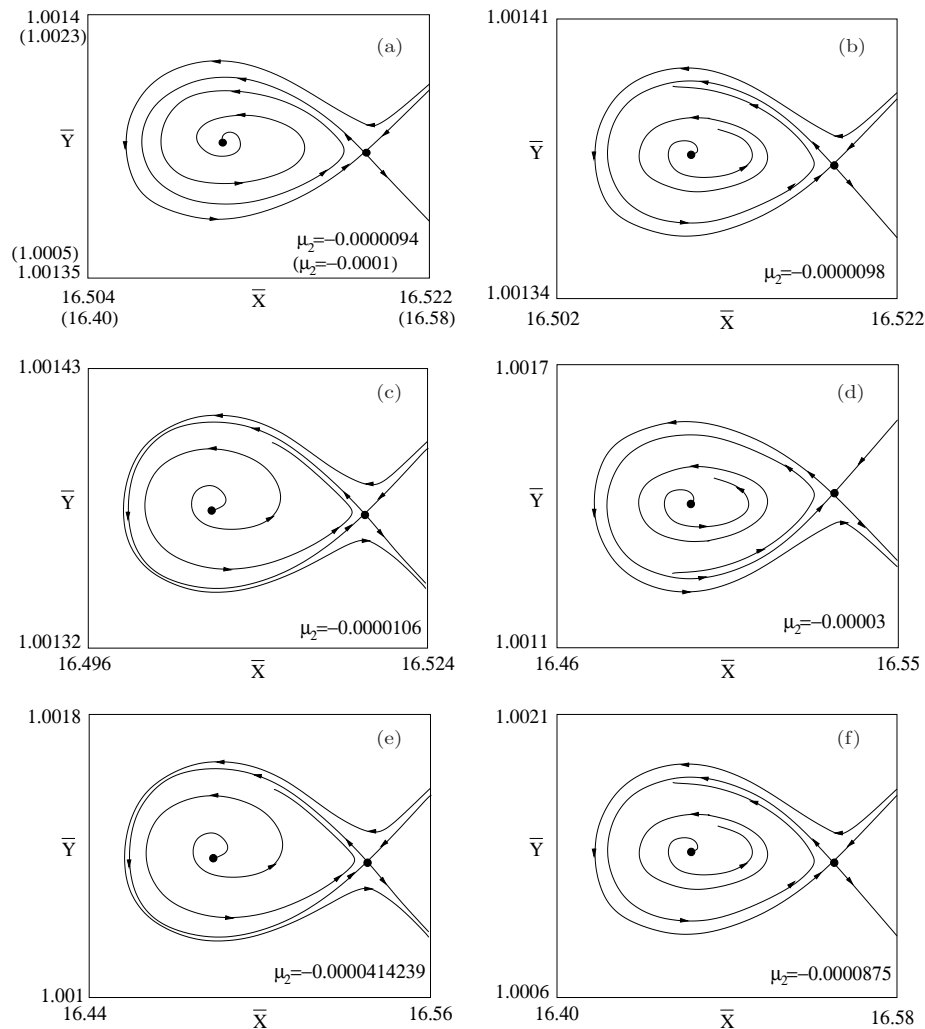


Figure 5.19: Simulations of system (5.1) when $B = 0.054$, $D = 0.057$, $A = 0.01487968$ for (a) $C = 0.07458837$ or $C = 0.07449777$, showing stable focus E_{1-} , (b) $C = 0.07458797$, showing unstable focus E_{1-} and a stable limit cycle, (c) $C = 0.07458717$, showing unstable focus E_{1-} and a stable homoclinic loop, (d) $C = 0.07456777$, showing unstable focus E_{1-} and a stable limit cycle, (e) $C = 0.0745563461$, showing unstable focus E_{1-} and a stable homoclinic loop, (f) $C = 0.07449777$, showing unstable focus E_{1-} and a stable limit cycle.

In the next example for $B = 0.054$, $D = 0.087$, we will see true simulating phase portraits, which clearly show the Hopf bifurcation and homoclinic bifurcation.

5.4.2 Case $B = 0.054$, $D = 0.087$

Now we turn to study the case $B = 0.054$, $D = 0.087$. As we have discussed, a particular difference between this case and previous case is that now the Hopf bifurcation near the BT_2 critical point is subcritical, and thus the bifurcating limit cycles are unstable. This difference can cause dramatically different meanings in the biological explanation of this phenomenon.

Since the solution procedure is similar to the previous case, we will skip some detailed

steps and only present the main results in the following. Using a series of transformations, similar to (5.28), (5.30) and (5.33), we obtain the following normal form:

$$\begin{aligned}\frac{dx_1}{d\tau_1} &= x_2, \\ \frac{dx_2}{d\tau_1} &= \beta_1 + \alpha\beta_2 x_1 + \beta_2 x_2 - x_1^2 - x_1 x_2,\end{aligned}\tag{5.46}$$

where $\alpha = \frac{4717}{5500}$, and the transformation for the parameters:

$$\begin{aligned}\mu_1 &= -\frac{65443353700213087530310106927}{79307833305870371766996170888000}\beta_1 + \frac{3000793340668563}{188911529837823194}\beta_2, \\ \mu_2 &= -\frac{7544255129488549703116068201513}{158615666611740743533992341776000}\beta_1 + \frac{12645552328645797}{377823059675646388}\beta_2.\end{aligned}\tag{5.47}$$

The solution formulae are the same as that given in (5.36). Again, we can similarly argue that $\beta_1 \geq 0$, and as a matter of fact, $H_1 \geq 0$ implies $\beta_1 \geq 0$, for which $x_{1+} > 0$ and $x_{1-} \leq 0$. The stability of these two equilibrium solutions are determined by

$$\begin{aligned}\text{Tr}^+ &= \beta_2 - x_{1+} = (1 - \frac{1}{2}\alpha)\beta_2 - \frac{1}{2}\sqrt{\alpha^2\beta_2^2 + 4\beta_1}, \\ \text{det}^+ &= -\alpha\beta_2 + 2x_{1+} = \sqrt{\alpha^2\beta_2^2 + 4\beta_1} > 0,\end{aligned}\tag{5.48}$$

and

$$\begin{aligned}\text{Tr}^- &= \beta_2 - x_{1-} = (1 - \frac{1}{2}\alpha)\beta_2 + \frac{1}{2}\sqrt{\alpha^2\beta_2^2 + 4\beta_1}, \\ \text{det}^- &= -\alpha\beta_2 + 2x_{1-} = -\sqrt{\alpha^2\beta_2^2 + 4\beta_1} < 0.\end{aligned}\tag{5.49}$$

These results indicate that $(x_{1-}, 0)$ is a saddle point, while $(x_{1+}, 0)$ is either a focus or node. The Hopf bifurcation near the BT_2 critical point is determined from $\text{Tr}^+ = 0$ as

$$\beta_1 = (1 - \alpha)\beta_2^2 = \frac{783}{5500}\beta_2^2,\tag{5.50}$$

and the bifurcation is subcritical, since the first focus value can be obtained as $v_1 = \frac{1}{\omega_c^2} > 0$. Similarly we can obtain the homoclinic bifurcation which occurs from the curve:

$$\text{Homo} : \quad \beta_1 = \frac{49}{25}\beta_2^2, \quad \beta_2 \geq 0.\tag{5.51}$$

The bifurcation set and corresponding phase portraits are depicted in Figure 5.20, which is quite different from the case $B = 0.054$, $D = 0.057$ (see Figure 5.17). Simulations based on the original system (5.1) for this case are shown in Figure 5.21, where the perturbation (μ_1, μ_2) , on the parameters A and C , take the following values:

$$(0.004, 0.0085), (0.004, 0.00807), (0.004, 0.0073813), (0.004, 0.007),$$

which represent four points on the same vertical green line in the bifurcation diagram, shown in Figure 5.16(b).

It is seen from Figure 5.21(a) that the phase portrait for the first perturbation, corresponding to a point above the Hopf bifurcation curve (see the blue curve in Figure 5.20), shows a unstable

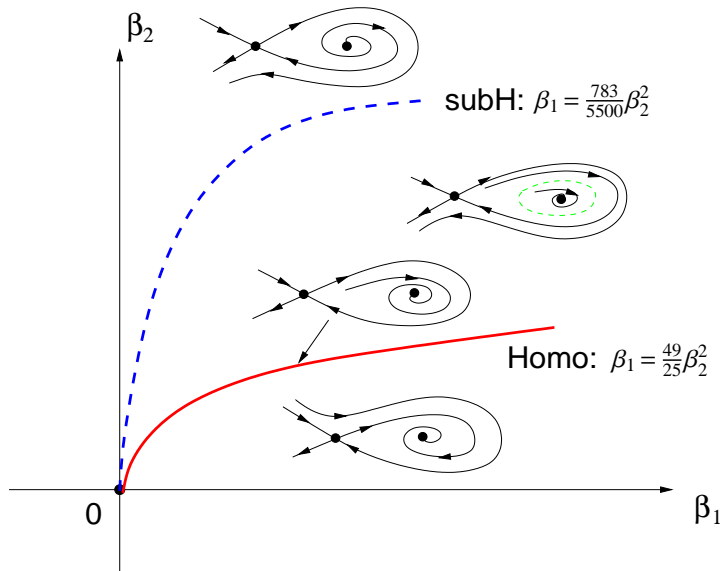


Figure 5.20: Bifurcation set and phase portraits of system (5.46).

focus E_{1-} and there exists one trajectory starting from the saddle point E_{1+} and converging to this focus as $\tau \rightarrow -\infty$. Figure 5.21(b) shows the phase portrait for the second perturbation, corresponding to a point between the Hopf bifurcation curve (the blue curve in Figure 5.20) and the homoclinic bifurcation curve (the red curve in Figure 5.20), shows an unstable limit cycle (see the green curve in Figure 5.20) and trajectories starting near this limit cycle either converge to the stable focus E_{1-} or to the stable node E_0 (which is not shown in Figures 5.20 and 5.21) as $\tau \rightarrow +\infty$. Figure 5.21(c) shows a homoclinic loop under the third perturbation, corresponding to a point on the homoclinic bifurcation curve, which encloses the stable focus E_{1-} , and all trajectories inside this homoclinic loop coverage to the focus as $\tau \rightarrow +\infty$. In fact, it can be shown that the saddle quantity, given in (5.49),

$$\text{Tr}_{\text{Homo}}^- = \left[1 - \frac{1}{2}\alpha + \frac{1}{2} \sqrt{\alpha^2 + \frac{196}{25}} \right],$$

is positive for $\beta_2 > 0$, implying that the homoclinic loop is unstable. Finally, Figure 5.21(d) shows a phase portrait for the fourth perturbation, corresponding to a point below the homoclinic bifurcation curve, which encloses the stable focus E_{1-} . It is seen from Figure 5.21 that the saddle connection before and after the homoclinic loop (Figure 5.21(c)) change the way to connect the focus or the limit cycle. Note that unlike the bifurcation shown in Figure 5.16(a) where there are two homoclinic loops which occur from the green line, here there is only one homoclinic loop since no more Hopf bifurcation happens when the parameter C is decreased to cross the Hopf critical line along the green line (see Figure 5.16(b)).

Summarizing the results obtained in this section we have the following theorem.

Theorem 5.4.1 *For system (5.1), when $B < D$ and $H_1 > 0$, there always exists Bogdanov-Takens bifurcation, which occurs from the precritical disease bifurcation solution, leading to homoclinic bifurcation near a Hopf bifurcation, with homoclinic loop being either stable or unstable.*

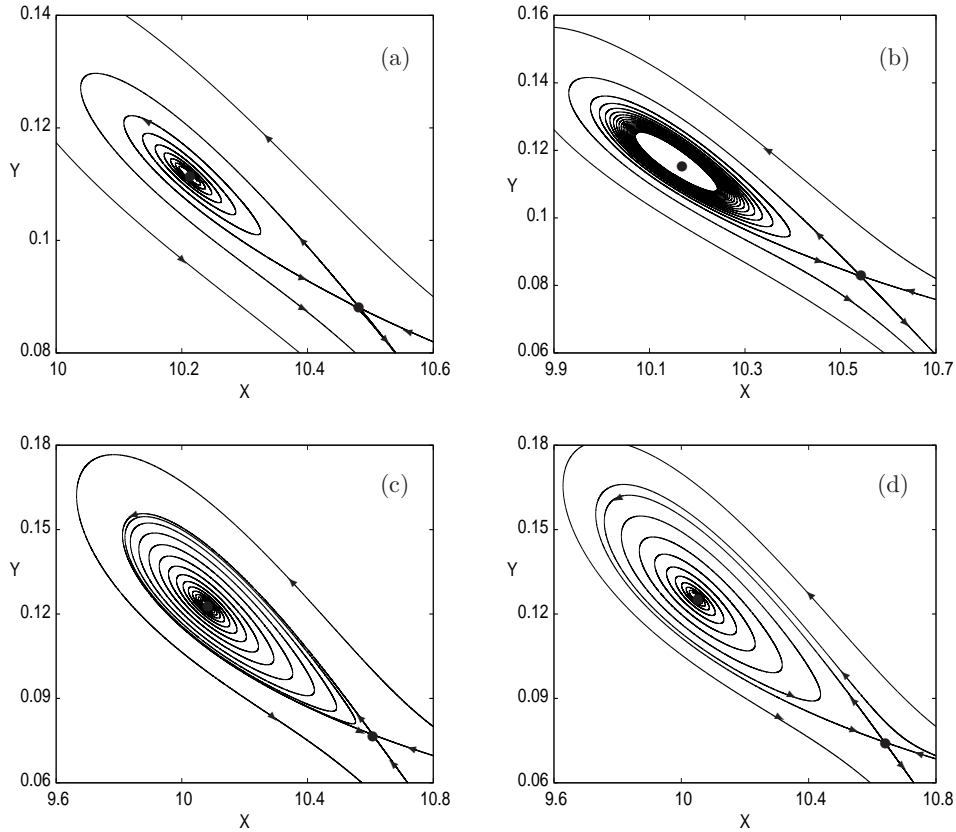


Figure 5.21: Simulations of system (5.1) when $B = 0.054$, $D = 0.087$ for (a) $A = 0.0569302656$, $C = 0.03330259146$, showing an unstable focus E_{1-} with one trajectory divergent to the saddle point E_{1+} , (b) $A = 0.0569302656$, $C = 0.0325959146$, showing stable focus E_{1-} enclosed by an unstable limit cycle, (c) $A = 0.0569302656$, $C = 0.0319072146$, showing a homoclinic loop enclosing a stable focus, and (d) $A = 0.0569302656$, $C = 0.0315259146$, showing convergence of the trajectory starting from the saddle point E_{1+} to the stable focus E_{1-} .

5.4.3 A new mechanism for generating blips

A detailed study for a 4-dimensional system has been given in [27, 28], shows a mechanism for generating the blips phenomenon, and four conditions are proposed in a hypothesis, which guarantee the existence of blips. In [27, 28], blips are also shown to exist in two 3-dimensional models as well as in the 2-dimensional model (5.1). An important condition for the existence of blips is Hopf bifurcation, which is the source of oscillation. Very recently, another mechanism has been identified in [29], which is also related to Hopf bifurcation. These two mechanisms have a common property that both of them generate oscillations with large changes in both amplitude and frequency, and they both appear on the post-critical disease bifurcation solution. It has also been noted that these two mechanisms have a fundamental difference: the former guarantees blips to occur near a transcritical bifurcation point; while the later yields blips far away from a transcritical bifurcation point, which are not guaranteed. The second mechanism needs further investigation.

In order to discuss a new mechanism of generating blips, in the following we list Hypothesis 1 from [27, 28], and propose a second Hypothesis based on the results obtained in [29].

Hypothesis 1 [27, 28] The following four conditions are needed for an in-host infection model to generate viral blips:

- (i) there exist at least two equilibrium solutions;
- (ii) there exists a transcritical bifurcation at an intersection of the two equilibrium solutions;
- (iii) there is a Hopf bifurcation which occurs from one of the equilibrium solutions; and
- (iv) large oscillations (or, more generally, global, persistent motions) can occur near the transcritical critical point.

Hypothesis 2 [29] The following four conditions are needed for an in-host infection model to generate viral blips: the conditions (i), (ii) and (iii) are the same as that given in Hypothesis 1; and

- (iv) large oscillations (or, more generally, global, persistent motions) can occur far away from the transcritical and Hopf critical points.

We use the bifurcation diagrams shown in Figures 5.22(a) and 5.22(b) (which are Figures 3.3(a) and 3.3(b) in [28]) to illustrate Hypothesis 1, and the bifurcation diagram in Figure 5.22(c) (which is Figure 3.1(a) in [29]) to explain Hypothesis 2, where R and A are state variables, B and α are parameters. E_0 and E_1 denote the disease-free and disease equilibrium solutions. The green lines indicate where the blip-like oscillations occur. It is clear from Figures 5.22(a) and 5.22(b) that the blips appear near the transcritical point, and may or may not appear near the Hopf critical point, where both E_0 and E_1 are unstable, illustrating condition (iv) in Hypothesis 1. Figure 5.22(c) (where the second Hopf critical point “Hopf₂” is outside the figure) shows that the blips occur far away from the transcritical and Hopf bifurcation points.

Through the study given in this section on the BT bifurcation, we have found a third mechanism for generating blips, due to the BT bifurcation, explained as follows. First of all, note that the trajectory starting from a point on the homoclinic loop will reach the saddle point either as $\tau \rightarrow +\infty$ or $\tau \rightarrow -\infty$. Therefore, it can be seen from Figure 5.17 that near the homoclinic bifurcation curve, for certain parameter values, the bifurcating stable limit cycles can be large close to the saddle separators and thus such a stable limit cycle will move extremely slowly near the saddle point but will move fast when it is away from the saddle point – giving rise to the blips phenomenon. A schematic bifurcation diagram for the case, which is depicted in Figure 5.19 when $B = 0.054$, $D = 0.057$, $A = 0.01487968$, is shown in Figure 5.22(d). Also note from Figures 5.20 and 5.21 that when the limit cycle inside the saddle separators is unstable, the trajectories starting near the unstable limit cycle may converge to the stable focus E_{1-} , or to the stable node E_0 but will take very long time since the solution will go through a route close to the saddle point though not generating blips in this case.

The big difference between the first two mechanisms and the new mechanism is that the first two mechanisms result in very large oscillations in both amplitude and frequency, while

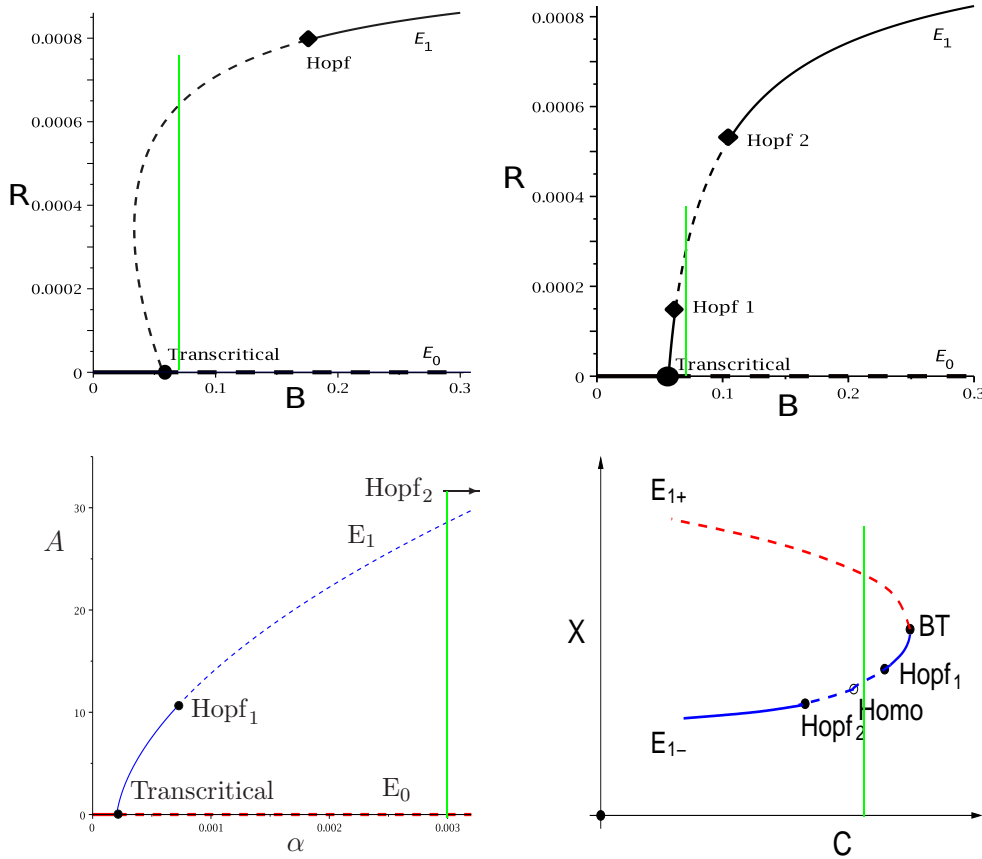


Figure 5.22: Bifurcation diagrams illustrating Hypotheses: (a) and (b) for Hypothesis 1, (c) for Hypothesis 2 and (d) Hypothesis 3.

the new mechanism only causes significant changes in frequency, but very little variation in the amplitude. The biological implication of the new mechanism is interesting and may explain some real situations, namely, in some situations a patient may not feel obvious changes nor will measurable changes in disease progression be apparent, but nonetheless the patient may be experiencing recurrent disease without any significant observation. In other situations, neither the infected individual nor the clinician may be able to detect whether the infection has been cured, since complete recovery may take an extremely long time. In both cases, the patient is in an uncertain situation. To describe these scenarios, we have the following hypothesis.

Hypothesis 3 The following four conditions are needed for an in-host infection model to generate viral blips or to take an extremely long time to recover (converge to the disease-free equilibrium): conditions (i), (ii) and (iii) are the same as that given in Hypothesis 1; and

- (iv) there exists Bogdanov-Takens bifurcation, leading to homoclinic loops near a Hopf bifurcation, which may yield blips with very small changes in amplitude, or extremely slow convergence to the disease-free equilibrium.

5.5 Conclusion and discussion

In this paper, we have given a detailed dynamical study of a 2-dimensional disease model, which can be used not only for in-host disease modelling, but also for epidemiologic modelling. We have shown that when the reproduction number, $R_0 = \frac{B}{D}$, is varied near $R_0 = 1$, the system exhibits rich dynamical behaviors, including equilibrium solutions which exchange their stability at the transcritical point $R_0 = 1$. Both Hopf and generalized Hopf bifurcations can occur regardless whether $R_0 < 1$ or $R_0 \geq 1$, which lead to bistability or even tristability. In particular, our study has indicated that when $R_0 < 1$, the system can have Bogdanov-Takens bifurcation leading to more complex dynamical behavior such as homoclinic orbit bifurcation. This special bifurcation may provide a new scenario/mechanism for generating recurrence or the viral blips phenomenon, summarized in Hypothesis 3.

Hypothesis 3 is completely different from Hypotheses 1 and 2, and may provide an explanation for interesting clinical phenomena. In many disease models, the concept of R_0 is straightforward, i.e. if $R_0 < 1$, the disease cannot invade or persist, and the disease only exists for $R_0 > 1$. In reality, disease dynamics are more complex, and our model indeed reflects this complexity. Hypothesis 3 allows for the possibility that even if control or therapy reduces R_0 below one, a disease may persist indefinitely with low level oscillations, or may die out, but with an extremely slow time course of decay. The possibility of disease persistence when $R_0 < 1$ is a feature of backward bifurcation [9, 30, 5, 4], an issue which we are investigating for this model and related disease models as well [30].

Mathematically, the most interesting dynamical behavior of our model is the Bogdanov-Takens bifurcation leading to homoclinic loops, which in turn provides a new mechanism for explaining a very different blips phenomenon. In particular, this phenomenon does not have obvious changes in the amplitude of the oscillating motion. This can only happen when $B < D$ (i.e. $R_0 < 1$). However, this condition, $B < D$, is not enough, the additional condition $H_1 \geq 0$, which guarantees the existence of disease equilibrium, E_1 , must also be satisfied. Intuitively, if $B < D$, then the epidemic cannot get started because near the disease-free equilibrium, E_0 , the behavior of the model is similar to that studied in [17], and thus no oscillation can occur with $R_0 < 1$. However, $H_1 \geq 0$, as mentioned in Remark 5.2.3, implies that the contact rate A exceeds its threshold such that the infected cells, denoted by Y , are sufficiently infectious such that the epidemic can sustain itself once started even if $B < D$. Therefore, this leads, after getting over an initial threshold, to potential bistable equilibrium solutions and even more complex dynamical behavior.

The ideas and methodologies presented in this paper can be used to analyze other types of in-host disease models as well as epidemiologic models. We hope that they can also be generalized to study functional differential systems (e.g. with time delays), or even other physical or engineering systems which exhibit similar “blips-like” phenomenon.

5.6 References

- [1] H. K. Alexander and L. M. Wahl. Self-tolerance and autoimmunity in a regulatory T cell model. *Bulletin of mathematical biology*, 73(1):33–71, 2011.
- [2] H. K. Altes, R. M. Ribeiro, and R. J. de Boer. The race between initial T-helper expan-

- sion and virus growth upon HIV infection influences polyclonality of the response and viral setpoint. *Proceedings of the Royal Society B: Biological Sciences*, 270(1522):1349–1358, 2003.
- [3] R. M. Anderson and R. M. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, 1992.
- [4] J. Arino, C. C. McCluskey, and P. van den Driessche. Global results for an epidemic model with vaccination that exhibits backward bifurcation. *SIAM Journal on Applied Mathematics*, 64(1):260–276, 2003.
- [5] K. W. Blayneh, A. B. Gumel, S. Lenhart, and T. Clayton. Backward bifurcation and optimal control in transmission dynamics of west nile virus. *Bulletin of Mathematical Biology*, 72(4):1006–1028, 2010.
- [6] C. J. Briggs and H. C. J. Godfray. The dynamics of insect-pathogen interactions in stage-structured populations. *American Naturalist*, pages 855–887, 1995.
- [7] C. Confavreux, S. Vukusic, T. Moreau, and P. Adeleine. Relapses and progression of disability in multiple sclerosis. *New England Journal of Medicine*, 343(20):1430–1438, 2000. PMID: 11078767.
- [8] W. R. Derrick and P. Driessche. A disease transmission model in a nonconstant population. *Journal of Mathematical Biology*, 31(5):495–512, 1993.
- [9] J. Dushoff, W. Huang, and C. Castillo-Chavez. Backwards bifurcations and catastrophe in simple models of fatal diseases. *Journal of mathematical biology*, 36(3):227–248, 1998.
- [10] E. M. Farber, R. H. Mullen, A. H. Jacobs, and L. Nall. Infantile psoriasis: a follow-up study. *Pediatric Dermatology*, 3:237–243, 1986.
- [11] D. M. Fergusson, L. J. Horwood, and F. T. Shannon. Early solid feeding and recurrent childhood eczema: A 10-year longitudinal study. *Pediatrics*, 86(4):541–546, 1990.
- [12] H. J. Girschick, C. Zimmer, G. Klaus, K. Darge, A. Dick, and H. Morbach. Chronic recurrent multifocal osteomyelitis: what is it and how should it be treated? *Nature Clinical Practice Rheumatology*, 3, 2007.
- [13] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Applied Mathematical Sciences. Springer-Verlag, New York, 4th edition, 1993.
- [14] H. W. Hethcote, M. A. Lewis, and P. Van Den Driessche. An epidemiological model with a delay and a nonlinear incidence rate. *Journal of mathematical biology*, 27(1):49–64, 1989.
- [15] H. W. Hethcote and P. Van den Driessche. Some epidemiological models with nonlinear incidence. *Journal of Mathematical Biology*, 29(3):271–287, 1991.

- [16] R. S. Iyer, M. M. Thapa, and F. S. Chew. Chronic recurrent multifocal osteomyelitis: review. *American Journal of Roentgenology*, 196(6 Suppl):S87S91, 2011.
- [17] A. Korobeinikov and P. K. Maini. Non-linear incidence and stability of infectious disease models. *Mathematical Medicine and Biology*, 22(2):113–128, 2005.
- [18] D. C. Krakauer and N. L. Komarova. Levels of selection in positive-strand virus dynamics. *Journal of Evolutionary Biology*, 16(1):64–73, 2003.
- [19] M. Y. Li and J. S. Muldowney. Global stability for the SEIR model in epidemiology. *Mathematical Biosciences*, 125(2):155–164, 1995.
- [20] W. Liu, H. W. Hethcote, and S. A. Levin. Dynamical behavior of epidemiological models with nonlinear incidence rates. *Journal of Mathematical Biology*, 25(4):359–380, 1987.
- [21] W. Liu, S. Levin, and Y. Iwasa. Influence of nonlinear incidence rates upon the behavior of SIRS epidemiological models. *Journal of Mathematical Biology*, 23(2):187–204, 1986.
- [22] D. D. Munro. Recurrent subacute discoid lupus erythematosus. *Proceedings of the Royal Society of Medicine*, 56(2):78–79, 1963.
- [23] M. A. Nowak and R. M. May. *Virus Dynamics*. Oxford University Press, New York, 2000.
- [24] S. Ruan and W. Wang. Dynamical behavior of an epidemic model with a nonlinear incidence rate. *Journal of Differential Equations*, 188(1):135 – 163, 2003.
- [25] R. D. van Gaalen and L. M. Wahl. Reconciling conflicting clinical studies of antioxidant supplementation as HIV therapy: a mathematical approach. *BMC Public Health*, 9(Suppl. 1):1–18, 2009.
- [26] P. Yu. Computation of normal forms via a perturbation technique. *Journal of Sound and Vibration*, 211(1):19–38, 1998.
- [27] W. Zhang, L. Wahl, and P. Yu. Conditions for transient viremia in deterministic in-host models: Viral blips need no exogenous trigger. *SIAM Journal on Applied Mathematics*, 73(2):853–881, 2013.
- [28] W. Zhang, L. Wahl, and P. Yu. Viral blips may not need a trigger: How transient viremia can arise in deterministic in-host models. *SIAM Review*, 56(1):127–155, 2014.
- [29] W. Zhang, L. Wahl, and P. Yu. Modelling and analysis of recurrent autoimmune disease. *SIAM Journal on Applied Mathematics (Submitted for publication)*, xx(xx):xx–xx, xxxx.
- [30] W. Zhang, P. Yu, and L. Wahl. Backward bifurcation underlies rich dynamics in simple disease models. *To be submitted for publication*, xx(xx):xx–xx, xxxx.

Chapter 6

Conclusion

In this thesis, the problem of recurrent disease in infection and autoimmune models is studied via the qualitative analysis of dynamical systems using bifurcation theory. Although previous models with triggers such as stochastic components or forcing terms can simulate the cycle of long remission and brief relapse, simple deterministic models also exhibit recurrence.

Recurrence in HIV infection is referred to as “viral blips”. A 4-dimensional HIV antioxidant-therapy model, which exhibits viral blips, is analysed. The first hypothesis consisting of four conditions for the emergence of viral blips is proposed, which guides the derivation of the simplest (2- and 3-dimensional) infection model producing viral blips. A complete parameter region for the 3-dimensional infection model exhibiting viral blips is identified. Further dynamical study is conducted on the simplest 2-dimensional infection model, and gives rise to two more blips-generating mechanisms: hypothesis 2 and 3. The first hypothesis describes the scenario in which two equilibrium solutions intersect at a transcritical bifurcation point, and a Hopf bifurcation occurs at the upper branch of the disease equilibrium. Blips appear when the bifurcation parameter is close to the transcritical bifurcation point, and located in the parameter region where both equilibrium solutions are unstable. The second hypothesis adds another blips-generating mechanism, i.e. that large oscillations (or, more generally, global, persistent motions) can occur far away from the transcritical and Hopf critical points. In the third hypothesis, the existence of a Bogdanov-Takens bifurcation is proposed, which leads to a homoclinic loop near a Hopf bifurcation. This scenario may yield blips with very little change in amplitude, or extremely slow convergence to the disease-free equilibrium. The relapse-remission cycle is also characteristic of many autoimmune diseases. An autoimmune model which includes the role of regulatory T cells is modified by adding the terminally differentiated regulatory T cell subclass. The dynamical behavior is altered. Thus, recurrence is displayed in the modified autoimmune model and can be explained by the second hypothesis. Recurrence in infection and autoimmune models can arise naturally from the dynamical behavior of the system, without stochastic stimulation or exogenous triggers.

From the viewpoint of mathematical modelling, the occurrence of blips in the (2- and 3- and 4-dimensional) infection model is attributed to the convex incidence rate, which is formed by an increasing and saturating infectivity function. The convex incidence rate represents a cooperative effect in infection progression, that is, the existing infection enhances the ability for new infection to become established. The convex incidence rate also induces backward bifurcation, which facilitates the appearance of Hopf bifurcation, and rich dynamical behaviors, such

as bistability, recurrence, and regular oscillation. Cooperative effects in autoimmune disease occur during the T cell regulation process, since HLA-DR⁻ regulatory T cells differentiate and proliferate, forming the terminally differentiated HLA-DR⁺ class, which shows more efficient regulating capability. The autoimmune model investigated here displays negative backward bifurcation, in which the turning point is located in the negative state variable space. With the help of additional state variable, the modified autoimmune model shows Hopf bifurcation and exhibits recurrence.

We note that the amplitudes and frequencies in the observed oscillating and recurrent motions are all constant, because all parameter values are fixed for deterministic systems. However, in reality parameters should be time-varying, rather than constant. Time-varying parameter values in deterministic systems can generate oscillations with varying amplitudes and phases, called “amplitude modulation” and “frequency modulation”, which are analogous to the variation from random perturbations in stochastic models. This is demonstrated in Figure 2.12 of Chapter 2.

Clearly, the models analysed in this thesis are extreme simplifications of the mechanisms considered, and more precise mechanisms and accurate models could be considered in future. Nevertheless, the main insight of this thesis is to demonstrate that recurrence in disease can be generated from the cooperative interplay of dynamic populations. Hypotheses proposed in this thesis will serve as a starting point for further research on recurrent phenomena in other physical systems.

Other mechanisms for recurrence also exist, such as the recurrent activation of latently-infected lymphocytes. The delay which is characteristic of latent infection can be modelled using delay differential equations (DDEs), which could also generate oscillation and even recurrent patterns. A study of recurrent disease using DDEs would be a clear possibility for future work.

Curriculum Vitae

Name: Wenjing Zhang

Post-Secondary Education and Degrees: Bachelor of Science, Information and Computational Science
Shandong Normal University
Jinan, Shandong, China, 2002 - 2006

Master of Science, Pure Mathematics
Dalian University of Technology
Dalian, Liaoning, China, 2006 - 2008

Doctor of Philosophy, Applied Mathematics
University of Western Ontario
London, Ontario, Canada, 2010 - 2014

Honours and Awards: The paper: *Conditions for transient viremia in deterministic in-host models: viral blips need no exogenous trigger*, was awarded a SIGEST paper award from the Society for Industrial and Applied Mathematics (SIAM). The award includes the publication of an invited article in the journal *SIAM Review*, 2014

The 2013 National Award for Outstanding
Self-financed Chinese Students Study Abroad, 2013

Young Researcher Travel Awards, AMMCS-2013, 2013

The third prize scholarship, 2005
The second prize scholarship, 2004
The specialized scholarship, 2003
Shandong Normal University, China

Related Work Teaching Assistant
Experience: The University of Western Ontario
2010 - 2014

Research Assistant
The University of Western Ontario
2010 - 2014

Lecturer
Ludong University, Yantai, China
2008-2010

Publications:

Articles: Wenjing Zhang, Lindi M. Wahl and Pei Yu
Backward bifurcation underlies rich dynamics in simple disease models.
To be submitted to Journal of Mathematical Biology.

Pei Yu, Wenjing Zhang and Lindi M. Wahl
Dynamical analysis of a 2-dimensional disease model with convex incidence.
To be submitted SIAM Journal on Applied Dynamical Systems.

Wenjing Zhang, Lindi M. Wahl and Pei Yu
Modelling and analysis of recurrent autoimmune disease.
Under review at SIAM Journal on Applied Mathematics.

Wenjing Zhang, Lindi M. Wahl and Pei Yu
Viral blips may not need a trigger: How transient viremia can arise in deterministic in-host models. *SIAM Review*, 56(1):127–155, 2014.

Wenjing Zhang, Lindi M. Wahl and Pei Yu
Conditions for transient viremia in deterministic in-host models:
Viral blips need no exogenous trigger.
SIAM Journal on Applied Mathematics, 73(2):853–881, 2013.