Backward Bifurcation and Reinfection in Mathematical Models of Tuberculosis

A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

Isaac Mwangi Wangari

BSc (mathematics)-University of Nairobi,
PGD (mathematical sciences)-African Institute for
Mathematical Sciences,
MSc (applied mathematics)-University of Kwazulu Natal

School of Science
College of Science, Engineering and Health
RMIT University

29 December 2017
Dedication

This PhD thesis is dedicated to my mother who supported me mentally and spiritually.
Declaration

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work, paid or unpaid, carried out by a third party is acknowledged; and, ethics procedures and guidelines have been followed.

Isaac Mwangi Wangari

1 December 2018
Acknowledgements

First, I profoundly thank my supervisors Prof. Lewi Stone and Dr. Stephen Davis for their great effort in encouraging and guiding me throughout my PhD period. Their invaluable inspiration and sound advice were of great help to my progress. I would like to thank Dr. James Trauer from Monash University, School of Health and Preventive Medicine, for his suggestions on how to improve chapter 5 of my thesis. Also I thank Dr. Simon Peter Johnstone-Robertson for his suggestions on how to debug some of my Matlab codes.

Finally, I acknowledge the support I received for my research through the award of a scholarship by RMIT University, in particular by the School of Science, College of Science, Engineering and Health.
Abstract

Mathematical models are widely used for understanding the transmission mechanisms and control of infectious diseases. Numerous infectious diseases such as those caused by bacterial and viral infections do not confer life long immunity after recovering from the first episode. Consequently, they are characterized by partial or complete loss of immunity and subsequent reinfection. This thesis explores the epidemiological implications of loss of immunity using simple and complex mathematical models. First, a simple basic model mimicking transmission mechanisms of tuberculosis (TB) is proposed with the aim of correcting problems that are often repeated by mathematical modellers when determining underlying bifurcation structures. Specifically, the model makes transparent the problems that may arise if one aggregates all the bifurcation parameters when computing backward bifurcation thresholds and structures. The backward bifurcation phenomenon is an important concept for public health and disease management. This is because backward bifurcation signals that disease will not be eliminated even when the basic reproduction number $R_0$ is decreased below unity; rather, for the disease to be eliminated, $R_0$ has to be reduced below another critical threshold. I provide conditions to find the threshold correctly.

Secondly, the simple basic TB model is extended to incorporate epidemiological and biological aspects pertinent to TB transmission such as recurrent TB, which is defined as a second episode of TB following successful recovery from a previous episode. I study the conditions for backward bifurcation in this extended model that features recurrent TB. Mathematical techniques based on the center manifold approach, are used to derive an exact backward bifurcation threshold. Furthermore, both analytical and numerical findings reveal that recurrent TB is capable of inducing a new and rare hysteresis effect where TB will persist when the basic reproduction number is below unity even though there is no backward bifurcation. Moreover, when the reinfection pathway among latently infected individuals is switched off, leaving only recurrent TB, the model analysis indicates that recurrent TB can independently induce a backward
bifurcation. However, this will only occur if recurrent TB transmission exceeds a certain threshold. Although this threshold seems to be relatively high when realistic parameters are used, it falls within the recent range estimated in the relevant literature.

The second TB model is extended by dividing the latent compartment into two: fast (early latent) and slow (late latent) latent compartments, to enhance realism. Individuals in both early and late compartments are subjected to treatment. The proposed TB model is used to investigate how heterogeneity in host susceptibility influences the effectiveness of treatment. It is found that making the assumption that individuals treated with preventive therapy and recovered individuals (previously treated for active TB) acquire equal levels of protection after initial infection, and are therefore reinfected at the same rate, may obscure dynamics that are imperative when designing intervention strategies. Comparison of reinfection rates between cohorts treated with preventive therapy and recovered individuals who were previously treated from active TB provides important epidemiological insights. That is, the reinfection parameter accounting for the relative rate of reinfection of the cohort treated with preventive therapy is the one that plays the key role in generating qualitative changes in TB dynamics. In contrast, the parameter accounting for the risk of reinfection among recovered individuals (previously treated for active TB) does not play a significant role. The study shows that preventive treatment during early latency is always beneficial regardless of the level of susceptibility to reinfection. And if patients have greater immunity following treatment for late latent infection, then treatment is again beneficial. However, if susceptibility increases following treatment for late latent infection, the effect of treatment depends on the epidemiological setting: (a) for (very) low burden settings, the effect on reactivation predominates and burden declines; (b) for high burden settings, the effect on reinfection predominates and burden increases. This is mostly observed between the two reinfection thresholds, RT2 and RT1, respectively associated with individuals being treated with preventive therapy and individuals with untreated late latent TB infection.

Finally, a mathematical model that examines how heroin addiction spreads in society is formulated. The model has many commonalities with the TB model. The global stability properties of the proposed model are analysed using both the Lyapunov direct method and the geometric approach by Li and Muldowney. It is shown that even for a four dimensional model, the use of two well known nonlinear stability techniques becomes nontrivial. When all the parameters of the model are accounted
for, it is difficult if not impossible, to design a Lyapunov function. Here I apply
the geometric approach to establish a global condition that accounts for all model
parameters. If the condition is satisfied, then heroin persistence within the community
is globally stable. However, if the global condition is not satisfied heroin users can
oscillate periodically in number. Numerical simulations are also presented to give a
more complete representation of the model dynamics. Sensitivity analysis performed
by Latin hypercube sampling (LHS) suggests that the effective contact rate in the
population, the relapse rate of heroin users undergoing treatment, and the extent of
saturation of heroin users, are the key mechanisms fuelling heroin epidemic proliferation.
However, in the long term, relapse of heroin users undergoing treatment back to a
heroin using career, has the most significant impact.
# Contents

Dedication i  
Declaration ii  
Acknowledgements iii  
Abstract iv  
List of Figures xi  
List of Tables xix  
1 Introduction 1  

Introduction 1  
  1.1 Background on epidemics 1  
    1.1.1 Birth of mathematical epidemiology 3  
  1.2 Disease incidence functions 5  
  1.3 Bifurcation structures 6  
    1.3.1 Forward bifurcation 6  
    1.3.2 Backward bifurcation phenomena 8  
    1.3.3 A review on backward bifurcation phenomena 10  
  1.4 Objective, motivation and outline of the thesis 11  
  1.5 Published work 15  

2 Mathematical preliminaries 16  
  2.1 Linear and non-linear systems equilibria 16
<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2 Stable, unstable solutions and bifurcations</td>
<td>17</td>
</tr>
<tr>
<td>2.3 Method for computing $R_0$ (next generation operator method)</td>
<td>18</td>
</tr>
<tr>
<td>2.4 Bifurcations</td>
<td>20</td>
</tr>
<tr>
<td>2.5 Lyapunov functions and Lasalle’s invariance principle</td>
<td>22</td>
</tr>
<tr>
<td>3 Backward bifurcation in epidemic models: problems arising with</td>
<td>24</td>
</tr>
<tr>
<td>aggregated bifurcation parameters</td>
<td></td>
</tr>
<tr>
<td>3.1 Chapter overview</td>
<td>24</td>
</tr>
<tr>
<td>3.2 Model construction</td>
<td>25</td>
</tr>
<tr>
<td>3.3 Mathematical analysis</td>
<td>27</td>
</tr>
<tr>
<td>3.3.1 Basic properties</td>
<td>27</td>
</tr>
<tr>
<td>3.3.2 Existence of backward bifurcation</td>
<td>29</td>
</tr>
<tr>
<td>3.4 The backward bifurcation threshold, $R_c$</td>
<td>34</td>
</tr>
<tr>
<td>3.4.1 Proof of existence of backward bifurcation for model equation (3.1)</td>
<td>36</td>
</tr>
<tr>
<td>3.4.2 Equilibrium expressed in terms of the force of infection</td>
<td>39</td>
</tr>
<tr>
<td>3.5 Summary of the chapter</td>
<td>42</td>
</tr>
<tr>
<td>4 Backward bifurcation and hysteresis in TB models incorporating</td>
<td>43</td>
</tr>
<tr>
<td>recurrent tuberculosis</td>
<td></td>
</tr>
<tr>
<td>4.1 Chapter overview</td>
<td>43</td>
</tr>
<tr>
<td>4.2 Introduction</td>
<td>44</td>
</tr>
<tr>
<td>4.2.1 Evidence of recurrent TB</td>
<td>45</td>
</tr>
<tr>
<td>4.3 Model of recurrent tuberculosis</td>
<td>46</td>
</tr>
<tr>
<td>4.4 Model analysis</td>
<td>53</td>
</tr>
<tr>
<td>4.4.1 Basic properties</td>
<td>53</td>
</tr>
<tr>
<td>4.4.2 Endemic equilibria</td>
<td>54</td>
</tr>
<tr>
<td>4.5 No recurrent TB</td>
<td>56</td>
</tr>
<tr>
<td>4.6 With recurrent TB</td>
<td>59</td>
</tr>
<tr>
<td>4.6.1 Existence of backward bifurcation reinfection threshold at $R_0 = 1$</td>
<td>61</td>
</tr>
<tr>
<td>4.6.2 Relation with models of Lipsitch &amp; Murray 2003 and Feng et al.</td>
<td>62</td>
</tr>
<tr>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>4.7 Hysteresis</td>
<td>64</td>
</tr>
<tr>
<td>4.8 No latent reinfection</td>
<td>67</td>
</tr>
<tr>
<td>4.9 Impact of recurrent TB</td>
<td>69</td>
</tr>
<tr>
<td>4.10 Summary of the chapter</td>
<td>70</td>
</tr>
</tbody>
</table>
5  Heterogeneity in host susceptibility to tuberculosis and its effect on public health interventions 74
5.1  Chapter overview ................................................. 74
5.2  Introduction .................................................. 75
5.3  Model description .............................................. 78
5.4  Basic reproduction number $R_0$ .................................. 84
5.5  Reinfection threshold ........................................... 84
   5.5.1  Homogeneous reinfection risk ($\sigma_1 = \sigma_2 = \sigma_3 < 1$) ............. 88
5.6  Heterogeneity .................................................. 88
   5.6.1  Other features: hysteresis .................................. 90
   5.6.2  No partial immunity ........................................ 90
5.7  Interventions that impact reinfection ................................ 92
   5.7.1  Effect of treating early LTBI .................................. 92
   5.7.2  Effect of treating late LTBI .................................... 93
5.8  Summary of the chapter ....................................... 100
   5.8.1  World Health Organization (WHO) TB burden estimates ............. 101
   5.8.2  Limitations and conclusion .................................. 102

6  Other related work: analysis of a heroin epidemic model with non-linear treatment function 106
6.1  Chapter overview ................................................. 106
6.2  Introduction .................................................. 107
6.3  Model formulation .............................................. 110
6.4  Basic properties and basic reproduction number ..................... 112
   6.4.1  Endemic equilibria ............................................. 117
   6.4.2  Global stability for heroin-free equilibrium when $\omega = 0$ ........ 118
   6.4.3  Computation of new threshold for heroin eradication $R_0^C$ ....... 122
6.5  Global stability ................................................. 124
   6.5.1  Global stability using the Lyapunov method (special case $\omega = 
   \sigma = \xi = 0$) .................................................. 124
   6.5.2  A geometric approach to global stability ........................... 127
6.6  Numerical examples ............................................. 135
   6.6.1  Uncertainty and sensitivity analysis ............................ 138
6.7  Summary of the chapter ....................................... 142
## Appendixes

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix A</td>
<td>Computation of $R_0$ for the TB model with recurrent TB</td>
<td>148</td>
</tr>
<tr>
<td>Appendix B</td>
<td>Theorem deduced when $\theta = 0$</td>
<td>149</td>
</tr>
<tr>
<td>Appendix C</td>
<td>Detailed derivation of backward bifurcation threshold $p_c$</td>
<td>150</td>
</tr>
<tr>
<td>Appendix D</td>
<td>Basic properties and computation of $R_0$ for model equation (5.1)</td>
<td>153</td>
</tr>
<tr>
<td>Appendix E</td>
<td>Interpretation for $R_0$ of model equation (5.1)</td>
<td>156</td>
</tr>
<tr>
<td>Appendix F</td>
<td>Global stability of DFE for a special case of model equation (5.1)</td>
<td>158</td>
</tr>
<tr>
<td>Appendix G</td>
<td>Existence of endemic equilibria for model (5.1)</td>
<td>160</td>
</tr>
<tr>
<td>Appendix H</td>
<td>Procedure for calculating reinfection threshold as derived in the literature</td>
<td>163</td>
</tr>
<tr>
<td>Appendix I</td>
<td>Computation of reinfection thresholds</td>
<td>166</td>
</tr>
<tr>
<td>Appendix J</td>
<td>Endemic equilibria of the submodel (5.3)</td>
<td>168</td>
</tr>
<tr>
<td>Appendix K</td>
<td>Derivation of $R_0$ for heroin epidemic model</td>
<td>169</td>
</tr>
<tr>
<td>Appendix L</td>
<td>Proof of existence of backward bifurcation for heroin epidemic model (6.1)</td>
<td>170</td>
</tr>
</tbody>
</table>

## Bibliography

<table>
<thead>
<tr>
<th>Bibliography</th>
<th>Page</th>
</tr>
</thead>
</table>
List of Figures

1.1 Qualitative illustration of forward bifurcation at the point $R_0 = 1$, where the vertical axis represents the equilibrium force of infection $\lambda^* = \beta I^*/N$. 7

1.2 Qualitative illustration of backward bifurcation at the point $R_0 = 1$, where the vertical axis represents the equilibrium force of infection $\lambda^* = \beta I^*/N$. 9

3.1 Schematic diagram for the model equation (3.1). The red curved dashed arrow indicates the reinfection process. 26

3.2 Backward bifurcation when all the parameters in $R_0$ are fixed at constant values but the aggregated $R_0$ is nonetheless varied. Parameters used are $\Lambda = 10$, $\mu = 0.016$, $r = 2$, $\mu_d = 0.4$, $k = 0.0005$, $p = 0.2$, $\beta = 0.036$. Black solid line represents stable equilibria while the red solid line represents unstable equilibria. 32

3.3 Backward bifurcation when one parameter $\beta$ in $R_0$ is varied. Parameters used are $\Lambda = 10$, $\mu = 0.016$, $r = 2$, $\mu_d = 0.4$, $k = 0.0005$, $p = 0.2$, $\beta \in [0.025, 0.175]$. Black solid lines represent stable equilibria while the red solid line represents an unstable equilibrium. 33

3.4 The red dashed line is a plot of $R_{c1}$ as a function of $p$ while the blue solid line illustrates a plot of the correct backward bifurcation threshold $R_{c2}$ as a function of $p$. Parameters used are $\Lambda = 10$, $\mu = 0.016$, $r = 2$, $\mu_d = 0.4$, $k = 0.0005$, $\beta = 0.0195$. The plot of $R_c$ as a function of reinfection is due to the fact that $R_c$ decreases as reinfection $p$ increases. $p_{\min}$ is the minimum value of exogenous reinfection that triggers backward bifurcation. 36
3.5 Backward bifurcation when endemic equilibria are expressed in terms of
the force of infection. (a) Backward bifurcation when all the parameters
in $R_0$ are fixed at constant values. Parameters used are $\mu = 0.01, r = 0.85, \mu_d = 0.1, k = 0.002316, p = 0.5, \beta = 4.5$. (b) Backward bifurcation
when one parameter in $R_0$ is varied. Parameters used are $\mu = 0.01, r = 0.85, \mu_d = 0.1, k = 0.002316, p = 0.5, \beta \in [3.6, 6]$. In both figures the
blue solid lines represent stable equilibria while red solid lines represent
unstable equilibria. 41

4.1 Schematic diagram of the main processes involved in TB infection
according to model equation (4.1). The red dotted curved arrows
represent reinfection pathways. 51

4.2 (a) Illustration of backward bifurcation when there is no recurrent TB
(i.e., $\theta = 0$). Parameters used remain as defined in Table 4.1 except
$p = 0.09 > p_c = 0.0658, k = 0.0002, q = 0.05, c = 25, \beta \in [0.4, 0.9]
and $\beta_c = 0.5099$ corresponding to $R_c = 0.8852$. In the figure the
blue solid lines represent the stable equilibria while the red dotted line
represents unstable equilibria. (b) A plot of the critical value of the
basic reproduction number, $R_c$, as a function of the level of exogenous
reinfection $p$. NBB and BB respectively denote no backward bifurcation
and backward bifurcation regions. That is, in the region denoted by
NBB the level of exogenous reinfection is too low to induce backward
bifurcation while in the region denoted by BB the level of exogenous
reinfection is sufficient to cause multiple positive endemic equilibria. 59

4.3 Plots the number of infectives at equilibrium as a function of $R_0$. In
both figures the blue solid lines represents stable equilibria while the
red dotted line represents an unstable equilibrium. (a) Shows forward
bifurcation with parameters $p = 0.06 < p_c = 0.0647, \sigma = 0.05, \theta = 0.3,
\mu = 0.0167, \mu_d = 0.1, r = 2, q = 0.05, \Lambda = 100, c = 60, k = 0.0002,$
$\beta \in [0.3, 0.8]$. (b) Shows backward bifurcation with the same as in (a)
except $p = 0.09 > p_c = 0.0647$. For a clear view all bifurcation structures
are plotted with semi-log axes. 65
4.4 Plots show the infectious population $I^*$ at equilibrium as a function of $R_0$. (a) Forward bifurcation with hysteresis where the multiple equilibria are strictly to the right of $R_0 = 1$. (b) Forward bifurcation with hysteresis where there are multiple equilibria to the left and to the right of $R_0 = 1$. Parameters used are $\mu = 0.0167, \mu_d = 0.1, k = 0.0002, \theta = 0.5, r = 2, q = 0.05, \sigma = 0.2, c = 60, \Lambda = 100, \beta \in [0.2, 0.8]$. In (a) $p = 0.057 < p_c = 0.0639$ and in (b) $p = 0.058 < p_c = 0.0639$. In both figures the blue solid lines represent stable equilibria while the red dotted lines represent unstable equilibria. For a clear view, all bifurcation structures are plotted with semi-log axes.

4.5 A scenario where the reinfection path $A$ is omitted (i.e., no reinfection of exposed individuals, $p = 0$). Parameters used are $\mu = 0.0167, \mu_d = 0.1, k = 0.0002, r = 2, q = 0.05, \sigma = 0.2, c = 60, \beta \in [0.5, 0.65]$. (a) A scenario where $\theta = 4 < \theta_c = 5$ and in this case there is only forward bifurcation at $R_0 = 1$. (b) and (c) show backward bifurcation and the corresponding $\theta$ values are respectively $\theta = 6 > \theta_c = 5$ and $\theta = 7.5 > \theta_c = 5$.

4.6 (a) The effect of recurrent TB on backward bifurcation due to the incorporation of reinfection pathways B and C. Parameter values are $\mu = 0.0167, \mu_d = 0.1, k = 0.0002, r = 2, q = 0.05, \sigma = 0.05, p = 0.09, c = 60, \Lambda = 100, \beta \in [0.45, 0.7]$. With no recurrent TB ($\theta = 0$), $p = 0.09 > p_c = 0.0659$ while with recurrent TB ($\theta = 0.3$) parameters are the same but $p_c$ is altered, i.e., $p = 0.09 > p_c = 0.0647$. The blue solid lines represent stable equilibria while the red dotted line represent an unstable equilibrium. (b) Contour plots of force of infection at equilibrium as a function of $\theta$ and $p$. The figure shows that increasing recurrent TB increases TB prevalence.
5.1 Schematic representation of the model. The square boxes represent classification of the general population into six mutually exclusive sub-populations, i.e., susceptibles $S(t)$, early latents $L_1(t)$, late latents $L_2(t)$, individuals treated with isoniazid preventive therapy $P(t)$, individuals with active TB $I(t)$, and recovered individuals $R(t)$. All arrows indicate either inflow or outflow or transition between compartments. Blue arrows illustrate transition of latently infected individuals as a result of treatment with IPT. Red dashed arrows show reinfection of late latently infected individuals, individuals treated with IPT and recovered individuals, respectively represented by $\sigma_i$, $i = 1, 2, 3$. 81

5.2 Bifurcation diagrams of equilibrium TB prevalence as a function of the transmission coefficient $\beta$. (a) In the absence of reinfection the epidemic threshold occurs at $R_0 = 1$. (b) In the presence of reinfection four thresholds emerges. That is the threshold that occurs at $R_0 = 1$ and the reinfection thresholds $RT_1$, $RT_2$ and $RT_3$. (c) Bifurcation diagram of the reinfection submodel (5.3). The red dotted vertical line marks the transcritical bifurcation point $R_0 = 1$, that is the endemic threshold, while the black vertical dashed lines mark the points where reinfection thresholds $RT_1$, $RT_2$ and $RT_3$ converge. Parameters used are the baseline values in Table 5.1. Reinfection susceptibility parameters are shown in the figures. In all figures a semi-logarithmic scale is used for a clear view. 85

5.3 Illustration of the relative importance of $\sigma_2$ and $\sigma_3$ to equilibrium dynamics. Parameters used are the baseline parameters in Table 5.1. (a) Impact of increasing $\sigma_2$ while both $\sigma_1$ and $\sigma_3$ are fixed. (b) Effect of increasing $\sigma_3$ while both $\sigma_1$ and $\sigma_2$ are fixed. 89
5.4 Exploration of the full range of plausible values of the relative risk of reinfection parameters. All panels show equilibrium prevalence as a function of the transmission coefficient $\beta$. From left to right, each column shows an increase in the relative risk of reinfection among recovered individuals $\sigma_3$, while from top to bottom each row shows increasing risk of reinfection among individuals treated with IPT $\sigma_2$. Other parameter values remain at baseline values in Table 5.1. The red dotted vertical lines mark the point where $R_0 = 1$. The dotted red segments of the endemic curves represent the unstable equilibria while the blue lines represent stable equilibria.

5.5 Emergence of hysteresis and backward bifurcation effects with changes to reinfection parameters. Parameters used are the baseline parameters in Table 5.1. (a) Occurrence of a hysteresis effect. (b) A scenario where all cohorts subject to reinfection have no immunity and are therefore infected at the same rate as the susceptible population. (c) A scenario where all post-infection cohorts have no partial immunity and have increased susceptibility to reinfection.

5.6 Impact of treating early latent individuals under different risks of susceptibility to reinfection. (a) Treatment of early LTBI (and of active disease) leads to increased susceptibility to reinfection compared to late latent infection. (b) The level of susceptibility to reinfection among late LTBI is lower than for both individuals treated with IPT and recovered individuals treated from active TB. Figures (c) and (d) respectively represent figures (a) and (b) with a linear scale y-axis. Treatment of late LTBI and individuals with active TB are respectively set at $\rho = 0.1$ and $\tau = 2$ while other parameters are as shown in Table 5.1.
5.7 Effect of treatment for late latent infection under the assumption that treatment decreases (panels (a) and (b)) and increases (panels (c) and (d)) susceptibility to reinfection, and assuming that treatment for active disease decreases (panels (a) and (c)) and increases (panels (b) and (d)) susceptibility to reinfection. Treatment of individuals with active TB is fixed to $\tau = 2$, and treatment of individuals with early latent TB is fixed to $\theta = 1$, while treatment of late latent individuals is introduced at different rates; $\rho = 0, 0.1, 1, 10$ and the limit as $\rho \to \infty$. Other parameters used are as shown in Table 5.1. (a) $\sigma_2, \sigma_3 < \sigma_1$. (b) $\sigma_2 < \sigma_1 < \sigma_3$. (c) $\sigma_3 < \sigma_1 < \sigma_2$. (d) $\sigma_1 < \sigma_2, \sigma_3$. In the figures the black dashed endemic lines represent the baseline case where there is no treatment of late LTBI (i.e., $\rho = 0$) while the blue dashed endemic lines represent immediate treatment for the entire population ($\rho \to \infty$). In both figures (a) and (b) the dashed red lines of the endemic curves represent unstable equilibria while blue solid lines represent stable equilibria. All figures are plotted using a semi-logarithmic scale.

5.8 Treatment of late latent infection under the assumption of no effect on susceptibility to reinfection (panels (a) and (b)) and increased susceptibility (panels (c) and (d)), with reactivation mechanisms present (panels (a) and (c)) and removed (panels (b) and (d)). Parameters are $\tau = 2$, $\theta = 1$, $\rho = 0, 0.5, 1$ and $\rho \to \infty$ while other parameters remain as in Table 5.1. The risks of reinfection are shown in each sub-figure.

5.9 Illustration of world TB burden by regions across the world obtained using global TB data [1]. Note that the size of the circle is proportional to the population size of the country.

5.10 Illustration of global TB burden by country using global TB report data [1].

6.1 Illustration of the qualitative features of backward bifurcation. The red dotted line represents the unstable equilibrium (i.e., unstable endemic equilibria and unstable heroin-free equilibrium) while the blue solid lines represent stable equilibria (i.e., stable endemic equilibria and stable heroin-free equilibrium).
6.2 A heroin epidemic model with a density-dependent incidence rate and a saturated treatment function. The blue solid arrows represent deaths either due to heroin or natural causes while the black solid arrows represent change of status from one compartment to another. 113

6.3 (a) and (b) represent bifurcations where drug users $U_1^*$ are not in treatment; equilibria are plotted as a function of $R_0$. The blue solid lines represent the stable equilibria while red solid lines represent unstable equilibria. (a) Represents forward bifurcation with parameters $\omega = 0.11 < \omega_c = 0.1156, \mu = 0.01, \beta = 0.001, \delta_1 = 0.002, \delta_2 = 0.001, \xi = 0.015, \alpha = 0.9, p = 0.467, \sigma = 0.1$ and $\Lambda \in [1, 3]$. (b) Represents backward bifurcation with parameters as in (a) except $\omega = 0.25 > \omega_c = 0.1156$. (c) The critical value $R_0^C$ as a function of the saturation parameter $\omega$. The black dotted line represents the threshold $\omega_c$ which if exceeded gives rise to a backward bifurcation. (d) The critical value $R_0^C$ as a function of treatment rate $\alpha$. 125

6.4 Illustration of the effect of increasing parameter $\omega$ that accounts for the saturation of heroin users. Here the parameters remain as in the caption of Figure 6.3 except $\Lambda = 2$ while $\omega$ is shown. The heroin eradication thresholds (i.e., $R_0^C$) corresponding to figures (a)-(d) are respectively $0.5204, 0.6038, 0.7314, 0.9131$. 136

6.5 (a) and (b) illustrate the dependence of the heroin epidemic extinction or persistence on the initial states provided to the model. Parameters are $\omega = 0.15, \beta = 0.001, \delta_1 = 0.002, \delta_2 = 0.001, \xi = 0.015, \mu = 0.01, \alpha = 0.9, p = 0.467, \sigma = 0.1, \Lambda = 1.5$. With these parameters the corresponding value of $R_0$ is 0.7506. (c) The impact of increasing parameter $\omega$ (the extent of saturation of treatment) on heroin users when $R_0 < 1$. Parameters used are $\beta = 0.001, \delta_1 = 0.002, \delta_2 = 0.001, \xi = 0.015, \mu = 0.01, \alpha = 0.9, p = 0.46, \sigma = 0.1, \Lambda = 1.5$ which correspond to $R_0 = 0.7427 < 1$. (d) The effect of treatment $\alpha$ on heroin users $U_1$ when all other parameters and initial conditions are fixed. 138

6.6 (a) Shows time series of susceptibles $S$ as a function of parameter $\omega$ that accounts for the extent of saturation of treatment when $R_0 > 1$. (b) Shows time series of drug users $U_1$ as a function of parameter $\omega$ when $R_0 > 1$. In both Figures, parameters used are the same as in Figure 6.5(a) except $p = 0.46, \Lambda = 3$ corresponding to $R_0 = 1.4855 > 1$. 139
6.7 Validity of the global stability condition. Figures (a), (b), (c) and (d) represent a scenario where condition (6.29) holds true and global stability is predicted. Figures (e), (f), (g) and (h) represent a scenario where condition (6.29) does not hold and oscillations are expected. . . . 140
6.8 The PRCC output for heroin users $U_1(t)$ not in treatment. . . . . . . . . 141
List of Tables

3.1 Description of model parameters and variables of model equation (3.1). 26
4.1 Description of variables and parameters of model (4.1). . . . . . . . . . . 52
4.2 Summary of the model equilibria. . . . . . . . . . . . . . . . . . . . . . 61
5.1 Parameters and definitions for model equation (5.1). . . . . . . . . . . . 83
6.1 Description of variables and parameters of model (6.1). . . . . . . . . . . 113
6.2 Parameter baseline values and ranges used in sensitivity analysis. . . . 142
Chapter 1

Introduction

1.1 Background on epidemics

The conceptualization of the existence of invisible creatures capable of transmitting disease can be traced back at least to the writings of Aristotle (384 BC-322 BC). This idea was further developed as a theory in the 16th century [2]. With the help of the first microscope, Leeuwenhoek (1632-1723) confirmed the existence of microorganisms and this led to the inception of the germ theory of disease by Jacob Henle (1809-1885) in 1840. In the nineteenth and the early part of the twentieth century the germ theory was later developed by Robert Koch (1843-1910), Joseph Lister (1827-1912) and Louis Pasteur (1827-1875). In the 21st century, the transmission mechanisms of many diseases is now well known, as is the operation of the human immune system. For example, diseases whose aetiological agent is a virus, such as influenza, measles, rubella (German measles) and chicken pox, are known to render immunity against reinfection. In contrast, diseases spread by bacteria such as tuberculosis, meningitis and gonorrhea render no immunity against reinfection [2].

A disease in a particular place and host population may either be categorized as epidemic or endemic. Loosely speaking, an epidemic might be described as an abrupt occurrence of a disease which infects an often substantial proportion of a population, possibly causing many deaths over a short period of time before vanishing. Endemic refers to diseases that persist within the population for a sustained and possibly indefinite period of time, usually only infecting a relatively small proportion of the population [3]. Moreover, when an epidemic occurs on a scale which crosses international boundaries, usually affecting a large number of people it is referred to as a pandemic
Pandemics are a common phenomenon in human history. For instance, in the 14th century (beginning in 1346-1350) the Black Death (bubonic plague) spread from Asia throughout Europe causing the death of about one third of the population of Europe. The disease remained chronic (endemic) for a period of about 300 years (for instance the great plague of London (1656-1666)), but then gradually disappeared from Europe. Over three hundred million people died from small pox in the 19th century. Another devastating pandemic was the 1918 influenza pandemic (Spanish Flu) which globally affected one third of the human population and resulted in the death of some 20 to 100 million people.

Epidemics continue to emerge and invade. For example, the human immuno-deficiency virus (HIV) which came to be known in 1980 as the causative agent of AIDS (Acquired Immuno-Deficiency virus) has killed approximately 25 million people globally and approximately 33 million people are currently living with HIV/AIDS. Annually, malaria infects about three hundred million people and kills approximately two million people. Infections such as cholera, hemorrhagic fevers and plague occasionally recur in some regions, while others (such as malaria, HIV/AIDS, mycobacterium tuberculosis, typhus, cholera, schistosomiasis, etc.) have become endemic in various locations worldwide.

There has been tremendous improvement and innovation in medical science over the late 20th and 21st century, but nonetheless infectious diseases continue to be a major contribution of high morbidity and mortality in human populations, especially in developing countries. Recent studies conducted in sub-Saharan Africa show that infectious diseases are responsible for more than half of human deaths. Consequently, infectious diseases impose heavy public health and sociol-economic burdens on populations that are least able to bear them. It is imperative to note that the infectious disease hazards go beyond human populations and affect domestic animals, wildlife and plant populations. A number of factors such as complex ecology, fast evolution in response to changing circumstances, and emergence of new pathogens ensure that infectious diseases will continue to pose serious challenges for the foreseeable future.

Often public health officials ask particular questions when there is an outbreak within a given community:

(i) How many people will be infected?
(ii) How many people require treatment?

(iii) For how long will the epidemic persist in the community?

(iv) Does infection confer immunity against a second episode?

(v) What is the appropriate intervention strategy (such as use of treatment, vaccination or quarantine) to minimise the impact of the epidemic?

Mathematical modelling acts as an invaluable tool in attempting to find answers to these questions. Historical time-series of epidemic outbreaks illustrate that infectious diseases can exhibit complex dynamics typical of non-linear systems. Mathematical models allow rigorous analysis of such complex dynamics, including:

(a) Prediction of epidemiological threshold phenomena (such as the basic reproduction number) and expected disease burden (morbidity, mortality, attack rate);

(b) Assessment and comparison of different prophylaxis (preventive) and therapeutic measures;

(c) Formulation and validation of theories, qualitatively determining sensitivities to changes in parameter values;

(d) Provision of insight on how an infectious disease spreads in the real world and how various complexities impact the dynamics;

(e) Identification of overall trends of an epidemic and making general forecasts.

1.1.1 Birth of mathematical epidemiology

The application of mathematical modelling in tracking transmission dynamics of infectious diseases was in use as early as the eighteenth century. An example is the pioneering work by Daniel Bernoulli on smallpox (1760) in which he developed a mathematical model to assess the impact of variolation against smallpox, thereby increasing individual life expectancy. His work incorporated the concept of differential mortality to enable him to approximate the rate of deaths attributable to a certain disease. Furthermore, his idea of differential mortality has been used to approximate disease death rates of past epidemics [2], for example, the influenza pandemic which occurred in 1918.
In the late nineteenth and twentieth centuries, both public-health physicians and biologists laid the foundations of mathematical epidemiology. Such persons include P.D. En’ko [11], W. H. Hamer [12], J. Brownlee [13], Sir R. A Ross [14], A. G. McKendrick and W.O Kermack [15, 16, 17]. After Bernoulli’s smallpox model, various mathematical models began to emerge. In 1906 Ronald Ross discovered that mosquitoes were the agents responsible for spreading malaria from human to human. He then developed a model to describe the spread of malaria [14]. From the model he deduced that by reducing the mosquito population in a region, malaria can be effectively controlled. His malaria model was the first example that clearly demonstrated the epidemic threshold concept, which became fundamental in epidemiology theory ever since [2]. Two decades later, Kermack and McKendrick developed the theory further and demonstrated that if the density of susceptible individuals exceeds a critical threshold then an epidemic is likely to occur, but below the critical threshold an epidemic cannot be triggered [15]. This theory, together with the principle of mass action, formed the cornerstone of modern mathematical biology.

The so-called “SIR model” equations used by Kermack and McKendrick to describe the transmission mechanisms of bubonic plague may be written down as [15]

\[
\begin{align*}
\frac{dS}{dt} &= -\frac{\beta SI}{N}, \\
\frac{dI}{dt} &= \frac{\beta SI}{N} - \alpha I, \\
\frac{dR}{dt} &= \alpha I.
\end{align*}
\] (1.1)

Here the total population, denoted by \(N(t)\), is classified into three mutually-exclusive disease states, with \(S(t)\) representing the susceptible population, \(I(t)\) the infected and \(R(t)\) the recovered or removed population. Hence, \(N(t) = S(t) + I(t) + R(t)\). The parameter \(\beta\) is the transmission coefficient (effective contact rate) and \(\alpha\) is the per-capita rate of recovery (or removal) for infected individuals. The interaction of two individuals, one susceptible (S) and another infected (I) may result in susceptible individuals becoming infected. Individuals infected are removed from the infectious state at recovery rate \(\alpha\). The behaviour of model equation (1.1) is entirely determined by its initial conditions and the deterministic rules which describe the development of the model; hence it is considered deterministic.
The SIR model equations (1.1) have been extended to incorporate important and realistic epidemiological and biological features such as loss of immunity, or extended to consider latent or exposed individuals who are not yet infectious and do not manifest disease symptoms. Such possibilities cannot be described by simple SIR models. Models then typically take the form of SIR, SIS, SIRS, SEIR, SEIRS, SEIRE or SEIRI where the notation E represents the population of individuals who have been exposed to infection and are not yet infectious. Note that non-italic symbols are used to define each infectious state while italic symbols are used to count the number of individuals in each state. Variables for compartment labels are italicized.

1.2 Disease incidence functions

In epidemic models incidence functions describe the rate at which new infections are generated. A general method for formulating infection incidence functions is described in [18]. Hence, following [18], let $\beta(N)$ represent the mean number of effective contacts per person per unit time. Now the force of infection $\lambda = \beta(N)I/N$ is the mean number of effective contacts between infectious and susceptible individuals per unit time. Hence, new cases of infectives are generated at a rate proportional to $\lambda S$. For a constant $\beta(N)$ (that is $\beta(N) = \beta$), $\lambda S$ is termed a frequency-dependent incidence function. The frequency-dependent incidence function reflects a situation where the number of contacts is independent of the size of the population [18, 19].

Moreover, if $\beta(N)$ is a function of the total population size (e.g., $\beta(N) = \beta N$) then $\lambda S$ is referred to as a density-dependent incidence function [20]. The density-dependent incidence function assumes that “as the size of the population increases so does the contact rate” [19]. Both frequency-dependent and density-dependent incidence functions are widely used in the modelling of infectious diseases [18, 21]. The preference of which incidence function to use is largely determined by the disease being modelled, although in some cases (as it will be seen) the choice has been dictated by mathematical tractability. In the literature some studies have pointed out that the frequency-dependent incidence formulation is more appropriate for human infections [22, 23].
1.3 Reproduction number and common bifurcation structures

The basic reproduction number, sometimes referred to as the reproduction ratio $R_0$, is one of the most celebrated quantities in the literature pertaining to mathematical epidemiology. The definition of $R_0$ is the average number of new infections that a typical infected individual gives rise to over the lifetime of the infection when introduced into a wholly susceptible population [24, 25]. Although the $R_0$ concept was formalized much earlier in the context of demographic theory, and despite numerous opportunities to cross over to other disciplines such as ecology and epidemiology [26], it was not until 1980 that it fully developed and its applicability was realized in theoretical epidemiology (see [25, 27]).

1.3.1 Forward bifurcation

The reproduction number is a fundamental epidemiological quantity since it determines whether an infection will be able on average to reproduce itself ($R_0 > 1$) in the population and spread, or not ($R_0 < 1$) and so die out. Characteristically, when $R_0$ is below unity, the introduction of a few infected individuals in a susceptible population will only lead to disease die-out as the infection is unable to reproduce itself or transmit through the population effectively. Conversely, when $R_0$ is above unity an epidemic may occur and long-term disease persistence is feasible. Classical epidemic models are therefore usually found to have two intrinsic equilibria: a disease free equilibrium (DFE); and a non-trivial endemic equilibrium point (EEP). By endemic is meant an equilibrium in which the number of infectives is greater than zero. The stability of these equilibria switch at the (transcritical) bifurcation point which occurs when $R_0 = 1$. Thus, the point $R_0 = 1$ defines an important threshold for understanding the transmission dynamics of infectious diseases. Figure 1.1 illustrates the more typical forward bifurcation by plotting the force of infection at equilibrium $\lambda^* = \beta I^*/N$ in a population as a function of $R_0$. (Here the star notation refers to equilibrium values.) The figure shows that a stable DFE in which $\lambda^* = 0$ exists when $R_0 < 1$. However, when crossing the threshold to a regime where $R_0 > 1$, there is a change of stability where the DFE becomes unstable and the endemic equilibrium point stabilizes.
Figure 1.1: Qualitative illustration of forward bifurcation at the point $R_0 = 1$, where the vertical axis represents the equilibrium force of infection $\lambda^* = \beta I^*/N$. 
1.3 Bifurcation structures

1.3.2 Backward bifurcation phenomena

For decades, it has been widely accepted that the condition $R_0 < 1$ is an essential requirement for the elimination of a disease. However, this viewpoint has been recently challenged with a number of theoretical studies demonstrating that the criterion may not always be sufficient. Instead, the phenomenon of \textit{backward bifurcation} offers a different interpretation since it shows that although $R_0 < 1$ and the DFE is stable, there might still be another stable endemic equilibrium coexisting simultaneously. Thus, even though $R_0 < 1$, a population may still be at an endemic equilibrium in which the disease persists indefinitely. When there are multiple stable equilibria coexisting simultaneously, the final equilibrium a population will reach depends on the initial conditions (in terms of numbers of individuals) of its sub-populations. Figure 1.2 provides a typical bifurcation diagram that shows the key features of a backward bifurcation. Note that three equilibria coexist when $R_0$ is in the range $0 < R_c < R_0 < 1$, where $R_c$ is a critical value, which in Figure 1.2 is $R_c = 0.83$. In this range, the “middle” equilibrium is unstable, while the other two outer equilibria (the DFE and the endemic equilibrium) are both stable. When $R_0 < R_c$, only the DFE exists and is stable.

Multiple coexisting equilibria can lead to interesting dynamics when, for example, $R_0$ varies slowly as can be understood from examining Figure 1.2. Suppose, for example, that the infected population is close to extinction ($I^* = 0$) (i.e., close to the DFE) and $R_0$ increases slowly. As soon as $R_0$ passes through the threshold point $R_0 = 1$ the number of infectives will jump from close to $I^* = 0$ ($\lambda = 0$) to the large positive endemic equilibrium $I^* > 0$ ($\lambda > 0$), as indicated by the arrow in Figure 1(b). Similarly, when the infective population sits close to the endemic equilibrium and $R_0$ reduces slowly through the threshold point $R_0 = 1$, rather than switching to the DFE, the infectives remain close and converge to the stable endemic equilibrium. When $R_0$ falls below the threshold $R_c$, the infective population jumps towards zero (the DFE), while the endemic equilibrium disappears. These jumps are a well known phenomenon in nonlinear dynamical systems where hysteresis effects can arise. Hysteresis refers to a scenario where multiple endemic equilibria coexist when $R_0 > 1$. 
Figure 1.2: Qualitative illustration of backward bifurcation at the point $R_0 = 1$, where the vertical axis represents the equilibrium force of infection $\lambda^* = \beta^* / N$. 
1.3 Bifurcation structures

1.3.3 A review on backward bifurcation phenomena

The epidemiological importance of backward bifurcation phenomena in public health management triggered a renewed interest amongst the mathematical modelling community [28]. In terms of epidemiological models, the phenomenon was first observed in an HIV model that segregated the sexually active population into $n$ mixing cohorts with differential risk levels of infection [29]. However, the backward bifurcation phenomena has since been discovered in many other disease contexts and has been associated with implementation of intervention strategies that are imperfect [30, 31], behavioral responses [29, 30, 31], epidemic models accounting for immunological factors [32, 33], cohort-structured models for fatal diseases [34, 35], incomplete immunity [30, 36, 37, 38, 39, 40], models with non-constant contact rates [41], vaccination [42, 43, 44, 45], infection-age structuring [43], vertical transmission and non-linear incidence [46], inadequate treatment resources [47, 48, 49, 50] and in standard incidence versus mass action contact models [51]. Numerous infectious diseases studied at the population and immunological level have been shown to be capable of exhibiting backward bifurcation phenomena. Pathogens include HIV as demonstrated in [51, 52, 53, 54], bovine respiratory syncytial virus amongst cattle [36, 55], TB [52, 56, 57, 58, 59, 60, 61, 62], HTLV-1 [46], dengue [51, 63], syphilis [64], West Nile virus [65, 66], hepatitis B and C [67, 68], H5N1 influenza [69], H1N1 influenza [70], malaria [35, 69, 71, 72, 73], Toxoplasma gondii [33] and echinococcus [74]. Backward bifurcation also occurs in epidemic models of drug abuse [75, 76].

1.3.4 Definition of reinfection TB, recurrent TB and endogenous reactivation

Reinfection TB (or exogenous reinfection) refers to a TB episode that results after being reinfected with a new TB strain from another infectious individual [56, 77]. Recurrent TB is defined as the emergence of a second episode of TB after the first episode has been successfully cured. It is important to note that there are two mechanisms by which recurrent TB can occur: (i) relapse with the previously responsible strain or (ii) exogenous reinfection from a new exposure [78, 79, 80, 81]. Note that reinfection of those who have latent state TB is not considered recurrent TB.
The difference between the two terminologies is that for reinfection TB one does not need to have recovered from the first episode while in the second case, for TB to be referred as recurrent TB, a person must have been successfully cured from the first episode of TB. Endogenous reactivation is defined as the reactivation of a preexisting dormant infection [82, 83]. Within the context of TB, endogenous reactivation occurs among latently infected individuals [56].

1.4 Objective, motivation and outline of the thesis

This thesis concerns mathematical models of infectious diseases especially those in which the phenomenon of immunity and loss of immunity arises. Many diseases are characterized by this behaviour including, for example, tuberculosis, influenza, chicken pox, and some sexually transmitted diseases. But also cultural phenomena, such as drug addiction, provide classic examples of this behaviour. In this thesis it is found that for nonlinear mathematical models, loss of immunity introduces different types of unusual dynamical behaviours. A key goal is to understand how loss of immunity generates backward bifurcations and other complex bi-stability phenomena. This thesis revolves around five main research questions:

(i) How do backward bifurcations arise in epidemic models, and how can one determine analytic thresholds for backward bifurcations?

(ii) How are backward bifurcation phenomena in models of tuberculosis impacted by loss of immunity and subsequent reinfection among recovered individuals, especially in light of recent new empirical evidence that loss of immunity and subsequent reinfection is quite high (approximately four times higher in comparison to rates of new TB [84])?

(iii) How does variability in risk of reinfection alter TB dynamics in a model accounting for heterogeneity in host susceptibility?

(iv) How does variability in risk of reinfection influence the effectiveness of treatment?

(v) How can one establish global stability properties of models that have complex non-linear treatment rates?

The first research question addresses two main problems that have been noted in the current mathematical literature regarding backward bifurcations. Backward bifurcations are generally studied by varying a bifurcation parameter which in epidemiological
models is usually the basic reproduction number $R_0$. However, it is often overlooked that the basic reproduction number $R_0$ is an aggregate of parameters in the model. Therefore, it will not be correct for one to simply vary the aggregate $R_0$ while leaving all model parameters constant as has happened many times in the literature (e.g., [21, 48, 52, 63, 85, 86, 87, 88, 89, 90]). Moreover, this research question emphasizes the correct approach for computing the critical value of the basic reproduction number $R_c$. This clarifies a great deal of confusion in the literature on how to compute the critical value. In this thesis these problems are addressed and resolved.

The second research question investigates the role of loss of immunity and subsequent reinfection amongst recovered individuals in a tuberculosis (TB) model. This research question was inspired by the recent understanding that despite tremendous improvements in TB treatment over recent years, individuals who have been adequately treated and assumed to have developed immunity against TB, are in fact still at risk of being infected [84, 91]. They can develop pulmonary recurrent TB, which is defined as an episode of TB following recovery from a previous episode. Recent estimates of the rates of recurrent TB across various regions point to a mean of 2300 cases per 100,000 person-years at twelve months after treatment. The recurrent TB rate can be significantly higher in high-incidence TB regions, with an average TB recurrence rate reaching 7850 per 100,000 person-years [92]. With modern technology, especially DNA fingerprinting techniques, it is possible to reveal whether a new episode of tuberculosis is a result of infection with the same strain as the previous one or is due to a different strain. A study conducted in areas of South Africa where TB incidence is high found that the majority of TB cases (approximately 77%) occur as a result of reinfection [84]. Furthermore, in the same study the rate of recurrent TB was found to be four times higher than that of new TB. This raised an important question regarding the underlying mechanisms [91]. Although, worldwide about 10-30% [84] of cases of TB are recurrent, the role of recurrent TB as far as the formation of backward bifurcation is concerned, is rarely if ever studied. Hence, the main objective in this research question will be to analyse how recurrent tuberculosis impacts the formation of backward bifurcation. An epidemiological model that provides a comprehensive description of the transmission pathways involved for recurrent tuberculosis, whereby cured individuals can become re-infected is proposed. The model incorporates progressive primary infection, exogenous reinfection, endogenous reactivation and recurrent TB as transmission mechanisms. Moreover, unlike other studies of TB dynamics that make use of frequency-dependent
transmission rates [56], the analysis conducted in this thesis provides exact analytic backward bifurcation threshold conditions without resorting to commonly applied approximations and simplifying assumptions. Lastly, by switching off reinfection of latently infected individuals it shall be shown that recurrent TB can independently induce backward bifurcation phenomena if it exceeds a certain threshold.

The TB model studied in research question two is extended by splitting latently infected individuals into two subgroups and used to investigate questions relating to heterogeneity in reinfection rates (i.e., cases (iii) and (iv)). Through clinical observation it has been revealed that variable rates of progression to active TB following infection exist [93]. Evidence, from the interferon-gamma release assay indicate that over the first 23 months following infection about 12.9% of infected individuals progress to active TB [93]. In contrast, after the high risk period the rate at which active TB occurs is relatively low and is approximately 5-10% over 20 years [94]. To account for this clinical observation past mathematical models devoted to tracking TB dynamics have incorporated two major pathways from susceptible to actively infected: fast and slow TB pathways. In such models a fraction of exposed susceptibles progress directly to the infective stage passing the latency compartment [56, 94, 95, 96]. Thus, a tuberculosis model that accounts for heterogeneity in host susceptibility by treatment status is proposed. The new model relaxes the assumption made in [97], that individuals treated with preventive therapy and recovered individuals previously treated from active TB are reinfected at the same rate. As will be shown in chapter 5, disregarding the assumption made in [97] leads to new epidemiological insights that are vital in deciphering the impact of treatment under different levels of susceptibility to reinfection.

The fifth research question investigates the implication of a nonlinear treatment rate in a heroin addiction model. A mathematical model is developed that examines how heroin addiction spreads in society. The model has similarities to epidemic equations for the spread of infectious diseases. The proposed model takes into account the treatment of heroin users by incorporating a realistic functional form that “saturates”, representing the limited availability of treatment. In these circumstances, bifurcation analysis reveals that the model has an intrinsic backward bifurcation whenever the saturation parameter is larger than a fixed threshold. The main objective of this model is to determine its stability properties. In the absence of backward bifurcations, Lyapunov functions can often be found and used to prove global stability. However,
in the presence of backward bifurcations, such Lyapunov functions may not exist or may be difficult to construct. Consequently, a geometric approach to global stability proposed in [98] is applied to provide conditions that ensure the system is globally asymptotically stable. Furthermore, numerical simulations are performed to verify theoretical findings.

The organization of the chapters in the thesis is as follows. Mathematical preliminaries that will be used in the thesis for deriving some important results are described in chapter 2. Two simple models for tuberculosis that will be used to investigate research question one are formulated in chapter 3. The models will be used to investigate problems that arise with aggregation of bifurcation parameters in epidemic models. In this chapter the equilibrium points are the main interest and therefore they shall be obtained under two scenarios: (i) a model with density-dependent incidence and (ii) a model with frequency-dependent incidence rates. A proof for the existence of backward bifurcation shall be given using a Center Manifold approach as developed by [58]. In chapter 4 the simple TB model in chapter 3 with frequency-dependent incidence is extended to incorporate new infection processes such as primary progression and recurrent TB. This model shall be used to answer question three which investigates the implication of recurrent TB in tuberculosis epidemics models as far as the phenomenon of backward bifurcation is concerned. Chapter 5 will be an extension of the model studied in chapter 4. The latently infected population is subdivided into two subgroups to represent high and low risk latent individuals. The high risk latent individuals shall be referred to as early latent, while the low risk individuals shall be referred to as late latent. Both early and late latent individuals shall be treated with preventive therapy. Thus, a new compartment shall be added to account for individuals who have received preventive therapy. To investigate the effect of heterogeneity, late latent individuals treated with preventive therapy and recovered individuals previously treated from active TB will be assumed to be susceptible to exogenous reinfection. Chapter 6 contains related work which involves stability analysis of a heroin addiction model with a non-linear treatment rate. The heroin model is based on an SIR epidemic model and thus has many features in common, including backward bifurcations. Finally, chapter 7 will outline the main contribution of this thesis as well as point out new directions to explore.
1.5 Published work

The work contained in some of the chapters in this thesis was submitted for publication, thus resulting in the following publications:

► Chapter 3 was published and can be accessed online from:

https://doi.org/10.1016/j.apm.2015.07.022;

► Chapter 4 was published and can be accessed online from:

https://doi.org/10.1371/journal.pone.0194256;

► Chapter 5 was published and can be accessed online from:

doi:10.1371/journal.pone.0206603;

► Chapter 6 includes other published work and can be accessed online from:

https://dx.doi.org/10.1155/2017/1953036.
Chapter 2

Mathematical preliminaries

This chapter is a summary of important mathematical concepts and methodologies that are applied throughout this thesis. The concepts given are mostly standard definitions and results as documented in the literature on mathematical theory. However, a few concepts are adopted from the readily available theory of mathematical modelling of infectious diseases.

2.1 Linear and non-linear systems equilibria

Given a system of ODEs (ordinary differential equations) as shown below:

\[ \dot{y} = g(y, t; \mu), \quad y \in U \subset \mathbb{R}^n, \quad t \in \mathbb{R}, \quad \mu \in V \subset \mathbb{R}^p, \]  

(2.1)

where \( U \) and \( V \) are respectively, open sets in \( \mathbb{R}^n \) and \( \mathbb{R}^p \), and \( \mu \) is a parameter. Then the right-hand side function \( g(y, t; \mu) \) of equation (2.1) is referred to as a vector field. ODEs which do not explicitly depend on time are referred to as autonomous differential equations while those that are explicitly dependent on time are called non-autonomous differential equations.

Consider the following general autonomous system

\[ \dot{y} = g(y), \quad y \in \mathbb{R}^n, \]  

(2.2)

then the following definitions, theorems and lemmas are stated:

Definition 2.1.1 The autonomous system (2.2) equilibrium solution is given by \( y = \bar{y} \in \mathbb{R}^n \), where \( g(\bar{y}) = 0 \). The vector or point \( \bar{y} \) is referred to as an equilibrium point.
Theorem 2.1.1 (Perko [99]). Fundamental existence-uniqueness theorem. Suppose $D$ is an open subset of $\mathbb{R}^n$ containing $y_0$ and assume that $g \in C^1(D)$. Then there exists $c > 0$ such that the initial value problem (IVP)

$$\dot{y} = g(y), \quad y(0) = y_0,$$

has a unique solution $y(t)$ on the interval $[-c, c]$.

Lemma 2.1.1 (Perko [99]). Let $D$ be an open subset of $\mathbb{R}^n$ and let $g : D \mapsto \mathbb{R}^n$. Then, if $g \in C^1(D)$, $g$ is locally Lipschitz on $D$.

Definition 2.1.2 The Jacobian matrix of $g$ at the equilibrium point $\bar{y}$, denoted by $Dg(\bar{y})$, is the matrix of partial derivatives of $g$ evaluated at $\bar{y}$ and can be obtained as

$$Dg(\bar{y}) = \begin{bmatrix} \frac{\partial g_1(\bar{y})}{\partial y_1} & \cdots & \frac{\partial g_m(\bar{y})}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m(\bar{y})}{\partial y_1} & \cdots & \frac{\partial g_m(\bar{y})}{\partial y_m} \end{bmatrix}. $$

Definition 2.1.3 Suppose $y = \bar{y}$ is an equilibrium solution of (2.2). Then $\bar{y}$ is called hyperbolic if none of the eigenvalues of $Dg(\bar{y})$ has zero real part. An equilibrium point that is not hyperbolic is called non-hyperbolic.

2.2 Stable, unstable solutions and bifurcations

Here the definitions and theorems that are relevant to analyzing the stability of an autonomous system are stated. Letting $\bar{y}(t)$ be any solution of the general autonomous system (2.2), then $\bar{y}(t)$ is considered stable if solutions starting “close” to $\bar{y}(t)$ at a given time remain close to $\bar{y}(t)$ for all later times. It is said to be asymptotically-stable if close solutions are attracted to $\bar{y}(t)$ as $t \to \infty$.

Definition 2.2.1 (Wiggins [100]). Let $\epsilon > 0$, then a solution $\bar{y}(t)$ is considered to be stable if there exists a $\delta = \delta(\epsilon) > 0$, such that for any solution $x(t)$ of (2.2), $|\bar{y}(t_0) - x(t_0)| < \delta$, $|\bar{y}(t) - x(t)| < \epsilon$ for every $t > t_0$, $t_0 \in \mathbb{R}$.

Definition 2.2.2 (Wiggins [100]). The solution $\bar{y}(t)$ is considered to be asymptotically stable if it has the following properties: (i) it is stable and (ii) a constant $c > 0$ exists such that any solution $x(t)$ of (2.2) fulfills conditions $|\bar{y}(t_0) - x(t_0)| < c$ and $\lim_{t \to \infty} |\bar{y}(t) - x(t)| = 0$. 
2.3 Method for computing $R_0$ (next generation operator method)

The basic reproduction number $R_0$ is defined as the average number of new infections that a typical infected individual gives rise to over the lifetime of the infection when introduced into a wholly susceptible population. van den Driessche and Watmough [25] developed a technique for calculating the basic reproduction number of disease transmission models. This method is called the next generation operator (NGO) method. The technique is also becoming increasingly popular in establishing the local asymptotic stability of the associated disease free equilibrium. Here the procedure as given in [41] is briefly outlined below.

Consider a disease transmission model that has non-negative initial conditions and can be expressed in terms of the following autonomous system:

$$
\dot{y}_i = f_i(y) = F_i(y) - V_i(y), \quad i = 1, \ldots, n,
$$

(2.3)

where $V_i = V_i^- - V_i^+$ and the functions fulfill the properties (A1)-(A5) stated below. In terms of disease transmission modelling $F_i(y)$ represent the rate at which new infections occur in compartment $i$, $V_i^+(y)$ represent the rate at which individuals flow into compartment $i$ by all other means and finally $V_i^-(y)$ represent the rate at which individuals flow out of compartment $i$ [25]. The functions $F_i(y), V_i^+(y), V_i^-(y)$ are assumed to be at least twice continuous-differentiable in each variable [25].

First note that $Y_s = \{y \geq 0|y_i = 0, \quad i = 1, \ldots, m\}$ define the set of all disease free states (disease free) of the model, where $y = (y_1, \ldots, y_n)^T$, $y_i \geq 0$ account for the number of individuals residing in respective compartments of the model. Properties
2.3 Method for computing $R_0$ (next generation operator method)

(A1)-(A5) are:

(A1) If $y_i \geq 0$, then $F_i, V_i^+, V_i^- \geq 0$ for $i = 1, \ldots, m$;
(A2) $F_i = 0$ if $i > m$;
(A3) If $y_i = 0$, then $V_i^- = 0$. That is if $y \in Y_s$ then $V_i^- = 0$ for $i = 1, \ldots, m$;
(A4) If $y \in Y_s$, then $F_i(y) = 0$ and $V_i^+(y) = 0$ for $i = 1, \ldots, m$;
(A5) If $F(y)$ is set to zero, then all eigenvalues of $DF(\bar{y})$ have negative real part.

**Definition 2.3.1 (Plemmons [101]).** M-Matrix. An $n \times n$ matrix $A$ is an M-Matrix if and only if every off-diagonal entry of $A$ is non-positive, the diagonal entries are all positive and the real parts of the eigenvalues of the matrix $A$ are non-negative.

**Lemma 2.3.1 (van den Driessche and Watmough [25]).** If $\bar{y}$ is the DFE of (2.3) and $f_i(y)$ fulfills conditions (A1)-(A5) then the derivatives $DF(\bar{y})$ and $DV(\bar{y})$ are partitioned as

$$DF(\bar{y}) = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix}, \quad DV(\bar{y}) = \begin{bmatrix} V & 0 \\ J_3 & J_4 \end{bmatrix}$$

where $F$ and $V$ are $m \times m$ matrices and can be obtained as

$$F = \left[ \frac{\partial F_i(\bar{y})}{\partial y_j} \right] \quad \text{and} \quad V = \left[ \frac{\partial V_i(\bar{y})}{\partial y_j} \right] \quad \text{with} \quad 1 \leq i, j \leq m.$$ 

Note that $F$ is non-negative and $V$ is a non-singular M-Matrix. $J_3$ and $J_4$ are matrices associated with the transition terms of the model and all eigenvalues of $J_4$ have positive real parts. Considering the above discussion and formulations the reproduction number is given as

$$R_0 = \rho(FV^{-1}),$$

where $\rho$ denotes the spectral radius. Consequently, Theorem 2.3.1 follows:

**Theorem 2.3.1 (van den Driessche and Watmough [25]).** Consider the system (2.3) representing disease transmission, where $f_i(y)$ fulfills the stated axioms (A1)-(A5). If $\bar{y}$ is a DFE of the model then $\bar{y}$ is locally asymptotically stable whenever $R_0 = \rho(FV^{-1}) < 1$ and unstable whenever $R_0 > 1$. 

2.4 Bifurcations

Typically all systems of equations are characterized by a set of variables and parameters. The qualitative structure of the solution of the system may be altered when certain parameter values within the system are varied within a given interval. These changes are called bifurcations. The parameter value where the bifurcation occurs is called a bifurcation value (or bifurcation point). The definition of a bifurcation can be given as follows:

Definition 2.4.1 (Gros [102]). Let
\[
\dot{y} = g(y; \mu), \quad y \in \mathbb{R}, \quad \mu \in \mathbb{R}
\] (2.4)
be a one-parameter family of one-dimensional ODEs. An equilibrium solution of (2.4) given by \((y, \mu) = (0, 0)\) is said to undergo bifurcation at \(\mu = 0\) if the flow for \(\mu\) near zero and \(y\) near zero is not qualitatively the same as the flow near \(y = 0\) at \(\mu = 0\).

There are several types of bifurcations in dynamical systems, including saddle-node, transcritical, pitchfork, backward, Bogdanov-Takens and Hopf bifurcations as described in [100]. Backward and forward bifurcations are pertinent to this thesis and have been described in chapter 1. Other types of bifurcations such as forward bifurcation with a hysteresis loop, sometimes referred to as the hysteresis effect are very rare in the existing literature on mathematical modelling of infectious diseases. Currently, there is an available theorem (as deduced in [58]) that is used to establish the existence of backward bifurcation phenomena. This theorem shall be applied in chapters 3, 4 and 5 for both simple and slightly complex models. Hence, it is important to state the theorem:

Theorem 2.4.1 (Castillo-Chavez and Song [58]). Consider the following general system of ordinary differential equations with a parameter \(\varphi\):
\[
\frac{dx}{dt} = f(x, \varphi), \quad f : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n \text{ and } f \in C(\mathbb{R}^n \times \mathbb{R}),
\] (2.5)
where \(0\) is an equilibrium point of the system (that is, \(f(0, \varphi) \equiv 0\) for all \(\varphi\)) and assume:

(i) \(A = D_x f(0, 0) = \left( \frac{\partial f}{\partial x} \right)(0, 0)\) is the linearization matrix of the (2.5) around the equilibrium \(0\) with \(\varphi\) evaluated at \(0\). Zero is a simple eigenvalue of \(A\) and other eigenvalues of \(A\) have negative real parts;
(ii) Matrix A has a right eigenvector \( w \) and a left eigenvector \( v \) (each corresponding to the zero eigenvalue).

Let \( f_k \) be the \( k \)th component of \( f \) and

\[
\begin{align*}
    a &= \sum_{k,i,j=1}^{n} v_k w_i w_j \frac{\partial^2 f_k}{\partial x_i \partial x_j}(0,0), \\
    b &= \sum_{k,i=1}^{n} v_k w_i \frac{\partial^2 f_k}{\partial x_i \partial \varphi}(0,0).
\end{align*}
\]

Then, the local dynamics of the system (2.5) around 0 are determined by the signs of \( a \) and \( b \):

(i) \( a > 0, b > 0 \). When \( \varphi < 0 \) with \( |\varphi| \ll 1 \), \( 0 \) is locally asymptotically stable and there exists a positive unstable equilibrium; when \( 0 < \varphi \ll 1 \), \( 0 \) is unstable and there exists a negative, locally asymptotically stable equilibrium;

(ii) \( a < 0, b < 0 \). When \( \varphi < 0 \) with \( |\varphi| \ll 1 \) \( 0 \) is unstable; when \( 0 < \varphi \ll 1 \), \( 0 \) is a locally asymptotically stable equilibrium, and there exists a positive unstable equilibrium;

(iii) \( a > 0, b < 0 \). When \( \varphi < 0 \) with \( |\varphi| \ll 1 \) \( 0 \) is unstable, and there exists a locally asymptotically stable negative equilibrium; when \( 0 < \varphi \ll 1 \), \( 0 \) is stable, and a positive unstable equilibrium appears;

(iv) \( a < 0, b > 0 \). When \( \varphi \) changes from negative to positive, \( 0 \) changes its stability from stable to unstable. Correspondingly a negative unstable equilibrium becomes positive and locally asymptotically stable.

In particular, if \( a > 0 \) and \( b > 0 \), then a backward bifurcation occurs at \( \varphi = 0 \).
2.5 Lyapunov functions and Lasalle’s invariance principle

Definition 2.5.1 A function $V : \mathbb{R} \to \mathbb{R}$ is said to be a positive-definite function if:

(i) $V(y) > 0$ for all $y \neq 0$;
(ii) $V(y) = 0$ if and only if $y = 0$.

Definition 2.5.2 Consider the following system

$$\dot{y} = g(y), \quad y \in \mathbb{R}^n. \quad (2.6)$$

Let $\bar{y}$ be an equilibrium solution of (2.6) and assume $V : U \to \mathbb{R}$ is a $C^1$ function defined on some neighbourhood $U$ of $\bar{y}$ such that:

(a) $V$ is positive definite;
(b) $\dot{V} \leq 0$ in $U \setminus \{\bar{y}\}$.

Then any function $V$ which satisfies the conditions (a) and (b) is called a Lyapunov function [100, 103].

Now the general Lyapunov function theorem is given as below:

Theorem 2.5.1 (Lasalle’s Invariance Principle [103]). Consider the system (2.6). Let $S = \{y \in U : \dot{V}(y) = 0\}$ and suppose $M$ is the largest invariant set of (2.6) in $S$. Now if $V$ is a Lyapunov function on $U$ and $\gamma^+(y_0)$ is a bounded orbit of (2.6) which lies in $S$, then the $\omega$-limit set of $\gamma^+(y_0)$ belongs to $M$ (that is $y(t,y_0) \to M$ as $t \to \infty$).

Corollary 2.5.1 If $V(y) \to \infty$ as $|y| \to \infty$ and $\dot{V} \leq 0$ on $\mathbb{R}^n$, then every solution of (2.6) is bounded and approaches the largest invariant set $M$ of (2.6) in the set where $V = 0$. To be precise if $M = \{0\}$, then the solution $y = 0$ is globally asymptotically stable (GAS).

Theorem 2.5.2 (Wiggins [100]). Suppose there is a continuously differentiable positive-definite and radially unbounded function $V : \mathbb{R}^n \to \mathbb{R}$, such that

$$\frac{\partial V}{\partial y}(y - \bar{y}).g(y) = \nabla V(y - \bar{y}).g(y) \leq W(y) \leq 0 \quad \forall \ y \in \mathbb{R}^n,$$
where $W(y)$ is any continuous function on $U$. Then $\bar{y}$ is said to be a globally stable equilibrium. The solution $y(t)$ converges to the largest invariant set $S$ contained in $E = \{x \in \mathbb{R}^n : W(y) = 0\}$. 
Chapter 3

Backward bifurcation in epidemic models: problems arising with aggregated bifurcation parameters

3.1 Chapter overview

Epidemiological models have become important tools for helping understand the qualitative dynamics controlling the spread of infectious diseases. Many of these models have strong non-linearities and therefore exhibit complex population dynamics and possess subtle bifurcation properties. Currently, there is a renewed interest in so-called “backward bifurcations” because of the unusual thresholds they introduce. In this chapter an overlooked problem that has arisen in the literature when calculating backward bifurcations, especially in the context of epidemic modelling is examined and resolved. In mathematical modelling a range of different epidemiological models have been found to exhibit backward bifurcation, including models that incorporate behavioural responses to perceived risks [30], vaccination [42, 104], multiple groups [29], vector-borne diseases [63] and exogenous reinfection [52, 56, 57, 58, 59] (just to mention a few). The presence of backward bifurcation is important in a practical sense because control programs must reduce the basic reproduction number, $R_0$, further than below unity to eliminate a disease. Thus, the problem addressed in this chapter stems from studies of backward bifurcation in the literature where there have been instances where authors illustrate the phenomena by varying $R_0$ without properly considering the fact that $R_0$ is an aggregate of parameters in the model. Hence, it would be incorrect to
3.2 Model construction

simply vary the aggregate $R_0$ while leaving all model parameters constant as has been
the practice in a number of important studies [21, 48, 52, 63, 85, 86, 87, 88, 89, 90].
In this chapter the two scenarios investigated include:

(i) For the incorrect approach, all parameters in the aggregate $R_0$ are fixed to
constant values, but $R_0$ is nevertheless varied as a bifurcation parameter;

(ii) In the correct approach, a key parameter in $R_0$ is allowed to vary, and hence $R_0$
itself varies and acts as a natural bifurcation parameter.

The main results will show how the outcomes of these two approaches are substantially
different. To explore the aforementioned problem a simple tuberculosis (TB) model that
incorporates reinfection as the parameter to induce backward bifurcation is examined,
although any other example can be used to exhibit the difference.

3.2 Model construction

The total population is partitioned into four sub-populations: susceptible (S); ex-
posed/latent individuals (E) (those who have been infected but do not manifest TB
symptoms and are not capable of infecting others); infectious individuals (I) (symp-
tomatic and infectious) and recovered (R) individuals (recovery due to treatment or
spontaneous cure). The total population denoted by $N(t)$ is given by

$$N(t) = S(t) + E(t) + I(t) + R(t).$$

The numbers of susceptible individuals increase by recruitment through births and
immigration at a rate $\Lambda$. Susceptibles who come into contact with infected individuals
move straight to the exposed E class but they are not themselves yet infective. The
susceptible population is thus diminished due to contact with infected individuals
at a rate $\beta SI$, where $\beta$ represents the per-capita effective contact rate of acquiring
TB bacteria. Concomitantly, the numbers in the exposed class increase at a rate
$\beta SI$. Progression to the infectious state occurs when an exposed individual harbors a
dormant infection that becomes active due to immune system destabilization. This
is the usual “slow TB” which can take years or decades before progression. Exposed
individuals move to the infected class I at rate $kE$. In addition, exposed individuals
can encounter infectious individuals (I) and be reinfected leading to an acceleration
into the infectious class at rate $p\beta EI$. The infected subpopulation is diminished when
individuals recover from TB due to treatment or spontaneous cure at rate $rI$ and disease induced death rate $\mu_d$. Finally, the recovered sub-population (R) is generated by recovery of infected individuals (at rate $rI$). The natural death rate decreases all classes at the same rate via the background mortality parameter $\mu$. A brief description of the model parameters and variables is given in Table 3.1 while a flow diagram is depicted in Figure 3.1.

Table 3.1: Description of model parameters and variables of model equation (3.1).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$</td>
<td>Population of susceptible individuals</td>
</tr>
<tr>
<td>$E$</td>
<td>Population of exposed (asymptomatic and non-infectious) individuals</td>
</tr>
<tr>
<td>$I$</td>
<td>Population with active TB (symptomatic and infectious)</td>
</tr>
<tr>
<td>$R$</td>
<td>Recovered individuals</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Per capita natural mortality rate</td>
</tr>
<tr>
<td>$\Lambda$</td>
<td>Recruitment rate into the population</td>
</tr>
<tr>
<td>$\mu_d$</td>
<td>Per capita disease-induced death rate</td>
</tr>
<tr>
<td>$k$</td>
<td>Per capita endogenous reactivation rate</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Transmission coefficient</td>
</tr>
<tr>
<td>$r$</td>
<td>Per capita recovery rate</td>
</tr>
<tr>
<td>$p$</td>
<td>Reinfection factor</td>
</tr>
</tbody>
</table>

Figure 3.1: Schematic diagram for the model equation (3.1). The red curved dashed arrow indicates the reinfection process.
The basic model governing change of status from one compartment to another is given by the following system of non-linear differential equations

\[
\begin{align*}
\frac{dS}{dt} &= \Lambda - \beta SI - \mu S, \\
\frac{dE}{dt} &= \beta SI - p\beta EI - (\mu + k)E, \\
\frac{dI}{dt} &= p\beta EI + kE - (\mu + r + \mu_d)I, \\
\frac{dR}{dt} &= rI - \mu R.
\end{align*}
\]

\[\text{(3.1)}\]

### 3.3 Mathematical analysis

#### 3.3.1 Basic properties

The model equation (3.1) monitors TB transmission dynamics in a human population, thus all the associated parameters are non-negative. Hence, the following non-negative result needs to be satisfied.

**Theorem 3.3.1** The state variables of the model are non-negative for all time. That is all the trajectories of the model equations with non-negative initial data will remain non-negative for all time \(t > 0\).

**Proof** Let \(t_1 = \sup\{t > 0 : S > 0, E > 0, I > 0, R > 0\}\). Now for \(t_1 > 0\) it follows from the first equation of model (3.1) that \(\frac{dS}{dt} = \Lambda - (\Psi + \mu)S\) (where \(\Psi = \beta I\)) which can be written as

\[
\frac{dS(t)}{dt} + (\Psi + \mu)S(t) = \Lambda > 0.
\]

\[\text{(3.2)}\]

The solution of (3.2) can be shown to be non-negative. Thus, rewriting (3.2) as

\[
\frac{d}{dt} \left[ S(t) \exp \left( \mu t + \int_0^t \Psi(\tau)d\tau \right) \right] = \Lambda \exp \left[ \mu t + \int_0^t \Psi(\tau)d\tau \right] > 0,
\]

\[\text{(3.3)}\]

and integrating both sides of (3.3) leads to

\[
S(t_1) \exp \left[ \mu t_1 + \int_0^{t_1} \Psi(\tau)d\tau \right] - S(0) = \int_0^{t_1} \Lambda \exp \left( \mu x + \int_0^x \Psi(\tau)d\tau \right) dx > 0.
\]

\[\text{(3.4)}\]
which can be rewritten as
\[
S(t_1) = S(0) \exp \left\{ -\mu t_1 - \int_0^{t_1} \Psi(\tau) d\tau \right\} \\
+ \left\{ \exp \left( -\mu t_1 - \int_0^{t_1} \Psi(\tau) d\tau \right) \right\} \int_0^{t_1} \Lambda \exp \left( \mu x + \int_0^x \Psi(\tau) d\tau \right) dx > 0.
\]
Thus, \( S(t_1) > 0 \). Similarly it can be shown that \( E(t) > 0 \), \( I(t) > 0 \) and \( R(t) > 0 \) for all \( t > 0 \). Theorem (3.3.1) can also be proven by applying the approach in Appendix A of [105].

**Theorem 3.3.2** The closed set
\[
\kappa = \left\{ (S(t), E(t), I(t), R(t)) \in \mathbb{R}^4_+ : S + E + I + R \leq \frac{\Lambda}{\mu} \right\}
\]
is positively invariant and attracting.

**Proof.** Summing all the equations of model system (3.1) yields
\[
\frac{dN}{dt} = \Lambda - \mu N - \mu_d I. \quad (3.5)
\]
Supposing \( \frac{dN}{dt} \leq \Lambda - \mu N \), then \( \frac{dN}{dt} \leq 0 \) if \( N \geq \frac{\Lambda}{\mu} \).
Using the standard comparison theorem [106] it can be shown that
\[
N(t) \leq N(0)e^{-\mu t} + \frac{\Lambda}{\mu}(1 - e^{-\mu t}). \quad (3.6)
\]
Note that \( N(t) \leq \frac{\Lambda}{\mu} \) if \( N(0) \leq \frac{\Lambda}{\mu} \). Consequently, the region \( \kappa \) is positively invariant. Moreover, for \( N(0) > \frac{\Lambda}{\mu} \) the solution \( N(t) \) may enter in the region \( \kappa \) infinitely many times or asymptotically approaches \( \frac{\Lambda}{\mu} \). Thus, all the solutions in \( \mathbb{R}^4_+ \) converge in the region \( \kappa \). Since the region \( \kappa \) is positively invariant the dynamics of all the trajectories generated by the model (3.1) are considered in \( \kappa \).

The disease free equilibrium of model (3.1) can be obtained by setting the right-hand terms of the model equation (3.1) to zero as
\[
\varepsilon_0 = (S_0, E_0, I_0, R_0) = \left( \frac{\Lambda}{\mu}, 0, 0, 0 \right).
\]
Following the next generation operator method (given in chapter 2), the local stability of the disease free equilibrium will be explored. Using similar notation as in [25], the non-negative matrix $F$ (representing new infections) and M-matrix $V$ (representing inflow and outflow) can respectively be obtained from model equation (3.1) as

$$ F = \begin{bmatrix} 0 & \frac{\beta \Lambda}{\mu} \\ 0 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} \mu + k & 0 \\ -k & \mu + r + \mu_d \end{bmatrix}. $$

Hence, it follows that the basic reproduction number is

$$ R_0 = \rho(FV^{-1}) = \frac{k\beta \Lambda}{\mu(\mu + k)(\mu + r + \mu_d)}, $$

where $\rho$ is the spectral radius.

The reproduction number $R_0$ measures the average number of new infections a single infected individual will generate when introduced into a wholly susceptible population.

**Lemma 3.3.1** The DFE of model equation (3.1) is locally asymptotically stable whenever $R_0 < 1$ and unstable whenever $R_0 > 1$.

The result deduced in Lemma 3.3.1 was established and proved in [25] and thus the proof is not repeated here. Lemma 3.3.1 implies that TB can be eliminated from the community if the initial sizes of the sub-populations of the model are in the basin of attraction of the disease free equilibrium $\varepsilon_0$. It is important to observe that the parameter $p$ which accounts for reinfection is not included in the basic reproduction number, yet the mechanism does increase the population of infectives. Thus, further analysis is needed to elucidate the role played by the reinfection process. Now the endemic equilibrium of model equation (3.1) and the associated bifurcation structure is established.

### 3.3.2 Existence of backward bifurcation

It is imperative to note that the precursor for the existence of the backward bifurcation is existence of multiple equilibria (both stable and unstable) when the associated threshold quantity $R_0$ is below unity. Let $\varepsilon_* = (S^*, E^*, I^*, R^*)$ be an arbitrary equilibrium of
model equation (3.1). Now equating the right-hand terms of model equation (3.1) to zero and solving leads to

\[ S^* = \frac{\Lambda}{\mu + \beta I^*}, \quad E^* = \frac{(\mu + r + \mu_d)I^*}{p\beta I^* + k}, \quad R^* = \frac{rI^*}{\mu}, \]

in terms of the number of infectives \( I^* \), where \( I^* \) can be obtained by solving the quadratic expression

\[ f(I^*) = c_2 I^{*2} + c_1 I^* + c_0 = 0, \quad (3.7) \]

where

\[ c_2 = p\beta^2(\mu + r + \mu_d), \]
\[ c_1 = (\mu + k + \mu p)(\mu + r + \mu_d)\beta - \beta^2 p\Lambda, \]
\[ c_0 = \mu(\mu + k)(\mu + r + \mu_d)(1 - R_0). \]

The endemic equilibria of model equation (3.1) are summarized through the following Theorem 3.3.3:

**Theorem 3.3.3** The model equation (3.1) has:

(i) A unique endemic equilibrium if \( c_1 < 0 \) and \( c_0 = 0 \) or the discriminant \( \Delta = c_1^2 - 4c_2c_0 = 0 \);

(ii) A unique endemic equilibrium if \( c_0 < 0 \);

(iii) Two positive endemic equilibria if \( c_1 < 0, c_0 > 0 \) and \( c_1^2 - 4c_2c_0 > 0 \);

(iv) No endemic equilibrium if \( c_1 > 0 \) and \( c_0 > 0 \).

**Proof.** It is easy to note that in the polynomial (4.12) \( c_2 \) is always positive and \( c_0 > 0 \) if \( R_0 < 1 \). For case (i) where \( c_1 < 0 \) and \( R_0 = 1 \) (i.e., \( c_0 = 0 \)) the quadratic equation \( f(I^*) \) reduces to \( c_2 I^* + c_1 = 0 \) and in this case the model equation (3.1) will have a unique positive endemic equilibrium if \( c_1 < 0 \) and no positive non-trivial equilibrium if \( c_1 \geq 0 \). From case (ii) where \( c_0 < 0 \) (that is \( R_0 > 1 \)) a unique endemic equilibrium exists since there is only one change of sign according to Descartes’s Rule of Signs. However, for case (iii) where \( c_1 < 0 \) and \( R_0 < 1 \) there is exactly two changes of sign indicating the existence of two non-trivial equilibria. For case (iv) where
3.3 Mathematical analysis

c_1 > 0 and R_0 < 1 there are no changes of sign, thus there is no endemic equilibrium in such a case. Hence, from case (iii) it is concluded that model equation (3.1) has a maximum of two endemic equilibria (I^*_1,2) when \( R_0 < 1, c_1 < 0 \) and \( c_1^2 - 4c_2c_0 > 0 \).

The above equilibrium analysis suggests a possibility of backward bifurcation since two endemic equilibria exist as exhibited by case (iii) when \( R_0 < 1 \), which is actually a necessary criterion for the occurrence of backward bifurcation phenomena.

The problem arises when retrieving these two non-trivial solutions (I^*_1,2) as a function of the aggregate parameter \( R_0 \). A number of studies [21, 48, 52, 63, 85, 86, 87, 88, 89, 90] incorrectly plot the solutions (I^*_1,2) by varying \( R_0 \) in equation (4.12), that is, without varying any parameter within \( R_0 \). Such a practice leads to an incorrect backward bifurcation diagram yielding a simple (though incorrect) parabolic shape such as shown in Figure 3.2. The correct approach for drawing the bifurcation diagram requires first choosing a proper specific model bifurcation parameter to vary, say the transmission rate \( \beta \). The bifurcation diagram can then be determined through varying this model bifurcation parameter. Once obtained, the bifurcation diagram can then be rescaled so that the x-axis is given in terms of the aggregated parameter \( R_0 \). For the particular model above, by varying the transmission rate \( \beta \) in the interval \( \beta \in [0.025, 0.175] \), Figure 3.3 is obtained. Using this approach the figure no longer appears parabolic in shape and has shifted to the left compared to Figure 3.2. Another outstanding difference between Figures 3.2 and 3.3 is the gap between the bifurcation branches. From 3.3 the gap between the bifurcation curves is wider than for Figure 3.2. This implies that the endemic equilibrium predicted when \( \beta \) is varied is higher than for the case where all parameters are kept constant i.e., the incorrect approach. Also, one can observe that the critical value of the basic reproduction number denoted by \( R_c \) where the backward bifurcation initiates (defined in more detail below) will be incorrect if estimated using the aggregated parameter method. (Hence, in the subsequent section comparison of the two methods, incorrect and correct approaches are analysed.)
3.3 Mathematical analysis

Figure 3.2: Backward bifurcation when all the parameters in $R_0$ are fixed at constant values but the aggregated $R_0$ is nonetheless varied. Parameters used are $\Lambda = 10,$ $\mu = 0.016,$ $r = 2,$ $\mu_d = 0.4,$ $k = 0.0005,$ $p = 0.2,$ $\beta = 0.036.$ Black solid line represents stable equilibria while the red solid line represents unstable equilibria.
Figure 3.3: Backward bifurcation when one parameter $\beta$ in $R_0$ is varied. Parameters used are $\Lambda = 10$, $\mu = 0.016$, $r = 2$, $\mu_d = 0.4$, $k = 0.0005$, $p = 0.2$, $\beta \in [0.025, 0.175]$. Black solid lines represent stable equilibria while the red solid line represents an unstable equilibrium.
3.4 The backward bifurcation threshold, $R_c$

As seen in the bifurcation diagram of Figure 3.3, there is a threshold quantity $R_c$ which is the value of $R_0$ where the two non-trivial endemic equilibria collide and annihilate each other, leaving only the disease free equilibria as the only stationary solution. For instance in Figure 3.3, this occurs at the threshold $R_c = 0.42$. If $R_0 < R_c$ then the only model equilibrium is the stable disease free equilibrium. Now the two approaches for computing $R_c$ [21, 86, 87, 88, 107] are explored.

(i) Incorrect aggregated parameter approach

Recall that this approach wrongly assumes that all parameters in $R_0$ are kept constant while $R_0$ may be varied to obtain the backward bifurcation diagram, as in Figure 3.2. If the discriminant ($\Delta = c_1^2 - 4c_2c_0$) of equation (4.12) is set to zero it is possible to obtain the critical point $R_{c1}$. This is just the value of $R_0$ where the stable and unstable endemic curves coincide, namely

$$R_{c1} = 1 - \frac{c_1^2}{4c_2\Omega}$$

where $\Omega = \mu(\mu + k)(\mu + r + \mu_d)$. Using the above equation, it is possible to investigate how the parameters that induce backward bifurcation affect $R_{c1}$. For example, a plot of $R_{c1}$ as a function of $p$ is shown in Figure 3.4.

(ii) Correct approach

The correct value for $R_c$ is obtained by first selecting a specific bifurcation parameter of the model. For illustration the parameter $\beta$ is chosen. By setting the discriminant $\Delta = 0$ and rearranging for the critical transmission rate $\beta_c$ yields

$$\beta_c = \frac{-\phi_1 \pm \sqrt{\phi_1^2 - 4\phi_2\phi_0}}{2\phi_2},$$

where

$$\begin{align*}
\phi_2 &= p^2\Lambda^2, \\
\phi_1 &= 4p\Lambda k(\mu + r + \mu_d) - 2p\Lambda(\mu + k + \mu p)(\mu + r + \mu_d), \\
\phi_0 &= (\mu + k + \mu p)^2(\mu + r + \mu_d)^2 - 4\mu p(\mu + r + \mu_d)^2(\mu + k).
\end{align*}$$
Replacing $\beta$ in $R_0$ with $\beta_c$ yields

$$R_{c2} = \left( \frac{k}{\mu(\mu + k)(\mu + r + \mu_d)} \right) \left( \frac{\sqrt{\phi_1^2 - 4\phi_2\phi_1} - \phi_1}{2\Lambda p^2} \right).$$

A plot of $R_{c2}$ as a function of $p$ is shown in Figure 3.4. The two methods of $R_c$ computation are graphically represented in Figure 3.4. A plot of $R_{c1}$ as a function of the reinfection parameter $p$ would suggest there is an optimum value of reinfection using the incorrect approach. However, a plot of $R_{c2}$ as a function of $p$ using the correct approach produces a totally different curve and shows no such optimum. Note that $p_{min}$, the minimum value of the reinfection parameter that induces bi-stability, may be calculated as $p_{min} = \left( \frac{k}{\mu} \right) \left( \frac{\mu + k}{\mu} \right)$ as found using the Center Manifold approach described in [58] (see section 3.4). Using the parameter values used to obtain Figure 3.4. The point $p_{min} = 0.0322$ corresponds to the value indicated at the arrow on Figure 3.4. The analysis that now follows in section 3.4.1 is important since it confirms the coexistence of a stable disease free equilibrium with two endemic equilibria where one is unstable and the other is stable as shown in Figure 3.3.
3.4 The backward bifurcation threshold, $R_c$

Figure 3.4: The red dashed line is a plot of $R_{c1}$ as a function of $p$ while the blue solid line illustrates a plot of the correct backward bifurcation threshold $R_{c2}$ as a function of $p$. Parameters used are $\Lambda = 10$, $\mu = 0.016$, $r = 2$, $\mu_d = 0.4$, $k = 0.0005$, $\beta = 0.0195$. The plot of $R_c$ as a function of reinfection is due to the fact that $R_c$ decreases as reinfection $p$ increases. $p_{\text{min}}$ is the minimum value of exogenous reinfection that triggers backward bifurcation.

3.4.1 Proof of existence of backward bifurcation for model equation (3.1)

Theorem 3.4.1 The model system (3.1) exhibits backward bifurcation whenever $p > p_{\text{min}}$ and no backward bifurcation otherwise.

Proof: Making use of the Centre Manifold approach as described in Castillo-Chavez and Song [58] (see chapter 2 Theorem 2.4.1) it is shown that

$$p_{\text{min}} = \left( \frac{k}{\mu} \right) \left( \frac{\mu + k}{\mu} \right),$$

indicated by the arrow in Figure 3.4, acts as a threshold that determines the positivity of the bifurcation coefficient $a$. 
To help explain the Center Manifold Theorem (2.4.1) it is convenient to transform the model variables of system (3.1) as follows: $x_1 = S, x_2 = E, x_3 = I, x_4 = R$ and $N = \sum_{j=1}^{4} x_j$. Now letting $X = (x_1, x_2, x_3, x_4)^T$ (where T denotes transpose) the model equation (3.1) can be written as $\frac{dX}{dt} = F(X)$ where $F = (f_1, f_2, f_3, f_4)^T$. Hence, it follows that

$$\begin{align*}
\frac{dx_1}{dt} &= \Lambda - \beta x_1 x_3 - \mu x_1 = f_1, \\
\frac{dx_2}{dt} &= \beta x_1 x_3 - p\beta x_2 x_3 - (\mu + k)x_2 = f_2, \\
\frac{dx_3}{dt} &= p\beta x_2 x_3 + kx_2 - (\mu + r + \mu_d)x_3 = f_3, \\
\frac{dx_4}{dt} &= rx_3 - \mu x_4 = f_4.
\end{align*}$$

(3.8)

The Jacobian matrix of the system (3.8) evaluated at the disease free equilibrium $P_0 = (\frac{\Lambda}{\mu}, 0, 0, 0)$ is obtained as

$$J = \begin{pmatrix} -\mu & 0 & \frac{-\beta \Lambda}{\mu} & 0 \\ 0 & - (\mu + k) & \frac{\beta \Lambda}{\mu} & 0 \\ 0 & k & - (\mu + r + \mu_d) & 0 \\ 0 & 0 & r & -\mu \end{pmatrix}.$$ 

At $R_0 = 1$ suppose $\beta$ is the bifurcation parameter, hence solving for $\beta$ from $R_0 = 1$ yields

$$\beta^* = \frac{\mu (\mu + k)(\mu + r + \mu_d)}{k \Lambda}.$$ 

With $\beta = \beta^*$ the transformed system (3.8) has a simple eigenvalue with zero real part and all other eigenvalues are negative (i.e., has a hyperbolic equilibrium point). Thus, the Center Manifold Theory can be applied [58] to investigate the dynamics of the transformed system (3.8) near $\beta = \beta^*$. It is possible to obtain the right eigenvectors of $J(\varepsilon_0)|_{\beta=\beta^*}$ which are denoted by $w = (w_1, w_2, w_3, w_4)^T$ where

$$w_1 = \frac{-\beta \Lambda w_3}{\mu^2}, w_2 = \frac{\beta \Lambda w_3}{\mu (\mu + k)}, w_3 = w_3 > 0, w_4 = \frac{rw_3}{\mu}.$$ 

Similarly it is easy to obtain the left eigenvectors denoted as $v = (v_1, v_2, v_3, v_4)^T$ where

$$v_1 = 0, v_2 = \frac{k v_3}{\mu + k}, v_3 = v_3 > 0, v_4 = 0.$$
Now, the associated bifurcation coefficients respectively denoted by \(a\) and \(b\) as described in Theorem 4.1 of [58] can be obtained. As indicated in Theorem 4.1 of [58] if bifurcation coefficients \(a\) and \(b\) are both non-negative then the system exhibits backward bifurcation where an unstable and a stable non-trivial equilibrium coexist with a stable disease free equilibrium.

**Computation of \(a\):** The transformed model equation (3.8) has the following non-vanishing partial derivatives of \(J\) evaluated at the disease free equilibrium \(\varepsilon_0\):

\[
\frac{\partial^2 f_1(0,0)}{\partial x_1 x_3} = -\beta, \quad \frac{\partial^2 f_2(0,0)}{\partial x_1 x_3} = \beta, \quad \frac{\partial^2 f_2(0,0)}{\partial x_2 x_3} = -p\beta, \quad \frac{\partial^2 f_3(0,0)}{\partial x_2 x_3} = p\beta.
\]

Hence,

\[
a = \sum_{k,i,j=1}^{4} v_k w_i w_j \frac{\partial^2 f_k(0,0)}{\partial x_i \partial x_j} = v_1 w_2 w_3 \frac{\partial^2 f_1(0,0)}{\partial x_1 \partial x_3} + v_2 w_1 w_3 \frac{\partial^2 f_2(0,0)}{\partial x_1 \partial x_3} + v_2 w_2 w_3 \frac{\partial^2 f_2(0,0)}{\partial x_2 \partial x_3} + v_3 w_2 w_3 \frac{\partial^2 f_3(0,0)}{\partial x_2 \partial x_3} = v_3 w_3 \Lambda \beta^2 \frac{(\mu + k)}{(\mu + k)^2} \left( \frac{p}{\mu} - \frac{k}{\mu} \right)^2 \mu \left( \frac{\mu + k}{\mu} \right) w_3.
\]

**Computation of \(b\):** Similarly, to compute the bifurcation coefficient \(b\), it can easily be shown that the non-vanishing partial derivatives associated with the Jacobian matrix \(J\) include

\[
\frac{\partial^2 f_1(0,0)}{\partial x_3 \partial \beta^*} = -\frac{\Lambda}{\mu}, \quad \frac{\partial^2 f_2(0,0)}{\partial x_3 \partial \beta^*} = \frac{\Lambda}{\mu}.
\]

so that,

\[
b = \sum_{k,i=1}^{4} v_k w_i \frac{\partial^2 f_k(0,0)}{\partial x_i \partial \beta^*} = v_1 w_3 \frac{\partial^2 f_1(0,0)}{\partial x_3 \partial \beta^*} + v_2 w_2 \frac{\partial^2 f_2(0,0)}{\partial x_3 \partial \beta^*} = k \mu \Lambda v_3 w_3 \frac{(\mu + k)}{\mu} > 0.
\]
Existence of multiple equilibria when $R_0 < 1$ is possible if the bifurcation coefficients $a$ and $b$ are positive. In fact $a > 0$ if and only if

$$p > p_{\text{min}} = \left( \frac{k}{\mu} \right) \left( \frac{\mu + k}{\mu} \right).$$

### 3.4.2 Equilibrium expressed in terms of the force of infection

It is important to note that some authors may choose to study backward bifurcations in epidemic models by expressing the endemic equilibria in terms of the force of infection [21, 86, 87, 88, 107] especially where the frequency-dependent transmission rate is used. This practice results in bifurcation diagrams that are different in shape to those seen in say Figure 2. Nevertheless, the same errors described here occur, no matter which way one chooses to plot the bifurcation diagram, if the aggregated parameter $R_0$ is not dealt with correctly.

I now reformulate model equation (3.1) with a frequency dependent transmission rate, as in [21, 86, 87, 88, 107], where the endemic equilibrium is given in terms of the force of infection $\lambda^* = \frac{\beta I^*}{N}$. Note that this is done by replacing $\beta I$ in model equation (3.1) with $\frac{\beta I}{N}$. Hence, the endemic equilibria expressed in terms of the force of infection are given as

$$S^{**} = \frac{\Lambda}{\lambda^* + \mu}, \quad E^{**} = \frac{\Lambda \lambda^*}{(p \lambda^* + \mu + k)(\lambda^* + \mu)}, \quad I^{**} = \frac{\Lambda \lambda^*(p \lambda^* + k)}{(p \lambda^* + \mu + k)(\lambda^* + \mu)(\mu + r + \mu_d)},$$

$$R^{**} = \frac{r \Lambda \lambda^*(p \lambda^* + \mu + k)}{\mu(p \lambda^* + \mu + k)(\lambda^* + \mu)(\mu + r + \mu_d)},$$

where $\lambda^*$ can be obtained by solving the following equation

$$g(\lambda^*) = A \lambda^2 + B \lambda^* + C = 0, \quad (3.9)$$

where

$$A = (\mu + r)p,$$

$$B = \mu p(\mu + r + \mu_d) + \mu(\mu + r + \mu_d) + k(\mu + r) - \mu p \beta,$$

$$C = \mu(\mu + k)(\mu + r + \mu_d)(1 - R_0).$$
3.4 The backward bifurcation threshold, $R_c$

Note that

$$R_0 = \frac{k\beta}{(\mu + k)(\mu + r + \mu_d)}.$$ 

Figures 3.5(a) and 3.5(b), respectively show a plot of the solution of equation (3.9) as a function of $R_0$ when all parameters in $R_0$ are fixed to constant values and when one parameter in $R_0$ is varied. There is a distinct difference between Figure 3.5(a) and Figure 3.5(b) as exhibited by the variation in their shape. Figure 3.5(a) where all parameters in $R_0$ are fixed to constant values represents the incorrect approach for obtaining backward bifurcation and therefore depict a parabolic shape. In fact Figure 3.5(a) resembles the plots of backward bifurcation produced in [21, 86, 87, 107] where all parameters in $R_0$ were kept constant. However, Figure 3.5(b) where $\beta \in [3.6, 6]$ is varied is the correct approach for obtaining backward bifurcation and in fact the parabolic shape is lost. Moreover, there is a difference in the gap between bifurcation curves with Figure 3.5(b) having a wider gap than Figure 3.5(a). In general varying at least one parameter in $R_0$, one allows the bifurcation curves to choose the colliding point but keeping parameters constant it is as if you have already determined the meeting point of the two bifurcation curves. Thus, keeping parameters constant in $R_0$ when obtaining backward bifurcation may result in either underestimating or overestimating the backward bifurcation threshold $R_c$. 
3.4 The backward bifurcation threshold, $R_c$

Figure 3.5: Backward bifurcation when endemic equilibria are expressed in terms of the force of infection. (a) Backward bifurcation when all the parameters in $R_0$ are fixed at constant values. Parameters used are $\mu = 0.01, r = 0.85, \mu_d = 0.1, k = 0.002316, p = 0.5, \beta = 4.5$. (b) Backward bifurcation when one parameter in $R_0$ is varied. Parameters used are $\mu = 0.01, r = 0.85, \mu_d = 0.1, k = 0.002316, p = 0.5, \beta \in [3.6, 6]$. In both figures the blue solid lines represent stable equilibria while red solid lines represent unstable equilibria.
3.5 Summary of the chapter

In this chapter it is shown that there is a profound difference in the backward bifurcation characteristics when the parameters in $R_0$ are fixed and $R_0$ is wrongly varied, as compared to validly varying a true model bifurcation parameter. Using the wrong approach, for constant values of $R_0$ a parabolic shape is observed. However, when a true model parameter in $R_0$ is varied this parabolic shape disappears. Moreover, if parameters in $R_0$ are incorrectly fixed researchers may overestimate or underestimate the backward bifurcation threshold $R_c$, below which $R_0$ needs to be reduced to eradicate the disease from the community. Since the ultimate goal of modelling is to give insight into disease dynamics, knowing the correct value of $R_c$ is important to public health. Thus, one has to be careful in plotting and calculating backward bifurcation, since if not done correctly errors can be introduced as in the studies [21, 48, 52, 63, 85, 86, 87, 88, 89, 90]. The correction shown here on how to obtain backward bifurcation in epidemic models will be useful for others investigating this interesting phenomena in the future.
Chapter 4

Backward bifurcation and hysteresis in TB models incorporating recurrent tuberculosis

4.1 Chapter overview

This chapter concerns modelling tuberculosis dynamics in a population giving careful attention to reinfection processes. Most of the previous TB modelling literature has been concerned only with the reinfection of individuals with latent TB i.e., reinfection of individuals whose TB is asymptomatic and non-infectious (see [56, 108, 109]). However, there are other important reinfection pathways that need to be dealt with. In particular, over the last decade it has now become appreciated that reinfection among individuals who have been successfully cured from the first episode of TB is substantial [77, 84, 110]. The second episode of TB which is referred to as recurrent TB often arises following treatment, because an individual who has recovered from a first episode of the disease does not necessarily gain permanent immunity to a second. Approximately, 10-30% of all cases of TB are due to recurrent tuberculosis [79] and multiple episodes are largely attributable to ineffective or poorly implemented tuberculosis control programs. Recent medical research concerning recurrent TB shows that individuals who have already had TB once are at a strongly increased risk of developing TB a second time through reinfection [84]. Moreover, the studies suggest that the incidence rate of TB
attributable to reinfection after successful treatment is four times higher than that of new TB [84].

Although recurrent TB is recognized as a serious problem it receives little attention [79]. The few studies that have attempted to model this form of TB (for instance see [56]) did not point out that recurrent TB can introduce new bifurcation structures besides the well known backward bifurcation structure. The contribution of this chapter is to assess how reinfection of recovered individuals can alter TB dynamics. In epidemic models the phenomenon of backward bifurcation is characterized by persistence of the disease when the associated basic reproduction number is below the endemic threshold. In this chapter, conditions that give rise to backward bifurcations are studied. Surprisingly, a rare hysteresis bifurcation may be observed where TB will continue to persist when the associated basic reproduction number is less than one, even though there is no backward bifurcation. Further, in this chapter the controversy raised by Lipsitch and Murray [111] regarding whether the phenomenon of backward bifurcation in TB models can occur in real life situations without compromising biological realism is revisited. It is revealed that when recurrent TB is accounted for, through incorporation of previously omitted reinfection pathways, then the phenomenon of backward bifurcation in TB epidemic models can occur in real life scenarios.

4.2 Introduction

In 1882 *Mycobacterium tuberculosis* was identified by Robert Koch as the aetiological agent responsible for tuberculosis (TB), yet despite all attempts to control its spread by modern medical science, it has become one of the most widespread and serious of all infectious diseases today [112]. The disease is transmitted from one person to another in tiny microscopic droplets when a person with pulmonary TB expels bacteria into the air by either coughing, sneezing, singing, laughing or other related activities that make use of airborne pathways. Amongst all infectious diseases, TB is one of the leading causes of death worldwide, and second only to HIV [1, 113]. Approximately a quarter of the global population harbours the TB bacteria and another eight to nine million new cases of tuberculosis emerge every year [114, 115]. Extraordinarily, TB is a treatable disease and can be prevented and cured through the use of prophylaxis and therapeutics for individuals with latent and clinically active TB respectively. Such treatment should in theory be an effective strategy for controlling the spread of TB. However, it has
failed in practice due to the inability to distribute sufficient drug treatment, usually in the form of antibiotics, to the world’s population combined with the difficulties of ensuring compliance to the required lengthy treatment program. Moreover, erratic treatment has led to the evolution of multi-drug resistant tuberculosis giving rise to the fear that TB may become an untreatable disease in the not too distant future. These problems, taken together, have led the WHO (World Health Organization) to formulate a post-2015 global “Stop TB Strategy” [112] to “end the global TB epidemic”.

A TB episode may have an exogenous or endogenous origin. Exogenous refers to a disease episode that results from recent exposure to some external infectious source (typically, contact with an infectious person). Endogenous designates situations where the individual is already harbouring the causative agent, under some healthy control by the immune system, which destabilizes and leads to disease. The pathogenesis of TB is characterized by the infection either remaining dormant, often for a long period that may last years, or progressing directly to active TB where clinical symptoms immediately manifest. The latter process is referred to as fast primary progression. The particular course of the disease depends on the host’s immune response towards the tubercle bacilli. Thus, the exposure to tubercle bacilli does not necessarily result in the manifestation of clinical forms of TB. Studies suggest that only 5-10% of individuals progress directly to the active stage after exposure to bacilli [116, 117]. The other component of the population of exposed individuals develop dormant TB and may remain latently infected possibly for the rest of their lifetime. However, destabilization of the immune system by the pathogen within the latently infected host can trigger endogenous reactivation, in which latent bacilli are reactivated and cause clinical *Mycobacterium tuberculosis*. The lifetime risk of a latently infected individual to progress to the infectious stage is approximately 5-10% [118].

### 4.2.1 Evidence of recurrent TB

Recurrent TB is defined as the emergence of a second episode of TB after the first episode has been successfully cured [77]. This often arises following treatment, because an individual that recovers from a first episode of the disease does not necessarily gain permanent immunity to a second. Approximately 10-30% of all cases of TB are due to recurrent tuberculosis [79], and multiple episodes are largely attributable to ineffective or poorly implemented tuberculosis control programs. Although recurrent
TB is recognized as a serious problem it receives little attention [79]. Through the use of advanced molecular fingerprinting techniques TB recurrence has been classified into two fundamental forms of infection: i) relapse of the original infecting strain, and ii) reinfection with a new strain of *Mycobacterium tuberculosis*. The role of reinfection and relapse to the overall burden of tuberculosis recurrence is not well understood and this has potential public health implications [119]. It is important to note that the system adopted by the WHO in recording and reporting TB cases does not differentiate between true relapse (i.e., reactivation of latent TB) and reinfection (exogenous acquisition of TB) and classifies relapse as any recurrence of TB [120]. This greatly affects the collection and analysis of data making it even more difficult to assess the specific role of reinfection. Studies conducted in [84] show that persons who had TB once are at a strongly increased risk of developing TB when reinfected. Moreover, the study suggests that the incidence rate of TB attributable to reinfection after successful treatment is four times higher than that of new TB [84].

There is concrete evidence that supports the view that exogenous reinfection forms an important source of recurrent TB among the successfully cured [56, 77, 84, 108, 109, 110]. Modern studies depict that exogenous reinfection is common in regions with a relatively high incidence of TB. A recent study conducted in Shanghai revealed that 61.5% of recurrent cases in a span of five years (1999 through 2004) were attributed to exogenous reinfection [110]. Thus, through detailed mathematical modelling, a model incorporating recurrent TB will now be formulated to investigate the impact of recurrent TB on the formation of backward bifurcations. The model will also be used to reveal new bifurcation structures that recurrent TB can induce.

### 4.3 Model of recurrent tuberculosis

Although there are numerous TB models that have attempted to include the recurrent TB pathway (for instance see [28, 56, 59, 85, 121, 122]), none have explored its role in inducing new bifurcation structures besides formation of backward bifurcation. Yang and Raimundo [59] investigated the impact of multiple infections and long latency on the dynamics of recurrent tuberculosis. Their results suggest that a backward bifurcation is expected to occur when a critical value of the disease incubation period is exceeded. However, in their analysis they assume that the reinfection of recovered
individuals is negligible, arguing that such a pathway increases non-linearity and makes the model mathematically intractable. The models of Feng et al. [56] and Kar and Mondal [85] were all based on the assumption that individuals went through a long latency period before TB reactivated to clinically active TB. However, based on the natural history of TB, fast primary progression of TB forms an important process through which symptomatic TB emerges [123] and needs to be included. Indeed, Feng et al. [56] incorporated a particular recurrent TB pathway in their model; however, their simplifying assumptions hindered further exploration with regard to its role in causing a backward bifurcation. The models developed by Gomes et al. [121] and Herrera et al. [122] attempted to incorporate exogenous reinfection, partial immunity to reinfection and primary progression. However, neither group examined how recurrent TB reinfection pathways could lead to a backward bifurcation; instead their main objective was to study the reinfection threshold. Hence, the main aim here is to study reinfection among recovered individuals and deduce the epidemiological implications, especially with regards to the possible formation of a backward bifurcation and other new bifurcation structures. For this purpose a TB model that includes fast primary progression and possible reinfection pathways is proposed.

The model is represented graphically in Figure 4.1 and assumes that every individual in the population belongs to one of four broad classes: susceptible individuals (S); exposed individuals (E) (these are infected individuals who are not able to transmit infection); individuals with active TB (I) (who manifest symptoms and are able to pass on the infection); and recovered/treated individuals (R). Individuals have the potential to move through these four classes, as for example in the loop $S \rightarrow E \rightarrow I \rightarrow R \rightarrow I$ or $S \rightarrow E \rightarrow I \rightarrow R \rightarrow E$, upon contact with the disease.

Susceptible individuals are generated by recruitment through births and immigration at a rate $\Lambda$. Upon effective contact with an infective, a small proportion $q$ of infected susceptible individuals follow the fast primary progression route (i.e., they move directly to the infective class) while the rest $(1 - q)$ move to the exposed class where they pass through a long latency period before reactivation and becoming infectious. Infected individuals who recover from the disease then move to the recovered class.

The model includes three different pathways for exogenous reinfection as shown in Figure 4.1:
(i) **Path A** = \( p\lambda E \), where exposed individuals (see [56]), become reinfected and progress to active TB;

(ii) **Path B** = \( (1-\sigma)\theta\lambda R \), where recovered individuals become reinfected and progress to the exposed sub-population;

(iii) **Path C** = \( \sigma\theta\lambda R \), where recovered individuals become reinfected and progress to active TB.

In the proposed model, the parameter \( p \) measures the degree of protection against TB among latently infected individuals and \( p\lambda E \) is the exogenous reinfection incidence rate. \( \sigma \) measures the probability of fast progression to the infectious class after reinfection. Following the studies of [56, 85, 121, 122, 124], \( p = 1 \) implies that the body does not render protection against exogenous reinfection while \( 0 < p < 1 \) implies the body is partially immune against exogenous reinfection (i.e., latent infection provides partial immunity against new infections) [28, 83]. Note that \( p > 1 \) would imply that an individual with latent TB infection has increased susceptibility to become newly infected compared to the susceptibility of the general population [28]. This would correspond to the results of studies which have found that recovered individuals are more likely to be susceptible to future TB infection than TB-naive individuals [84].

The model parameter \( \theta \) (0 < \( \theta \) < 1) quantifies the amount of exogenous reinfection among TB recovered individuals via paths B and C, while \( \sigma \) represents the probability of fast progression after reinfection amongst recovered individuals. \( \theta < 1 \) indicates that recovered individuals have acquired some degree of partial protective immunity to TB while \( \theta > 1 \) indicates increased susceptibility.

The full set of model equations is given in terms of the rates of change of each of the sub-populations S, E, I, and R, namely:
4.3 Model of recurrent tuberculosis

\[
\begin{align*}
\frac{dS}{dt} &= \Lambda - \lambda S - \mu S, \\
\frac{dE}{dt} &= (1 - q)\lambda S + (1 - \sigma)\theta \lambda R - p\lambda E - (k + \mu)E, \\
\frac{dI}{dt} &= q\lambda S + \sigma \theta \lambda R + p\lambda E + kE - (\mu + r + \mu_d)I, \\
\frac{dR}{dt} &= rI - \theta \lambda R - \mu R.
\end{align*}
\] (4.1)

Note that the total population at time \( t \) is given by \( N(t) \),

\[ N(t) = S(t) + E(t) + I(t) + R(t). \]

The model assumes a frequency-dependent incidence rate. Here the convention of working with the so-called “force of infection” \( \lambda \) defined as

\[ \lambda = \frac{c\beta I(t)}{N(t)}, \] (4.2)

where \( c \) represents the host-host contact rate is used. The parameter \( \beta \) is the probability of a contact being infectious [125] and \( I(t)/N(t) \) denotes the likelihood that the encounter is with an individual with active TB [126]. That is, individuals become infected when they come in contact with infected individuals regardless of the size of the human population [19, 127]. This form of incidence is considered to be more appropriate for infections in human populations [19]. It is important to note that although a large number of the previous TB models utilised frequency-dependent incidence rates, most assume that the disease-induced death rate \( (\mu_d) \) is negligible for the purpose of mathematical simplification. Indeed the analysis becomes mathematically difficult or even intractable without making such assumptions, since otherwise the total population of the model \( N(t) \) might never remain constant. Both [56] and [117] resorted to this assumption, even though this was not true of their more general model formulation. In the analysis here no such assumptions or simplifications are made and the total population is considered to be varying in time.

In more detail, as equation (4.1) shows, new susceptible individuals are generated by recruitment through births and immigration at a rate \( \Lambda \). Susceptible individuals move to the latently infected class \( E \) upon an effective contact with an infected individual
4.3 Model of recurrent tuberculosis

(I) at a rate $\lambda I$. The exposed sub-population ($E$) increases with the infection of susceptible individuals at rate $(1 - q)\lambda S$ and reinfection (recurrent TB) of recovered individuals at a rate $(1 - \sigma)\theta R$. It decreases by exogenous reinfection ($p\lambda EI$), and endogenous reactivation ($k E$) upon which exposed individuals move to the infectious class. The infected sub-population is generated by fast primary progression of TB susceptibles ($\lambda IG$), exogenous reinfection amongst the exposed sub-population (at rate $\lambda E$), exogenous reinfection of recovered individuals ($\sigma\theta R$) and endogenous reactivation ($k E$). The sub-population is decreased by per-capita recovery due to treatment $r I$, and by disease induced death at rate $\mu_d I$. Finally the recovered sub-population ($R$) is generated by the recovery of infected individuals and the loss due to exogenous reinfection (at rate $\theta R$). Note that natural death affects all classes of the population at the same per-capita rate via the mortality parameter $\mu$.

Note that the proposed model assumes all immigrants are susceptible and enter the population at rate $\Lambda$. While this excludes the realistic possibility of immigration of infectives (either latent individuals or individuals with active TB), it helps untangle the conditions that result in backward bifurcation. Previous modelling studies have already demonstrated that the immigration of infectives is a pathway that can trigger backward bifurcations [128, 129]. Since the aim of the proposed model is to exclusively investigate how reinfection after recovery influences TB dynamics, in particular the phenomenon of backward bifurcation, recruitment of infected individuals is removed. Otherwise the analysis is mathematically intractable. Moreover, it is important to justify the choice of a single latent compartment instead of two (i.e., slow and fast TB progression) as the modelling community is divided on how best to model TB dynamics. In this chapter the choice of a single latent compartment instead of two was based on the available literature on TB dynamics. Progression of active TB is not uniform, as some infected individuals are more likely to progress to active TB than others. A number of models incorporating long and variable rates of progression have been constructed and analysed (for instance see [56, 94, 130]). Feng et al. [131] investigated the impact of variability in latency using arbitrary continuous distributions and found that such generalization did not result in qualitative differences in terms of the model dynamics. Before the Feng et al. [56] study, Blower et al. [94] formulated a differential equation model with two latent cohorts: one cohort consisted of those who rapidly develop TB after primary infection; while the second cohort involved individuals who develop the infection slowly through endogenous reactivation. Feng et
al. [131] also showed that the artificial divisions (as in Blower et al. [94]) play no role in the qualitative dynamics.

Secondly, a single latent compartment over two latent compartments (fast and slow TB progression) was chosen for the purpose of mathematical tractability. The increased number of reinfection pathways needed with two compartments would add a considerable degree of model complexity, and become very difficult to analyse. Moreover, the original study of [56], where the role of reinfection in inducing backward bifurcation was first identified, consisted of a single latent compartment. The paper [111] that criticised Feng et al. [56] also had a single latent compartment. The goal of this chapter is to investigate how recurrent TB (reinfection of recovered individuals) can impact the backward bifurcation phenomenon found by Feng et al. [56], thus the need to keep the same single latent compartment structure.

Table 4.1 provides a detailed list and description of the model parameters as well as typical parameter values used here, as obtained from the literature.

Figure 4.1: Schematic diagram of the main processes involved in TB infection according to model equation (4.1). The red dotted curved arrows represent reinfection pathways.
4.3 Model of recurrent tuberculosis

Table 4.1: Description of variables and parameters of model (4.1).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Interpretation</th>
<th>Nominal value</th>
<th>Sources</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$</td>
<td>Susceptible sub-population</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E$</td>
<td>Asymptomatic and non-infectious individuals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$I$</td>
<td>Symptomatic and infectious individuals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R$</td>
<td>Recovered individuals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu$</td>
<td>Per-capita natural death rate</td>
<td>0.016</td>
<td>[132]</td>
<td>$year^{-1}$</td>
</tr>
<tr>
<td>$\mu_d$</td>
<td>Per-capita disease induced death rate</td>
<td>0.1</td>
<td>[56]</td>
<td>$year^{-1}$</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Recurrent TB due to reinfection</td>
<td>1.61 [1.61, 7.79]</td>
<td>[28, 84, 91]</td>
<td></td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Probability of fast progression after reinfection</td>
<td>0.05-0.1</td>
<td>[28]</td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>Per-capita recovery rate</td>
<td>2</td>
<td>[56, 133]</td>
<td>$year^{-1}$</td>
</tr>
<tr>
<td>$p$</td>
<td>Exogeneous re-infection</td>
<td>0.25 [0, 1]</td>
<td>[56, 132, 134]</td>
<td></td>
</tr>
<tr>
<td>$q$</td>
<td>Primary progression proportion (progress to active TB soon after infection)</td>
<td>0.05</td>
<td>[56, 85, 117]</td>
<td>$year^{-1}$</td>
</tr>
<tr>
<td>$k$</td>
<td>Per-capita endogeneous reactivation rate</td>
<td>0.0002</td>
<td>[28, 91, 97]</td>
<td>$year^{-1}$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Probability of becoming infected per contact</td>
<td>variable</td>
<td>estimated</td>
<td></td>
</tr>
<tr>
<td>$c$</td>
<td>Mean number of contacts</td>
<td>variable</td>
<td>estimated</td>
<td>$year^{-1}$</td>
</tr>
<tr>
<td>$\Lambda$</td>
<td>Recruitment rate</td>
<td>100</td>
<td>[135]</td>
<td>$year^{-1}$</td>
</tr>
</tbody>
</table>
4.4 Model analysis

4.4.1 Basic properties

Following the methods in [63] (and also the proof in chapter 3), it is not difficult to prove that when all model parameters are nonnegative the state variables \( S(t), I(t), E(t) \) and \( R(t) \) are all positive for all time \( t \) provided they are positive initially.

**Theorem 4.4.1** The region

\[
\mathcal{N} = \left\{ (S, E, I, R) \in \mathbb{R}_+^4 : S + E + I + R \leq \frac{\Lambda}{\mu} \right\}
\]

is positively invariant and attracting with respect to the model equation (4.1) with initial conditions in \( \mathbb{R}_+^4 \).

**Proof.** The proof of Theorem 4.4.1 is similar to the proof of Theorem 3.3.2 in chapter 3 and therefore is not repeated here.

The equilibria of the model are found by setting the rates of all variables in the left-hand side of equation (4.1) to zero. Clearly the equations have an intrinsic disease-free equilibrium (DFE) given by

\[
(S, E, I, R) = P_0 = \left( \frac{\Lambda}{\mu}, 0, 0, 0 \right).
\]

The stability of the DFE is controlled by the basic reproduction number \( R_0 \) which represents the average number of new infections generated by an infected individual when introduced into an entirely susceptible population. \( R_0 \) may be determined using the next generation operator method (see [136]) as shown in Appendix A, where it is found that

\[
R_0 = \frac{c \beta(k + q \mu)}{(\mu + k)(\mu + r + \mu_d)}. \tag{4.3}
\]

It is possible to decouple the expression for \( R_0 \) to account for slow TB progression and fast primary progression

\[
R_0 = \left( (1 - q) \frac{c \beta}{\mu + r + \mu_d} \right) \left( \frac{k}{\mu + k} \right) + \left( qc \beta \frac{1}{\mu + r + \mu_d} \right).
\]

Slow TB \( R_0 \)

Fast TB \( R_0 \)
The slow TB component of $R_0$ can be obtained by observing that the average infectious period is given as $\frac{1}{\mu + r + \mu_d}$ and the probability of progressing from the latent compartment to the infective class is given as $\left(\frac{k}{\mu + k}\right)$. Thus, the average time an individual who starts in the latent compartment is expected to spend in the infectious compartment is $\left(\frac{1}{\mu + r + \mu_d}\right) \times \left(\frac{k}{\mu + k}\right)$. Multiplying this average time by $(1 - q)c\beta$ yields the slow TB $R_0$. Moreover, multiplying the mean infectious period $\left(\frac{1}{\mu + r + \mu_d}\right)$ with $qc\beta$ yields the fast TB component of $R_0$.

An important result is the following Theorem 4.4.2:

**Theorem 4.4.2** Provided $R_0 < 1$, the DFE of the model (4.1) is locally asymptotically stable, otherwise it is unstable.

This general result has been reviewed in [25], and hence Theorem 4.4.2 will not be proved here. The theorem implies that it is possible to eradicate the disease from the community when $R_0 < 1$ if the initial sizes of the sub-populations of model (4.1) are in the basin of attraction of the disease free equilibrium.

Interestingly, the formula for the basic reproduction number (4.3) does not include the reinfection parameters $p$ and $\theta$ despite the fact that these terms should contribute significantly to the emergence of new cases of TB infection. Hence, this already suggests that $R_0$ alone is unable to completely quantify some key dynamical features of the TB epidemic, and is in fact the first sign that a backward bifurcation might be involved. It will emerge that the reinfection parameters $p$ and $\theta$ do play an important role and are responsible for the presence of the backward bifurcation intrinsic to this model.

### 4.4.2 Endemic equilibria

First the model’s endemic equilibrium points are identified being mindful that there may in fact be several such points coexisting simultaneously. As before, to find the endemic equilibria $(S^*, E^*, I^*, R^*)$ the rate equations (4.1) are set to zero and solved for the equilibrium quantities $S^*, E^*, I^*$ and $R^*$ in terms of the force of infection $\lambda$. 
4.4 Model analysis

This gives

\[ S^* = \frac{\Lambda}{\lambda + \mu}, \]
\[ E^* = \frac{(1 - q)\lambda\lambda(\mu + \mu_d)(\theta\lambda + \mu) + (1 - q)\lambda\lambda r\mu + (1 - \sigma)\theta r \lambda^2}{\varpi}, \]
\[ I^* = \frac{q\lambda\lambda(\theta\lambda + \mu)(p\lambda + k + \mu) + (1 - q)\lambda\lambda(\theta\lambda + \mu)(p\lambda + k)}{\varpi}, \]
\[ R^* = \frac{rq\lambda\lambda\mu + r\lambda\lambda(p\lambda + k)}{\varpi}, \]

where

\[ \varpi = (\mu + \mu_d)(\theta\lambda + \mu)(\lambda + \mu)(p\lambda + k + \mu) + r\theta\lambda\mu(1 - \sigma)(\lambda + \mu) \]
\[ + r\mu(\lambda + \mu)(p\lambda + k + \mu). \]

Substituting these equilibrium quantities (i.e., expressions (4.4)) into the force of infection equation (4.2) yields

\[ P(\lambda) = \lambda(a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0) = 0, \] (4.5)

where

\[ a_3 = \theta p, \ a_2 = c\theta p(\beta_0 - \beta), \ a_1 = c[q\theta\mu + \theta k + \mu p] (\beta_1 - \beta), \]
\[ a_0 = c\mu(k + \mu q)(\beta_R - \beta), \] (4.6)
and

\[
\beta_0 = \frac{\theta p(\mu + \mu_d) + (1 - q)(\mu + \mu_d)\theta + (1 - \sigma)\theta r + \theta(k + \mu q) + p(\mu + r)}{c\theta p},
\]

(4.7)

\[
\beta_1 = \frac{\Xi}{c(q\theta \mu + \theta k + \mu p)},
\]

(4.8)

\[
\beta_R = \frac{(\mu + r + \mu_d)(\mu + k)}{c(k + \mu q)},
\]

(4.9)

where

\[
\Xi = \theta(\mu + \mu_d)(\mu + k) + \mu p(\mu + r + \mu_d) + r\theta \mu(1 - \sigma) + (1 - q)(\mu + \mu_d)\mu + r(\mu + k) + \mu(k + \mu q).
\]

One notes from equation (4.9) that \(R_0 = \frac{\beta}{\beta_R}\), and also that the root \(\lambda = 0\) corresponds to the DFE, where \(I^* = 0\). Now the roots of the cubic equation

\[
P_1(\lambda) = a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 = 0
\]

(4.11)

substituted in \(S^*, E^*, I^*, R^*\) yield the endemic equilibrium for any specific set of model parameters.

### 4.5 No recurrent TB (i.e., \(\theta = 0\)): quadratic \(P_2(\lambda)\)

First an important particular parameter subset in which \(\theta = 0\) is examined, that is, in the absence of recovered individuals becoming reinfected. This model has the same basic reproduction number as for the case \(\sigma, \theta > 0\), as \(\sigma\) and \(\theta\) do not appear in the \(R_0\) expression (4.3). For \(\sigma = \theta = 0\) the third degree polynomial equation (4.11) collapses to the quadratic

\[
P_2(\lambda) = c_2\lambda^2 + c_1\lambda + c_0 = 0,
\]

(4.12)
In the above, it is observed that $c_0 < 0$ corresponds to $R_0 > 1$ and vice versa. Thus, in the absence of recurrent TB it can be deduced that $c_1 < 0$, $R_0 < 1$ and $c_1^2 - 4c_2c_0 > 0$ (see Theorem B.2.1 in Appendix B) indicate conditions for the existence of a backward bifurcation, based on the roots of the quadratic equation (4.12).

Note that case (iii) of Theorem B.2.1 in Appendix B stipulates the condition that $\Delta = c_1^2 - 4c_2c_0 > 0$ which means there are two real positive endemic equilibria as required for a backward bifurcation to appear. In fact $\Delta = 0$ provides the critical point for the backward bifurcation where the two positive endemic equilibria collide and annihilate each other leaving the DFE as the only equilibrium. By setting $\Delta = 0$, the critical value of the transmission coefficient denoted by $\beta_c$ is obtained. For mathematical convenience let

$$c_1 = \phi_1 - \phi_2 \beta, \ c_0 = \phi_3 - \phi_4 \beta, \ \phi_1 = (\mu + r + \mu_d)\mu p + (\mu + k)(\mu + r) + \mu \mu_d (1 - q),$$

$$\phi_2 = c\mu p, \ \phi_3 = (\mu + r + \mu_d)(\mu + k), \ \phi_4 = c\mu (k + \mu q).$$

Now the discriminant $\Delta(\beta)$ may be expressed in terms of $\beta$. Let $\beta_c$ be the critical value of $\beta$ for which the discriminant equals zero i.e.,

$$\Delta(\beta_c) = \phi_2^2 \beta_c^2 + 2(2c_2 \phi_4 - \phi_1 \phi_2) \beta_c + (\phi_1 - 4c_2 \phi_3) = 0.$$

Some algebraic rearrangement gives

$$\beta_c = \frac{(\phi_1 \phi_2 - 2c_2 \phi_4) + 2\sqrt{c_2^2 \phi_4^2 + c_2 \phi_2 (\phi_2 \phi_3 - \phi_1 \phi_4)}}{\phi_2^2}.$$  \hspace{1cm} (4.13)
The critical value of the basic reproduction number denoted by $R_c$ is obtained by replacing the parameter $\beta$ in $R_0$ with $\beta_c$ which yields

$$R_c = \left( \frac{c(k + \mu q)}{\mu + k} \right) \left( \frac{(\phi_1 \phi_2 - 2c_2 \phi_4) + 2\sqrt{c_2^2 \phi_4^2 + c_2 \phi_2 (\phi_2 \phi_3 - \phi_1 \phi_4)}}{\phi_2} \right)$$

(4.14)

where the right-hand term in large brackets is just $\beta_c$.

In fact $R_c$ defines a sub-threshold domain of bistable equilibria of the model system (4.1) in the sense that within the region $R_c < R_0 < 1$ the model equation (4.1) has two positive endemic equilibria simultaneously existing with a stable disease free equilibrium. Thus, the backward bifurcation for equation (4.1) occurs for values of the basic reproduction number $R_0$ that lie between $R_c < R_0 < 1$. The associated backward bifurcation for the model without the reinfection pathways $A$ and $B$ (i.e., $\sigma = \theta = 0$) shown in Figure 4.2(a) is obtained by plotting $\lambda$ as a function of $\beta$. Figure 4.2(a) shows that model (4.1) has a disease free equilibrium which corresponds to $\lambda = 0$ and two non-trivial endemic equilibria which, according to numerical simulations, one is locally asymptotically stable (LAS) and the other is unstable (saddle). Applying the Center Manifold Theory it is possible to examine the stability and coexistence of these three equilibria (see proof in Appendix C). Coexistence of two positive equilibria when $R_0 < 1$ confirms that the model exhibits the phenomenon of backward bifurcation for $R_0 < 1$.

It can be seen from equations (4.13) and (4.14), that when the exogenous reinfection parameter $p$ increases, $R_c$ decreases, and when exogenous reinfection parameter decreases, $R_c$ increases. Hence, it is deduced from expression (4.14) that $R_c$ is inversely proportional to the level of exogenous reinfection $p$. This observation is confirmed by Figure 4.2(b) which illustrates the effect of increasing exogenous reinfection $p$ on $R_c$. That is, with low values of exogenous reinfection the critical value $R_c$ is high implying that the extent of the backward bifurcation regime becomes smaller as $R_c$ becomes closer to unity ($R_0 = 1$). The threshold implies that TB can be eliminated from the community if the basic reproduction number is maintained below $R_c$ (i.e., $R_0 < R_c$).

More formally the following lemma is stated:

**Lemma 4.5.1** For model equation (4.1), when $\sigma = \theta = 0$,
(i) If $R_0 > 1$ the model has one positive endemic equilibrium point,

(ii) If $R_c < R_0 < 1$ the model has two positive endemic equilibria,

(iii) If $R_0 < R_c$ the model has only a disease free equilibrium.

Figure 4.2: (a) Illustration of backward bifurcation when there is no recurrent TB (i.e., $\theta = 0$). Parameters used remain as defined in Table 4.1 except $p = 0.09 > p_c = 0.0658$, $k = 0.0002$, $q = 0.05$, $c = 25$, $\beta \in [0.4, 0.9]$ and $\beta_c = 0.5099$ corresponding to $R_c = 0.8852$. In the figure the blue solid lines represent the stable equilibria while the red dotted line represents unstable equilibria. (b) A plot of the critical value of the basic reproduction number, $R_c$, as a function of the level of exogenous reinfection $p$. NBB and BB respectively denote no backward bifurcation and backward bifurcation regions. That is, in the region denoted by NBB the level of exogenous reinfection is too low to induce backward bifurcation while in the region denoted by BB the level of exogenous reinfection is sufficient to cause multiple positive endemic equilibria.

4.6 Recurrent TB: model with all reinfection pathways ($A$, $B$ and $C$) $\theta > 0$; cubic $P_1(\lambda)$

Now returning to the fully general model with all parameters $p, \sigma, \theta$ positive. Recall that the sign of the roots of the cubic polynomial (equation (4.5)) $P_1(\lambda)$, tell us the signs of the equilibrium populations for the number of infected individuals (via
equation (4.2)). Observe that the coefficients in the cubic polynomial \( a_3, a_2, a_1 \) and \( a_0 \) (see equation (4.6)) are all real numbers. For any non-negative model parameters, \( a_3 \) is always positive while \( a_2, a_1 \) and \( a_0 \) can be either positive, zero or negative depending on \( \beta_0, \beta_1 \) and \( \beta_R \), respectively (see equations (4.7)-(4.9)). A comprehensive analysis of the roots of the cubic (4.11) may be carried out using Descarte’s Rule of Signs [137] (see Table 4.2). A simpler more intuitive approach is to examine the roots at the transcritical bifurcation point \( R_0 = 1 \). Such an analysis gives an understanding of the type of bifurcation that is likely to occur in the vicinity of \( R_0 = 1 \), and provides conclusions that coincide with the more detailed analysis based on Descarte’s Rule of Signs [137]. Conveniently when \( R_0 = 1 \) the cubic polynomial (4.11) reduces to the quadratic equation

\[ f(\lambda) = a_3 \lambda^2 + a_2 \lambda + a_1 = 0. \]

A simple study of the roots shows that if \( a_2 < 0 \) and \( a_1 < 0 \) (i.e., if \( \beta > \beta_0 \) and \( \beta > \beta_1 \)) or if \( \beta_1 < \beta < \beta_0 \), the quadratic equation has one positive endemic equilibrium. As seen in Figure 4.3(b), this is the signature of a backward bifurcation. Namely, when \( R_0 \) is slightly below unity, the model equation (4.1) has two positive endemic equilibria but only one when \( R_0 \geq 1 \). Furthermore, if \( a_2 > 0 \) and \( a_1 > 0 \) (i.e., \( \beta < \beta_0 \) and \( \beta < \beta_1 \)) then at the point \( R_0 = 1 \), the model has no positive endemic equilibria. This characteristic is indicative of forward bifurcation as seen in Figure 4.3(a). However, if \( R_0 \) is increased slightly above unity then model (4.1) has one positive endemic equilibrium point. It can be shown that if \( \beta_0 < \beta < \beta_1 \), then the reduced equation has two positive real roots, indicating that the model equation (4.1) exhibits hysteresis (see Figures 4.4(a) and 4.4(b)), as will be discussed in detail shortly. This discussion based on the point \( R_0 = 1 \) is summarized in Lemma 4.6.1:

**Lemma 4.6.1** At the point \( R_0 = 1 \), where \( \beta = \beta_R \) the model equation (4.1) has:

(i) Two positive endemic equilibria if \( \beta_0 < \beta_R < \beta_1 \), which signals hysteresis;

(ii) One positive endemic equilibrium if \( \beta_R > \beta_0 \) and \( \beta_R > \beta_1 \), or if \( \beta_1 < \beta_R < \beta_0 \), either of which signals backward bifurcation;

(iii) No positive endemic equilibria if \( \beta_R < \beta_0 \) and \( \beta_R < \beta_1 \), which signals forward bifurcation.
Table 4.2: Summary of the model equilibria.

<table>
<thead>
<tr>
<th>Range of $R_0$</th>
<th>Conditions</th>
<th>Equilibria of model system (4.1)</th>
<th>Type of bifurcation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_0 = 1$</td>
<td>$\beta &gt; \beta_0, \beta &gt; \beta_1$</td>
<td>One positive endemic equilibrium</td>
<td>Forward bifurcation</td>
</tr>
<tr>
<td></td>
<td>$\beta_1 &lt; \beta &lt; \beta_0$</td>
<td>One positive endemic equilibrium</td>
<td>Forward bifurcation</td>
</tr>
<tr>
<td></td>
<td>$\beta_0 &lt; \beta &lt; \beta_1$</td>
<td>Two positive endemic equilibria</td>
<td>Associated to hysteresis</td>
</tr>
<tr>
<td>$R_0 &gt; 1$</td>
<td>$\beta \geq \beta_0, \beta \geq \beta_1$</td>
<td>One positive endemic equilibrium</td>
<td>Forward bifurcation</td>
</tr>
<tr>
<td></td>
<td>$\beta \leq \beta_0, \beta \leq \beta_1$</td>
<td>One positive endemic equilibrium</td>
<td>Forward bifurcation</td>
</tr>
<tr>
<td></td>
<td>$\beta_0 &lt; \beta &lt; \beta_1$</td>
<td>Three positive endemic equilibria</td>
<td>Hysteresis</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>Two positive endemic equilibria</td>
<td>Backward bifurcation</td>
</tr>
<tr>
<td></td>
<td>$\beta &gt; \beta_0, \beta &gt; \beta_1$</td>
<td>Two positive endemic equilibria</td>
<td>Backward bifurcation</td>
</tr>
<tr>
<td></td>
<td>$\beta_1 \leq \beta &lt; \beta_0$</td>
<td>Two positive endemic equilibria</td>
<td>Backward bifurcation</td>
</tr>
<tr>
<td></td>
<td>$\beta &lt; \beta_0, \beta &lt; \beta_1$</td>
<td>No positive endemic equilibria</td>
<td>Associated to forward bifurcation</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.6.1 Existence of backward bifurcation reinfection threshold at $R_0 = 1$

The concept of existence of a critical point at $R_0 = 1$ was fully developed by [58]. For the general model (4.1), the critical value of exogenous reinfection $p_c$ required to allow the formation of a backward bifurcation at $R_0 = 1$ is determined.

Define

$$p_c = \frac{(k + \mu q)}{\mu(1 - q)} \left( \frac{\mu(1 - q)(\mu + r + \mu_d) + (\mu + r)(k + \mu q)}{\mu(\mu + r + \mu_d)} - F_r \right), \quad (4.15)$$

where

$$F_r = \frac{r \theta (k + \sigma \mu)}{\mu(\mu + r + \mu_d)}. \quad (4.16)$$

The expression $F_r$ (equation (4.16)) is associated with recurrent TB since it is the only term containing the reinfection parameter $\theta$. In the absence of recurrent TB due to reinfection (i.e., $\theta = 0$) the expression $F_r$ reduces to zero. Now define the following Lemma 4.6.2:

**Lemma 4.6.2** The model system (4.1) at $R_0 = 1$ exhibits

(i) Backward bifurcation whenever $p > p_c$,

(ii) Forward bifurcation whenever $p < p_c$. 

To see this, first note that referring to Figure 4.3(b), it is apparent that for a model to exhibit backward bifurcation it is required that, at $R_0 = 1$, there is a single positive endemic equilibrium. Note again that when $R_0 = 1$ (where $a_0 = 0$), the equilibria relate to the roots of the quadratic $f(\lambda) = a_3\lambda^2 + a_2\lambda + a_1 = 0$. The roots are

$$\lambda = \left( -a_2 \pm \sqrt{a_2^2 - 4a_1a_3} \right) / 2a_3 \quad \text{where} \quad a_3 > 0. \tag{4.17}$$

The point of interest is the point where the bifurcation structure changes from forward to backward. But when $R_0 = 1$ ($\beta = \beta_R$) this is just the point where the positive endemic equilibrium vanishes to zero, and from equation (4.17) this must occur when $a_1 = 0$ ($\beta = \beta_1$); or equivalently $\beta_1 = \beta_R$. Also note that $\beta_1$ is a function of $p$ (i.e., $\beta_1 = \beta_1(p)$) while $\beta_R$ is not. Hence, equating equations (4.8) and (4.9) for $\beta_1(p)$ and $\beta_R$ and solving yields the critical value $p_c$ for which $\beta_1(p_c) = \beta_R$. After some algebraic manipulation it is found that the required backward bifurcation reinfection threshold is given by equation (4.15). A more detailed derivation of $p_c$ can be found in Appendix C, which takes into account the stability of the equilibria through the use of Center Manifold Theory.

Numerically, Lemma 4.6.2 is illustrated by Figures 4.3(a) and 4.3(b) which respectively show that the model equation (4.1) has forward bifurcation when $p < p_c$ and backward bifurcation when $p > p_c$.

**4.6.2 Relation with models of Lipsitch & Murray 2003 and Feng et al. 2000**

It is interesting to compare the critical point $p_c$ for the reinfection threshold given in equation (4.15) with that found by Feng et al. [56] in their much simpler but still important model. First note that Feng et al. [56] do not include fast primary progression which is equivalent to setting $q = 0$ (see Figure (4.1)). They also assume that the exogenous reinfection rates are $\theta = 1, \sigma = 0$ and that the disease induced death rate is negligible with $\mu_d = 0$. Under these conditions Feng et al. [56] find that backward bifurcations occur only if $p > P_{Feng}$, where

$$P_{Feng} = \frac{k}{\mu} \left( 1 + \frac{k}{\mu + r} \right) \approx \frac{k}{\mu},$$
and the latter approximation assumes that $k \ll r$.

In their controversial paper Lipsitch and Murray [111] argued that in the real world recovered individuals gain immunity to reinfection, and thus reinfection among exposed individuals must be less than the probability of progressing to the infectious stage of TB. They showed that this implies

$$p < P_L = k / (\mu + k) \approx \frac{k}{\mu}$$

given that $k \ll \mu$. As such, Lipsitch and Murray [111] argued that backward bifurcations should not be expected in the real world. The extended model has some interesting insights with regard to these studies. First note that after inclusion of the assumptions made by Feng et al. [56] in model equation (4.1) (e.g., setting $q = 0$), the threshold $p_c$ for backward bifurcation (equation (4.15)) simplifies to

$$p_c \approx P_{Feng}.$$  (4.18)

That is, backward bifurcations are possible only when $p > p_c \approx P_{Feng}$, and thus the result of Feng et al. [56] is retrieved.

Consider now the extended model with more realistic infection pathways ($q > 0, \sigma > 0$), but still assuming that $\mu_d = 0$, $k \ll \mu$ and $\mu \ll r$. For this approximation $F_r \approx \sigma \theta$, and the backward bifurcation threshold equation (4.15) simplifies to

$$p_c \approx \left( q + \frac{k}{\mu} \right) \left( 1 + \frac{k}{\mu} - \sigma \theta \right) \approx \left( \frac{k}{\mu} + q \right) (1 - \sigma \theta).$$  (4.19)

It is interesting to compare this result to the threshold $P_L$ [111] discussed above. If recovered individuals gain high immunity from having been infected, then the reinfection pathway $R \rightarrow I$ is relatively small (see Figure 4.1), with $\sigma \theta \ll 1$. In this case equation (4.19) shows that $p_c \approx P_L + q$. This is similar to the Lipsitch and Murray [111] criterion, and suggests that backward bifurcation is unlikely to occur in the real world, if as Lipsitch and Murray [111] claim that in reality $p < P_L$. However, if recovered individuals gain only mild immunity against reinfection, then the reinfection pathway $R \rightarrow I$ and $\sigma \theta$ can be relatively large. Note that $\theta$ can be greater than unity.
4.7 Hysteresis

as suggested by Verver et al. [84]. Gomes et al. [91] estimate $\theta$ to be in the range [1.61, 7.79]. In this situation it is quite possible that $p_c < P_L$. Hence, with relatively low immunity amongst recovered individuals [84] backward bifurcation can occur despite the fact that the Lipsitch and Murray [111] prediction would predict otherwise. This does not mean that Lipsitch and Murray [111] have erred, but that their result may need modifications when discussing the presence of more complex reinfection pathways. In fact even just for the simplified reinfection pathways of the original Feng et al. [56] model, the validity of the Lipsitch and Murray [111] argument has recently been called into question given the difficulties of comparisons with real world processes [126].

In the more recent literature, numerous studies have pointed out that initial infection may not confer protection against exogenous reinfection especially in high-risk populations [138, 139]. Some studies have demonstrated that it is possible for exogenous reinfection to outweigh endogenous reactivation [140, 141]. This is supported by the fact that the majority of new TB cases (about 90%) occur as a result of reinfection rather than endogenous reactivation [134, 140, 141, 142]. In this situation, $p_c > k/\mu k)$, and backward bifurcations can occur even according to the Lipsitch and Murray [111] criteria. Other distinguished medical research shows that reinfection rates after successful treatment are much higher than rates of new TB; sometimes approximately four times higher [84, 91]. Furthermore, similar to vaccine conferred immunity, the protection rendered by latent TB infection wanes with time and it is uncertain whether latent infections would provide a similar immunity decades after the first episode of TB [134].

4.7 Hysteresis

For the usual forward bifurcation (for example, Figure 4.3(a)), a model has two locally stable branches at the transcritical point $R_0 = 1$: i) an infection free equilibrium that is locally asymptotically stable when $R_0 < 1$; ii) an endemic equilibrium which is stable for $R_0 > 1$. However, this scenario where there is only one endemic equilibrium when $R_0 > 1$ may not always be the case. For example, Reluga et al. [39] noted in their study of epidemic models with structured immunity that it is possible that more than one endemic equilibria may coexist even though the basic reproduction number is greater than one. This leads to an unusual phenomenon of forward bifurcation with
Figure 4.3: Plots the number of infectives at equilibrium as a function of $R_0$. In both figures the blue solid lines represents stable equilibria while the red dotted line represents an unstable equilibrium. (a) Shows forward bifurcation with parameters $p = 0.06 < p_c = 0.0647$, $\sigma = 0.05$, $\theta = 0.3$, $\mu = 0.0167$, $\mu_d = 0.1$, $r = 2$, $q = 0.05$, $\Lambda = 100$, $c = 60$, $k = 0.0002$, $\beta \in [0.3, 0.8]$. (b) Shows backward bifurcation with the same as in (a) except $p = 0.09 > p_c = 0.0647$. For a clear view all bifurcation structures are plotted with semi-log axes.
4.7 Hysteresis

Figure 4.4: Plots show the infectious population $I^*$ at equilibrium as a function of $R_0$. (a) Forward bifurcation with hysteresis where the multiple equilibria are strictly to the right of $R_0 = 1$. (b) Forward bifurcation with hysteresis where there are multiple equilibria to the left and to the right of $R_0 = 1$. Parameters used are $\mu = 0.0167, \mu_d = 0.1, k = 0.0002, \theta = 0.5, r = 2, q = 0.05, \sigma = 0.2, c = 60, \Lambda = 100, \beta \in [0.2, 0.8]$. In (a) $p = 0.057 < p_c = 0.0639$ and in (b) $p = 0.058 < p_c = 0.0639$. In both figures the blue solid lines represent stable equilibria while the red dotted lines represent unstable equilibria. For a clear view, all bifurcation structures are plotted with semi-log axes.
hysteresis, which can be triggered in TB model equation (4.1) when reinfection is taken into account. Thus equation (4.1) exhibits a hysteresis effect where multiple endemic equilibria coexist when $R_0 > 1$, as shown in Figure 4.4(a). The two “outer” equilibria are stable while the interior equilibrium (dotted red line) is unstable. Table 4.2 clarifies that three endemic equilibria coexist for $R_0 > 1$ if $\beta_0 < \beta < \beta_1$. For some parameter regimes with hysteresis, the endemic equilibria may also be found in the region where $R_0 < 1$ and where disease is not expected, as shown in Figure 4.4(b). In this scenario, (similar to a backward bifurcation) TB persists for $R_0 < 1$ even though the bifurcation at $R_0 = 1$ is a forward bifurcation. So far no other epidemic modelling study is observing this feature. This feature where hysteresis loops shifts to the left, thus crossing the epidemic threshold $R_0 = 1$ has epidemiological implications in that, although there is no backward bifurcation phenomena (as ascertained by the fact that the hysteresis effect occurs when $p < p_c$) policy makers and clinicians need to reduce the basic reproduction number below another threshold to eradicate TB. That is reducing $R_0$ below unity will be necessary but not sufficient in eradicating TB within the community.

It is important to note that hysteresis effects appear to occur for a narrow range of parameters values of $R_0$ when realistic model parameter values are used. However, the hysteresis effect can be wider for other parameter values. Hence, the occurrence of hysteresis here within a narrow range of $R_0$ values is because the parameter values selected are within a plausible and realistic range.

4.8 No reinfection path A, (i.e., $p = 0$, $\theta > 0$)

The case when there is no reinfection among recovered individuals but reinfection of exposed individuals (i.e., $p = 0$ and $\theta = 0$) was studied by Kar and Mondal [85], where an exogenous reinfection threshold was established. Thus, this section focuses on the scenario when there is no exogenous reinfection among exposed individuals (i.e., $p = 0$, reinfection path A is omitted). The goal is to determine whether recurrent TB due to reinfection of recovered individuals only ($\theta > 0$) can induce the phenomenon of backward bifurcation.
In model system (4.1) setting \( p = 0 \) yields the following subsystem:

\[
\begin{align*}
\frac{dS}{dt} &= \Lambda - \lambda S - \mu S, \\
\frac{dE}{dt} &= (1 - q)\lambda S + (1 - \sigma)\theta \lambda R - (\mu + k)E, \\
\frac{dI}{dt} &= q\lambda S + \sigma \theta \lambda R + kE - (\mu + r + \mu_d)I, \\
\frac{dR}{dt} &= rI - \theta \lambda R - \mu R.
\end{align*}
\]

(4.20)

The model system (4.20) has equilibrium points \((S^*, E^*, I^*, R^*)\) that can be expressed in terms of the force of infection \( \lambda^* \), obtained by solving

\[
P_3(\lambda^*) = b_2\lambda^{*2} + b_1\lambda^* + b_0 = 0,
\]

(4.21)

where

\[
\begin{align*}
b_2 &= (1 - q)(\mu + \mu_d)\theta + (1 - \sigma)\theta r + \theta (k + \mu q), \\
b_1 &= \theta (\mu + \mu_d)(\mu + k) + r\theta \mu (1 - \sigma) + (1 - q)(\mu + \mu_d)\mu + r(\mu + k) + \mu (k + \mu q) \\
&\quad - c\beta(q\theta \mu + \theta k), \\
b_0 &= \mu(\mu + r + \mu_d)(\mu + k)(1 - R_0).
\end{align*}
\]

Note that the subsystem (4.20) has the same basic reproduction number as model (4.1).

It is easy to see that \( R_0 = 1 \) implies \( b_0 = 0 \). Thus, the following equality is satisfied when \( R_0 = 1 \):

\[
(\mu + r + \mu_d)(\mu + k) = \beta c(k + \mu q).
\]

(4.22)

This combined with the condition \( b_1 < 0 \), which is necessary for backward bifurcation to occur, and with some algebraic manipulation leads to

\[
\theta > \frac{(1 - \sigma)(\mu + \mu_d)\mu + r(\mu + k) + \mu (k + \mu q)}{r(k + \mu \sigma)} \triangleq \theta_c
\]

(4.23)

as the required threshold for backward bifurcation.

**Theorem 4.8.1** The model subsystem (4.20) at \( R_0 = 1 \) has:
(i) **Backward bifurcation if** $\theta > \theta_c$;

(ii) **Forward bifurcation if** $\theta < \theta_c$.

Furthermore, a similar condition to (4.23) can be obtained by setting $p = 0$ in the Center Manifold results given in Appendix C equation (C2), hence corroborating Theorem 4.8.1. Thus, if $\theta > \theta_c = 5$ model (4.20) will exhibit backward bifurcation. However, if $\theta < \theta_c = 5$ model (4.20) does not exhibit backward bifurcation (i.e., has only forward bifurcation).

With the existing evidence that recovered individuals have increased susceptibility to reinfection, that is four times higher than that of new TB [28, 84], Theorem 4.8.1 suggests that the contribution of recurrent TB in the general TB burden can significantly alter TB dynamics, especially in a scenario where recurrent TB independently triggers the phenomenon of backward bifurcation.

It is important to note that although previous TB models have attempted to incorporate recurrent TB pathways they do not investigate whether recurrent TB alone can trigger bi-stability, but rather concentrate on backward bifurcation caused by exogenous reinfection of exposed individuals (i.e., $p > 0$). Selecting values of $\theta$ from the estimated interval, i.e., $\theta \in [1.61, 7.79]$ (see [91]), the threshold given in Theorem 4.8.1 is verified. Figure 4.5(a) is a bifurcation diagram corresponding to the case $\theta < \theta_c$ and indicates a forward bifurcation, as predicted by Theorem 4.8.1 case (ii). Similarly, Figures 4.5(b) and 4.5(c) indicate backward bifurcation since $\theta > \theta_c$ as predicted by Theorem 4.8.1 case (i).

### 4.9 Impact of incorporating recurrent TB parameters

Recall that recurrent TB due to reinfection is denoted by reinfection pathways $B$ and $C$ in Figure (4.1). Given that reinfection path $A$ does induce backward bifurcation when $p > p_c$ it is of interest to investigate how the additional recurrent reinfection paths $B$ and $C$ can impact the backward bifurcation. In Figure 4.6(a) the force of infection ($\lambda$) is plotted as a function of the basic reproduction number $R_0$ for scenarios with different recurrent TB contributions. The figure shows that recurrent TB due to
4.10 Summary of the chapter

van Rie et al. [138] wrote: “For decades it has been assumed that postprimary tuberculosis is usually caused by reactivation of endogenous infection rather than by a new, exogenous infection.” Until just before the turn of the century the role of exogenous reinfection in the transmission of TB was usually believed to be minimal.
Figure 4.6: (a) The effect of recurrent TB on backward bifurcation due to the incorporation of reinfection pathways B and C. Parameter values are $\mu = 0.0167, \mu_d = 0.1, k = 0.0002, r = 2, q = 0.05, \sigma = 0.05, p = 0.09, c = 60, \Lambda = 100, \beta \in [0.45, 0.7]$. With no recurrent TB ($\theta = 0$), $p = 0.09 > p_c = 0.0659$ while with recurrent TB ($\theta = 0.3$) parameters are the same but $p_c$ is altered, i.e., $p = 0.09 > p_c = 0.0647$. The blue solid lines represent stable equilibria while the red dotted line represent unstable equilibrium. (b) Contour plots of force of infection at equilibrium as a function of $\theta$ and $p$. The figure shows that increasing recurrent TB increases TB prevalence.

However, these views are no longer considered accurate and understanding of the role of exogenous reinfection has been completely revised. Warren et al. [143] refuted the unitary concept of pathogenesis of tuberculosis proposed in the 1960s, that tuberculosis results from a single infection with a single *Mycobacterium tuberculosis* strain, and such infections were thought to confer protective immunity against exogenous reinfection. Thus, exogenous reinfection was thought to be uncommon. Murray and Cohen [142] found that their data for exogenous reinfection among US immigrants strongly suggested that reinfection likely plays a major role in high-incidence TB areas.

For the proposed TB model the condition $p > p_c > q$ is necessary for backward bifurcation to occur if reinfection of recovered individuals is not considered (i.e., $F_r = 0$). However, if reinfection of recovered individuals is accounted for the condition $p > p_c > q$ does not necessarily have to hold. This is because reinfection of recovered individuals can independently induce backward bifurcation phenomenon. Lipsitch
and Murray [111] argued that backward bifurcation should not occur when taking into account biologically realistic parameter values. Their argument was built on the premise that exogenous reinfection among exposed individuals should be less than the probability of progressing to clinically active TB due to endogenous reactivation, which translates in mathematics to $p < P_L = k/(\mu + k)$. But Feng et al. [56] found that backward bifurcation can only take place if $p > P_{Feng} > P_L$. For this reason, Lipsitch and Murray [111] concluded that backward bifurcations are unlikely to be relevant in the context of TB despite the non-existence of data to support their claim. The argument of Lipsitch and Murray [111] was based on the Feng et al. [56] TB model which, as mentioned, failed to incorporate key TB pathways such as primary progression and recurrent TB due to reinfection where some recovered/treated individuals revert directly to the infective stage. Yet, these pathways are critical to TB epidemiology, an aspect that was pointed out by Lipsitch and Murray [111] as a weakness of the Feng et al. [56] model, and they called for further research that would account for these omitted pathways. However, as has been pointed out, modern research no longer seems to support the argument developed in Lipsitch and Murray [111] since exogenous reinfection among individuals with latent TB can outweigh endogenous reactivation [140, 141]. And this implies that the protection provided by latent TB infection is not strong enough to prevent individuals becoming reinfected. This is supported by the fact that the majority of new TB cases (about ninety percent) are a result of reinfection rather than endogenous reactivation [84, 140, 141, 142]. Moreover, studies of TB reinfection in high-HIV burden countries have demonstrated the strong negative impact of HIV on immunity, which is likely to outweigh the possible protection conferred by latent infection [134]. Under this scenario, backward bifurcation is likely to occur since reinfection will be greater than endogenous reactivation (i.e., $p > k/(\mu + k)$).

The analysis conducted here has shown that when exogenous reinfection is significant (resulting in a relatively small $p_c$), as is currently understood to be not atypical, backward bifurcation can indeed occur for relatively low values of the reinfection parameter $p$. Furthermore, it is observed that if the exogenous reinfection path $A$ is omitted, which is required to cause backward bifurcation in the study of Feng et al. [56], then recurrent TB alone (i.e., paths $B$ and $C$ in model equation (4.1)) may still yield backward bifurcations. For instance, if $p = 0$ (thereby omitting pathway $A$), and the recurrent TB rate parameter $\theta$ exceeds a certain threshold $\theta > \theta_c$, backward bifurcation can occur (see Figures 4.5(b) and 4.5(c)). According to the available literature on
TB, this result has not been observed in previous TB modelling studies. In addition, recurrent TB can induce forward bifurcation with hysteresis, rather than just the usual forward bifurcation. Interestingly, the hysteresis loop depends on reinfection parameters, and this may lead to an unusual scenario where the hysteresis crosses the threshold $R_0 = 1$ thus entering the region where only backward bifurcation is expected (see Figure 4.4(b)). Epidemiologically this implies that TB can persist in the community when $R_0 < 1$ even though there is only a forward bifurcation.

In the literature it has been observed that individuals who previously have had active TB and who were successfully treated are more likely to gain active TB another time [84]. However there is little understanding as to why this occurs. Gomes et al. [91] suggested two alternative mechanisms a) previous infection increases susceptibility of individuals to reinfection; b) population heterogeneity, in which some individuals are more at-risk than others, might lead to this conclusion. In this chapter, the former possibility is almost exclusively explored. However the latter possibility may also be at play as examined by Gomes et al. [91]. In the latter case, heterogeneity would be less likely to create changes in reinfection parameters such as $\theta$, and thus might not lead to the same bifurcation phenomena found in this study. Nevertheless, a wealth of theoretical studies suggest that heterogeneous infection processes are often involved in creating backward bifurcations, and thus complex dynamical phenomena can be expected as exhibited in [31, 55, 144].

In future work, it would be interesting to consider the possibility that some infected individuals who are treated do not become completely cured. This situation would lead to another cohort of individuals characterised by the fact that treatment has failed, and thus require adding an extra compartment which distinguishes complete recovery and incomplete recovery (of individuals who are infectious). The extra infectives from the latter compartment will tend to increase TB prevalence and thus widen the bifurcation curves. While this possibility is of importance, it falls outside the main scope of the present chapter, and would lead to a model that is very difficult to analyse mathematically.
Chapter 5

Heterogeneity in host susceptibility to tuberculosis and its effect on public health interventions

5.1 Chapter overview

In this chapter a tuberculosis (TB) model that accounts for heterogeneity in host susceptibility to tuberculosis is proposed, with the aim of investigating the implications this may have for the effectiveness of public health interventions. The model examines the possibility that recovered individuals treated from active TB and individuals treated with preventive therapy acquire different levels of immunity. This contrasts with recent studies that assume the two cohorts acquire the same level of immunity, and therefore both groups are reinfected at the same rate. The analysis presented here examines the impact of this assumption when designing intervention strategies. Comparison of reinfection rates between cohorts treated with preventive therapy and recovered individuals who were previously treated for active TB provides important epidemiological insights. It is found that the reinfection rate of the cohort treated with preventive therapy is the one that plays the key role in qualitative changes in TB dynamics. By contrast, the reinfection rate of recovered individuals (previously treated from active TB) plays a minor role.

Moreover, the study shows that preventive treatment of individuals during early latency is always beneficial regardless of the level of susceptibility to reinfection; the
only notable difference being the magnitude of reduction in the TB burden. Greater impacts occur when the risk of reinfection is relatively low. Further, if patients have greater immunity following treatment for late latent infection, then treatment is again beneficial. However, if susceptibility increases following treatment for late latent infection, the effect of treatment depends on the epidemiological setting. That is: (a) in (very) low burden settings, the effect on reactivation predominates and the burden declines with treatment; (b) in moderate to high burden settings the effect of reinfection predominates and burden increases with treatment. The effect is most dominant between the two reinfection thresholds, RT2 and RT1, respectively associated with individuals being treated with preventive therapy and individuals with untreated late latent TB infection, as explained in more detail below.

5.2 Introduction

After coming into contact with *Mtb*, individuals may progress directly to active infectious disease (primary progression) or enter a state of latent *Mtb* infection (LTBI) from which they may develop active disease after a variable period of time through “reactivation”. This pattern is consistent with epidemiological evidence indicating that the risk of active TB is highest in the first five years from exposure and declines thereafter, with the highest risk period being immediately after infection [134]. The risk of reinfection or superinfection with further episodes of exposure to *Mtb* is unclear, and although there is likely to be some degree of immunity to subsequent infections, little is known about the extent of protection [77, 145]. Models emphasize that understanding the degree of reduction of TB risk following previous infection in comparison to primary infection is critical to understanding the epidemiology of TB [134]. For example, following the introduction of a drug-resistant *Mtb* strain into a population where TB burden is high, the proliferation of the strain may be hampered by the size of the effective susceptible population, which may be largely determined by the level of immunity among individuals with LTBI [146]. As a consequence, if latent infection provides sufficient protection against future infection, then the rate of infection with the resistant strain will fall, markedly curtailing the TB epidemic. However, issues such as disparities in infection rates between communities burdened with human immuno-deficiency virus (HIV) make it difficult to study reinfection directly, as the detrimental impact of HIV on immunity surpasses immunity provided by latent infection [134]. There have been past attempts to estimate the risk of reinfection amongst latently infected individuals,
including through population models such as [140, 147, 148, 149], which have shown risk reductions ranging from 41% to 81%. Together, these studies suggest that partial protection (from TB) is provided against future episodes of disease.

Besides reinfection of LTBI, individuals who have had active TB but have recovered, are also at risk of reinfection. For this reason, many models incorporate a compartment accounting for recovered individuals who remain susceptible to further episodes of TB (recurrent TB). It is important to note that there are two mechanisms by which recurrent TB can occur: (i) relapse with the previously responsible strain or (ii) re-infection from a new strain of TB. The latter contribution of exogenous reinfection with \( Mtb \) (in comparison to the endogenous reactivation of LTBI) to recurrent TB is a subject that is still debated as the two mechanisms cannot be easily disentangled [91]. However, advances in clinical medicine and gene technology, such as DNA fingerprinting techniques, can now distinguish the first episode of TB from the second [78, 79, 80, 81]. Further, these techniques can determine whether a new episode of TB is caused by infection with the same strain as previously or a newly encountered strain, enabling classification of TB episodes as either relapse or reinfection, respectively. However, there is no consensus on whether recovered/treated individuals should be assigned a higher, lower or equivalent rate of infection in comparison to either latently infected or to uninfected individuals (susceptible). This raises the important question of how different levels of susceptibility across a population may interact to affect \( Mtb \) transmission dynamics. Some different approaches to exploring the impact of rates of recurrent TB adopted in the past include: assuming recovered individuals have no risk of reinfection [130, 150]; assuming relapse is responsible for all recurrent cases [94]; assuming equal risk of reinfection as for latently infected individuals [151]; assuming recovered individuals have equal rates of reinfection as for susceptible individuals [152]; incorporating both reinfection and relapse pathways after treatment [153]. Therefore, there is no consensus on whether recovered individuals have no risk of future infection, reduced risk, equal risk, or increased risk.

A previous review of recurrent TB episodes revealed that the proportion of recurrent cases that were due to subsequent infection with a new strain as opposed to relapse with the same strain varied markedly from 0% to 100% [119]. The review emphasized that relapse and reinfection should be treated as separate mechanisms and the two mechanisms are likely to be responsible for the extent of variability in results. According
to [138, 154, 155, 156], rates of reinfection after successful treatment have been found to be variable in highly endemic regions, which likely reflects the degree of continuing exposure after treatment. Estimates of rates of recurrent TB in various settings often reach several thousand per 100,000 person-years, including estimates as high as 7850 per 100,000 persons-years [92]. A meta-analysis of such studies found that reinfection rates after successful treatment are higher than the background rate of TB in the community [84].

Currently, drugs are available that can be used to treat both individuals with LTBI and individuals with active TB, with the two most important first-line drugs being isoniazid and rifampicin. These two medications are effective in the treatment of active TB disease and as preventive therapy for patients who have previously been infected but are yet to manifest symptoms. Isoniazid preventive therapy (IPT) is the most commonly used preventive regimen globally and has established efficacy in dramatically reducing a patient’s future risk of progression to active TB [157]. Past case studies of isoniazid preventive therapy (IPT) among latently TB infected individuals (conducted in South Africa gold mines) suggested that IPT is effective at the individual level, significantly reducing the risk of subsequent disease. However, the effect of IPT may be lost immediately when treatment is discontinued, which led the authors to conclude that the role of IPT at the population level is unclear. However, they also called for further research, since the effectiveness at the population level may have been compromised by a number of factors, such as post-treatment reinfection of miners or inadequately treated LTBI [158]. Other factors such as a high prevalence of HIV and silicosis, which are known to be strong risk factors for tuberculosis, may have also influenced the population level effect of IPT. Therefore, for IPT to be effective, it may need to be administered continuously amongst individuals at highest risk of TB.

Although, previous studies have considered population-level heterogeneity in susceptibility to reinfection between previously treated and latently infected persons [91, 97, 159], no previous work has considered differential susceptibility across all four possible exposure and treatment histories (i.e., fully susceptible, LTBI, treated LTBI and treated TB disease), together with the population level impact of all relevant public health interventions. Moreover, it is highly likely that the levels of susceptibility of the two previously treated populations differ considerably, given the likelihood that those treated for latent infection may retain some of the considerable immunological
5.3 Model description

Following contact with *Mtb* an individual may develop TB disease as a result of one of three possible routes. These are fast primary progression after a recent infection, endogenous reactivation of LTBI and exogenous reinfection of a previously infected individual [160]. Here a deterministic mathematical model of the transmission of *Mtb*, taking into consideration the treatment of latently infected individuals with IPT is developed. Numerous infectious diseases demonstrate considerable latent periods during which an individual harbours the disease but does not manifest symptoms and is not infectious. A key feature of TB is its long latency period. This characteristic has crucial epidemiological implications [94], and thus most mathematical models of *Mtb* transmission in the literature incorporate latent compartments [161]. Through clinical observation it has been noted that following infection with TB, different rates of progression to active TB exist and that these rates decrease with time from infection. For example, 12.9% of patients with infection confirmed with interferon-gamma release assays following exposure to a smear-positive index case progressed to active TB in 23 months [162]. By contrast, after the initial high risk period, the rate at which reactivation TB occurs is relatively low and is estimated at 5-10% over 20 years [94]. To account for these marked differences, past mathematical models devoted to tracking TB dynamics have incorporated two major pathways from susceptible to actively infected: fast and slow TB progression. In such models, a fraction of exposed susceptibles progresses directly to active TB, bypassing the latency compartment [56, 94, 146, 163]. This modelling method enables a slight modification of the standard exponential function that governs time spent in the exposed compartment [164]. Other approaches include employing a stepwise reduction in the rate of progression occurring five years after exposure [140] or an arbitrary distribution of the latent period [56].

In recent TB transmission models, compartments for both early and late latency are increasingly utilized to account for high and low risk periods following infection.
5.3 Model description

[91, 97, 165, 166, 167]. In such compartmental configurations all individuals progress to the early latent compartment following infection, after which a fraction may progress to infectious TB while the remainder transit to the low-risk late latent compartment [150, 167, 168, 169]. In consideration of the above discussion, the present study stratifies latent Mtb infection into two cohorts: a cohort at high risk of developing active TB, which is referred to as early LTBI, and a later stage of individuals with low risk for developing active TB, which shall be referred to as late LTBI. Therefore, the overall population is partitioned into six mutually exclusive classes: susceptible S which comprises individuals who have not come into contact with tuberculosis; early latently infected L1 which represents individuals who have recently been infected with Mtb (generally within a period of less than two years); late latently infected L2 which represent individuals with persistent latent TB who have contained TB infection and whose TB infection remains inactive; infectives I which represents individuals with active TB and are capable of infecting others; P which represents individuals who are being or who have been treated with isoniazid preventive therapy; recovered R which represents individuals who were previously infected and have been successfully treated. The total population is assumed to be large enough to be modelled deterministically and random mixing is assumed.

For the sake of mathematical tractability, here it is assumed that the birth rate compensates for TB-induced and background mortality (similar to the simplification used in some of the classical studies in the field, as for example in some of the key studies of Blower et al. [94, 117, 130, 167] and Dye et al. [170]). Thus, \( \lambda = \mu + dI \) is the recruitment rate and all state variables are expressed as a fraction of the total population. The susceptible population comprises of individuals who enter into this compartment at a rate \( \lambda \) and they diminish as individuals are infected with Mtb at a density-dependent infection rate \( \beta I \), where \( \beta \) is the transmission coefficient. Newly infected individuals enter the early latent compartment L1 and it is assumed that a proportion of individuals in the early latent compartment are detected following screening for TB and are treated with IPT at rate \( \theta \), progressing to compartment P. A proportion of individuals in the early latent compartment progress to the active TB compartment I at a rate \( f \phi \), while the remaining proportion proceeds to the late latent compartment at a rate \( (1 - f) \phi \).
Individuals in the late latent compartment may also receive IPT and thus progress to compartment $P$ at a rate $\rho$. Furthermore, individuals in the late latent compartment can transit into the infectious compartment $I$ due to endogenous reactivation of their latent TB at a rate $\eta$.

Only persons in the I compartment are infectious, and as such compartments $L_1$, $L_2$ and $P$ do not contribute to the force of infection. Therefore, the infectious compartment is generated by fast progression of TB, endogenous reactivation from late latency and relapse of recovered individuals at a rate $\omega$. The subpopulation is diminished when individuals are successfully treated at rate $\tau$ or as a result of spontaneous recovery (self cure) at a rate $\alpha$.

Previously infected individuals may be fully susceptible to exogenous reinfection and infected at the same rate as the susceptible population ($S(t)$), or partially immune or have no immunity against reinfection. Consequently, late latently infected individuals, individuals treated with IPT and recovered individuals are reinfected at rates $\sigma_i \beta$ (where $i=1,2,3$), respectively, with $\sigma_i \in [0, 1]$ ($i = 1, 2, 3$) accounting for partial immunity against exogenous reinfection. Note that $\sigma_i = 1$ ($i = 1, 2, 3$) corresponds to a scenario where late LTBI, treated LTBI and recovered individuals are infected at the same rate as susceptible individuals, while $\sigma_i > 1$ ($i = 1, 2, 3$) implies that all post-infection cohorts have increased susceptibility to reinfection in comparison to susceptible individuals. This would also correspond to some past studies which have shown that individuals who have recovered from TB infection are more susceptible to future infection and in such a scenario $\sigma_i > 1$, $i = 1, 2, 3$ [84].

All individuals experience natural death at a constant rate $\mu$, except infectious individuals who suffer an additional TB-induced death at rate $d$. Transitions between compartments are shown diagrammatically in Figure 5.1. Combining the aforementioned assumptions, the following system of nonlinear ordinary differential equations
Figure 5.1: Schematic representation of the model. The square boxes represent classification of the general population into six mutually exclusive subpopulations, i.e., susceptibles $S(t)$, early latents $L_1(t)$, late latents $L_2(t)$, individuals treated with isoniazid preventive therapy $P(t)$, individuals with active TB $I(t)$, and recovered individuals $R(t)$. All arrows indicate either inflow or outflow or transition between compartments. Blue arrows illustrate transition of latently infected individuals as a result of treatment with IPT. Red dashed arrows show reinfection of late latently infected individuals, individuals treated with IPT and recovered individuals, respectively represented by $\sigma_i$, $i = 1, 2, 3$.

govern the model:

\[
\begin{align*}
\frac{dS}{dt} &= \lambda - \mu S - \beta IS, \\
\frac{dL_1}{dt} &= \sigma_1 \beta I L_2 + \sigma_2 \beta I P + \sigma_3 \beta I R + \beta IS - (\theta + \mu + \phi) L_1, \\
\frac{dL_2}{dt} &= (1 - f)\phi L_1 - (\mu + \eta + \rho + \sigma_1 \beta I) L_2, \\
\frac{dI}{dt} &= \phi f L_1 + \eta L_2 + \omega R - (\mu + d + \tau + \alpha) I, \\
\frac{dP}{dt} &= \theta L_1 + \rho L_2 - (\mu + \sigma_2 \beta I) P, \\
\frac{dR}{dt} &= (\tau + \alpha) I - (\mu + \omega + \sigma_3 \beta I) R.
\end{align*}
\]
The proposed model equations are different from the recently published model by Ragonnet et al. [166], in that each reinfection pathway is explicitly distinguished and individuals treated with IPT are not distinguished according to their time since infection. Moreover, since both early and late latently infected individuals are treated with isoniazid preventive therapy, instead of having two compartments for each as in [166], in this study the two compartments are coalesced into a single compartment for parsimony. Another important paper [97] incorporated treatment of early and late latent individuals but assumed that individuals treated with IPT and recovered individuals have identical risks of reinfection after recovery (i.e., $\sigma_2 = \sigma_3$). In the present study this assumption is relaxed by adding another compartment of individuals treated with IPT, so that the risk of exogenous reinfection can be varied between late latently infected individuals, individuals treated with IPT and recovered individuals. The motivation behind this is that there is a reasonable estimate of the value of $\sigma_1$ (see [91, 97, 121, 134, 171, 172]), whereas $\sigma_2$ and $\sigma_3$ are highly uncertain.

The parameter values used in investigating the aforementioned objectives are selected from the relevant literature on TB epidemic models. The natural death rate $\mu$ is set to correspond to an average lifespan of 70-80 years [97]. From [173], the duration of TB from the first onset of TB symptoms to treatment or death is approximately three years. Consequently, both parameter $d$ and $\alpha$ are estimated by assuming that $d + \alpha = 1/3$ and $2d \approx \alpha$. Thus, $d$ is taken as $d = 1/9 \approx 0.1$. From evidence that about 5-10% of the infected population manifest active TB shortly after infection [82, 174], parameter $f$ is set to 0.05-0.1. Parameter $\phi$ is selected from a range of values $\phi \in [1.5, 12]$ [82, 91, 97, 166, 167]. The rate of endogenous reactivation among untreated late latent individuals is taken as $\eta = 0.0002$ per year, relapse among those who were previously cured through either therapeutic interventions or spontaneous cure is set to $\omega = 0.00002$ per year; both adopted from [140, 147]. The relative risk of reinfection among untreated late latent individuals, $\sigma_1$ is fixed at 0.25 as in [97, 121], with the justification that it agrees with the maximum level of immunity rendered by BCG (Bacille Calmette-Guérin) vaccination [175] (although the effects of varying this parameter from its baseline value are explored in detail below). The parameter $\sigma_2$ corresponds to the relative risk of reinfection among individuals treated with IPT, while $\sigma_3$ corresponds to the relative risk of reinfection among recovered individuals. Exploring the effects of varying these highly uncertain parameters (including their epidemiological effects and their influence on the effectiveness of public health interventions) is the primary
5.3 Model description

The purpose of this study. The baseline parameter value for therapeutic intervention among individuals manifesting TB symptoms is set at $\tau = 2$ per year, which corresponds to a mean duration of infectiousness of six months [97] (which implicitly assumes that the R compartment incorporates those currently under treatment for active disease). Last, the transmission coefficient $\beta$ is varied over a wide range. A summary of the parameters and their respective values are shown in Table 5.1.

Table 5.1: Parameters and definitions for model equation (5.1).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>baseline value</th>
<th>range</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>Transmission coefficient</td>
<td>–</td>
<td>0-500 yr$^{-1}$</td>
<td>[91, 97]</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Natural death rate</td>
<td>1/70 yr$^{-1}$</td>
<td>–</td>
<td>[91, 97, 166, 167]</td>
</tr>
<tr>
<td>$d$</td>
<td>TB-induced death rate</td>
<td>0.1 yr$^{-1}$</td>
<td>–</td>
<td>[56, 173]</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Rate at which infected individuals exit early latent compartment $L_1$</td>
<td>12 yr$^{-1}$</td>
<td>1.5-12</td>
<td>[91, 97, 166, 167]</td>
</tr>
<tr>
<td>$f$</td>
<td>Fraction of TB infected population that progress to active TB soon after infection</td>
<td>0.05</td>
<td>0.05-0.1</td>
<td>[174]</td>
</tr>
<tr>
<td>$\eta$</td>
<td>Rate of endogenous reactivation for late latents</td>
<td>0.0002 yr$^{-1}$</td>
<td>–</td>
<td>[151, 176]</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Treatment rate of active TB</td>
<td>2 yr$^{-1}$</td>
<td>–</td>
<td>[91, 97, 177]</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Spontaneous cure/self cure</td>
<td>2/9 yr$^{-1}$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Treatment rate of LTBI $L_1$ with IPT</td>
<td>1</td>
<td>variable</td>
<td>[97]</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Treatment rate of LTBI $L_2$ with IPT</td>
<td>0.1</td>
<td>variable</td>
<td>[97]</td>
</tr>
<tr>
<td>$\omega$</td>
<td>Rate of relapse following recovery</td>
<td>0.00002 yr$^{-1}$</td>
<td>–</td>
<td>[140, 147]</td>
</tr>
</tbody>
</table>

Levels of susceptibility

$\sigma_1$: Multiplier for exogenous reinfection for latent $L_1$, 0.25, 0.25-1, [91, 134]
$\sigma_2$: Multiplier for exogenous reinfection for population treated with IPT, 0.5, 0.25-2, [84, 91, 134]
$\sigma_3$: Multiplier for exogenous reinfection for recovered population, 0.5, 0.25-2, [84, 91, 134]
5.4 Basic reproduction number $R_0$

In epidemic theory the basic reproduction number, denoted by $R_0$, is one of the most important model quantities, given its ability to predict the triggering of an epidemic. $R_0$ is defined as the number of secondary infections that would occur when a single infectious individual is introduced into an entirely susceptible population, and considered over the lifetime of the disease. Following the method in [25] the basic reproduction number for the model system (5.1) as computed in Appendix D is given as:

$$R_0 = \frac{\beta(\mu + \omega)\phi(f(\mu + \rho) + \eta)}{((\mu + d)(\mu + \omega) + \mu(\tau + \alpha))(\mu + \eta + \rho)(\theta + \mu + \phi)}.$$  (5.2)

(See Appendix E for biological interpretation of this expression (5.2) for $R_0$.) In general it is known that a value of $R_0 < 1$ implies that each individual is only able to infect less than one individual on average, such that the disease will die out, whereas a value of $R_0 > 1$ implies that each individual is able to infect more than one individual and that endemic disease will persist within the population. Hence, $R_0 = 1$ is a crucial epidemic threshold in determining the epidemic trajectory.

Figure 5.2(a) illustrates that in the complete absence of reinfection pathways ($\sigma_1 = \sigma_2 = \sigma_3 = 0$), the model dynamics are quite simple. When $R_0 < 1$ there is a disease free equilibrium ($I^* = 0$), or DFE, and when $R_0 > 1$ there is one endemic equilibrium. (Note the logarithmic scale in the Figure which hides the DFE). For this scenario (i.e., no reinfection pathways), the DFE is shown in Appendix F to be globally asymptotically stable (g.a.s). The epidemiological implication of DFE being g.a.s is that TB will be eliminated from the community if the threshold quantity $R_0$ is maintained (or decreased) to a value below unity.

5.5 Reinfection threshold

With the introduction of reinfection pathways, the non-linear dynamics of model equation (5.1) yield a different bifurcation structure (see [121]). Figure 5.2(b) which is obtained by plotting $I^*$ versus the transmission coefficient $\beta$ for the full model (equation (5.1)), illustrates the existence of a low endemic region (low $I^*$ for $\beta < 200$) and a high endemic region (high $I^*$ for $\beta > 200$). The red dotted line marks the endemic threshold.
5.5 Reinfection threshold

which occurs at the point $R_0 = 1$. However, besides the endemic threshold at $R_0 = 1$, there are other “reinfection thresholds” (here when $\beta \approx 200$) which play a critical role in determining endemic equilibria [97, 121]. A reinfection threshold is said to occur in a model when rates of reinfection are just sufficient to maintain an endemic disease in the absence of contributions from other pathways (i.e., primary infection, $\beta IS = 0$ and reactivation mechanisms, $\eta = \omega = 0$). Often, disease prevalence increases by two orders of magnitude when transmission increases across the reinfection threshold [121, 171], as seen in Figure 5.2(b). The reinfection threshold partitions the transmissibility axis into two regions: the low and high endemic regions. The studies of [121, 171] introduced a technique that is helpful for mathematically identifying reinfection thresholds (see Appendix H).

![Bifurcation diagrams](image)

**Figure 5.2:** Bifurcation diagrams of equilibrium TB prevalence as a function of the transmission coefficient $\beta$. (a) In the absence of reinfection the epidemic threshold occurs at $R_0 = 1$. (b) In the presence of reinfection four thresholds emerges. That is the threshold that occurs at $R_0 = 1$ and the reinfection thresholds RT1, RT2 and RT3. (c) Bifurcation diagram of the reinfection submodel (5.3). The red dotted vertical line marks the transcritical bifurcation point $R_0 = 1$, that is the endemic threshold, while the black vertical dashed lines mark the points where reinfection thresholds RT1, RT2 and RT3 converge. Parameters used are the baseline values in Table 5.1. Reinfection susceptibility parameters are shown in the figures. In all figures a semi-logarithmic scale is used for a clear view.

The reinfection thresholds of model (5.1) are approximated by analysing special submodels that distinguish reinfection from other transmission processes such as primary infection, reactivation and relapse [178]. The first step is to set primary infection,
endogenous reactivation and relapse to zero (i.e., $\beta IS = 0$ and $\eta = \omega = 0$). It is then possible to approximate three reinfection thresholds from the following respective submodels:

(i) RT1: The threshold due to reinfection during late latency ($L_2$). In this scenario reinfection of recovered individuals and those previously treated for latent infection are switched off (i.e., $R(t) = P(t) = 0$);

(ii) RT2: The threshold due to reinfection of individuals previously treated for LTBI ($P$). In this scenario, reinfection of recovered individuals and those with latent infection are switched off (i.e., $R(t) = L_2(t) = 0$);

(iii) RT3: The threshold due to reinfection of recovered individuals ($R$). In this scenario, reinfection of latently infected individuals and those previously treated for LTBI are switched off (i.e., $L_2(t) = P(t) = 0$).

The following example gives calculations for finding the first reinfection threshold RT1 in model (5.1) using the procedure outlined in [178] (see Appendix H). As mentioned, reactivation and primary infection mechanisms are set to zero ($\beta IS = 0$ and $\eta = \omega = 0$), and post-infection levels of population immunity risk are assumed to be homogeneous or equal (that is $\sigma_1 = \sigma_2 = \sigma_3$). In this configuration, the rates of infection of compartments $L_2, P$ and $R$ become equivalent, so that it is possible to merge these three compartments. The reinfection submodel is

$$
\frac{d(L_2 + P + R)}{dt} = \lambda + (1 - f)\phi L_1 + (\tau + \alpha)I + \theta L_1 - \mu(L_2 + P + R) - \sigma_1 \beta (L_2 + P + R),
$$

$$
\frac{dL_1}{dt} = \sigma_1 \beta (L_2 + P + R) - (\theta + \mu + \phi)L_1,
$$

$$
\frac{dI}{dt} = \phi f L_1 - (\mu + d + \tau + \alpha)I.
$$

The Jacobian matrix of the reinfection submodel (5.3) evaluated at the disease free equilibrium $(1, 0, 0)$ is then

$$
J_R = \begin{pmatrix}
-\mu & (1 - f)\phi + \theta & (d + \tau + \alpha) - \sigma_1 \beta \\
0 & -(\theta + \mu + \phi) & \sigma_1 \beta \\
0 & \phi f & -(\mu + d + \tau + \alpha)
\end{pmatrix}.
$$
Setting the determinant of the Jacobian matrix (5.4) to zero and evaluating $\beta$ in terms of model parameters yields the critical value of the first reinfection threshold,

$$\beta = \frac{1}{\phi f} \left( \frac{1}{\sigma_1} \left( \theta + \mu + \phi \left( \mu + d + \tau + \alpha \right) \right) \right) = RT1. \quad (5.5)$$

Parameters are taken to be $\phi = 12$ and $\theta = 1$ while others remain as shown in Table 5.1. This yields $RT1 \approx 201$ (see Figures 5.2(b) and 5.2(c) where RT1 is marked with a black dotted line). The reinfection threshold expressed in terms of $R_0$ (this is $R_0$ for the full model) is obtained by substituting (5.5) into equation (5.2), leading to

$$R_0^{RT1} = \frac{1}{\sigma_1 f \phi} \frac{(\mu + d + \tau + \alpha)(\mu + \omega)(f \phi (\mu + \rho) + \eta \phi)}{\left(\mu + d + \tau + \alpha\right)(\mu + \eta + \rho)}. \quad (5.6)$$

Note that, if the reactivation and relapse mechanisms are now set equal to zero ($\omega = \eta = 0$) then expression (5.6) reduces to $R_0 \approx 1/\sigma_1$ which is equivalent to the simplest form of reinfection threshold in terms of $R_0$, as originally obtained by [121].

The other reinfection thresholds $RT2$ and $RT3$ are computed similarly as shown in Appendix I. The equilibrium of the reinfection submodel (5.3) can be easily obtained by setting the right-hand terms to zero and evaluating for $I^*$ as

$$I^* = \frac{f \phi \left[ \beta - \frac{(\theta + \mu + \phi \left( \mu + d + \tau + \alpha \right))}{\sigma_1 f \phi} \right]}{\beta \left[ f \phi + (\mu + d + \tau + \alpha) \right]}. \quad (5.7)$$

(See Appendix J for other steady states.)

It is interesting to see how well the submodel approximates the behaviour of the full model in the vicinity of the reinfection threshold $RT1$. Figure 5.2(c) plots the equilibrium $I^*$ for the submodel equations as a function of $\beta$. Above $\beta = 201$ there is a positive endemic equilibrium but below this value only the disease-free equilibrium is present. Thus, the theoretically predicted reinfection threshold for the submodel is confirmed to be $RT1 = 201$.

The success of the prediction for $RT1$ can be gauged by returning to Figure 5.2(b). The high endemic zone occurs when $\beta \gtrsim 200$, that is there is a transition from low to high TB burden with the proportion of active TB increasing by about two orders of magnitude when $\beta \gtrsim 200$ (as can also be observed in [97, 121]). That is, there is a
100-fold change in prevalence associated with a very small change in $\beta$ value.

### 5.5.1 Homogeneous reinfection risk ($\sigma_1 = \sigma_2 = \sigma_3 < 1$)

First, it is useful to reflect on how the reinfection parameters $\sigma_i$ ($i = 1, 2, 3$) should be interpreted by re-examining a typical infection term $\sigma_i \beta$ in equation (5.1). Note that when the risk of reinfection parameter $\sigma_i < 1$ ($i = 1, 2, 3$), it corresponds to a scenario where individuals susceptible to reinfection have partial immunity, while $\sigma_i = 1$ ($i = 1, 2, 3$) corresponds to susceptibility to infection being the same as for a typical susceptible individual. Shortly a scenario where the risk of reinfection parameters can in some cases be greater than unity $\sigma_i > 1$ will also be discussed. That is, individuals who have already been infected have the same or even higher risks of reinfection, when compared with a typical susceptible individual who has never been infected. This indicates that individuals may have increased susceptibility to tuberculosis and is biologically plausible, e.g., due to local tissue damage to the respiratory tract impairing innate immunity. Supposing $\sigma_i < 1$ ($i = 1, 2, 3$), the bifurcation diagram in Figure 5.2(b) indicates a scenario where all cohorts susceptible to reinfection have equal risk of reinfection (that is reinfection risk parameters are set to $\sigma_1 = \sigma_2 = \sigma_3 = 0.25$). When interpreting Figure 5.2(b) recall that the vertical axis representing $I^*$ has a logarithmic scale. The reinfection thresholds $RT_1 = RT_2 = RT_3 \approx 200$ divide the transmissibility axis into low and high endemic regions. For $\beta \lesssim 50$ (equivalently $R_0 < 1$), $I^* = 0$. For $50 \lesssim \beta \lesssim 200$ then $I^* \approx 10^{-3}$, when $\beta \approx 200$ then $I^* \approx 10^{-2}$ and when $\beta > 200$ then $I^* \to 10^{-1}$.

### 5.6 Heterogeneity in susceptibility to reinfection

Now the effects of reinfection parameters $\sigma_2$ and $\sigma_3$ on TB dynamics are investigated. Suppose that $\sigma_1$ is fixed to 0.25 while either $\sigma_2$ or $\sigma_3$ are selected such that they are equal or greater than $\sigma_1$. Letting $\sigma_1 = 0.25, \sigma_2 = 0.5, \sigma_3 = 0.25$ results in Figure 5.3(a) which shows that increasing $\sigma_2$ results in an increase in TB prevalence. However, by setting $\sigma_1 = 0.25, \sigma_2 = 0.25$ and $\sigma_3 = 0.5$ results in Figure 5.3(b) which shows that $\sigma_3$ has little effect on TB prevalence in comparison to the same increase in $\sigma_2$ (see Figure 5.3(a)).
Figure 5.3: Illustration of the relative importance of $\sigma_2$ and $\sigma_3$ to equilibrium dynamics. Parameters used are the baseline parameters in Table 5.1. (a) Impact of increasing $\sigma_2$ while both $\sigma_1$ and $\sigma_3$ are fixed. (b) Effect of increasing $\sigma_3$ while both $\sigma_1$ and $\sigma_2$ are fixed.

Further, model dynamics are explored over a wider range of reinfection parameters by again modifying the relative risks of reinfection among individuals treated with IPT ($\sigma_2$) and those previously recovered from active TB ($\sigma_3$). The relative rate of reinfection among LTBI is fixed to $\sigma_1 = 0.25$, consistent with the pertinent literature described above and with [97, 121, 171]. The remaining two risk of reinfection parameters, ($\sigma_2$ and $\sigma_3$) are varied and may take values of 0.125, 0.50 and 1.5, thereby creating heterogeneity in susceptibility to TB transmission. From left to right the three columns of panels in Figure 5.4 show an increasing risk of reinfection among recovered individuals ($\sigma_3$) while from top to bottom each row of the figure shows an increasing risk of reinfection among individuals treated with IPT ($\sigma_2$).

It is observed that as $\sigma_2$ increases (i.e., moving from top to bottom of each column in Figure 5.4), there is a structural change in the bifurcation curve. As $\sigma_2$ increases (from top to bottom) TB prevalence rises. In contrast, within each row of panel 5.4 there is no significant qualitative change in the bifurcation structure as $\sigma_3$ is varied (parameter values $\sigma_3 = 0.125, 0.50, 1.50$) i.e., moving from left to right.
Moreover, considering a scenario where either individuals treated with IPT or recovered individuals (or both) have a significant loss of immunity by readjusting $\sigma_2$ and $\sigma_3$ such that they can take values greater than one, results in bi-stability phenomena whereby TB can be endemic below the threshold $R_0 = 1$. However, the occurrence of backward bifurcation is attributed to $\sigma_2$ and not $\sigma_3$ as illustrated in the last row of panels of Figure 5.4 where backward bifurcation sets in when $\sigma_2$ is greater than one. Further simulations show that $\sigma_3$ has minimal effect as can be observed in Figures 5.4(c) and 5.4(f) appearing in the last column of panel Figure 5.4. However, it is important to note that at the present time, the lack of sensitivity of model dynamics to $\sigma_3$ is unclear and is a subject that will be examined in future research.

The observation made here regarding $\sigma_2$ implies that the reinfection parameter accounting for reinfection among individuals treated with IPT plays a key role in determining TB dynamics.

### 5.6.1 Other features: hysteresis

Besides the backward bifurcation phenomena observed in the last row of Figure 5.4, the proposed model equation (5.1) exhibits hysteresis effects. Hysteresis is a phenomenon whereby multiple equilibria, both stable and unstable occur simultaneously above the epidemic threshold $R_0 = 1$ (see [39]). Selecting a set of parameters such that both $\sigma_2 = \sigma_3 = 0.125$, and $\sigma_1 = 0.25$ while other parameters remain as the baseline values shown in Table 5.1, results in Figure 5.5(a) which shows a hysteresis phenomenon. In Figure 5.5(a) the unstable equilibrium is marked by a red dotted line that separates two stable equilibria: low endemic and high endemic. Rather similar to a backward bifurcation, there can be jumps between the two stable equilibria. In the regime where the contact rate is approximately $\beta = 253$ there is a low endemic equilibrium. A small epidemiological change, such as a slight rise in $\beta$, (which pushes $I$ above the unstable equilibrium marked by the red dashed line) may trigger a jump to the high endemic equilibrium.

### 5.6.2 No partial immunity

As discussed, the degree of protection conferred by initial infection with TB is still controversial. Gomes et al. [91] and Verver et al. [84] suggest that the risk of reinfection parameters $\sigma_1, \sigma_2, \sigma_3$ can in some cases be close to or greater than unity, so that
Figure 5.4: Exploration of the full range of plausible values of the relative risk of reinfection parameters. All panels show equilibrium prevalence as a function of the transmission coefficient $\beta$. From left to right, each column shows an increase in the relative risk of reinfection among recovered individuals $\sigma_3$, while from top to bottom each row shows increasing risk of reinfection among individuals treated with IPT $\sigma_2$. Other parameter values remain at baseline values in Table 5.1. The red dotted vertical lines mark the point where $R_0 = 1$. The dotted red segments of the endemic curves represent the unstable equilibria while the blue lines represent stable equilibria.

individuals who have already been infected may have the same or even higher risks of reinfection as compared to typical susceptible individuals who have never been infected. As discussed above, this situation seems less intuitive, but remains plausible.
5.7 Interventions that impact reinfection

5.7.1 Effect of treating early LTBI

The effect of treating early LTBI for different levels of susceptibility to reinfection is examined. Recall that treatment of early latent TB with preventive therapy (IPT)
is modelled by the parameter $\theta$ which is the treatment rate of early LTBI $L_1$. First, assume that the relative rate of reinfection among late LTBI ($\sigma_1$) is less than the levels of reinfection of both recovered individuals ($\sigma_3$) and individuals treated with IPT ($\sigma_2$). Consider risk of reinfection parameters $\sigma_2 = \sigma_3 = 0.5$ and $\sigma_1 = 0.25$. Figure 5.6(a) illustrates model dynamics for different values of the treatment parameter $\theta$, and shows that treatment of early latent TB decreases TB prevalence regardless of the higher risk of reinfection (i.e., $\sigma_2 > \sigma_1$ and $\sigma_3 > \sigma_1$). Now consider a scenario where the level of susceptibility to reinfection of late LTBI is high in comparison to the levels of susceptibility to reinfection of both recovered and individuals treated with IPT. Selecting $\sigma_2 = \sigma_3 = 0.125$ and $\sigma_1 = 0.25$, yields the bifurcation diagram seen in Figure 5.6(b). Again it is observed that treatment of early LTBI via $\theta$ decreases TB prevalence. Note that the magnitude of TB reduction is relatively stronger when $\theta > 1$. These same results are plotted with linear scales in Figures 5.6(c) and 5.6(d).

### 5.7.2 Effect of treating late LTBI

It is important to note that Gomes et al. [97] investigated a scenario where intervention (i.e., treatment of late LTBI) is assumed to increase or decrease risks of reinfection i.e., the values of $\sigma_i$. However, their study did not distinguish between individuals treated with IPT and individuals previously assumed to have recovered due to antibiotic treatment or self-cure. As stated above, the present study distinguishes individuals who have recovered from active TB from those who are being or have been treated with IPT. Thus, contrary to Gomes et al. [97] where only two groups that were subject to reinfection were considered, the model presented here has three such cohorts. Distinguishing between individuals treated for active TB and individuals treated with IPT provides a more comprehensive analysis of treating late LTBI when the population is subjected to different levels of reinfection.

Recall that the parameter $\rho$ represents treatment of late LTBI with IPT. Thus, bifurcation diagrams obtained for different levels of reinfection of individuals treated with IPT (modified by $\sigma_2$) and recovered individuals (modified by $\sigma_3$) to reinfection of late LTBI (modified by $\sigma_1$) are compared. Panels are again presented with $\sigma_2$ increasing from top to bottom while $\sigma_3$ increases from left to right (see Figure 5.7). First assuming intervention decreases the level of susceptibility to reinfection it is assumed that both recovered individuals and individuals treated with IPT have a lower rate of reinfection in comparison to late LTBI. By letting $\sigma_2 = \sigma_3 = 0.125$ while...
5.7 Interventions that impact reinfection

Figure 5.6: Impact of treating early latent individuals under different risks of susceptibility to reinfection. (a) Treatment of early LTBI (and of active disease) leads to increased susceptibility to reinfection compared to late latent infection. (b) The level of susceptibility to reinfection among late LTBI is lower than for both individuals treated with IPT and recovered individuals treated from active TB. Figures (c) and (d) respectively represent figures (a) and (b) with a linear scale y-axis. Treatment of late LTBI and individuals with active TB are respectively set at $\rho = 0.1$ and $\tau = 2$ while other parameters are as shown in Table 5.1.

$\sigma_1 = 0.25$, Figure 5.7(a) is obtained. The baseline scenario is obtained by setting $\rho = 0$ while the extreme case is illustrated by assuming $\rho \to \infty$. Figure 5.7(a) illustrates
that in a scenario where \( \sigma_2, \sigma_3 < \sigma_1 \) treatment of late LTBI reduces TB prevalence. This decrease in TB prevalence is largely attributable to the general reduction in susceptibility to reinfection (small \( \sigma_2 \) and \( \sigma_3 \)). In addition, Figure 5.7(a) indicates the existence of bi-stable equilibria (hysteresis effect) in which stable equilibria are separated by an unstable equilibrium (dashed red line). Performing another simple numerical experiment can shed some light on the differences between reinfection parameters \( \sigma_2 \) and \( \sigma_3 \). Hence, assuming \( \sigma_2 < \sigma_1 \) and \( \sigma_3 > \sigma_1 \), and plotting TB prevalence as a function of \( \beta \) results in Figure 5.7(b), which is almost the same as Figure 5.7(a). That is, treatment of late LTBI is beneficial despite \( \sigma_3 = 0.5 \) being double \( \sigma_1 \). These results imply again that \( \sigma_3 \) is unimportant, while \( \sigma_2 \) is the main parameter of interest. The respective reinfection thresholds associated with individuals treated with IPT (RT2), recovered individuals (RT3) and late LTBI (RT1) are marked on Figure 5.7 with black dotted lines. Considering the regions bounded by reinfection thresholds, it is clear from Figure 5.7(b) that treatment has the most beneficial impact within the region bounded by RT1 and RT2, and the position of RT3 has little influence.

Secondly, another set of risk of reinfection parameters is considered such that the rate of reinfection among recovered individuals (\( \sigma_3 \)) is less than the rate of reinfection of late LTBI (\( \sigma_1 \)) while the level of reinfection among individuals treated with IPT (\( \sigma_2 \)) is relatively high (\( \sigma_1 = 0.25, \sigma_2 = 0.5 \) and \( \sigma_3 = 0.125 \)). This set of parameters leads to Figure 5.7(c) which shows that treatment of late LTBI now leads to an increase in TB prevalence. It is evident from Figure 5.7(c) that the reinfection thresholds RT2 and RT1 bound the parameter space where treatment of late LTBI has most impact. Outside this region treatment has a minor effect.

Finally, assuming susceptibility to reinfection among individuals treated with IPT and recovered individuals increases after treatment, results in the findings presented in Figure 5.7(d) where \( \sigma_2 = \sigma_3 = 0.50 \) while \( \sigma_1 = 0.25 \). Figure 5.7(d) again shows that increasing treatment of late LTBI may lead to an increase in TB prevalence when \( \sigma_2, \sigma_3 > \sigma_1 \). Note that the greatest increase in TB prevalence also occurs between the reinfection thresholds RT2 (=RT3) and RT1. This implies that the relative magnitude of risk of reinfection, \( \sigma_1 \) compared to \( \sigma_2 \), is vital in determining whether treatment will increase or decrease TB prevalence.
5.7 Interventions that impact reinfection

The previously described study [97] investigated the effect of treating late latent TB under two assumptions: a) susceptibility to reinfection increases after treatment of late LTBI, and b) susceptibility to reinfection decreases after treatment of late LTBI. They found that if treatment of late LTBI increases the risk of reinfection of recovered individuals $\sigma_2$ then TB prevalences increases, while if treatment of LTBI is assumed to decrease risk of reinfection of recovered individuals $\sigma_2$, then TB prevalence decreases [97]. As mentioned above, in Gomes et al. [97] both individuals treated for active TB and individuals treated with preventive therapy were assumed to be indistinguishable and therefore classified as one cohort of recovered individuals.

However, it is important to identify the precise epidemiological parameter responsible for the observed changes in qualitative dynamics. The findings in Figure 5.7 suggest that distinguishing the relative susceptibility to reinfection between individuals treated with IPT from those who have been cured from active TB is important to elucidate the complex dynamics of TB. As observed in the formulated model (5.1), $\sigma_3$ is essentially redundant when it comes to predicting whether treatment will decrease or increase the prevalence of TB. Hence, epidemiological studies interested in understanding the impact of preventive therapy (IPT) should focus on quantifying the risk of reinfection among persons treated with IPT in comparison to latently infected individuals (modified by $\sigma_2$).

Further, it is important to investigate a scenario where late latently infected individuals have the same level of reinfection as individuals treated with IPT and recovered individuals. Thus, considering $\sigma_1 = \sigma_2 = \sigma_3 = 0.25$ while varying treatment results in Figure 5.8(a) which depicts that treatment has a positive impact between the region bounded by $R_0 = 1$ and $RT_1 (=RT_2=RT_3)$ and minimal impact above the reinfection thresholds $RT_1 (=RT_2=RT_3)$. Figure 5.8(b) is obtained using the same parameter values as Figure 5.8(a) except that the reactivation mechanisms are switched off (i.e., $\omega = \eta = 0$). It is observed that all the endemic curves under different treatment values merge. However, further exploration of a scenario where reactivation mechanisms are neglected while levels of reinfection are unequal (i.e., $\sigma_1 < \sigma_2 = \sigma_3$) results in Figure 5.8(d) which is qualitatively similar to when reactivation mechanisms are included (see Figure 5.8(c)). This observation suggests that reactivation pathways do not play a significant role when it comes to determining the outcome of treatment of late LTBI; rather it is the rate of reinfection and particularly reinfection of individuals treated...
5.7 Interventions that impact reinfection

with IPT that greatly influence treatment outcome.

Within the medical literature the endogenous reactivation mechanism (i.e., parameter $\eta$) is well known to be very small as it takes over 20 years for an individual with latent TB to develop active TB following initial infection [118]. Moreover, the lifetime risk of a latently infected individual to progress to the infectious stage is approximately 5-10% [118]. Thus the choice of parameter $\eta = 0.0002yr^{-1}$ is in line with the current TB literature (see [97, 140, 147, 177]). Hence, in a scenario where I compared $\eta = 0.0002yr^{-1}$ with $\eta = 0$ it is because selecting a value that is above $0.0002yr^{-1}$ would be choosing a value that is outside the estimated range. See also Gomes et al. [97] where they considered different levels of reinfections under two scenarios: $\eta = 0.0002yr^{-1}$ and $\eta = 0$. On the other hand reinfection is known to be much higher in comparison to endogenous reactivation. The available literature on TB supposes that the majority of new TB cases (about 90%) occur as a result of reinfection rather than endogenous reactivation [134, 140, 141, 142]. Thus, reinfection is the process that has most influence on TB dynamics. Moreover, setting $\eta = 0$ was a simplifying assumption to check whether the results deduced regarding impact of reinfection in the presence of preventive treatment hold.
Figure 5.7: Effect of treatment for late latent infection under the assumption that treatment decreases (panels (a) and (b)) and increases (panels (c) and (d)) susceptibility to reinfection, and assuming that treatment for active disease decreases (panels (a) and (c)) and increases (panels (b) and (d)) susceptibility to reinfection. Treatment of individuals with active TB is fixed to $\tau = 2$, and treatment of individuals with early latent TB is fixed to $\theta = 1$, while treatment of late latent individuals is introduced at different rates; $\rho = 0, 0.1, 1, 10$ and the limit as $\rho \to \infty$. Other parameters used are as shown in Table 5.1. (a) $\sigma_2, \sigma_3 < \sigma_1$. (b) $\sigma_2 < \sigma_1 < \sigma_3$. (c) $\sigma_3 < \sigma_1 < \sigma_2$. (d) $\sigma_1 < \sigma_2, \sigma_3$. In the figures the black dashed endemic lines represent the baseline case where there is no treatment of late LTBI (i.e., $\rho = 0$) while the blue dashed endemic lines represent immediate treatment for the entire population ($\rho \to \infty$). In both figures (a) and (b) the dashed red lines of the endemic curves represent unstable equilibria while blue solid lines represent stable equilibria. All figures are plotted using a semi-logarithmic scale.
5.7 Interventions that impact reinfection

Figure 5.8: Treatment of late latent infection under the assumption of no effect on susceptibility to reinfection (panels (a) and (b)) and increased susceptibility (panels (c) and (d)), with reactivation mechanisms present (panels (a) and (c)) and removed (panels (b) and (d)). Parameters are $\tau = 2$, $\theta = 1$, $\rho = 0, 0.5, 1$ and $\rho \to \infty$ while other parameters remain as in Table 5.1. The risks of reinfection are shown in each sub-figure.
5.8 Summary of the chapter

In this chapter, a mathematical TB model accounting for heterogeneity in susceptibility to reinfection has been proposed. Analysis of the model yielded the following results:

(i) It was found that the risk of reinfection among individuals treated with IPT (i.e., $\sigma_2$) plays a central role in qualitative changes in model dynamics, particularly in shifting TB prevalence between low and high burden values. In contrast, the risk of reinfection (i.e., $\sigma_3$) among the recovered cohort who were previously treated from active TB (or self-cure) plays an insignificant role in terms of the qualitative dynamics of the model. See Figure 5.4;

(ii) Treatment of early latent infection is shown to be always beneficial irrespective of the level of reinfections among cohorts subject to reinfection. The benefit is strongest if individuals treated with IPT and recovered individuals previously treated from active TB, have a relatively low risk of reinfection i.e., compared to late LTBI. That is, when $\sigma_2 < \sigma_1$ and $\sigma_3 < \sigma_1$; see Figure 5.6(b) or Figure 5.6(d) (given in linear scale);

(iii) Similar to previous findings [97], the assumption that treatment decreases the risks of reinfection among both cohorts of individuals treated with IPT and individuals recovered from active TB was considered. Under this assumption, treatment of late latency TB individuals has a positive impact (see Figure 5.7(a)) and therefore may be highly synergistic with other interventions;

(iv) Alternatively, the assumption that treatment of late latent infection increases the risk of reinfection among individuals treated with IPT and recovered individuals was also considered. This yields contrasting results to case (iii) above. That is, treatment of late LTBI increases TB prevalence, although treatment is more detrimental above the reinfection threshold, particularly in an intermediate prevalence zone lying between RT2(=RT3) and RT1. See Figure 5.7(d);

(v) Assume now that treatment of late latently infected individuals increases the risk of reinfection of individuals treated with IPT (i.e., $\sigma_2 > \sigma_1$) but decreases the risk of reinfection among recovered individuals (i.e., $\sigma_3 < \sigma_1$). The results still show that increasing treatment of late LTBI increases the prevalence of active TB (see Figure 5.7(c)). This observation suggests that the parameter $\sigma_2$ associated with reinfection of individuals treated with preventive therapy is the one that
5.8 Summary of the chapter

determines whether treatment of late latency will decrease or increase prevalence of active TB. Hence, $\sigma_2$ is a key epidemiological parameter;

(vi) It is observed that reactivation mechanisms (in particular reactivation from late latent infection ($\eta$) and from recovered individuals ($\omega$)) play a minimal role in determining treatment outcomes. See Figures 5.8(a) and 5.8(b);

(vii) Finally, reinfection of previously infected persons can lead to unusual dynamics, such as backward bifurcation and hysteresis effects. This is epidemiologically important because it could lead to extreme changes in disease burden following relatively minor epidemiological changes.

The conclusions made in this chapter show differences to previous work [97] because of the different form of the proposed model and its more realistic structure. This previous study, Gomes et al. [97], assumed that individuals who were previously treated for active TB and individuals treated with preventive therapy are not differentiable and therefore they can be treated as a single group. In particular, $\sigma_2$ and $\sigma_3$ were coupled as a single parameter, obscuring the understanding of the role of each individual parameter. However, this study shows that relaxing this assumption yields new epidemiological findings. This follows from the fact that it is impossible to tell from Gomes et al. [97] whether an increase in prevalence of active TB is attributed to reinfection of individuals being treated or who have been treated with IPT, or due to reinfection of individuals treated from active TB. Consequently, this study refines the results of Gomes et al. [97] by pointing out the precise parameter attributed to an increase or decrease in prevalence of active TB as treatment increases. Thus, epidemiological studies with the ability to quantify $\sigma_2$ accurately would be important in shedding light on TB dynamics. Currently little is known about the true value of $\sigma_2$ which is associated with reinfection of individuals being treated or who have been treated with preventive therapy (IPT).

5.8.1 World Health Organization (WHO) TB burden estimates

TB is present in every region and country of the world but its distribution varies greatly with the most highly endemic countries reporting rates of disease around 1000 per 100,000 per year, while the least endemic countries have rates as low as 5 per 100,000 per year. There can be little doubt that this observation of 200-fold differences in
5.8 Summary of the chapter
disease burden relates in part to heterogeneity in socio-economic development, living conditions, prevalence of comorbidities and the strength of health systems. However, given such a huge gulf in disease rates, additional factors may well be at work. Here it is postulated that decreased susceptibility to reinfection in comparison to first infection acts to create a threshold effect, which can lead to a 100-fold increase in burden once crossed. Similarly, disease may be considerably easier to control once prevalence has dropped below the reinfection threshold and entered the low endemic, controllable zone. That is, while socio-economic development and improvements in treatment programs could explain gradual decreases in burden, this additional phenomenon may help to explain more dramatic shifts. For example, the recent rapid declines in TB burden in China and other countries of East Asia could be partly attributable to this threshold effect.

Empirical evidence for the role of reinfection heterogeneity is difficult to find, given the difficulty in obtaining high-quality data on TB burden relative to the slow speed with which the epidemic evolves. However, many regions of the world appear to show significant divides between high and low burden countries (Figure 5.9). Although, this is not clearly apparent in all regions and an overall threshold is not evident (Figure 5.10), this grouping of countries is arguably seen in current WHO data [1]. The absence of a clear divide could relate to factors such as comorbidities (e.g., HIV infection), differences in health systems and socio-economic development, as well as the fact that TB transmission frequently occurs over a much smaller scale than a nation state. Therefore, the reported overall burden for individual countries actually represents the summation of many heterogeneous sub-epidemics, particularly for large countries such as China and India.

5.8.2 Limitations and conclusion
Mathematical modelling is an important tool for epidemiologists since it provides insights into underlying processes where empirical epidemiological observations generally cannot. With the proposed model, the key parameter influencing preventive treatment outcomes is identified. It was found that the reinfection parameter accounting for reinfection of individuals treated with IPT ($\sigma_2$), and not the parameter accounting for reinfection of recovered individuals ($\sigma_3$), alters treatment outcome. Further, changes in $\sigma_2$ can either be beneficial or detrimental when there are treatment programs. The need to quantify the parameter is important if epidemiologists are to accurately estimate
its effect on TB dynamics. Moreover, it is observed that reactivation mechanisms (in particular reactivation from late latent infection and from recovered individuals) play a minimal role in determining treatment outcomes.

Moreover, it is imperative to emphasize that the result obtained regarding unimportance of $\sigma_3$ is not related to relative equilibrium sizes e.g., $R^*$ and $P^*$. Reason being that different parameter sets were explored and results deduced lead to the same conclusion. A clear illustration can be observed in the Figure 5.3 where the relative importance of $\sigma_2$ and $\sigma_3$ on equilibrium prevalence are compared by the equilibrium prevalence curve where $\sigma_1 = \sigma_2 = \sigma_3 = 0.25$. In Figure 5.3(a) $\sigma_2$ is increased to 0.5 while both $\sigma_1$ and $\sigma_3$ are fixed to 0.25. In such set of parameters $\sigma_2$ is observed to increase the equilibrium prevalence $I^*$ relative to the baseline curve. On the other hand when $\sigma_3$ is increased to 0.50 while both $\sigma_1$ and $\sigma_2$ are fixed to 0.25 the equilibrium curve is almost merging with the baseline curve (see Figure 5.3(b)). This stresses that, the unimportance of $\sigma_3$ is not driven by different choice of parameter values. Furthermore, a broad range of parameter values for both $\sigma_2$ and $\sigma_3$ were also considered. In the panel Figure 5.4, parameter $\sigma_3$ was assumed to increase from left to right taking values 0.125, 0.50 and 1.50. And it is clearly observed that, there is no significant change in bifurcation structure as $\sigma_3$ is varied from left to right. This is contrary to when $\sigma_2$ is varied through similar values, where it is observed that the bifurcation structure changes from hysteresis to forward and finally to backward.

The proposed model has some limitations that can be addressed in future studies. In the proposed model, individuals being treated from early and late latent compartments are coalesced into a single compartment so as to simplify the model (although the effect of distinguishing these groups following preventive treatment has been previously explored [166]). This could be addressed in conjunction with the approach for distinguishing four susceptibility categories as studied here. Finally, incomplete efficacy of preventive treatment could be reasonably considered and explored with the proposed model as a simple reduction in treatment rates.
Figure 5.9: Illustration of world TB burden by regions across the world obtained using global TB data [1]. Note that the size of the circle is proportional to the population size of the country.
Figure 5.10: Illustration of global TB burden by country using global TB report data [1].
Chapter 6

Other related work: analysis of a heroin epidemic model with nonlinear treatment function

6.1 Chapter overview

A mathematical model is developed that examines how heroin addiction spreads in society. The model has similarities to epidemic equations that study the spread of infectious diseases. The model takes into account the treatment of heroin users by incorporating a realistic functional form that “saturates” representing the limited availability of treatment. In these circumstances, bifurcation analysis reveals that the model has an intrinsic backward bifurcation whenever the saturation parameter is larger than a fixed threshold. Since the objective of this chapter is to study the proposed model’s global stability it is important to note that in the absence of backward bifurcations, Lyapunov functions can often be found and used to prove global stability. However, in the presence of backward bifurcations, such Lyapunov functions may not exist or may be difficult to construct. In this chapter a geometric approach to global stability is applied to derive conditions that ensure the system is globally asymptotically stable. Numerical simulations are also presented to give a more complete representation of the model dynamics. Sensitivity analysis performed by Latin hypercube sampling (LHS) suggests that the effective contact rate in the population, the relapse rate of heroin users undergoing treatment, and the extent of saturation of heroin users are mechanisms fuelling heroin epidemic proliferation. However, in the long term relapse
of heroin users undergoing treatment, i.e., going back to a heroin using career, has slightly higher impact on heroin usage than effective contact rate.

6.2 Introduction

In 1897, Germany’s Bayer pharmaceutical company synthesised heroin, and soon after marketed the product as a non-addictive miracle drug, for use as a cough syrup and pain reliever [179]. Cough medicine was in fact in high demand, since tuberculosis and pneumonia were fast-spreading diseases of the time. As such, the miracle drug heroin was rapidly disseminated across the globe. Fast forward to today, it is known that addiction to heroin is an extremely common phenomena among heroin users; some 23% of individuals who consume the drug become dependent on it. World wide, many countries are affected by the heroin drug-trafficking industry and its growing number of users. America is currently in the midst of another heroin epidemic [180] with approximately 700,000 Americans using heroin in the past year [180]. The number of people using heroin for the first time is increasing at an alarming rate, with > 150,000 Americans engaging in heroin use in 2012, which is almost double that recorded in 2006 [180]. Heroin also leads to other diseases and is considered a major pathway responsible for fuelling proliferation of Human immunodeficiency virus (HIV) and Hepatitis B and C virus (HBV, HCV) [181, 182].

The development of heroin habituation and addiction has similar characteristics to an epidemic, in terms of its disturbingly contagious spread through a susceptible population. In the last decades, a whole range of mathematical models have been developed to forecast how diseases spread in time and space, and how they can be controlled. Recently, the same mathematical modelling techniques have been extended for the purpose of understanding and combating drug addiction problems. The aim of the present study is to propose a novel heroin epidemic model and make use of it to study issues arising with treatment as well as establish conditions that may signal heroin persistence within the community.

The ultimate goal of mathematical epidemiology is to understand how to control and eliminate infectious diseases and these ideas have a place for also dealing with social problems. In epidemic theory the basic reproduction number, usually denoted by $R_0$ is one of the most important concepts, given its ability to predict the course of an epidemic.
6.2 Introduction

It will also prove invaluable in the study of heroin dynamics in society. $R_0$ is defined as the number of secondary infections that are likely to occur when a single infectious individual is introduced into an entirely susceptible population [183]. Until recently, it has been widely accepted that the condition $R_0 < 1$ is an essential requirement for the eradication of a disease. However, this viewpoint has been challenged with a number of theoretical studies demonstrating that this criterion may not always be sufficient. Instead, the phenomenon of backward bifurcation offers a different interpretation since it shows that although the basic reproduction number is below unity and the infection free equilibrium is stable, there might still be another stable endemic equilibrium and unstable endemic equilibrium coexisting simultaneously. Thus even though $R_0 < 1$, a population may still reside at an endemic equilibrium in which the disease persists indefinitely. In a scenario where multiple equilibria concurrently exist the extinction or persistence of an epidemic is dependent on the initially infected size of subpopulations. The qualitative features of backward bifurcation are illustrated by Figure 6.1.

A variety of behavioural and pharmacological medications can be administered to effectively treat heroin addiction. The side effects associated to quitting using heroin (such as pain, diarrhoea, nausea and vomiting) are very severe and very often compel heroin addicts to relapse. To prevent such cases there are available medications that can be administered during the detoxification stage to relieve craving and physical symptoms. A number of studies have established that pharmacological therapy has positive impact in facilitating drug addicts to remain in treatment programs. Furthermore, it has been noted that during addiction treatment there is a decline in drug consumption, infectious disease transmission and crime rates [180]. In this chapter a model incorporating a saturated treatment function is proposed and threshold conditions that indicate when heroin is able to persist within a community are derived. Besides incorporation of a saturated treatment function, the proposed heroin model will also include an extra class of individuals, namely those who have been successfully treated. This class has been neglected in previous heroin epidemic models [184, 185, 186, 187]. Much of the work in this chapter will focus on exploring the conditions for global stability of the heroin model with treatment. The analysis deals with global stability of a heroin model with a density-dependent incidence rate, “self cure”, relapse and saturated treatment function using the Bendixson criterion.
With this in mind, the SIR (Susceptible-Infectives-Recovered) model by Wang et al. [186] is extended to represent a heroin epidemic model for which global stability properties are investigated. The proposed model will be used to study global stability for the non-trivial equilibrium states by using two distinct approaches: The Lyapunov direct method and Li and Muldowney’s [188, 189] geometric approach to global stability. It is with no doubt that the famous Lyapunov direct method is a powerful tool for nonlinear stability analysis [190]. One of the main advantages of Lyapunov’s direct technique is that it is directly applicable to nonlinear systems [191]. However, a major challenge is that it requires an auxiliary function which is often hard to construct. And this difficulty is exacerbated especially if the model exhibits backward bifurcation phenomena because Lyapunov functions for such models may not exist. To address these difficulties another powerful tool; the geometric technique due to Li and Muldowney was developed in the mid nineties [98, 191, 192]. Their method involves a generalization of Bendixson’s criterion to systems of any finite dimension and applies compound matrices. Presently, this method has gained popularity due to its vast range of applications, in particular to mathematical models that are of biological interest. Although this method is mainly applied in epidemic models (for instance see [188, 189, 193, 194, 195, 196]) its use can be found in other population dynamics contexts (see [197]). It has been shown in [189] that the geometric technique is more appropriate for mathematical models of SEIR-like structure since their analysis can be easily reduced to a three dimensional system. Nevertheless, the method has been extended to four dimensional systems that may be difficult to reduce, though applications to four dimensional systems are rare because the procedure becomes mathematically involved when \( n \geq 4 \). Examples can be found in the work of Ballyk and coworkers who applied compound matrices to a four dimensional population model [198] and also by Gumel and coworkers [199] who studied a SVEIR (Susceptible-Vaccinated-Exposed-Infected-Recovered) model of severe acute respiratory syndrome (SARS) epidemic spread.

The four dimensional model studied here can be reduced to a three dimensional system. Both the Lyapunov direct method and the geometric approach are applied to investigate global properties of a four dimensional heroin epidemic model. The Lyapunov direct method will be applied in a special case, where the parameter that triggers bi-stability phenomena is switched off. On the other hand the geometric approach will be applied in the general model where all parameters are present including the one that causes
bi-stability. Here the procedure in [189, 191] is followed to obtain sufficient conditions for global stability.

![Backward Bifurcation Diagram](image)

**Figure 6.1**: Illustration of the qualitative features of backward bifurcation. The red dotted line represents the unstable equilibrium (i.e., unstable endemic equilibria and unstable heroin-free equilibrium) while the blue solid lines represent stable equilibria (i.e., stable endemic equilibria and stable heroin-free equilibrium).

### 6.3 Model formulation

In the spirit of the SIR (Susceptible-Infected-Recovered) model in the literature (i.e., [88]), a heroin epidemic model is formulated based on the assumption that heroin use follows a process that can be modelled similarly to infectious diseases [200, 201]. The general population is stratified into four mutually exclusive classes, namely susceptibles (S), individuals successfully treated from heroin use (U₃), heroin users undergoing treatment (U₂) and heroin users not in treatment (U₁). The proposed heroin epidemic model is based on key assumptions which include:
6.3 Model formulation

- Uniform mixing: individuals in the above mentioned classes freely interact with each other;
- Individuals undergoing treatment are still often using drugs [202];
- Heroin users in treatment relapse to heroin users not in treatment as a result of the self decision to terminate treatment [203];
- Heroin users in treatment do not infect susceptibles;
- Individuals who have been successfully cured are not susceptible to relapse, either through a simple per-capita flow or as a result of an interaction with existing heroin users.

These are simplifying assumptions that are made largely to make the model analytically tractable. Given these assumptions the heroin model may be described by the processes illustrated in Figure 6.2, which can be written in terms of the following set of equations:

\[
\begin{align*}
\frac{dS}{dt} &= \Lambda - \beta U_1 S - \mu S, \\
\frac{dU_1}{dt} &= \beta U_1 S + p U_2 - (\mu + \delta_1 + \xi) U_1 - T(U_1), \\
\frac{dU_2}{dt} &= T(U_1) - (p + \sigma + \delta_2 + \mu) U_2, \\
\frac{dU_3}{dt} &= \sigma U_2 + \xi U_1 - \mu U_3.
\end{align*}
\]

(6.1)

In brief, the susceptible sub-population \( S(t) \) is generated at a constant rate through immigration and birth at rate \( \Lambda \). Some susceptible individuals who come into contact with heroin users \( U_1(t) \) may begin to use heroin. Hence, the susceptible population is diminished due to contact with heroin users at rate \( \beta U_1 S \), while heroin users increase at the same rate. Heroin users also increase when those undergoing treatment relapse at rate \( p U_2 \), and return to their heroin using lifestyle. Heroin users reduce in number as a result of treatment which is represented by the treatment function \( T(U_1) \). Moreover, the user subpopulation is reduced by heroin-induced death at rate \( \delta_1 U_1 \) as well as a result of the self decision to cease using heroin (also referred as “self-cure”) at rate \( \xi U_1 \). Individuals undergoing treatment are diminished through relapse to heroin using at rate \( p U_2 \), heroin-induced death at rate \( \delta_2 U_2 \) and successful treatment at rate \( \sigma U_2 \). Finally, the recovered/successfully treated subpopulation \( U_3(t) \) is generated when heroin users undergoing treatment are successfully cured and also through “self-cure”. All
subpopulations are decreased by natural death via the background mortality parameter \( \mu \).

Heroin epidemic models studied to date [184, 185, 186, 187] assume the classical view that the treatment rate of the infective population should be proportional to the number of infective individuals [23]. This view was criticised during the SARS (Severe Acute Respiratory Syndrome) outbreaks in 2003. The dramatic increase of SARS cases in Beijing challenged the normal public-health system because it was only possible to treat a limited number of SARS patients at a given time. The experience with SARS sparked a renewed interest among modellers to investigate the implication of the capacity of the health-care system. Wang and Ruan [204] considered an SIR epidemic model and assumed a Heaviside treatment function while Wang [205] restudied the same SIR model but assumed a piecewise linear treatment function. Here it is assumed that the heroin users \( U_1(t) \) receive treatment based on the following more general saturated treatment function:

\[
T(U_1) \triangleq \frac{\alpha U_1}{1 + \omega U_1},
\]

where \( \alpha \) is positive and \( \omega \) is nonnegative. In the proposed model the parameter \( \omega \) accounts for the extent of saturation of heroin users. Note that for small \( U_1 \) the treatment function reduces to \( T(U_1) \approx \alpha U_1 \) while for large \( U_1 \) it reduces to \( T(U_1) \approx \alpha/\omega \) which actually characterizes the saturated phenomena of the treatment. Further, if \( \omega = 0 \), the treatment function becomes \( T(U_1) = \alpha U_1 \) which is the usual linear treatment rate. \( 1/(1 + \omega U_1) \) is a measure of inhibition due to a saturation of heroin users who are usually too many to be dealt with given the limited available treatment.

A summary of the model variables and parameters is given in Table 6.1.

### 6.4 Basic properties and basic reproduction number

Since the proposed model involves a human population, the model must be able to ensure that all the associated parameters and the state variables \( S, U_3, U_2, U_1 \) are nonnegative for all time \( t > 0 \). Hence, the following result:
Figure 6.2: A heroin epidemic model with a density-dependent incidence rate and a saturated treatment function. The blue solid arrows represent deaths either due to heroin or natural causes while the black solid arrows represent change of status from one compartment to another.

Table 6.1: Description of variables and parameters of model (6.1).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$</td>
<td>Number of susceptible individuals at time $t$</td>
</tr>
<tr>
<td>$U_3$</td>
<td>Number of heroin users who have been successfully treated from heroin use, as well as individuals who have voluntarily stopped using heroin (and have withdrawal symptoms) at time $t$</td>
</tr>
<tr>
<td>$U_2$</td>
<td>Number of heroin users undergoing treatment at time $t$</td>
</tr>
<tr>
<td>$U_1$</td>
<td>Number of drug users not undergoing treatment at time $t$ i.e., the initial and relapsed heroin users</td>
</tr>
<tr>
<td>$N$</td>
<td>Total population at time $t$ ($N = S(t) + U_3(t) + U_2(t) + U_1(t)$)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Lambda$</td>
<td>Recruitment rate of individuals in the general population entering the susceptible population</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Effective contact rate</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Death rate due to natural causes</td>
</tr>
<tr>
<td>$p$</td>
<td>Rate at which individuals undergoing treatment relapse to heroin use</td>
</tr>
<tr>
<td>$\xi$</td>
<td>\textit{&quot;Self cure&quot;} rate at which heroin users stop using heroin and join the successfully cured class of individuals not taking heroin</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>Heroin-related death rate of heroin users not in treatment</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>Heroin-related death rate of individuals undergoing treatment</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Rate at which heroin users in treatment are successfully cured (i.e., completely detoxicated) from heroin use</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Rate at which heroin users are treated</td>
</tr>
<tr>
<td>$\omega$</td>
<td>The extent of saturation of heroin users within the community</td>
</tr>
</tbody>
</table>
6.4 Basic properties and basic reproduction number

**Theorem 6.4.1** Let the initial conditions supplied to model (6.1) be such that $S(0) > 0, U_3(0) > 0, U_2(0) > 0$ and $U_1(0) > 0$. Then the trajectories $(S(t), U_3(t), U_2(t), U_1(t))$ of the model (6.1), with nonnegative initial conditions, will remain nonnegative for all time $t$.

**Proof.** Let $t_1 = \sup\{t > 0 : S(t) > 0, U_3(t) > 0, U_2(t) > 0, U_1(t) > 0\} > 0$. Now from the first equation of model (6.1) it follows that

$$\frac{dS}{dt} = \Lambda - \beta U_1 S - \mu S = \Lambda - \Phi S - \mu S \quad \text{(where} \quad \Phi = \beta U_1),$$

which can be written as

$$\frac{d}{dt} \left\{ S(t) \exp\left[ \mu t + \int_0^t \Phi(\tau)d\tau \right] \right\} = \Lambda \left\{ \exp\left[ \mu t + \int_0^t \Phi(\tau)d\tau \right] \right\}.$$

Hence,

$$S(t_1) \exp\left[ \mu t_1 + \int_0^{t_1} \Phi(\tau)d\tau \right] - S(0) = \int_0^{t_1} \Lambda \left\{ \exp\left[ \mu x + \int_0^x \Phi(\tau)d\tau \right] \right\} dx,$$

so that

$$S(t_1) = S(0) \exp\left[ -\mu t_1 - \int_0^{t_1} \Phi(\tau)d\tau \right]$$

$$+ \left\{ \exp\left[ -\mu t_1 - \int_0^{t_1} \Phi(\tau)d\tau \right] \right\} \times \int_0^{t_1} \Lambda \left\{ \exp\left[ \mu x + \int_0^x \Phi(\tau)d\tau \right] \right\} dx > 0.$$

Following a similar procedure it can be shown that $U_3(t) > 0, U_2(t) > 0, U_1(t) > 0$ for all time $t > 0$. Thus, all trajectories of model (6.1) remain positive for all non-negative initial conditions, as required. □

Now the region where model (6.1) is considered to be biologically feasible is established. Summing all the equations of the basic model (6.1) yields

$$\frac{dN(t)}{dt} = \Lambda - \mu N(t) - \delta_1 U_1(t) - \delta_2 U_2(t).$$

(6.2)
Considering that \(0 < U_1(t) < N(t), 0 < U_2(t) < N(t)\) and letting \(\bar{\delta} = \max\{\delta_1, \delta_2\}\), it follows from equation (6.2) that
\[
\Lambda - (\mu + 2\bar{\delta})N(t) \leq \frac{dN(t)}{dt} \leq \Lambda - \mu N(t).
\]
Therefore
\[
\frac{\Lambda}{\mu + 2\bar{\delta}} \leq \liminf_{t \to \infty} N(t) \leq \limsup_{t \to \infty} N(t) \leq \frac{\Lambda}{\mu}
\]
such that \(\limsup_{t \to \infty} N(t) \leq \frac{\Lambda}{\mu}\).

**Theorem 6.4.2** The closed set
\[
\mathcal{U} = \left\{ (S, U_3, U_2, U_1) : 0 \leq S, U_3, U_2, U_1; S + U_3 + U_2 + U_1 \leq \frac{\Lambda}{\mu} \right\}
\]
is positively invariant and absorbing with respect to the set of nonlinear differential equations (6.1).

**Proof.** Here it is shown that the feasible solutions of model (6.1) are uniformly bounded in the region \(\mathcal{U}\). Suppose \(S, U_3, U_2, U_1\) is any solution of the system (6.1) supplied with nonnegative initial conditions. Then it is straightforward to note that the total population \(N\) satisfies the inequality
\[
\frac{dN}{dt} = \Lambda - \mu N - \delta_1 U_1 - \delta_2 U_2 \leq \Lambda - \mu N.
\]  
(6.3)
From (6.3) it follows that \(\frac{dN}{dt} \leq \Lambda - \mu N\) which implies \(\frac{dN}{dt} \leq 0\) if \(N \geq \frac{\Lambda}{\mu}\). The standard comparison theorem [106] can be used to deduce that \(N(t) \leq N(0)e^{-\mu t} + \frac{\Lambda}{\mu}(1 - e^{-\mu t})\). In particular \(N(t) \leq \frac{\Lambda}{\mu}\) if \(N(0) \leq \frac{\Lambda}{\mu}\) for all \(t > 0\). Thus, under the flow induced by system (6.1), the region \(\mathcal{U}\) is positively invariant. Furthermore, for \(N(0) > \frac{\Lambda}{\mu}\) the trajectory solution \(N(t)\) enters either in the region \(\mathcal{U}\) in finite time or asymptotically approaches \(\frac{\Lambda}{\mu}\). Thus, in the region \(\mathcal{U}\), model (6.1) is said to be mathematically and epidemiologically well posed [20] and the solution of all the trajectories generated by
model (6.1) are considered to be in the biologically feasible region \( \mathcal{D} \). □

Clearly system (6.1) has an intrinsic heroin free equilibrium (HFE) given by \( D^0 = (S_0, 0, 0, 0) \), a scenario representing a heroin free state in the community. \( S_0 = \frac{\Lambda}{\mu} \) represents the number of susceptibles when no one is using heroin. The basic reproduction number denoted by \( R_0 \) is defined as the number of secondary infections that are likely to be triggered by a single infectious individual when introduced into a wholly susceptible population [20]. Here \( R_0 \) is interpreted as the mean number of secondary cases of heroin users generated by a typical heroin user not in treatment during his/her duration of heroin use in a population of potential drug users.

To obtain the basic reproduction number it can be easily observed that the average time an individual spends as a heroin user without treatment is \( T_0 = \frac{1}{\mu + \delta_1 + \xi + \alpha} \) and the probability of surviving this compartment and moving to the treatment compartment is \( T_1 = \frac{\alpha}{\mu + \delta_1 + \xi + \alpha} \). Now the probability of surviving treatment and then returning to the heroin users class is \( T_2 = \frac{p}{p + \sigma + \delta_2 + \mu} \). Thus, the total average time spent as a heroin user (not in the treatment compartment) on multiple passes, can be obtained as

\[
T = T_0[1 + T_1T_2 + (T_1T_2)^2 + \ldots].
\] (6.4)

Clearly, the terms inside the square brackets in (6.4) constitute a geometric sequence (see Appendix K for detailed derivation) and therefore expression (6.4) can be written as

\[
T = \frac{(p + \sigma + \delta_2 + \mu)}{\mu + \delta_1 + \xi + \alpha(p + \delta_2 + \sigma + \mu) + \alpha(\sigma + \delta_2 + \mu)}.
\] (6.5)

Multiplying (6.5) with the effective contact rate \( \beta \) and the average recruitment rate \( \frac{\Lambda}{\mu} \) yields a heroin basic reproduction number as

\[
R_0 = \frac{\beta \Lambda(p + \delta_2 + \mu + \sigma)}{\mu \alpha(\mu + \delta_2 + \sigma) + \mu(\mu + \delta_1 + \xi)(p + \delta_2 + \mu + \sigma)}.
\] (6.6)

It is easy to observe that \( R_0 \) is inversely proportional to treatment \( \alpha \), which implies that if treatment rate is maintained sufficiently high it can control a heroin epidemic (by reducing \( R_0 \) to less than one). However, as it will be seen later, when parameter \( \omega \)
(representing the extent of saturation of heroin users) is accounted for, this control is no longer guaranteed.

**Theorem 6.4.3** The HFE is locally asymptotically stable provided \( R_0 < 1 \), otherwise it is unstable.

This general result has been reviewed in [25] and thus not proved again here. The theorem implies that heroin users will disappear from the community when \( R_0 < 1 \) if the initial sizes of the sub-populations of system (6.1) are in the basin of attraction of the heroin free equilibrium.

**Remark 1** It is instructive to note that the basic reproduction number does not include the parameter \( \omega \) that accounts for the extent of saturation of heroin users. In what follows, the endemic equilibria of the model, where the parameter \( \omega \) plays a key role in the emergence of bi-stability, are analysed.

### 6.4.1 Endemic equilibria

Within the context of the proposed heroin model the endemic equilibrium refers to a state when heroin addiction is maintained over long time-scales in the population. Therefore \( S^*, U_3^*, U_2^*, U_1^* > 0 \) holds. To obtain the endemic equilibria \( E^* = (S^*, U_3^*, U_2^*, U_1^*) \) equations (6.1) are set to zero and equilibrium quantities \( S^*, U_3^* \) and \( U_2^* \) are solved in terms of \( U_1^* \). That is

\[
S^* = \frac{\Lambda}{\beta U_1^* + \mu},
\]

\[
U_3^* = \frac{\sigma \alpha U_1^* + \xi(p + \sigma + \delta_2 + \mu)(1 + \omega U_1^*)U_1^*}{\mu(p + \sigma + \delta_2 + \mu)(1 + \omega U_1^*)},
\]

\[
U_2^* = \frac{\alpha U_1^*}{(p + \sigma + \delta_2 + \mu)(1 + \omega U_1^*)}.
\]

Substituting (6.7) into the second equation of (6.1) and factoring out the solution \( U_1^* = 0 \) yields

\[
f(U_1^*) = AU_1^{*2} + BU_1^* + C = 0,
\]

(6.8)
where
\[
A = (\mu + \delta_1 + \xi)(p + \sigma + \delta_2 + \mu)\beta \omega,
\]
\[
B = (\mu + \delta_1 + \xi)(p + \sigma + \delta_2 + \mu)\beta + \mu \omega (\mu + \delta_1 + \xi)(p + \sigma + \delta_2 + \mu)
\]
\[
+ \alpha(\sigma + \delta_2 + \mu)\beta - \beta \Lambda \omega (p + \sigma + \delta_2 + \mu),
\]
\[
C = [(\mu + \delta_1 + \xi)(p + \sigma + \delta_2 + \mu)\mu + \alpha \mu (\sigma + \delta_2 + \mu)](1 - R_0).
\]

The quadratic equation (6.8) can be analysed to investigate the existence of multiple equilibria when the basic reproduction number is below unity.

If the parameter that accounts for the extent of saturation of heroin users in model (6.1) is excluded, that is \( \omega = 0 \), (6.8) reduces to a linear equation
\[
U_1^* \hat{B} + C = 0,
\]
where
\[
\hat{B} = (\mu + \delta_1 + \xi)(p + \sigma + \delta_2 + \mu)\beta + \alpha (\sigma + \delta_2 + \mu)\beta
\]
so that model (6.1) has the unique solution
\[
U_1^* = \frac{-C}{\hat{B}} = \frac{\mu}{\beta}(R_0 - 1)
\]
which is nonnegative if and only if \( R_0 > 1 \). Hence, if \( \omega = 0 \), model (6.1) has a unique endemic equilibrium whenever \( R_0 > 1 \) and this equilibrium approaches zero as \( R_0 \) tends to one \( (R_0 \to 1+) \) because \( C \to 0 \). But there is no positive endemic equilibria if \( R_0 < 1 \). These results are summarized in the following Lemma:

**Lemma 6.4.1** The heroin epidemic model (6.1) when \( \omega = 0 \) has a unique positive (endemic) equilibrium \( E_1^* = (S^*, U_3^*, U_2^*, U_1^*) \) whenever \( R_0 > 1 \) and no positive endemic equilibrium otherwise.

In what follows the global stability for both the HFE and the unique endemic equilibrium \( E_1^* \) for the case \( \omega = 0 \) is investigated.

### 6.4.2 Global stability for heroin-free equilibrium when \( \omega = 0 \)

To investigate global stability the method presented by Castillo-Chavez et al. [206] is applied. First let \( \mathcal{X} = (S, U_3) \) and \( \mathcal{Y} = (U_2, U_1) \) with \( \mathcal{X} \in \mathbb{R}^2 \) representing the number
of individuals not using heroin and \( Y \in \mathbb{R}^2 \) representing the number of individuals using heroin (i.e., heroin users in treatment and heroin users not in treatment). Now suppose

\[
\begin{align*}
X' &= F(X, Y), \\
Y' &= G(X, Y), \\
G(X, 0) &= 0,
\end{align*}
\]

where \( X' \) and \( Y' \) denote differentiation with respect to time. The HFE is now denoted by \( D^0 = (X^0, 0) \), where \( X^0 = (S^0, 0) \). The following conditions (H1) and (H2) have to be met to guarantee global asymptotic stability:

- **(H1)** For \( X' = F(X, 0) \), \( X^0 \) is globally asymptotically stable (g.a.s);
- **(H2)** \( G(X, Y) = BY - \hat{G}(X, Y) \), where \( \hat{G}(X, Y) \geq 0 \), for \((X, Y) \in \mathcal{U}\).

\( B = D_2 G(X, 0) \) and \( \mathcal{U} \) is the region where the model (6.1) is biologically realistic. Then, Castillo-Chavez et al. [58] have shown that the following Lemma is satisfied.

**Lemma 6.4.2** The fixed point \( D^0 = (X^0, 0) \) is a g.a.s equilibrium of model (6.1) provided that \( R_0 < 1 \) (locally asymptotically stable) and that assumptions (H1) and (H2) hold.

Now consider the following Theorem:

**Theorem 6.4.4** Suppose \( R_0 < 1 \). Then the HFE \( D^0 \) is g.a.s.

**Proof** Let \( X = (S, U_3) \) and \( Y = (U_2, U_1) \), and

\[
D^0 = (X^0, 0) \text{ where } X^0 = \left( \frac{\Lambda}{\mu}, 0 \right).
\]

Then,

\[
X' = F(X, Y) = \begin{pmatrix} \Lambda - \beta U_1 S - \mu S \\ \sigma U_2 + \xi U_1 - \mu U_3 \end{pmatrix}.
\]

It is straightforward to see that at the heroin free equilibrium (HFE) \( S = S_0 = \frac{\Lambda}{\mu} \), \( F(X^0, 0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \). Thus,

\[
X' = F(X, 0) = \begin{pmatrix} \Lambda - \mu S \\ -\mu U_3 \end{pmatrix}.
\]
Now, as $t \to \infty$, $X \to X^0$. Hence, $X^0$ is globally asymptotically stable (i.e., condition (H1) is satisfied).

Now consider
\[
G(X, Y) = \begin{pmatrix}
- (p + \sigma + \delta_2 + \mu) & \alpha \\
p & \beta S_0 - (\mu + \delta_1 + \xi + \alpha)
\end{pmatrix}
\begin{pmatrix}
U_2 \\
U_1
\end{pmatrix} - \begin{pmatrix}
0 \\
\beta U_1 (S_0 - S)
\end{pmatrix}
\]
so that,
\[
B = \begin{pmatrix}
- (p + \sigma + \delta_2 + \mu) & \alpha \\
p & \beta S_0 - (\mu + \delta_1 + \xi)
\end{pmatrix}
\text{ and } \hat{G}(X, Y) = \begin{pmatrix}
0 \\
\beta U_1 (S_0 - S)
\end{pmatrix}.
\]

Since, the total population is bounded by $N = S + U_3 + U_2 + U_1 \leq \frac{\Lambda}{\mu}$, it follows that $S \leq N \leq \frac{\Lambda}{\mu}$. Thus, $\hat{G}(X, Y) \geq 0$, which now implies that conditions (H1) and (H2) are satisfied. Consequently by Lemma 6.4.2 the fixed point $D^0$ is globally asymptotically stable when $R_0 < 1$, which indicates non-existence of multiple non-trivial equilibria when $\omega = 0$. The epidemiological implication of HFE being g.a.s is that any heroin epidemic will be eliminated from the community if the threshold quantity $R_0$ is decreased to (or maintained at) a value below unity. □

Now for $\omega > 0$, Theorem 6.4.5 follows.

**Theorem 6.4.5** For $\omega > 0$ model (6.1) has:

(i) A unique positive endemic equilibrium if $B < 0$ and either $C = 0$ or $B^2 - 4AC = 0$;

(ii) A unique positive endemic equilibrium if $C < 0$ (i.e., $R_0 > 1$) and $B < 0$;

(iii) Two positive endemic equilibria if $C > 0$, $B < 0$ and $B^2 - 4AC > 0$;

(iv) No positive endemic equilibrium if $B > 0$ and either $C > 0$ or $B^2 < 4AC$.

The theorem may be proved as follows. It is obvious to note that in the quadratic equation (6.8) $A$ is always positive and $C$ is either positive or negative depending on whether the basic reproduction number is less than or greater than one respectively.

For case (i) where $B < 0$ and $C = 0$ (i.e., $R_0 = 1$) then equation (6.8) becomes linear $AU_1^* + B = 0$ and has a unique nonzero solution $U_1^* = -B/A$ which is positive if $B < 0$.
and negative if $B > 0$. Referring to equation (6.7) it is easy to see that if $U_1^*$ is unique then so are $S^*, U_2^*$ and $U_3^*$.

For case (ii) where $C < 0$ (that is $R_0 > 1$) and $B < 0$ equation (6.8) is quadratic and according to Descarte’s Rule of Signs (see [207]), (6.8) has one change of signs indicating (6.8) has a unique positive root and therefore there is a unique endemic equilibrium.

In case (iii) where $B < 0$ there is a nonnegative endemic equilibrium at $R_0 = 1$. However, because equation (6.8) is quadratic and since the equilibrium is continuously determined by $R_0$ then there must be an interval to the left of $R_0 = 1$ on which two nonnegative equilibria coexist. That is

$$U_{1,1}^* = -B - \sqrt{B^2 - 4AC} \quad 2A, \quad U_{1,2}^* = -B + \sqrt{B^2 - 4AC} \quad 2A. \quad (6.9)$$

For case (iv) where $B > 0$ and $C > 0$ or $B^2 < 4AC$ equation (6.8) has no positive real root as can be seen in equation (6.9), implying non-existence of a positive endemic equilibrium.

Case (iii) suggests that model (6.1) exhibits the phenomenon of backward bifurcation since the classical requirement for the occurrence of the phenomenon of backward bifurcation is satisfied, that is the existence of multiple equilibria when the basic reproduction number is less than one. Thus, the following Theorem 6.4.6:

**Theorem 6.4.6** Model (6.1) has a backward bifurcation at $R_0 = 1$ if and only if $B < 0$ (i.e., $\omega > \omega_c$).

**Proof.** Consider equation (6.8) $f(U_1^*) = AU_1^{*2} + BU_1^* + C = 0$. Note that at $R_0 = 1$, $C = 0$ implies the graph $f(U_1^*)$ passes through the origin. If $B < 0$ it follows that $f(U_1^*) = 0$ has a nonnegative root. Since $f(U_1^*)$ is a continuous function of $C$, if $C$ is increased such that $C > 0$, there is some open interval of $C$ say $(0, \psi)$ on which $f(U_1^*) = 0$ has two nonnegative roots. That is, there exists two nonnegative endemic equilibria when $R_0 < 1$. This is indeed true since case (iv) of Theorem 6.4.5 has already shown that for $B \geq 0$ model (6.1) does not have positive real roots when $R_0 < 1$. Note that at $R_0 = 1$, $C = 0$ the following equality holds

$$(\mu + \delta_1 + \xi)(p + \sigma + \delta_2 + \mu)\mu + \alpha \mu (\mu + \delta_2 + \sigma) = (p + \sigma + \delta_2 + \mu)\beta \Lambda. \quad (6.10)$$
6.4 Basic properties and basic reproduction number

This together with condition $B < 0$ implies

$$
\omega > \frac{(\mu + \delta_1 + \xi)(p + \sigma + \delta_2 + \mu)\beta + \alpha(\sigma + \delta_2 + \mu)\beta}{\mu\alpha(\sigma + \delta_2 + \mu)} \triangleq \omega_c. \square
$$

Thus, the phenomenon of backward bifurcation (referring to case (iii) a situation where there are two endemic equilibria) occurs at the left of $R_0 = 1$ if and only if condition (6.11) is satisfied. This suggests that backward bifurcation will only occur if the parameter $\omega$ that accounts for the extent of saturation of heroin users exceeds a certain threshold (i.e., $\omega > \omega_c$). However, if $\omega < \omega_c$ backward bifurcation cannot occur.

Thus, the parameter $\omega$ plays a critical role in the formation of backward bifurcation for model (6.1). It is instructive to note that similar results as the one shown in inequality (6.11) can be obtained by Center Manifold Theory (see Appendix L), where it is emphasized that if $\omega > \omega_c$ the bifurcation coefficient $a$ is positive indicating that the model system (6.1) undergoes the phenomenon of backward bifurcation. The epidemiological implication of backward bifurcation is that although it is necessary to reduce the basic reproduction number below one it is not sufficient to eradicate a heroin epidemic, rather $R_0$ should be reduced further. In fact, $R_0$ should be reduced below a certain threshold which shall be denoted by $R^C_0$ (see Figure 6.3(b)).

6.4.3 Computation of new threshold for heroin eradication $R^C_0$

Here the critical value of the basic reproduction number where the two non-trivial endemic equilibria (both stable and unstable) collide and annihilate each other leaving only the heroin-free equilibrium point as the stationary solution is computed. This is $R^C_0$ in Figure 6.3(b). $\Lambda$ is chosen as the parameter of backward bifurcation. Note that in case (iii) of Theorem 6.4.5 equation (6.8) has nonnegative roots corresponding to two endemic equilibria if and only if $C > 0$ (i.e., $R_0 < 1$) and $B < 0$, $B^2 > 4AC$. It follows that if $B = -2\sqrt{AC}$ equation (6.8) has one nonnegative root $\frac{B}{2A}$. Supposing there is backward bifurcation at $R_0 = 1$, then there are two endemic equilibria for an interval of the basic reproduction number starting from a threshold $R^C_0$ defined by $B = -2\sqrt{AC}$ to a point where $R_0 = 1$. To obtain the threshold $R^C_0$ which is often referred to as the critical value of the basic reproduction number, one needs to replace the values of $A$, $B$ and $C$ in the equality $B^2 = 4AC$ to obtain a quadratic equation in terms of
6.4 Basic properties and basic reproduction number

Λ. For mathematical tractability redefine coefficients B and C as $B = B_1 - B_2 \Lambda$ and $C = C_1 - C_2 \Lambda$, where

\begin{align*}
B_1 &= (\mu + \delta_1 + \xi)(p + \sigma + \delta_2 + \mu)\beta + \mu\omega(\mu + \delta_1 + \xi)(p + \sigma + \delta_2 + \mu) + \alpha(\sigma + \delta_2 + \mu)\beta, \\
B_2 &= \beta\omega(p + \sigma + \delta_2 + \mu), \\
C_1 &= \mu(\mu + \delta_1 + \xi)(p + \sigma + \delta_2 + \mu) + \alpha\mu(\sigma + \delta_2 + \mu), \\
C_2 &= \beta(p + \sigma + \delta_2 + \mu), \\
B_1, B_2, C_1, C_2 &> 0.
\end{align*}

Note that $A > 0$ remains as previously defined in equation (6.8). Now the quadratic equation in terms of $\Lambda$ can be obtained as

$$B_2^2 \Lambda^2 + (4AC_2 - 2B_1B_2)\Lambda + (B_2^1 - 4AC_1) = 0.$$ 

Since the scenario where $B < 0$ and $R_0 < 1$ is being considered it follows that $C_1B_2 > C_2B_1$ and thus there is just the single solution

$$\Lambda^c = \frac{B_1B_2 - 2AC_2 + 2\sqrt{A^2C_2^2 + AB_2(C_1B_2 - C_2B_1)}}{B_2^2} > 0.$$ 

Thus, the critical value of basic reproduction number (i.e., the new threshold for heroin eradication), $R_0^C$, is given as

$$R_0^C = \frac{\beta\Lambda^c(p + \delta_2 + \mu + \sigma)}{\mu\alpha(\mu + \delta_2 + \sigma) + \mu(\mu + \delta_1 + \xi)(p + \delta_2 + \mu + \sigma)}.$$ 

Consequently, from the above analysis of computation of threshold for heroin eradication the following Lemma is deduced:

**Lemma 6.4.3**  
(a) If $R_0 > 1$, then model (6.1) has a unique endemic equilibrium point $E^*$. In this case a heroin epidemic will persist in the community;

(b) If $R_0^C < R_0 < 1$, then model (6.1) has two endemic equilibria $\bar{E}_1$ and $\bar{E}_2$, and signals that model (6.1) has backward bifurcation;

(c) If $R_0 < R_0^C < 1$, then model (6.1) has only the heroin-free equilibrium point $D^0$ and in this case heroin users will disappear.

Figure 6.3 exhibits typical bifurcation diagrams for model (6.1). To obtain the graphs the recruitment rate $\Lambda$ is varied while other parameter values are held fixed. The
parameters used for the numerical simulation that leads to Figure 6.3(a) include \( \omega = 0.11, \mu = 0.01, \beta = 0.001, \delta_1 = 0.002, \delta_2 = 0.001, \xi = 0.015, \alpha = 0.9, p = 0.467, \sigma = 0.1 \) and \( 1 \leq \Lambda \leq 3 \). Figure 6.3(a) represents the forward bifurcation scenario where if \( R_0 < 1 \) the heroin-free equilibrium is globally asymptotically stable while when \( R_0 > 1 \) the heroin epidemic can persist. However, as noted from Figure 6.3(b) increasing parameter \( \omega \) from \( \omega = 0.11 \) to \( \omega = 0.25 \) such that \( \omega > \omega_c \), a heroin epidemic can persist once established for a range of \( R_0 \) values that are below unity which indicates the occurrence of backward bifurcation. This implies reducing \( R_0 \) below one will not necessarily be sufficient for eradication of heroin usage from the community. If \( R_0 \) is sufficiently decreased such that \( R_0 < R_0^C \) the positive equilibrium no longer exists and heroin usage will cease to thrive and eventually fall from its relatively high endemic level to the heroin-free equilibrium. From Figure 6.3(b) it is observed that when \( R_0^C \leq R_0 \leq 1 \) there is a stable endemic equilibrium, an unstable endemic equilibrium and a stable heroin-free equilibrium. When \( R_0 > 1 \) there is only one stable endemic equilibrium. Figure 6.3(c) shows the effect of the saturation parameter \( \omega \) on \( R_0^C \), namely that increasing \( \omega \) decreases \( R_0^C \). Figure 6.3(d) shows that increasing the treatment rate \( \alpha \) increases \( R_0^C \) which epidemiologically implies that high cure rates of heroin users can lead to shrinking of the backward bifurcation regime.

6.5 Global stability

According to Theorem 6.4.5 model (6.1) may have multiple equilibria when \( R_0 < 1 \) and a unique endemic equilibrium whenever \( R_0 > 1 \). First, global stability of the endemic equilibrium of model (6.1) is investigated for a special case i.e., when \( \omega = \xi = \sigma = 0 \), using the Lyapunov direct method and later proven for the general model (i.e., \( \omega, \sigma, \xi > 0 \)) using a geometric approach.

6.5.1 Global stability using the Lyapunov method (special case \( \omega = \sigma = \xi = 0 \))

Lyapunov functions have previously been used in proving global stability of epidemic models for instance see [208, 209, 210, 211]. Now consider the following Theorem:

**Theorem 6.5.1** If \( \omega = \sigma = \xi = 0 \) the unique endemic equilibrium \( E_1^* \) of model (6.1) is globally asymptotically stable in the interior of \( \Omega \) if \( R_0 > 1 \).
6.5 Global stability

Figure 6.3: (a) and (b) represent bifurcations where drug users $U_1^*$ are not in treatment; equilibria are plotted as a function of $R_0$. The blue solid lines represent the stable equilibria while red solid lines represent unstable equilibria. (a) Represents forward bifurcation with parameters $\omega = 0.11 < \omega_c = 0.1156, \mu = 0.01, \beta = 0.001, \delta_1 = 0.002, \delta_2 = 0.001, \xi = 0.015, \alpha = 0.9, p = 0.467, \sigma = 0.1$ and $\Lambda \in [1, 3]$. (b) Represents backward bifurcation with parameters as in (a) except $\omega = 0.25 > \omega_c = 0.1156$. (c) The critical value $R_0^C$ as a function of the saturation parameter $\omega$. The black dotted line represents the threshold $\omega_c$ which if exceeded gives rise to a backward bifurcation. (d) The critical value $R_0^C$ as a function of treatment rate $\alpha$. 
Proof. Defining the following Lyapunov candidate function,

\[ W(S,U_1,U_2) = \frac{(S - S^*)^2}{2S^*} + \left(U_1 - U_1^* - U_1^* \ln \frac{U_1}{U_1^*}\right) + \frac{pU_2^*}{\alpha U_1^*} \left(U_2 - U_2^* - U_2^* \ln \frac{U_2}{U_2^*}\right), \]

(6.12)

the time derivative of \( W(S,U_1,U_2) \) along the solutions of system (6.1) is

\[ W'(S,U_1,U_2) = \left(\frac{S - S^*}{S^*}\right) \frac{dS}{dt} + \left(\frac{U_1 - U_1^*}{U_1}\right) \frac{dU_1}{dt} + \frac{pU_2^*}{\alpha U_1^*} \left(1 - \frac{U_2^*}{U_2}\right) \frac{dU_2}{dt} \]

\[ = \left(\frac{S - S^*}{S^*}\right) [\Lambda - \beta U_1 S - \mu S] \]

\[ + \left(\frac{U_1 - U_1^*}{U_1}\right) [\beta U_1 S + pU_2 - (\mu + \delta_1 + \alpha) U_1] \]

\[ + \frac{pU_2^*}{\alpha U_1^*} \left(1 - \frac{U_2^*}{U_2}\right) [\alpha U_1 - (p + \delta_2 + \mu) U_2]. \]

(6.13)

Because \((S^*,U_2^*,U_1^*)\) is an endemic steady point of model (6.1) when \(\omega = \xi = \sigma = 0\), then it follows that

\[ \Lambda = \beta U_1^* S^* + \mu S^*, \quad (p + \delta_2 + \mu) = \frac{\alpha U_1^*}{U_2^*}, \quad (\mu + \delta_1 + \alpha) = \beta S^* + \frac{pU_2^*}{U_1^*}. \]

(6.14)

Using (6.14) in equation (6.13) yields

\[ W'(S,U_1,U_2) = - \left(\frac{S - S^*}{S^*}\right) \left(\mu (S - S^*) + \beta (SU_1 - S^* U_1^*)\right) \]

\[ + \left(U_1 - U_1^*\right) \left(\beta (S - S^*) + p \left(\frac{U_2}{U_1} - \frac{U_2^*}{U_1^*}\right)\right) \]

(6.15)

\[ + \frac{pU_2^*}{\alpha U_1^*} \left(1 - \frac{U_2^*}{U_2}\right) \left(\alpha U_1 - \alpha U_2 U_1^* U_2^*\right) \left(\alpha U_1 - \frac{\alpha U_2 U_1^*}{U_2^*}\right) \right]. \]

Note that

\[ SU_1 - S^* U_1^* = S^* (U_1 - U_1^*) + U_1 (S - S^*). \]

Replacing the above equality in equation (6.15) results in

\[ W'(S,U_1,U_2) = - \left(\frac{(S - S^*)^2}{S^*}\right) \left(\mu + \beta U_1\right) + pU_2^* \left(2 - \frac{U_1 U_2}{U_1 U_2^*} - \frac{U_2 U_1}{U_2 U_1^*}\right) \]

\[ = - \left(\frac{(S - S^*)^2}{S^*}\right) \left(\mu + \beta U_1\right) - pU_2^* \left[\sqrt{\frac{U_1 U_2}{U_1 U_2^*}} - \sqrt{\frac{U_2 U_1}{U_2 U_1^*}}\right]^2. \]
6.5 Global stability

Hence, \( W'(S, U_1, U_2) \leq 0 \) for all \( S, U_1, U_2 > 0 \). Hence, the heroin endemic equilibrium \( E_1^* \) is stable and \( W'(S, U_1, U_2) = 0 \) if and only if \( S = S^*, U_2 = U_2^*, U_1 = U_1^* \). The largest compact invariant set when \( \omega = \xi = \sigma = 0 \) in \( \{(S^*, U_1^*, U_2^*) \in \mathbb{R}_+^3\} \) is the singleton \( \{E_1^*\} \). Therefore, by LaSalle’s invariance principle the endemic steady state \( E_1^* \) is globally asymptotically stable in the interior of \( \mathbb{R}_+^3 \). □

The previous global stability analysis was only relevant for a very specific case. In the subsequent subsection the geometric approach by Li and Muldowney [98, 191, 212] is used to obtain sufficient conditions that ensure the unique endemic equilibrium is globally asymptotically stable for a wide range of parameter values.

6.5.2 A geometric approach to global stability

For the general model global stability is investigated using the Li and Muldowney [98, 191, 212] generalizations of the Poincare-Bendixson approach for systems of \( n > 2 \) ordinary differential equations. This criterion is sometimes referred to as a geometric approach to global stability [193, 213].

To apply the geometric approach on model (6.1) consider the autonomous dynamical system

\[
\frac{dy}{dt} = f(y)
\]

(6.16)

where \( f = (f_1, f_2, f_3, f_4)^T \) and \( f_1, f_2, f_3, f_4 \) represent the right-hand side of system (6.1), respectively. First, the general mathematical framework of the procedure developed in Li and Muldowney [98, 189] is outlined.

Suppose the map \( y \mapsto f(y) \) is a \( C^1 \) function for \( y \) in an open subset \( D \subset \mathbb{R}^n \) and consider the following autonomous dynamical system

\[
y' = f(y).
\]

(6.16)

Let \( y(t, y_0) \) be the solution to equation (6.16) satisfying \( y(0, y_0) = y_0 \). Now the following basic assumptions are made:

(H3) \( D \) is simply connected;

(H4) There exists a compact absorbing set \( K \subset D \);

(H5) Equation (6.16) has a unique equilibrium \( y^* \) in \( D \).

Now under the stated assumptions (H3)-(H5), \( y^* \) is said to be globally stable in \( D \) if it is locally stable and all trajectories in \( D \) converge to the same equilibrium \( y^* \). That is system (6.16) has no non-constant periodic solutions. It is important to mention that global stability can be tested by the Bendixson criteria. For \( n \geq 2 \) a Bendixson
criterion refers to a condition satisfied by field $f$ which precludes the existence of non-constant periodic solutions of equation (6.16). When $n = 2$, (i.e., the planar case) the classical results (Poincaré-Bendixson theorem and Dulac criteria, see [103]) adequately provide such global conditions. For $n \geq 3$ a remarkable approach for proving global stability is given by the work of Li and Muldowney [98, 189, 191]. They showed that if conditions (H3)-(H5) hold and differential equation (6.16) fulfills a Bendixson criterion that is robust under $C^1$ local $\epsilon$-perturbations of $f$ at all non-equilibrium non-wandering points for system (6.16), then $y^*$ is globally stable in $D$ provided it is stable.

Now the new Bendixson criterion based on the use of the Lozinskiĭ measure as developed in [98] is stated. Consider the differential equation (6.16) under the stated assumptions (H3)-(H5). Let $y \mapsto P(y)^{(n/2)} \times (n/2)$ be a matrix-valued function which is $C^1$ for $y \in D$ and consider

$$A = P_f P^{-1} + PJ^{[2]} P^{-1} \quad (6.17)$$

where $P_f$ is the directional derivative of $P$ in the direction of the vector field $f$ in system (6.16) and is defined as

$$(p_{i,j}(y))_f = (\partial p_{i,j}(y)/\partial x)^T f(y) = \nabla p_{i,j} \cdot f(y). \quad (6.18)$$

Next, let $J^{[2]}$ represent the second additive compound matrix of $J$ (where $J(y) = Df(y)$). In Muldowney [214] the relation of compound matrices to differential equations is established. It is shown that for an arbitrary $n \times n$ matrix $J = (J_{i,j})$, $J^{[2]}$ is an $\binom{n}{2} \times \binom{n}{2}$ matrix. Now define the following quantity

$$\bar{q}_2 = \limsup_{t \to \infty} \sup_{y_0 \in U} \frac{1}{t} \int_0^t \rho(A(s, y_0)) ds \quad (6.19)$$

1 A function $g \in C^1(D \to \mathbb{R}^n)$ is called a $C^1$ local $\epsilon$-perturbation of $f$ at $y_0 \in D$ if there exists an open neighbourhood $U$ of $y_0$ in $D$ such that the support $\text{supp}(f - g) \subset U$ and $|f - g|_{C^1} < \epsilon$, where $|f - g|_{C^1} = \sup \{|f(y) - g(y)| + |f_y(y) - g_y(y)| : y \in D\}$.

2 A point $y_0 \in D$ is said to be non-wandering for system (6.16) if for any neighbourhood $U$ of $y_0$ in $D$ there exists arbitrary large $t$ such that $U \cap y(t, U) \neq \emptyset$. As an example, any equilibrium, alpha limit point, or omega limit point is non-wandering.
where $\rho(A)$ is the Lozinskiĭ measure of $A$ with respect to vector norm $\| \cdot \|$ in $\mathbb{R}^N$, $N = \binom{n}{2}$ and $\rho(A)$ is defined as

$$\rho(A) = \lim_{h \to 0^+} \frac{|1 + hA| - 1}{h} \quad (6.20)$$

(see [215, 216]). In Li and Muldowney [98] it is proved that if conditions (H3) and (H4) are satisfied then $\bar{q}_2 < 0$ indicates that there are no orbits giving rise to a simple closed rectifiable curve in $D$ that is invariant for system (6.16) (that is periodic orbits, homoclinic orbits and heteroclinic cycles). Furthermore, it has been demonstrated by Li and Muldowney [98] that under the stated assumptions (H3)-(H5), the quantity $\bar{q}_2 < 0$ implies the local stability of equilibrium point $y^*$. As a result the following is true:

**Theorem 6.5.2 (Li and Muldowney [98]).** Assuming that conditions (H3)-(H5) hold the equilibrium point $y^*$ is globally asymptotically stable in $D$ if a function $P(y)$ and a Lozinskiĭ measure $\rho$ exists such that quantity the $\bar{q}_2 < 0$.

Observe that whenever $R_0 > 1$, there exists a unique and positive endemic equilibrium $E^*$ (see Lemma 6.4.3) for model system (6.1). The method outlined above requires that (i) the endemic equilibrium $E^*$ is unique in the interior of $\mathcal{U}$ (i.e., condition H5 holds) and (ii) in the interior of $\mathcal{U}$ there exists an absorbing compact set (condition H4 holds). The heroin model studied here with the assumption that $R_0 > 1$ fulfills conditions H4-H5. It is easy to prove that when $R_0 > 1$, the heroin free equilibrium $D^0$ is unstable (see Theorem 6.4.3). The instability of the heroin free equilibrium $D^0$ combined with $D^0 \in \partial D$ (where $\partial D$ defines the boundary of the region $\mathcal{U}$) signals uniform persistence [217]. That is, there exists a positive constant $c_0 > 0$ such that for every solution $(S(t), U_1(t), U_2(t), U_3(t))$ of system (6.1) with $(S(0), U_1(0), U_2(0), U_3(0))$ in the interior of the biologically feasible region $\mathcal{U}$ satisfies

$$\lim_{t \to \infty} \inf |(S(t), U_1(t), U_2(t), U_3(t))| \geq c_0.$$

Because of the boundedness of the region $\mathcal{U}$, uniform persistence is equivalent to the existence of a compact set in the interior of $\mathcal{U}$ which is absorbing for (6.1) (see [218]). Hence, condition (H4) is satisfied. Also it is shown that whenever $R_0 > 1$ the model system (6.1) has only one equilibrium $E^*$ in the interior of $\mathcal{U}$, so that condition H5 is verified. Now for the heroin model system (6.1) the task involves verifying the Bendixson criterion (6.28). Note that the variable $U_3$ does not affect the first, second
or third equation of system (6.1). Thus, the fourth equation can be dropped from the analysis leading to the following subsystem:

\[
\begin{align*}
\frac{dS}{dt} &= \Lambda - \beta U_1 S - \mu S, \\
\frac{dU_1}{dt} &= \beta U_1 S + p U_2 - (\mu + \delta_1 + \xi) U_1 - \frac{\alpha U_1}{1 + \omega U_1}, \\
\frac{dU_2}{dt} &= \frac{\alpha U_1}{1 + \omega U_1} - (p + \sigma + \delta_2 + \mu) U_2.
\end{align*}
\]  

(6.21)

The Jacobian matrix of subsystem (6.21) is found to be:

\[
J = \begin{pmatrix}
-(\beta U_1 + \mu) & -\beta S & 0 \\
\beta U_1 & \beta S - (\mu + \delta_1 + \xi) - \frac{\alpha}{(1 + \omega U_1)^2} & p \\
0 & \frac{\alpha}{(1 + \omega U_1)^2} & -(p + \sigma + \delta_2 + \mu)
\end{pmatrix}.
\]

In working with Theorem 6.5.2 one needs to make use of additive compound matrices. For an arbitrary \(3 \times 3\) matrix \(B\), the second additive compound matrix \(B^{[2]}\) is defined as:

\[
B = \begin{pmatrix}
b_{11} & b_{12} & b_{13} \\
b_{21} & b_{22} & b_{23} \\
b_{31} & b_{32} & b_{33}
\end{pmatrix}, \quad \text{and} \quad B^{[2]} = \begin{pmatrix}
b_{11} + b_{22} & b_{23} & -b_{13} \\
b_{32} & b_{11} + b_{33} & b_{12} \\
-b_{31} & b_{21} & b_{22} + b_{33}
\end{pmatrix}.
\]

Thus, the second additive compound matrix of Jacobian matrix \(J\) of system (6.21) is given as

\[
J^{[2]} = \begin{pmatrix}
J_{11} & p & 0 \\
\frac{\alpha}{(1 + \omega U_1)^2} & J_{22} & -\beta S \\
0 & \beta U_1 & J_{33}
\end{pmatrix},
\]
where

\[
J_{11} = -\beta U_1 - (2\mu + \delta_1 + \xi) - \frac{\alpha}{(1 + \omega U_1)^2} + \beta S,
\]

\[
J_{22} = -\beta U_2 - (2\mu + p + \sigma + \delta_2),
\]

\[
J_{33} = -\frac{\alpha}{(1 + \omega U_1)^2} - (2\mu + p + \sigma + \delta_1 + \delta_2 + \xi) + \beta S.
\]

For the model system (6.21) a suitable vector norm \(|\cdot|\) in \(\mathbb{R}^3\) and a \(3 \times 3\) matrix-valued function \(P(y)\) is given by:

\[
P(S, U_1, U_2) = \begin{pmatrix}
1 & 0 & 0 \\
0 & U_1 & 0 \\
0 & U_2 & 0
\end{pmatrix}.
\]

Thus

\[
P J[2] P^{-1} = \begin{pmatrix}
0 & 0 & 0 \\
0 & \frac{U_1'}{U_1} - \frac{U_2'}{U_2} & 0 \\
0 & 0 & \frac{U_1'}{U_1} - \frac{U_2'}{U_2}
\end{pmatrix}
\]

and

\[
P J[2] P^{-1} = \begin{pmatrix}
J_{11} & \frac{U_1'}{U_1} & 0 \\
\frac{U_2'}{U_1} & J_{22} & -\beta S \\
0 & \beta U_1 & J_{33}
\end{pmatrix}.
\]

Note that upper prime (') denotes differentiation with respect to time. Thus, \(A = P J P^{-1} + P J[2] P^{-1}\) can be obtained as

\[
A = \begin{pmatrix}
J_{11} & \frac{U_1'}{U_1} & 0 \\
\frac{U_2'}{U_1} & J_{22} + \frac{U_1'}{U_1} - \frac{U_2'}{U_2} & -\beta S \\
0 & \beta U_1 & J_{33} + \frac{U_1'}{U_1} - \frac{U_2'}{U_2}
\end{pmatrix}.
\]
6.5 Global stability

It is helpful to write matrix $A$ in block form as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where

$$A_{11} = -\beta U_1 - (2\mu + \delta_1 + \xi) - \frac{\alpha}{(1 + \omega U_1)^2} + \beta S,$$

$$A_{12} = \begin{bmatrix} \frac{\alpha U_1}{(1 + \omega U_1)^2 U_2} \\ 0 \end{bmatrix},$$

$$A_{21} = \begin{bmatrix} \frac{\alpha U_1}{(1 + \omega U_1)^2 U_2} \\ 0 \end{bmatrix}^T,$$

$$A_{22} = \begin{bmatrix} J_{22} + \frac{U_1'}{U_1} - \frac{U_2'}{U_2} & -\beta S \\ \beta U_1 & J_{33} + \frac{U_1'}{U_1} - \frac{U_2'}{U_2} \end{bmatrix}.$$

Following Li and Muldowney [98], let $(u, v, w)$ represent the vectors in $\mathbb{R}^3 \cong \mathbb{R}^3$. Now for the norm $|\cdot|$ in $\mathbb{R}^3$ select

$$|(u, v, w)| = \max\{|u|, |v| + |w|\},$$

and let $\rho$ represent the Lozinskiǐ measure with respect to this norm. Applying the method of approximating the $\rho(A)$ as given in [216] leads to

$$\rho(A) \leq \sup \{g_1, g_2\},$$

where

$$g_1 = \rho_1(A_{11}) + |A_{12}|,$$

$$g_2 = |A_{21}| + \rho_1(A_{22}).$$

(6.22)

Here $|A_{12}|$ and $|A_{21}|$ are operator norms of $A_{12}$ and $A_{21}$ with respect to the $l_1$ vector norm, where they are both regarded as mapping from $\mathbb{R}^2$ to $\mathbb{R}$. $\rho_1(A_{22})$ represents the Lozinskiǐ measure of the $2 \times 2$ matrix $A_{22}$ with respect to the $l_1$ norm in $\mathbb{R}^2$. To obtain $\rho_1(A_{22})$ sum the absolute value of the off-diagonal elements to the diagonal one in each column of $A_{22}$ and then take the maximum of two sums. Assuming that

$$\frac{1}{2} \left( \delta_1 + \xi + \frac{\alpha}{(1 + \omega U_1)^2} \right) > \beta S,$$

(6.23)
Numerical simulations show that the inequality (6.23) holds at the parameter ranges of interest. However, it remains for future work to prove this is true in general. It follows that,

$$\rho_1(A_{11}) = -\beta U_1 - (2\mu + \delta_1 + \xi) - \frac{\alpha}{(1 + \omega U_1)^2} + \beta S;$$

$$|A_{12}| = \max \left\{ \frac{pU_2}{U_1}, 0 \right\} = \frac{pU_2}{U_1};$$

$$|A_{21}| = \max \left\{ \frac{\alpha U_1}{(1 + \omega U_1)^2 U_2}, 0 \right\} = \frac{\alpha U_1}{(1 + \omega U_1)^2 U_2};$$

$$\rho_1(A_{22}) = \max \left\{ \frac{U_1'}{U_1'} - \frac{U_2'}{U_2'} - (2\mu + p + \sigma + \delta_2), \frac{U_1'}{U_1'} - \frac{U_2'}{U_2'} - (2\mu + p + \sigma + \delta_2) - \delta_1 - \xi - \alpha \omega \frac{U_1}{(1 + \omega U_1)^2} + 2\beta S \right\} = \frac{U_1'}{U_1} - \frac{U_2'}{U_2} - (2\mu + p + \sigma + \delta_2).$$

Thus $g_1$ and $g_2$ are respectively

$$g_1 = \rho_1(A_{11}) + |A_{12}| = \beta S + \frac{pU_2}{U_1} - \beta U_1 - \frac{\alpha}{(1 + \omega U_1)^2} - (2\mu + \delta_1 + \xi), \quad (6.24)$$

$$g_2 = |A_{21}| + \rho_1(A_{22}) = \frac{\alpha U_1}{(1 + \omega U_1)^2 U_2} + \frac{U_1'}{U_1} - \frac{U_2'}{U_2} - (2\mu + p + \sigma + \delta_2). \quad (6.25)$$

Now from the second and third equation of (6.21) it is easy to obtain

$$\frac{U_1'}{U_1} = \beta S + \frac{pU_2}{U_1} - \frac{\alpha}{(1 + \omega U_1)} - (\mu + \delta_1 + \xi), \quad (6.26)$$

$$\frac{U_2'}{U_2} = \frac{\alpha U_1}{(1 + \omega U_1)U_2} - (p + \sigma + \delta_2 + \mu). \quad (6.27)$$

Substituting (6.26) into (6.24) and (6.27) into (6.25) leads to

$$g_1 = \frac{U_1'}{U_1} - \mu - \beta U_1 + \frac{\alpha \omega U_1}{(1 + \omega U_1)^2} \leq \frac{U_1'}{U_1} - \mu + \frac{\alpha \omega U_1}{(1 + \omega U_1)^2},$$

$$g_2 = \frac{U_1'}{U_1} - \mu - \frac{\alpha \omega U_2}{(1 + \omega U_1)^2 U_2} \leq \frac{U_1'}{U_1} - \mu.$$

Now based on the definition of the method of approximating the Lozinskié measure $\rho(A)$ as given in Martin [216], the supremum of both $g_1$ and $g_2$, can be approximated.
Hence,

\[ \rho(A) \leq \sup(g_1, g_2) \]

\[ \leq \sup \left\{ \frac{U'_1}{U_1} - \mu + \frac{\alpha \omega U_1}{(1 + \omega U_1)^2}, \frac{U'_1}{U_1} - \mu \right\} \]

\[ = \left( \frac{U'_1}{U_1} - \mu \right) + \sup \left\{ \frac{\alpha \omega U_1}{(1 + \omega U_1)^2}, 0 \right\} \]

\[ = \frac{U'_1}{U_1} - \mu + \frac{\alpha \omega U_1}{(1 + \omega U_1)^2}. \]

Thus the inequality

\[ \rho(A) \leq \frac{U'_1}{U_1} - \mu + \frac{\alpha \omega U_1}{(1 + \omega U_1)^2}. \]

Now the next step involves substituting \( \rho(A) \) into

\[ \bar{q}_2 = \lim_{t \to \infty} \sup_{y_0 \in \Omega} \frac{1}{t} \int_0^t \rho(A(s, y_0)) ds, \quad (6.28) \]

and deducing whether \( \bar{q}_2 < 0 \). And if the inequality \( \bar{q}_2 < 0 \) does not hold then a condition that leads to \( \bar{q}_2 < 0 \) being fulfilled is established.

Considering uniform persistence there exists a \( c_0 > 0 \) and \( T > 0 \) such that \( t > T \) implies

\[ S(t) \geq c_0, \quad U_1(t) \geq c_0, \quad U_2(t) \geq c_0, \quad \text{and} \quad U_3(t) \geq c_0. \]

Now by letting \( \Gamma_1 = \frac{\alpha \omega c_0}{(1 + \omega c_0)^2} \) and \( \Gamma_2 = \mu \) the following claim is made:

If

\[ \Gamma_1 < \Gamma_2, \quad (6.29) \]
then
\[ \rho(A) \leq \frac{U_1'}{U_1} - \tilde{V}, \] (6.30)

where
\[ \tilde{V} = \mu - \frac{\alpha \omega c_0}{(1 + \omega c_0)^2} > 0. \]

Now, for \( t > T \) it can be deduced that
\[ \bar{q}_2 = \limsup_{t \to \infty} \frac{1}{t} \int_0^t \rho(A) ds \]
\[ = \limsup_{t \to \infty} \frac{1}{t} \int_0^T \rho(A) ds + \limsup_{t \to \infty} \frac{1}{t} \int_T^t \rho(A) ds \]
\[ \leq \limsup_{t \to \infty} \frac{1}{t} \log \frac{U_1(t)}{U_1(T)} + \limsup_{t \to \infty} \frac{1}{t} \int_0^T \rho(A) ds - \limsup_{t \to \infty} \tilde{V} \frac{t - T}{t} \] (6.31)
\[ < 0 \text{ when } \tilde{V} > 0, \]

and thus, the Bendixson criterion given by equation (6.28) is verified. Note that with the stated condition \( \Gamma_1 < \Gamma_2, \tilde{V} > 0 \) the limits of the first and second terms of the equation (6.31) approach zero while the limit of the final term approaches \(-\tilde{V}\). Thus, \( \bar{q}_2 < 0 \). Since, \( \bar{q}_2 < 0 \) if and only if condition (6.29) holds true, the following Theorem 6.5.3 is established:

**Theorem 6.5.3** Provided \( R_0 > 1 \), and if \( \Gamma_1 < \Gamma_2 \) then the system (6.1) has a unique endemic equilibrium \( E^* \) which is globally asymptotically stable with respect to solutions of (6.1) originating in the interior of \( \mathcal{O} \). \( \Box \)

The validity of Theorem 6.5.3 will be shortly verified numerically.

### 6.6 Numerical examples

In this section numerical simulations of the heroin epidemic model are presented to support theoretical findings. Figure 6.4 which shows backward bifurcation is obtained
by plotting the equilibrium number of heroin users as a function of $R_0$. The figures present a scenario where $R_0$ is varied via parameter $\beta$ (i.e., $0.0005 \leq \beta \leq 0.0015$) and other parameters are fixed. Figures 6.4(a)-6.4(d) show that increasing $\omega$ leads to the expansion of the region of bistability while decreasing $\omega$ results into contraction of the bistability region. The heroin eradication threshold (also referred to as critical reproduction number) $R_0^C$ shifts from right to left when $\omega$ increases and vice versa when $\omega$ decreases. High values of $\omega$ imply there is not enough treatment available for a large population of heroin users, favouring a situation where there will always be heroin users within the community even though $R_0 < 1$.

![Figure 6.4: Illustration of the effect of increasing parameter $\omega$ that accounts for the saturation of heroin users. Here the parameters remain as in the caption of Figure 6.3 except $\Lambda = 2$ while $\omega$ is shown. The heroin eradication thresholds (i.e., $R_0^C$) corresponding to figures (a)-(d) are respectively 0.5204, 0.6038, 0.7314, 0.9131.](image)

Figures 6.5(a) and 6.5(b) exhibit the time course of the heroin endemic in a parameter regime where there is a backward bifurcation. In both figures 6.5(a) and 6.5(b) $R_0 = 0.7506 < 1$. The figures show the dependence of heroin usage on the size of the initial conditions supplied to the system, which is a common characteristic of models that have a bi-stability region. If the model is supplied with initial conditions that are below the unstable curve (see the red solid line on Figure 6.3(b)) the solution trajectories are attracted to the heroin free equilibrium while if initial conditions are chosen such that they are above the unstable curve, then the solution trajectories are attracted to a stable non-trivial equilibrium. Thus, in the case where there is backward bifurcation, the initial number of people engaging in heroin use govern the course of the heroin epidemic.

Figures 6.5(c) shows the time course of the number of heroin users when $\omega$ (that accounts for the extent of saturation of heroin users) is varied, while the initial states...
and all other parameter values are fixed to constant values. It can be seen that not all values of $\omega$ will trigger rapid growth towards an endemic equilibrium when $R_0 < 1$. Indeed, parameter $\omega$ has to exceed a certain fixed threshold $\omega_c$, hence supporting theoretical findings, that a non-zero equilibria when $R_0 < 1$ can only be maintained when $\omega$ is greater than $\omega_c$ (see equation (6.11)). Figure 6.5(d) shows the effect of treatment $\alpha$ on heroin users. High treatment leads to a steady decline of heroin users.

Figure 6.6 presents a scenario where $R_0 > 1$. In this scenario it is expected that when a heroin user enters a heroin-free community there will be rapid growth of heroin users until a globally stable equilibrium point is reached. Recalling that parameter $\omega$ does not appear in $R_0$ it nevertheless does affect the model dynamics. The impact of $\omega$ when $R_0 > 1$ is different from the case where $R_0 < 1$. For $R_0 < 1$ it plays a key role in inducing bi-stability while for $R_0 > 1$ the parameter $\omega$ impacts the heroin dynamics by determining the time taken for an epidemic to occur. For relatively high $\omega$ values there is a sudden decrease in the susceptible subpopulation while for relatively low values of $\omega$ there is a gradual decrease in the susceptible subpopulation. Moreover, Figure 6.6(b) depicts that for any given value of $\omega$ the heroin users gradually approach a stable endemic equilibrium point. The only striking difference is the time taken to reach the heroin endemic equilibrium. At high values of $\omega$ the heroin endemic will rapidly approach an equilibrium.

Now the global stability condition obtained using the geometric approach is verified using the following parameter values:

(i) $\beta = 0.001, \delta_1 = 0.002, \delta_2 = 0.001, \xi = 0.015, \mu = 0.01, \alpha = 0.9, p = 0.467, \sigma = 0.1, \omega = 2, \Lambda = 3, c_0 = 50$. With these parameters the corresponding $R_0 = 1.5012 > 1$ and $\Gamma_1 = 0.0088 < \Gamma_2 = 0.0100$. In this scenario condition (6.29) is satisfied and the model system (6.1) should be globally asymptotically stable. Figures 6.7(a), 6.7(b), 6.7(c) and 6.7(d) show existence of an apparently stable equilibrium;

(ii) Using the same set of parameter values as case (i) except $\omega = 0.05$, leads to $\Gamma_1 = 0.1837 > \Gamma_2 = 0.0100$. In this case the asymptotic stability condition is not satisfied and unsurprisingly model system (6.1) has periodic solutions as shown in figures 6.7(e), 6.7(f), 6.7(g) and 6.7(h). The epidemiological interpretation of this is that the heroin epidemic will fluctuate between low and high endemic levels. The cycles are induced by time delays in the transmission processes.
6.6 Numerical examples

Figure 6.5: (a) and (b) illustrate the dependence of the heroin epidemic extinction or persistence on the initial states provided to the model. Parameters are $\omega = 0.15$, $\beta = 0.001$, $\delta_1 = 0.002$, $\delta_2 = 0.001$, $\xi = 0.015$, $\mu = 0.01$, $\alpha = 0.9$, $p = 0.467$, $\sigma = 0.1$, $\Lambda = 1.5$. With these parameters the corresponding value of $R_0$ is 0.7506. (c) The impact of increasing parameter $\omega$ (the extent of saturation of treatment) on heroin users when $R_0 < 1$. Parameters used are $\beta = 0.001$, $\delta_1 = 0.002$, $\delta_2 = 0.001$, $\xi = 0.015$, $\mu = 0.01$, $\alpha = 0.9$, $p = 0.46$, $\sigma = 0.1$, $\Lambda = 1.5$ which correspond to $R_0 = 0.7427 < 1$. (d) The effect of treatment $\alpha$ on heroin users $U_1$ when all other parameters and initial conditions are fixed.

6.6.1 Uncertainty and sensitivity analysis

Sensitivity analysis is conducted so as to identify critical inputs of the proposed heroin epidemic model as well as gain insights on how input uncertainty influences the model.
Figure 6.6: (a) Shows time series of susceptibles $S$ as a function of parameter $\omega$ that accounts for the extent of saturation of treatment when $R_0 > 1$. (b) Shows time series of drug users $U_1$ as a function of parameter $\omega$ when $R_0 > 1$. In both Figures, parameters used are the same as in Figure 6.5(a) except $p = 0.46, \Lambda = 3$ corresponding to $R_0 = 1.4855 > 1$.

outcome [219]. To achieve this the Latin hypercube sampling (LHS) technique which provides a comprehensive method of assessing model sensitivity to parameters over a multi-dimensional parameter space is used. One of the advantages of using LHS is that it requires fewer samples of parameters than simple random sampling to achieve the same accuracy (see [219] and references therein for in-depth discussion of LHS).

In the proposed heroin epidemic model the LHS technique is important due to the relatively large uncertainty in the model parameter estimates that have been used. The technique works in combination with the partial rank correlation coefficient (PRCC) which estimates the sign and strength of the relationship that exists between each model parameter and any specified output [220, 221]. The PRCC values are bounded between 1 and -1, with a PRCC value close to 1(-1) indicating very strong positive (negative) correlation. The relative importance of the model parameters can be directly evaluated by comparing the values of the PRCC [221]. The uncertainty and sensitivity analysis using the LHS technique involves first selecting a baseline value and a range for each parameter of the heroin epidemic model (6.1) (see Table 6.2), and then performing multiple runs for a given outcome variable or response function. To enhance accuracy, 1500 random samples of parameter values were used for the sensitivity analysis and
6.6 Numerical examples

Figure 6.7: Validity of the global stability condition. Figures (a), (b), (c) and (d) represent a scenario where condition (6.29) holds true and global stability is predicted. Figures (e), (f), (g) and (h) represent a scenario where condition (6.29) does not hold and oscillations are expected.

significant levels set for $p_{value} < 0.05$.

Figure 6.8 displays the sensitivity analysis results for the number of heroin users not in treatment $U_1(t)$. It is straightforward to see that recruitment rate $\Lambda$, effective contact rate $\beta$, relapsing rate of heroin users in treatment to heroin users not in treatment $p$, and saturation parameter $\omega$ are positively correlated while natural death $\mu$, treatment rate $\alpha$, heroin-induced death rates ($\delta_1, \delta_2$), "self cure" rate $\xi$ and successful recovery rate of heroin users in treatment $\sigma$ are negatively correlated. Amongst the positively correlated PRCC values the parameters $\beta, p$ and $\omega$ are strongly positively correlated with the number of heroin users not in treatment as evidenced by their high PRCC values. However, at time point year 15 the effective contact rate $\beta$ has a slightly higher PRCC value than relapsing rate $p$ suggesting that during the initial stage of a heroin epidemic effective contact between heroin users $U_1(t)$ and susceptibles significantly contributes to the emergence of a heroin epidemic. On the other hand, at time point year 30 the situation observed at time point year 15 is reversed. That is relapsing of heroin users in treatment has a slightly higher PRCC value than the effective
contact rate $\beta$. Hence, in the long-term, relapsing of heroin users in treatment back to heroin use also plays a role in ensuring that there will always be heroin users within the community. Thus, in attempting to control heroin usage within the community, measures that ensure heroin users undergoing treatment do not relapse should be of great importance. The extent of saturation of heroin users as a result of failure to treat heroin users promptly, which is accounted by parameter $\omega$, also contributes to sustaining a heroin epidemic. As suggested by the strongly negatively correlated PRCC value of parameter $\sigma$, ensuring heroin users in treatment are successfully treated (that is they do not relapse) can substantially reduce the sub-population of heroin users. In general the sensitivity analysis results suggest that to combat a heroin epidemic, policy makers and clinicians should target the effective contact rate $\beta$, relapsing rate $p$, and the extent of saturation of treatment for heroin users.

![Figure 6.8](image.png)

**Figure 6.8:** The PRCC output for heroin users $U_1(t)$ not in treatment.
Table 6.2: Parameter baseline values and ranges used in sensitivity analysis.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>baseline value</th>
<th>range</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Lambda$</td>
<td>2</td>
<td>1-5</td>
<td>assumed</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.001</td>
<td>0.0005-0.015</td>
<td>[185]</td>
</tr>
<tr>
<td>$\mu$</td>
<td>1/80</td>
<td>0.01125-0.01375</td>
<td>[222]</td>
</tr>
<tr>
<td>$p$</td>
<td>0.467</td>
<td>0.1-0.8</td>
<td>[185]</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.5</td>
<td>0.2-0.95</td>
<td></td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.002</td>
<td>0.0008-0.0025</td>
<td>assumed</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>0.001</td>
<td>0.00095-0.002</td>
<td>assumed</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.5</td>
<td>0.05-0.5</td>
<td>assumed</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.1</td>
<td>0.1-0.7</td>
<td>assumed</td>
</tr>
<tr>
<td>$\omega$</td>
<td>0.01</td>
<td>0.008-0.25</td>
<td>[88]</td>
</tr>
</tbody>
</table>

6.7 Summary of the chapter

In this chapter a heroin epidemic model with a density-dependent incidence and saturated treatment function was formulated. The threshold parameter $R_0$ usually referred to as the basic reproduction number plays a key role in the prediction of disease persistence or extinction. Epidemiologically, when $R_0$ exceeds one an epidemic persists and if it is below unity the disease will die out. This classical viewpoint has recently been challenged by many researchers since it is not always true a disease will disappear if $R_0$ is decreased below one. In the present heroin epidemic model the analytical results indicate that $R_0 = 1$ is indeed the threshold when the parameter $\omega = 0$. However, when a saturated treatment function (i.e., $\omega > 0$) rather than a linear treatment rate is used, the heroin model exhibits the phenomenon of backward bifurcation where a heroin-free equilibrium and two non-trivial equilibria co-exist even though the basic reproduction number is below unity (see Theorem 6.4.5 case (iii)). The appearance of backward bifurcation indicates that it is not sufficient to decrease the basic reproduction number below unity for the eradication of heroin users within the community. Thus, to effectively control the spread of heroin use one has to reduce $R_0$ below another threshold referred to as the critical value of the basic reproduction number $R_0^C$. That is, heroin users can be eradicated if $R_0 < R_0^C < 1$.

It is important to note that although the parameter $\omega$ might be present in the model, not every value of $\omega$ will lead to bi-stability. Instead $\omega$ has to be greater than a certain threshold $\omega_c$ which is an aggregate of model parameters (see equation (6.11)). In general both analytical (see Appendix L for a Center Manifold approach) and numerical
results suggest that the saturation parameter $\omega$ is responsible for backward bifurcation. Failure to intervene before heroin users have accumulated in the community will lead to a situation where a heroin epidemic can continue to exist even though the basic reproduction number is below one. Improvement of existing medical technology as well as channelling sufficient resources could significantly facilitate early intervention by ensuring that heroin users receive treatment promptly.

In addition, global stability properties using both the Lyapunov direct method and geometric approach by Li and Muldowney have been investigated. It is important to note that even for a four dimensional model, the use of the two nonlinear stability techniques becomes nontrivial. In fact when all the parameters of the model are accounted for, it is difficult if not impossible, to design a Lyapunov function. Using the geometric approach a global condition that accounts for all parameters is established. If the condition is satisfied, it signals heroin persistence within the community is globally stable. However, if the global condition is not satisfied heroin users can oscillate periodically in number (see Figures 6.7(e), 6.7(f), 6.7(g) and 6.7(h)). Moreover, sensitivity and uncertainty analysis using LHS results indicate that the effective contact rate between susceptibles and heroin users $\beta$, the relapsing rate of heroin users in treatment $p$ and the extent of saturation of heroin users $\omega$ are the parameters which contribute to persistence of heroin use within the community.
Chapter 7

Conclusion and future work

This thesis mainly focussed on building epidemiological mathematical models to enhance understanding of the transmission dynamics of tuberculosis in a human population. Incorporating biological and epidemiological aspects pertinent for TB transmission, such as recurrent TB, led to a number of new insights. The contributions emanating from the findings in this thesis are respectively given in chapter 3, chapter 4 and chapter 5.

The first question as given in detail in chapter 3 involved correcting an error that is repeatedly found in epidemic models, in particular models that exhibit the phenomenon of backward bifurcation. Understanding the phenomenon of backward bifurcation is imperative because it predicts that disease will continue to spread even when the basic reproduction number is below the epidemic threshold $R_0 = 1$. Thus, obtaining the backward bifurcation structure correctly is important to avoid overestimating or underestimating disease prevalence. Moreover, this chapter clarifies the correct method of computing the critical value of $R_0$, which is the threshold below which the reproduction number needs to be reduced in order to eliminate disease. Comparison of the correct non-aggregated parameter approach with the incorrect aggregated parameter approach demonstrates that the two methods are substantially different. This is a counter-intuitive result given numerous mathematical models have used the aggregated parameter approach to compute both bifurcation structures as well as critical values for $R_0$.

The second question (studied in chapter 4) extended the simple TB model developed in chapter 3 by incorporating realistic TB pathways accounting for recurrent TB. In this
chapter a new model is formulated with the aim of understanding how recurrent TB can alter TB dynamics. Applying the Center Manifold Theory it is shown that the model exhibits the phenomenon of backward bifurcation where unstable and stable equilibria coexist when the associated reproduction number is less than one. The backward bifurcation threshold is obtained using analytical analysis as well as confirmed using the Center Manifold approach. The backward bifurcation threshold so obtained is used to stress the importance of not ignoring backward bifurcation phenomena, in particular in the context of TB epidemiology. This is as a result of the past misconception that backward bifurcation cannot occur in the real world without compromising biological realism, especially in tuberculosis as a case study [111]. In the past fifteen years this was understood to be the case as suggested by Lipsitch and coworkers [111]. In this thesis the misconception is revisited and it is shown that with the incorporation of recurrent TB pathways, now known to be common especially in high TB burden settings, the phenomenon of backward bifurcation can occur even without compromising biological realism.

The chapter also reveals a rare hysteresis effect even where $R_0$ is below unity. This bi-stability phenomena has not been mentioned in any of the pertinent TB literature. The findings in chapter 4 are imperative in TB epidemiology in that the results emphasize how recurrent TB can alter the well-known TB dynamics. Hence, not taking into account the contribution of recurrent TB to TB burden can lead to clinicians and policy makers formulating decisions for combating TB that will be ineffective.

Chapter 5 was inspired by the fact that although previous studies have considered population-level heterogeneity in susceptibility to reinfection between previously treated and latently infected persons [91, 97, 159], it has only been in a limited way. No previous work has considered differential susceptibility across all possible exposure and treatment histories (i.e., LTBI, treated LTBI and treated TB disease), along with the population level impact of all relevant public health interventions. Thus, the chapter attempted to shed some light on two key questions: (a) how variability in risk of reinfection could alter TB dynamics in a model accounting for heterogeneity in host susceptibility and (b) how this variability in risk of reinfection influences the effectiveness of public health interventions. The results deduced in the chapter can help with our understanding and projection of TB epidemiology. The findings in chapter 5 show that distinguishing individuals treated with preventive therapy from
those who were previously treated from active TB can yield new epidemiological insights.

Finally, chapter 6 investigated a heroin epidemic model with a saturated treatment function. The analysis of the model shows that when a saturated treatment function rather than a linear treatment rate is used, the heroin model exhibits the phenomenon of backward bifurcation where a heroin-free equilibrium and two non-trivial equilibria co-exist even though the basic reproduction number is below unity. In such a situation, to eliminate heroin addiction from the community, the basic reproduction number has to be reduced further below unity such that it is below another new threshold referred to as the critical value of the basic reproduction number. Further analysis of the global stability using both the Lyapunov direct method and the geometric approach was conducted. It is noted that even for a four dimensional model, the use of the two nonlinear stability techniques becomes nontrivial. Applying the geometric approach a global condition for stability that accounts for all parameters was established. The condition so obtained was tested numerically and it was observed that heroin persistence within the community is globally stable if the condition is satisfied. However, if the global condition is not satisfied heroin users can oscillate periodically in number.

7.0.1 Future work

The study conducted in this thesis has managed to shed some light on the spreading mechanisms of TB. In particular, on the role of recurrent TB, as well as the role of heterogeneity in host susceptibility to TB, and on the effectiveness of public health interventions. This thesis can be extended by considering the following aspects:

(i) Establishing the local and global dynamics of the endemic equilibria of the model studied in chapter 4;

(ii) Rigorous stability analysis of the model studied in chapter 5, in particular local and global stability;

(iii) Investigating optimal control of preventive therapy among early and late LTBI that could enable the public health in designing and implementing effective measures in controlling the spread of TB in a population;

(iv) Testing the results of Lipsitch and Murray [111] and results obtained in chapter 4 using real data could be important in projecting TB dynamics;
(v) In chapter 5, it is important to note that at the present time the lack of sensitivity of model dynamics to $\sigma_3$ is unclear and is a subject worth exploring in future research;

(vi) Attempting to estimate parameters of the model studied in chapter 5. Unfortunately current datasets are limited for this purpose.
Appendices

Appendix A: Computation of $R_0$ for the TB model with recurrent TB

This appendix relates to chapter 4.

Using the method described in chapter 2, the non-negative matrix $F$ and non-singular M-matrix $V$ are respectively given as

$$F = \begin{pmatrix} 0 & (1-q)c\beta \\ 0 & q\beta \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} \mu + k & 0 \\ -k & \mu + r + \mu_d \end{pmatrix}.$$ 

Hence,

$$FV^{-1} = \begin{pmatrix} (1-q)c\beta k & (1-q)c\beta \\ \frac{(1-q)c\beta}{(\mu+k)(\mu+r+\mu_d)} & \frac{(1-q)c\beta}{q\beta} \end{pmatrix} \frac{(1-q)c\beta}{q\beta}$$

and the basic reproduction number is now given as

$$R_0 = \rho(FV^{-1}) = \frac{c\beta(k + q\mu)}{(\mu + k)(\mu + r + \mu_d)}, \quad (A1)$$

where $\rho$ is the spectral radius.
Appendix B: Theorem deduced when $\theta = 0$

This appendix relates to chapter 4.

**Theorem B.2.1** The model (4.1) when $\sigma = \theta = 0$ has:

(i) A unique positive endemic equilibria if $c_1 < 0$ and $R_0 = 1$, or if the discriminant
$\Delta = c_1^2 - 4c_2c_0 = 0$;

(ii) A unique positive endemic equilibria if $R_0 > 1$;

(iii) Two positive endemic equilibria if the three conditions hold: $c_1 < 0$, $R_0 < 1$ and
$\Delta = c_1^2 - 4c_2c_0 > 0$, and thus backward bifurcation;

(iv) No positive endemic equilibria if $c_1 > 0$ and $R_0 \leq 1$.

**Proof** It is easy to note that in polynomial (4.12) $c_2$ is always positive. Also $c_0 > 0$ if $R_0 < 1$. For case (i) where $c_1 < 0$ and $R_0 = 1$ (i.e., $c_0 = 0$) the quadratic equation (4.12) reduces to $P_2(\lambda) = c_2\lambda + c_1 = 0$ and in this case the model equation (4.1) will have a unique positive endemic equilibrium since $c_1 < 0$. (Note there would be no positive equilibrium if $c_1 \geq 0$.) Moreover, it is instructive to note that in case (i) if the discriminant $\Delta = 0$, then the quadratic equation (4.12) has the repeated root $\lambda = -\frac{c_1}{2c_2}$.

In such a case model equation (4.1) has a unique positive endemic equilibrium if $c_1 < 0$, no positive endemic if $c_1 > 0$ and DFE if $c_1 = 0$ (i.e., $\lambda = 0$ which corresponds to DFE). Similarly, the remaining cases in the proof follow directly from the trivial properties of the roots of quadratic polynomials.
Appendix C: Detailed derivation of backward bifurcation threshold $p_c$

This appendix relates to chapter 4.

**Proof** The proof employs Center Manifold approach as exhibited in Center Manifold Theorem from Castillo-Chavez and Song [58]. For simplification and understanding of the Center Manifold Theorem, it is convenient to transform the model variables of system (4.1) as follows: $x_1 = S, x_2 = E, x_3 = I, x_4 = R$ and $N = \sum_{j=1}^{4} x_j$. Now letting $X = (x_1, x_2, x_3, x_4)^T$ (T denote transpose) the model system (4.1) can be written as $\frac{dX}{dt} = F(X)$ where $F = (f_1, f_2, f_3, f_4)^T$. Hence,

\[
\begin{align*}
\frac{dx_1}{dt} &= \Lambda - \frac{\beta cx_1x_3}{x_1 + x_2 + x_3 + x_4} - \mu x_1 = f_1, \\
\frac{dx_2}{dt} &= \frac{(1-q)\beta cx_1x_3}{x_1 + x_2 + x_3 + x_4} + \frac{(1-\sigma)\theta \beta cx_3x_4}{x_1 + x_2 + x_3 + x_4} - \frac{p\beta cx_2x_3}{x_1 + x_2 + x_3 + x_4} - (\mu + k)x_2 = f_2, \\
\frac{dx_3}{dt} &= \frac{q\beta cx_1x_3}{x_1 + x_2 + x_3 + x_4} + \frac{\sigma\theta \beta cx_3x_4}{x_1 + x_2 + x_3 + x_4} + \frac{p\beta cx_2x_3}{x_1 + x_2 + x_3 + x_4} \\
&\quad + kx_2 - (\mu + r + \mu_d)x_3 = f_3, \\
\frac{dx_4}{dt} &= rx_3 - \frac{\theta \beta cx_3x_4}{x_1 + x_2 + x_3 + x_4} - \mu x_4 = f_4.
\end{align*}
\]

Now choosing $\beta_c = \tilde{\beta}$ as the bifurcation parameter and considering that at $R_0 = 1$, $\tilde{\beta} = \beta^* = \frac{(\mu + k)(\mu + r + \mu_d)}{(k + \mu q)}$, the Jacobian matrix of the system (C1) evaluated at the disease free equilibrium is

\[
H = \begin{pmatrix}
-\mu & 0 & \beta^* & 0 \\
0 & -(\mu + k) & (1-q)\beta^* & 0 \\
0 & k & q\beta^* - (\mu + r + \mu_d) & 0 \\
0 & 0 & r & -\mu
\end{pmatrix}.
\]

With $\tilde{\beta} = \beta^*$ the transformed system (C1) has a simple eigenvalue with zero real part and all other eigenvalues are negative (i.e., has a hyperbolic equilibrium point). Thus, the Center Manifold Theory [58, 223] can be applied to investigate the dynamics of the transformed system (C1) near $\tilde{\beta} = \beta^*$. It is possible to obtain the right eigenvectors of...
\(H(P_0)\) which are denoted by \(w = (w_1, w_2, w_3, w_4)^T\), where

\[ w_1 = \frac{-\beta^*w_3}{\mu}, w_2 = \frac{(1-q)\beta^*w_3}{(\mu + k)}, w_4 = \frac{rw_3}{\mu}, w_3 = w_3 > 0. \]

Similarly, the left eigenvectors of \(H(P_0)\) are obtained and are denoted by \(v = (v_1, v_2, v_3, v_4)\), where

\[ v_1 = 0, \quad v_2 = \frac{k v_3}{\mu + k}, \quad v_3 = v_3 > 0, \quad v_4 = 0. \]

Now the associated bifurcation coefficients, \(a\) and \(b\) as described in Theorem 4.1 of [58] can be obtained. For the purpose of clarity Theorem 4.1 of [58] stated in the mathematical preliminaries is used (see chapter 2).

**Computation of \(a\).**

The transformed model system (C1) has the following non-vanishing partial derivatives of \(H\) evaluated at disease free equilibrium,

\[
\frac{\partial^2 f_2}{\partial x_2 \partial x_3} = -\frac{2(1-q)\beta^*\mu}{\Lambda}, \quad \frac{\partial^2 f_2}{\partial x_3 \partial x_4} = -\frac{2(1-q)\beta^*\mu}{\Lambda} + \frac{2(1-\sigma)\theta \beta^*\mu}{\Lambda},
\]

\[
\frac{\partial^2 f_3}{\partial x_2 \partial x_3} = -\frac{2q\beta^*\mu}{\Lambda} + \frac{2p\beta^*\mu}{\Lambda}, \quad \frac{\partial^2 f_3}{\partial x_3 \partial x_4} = \frac{-2q\beta^*\mu}{\Lambda} + \frac{2\sigma \beta^*\mu}{\Lambda},
\]

\[
\frac{\partial^2 f_2}{\partial x_3 \partial x_3} = -\frac{2(1-q)\beta^*\mu}{\Lambda}, \quad \frac{\partial^2 f_3}{\partial x_3 \partial x_3} = \frac{-2q\beta^*\mu}{\Lambda}.
\]

Hence,

\[
a = \sum_{k,i,j=1}^{4} v_k w_i w_j \frac{\partial^2 f_k(0,0)}{\partial x_i \partial x_j} = v_2 w_2 w_3 \frac{\partial^2 f_2(0,0)}{\partial x_2 \partial x_3} + v_2 w_3 w_4 \frac{\partial^2 f_2(0,0)}{\partial x_3 \partial x_4} + v_3 w_3 w_3 \frac{\partial^2 f_3(0,0)}{\partial x_2 \partial x_3} + v_3 w_3 w_4 \frac{\partial^2 f_3(0,0)}{\partial x_3 \partial x_4}
\]

\[+ v_2 w_3 w_3 \frac{\partial^2 f_2(0,0)}{\partial x_3 \partial x_3} + v_3 w_3 w_3 \frac{\partial^2 f_3(0,0)}{\partial x_3 \partial x_3}
\]

\[
= \frac{2\mu^2(1-q)\beta^2 v_3 w_3^2}{\Lambda(\mu + k)} \left( p - (k + \mu q) \left( \frac{\mu(1-q)(\mu+r+\mu_d)+(\mu+r)(k+\mu q)}{\mu^2(1-q)(\mu+r+\mu_d)} \right) + \frac{r\theta(k+r+\mu_d)}{(\mu+r+\mu_d)\mu^2(1-q)} \right)
\]

\[= \frac{2\mu^2(1-q)\beta^2 v_3 w_3^2}{\Lambda(\mu + k)} (p - p_c). \]
Computation of $b$

The sign of the bifurcation parameter $b$ is associated with the following non-vanishing partial derivatives of $F$, also evaluated at disease free equilibrium:

$$\frac{\partial^2 f_2}{\partial x_3 \partial \beta^*} = (1 - q), \frac{\partial^2 f_3}{\partial x_3 \partial \beta^*} = q.$$

Now

$$b = \sum_{k, i=1}^{4} v_kw_i \frac{\partial^2 f_k(0, 0)}{\partial x_i \partial \beta^*}$$

$$= v_2w_3 \frac{\partial^2 f_2(0, 0)}{\partial x_3 \partial \beta^*} + v_3w_3 \frac{\partial^2 f_3(0, 0)}{\partial x_3 \partial \beta^*}$$

$$= v_3w_3 \frac{(k + \mu q)}{(\mu + k)} > 0.$$

The eigenvectors $v_3$ and $w_3$ are positive. The bifurcation coefficient $b$ is always positive. From Theorem 2.4.1 the model system (C1) will exhibit backward bifurcation phenomena if the bifurcation coefficient $a$ defined by (C2) is positive. It is observed from equation (C2) that the positivity of $a$ is entirely dependent on the level of exogenous reinfection parameter $p$. This suggests existence of a bifurcation threshold below which backward bifurcation disappears and above which bi-stability phenomena occurs. After algebraic manipulation it can be shown that the bifurcation coefficient $a > 0$ whenever

$$p > p_c = (k + \mu q) \left( \frac{\mu(1 - q)(\mu + r + \mu d) + (\mu + r)(k + \mu q)}{\mu^2(1 - q)(\mu + r + \mu d)} - \frac{r \theta(k + \sigma \mu)}{\mu^2(1 - q)(\mu + r + \mu d)} \right)$$

$$= \frac{k + \mu q}{\mu(1 - q)} \left( \frac{\mu(1 - q)(\mu + r + \mu d) + (\mu + r)(k + \mu q)}{\mu(\mu + r + \mu d)} - \theta \right).$$
Appendix D: Basic properties and computation of $R_0$ for model equation (5.1)

This appendix relates to chapter 5.

Note that all model parameters and variables of model equation (5.1) are considered to be non-negative since the model represents a human population. First a biologically and epidemiologically feasible region is established through the following theorem:

**Theorem D.4.1** The region

$$\Lambda := \{(S, L_1, L_2, I, P, R) \in \mathbb{R}_+^6 : S + L_1 + L_2 + I + P + R \leq 1\}$$

is positively-invariant and absorbing with respect to the model equation (5.1) with initial conditions in $\mathbb{R}_+^6$.

**Proof:** The proof involves showing that the feasible solution of model equation (5.1) are uniformly bounded in the region $\Lambda$. Supposing $S, L_1, L_2, I, P, R$ is any solution of model equation (5.1) with positive initial conditions, then the total population fulfils the following inequality

$$\frac{dN}{dt} \leq \mu - \mu N.$$  \hspace{1cm} (D1)

It follows from (D1) that $\frac{dN}{dt} \leq \mu - \mu N$ implies that $\frac{dN}{dt} \leq 0$ if $N \geq 1$. Applying the standard comparison theorem [106] it can be shown that

$$N(t) \leq N(0)e^{-\mu t} + (1 - e^{-\mu t}).$$

Specifically, $N(t) \leq 1$, if $N(0) \leq 1$ for all $t > 0$. Hence, under the flow induced by equation (5.1), the region $\Lambda$ is positively invariant. Moreover, for $N(0) > 1$, the trajectory solutions $N(t)$ enters either in the region $\Lambda$ finite time or asymptotically approaches 1. Thus, in the region $\Lambda$ model equation (5.1) is considered to be mathematically and epidemiologically well posed [20] and the solution of all the trajectories generated by model (5.1) are in a biologically feasible region $\Lambda$.

The model system (5.1) has two important steady states: the disease free equilibrium and the non-trivial endemic equilibria where TB is expected to persist. By setting
all the infected states of model (5.1) to zero (i.e., \( L_1, L_2, I, P, R = 0 \)) the disease free equilibrium denoted by \( Y_0 \) is given as

\[
Y_0 = (1, 0, 0, 0, 0, 0).
\]

To compute the basic reproduction number the next generation operator (NGO) approach developed by [25] (described in detail in chapter 2) is applied. In order to use the method it is important to distinguish new infections in each class from all other changes in population. In model (5.1) the infected classes are \( L_1, L_2, I, P, R \). The system (5.1) can be written as

\[
\dot{Y} = F(Y) - V(Y),
\]

where \( Y = (S, L_1, L_2, I, P, R) \). \( F(Y) \) represents the rate of appearance of new infections and \( V \) represents the rate of transition from one compartment to another. Thus,

\[
F(Y) = (0, \beta IS, 0, 0, 0, 0)^T.
\]

The model system (5.1) has an intrinsic disease free equilibrium given as \( Y_0 = (1, 0, 0, 0, 0, 0) \). Now the derivatives of \( F(Y) \) and \( V(Y) \) with respect to the infected compartment, evaluated at disease free equilibrium can respectively, be obtained as

\[
F = \begin{bmatrix}
0 & 0 & \beta & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix},
V = \begin{bmatrix}
k_1 & 0 & 0 & 0 & 0 \\
-(1-f)\phi & k_2 & 0 & 0 & 0 \\
-\phi f & -\eta & k_3 & 0 & -\omega \\
-\theta & -\rho & 0 & \mu & 0 \\
0 & 0 & -(\tau + \alpha) & 0 & k_4 \\
\end{bmatrix}
\]

with

\[
V^{-1} = \begin{bmatrix}
\frac{1}{k_1} & 0 & 0 & 0 & 0 \\
\frac{1}{k_2} & \frac{1}{k_2} & 0 & 0 & 0 \\
0 & k_{32} & k_{33} & 0 & k_{35} \\
0 & \frac{1}{\mu \kappa_2} & 0 & \frac{1}{\mu} & 0 \\
0 & k_{52} & k_{53} & 0 & k_{55} \\
\end{bmatrix},
\]
where

\[ k_1 = (\theta + \mu + \phi), \]
\[ k_2 = (\mu + \eta + \rho), \]
\[ k_3 = (\mu + d + \tau + \alpha), \]
\[ k_4 = (\mu + \omega), \]
\[ k_{21} = \frac{(1 - f)\phi}{k_1 k_2}, \]
\[ k_{31} = \frac{(f \phi(\mu + \rho) + \eta \phi)k_4}{k_1 k_2((\mu + d)k_4 + \mu(\tau + \alpha))}, \]
\[ k_{32} = \frac{\eta k_4}{k_2((\mu + d)k_4 + \mu(\tau + \alpha))}, \]
\[ k_{33} = \frac{k_4}{(\mu + d)k_4 + \mu(\tau + \alpha)}, \]
\[ k_{35} = \frac{\omega}{(\mu + d)k_4 + \mu(\tau + \alpha)}, \]
\[ k_{41} = \frac{k_3(\mu \theta k_1 + \mu \rho \phi(1 - f)) + (\mu + d)\omega(\theta k_2 + \rho \phi(1 - f))}{\mu k_3 k_2((\mu + d)k_4 + \mu(\tau + \alpha))}, \]
\[ k_{51} = \frac{\phi(\tau + \alpha)(f(\mu + \rho) + \eta)}{k_1 k_2((\mu + d)k_4 + \mu(\tau + \alpha))}, \]
\[ k_{52} = \frac{\eta(\tau + \alpha)}{k_2((\mu + d)k_4 + \mu(\tau + \alpha))}, \]
\[ k_{53} = \frac{(\tau + \alpha)}{((\mu + d)k_4 + \mu(\tau + \alpha))}, \]
\[ k_{55} = \frac{k_3}{((\mu + d)k_4 + \mu(\tau + \alpha))}. \]

Following van den Driessche and Watmough [25] the basic reproduction number is defined as the spectral radius of the next generation matrix, \( FV^{-1} \) (i.e., \( R_0 = \bar{\rho}(FV^{-1}) \)), where \( \bar{\rho} \) denote the spectral radius, which for the model system (5.1) is given as

\[ R_0 = \frac{\beta(\mu + \omega)(f \phi(\mu + \rho) + \eta \phi)}{((\mu + d)k_4 + \mu(\tau + \alpha))(\mu + \eta + \rho)(\theta + \mu + \phi)}. \]
Appendix E: Interpretation for $R_0$ of model equation (5.1)

This appendix relates to chapter 5.

The expression for basic reproduction number $R_0$ as given above in equation (D3) can be decoupled to be expressed as a combination of various pathways leading to active TB. First starting with a simpler scenario where treated active TB individuals do not reactivate. That is by setting $\omega = 0$, the $R_0$ expression (D3) can be rewritten as

$$R'_0 = \left( \frac{\beta}{\mu + d + \tau + \alpha} \right) \left[ \left( \frac{f \phi}{\mu + \theta + \phi} \right) + \left( \frac{(1-f) \phi}{\mu + \theta + \phi} \right) \left( \frac{\eta}{\mu + \eta + \rho} \right) \right]. \quad (E1)$$

Each factor in $R'_0$ has an epidemiological interpretation as follows:

(i) $\frac{\beta}{\mu + d + \tau + \alpha}$ represent the average number of secondary cases an individual with active TB produces;

(ii) $\left( \frac{f \phi}{\mu + \theta + \phi} \right)$ reflect that individuals progress toward active TB through the early latent compartment;

(iii) $\left( \frac{(1-f) \phi}{\mu + \theta + \phi} \right) \left( \frac{\eta}{\mu + \eta + \rho} \right)$, reflect that individuals progress to active TB through late latent compartment. Note that $\left( \frac{(1-f) \phi}{\mu + \theta + \phi} \right)$ account for the fraction that proceed to late latent class while $\left( \frac{\eta}{\mu + \eta + \rho} \right)$ is the probability of surviving the late latent compartment and progressing to active TB.

Now, considering that $\omega > 0$, recovered individuals have an additional chance to progress to active TB that is independent of re-exposures. This additional contribution
for the infectious period is given by the term

\[ \Upsilon = 1 + \frac{(\tau + \alpha)}{\mu + d + \tau + \alpha \mu + \omega} \omega \left( \frac{(\tau + \alpha)}{\mu + d + \tau + \alpha \mu + \omega} \right)^2 + \cdots \]

\[ = \frac{1}{1 - \frac{(\tau + \alpha)}{\mu + d + \tau + \alpha \mu + \omega} \omega} \]

\[ = \frac{(\mu + d + \tau + \alpha)(\mu + \omega)}{\mu(\mu + d + \tau + \alpha) + \omega(\mu + d)} \]  \hspace{1cm} \text{(E2)}

that results from the innumerable chances the infected individual has to repeat this event. The full expression for \( R_0 \) is then obtained by the product of equations (E1) and (E2). That is

\[ R_0 = \Upsilon R'_0. \]

The following result is deduced from Theorem 2 of [25].

**Lemma E.5.1** The DFE \( Y_0 \) of model equation (5.1) is locally asymptotically stable (LAS) whenever \( R_0 \) is below one and unstable whenever \( R_0 \) exceeds one.

In general it is known that a value of \( R_0 < 1 \) implies that each individual is only able to infect less than one individual on average, such that the disease will die out, whereas a value of \( R_0 > 1 \) implies that each individual is able to infect more than one individuals and that endemic disease will persist within the population. Hence, \( R_0 = 1 \) is a crucial endemic threshold in determining the epidemic trajectory. In what follows, it is shown that if all cohorts subject to reinfection have complete immunity (i.e., \( \sigma_i = 0, i = 1, 2, 3 \)) then the elimination of TB from the community is independent of the initial sizes of the sub-population. Consequently, the DFE is shown to be globally asymptotically stable.
Appendix F: Global stability of DFE for a special case of model equation (5.1)

This appendix relates to chapter 5.

The global stability of the DFE of model equation (5.1) is analysed for a special case. That is, in a case where all post-infection cohorts are not reinfected, and hence $\sigma_i = 0$, ($i = 1, 2, 3$). Applying a Theorem from [224], the global stability of DFE $Y_0$ (for a special case) is investigated when the reproduction number is less than one.

**Theorem F.6.1** Let $X = S$ and $Y = (L_1, L_2 I, P, R)$ where $X \in \mathbb{R}^1$ and $Y \in \mathbb{R}^5$ represent uninfected population and infected population, respectively. Further, let

\[
X' = F(X, Y), \\
Y' = G(X, Y), \quad G(X, 0) = 0,
\]

where $X'$ and $Y'$ denote differentiation with respect to time. The DFE can be rewritten as $Y_0 = (X^*, 0)$ where $X^* = 1$. Now to guarantee a local asymptotic stability the following conditions (i) and (ii) have to be fulfilled:

(H1) For $X' = F(X, 0)$, $D_0$ is globally asymptotically stable (g.a.s);

(H2) $G(X, Y) = AY - \hat{G}(X, Y)$ where $\hat{G}(X, Y) \geq 0$, for $(X, Y) \in \Lambda$, $A = D_Y G(X^*, 0)$.

**Lemma F.6.1** For $\sigma_i = 0$, the fixed point $Y_0 = (X^*, 0)$ is a g.a.s equilibrium of model system (5.1) provided $R_0 < 1$ and that conditions (i) and (ii) hold.

**Proof:** Let $X = S$, and $Y = (L_1, L_2, I, P, R)$ and $Y_0 = (X^*, 0)$, where $X^* = 1$.

Then $X' = F(X, Y) = (\mu + dI - \mu S - \beta IS)$. At DFE $S = 1$ so $F(X, 0) = (\mu - \mu S) = 0$.

Hence,

\[
X' = F(X, 0) = (\mu - \mu S) = 0.
\]

As $t \to \infty$, $X \to X^*$. This implies that the DFE $Y_0$ is g.a.s (that is condition (H1) is fulfilled).
Now consider

$$G(X, Y) = AY - \tilde{G}(X, Y),$$

$$A = \begin{bmatrix}
-(\theta + \mu + \phi) & 0 & \beta X^* & 0 & 0 \\
(1 - f)\phi & -(\mu + \eta + \rho) & 0 & 0 & 0 \\
\phi f & \eta & -(\mu + \eta + \rho) & 0 & \omega \\
\theta & \rho & 0 & -\mu & 0 \\
0 & 0 & (\tau + \alpha) & 0 & -(\mu + \omega)
\end{bmatrix}.$$ 

Then,

$$\tilde{G}(X, Y) = \begin{bmatrix}
\tilde{G}_1(X, Y) \\
\tilde{G}_2(X, Y) \\
\tilde{G}_3(X, Y) \\
\tilde{G}_4(X, Y) \\
\tilde{G}_5(X, Y)
\end{bmatrix} = \begin{bmatrix}
\beta I(X^* - S) \\
0 \\
0 \\
0 \\
0
\end{bmatrix}.$$

Given that the total population is bounded such that $N = S + L_1 + L_2 + I + P + R \leq X^* = 1$, it follows that $S \leq N \leq 1$. Hence $\tilde{G}(X, Y) \geq 0$ which now imply that condition (H2) is satisfied. Consequently, by Lemma F.6.1 $Y_0 = (X^*, 0, 0, 0, 0, 0)$ is globally asymptotically stable whenever $R_0 < 1$ and $\sigma_i = 0 (i = 1, 2, 3)$. The epidemiologically implication of DFE being g.a.s is that TB will be eliminated from the community if the threshold quantity $R_0$ is maintained (or decreased) to a value below unity.
Appendix G: Existence of endemic equilibria for model (5.1)

This appendix relates to chapter 5.

Due to the complex nature of the model system (5.1) the steady states cannot be explicitly expressed in terms of model parameters. Consequently, all steady states are written in terms of $I^*$. Namely

$$ S^* = \frac{\mu + dI^*}{\mu + \beta I^*}, $$

$$ L^*_1 = \frac{(\mu + \eta + \rho + \sigma_1 \beta I^*)(\mu + d + \tau + \alpha)(\mu + \sigma_3 \beta I^*) + \omega (\mu + d)I^*}{(\mu + \omega + \sigma_3 \beta I^*)(f \phi (\mu + \rho + \sigma_1 \beta I^*) + \eta \phi)}, $$

$$ L^*_2 = \frac{(1 - f) \phi (\mu + d + \tau + \alpha)(\mu + \sigma_3 \beta I^*) + \omega (\mu + d)I^*}{(\mu + \omega + \sigma_3 \beta I^*)(f \phi (\mu + \rho + \sigma_1 \beta I^*) + \eta \phi)}, $$

$$ P^* = \frac{((\mu + d + \tau + \alpha)(\mu + \sigma_3 \beta I^*) + \omega (\mu + d))(\theta (\mu + \eta + \rho + \sigma_1 \beta I^*) + \rho (1 - f) \phi I^*)}{(\mu + \sigma_2 \beta I^*)(\mu + \omega + \sigma_3 \beta I^*)(f \phi (\mu + \rho + \sigma_1 \beta I^*) + \eta \phi)}, $$

$$ R^* = \frac{(\tau + \alpha)I^*}{\mu + \omega + \sigma_3 \beta I^*}. $$

$I^*$ can be solved from the following polynomial:

$$ g(I^*) = I^*(d_4 I^{*4} + d_3 I^{*3} + d_2 I^{*2} + d_1 I^* + d_0) = 0, $$
where
\[
d_4 = -\sigma_1\sigma_2\sigma_3\mu(f\phi + \mu + d + \tau + \alpha)\beta^4,
\]
\[
d_3 = \sigma_1\sigma_2^3\mu(1 - f)\phi(\mu + d + \tau + \alpha) + \sigma_1\sigma_2\sigma_3\mu(1 - f)\phi\beta^3(\mu + d + \tau + \alpha)
\]
\[
\quad + \sigma_1\sigma_3\mu\beta^3(1 - f)\phi(\mu + d + \tau + \alpha) + \sigma_1\sigma_2\beta^3(1 - f)\phi\omega(\mu + d)
\]
\[
\quad + \sigma_1\sigma_2\sigma_3\mu\beta^3\theta(\mu + d + \tau + \alpha)
\]
\[
\quad + \sigma_2\sigma_3\beta^3\theta(\mu + \eta + \rho)(\mu + d + \tau + \alpha) + \sigma_1\sigma_2\theta\beta^3\mu(\mu + d + \tau + \alpha)
\]
\[
\quad + \sigma_1\sigma_2\theta^3\omega(\mu + d) + \sigma_2\sigma_3\rho(1 - f)\phi\beta^3(\mu + d + \tau + \alpha) + \sigma_1\sigma_2\sigma_3\beta^3(\tau + \alpha)f\phi
\]
\[
\quad + \sigma_1\sigma_3\mu\beta^3(\tau + \alpha)\phi + \sigma_2\sigma_3\beta^3(\tau + \alpha)(\phi(\mu + \rho) + \eta\phi) + \sigma_1\sigma_3\mu\beta^3 f\phi
\]
\[
\quad + \sigma_1\sigma_2d(\mu + \omega)\beta^3 f\phi + \sigma_2\sigma_3\beta^3(\phi(\mu + \rho) + \eta\phi)
\]
\[
\quad + \sigma_1\sigma_2\sigma_3\mu\beta^3(\theta + \mu + \phi)\beta^3(\mu + d + \tau + \alpha)
\]
\[
\quad + \sigma_1\sigma_2\beta^3(\mu + d + \tau + \alpha)(\theta + \mu + \phi)
\]
\[
\quad - \sigma_2\sigma_3\beta^3(\mu + \eta + \rho)(\mu + d + \tau + \alpha)(\theta + \mu + \phi)
\]
\[
\quad - \sigma_1\sigma_2\beta^3(\theta + \mu + \phi)(\mu + d + \tau + \alpha)
\]
\[
\quad - \sigma_1\sigma_2\beta^3(\omega(\mu + d)(\theta + \mu + \phi),
\]
\[
d_2 = \sigma_1\sigma_2\mu^2\beta^2(1 - f)\phi(\mu + d + \tau + \alpha) + \sigma_1\beta^2\mu^2(1 - f)\phi(\mu + d + \tau + \alpha)
\]
\[
\quad + \sigma_1\sigma_3\beta^2(1 - f)\phi\mu^2(\mu + d + \tau + \alpha) + \sigma_1\sigma_2\beta^2(1 - f)\phi\omega\mu(\mu + d)
\]
\[
\quad + \sigma_1\beta^2(1 - f)\phi\omega\mu(\mu + d) + \sigma_2\sigma_3\mu\beta^2(\mu + \eta + \rho)(\mu + d + \tau + \alpha)
\]
\[
\quad + \sigma_1\sigma_2\beta^2\mu^2\beta(\mu + d + \tau + \alpha) + \sigma_1\sigma_2\mu\beta^2\beta(\mu + d + \tau + \alpha)
\]
\[
\quad + \sigma_2\sigma_3\beta^2\mu^2(1 - f)\phi(\mu + d +\tau + \alpha) + \sigma_2\beta^2\phi(\mu + \eta + \rho)(\mu + d + \tau + \alpha)
\]
\[
\quad + \sigma_2\beta^2\phi(\mu + \eta + \rho)(\omega(\mu + d) + \sigma_2\beta^2(\rho(1 - f)\phi\mu(\mu + d + \tau + \alpha)
\]
\[
\quad + \sigma_2\beta^2(1 - f)\phi\omega(\mu + d) + \sigma_1\sigma_3\beta^2\mu^2(\tau + \alpha)f\phi
\]
\[
\quad + \sigma_3\beta^2(\tau + \alpha)\mu(\phi(\mu + \rho) + \eta\phi) + \sigma_1\sigma_3\mu^2\beta^3 f\phi
\]
\[
\quad + \sigma_1\sigma_2\mu(\mu + \omega)\beta^3 f\phi + \sigma_2\sigma_3\beta^3(\phi(\mu + \rho) + \eta\phi) + \sigma_1\beta^2 d(\mu + \omega)f\phi
\]
\[
\quad + \sigma_3\beta^2 d(\phi(\mu + \rho) + \eta\phi) + \sigma_2 d\beta^2(\mu + \omega)(f\phi(\mu + \rho) + \eta\phi)
\]
\[
\quad + \sigma_2\sigma_3\beta^2(\tau + \alpha)(\phi(\mu + \rho) + \eta\phi) - \sigma_1\sigma_3\beta^2\beta^2(\mu + d + \tau + \alpha)(\mu + \phi + \theta)
\]
\[
\quad - \sigma_2\beta^2\mu(\theta + \mu + \phi)(\mu + \eta + \rho)(\mu + d + \tau + \alpha)
\]
\[
\quad - \sigma_1\sigma_2\beta^2(\phi(\mu + \rho))(\mu + d + \tau + \alpha) - \sigma_1\sigma_2\beta^2(\omega(\mu + d)
\]
\[
\quad - \sigma_3\beta^2(\mu + \theta + \phi)(\mu + \eta + \rho)(\mu + d + \tau + \alpha)
\]
\[
\quad - \sigma_1\beta^2(\phi(\mu + \rho))(\mu + d + \tau + \alpha) - \sigma_2\beta^2(\phi(\mu + \rho))(\mu + \eta + \rho)(\mu + d + \tau + \alpha)
\]
\[
\quad - \sigma_1\beta^2(\phi(\mu + \rho))(\mu + d + \tau + \alpha) - \sigma_2\beta^2(\phi(\mu + \rho))(\mu + \eta + \rho)(\mu + d + \tau + \alpha)
\]
\[
\quad - \sigma_1\beta^2(\phi(\mu + \rho))(\mu + d + \tau + \alpha) - \sigma_2\beta^2(\phi(\mu + \rho))(\mu + \eta + \rho)(\mu + d + \tau + \alpha),
\]
\[ d_1 = \sigma_1 \beta (1 - f) \phi \mu^3 (\mu + d + \tau + \alpha) + \sigma_1 \beta (1 - f) \phi (\mu + d) \omega \mu^2 \]
\[ + \sigma_2 \beta \mu^2 \theta (\mu + \eta + \rho) (\mu + d + \tau + \alpha) + \sigma_2 \beta \mu \theta (\mu + \eta + \rho) \omega (\mu + d) \]
\[ + \sigma_2 \beta \mu^2 \rho (1 - f) \phi (\mu + d + \tau + \alpha) + \sigma_2 \beta \mu \rho (1 - f) \phi \omega (\mu + d) \]
\[ + \sigma_3 \beta (\tau + \alpha) \mu^2 (f \phi (\mu + \rho) + \eta \phi) + \sigma_1 \beta^2 \mu^2 (\mu + \omega) f \phi \]
\[ + \sigma_3 \beta^2 \mu^2 (f \phi (\mu + \rho) + \eta \phi) + \sigma_2 \beta^2 \mu (\mu + \omega) (f \phi (\mu + \rho) + \eta \phi) \]
\[ + \beta d \mu (\mu + \omega) (f \phi (\mu + \rho) + \eta \phi) - \sigma_3 \beta^2 \mu^2 (\theta + \mu + \phi) (\mu + \eta + \rho) (\mu + d + \tau + \alpha) \]
\[ - \sigma_1 \beta^3 (\theta + \mu + \phi) (\mu + d + \tau + \alpha) - \sigma_1 \beta^2 (\theta + \mu + \phi) \omega (\mu + d) \]
\[ - \sigma_2 \beta^2 (\theta + \mu + \phi) (\mu + \eta + \rho) (\mu + d + \tau + \alpha) \]
\[ - \beta \mu^2 (\theta + \mu + \phi) (\mu + \eta + \rho) (\mu + d + \tau + \alpha) \]
\[ - \beta \mu \omega (\theta + \mu + \phi) (\mu + d) (\mu + \eta + \rho) , \]

\[ d_0 = \beta \mu^2 (\mu + \omega) (f \phi (\mu + \rho) + \eta \phi) \]
\[ - \mu^2 (\theta + \mu + \phi) (\mu + \eta + \rho) ((\mu + \omega) (\mu + d) + \mu (\tau + \alpha)) . \]
Appendix H: Procedure for calculating reinfection threshold as derived in the literature

This appendix relates to chapter 5.

For any given compartmental model with reinfection mechanisms, it is possible to define the reinfection submodel and subsequently reinfection threshold RT. According to [97, 121, 178] the RT is interpreted as the transmission potential which if exceeded reinfection can sustain transmission in a partially immunized population. To describe the procedure (adopted from [178]) involved in obtaining reinfection submodel in which RT is obtained consider the following systems of equations:

\[ x'_i = f_i(x), \quad i = 1, \ldots, n + m + k, \quad (H1) \]

where \( x = (x_1, \ldots, x_{n+m+k})^t \) with each \( x_i \geq 0 \) representing the proportion of individuals in each compartment. The equations are arranged such that the first \( n \) equations correspond to the partially immunized classes, followed by the \( m \) remaining classes except for the \( k \) which represent the totally susceptible classes that are classified last. Differentiation between infected, partially immunized and susceptible compartments should be deduced from the epidemiological description of the model and cannot not be determined from the structure of the model equations alone. Hence, the first step involves removing the totally susceptible classes and adding the recruitment rate into the compartments that are partially immunized. Let \( y = (x_1, \ldots, x_{n+m}) \). Now defining

\[ \tilde{g}_i(y) = \tilde{f}_i(y, 0) + r_i(y, 0), \quad (H2) \]

for \( i = 1, \ldots, n + m \), where \( r_i \) is the recruitment rate verifying \( r_i = 0 \) for \( i = n + 1, \ldots, n + m \) and \( \tilde{f}_i \) is given by \( f_i \) with the rate of transfer of individuals into the last \( k \) compartments set to zero. Specifically this last condition indicate that the terms accounting for temporary immunity are excluded. Again, it is important to mention that the recruitment functions \( r_i \) depend on the epidemiological interpretation of the model. Supposing there is one compartment subject to reinfection, then the \( r_i \) is the sum of all recruitment rates of the model, which correspond to the total inflow of individuals into the model at each time step. On one hand, it is possible to have multiple compartments subject to reinfection with varied protection or infection progression rates. In case of this scenario, for each of the compartments subject to
reinfection respective reinfection submodels should be considered, hence leading to several reinfection thresholds. For each submodel, there should be only one or a matching number of reinfection compartments of interest at a time. Thus, in such case the recruitment function has to be defined as for the simpler cases. Consequently, the reinfection submodel is given by the following differential equations of \( n+m \) dimension:

\[
y'_i = g_i(y) = g^+_i(y) - g^-_i(y), \quad i = 1, \cdots, n+m, \tag{H3}
\]

where \( y = (y_1, \cdots, y_{n+m})' \) and \( g^+_i \) and \( g^-_i \) are the rates representing inflow and outflow of individuals in compartment \( i \), respectively. Note that since each function represents a directed transfer of individuals proportion, they are all nonnegative. Hence, the following:

(i) If \( y_i \geq 0 \), then \( g^+_i, g^-_i \geq 0 \), for \( i = 1, \cdots, n+m \). On one hand if a compartment is empty there can be no transfer out of the compartment;

(ii) If \( y_i = 0 \), then \( g^-_i = 0 \), for \( i = 1, \cdots, n+m \). Considering that the disease transmission model given by (H3) with \( g_i \) fulfilling conditions (i) and (ii), then the non-negative cone \( (y_i \geq 0, i = 1, \cdots, n+m) \) is forward invariant. According to Theorems 1.1.8 and 1.1.9 of [100] for each non-negative initial condition there exist a unique nonnegative solution of model equation (H3). Letting \( Y_0 \) represent the set of all disease free states, that is \( Y_0 = \{y \geq 0 : y_i = 0, i = n+1, \cdots, n+m\} \). For the disease free subspace \( Y_0 \) to be invariant one has to assume that if the population is free of disease then the population will remain disease free. No (density-independent) immigration of infectives is allowed. This property is stated as follows;

(iii) If \( y \in Y_0 \), then \( g^+_i = 0 \), for \( i = n+1, \cdots, n+m \).

This condition means specifically that some immune processes should be set to zero in the reinfection submodel.

Now after constructing the reinfection submodel (H3) and supposing the submodel fulfil conditions (i)-(iii). If the reinfection submodel undergo bifurcation through the transmission parameter \( \beta \), then the bifurcation will correspond to the reinfection threshold of the full model. Considering the following linearized system:

\[
y' = Dg(y_0)(y - y_0), \tag{H4}
\]
where $Dg$ is the Jacobian matrix evaluated at the disease free state $y_0 \in Y_0$. The bifurcation point for $y = y_0$ can be easily obtained by setting the determinant of $Dg$ to zero and solving the respective equation for $\beta$. Let $\beta^{RT}$ denote the solution, then $\beta^{RT}$ correspond to the point where the stability of the disease free equilibrium occur. Now to obtain the reinfection threshold in terms of $R_0$, replace $\beta$ by $\beta^{RT}$ in the formula for $R_0$. 
Appendix I: Computation of reinfection thresholds

This appendix relates to chapter 5.

Following case (i), case (ii) and case (iii) as shown in the text of chapter 5, the respective reinfection submodels of model equation (5.1) can be obtained as

\[
\frac{dL_2}{dt} = \lambda + (1 - f)\phi L_1 - (\mu + \rho + \sigma_1 \beta I)L_2,
\]

\[
\frac{dL_1}{dt} = \sigma_1 \beta I L_2 - (\theta + \mu + \phi)L_1,
\]

\[
\frac{dI}{dt} = \phi f L_1 - (\mu + d + \tau + \alpha)I. (I1)
\]

\[
\frac{dP}{dt} = \lambda + \theta L_1 - (\mu + \sigma_2 \beta I)P,
\]

\[
\frac{dL_1}{dt} = \sigma_2 \beta I P - (\theta + \mu + \phi)L_1,
\]

\[
\frac{dI}{dt} = \phi f L_1 - (\mu + d + \tau + \alpha)I. (I2)
\]

\[
\frac{dR}{dt} = \lambda + (\tau + \alpha)I - (\mu + \sigma_3 \beta I)R,
\]

\[
\frac{dL_1}{dt} = \sigma_3 \beta I R - (\theta + \mu + \phi)L_1,
\]

\[
\frac{dI}{dt} = \phi f L_1 - (\mu + d + \tau + \alpha)I. (I3)
\]

The reinfection threshold for the first reinfection submodel (I1) can be computed following the procedure followed for the reinfection submodel for a scenario where \(\sigma_1 = \sigma_2 = \sigma_3\), (see the text) and it can be found that the reinfection submodel sustains an endemic equilibrium when transmission is above the critical value given in (5.5). Consequently, the reinfection threshold written in terms of \(R_0\) is similar to equation (5.6). The other, reinfection thresholds, RT2 and RT3 are obtained from their respective reinfection submodels (9) and (10) as

\[
\beta = \frac{1}{\sigma_2} \frac{(\theta + \mu + \phi)(\mu + d + \tau + \alpha)}{\phi f} = RT2,
\]

\[
\beta = \frac{1}{\sigma_3} \frac{(\theta + \mu + \phi)(\mu + d + \tau + \alpha)}{\phi f} = RT3.
\]
Similar to RT1, both RT2 and RT3 represent a transmission rate above which disease prevalences steeply increases. The reinfection thresholds RT2 and RT3 expressed in terms of $R_0$ are respectively, given as

$$
R_{0}^{RT2} = \frac{1}{\sigma_2 f\phi((\mu + d)(\mu + \omega) + \mu(\tau + \alpha))(\mu + \eta + \rho)} \frac{(\mu + d + \tau + \alpha)(\mu + \omega)(f\phi(\mu + \rho) + \eta\phi)}{f\phi((\mu + d)(\mu + \omega) + \mu(\tau + \alpha))(\mu + \eta + \rho)}. \tag{14}
$$

$$
R_{0}^{RT3} = \frac{1}{\sigma_3 f\phi((\mu + d)(\mu + \omega) + \mu(\tau + \alpha))(\mu + \eta + \rho)} \frac{(\mu + d + \tau + \alpha)(\mu + \omega)(f\phi(\mu + \rho) + \eta\phi)}{f\phi((\mu + d)(\mu + \omega) + \mu(\tau + \alpha))(\mu + \eta + \rho)}. \tag{15}
$$
Appendix J: Endemic equilibria of the submodel (5.3)

This appendix relates to chapter 5.

Let $H^* = L_2^* + P^* + R^*$ then, the equilibrium of the reinfection submodel (5.3) can be easily obtained by setting the right-hand terms to zero and evaluating for $H^*, L_1^*$ and $I^*$ as

\[
H^* = \frac{(\theta + \mu + \phi)(\mu + d + \tau + \alpha)}{f \phi \sigma_1 \beta}, \quad (J1)
\]

\[
L_1^* = \frac{(\mu + d + \tau + \alpha)I^*}{f \phi}, \quad (J2)
\]

\[
I^* = \frac{f \phi \left[ \beta - \frac{(\theta + \mu + \phi)(\mu + d + \tau + \alpha)}{\sigma_1 \phi} \right]}{\beta [f \phi + (\mu + d + \tau + \alpha)]}. \quad (J3)
\]
Appendix K: Derivation of $R_0$ for heroin epidemic model

This appendix relates to chapter 6.

Generally the geometric sequence is given as $\{a, ar, ar^2, ar^3, \cdots\}$ and the sum of a certain number of terms of the geometric sequence is given as $S_n = \frac{a(1 - r^n)}{1 - r}$ where $S_n$ is the sum of $n$ terms ($n^{th}$ partial sum), $a$ is the first term and $r$ is the common ratio. Now considering the geometric sequence from expression (6.4).

Note that $a = 1$ and $r = \frac{p\alpha}{(\mu + \delta_1 + \xi + \alpha)(p + \sigma + \delta_2 + \mu)} < 1$.

$$S_n = \frac{a(1 - r^n)}{1 - r} = \frac{a}{1 - r} \text{ since } r < 1.$$  \hspace{1cm} (K1)

Substituting $a$ and $r$ in (K1) yield

$$S_n = \frac{(\mu + \delta_1 + \xi + \alpha)(p + \sigma + \delta_2 + \mu)}{(\mu + \delta_1 + \xi)(p + \sigma + \delta_2 + \mu) + \alpha(\sigma + \delta_2 + \mu)}. \hspace{1cm} (K2)$$

Multiplying (K2) by $\frac{1}{\mu + \delta_1 + \xi + \alpha}$ gives the required expression

$$\frac{p + \sigma + \delta_2 + \mu}{(\mu + \delta_1 + \xi)(p + \sigma + \delta_2 + \mu) + \alpha(\sigma + \delta_2 + \mu)}, \hspace{1cm} (K3)$$

which if multiplied by the effective contact rate $\beta$ and the average recruitment rate $\frac{\Lambda}{\mu}$ yields the basic reproduction number $R_0$. 
Appendix L: Proof of existence of backward bifurcation for heroin epidemic model (6.1)

This appendix relates to chapter 6.

**Proof.** To prove existence of backward bifurcation in model equation (6.1) the Center Manifold approach as outlined by Castillo-Chavez and Song in [58] is used. First for clarity and understanding of the Center Manifold Theory the model equation (6.1) variables are transformed as follows:

\[ y_1 = S, \quad y_2 = U_1, \quad y_3 = U_2, \quad y_4 = U_3 \]

and the total population \( N = \sum_{j=1}^{4} y_j \). Define \( Y = (y_1, y_2, y_3, y_4)^T \) (T denote transpose), such that the model equation (6.1) can be rewritten as \( \frac{dY}{dt} = F(y) \) where \( F = (f_1, f_2, f_3, f_4) \).

Hence,

\[
\begin{align*}
\frac{dy_1}{dt} &= f_1 = \Lambda - \beta y_1 y_2 - \mu y_1, \\
\frac{dy_2}{dt} &= f_2 = \beta y_1 y_2 + py_3 - (\mu + \delta_1 + \xi) y_2 - \frac{\alpha y_2}{1 + \omega y_2}, \\
\frac{dy_3}{dt} &= f_3 = \frac{\alpha y_2}{1 + \omega y_2} - (p + \sigma + \delta_2 + \mu) y_3, \\
\frac{dy_4}{dt} &= f_4 = \sigma y_3 + \xi y_2 - \mu y_4.
\end{align*}
\]

(L1)

Now let \( \beta = \beta^* \) be the bifurcation parameter. Observe that at \( R_0 = 1 \),

\[ \beta = \beta^* = \frac{\alpha (\mu + \delta_2 + \sigma) + (\mu + \delta_1 + \xi)(p + \delta_2 + \mu + \sigma)}{y_1^* (p + \delta_2 + \mu + \sigma)}, \]

where \( y_1^* = \frac{\Lambda}{\mu} = S_0 \). With \( \beta = \beta^* \) the transformed model equation (L1) has a simple eigenvalue with zero real part and all other eigenvalues are negative (that is has a hyperbolic equilibrium point). Thus, Center Manifold Theory can be applied to investigate the local dynamics of the transformed system (L1) near \( \beta = \beta^* \). Now the Jacobian matrix of the transformed system evaluated at heroin free equilibrium HFE is obtained as

\[
J_{HFE} = \begin{pmatrix}
-\mu & -\beta S_0 & 0 & 0 \\
0 & \beta S_0 - (\mu + \delta_1 + \xi + \alpha) & p & 0 \\
0 & \alpha & -\frac{(p + \sigma + \delta_2 + \mu)}{\sigma} & 0 \\
0 & \xi & \sigma & -\mu
\end{pmatrix}.
\]
It is easy to obtain the right eigenvectors of this Jacobian matrix as \( \bar{V} = (\bar{v}_1, \bar{v}_2, \bar{v}_3, \bar{v}_4)^T \), where

\[
\begin{pmatrix}
\bar{v}_1 \\
\bar{v}_2 \\
\bar{v}_3 \\
\bar{v}_4 \\
\end{pmatrix}
= 
\begin{pmatrix}
\frac{-\beta S_0}{\mu} \\
1 \\
\frac{\alpha}{p+\sigma+\delta_2+\mu} \\
\frac{\xi(p+\sigma+\delta_2+\mu)+\sigma\alpha}{\mu(p+\sigma+\delta_2+\mu)} \\
\end{pmatrix}^T \bar{v}_2.
\]

\( \bar{v}_2 > 0 \). Similarly, it is possible to obtain the left eigenvectors which are denoted by \( \tilde{W} = (\tilde{w}_1, \tilde{w}_2, \tilde{w}_3, \tilde{w}_4) \), so that \( \tilde{w}_1 = 0 \), \( \tilde{w}_2 = \tilde{w}_2 > 0 \), \( \tilde{w}_3 = \frac{p\tilde{w}_2}{p+\sigma+\delta_2+\mu} \), \( \tilde{w}_4 = 0 \).

Now proceeding to obtain the bifurcation coefficients \( a \) and \( b \) as defined in Theorem 4.1 of [58].

**Calculation of coefficient \( a \)**

First the non-vanishing partial derivatives of the transformed model (L1) evaluated at heroin free equilibrium are obtained as

\[
\frac{\partial^2 f_1(0,0)}{\partial y_1 \partial y_2} = -\beta^*, \quad \frac{\partial^2 f_2(0,0)}{\partial y_1 \partial y_2} = \beta^*, \quad \frac{\partial^2 f_2(0,0)}{\partial^2 y_2^2} = 2\omega\alpha, \quad \frac{\partial^2 f_3(0,0)}{\partial^2 y_2^2} = -2\omega\alpha,
\]

so that

\[
a = \sum_{k,i,j=1}^{4} \tilde{w}_k \tilde{v}_i \tilde{v}_j \frac{\partial^2 f_k(0,0)}{\partial y_i \partial y_j}
= \tilde{w}_2 \tilde{v}_1 \tilde{v}_2 \frac{\partial^2 f_2(0,0)}{\partial y_1 \partial y_2} + \tilde{w}_2 \tilde{v}_2 \tilde{v}_2 \frac{\partial^2 f_2(0,0)}{\partial^2 y_2^2} + \tilde{w}_3 \tilde{v}_2 \tilde{v}_2 \frac{\partial^2 f_3(0,0)}{\partial^2 y_2^2}
= 2\tilde{w}_2 \tilde{v}_2^2 \left[ \frac{\mu \omega \alpha (\sigma + \delta_2 + \mu) - \beta^* (\alpha (\mu + \delta_2 + \sigma) + (\mu + \delta_1 + \xi)(p + \delta_2 + \mu + \sigma))}{\mu (p + \sigma + \delta_2 + \mu)} \right]
= \frac{2\tilde{w}_2 \tilde{v}_2^2 \alpha (\sigma + \delta_2 + \mu)}{p + \sigma + \delta_2 + \mu} [\omega - \omega_c],
\]
where $\omega_c$ remains as previously defined in equation (6.11).

**Calculation of coefficient $b$**

The bifurcation coefficient $b$ is obtained as

$$b = \sum_{k,i=1}^{4} \tilde{w}_k \tilde{v}_i \frac{\partial^2 f_k(0,0)}{\partial y_i \partial \beta^*}$$

$$= \frac{\tilde{w}_2 \tilde{v}_2 \Lambda}{\mu} > 0.$$

According to Theorem 4.1 of [58] if both bifurcation coefficients $a$ and $b$ are positive then model (6.1) will exhibit backward bifurcation. Observe that $b$ is always positive while $a > 0$ if and only if $\omega > \omega_c$. Thus, if $\omega > \omega_c$ then model (6.1) will exhibit the phenomenon of backward bifurcation.
Bibliography


