Mathematical aspects of infectious disease dynamics

Een wiskundige blik op de dynamica van besmettelijke ziekten
(met een samenvatting in het Nederlands)

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. W. H. Gispen, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 5 september 2007 des middags te 12.45 uur

door

BARBARA BOLDIN

geboren op 5 december 1976 te Jesenice in Slovenië

Promotoren: prof. dr. O. Diekmann
prof. dr. M. J. M. Bonten

Dit proefschrift werd mede mogelijk gemaakt met financiële steun van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO-CLS project no. 635.100.002).

Mathematical aspects of infectious disease dynamics

Barbara Boldin

Mathematics Subject Classification (2000): 37C10, 37G10, 37N25, 92D15, 92D25, 92D30, 92D40

ISBN: 978-90-9022069-7

Printed by PrintPartners Ipskamp, Enschede

“Would you tell me, please, which way I ought to go from here?”
“That depends a good deal on where you want to get to,” said the Cat.
“I don’t much care where ...” said Alice.
“Then it doesn’t matter which way you go,” said the Cat.
“... so long as I get somewhere,” Alice added as an explanation.
“Oh, you’re sure to do that,” said the Cat, “if you only walk long enough.”

from *Alice’s Adventures in Wonderland* by Lewis Carroll

Contents

Chapter 1

Introduction

1.1 History, motivation and summary

This thesis is about model formulation, analysis and interpretation triggered by four questions arising from biology and medicine, thereby placing the thesis in the field of mathematical biology. For the most part, the thesis in fact deals with various aspects of infectious disease dynamics.

Several notions pervade the many areas of mathematical biology, including mathematical epidemiology. One of these widespread notions is the *basic reproduction ratio*, commonly denoted by $\mathcal{R}_0$. In biological terms, $\mathcal{R}_0$ is defined as the

expected number of offspring a 'typical' individual has in all of its life.

Or, in the context of infectious disease dynamics, $\mathcal{R}_0$ stands for the

expected number of secondary infections caused by a 'typical' primary infection in an otherwise infection free population.

The basic reproduction ratio is a threshold quantity. That is, if, for instance an infectious agent is introduced into an otherwise infection free environment and we describe the spread of the disease by a deterministic model, then according to this model the disease will start spreading if the basic reproduction ratio is larger than 1 and will die out when $\mathcal{R}_0 < 1$. The latter remains true also in a stochastic setting, while the initially introduced infections may die out in the initial stages of an invasion (while they are still rare) even when their basic reproduction ratio exceeds one (see [59] and Chapter 3 of this thesis).

The question of how the 'typical' in the biological definition should be translated into the calculation of $\mathcal{R}_0$ for a given model remained an open question for a

long time. It was only in 1990 that the basic reproduction ratio was, with the work of Diekmann, Heesterbeek and Metz [36], given a precise mathematical definition as being the dominant eigenvalue of the so called *next generation operator* (cf. Chapter 2).

The underlying idea of $\mathcal{R}_0$ and its threshold property, however, were noticed long before that in the work of the British medical doctor Ronald Ross on the spread of malaria. In 1902, Ross received the Nobel prize in medicine for his discovery that malaria was caused by a protozoan parasite and transmitted via Anopheles mosquitos and not via 'bad' air, as is suggested by the name. In his later work [95] Ross showed that there exists a quantity which, being suppressed below one, guarantees disappearance of malaria from the area. He furthermore observed that this quantity depends on the ratio of the mosquito : human density, thus overthrowing the then common belief that the disease wasn't related to the abundance of mosquitos in the area.

Around the same time, the American biologist and actuary Alfred J. Lotka investigated the dynamics of an age structured population [76]. He noted that, if the survival function $\mathcal{F}$ and the fecundity function β are known, then the rate of natural increase of the population r is obtained as a solution of

$$1 = \int_0^\infty e^{-ra}\beta(a)\mathcal{F}(a)da.$$

This equation was, in fact, derived already by Euler in 1760 in his work on determination of annuities [44]. However, Lotka and his coworker Sharpe took it one step further (at least as far as the biological insight is concerned) by observing (i) that the age distribution asymptotically behaves like Ce^{rt} for some constant C and (ii) that

$$\begin{aligned} r < 0 &\iff \int_0^\infty \beta(a)\mathcal{F}(a)da < 1, \\ r > 0 &\iff \int_0^\infty \beta(a)\mathcal{F}(a)da > 1. \end{aligned} \tag{1.1}$$

Lotka and Sharpe did not give a name to the quantity $\int_0^\infty \beta(a)\mathcal{F}(a)da$, nor did they offer any interpretation [99]. Recalling, however, that $\beta(a)$ and $\mathcal{F}(a)$ denote, respectively, the survival and the fecundity function, a moment of reflection brings about that $\int_0^\infty \beta(a)\mathcal{F}(a)da$ equals the basic reproduction ratio $\mathcal{R}_0$.

The parameter r goes by the name of the *intrinsic growth rate* or the *Malthusian parameter*. As opposed to $\mathcal{R}_0$ that measures population growth on a generation basis, the Malthusian parameter measures growth on a real time basis. The set of inequalities in (1.1) gives the known relation between the two.

The so called *invasibility question* (i.e. the question of whether a newly introduced population is able to settle in the existing community) can be answered

in terms of the basic reproduction ratio in several biological settings, not only when it comes to introductions of infectious agents into susceptible populations. In general we can say that when the process of invasion is described by a deterministic model, the newly introduced population meets with success when its basic reproduction ratio exceeds 1 and fails when its $\mathcal{R}_0$ is smaller than 1. What exactly happens when $\mathcal{R}_0$ passes the critical value 1?

It is this question that is central in the beginning of Chapter 2. We focus on populations characterized by finitely many characteristics (such as, for instance, sex, age class, stage in a progression of the disease, etc.) and assume that the dynamics is described as a deterministic, continuous, or discrete time process, either in the form of a parametrized system of differential equations

$$\frac{dx}{dt} = f(x, \mu), \tag{1.2a}$$

or in the form of a parametrized map

$$x \mapsto f(x, \mu), \tag{1.2b}$$

with the vector $x \in \mathbb{R}^k$ describing the population state and the vector $\mu \in \mathbb{R}^p$ of (in principle known) parameter values.

In the context of infectious disease dynamics, the problem was studied already in [105]. In the present study, we formulate the problem in a more general setting, thus making the results applicable to various other biological scenarios, such as, for instance, to study an introduction of a predator that preys on the resident community, introduction of a population that competes for resources with the resident, etc. We argue that population invasion models, regardless of the biology that underlies them, take a specific form, which significantly simplifies the analysis. This particular form of population invasion models implies that the transition of $\mathcal{R}_0$ through the value 1 corresponds to a transcritical bifurcation. From a purely mathematical point of view there is but one type of a transcrical bifurcation [56, 108]. In the biological context, however, there is a natural notion of positivity: the population sizes contained in $x \in \mathbb{R}^k$ must be, in order to be biologically meaningful, nonnegative.

In the case of a transcritical bifurcation, in particular, we must thus distinguish two cases according to whether the biologically meaningful branch of nontrivial equlibria exists to the left of $\mathcal{R}_0 = 1$, i.e. on some interval $(1 - \delta, 1)$ with $\delta > 0$, or to the right of $\mathcal{R}_0 = 1$, i.e. on some interval $(1, 1 + \delta)$ (see Figure 1.1). Accordingly, we call the former a *subcritical* or *backward* bifurcation and the latter a *supercritical* bifurcation and derive a simple formula for the direction of bifurcation from the 'residents only' steady state, which enables us to distinguish between the two qualitatively very different outcomes.

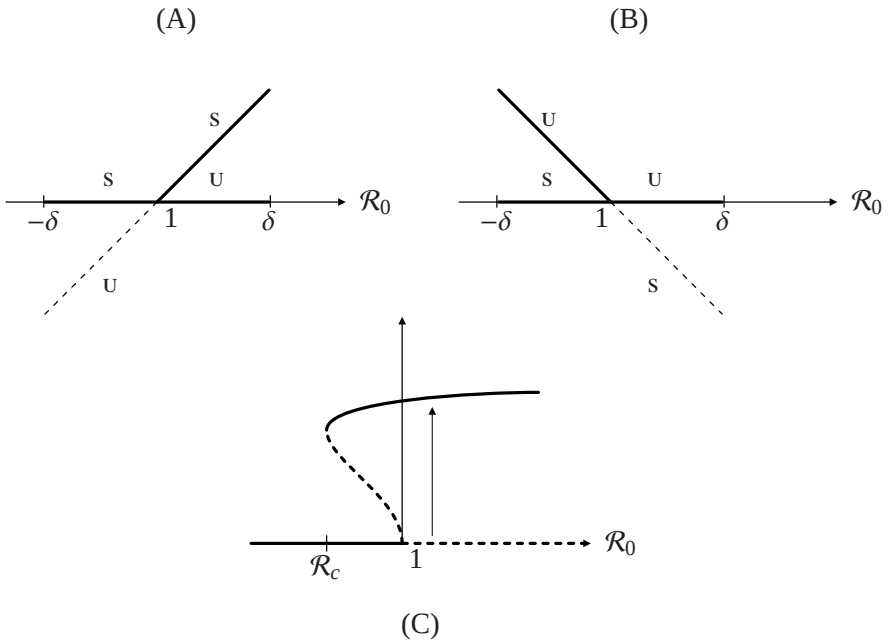

Figure 1.1: (A) Supercritical bifurcation. (B) Subcritical bifurcation. In both figures, the branch of nonnegative steady states is denoted by a solid line. A dashed line represents steady states with negative components. Stability of equilibria is indicated by **s** (for stable) and **u** (for unstable). (C) Catastrophic transition. Here, the unstable part of the branch is depicted by a dashed, stable by a full line.

The question of which of the two bifurcations occurs in a particular biological process, or better, in a particular model we have chosen to describe the process with, is interesting in many cases, but is of particular relevance for control of infectious diseases. Suppose a disease with an $\mathcal{R}_0$ greater than one is introduced into a susceptible population. The number of infected individuals begins to increase. When control measures are implemented that reduce the value of $\mathcal{R}_0$ below one and the system undergoes a supercritical bifurcation at $\mathcal{R}_0 = 1$, then these control measures are sufficient to eradicate the disease from the population. In the malaria model of Ronald Ross, the bifurcation at $\mathcal{R}_0 = 1$ is supercritical, which enabled Ross to arrive at his conclusions.

In the subcritical case, however, it often occurs that the unstable nontrivial branch that exists (locally) for $\mathcal{R}_0 < 1$, bents forward in a saddle node bifurcation, as is illustrated in Figure 1.1 C. In such a case, the epidemic, first of all, reaches (often significantly) higher endemic levels, and secondly, the control measures that push $\mathcal{R}_0$ slightly below 1 are not sufficient to eliminate the disease. Only when $\mathcal{R}_0$ is reduced to a certain level (in Figure 1.1 C denoted by $\mathcal{R}_c$), the endemic

branch becomes unstable, the population of infecteds 'collapses' to zero and the disease vanishes from the population.

◇ ◇ ◇

Chapter 2 deals with dynamics on a so called *ecological time scale*. This means that we consider the parameters in the model (specifying, for instance, reproduction, survival and maturation rates/probabilities) as given and investigate the course of the population state $x(t)$ in time. But how have these parameters evolved themselves? Can we suggest mechanisms that drive their dynamics and understand how they evolve in time?

The exciting and fairly new area of research called Adaptive Dynamics (initiated in the 1990s by Metz and Geritz [49, 51]; see also [29, 30] for an introduction to the subject and a more extensive list of references) investigates this aspect under the assumption of *separation of time scales*, explaining the dynamics of the individuals' characteristics as a process of mutation and selection on a much larger *evolutionary time scale*.

Suppose that we single out one, or more, characteristics that are subject to evolutionary change (these characteristics are in the context of Adaptive Dynamics also called *traits* or *strategies*). Separation of times scales means that we assume that the mutations occur on a time scale that is much larger compared to the time scale of convergence to an ecological attractor. Accordingly, the resident population (that is, a population with a particular trait) is assumed to reside at an ecological attractor (that can either be a steady state or a dynamic attractor). One then derives the basic reproduction ratio of a mutant in the resident population, or, measured on a real time (instead of on a generation) basis, the rate of increase of the rare mutant in the environment set by the residents. The intrinsic growth rate is in the context of Adaptive Dynamics also called the *invasion exponent* and enables us to determine whether the invasion of a mutant trait is successful or not.

In this spirit we investigate in Chapter 3 the evolution of chronic infectious diseases or, more precisely, the evolution of the within-host reproduction rate of a pathogen that causes such a disease.

Parasites reproduce inside a host, but need to, in order to keep the reproduction cycle going, be transmitted to another host. Parasites thus reproduce at two different levels and are subject to natural selection at each of these. But while an increased reproduction within a host may enhance transmission at the host population level, it may also harm the host (by increasing host's mortality) and consequently decrease transmission to other hosts. So, naturally, the following question arises: how are these two tendencies balanced in the course of evolution and to what extent do the two levels of reproduction influence the evolutionary outcome?

Building on the work of Gilchrist and Coombs [52], we relate the between-host transmission to the dynamics inside an infected host. We introduce two models, a model describing the within-host pathogen dynamics and a simple SI model that captures the dynamics at the host population level. Natural selection is assumed to act only on one trait, namely, on the value of the pathogen within-host reproduction rate. We relate the pathogen's between-host characteristics, such as transmissibility and host mortality, to the dynamics within a host and model introductions of mutant traits in an already infected host as superinfections. In the spirit of [84] we introduce a superinfection function $\phi = \phi(p, q)$, giving the probability with which pathogens with trait q (i.e., pathogens that reproduce inside a host at a rate q), upon transmission to a host already infected by trait p, 'take over' the host.

The outcome of evolution depends heavily on the behaviour of the function $q \mapsto \phi(p, q)$ in the neighbourhood of $q = p$. We demonstrate this by considering three cases according to whether the function $q \to \phi(p, q)$ (i) has a jump discontinuity, (ii) is continuous, but not differentiable, or (iii) is differentiable in $q = p$. We find that, in the first case, the within-host selection dominates and, with the particular within-host model used in the study, the continuously stable strategies are found by maximization of the pathogen's within-host reproduction ratio $\mathcal{R}_0^w$. In case (iii), it is the transmission to susceptible hosts that dominates the evolution to the extent that the singular strategies are obtained as the extrema of $\mathcal{R}_0$. In case (ii), both forms of reproduction contribute to the value of a singular trait. We argue that case (ii) is the biologically most relevant case. Indeed, if ϕ is derived from a branching process variant of the submodel for the within-host interaction of pathogens and target cells, the superinfection functions fall under case (ii).

Even though it appears at first sight that in case (i) the evolutionary dynamics is essentially the same as evolution within a single infected host, we demonstrate that this is in fact not the case. In particular, while we find optimization in the course of evolution inside a single infected host, there is no optimization principle when superinfections are taken into account. In fact, we find that there is no optimization principle for the superinfection model, regardless of the smoothness of the superinfection function.

When the singular traits can not be characterized by way of an optimization principle, the skew symmetry of the pairwise invasibility plots is lost and the PIPs make it clear that we may encounter so called protected dimorphisms. Whether such dimorphisms bear any evolutionary significance relies on the (lack of) their convergence. And while converging dimorphisms have no evolutionary significance, diverging dimorphisms are important since they may lead to coexistence of (at least) two distinct traits on the evolutionary time scale. By way of numerical examples, we show that the superinfection model presented in Chapter 3, allows for converging, as well as diverging dimorphisms.

◇ ◇ ◇

A majority of models describing the spread of an infectious agent possess a disease free equilibrium in which there are no infected individuals present in the population. For most cases, this is, of course, natural (imagine, for instance, a model that describes the spread of measles in a community or a model describing the spread of classical swine fever on a pig farm etc.). Hospital settings, however, are an example of a setting in which such an equilibrium in general won't exist. Consequently, the basic reproduction ratio is in such a setting not defined and we can not make use of it to investigate the relative effects of different control measures.

In Chapter 4 we present one model that describes the spread of pathogens in intensive care units (ICUs) and does (in a generic case) not exhibit a disease free steady state. Our study aims to determine the relative effects of barrier precautions (use of gloves, gowns, etc.) and antibiotic prophylaxis on the prevalence of colonization in ICUs and can be applied to several pathogens commonly found in ICUs, such as *Pseudomonas Aeruginosa*, enteric *Gram-negative* bacteria, MRSA and *enterococci*.

It is known [9] that nosocomial bacterial infections in critically ill patients are generally preceded by asymptomatic carriage (i.e. colonization) at one, or even several, body sites such as the skin, the gastro-intestinal and the respiratory tract. Different routes of transmission between the colonized sites create a complex epidemiology, which is additionally complicated by the smallness of the patient population size and the rapid patient turnover, characteristic for intensive care units. Naturally occuring large fluctuations in the prevalence of colonization make it very difficult to determine the efficacy of control measures that aim to reduce the prevalence of bacteria in ICUs.

Theoretical models can often sharpen our intuition through carefully designed thought experiments. In this spirit, we introduce and investigate two models that incorporate the fact that colonization may occur at multiple body sites. Due to the small size of ICU population, it seems natural to model the dynamics in the ICU as a stochastic process. However, exactly this randomness makes it very difficult to draw conclusions about the efficacy of control strategies. We thus formulate a Markov chain model and, for comparison, also a deterministic model and investigate the effects of barrier precautions (improved hygiene, use of gloves and gowns etc.) and of administration of non-absorbable antibiotics on the prevalence of colonization in ICUs. We find that the effect of the controversial, though widely used, antibiotic prophylaxis can only be substantial if the patient-to-patient transmission has already been reduced to a 'subcritical' level by barrier precautions. Taking into account that the very use of antibiotics may increase the selection for resistant strains and may thereby only add to the ever increasing problem of an-

tibiotic resistance, our findings represent a firm theoretical argument against the routine use of topical antimicrobial prophylaxis for infection control.

◇ ◇ ◇

The contents of Chapters 2, 3 and 4 are completed manuscripts that have already been either published or accepted for publication (see Section 1.3). Chapter 5, on the other hand, is work in progress.

The aim of Chapter 5 is to contribute to understanding of the dynamics of enterotoxigenic *Escherischia coli* (ETEC) in piglets. ETEC is one of the causative agents of the post-weaning diarrhoea (PWD), a disease that occurs commonly in piglets during the first two weeks after weaning. It is known [48] that ETEC can adhere to the microvilli of the small intestine enterocytes. When attached, bacteria produce enterotoxins that act locally on the enterocytes, resulting in increased secretion of fluid and $NaHCO_3$, which can lead to diarrhoea, dehydration, acidosis or even death.

To this day, several experimental studies have been conducted in order to obtain some insight on how various factors (such as weaning age, litter size, etc.) contribute to the occurrence and the clinical symptoms of PWD. The recent field work performed by P. Geenen [48] furthermore provided some data on the shedding patterns of piglets and transmission of ETEC between piglets. However, the real understanding of the dynamics of ETEC in piglets is still missing. There are many questions to be answered, both regarding the dynamics of ETEC within one host and at the (piglet) population level. Since the latter requires information about the former, we begin our investigation of this disease by setting up and analysing a model that describes the microbial dynamics in the intestine of a single infected piglet.

Even though the long term objective is to study more complicated models of within-host dynamics, we begin with a very simple linear model. We describe the intestine as a cylindrical tube and consider two types of bacteria in the intestine: the ones moving downstream and the ones attached to the wall of the intestine. Bacteria can at any point in time change their state and we assume that bacteria in both states are able to reproduce.

We focus in particular on the invasibility question for two scenarios. To begin with, we determine the Malthusian parameter for the so called single infection case, meaning that a piglet is infected by a single dose of ETEC at time $t = 0$ and no bacteria enter its intestine after $t = 0$. Since it is likely that the infected piglet comes into contact with its faeces that contain the bacteria, we furthermore investigate the conditions for persistence when the piglet is reinfected with a fraction of the shed bacteria.

In the spirit of Lotka's demographical model, we furthermore investigate whether there exists a stable bacterial distribution, both for the single infection

and the reinfection case. While in the single infection case there is no convergence to a stable distribution, we find that such a stable distribution exists in the reinfection case, provided that bacteria in the intestine are, in a sense, well mixed (we postpone the precise explanation of the term 'well mixed' until Chapter 5).

◇ ◇ ◇

The biological motivation in the four problems and the tools used for their analysis are quite diverse. We thus present the results of their investigation in the form of four self-contained chapters. In each of these chapters, we shall along the way also introduce (hopefully) all the necessary notions and known results needed for understanding of our work. But before we begin with the main part of the thesis, some words on what the future may hold.

1.2 Outlook

The results of the thesis provide a basis and offer several possibilities for further research. For instance, results of Chapter 2 apply when the resident populations reside in a steady state, but the same questions can of course be posed when the resident community resides in a dynamic attractor.

The within-host modeling performed in Chapter 5 can provide useful information for studying the dynamics of ETEC at the piglet population level. For instance, it seems reasonable to assume that piglet's infectiousness is related (or, in the simplest case, proportional) to the amount of shed bacteria, something for which we derive an explicit expression in Chapter 5.

Some of the prospects for further work are also included in the concluding parts of the chapters. Here we wish to discuss in more detail how the models and tools of Chapter 3 can provide a basis for future study of the evolution of infectious diseases.

In Chapter 3 we investigate the evolutionary dynamics of a chronic infectious disease. Along with the work of Gilchrist and Coombs [52] and Alizon [1], the model of Chapter 3 represents one of the first models in which the evolution of an infectious disease is studied while explicitly incorporating a sketch of within-host pathogen dynamics into a model for the spread of the disease at the host population level.

In our work, we focus on chronic diseases that are at the host population level described by an SI model. Hence, once a host is infected, it retains the infection until death.

The within-host dynamics is described by the following set of differential equations,

$$\begin{aligned}\frac{dT}{dt} &= \lambda - kVT - dT\\ \frac{dT^*}{dt} &= kVT - (\mu(p) + d)T^*\\ \frac{dV}{dt} &= pT^* - kVT - cV.\end{aligned} \tag{1.3}$$

Here, T, T^* and V denote, respectively, the number of uninfected target cells, infected target cells and free pathogens. The parameters λ and d denote, respectively, the growth and the per capita death rate of target cells in the absence of the infectious agent. The mass action term kVT models how infection of target cells occurs: it states that the rate at which pathogens find uninfected target cells, bind to their surface and/or enter the target cell, is proportional (with proportionality constant k) to the abundance of both the uninfected target cells as well as the free pathogens. Infected target cells produce new pathogens at a rate p and experience an increased (when compared to uninfected cells) death rate $\mu(p) + d$. Free pathogens die at a constant per capita rate c.

In order not to isolate the within-host evolution from the one occuring at the host population level, we must allow individuals to be reinfected (indeed, if variations in traits created by mutation inside one host can trigger selection, then this should also be possible for mutants introduced from the outside). In Chapter 3 we choose to model these reinfections as superinfections. However, at the cell population level, we do not allow for reinfections. One possibility for future work is thus to expand (1.3) to allow for the fact that target cells can be infected by different pathogen traits.

But perhaps more interesting, and certainly more challenging, is the question of incorporating some description of the immune system. In an SI setting that we consider in Chapter 3, the host or, rather, its immune system, fails to tackle the infection. The effect of immune response in (1.3), be that antibody or cellular immune response, could perhaps be interpreted as being contained in the increase of mortality of infected target cells, but such a response is not sufficient to eliminate the infection in a host.

The immune system is a very complicated system and its response to pathogens is in many cases poorly understood. To incorporate an explicit account of such a system, one should thus first start with simple suggestions of its working and increase the complexity only once the simplest models have been understood.

By introducing a new variable representing the abundance of immune effectors, say $I(t)$, the simplest equation for the dynamics of the immune system is perhaps the equation

$$\frac{dI}{dt} = \gamma - d_I I,$$

where γ is assumed to be strictly positive when the infectious agent is present and zero otherwise. In the absence of an infection, the number of immune effectors exponentially decreases to zero at a rate d_I.

We furthermore rewrite the equation for the dynamics of infected target cells as

$$\frac{dT^*}{dt} = kVT - (\mu(p) + d)T^* - \kappa(p)T^*I,$$

to describe the situation where the effectors kill the infected target cells at a rate $\kappa(p)$.

It can be seen that, with such an immune response, the infection will persist if the pathogen's within-host reproduction ratio in the presence of the immune response is larger than one (but the steady state level of free pathogens as well as infected target cells will be lower than in the case of no immune response) and will be cleared when the within-host basic reproduction ratio in the presence of the immune response is smaller than one.

Alternatively, one could suggest

$$\frac{dI}{dt} = \gamma T^* - d_I I,$$

to describe the fact that the growth of immune effectors is proportional to the abundance of infected target cells (but not their own abundance) or

$$\frac{dI}{dt} = \gamma T^* I - d_I I,$$

to incorporate also the dependence of the growth on the abundance of the immune effectors. There are, of course, several other options to describe the interactions of immune effectors with infected target cells.

In Chapter 3, the analysis is significantly simplified by double separation of times scales: (i) to make use of the tools of Adaptive Dynamics we first assume that mutations are rare on the ecological time scale and (ii) we assume that the within-host processes are fast compared to processes at the host population level. Thus

$$\underbrace{\text{within-host time scale} \ll \text{between-host time scale}}_{\text{ecological time scale}} \ll \underbrace{\text{time scale of mutations.}}_{\text{evolutionary time scale}}$$

When the immune response is explicitly included into the analysis, but is not strong enough to eliminate the infection, then the analysis, and perhaps also the outcome of evolution, should not be that much different from the analysis of Chapter 3.

However, if the immune system can clear the infection, the within-host processes are much more intertwined with the processes at the host population level (note also that the model describing the dynamics at the host population has to be altered), which makes the analysis much more challenging.

1.3 Publications

The contents of this thesis relate to the following publications:

(1) Chapter 2 is adapted with minor changes from
B. BOLDIN.: *Introducing a population into a steady community: the critical case, the center manifold and the direction of bifurcation,* SIAM J. Appl. Math, Vol. 66 (2006), pp. 1424 - 1453,

(2) Chapter 3 is adapted with minor changes from
B. BOLDIN, O. DIEKMANN.: *Superinfections can induce evolutionarily stable coexistence of pathogens*. To appear in Journal of Mathematical Biology,

(3) Chapter 4 is adapted with minor changes from
B. BOLDIN, M. J. M. BONTEN, O. DIEKMANN.: *Relative effects of barrier precautions and topical antibiotics on nosocomial bacterial transmission: results of multi-compartment models,* Bulletin of Mathematical Biology (2007). Available in electronic form: DOI:10.1007/s11538-007-9205-1,

(4) Chapter 5 is an unpublished manuscript,
B. BOLDIN.: *Persistence and spread of enterotoxigenic Escherichia coli in the gastro-intestinal tract of piglets.*

Chapter 2

Introducing a population into a steady community: the critical case, the center manifold and the direction of bifurcation

2.1 Motivation and aims

Let us begin with an (overly simple) example to motivate the ideas and the goals of this chapter.

Consider a population (of, say, plants, cells, rabbits, etc.) that grows logistically with intrinsic growth rate r and carrying capacity K. If we denote by $x(t)$ the size of this population, then $x(t)$ obeys the differential equation

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{K}\right).$$

Imagine now that a second population is introduced, one which preys on the resident community (for instance, herbivorous animals, viruses, foxes). Let us denote by $y(t)$ the size of the predator population and assume that (i) in the absence of prey, predators die at per capita rate d (prey is thus essential for predators' survival) and (ii) when both populations are present, prey is captured at per capita rate proportional to the size of the predator population (with proportionality constant c) and predators are born according to mass action principle with proportionality constant b.

These assumptions give rise to the following variation of the classical Lotka-Volterra model [42]

$$\begin{aligned} \frac{dx}{dt} &= rx\left(1-\frac{x}{K}\right) - cxy \\ \frac{dy}{dt} &= bxy - dy. \end{aligned} \tag{2.1}$$

One of the simplest questions one might ask is the following.

Suppose that a new population (predators) is introduced into the resident (prey) community. Under which conditions will the newly introduced population be able to grow and when will the dynamics lead to its extinction?

The question of invasibility is one of the basic questions of mathematical biology and the literature that deals with it is vast and diverse. We could roughly group the biological settings into the following categories.

I. In *ecology* one is studying an introduction of (i) a population of predators that forages on the resident prey community [19, 26, 42, 62, 65, 77, 80, 87, 103], or (ii) a population that competes for resources with the resident community (for examples of competition models see [19, 42, 62, 80, 100]).

   While it is common that the newly introduced population is also a new species, i.e., one that is not present in the resident community, there are also examples in which one is interested in the ability of what we shall call a *reproductively isolated subpopulation* of one species to be able to settle among individuals of another subpopulation of the same species. We shall define the precise meaning of the term reproductive isolation later on. The reader may at this point have in mind, for instance, studying interactions (say, competition for shared resources) among different year classes of semelparous species [12, 18, 20, 21] or among different morphs in size structured populations [13, 14].

II. In a branch of the theory of evolution called *adaptive dynamics* one is investigating the ability of a rare mutant phenotype to invade the environment set by the resident community (see [29, 30, 49, 61] and the references therein).

III. *Epidemiology of infectious diseases* is concerned with introductions of infectious pathogens into susceptible populations [35, 40, 41, 54, 103, 105].

Let us return briefly to (2.1) and assume that, prior to an invasion, the size of the prey population takes the (nontrivial) equilibrium value, $x = K$. Since in the very first stages of an invasion the depletion of they prey due to predation can be ignored (at least in this deterministic description, that is), we can rewrite the equation for the predator dynamics in the initial stages of an invasion as

$$\dot{y} = (bK - d)y.$$

Hence, the predators will increase in numbers if the per capita growth rate $bK - d$ is positive, and will die out when $bK - d < 0$. We can alternatively formulate the answer in terms of the predators' *basic reproduction ratio*, $\mathcal{R}_0$. The predators' basic reproduction ratio is defined as the expected number of offspring born to one predator that is introduced into a steady prey population. In this case, $\mathcal{R}_0$ can easily be determined. Namely, the predators live on average $\frac{1}{d}$ units of time, in which they produce bK offspring and so $\mathcal{R}_0 = \frac{bK}{d}$. Hence, the predators grow if $\mathcal{R}_0 > 1$ and go extinct if $\mathcal{R}_0 < 1$.

This result holds, in fact, in a much more general setting: when the resident community is at a stable equilibrium and we describe the process of invasion by a deterministic model we can, regardless of the biological background, answer the invasibility question in terms of the basic reproduction ratio of the invading population, $\mathcal{R}_0$, as follows: if $\mathcal{R}_0 < 1$ the invading population will go extinct, while the invader will settle in the community if $\mathcal{R}_0 > 1$ [34, 35].

What happens when $\mathcal{R}_0$ passes the critical value one? The answer can be formulated mathematically or biologically. In mathematical terms one says that a *transcritical bifurcation* of a steady state and an *exchange of stability* take place [16, 17, 56, 108]. From a strictly mathematical point of view there is but one generic type of transcritical bifurcation. But when it comes to seeing the results from a biologist's point of view one must realize that a steady state is only meaningful when all its components are nonnegative, in particular those corresponding to the invading population. In many models the latter requirement is only fulfilled when $\mathcal{R}_0 > 1$. The bifurcation is then called *supercritical* or *forward* or, also, *soft* or *smooth* since the size of the invading population remains small when $\mathcal{R}_0 - 1$ is positive but small. In other models, however, the positivity requirement is fulfilled only for $\mathcal{R}_0 < 1$ and one then speaks of a *subcritical* or *backward* bifurcation.

While in the case of a supercritical bifurcation the invasion fails when $\mathcal{R}_0$ of the invading population falls below one, the invader can meet with success even if $\mathcal{R}_0 < 1$ (when introduced in sufficiently large quantities) when the bifurcation is backward. Moreover, even when the invader is introduced in small quantities a small perturbation of $\mathcal{R}_0$ to a value greater than one can in a subcritical case lead to a rather large invader population size. This phenomenon is sometimes called *catastrophic transition* (see Figure 2.2).

Clearly then, it is important to be able to tell which of the two cases applies in any given situation and the purpose of this chapter is to provide the reader with a simple criterion to distinguish between the two scenarios.

The invasibility question is, of course, meaningful also when the resident community resides in a dynamic attractor. This situation is, however, outside the scope of this chapter: we shall assume throughout that, prior to an invasion, the resident community resides in a stable equilibrium. Moreover, we shall consider only

communities whose members differ in a finite number of characteristics (such as, for instance, sex, age class, etc.). These characteristics are in the context of population models often called *i- states* [34, 35], with *i* standing for individual. Although a finite number of *i-* states is, of course, a restriction, we believe for two reasons that the results presented in this chapter apply to a broad variety of models. Firstly, a lot of population models in fact deal with finite dimensional state spaces and secondly, even if the state space is infinite dimensional to begin with, one often makes approximations to make the problem computationaly manageable.

Ideally, *i-* states should capture precisely the features that are relevant for the description of the process one is studying and are hence to be considered for each problem separately. In a general setting we assume that the community is divided into $m + n$ subpopulations, of which m subpopulations constitute the invading population and the remaining n make up the resident community. We denote by

$$\mathcal{Y} = \{(y_1, ..., y_m)\ ;\ y_j \geq 0 \ \text{ for } j = 1, ..., m\} = \mathbb{R}^m_+$$

the population state space (p- state space) of the invading population (i.e., for each $j \in \{1, ..., m\}$ we denote by y_j the number (fraction/density) of individuals in the j- th subpopulation) and by

$$\mathcal{Z} = \{(z_1, ..., z_n)\ ;\ z_j \geq 0 \ \text{ for } j = 1, ..., n\} = \mathbb{R}^n_+$$

the community state space of the resident community. The c- state space of the joint community will be written as $\mathcal{Y} \times \mathcal{Z}$.

Now let $(y(t), z(t))$ denote the community state at time t, where time is measured from some conveniently chosen point on. The dynamics of $(y(t), z(t))$ in time often depends not only on the present community state, but also on a number of parameters, such as per capita death rates, birth rates, etc., and, quite commonly, population models involve more than one parameter.

The aim of this chapter is to study the ability of a newly introduced population to invade the existing community for the case when its basic reproduction ratio is near one and to derive a formula for the direction of bifurcation from a 'residents only' steady state. We shall therefore later on concentrate on one distinguished parameter, which we call the bifurcation parameter (i.e., we shall keep all parameters but one fixed).

With this in mind we already at this point include but one (real) parameter μ and assume that the process we study is either a continuous time process described by a parametrized system of ordinary differential equations

$$\begin{aligned} \dot{y} &= g(y, z, \mu) \\ \dot{z} &= h(y, z, \mu) \qquad y \in \mathbb{R}^m, z \in \mathbb{R}^n, \mu \in \mathbb{R}, \end{aligned} \tag{2.2a}$$

or a discrete time process described by a parametrized map

$$\begin{aligned} y &\mapsto g(y,z,\mu) \\ z &\mapsto h(y,z,\mu) \qquad\qquad y \in \mathbb{R}^m, z \in \mathbb{R}^n, \mu \in \mathbb{R}. \end{aligned} \tag{2.2b}$$

If we consider a steady state of (2.2a) (or (2.2b)) in which the invading population is not present (these steady states lie on the boundary of the c- state space) and study the effect of perturbations corresponding to an introduction (in small quantities) of the missing population, we find that such an equilibrium is locally asymptotically stable when $\mathcal{R}_0$ of the invading population is below one, and unstable when $\mathcal{R}_0$ exceeds one. Moreover, stability can in these two cases be inferred from the linearization of (2.2) around the steady state.

The Perron - Frobenius theory of nonnegative matrices [7, 72], which applies to problems in population dynamics, leads us to the observation that the critical case, i.e., the case when $\mathcal{R}_0 = 1$, corresponds to the situation when (i) the linearization of (2.2a) around the steady state yields a zero eigenvalue and (ii) the linearization of (2.2b) around the steady state yields an eigenvalue one. In other words, when $\mathcal{R}_0 = 1$ we are dealing with nonhyperbolic steady states and it is well known [56, 108] that the stability of nonhyperbolic equilibria cannot be determined by linearization alone.

Several papers ([26, 33, 40, 41, 54, 68, 77, 87, 105] to name but a few) deal with this situation in the context of population models, most of them (with the exception of [33]) treating special cases or restricting to models describing the spread of infectious diseases.

In this chapter we study the critical case for general finite dimensional population models. We argue that an introduction of either one new species or a reproductively isolated subpopulation of one of the existing species yields a property of (2.2) that significantly simplifies the center manifold analysis. More precisely, (2.2a) will be shown to be of the form

$$\begin{aligned} \dot{y} &= G(y,z,\mu)y \\ \dot{z} &= h(y,z,\mu) \qquad\qquad y \in \mathbb{R}^m, z \in \mathbb{R}^n, \mu \in \mathbb{R}. \end{aligned} \tag{2.3}$$

An analogous decomposition can be obtained for parametrized maps in (2.2b).

This will lead us to the observation that an introduction of a population whose basic reproduction ratio is close to one corresponds to a transcritical bifurcation of a steady state of (2.2) in which only the resident populations are present. In order to obtain the direction of bifurcation from such a steady state only the first derivatives of G and h are needed. This reduction of the order of the derivatives needed (in general, second order derivatives are needed) is, of course, most useful when one is dealing with large systems.

We will also see that among those bifurcation parameters for which

$$\begin{cases} \mu < 0 \iff \mathcal{R}_0 < 1 \\ \mu = 0 \iff \mathcal{R}_0 = 1 \end{cases}$$

holds on some neighborhood of $\mu = 0$ and the crossing of the point $\mathcal{R}_0 = 1$ occurs at a nonzero 'speed', we obtain the same direction of bifurcation.

Moreover, we will show how G in (2.3) can be obtained by only considering the basic modeling ingredients, such as birth, growth and survival rates - an approach that might be of interest to more biologically inclined readers.

The chapter is structured as follows. In Section 2.2 we study continuous time population invasion models described by (2.3). Section 2.3 is devoted to justifying the use of this particular form of models. We argue that this form is characteristic of population invasion models. It appears in all biological scenarios mentioned in the beginning of this Introduction and hence allows us to make a uniform study of the invasibility question for ecological, adaptive dynamics and disease transmission models. We also show how G is obtained from basic modeling ingredients. Population models in discrete time are the theme of Section 2.4. Section 2.5 provides some interpretation of the assumptions made in previous sections and draws attention to the link between continuous and discrete time population models. In Section 2.6 we give some examples to illustrate the theory of the preceding sections. Throughout the text, the reader will also encounter boxes, in which we collect some basic definitions and results regarding physiologically structured population models. We close this first chapter with an Appendix, in which we put the notions of a *population* and of a *reproductively isolated subpopulation* in a more mathematical setting.

2.2 Population invasion models in continuous time

We begin our study of continuous time population models by recalling the decomposition of the community state space

$$\mathcal{Y} \times \mathcal{Z} = \mathbb{R}^m_+ \times \mathbb{R}^n_+,$$

where $\mathcal{Y} = \mathbb{R}^m_+$ denotes the state space of the invading population and $\mathcal{Z} = \mathbb{R}^n_+$ the state space of the resident community. The processes we study in this section are continuous time processes described by

$$\begin{aligned} \dot{y} &= G(y, z, \mu)y \\ \dot{z} &= h(y, z, \mu) \qquad\qquad y \in \mathbb{R}^m, z \in \mathbb{R}^n, \mu \in \mathbb{R}, \end{aligned} \tag{2.4}$$

where we shall furthermore assume that $G \in M_{m\times m}(C^1(\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}, \mathbb{R}))$ and $h \in C^1(\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}, \mathbb{R}^n)$.

The form (2.4) is characteristic of continuous time population invasion models and one with which experienced modelers may already be familiar. We shall at this point take for granted that the systems indeed take this form, but we will return to this subject in the next section and explain, in both mathematical and biological terms, why and how this form is obtained.

By writing

$$x = \begin{bmatrix} y \\ z \end{bmatrix} \qquad f = \begin{bmatrix} Gy \\ h \end{bmatrix} \tag{2.5}$$

we shall write (2.4) as $\dot{x} = f(x, \mu)$ and also use (2.2a) whenever this notation will be more convenient.

Consider now an equilibrium of (2.4) of the form $e = (0, z_0)$ for some $z_0 \in \mathcal{Z}$, i.e. a steady state in which only the resident populations are present. In general, (2.4) can have more than one steady state of this form and these steady states may also depend on μ. So let us note this dependence on μ explicitly: $e(\mu) = (0, z_0(\mu))$ with $z_0(\mu) \in \mathcal{Z}$.

To study the linearized stability of $e(\mu)$ we write

$$Df((0, z_0(\mu), \mu)) = \begin{bmatrix} G(e(\mu), \mu) & 0 \\ h_y(e(\mu), \mu) & h_z(e(\mu), \mu) \end{bmatrix},$$

where

$$h_y(y, z, \mu) = \frac{\partial h(y, z, \mu)}{\partial y} \quad \text{and} \quad h_z(y, z, \mu) = \frac{\partial h(y, z, \mu)}{\partial z}. \tag{2.6}$$

The spectrum of $Df(e(\mu), \mu)$ can thus be written as

$$\sigma(Df(e(\mu), \mu)) = \sigma(G(e(\mu), \mu)) \cup \sigma(h_z(e(\mu), \mu)).$$

The next assumption we shall make is that the equilibrium $e(\mu) = (0, z_0(\mu))$ is internally asymptotically stable, i.e., it is asymptotically stable under perturbations within the invariant subspace $\{0\}^m \times \mathcal{Z}$. In other words, as long as no new populations are introduced, the steady state of the resident community, $z_0(\mu)$, is locally asymptotically stable. We shall make the slightly stronger assumption that the stability can be inferred from the linearization. In mathematical terms this means that we assume

$\mathbf{A_1}$: if $\lambda \in \sigma(h_z(e(\mu), \mu))$ then $Re(\lambda) < 0$.

The spectrum of $G(e(\mu), \mu)$ hence completely determines the linearized stability of the steady state $e(\mu)$.

For the existence and uniqueness assertions that follow we actually only need 'internal hyperbolicity', i.e. that $Re(\lambda) \neq 0$ for any $\lambda \in \sigma(h_z(e(\mu), \mu))$. Assumption A_1 will allow us to make more detailed stability assertions, which are known as the *Principle of the Exchange of Stability* [16, 17, 56].

Now, we would like to know whether the newly introduced population is able to settle in the community. As mentioned in the Introduction, the answer is *no* when the basic reproduction ratio of the invading population is below one and *yes* when $\mathcal{R}_0$ of the invading population exceeds one.

The basic reproduction ratio is by definition the spectral radius of the next generation matrix (cf. Box 2.2). All the modeling ingredients needed to write down the next generation matrix in the context of the model given by (2.4) are contained in G (remember that $\mathcal{R}_0$ of the invading population is the one we need) and it is known [35, 105] that $\mathcal{R}_0$ of the newly introduced population relates to the spectral bound of $G(e(\mu), \mu)$ in the following way,

$$\begin{cases} s(G(e(\mu),\mu)) < 0 & \Longleftrightarrow \ \mathcal{R}_0 < 1 \\ s(G(e(\mu),\mu)) = 0 & \Longleftrightarrow \ \mathcal{R}_0 = 1 \end{cases}$$

where the spectral bound, $s(.)$, is given by

$$s(A) = \max\{Re(\lambda); \lambda \in \sigma(A)\}.$$

Since the next generation matrix is a nonnegative matrix we can apply the Perron - Frobenius theory [7] to conclude that $\mathcal{R}_0$ is an eigenvalue with a corresponding nonnegative eigenvector. The dominant eigenvalue is often called the *transversal eigenvalue* and if it exceeds 1 we say that the newly introduced population is able to *invade successfully*. If $\mathcal{R}_0$ is below one, the invasion of the newly introduced population is doomed to fail.

The interesting situation to consider is hence the situation when the parameter μ is such that $s(G(e(\mu), \mu)) = 0$, the case where linearization around the steady state does not yet answer the question of invasibility.

Now, in many models the computation of the basic reproduction ratio $\mathcal{R}_0$ and certainly the spectral bound $s(G(e(\mu), \mu))$ yields complicated functions of parameters that we may not be able to express explicitly. We therefore choose a bifurcation parameter μ with the following properties,

$$\mathbf{A_2}: \quad \begin{cases} \mu < 0 \iff s(G(e(\mu),\mu)) < 0 \iff \mathcal{R}_0 < 1 \\ \mu = 0 \iff s(G(e(\mu),\mu)) = 0 \iff \mathcal{R}_0 = 1. \end{cases}$$

The results that follow are based on local information only. It therefore suffices that A_2 holds on some neighbourhood of $\mu = 0$.

Assumption A_2 means that the function $\mu \mapsto s(G(e(\mu), \mu))$ crosses the origin. We shall furthermore assume that this crossing occurs at a nonzero 'speed',

$\mathbf{A_3}: \frac{d}{d\mu} s(G(e(\mu),\mu))\Big|_{\mu=0} > 0.$

We now denote by e an equilibrium that corresponds to $\mathcal{R}_0 = 1$, i.e. $e = e(0)$,

define $e' = e'(0)$ and also shorten the notation by defining

$$H_y = h_y(e,0), \quad H_z = h_z(e,0), \quad G_0 = G(e,0). \tag{2.7}$$

Denoting by E^c the center subspace of G_0 we shall furthermore assume that

$\mathbf{A_4}$: $\dim E^c = 1$.

We have already given the interpretation behind the first three assumptions. We shall return to this last assumption in Section 2.5 and explain in more detail which biological requirements are sufficient in order for A_4 to hold. Let us only remark that, in systems that arise from modeling population dynamics, the matrix G_0 will be a matrix with nonnegative off-diagonal entries and hence the Perron - Frobenius theory guarantees that A_4 is satisfied when G_0 is irreducible.

Before stating the main result we make the following observation that will be useful later on.

Lemma 2.1. *Let $\mu \mapsto G(e(\mu),\mu) \in C^1(\mathbb{R}, \mathbb{R}^{m\times m})$, assume A_2 and A_4 and let w and v denote respectively the left and the right eigenvector of G_0 corresponding to eigenvalue zero, normalized, so that $v \cdot w = 1$. Then*

$$\tfrac{d}{d\mu} s(G(e(\mu),\mu))\Big|_{\mu=0} = w \cdot \Big(D_x G(e,0)e' + D_\mu G(e,0)\Big)v. \tag{2.8}$$

Proof. According to the Implicit Function Theorem there exists a neighborhood of $\mu = 0$, say U, on which a branch of eigenvalues of $G(e(\mu),\mu)$ is defined. That is,

$$G(e(\mu),\mu)v(\mu) = \lambda(\mu)v(\mu) \tag{2.9}$$

for $\mu \in U$, and, since $\mu \mapsto G(e(\mu),\mu) \in C^1(\mathbb{R}, \mathbb{R}^{m\times m})$ we have $\mu \mapsto \lambda(\mu) \in C^1(U,\mathbb{R})$. Moreover, $\mu \mapsto v(\mu) \in C^1(U,\mathbb{R}^m)$. Differentiation of (2.9) with respect to μ yields

$$\Big(\frac{\partial G}{\partial e}e'(\mu) + \frac{\partial G}{\partial \mu}\Big)v(\mu) + Gv'(\mu) = \lambda'(\mu)v(\mu) + \lambda(\mu)v'(\mu). \tag{2.10}$$

Since zero is also the spectral bound of G_0 and the spectral bound $s(G(e(\mu),\mu))$ is a continuous function of μ we have that $\lambda(\mu) = s(G(e(\mu),\mu))$ in some neighborhood of $\mu = 0$. By taking $\mu = 0$ in (2.10) and taking into account that $\lambda(0) = 0$ we obtain

$$\tfrac{d}{d\mu} s(G(e(\mu),\mu))\Big|_{\mu=0} v = \Big(D_x G(e,0)e' + D_\mu G(e,0)\Big)v + G_0 v'(\mu)\Big|_{\mu=0},$$

which brings us, after premultiplication by w on both sides, to (2.8). □

We can now prove the following

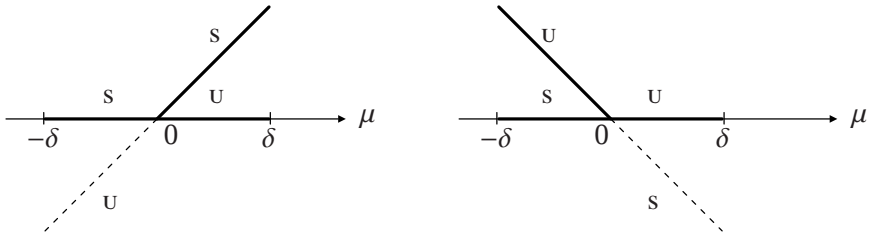

Figure 2.1: Left: supercritical bifurcation. Right: subcritical bifurcation. The branch of nonnegative steady states is denoted by a solid line. A dashed line represents steady states with negative components. Stability of equilibria is indicated by **s** (for stable) and **u** (for unstable).

Theorem 2.1. *Consider a population model described by* (2.4) *and let* $e(\mu) = (0, z_0(\mu))$ *be a steady state of* (2.4). *Assume that* A_1, A_2, A_3 *and* A_4 *hold. Furthermore assume that* $\mu \mapsto e(\mu) \in C^1(\mathbb{R}, \mathbb{R}^{n+m})$ *and denote by* e *the steady state that corresponds to* $\mathcal{R}_0 = 1$, *i.e.* $e = e(0)$ *and by* $e' = e'(0)$. *Let* G_0, H_y *and* H_z *be as in* (2.7) *and let* w *and* v *denote, respectively, the left and the right eigenvector of* G_0 *corresponding to eigenvalue zero, normalized, so that* $v \cdot w = 1$. *Let*

$$\begin{aligned} M = &\sum_{i,j,k=1,\dots m} w_i \Big(\frac{\partial G_{ij}(e,0)}{\partial y_k} + \frac{\partial G_{ik}(e,0)}{\partial y_j} \Big) v_j v_k - \\ &- 2 \sum_{\substack{i,j=1\dots m \\ k=1\dots,n}} w_i \frac{\partial G_{ij}(e,0)}{\partial z_k} v_j (H_z^{-1} H_y v)_k. \end{aligned} \tag{2.11}$$

There exists a $\delta > 0$ *such that*

(i) if $M < 0$ *there is a branch* $\mu \mapsto (y(\mu), z(\mu))$, *defined for* $\mu \in (0, \delta)$, *of positive, locally asymptotically stable steady states of* (2.4).

(ii) if $M > 0$ *there is a branch* $\mu \mapsto (y(\mu), z(\mu))$, *defined for* $\mu \in (-\delta, 0)$, *of positive, unstable steady states of* (2.4).

In other words, there exists, in a neighborhood of $\mu = 0$, *a branch of nontrivial, positive (and hence biologically meaningful) steady states of* (2.4) *and* M *tells us about its initial slope. The former case, case* (*i*), *is often referred to as a* **supercritical** *bifurcation and the latter, case* (*ii*), *as a* **subcritical** *or* **backward** *bifurcation (cf. Figure 2.1).*

At this point the following remark regarding the terminology is in place.

Remark 2.1. As already mentioned in the Introduction, the resulting bifurcations are the so called *transcritical bifurcations*. They correspond to an intersection of two branches of equilibria, the trivial and the nontrivial one, at $\mu = 0$, where the branches exchange stability. In contrast with the purely mathematical point of view where these two transitions are qualitatively the same, we need to distinguish between the two in the biological context since in that case only the nonnegative equilibria are of any relevance.

Remark 2.2. Note that only the first order derivatives of G and h are needed to determine the direction of bifurcation from e. Moreover, the expression M for the direction of bifurcation is independent of the bifurcation parameter but for the restrictions A_2 and A_3. In other words, provided that A_2 and A_3 are satisfied, we obtain the same direction of bifurcation for any choice of the bifurcation parameter.

The principle of the exchange of stability guarantees that the biologically meaningful, nontrivial bifurcating branch consists of stable equilibria in the supercritical case and of unstable equilibria in the subcritical case. The stable manifold of an unstable equilibrium then serves as a separatrix between the domains of attraction of the 'residents only' steady state and some other attractor (frequently the same branch bent forwards in a saddle node bifurcation).

Imagine now that an invader is introduced in small quantities into the resident community. In both, the supercritical and the subcritical case, the invasion will fail if the basic reproduction ratio of the invader falls below one. In the supercritical case, the invader will be successful when its basic reproduction ratio exceeds one, but its population size will be small when $\mathcal{R}_0 - 1$ is small. Because of this smooth transition one sometimes calls this bifurcation *soft* or *smooth*. In the subcritical case, on the other hand, a small introduction of the invading population with $\mathcal{R}_0 - 1$ small but positive leads to a large invader population size. Accordingly, one also calls this bifurcation *hard* or *catastrophic*. Moreover, the invader can meet with success, despite $R_0 < 1$, if it is introduced in sufficiently large quantities. Catastrophic transition is illustrated in Figure 2.2, where the unstable equilibria are depicted by a dashed, stable by a solid line.

We could then restate Theorem 2.1 as follows. When a new population, an invader, is successfully introduced into the community, we observe one of the following: either

(i) a smooth change to a positive, but small invader population size or

(ii) a sudden, catastrophic transition to a rather large invader population size.

When all the assumptions of Theorem 2.1 are met, the sign of M in (2.11) determines which of the two scenarios we will observe in a concrete situation.

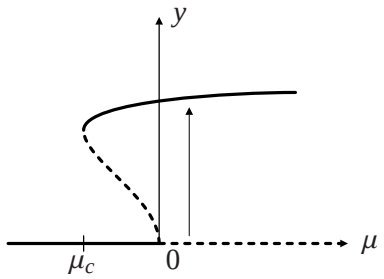

Figure 2.2: Catastrophic transition

We now prove Theorem 2.1.

Proof. We have $G_0 v = 0, w^T G_0 = 0$ and $v \cdot w = 1$. According to A_1, the matrix H_z is invertible. The left and the right zero eigenvectors of $Df(e,0)$, we shall denote them by W and V, are then of the form

$$W = \begin{bmatrix} w \\ 0 \end{bmatrix}, \qquad V = \begin{bmatrix} v \\ -H_z^{-1} H_y v \end{bmatrix}. \tag{2.12}$$

Moreover, $V \cdot W = 1$. By A_1 and A_4 the dimension of the center linear subspace equals one and the subspace is spanned by V.

We now take the (generalized) right eigenvectors of $Df(e,0)$ for the basis of $\mathbb{R}^{m+n}$. It is known that the right (generalized) eigenvectors of $Df(e,0)$ that correspond to nonzero eigenvalues are orthogonal to W.

The center manifold theory [56, 108] states that the center manifold of the equilibrium e, denoted by $\mathcal{M}^c(e)$, can be (locally) parametrised by μ and a real variable u as

$$\mathcal{M}^c(e) = \{(x,\mu);\ x = e(\mu) + uV + \Phi(u,\mu)\},$$

where $\Phi(.)$ is defined on some neighbourhood of the origin. Moreover, $\Phi(0,0) = D\Phi(0,0) = 0$ and $W \cdot \Phi(u,\mu) = 0$ for every u and μ. The center manifold is furthermore invariant under (2.4), that is,

$$\dot{x} = \dot{u}V + \dot{\Phi}(u,\mu) = f(x,\mu) = f(e(\mu) + uV + \Phi(u,\mu),\mu).$$

Since $W \cdot \frac{d}{dt}(\Phi(u,\mu)) = \frac{d}{dt}(W \cdot \Phi(u,\mu)) = 0$ and $V \cdot W = 1$, the inner product with W yields

$$\dot{u} = W \cdot f(e(\mu) + uV + \Phi(u,\mu),\mu) = w \cdot g(e(\mu) + uV + \Phi(u,\mu),\mu),$$

where we have used (2.12) in the last equality. Using the Taylor series expansion around $(e,0)$ we can continue as follows

$$\begin{aligned} \dot{u} &= w \cdot g(e,0) + w \cdot D_x g(e,0)(e(\mu) - e + uV + \Phi(u,\mu)) + w \cdot D_\mu g(e,0)\mu + \\ &+ \tfrac{1}{2} w \cdot D_{\mu\mu} g(e,0)\mu^2 + \tfrac{1}{2} w \cdot D_{xx} g(e,0)(e(\mu) - e + uV + \Phi(u,\mu))^2 + \\ &+ w \cdot D_{\mu x} g(e,0)(e(\mu) - e + uV + \Phi(u,\mu))\mu + O(3), \end{aligned}$$

where $O(3)$ contains the terms of third and higher order in u and μ.
Now, since e is an equilibrium of (2.4), the first term equals zero. So does the second because $w^T G_0 = 0$ and $D_z g(e,0) = 0$. Since $g = Gy$, the third and the fourth term also equal zero. Hence

$$\begin{aligned}\dot{u} &= \tfrac{1}{2} w \cdot D_{xx} g(e,0)(e(\mu) - e + uV + \Phi(u,\mu))^2 \\ &+ w \cdot D_{\mu x} g(e,0)(e(\mu) - e + uV + \Phi(u,\mu))\mu + O(3).\end{aligned}$$

By writing $e(\mu) = e + e'(0)\mu + O(2)$ and taking into account that Φ has no constant and linear terms in u and μ we can continue with

$$\begin{aligned}\dot{u} &= w \cdot \Big(\tfrac{1}{2} D_{xx} g(e,0) e'^2 + D_{\mu x} g(e,0) e'\Big)\mu^2 \\ &+ w \cdot \Big(D_{xx} g(e,0) e' V + D_{\mu x} g(e,0) V\Big)\mu u + \tfrac{1}{2} w \cdot D_{xx} g(e,0) u^2 V^2 + O(3) \\ &= w \cdot \Big(D_{xx} g(e,0) e' V + D_{\mu x} g(e,0) V\Big)\mu u + \tfrac{1}{2} w \cdot D_{xx} g(e,0) u^2 V^2 + O(3),\end{aligned}$$

where we have in the last equality taken into account that the first m components of $e(\mu)$ equal zero and the fact that $g = Gy$ implies $D_{zz} g(e,0) = D_{z\mu} g(e,0) = 0$. Moreover, the special form of g then gives us

$$\dot{u} = \mu u w \cdot \Big(D_x G(e,0) e' + D_\mu G(e,0)\Big) v + \tfrac{1}{2} w \cdot D_x G(e,0) u^2 V^2 + O(3),$$

which, by denoting

$$N = w \cdot \Big(D_x G(e,0) e' + D_\mu G(e,0)\Big) v$$

and using (2.11), (2.12), becomes

$$\dot{u} = \mu N u + \tfrac{1}{2} M u^2 + O(3). \tag{2.13}$$

Note that, according to Lemma 2.1, $N = \frac{d}{d\mu} s(G(e(\mu),\mu))\Big|_{\mu=0}$ and so, by assumption A_3, $N \neq 0$.

Now, the center manifold theory also states that the stability of the steady state under the initial system is determined by its stability under the restriction of the system to the center manifold. This restriction is now given in (2.13). For u and μ close to zero we can neglect the higher order terms that are collected in $O(3)$. The nontrivial steady state solutions of (2.13) that are near the origin are then close to the line $u = -2\mu N M^{-1}$, assuming, of course, that $M \neq 0$. By assumption, N is strictly positive.

Our assumptions were that the equilibrium e is locally stable for $\mu < 0$ and unstable when $\mu > 0$. This steady state corresponds to $u = 0$. The local stability analysis shows that the nontrivial steady states are locally stable when $\mu > 0$ and unstable when $\mu < 0$. We shall see in the following section that we can choose

the eigenvectors v and w so that all their components are nonnegative. Hence, the steady states of (2.4) that correspond to nontrivial equilibria of (2.13) can only be biologically meaningful when either $M < 0$ and $\mu > 0$ or $M > 0$ and $\mu < 0$. Of course, when M equals zero, higher order terms of the Taylor expansion need to be taken into account in order to obtain some information about the nontrivial equilibria of (2.13). □

The determination of the direction of bifurcation simplifies in a number of cases. For example, as mentioned before we can choose the eigenvectors v and w so that all their components are nonnegative. The sign of M can hence sometimes be determined without explicitly calculating the eigenvectors. In the remarks that follow we describe a couple of situations in which further simplifications can be made.

Remark 2.3. One situation in which the expression for the direction of bifurcation can be further simplified is when (2.4) describes the spread of an infectious disease. Introduction of an infectious agent into the community of hosts results in a redistribution of hosts into new compartments of individuals, such as, for example, latent or infectious individuals. A quite common assumption is that the population of hosts has reached an invariant attracting affine set (the reader can find two such examples in Section 2.6), which means that we can eliminate one of the variables. In the case when the population of susceptible hosts is homogeneous (i.e., $n = 1$) we can, by choosing to eliminate the variable corresponding to the susceptible subpopulation (z) redefine G (which is now a function of y only and will be denoted by $\hat{G}$) and arrive at

$$M = \sum_{i,j,k=1,\dots m} w_i \Big(\frac{\partial \hat{G}_{ij}(e,0)}{\partial y_k} + \frac{\partial \hat{G}_{ik}(e,0)}{\partial y_j} \Big) v_j v_k.$$

Remark 2.4. Another circumstance that allows for simplification of (2.11) is when the newly introduced population is homogeneous, i.e. $m = 1$. We can then choose $v = w = 1$ and the expression for the direction of bifurcation becomes

$$\tfrac{1}{2} M = \frac{\partial G(e,0)}{\partial y} - \sum_{k=1\dots,n} \frac{\partial G(e,0)}{\partial z_k} (H_z^{-1} H_y)_k.$$

The reader can find two examples in this spirit in Section 2.6.

2.3 On the characteristic form of population invasion models

The purpose of this section in twofold. We first fulfill the promise made in the previous section and show that assuming that the population invasion models in continuous time have the form (2.4) did not confine our study to a certain subclass of population invasion models. We will see that population invasion models, regardless of the biology that underlies them, indeed have a distinctive form, of which the continuous time version is given in (2.4). Once this part is established we shall provide the reader with a way of obtaining G by only considering the basic modeling ingredients, such as birth, survival and reproduction rates.

So let us suppose that the process of invasion is described by the more general (2.2) with $g \in C^2(\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}, \mathbb{R}^m)$ and $h \in C^1(\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}, \mathbb{R}^n)$.

In ecology and adaptive dynamics we consider invasions of either one new species or a reproductively isolated subpopulation of one of the already present species and we have by the very definiton of reproductive isolation (cf. Appendix) that $g(0, z, \mu) = 0$ for every $z \in \mathcal{Z}, \mu \in \mathbb{R}$ and also $h(y, 0, \mu) = 0$ for every $y \in \mathcal{Y}, \mu \in \mathbb{R}$.

On the other hand, when (2.2) describes the spread of an infectious disease in a population of susceptible hosts, a slight modification of terminology is needed. Namely, when an infectious agent is introduced in a population of susceptible hosts we indeed introduce another species, i.e. the pathogen. However, from that point on we are, on a population level, interested in how this agent spreads in the population of hosts. In this case, therefore, $\mathcal{Y}$ captures the subpopulations of hosts (i.e. members of the resident community) that carry the agent (i.e. the invading species). Since susceptible individuals don't have infected offspring we have $g(0, z, \mu) = 0$ for every $z \in \mathcal{Z}, \mu \in \mathbb{R}$. But since infected individuals (that belong to $\mathcal{Y}$) may become susceptible again (i.e. enter $\mathcal{Z}$) once they get rid of the infection or they might have susceptible offspring, the subspace of the invading community, $\mathcal{Y} \times \{0\}^n$, may not be invariant under (2.2).

In any case we can say the following: since individuals in $\mathcal{Z}$ don't have offspring in $\mathcal{Y}$, the subspace of the resident community, $\{0\}^m \times \mathcal{Z}$, remains invariant under (2.2). That is,

$$g(0, z, \mu) = 0 \quad \text{for every} \quad z \in \mathcal{Z}, \mu \in \mathbb{R}. \tag{2.14}$$

Hence the following known result, due to Hadamard, can be used.

Lemma 2.2. *Let $f = (f_1, ..., f_k)^T \in C^r(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^k)$ for some $r \in \mathbb{N}$ be such that $f(0, y) = 0$ for every $y \in \mathbb{R}^n$. There exists an $F \in C^{r-1}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^{k \times m})$ such that*

$$f(x, y) = F(x, y)x, \quad x = (x_1, ..., x_m)^T.$$

The proof of this result can be found in [38].

The property (2.14) therefore yields a matrix $G \in M_{m\times m}(C^1(\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R},\mathbb{R}))$ such that g in (2.2) can be written as

$$g(y,z,\mu) = G(y,z,\mu)y, \qquad y = (y_1,...,y_m)^T. \tag{2.15}$$

Note that this decomposition is in general not unique as the following simple example shows.

Example 2.1. Take $y = (y_1, y_2)$ and $g = y_1y_2$. Then $G = [y_2, 0]$ and $G = [0, y_1]$ yield two possible decompositions.

This nonuniqueness will, however, not affect our results - for our purposes, any correct decomposition will do. We shall nevertheless now point out one, a more interpretation motivated (and hence perhaps mainly of use to more biologically inclined readers), way of obtaining G in (2.15).

The key to this decomposition is the so called *environmental condition* [32, 34]. The defining property of an environmental condition is that individuals are independent of one another (and hence the equations are linear) when this condition is prescribed as a function of time. We then view (2.2) as linear systems together with feedback equations that tell us how, in turn, the environment is influenced by the population size and composition. In general, the environment is set by all subpopulations involved. The environmental condition will hence be a function of $x = (y, z)$.

Readers that find this general definition a bit unclear are encouraged to have a look at Box 2.1, where we explain the notion of an environmental condition by way of a simple example.

In order to arrive directly at the desired decomposition of g in (2.2a) we first separate the reproduction in $\mathcal{Y}$ from all other processes.

Since individuals in $\mathcal{Z}$ don't have offspring in $\mathcal{Y}$ the invading population completely determines the reproduction in $\mathcal{Y}$. To describe it we define the following matrix,

$$P_{ij}(x) : = \textit{the rate at which individuals with birth state } i \textit{ are born to an individual with state } j\textit{, given a constant environmental condition } x \in \mathcal{X}.$$

What remains is to describe other processes, namely maturation and survival. We consider the dynamics of an individual's state after birth as a Markov process on the set of i- states, where the probabilities of changing a state are again determined by the environmental condition $x \in \mathcal{Y}\times\mathcal{Z}$. We define the matrix $Q(x)$ as follows:

Box 2.1. Environmental condition

The defining property of an **environmental condition** [32, 33, 34] (also called *environmental variable* or *input*) is that individuals are independent of one another when this condition is prescribed as a function of time.
The notion of the environmental condition can perhaps most easily be clarified by way of examples. So let us consider the following competition model, first used by Gause in 1934 [63],

$$\begin{aligned} \dot{N}_1 &= r_1 N_1 \frac{K_1 - (N_1 + \alpha_2 N_2)}{K_1} \\ \dot{N}_2 &= r_2 N_2 \frac{K_2 - (N_2 + \alpha_1 N_1)}{K_2} \end{aligned}$$

Here, N_1 and N_2 denote the sizes of two competing populations. When population j ($j = 1, 2$) is absent, population $i \neq j$ grows logistically with intrinsic growth rate r_i and carrying capacity K_i. The parameter α_i ($i = 1, 2$) represents the intensity of the negative influence that population i has (due to competition) on the other population.
In this case both populations influence the rates at which these two populations interact. Hence, defining the environmental variable I as $I = (I_1, I_2) = (N_1, N_2)$, we can rewrite the equations so that all interactions are expressed in terms of the environmental condition I,

$$\begin{aligned} \dot{N}_1 &= r_1 N_1 \frac{K_1 - (I_1 + \alpha_2 I_2)}{K_1} \\ \dot{N}_2 &= r_2 N_2 \frac{K_2 - (I_2 + \alpha_1 I_1)}{K_2}, \end{aligned}$$

And indeed one sees that when I is prescribed as a function of time the individuals act independently of one another, i.e., the equations are linear.

$$Q_{ij}(x) : = \begin{cases} \textit{the rate of leaving state } j \textit{ to go to state } i & ; i \neq j \\ -\textit{ the rate of leaving state } j & ; i = j, \end{cases}$$

given the environmental condition x.

Hence, the off-diagonal elements of Q describe the changes of states as long as the individual remains alive and does not move to $\mathcal{Z}$, while the diagonal elements denote the rate of leaving the state, either by leaving to another state in $\mathcal{Y}$, to $\mathcal{Z}$ or by death. By taking into account all the processes we can now write the matrix G as

$$G = P + Q. \tag{2.16}$$

This decomposition has, apart from offering biological interpretation, another advantage. Since the off-diagonal elements of G in (2.16) are nonnegative we can apply the theory of nonnegative matrices [7] to see that we can indeed choose the (left and the right) eigenvector of G_0 in (2.7) corresponding to eigenvalue zero to be nonnegative. If the off-diagonal elements of G_0 are strictly positive, the eigenvectors can be chosen to be strictly positive. In many cases this observation makes it a lot easier to determine the sign of M in (2.11).

Now, the above definition of the matrices P and Q leads to the decomposition of g for continuous time models. The discrete time analogue can be obtained in a similar manner, by replacing the rates with appropriate probabilities. In the discrete time case, however, the matrix $G = Q + P$ will be a positive matrix.

Note also that one could make a similar 'per capita' description for the resident populations. For our purposes, however, this description is irrelevant since we are only interested in the c- states of the resident community that are not close to zero. It might help, though, when one wants to find $z_0(\mu)$ [33].

2.4 Population invasion models in discrete time

Sometimes the nature of the problem, the available data or some other reason makes it more convenient to formulate a discrete time population model. Since the linearization theorem of Hartman and Grobman [56, 108] and the center manifold theory apply for discrete time dynamical systems generated by maps as well as for flows generated by vector fields we can reformulate the problem and the results to hold for population invasion models in discrete time.

In the same way as before we decompose the population state space

$$\mathcal{Y} \times \mathcal{Z} = \mathbb{R}^m_+ \times \mathbb{R}^n_+,$$

so that $\mathcal{Y}$ denotes the population state space of the newly introduced population and $\mathcal{Z}$ the state space of the resident community. We shall now study processes described by parametrized maps

$$\begin{aligned} y &\mapsto G(y,z,\mu)y \\ z &\mapsto h(y,z,\mu) \end{aligned} \qquad y \in \mathbb{R}^m, z \in \mathbb{R}^n, \mu \in \mathbb{R}, \tag{2.17}$$

where we assume that $G \in M_{m\times m}(C^1(\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}, \mathbb{R}))$, $h \in C^1(\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}, \mathbb{R}^n)$. We shall also use the notation in (2.2b) or (2.5) to write the map (2.17) as $x \mapsto f(x,\mu)$ whenever this notation will be more convenient.

Suppose now that we have a steady state, that is, a fixed point of (2.17) of the form $e(\mu) = (0, z_0(\mu))$ for some $z_0(\mu) \in \mathcal{Z}$. The associated linear map is then

given by

$$Df(e(\mu),\mu) = \begin{bmatrix} G(e(\mu),\mu) & 0 \\ h_y(e(\mu),\mu) & h_z(e(\mu),\mu) \end{bmatrix}.$$

Hence

$$\sigma(Df(e(\mu),\mu)) = \sigma(G(e(\mu),\mu)) \cup \sigma(h_z(e(\mu),\mu)).$$

We shall again assume that the steady state $e(\mu)$ is internally asymptotically stable, i.e., that it is asymptotically stable under perturbations within the invariant subspace $\{0\}^m \times \mathcal{Z}$ and that this can be inferred from the linearization. In the discrete time setting this means that we assume

B₁ : if $\lambda \in \sigma(h_z(e(\mu),\mu))$ then $|\lambda| < 1$.

The spectrum of $G(e(\mu),\mu)$ hence determines the linearized stability of $e(\mu)$. Again, the theory of nonnegative matrices tells us that the spectral radius of G is an eigenvalue and that the corresponding eigenvector can be chosen to be nonnegative. The interesting case to consider is therefore when the parameter μ is such that $G(e(\mu),\mu)$ has an eigenvalue one, a situation where the linearization alone does not tell us whether the newly introduced population is able to settle in the community.

We now again take for a bifurcation parameter some parameter μ, such that the fixed points of (2.17) of the form $e(\mu) = (0, z_0(\mu))$ are linearly stable for $\mu < 0$ and unstable when $\mu > 0$. Thus

$$\mathbf{B_2}: \quad \begin{cases} \mu < 0 \iff r(G(e(\mu),\mu)) < 1 \iff \mathcal{R}_0 < 1 \\ \mu = 0 \iff r(G(e(\mu),\mu)) = 1 \iff \mathcal{R}_0 = 1 \end{cases}$$

where $r(.)$ denotes the spectral radius. We refer the reader to [19, 72] for the justification of the equivalence between $r(.)$ and $\mathcal{R}_0$. Since the results that follow rely upon local information only, it suffices that B_2 holds in some neighbourhood of $\mu = 0$.

Assumption B_2 tells us that the function $\mu \mapsto r(G(e(\mu),\mu))$ crosses the point $(\mu, r(.)) = (0, 1)$. We shall again assume that this crossing occurs at nonzero 'speed', i.e.,

B₃ : $\frac{d}{d\mu} r(G(e(\mu),\mu))\big|_{\mu=0} > 0$.

Let e denote the equilibrium that corresponds to $\mathcal{R}_0 = 1$, i.e., $e = e(0)$ and let $e' = e(0)$. We shall furthermore use the notation introduced in (2.7). Denoting by E^c the center subspace of G_0 we shall furthermore assume that

B₄ : $\dim E^c = 1$

and refer the reader to the next section for the biological interpretation of this assumption.

We can now write the discrete time analogue of Theorem 2.1.

Theorem 2.2. *Consider a population model described by* (2.17) *and let* $e(\mu) = (0, z_0(\mu))$ *be a steady state of* (2.17). *Assume that* B_1, B_2, B_3 *and* B_4 *hold. Furthermore assume that* $\mu \mapsto e(\mu) \in C^1(\mathbb{R}, \mathbb{R}^{n+m})$ *and denote by* e *the steady state that corresponds to* $\mathcal{R}_0 = 1$, *i.e.,* $e = e(0)$ *and by* $e' = e'(0)$. *Let* G_0, H_y *and* H_z *be as in* (2.7) *and let* w *and* v *denote, respectively, the left and the right eigenvector of* G_0 *corresponding to eigenvalue one, normalized, so that* $v \cdot w = 1$. *Let*

$$\begin{aligned} M = &\sum_{i,j,k=1,\dots m} w_i \Big(\frac{\partial G_{ij}(e,0)}{\partial y_k} + \frac{\partial G_{ik}(e,0)}{\partial y_j} \Big) v_j v_k - \\ &- 2 \sum_{\substack{i,j=1\dots m \\ k=1\dots,n}} w_i \frac{\partial G_{ij}(e,0)}{\partial z_k} v_j ((I - H_z)^{-1} H_y v)_k \end{aligned} \tag{2.18}$$

There exists a $\delta > 0$ *such that*

(i) if $M < 0$ *there is a branch* $\mu \mapsto (y(\mu), z(\mu))$, *defined for* $\mu \in (0, \delta)$, *of positive, locally asymptotically stable steady states of* (2.17). *In other words, the bifurcation is supercritical.*

(ii) if $M > 0$ *there is a branch* $\mu \mapsto (y(\mu), z(\mu))$, *defined for* $\mu \in (-\delta, 0)$, *of positive, unstable steady states of* (2.17). *That is, the bifurcation is subcritical.*

Remark 2.5. We again see that only the first order derivatives of G and h are needed to determine the direction of bifurcation from e. Moreover, the expression for the direction of bifurcation is independent of the bifurcation parameter but for the restrictions B_2 and B_3. In other words, provided that all the assumptions of Theorem 2.2 are satisfied, we obtain the same direction of bifurcation for any bifurcation parameter.

Remark 2.6. For some further remarks on the terminology and on the interpretation of the results of Theorem 2.2 in biological terms we refer the reader to the remarks made after Theorem 2.1.

Proof. We have $G_0 v = v, w^T G_0 = w^T$ and $v \cdot w = 1$. According to B_1, the matrix $I - H_z$ is invertible. We can then calculate the left and the right eigenvector of $Df(e, 0)$ corresponding to eigenvalue one, we shall denote them by W and V, and find that

$$W = \begin{bmatrix} w \\ 0 \end{bmatrix} \qquad V = \begin{bmatrix} v \\ (I - H_z)^{-1} H_y v \end{bmatrix}. \tag{2.19}$$

Moreover, $V \cdot W = 1$.

By B_1 and B_4 the dimension of the center linear subspace equals one and the subspace is spanned by V. We take the set of (generalized) eigenvectors of

$Df(e,0)$ for the basis of $\mathbb{R}^{m+n}$. The right eigenvectors of $Df(e,0)$ that correspond to eigenvalues different from one are orthogonal to W.

The center manifold theory states that there exists a center manifold of the equilibrium e, denoted by $\mathcal{M}^c(e)$, that can be locally parametrised by μ and a real variable u as

$$\mathcal{M}^c(e) = \{(x,\mu);\ x = e(\mu) + uV + \Phi(u,\mu)\},$$

where Φ is defined on some neighbourhood of the origin. Moreover, $\Phi(0,0) = D\Phi(0,0) = 0$ and $W \cdot \Phi(u,\mu) = 0$ for every u and μ. Since the center manifold is also invariant under (2.17) we have

$$\begin{aligned} x(k+1) &= e(\mu) + u(k+1)V + \Phi(u(k+1),\mu) \\ &= f(x(k),\mu) \\ &= f(e(\mu) + u(k)V + \Phi(u(k),\mu),\mu). \end{aligned}$$

We calculate the inner product with W, take into account that $W \cdot e(.) = 0$, $W \cdot V = 1$ and $W \cdot \Phi(.) = 0$ and obtain

$$u(k+1) = w \cdot g(e(\mu) + u(k)V + \Phi(u(k),\mu),\mu).$$

Written differently, the restriction of (2.17) to the center manifold is given by the map

$$u \mapsto w \cdot g(e(\mu) + uV + \Phi(u,\mu),\mu).$$

Using the Taylor series we can now write

$$\begin{aligned} u \ \mapsto\ & w \cdot g(e,0) + w \cdot D_x g(e,0)(e(\mu) - e + uV + \Phi(u,\mu)) + w \cdot D_\mu g(e,0)\mu \\ +\ & \tfrac{1}{2} w \cdot D_{xx} g(e,0)(e(\mu) - e + uV + \Phi(u,\mu))^2 + \tfrac{1}{2} w \cdot D_{\mu\mu} g(e,0)\mu^2 \\ +\ & w \cdot D_{\mu x} g(e,0)(e(\mu) - e + uV + \Phi(u,\mu))\mu + O(3) \end{aligned}$$

where $O(3)$ denotes third and higher order terms in u and μ.

Now, the first term equals zero since e is a fixed point of f and therefore $g(e,0) = 0$. The second term equals u since $w^T G_0 = w^T$, $W \cdot e(\mu) = W \cdot e = 0$, $W \cdot V = 1$ and $W \cdot \Phi(.) = 0$.

Since $g = Gy$, the third and the fourth term are also equal to zero. Furthermore, by writing $e(\mu) = e(0) + e'(0)\mu + O(2)$, taking into account that Φ has no constant and no linear terms in u and μ and noting that $W^T = (w,0)^T$ we are left with

$$\begin{aligned} u \ \mapsto\ & u + w \cdot \left(\tfrac{1}{2} D_{xx} g(e,0) e'^2 + D_{\mu x} g(e,0) e'\right)\mu^2 \\ +\ & w \cdot \left(D_{xx} g(e,0) e' V + D_{\mu x} g(e,0) V\right)\mu u \\ +\ & \tfrac{1}{2} w \cdot D_{xx} g(e,0) u^2 V^2 + O(3), \end{aligned}$$

which, by writing

$$N = w \cdot \Big(D_x G(e,0)e' + D_\mu G(e,0)\Big)v,$$

taking into account that $g = Gy$, (2.18) and the fact that the first m components of e equal zero, becomes

$$u \mapsto u + \tfrac{1}{2}Mu^2 + \mu Nu + O(3). \tag{2.20}$$

Similar reasoning as in Lemma 2.1 establishes that assumptions B_2 and B_4 lead to

$$\tfrac{d}{d\mu} r(G(e(\mu),\mu))\Big|_{\mu=0} = w \cdot \Big(D_x G(e,0)e' + D_\mu G(e,0)\Big)v \tag{2.21}$$

and so by (2.21) and assumption B_4, $N > 0$.

For u and μ close to zero we can neglect the higher order terms that are collected in $O(3)$ and look for fixed points of $u \mapsto u + \frac{1}{2}Mu^2 + \mu Nu$. Nonzero fixed points are then near the line given by $u = -2\mu NM^{-1}$.

Our assumptions were that the steady state e is locally stable for $\mu < 0$ and unstable when $\mu > 0$. This steady state corresponds to $u = 0$. The local stability analysis yields that the nontrivial steady states are locally stable when $\mu > 0$ and unstable when $\mu < 0$. As we have argued in Section 2.2, we can choose the eigenvectors v and w so that all their components are nonnegative. Hence, the steady states of (2.17) that correspond to nontrivial equilibria of (2.20) can only be biologically meaningful when either $M < 0$ and $\mu > 0$ or $M > 0$ and $\mu < 0$. If M is zero then higher order terms of the Taylor expansion need to be taken into account in order to obtain some information about the nontrivial equilibria of (2.20). □

All the situations mentioned at the end of Section 2.2 that lead to a simplified formula for the direction of bifurcation, occur, of course, also in the discrete time setting. However, modifying the obtained formulas for M to apply for discrete time models is a rather straightforward matter and we therefore omit it.

2.5 On the basic reproduction ratio. The case $\mathcal{R}_0 = 1$

The aim of this section is to offer some interpretation of the assumptions made in previous sections. In order to do this we shall state some known results and refer the reader that is interested in their proofs to the literature. Two basic notions, the one of the next generation matrix and the other of $\mathcal{R}_0$ are also defined in Box 2.2.

The basic reproduction ratio $\mathcal{R}_0$ is defined as the expected number of offspring an 'average' individual has in all of its life and is mathematically expressed as the spectral radius of the next generation operator [34, 35, 105]. The key to the

calculation of $\mathcal{R}_0$ of the newly introduced population in the context of the model is a decomposition of g which separates reproduction in $\mathcal{Y}$ from other transitions (for instance, separates new infections from progression of the disease to another stage), as was already done in Section 2.3.
We then write

$$g(y,z) = (P(y,z) + Q(y,z))y,$$

where P and Q are as in Section 2.3.

Let us focus on the continuous time case. If we now denote by $e = (0, z_0)$ a steady state of the system and define $\mathcal{P} = P(e)$ and $\mathcal{Q} = Q(e)$, then $\mathcal{P}$ is a nonnegative matrix and $-\mathcal{Q}$ is a nonsingular M-matrix [7, 105]. Hence, $\mathcal{Q}$ is invertible and $-\mathcal{Q}^{-1}$ is nonnegative. Moreover, the entries of Q come with an interpretation, namely, the element $-\mathcal{Q}^{-1}_{jk}$ equals the time an individual born with i- state k is expected to spend in state j [7, 35, 105]. The matrix $-\mathcal{Q}^{-1}$ describes individual's i- state dynamics. The matrix $\mathcal{P}$, on the other hand, describes the reproduction and so the product $-\mathcal{P}\mathcal{Q}^{-1}$ yields the next generation matrix*. By definition, $\mathcal{R}_0$ equals its spectral radius. One can also prove [7, 35, 105] that the following holds

$$\begin{cases} \mathcal{R}_0 = r(-\mathcal{P}\mathcal{Q}^{-1}) < 1 \iff s(\mathcal{P}+\mathcal{Q}) = s(G_0) < 0 \\ \mathcal{R}_0 = r(-\mathcal{P}\mathcal{Q}^{-1}) = 1 \iff s(\mathcal{P}+\mathcal{Q}) = s(G_0) = 0, \end{cases} \tag{2.22}$$

where r denotes the spectral radius and s the spectral bound.

Now, it is reasonable to assume irreducibility of the next generation matrix (cf. Appendix) . This guarantees that the spectral radius is an algebraically simple eigenvalue and that we can choose a strictly positive corresponding eigenvector [7]. In biological terms, the assumption of irreducibility of the next generation matrix means that the invading subpopulations are well mixed, that is, for every pair of i- states j and k the individuals of the j-th subpopulation will eventually have offspring in the k-th subpopulation.

Under a more strict condition, namely the primitivity [7] of the next generation matrix, the modulus of the spectral radius is strictly greater than the modulus of any other eigenvalue of $-\mathcal{P}\mathcal{Q}^{-1}$. In biological terms the assumption of primitivity means that we require that, from some generation on, individuals with birth state j can have offspring with birth state k for any two conceivable i- states at birth j and k.

We have assumed in sections 2.2 and 2.4 that the dimension of the center subspace of G_0 equals one. Now, in the discrete time setting the matrix G_0 is a nonnegative matrix. Its primitivity therefore guarantees that the assumption B_4 is

*For the continuous time case, while for the discrete time decomposition $G = Q + P$, the next generation matrix is given by $P(I - Q)^{-1}$ [72]

Box 2.2. The next generation matrix and $\mathcal{R}_0$

The basic reproduction ratio, $\mathcal{R}_0$, is defined as the expected number of offspring that an 'average' individual has in all of its life. For some very simple examples (such as for instance (2.1)), the basic reproduction ratio is easy to compute. When the populations are not homogeneous, however, things become more complicated. In particular, the right way of finding this 'average' individual is not a priori clear.
The mathematical definition of $\mathcal{R}_0$ for general physiologically structured population models first appeared around 1990 (see [31, 35, 36]) and it goes as follows. First, one needs to find the set of conceivable *i*- states at birth. We then compute the **next generation matrix**, which we denote by R and which is defined as follows

$$R_{ij}(I) := \begin{array}{l} \textit{the expected number of offspring with birth state } i, \\ \textit{born to one individual that was born with state } j, \\ \textit{given a constant environmental condition } I. \end{array}$$

The basic reproduction ratio, $\mathcal{R}_0(I)$, is by definition [31, 32, 34, 35, 36] the spectral radius of $R(I)$.
Let us, as a very simple example, consider an infectious disease that spreads in a population of hosts via vectors. For instance, one may have malaria in mind, with humans being hosts and mosquitos being the vectors. Let us label the hosts as having *i*- state 1 and vectors as having *i*- state 2. Since hosts cannot infect hosts directly and vectors also cannot infect vectors, the next generation matrix takes the form

$$R = \begin{bmatrix} 0 & r_{12} \\ r_{21} & 0 \end{bmatrix},$$

where r_{12} and r_{21} denote, respectively, the expected number of vector-host and host-vector transmissions. Then $\mathcal{R}_0 = \sqrt{r_{12}r_{21}}$.
For more examples of $\mathcal{R}_0$ computations see Section 2.6 and [35, 103, 105].

satisfied. In contrast with the discrete time setting we know that G_0 in the continuous time case is a nonnegative off - diagonal matrix. The Perron - Frobenius theory then tells us that already the assumption of an irreducible G_0 guarantees that A_4 holds, i.e., the zero eigenvalue is an algebraically simple eigenvalue and all other eigenvalues have strictly negative real parts. In biological terms an irreducible G_0 means that for every pair of *i*- states j and k ($j \neq k$) we will eventually observe an inflow of individuals of the j- th subpopulation to the k- th subpopulation.

We have seen that the fact that the linearization theorem of Hartman and Grob-

man and the center manifold theory run for discrete time dynamical systems generated by maps parallel to the one for flows generated by vector fields, allows us to formulate the results for both settings. The basic reproduction ratio, $\mathcal{R}_0$, provides, due to relation (2.22), a further connection and links the continuous time results directly to corresponding discrete time results.

2.6 Examples

In this section we present four examples to illustrate the theory presented in previous sections.

In the first example we study a continuous time model describing the dynamics in a community in which the predator selectively forages on a stage structured prey. This example is motivated by the work of de Roos et al. [26] and, hopefully, demonstrates how little effort is needed to study the occurrence of subcritical equilibria for a class of models, in this particular case, models obtained by varying the preference of the predators.

The second example is a discrete time model describing the life cycle of biennials. This example is inspired by the work of Davydova et al. [20, 21] and demonstrates how the theory can also be applied to study reproductively isolated subpopulations of the same species to see whether one missing year class is, after being introduced, able to settle among the existing year classes.

The last two examples are simple continuous time epidemic models related to the author's other work, namely, modelling the spread of infectious agents that can reside at several different parts of the host's body. Though very simple in the first place, they illustrate how the determination of the direction of bifurcation can be further simplified by (i) assuming that the total population size has reached an equilibrium and (ii) the fact that the eigenvectors in question can be chosen to be nonnegative (see Remark 2.3).

After determining the direction of bifurcation from a 'residents only' steady state we shall in all of these examples write some interpretation of the results for the problem at hand. We have, however, already in the remarks after Theorem 2.1 described in biological terms what can in general be said about an invasion given that we know the direction of bifurcation. We shall therefore not repeat these general facts in the examples and rather refer the reader to Section 2.2.

Example 2.2. In this first example we study a continuous time model describing interactions in a community that consists of a stage structured prey population and a population of predators that preys exclusively on one of the prey stages.

Suppose that the prey is divided into three stages - juveniles, subadults and adults, and let their numbers be denoted respectively by J, S and A. The number of the predators will be denoted by P. We describe the dynamics of the predator

population that forages exclusively on the adult stages of prey by the following differential equation,

$$\frac{dP}{dt} = (\phi f(A) - \nu)P. \tag{2.23}$$

Here, ϕ indicates the conversion efficiency of prey biomass into newborn predators, ν denotes the per capita death rate of the predators and $f(.)$ stands for the predator functional response (for example, Holling type 2 or Holling type 3 response). In what follows, the function f will not be specified; we shall only assume that it is an increasing function of the abundance of the adult prey.

We shall take a closer look at the situation in which the regulation of the prey population takes place within the subadult stage. We describe the dynamics of the prey population by the following system of differential equations,

$$\begin{aligned} \frac{dJ}{dt} &= \beta A - (\rho + \mu_J)J \\ \frac{dS}{dt} &= \rho J - (\pi(S) + \mu_S(S))S \\ \frac{dA}{dt} &= \pi(S)S - \mu_A A - f(A)P. \end{aligned} \tag{2.24}$$

Here, the parameters have the following meaning: β denotes the adult fecundity, ρ the maturation rate from the juvenile to the subadult stage and μ_J the per capita death rate of the juveniles. Functions $\pi(S)$ and $\mu_S(S)$ denote, respectively, the (possibly density dependent) maturation rate of subadults into adults and per capita death rate of the subadults. The per capita death rate of the adult prey in the absence of predators is denoted by μ_A.

Regulation of the subadult prey population through maturation and/or mortality can occur if the maturation rate $\pi(.)$ decreases and/or the mortality rate $\mu_S(.)$ increases with an increase in the abundance of subadults. We shall therefore assume that $\pi(.)$ is a nonincreasing and $\mu_S(.)$ a nondecreasing function of S and exclude the situation in which the derivatives of both vanish in some point.

We now first calculate the steady states of (2.24) in the absence of the predators. We obtain (the steady state values are denoted by *)

$$\begin{aligned} J^* &= \frac{\beta A^*}{\mu_J + \rho} = \frac{\beta\pi(S^*)S^*}{\mu_A(\mu_J + \rho)}, \\ A^* &= \frac{\pi(S^*)S^*}{\mu_A} \end{aligned}$$

as the steady state numbers of the juvenile and adult prey and the following equilibrium equation for the (nontrivial) steady state value of the abundance of the subadult prey,

$$\frac{\rho\beta\pi(S^*)}{\mu_A(\mu_J + \rho)} = \pi(S^*) + \mu_S(S^*). \tag{2.25}$$

Now, in our previous notation we would have $y = P$ and $(z_1, z_2, z_3) = (J, S, A)$ and so $G = \phi f(A) - \nu$. Since the predator population is homogeneous, we can take $w = v = 1$. Furthermore, we consider the case when the basic reproduction ratio of the predators equals one, i.e.,

$$\mathcal{R}_0 = \frac{\phi f(A^*)}{\nu} = 1$$

and so $f(A^*) = \frac{\nu}{\phi}$.

Since G is a function of A only we have

$$M = -2G'(A^*)(H_z^{-1}H_y)_3, \tag{2.26}$$

where

$$H_y = \begin{bmatrix} 0 \\ 0 \\ -f(A^*) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\frac{\nu}{\phi} \end{bmatrix}$$

and

$$H_z = \begin{bmatrix} -(\rho + \mu_J) & 0 & \beta \\ \rho & -(\pi(S)S + \mu_S(S)S)'\big|_{S=S^*} & 0 \\ 0 & (\pi(S)S)'\big|_{S=S^*} & -\mu_A \end{bmatrix}.$$

Since only the last component of H_y is nonzero and we only need the third component of $H_z^{-1}H_y$ it suffices to calculate $(H_z^{-1})_{33}$. We have

$$(H_z^{-1})_{33} = \frac{1}{\det H_z}(\rho + \mu_J)(\pi(S)S + \mu_S(S)S)'\big|_{S=S^*}$$

and we can now rewrite (2.26) as

$$M = 2\nu(\rho + \mu_J)f'(A^*)(\det H_z)^{-1}(\pi(S)S + \mu_S(S)S)'\big|_{S=S^*}. \tag{2.27}$$

The parameters ν and $(\rho + \mu_J)$ are positive. According to our assumptions, so is $f'(A^*)$. Using (2.25) we can furthermore see that

$$\det H_z = \frac{\beta\rho S^*(\pi'(S^*)\mu_S(S^*) - \pi(S^*)\mu_S'(S^*))}{\pi(S^*) + \mu_S(S^*)},$$

which is, by our assumptions on $\pi(.)$ and $\mu_S(.)$, strictly negative. We have therefore arrived at the fact that

$$\text{sign } M = -\text{sign } (\pi(S)S + \mu_S(S)S)'\big|_{S=S^*},$$

as was also found in [26].

We have assumed that μ_S is a nondecreasing function of the subadult abundance. The function $\mu_S S$ is therefore an increasing function. In more biological

terms we could therefore say that a subcritical bifurcation can only occur when an overcompensation in the total maturation rate $\pi(S)S$ takes place, i.e. for certain values of S, an increase in the abundance of subadults actually decreases the total maturation rate, and that this overcompensation is sufficiently strong.

In [26] the authors also studied the cases when the predator forages exclusively on either the juvenile or the subadult prey and found that the emergent Allee effect can occur (with a suitable overcompensation in the regulation) when the predators forage on one of the nonregulating stages of the prey population and can never occur when they forage on the regulating stage.

Hopefully, this example shows how little effort it would take to, with the tools that we have developed in the previous sections, consider these and also many other situations of interest.

Example 2.3. In this example we consider a community of strict biennials, that is, a community that consists of two age classes, with only the eldest class reproducing. Time will in this case be measured in years. We shall label the two classes by indices 0 and 1, the 0 denoting the subpopulation of individuals that have not reached age one and 1 the subpopulation of one year old individuals. If individuals survive till the end of their second year they reproduce and die.

Survival and reproduction rates are described in terms of an environmental condition I, which will be taken to be the weighed sum of the two populations. More precisely, if $x_j(t)$ denotes the number (or the density) of j-year old individuals ($j = 0, 1$) at time t, we take

$$I(t) = c_0 x_0(t) + c_1 x_1(t).$$

The weights c_0 and c_1 are also called the *impacts* of the corresponding age classes. Now let us denote by $F_0(I(t))$ the probability of surviving the first year and by $F_1(I(t))$ the number of offspring of an individual that survives till the end of its second year. Since increasing the value of I means worsening the conditions for both classes, the functions F_0 and F_1 are decreasing functions. We shall also assume that they are differentiable at least once.

We can now formulate the following discrete time model

$$\begin{aligned} x_0(t+1) &= F_1(I(t))x_1(t) \\ x_1(t+1) &= F_0(I(t))x_0(t). \end{aligned}$$

The functions F_i are also called *sensitivities to the environment* and the index specifies how this sensitivity depends on age. Typical examples of sensitivity functions are the so called

(i) *Ricker family* where $F_i(I) = a_i e^{-b_i I}$ and

(ii) *Beverton - Holt family* where $F_i(I) = a_i(1 + b_i I)^{-1}$.

In order to illustrate the theory on this example we first compute the Full Life Cycle map, that is, we apply the map

$$\begin{bmatrix} x_0(t) \\ x_1(t) \end{bmatrix} \mapsto \begin{bmatrix} 0 & F_1(I(t)) \\ F_0(I(t)) & 0 \end{bmatrix} \begin{bmatrix} x_0(t) \\ x_1(t) \end{bmatrix}$$

twice to obtain the community state after a two year time interval. We have

$$\begin{aligned} x_0(t+2) &= F_1(I(t+1))x_1(t+1) \\ &= F_1\Big(c_0F_1(I(t))x_1(t) + c_1F_0(I(t))x_0(t)\Big)F_0(I(t))x_0(t) \\ &= F_1(I_1(t))F_0(I(t))x_0(t), \end{aligned}$$

where

$$I_1(t) := c_0F_1(I(t))x_1(t) + c_1F_0(I(t))x_0(t)$$

denotes the environmental condition in the second year.

The Full Life Cycle map is then given by

$$\begin{aligned} x_0(t+2) &= F_1(I_1(t))F_0(I(t))x_0(t) \\ x_1(t+2) &= F_0(I_1(t))F_1(I(t))x_1(t). \end{aligned}$$

Now let us assume that only individuals with label zero are present in every second year and that the population is in a steady state, say x_0^*. This means that

$$F_1(c_1F_0(c_0x_0^*)x_0^*)F_0(c_0x_0^*) = 1. \tag{2.28}$$

Furthermore, the assumption that the basic reproduction ratio of invaders, i.e., individuals with label one equals one, translates into

$$F_0(c_1F_0(c_0x_0^*)x_0^*)F_1(c_0x_0^*) = 1. \tag{2.29}$$

In our previous notation we have

$$\begin{aligned} G &= F_0(I_1(t))F_1(I(t)) \\ h &= F_1(I_1(t))F_0(I(t))x_0(t). \end{aligned}$$

Let us now compute the required derivatives for the case where both sensitivity functions belong to the Ricker family. We obtain

$$\begin{aligned} \frac{\partial G}{\partial x_1} &= F_0(I_1)F_1(I)\Big(b_0b_1c_0c_1F_1(I)x_1 + b_0^2c_1^2F_0(I)x_0 - b_0c_0F_1(I) - b_1c_1\Big) \\ \frac{\partial G}{\partial x_0} &= F_0(I_1)F_1(I)\Big(b_0b_1c_0^2F_1(I)x_1 + b_0^2c_0c_1F_0(I)x_0 - b_0c_1F_0(I) - b_1c_0\Big) \\ \frac{\partial h}{\partial x_1} &= F_0(I_1)F_1(I)x_0\Big(b_1^2c_0c_1F_1(I)x_1 + b_0b_1c_1^2F_0(I)x_0 - b_1c_0F_1(I) - b_0c_1\Big) \\ \frac{\partial h}{\partial x_0} &= F_0(I_1)F_1(I)\Big(x_0(b_1^2c_0^2F_1(I)x_1 + b_0b_1c_0c_1F_0(I)x_0 - b_1c_1F_0(I) - b_0c_0) + 1\Big). \end{aligned}$$

We evaluate these derivatives in $x_0 = x_0^*, x_1 = 0$, take (2.28) and (2.29) into account and denote the results respectively by $\mathcal{G}_1, \mathcal{G}_0, \mathcal{H}_1$ and $\mathcal{H}_0$. We arrive at

$$\begin{aligned}
\mathcal{G}_1 &= b_0^2 c_1^2 F_0(c_0 x_0^*) x_0^* - b_0 c_0 F_1(c_0 x_0^*) - b_1 c_1 \\
\mathcal{G}_0 &= b_0^2 c_0 c_1 F_0(c_0 x_0^*) x_0^* - b_0 c_1 F_0(c_0 x_0^*) - b_1 c_0 \\
\mathcal{H}_1 &= x_0^* \Big(b_0 b_1 c_1^2 F_0(c_0 x_0^*) x_0^* - b_1 c_0 F_1(c_0 x_0^*) - b_0 c_1 \Big) \\
\mathcal{H}_0 &= x_0^* \Big(b_0 b_1 c_0 c_1 F_0(c_0 x_0^*) x_0^* - b_1 c_1 F_0(c_0 x_0^*) - b_0 c_0 \Big) + 1,
\end{aligned}$$

Now, equalities (2.28) and (2.29) in the Ricker case imply that

$$b_0 = b_1 \qquad \text{or} \qquad \Big(F_0(c_0 x_0^*) = \tfrac{c_0}{c_1} \quad \text{and} \quad F_1(c_0 x_0^*) = \tfrac{c_1}{c_0} \Big). \tag{2.30}$$

Moreover, we can take $v = w = 1$ in (2.18). The expression for the direction of bifurcation then translates into

$$M = 2\mathcal{G}_1 - 2\mathcal{G}_0 \mathcal{H}_1 \Big(\mathcal{H}_0 - 1 \Big)^{-1}$$

and one can quickly see, by taking (2.30) into account, that the bifurcation is vertical (i.e., $M = 0$), as was also found in [20].

In [20, 21] it was actually shown that the bifurcation is vertical in the stronger sense that a family of period two points exists for exactly the critical parameter combination.

Example 2.4. Consider an infectious disease that spreads in a population of hosts that are susceptible to this infection and assume that there are two parts of the body (the same two parts for all individuals) that can become infected. We shall assume that one of these two parts, let us name it part one, is necessarily the part where an individual's first infection occurs. Once infected at part one, the infection can spread on by endogenous transmission to part two. We shall use the following notation and assumptions:

(i) β_1 denotes the rate at which one individual that is infected at part one infects a susceptible individual, β_{12} the rate at which one individual that is infected at both parts infects a susceptible individual,

(ii) α denotes the rate of endogenous transmission of individual's infection from part one to part two,

(iii) infected individuals become infectious at the moment of infection,

(iv) infected individuals retain their infection(s) until death,

(v) the death rate is the same for all individuals and is denoted by μ,

(vi) the population birth rate is denoted by λ,

(vii) all newborns are susceptible.

Now let S denote the number of susceptible individuals, I_1 the number of those infected at part one and I_{12} the number of individuals infected at both parts. If the sizes of all subpopulations are large, we can write the following system of differential equations to describe the dynamics,

$$\begin{aligned} \frac{dS}{dt} &= \lambda - \beta_1 I_1 S - \beta_{12} I_{12} S - \mu S \\ \frac{dI_1}{dt} &= \beta_1 I_1 S + \beta_{12} I_{12} S - (\alpha + \mu) I_1 \\ \frac{dI_{12}}{dt} &= \alpha I_1 - \mu I_{12}. \end{aligned} \tag{2.31}$$

Put into our previous notation we have

$$y = (I_1, I_{12}) \in \mathbb{R}^2_+, \quad z = S \in \mathbb{R}_+$$

and the disease free steady state is $e = (0, 0, \frac{\lambda}{\mu})$.
Now, we have only one *i*- state at birth in this case - all individuals are born (from an epidemiological point of view) by acquiring the infection at part one. Each individual that is infected at part one is expected to retain (only) this infection for time $1/(\mu + \alpha)$. In this time it is expected to infect $\beta_1 \lambda/\mu$ individuals (assuming, of course, that it is introduced into an otherwise infection free environment). With probability $\alpha/(\alpha + \mu)$ the individual also becomes infected at part two. It is then expected to remain as such for time $1/\mu$ and in that time infects on average $\beta_{12}\lambda/\mu$ susceptibles. The basic reproduction ratio (i.e., the expected number of new infections caused by an infected individual that is introduced into a completely susceptible population, in all of its infectious period) hence equals

$$\mathcal{R}_0 = \frac{\lambda \beta_1}{\mu(\mu + \alpha)} + \frac{\lambda \alpha \beta_{12}}{\mu^2(\mu + \alpha)}.$$

The elaboration of the direction of the bifurcation can be further simplified in this case if we assume that the total population size has reached an equilibrium. The size of the whole population is then λ/μ and we can eliminate one of the equations in (2.31). By choosing to eliminate the first and replacing S by $\lambda/\mu - I_1 - I_{12}$ in the other two equations we see that we don't need to compute H_y and H_z in (2.11).
To compute the direction of bifurcation we write

$$G = \begin{bmatrix} \beta_1 S - \alpha - \mu & \beta_{12} S \\ \alpha & -\mu \end{bmatrix}$$

with $S = \frac{\lambda}{\mu} - I_1 - I_{12}$. Then

$$\frac{\partial G_{2i}}{\partial y_j} = 0 \quad \text{for } i, j = 1, 2$$

and

$$\begin{aligned} \frac{\partial G_{11}(e,0)}{\partial y_1} &= -\beta_1, & \frac{\partial G_{11}(e,0)}{\partial y_2} &= -\beta_1 \\ \frac{\partial G_{12}(e,0)}{\partial y_1} &= -\beta_{12}, & \frac{\partial G_{12}(e,0)}{\partial y_2} &= -\beta_{12}. \end{aligned}$$

Hence

$$M = w_1(-2\beta_1 v_1^2 - 2(\beta_1 + \beta_{12})v_1 v_2 - 2\beta_{12} v_2^2).$$

Since the off-diagonal elements of G_0 are strictly positive we can choose w to be strictly positive. Since v can always be chosen to be nonnegative we see that M is negative and the bifurcation is supercritical. That is, control measures with which we will decrease the value of $\mathcal{R}_0$ below one will allow us to eradicate the disease, while the infection will spread further as long as $\mathcal{R}_0$ stays above one.

The attentive reader must have noticed that we have not specified the bifurcation parameter. As explained in Section 2.2 we obtain the same direction of bifurcation for all bifurcation parameters, provided that the assumptions of Theorem 2.1 are satisfied. That they are indeed satisfied in this case can easily be verified and we leave the details to the reader.

Example 2.5. Consider again an infectious agent that spreads in the population of susceptibles and suppose again that there are two different parts of the body (the same for all individuals) at which a susceptible can become infected. These are the additional assumptions and the notation:

(i) susceptibility and infectivity of an individual have independent influences on the rate of transmission. Susceptibility to infection does not change if one is already infected at the other part of the body. That way we can write the rate at which someone who is itself infected at $\mathcal{J} \subseteq \{1, 2\}$, infects someone at part j as $b_j B_{\mathcal{J}}$,

(ii) once infected at one of the parts, individuals can only become infected at the other part by another cross transmission,

(iii) infected individuals become infectious at the moment of infection,

(iv) infected individuals retain their infection(s) until death,

(v) the death rate is the same for all individuals and is denoted by μ,

(vi) the population birth rate is denoted by λ,

(vii) all newborns are susceptible.

Let S denote the number of susceptibles and I_1, I_2, I_{12} the number of infected individuals that carry an infection respectively at the first, second and both parts of the body.

If we assume that the sizes of all subpopulations are large, we can describe the dynamics with the following system of differential equations,

$$\begin{aligned}
\frac{dS}{dt} &= \lambda - \Big(\mu + (b_1+b_2)B_1I_1 + (b_1+b_2)B_2I_2 + (b_1+b_2)B_{12}I_{12}\Big)S,\\
\frac{dI_1}{dt} &= b_1S(B_1I_1 + B_2I_2 + B_{12}I_{12}) - b_2I_1(B_1I_1 + B_2I_2 + B_{12}I_{12}) - \mu I_1,\\
\frac{dI_2}{dt} &= b_2S(B_1I_1 + B_2I_2 + B_{12}I_{12}) - b_1I_2(B_1I_1 + B_2I_2 + B_{12}I_{12}) - \mu I_2,\\
\frac{dI_{12}}{dt} &= (b_2I_1 + b_1I_2)(B_1I_1 + B_2I_2 + B_{12}I_{12}) - \mu I_{12}.
\end{aligned}$$

Put into our previous notation we have

$$y = (I_1, I_2, I_{12}), \quad z = S$$

and the disease free equilibrium is $e = (0, 0, 0, \frac{\lambda}{\mu})$.

There are two *i*- states at birth in this case, i.e., becoming first infected at part one being the first *i*- state and becoming first infected at part two being the second. We label these two birth states with 1 and 2, respectively. The next generation matrix is hence a 2×2 matrix, which, written for the case when an infected individual is introduced into a 'virgin' environment, takes the form

$$R = \frac{\lambda}{\mu^2}\begin{bmatrix} b_1B_1 & b_1B_2 \\ b_2B_1 & b_2B_2 \end{bmatrix}.$$

The basic reproduction ratio $\mathcal{R}_0$ equals its dominant eigenvalue. Since R is a matrix of rank one, $\mathcal{R}_0$ equals its trace,

$$\mathcal{R}_0 = \frac{\lambda}{\mu^2}\Big(b_1B_1 + b_2B_2\Big).$$

We shall again assume that the total population has reached an equilibrium and eliminate S by taking $S = \lambda/\mu - I_1 - I_2 - I_{12}$.
To compute the direction of bifurcation we take

$$G = \begin{bmatrix} b_1B_1S - b_2Y - \mu & b_1B_2S & b_1B_{12}S \\ b_2B_1S & b_2B_2S - b_1Y - \mu & b_2B_{12}S \\ b_2Y & b_1Y & -\mu \end{bmatrix},$$

with $Y = B_1I_1 + B_2I_2 + B_{12}I_{12}$ and $S = \lambda/\mu - I_1 - I_2 - I_{12}$.

As was the case in the previous example, all bifurcation parameters yield the same direction of bifurcation from the disease free steady state. However, we cannot tell the sign of M right away. We hence compute the left and the right zero eigenvectors of G_0 and obtain

$$w = \begin{bmatrix} B_1 \\ B_2 \\ B_{12} \end{bmatrix}, \qquad v = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}.$$

Then $M = M_1 + M_2 + M_3$ with

$$\begin{aligned} M_1 &= B_1(-2(b_2+b_1)b_1^2B_1 - 2(b_1B_2 + b_2B_2 + b_1B_1)b_1b_2 - 2b_1b_2^2B_2), \\ M_2 &= B_2(-2(b_2+b_1)b_2^2B_2 - 2(b_2B_1 + b_2B_2 + b_1B_1)b_1b_2 - 2b_2b_1^2B_1), \\ M_3 &= B_{12}(2b_1^2b_2B_1 + 2(b_1B_1 + b_2B_2)b_1b_2 + 2b_1b_2^2B_2). \end{aligned}$$

In contrast with the previous example, bifurcation from a disease free steady state may not always be supercritical. However, if we don't expect any 'amplification' of individual's infectiousness by multiple infected parts, that is, if we assume that

$$B_{12} \le B_1 + B_2,$$

one would expect a supercritical bifurcation and a simple computation (which we leave to the reader) indeed shows that in such a case $M < 0$. As in the previous case, control measures that suppress $\mathcal{R}_0$ below one are sufficient to eradicate the disease.

2.7 Appendix. On the notion of a species, a population and of a reproductively isolated subpopulation

In the main part of this chapter we have used terms such as *population* and *reproductively isolated subpopulation* in a vague, intuitive way. The aim of this section is to describe these notions in mathematical terms (here we are inspired by an unpublished note of Mats Gyllenberg [57]).

First, the following definition.

Definition 2.1. *A square matrix A is* **reducible** *if there exists a permutation matrix P such that*

$$P^{-1}AP = \begin{bmatrix} A_1 & 0 \\ B & A_2 \end{bmatrix}.$$

A matrix that is not reducible is **irreducible**.

Following [57] *we shall call a matrix A* **decomposable** *if the permutation matrix P can be chosen so that* $B = 0$. *A matrix that is not decomposable is* **indecomposable**.

Now, if an element of the next generation matrix, say $R_{ij}(I)$, is strictly positive for some environmental condition I, then, by definition, individuals with birth state j can have offspring with birth state i in this environment. Or, equivalently, the predecessors of individuals with birth state i may, in the environment I, be individuals with birth state j.

If $R(I)$ is indecomposable for some environmental condition I, then all the i-states at birth are in this environment related by either ancestry or descent and hence belong to one species. On the other hand, if the next generation matrix is decomposable for some environmental condition I then the set of i- states at birth partitions into (at least) two disjoint sets of birth states that are not reproductively connected in I.

We speak of *reproductive isolation* of two sets of i- states at birth (and of reproductive isolation of the corresponding subpopulations) when these two sets are reproductively isolated in any conceivable environment. Two sets of i- states at birth (and the corresponding subpopulations) that are not reproductively isolated are *reproductively connected*.

A *population* is a collection of subpopulations that are reproductively connected and are at the same time the maximal connected collection in the sense that they are reproductively isolated from every subpopulation that is not included in the collection.

We now make these terms more precise and make the following, almost mathematical definition (where 'almost' refers to the unspecified 'conceivable' below).

Definition 2.2. *Consider a finite set* $\mathcal{J}$ *of i- states at birth and let* $R(.)$ *denote the corresponding next generation matrix. We say that*

(i) the set $\mathcal{J}$ *of i- states at birth is* **reproductively connected** *if there exists a conceivable environmental condition I in which the corresponding next generation matrix* $R(I)$ *is indecomposable.*

(ii) If the set $\mathcal{J}$ *of i- states at birth is not reproductively connected, it consists of (at least) two* **reproductively isolated** *subsets of i- states at birth. In other words, if* $\mathcal{J}$ *is not reproductively connected, the matrix* $R(I)$ *is decomposable for every conceivable environmental condition I. Note, however, that, in principle, different environmental conditions may yield a different number of blocks in the next generation matrix.*

(iii) Let $\mathcal{J}_1 \subseteq \mathcal{J}$. *We say that individuals with i- states at birth in* $\mathcal{J}_1$ *form a* **population** *if*

(a) $\mathcal{J}_1$ *is reproductively connected and*

(b) if $\mathcal{J}_1 \subseteq \mathcal{J}_2 \subseteq \mathcal{J}$ *and* $\mathcal{J}_2$ *is reproductively connected then* $\mathcal{J}_1 = \mathcal{J}_2$.

Reproductive isolation is certainly a property that underlies the concept of species, i.e., two different species are reproductively isolated. Reproductive isolation alone, however, is not sufficient to deduce that we actually observe different species. Think of, for example, two groups of individuals that belong to the same species but live on areas that are not connected, say, on two different continents.

Another display of this fact would be the semelparous species, species that reproduce only once in their lives and die afterwards. Suppose that we observe a community whose individuals live for a fixed maximum length of time, say, l years. If individuals survive till the end of l-th year they reproduce and die.

We could characterize individuals by the year of their birth. Instead of doing so, we split the whole community into year classes according to the year of birth (modulo l). For instance, if $l = 2$, we divide the community into two year classes, one consisting of individuals that were born in odd numbered years, and the other of individuals born in even numbered years. Different year classes are reproductively isolated subpopulations of the same species that interact (for example, compete for food, etc.) and we can study whether a missing year class is, after being introduced into the community, able to settle among the existing year classes. The reader can find one example in this spirit in Section 2.6.

Consider now a community, consisting of several species perhaps, whose individuals are characterized by finitely many i- states. We introduce one new population and assume that the set of conceivable i- states of this population is also finite. We find the set of all possible i-states at birth and write the next generation matrix of the combined community, which is for every conceivable environmental condition I of the form

$$R(I) = \begin{bmatrix} R(I)_{\text{new}} & 0 \\ 0 & R(I)_{\text{old}} \end{bmatrix}.$$

Since the resident community might consist of several populations, the matrix $R(I)_{\text{old}}$ may be decomposed further into indecomposable blocks. However, finding a way to deduce the number of reproductively isolated subpopulations from the next generation matrix and recognizing the set of i-states that constitute a population is not our aim here. We therefore refrain from these further decompositions.

Chapter 3

Superinfections can induce evolutionarily stable coexistence of pathogens

3.1 Introduction

Evolution of virulence is an intriguing topic that has been on the minds of scientists for many decades. It was believed for a long time that all pathogens would eventually evolve to be benign to their hosts. The words of R. Dubos (1965) reflect this, in that time widely accepted, idea: "Given enough time, a state of peaceful coexistence eventually becomes established between any host and parasite."

This belief, sometimes termed *conventional wisdom*, was rejected only very recently (the first ideas and models can be found in [2, 45, 46, 69, 70, 71, 79]), when it was realised that it was based on the misconception that natural selection favours what is best for a species as a whole.

If, on the other hand, we take as a starting point that natural selection acts on finer levels, such as for example the parasite's reproductiveness, then some ambiguity arises in the case of microparasites (bacteria, viruses, protozoa and fungi) that reproduce within a host, but need to be transmitted at some stage to another host to keep the reproduction cycle going. These parasites reproduce at two different levels and are subject to natural selection at each of these.

For instance, imagine a simple within-host scenario in which pathogens compete for only one resource (i.e., one particular kind of cells that we shall call 'target' cells), and where the pathogen's reproduction rate is the only trait subject to natural selection. In such a case, the pessimization principle [30, 86] applies:

optimal or, more precisely, continuously stable values of the pathogen's reproduction rate are those that minimize the availability of uninfected target cells within the single host that is being considered.

Imagine, on the other hand, that the spread of an infectious agent in a population of susceptible hosts is described by a simple SI model,

$$\begin{aligned} \frac{dS}{dt} &= b - \beta S I - \delta S \\ \frac{dI}{dt} &= \beta S I - (\alpha + \delta) I, \end{aligned} \tag{3.1}$$

where S and I denote, respectively, the number of susceptible and infected (and also infectious) individuals, b stands for the population birth rate, β is the transmission rate and the parameters α and δ denote, respectively, the disease induced mortality and the death rate related to causes other than the disease. The disease induced death rate α is what a majority of the literature refers to as *virulence* [3, 29, 30, 46, 52, 84, 88, 92, 93].

If we now consider the virulence α as the only evolving trait and assume that infection with pathogens of one trait provides individuals with complete immunity to infections with different traits, then the pessimization principle implies that, at the host population level, the evolutionary winner is the trait that (locally) minimizes the steady state value of the number of susceptible hosts, $\hat{S}$. Since $\hat{S} = \frac{b}{\delta \mathcal{R}_0}$, we can equivalently say that natural selection tends to (locally) maximize the pathogen's basic reproduction ratio $\mathcal{R}_0$ (i.e. the expected number of new infections, caused by one infected individual that is introduced into a disease free environment), which is in this case given by

$$\mathcal{R}_0 = \frac{b}{\delta} \cdot \frac{\beta}{\alpha + \delta} .$$

If b, β and δ are kept constant, then the basic reproduction ratio increases when α decreases and so the pathogens indeed evolve to be avirulent, just as the conventional wisdom predicts (here we exclude the possibility of mutualism, i.e., α cannot take negative values). This seems, however, a very unlikely scenario: virulence as well as transmissibility are in reality likely to be related to within-host characteristics, such as the rate at which pathogens reproduce inside a host or the pathogen load. But while an increased reproduction within a host may enhance transmissibility β, it may also harm the host by increasing α. In other words, there is a trade-off between pathogen production and transmissibility on the one side and virulence on the other.

In the last years several papers and books [1, 4, 22, 29, 46, 47, 52, 67, 71, 84, 88, 89, 92, 93] appeared in which the evolution of virulence was studied,

while taking into account such trade-offs. Most of them (with the exception of [1, 52]), however, kept the within-host characteristics implicit and only assumed that the transmission parameter β is somehow directly related to virulence. In other words, it is assumed that $\beta = \beta(\alpha)$.

In [52], the very paper that inspired this work, both the within-host and the between-host dynamics were made explicit. The authors considered the rate at which pathogens are produced inside a host as the only trait subject to natural selection and determined the continuously stable strategies (CSSs) for selection at the within-host, as well as for selection at the host population level. However, even though the between-host model was related to the dynamics within a host, the authors used a single infection model to describe the between-host dynamics. According to this model, no host can harbour more than one pathogen trait, which directly excludes re-infection induced evolution at the within-host level. The two levels at which natural selection works thus remain separated and, as a consequence, natural selection at these two levels may appear to be in conflict, i.e. the evolutionarily stable trait on one level may differ from the evolutionarily stable trait on the other.

In reality, natural selection doesn't act exclusively at any of these two levels. When random mutations of the pathogen evoke variation and, subsequently, selection within one individual, this also influences the evolutionary dynamics at the host population level (say, by influencing transmission) which, in turn, may have an effect on the within-host dynamics. In order to study the evolution of infectious diseases in a realistic manner, we must therefore consider the two levels as being coupled: we must relate the between-host dynamics to within-host characteristics and furthermore take into account that hosts may harbour more than one pathogen trait, either at one and the same time or consecutively.

Multiple infections, which may result either from additional transmissions or from random mutations of the pathogens inside a host, can be modeled in different ways, depending also on the within-host model used. The particular within-host model we use in this chapter doesn't allow for steady coexistence of different traits. In other words, there may be a period in which several traits are present within a host, but one of them will eventually outcompete the others. *Coinfection* models take the period of coexistence explicitly into account. *Superinfection* models, on the other hand, assume that, if the new trait takes over, it does so immediately. Both ways of modeling the additional infections have advantages as well as downsides. One can rightfully argue that superinfection models are less realistic compared to coinfection models since there will always be a period of time, however short, in which several traits will be present inside a host. But the added realism in coinfection models entails specification of the pathogen level immediately after transmission took place and there is, in general, hardly any biological information that can be translated into such a specification in the context

of caricatural models of within-host dynamics.

In this chapter we model multiple infections as superinfections. We take the SI model underlying (3.1) as the basis for the description of the spread of an infectious agent in a population of hosts. We relate the transmissibility β and the disease induced death rate α to within-host dynamics. By assuming that the dynamics within a host is fast compared to the dynamics at the host population level, β and α will actually depend on the steady state values of the within-host variables. Moreover, this difference in time scales of within- and between-host processes implies that when more than one trait is present inside a host, the best within-host competitor immediately eliminates all other traits. Multiple infections thus indeed manifest themselves as superinfections.

We assume throughout that mutations are rare on the time scale of transmission and demography. We thus adopt the Adaptive Dynamics point of view and use some standard terminology of this field throughout the chapter. Readers that are not familiar with this terminology are advised to consult Boxes 3.1 and 3.2, where we provide the definitions.

The outline of the chapter is as follows. In Section 3.2 we present the two basic ingredients of the main superinfection model in Section 3.3, i.e., the within-host and the single infection between-host model. We single out the pathogen's within-host reproduction rate p as the only evolving trait and characterize the continuously stable values of p for the evolution at each of the two isolated levels.

In Section 3.3 we show that the singular strategies obtained at the two isolated levels can be seen as two extreme ends of the possible evolutionary outcomes of the combined selection. The precise value of a singular strategy, however, will depend on the way superinfections are incorporated into the model. Following [84], we introduce a *superinfection function* $\phi(p, q)$, describing the ability of pathogens with trait q to 'take over' a host that is already infected by pathogens with trait p, and show that when the function $q \mapsto \phi(p, q)$

(i) has a jump discontinuity in $q = p$, then the continuously stable strategies coincide with the CSSs at the within-host level,

(ii) is continuous, but not differentiable, in $q = p$, then both levels contribute to the value of a singular strategy. In particular, when the CSSs at the two isolated levels are unique, the convergence stable strategy lies in-between the within-host CSS and the CSS of the single infection model.

(iii) is differentiable in $q = p$, then the singular strategies are the same as the ones given by the single infection model,

For the ease of formulation we sometimes use the terms **jump**, **mechanistic**, and **smooth** to describe, respectively, case (i), (ii) and (iii). This terminology is

justified in Section 3.6, where we model in more detail the initial stages of the introduction of a mutant trait in a single infected host. Since the mutant trait is then likely to be present only in small quantities, we model the invasion as a stochastic birth-and-death process. We find that the smoothness of the superinfection function relates to the ability of the mutant trait to survive the initial phase of low abundance in an already infected host, which itself depends on the starting dose of the mutant.

In Section 3.4 we observe that singular strategies can no longer be characterized by way of an optimization principle when superinfections are taken into account. Among other things, this implies that the superinfection model allows for coexistence of pathogen traits at the host population level. While epidemiological coexistence (namely, coexistence on the time scale of transmission and demography) comes as no surprise (in fact, many previous papers in which superinfection models were studied [60, 84, 88, 92, 93, 102] have also encountered such coexistence), the fact that coexistence can be maintained on the evolutionary time scale is a bit surprising (cf. [93] and Section 3.7) and has, to our knowledge, in the context of a superinfection model only been found in [88].

The results on evolutionary coexistence are collected in Section 3.5, where we furthermore demonstrate that the existence of branching points (which can lead to evolutionary coexistence), is promoted by a high transmission dose. The branching points we encounter in numerical experiments are of an unusual, asymmetric type.

Even though the models we use are very caricatural, the results of this chapter highlight the following important point: in order to understand the evolution of infectious diseases one needs to form a sound understanding of how the host mediates interactions between slightly different strains that infect the host, either simultaneously or consecutively. Indeed, the (non)smoothness of the superinfection function has a huge impact on the outcome of evolution and hence the details matter!

Some concluding remarks are collected in Section 3.7.

Admittedly, we have chosen a very particular setting, namely, an SI model which assumes that individuals retain the infection until death, but that, perhaps due to reinfections or random mutations that take place inside the host, the pathogens' trait may change in the course of the infection. Although this shortens significantly the list of infectious diseases that are covered, we believe that this setting is worth studying for two reasons. Firstly, the assumptions made here provide an acceptable caricature for many relevant pathogens such as HIV, Hepatitis B and Hepatitis C virus. Secondly, the ideas and the framework presented in this chapter can be extended to study evolutionary dynamics in other settings that incorporate different dynamics at the within-host, as well as at the host population level (SEI, SIR, etc. models, different incidence rates etc.).

3.2 Preliminaries

The aim of this section is to introduce two models, the within-host and the single infection between-host model, that will constitute the building blocks of the main model in the next section.

3.2.1 The within-host dynamics

To describe the dynamics within one host we use the following system of ODEs,

$$\frac{dT}{dt} = \lambda - kVT - dT \tag{3.2a}$$

$$\frac{dT^*}{dt} = kVT - (\mu(p) + d)T^* \tag{3.2b}$$

$$\frac{dV}{dt} = pT^* - kVT - cV. \tag{3.2c}$$

The three variables in (3.2), T, T^* and V represent, respectively, the number of uninfected and infected target cells and the number of free pathogens. The system (3.2) corresponds to the following scenario.

In the absence of the infectious agent, target cells are produced at a constant rate λ and die at a constant per capita rate d. When the host is infected, free pathogens inside a host die at per capita rate c. The mass action term kVT is used to model the process of infection within a host: it says that the rate at which pathogens find uninfected target cells, successfully bind to the surface of the cell and/or enter the target cell, is proportional to the product of the numbers of uninfected target cells and free pathogens. Upon infection, the uninfected target cell and the pathogen that infected it, form an infected cell. Hence, the term kVT is substracted from (3.2a) and (3.2c) and added in (3.2b). Infected target cells produce free pathogens at a rate p. This production comes at a cost, namely, it increases the death rate of infected target cells by $\mu(p)$. We shall assume that $\mu \in C^2(\mathbb{R})$ and that it is a nonnegative, increasing function of the production rate p.

Remark 3.1. According to (3.2), a free pathogen infects individual cells of the host. The model is thus certainly a meaningful description for viruses (that, after binding to a cell, enter, or at least inject their genetic material, into the host cell [89]), but not so for bacteria and fungi that live in the interstitial fluid.

Remark 3.2. System (3.2) is one of the models commonly used in the literature to describe the virus dynamics within a host [25, 91]. According to (3.2), a free pathogen disappears as a free particle at the moment it infects a target cell. A variation of this model, which is also used quite often [52, 89], neglects the term

$-kVT$ in (3.2c). The reasoning behind this omission is that the number of pathogens is much higher than the number of target cells and while the term $-kVT$ is relevant in (3.2a), it is negligible (compared to other terms) in (3.2c). While this simplified model was shown to yield a very good fit for, for instance, some stages of HIV progression, it may not do so in the very first stages of an infection, when the pathogens are rare. Since we are in this chapter mainly interested in the ability of mutant traits to invade the resident pathogen population (and so we focus precisely on the initial stages of the introduction of a mutant trait) we shall work with the full model (3.2). Note, however, that omission of the term $-kVT$ in (3.2c) has one important modeling consequence. Namely, if we leave out the term $-kVT$ in (3.2c), the free pathogen is not 'lost' upon infecting a target cell and can therefore infect more than one cell. This observation will be of great importance in Section 3.6, where we shall describe the initial stages of a reinfection as a stochastic birth-and-death process.

System (3.2) has two equilibria: the infection free steady state in which there are no pathogens and no infected target cells,

$$\hat{V} = \hat{T}^* = 0 \quad \text{and} \quad \hat{T} = \frac{\lambda}{d} =: T_0 \tag{3.3}$$

and a nontrivial equilibrium given by

$$\hat{T} = \frac{c}{k(\mathcal{B}_0(p) - 1)} \tag{3.4a}$$

$$\hat{T}^* = \frac{\mathcal{B}_0(p)}{p}\Big(\lambda - \frac{cd}{k(\mathcal{B}_0(p) - 1)}\Big) \tag{3.4b}$$

$$\hat{V} = \frac{\lambda}{c}\Big(\mathcal{B}_0(p) - 1\Big) - \frac{d}{k}. \tag{3.4c}$$

Here, $\mathcal{B}_0$ stands for the so called *burst size*, i.e. the expected number of pathogens produced by one infected target cell. If the pathogen production rate of the infected target cell equals p, then

$$\mathcal{B}_0(p) = \frac{p}{\mu(p) + d}.$$

The nontrivial steady state given by (3.4) is biologically meaningful only when all three components in (3.4) are positive. The first, $\hat{T}$, is positive when the burst size exceeds one. If we then rewrite (3.4b) (and (3.4c)), we find that the other two components are strictly positive only when the pathogen's within-host reproduction ratio is larger than one. The within-host basic reproduction ratio of a pathogen, $\mathcal{R}_0^w$ (the superscript w serves to distinguish it from the pathogen's basic

reproduction ratio at the host population level), is defined as the expected number of new pathogens produced by a single pathogen introduced into a virgin cell environment. Since free pathogens need to enter uninfected target cells in order to reproduce and since the probability with which the pathogen enters a target cell in a virgin environment equals $\frac{k\lambda}{k\lambda+dc}$, the within-host basic reproduction ratio of a pathogen with trait p equals

$$\mathcal{R}_0^w(p) = \frac{k\lambda}{k\lambda + dc}\mathcal{B}_0(p).$$

The nontrivial equilibrium is thus biologically meaningful only when $\mathcal{R}_0^w(p) > 1$. When it exists, it is also locally asymptotically stable, while the infection free steady state is unstable in that case (see [25] for a global stability result and also [85] for local stability of variants of (3.2)).

We shall throughout the chapter consider the rate of pathogen production p as the only trait subject to natural selection. All the other parameters in the within-host model will be kept constant throughout. Furthermore, we shall sometimes omit the word 'pathogen' and simply describe hosts or cells as being 'infected by a certain trait'.

We assume for simplicity that, if a target cell is infected with pathogens of one trait, it is protected from further infections. In other words, we do not consider superinfections or coinfections at the cell level. Since this assumption implies that the pathogens compete within a host for only one resource, viz. uninfected target cells, the evolutionary dynamics at the within-host level is very simple. Namely, when a mutant trait, say q, is introduced into a host where the trait p is resident, the mutant is successful (according to the deterministic model) if and only if it exploits the resource better than the resident, i.e. when $\hat{T}(q) < \hat{T}(p)$. This is sometimes called the *pessimization principle*: when the environment the pathogens experience is one dimensional and an invasion results in competitive exclusion of one of the traits, then natural selection necessarily leads to the worst possible environment [30, 50, 86] (note, incidentally, that minimization of $\hat{T}$ is equivalent to maximization of $\mathcal{B}_0$ and also to maximization of $\mathcal{R}_0^w$).

The precise conclusions, however, will depend on the trade-off function $\mu(p)$. In the following two examples we present the two most commonly used trade-off relations.

Example 3.1. Let μ be a concave function of the form $\mu(p) = \frac{ap}{p+b}$. This makes $\hat{T}(p)$ a strictly decreasing function. The requirement $\mathcal{R}_0^w > 1$, which can be rewritten as $\hat{T}(p) < T_0$ (see (3.3) and (3.4a)), hence gives a lower bound for the admissible values of p, denoted by $p_{\min}$, but there exists no upper bound for p (see Figure 3.1). In this case the minimum of $\hat{T}(p)$ is not reached for finite p. In practice, however, it is likely that there are physiological constraints on the pathogen

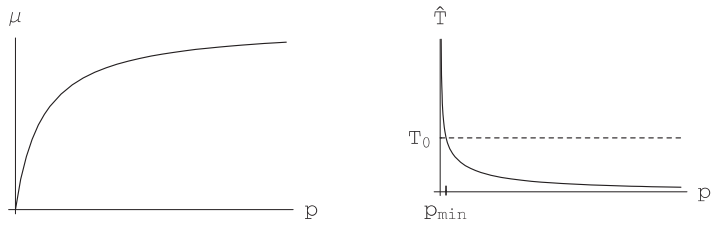

Figure 3.1: An example of a concave trade-off function $\mu(p)$ and the corresponding $\hat{T}(p)$.

production rate p. In such a case then, natural selection will drive p towards the physiological maximum, $p_{\max}$.

Example 3.2. Let us now take $\mu(p) + d = de^{ap}$. Requiring that $0 < \hat{T}(p) < T_0$, we find a lower, as well as an upper bound, for feasible values of the pathogen production rate (see Figure 3.2). The minimal value of $\hat{T}(p)$ is now obtained for some intermediate value $p^* \in (p_{\min}, p_{\max})$.

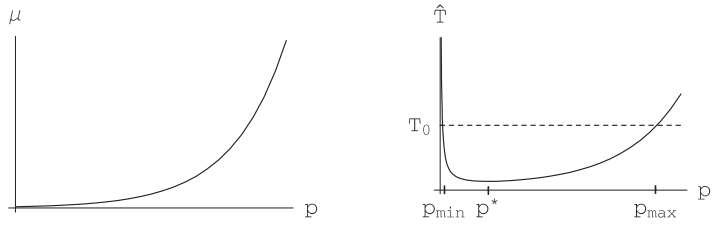

Figure 3.2: An example of a convex trade-off function $\mu(p)$ and the corresponding $\hat{T}(p)$.

In general, of course, the trade-off function μ may be more complicated and there may be more than one CSS (cf. Box 3.1 for the definition of a CSS) for the within-host selection. We shall in any case assume that

- there exists a finite interval $[p_{\min}, p_{\max}]$ of feasible values of p,
- the function $\hat{T}(p)$ is nowhere locally constant and
- there are no accumulation points of the extrema of $\hat{T}(p)$.

If we denote the set of all within-host CSSs by C^*, then $p^* \in C^*$ is either one of the boundary points, $p^* \in \{p_{\min}, p_{\max}\}$, or a point in $(p_{\min}, p_{\max})$ that, in a

Box 3.1. Some basic notions of Adaptive Dynamics

The **invasion exponent** $\sigma_p(q)$ is defined as the growth rate of a mutant population with trait q in the environment set by the resident population with trait p.
If the invasion exponent is differentiable as a function of q in the point $q = p$, then the sign of the **selection gradient**

$$\frac{\partial \sigma_p(q)}{\partial q}\Big|_{q=p}$$

determines the direction of evolution from the resident trait p. If it is positive, the trait will (at least locally) increase in the course of evolution and if it is negative, the trait will (locally) decrease in the course of evolution. **Singular strategies** are trait values in which the selection gradient vanishes, i.e. when

$$\frac{\partial \sigma_p(q)}{\partial q}\Big|_{q=p} = 0.$$

A singular trait $\overline{p}$ is called **convergence stable**, if a nearby strategy can be invaded (only) by traits that are nearer to $\overline{p}$. That is, if $p < \overline{p}$, then $s_p(q) > 0$ for $p < q < \overline{p}$, while for $p > \overline{p}$ the invasion exponent is positive when $\overline{p} < q < p$. Convergence stable strategies are thus (local) attractors for monomorphic evolutionary dynamics.
An **evolutionarily stable strategy** (**ESS**) is a strategy that cannot be invaded by neighbouring traits. That is, $\overline{p}$ is an ESS if $s_{\overline{p}}(q) < 0$ for $q \in (\overline{p} - \varepsilon, \overline{p} + \varepsilon)$ with some $\varepsilon > 0$. Despite the enticing 'stable' in its name, an ESS may not be an evolutionary attractor. If it is, it is called a **continuously stable strategy** (**CSS**).
An **evolutionary branching point** is a singular strategy that is convergent stable, but not an ESS.
If the invasion exponent is differentiable twice, then the second partial derivatives allow us to classify the singular points. In particular, if

$$\frac{\partial^2 \sigma_p(q)}{\partial q^2}\Big|_{q=p} < 0,$$

the point $q = p$ is an ESS. If

$$\frac{\partial^2 \sigma_p(q)}{\partial q^2}\Big|_{q=p} < \frac{\partial^2 \sigma_p(q)}{\partial p^2}\Big|_{q=p},$$

then p is convergence stable. A singular point for which

$$\frac{\partial^2 \sigma_p(q)}{\partial p^2}\Big|_{q=p} > \frac{\partial^2 \sigma_p(q)}{\partial q^2}\Big|_{q=p} > 0$$

is a branching point.

generic case, satisfies

$$\left.\frac{d\hat{T}(q)}{dq}\right|_{q=p^*} = 0 \quad \text{and} \quad \left.\frac{d^2\hat{T}(q)}{dq^2}\right|_{q=p^*} > 0.$$

Note that we have assumed that $\hat{T}(p)$ is differentiable twice and so we can indeed compute and characterize singular points for the within-host evolution by way of differentiation.

3.2.2 The single infection between-host model

When studying the spread of an infectious agent at the host population level, one can make several different assumptions concerning the way in which traits interact within one host. The *single infection model* [52], for instance, assumes complete cross-immunity between different traits. That is, hosts infected by some trait p are completely protected from further infections by other traits. In our opinion, this assumption is inconsistent with acknowledging that within-host selection may take place. Indeed, if an internally produced mutant may take over, then why would that be impossible for a competitor introduced from outside?

We shall nevertheless now formulate this single infection model as we shall need it (or, more precisely, the singular traits it admits) in the next section.

We make the following assumptions regarding the dynamics at the host population level. In a disease free environment, individuals are born at a rate b and die at per capita rate δ. Individuals become infectious at the moment they are infected and retain the infection until death, which now occurs at an increased per capita rate $\alpha+\delta$. Transmission occurs according to mass action with proportionality constant β, which incorporates both the rate at which a susceptible individual comes into contact with an infectious host and the probability with which a contact results in a transmission.

We shall assume that the parameters b and δ, which capture the population dynamics in the absence of the disease, are constant. The disease induced mortality α and the transmission parameter β, however, will depend on the infection status of the host. For instance, one can imagine that hosts that carry more pathogens may be more infectious than the ones in which the amount of pathogens is lower. Similarly, α may depend on the parasite abundance in the host but perhaps also on the amount of uninfected target cells (which seems likely at least in the case of HIV, where the virus attacks the very cells that should serve to protect the host). We shall therefore assume that α and β depend on the production rate p, but that this dependence is expressed mechanistically in terms of the within-host variables, T, T^* and V. Since we have furthermore assumed that the dynamics

within a host is fast compared to the dynamics at the population level, we assume that α and β are in fact functions of the steady state values $\hat{T}, \hat{T}^*$ and $\hat{V}$, say,

$$\alpha(p) = A(\hat{T}(p), \hat{T}^*(p), \hat{V}(p)), \qquad \beta(p) = B(\hat{T}(p), \hat{T}^*(p), \hat{V}(p))$$

and assume that $A, B \in C^2(\mathbb{R}^3, \mathbb{R})$.

If we now assume that two traits, say, p and q, circulate in the population, then the above assumptions give rise to the following system

$$\begin{aligned}
\frac{dS}{dt} &= b - \beta(p)S\,I_p - \beta(q)S\,I_q - \delta S \\
\frac{dI_p}{dt} &= \beta(p)S\,I_p - (\alpha(p) + \delta)I_p \\
\frac{dI_q}{dt} &= \beta(q)S\,I_q - (\alpha(q) + \delta)I_q,
\end{aligned}$$

where S denotes the number of susceptible individuals and I_p, I_q denote the number of individuals infected by, respectively, trait p and trait q. Note that there are, due to complete cross-immunity between the traits, no doubly infected individuals.

In the absence of trait q, the system admits two equilibria: the disease free steady state

$$\hat{S}(p) = \frac{b}{\delta} =: S_0, \quad \hat{I}_p(p) = 0,$$

and the nontrivial equilibrium, which is biologically meaningful only when

$$\mathcal{R}_0(p) = \frac{b}{\delta} \cdot \frac{\beta(p)}{\alpha(p) + \delta} > 1.$$

When the endemic steady state exists, it is also a global attractor (cf. [35], Exercise 3.11).

Since selection at the host population level occurs in this case solely through the subpopulation of susceptible hosts, we already know that the pessimization principle will apply and that the basic reproduction ratio $\mathcal{R}_0(p)$ will be (locally) maximized in the course of evolution. However, with our minds already on the following section, we now reformulate the CSS conditions in terms of the *invasion exponent*, $s_p(q)$, which denotes the growth rate of the subpopulation of hosts infected by trait q, when introduced into a steady resident population in which (only) trait p is present. In this case

$$s_p(q) = \beta(q)\hat{S}(p) - (\alpha(q) + \delta), \tag{3.5}$$

with

$$\hat{S}(p) = \frac{\alpha(p) + \delta}{\beta(p)}. \tag{3.6}$$

The number of local maxima of $\mathcal{R}_0$ will depend on the choice of trade-off functions α and β and the value of δ. These choices will, along with the condition $0 < \hat{S}(p) < S_0$, pose restrictions on the values of p that guarantee persistence at the host population level. Taking these, along with similar restrictions at the within-host level (see the previous subsection) into account, we obtain the set of feasible values of the pathogen production rate p.

If we denote by $C^\bullet$ the set of CSS values for selection at the between-host level (in the context of the single infection model), then $p^\bullet \in C^\bullet$ either lies on the boundary of the domain, or is a point in the interior, which in a generic case satisfies

$$\left.\frac{\partial s_{p^\bullet}(q)}{\partial q}\right|_{q=p^\bullet} = 0 \quad \text{and} \quad \left.\frac{\partial^2 s_{p^\bullet}(q)}{\partial q^2}\right|_{q=p^\bullet} < 0.$$

3.3 The superinfection model

If infected individuals are not fully protected from colonization by novel traits, there are several ways in which we can incorporate multiple infections into the between-host model. As already written in the introduction, we treat all reinfections as superinfections by assuming that the within-host dynamics is fast compared to population dynamics. Since the within-host model we use does not allow for coexistence of different traits, this means that when a mutant trait is introduced (in one host) into a monomorphic pathogen population, the invasion is either unsuccessful, or it results in an immediate trait substitution within the host. The single infection model is therefore replaced by the following, so called *superinfection model*,

$$\begin{aligned}
\frac{dS}{dt} &= b - \beta(p)S I_p - \beta(q)S I_q - \delta S \\
\frac{dI_p}{dt} &= \beta(p)S I_p - (\alpha(p) + \delta)I_p + \Phi(q,p)I_p I_q \\
\frac{dI_q}{dt} &= \beta(q)S I_q - (\alpha(q) + \delta)I_q + \Phi(p,q)I_p I_q,
\end{aligned} \tag{3.7}$$

where

$$\Phi(p,q) = \beta(q)\phi(p,q) - \beta(p)\phi(q,p) \tag{3.8}$$

and where the **superinfection function** $\phi(p,q)$ describes the ability of the pathogens with trait q to 'take over' a host that is already infected by trait p. More

precisely, we define

$\phi(p,q)$:= *the probability with which the trait q, upon transmission to a host already infected by trait p, eliminates p and thus takes over the host.*

Since the ability of the invading trait q to grow in a host that is already infected by trait p is completely determined by $\hat{T}(p)$ and $\hat{T}(q)$, we shall actually deal only with functions $\phi(p,q)$ that can be written as functions of $\hat{T}(p)$ and $\hat{T}(q)$, say,

$$\phi(p,q) = \psi(\hat{T}(p), \hat{T}(q)). \tag{3.9}$$

Suppose now that the newly introduced trait q does worse at the within-host level than the resident trait p, i.e. $\hat{T}(p) < \hat{T}(q)$. Trait q is thus unable to reinfect a host infected by p. And since also the traits that reduce $\hat{T}$ to exactly the same level as the resident cannot grow (cf. Section 3.6), we shall assume that ψ is a nonnegative function that satisfies

H$_1$. $\psi(x,y) = 0$ whenever $y \geq x$.

To determine whether a trait q will grow when introduced into the resident population of hosts with trait p, we compute the invasion exponent, which we shall now denote by $r_p(q)$, and which in this case equals

$$\begin{aligned} r_p(q) &= \beta(q)\hat{S}(p) - (\alpha(q)+\delta) + \Phi(p,q)\hat{I}(p) \\ &= s_p(q) + \hat{I}(p)(\beta(q)\psi(\hat{T}(p),\hat{T}(q)) - \beta(p)\psi(\hat{T}(q),\hat{T}(p))), \end{aligned} \tag{3.10}$$

where s denotes the invasion exponent for the single infection model (given in (3.5) and (3.6)) and

$$\hat{I}(p) = \frac{b}{\alpha(p)+\delta} - \frac{\delta}{\beta(p)} = \frac{b}{\beta(p)\hat{S}(p)} - \frac{\delta}{\beta(p)}$$

denotes the equilibrium value of I in the absence of trait q.

The evolutionary dynamics of p and, in particular, the corresponding singular strategies, depend heavily on the behaviour of $\psi(x,y)$ when $y \uparrow x$. We shall consider the following three possibilities for ψ as a function of y: (i) the '**jump**' case, i.e. when $y \mapsto \psi(x,y)$ has a jump discontinuity in the point $y = x$, (ii) the '**mechanistic**' case, i.e. when $y \mapsto \psi(x,y)$ is continuous, but not differentiable, in $y = x$ and (iii) the '**smooth**' case when $y \mapsto \psi(x,y)$ is differentiable in $y = x$ (cf. Figure 3.3).

The remaining part of this section is devoted to showing that these three classes of superinfection functions determine three very different ways in which natural selection can work out. We shall not, however, be concerned at this point with details on how the behaviour of $\psi(x,y)$ for $y \uparrow x$ reflects the way in which

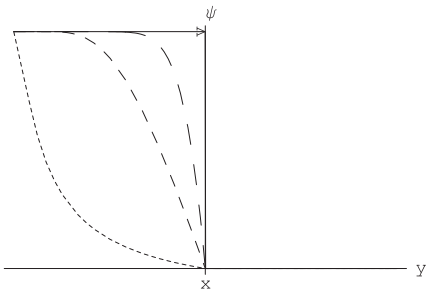

Figure 3.3: For every function ψ we have $\psi(x,y) = 0$ whenever $y \geq x$. For $y < x$, the solid line represents the 'jump' case, the two dashed lines in the middle represent the 'mechanistic' case and the bottom dashed line represents the 'smooth' case.

interactions of different traits within one host are modeled. This will be our task in Section 3.6. We furthermore emphasize that the analysis of this section is local analysis, based on the assumption that mutations are not only rare on the time scale of transmission and demography, but also 'sufficiently' small. In other words, we assume that the mutant trait q is 'sufficiently' close to the resident trait p. Some observations regarding the global evolutionary dynamics are collected in Sections 3.4 and 3.5.

So let us denote by $C^{\bullet *}$ the set of continuously stable strategies of the superinfection model and begin with case (i).

3.3.1 The 'jump' case

As an example of a discontinuous superinfection function, one may have the following in mind.

Example 3.3. Let

$$\phi(p,q) = \begin{cases} 1, & \hat{T}(q) < \hat{T}(p) \\ 0, & \text{otherwise} \end{cases} \tag{3.11}$$

In this case, superinfections occur in accordance with the deterministic description: the invading trait q is successful if and only if it is able to win the internal competition with the resident trait p. In particular, ϕ does not distinguish between winning strategies: traits that are only slightly better at the within-host level have the same advantage over the resident than significantly better within-host strategies.

So let us assume that ψ is a nonnegative function that satisfies $\mathbf{H}_1$ and has a jump discontinuity in (x,x). Because of this discontinuity, we cannot find and

characterize singular strategies by way of differentiation. Instead, a simple observation of orders of magnitude will show that (i) singular strategies coincide with the ones given by the within-host model and (ii) their 'character' is the same as in the context of the within-host model. In particular, the continuously stable strategies coincide with the CSSs of the within-host model. In symbols, $C^{\bullet *} = C^*$.

To see that this is indeed the case, we first consider $p^* \in C^*$. Since the within-host CSS p^* (locally) minimizes $\hat{T}(p)$, we can find a neighbourhood of p^*, say U, such that for $q \in U \setminus \{p^*\}$ we have $\hat{T}(q) > \hat{T}(p^*)$. So if $q \in U$ then

$$r_{p^*}(q) = s_{p^*}(q) - \beta(p^*)\hat{I}(p^*)\phi(q, p^*).$$

Since $s_p(q)$ is differentiable as a function of q in the point $q = p$ and $s_p(p) = 0$, we have $s_{p^*}(q) = O(\varepsilon)$ for $q = p^* + O(\varepsilon)$. The term $\beta(p^*)\hat{I}(p^*)\phi(q, p^*)$, on the other hand, is $O(1)$ and so $r_{p^*}(q) < 0$ when $q = p^* + O(\varepsilon) \in U$. That is, p^* is an ESS.

To see that p^* is also an evolutionary attractor we first take $\overline{p} < p^*$ such that $\hat{T}'(q) \leq 0$ for $q \in [\overline{p}, p^*]$. For $p \in (\overline{p}, p^*)$ and $\varepsilon > 0$ small enough so that $\overline{p} < q_- = p - \varepsilon$ we have $\hat{T}(q_-) > \hat{T}(p)$. Similarly, for $\varepsilon > 0$ such that $q_+ = p + \varepsilon < p^*$ we have $\hat{T}(p) > \hat{T}(q_+)$. Hence, traits that move closer to p^* do better at the within-host level than the residents, while the traits that move away from p^* do worse. We therefore have

$$\begin{aligned} r_p(q_+) &= s_p(q_+) + \beta(q_+)\hat{I}(p)\phi(p, q_+), \\ r_p(q_-) &= s_p(q_-) - \beta(p)\hat{I}(p)\phi(q_-, p) \end{aligned}$$

and with the same comparison of orders of magnitude of the two terms on the RHS we find that p^* is locally attracting from the left.

In a similar fashion we find that p^* is (locally) attracting from the right and so p^* is indeed a CSS. Furthermore, these last arguments also show that (i) there can be no evolutionarily stable strategies but the elements of C^*: if $p \notin C^*$ then $r_p(q)$ will be positive, for ε small enough, on at least one of the intervals $(p-\varepsilon, p)$ or $(p, p + \varepsilon)$ and (ii) there are no other singular strategies but the extrema of $\hat{T}$. Moreover, in the case when $C^{\bullet *}$ contains more than one point, it is the traits that (locally) maximize $\hat{T}$ that separate the domains of attraction, exactly as is the case for within-host evolution.

We conclude this subsection with two examples.

Example 3.4. Let us take $\lambda = 1, k = 10, c = 0.1, d = 1$ in the within-host model and $\delta = 0.1, b = 2$ for the between-host model. We furthermore choose

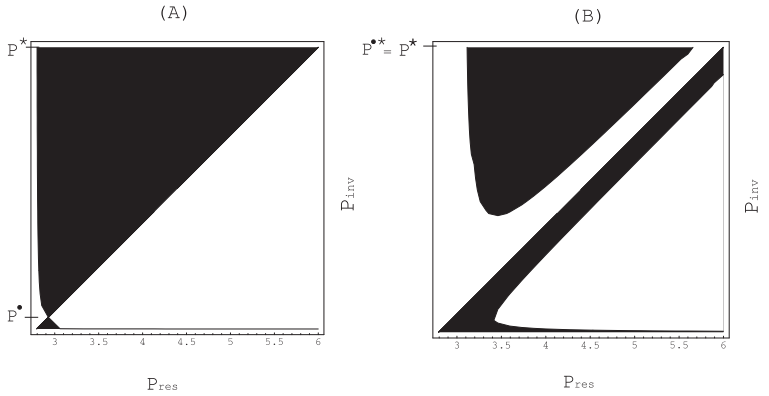

Figure 3.4: (A) PIP for the single infection model, along with the within-host CSS, $p^* = p_{\max}$. (B) PIP for the superinfection model with ϕ as in (3.11). The white and the black areas correspond to the regions in $(p_{\text{res}}, p_{\text{inv}})$ space that yield, respectively, positive and negative values of the invasion exponent.

$\mu(p) = \frac{p}{p+1}$ and

$$\alpha(p) = \tfrac{k}{\lambda}\hat{V}(p),$$
$$\beta(p) = \hat{V}(p) + \hat{T}^*(p).$$

In this case, $\hat{T}(p)$ is a strictly decreasing function. Taking into account the restriction $0 < \hat{T}(p) < T_0$, we therefore find a lower bound, $p_{\min}$, for the admissible trait values at the within-host level, but the model gives no upper bound on p (see also Example 3.1). We assume, however, that there is a physiological restriction $p_{\max} = 6$ on the values of the pathogen production rate. The condition $0 < \hat{S}(p) < S_0$ yields a lower bound $p'_{\min}$, but no upper bound. Since in this case $p_{\min} < p'_{\min} < p_{\max}$, we plot the pairwise invasibility plots on $[p'_{\min}, p_{\max}] \times [p'_{\min}, p_{\max}]$. Figure 3.4 shows two pairwise invasibility plots (cf. Box 3.2 on how PIPs are constructed): in (A) we see the PIP for the single infection model, while (B) shows the PIP for the superinfection model, where the superinfection function is the discontinuous function given in (3.11). The white and the black areas correspond to the regions in (p, q) space that yield, respectively, positive and negative values of the invasion exponent. The single infection model in this case yields a unique CSS, $p^\bullet \in (p'_{\min}, p_{\max})$, while the evolution at the within-host level drives p towards the physiological maximum, $p^* = p_{\max}$. We see that, when superinfections are modeled with a discontinuous superinfection function, p^* is indeed the only CSS at the host population level, i.e. $p^{\bullet *} = p^*$.

Box 3.2. Pairwise invasibility plots

Pairwise invasibility plots (**PIPs**) are a handy way of representing graphically the ability of a mutant trait to grow in the resident community. A PIP is constructed by plotting the sign of the invasion exponent $\sigma_p(q)$ for all feasible pairs (p, q) of (resident, mutant) trait values.

If the resident population is at a stable equilibrium, then $\sigma_p(p) = 0$ and so the zero contour lines contain at least the main diagonal. The shapes of other zero contour lines, if there are any, contain important information about the course of evolution. In particular, singular points are found as intersections of zero contour lines with the main diagonal. If we now imagine that black and white regions in the PIP represent the regions where the invasion exponent is, respectively, negative and positive, then the 'character' of a singular strategy can easily be recognized from a pairwise invasibility plot. Namely, if $\overline{p}$ is to be an ESS, and hence uninvadable by the neighbouring strategies, the straight vertical line through $(\overline{p}, \overline{p})$ must lie, at least locally, in the region where the invasion exponent is negative, i.e. in the black region. The singular trait $\overline{p}$ is convergence stable when the regions left of $(\overline{p}, \overline{p})$ are, at least close to the diagonal, white above the diagonal and black below the diagonal (i.e. $\overline{p}$ is locally attracting from the left), while the regions right of the point $(\overline{p}, \overline{p})$ are (at least close to the diagonal) black above the diagonal and white below the diagonal (in other words, $\overline{p}$ is locally attracting from the right).

Example 3.5. We now take the same parameter values and trade-off functions α and β as in the previous example, only now $\mu(p) + d = de^{0.2p}$. Taking into account the restrictions $0 < \hat{T}(p) < T_0$ and $0 < \hat{S}(p) < S_0$, we first find the intervals of feasible values of p, $[p_{\min}, p_{\max}]$ and $[p'_{\min}, p'_{\max}]$, respectively. For the chosen trade-off functions μ, α and β and the chosen parameter values we find that $p_{\min} < p'_{\min}$ and $p'_{\max} < p_{\max}$ and so we plot the pairwise invasibility plots on $[p'_{\min}, p'_{\max}] \times [p'_{\min}, p'_{\max}]$.

In this case, p^* is unique (and will hence be a global attractor for the within-host evolution) and takes some intermediate value. The single infection model, however, admits two CSSs. The pairwise invasibility plots for this example are shown in Figure 3.5. Figure 3.5A shows the PIP for the single infection model, along with p^*. With a discontinuous superinfection function, p^* becomes the only CSS also at the host population level, as is demonstrated in Figure 3.5B.

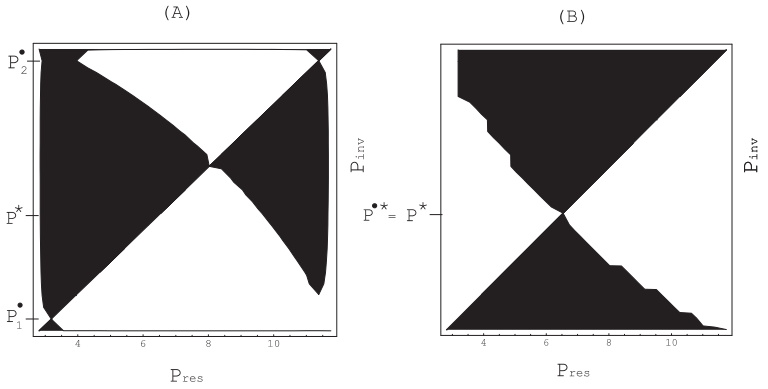

Figure 3.5: Pairwise invasibility plots for (A) the single infection model (with p^* also shown) and (B) the superinfection model with ϕ as in (3.11). The white and the black areas correspond to the regions in $(p_{\text{res}}, p_{\text{inv}})$ space that yield, respectively, positive and negative values of the invasion exponent.

3.3.2 The 'mechanistic' case

Suppose now that $\psi(x, y)$ is a nonnegative function that satisfies $\mathbf{H}_1$ and is furthermore continuous as a function of y in the point $y = x$.

Since $\psi(x, y) = 0$ whenever $x \leq y$, the function $y \mapsto \psi(x, y)$ is differentiable from the right in $y = x$. In fact,

$$\lim_{y \downarrow x} \frac{\psi(x, y)}{y - x} = 0.$$

We shall now furthermore assume that

$\mathbf{H}_2$. The limits

$$\lim_{y \uparrow x} \frac{\psi(x, y)}{y - x} \quad \text{and} \quad -\lim_{y \downarrow x} \frac{\psi(y, x)}{y - x}$$

exist, are either negative or zero, and are for any feasible x equal to each other. We denote them by $D(x)$.

Note that, when $D(x) = 0$ for some x, the function $y \mapsto \psi(x, y)$ is differentiable in the point $y = x$. In this subsection we consider the case when $D(x)$ is not identically zero, leaving the case $D = 0$ for the next subsection.

The reason for posing the additional assumption $\mathbf{H}_2$ on ψ is twofold. Firstly, we believe that this condition poses no real restriction, as we shall see in Section 3.6 that this assumption is fulfilled for the most natural choices of superinfection functions, i.e., the superinfection functions obtained when modeling the initial stages of a superinfection as a stochastic birth-and-death process. Secondly, we

shall now prove that with this additional assumption, the function $\Phi(p,q)$ in (3.8) is differentiable as a function of q in the point $q = p$, which means that we can, at least in principle, determine the singular strategies analytically.

Let us mention, however, that we have also performed a series of numerical experiments (not shown here) with continuous superinfection functions that do not satisfy assumption $\mathbf{H}_2$, and found that all the results that we shall present below (apart from, of course, analytical characterization of singular points) hold also for such functions.

Lemma 3.1. *Let $y \mapsto \psi(x,y)$ be continuous in $y = x$. Furthermore, assume that ψ satisfies $\mathbf{H}_1$ and $\mathbf{H}_2$ and let*

$$D(x) = \lim_{y \uparrow x} \frac{\psi(x,y)}{y-x} = -\lim_{y \downarrow x} \frac{\psi(y,x)}{y-x}. \tag{3.12}$$

The function

$$\Phi(p,q) = \beta(q)\psi(\hat{T}(p),\hat{T}(q)) - \beta(p)\psi(\hat{T}(q),\hat{T}(p))$$

is differentiable as a function of q in the point $q = p$.

Proof. Let us verify that the limits

$$\lim_{\varepsilon \downarrow 0} \frac{\Phi(p,p-\varepsilon)}{-\varepsilon} \quad \text{and} \quad \lim_{\varepsilon \downarrow 0} \frac{\Phi(p,p+\varepsilon)}{\varepsilon}$$

exist and are equal.

We shall only consider the case when $\hat{T}'(p) > 0$, i.e., in a neighbourhood of p, the potentially successful invaders are precisely the ones that lie to the left of p. The other cases are considered in a similar manner.

In this case, there exists an interval, say (p,p'), so that for every $\varepsilon > 0$ for which $p < p+\varepsilon < p'$ we have $\hat{T}(p) < \hat{T}(p+\varepsilon) < \hat{T}(p')$ and so

$$\begin{aligned}\Phi(p,p+\varepsilon) &= \beta(p+\varepsilon)\psi(\hat{T}(p),\hat{T}(p+\varepsilon)) - \beta(p)\psi(\hat{T}(p+\varepsilon),\hat{T}(p)) \\ &= -\beta(p)\psi(\hat{T}(p+\varepsilon),\hat{T}(p)).\end{aligned}$$

Hence

$$\lim_{\varepsilon \downarrow 0} \frac{\Phi(p,p+\varepsilon)}{\varepsilon} = \beta(p)D(\hat{T}(p))\frac{d\hat{T}}{dq}\Big|_{q=p}.$$

Furthermore, there exists an interval (p'',p) so that for every $\varepsilon > 0$ such that $p'' < p-\varepsilon < p$ we have $\hat{T}(p'') < \hat{T}(p-\varepsilon) < \hat{T}(p)$, and so

$$\begin{aligned}\Phi(p,p-\varepsilon) &= \beta(p-\varepsilon)\psi(\hat{T}(p),\hat{T}(p-\varepsilon)) - \beta(p)\psi(\hat{T}(p-\varepsilon),\hat{T}(p)) \\ &= \beta(p-\varepsilon)\psi(\hat{T}(p),\hat{T}(p-\varepsilon)).\end{aligned}$$

Hence

$$\lim_{\varepsilon\downarrow 0}\frac{\Phi(p,p-\varepsilon)}{-\varepsilon}=\beta(p)D(\hat{T}(p))\frac{d\hat{T}}{dq}\Big|_{q=p}.$$

□

We can thus compute the selection gradient and find that

$$\begin{aligned}\frac{\partial r_p(q)}{\partial q}\Big|_{q=p}&=\frac{\partial s_p(q)}{\partial q}\Big|_{q=p}+\beta(p)D(\hat{T}(p))\hat{I}(p)\frac{d\hat{T}}{dq}\Big|_{q=p}\\&=-\beta(p)\Big[\frac{d\hat{S}}{dq}\Big|_{q=p}-D(\hat{T}(p))\hat{I}(p)\frac{d\hat{T}}{dq}\Big|_{q=p}\Big],\end{aligned}\tag{3.13}$$

where we have used in the second line that

$$\frac{\partial s_p(q)}{\partial p}+\frac{\partial s_p(q)}{\partial q}=0$$

holds on the line $q=p$.

Singular points are obtained by putting $\frac{\partial r_p(q)}{\partial q}\big|_{q=p}=0$.

If $D(\hat{T}(p))=0$ for some feasible trait value p, then p is a singular strategy if and only if it is a singular strategy for the selection at the between-host level in the context of the single infection model. If $D(\hat{T}(p))\neq 0$, then p can only be a singular strategy when (i) it is both a singular strategy for selection at the between-host level (as given by the single infection model) and for the within-host selection or (ii) selection at the within-host level works in a different direction than selection at the between-host level (in the context of the single infection model). Indeed, if selection tends to increase p at both levels then $\hat{S}'(p)<0$ and $\hat{T}'(p)<0$, which gives a positive selection gradient in $q=p$. Similarly, there can be no singular points in regions where both selection pressures tend to decrease p. In particular, (3.13) implies that when p^* and $p^\bullet$ are unique, the convergence stable singular strategy can only lie inbetween the within-host CSS and the CSS obtained with the single infection model (cf. Example 3.6).

Remark 3.3. Note that when $\frac{d\hat{S}}{dq}\big|_{q=p}=0$, then

$$\operatorname{sign}\frac{\partial r_p(q)}{\partial q}\Big|_{q=p}=-\operatorname{sign}\frac{d\hat{T}}{dq}\Big|_{q=p},$$

which can be interpreted as meaning that superinfection leads to an increased virulence: indeed, if $\hat{S}$ has a unique minimum, $p=p^\bullet$, then evolution ends in $p^\bullet$ if superinfections are ignored, but when we take superinfections into account, we move from $p^\bullet$ in the direction of decreasing $\hat{T}$ and, for reasonable choices of how α depends on the internal state $(\hat{T},\hat{T}^*,\hat{V})$, this means increasing α. The increased CSS value of virulence as a consequence of superinfections has also been observed by others [84, 88, 92, 93].

In contrast with the single infection case and the 'jump' case, however, we may encounter convergence stable singular traits that are invadable, viz. branching points (cf. Figure 3.7D). Such traits turn out to be very interesting for the course of evolution since they evoke evolutionary branching and can, subsequently, lead to evolutionary coexistence of pathogen traits [51, 81]. We shall return to this important point in Section 3.5.

Let us now present two examples. In both of the examples we take superinfection functions with

$$\psi_n(x,y) = \begin{cases} 1 - \left(1 - \dfrac{kx}{c+kx} + \dfrac{ky}{c+ky}\right)^n, & y < x \\ 0, & \text{otherwise} \end{cases} \tag{3.14}$$

for some $n \in \mathbb{N}$. As we shall see in Section 3.6, these are the functions obtained when we describe the initial stages of a superinfection as a stochastic birth-and-death process.

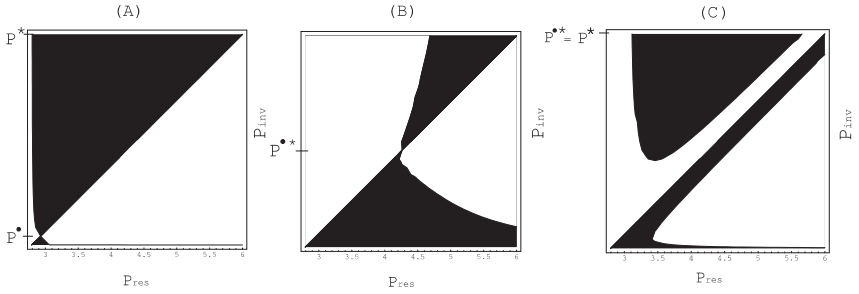

Figure 3.6: Pairwise invasibility plots for (A) single infection model (B) the 'mechanistic' case and (C) the 'jump' case. The CSS for the mechanistic case, $p^{\bullet *}$, lies inbetween $p^{\bullet}$ and p^*.

Example 3.6. In this example we basically repeat the experiment from Example 3.4, only this time with a continuous superinfection function. In Figure 3.6 we plot three PIPs: for comparison, (A) and (C) show, respectively, the PIP for the single infection model and the superinfection model with a discontinuous superinfection function (i.e. the 'jump' case), while (B) shows the PIP when the superinfection function is constructed with ψ_n in (3.14), taking $n = 100$. In this case, the convergence stable singular strategy $p^{\bullet *}$ is also an ESS, hence a CSS.

Example 3.7. We now take the same parameter values and trade-off functions as in Example 3.5. In Figure 3.7 we plot six pairwise invasibility plots. For comparison, we include the PIP for the single infection model (Figure 3.7A) and the PIP for the 'jump' case (Figure 3.7F). In Figures 3.7 B-E we construct superinfection

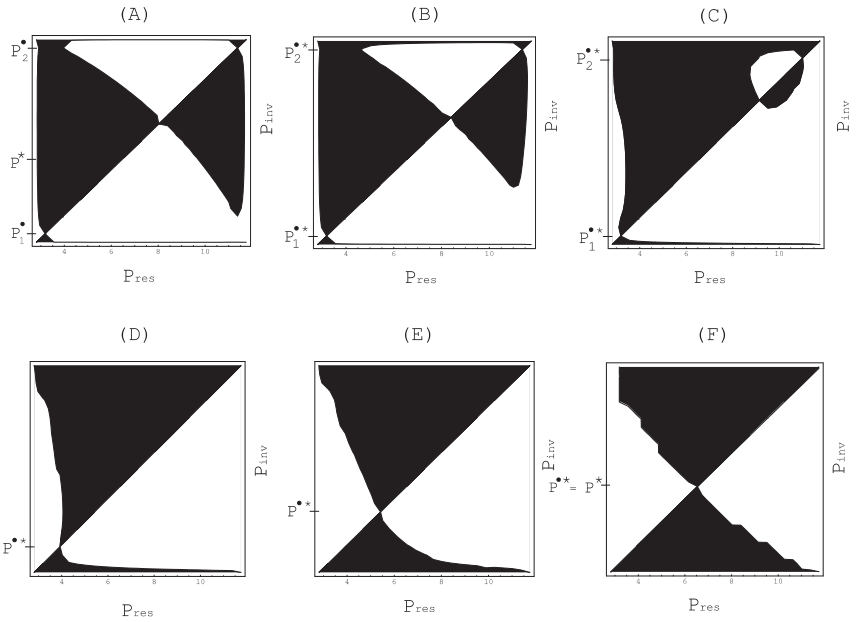

Figure 3.7: Pairwise invasibility plots for (A) single infection model (B)-(E) superinfection model with continuous superinfection functions derived from (3.14) with, respectively, $n = 10, 50, 100$ and 1000, (F) the 'jump' case.

functions with ψ_n given by (3.14) with, respectively, $n = 10, 50, 100$ and 1000. While in Figures 3.7B and 3.7C the convergence stable singular traits are also ESSs, this is no longer the case in Figures 3.7D and 3.7E.

3.3.3 The 'smooth' case

We have seen that, in the 'jump' case, a slightly better within-host competitor has a large advantage at the host population level and, as a consequence, singular strategies at the host population level coincide with the singular strategies of the within-host model.

By contrast, when $D(x)$ in (3.12) is equal to zero for every feasible x, we find that

$$\frac{\partial r_p(q)}{\partial q}\Big|_{q=p} = \frac{\partial s_p(q)}{\partial q}\Big|_{q=p}, \tag{3.15}$$

or, in words, the selection gradient is exactly the same as when the possibility of superinfections is ignored.

Equality (3.15) then implies the following: (i) singular strategies of the superinfection model coincide with the ones given by the single infection model,

(ii) singular strategies that are (local) evolutionary attractors in the context of the single infection model are also attractors for the monomorphic dynamics in the context of the superinfection model and (iii) singular traits that are (local) repellors in the context of the single infection model remain repellors also when superinfections are taken into account.

In those points in which the selection gradient vanishes, (3.15) offers no insight into the course of evolution. And while the convergence stable strategies are necessarily ESSs in the context of the single infection model, this may no longer be the case for the superinfection model. We may, in fact, encounter evolutionary branching points, which is impossible in the context of the single infection model. Note, however, that when $y \mapsto \phi(x, y)$ and $y \mapsto \phi(y, x)$ are differentiable twice in $y = x$, these derivatives are necessarily zero and hence the singular traits have precisely the same character as in the context of the single infection model. In particular, convergence stable singular traits are also uninvadable.

3.4 An interlude: dimension of the environment, (non)existence of an optimization principle and mutual invasibility

Once the singular strategies have been determined, one needs to establish their properties (such as convergence stability and (un)invadability) to determine the possible evolutionary scenarios.

For selection at the two isolated levels, the pessimization principle implies that the convergence stable strategies are necessarily uninvadable, while evolutionary repellors are invadable. In Section 3.3.1 we showed that this is also the case if the superinfection function is discontinuous, i.e. when a slightly better within-host competitor has a large advantage over the resident. One might thus be tempted to believe that an optimization (or pessimization) principle exists, at least in the 'jump' case, also for the superinfection model. This is, in fact, not so.

The aim of this section is to show that singular traits can not be characterized by an optimization (and hence also not by a pessimization) principle in the context of the superinfection model. Along the way we shall also introduce some notions that will be useful in the following section.

When infected individuals are fully protected from further infections, selection at the host population level occurs entirely through the size of the subpopulation of susceptible individuals. As a consequence, evolution drives the within-host reproduction rate in the direction that will push the pathogen's between-host basic reproduction ratio $\mathcal{R}_0$ towards a local maximum.

If, on the other hand, infected hosts can be reinfected, then the environment

the pathogens experience is no longer one dimensional as the pathogens can, conditionally on their capacity to win the within-host competition, also superinfect an already infected host. When the dimension of the environment is larger than one we can, in general, no longer expect that we will find optimization in the course of evolution [82], nor can we be sure that the pathogen population remains monomorphic at the host population level.

To see that the singular traits can no longer be characterized by way of an optimization principle when superinfections are taken into account, we recall the results of [82], stating that an optimization principle exists if and only if the environment *acts* in a one dimensional manner. More precisely, let us denote by $\mathbb{X}$ the set of feasible trait values, by $\mathbb{E}$ the set of all feasible environmental conditions and by $\rho(x, E)$ the invasion fitness of the trait $x \in \mathbb{X}$ in an environment $E \in \mathbb{E}$. Metz et al. [82] showed that an optimization principle exists if and only if there exist a function $\theta : \mathbb{E} \to \mathbb{R}$ and a function $\eta : \mathbb{X} \times \mathbb{R} \to \mathbb{R}$, decreasing in its second variable, such that

$$\operatorname{sign} \rho(x, E) = \operatorname{sign} \eta(x, \theta(E))$$

for every $x \in \mathbb{X}$, $E \in \mathbb{E}$.

The (non)existence of an optimization principle can be recognized from a pairwise invasibility plot. Namely, if such a principle exists, the PIPs must be skew symmetric [28] (meaning they are invariant under reflection across the main diagonal and simultaneous exchange of the two signs/colors, in our case black and white). The examples of Section 3.3 (in particular, Figures 3.4B, 3.5B, 3.6B, 3.6C and 3.7B-F) thus clearly show that the outcome of evolution can not be determined by an optimization principle when superinfections are possible (these examples in fact only show that there is no optimization in the 'jump' and the 'mechanistic' case. Skew symmetry is lost also in the 'smooth' case, but since this case is biologically not as relevant as the other two, we do not include the PIPs).

The lack of skew symmetry implies that we encounter pairs (p, q) for which either $r_p(q) > 0$ and $r_q(p) > 0$ or $r_p(q) < 0$ and $r_q(p) < 0$. In the former case, the traits are said to be *mutually invadable*, while they are *mutually uninvadable* in the latter. In words, if p and q are mutually invadable, trait p is able to invade q when q is rare, but is not able to outcompete it since q can also invade p when p is rare. In such a case, a so called *protected dimorphism* arises.

Mutual (un)invadability can easily be read off from a pairwise invasibility plot. Namely, interchanging p and q corresponds geometrically to reflection across the main diagonal. If we plot the reflected PIP along with the original one, then the regions where both $r_p(q) > 0$ and $r_q(p) > 0$ correspond to the regions where both PIPs are white, while p and q are mutually uninvadable if both $r_p(q)$ and $r_q(p)$ fall into the black region.

3.5 Coexistence of pathogen traits on the evolutionary time scale

We have thus established that coexistence of pathogens at the host population level is possible when superinfections are taken into account. The evolutionary significance of dimorphisms (or, more generally, polymorphisms) relies on whether they are maintained also on the evolutionary time scale.

The aim of this section is to demonstrate that the superinfection model can, in fact, sustain also evolutionary coexistence. In principle, two questions need to be answered. First, if the pathogen population is monomorphic to begin with, can polymorphisms be created and maintained in the course of evolution? And second, if the pathogen population level is polymorphic to begin with, can polymorphisms stand the test of evolutionary persistence?

Retaining the assumption of Section 3.3 that mutations are small (and rare), we begin with the first question. We shall in fact only deal with the 'mechanistic' and the 'jump' case since these two cases have clear biological interpretation. Let us only remark that evolutionary coexistence was found by way of numerical experiments also in the 'smooth' case.

3.5.1 Evolutionary branching points

When the trait of an initially monomorphic pathogen population evolves, by a series of trait substitutions, into a neighbourhood of a convergence stable singular trait around which coexistence is possible, the population becomes dimorphic. If the convergence stable singular trait is also uninvadable (and hence a CSS), dimorphisms occur only as transients in the evolution towards a monomorphic pathogen population and hence have no evolutionary significance. Such dimorphisms are also called *converging* dimorphisms [51, 81].

If, on the other hand, the population becomes dimorphic in a neighbourhood of a *branching point*, i.e. a singular trait that is convergence stable and invadable, the dimorphisms can not be dismissed: the two traits become more and more distinct in the course of evolution (at least for a while, that is) and can lead to coexistence of two (or more, since further branching may occur) pathogen traits on the evolutionary time scale. One also speaks of *diverging dimorphisms* [51, 81]. In the long run, the population may end up being monomorphic by way of extinction of one of the branches or an attractor coalition of two, or more, traits may develop.

In the 'jump' case, the singular strategies are the same, and furthermore have the same character, as the ones obtained with the within-host model. And since the within-host model does not allow for branching points, neither does the superinfection model in that case. Hence, if the pathogen population is initially

monomorphic, the outcome of evolution at the host population level will be the same as the outcome in a single infected host, only that perhaps the CSS is at the host population level approached also via polymorphisms, and not only through trait substitutions, as is the case within one host. However, as we will see later on, existing dimorphisms can be maintained in the course of evolution provided that some additional conditions are satisfied.

By contrast, the (biologically most relevant) mechanistic superinfection model can induce branching points. Now, for a singular trait $p^{\bullet *}$ to be uninvadable, the invasion exponent has to be negative at both sides of $p^{\bullet *}$, i.e., the function $q \mapsto r_{p^{\bullet *}}(q)$ has a strict local maximum in $q = p^{\bullet *}$. If the function $q \mapsto r_{p^{\bullet *}}(q)$ belongs to $C^2(\mathbb{R})$, then a sufficient condition for uninvadability is that the second derivative evaluated in $q = p^{\bullet *}$ is strictly negative. In the present case, the second derivative in the point $q = p^{\bullet *}$ need not exist. We can compute the right, as well as the left second derivative, and when both are negative, the singular point will be an ESS and hence a CSS when it is also convergence stable. Let us remark that we have computed these derivatives, but found that they offer no clear road to meaningful biological interpretation. Since they furthermore incorporate second derivatives of $\hat{S}$ and $\hat{T}$, it seems almost impossible to make any general conclusions (remember that, in the 'mechanistic' case, the singular trait will generically lie inbetween the extrema of $\hat{S}$ and $\hat{T}$) as to when the singular trait will be uninvadable (or convergence stable, for that matter). We hence chose not to include these explicit expressions in the presentation.

Generically, branching points are convergence stable singular traits $p^{\bullet *}$ for which $q \mapsto r_{p^{\bullet *}}(q)$ has a strict local minimum in $q = p^{\bullet *}$. That is, in a generic case, the left as well as the right second derivative of $q \mapsto r_{p^{\bullet *}}(q)$ in $q = p^{\bullet *}$ will be strictly positive. But in the present 'mechanistic' situation, it is not unusual for the left and the right second derivatives to have different signs! That this indeed does happen is demonstrated in Figure 3.8C. In such a case, evolution does not stop when a sequence of trait substitutions brings the trait into a vicinity of $p^{\bullet *}$. Just as in the case of a minimum of $q \mapsto r_{p^{\bullet *}}(q)$, mutual invasibility may lead to coexistence, but now (cf. Figure 3.9) there is asymmetry, with very little scope for one trait (in this case the lower trait) when compared to the scope of the other (upper) trait.

When a population undergoes branching, the invasion exponent $r_p(q)$ is no longer useful as it presupposes a monomorphic population. Instead, we replace it by $r_{p_1,p_2}(q)$, which denotes the growth rate of the population with trait q, when introduced into a steady host population in which traits p_1 and p_2 are resident. Note that this formalism needs to be extended again if the population undergoes further branching. It is known that the dimension of the environment sets an upper bound on the number of pathogen traits that can coexist in a steady state [33]. When superinfections are taken into account, the environment is, in principle,

infinite dimensional and given by $\{\hat{S}, \hat{I}_p\}$ with p in the domain of feasible trait values. Moreover, there may exist other attractors and not only steady states. We shall not go any further into the attractors of (3.7) and their stability at this point: the aim of this section is merely to demonstrate that evolutionarily stable coexistence of two pathogen traits is possible in the context of the superinfection model.

We therefore focus on the case with two resident traits and compute the invasion exponent $r_{p_1,p_2}(q)$, which takes the form

$$r_{p_1,p_2}(q) = \beta(q)\hat{S}(p_1,p_2) - (\alpha(q)+\delta) + \Phi(p_1,q)\hat{I}_{p_1}(p_1,p_2) + \Phi(p_2,q)\hat{I}_{p_2}(p_1,p_2), \tag{3.16}$$

where $\hat{S}(p_1,p_2)$ denotes the steady state value of the susceptible population when traits p_1, p_2 are present in the population and $\hat{I}_{p_1}, \hat{I}_{p_2}$ stand for, respectively, the nontrivial equilibrium values of the number of hosts infected by p_1 and p_2.

Using (3.7) we find that

$$\hat{S}(p_1,p_2) = \frac{b\Phi(p_1,p_2)}{\delta\Phi(p_1,p_2) + \beta(p_1)\beta(p_2)(\hat{S}(p_2) - \hat{S}(p_1))} \tag{3.17a}$$

$$\hat{I}_{p_1}(p_1,p_2) = \frac{\beta(p_2)}{\Phi(p_1,p_2)}\Big(\hat{S}(p_2) - \hat{S}(p_1,p_2)\Big) \tag{3.17b}$$

$$\hat{I}_{p_2}(p_1,p_2) = \frac{\beta(p_1)}{\Phi(p_1,p_2)}\Big(\hat{S}(p_1,p_2) - \hat{S}(p_1)\Big), \tag{3.17c}$$

where $\hat{S}(p) = \frac{\alpha(p)+\delta}{\beta(p)}$.

Without any loss of generality we can assume that $\hat{T}(p_1) > \hat{T}(p_2)$. Hence $\Phi(p_1,p_2) = \beta(p_2)\phi(p_1,p_2)$ and (3.17) specify a feasible (i.e., nonnegative and nontrivial) steady state if and only if

$$\hat{S}(p_1) < \hat{S}(p_1,p_2) < \hat{S}(p_2). \tag{3.18}$$

Recalling that the boundary steady states are given by

$$\hat{S}(p) = \frac{\alpha(p)+\delta}{\beta(p)} \quad \text{and} \quad \hat{I}(p) = \frac{1}{\beta(p)}\Big(\frac{b}{\hat{S}(p)} - \delta\Big),$$

we find that p_1 can invade the p_2-only steady state when $\hat{S}(p_1,p_2) < \hat{S}(p_2)$, while vice versa, p_2 can invade the p_1-only steady state when $\hat{S}(p_1) < \hat{S}(p_1,p_2)$. In other words, steady coexistence of two traits is possible if and only if the traits are mutually invadable. A necessary condition for this is that $\hat{S}(p_1) < \hat{S}(p_2)$ (that is, given that the trait p_2 wins the internal competition, p_2 must win the single infection between-host competition), while a sufficient condition is given in (3.18).

In a similar manner as finding the singular points for the monomorphic dynamics, we can determine the (internal) singular coalitions by putting

$$\frac{d}{dq} r_{p_1,p_2}(q)\Big|_{q=p_1} = \frac{d}{dq} r_{p_1,p_2}(q)\Big|_{q=p_2} = 0.$$

Note that we can depict such coalitions from a pairwise invasibility plot. Namely, each of the conditions $\frac{d}{dq} r_{p_1,p_2}(q)|_{q=p_1} = 0$ and $\frac{d}{dq} r_{p_1,p_2}(q)|_{q=p_2} = 0$ represents a curve in a (p, q) plane. If the curves intersect in the white area, where coexistence is possible, we obtain a feasible singular coexistence. The (un)invadability of a singular coalition can be (relatively) easily verified. Namely, a singular coalition is evolutionarily stable if and only if the traits that form it are uninvadable. Convergence stability, however, is not so straightforward, but the reader can find some results in [78]. We shall refrain from analytic treatment of singular coalitions and proceed by way of numerical examples.

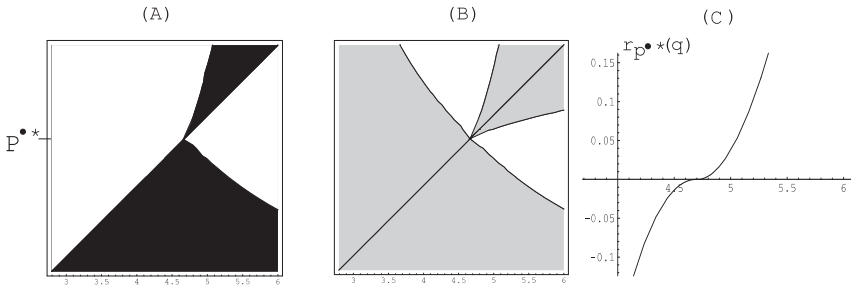

Figure 3.8: (A) The PIP corresponding to the superinfection model with $\psi = \psi_{200}$ in (3.14). (B) White areas correspond to the regions in (p, q) space in which p and q are mutually invadable. Grey regions are the areas where $r_p(q)$ and $r_q(p)$ have opposite signs. (C) The graph of $r_{p^{\bullet *}}(q)$.

Example 3.8. We take the parameter values and the trade-off functions as in Example 3.4 and investigate numerically whether evolutionary branching points exist if the superinfection function is given by (3.14) for some $n \in \mathbb{N}$. In this case, the function $p \mapsto \hat{T}(p)$ is a strictly decreasing function, which implies that the within-host CSS will take the value of the physiological restriction on trait values, in this case taken to be $p_{\max} = p^* = 6$.

We found that for low values of n (that is, when the host is reinfected by a relatively low dose) the convergence stable singular strategy will also be an ESS and will hence give rise only to converging dimorphisms. When n becomes larger, the convergence stable steady state becomes invadable, which can give rise to evolutionary coexistence of two distinct traits at the host population level. This is demonstrated in Figure 3.8, where $n = 200$. The convergence stable

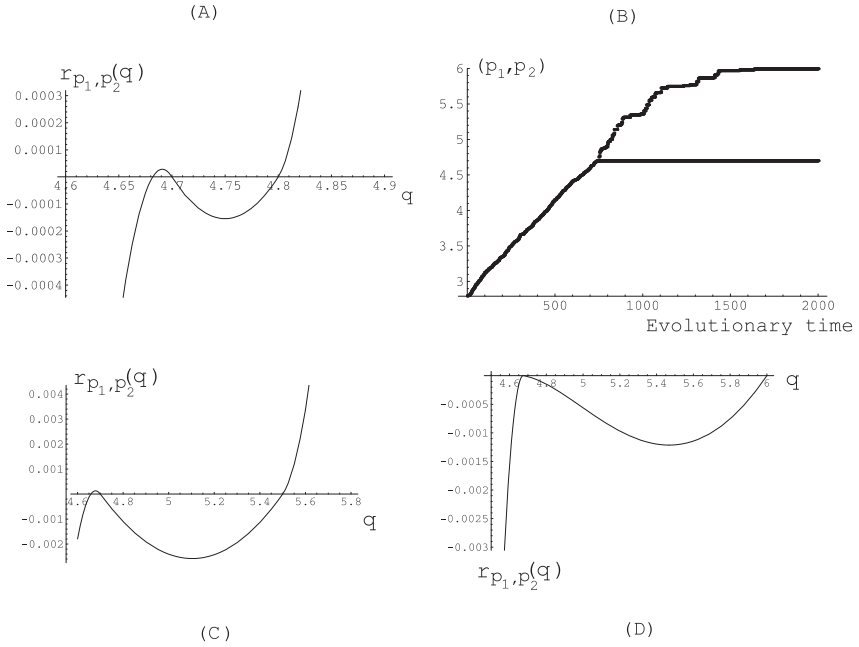

Figure 3.9: (A) The graph of $r_{p_1,p_2}(q)$ with the initial dimorphism $(p_1, p_2) = (4.7, 4.8)$ shows that (p_1, p_2) can be invaded by a small mutation when $q < p_1$ or $q > p_2$. (B) The simulated evolutionary tree shows evolutionary branching leading to two distinct pathogen subpopulations. (C) The graph of $r_{4.7,5.5}(q)$ (D) The graph of $r_{4.675,6}(q)$.

strategy for the monomorphic dynamics was found numerically and takes the value $p^{\bullet *} = 4.69345$. Note that $p^{\bullet *}$ in Figure 3.8 is a branching point, however, of an unusual type as it is invadable from above, but not from below.

The simulated evolutionary tree in Figure 3.9B shows monomorphic evolution towards $p^{\bullet *}$, after which branching occurs. In general, branching does not guarantee that we will find coexistence on the evolutionary time scale since one branch may go extinct and the other may evolve towards a monomorphic ESS. However, extinction after branching was not encountered in this case. We found that the within-host CSS, p^* (which is in this case at the physiological maximum, cf. Example 4), coexists at the host population level with trait $p = 4.675$.

Figures 3.9A and 3.9C show the graphs of $q \mapsto r_{p_1,p_2}(q)$ for two choices of resident coalitions (p_1, p_2): Figure A shows the initial dimorphism and the invasion ability of neighbouring traits, while Figure C is taken later on in the evolutionary time with $(p_1, p_2) = (4.7, 5.5)$. Note that the scope for successful invaders around the lower trait is indeed very small, compared to the scope for successful invaders around the upper trait. Figure 3.9B shows the result of a

single simulation. Note also that the lower branch appears to be constant but is fact not so since $p^{\bullet *} = 4.69345$, but the lower trait in the evolutionarily stable coexistence takes the value $p = 4.675$. We have performed a large number of simulations, which suggest that the coalition $(4.675, 6)$ is convergence stable as well as uninvadable (the uninvadability is also seen in Figure 3.9D).

3.5.2 Maintenance of dimorphisms through the evolutionary time

We have demonstrated that the mechanistic case allows for evolutionary branching as well as evolutionary coexistence. In fact, so does the 'smooth' case, but since this case is not as biologically relevant as the other two, we do not give it much attention. We have furthermore shown that the 'jump' case does not induce branching points; natural selection will in this case drive p towards a local minimum of $\hat{T}(p)$, but perhaps also via convergent polymorphisms and not only through trait substitutions.

When there is no limitation on the size of the mutation (apart from, of course, that the traits remains in the biologically feasible interval), or when the pathogen population is polymorphic to begin with, branching points are neither necessary nor sufficient for maintenance of polymorphisms. In such a case, evolutionary coexistence becomes possible also in the 'jump' case. The only candidates for stable coalitions, however, are the minima of $\hat{T}$ that are mutually invadable (a necessary condition for such coexistence is therefore that $p \mapsto \hat{T}(p)$ has more than one local minimum). This can be seen if we use (3.17b) to rewrite (3.16) as

$$r_{p_1,p_2}(q) = s_{p_2}(q) - \left(\tfrac{\beta(q)}{\beta(p_2)}\Phi(p_1,p_2) - \Phi(p_1,q)\right)\hat{I}_{p_1}(p_1,p_2) + \Phi(p_2,q)\hat{I}_{p_2}(p_1,p_2) \tag{3.19a}$$

and (3.17c) to write

$$r_{p_1,p_2}(q) = s_{p_1}(q) + \left(\tfrac{\beta(q)}{\beta(p_1)}\Phi(p_1,p_2) + \Phi(p_2,q)\right)\hat{I}_{p_2}(p_1,p_2) + \Phi(p_1,q)\hat{I}_{p_1}(p_1,p_2). \tag{3.19b}$$

With the same reasoning as in Section 3.3.1 we can then show that the only attracting coalitions are the ones consisting of two local minima of $\hat{T}(p)$: if the mutations are small, i.e., lie in $(p_1 - \varepsilon, p_1 + \varepsilon) \cup (p_2 - \varepsilon, p_2 + \varepsilon)$ for some $\varepsilon > 0$, then each successful mutation will bring either p_1 or p_2 closer to a local minimum. Note, however, that in a generic case we will have either $\hat{T}(p_1) < \hat{T}(p_2)$ or $\hat{T}(p_2) < \hat{T}(p_1)$, but the better of the two at the within-host level will not outcompete the other at the host population level, since the terms $s_{p_1}(q)$ and $s_{p_2}(q)$ can only be neglected in small enough neighbourhoods.

Finally, we observe that Nowak and May showed in [88] that, in the 'jump' case, many traits can coexist on the epidemiological as well as evolutionary time

scale, while here we find that the number of local minima of $\hat{T}$ sets an upper bound for the number of traits that can coexist on the evolutionary time scale. When mutations are uniformly distributed (rather than being small) Nowak and May even obtain a continuum of traits on the evolutionary time scale by a balance between evolution towards increased virulence and generation (by mutation) of low virulence strains.

3.6 The dynamics in the initial stages of a superinfection

In the initial stages of a superinfection, the invading trait q is likely to be present only in small quantities. Hence, even when $\hat{T}(q) < \hat{T}(p)$ (and so the newly introduced trait has the potential to outcompete the resident trait), trait q may go extinct due to demographic stochasticity in the initial stages of a superinfection, when it is still rare.

In this section we supplement the deterministic model by describing the initial stages of an invasion as a stochastic birth-and-death process. This will lead us to continuous (but not differentiable) superinfection functions, hence the word 'mechanistic' for this case.

So let us begin by assuming that only one free pathogen with trait q is introduced into a host already infected by trait p. If we assume that the trait p resides in a stable equilibrium, then the new trait q is introduced into an environment given by the steady state value of $\hat{T}(p)$,

$$\hat{T}(p) = \frac{c}{k(\mathcal{B}_0(p) - 1)}. \tag{3.20}$$

The probability with which the clan of this initially introduced pathogen survives in an already infected host, is given as the smallest fixed point of a generating function [59]. In order to compute it, we must first derive the probabilities π_n, with which one free pathogen with trait q will produce n new pathogens.

Now, in order to reproduce, a pathogen must bind to an uninfected target cell. This happens with probability

$$\frac{k\hat{T}(p)}{k\hat{T}(p) + c}.$$

When the pathogen enters a target cell, its survival relies on the survival of the target cell that hosts it. The life span of a target cell infected with trait q is exponentially distributed with parameter $(\mu(q) + d)$. The infected target cell produces free pathogens according to a Poisson process with parameter q. So the probability density that an infected target cell lives t units of time and in that time produces n offspring equals

$$(\mu(q) + d)e^{-(\mu(q)+d)t}e^{-qt}\frac{q^n t^n}{n!}.$$

Accounting for all possible times t, we arrive at the following expression for π_n,

$$\pi_n = \frac{k\hat{T}(p)}{k\hat{T}(p)+c} \int_0^\infty (\mu(q)+d)e^{-(\mu(q)+d)t} e^{-qt} \frac{q^n t^n}{n!} dt, \tag{3.21}$$

which is valid for $n \geq 1$. The probability of having no offspring at all is, however, not given by π_0, as the pathogen may never reproduce simply because it never enters an uninfected target cell. Since the probability with which the pathogen dies before it binds to a cell equals $\frac{c}{k\hat{T}(p)+c}$, we thus obtain the following generating function $G(z)$,

$$G(z) = \frac{c}{k\hat{T}(p)+c} + \sum_{n=0}^{\infty} \pi_n z^n,$$

which, by using (3.21) and interchanging the order of summation and integration, can be written as

$$G(z) = \frac{c}{k\hat{T}(p)+c} + \frac{k\hat{T}(p)}{k\hat{T}(p)+c} \cdot \frac{1}{1+\mathcal{B}_0(q)(1-z)}. \tag{3.22}$$

The function G has the following properties:

(i) $G(0) = \frac{c}{k\hat{T}(p)+c} + \pi_0$,

(ii) $G(1) = 1$,

(iii) $G'(1) = \frac{k\hat{T}(p)}{c+k\hat{T}(p)} \mathcal{B}_0(q) = \mathcal{R}_0^w(q)(\hat{T}(p))$,

(iv) $G'(z) > 0$,

(v) $G''(z) > 0$.

Let us assume that $G(0) > 0$. One can then see that these properties guarantee that G has at most two fixed points, one of which is $z = 1$. The probability with which the clan of the invading pathogen goes extinct is given as the smallest solution of $G(z) = z$ (see [35, 59] for details). Whether the second solution lies in $[0, 1]$ (and hence gives a meaningful value for a probability), depends on the value of $G'(1)$, which equals the invaders within-host reproduction ratio in the environment set by the resident, $\mathcal{R}_0^w(q)(\hat{T}(p))$. If $\mathcal{R}_0^w(q)(\hat{T}(p)) \leq 1$, then the second solution is above 1 and so the clan will go extinct with certainty. If, on the other hand, $\mathcal{R}_0^w(q)(\hat{T}(p)) > 1$, the invasion will be successful with nonzero probability.

Let $P_1(p, q)$ denote the probability of extinction of trait q, following an introduction of a single free pathogen into an environment set by the resident trait p. Using (3.22) and (3.20) we obtain

$$P_1(p,q) = \min\Big\{1, \frac{c}{c+k\hat{T}(p)} + \frac{1}{\mathcal{B}_0(q)}\Big\} = \min\Big\{1, 1 - \frac{1}{\mathcal{B}_0(p)} + \frac{1}{\mathcal{B}_0(q)}\Big\}.$$

Note that, having rewritten $P_1(p,q)$ in terms of burst sizes only, we can now also say that the invading trait has a nonzero probability of success only when its burst size exceeds the burst size of the resident trait p. We also observe that (i) $P_1(p,p) = 1$, as it should be since the resident trait resides at a stable equilibrium, and (ii) when $\mathcal{B}_0(q) \to \infty$, the invading trait will survive with certainty, provided that the initially introduced pathogen makes it to an uninfected target cell. The probability of extinction must therefore equal the probability with which the pathogen dies before it enters a target cell. And indeed we find that

$$\lim_{\mathcal{B}_0(q)\to\infty} P_1(p,q) = \frac{c}{c + k\hat{T}(p)}.$$

Let now $\phi_1(p,q)$ denote the complementary probability that the clan of one free pathogen with trait q survives in the environment set by the resident trait. This is then given by

$$\phi_1(p,q) = \begin{cases} \dfrac{1}{\mathcal{B}_0(p)} - \dfrac{1}{\mathcal{B}_0(q)}, & \mathcal{B}_0(p) < \mathcal{B}_0(q) \\ 0, & \text{otherwise.} \end{cases} \tag{3.23}$$

When n particles are introduced, therefore, the probability of survival equals

$$\phi_n(p,q) = \begin{cases} 1 - P_1^n(p,q), & \mathcal{B}_0(p) < \mathcal{B}_0(q) \\ 0, & \text{otherwise.} \end{cases} \tag{3.24}$$

or, to continue the thread of previous sections, rewritten in terms of $\hat{T}(p)$ and $\hat{T}(q)$,

$$\phi_n(p,q) = \begin{cases} 1 - \left(1 - \dfrac{k\hat{T}(p)}{c + k\hat{T}(p)} + \dfrac{k\hat{T}(q)}{c + k\hat{T}(q)}\right)^n, & \hat{T}(q) < \hat{T}(p) \\ 0, & \text{otherwise.} \end{cases} \tag{3.25}$$

In the limit, when the number of initially introduced pathogens approaches infinity we have

$$\lim_{n\to\infty} \phi_n(p,q) = \begin{cases} 1, & \hat{T}(q) < \hat{T}(p) \\ 0, & \text{otherwise.} \end{cases}$$

That is, when a large number of pathogens with trait q is introduced, the deterministic description gives the full story: if the newly introduced trait goes extinct, it is because it loses the competition within the host and not due to bad luck while it is still rare.

Note that this description makes the superinfection functions $\phi = \phi_n(p, q)$ continuous (but not differentiable) as a function of q in the point $q = p$. The fact that they are increasing as functions of $\mathcal{B}_0(q)$, implies that the traits that significantly increase the burst size (i.e., the ones that significantly reduce the steady state level of uninfected target cells within the host) also have a better chance of surviving in the host than the traits which are only slightly better within-host competitors than the resident. When the number of initially introduced mutants goes to infinity, we obtain a superinfection function with a jump discontinuity in $q = p$, which furthermore makes no distinction among the winning strategies.

3.7 Concluding remarks

The individual host is a cul-de-sac for parasites that reproduce within it and hence parasite persistence relies on host-to-host transmission. When the risk that competitors enter the same host is negligible, this need for transmission sets an optimal level for within-host prudence. But if the same host can be reinfected by another strain, the parasites are confronted with a version of the milker-killer dilemma [104].

In order to evaluate the balance of the various evolutionary forces one needs, to begin with, a somewhat detailed description of the status of an infected host and how this status changes in the course of time. In principle, the status should incorporate all relevant information about parasite burden, immune response and harm done. Pragmatic first steps are usually based on very caricatural descriptions in terms of just a few variables, with dynamics generated by very simple deterministic rules. The link to the host population level is then made by specifying how the mortality of the host and the transmissibility of the parasite (upon contact of the host with another, susceptible, host) depend on the status. Various authors use terms like 'nested' or 'embedded' models to indicate that a submodel for within-host dynamics is used as a building block for an epidemic model at the host population level (see [1] for a survey and references).

In order to incorporate the effects of reinfection by another strain, one needs to address a number of rather subtle issues, having to do with the 'dose' of parasites that enters the host at transmission. A first point is that demographic stochasticity may play a major role right away, simply because the dose is small. A second point is that there is no clue as to the choice of an initial value for the deterministic variable representing the new strain in the host status.

In this chapter we have, building on the work of Gilchrist and Coombs [52], adopted a consistent series of simplifying assumptions to obtain a tractable model. First of all, we assumed that mutations are so rare at the epidemiological time scale that we can use the Adaptive Dynamics description of evolution as a

sequence of trait substitutions and, possibly, branching. We consider an SI model, i.e., we ignore an immune response that clears the infection. This allows us to assume that within-host dynamics is so fast compared to (the time scale of) the host contact process and host demography, that we can neglect transients and focus on (steady state) attractors. Since the within-host model exhibits competitive exclusion, an added bonus is that, in the deterministic description, we have superinfection rather than coinfection (i.e., a newly arriving strain fails to have any effect if it is a weaker competitor, while outcompeting the resident strain instantaneously if it is a better competitor). This no-yes dichotomy is shaded down by demographic stochasticity, in the sense that 'no' remains 'no' but 'yes' becomes 'possibly', with the probability depending on the degree of competitive superiority. We use the branching process corresponding to the linearized within-host model to compute how this probability of successful take-over depends on both the difference in competitive ability and the dose (Mosquera and Adler [84] introduced this idea in the present context in their Appendix B, but, curiously, did not elaborate the link between the branching process and the mechanistic within-host submodel in their Appendix A. Note also that Pugliese [92, 93] questions the absoluteness of the 'no').

Clearly all of these assumptions are open for debate (see [22, 97] for inspiration). They are made in order to make progress, not as a terminus. The main conclusions that we derived from these assumptions are

(i) superinfection leads to increased virulence and

(ii) superinfection may lead to evolutionarily stable coexistence of (at least) two strains.

While the first is exactly what one would expect, the second is a bit surprising, given the fact that Pugliese [92, 93] found converging rather than diverging dimorphisms. It appears that the mechanistic within-host model leads to a richer repertoire of adaptive dynamics than a phenomenological trade-off between transmissibility and host mortality. We are thus inclined to answer the rhetoric question in the title of the stimulating paper [47] with 'yes', as did the authors!

Various attempts of devising meaningful, yet tractable, within-host submodels have been made [1, 4, 29, 47, 52, 64, 83, 89, 101] and it has been clarified how superinfection can be seen as a limiting case of coinfection [29, 84]. In our view, it is now a great challenge to incorporate immunity. If the infectious period has a finite length due to the fact that the parasites are ultimately eliminated by the immune system, it becomes much more difficult, if not impossible, to postulate time scale differences that simplify the analysis. In addition, immunity has a 'long term memory' effect [53, 55]. Will it be possible to determine the influence of all these factors on the evolution of infectious diseases?

Chapter 4

Relative effects of barrier precautions and topical antibiotics on nosocomial bacterial transmission

4.1 Introduction

Infections with antibiotic-resistant microorganisms frequently occur in critically ill patients. In almost all cases, infection is preceded by asymptomatic carriage, i.e. colonization. Although different pathogens have different preferential colonization sites, several body sites may eventually become colonized, even those that are protected in healthy individuals. After one week of hospitalization, colonization of the respiratory tract, gastro-intestinal tract and skin with potentially pathogenic microorganisms occurs in almost all critically ill patients [9].

Intensive care units (ICUs) generally hold a small number of patients (typically 5-50), with an average length of stay that may very well be less than one week. Several routes of initiation of colonization can be distinguished. Patients may become colonized during their stay in the ICU through *cross-transmission*, a process that depends on the colonization pressure, that is, the number/proportion of other patients already being colonized [10]. In some patients susceptible microorganisms develop antibiotic resistance, either through specific mutations or through horizontal transfer of resistance genes. These events usually occur under antibiotic exposure and this mode of acquisition of resistance has been labeled *endogenous* acquisition of resistance. Finally, patients may be admitted to an

ICU while being colonized, either at a level that is detectable by microbiological cultures, or in bacterial amounts that do not (yet) exceed the detection limits of sampling tests. In the latter case, selection through antibiotics may ultimately reveal 'new' cases of colonization. The smallness of ICU populations, rapid patient turnover and different routes of transmission create a complex epidemiology, with naturally occurring large fluctuations in the prevalence of colonization [8].

Several infection control strategies can be used to reduce the prevalence of colonization in such a setting. Barrier precautions (e.g. hand disinfection, use of gloves and gowns, isolation of patients) aim to reduce cross-transmission (i.e. patient-to-patient transmission) and, although to a much smaller extent, endogenous transmission (transmission of microorganisms from one part of a patient's body to another, as yet uncolonized, part).

Administration of antibiotics, on the other hand, has many different facets. When administered intravenously or topically, antibiotics may serve as a way to reduce the prevalence of colonization in ICUs, either by reducing or by completely eradicating the susceptible flora, thereby also preventing the development of resistance [23, 106]. Yet, the very use of antibiotics may also increase the selection pressure and the emergence of resistant strains and may thereby only be adding to the ever increasing problem of antibiotic resistance.

Enteral administration of nonabsorbable antibiotics, as implemented in *selective decontamination of the digestive tract* (SDD), is highly controversial: topical intestinal application of non-absorbable antibiotics in ICU patients was part of a successful strategy to control an outbreak of multi-resistant *Gram negative* bacteria in a French ICU [11], but was associated with increased resistance in some other ICUs [74, 107]. In all successful interventions, SDD was presumably instrumental in reducing the spread in the ICU [23, 106]. However, in all such interventions barrier precautions and antibiotic use were introduced simultaneously, but the relative importance of the two control measures has never been determined.

In this chapter we present a theoretical setting to investigate the relative effects of different interventions on the prevalence of colonization in ICUs. The small population sizes and rapid patient turnover that are typical for intensive care units, make it natural to formulate and study a stochastic model. But for precisely the same reasons, one needs to be very careful when drawing conclusions regarding the efficacy of different control measures. Namely, if a certain intervention is implemented for some time and this intervention appears to be successful, to what extent is the reduction of colonization a result of random fluctuations and to what extent is it a result of the control measure? We therefore also formulated a deterministic counterpart of the stochastic model and compared the conclusions regarding the relative effects of different interventions with the ones drawn from the Markov chain model.

We present both models and the results in Sections 4.3 and 4.4. But to begin with, we now first describe the epidemiology and patient dynamics in intensive care units.

4.2 The epidemiology and the dynamics in the ICU

Our study does not focus on any pathogen in particular. The microorganism we have in mind is characterized by the following properties: it is a pathogen widely found in ICUs, it can colonize several sites of a patient's body (including the gastro-intestinal tract) and is resistant to some antibiotics, but is susceptible to the intestinally administered nonabsorbable antibiotics. *Pseudomonas aeruginosa*, enteric *Gram-negative* bacteria, methicillin-resistant *Staphylococcus aureus* (MRSA) and *enterococci* are all examples of such a pathogen.

We shall, however, single out three body sites that can become colonized: the skin, the gut and the lungs, and label the patients in the ICU according to the parts of their bodies that are colonized at the time. Each of the body sites is assumed to be either colonized or not (in other words, we neglect the fact that individuals may differ in the bacterial load they carry at various sites) and so the label GS, for example, indicates that the individual is colonized in the gut and on the skin, but not (yet) in the lungs. In this way we can at any point in time characterize each patient in the ICU by one of the following eight labels: 0, G, S, L, GS, GL, SL, GSL.

Skin colonization may lead to contamination of health care workers' hands during patient care and is a prerequisite for cross-transmission. We shall use the term 'skin colonization' throughout the chapter and describe by this term any situation that may lead to cross-transmission (e.g. a contaminated respiratory tubing system or suffering from diarrhoea). In fact, colonization of the lungs/intestines is characterized as one that cannot be transmitted through contaminated hands. It is, however, possible that a patient colonized in the lungs or in the gut transmits the bacteria to the skin by endogenous transmission.

The lungs, the gastro-intestinal tract and the skin are certainly sites frequently colonized by one, or even several pathogens, in ICUs [9]. But more importantly, these three sites capture precisely the three substantially different manners of colonization in view of the possible control measures: patients colonized on the skin are infectious and hence an immediate threat to other patients (and health care workers). We can reduce this threat by applying barrier precautions. Patients colonized in the gut/lungs are not yet infectious but may become so through endogenous transmission (e.g. developing diarrhoea or having a respiratory tube inserted) and while reducing the chance of bacteria spreading from the gut to the skin can be achieved by barrier precautions as well as antibiotics, preventing

the bacteria to spread from the lungs to the skin can only be done by improved hygiene. These three body sites hence capture the essence and any additional colonization site would mainly only complicate the notation as well as the computations without adding new qualitative aspects.

4.2.1 Transmission

We have already indicated in the previous subsection the possible transmission routes. Here we describe them again in order to also introduce the notation and some additional assumptions. We distinguish the following transmissions:

I. *Cross-transmission*, meaning that a patient colonized on the skin transmits the infection, via health care workers, to another patient's skin. All individuals not yet colonized on the skin are assumed to be equally susceptible to an infection (i.e. possible colonization of other body sites does not influence their susceptibility). Infectiousness of patients colonized on the skin is assumed to be independent of the colonization status of their other parts. We denote the (constant) rate at which an individual colonized (at least) on the skin infects type 0, G, L, GL by β.

II. *Endogenous transmission*, where a patient spreads the bacteria from one of the colonized sites to one of its as yet uncolonized parts. The following internal transmissions are possible:

   - from the gut to the skin,
   - from the skin to the gut or to the lungs,
   - from the lungs (possibly via the respiratory tube) to the skin.

   The list of all possible endogenous transmissions and the (constant) rates at which they occur is as follows:

$$\begin{aligned} G &\xrightarrow{\alpha_{GS}} GS, & GL &\xrightarrow{\alpha_{GS}} GSL,\\ L &\xrightarrow{\alpha_{LS}} SL, & GL &\xrightarrow{\alpha_{LS}} GSL,\\ S &\xrightarrow{\alpha_{SG}} GS, & SL &\xrightarrow{\alpha_{SG}} GSL,\\ S &\xrightarrow{\alpha_{SL}} SL, & GS &\xrightarrow{\alpha_{SL}} GSL. \end{aligned}$$

4.2.2 Admission of new patients

Individuals are admitted to the ICU at a constant rate λ. Newly admitted patients can be colonized at any of the three body parts. We shall assume that the

probabilities with which a newly admitted individual is of a certain type are constant in time (one can for example imagine that the prevalence of this pathogen in the community at large is stationary and that the individuals admitted to the ICU are sampled randomly from the outside community) and we denote them by $p_0, p_G, p_S, p_L, p_{GS}, p_{GL}, p_{SL}$ and p_{GSL} and require that

$$p_0 + p_G + p_S + p_L + p_{GS} + p_{GL} + p_{SL} + p_{GSL} = 1.$$

4.2.3 Discharge/death

Individuals are removed from the ICU (either due to discharge or death) with a constant probability per unit of time μ, which is independent of the patient's colonization status. The numbers λ and μ are related to each other via the size N^* of the ICU as follows: $\lambda = \mu N^*$ (see Appendix A).

4.3 The modeling

The typically low number of patients in ICUs makes a stochastic model a more suitable and natural description of the process than a deterministic one. We shall nevertheless first formulate a deterministic model (for the reasons mentioned in the introduction) and then later on compare some of its features to the results of the simulations of the stochastic version of the model.

4.3.1 The deterministic model

If the population sizes were large, we could describe the dynamics in the ICU by the following set of ordinary differential equations:

$$\left.\begin{aligned}
\dot{x}_0 &= \mu p_0 - (\sigma I + \mu)x_0,\\
\dot{x}_G &= \mu p_G - (\sigma I + \alpha_{GS} + \mu)x_G,\\
\dot{x}_S &= \mu p_S + \sigma I x_0 - (\alpha_{SG} + \alpha_{SL} + \mu)x_S,\\
\dot{x}_L &= \mu p_L - (\sigma I + \alpha_{LS} + \mu)x_L,\\
\dot{x}_{GS} &= \mu p_{GS} + (\sigma I + \alpha_{GS})x_G + \alpha_{SG}x_S - (\alpha_{SL} + \mu)x_{GS},\\
\dot{x}_{GL} &= \mu p_{GL} - (\sigma I + \alpha_{GS} + \alpha_{LS} + \mu)x_{GL},\\
\dot{x}_{SL} &= \mu p_{SL} + (\sigma I + \alpha_{LS})x_L + \alpha_{SL}x_S - (\alpha_{SG} + \mu)x_{SL},\\
\dot{x}_{GSL} &= \mu p_{GSL} + (\sigma I + \alpha_{LS} + \alpha_{GS})x_{GL} + \alpha_{SL}x_{GS} + \alpha_{SG}x_{SL} - \mu x_{GSL}.
\end{aligned}\right\} \tag{4.1}$$

Here, x_j with $j \in \{0, G, S, L, GS, GL, SL, GSL\}$ represent the fractions of the corresponding subpopulations (indicated by the indices) and

$$I = x_S + x_{GS} + x_{SL} + x_{GSL} \tag{4.2}$$

denotes the proportion of individuals that are colonized on the skin and can hence (immediately) cause new cases of colonization. The new parameter σ can be expressed in terms of the 'old' ones as $\sigma = \frac{\beta\lambda}{\mu}$.

The system (4.1) is a scaled version of the one written in terms of actual numbers (as opposed to fractions) and the interested reader can find all the details in Appendix A.

For reader's convenience we depict the possible changes in an individual's colonization status and the rates at which they occur in the form of the diagram in Figure 4.1.

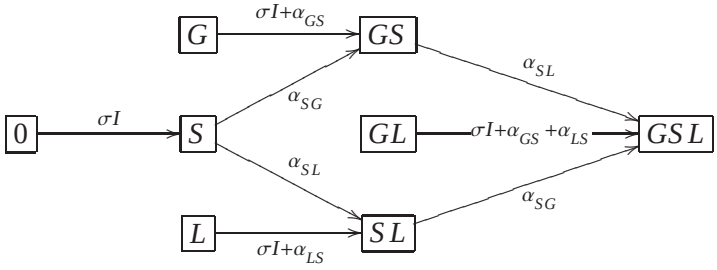

Figure 4.1: Dynamics of an individual's colonization status.

Equilibria and their stability

In Appendix A we show that the equilibrium value of I is obtained as a fixed point of the function F given by

$$F(I) = p + \frac{p_0\sigma I}{\sigma I + \mu} + \frac{p_L(\sigma I + \alpha_{LS})}{\sigma I + \alpha_{LS} + \mu} + \frac{p_G(\sigma I + \alpha_{GS})}{\sigma I + \alpha_{GS} + \mu} + \frac{p_{GL}(\sigma I + \alpha_{GS} + \alpha_{LS})}{\sigma I + \alpha_{GS} + \alpha_{LS} + \mu} \quad (4.3)$$

where

$$p = p_S + p_{GS} + p_{SL} + p_{GSL} \quad (4.4)$$

denotes the probability that a patient is colonized on the skin already on admission.

When $F(0) > 0$, the system (4.1) has a unique positive equilibrium and this equilibrium is locally asymptotically stable.

In the case $F(0) = 0$ (for instance, when all newly admitted individuals are free of the infectious agent in question, i.e., $p_0 = 1$), the system admits a 'disease

free' equilibrium, by which we mean an equilibrium consisting of only the subpopulations that cannot cause new infections. Whether the system also admits a nontrivial steady state, relies on the value of the basic reproduction ratio $\mathcal{R}_0$ [35], which is in the case $F(0) = 0$ (and only in this case) well defined and is equal to $\frac{\sigma}{\mu}$. Indeed, if, for instance, $p_0 = 1$ (see Appendix A for the derivation of $\mathcal{R}_0$ for other cases in which $F(0) = 0$) and one individual colonized on the skin is introduced into an otherwise infection free ICU, then this individual is expected to remain in the ICU for $\frac{1}{\mu}$ units of time and is in this time expected to infect the skin of $\frac{\sigma \cdot 1}{\mu}$ patients.

If $\mathcal{R}_0 \leq 1$, the infection free steady state is the only steady state of (4.1) and it is locally stable. When $\mathcal{R}_0 > 1$ there exists also a unique nontrivial equilibrium, which is then locally stable, while the infection free steady state is in that case unstable.

The proof of these claims is rather technical and can be found in Appendix A.

4.3.2 The discrete time Markov chain model

In addition to the deterministic model, we formulate a discrete time Markov chain model, which we investigate by way of simulations. The outcomes of simulations will be presented later on, here we only briefly describe the setting.

In all the simulations, the reported ICU is assumed to hold 20 beds, all of which were assumed to be occupied at all times. The process was simulated as a discrete time process, taking one day as the unit of time. We have in fact simulated the process for a number of different choices of ICU size. We found that, even though the outcomes (say, colonization levels) are different, our main conclusions do not rely on the size of the intensive care unit.

The epidemiology (i.e. the colonization sites and the transmission routes) has already been described in Section 4.2. We now furthermore assume that the discharge and admission take place at one particular time each day. In this way, the only thing that changes in the ICU in the course of one day are the patients' colonization statuses.

The length of stay of each patient in the ICU is taken to be exponentially distributed with parameter μ. After the discharge has taken place, the available beds are immediately filled with new patients. The colonization statuses (i.e. labels) of newly admitted patients are distributed according to the admission probabilities p_j. The endogenous transmissions are taken to be exponentially distributed with the rates that were introduced in Section 4.2.

Patient-to-patient transmission, on the other hand, is a density dependent process. The cross-transmission rate is proportional to the number of patients in the

ICU that are colonized on the skin and is given by

$$\frac{\sigma * \text{the number of patients colonized on the skin}}{\text{the number of patients in the ward} - 1}.$$

Note that this rate may change in time. We therefore update it on a daily basis, after each round of discharge and admission. Lastly, transmissions are assumed to be independent of one another.

4.4 Infection control measures

Administration of topical, nonabsorbable antibiotics has no effect on the colonization of the skin or the lungs. It can, however, have two different effects on the individual's gastro-intestinal tract: (i) it either reduces the rate of endogenous transmission of bacteria from the gut to the skin (although bacterial loads were not explicitly included, one can imagine that the lower rate is a result of the reduction in the bacterial load due to antibiotic therapy), but the patient remains colonized until discharge, or, (ii) colonization can be completely eradicated before the patient leaves the ICU. Both of these scenarios were investigated but since the outcomes were very similar we shall only present the results of the former scenario. Let us only remark that the latter requires a modification of both models to take into account the possibility that the intestinal carriage is successfully eliminated during patient's ICU stay.

Barrier precautions, on the other hand, can have an effect on several parameters. Primarily, they aim to reduce the cross-transmission parameter σ, but improved hygiene during patient care may also lower the rates α_{LS} and α_{GS}. Furthermore, isolation of patients on admission or disinfection of patients' skin on admission may lower the value of $p = p_S + p_{GS} + p_{SL} + p_{GSL}$.

In the deterministic setting we were in particular interested in the sensitivity of the equilibrium value of skin colonizations, I^*, to different parameters. The steady state value of I is given implicitly as a fixed point of the function F (cf. Appendix A), given in (4.3).

The sensitivity of I^* to a certain parameter, say θ,

$$S_{I^*,\theta} = \frac{\partial I^*}{\partial \theta}$$

can be determined by using the implicit function theorem as

$$S_{I^*,\theta} = \frac{\frac{\partial F}{\partial \theta}}{1 - \frac{\partial F}{\partial I^*}},$$

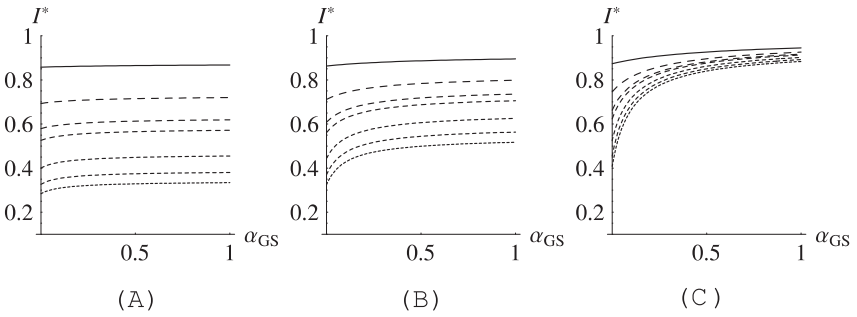

Figure 4.2: I^* plotted as a function of α_{GS}. In all three figures $\mu = \frac{1}{10}$, $\alpha_{LS} = \frac{1}{2}$ and, from the top to the bottom, $\sigma = \frac{1}{2}, \frac{1}{5}, \frac{1}{8}, \frac{1}{10}, \frac{1}{20}, \frac{1}{50}, 0$. In (A) $p_j = 0.05$ for $j \neq 0$, (B) $p_G = 0.2, p_{GL} = 0.1, p_j = 0.05$ for $j \neq \{0, G, GL\}$ and in (C) $p_G = 0.5, p_{GL} = 0.2, p_j = 0.05$ for $j \neq \{G, GL\}$.

but since explicit computations lead to very cumbersome expressions that do not offer straightforward ways of comparing the sensitivities of I^* to different parameters, we proceed with numerical examples.

We performed a series of numerical experiments for both models, aiming to determine the relative importance of different interventions. In particular, we tried to identify circumstances in which antibiotic therapy has substantial effects and should be used as a control measure to reduce the prevalence of colonization in an ICU. In the next section we present a couple of representative figures, along with the interpretation and the conclusions concerning implementation of control measures in ICUs. Some additional figures can be found in Appendix B.

4.4.1 Results

In reality some parameters, such as the discharge rate μ and the fraction of skin colonizations on admission p, are difficult to influence. We will therefore keep μ and p constant in the examples presented in this section, taking the same values for the deterministic and the stochastic model: $\mu = 0.1$, $p = 0.2$. The reader can find some additional figures, in which the effects of reducing p (which would correspond either to increased isolation on admission of patients that are colonized on the skin or to improved hygiene on admission) are displayed, in Appendix B.

Studying the deterministic model, we found that the efficacy of topical antimicrobial prophylaxis relies heavily on (i) the level of cross-transmission in the ICU and (ii) the proportion of individuals that are colonized in the gut on admission. This can be seen from Figure 4.2 in which three scenarios are presented; (A) low, (B) intermediate and (C) high level of intestinal colonization on admission. We observe that

- in Figure 4.2A, the percentage of individuals that are colonized in the gut on admission is very low (15%, including 5% of the total that are also already colonized on the skin). In this case, antibiotic prophylaxis may, of course, be beneficial at the individual level, but has virtually no effect on the endemic level of skin colonizations, regardless of the level of cross-transmission in the ICU.

- in Figure 4.2B, the proportion of individuals that carry the pathogen in the gut on admission is 35%, including 5% of the total that are also colonized on the skin. Here, reducing α_{GS} to zero has again almost no effect on I^* when the level of cross transmission in the ICU is high (i.e. when $(1-p)\sigma/\mu > 1$; see the top two curves), but when the cross-transmission rate is reduced so that on average the patients are not expected to transmit the infection to another patient during their stay in the ICU (the bottom curves), antibiotic prophylaxis starts to have an effect on I^*. In the case $\sigma = 0$ (i.e. when there is no patient-to-patient transmission at all in the ward), antibiotic treatment has the capacity to reduce the prevalence of colonization by about a third (prevalence is reduced from 50% to 35%; see the bottom curve in Figure 4.2B).

- antibiotic prophylaxis can, at the population level, only really be beneficial when a large proportion of patients is amenable for treatment. However, even then, the level of cross-transmission in the ICU determines the efficacy of antibiotics. We observe from Figure 4.2C that even though as much as 70% of the newly admitted patients are colonized in the gut on admission (and not yet on the skin), reducing α_{GS} to zero has almost no effect on I^* when $(1-p)\sigma/\mu > 1$ (see the top curve in Figure 4.2C). Only when $(1-p)\sigma/\mu$ is brought below one (by applying barrier precautions), do antibiotics start to have an effect on the endemic level of skin colonizations and in such a case we can observe a two-fold reduction in I^* (bottom three curves).

Figure 4.2 also demonstrates the sensitivity of I^* to the cross-transmission parameter σ. We see from Figures 4.2A and 4.2B that, while antibiotics have practically no effect on the endemic level of skin colonizations, barrier precautions that aim to reduce σ have the potential to cause a two- or even three-fold reduction in the prevalence: in 4.2A, the prevalence is reduced from 85% to 30%, regardless of the value of α_{GS}, while in 4.2B the reduction is on average 45%.

Only in the case when the proportion of individuals colonized in the gut on admission and the risk of transmitting the infection from the gut to the skin are high (Figure 4.2C), is the beneficial effect of barrier precautions less visible. In this case, antibiotic prophylaxis can significantly reduce the prevalence of skin

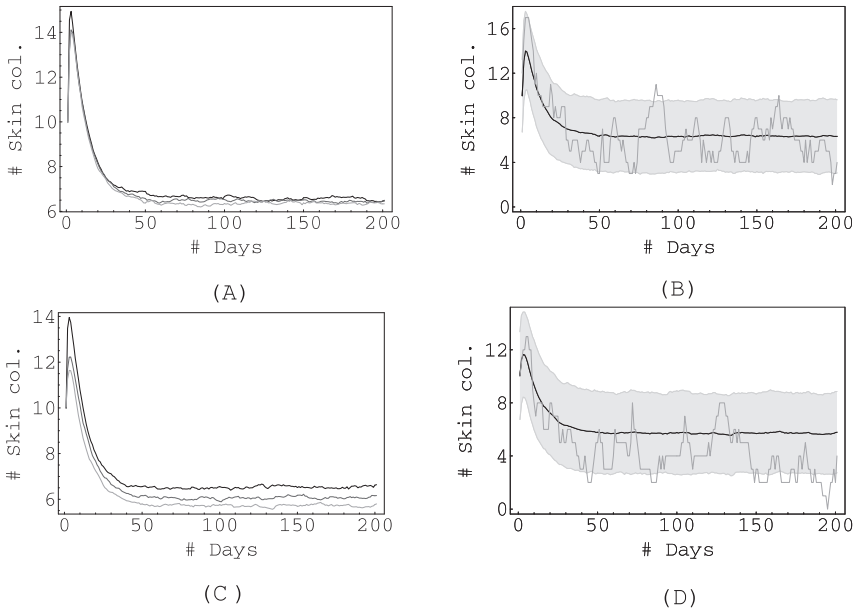

Figure 4.3: $p_j = 0.05$ for $j \neq 0$, $\alpha_{SG} = 0.5$, $\alpha_{LS} = 0.5$, $\alpha_{SL} = 0.2$. In (A) $\sigma = 1$ and $\alpha_{GS} = 1$ (black), $\alpha_{GS} = 0.1$ (dark grey), $\alpha_{GS} = 0.01$ (light grey). In (C) $\sigma = 0.01$ and $\alpha_{GS} = 1$ (black), $\alpha_{GS} = 0.1$ (dark grey), $\alpha_{GS} = 0.01$ (light grey). The right columns (figures (B) and (D)) show the mean of 1000 simulations, corresponding to $\alpha_{GS} = 0.01$, along with the shaded area containing 90% of the simulations, and the result of a single simulation.

colonizations in the ICU, but, as mentioned before, the level of cross-transmission has to be reduced first. In such a case, therefore, a combination of the two control measures would bring the best results.

A similar experiment was performed taking $\alpha_{LS} = \frac{1}{15}$ (and so, while in the scenario depicted in Figure 4.2 the patients are expected to transmit the bacteria from the lungs to the skin during their stay in the ICU, they are not expected to do so when $\alpha_{LS} = \frac{1}{15}$). The results are very similar to the ones in Figure 4.2 and will not be shown here, but the interested reader can find them in Appendix B (Figure 4.6).

We now shift our attention to the results of the simulations. In all the figures presented here, the process was simulated 1000 times for a period of 200 days, starting with a randomly chosen initial condition. Just as in the deterministic case, we study three scenarios of intestinal colonization on admission: in Figure 4.3 the proportion of individuals admitted with colonized gut is low, in Figure 4.4 it is intermediate and it is high in Figure 4.5.

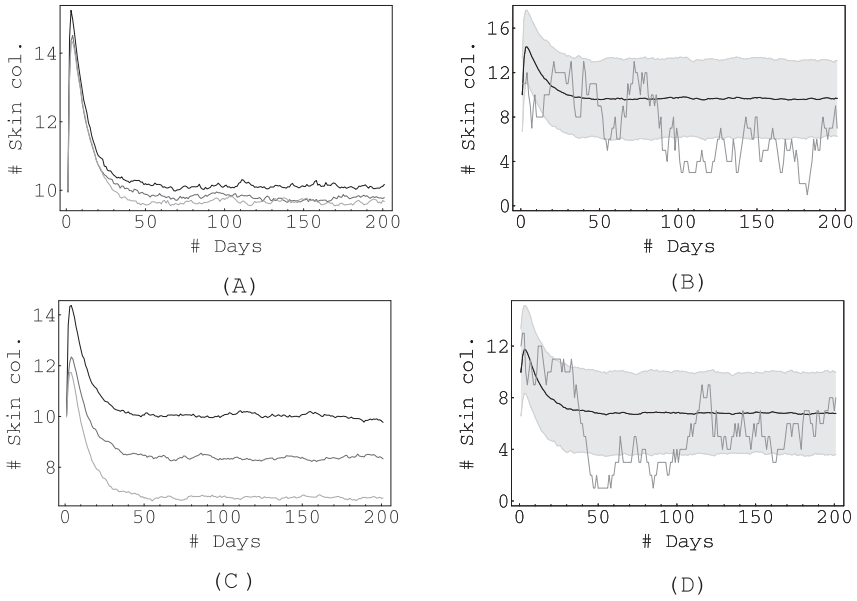

Figure 4.4: $p_G = 0.2$, $p_{GL} = 0.1$, $p_j = 0.05$ for $j \neq 0, G, GL$. $\alpha_{SG} = 0.5$, $\alpha_{LS} = 0.5$, $\alpha_{SL} = 0.2$. Top: $\sigma = 1$, bottom $\sigma = 0.01$. Left column: $\alpha_{GS} = 1$ (black), $\alpha_{GS} = 0.1$ (dark grey), $\alpha_{GS} = 0.01$ (light grey). The right columns (figures (B) and (D)) show the mean of 1000 simulations, corresponding to $\alpha_{GS} = 0.01$, along with the shaded area containing 90% of the simulations, and the result of a single simulation.

The left columns (figures 4.3A,C , 4.4A,C and 4.5A,C) show the mean of 1000 simulations for three values of the endogenous rate α_{GS}. The right columns (figures 4.3B,D , 4.4B,D and 4.5B,D) contain the bottom curve of the left column (i.e. the one corresponding to the lowest α_{GS}), the result of a single simulation and also a shaded grey area corresponding to the area containing 90 % of all simulations.

The simulations support the conclusions obtained from the deterministic model: when the admission of intestinal colonizations to the ICU is low (Figure 4.3), antibiotic prophylaxis has almost no effect on the mean level of the prevalence, regardless of the level of cross-transmission in the unit. This is also the case when the admission of intestinal colonizations is somewhat higher (see the top row in Figure 4.4). However, lowering the rate α_{GS} does start to have an effect on the mean prevalence if the cross-transmission is reduced to almost zero (see Figures 4.4A and 4.4C): while for $\sigma = 1$ half of the patients in the unit are colonized on average (regardless of the value of α_{GS}), this number drops from 10 to 7 when $\sigma = 0.01$ and we reduce α_{GS} from 1 to 0.01. The grey area (that contains 90 % of

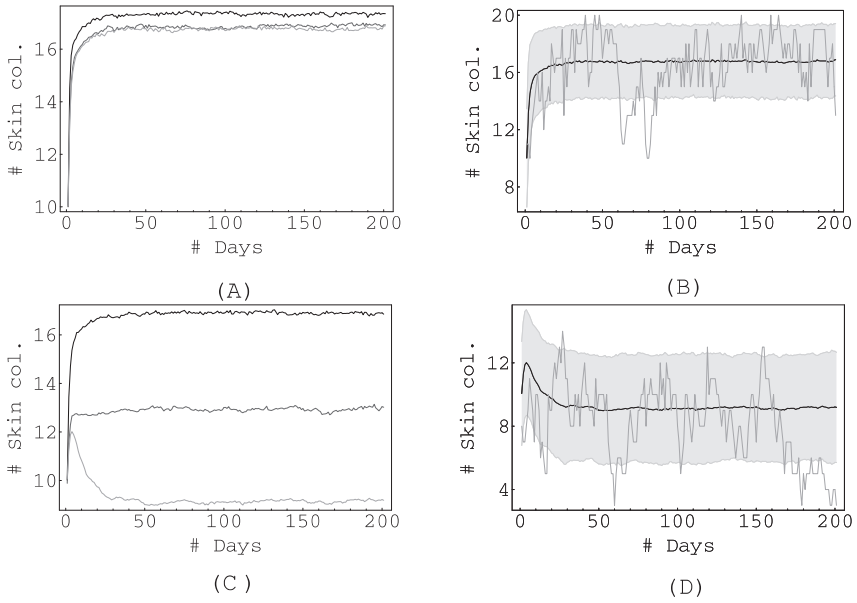

Figure 4.5: $p_G = 0.5$, $p_{GL} = 0.2$, $p_j = 0.05$ for $j \neq 0, G, GL$. $\alpha_{SG} = 0.5$, $\alpha_{LS} = 0.5$, $\alpha_{SL} = 0.2$. Top: $\sigma = 1$, bottom $\sigma = 0.01$. Left column: $\alpha_{GS} = 1$ (black), $\alpha_{GS} = 0.1$ (dark grey), $\alpha_{GS} = 0.01$ (light grey). The right columns (figures (B) and (D)) show the mean of 1000 simulations, corresponding to $\alpha_{GS} = 0.01$, along with the shaded area containing 90% of the simulations, and the result of a single simulation.

all simulations), shifts from 10 ± 3 (not shown) for $\alpha_{GS} = 1$ to 7 ± 3 for $\alpha_{GS} = 0.01$.

Figure 4.5 again shows that maintaining proper hygiene in the ICU is necessary if the antibiotics are to have an effect at all: in 4.5A, when the cross-transmission is high, antibiotics have no effect: only when σ is considerably reduced do the antibiotics have an effect on the population level. This effect is now, due to the large proportion of newly admitted individuals with intestinal colonization, considerable: on average, the number of skin colonizations in the ICU is reduced from 17 ± 3 to 9 ± 3.

When focusing on averages, therefore, the deterministic and the stochastic model echo the same message: antibiotic therapy as a means to reduce the prevalence of colonization in the ICU should only be used when a relatively high proportion of patients are colonized in the intestines on admission and furthermore the barrier precautions have already been applied to make sure that the occurrence of patient-to-patient transmission in the ICU is as low as possible. Otherwise, the benefits of antibiotic treatment may be high at the individual level, but the effect at the population level will be negligible.

Due to the low number of patients in ICUs and the rapid turnover, stochastic effects may lead to considerable fluctuations in the prevalence of skin colonization in the ICU. In 10% of the simulations that fall out of the grey area, we may observe long periods in which the prevalence remains below average, even decreases, for long periods of time before it increases again. In observing a single run (or, in real life, a single intervention in one ICU) one could, therefore, be tempted to believe that intestinal application of antibiotics has significant benefits even in the case of low/intermediate admission of intestinal colonization. The deterministic model and the figures representing the means of 1000 simulations show that this is not the case: if such an intervention seems to be successful in an ICU, it may very well be a matter of chance rather than due to real benefits of antibiotic prophylaxis.

4.5 Discussion

Complex pathogen dynamics, small patient populations and rapid patient turnover make the observation of the efficacy of different infection control measures in ICUs very difficult. Barrier precautions and intestinal administration of topical antibiotics are two measures that can be applied to reduce the prevalence of antibiotic resistant pathogens. Yet, prophylactic use of topical antibiotics is highly controversial [8, 9, 10] as it may lead to increased occurrence of resistant strains. Moreover, different control measures are usually implemented simultaneously, precluding the observation of the effects of individual measures. To amend for the lack of theoretical background in this matter, we formulated a deterministic and a discrete time Markov chain model to describe the dynamics of antibiotic resistant bacteria in ICUs. We demonstrated that intestinal application of nonabsorbable antibiotics, with the aim of reducing antibiotic resistance at the population level, can only be effective when (i) a high percentage of newly admitted patients is colonized in the intestines and (ii) cross-transmission rates in the ward are low. Our findings, therefore, represent a firm theoretical argument against the routine use of topical antimicrobial prophylaxis for infection control.

Our study can be applied to several pathogens commonly found in ICUs, such as *Pseudomonas aeruginosa*, MRSA, *enterococci* and enteric *Gram-negative* bacteria. Simultaneous colonization at multiple body sites with such pathogens is universal among ICU-patients [9]. Pathogen dynamics, therefore, includes both the within-host and patient-to-patient transmission. From a conceptual point of view, these colonization sites can be divided in internal and external compartments, distinguished by the capacity to act as a source for cross-transmission. The model used in this study was based on this concept, with the skin as the representative for all body sites that can be contacted by health care workers, and

thus be a source for cross-transmission. For the internal compartments we chose the lungs and the gastro-intestinal tract.

Under the vocal cords, the respiratory tract is usually sterile. Colonization frequently occurs in ICU patients as the physiological defence mechanisms against aspirated microorganisms are breached because of intubation. Moreover, the commensal bacterial flora of the upper respiratory tract, susceptible to many antibiotics, rapidly disappears after ICU-admission, partly because of antibiotics. Subsequently, the upper and the lower respiratory tract become colonized with the typical hospital-acquired bacterial flora, which is often resistant to many antibiotics (e.g. *Pseudomonas aeruginosa*, *Enterobacteriaceae* and MRSA). From the respiratory tract pathogens may reach the external compartment (where they can contaminate health care workers' hands) during patient care. Eradication of bacterial colonization in the oropharyngeal cavity (part of the upper respiratory tract) can be achieved with topical antimicrobial agents [6]. Yet, this approach does not completely eradicate colonization of the lower respiratory tract. In our analyses we, therefore, considered lung colonization as a constant source from where bacteria could reach the external compartment.

The intestinal tract is considered the most abundant source of pathogens causing hospital-acquired infections. As in the respiratory tract, the commensal bacterial flora, susceptible to many antibiotics, is rapidly replaced after ICU-admission by the typical hospital-acquired bacterial flora. In contrast to the respiratory tract, though, eradication of this flora is achievable with non-absorbable antibiotics.

Because of the constantly increasing prevalence of antibiotic resistance, prevention of transmission of such pathogens becomes more and more important. Barrier precautions are essential in infection prevention in hospitals. In clinical practice, such measures include optimal hand hygiene after patient contact, placing a patient in a single-bed room and approaching the patient only when wearing gloves and gowns. Barrier precautions alone have, however, frequently been insufficient to prevent the spread of antibiotic resistant pathogens and additional preventive measures are needed, especially in ICUs.

Eradication of bacterial colonization could be a very effective control measure. For most antibiotic-resistant bacteria, colonization, once established, persists until patient discharge from the ICU. After that, re-establishment of the commensal flora occurs, possibly because of reduced antibiotic selection pressure. Several eradication strategies have been implemented, though frequently without success [75]. The concept of nonabsorbable antibiotics applied into the intestinal tract, usually through a nasogastric tube, has met considerable enthusiasm [106]. It has been tested frequently as part of SDD, in which intestinal decontamination is combined with oropharyngeal decontamination and a short course of intravenous antibiotic prophylaxis. The primary goal of SDD was to prevent the development of ICU-acquired respiratory tract infections. After more

than forty randomized trials evaluating the effects of SDD, there is strong evidence that SDD indeed reduces incidences of ICU-acquired respiratory tract infections [106]. Moreover, ICU-survival among patients receiving SDD increased according to three recent studies [23, 24, 66]. The effects of SDD on the prevalence of antibiotic resistant bacteria, however, are controversial: SDD enhanced preexisting high antibiotic resistant prevalence levels [73], whereas SDD reduced such prevalence levels in units with low baseline prevalences [23]. It is generally assumed that the intestinal component of SDD has the largest effect on increasing or decreasing unit-wide prevalence levels of antibiotic resistance. This assumption is supported by the reported effects of topical antimicrobial prophylaxis, as in SDD, as a measure for controlling outbreaks of antibiotic resistance [11]. Therefore, the use of SDD has been advocated as a routine measure to control antibiotic resistance in ICUs.

Our findings, to some extent, support this concept. Yet, optimal use of barrier precautions and a relatively high admission rate of intestinal carriage of resistant bacteria are prerequisites for the efficacy of intestinal decontamination. In most ICUs a considerable part of the patients is directly admitted from the community (such as trauma patients) and the admission prevalence of colonization with resistant bacteria will be low. To conclude, the results of this study, therefore militate against the routine use of topical antimicrobial prophylaxis for infection control.

4.6 Appendix A. The deterministic model revisited

The aim of this part is to establish the number of equilibria of the system (4.1) and their stability. But first, we pay attention to the formulation of the model.

Let $N_0, N_G, N_S, N_L, N_{GS}, N_{GL}, N_{SL}, N_{GSL}$ denote the sizes of the subpopulations indicated by the indices. If the population sizes were large, we could describe the dynamics with the following system of ODEs:

$$\begin{aligned}
\dot{N}_0 &= \lambda p_0 - \beta(N_S + N_{GS} + N_{SL} + N_{GSL})N_0 - \mu N_0,\\
\dot{N}_G &= \lambda p_G - \beta(N_S + N_{GS} + N_{SL} + N_{GSL})N_G - (\alpha_{GS} + \mu)N_G,\\
\dot{N}_S &= \lambda p_S + \beta(N_S + N_{GS} + N_{SL} + N_{GSL})N_0 - (\alpha_{SG} + \alpha_{SL} + \mu)N_S,\\
\dot{N}_L &= \lambda p_L - \beta(N_S + N_{GS} + N_{SL} + N_{GSL})N_L - (\alpha_{LS} + \mu)N_L,
\end{aligned}$$

$$\begin{aligned}
\dot{N}_{GS} &= \lambda p_{GS} + \beta(N_S + N_{GS} + N_{SL} + N_{GSL})N_G - (\alpha_{SL} + \mu)N_{GS} + \\
&+ \alpha_{GS} N_G + \alpha_{SG} N_S, \\
\dot{N}_{GL} &= \lambda p_{GL} - \beta(N_S + N_{GS} + N_{SL} + N_{GSL})N_{GL} - (\alpha_{GS} + \alpha_{LS} + \mu)N_{GL}, \\
\dot{N}_{SL} &= \lambda p_{SL} + \beta(N_S + N_{GS} + N_{SL} + N_{GSL})N_L - (\alpha_{SG} + \mu)N_{SL} + \\
&+ \alpha_{SL} N_S + \alpha_{LS} N_L, \\
\dot{N}_{GSL} &= \lambda p_{GSL} + \beta(N_S + N_{GS} + N_{SL} + N_{GSL})N_{GL} + (\alpha_{LS} + \alpha_{GS})N_{GL} + \\
&+ \alpha_{SL} N_{GS} + \alpha_{SG} N_{SL} - \mu N_{GSL}.
\end{aligned}$$

The total population size,

$$N(t) = N_0(t) + N_G(t) + N_S(t) + N_L(t) + N_{GS}(t) + N_{GL}(t) + N_{SL}(t) + N_{GSL}(t)$$

changes according to the differential equation

$$\dot{N} = \lambda - \mu N$$

and so the equilibrium $N^* = \frac{\lambda}{\mu}$ is globally asymptotically stable. If we define $\sigma = \frac{\lambda\beta}{\mu}$,

$$x_j = \frac{N_j}{N^*}, \quad j \in \{0, G, S, L, GS, GL, SL, GSL\}$$

and

$$I := x_S + x_{GS} + x_{LS} + x_{GSL},$$

we can indeed rewrite the system as (4.1).

To study the steady states of (4.1), we first express the steady state values of $x_j,\ j \in \{0, G, S, L, GS, GL, SL, GSL\}$ in terms of I^* (the steady state values will be indicated by a superscript *).

We have

$$\begin{cases}
x_0^* = \dfrac{\mu p_0}{\sigma I^* + \mu} \\
x_G^* = \dfrac{\mu p_G}{\sigma I^* + \alpha_{GS} + \mu} \\
x_L^* = \dfrac{\mu p_L}{\sigma I^* + \alpha_{LS} + \mu} \\
x_{GL}^* = \dfrac{\mu p_{GL}}{\sigma I^* + \alpha_{GS} + \alpha_{LS} + \mu},
\end{cases} \tag{4.5a}$$

as the steady state values of four variables, and, using these, we can write the remaining four as

$$\begin{cases} x_S^* = \dfrac{\mu p_S}{\alpha_{SG}+\alpha_{SL}+\mu} + \dfrac{\sigma I^*}{\alpha_{SG}+\alpha_{SL}+\mu} x_0^* \\ x_{GS}^* = \dfrac{\mu p_{GS}}{\alpha_{SL}+\mu} + \dfrac{\sigma I^* + \alpha_{GS}}{\alpha_{SL}+\mu} x_G^* + \dfrac{\alpha_{SG}}{\alpha_{SL}+\mu} x_S^* \\ x_{SL}^* = \dfrac{\mu p_{SL}}{\alpha_{SG}+\mu} + \dfrac{\sigma I^* + \alpha_{LS}}{\alpha_{SG}+\mu} x_L^* + \dfrac{\alpha_{SL}}{\alpha_{SG}+\mu} x_S^* \\ x_{GSL}^* = p_{GSL} + \dfrac{\sigma I^* + \alpha_{LS} + \alpha_{GS}}{\mu} x_{GL}^* + \dfrac{\alpha_{SL}}{\mu} x_{GS}^* + \dfrac{\alpha_{SG}}{\mu} x_{SL}^*, \end{cases} \tag{4.5b}$$

which are explicit expressions in terms of I^* once we insert (4.5a).

The steady state value I^* satisfies

$$I^* = x_S^* + x_{GS}^* + x_{SL}^* + x_{GSL}^* =: F(I^*),$$

where $x_S^*, x_{GS}^*, x_{SL}^*$ and x_{GSL}^* are given in (4.5b). Hence, we need to investigate the fixed points of F.

After some manipulation we find that

$$F(I) = p + \frac{p_0 \sigma I}{\sigma I + \mu} + \frac{p_L(\sigma I + \alpha_{LS})}{\sigma I + \alpha_{LS} + \mu} + \frac{p_G(\sigma I + \alpha_{GS})}{\sigma I + \alpha_{GS} + \mu} + \frac{p_{GL}(\sigma I + \alpha_{GS} + \alpha_{LS})}{\sigma I + \alpha_{GS} + \alpha_{LS} + \mu},$$

where $p = p_S + p_{GS} + p_{SL} + p_{GSL}$ and the function F has the following properties:

1. $F(0) \geq 0$,
2. $F'(I) \geq 0$ for $I \geq 0$
3. $\lim_{I\to\infty} F(I) = 1$ and
4. $F''(I) \leq 0$ for $I \geq 0$.

These properties ensure that, when $F(0) > 0$, the function F has a unique, strictly positive fixed point I^*. The system (4.1) then has a unique strictly positive steady state. Once I^* is determined, we can calculate the steady state values of all state variables using (4.5a) and (4.5b).

Suppose now that $F(0) = 0$. If we assume that $\alpha_{GS} > 0$ and $\alpha_{LS} > 0$ then $F(0) = 0$ precisely when all newly introduced individuals are susceptible. In this case the value of $F'(0)$ determines whether there will exist a nontrivial steady state. If $F'(0) = \frac{\sigma}{\mu} = \mathcal{R}_0 > 1$ then we have a unique, strictly positive steady state and if $\mathcal{R}_0 \leq 1$ then there is only the trivial steady state.

What remains is the possibility that at least one of α_{GS}, α_{LS} equals zero. Let us only consider the case $\alpha_{GS} = 0, \alpha_{LS} > 0$ (the other two options are treated in

the same way). Then $F(0) = 0$ precisely when $p_0 + p_G = 1$, that is, all newly introduced patients are either not colonized at all or carry the bacteria in their gut. But since $\alpha_{GS} = 0$ they can not spread the bacteria to their skin and we can treat individuals in x_0, x_G as a heterogeneous susceptible population. Hence, $\mathcal{R}_0 = p_0\frac{\sigma}{\mu} + p_G\frac{\sigma}{\mu} = \frac{\sigma}{\mu} = F'(0)$ and an endemic equilibrium exists only when $\mathcal{R}_0 > 1$.

Let us now focus on the case when $F(0) > 0$ and show that the (unique) equilibrium is locally asymptotically stable. In a similar fashion we can in fact also prove the stability statements for the case $F(0) = 0$.

To this end let

$$x^* = (x_0^*, x_G^*, x_S^*, x_L^*, x_{GS}^*, x_{GL}^*, x_{SL}^*, x_{GSL}^*)$$

denote the (unique) equilibrium of (4.1). The Jacobi matrix evaluated in x^* is denoted by J and has the following form

$$J = \begin{pmatrix} J_1 & 0 & -\sigma x_0^* & 0 & -\sigma x_0^* & 0 & -\sigma x_0^* & -\sigma x_0^* \\ 0 & J_2 & -\sigma x_G^* & 0 & -\sigma x_G^* & 0 & -\sigma x_G^* & -\sigma x_G^* \\ \sigma I^* & 0 & J_3 & 0 & \sigma x_0^* & 0 & \sigma x_0^* & \sigma x_0^* \\ 0 & 0 & -\sigma x_L^* & J_4 & -\sigma x_L^* & 0 & -\sigma x_L^* & -\sigma x_L^* \\ 0 & \sigma I^* + \alpha_{GS} & \sigma x_G^* + \alpha_{SG} & 0 & J_5 & 0 & \sigma x_G^* & \sigma x_G^* \\ 0 & 0 & -\sigma x_{GL}^* & 0 & -\sigma x_{GL}^* & J_6 & -\sigma x_{GL}^* & -\sigma x_{GL}^* \\ 0 & 0 & \sigma x_L^* + \alpha_{SL} & \sigma I^* + \alpha_{LS} & \sigma x_L^* & 0 & J_7 & \sigma x_L^* \\ 0 & 0 & \sigma x_{GL}^* & 0 & \sigma x_{GL}^* + \alpha_{SL} & \sigma I^* + \alpha_{LS} + \alpha_{GS} & \sigma x_{GL}^* + \alpha_{SG} & J_8 \end{pmatrix}$$

where

$$J_1 = -(\sigma I^* + \mu), \tag{4.6a}$$
$$J_2 = -(\sigma I^* + \alpha_{GS} + \mu), \tag{4.6b}$$
$$J_3 = \sigma x_0^* - \alpha_{SG} - \alpha_{SL} - \mu, \tag{4.6c}$$
$$J_4 = -(\sigma I^* + \alpha_{LS} + \mu), \tag{4.6d}$$
$$J_5 = \sigma x_G^* - \alpha_{SL} - \mu, \tag{4.6e}$$
$$J_6 = -(\sigma I^* + \alpha_{GS} + \alpha_{LS} + \mu), \tag{4.6f}$$
$$J_7 = \sigma x_L^* - \alpha_{SG} - \mu, \tag{4.6g}$$
$$J_8 = \sigma x_{GL}^* - \mu. \tag{4.6h}$$

Let us define

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 & -1 & 1 \end{pmatrix}$$

and calculate $K = S^{-1}JS$. We obtain

$$K = \begin{pmatrix} K_1 & 0 & 0 & 0 & 0 & 0 & 0 & -\sigma x_0^* \\ 0 & K_2 & 0 & 0 & 0 & 0 & 0 & -\sigma x_G^* \\ \sigma I^* & 0 & K_3 & 0 & 0 & 0 & 0 & \sigma x_0^* \\ 0 & 0 & 0 & K_4 & 0 & 0 & 0 & -\sigma x_L^* \\ 0 & \sigma I^* + \alpha_{GS} & \alpha_{SG} & 0 & K_5 & 0 & 0 & \sigma x_G^* \\ 0 & 0 & 0 & 0 & 0 & K_6 & 0 & -\sigma x_{GL}^* \\ 0 & 0 & \alpha_{SL} & \sigma I^* + \alpha_{LS} & 0 & 0 & K_7 & \sigma x_L^* \\ \sigma I^* & \sigma I^* + \alpha_{GS} & 0 & \sigma I^* + \alpha_{LS} & 0 & \sigma I^* + \alpha_{LS} + \alpha_{GS} & 0 & K_8 \end{pmatrix}$$

where

$$K_i = J_i \quad \text{for } i = 1, 2, 4, 6 \tag{4.7}$$

and

$$K_3 = -(\alpha_{SG} + \alpha_{SL} + \mu), \tag{4.8a}$$

$$K_5 = -(\alpha_{SL} + \mu), \tag{4.8b}$$

$$K_7 = -(\alpha_{SG} + \mu), \tag{4.8c}$$

$$K_8 = -\mu + \sigma(x_0^* + x_G^* + x_L^* + x_{GL}^*). \tag{4.8d}$$

Now, it is clear from the form of K that K_3, K_5 and K_7 are eigenvalues of K (and therefore of J, since K and J are related by a similarity transformation) and from (4.8a) - (4.8c) that they are strictly negative. Our aim now is to show that all other eigenvalues of K have negative real parts as well. In fact, we will see that all the eigenvalues of K are real.

So let ϕ be an eigenvalue of K and let $v^T = (v_1, ..., v_8)$ be a corresponding eigenvector. We have the following relations:

$$0 = (K_1 - \phi)v_1 - \sigma x_0^* v_8, \tag{4.9a}$$

$$0 = (K_2 - \phi)v_2 - \sigma x_G^* v_8, \tag{4.9b}$$

$$0 = (K_4 - \phi)v_4 - \sigma x_L^* v_8, \tag{4.9c}$$

$$0 = (K_6 - \phi)v_6 - \sigma x_{GL}^* v_8, \tag{4.9d}$$

$$0 = \sigma I^* v_1 + (\sigma I^* + \alpha_{GS})v_2 + (\sigma I^* + \alpha_{LS})v_4 + (\sigma I^* + \alpha_{LS} + \alpha_{GS})v_6 + (K_8 - \phi)v_8 \tag{4.9e}$$

and

$$0 = \sigma I^* v_1 + (K_3 - \phi)v_3 + \sigma x_0^* v_8, \tag{4.9f}$$
$$0 = (\sigma I^* + \alpha_{GS})v_2 + \alpha_{SG} v_3 + (K_5 - \phi)v_5 + \sigma x_G^* v_8, \tag{4.9g}$$
$$0 = \alpha_{SL} v_3 + (\sigma I^* + \alpha_{LS})v_4 + (K_7 - \phi)v_7 + \sigma x_L^* v_8. \tag{4.9h}$$

We first consider the following situation:

CASE 1. $\alpha_{GS} > 0, \alpha_{LS} > 0, \alpha_{GS} \neq \alpha_{LS}$.

Let us first see that in this case the spectrum of K cannot contain K_i for $i \in \{1, 2, 4, 6\}$.

We show this by contradiction. Suppose that $\phi = K_1$. Since $x_0^* \neq 0$, $K_1 \neq K_2, K_4, K_6$ and $I^* \neq 0$ we see from (4.9a) - (4.9e) that $v_1 = v_2 = v_4 = v_6 = v_8 = 0$ and then from (4.9f) - (4.9h) that also $v_3 = v_5 = v_7$ and so $v = 0$, which cannot be the case since v is an eigenvector.

In the same way we can see that neither of K_2, K_4, K_6 lies in the spectrum of K. We can hence use (4.9a) - (4.9d) to express v_1, v_2, v_4, v_6 in terms of v_8. We then obtain from (4.9e) that $v_8 = 0$ or

$$\begin{aligned}\frac{\sigma^2 I^* x_0^*}{K_1 - \phi} + \frac{\sigma(\sigma I^* + \alpha_{GS})x_G^*}{K_2 - \phi} + \frac{\sigma(\sigma I^* + \alpha_{LS})x_L^*}{K_4 - \phi} \\ + \frac{\sigma(\sigma I^* + \alpha_{GS} + \alpha_{LS})x_{GL}^*}{K_6 - \phi} + K_8 - \phi = 0.\end{aligned} \tag{4.10}$$

Let us first assume that $v_8 \neq 0$. Then, using (4.6a), (4.6b), (4.6d), (4.6f), (4.7) and (4.8d) we can rewrite (4.10) as

$$(\phi + \mu)\Big(\frac{\sigma x_0^*}{-K_1 + \phi} + \frac{\sigma x_G^*}{-K_2 + \phi} + \frac{\sigma x_L^*}{-K_4 + \phi} + \frac{\sigma x_{GL}^*}{-K_6 + \phi}\Big) - (\phi + \mu) = 0. \tag{4.11}$$

Hence, either $\phi = -\mu < 0$, or

$$\frac{\sigma x_0^* I^*}{-K_1 + \phi} + \frac{\sigma x_G^* I^*}{-K_2 + \phi} + \frac{\sigma x_L^* I^*}{-K_4 + \phi} + \frac{\sigma x_{GL}^* I^*}{-K_6 + \phi} - I^* = 0. \tag{4.12}$$

Let us now write ϕ in (4.12) as $\phi = \nu + i\kappa$. Separating the real and the imaginary component, we see that $\kappa = 0$ or

$$\frac{\sigma x_0^*}{(-K_1 + \nu)^2 + \kappa^2} + \frac{\sigma x_G^*}{(-K_2 + \nu)^2 + \kappa^2} + \frac{\sigma x_L^*}{(-K_4 + \nu)^2 + \kappa^2} + \frac{\sigma x_{GL}^*}{(-K_6 + \nu)^2 + \kappa^2} = 0, \tag{4.13}$$

but since the steady state values of $x_0^*, x_G^*, x_L^*, x_{GL}^*$ are strictly positive we see that necessarily $\kappa = 0$.

Suppose now that $\nu = \phi \geq 0$. We can then estimate the left hand side of (4.12) as follows

$$\begin{aligned}
\ldots &< \frac{\sigma x_0^* I^*}{-K_1} + \frac{\sigma x_G^* I^*}{-K_2} + \frac{\sigma x_L^* I^*}{-K_4} + \frac{\sigma x_{GL}^* I^*}{-K_6} - I^* \\
&< \frac{\sigma p_0 I^*}{-K_1} + \frac{\sigma p_G^* I^*}{-K_2} + \frac{\sigma p_L I^*}{-K_4} + \frac{\sigma p_{GL} I^*}{-K_6} - I^* \\
&< p + \frac{p_0 \sigma I^*}{-K_1} + \frac{p_G(\sigma I^* + \alpha_{GS})}{-K_2} + \frac{p_L(\sigma I^* + \alpha_{LS})}{-K_4} + \frac{p_{GL}(\sigma I^* + \alpha_{GS} + \alpha_{LS})}{-K_6} - I^* \\
&= F(I^*) - I^* \\
&= 0,
\end{aligned}$$

where we have used in the second line that $x_j^* < p_j$ for $j \in \{0, G, L, GL\}$ and in the last that I^* is a fixed point of F.

Hence, in order for ϕ to be a solution of (4.12), ϕ has to be strictly negative.

Now, all that remains to be considered in this first case are the eigenvalues of K for which the corresponding eigenvector is such that $v_8 = 0$. Since we have already established that $\phi \neq K_1, K_2, K_4, K_6$ we obtain from (4.9a) - (4.9d) that also $v_1 = v_2 = v_4 = v_6 = 0$. We can then rewrite (4.9f) - (4.9h) as

$$0 = (K_3 - \phi)v_3, \tag{4.14a}$$

$$0 = \alpha_{SG} v_3 + (K_5 - \phi)v_5, \tag{4.14b}$$

$$0 = \alpha_{SL} v_3 + (K_7 - \phi)v_7 \tag{4.14c}$$

and observe that, since not all three of v_3, v_5, v_7 can be zero, the remaining three eigenvalues of K are K_3, K_5 and K_7, as we have already observed in the beginning.

Since K_3, K_5 and K_7 are all strictly negative we have now shown that in the case when $\alpha_{GS} > 0, \alpha_{LS} > 0$ and $\alpha_{GS} \neq \alpha_{LS}$ all the eigenvalues of K are real and strictly negative and so the steady state x^* is locally asymptotically stable.

To complete the proof, we have to consider the situation when the assumptions of the Case 1 are not met. This can happen in the following ways:

CASE 2. $\alpha_{GS} > 0, \alpha_{LS} > 0, \alpha_{GS} = \alpha_{LS}$,

CASE 3. $\alpha_{GS} = 0, \alpha_{LS} > 0$,

CASE 4. $\alpha_{GS} > 0, \alpha_{LS} = 0$,

CASE 5. $\alpha_{GS} = 0, \alpha_{LS} = 0$.

Since all these cases are handled in a similar manner, we shall give details for only one of them, namely, Case 2.

In the same way as before, we observe that K_1 and K_6 cannot be in the spectrum of K. In this case, however $K_2(= K_4)$ can be an eigenvalue and we will see that this necessarily is the case.

Let us first consider an eigenvalue of K, say ϕ, such that $\phi \neq K_2$. We can then express v_1, v_2, v_4 and v_6 in terms of v_8 and just as in Case 1 obtain that $v_8 = 0$, $\phi = -\mu$ or that ϕ satisfies (4.12). Since the conclusion from Case 1, that in the latter case necessarily $\phi < 0$, relied nowhere on the fact that $\alpha_{GS} \neq \alpha_{LS}$ we can carry it over right away to Case 2. Note however that, since $K_2 = K_4$, we obtain a polynomial of degree 3 and hence three (strictly negative) roots. The 'missing' eigenvalue can only be equal to K_2, which is strictly negative and so we again conclude that x^* is locally asymptotically stable.

This completes the proof.

4.7 Appendix B. Additional figures

In Figure 4.6 we show the results of the same experiment as described in Section 4.4, only this time with $\alpha_{LS} = \frac{1}{15}$.

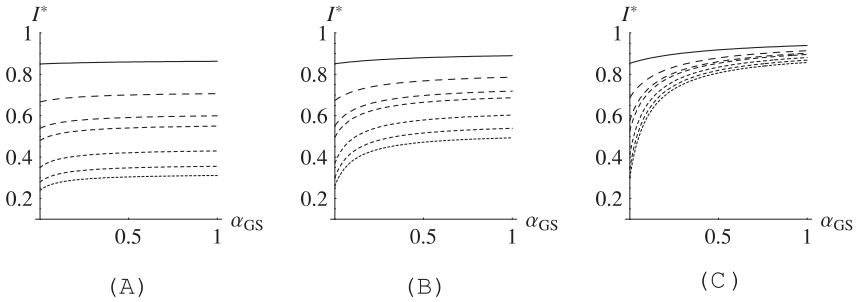

Figure 4.6: The endemic level of skin colonizations plotted as a function of α_{GS}. In all three figures $\mu = \frac{1}{10}$, $\alpha_{LS} = \frac{1}{15}$ and, from the top to the bottom, $\sigma = \frac{1}{2}, \frac{1}{5}, \frac{1}{8}, \frac{1}{10}, \frac{1}{20}, \frac{1}{50}, 0$. In (a) $p_j = 0.05$ for $j \neq 0$, (b) $p_G = 0.2, p_{GL} = 0.1, p_j = 0.05$ for $j \neq \{0, G, GL\}$ and in (c) $p_G = 0.5, p_{GL} = 0.2, p_j = 0.05$ for $j \neq \{G, GL\}$.

Figures 4.7 and 4.8 show the effects of reducing the proportion of individuals that are colonized on the skin already on admission.

We begin at the right end point of the p-axis, with a certain starting distribution $\{p_j\}$, $j \in \{0, G, S, L, GS, GL, SL, GSL\}$. Reductions in $p = p_S + p_{GS} + p_{SL} + p_{GSL}$ are uniform in the four compartments S, GS, SL, GSL and yield a uniform increase in $0, G, L, GL$.

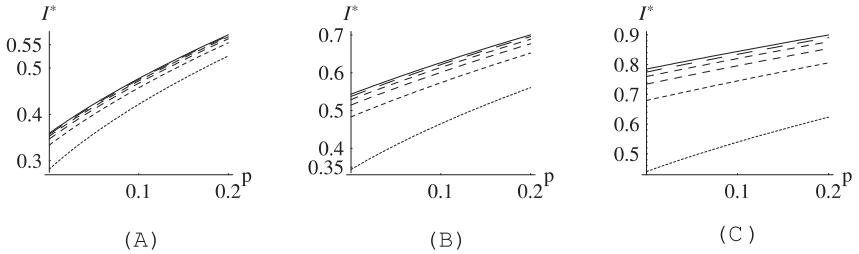

Figure 4.7: In all three figures $\sigma = 0.1$ and, from top to bottom: $\alpha_{GS} = 1, 0.8, 0.6, 0.4, 0.2, 0$. The starting probabilities at the right end point of the p-axis are (*a*) $p_j = 0.05$ for $j \neq 0$, (*b*) $p_G = 0.2, p_{GL} = 0.1, p_j = 0.05$ for $j \neq \{0, G, GL\}$ and in (*c*) $p_G = 0.5, p_{GL} = 0.2, p_j = 0.05$ for $j \neq \{G, GL\}$.

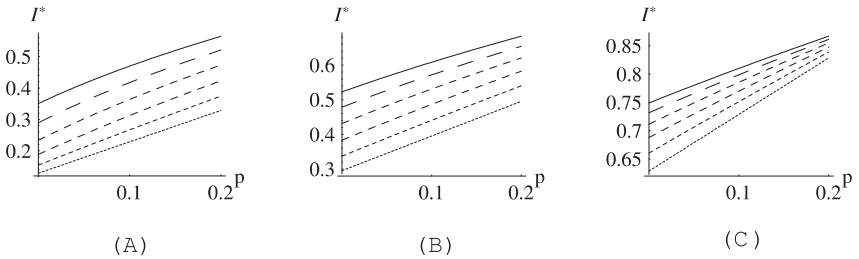

Figure 4.8: In all three figures $\alpha_{GS} = 0.5$ and, from top to bottom: $\sigma = 0.1, 0.08, 0.06, 0.04, 0.02, 0$. The starting probabilities at the right end point of the p-axis are (*a*) $p_j = 0.05$ for $j \neq 0$, (*b*) $p_G = 0.2, p_{GL} = 0.1, p_j = 0.05$ for $j \neq \{0, G, GL\}$ and in (*c*) $p_G = 0.5, p_{GL} = 0.2, p_j = 0.05$ for $j \neq \{G, GL\}$.

Chapter 5

Persistence and spread of enterotoxigenic *Escherichia coli* in the gastro-intestinal tract of piglets

5.1 Introduction

Post-weaning diarrhoea (PWD) is a multi faceted disease occurring in piglets in the first two weeks after weaning. One of the causative agents of this disease is enterotoxigenic *Escherichia coli* (ETEC) expressing $F4$ fimbriae ($F4$+ *E. coli*). ETEC can adhere to the microvilli of the small intestine enterocytes. When attached, bacteria produce enterotoxins, which act locally on the enterocytes, resulting in increased secretion of fluid and $NaHCO_3$, leading to diarrhoea, dehydration, acidosis or even death [48].

To this day, several empirical studies have investigated how the occurrence and the clinical symptoms of PWD depend on different environmental factors, such as, for instance, litter size, weaning age and temperature. In the recent field work of P. Geenen [48], data had been obtained on the shedding patterns of piglets and transmission of ETEC. The study furthermore contains some statistical analysis of the obtained data. However, there exist (to our knowledge) no theoretical studies of the dynamics of ETEC and the real insight into this disease is still lacking.

The aim of this chapter is to contribute to understanding of the dynamics of ETEC in piglets. In order to study the spread of ETEC at the (piglet) population level, one needs to be able to determine some individual characteristics, such as individual's infectiousness, and then 'lift' the dynamics at the individual level to the dynamics at the (piglet) population level. We thus start by modeling the

dynamics of ETEC in a single infected piglet. Even though our long term objective is to study more complicated models of within-host dynamics, we begin by investigating a simple linear model.

We shall in particular investigate two aspects of the within-host dynamics. Firstly, we aim to establish the conditions that guarantee growth of bacteria in the intestine and specify when ETEC will not be able to persist. Secondly, we investigate whether we observe convergence to a so called stable bacterial distribution (i.e. asynchronous exponential growth [37, 58]). We address these two questions for the case when (i) the piglet is infected with a dose of ETEC at time $t = 0$ and there is no inflow of bacteria into its intestine after $t = 0$ (we accordingly call this case the *single infection case*) and (ii) a proportion of the excreted *E-coli* is reintroduced into the intestine (here we have in mind that piglets come into contact with faeces that contain the bacteria). We call this scenario the *(self)reinfection case*.

The chapter is structured as follows. The model describing the microbial growth and movement in the intestine of a single piglet is introduced in Section 5.2. In Section 5.3 we investigate the single infection case and Section 5.4 deals with the (self)reinfection case. Some concluding remarks are collected in Section 5.5.

5.2 The model

We begin our study of the within-host dynamics of ETEC by describing the setting and introducing the notation.

The intestine is represented by a cylindrical, but not necessarily circular, tube of length L, constant cross-sectional area A and of constant circumference C. Two types of bacteria are considered: the ones attached to the wall of the intestine and free bacteria that move downstream. We assume that on every cross-section, the attached, as well as the free bacteria are uniformly distributed. Hence, if $u(\tau,\xi)$ and $v(\tau,\xi)$ denote, respectively, the distribution of attached and free bacteria at time τ, then for $0 \le \xi_1 < \xi_2 \le L$ the expressions

$$C\int_{\xi_1}^{\xi_2} u(\tau,\xi)d\xi \quad \text{and} \quad A\int_{\xi_1}^{\xi_2} v(\tau,\xi)d\xi$$

represent, respectively, the number of attached and free bacteria at time τ with positions between ξ_1 and ξ_2.

We assume that free bacteria move downstream with (at this stage assumed constant) velocity g. They reproduce at a constant rate r_2 and attach to the wall of the intestine at a constant rate b.

Notation	Meaning
A	cross - sectional area of the intestine
C	circumference of the intestine
L	length of the intestine
$Cu(\tau,\xi)$	per unit of length density of the attached bacteria
$Av(\tau,\xi)$	per unit of length density of free bacteria
a	wall detachment rate
b	wall attachment rate
g	stream velocity
r_1	reproduction rate of attached bacteria
r_2	reproduction rate of free bacteria
p_1	probability with which the offspring of an attached bacterium is attached
p_2	probability with which the offspring of an attached bacterium is free

Table 5.1: Parameters and their description

Attached bacteria reproduce at a rate r_1. Their offspring is either attached or free; the former occurs with a constant probability p_1 and the latter with a constant probability p_2. The rate at which attached bacteria detach from the wall is denoted by a.

If we define

$$r_{11} := r_1 p_1 \qquad \text{and} \qquad r_{12} := r_1 p_2, \tag{5.1}$$

then the above assumptions lead to the following (linear) system of PDEs,

$$\begin{aligned} u_\tau &= (r_{11} - a)u + \tfrac{bA}{C}v \\ v_\tau &= (r_{12} + \tfrac{aC}{A})u + (r_2 - b)v - g v_\xi. \end{aligned} \tag{5.2}$$

For the reader's convenience, we describe the notation once more in Table 5.1.

Let us first rescale space and time so that both the length of the intestine and the stream velocity equal one. In other words, we introduce new independent (and dimensionless) variables t and x defined by

$$t := \frac{g\tau}{L} \quad \text{and} \quad x := \frac{\xi}{L} \tag{5.3}$$

and rewrite (5.2) as

$$\begin{aligned} u_t &= \alpha_{11}u + \alpha_{12}v \\ v_t &= \alpha_{21}u + \alpha_{22}v - v_x, \end{aligned} \tag{5.4}$$

with

$$\alpha_{11} = \frac{(r_{11} - a)L}{g}, \quad \alpha_{12} = \frac{bAL}{gC}, \quad \alpha_{21} = \frac{(r_{12}A + aC)L}{gA}, \quad \alpha_{22} = \frac{(r_2 - b)L}{g}. \tag{5.5}$$

Note that, while α_{12} and α_{21} must be nonnegative in order to be biologically meaningful, α_{11} and α_{22} can, in principle, take any value. We thus make the following

Assumption. $\alpha_{12} \geq 0$ and $\alpha_{21} \geq 0$.

5.3 Persistence and spread of bacteria following a single introduction

Suppose that bacteria are introduced into an infection free intestine. Our first aim is to find the conditions under which the bacteria will be able to persist in the intestine and find out when their washout is inevitable.

We hence take

$$\begin{aligned} u(0, x) &= 0, \\ v(0, x) &= 0, \qquad \text{for } x \in (0, 1) \end{aligned} \tag{5.6}$$

as the initial condition that corresponds to an infection free intestine at time $\tau = 0$, and consider a sudden introduction of bacteria at $\tau = 0$ and $\xi = 0$. Mathematically we express this by choosing for the boundary condition the Dirac measure $gAv(\tau, 0) = \delta_0(\tau)$. Since $c\delta_0(cz) = \delta_0(z)$, we can use (5.3) to prescribe the flux into the intestine as a function of t as,

$$LAv(t, 0) = \delta_0(t). \tag{5.7}$$

Remark 5.1. The boundary condition (5.7) requires that we look for a solution of (5.4), which is not a function in the traditional sense of the word, but rather a *generalized function*, i.e. a *distribution* [94, 96]. Here we only briefly give the idea behind generalized functions and describe in what sense the differentiation in (5.4) should be understood. We refer the reader to [94] for a more extensive treatment of the subject.

Let Ω be a nonempty set in $\mathbb{R}^n$. If $f(x)$ is a function on Ω, we can define a linear mapping

$$\phi \mapsto \int_\Omega f(x)\phi(x)dx \tag{5.8}$$

for a suitable class of functions, which are called *test functions*. When the linear map in (5.8) satisfies some continuity conditions (see [94], Chapter 5, Definition 5.7), it is called a *distribution* or a *generalized function*. Test functions are

required to vanish near the boundary of Ω, and the derivative $\frac{\partial f}{\partial x_j}$ can then be defined as the mapping

$$\phi \mapsto -\int_\Omega f(x)\frac{\partial \phi(x)}{\partial x_j}dx.$$

We thus require no differentiability of f in the usual sense; the only requirement is that ϕ is differentiable and we hence choose the test functions to be as smooth as needed. Note, however, that, if the function f is differentiable in the usual sense, then this definition agrees with the classical derivative, as can be seen by integration by parts.

Let us now focus on obtaining a solution of (5.4). One way of obtaining the solution of (5.4) that satisfies (5.6) and (5.7) is to employ the Laplace transform of generalized functions. Although this is not too difficult, we choose to take another route here, namely, we consider smaller, easier to handle building blocks that eventually lead to the explicit solution of the problem and moreover allow for more interpretation.

The idea is the following. Initially, only free bacteria are introduced at $x = 0$ and there are no bacteria present inside the intestine. The equation

$$v_t^{(1)} = \alpha_{22}v^{(1)} - v_x^{(1)}, \tag{5.9a}$$

with conditions

$$\begin{aligned} v^{(1)}(0,x) &= 0, \qquad x \in (0,1) \\ LAv^{(1)}(t,0) &= \delta_0(t), \end{aligned} \tag{5.9b}$$

hence describes the dynamics of the initial free bacteria and the part of their complete offspring (not only the first generation) that never attaches to the wall nor has any ancestors that ever attached to the wall.

Let $u^{(1)}(t,x)$ denote the distribution of the 'first generation' of attached bacteria. The function $u^{(1)}(t,x)$ satisfies the equation

$$u_t^{(1)} = \alpha_{11}u^{(1)} + \alpha_{12}v^{(1)} \tag{5.10a}$$

with initial condition

$$u^{(1)}(0,x) = 0, \quad x \in (0,1). \tag{5.10b}$$

After (5.9) has been solved, (5.10) represents a simple linear ordinary differential equation. When the solution of (5.9) and (5.10) is found, we can proceed

inductively to obtain the distribution of the free bacteria produced after the n-th detachment from the wall and before the $(n+1)$- th attachment to the wall, $v^{(n)}(t,x)$, and the distribution of the attached bacteria that are produced after the n-th stop on the wall and before the next detachment from the wall, $u^{(n)}(t,x)$, as the solutions of the equations

$$v_t^{(n+1)} = \alpha_{21}u^{(n)} + \alpha_{22}v^{(n+1)} - v_x^{(n+1)} \tag{5.11a}$$

$$u_t^{(n+1)} = \alpha_{11}u^{(n+1)} + \alpha_{12}v^{(n+1)} \tag{5.11b}$$

with $n \in \mathbb{N}$ and where we furthermore require that for every $n > 1$

$$\begin{aligned} u^{(n)}(0,x) = v^{(n)}(0,x) = 0, \qquad x \in (0,1) \\ v^{(n)}(t,0) = 0. \end{aligned} \tag{5.11c}$$

The solution of (5.4) that satisfies (5.6) and (5.7) is then obtained by summing the contributions of all 'generations',

$$u(t,x) = \sum_{n=1}^{\infty} u^{(n)}(t,x), \tag{5.12a}$$

$$v(t,x) = \sum_{n=1}^{\infty} v^{(n)}(t,x). \tag{5.12b}$$

Let us therefore begin by finding the solution of (5.9). To do so we introduce a new independent variable and instead of (t,x) work with (s,x), where s is the 'traveling' coordinate defined by

$$s := t - x. \tag{5.13}$$

With this substitution, (5.9a) becomes an ordinary differential equation (we thus employ the so called *method of characteristics* [94]). Taking into account (5.9b), we find that the solution takes the form

$$v^{(1)}(t,x) = \tfrac{1}{LA}\delta_0(t-x)e^{\alpha_{22}x}. \tag{5.14}$$

Having obtained (5.14), we can now solve (5.10). We find that

$$u^{(1)}(t,x) = \tfrac{1}{LA}\alpha_{12}e^{(\alpha_{22}-\alpha_{11})x}e^{\alpha_{11}t}H(t-x), \tag{5.15}$$

where H denotes the Heaviside function,

$$H(x) = \begin{cases} 1\ ;\ x > 0 \\ 0\ ;\ x \leq 0\ . \end{cases}$$

The second generation of free bacteria, $v^{(2)}(t,x)$, is again obtained by employing the change of variables given in (5.13). Taking into account (5.11c), we find that

$$v^{(2)}(t,x) = \tfrac{1}{LA}\alpha_{12}\alpha_{21}xe^{(\alpha_{22}-\alpha_{11})x}e^{\alpha_{11}t}H(t-x). \tag{5.16}$$

We then proceed in the same manner to find the solution of (5.11) for $n \geq 1$. Using (5.14), (5.15), (5.16) and the fact that $H'(x) = \delta_0(x)$ it is simple to check by induction on n that for every $n \in \mathbb{N}$

$$u^{(n)}(t,x) = \tfrac{1}{LA}\alpha_{12}^{n}\alpha_{21}^{n-1}\tfrac{x^{n-1}}{(n-1)!}\tfrac{(t-x)^{n-1}}{(n-1)!}H(t-x)e^{(\alpha_{22}-\alpha_{11})x}e^{\alpha_{11}t} \tag{5.17}$$

gives the distribution of the n-th generation of attached bacteria, while (5.14) along with

$$v^{(n)}(t,x) = \tfrac{1}{LA}\alpha_{12}^{n-1}\alpha_{21}^{n-1}\tfrac{x^{n-1}}{(n-1)!}\tfrac{(t-x)^{n-2}}{(n-2)!}H(t-x)e^{(\alpha_{22}-\alpha_{11})x}e^{\alpha_{11}t} \tag{5.18}$$

for $n > 1$ represent the distributions of different generations of free bacteria.

According to (5.12) we obtain $u(t,x)$ and $v(t,x)$ by summation of, respectively, $u^{(n)}$ and $v^{(n)}$ on n. In both sums we recognize modified Bessel functions of the first kind,

$$I_\mu(z) = \Big(\frac{z}{2}\Big)^{\mu}\sum_{n=0}^{\infty}\frac{z^{2n}}{2^{2n}n!\Gamma(n+\mu+1)},$$

and we can now write $u(t,x)$ and $v(t,x)$ as

$$u(t,x) = \tfrac{1}{LA}\alpha_{12}e^{(\alpha_{22}-\alpha_{11})x}e^{\alpha_{11}t}I_0\Big(2\sqrt{\alpha_{12}\alpha_{21}x(t-x)}\Big)H(t-x), \tag{5.19a}$$

and

$$\begin{aligned} v(t,x) = {} & \tfrac{1}{LA}\delta_0(t-x)e^{\alpha_{22}x} + \\ & + \tfrac{1}{LA}\alpha_{12}e^{(\alpha_{22}-\alpha_{11})x}e^{\alpha_{11}t}\sqrt{\frac{\alpha_{12}\alpha_{21}}{t-x}}I_1\Big(2\sqrt{\alpha_{12}\alpha_{21}x(t-x)}\Big)H(t-x). \end{aligned} \tag{5.19b}$$

Note that, since the convergence radius of I_μ is infinite, (5.19a) and (5.19b) give for every $x \in [0,1]$ and every $t \geq 0$ a well defined solution of the problem.

To study the behaviour of the solution for $t \to \infty$ we note that for a fixed μ and $z \gg \mu$ we have

$$I_\mu(z) \sim \frac{e^z}{\sqrt{2\pi z}}.$$

It is then clear that the sign of α_{11} determines whether bacteria persist or not. Since $\alpha_{11} = (r_{11} - a)L/g$ we conclude the following (see also Figures 5.1 - 5.3):

(i) if $r_{11} \geq a$, i.e., if the reproduction of bacteria on the wall exceeds, or equals, the rate at which bacteria detach from the wall, then bacteria in the intestine persist and

(ii) if $r_{11} < a$, bacteria are washed out.

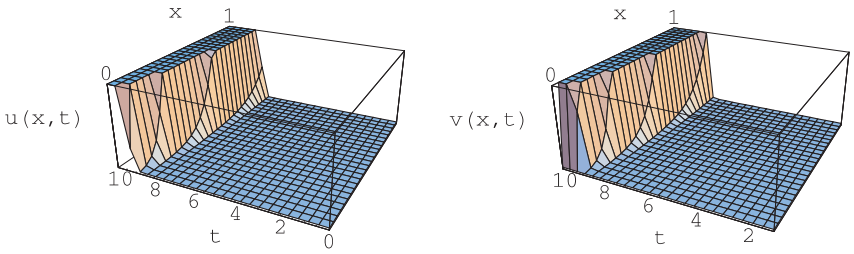

Figure 5.1: The case $r_{11} - a > 0$. Bacteria in the intestine persist.

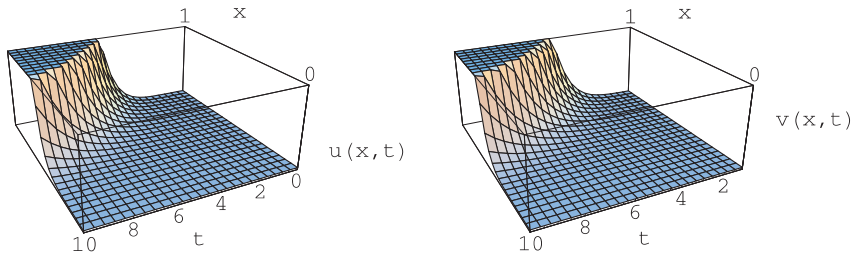

Figure 5.2: The case $r_{11} - a = 0$. Bacteria persist also in this case.

5.4 Persistence and spread of ETEC in the case of (self)reinfection

The aim of this section is to investigate the dynamics in the intestine when a proportion, say ε, of the excreted bacteria is reintroduced into the individual's gut. We therefore study (5.4) with initial conditions

$$u(0, x) = u_0(x),$$
$$v(0, x) = v_0(x),$$

and the boundary condition

$$v(t, 0) = \varepsilon v(t, 1),$$

which ensures that our requirement is met: a fraction ε of the bacteria shed at $x = 1$ is reintroduced into the intestine. The initial conditions are biologically meaningful if $u_0(x)$ and $v_0(x)$ are nonnegative functions such that the total amount of the attached, as well as the free, bacteria in the intestine is finite, i.e.,

$$C \int_0^1 u_0(x)dx < \infty \qquad \text{and} \qquad A \int_0^1 v_0(x)dx < \infty.$$

We thus require that $0 \le u_0 \in \mathcal{L}^1[0, 1]$ and $0 \le v_0 \in \mathcal{L}^1[0, 1]$.

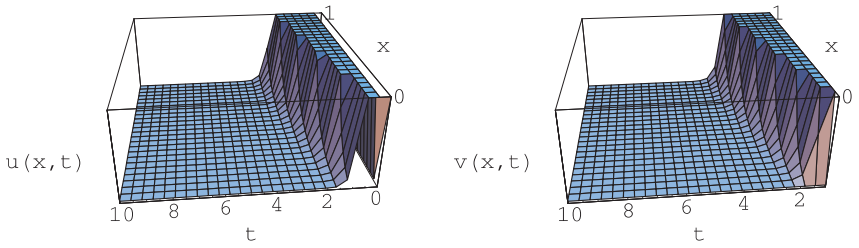

Figure 5.3: The case $r_{11} - a < 0$. Bacteria are washed out.

Our first step is to introduce the following transformation. We define $U(t, x)$ and $V(t, x)$ as

$$\begin{aligned} U(t, x) &:= \varepsilon^x u(t, x) \\ V(t, x) &:= \varepsilon^x v(t, x), \end{aligned} \tag{5.20}$$

which leads to the following evolution problem (EP)

$$(EP) \quad \begin{cases} U_t = \alpha_{11} U + \alpha_{12} V \\ V_t = \alpha_{21} U + (\alpha_{22} + \ln \varepsilon) V - V_x \\ U(0, x) = U_0(x), V(0, x) = V_0(x) \\ V(t, 0) = V(t, 1), \end{cases} \tag{5.21}$$

where, by definition, $U_0 = \varepsilon^x u_0$ and $V_0 = \varepsilon^x v_0$.

With this scaling, the bacteria shed at $x = 1$ are all reintroduced at $x = 0$, i.e. $V(t, 0) = V(t, 1)$ and we can thus view (5.21) with the domain for x being a circle instead of an interval (see also Remark 5.3).

To obtain information about the existence and the qualitative properties of a solution of (5.21), we employ some known results from the theory of (positive) one-parameter semigroups. Since our goal is to look for solutions of (5.21) that are integrable functions on $[0, 1]$, we first make the following definition.

Let $\mathcal{X}$ denote the Banach space $\mathcal{L}^1[0, 1] \times \mathcal{L}^1[0, 1]$ and let A be the (unbounded) operator on $\mathcal{X}$ defined by

$$\begin{aligned} A \begin{pmatrix} f \\ g \end{pmatrix} &= \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} + \ln \varepsilon - \frac{d}{dx} \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix} \qquad \text{with} \\ \mathcal{D}(A) &= \{(f, g) \in \mathcal{X} \mid g \text{ is absolutely continuous}, \ g(0) = g(1)\}. \end{aligned} \tag{5.22}$$

Remark 5.2. If (f, g) is to be a solution of (5.21), the function g needs to be differentiable (in the sense of distributions). Elements of $\mathcal{L}^1[0, 1]$ that are differentiable in the sense of distributions form the so called Sobolev space $\mathcal{W}^{1,1}[0, 1]$.

It turns out [43, 94] that $g \in \mathcal{W}^{1,1}[0,1]$ precisely when g is absolutely continuous, i.e. of the form $g(x) = g(0) + \int_0^x h(y)dy$ for some $h \in \mathcal{L}^1[0,1]$.

We can thus reformulate (5.21) as the abstract Cauchy problem (ACP)

$$(ACP) \quad \begin{cases} \begin{pmatrix} \dot{\phi}(t) \\ \dot{\psi}(t) \end{pmatrix} = A \begin{pmatrix} \phi(t) \\ \psi(t) \end{pmatrix} \\ \phi(0) = \phi_0, \\ \psi(0) = \psi_0. \end{cases} \tag{5.23}$$

Our aim is now to show that the operator A generates a strongly continuous semigroup $\{T(t)\}_{t\geq 0}$, which will, by

$$\begin{pmatrix} \phi(t) \\ \psi(t) \end{pmatrix} = T(t) \begin{pmatrix} \phi_0 \\ \psi_0 \end{pmatrix}$$

give the solution of the problem for every $(\phi_0, \psi_0) \in \mathcal{D}(A)$.

Note that A is a closed, densely defined linear operator, which is necessary if A is to generate a strongly continuous semigroup (cf. [43], Chapter II, Theorem 1.4).

5.4.1 Existence and uniqueness of a solution

Before embarking on the existence of a solution of the evolution problem given in (5.21), we make the following definitions.

Definition 5.1. *A family $\{T(t)\}_{t\geq 0}$ of bounded linear operators on a Banach space X is called a* **strongly continuous (one-parameter) semigroup** *if it satisfies the functional equation*

$$\begin{aligned} T(t+s) &= T(t)T(s) \quad \text{for all} \quad t, s \geq 0, \\ T(0) &= I \end{aligned}$$

and if furthermore the maps

$$\begin{aligned} \mathbb{R}_+ &\to X \\ t &\mapsto T(t)x \end{aligned}$$

are continuous for every $x \in X$.

An important (in problems arising from biology) subclass are the so called positive semigroups. To introduce positivity, one first needs a notion of an ordering on a Banach space. We thus first introduce the notion of a Banach lattice [98].

Definition 5.2. *Let $(X, \leq)$ be an ordered vector space (where the binary relation "$\leq$" is assumed to be reflexive, antisymmetric and transitive), that is, a vector space in which the following holds*

$$\begin{aligned} f \leq g \Rightarrow\ & f + h \leq g + h \quad \text{for all} \quad f, g, h \in X \\ f \leq g \Rightarrow\ & \lambda f \leq \lambda g \qquad \text{for all} \quad f, g \in X,\ \lambda \in \mathbb{R}_+. \end{aligned}$$

The vector space $(X, \leq)$ is called a **vector lattice**, *if for every $f, g \in (X, \leq)$ the elements $\sup\{f, g\}$ and $\inf\{f, g\}$ belong to $(X, \leq)$. A* **Banach lattice** *is a vector lattice that is also a Banach space equipped with a norm $\|.\|$, for which the monotonicity condition*

$$f \leq g \Rightarrow \|f\| \leq \|g\|$$

holds for all $f, g \in (X, \leq)$.

Note that $\mathcal{L}^1[0,1]$ is a Banach lattice with the ordering $f \leq g \iff f(x) \leq g(x)$ a.e. and the standard norm on $\mathcal{L}^1[0,1]$, defined as $\|f\| = \int_0^1 |f(x)|dx$.

Definition 5.3. *A strongly continuous semigroup $\{T(t)\}_{t\geq 0}$ on a Banach lattice is called* **positive**, *if every operator $T(t)$ is positive, i.e., if*

$$0 \leq f \in X \quad \text{implies} \quad 0 \leq T(t)f \quad \text{for every} \ \ t \geq 0.$$

Definition 5.4. *Let $\{T(t)\}_{t\geq 0}$ be a strongly continuous semigroup on a Banach space X. The* **(infinitesimal) generator** *of the semigroup is the operator A defined by*

$$Af = \lim_{h \downarrow 0} \frac{T(h)f - f}{h}$$

and the domain of A is the set of all $f \in X$ for which this limit exists.

By showing that the operator given in (5.22) generates a strongly continuous semigroup we can then use the following result (see [43], Chapter II, Proposition 6.2) to establish existence and uniqueness of a classical solution of EP.

Theorem 5.1. *Let $(A, \mathcal{D}(A))$ be the generator of a strongly continuous semigroup $\{T(t)\}_{t\geq 0}$. Then, for every $f \in \mathcal{D}(A)$, the function*

$$t \mapsto T(t)f$$

is the unique classical solution of ACP.

One way of showing that A indeed generates a strongly continuous semigroup of bounded linear operators on $\mathcal{X}$, is to employ the Hille-Yosida theorem ([43], Chapter II, Theorem 3.5). The conditions of the Hille - Yosida theorem are often

very difficult to check. We thus prefer another approach here, namely, we construct the generator and the corresponding semigroup by decoupling the processes taking place in the intestine (the idea of decoupling the 'movement' and 'growth' in the semigroup framework goes back to 1984 with the work of Diekmann et al. on the stability of the cell size distribution [37]).

We first write A as

$$A = A_0 + B$$

where

$$A_0 \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} \alpha_{11} & 0 \\ 0 & \alpha_{22} + \ln \varepsilon - \frac{d}{dx} \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix} \tag{5.24a}$$

with $\mathcal{D}(A_0) = \mathcal{D}(A)$ and

$$B \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} 0 & \alpha_{12} \\ \alpha_{21} & 0 \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix}. \tag{5.24b}$$

Note that the operator A_0 is an unbounded operator whereas B is a bounded operator (defined on $\mathcal{X}$), which we consider as a perturbation of A_0.

Clearly, the diagonal operator $(A_0, \mathcal{D}(A))$ generates a positive, strongly continuous semigroup $\{T_0(t)\}_{t\geq 0}$, which is given explicitly by

$$T_0(t) \begin{pmatrix} f \\ g \end{pmatrix}(x) = \begin{pmatrix} e^{\alpha_{11} t} f(x) \\ e^{(\alpha_{22} + \ln \varepsilon)t} g(x - t + k) \end{pmatrix} \tag{5.25}$$

with the unique $k \in \mathbb{N}$ for which $x - t + k \in [0, 1)$. Namely, since according to A_0 bacteria do not change their state after birth (i.e., there is no attaching or falling off of the intestine wall) the distribution of the attached bacteria at time t is obtained simply by multiplying the initial distribution by $e^{\alpha_{11} t}$, while the free bacteria at time t and position x with $x - t \geq 0$ either must have been at position $x - t$ at time 0, or are (moving) descendants of (free) bacteria that were at position $x - t$ at time 0. If $x - t < 0$, then bacteria must have reentered the intestine at least once. In such a case, the position and the number of reentries k are determined so that $x - t + k \in [0, 1)$.

Remark 5.3. Note that the introduction of new dependent variables in (5.20) made sure that for $(f, g) \in \mathcal{D}(A)$ we have $g(0) = g(1)$. Identifying the points $x = 0$ and $x = 1$, we can then look at the operator $(A_0, \mathcal{D}(A))$ as an operator describing the reproduction and movement of bacteria on a circle

$$\Gamma = \{z \in \mathbb{C}, |z| = 1\}$$

and with

$$\tilde{f} : \Gamma \to \mathbb{R}$$
$$\tilde{f}(e^{2\pi ix}) := f(x)$$

rewrite the semigroup $\{T_0(t)\}_{t\geq 0}$ as

$$T_0(t)\begin{pmatrix} \tilde{f} \\ \tilde{g} \end{pmatrix}(e^{2\pi ix}) = \begin{pmatrix} e^{\alpha_{11}t}\tilde{f}(e^{2\pi ix}) \\ e^{(\alpha_{22}+\ln\varepsilon)t}\tilde{g}(e^{2\pi i(x-t)}) \end{pmatrix}. \tag{5.26}$$

Since B is a bounded operator we can use the following Theorem (see [43], Chapter III, Theorem 1.3) to show that also the operator $(A, \mathcal{D}(A))$ generates a strongly continuous semigroup on $\mathcal{X}$. We shall denote this semigroup by $\{T(t)\}_{t\geq 0}$.

Theorem 5.2. *(Bounded Perturbation Theorem) Let $(A_0, \mathcal{D}(A_0))$ be the generator of a strongly continuous semigroup $\{T_0(t)\}_{t\geq 0}$ on a Banach space $\mathcal{X}$ satisfying*

$$\|T_0(t)\| \leq Me^{\omega t}$$

for all $t \geq 0$ and some $\omega \in \mathbb{R}$, $M \geq 1$. If $B \in \mathcal{L}(\mathcal{X})$ then

$$A = A_0 + B \quad \textit{with} \quad \mathcal{D}(A) = \mathcal{D}(A_0)$$

generates a strongly continuous semigroup $\{T(t)\}_{t\geq 0}$ on $\mathcal{X}$ that satisfies

$$\|T(t)\| \leq Me^{(\omega + M\|B\|)t}$$

for every $t \geq 0$.

According to [43] (Chapter III, Theorem 1.10), we can construct $\{T(t)\}_{t\geq 0}$ as

$$T(t) = \sum_{k=0}^{\infty} T_k(t) \tag{5.27}$$

where T_0 is given in (5.26) and the operators T_k are for $k > 0$ obtained by

$$T_{k+1}(t) := \int_0^t T_0(t-s)BT_k(s)ds. \tag{5.28}$$

For $k = 0$, $T_0(s)(f, g)^T$ gives the distribution of attached and free bacteria at time s, given that none of the initial bacteria (the distribution of which is given by $(f, g)^T$) and none of their offspring change their state in the time interval

$[0, s]$. Applying B corresponds to the bacteria changing their states and hence $T_1(t)(f, g)^T$ describes the distribution of the first generation of attached and free bacteria in exactly the same sense as we have described it in Section 5.3. Similar interpretation can be obtained for operators T_{k+1} with $k > 0$ and we can thus, by using (5.27), (5.28) and Theorem 5.1 to write the solution of the ACP, obtain the generation expansion in the same sense as in Section 5.3.

Theorem 5.1 guarantees existence and uniqueness of the solution of the ACP and accordingly also of the evolution problem. Our next aim is to investigate the qualitative properties of the solution.

Using the generation expansion, we can compute the solution for finite times t, but this alone gives us no information about the behaviour of the solution for $t \to \infty$. In order to study the asymptotic behaviour of the solution, we shall determine the *growth bound* of the semigroup defined as

$$\omega_0(T(t)) = \inf\{\omega : \exists C < \infty \text{ so that } \|T(t)\| \leq Ce^{\omega t} \text{ for all } t \geq 0\}.$$

Since the semigroup is completely characterized by its generator, one sometimes writes $\omega_0(A)$ instead of $\omega_0(T(t))$.

We shall in fact show that the growth bound equals the spectral bound of the generator, defined as

$$s(A) = \sup\{\mathrm{Re}(\lambda) \mid \lambda \in \sigma(A)\}. \tag{5.29}$$

We shall do so by using the known result (cf. [15], Chapter 8, Theorem 8.6) that

$$\omega_0(A) = \max\{s(A), \omega_{\mathrm{ess}}(A)\}, \tag{5.30}$$

where ω_{ess} denotes the *essential growth bound* (introduced shortly). Let us thus begin by determining the spectrum of A.

5.4.2 The spectrum of the infinitesimal generator

Let us begin by recalling the notion of the spectrum and of the resolvent set.

Definition 5.5. *Let X be a Banach space and let $(L, \mathcal{D}(L))$ be a linear operator on X. The* **resolvent set** *$\rho(L)$ is a set of all $\lambda \in \mathbb{C}$, for which the operator $(\lambda I - L)^{-1}$ exists and is bounded. The complement of the resolvent set $\sigma(L) = \mathbb{C} \setminus \rho(L)$ is called the* **spectrum** *of L. The spectrum can be divided into three (disjoint) sets, (i) the* **point spectrum** *$\sigma_p(L)$, the* **continuous spectrum** *$\sigma_c(L)$ and the* **residual spectrum** *$\sigma_r(L)$ characterized as follows:*

$$\sigma_p(L) = \{\lambda;\ \lambda I - L \text{ is not injective}\}$$
$$\sigma_c(L) = \{\lambda;\ \lambda I - L \text{ is injective and has a dense range}\}$$
$$\sigma_r(L) = \{\lambda;\ \lambda I - L \text{ is injective and its range is not dense in } X\}$$

The elements of the point spectrum are also called **eigenvalues** *and* $f \in \mathcal{D}(L)$ *for which* $Lf = \lambda f$ *is called an* **eigenfunction** *corresponding to eigenvalue* λ.

The **spectral radius** of an operator L, $r(L)$, is defined as

$$r(L) = \sup\{|\lambda| \; ; \; \lambda \in \sigma(L)\}$$

and the **boundary spectrum**, σ^+, is the set

$$\sigma^+(L) = \{\lambda \; ; \; |\lambda| = r(L)\}.$$

With these definitions in mind we can now characterize the spectrum of the generator $(A, \mathcal{D}(A))$. We distinguish two cases. Let us begin with

Case I. $\alpha_{12} \neq 0, \alpha_{21} \neq 0$.

and first determine the point spectrum of A in (5.22).

The point spectrum of $(A, \mathcal{D}(A))$

By definition, $\lambda \in \sigma_p(A)$ if there exists a nonzero $(f, g) \in \mathcal{D}(A)$ such that

$$(\alpha_{11} - \lambda)f + \alpha_{12}g = 0 \tag{5.31a}$$

$$\alpha_{21}f + (\alpha_{22} + \ln\varepsilon - \lambda)g - g' = 0. \tag{5.31b}$$

We first observe that if $\lambda \in \sigma_p(A)$ then $\lambda \neq \alpha_{11}$. Were that not the case, we would first obtain from (5.31a) that $g = 0$ and then from (5.31b) that also $f = 0$, which cannot be the case.

Now, if $\lambda \neq \alpha_{11}$, we can express g in (5.31a) in terms of f and rewrite (5.31b) as

$$g' = \Big[\frac{\alpha_{12}\alpha_{21}}{\lambda - \alpha_{11}} + \alpha_{22} + \ln\varepsilon - \lambda\Big]g. \tag{5.32}$$

Taking into account the restriction $g(0) = g(1)$ we find that λ is an eigenvalue of A precisely when either $g = 0$ (but since this would lead to $f = 0$ we can dismiss this option) or when λ satisfies the equation

$$\chi(\lambda) := \frac{\alpha_{12}\alpha_{21}}{\lambda - \alpha_{11}} + \alpha_{22} + \ln\varepsilon - \lambda = 2k\pi i \tag{5.33}$$

for some $k \in \mathbb{Z}$. We call (5.33) the *characteristic equation*.

By writing $\lambda = \mu + i\nu$ we can rewrite (5.33) as

$$\frac{\alpha_{12}\alpha_{21}(\mu - \alpha_{11})}{(\mu - \alpha_{11})^2 + \nu^2} + \alpha_{22} + \ln\varepsilon - \mu = 0, \tag{5.34a}$$

$$\nu\Big[\frac{\alpha_{12}\alpha_{21}}{(\mu - \alpha_{11})^2 + \nu^2} + 1\Big] = 2k\pi \quad \text{with} \quad k \in \mathbb{Z}. \tag{5.34b}$$

Condition (5.34b) implies that when $k \neq 0$ the eigenvalues are truly complex. It is also clear that if (μ, ν) is a solution of (5.34a) and (5.34b), then so is $(\mu, -\nu)$; solutions of (5.34a) and (5.34b) corresponding to a pair $-k, k$ for $k \neq 0$ form conjugated pairs. If $k = 0$ then (5.34b) implies that the eigenvalues are real. We thus find that $k = 0$ yields (the only) two real eigenvalues of A and that these are given by

$$2\lambda_0^{\pm} = \alpha_{11} + \alpha_{22} + \ln\varepsilon \pm \sqrt{(\alpha_{11} - \alpha_{22} - \ln\varepsilon)^2 + 4\alpha_{12}\alpha_{21}}. \tag{5.35}$$

Conditions (5.34a) and (5.34b) yield a countable set of eigenvalues of A (here we can either use the fact that the roots of an analytic function are isolated points or Bezout's theorem about the number of intersections of algebraic curves). Since we will in the end only be interested in the spectral bound of A we refrain from giving explicit expressions for the complex eigenvalues and only show that no eigenvalue of A has a real part larger than λ_0^+.

So let us assume that the opposite holds, i.e. that there exists an eigenvalue $\lambda_k = \mu_k + \nu_k$ such that $\mu_k > \lambda_0^+$.

We first observe that

$$\lambda_0^+ > \alpha_{11}$$

and also that $\lambda_0^+ > \alpha_{22} + \ln\varepsilon$ (this latter will be useful later on).

Consider now (5.34a) for λ_0^+ and μ_k. Since $\mu_k > \lambda_0^+ > \alpha_{11}$, we can do the following estimate in (5.34a),

$$\begin{aligned}\frac{\alpha_{12}\alpha_{21}}{\lambda_0^+ - \alpha_{11}} + \alpha_{22} + \ln\varepsilon - \lambda_0^+ = 0 \\ &= \frac{\alpha_{12}\alpha_{21}(\mu_k - \alpha_{11})}{(\mu_k - \alpha_{11})^2 + \nu_k^2} + \alpha_{22} + \ln\varepsilon - \mu_k \\ &< \frac{\alpha_{12}\alpha_{21}}{\mu_k - \alpha_{11}} + \alpha_{22} + \ln\varepsilon - \mu_k\end{aligned}$$

and hence

$$\mu_k - \lambda_0^+ < \frac{\alpha_{12}\alpha_{21}(\lambda_0^+ - \mu_k)}{(\mu_k - \alpha_{11})(\lambda_0^+ - \alpha_{11})},$$

which brings us at a contradiction. Hence, no eigenvalue of A has a real part larger than λ_0^+.

The continuous spectrum of $(A, \mathcal{D}(A))$

We have seen that, if $\alpha_{12} \neq 0$ and $\alpha_{21} \neq 0$, then α_{11} cannot lie in the point spectrum of A. We will now show that α_{11} in fact belongs to the continuous spectrum, $\sigma_c(A)$. To this end, let $(f, g) \in \mathcal{D}(A)$ and $(F, G) \in \mathcal{X}$ such that

$$(A - \alpha_{11}I)\begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} F \\ G \end{pmatrix},$$

which we rewrite as

$$\alpha_{12} g = F \tag{5.36a}$$

$$\alpha_{21} f + (\alpha_{22} - \alpha_{11} + \ln \varepsilon) g - g' = G \tag{5.36b}$$

Note first that, since g needs to be absolutely continuous, the range of $A - \alpha_{11} I$ cannot be all of $\mathcal{X}$. Since $g(0) = g(1)$ we also require that $F(0) = F(1)$. The domain of the resolvent $R_{\alpha_{11}}(A) = (A - \alpha_{11} I)^{-1}$, i.e. the set $\{(F, G) \in \mathcal{X} | F \in \mathcal{W}^{1,1}[0,1], F(0) = F(1)\}$, is dense in $\mathcal{X}$ and the resolvent can, by using (5.36a) and (5.36b), explicitly be written as

$$(A - \alpha_{11} I)^{-1} \begin{pmatrix} F \\ G \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix},$$

with

$$g = \frac{1}{\alpha_{12}} F, \tag{5.37a}$$

$$f = \frac{1}{\alpha_{21}} \Big[G - (\alpha_{22} - \alpha_{11} + \ln \varepsilon) \frac{1}{\alpha_{12}} F + \frac{1}{\alpha_{12}} F' \Big]. \tag{5.37b}$$

Clearly, the resolvent is not bounded and so α_{11} indeed belongs to the continuous spectrum of A, $\sigma_c(A)$.

The resolvent set of $(A, \mathcal{D}(A))$

We will now show that if λ is not an eigenvalue of A and not equal to α_{11}, then λ belongs to the resolvent set, $\rho(A)$. So let us take such a λ and let for $(f, g) \in \mathcal{D}(A), (F, G) \in \mathcal{X}$

$$(A - \lambda I) \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} F \\ G \end{pmatrix}.$$

We write this in detail as

$$(\alpha_{11} - \lambda) f + \alpha_{12} g = F, \tag{5.38a}$$

$$\alpha_{21} f + (\alpha_{22} + \ln \varepsilon - \lambda) g - g' = G. \tag{5.38b}$$

Since $\lambda \neq \alpha_{11}$ we can use the first equation to express f in terms of g as

$$f = \frac{1}{\alpha_{11} - \lambda} \Big[F - \alpha_{12} g \Big] \tag{5.39}$$

and next obtain from the second that

$$\Big[\alpha_{22} + \ln \varepsilon - \lambda - \frac{\alpha_{12} \alpha_{21}}{\alpha_{11} - \lambda} - \frac{d}{dx} \Big] g = G - \frac{\alpha_{21}}{\alpha_{11} - \lambda} F. \tag{5.40}$$

Since λ is not an eigenvalue of A, the operator defined by the left hand side of (5.40) has a trivial nullspace. With

$$\chi(\lambda) = \alpha_{22} + \ln\varepsilon - \lambda - \frac{\alpha_{12}\alpha_{21}}{\alpha_{11} - \lambda}$$

we can explicitly write

$$g(x) = ce^{\chi(\lambda)x} - e^{\chi(\lambda)x}\int_0^x e^{-\chi(\lambda)y}\Big(G(y) - \frac{\alpha_{21}}{\alpha_{11} - \lambda}F(y)\Big)dy \tag{5.41}$$

and determine c so that $g(0) = g(1)$,

$$c = \frac{e^{\chi(\lambda)}\int_0^1 e^{-\chi(\lambda)y}(G(y) - \frac{\alpha_{21}}{\alpha_{11}-\lambda}F(y))dy}{e^{\chi(\lambda)} - 1}.$$

It is then clear that (5.39) along with (5.41) defines, by way of $(F, G) \mapsto (f, g)$, a bounded operator $R_\lambda(A)$ on $\mathcal{X}$.

Hence, since $\lambda_0^+ > \alpha_{11}$, we obtain in this first case that

$$2s(A) = 2\lambda_0^+ = \alpha_{11} + \alpha_{22} + \ln\varepsilon + \sqrt{(\alpha_{11} - \alpha_{22} - \ln\varepsilon)^2 + 4\alpha_{12}\alpha_{21}}$$

and that $s(A)$ is an eigenvalue. The spectral bound is in a fact an algebraically simple eigenvalue and this fact will play an important role later on when we discuss convergence to a stable bacterial distribution. Let us see that this is indeed the case.

The eigenvalue λ is an *algebraically simple eigenvalue* if the dimension of the generalized eigenspace of A associated with λ, defined as

$$\cup_{k=1}^{\infty}\mathcal{N}(A - \lambda I)^k,$$

equals one. Here, $\mathcal{N}(.)$ denotes the nullspace.

Let us show that λ is an algebraically simple eigenvalue precisely when it is a simple root of the characteristic equation.

Lemma 5.1. *λ is an algebraically simple eigenvalue of A if and only if $\chi'(\lambda) \neq 0$, where χ is defined in* (5.33).

Proof. We first note that $(A - \lambda I)^2$ is given by

$$\begin{pmatrix} (\alpha_{11} - \lambda)^2 + \alpha_{12}\alpha_{21} & \alpha_{12}(\alpha_{11} - \lambda + \alpha_{22} + \ln\varepsilon - \lambda - \frac{d}{dx}) \\ \alpha_{21}(\alpha_{11} - \lambda + \alpha_{22} + \ln\varepsilon - \lambda - \frac{d}{dx}) & \alpha_{12}\alpha_{21} + (\alpha_{22} + \ln\varepsilon - \lambda - \frac{d}{dx})^2 \end{pmatrix}$$

and that

$$\chi'(\lambda) \neq 0 \iff (\alpha_{11} - \lambda)^2 + \alpha_{12}\alpha_{21} \neq 0.$$

Let us first assume that $\chi'(\lambda) = 0$ and let $(f, g)^T \in \mathcal{D}(A^2)$ be such that $(A - \lambda I)^2(f, g)^T = 0$. We can rewrite this as

$$0 = \Big((\alpha_{11} - \lambda)^2 + \alpha_{12}\alpha_{21}\Big) f + \alpha_{12}\Big(\alpha_{11} - \lambda + \alpha_{22} + \ln\varepsilon - \lambda - \tfrac{d}{dx}\Big) g \tag{5.42a}$$

$$0 = \alpha_{21}\Big(\alpha_{11} - \lambda + \alpha_{22} + \ln\varepsilon - \lambda - \tfrac{d}{dx}\Big) f + \Big(\alpha_{12}\alpha_{21} + (\alpha_{22} + \ln\varepsilon - \lambda - \tfrac{d}{dx})^2\Big) g. \tag{5.42b}$$

If $(\alpha_{11} - \lambda)^2 + \alpha_{12}\alpha_{21} = 0$ and $\alpha_{12} \neq 0$, then the first equation coincides with equation (5.32). It is then straightforward to see that, for $(f, g)^T \in \mathcal{N}(A - \lambda I)$, also $(0, g)^T \in \mathcal{N}(A - \lambda I)^2$, while no $(f, g)^T$ with $f = 0$ lies in $\mathcal{N}(A - \lambda I)$. The dimension of $\mathcal{N}(A - \lambda I)^2$ is thus strictly larger than the dimension of $\mathcal{N}(A - \lambda I)$ and the eigenvalue λ is not simple.

Suppose now that $\chi'(\lambda) \neq 0$. We know from (5.31a), (5.31b) and (5.33) that

$$\begin{pmatrix} F \\ G \end{pmatrix} \in \mathcal{N}(A - \lambda I) \iff \begin{pmatrix} F \\ G \end{pmatrix} = e^{2k\pi i x} \begin{pmatrix} \frac{\alpha_{12}}{\lambda - \alpha_{11}} \\ 1 \end{pmatrix}$$

for some $k \in \mathbb{Z}$. Hence,

$$(A - \lambda I) \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} F \\ G \end{pmatrix}$$

for some $(f, g)^T \in \mathcal{D}(A)$ requires that

$$2k\pi i g - g' = e^{2k\pi i x}\Big(1 + \frac{\alpha_{12}\alpha_{21}}{(\alpha_{11} - \lambda)^2}\Big)$$

and so

$$g(x) = g(0)e^{2k\pi i x} - \Big(1 + \frac{\alpha_{12}\alpha_{21}}{(\alpha_{11} - \lambda)^2}\Big) x e^{2k\pi i x},$$

which does not satisfy the condition $g(0) = g(1)$ if $\chi'(\lambda) \neq 0$. □

Case II. When at least one of α_{12} or α_{21} is equal to zero, then the operator A is either upper or lower diagonal or, when both α_{12} and α_{21} are zero, diagonal. In any of these cases, α_{11} belongs to the point spectrum (indeed, $A - \alpha_{11} I$ has a nontrivial nullspace in all three cases). Moreover, the second set of eigenvalues is obtained by taking

$$-g' + (\alpha_{22} + \ln\varepsilon - \lambda) g = 0$$

which, taking into account the condition $g(0) = g(1)$, requires that

$$\lambda = \alpha_{22} + \ln\varepsilon + 2\pi k i$$

for some $k \in \mathbb{Z}$.

In the same way as we have done in Case I we can show that if $\lambda \neq \alpha_{11}$ and $\lambda \notin \{\alpha_{22} + \ln\varepsilon + 2\pi k i \mid k \in \mathbb{Z}\}$, then λ lies in the resolvent set.

Since there is now no relation between α_{11} and the rest of the spectrum, we obtain

$$s(A) = \max\{\alpha_{11}, \alpha_{22} + \ln \varepsilon\}.$$

It is straightforward to see that $s(A)$ is also in this case an algebraically simple eigenvalue.

5.4.3 The Browder essential spectrum

In order to show that the growth bound of the semigroup equals the spectral bound of the generator we shall now show that the *essential growth bound* of A is strictly smaller than $s(A)$. According to (5.30), it then follows that $\omega_0(A) = s(A)$.

The fact that the spectral bound of A equals the growth bound of the semigroup in fact holds for every positive semigroup on an $\mathcal{L}^1$ space and follows directly from Derndingers's result in [27]. The information about the essential growth bound will allow us to make some further statements about the convergence to a stable distribution (cf. subsection 5.4.4).

So let us first introduce the notion of the Browder essential spectrum.

Definition 5.6. *The complex number λ belongs to the* **Browder essential spectrum** *of the bounded operator L, σ_{ess}, if at least one of the following holds:*

(i) *λ is a limit point of $\sigma(L)$,*

(ii) *$\mathcal{R}(\lambda I - L)$ is not closed,*

(iii) *$\cup_{k\geq 0}\mathcal{N}(\lambda I - L)^k$ is infinite dimensional.*

It can be shown that if $\lambda \in \sigma(L) \setminus \sigma_{\text{ess}}(L)$, then $\lambda \in \sigma_p(L)$. These are the so called normal eigenvalues.

In order to define the essential growth bound, we introduce the notion of a **measure of non-compactness**.

Definition 5.7. *For a bounded subset $\Omega \in X$, we define the measure of noncompactness $\alpha(\Omega)$ by*

$$\alpha(\Omega) = \inf\{d > 0 : \textit{there exist finitely many subsets } \Omega_1, ..., \Omega_n \textit{ of } X \textit{ such that } diam(\Omega_k) \leq d \textit{ and } \Omega \subseteq \cup_{k=1}^{n}\Omega_k\}.$$

where the diameter of the set Ω, $diam(\Omega)$, is defined by

$$diam(\Omega) = \sup\{\|x - y\| : x, y \in \Omega\}.$$

For a bounded operator L we define the measure of non-compactness $|L|_\alpha$ by

$$|L|_\alpha = \inf\{k > 0 : \alpha(L\Omega) \leq k\alpha(\Omega) \textit{ for every bounded set } \Omega \textit{ of } X\}.$$

Essential growth bound of the generator A of a strongly continuous semigroup $\{T(t)\}_{t\geq 0}$, $\omega_{\text{ess}}(A)$, is given by

$$\omega_{\text{ess}}(A) = \inf_{t>0} \frac{1}{t} \log |T(t)|_\alpha = \lim_{t\to\infty} \frac{1}{t} \log |T(t)|_\alpha. \tag{5.43}$$

One can see that $|T(t)|_\alpha \leq \|T(t)\|$. This implies that the essential growth bound never exceeds the growth bound, i.e.

$$\omega_{\text{ess}}(A) \leq \omega_0(A).$$

To see that $\omega_{\text{ess}}(A) < \omega_0(A)$, we recall the generation expansion

$$T(t) = T_0 + \sum_{k=1}^{\infty} T_k(t).$$

The first generation, T_1, is according to (5.28) given by

$$T_1(t) = \int_0^t T_0(t-s)BT_0(s)ds,$$

which, using (5.26) can explicitly be written as

$$T_1(t)\begin{pmatrix} f \\ g \end{pmatrix}(x) = \begin{pmatrix} \int_0^t \alpha_{12} e^{\alpha_{11}(t-s)} e^{(\alpha_{22}+\ln\varepsilon)s} \tilde{g}(e^{2\pi i(x-s)})ds \\ \int_0^t \alpha_{21} e^{(\alpha_{22}+\ln\varepsilon)(t-s)} e^{\alpha_{11}s} \tilde{f}(e^{2\pi i(x-(t-s))}ds \end{pmatrix} \tag{5.44}$$

According to [39], $T_1(t)$ is for every $t \geq 0$ a compact operator. Using the fact that compact operators form an ideal in the space of bounded linear operators, we can show (i) by induction on k that also every subsequent generation operator T_k, obtained by (5.28) is compact and (ii) every finite sum $\sum_{k=1}^n T_k$ is a compact operator. But since the ideal of compact operators is a closed one, also the operator $\sum_{k=1}^\infty T_k(t)$ is compact.

Since we furthermore know that for every bounded operator L and a compact operator K

$$|L + K|_\alpha = |L|_\alpha,$$

we have

$$\omega_{\text{ess}}(A) = \omega_{\text{ess}}(A_0).$$

Hence

$$\omega_{\text{ess}}(A) = \omega_{\text{ess}}(A_0) \leq \omega_0(A_0) = \max\{\alpha_{11}, \alpha_{22} + \ln\varepsilon\} \leq s(A)$$

and the last inequality is strict in Case I when α_{12} and α_{21} are strictly positive. According to (5.30),

$$\omega_0(A) = s(A).$$

5.4.4 Asynchronous exponential growth

The aim of this section is to establish whether the bacterial population in the intestine experiences the so called asynchronous exponential growth.

A population is said to have *asynchronous exponential growth*, if there exist an operator P with a one dimensional range and $\lambda > 0$ such that for every $\phi \in \mathcal{X}$

$$\lim_{t\to\infty} e^{-\lambda t}T(t)\phi = P\phi.$$

To see whether such convergence of the distribution of bacteria in the intestine takes place we again treat cases I and II separately. We begin with

Case I. We have $\omega_{\text{ess}}(A) < \omega_0(A) = s(A)$ and we know that $s(A)$ is a simple eigenvalue. So let $\delta > 0$ be such that

(i) $\omega_{\text{ess}}(A) \le s(A) - \delta$ and

(ii) Re $\lambda \le s(A) - \delta$ for every $\lambda \in \sigma(A)$, $\lambda \ne s(A)$.

We can decompose the space $\mathcal{X}$ as follows (see [15], Theorem A.3.1),

$$\mathcal{X} = \mathcal{N}(s(A)I - A) \oplus \mathcal{R}(s(A)I - A),$$

where $\mathcal{N}(.)$ and $\mathcal{R}(.)$ denote, respectively, the null space and the range. Both subspaces are invariant under the action of the semigroup $\{T(t)\}_{t\ge 0}$. Let us denote by P the projection on $\mathcal{N}(s(A)I - A)$. Note that the operator P commutes with the operators $T(t)$.

The action of the semigroup on the subspace $\mathcal{N}(s(A)I - A)$ is very simple, namely,

$$T(t)P = PT(t) = e^{s(A)t}P.$$

The restriction of the semigroup to the invariant subspace $\mathcal{R}(s(A)I - A)$, on the other hand, is given by $\{T(t)(I - P)\}_{t\ge 0}$ and is generated by the restriction of A to $\mathcal{R}(s(A)I - A)$. Let us denote this restriction by $\tilde{A}$. The spectrum of $\tilde{A}$ is the set $\sigma(A) \setminus \{s(A)\}$. Hence

$$s(\tilde{A}) \le s(A) - \delta.$$

Clearly

$$\omega_{\text{ess}}(\tilde{A}) \le \omega_{\text{ess}}(A) \le s(A) - \delta$$

and hence

$$\omega_0(\tilde{A}) = \max\{s(\tilde{A}), \omega_{\text{ess}}(\tilde{A})\} \le s(A) - \delta.$$

Thus, there exists a constant $M(\delta)$ such that for every $\phi \in \mathcal{X}$

$$\|T(t)(I - P)\phi\| \le M(\delta)e^{s(A)-\delta}\|\phi\|.$$

By writing $\phi = P\phi + (I - P)\phi$ we hence obtain

$$\|e^{-s(A)t}T(t)\phi - P\phi\| \leq M(\delta)e^{-\delta t}\|\phi\|.$$

By Theorem 8.17 of [15] there in fact exists an element $\phi_0 \in \mathcal{X}^+$ and a strictly positive functional $\phi_0^* \in (\mathcal{X}^*)^+$ such that $A\phi_0 = s(A)\phi_0$, $A^*\phi_0^* = s(A)\phi_0^*$, $< \phi_0, \phi_0^* >= 1$ and the projection is given by

$$P\phi =< \phi, \phi_0^* > \phi_0.$$

For an arbitrary initial distribution ϕ, therefore, the solution behaves as

$$e^{s(A)t}P\phi = e^{s(A)t} < \phi, \phi_0^* > \phi_0$$

when t goes to infinity. The dominant eigenvalue, $s(A)$, is accordingly also called the *Malthusian parameter*.

Case II. We now have $s(A) = \max\{\alpha_{11}, \alpha_{22} + \ln\varepsilon\}$. If $\alpha_{22} + \ln\varepsilon < \alpha_{11}$ we can, in the same way as before, decompose the space $\mathcal{X}$ to establish convergence to a steady distribution. If, on the other hand $\alpha_{11} \leq \alpha_{22} + \ln\varepsilon$, the eigenspace corresponding to $s(A)$ is no longer one dimensional (remember that the eigenvalues of A include the set $\{\alpha_{22} + \ln\varepsilon + 2k\pi i \mid k \in \mathbb{Z}\}$). The bacterial population will in the long run still grow (or decay) like $e^{s(A)t}$, but we will not observe convergence of the distribution of bacteria in the intestine. In such a case we observe the so called *merry-go-round* behaviour.

5.4.5 The effects of (self)reinfection on the intrinsic growth rate of ETEC

Now that the conditions for persistence have been established, it is time to discuss the effects of reinfection on the intrinsic growth rate.

We have shown in Section 5.3 that the parameter α_{11} determines whether the bacteria in the intestine persist after a single introduction. When a fraction ε of the excreted bacteria is reintroduced into the intestine, the intrinsic growth rate of bacteria is given by the spectral bound of the operator A given in (5.22). We have seen that $s(A)$ is, in both Case I and II, an eigenvalue of A. In particular, in

Case I. The spectral bound is given by

$$2s(A) = \alpha_{11} + \alpha_{22} + \ln\varepsilon + \sqrt{(\alpha_{11} - \alpha_{22} - \ln\varepsilon)^2 + 4\alpha_{12}\alpha_{21}} \tag{5.45}$$

and since in this case $s(A) > \alpha_{11}$, reintroduction always enhances the growth.

It is simple to check that $s(A)$ is an increasing function of ε. One can also see that

$$\lim_{\varepsilon \to 0} s(A)(\varepsilon) = \alpha_{11}, \tag{5.46}$$

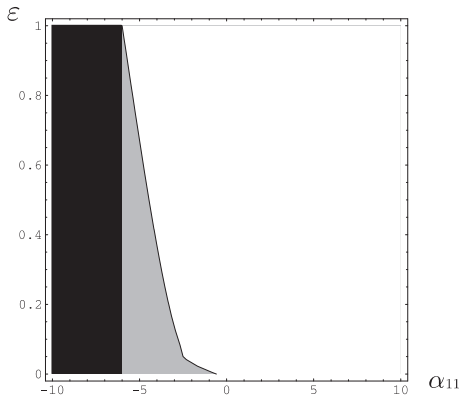

Figure 5.4: The figure represents the sign of the Malthusian parameter $s(A)$ as a function of ε (the proportion of reintroduced bacteria) and the parameter α_{11} that determines persistence in the 'single infection' case. We furthermore take: $\alpha_{12} = 3, \alpha_{21} = 4, \alpha_{22} = -2$. Bacteria persist in the white region and go extinct in grey and black area.

as is to be expected and that

$$\lim_{\varepsilon \to 1} s(A)(\varepsilon) = s(A)(1) = \tfrac{1}{2}(\alpha_{11} + \alpha_{22} + \sqrt{(\alpha_{11} - \alpha_{22})^2 + 4\alpha_{12}\alpha_{21}}) \qquad (5.47)$$

We thus have the following three possibilities.

CASE IA. $\alpha_{11} \geq 0$.

In this case bacteria persist already following a single introduction and reintroduction only enhances their growth.

CASE IB. $\alpha_{11} < 0, s(A)(1) > 0$.

Bacteria are washed out following a single introduction of ETEC. When a sufficient proportion of excreted bacteria is reintroduced into the intestine, the bacteria can persist. This 'sufficient' proportion is obtained by calculating ε_0 such that $s(A)(\varepsilon_0) = 0$. Every $\varepsilon \in (\varepsilon_0, 1]$ then guarantees growth.

CASE IC. $\alpha_{11} < 0, s(A)(1) < 0$.

In this case bacteria cannot persist even when they are all reintroduced back into the intestine.

Figure 5.4 shows the three regions by plotting $s(A)$ as a function of ε and α_{11} for following parameter values: $\alpha_{12} = 3, \alpha_{21} = 4$ and $\alpha_{22} = -2$ and the domains of ε and α_{11} taken to be, respectively, intervals $[0, 1]$ and $[-10, 10]$. The curve separating the white and the grey area represents the curve $s(A)(\varepsilon, \alpha_{11}) = 0$, i.e. the curve that separates the domains of extinction and asynchronous exponential growth. The white region corresponds to the area in which bacteria persist. Since

$s(A)(\varepsilon) > \alpha_{11}$, this region always includes the area in which α_{11} is positive. The black region depicts the region in which bacteria don't stand a chance even when they are all reintroduced into the intestine. Also in the grey area the bacteria go extinct. However, assuming the same value of α_{11}, bacteria can persist if ε is increased so that $s(A)(\alpha_{11}, \varepsilon)$ crosses the curve $s(A)(\alpha_{11}, \varepsilon) = 0$ and enters the white region.

Case II. The spectral bound is now

$$s(A) = \max\{\alpha_{11}, \alpha_{22} + \ln \varepsilon\}.$$

If $\alpha_{11} > \alpha_{22} + \ln \varepsilon$, then bacteria persist if and only if they persist after a single introduction. If $\alpha_{11} \leq \alpha_{22} + \ln \varepsilon$, we again have three possibilities.

CASE IIA. $0 < \alpha_{11} \leq \alpha_{22} + \ln \varepsilon$.

Bacteria persist already following a single introduction and reinfection only enhances the growth.

CASE IIB. $\alpha_{11} < 0 < \alpha_{22} + \ln \varepsilon$

Bacteria are washed out after a single introduction, but a proportion ε of reintroduced bacteria is sufficient to guarantee persistence.

CASE IIC. $\alpha_{11} \leq \alpha_{22} + \ln \varepsilon < 0$.

Bacteria cannot persist in the long run, even if a fraction ε is reintroduced in the intestine. In the limiting case $\varepsilon \to 1$ (i.e. when $\alpha_{11} \leq \alpha_{22} < 0$), bacteria don't stand a chance even when they are all reintroduced.

5.5 Conclusions and outlook

The aim of this chapter was to investigate the dynamics of enterotoxigenic *Escherichia coli* (ETEC) in the gastro-intestinal tract of a single piglet. ETEC can lead to post-weaning diarrhoea, a disease that often results in severe deterioration or even death in newly weaned piglets.

We formulated a simple model describing bacterial growth and movement in the intestine of a single infected piglet. In other words, we focused on the within-host dynamics.

The intestine is represented by a cylindrical tube of constant cross-sectional area and constant circumference. According to our model, bacteria can at any point in time be in one of the following two states: either attached to the wall of the intestine or moving downstream with (constant) speed. Bacteria change their state (i.e., either attach to the wall or fall off of the intestine wall) at a constant rate. We furthermore assume that both the attached, as well as the free, bacteria can reproduce. However, while the offspring of free bacteria is necessarily free

(at least initially, that is), offspring of the attached bacteria can be either attached or free.

We have in particular focused on the condition(s) that guarantee growth or extinction of ETEC. This condition turns out to involve a single number, called the *intrinsic growth rate* or the *Malthusian parameter*. In the case of persistence, we have also investigated whether the bacterial population experiences the so called *asynchronous exponential growth*, which means that the distribution of bacteria in the intestine, given by

$$\begin{pmatrix} u(t,x) \\ v(t,x) \end{pmatrix}$$

behaves asymptotically (i.e. for $t \to \infty$) as

$$\begin{pmatrix} u(t,x) \\ v(t,x) \end{pmatrix} \sim \; Ce^{rt} \begin{pmatrix} \tilde{u}(x) \\ \tilde{v}(x) \end{pmatrix}$$

where r is the Malthusian parameter, $(\tilde{u}(x), \tilde{v}(x))$ is the so called stable distribution and C is a constant.

To begin with, we investigated a 'single infection' case, i.e. the case when some bacteria are introduced into the intestine at time $t = 0$ and there is no inflow of bacteria into the intestine after $t = 0$. Since piglets are likely to come into contact with the shed bacteria contained in the faeces, it seems relevant to also investigate the conditions for growth/extinction for the case when a proportion ε of the excreted bacteria is reintroduced into the intestine. Of course, if bacteria can persist following a single introduction, their persistence is guaranteed also when the piglet can be reinfected.

The question of how the intrinsic growth rate depends on ε is particularly relevant for control of ETEC. Indeed, since the proportion of reintroduced bacteria depends on the amount of shed bacteria, maintaining a high level of hygiene on pig farms can reduce ε and may thus prevent the bacteria to settle in an individual piglet.

In the single infection case it is clear that, since no bacteria are introduced after the initial dose at $t = 0$, bacteria must, in order to persist, establish themselves on the wall of the intestine. And indeed, we found, by constructing the explicit solution of the problem, that the bacteria persist precisely when the growth rate on the wall exceeds the rate at which bacteria detach from the wall. It is also clear from the explicit solution in (5.19a), (5.19b), that there is (in the case of growth) no convergence to a stable bacterial distribution in the intestine.

To find the conditions for persistence in the case a piglet can be reinfected, we approached the problem in a different manner, namely, we reformulated the problem as an abstract Cauchy problem and employed the theory of positive semigroups. We found that, even though the spectrum of the (generator of the)

semigroup does not consist solely of point spectrum, finding the condition for persistence boils down to finding the dominant eigenvalue of the generator. We furthermore showed that when bacteria in the two states (attached and free) are, in a sense, well mixed, then the bacterial population in the intestine experiences asynchronous exponential growth.

The two cases, or, rather, the approaches with which we studied them, are related in the following way. First of all, having obtained the fundamental solution (that is, the solution in the 'single infection' case) we can construct the solution for any boundary condition $v(t,0) = v_0(t)$. Indeed, since we are dealing with a linear system, we can think of $v_0(t)$ as a superposition of impulses given by Dirac measures,

$$v(t) = \int_0^t \delta_0(t-s)v(s)ds.$$

On the other hand, we can similarly formulate the 'single infection' problem as an abstract Cauchy problem and study it by employing semigroup theory. Note however, that the operators should then be defined on the Banach space of measures $\mathcal{M}([0,1])$ (which is known to be a Banach lattice when equipped with the total variation norm and the following order relation,

$$m_1 \leq m_2 \iff m_1(\omega) \leq m_2(\omega)$$

for every measurable set $\omega \subseteq [0,1]$).

There are, to the best of our knowledge, no models in the literature that describe the dynamics of enterotoxigenic *E-coli* in piglets. We can, however, draw some parallels with the work of others on related subjects. In particular, we wish to mention here the papers of Gyllenberg and Webb [58], Pachepsky et al. [90] and Ballyk and Smith [5].

In [58], the authors considered a model for cell growth incorporating the fact that cells can either be 'normal' (meaning that they progress in size, as well as in age), or quiescent (meaning they get old, but do not increase in size). Cells can change the state they are in at any point in time and so, in a sense, their 'normal' resembles our 'free' (free bacteria can move as well as reproduce), and their 'quiescent' bears likeness to our 'attached' (attached bacteria do not move, but can reproduce). The model of Gyllenberg and Webb is more elaborated than ours, taking into account age and size dependent rates, but their work also demonstrates that populations can avoid extinction by transferring cells from normal (or, in our case, free) to the quiescent state (attached).

The model of Pachepsky et al. [90] resembles the model of this chapter in many respects. The authors investigate the so called 'drift paradox', namely the question of how some populations, say aquatic organisms in rivers, can persist, even though they are subjected to strong, mostly unidirectional, flow. As we in

the present study, the authors in [90] consider attached and floating organisms. The model in [90] furthermore includes the random movement of free organisms (that is modeled by an additional $D\frac{\partial^2}{\partial x^2}$ in the free compartment), however, it does not allow for reproduction of free organisms. Two conditions are found which guarantee persistence. The first coincides with the one found in the present study: a population can persist if the rate of reproduction on the wall (or, benthos) exceeds the rate at which the organisms enter the stream. However, when the model also includes random movement, organisms can persist even when the reproduction on the wall is not sufficient to guarantee persistence in the absence of such movement. The second condition includes other parameters and in fact involves a minimum length of the domain that is required for persistence.

In [58] as well as in [90] the question of reintroduction does not have much, if at all any, biological significance. The question of reintroduction is, however, relevant and in fact investigated in the paper of Ballyk and Smith [5], along with the question of persistence in a single introduction case. Along with the equations for attached and floating bacteria, the model in [5] includes a third equation for the dynamics in the external pool of bacteria, from which bacteria can be introduced into the gut. The conditions for growth/extinction for the 'no inflow' case are similar to the ones presented in [90]. The authors in [5] furthermore show that persistence is always guaranteed when the bacteria are reintroduced at a constant rate.

In the present study, we have investigated the dynamics of ETEC in a single piglet. With 'reintroduction' we thus have in mind that selfreinfection occurs when a piglet comes into contact with its own faeces. Instead of assuming a constant source of reintroduction, we have assumed that the fraction ε of shed bacteria reenter the intestine. We found that (i) reintroduction enhances the growth of bacteria in a piglet and (ii) the Malthusian parameter depends on ε and accordingly reinfection of a proportion ε of the shed bacteria does not always guarantee persistence. The external pool of bacteria modeled with an additional variable seems, however, a good idea to describe the source of (re)infection at the (piglet) population level.

Our study can be seen as a starting point for investigating more complicated models of within-host dynamics as well as dynamics at the piglet population level. Much remains to be done.

At the individual level, a more realistic within-host model should incorporate many other aspects. For instance, the movement in the intestine almost certainly does not only include unidirectional downstream movement, but also random movement, which we can model as diffusion. A more realistic model should furthermore incorporate the effect of crowding of bacteria on the wall of the intestine. According to our model, offspring of the attached bacteria may not find its place on the wall, which can be seen as a consequence of crowding, but there

is no saturation effect included - all free bacteria can attach to the wall at the same constant rate b. Moreover, one could incorporate a position dependent speed $g(\xi)$, position dependent rates a, b, etc. We plan to study more elaborated models in the future.

Another aspect not investigated in the present study is modeling the population dynamics of ETEC. To study the spread of ETEC in a population of piglets one needs to determine some individual characteristics. Within-host modeling, as performed in this chapter, can provide such information. We may, for instance, assume that infectiousness of an individual piglet in time, say $\mathcal{A}(t)$, is proportional to the amount of shed bacteria, i.e.,

$$\mathcal{A}(t) = cv(t, L).$$

What remains is to describe contact structure and transmission of ETEC. We plan to expand our study in the near future.

Bibliography

[1] S. Alizon, *Parasite virulence evolution: insights from embedded models*, PhD thesis, University of Paris, France, 2006.

[2] R. M. Anderson and R. M. May, *Coevolution of hosts and parasites*, Parasitology, 85 (1982), pp. 411–426.

[3] ——, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, Oxford, UK, 1991.

[4] R. Antia, B. R. Levin, and R. M. May, *Within-host population dynamics and the evolution and maintenance of macroparasite virulence*, Am. Nat., 144 (1994), pp. 457–472.

[5] M. Ballyk and H. Smith, *A model of microbial growth in a plug flow reactor with wall attachment*, Math. Biosci., 158 (1999), pp. 95–126.

[6] D. Bergmans, M. Bonten, C. Gaillard, J. Palling, S. van der Geest, F. van Tiel, A. Beysens, P. de Leeuw, and E. Stobberingh, *Prevention of ventilator-associated pneumonia by oral decontamination: a prospective, randomized, double-blind, placebo-controlled study*, Am. J. Respir. Crit. Care Med., 163 (3) (2001), pp. 382–388.

[7] A. Berman and R. J. Plemmons, *Nonnegative matrices in the mathematical sciences*, vol. 9 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994. Revised reprint of the 1979 original.

[8] M. Bonten, D. Austin, and M. Lipsitch, *Understanding the spread of antibiotic resistant pathogens in hospitals: mathematical models as tools for control*, Clin. Infect. Dis., 33 (2001), pp. 1739–1746.

[9] M. Bonten, D. Bergmans, A. Ambergen, P. de Leeuw, S. van der Geest, E. Stobberingh, and C. Gaillard, *Risk factors for pneumonia, and colonization of respiratory tract and stomach in mechanically ventilated icu patients*, Am. J. Respir. Crit. Care Med., 154 (1996), pp. 1339–1346.

[10] M. Bonten, S. Slaughter, A. Ambergen, M. Hayden, J. van Voorhis, C. Nathan, and R. Weinstein, *The role of "colonization pressure" in the spread of vancomycin-resistant enterococci: an important infection control variable*, Arch. Intern. Med., 158 (10) (1998), pp. 1127–1132.

[11] C. Brun-Buisson, P. Legrand, A. Rauss, C. Richard, F. Montravers, M. Besbes, J. Meakins, C. Soussy, and F. Lemaire, *Intestinal decontamination for control of nosocomial multiresistant gram-negative bacilli. Study of an outbreak in an intensive care unit.*, Ann. Intern. Med., 110 (11) (1989), pp. 873–881.

[12] M. G. Bulmer, *Periodical insects*, Am. Nat., 111 (1997), pp. 1099–1117.

[13] D. Claessen, *Dwarfs and giants. The dynamic interplay of size-dependent cannibalism and competition*, UvA-IBED, 2002.

[14] D. Claessen and A. de Roos, *Bistability in a size-structured population model of cannibalistic fish - a continuation study*, Theor. Pop. Biol., 64 (2003), pp. 49–65.

[15] P. Clément, H. J. A. M. Heijmans, S. Angenent, C. J. van Duijn, and B. de Pagter, *One-parameter semigroups*, vol. 5 of CWI Monographs, North-Holland Publishing Co., Amsterdam, 1987.

[16] M. G. Crandall and P. H. Rabinowitz, *Bifurcation from simple eigenvalues*, J. Functional Analysis, 8 (1971), pp. 321–340.

[17] ——, *The principle of exchange of stability*, in Dynamical systems (Proc. Internat. Sympos., Univ. Florida, Gainesville, Fla., 1976), Academic Press, New York, 1977, pp. 27–41.

[18] J. Cushing and J. Li, *Intra-specific competition and density dependent juvenile growth*, Bull. Math. Biol., 54 (1992), pp. 503–519.

[19] J. M. Cushing, *An introduction to structured population dynamics*, vol. 71 of CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998.

[20] N. Davydova, *Old and young. Can they coexist?*, PhD Thesis, University of Utrecht, The Netherlands, 2004.

[21] N. V. Davydova, O. Diekmann, and S. A. van Gils, *Year class coexistence or competitive exclusion for strict biennials?*, J. Math. Biol., 46 (2003), pp. 95–131.

[22] T. Day and S. R. Proulx, *A general theory for the evolutionary dynamics of virulence*, Am. Nat., 163 (2004), pp. 40–63.

[23] E. de Jonge, M. Schultz, L. Spanjaard, P. Bossuyt, M. Vroom, J. Dankert, and J. Keseciglou, *Effects of selective decontamination of digestive tract on mortality and acquisition of resistant bacteria in intensive care: a randomised controlled trial.*, Lancet, 362 (9389) (2003), pp. 1011–1016.

[24] M. de la Cal, E. Cerda, P. Garcia-Hierro, H. van Saene, D. Gomez-Santos, E. Negro, and J. Lorente, *Survival benefit in critically ill burned patients receiving selective decontamination of the digestive tract: a randomized, placebo-controlled, double-blind trial.*, Ann. Surg., 241 (3) (2005), pp. 424–430.

[25] P. De Leenheer and H. L. Smith, *Virus dynamics: a global analysis.*, SIAM J. Appl. Math., 63 (2003), pp. 1313–1327.

[26] A. de Roos, L. Persson, and H. Thieme, *Emergent allee effects in top predators feeding on structured prey populations*, Proc. R. Soc. Lond. B, 270 (2003), pp. 611–618.

[27] R. Derndinger, *Über das spektrum positiver generatoren*, Mathematische Zeitschrift, 172 (1980), pp. 281–293.

[28] U. Dieckmann and J. A. J. Metz, *Surprising evolutionary predictions from enhanced ecological realism.*, Theor. Popul. Biol., 69 (2006), pp. 263–281.

[29] U. Dieckmann, J. A. J. Metz, M. W. Sabelis, and K. Sigmund, *Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management*, Cambridge studies in Adaptive Dynamics, Cambridge University Press, 2002.

[30] O. Diekmann, *A beginner's guide to adaptive dynamics*, in Mathematical modelling of population dynamics, vol. 63 of Banach Center Publ., Polish Acad. Sci., Warsaw, 2004, pp. 47–86.

[31] O. Diekmann, K. Dietz, and J. Heesterbeek, *The basic reproduction ratio for sexually transmitted diseases, Part I - Theoretical considerations*, Math. Biosci., 107 (1991), pp. 325–339.

[32] O. Diekmann, M. Gyllenberg, H. Huang, M. Kirkilionis, J. A. J. Metz, and H. R. Thieme, *On the formulation and analysis of general deterministic structured population models. II. Nonlinear theory*, J. Math. Biol., 43 (2001), pp. 157–189.

[33] O. Diekmann, M. Gyllenberg, and J. A. J. Metz, *Steady state analysis of structured population models*, Theor. Pop. Biol., 63 (2003), pp. 309–338.

[34] O. Diekmann, M. Gyllenberg, J. A. J. Metz, and H. R. Thieme, *On the formulation and analysis of general deterministic structured population models. I. Linear theory*, J. Math. Biol., 36 (1998), pp. 349–388.

[35] O. Diekmann and J. A. P. Heesterbeek, *Mathematical epidemiology of infectious diseases*, Wiley Series in Mathematical and Computational Biology, John Wiley & Sons Ltd., Chichester, 2000. Model building, analysis and interpretation.

[36] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz, *On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations*, J. Math. Biol., 28 (1990), pp. 365–382.

[37] O. Diekmann, H. J. A. M. Heijmans, and H. R. Thieme, *On the stability of the cell size distribution*, J. Math. Biol., 19 (1984), pp. 227–248.

[38] J. J. Duistermaat and J. A. C. Kolk, *Multidimensional real analysis. I. Differentiation*, vol. 86 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 2004. Translated from the Dutch by J. P. van Braam Houckgeest.

[39] N. Dunford and J. T. Schwartz, *Linear Operators. I. General Theory*, With the assistance of W. G. Bade and R. G. Bartle. Pure and Applied Mathematics, Vol. 7, Interscience Publishers, Inc., New York, 1958.

[40] J. Dushoff, *Incorporating immunological ideas in epidemiological models*, J. Theor. Biol., 180 (1996), pp. 181–187.

[41] J. Dushoff, W. Huang, and C. Castillo-Chavez, *Backwards bifurcations and catastrophe in simple models of fatal diseases*, J. Math. Biol., 36 (1998), pp. 227–248.

[42] L. Edelstein-Keshet, *Mathematical models in biology*, The Random House, Birkhauser Mathematics Series, Random House Inc., New York, 1988.

[43] K.-J. Engel and R. Nagel, *One-parameter semigroups for linear evolution equations*, vol. 194 of Graduate Texts in Mathematics, Springer-Verlag, New York, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.

[44] L. Euler, *Recherches générales sur la mortalité et la multiplication du genre humain*, Mémoires de l'Académie Royale des Sciences et Belles Lettres, XVI (1760), pp. 144–164.

[45] P. W. Ewald, *Host-parasite relations, vectors, and the evolution of disease severity*, Ann. Rev. Ecol. Syst., 14 (1983), pp. 465–485.

[46] ——, *Evolution of infectious disease*, Oxford University Press, Oxford, UK, 1994.

[47] V. V. Ganusov and R. Antia, *Trade-offs and the evolution of virulence of microparasites: do details matter?*, Theor. Popul. Biol., 64 (2003), pp. 211–220.

[48] P. Geenen, *F4+ Escherichia coli in piglets: effect of host characteristics on population dynamics*, PhD Thesis, Universtity Utrecht, The Netherlands, 2005.

[49] S. Geritz, J. Metz, E. Kisdi, and G. Meszena, *The dynamics of adaptation and evolutionary branching*, Phys. Rev. Letters, 78 (1997), pp. 2024–2027.

[50] S. A. H. Geritz, *Resident-invader dynamics and the coexistence of similar strategies.*, J. Math. Biol., 50 (2005), pp. 67–82.

[51] S. A. H. Geritz, E. Kisdi, G. Meszena, and J. A. J. Metz, *Evolutionary singular strategies and the adaptive growth and branching of the evolutionary tree*, Evol. Ecol., 12 (1998), pp. 35–57.

[52] M. A. Gilchrist and D. Coombs, *Evolution of virulence: Interdependence, constraints and selection using nested models*, Theor. Pop. Biol., 69 (2006), pp. 145–153.

[53] G. M. Gomes and G. F. Medley, *Dynamics of multiple strains of infectious agents coupled by cross-immunity: A comparison of models.* Castillo-Chavez, Carlos (ed.) et al., Mathematical approaches for emerging and reemerging infectious diseases: Models, methods, and theory. Proceedings of a workshop, integral part of the IMA program on mathematics in biology. New York, NY: Springer. IMA Vol. Math. Appl. 126, 171-191, 2002.

[54] D. Greenhalgh, O. Diekmann, and M. C. M. de Jong, *Subcritical endemic steady states in mathematical models for animal infections with incomplete immunity*, Math. Biosci., 165 (2000), pp. 1–25.

[55] B. T. Grenfell, O. G. Pybus, J. R. Gog, J. L. N. Wood, J. M. Daly, J. A. Mumford, and E. C. Holmes, *Unifying the epidemiological and evolutionary dynamics of pathogens*, Science, 303 (2004), pp. 327–332.

[56] J. Guckenheimer and P. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, vol. 42 of Applied Mathematical Sciences, Springer-Verlag, New York, 1983.

[57] M. Gyllenberg, *On the definition of species in structured population model*, Unpublished manuscript, (2001).

[58] M. Gyllenberg and G. F. Webb, *Age-size structure in population with quiescence*, Math. Biosci., 86 (1987), pp. 67–95.

[59] P. Haccou, P. Jagers, and V. Vatutin, *Branching processes: variation, growth, and extinction of populations*, Cambridge studies in Adaptive Dynamics, Cambridge University Press, 2005.

[60] M. E. Hochberg and R. D. Holt, *The coexistence of competing parasites. I. The role of cross-species infection*, Am. Nat., 136 (1990), pp. 517–541.

[61] J. Hofbauer and K. Sigmund, *The theory of evolution and dynamical systems*, vol. 7 of London Mathematical Society Student Texts, Cambridge University Press, Cambridge, 1988. Mathematical aspects of selection, Translated from the German.

[62] G. E. Hutchinson, *An introduction to population ecology*, Yale University Press, New Haven, Conn., 1978.

[63] S. E. Kingsland, *Modeling nature. Episodes in the history of population ecology*, The University of Chicago Press, Chicago, 1985.

[64] D. Klinkenberg and J. A. P. Heesterbeek, *A simple model for the within-host dynamics of a protozoan parasite*, Proc. Roy. Soc. B, 272 (2005), pp. 593–600.

[65] B. Kooi, M. P. Boer, and S. Kooijman, *Resistance of a food chain to invasion by a top predator*, Math. Biosci., 157 (1999), pp. 217–236. Deterministic models with applications in population dynamics and other fields of biology (Sofia, 1997).

[66] W. Krueger, F. Lenhart, G. Neeser, G. Ruckdeshel, H. Schreckhase, H. Eissner, H. Forst, J. Eckart, K. Peter, and K. Unertl, *Influence of combined intravenous and topical antibiotic prophylaxis on the incidence of infections, organ dysfunctions, and mortality in critically ill surgical patients: a prospective, stratified, randomized, double-blind, placebo-controlled clinical trial.*, Am. J. Respir. Crit. Care Med., 166 (8) (2002), pp. 1029–1037.

[67] R. E. Lenski and R. M. May, *The evolution of virulence in parasites and pathogens: reconciliation between the competing hypotheses*, J. Theor. Biol., 169 (1994), pp. 253–265.

[68] S. Lessard and S. Karlin, *A criterion for stability-instability at fixation states involving an eigenvalue one with applications in population genetics*, Theoret. Population Biol., 22 (1982), pp. 108–126.

[69] S. A. Levin, *Coevolution*. In H. Freedman and C. Strobeck, eds.: Population Biology. Lecture notes in biomathematics 52, pp. 328–334, 1983.

[70] ——, *Some approaches to the modelling of coevolutionary interactions.* In M. Nitecki, eds. Coevolution, pp. 21–65, 1983.

[71] S. A. Levin and D. Pimentel, *Selection of intermediate rates of increase in parasite-host systems*, Am. Nat., 117 (1981), pp. 308–315.

[72] C.-K. Li and H. Schneider, *Applications of Perron-Frobenius theory to population dynamics*, J. Math. Biol., 44 (2002), pp. 450–462.

[73] W. Lingnau, J. Berger, F. Javorsky, M. Fille, F. Allerberger, and H. Benzer, *Changing bacterial ecology during a five-year period of selective intestinal decontamination.*, J. Hosp. Infect., 39 (3) (1998), pp. 195–206.

[74] W. Lingnau, J. Berger, F. Javorsky, P. Lejeune, N. Mutz, and H. Benzer, *Selective intestinal decontamination in multiple trauma patients: prospective, controlled trial.*, J. Trauma, 42(4) (1997), pp. 687–694.

[75] M. Loeb, C. Main, C. Walker-Dilks, and A. Eady, *Antimicrobial drugs for treating methicillin-resistant staphylococcus aureus colonization.*, Cochrane Database Syst. Rev., 4:CD003340 (2003).

[76] A. J. Lotka, *Relation between birth rates and death rates*, Science, 26 (1907), pp. 21–22.

[77] K. Magnusson, *Destabilizing effect of cannibalism on a structured predator-prey system*, Math. Biosci., 155 (1999), pp. 61–75.

[78] C. Matessi and C. D. Pasquale, *Long-term evolution of multilocus traits*, J. Math. Biol., 34 (1996), pp. 613–653.

[79] R. M. May and R. M. Anderson, *Epidemiology and genetics in the coevolution of parasites and hosts*, Proc. Roy. Soc. Lond. B Bio, 219 (1983), pp. 281–313.

[80] J. Metz and O. Diekmann, *The dynamics of physiologically structured populations*, vol. 68 of Lecture notes in biomathematics, Springer, Berlin, 1986.

[81] J. A. J. Metz, S. A. H. Geritz, G. Meszéna, F. J. A. Jacobs, and J. S. van Heerwaarden, *Adaptive dynamics, a geometrical study of the consequences of nearly faithful reproduction.* van Strien, S. J. (ed.) et al., Stochastic and spatial structures of dynamical systems. Proceedings of the meeting, Amsterdam, Netherlands, January 1995. Amsterdam: North-Holland. Verh. Afd. Natuurkd., 1. Reeks, K. Ned. Akad. Wet. 45, 183-231, 1996.

[82] J. A. J. Metz, S. D. Mylius, and O. Diekmann, *When does evolution optimise? On the relation between types of density dependence and evolutionarily stable life histories.* IIASA working paper WP-96-04. Available online at: http://www.iiasa.ac.at/cgi-bin/pubsrch?WP96004, 1996.

[83] L. A. Meyers, B. R. Levin, A. R. Richardson, and I. Stojiljkovic, *Epidemiology, hypermutation, within-host evolution and the virulence of neisseria meningitidis*, Proc. Roy. Soc. Lond. B, 270 (2003), pp. 1667–1677.

[84] J. Mosquera and F. R. Adler, *Evolution of virulence: a unified framework for coinfection and superinfection*, J. Theor. Biol., 195 (1998), pp. 293–313.

[85] A. Murase, T. Sasaki, and T. Kajiwara, *Stability analysis of pathogen-immune interaction dynamics.*, J. Math. Biol., 51 (2005), pp. 247–267.

[86] S. D. Mylius and O. Diekmann, *On evolutionarily stable life histories, optimization and the need to be specific about density dependence*, Oikos, 74 (1995), pp. 218–224.

[87] M. G. Neubert and M. Kot, *The subcritical collapse of predator populations in discrete-time predator-prey models*, Math. Biosci., 110 (1992), pp. 45–66.

[88] M. A. Nowak and R. M. May, *Superinfection and the evolution of parasite virulence*, Proc. Roy. Soc. Lond. B, 255 (1994), pp. 81–89.

[89] ——, *Virus Dynamics: Mathematical principles of immunology and virology*, Oxford University Press, Oxford, UK, 2000.

[90] E. Pachepsky, F. Lutscher, R. M. Nisbet, and M. A. Lewis, *Persistence, spread and the drift paradox*, Theor. Popul. Biol., 67 (2005), pp. 61–73.

[91] A. S. Perelson, D. E. Kirschner, and R. de Boer, *Dynamics of HIV infection of $CD4^+$ T cells.*, Math. Biosci., 114 (1993), pp. 81–125.

[92] A. Pugliese, *Evolutionary dynamics of virulence*. Available online at: http://www.science.unitn.it/ pugliese/.

[93] A. Pugliese, *On the evolutionary coexistence of parasite strains*, Math. Biosci., 177/178 (2002), pp. 355–375.

[94] M. Renardy and R. C. Rogers, *An introduction to partial differential equations*, vol. 13 of Texts in Applied Mathematics, Springer-Verlag, New York, second ed., 2004.

[95] R. Ross, *Some quantitative studies in epidemiology*, Nature, 87 (1911), pp. 466–467.

[96] W. Rudin, *Functional analysis*, International Series in Pure and Applied Mathematics, McGraw-Hill Inc., New York, second ed., 1991.

[97] J. Saldaña, S. F. Elena, and R. V. Solé, *Coinfection and superinfection in RNA virus population: a selection-mutation model*, Math. Biosci., 183 (2003), pp. 135–160.

[98] H. H. Schaefer, *Banach Lattices and Positive Operators*, Springer-Verlag, New York, 1974.

[99] F. R. Sharpe and A. J. Lotka, *A problem in age distribution*, Philosophical Magazine, 21 (1911), pp. 435–438.

[100] H. L. Smith and P. Waltman, *The theory of the chemostat*, vol. 13 of Cambridge Studies in Mathematical Biology, Cambridge University Press, Cambridge, 1995. Dynamics of microbial competition.

[101] V. H. Smith and R. D. Holt, *Resource competition and within-host disease dynamics*, Tree, 11 (1996), pp. 386–389.

[102] H. R. Thieme, *Pathogen competition and coexistence and the evolution of virulence*. To appear in: Mathematics for Life Sciences and Medicine, Springer.

[103] H. R. Thieme, *Mathematics in population biology*, Princeton Series in Theoretical and Computational Biology, Princeton University Press, Princeton, NJ, 2003.

[104] M. van Baalen and M. W. Sabelis, *The milker-killer dilemma and spatially structured predator-prey interactions*, Oikos, 74 (1995), pp. 391–400.

[105] P. van den Driessche and J. Watmough, *Reproduction numbers and subthreshold endemic equilibria for compartmental models of disease transmission*, Math. Biosci., 180 (2002), pp. 29–48. John A. Jacquez memorial volume.

[106] H. van Saene, A. Petros, G. Ramsay, and D. Baxby, *All great truths are iconoclastic: selective decontamination of the digestive tract moves from heresy to level 1 truth.*, Intensive Care Med., 29 (5) (2003), pp. 677–690.

[107] C. Verwaest, J. Verhaegen, P. Ferdinande, M. Schetz, G. van der Berghe, and P. Lauwers, *Randomized, controlled trial of selective digestive decontamination in 600 mechanically ventilated patients in a multidisciplinary intensive care unit.*, Crit. Care Med., 25 (1) (1997), pp. 63–71.

[108] S. Wiggins, *Introduction to applied nonlinear dynamical systems and chaos*, vol. 2 of Texts in Applied Mathematics, Springer-Verlag, New York, second ed., 2003.

Samenvatting

Dit proefschrift gaat over vier vragen met biologische of medische achtergrond. Modellen worden geformuleerd, geanalyseerd en vervolgens worden de antwoorden geïnterpreteerd. Het grootste gedeelte van het proefschrift handelt over verschillende aspecten van de dynamica van infectieziekten.

Stel dat er een nieuwe populatie geïntroduceerd wordt in een stabiele bestaande populatie. Hierbij kan gedacht worden aan een roofdier, dat de huidige populatie als prooi heeft, of aan een infectieziekte in een vatbare populatie. Als de reproductie ratio, aangegeven door $\mathcal{R}_0$, voor de binnenkomende populatie groter dan 1 is, dan kan de nieuwe populatie gaan groeien in de bestaande situatie. Dit is onmogelijk als $\mathcal{R}_0 < 1$. Onze vraag is wat er gebeurt als $\mathcal{R}_0$ de kritische waarde van 1 passeert?

In Hoofdstuk 2 geven we een antwoord op deze vraag, onder de aanname dat de individuen gekarakteriseerd kunnen worden door een eindig aantal eigenschappen (zoals bijvoorbeeld, geslacht, leeftijdscategorie, stadium van de voortgang van de ziekte, enz.) en dat de dynamica wordt beschreven als een deterministisch proces, of in de vorm van een geparameteriseerd stelsel van differentiaalvergelijkingen of als een geparameteriseerde afbeelding. We zien dat de populatie-invasie modellen (ongeacht de achterliggende biologie) specifiek gedrag vertonen, waaruit we afleiden dat het passeren van $\mathcal{R}_0 = 1$ overeenkomt met een trans-kritische bifurcatie. Vanuit wiskundig oogpunt is er maar één generiek type trans-kritische bifurcatie. In de biologische context moeten we daarentegen onderscheid maken tussen twee gevallen, afhankelijk van of de positieve tak van de evenwichten sub-kritisch is of super-kritisch. We geven een eenvoudige formule die ons in staat stelt de twee, kwalitatief zeer verschillende, scenario's te onderscheiden.

In Hoofdstuk 3 beschouwen we de evolutie van een chronisch infectieziekte. Pathogenen reproduceren en zijn onderhevig aan natuurlijke selectie op verschillende, maar nauw verbonden, niveaus. Bijvoorbeeld, pathogeen plant zich voort

en strijdt om de benodigde bouwstoffen in een gegeven geïnfecteerde gastheer, maar ook op het niveau van de populatie concurreren pathogenen om de vatbare gastheren. Ondanks dat een toenemende reproductie binnen een gastheer de verspreiding van het pathogeen op populatie niveau in eerste instantie zal doen toenemen, kan deze toenemende reproductie ook de gastheer schaden en de oorzaak zijn van stijgende sterfte. Zo rijst de volgende vraag: hoe verhouden deze twee tendensen zich tot elkaar gedurende de evolutie en in welke mate beïnvloeden de twee niveaus van voortplanting het resultaat van de evolutie? In Hoofdstuk 3 gaan we op deze vraag in, door de dynamica op het gastheer-populatie niveau van een SI (vatbaar - infectieus) model met de dynamica op het binnen-gastheer niveau te combineren. We combineren de twee niveaus door de mogelijkheid van super-infectie in het model op te nemen en we bestuderen de evolutie van de binnen-gastheer voortplantingssnelheid van het pathogeen.

In Hoofdstuk 4 beschouwen we de dynamica van pathogenen, die regelmatig op intensive care afdelingen (ICUs) gevonden worden. We richten ons in het bijzonder op de evaluatie van het relatieve effect van maatregelen die bedoeld zijn om de prevalentie van kolonisatie in ICUs te verminderen. Infecties met ziekenhuisbacteriën worden in het algemeen voorafgegaan door asymptomatisch dragerschap op één of meerdere delen van het lichaam, zoals de huid, het maag-darmkanaal en de luchtwegen. De dynamica van het pathogeen is dus afhankelijk van binnen-gastheer transmissie en van de verspreiding tussen patiënten. De verschillende manieren van verspreiding zorgen voor een ingewikkelde epidemiologie, die nog verder gecompliceerd wordt door de snelle doorstroom van patiënten en de relatief kleine populatie groottes in ICUs. In Hoofdstuk 4 introduceren we een model, dat het feit dat een persoon op meerdere lichaamsdelen gekoloniseerd kan zijn, meeneemt. We onderzoeken het relatieve effect van preventiemaatregelen en profylactische antibiotica op de prevalentie van kolonisatie. Ons onderzoek kan worden toegepast op verschillende pathogenen, die gevonden worden in ICUs, zoals *Pseudomonas Aeruginosa*, MRSA, *enterococci* en enterische *Gram-negatieve* bacteriën.

Hoofdstuk 5 handelt over de binnen-gastheer dynamica van enterotoxigene *Escherichia coli* (ETEC) in biggetjes. ETEC is één van de veroorzakers van "post-weaning" diarree (PWD), een ziekte die regelmatig voorkomt bij biggetjes gedurende de eerste twee weken na het spenen en die kan leiden tot ernstige gezondheidsproblemen of zelfs de dood. We introduceren en bestuderen een eenvoudig model, dat de dynamica in de darmen van een geïnfecteerd biggetje op microbieel niveau beschrijft. We richten ons op de randvoorwaarden die nodig zijn voor het overleven van ETEC in twee scenario's. Om te beginnen bepalen we de Malthusiaanse parameter voor het zogenaamde 'één infectie geval', waar een biggetje geïnfecteerd wordt met een enkele dosis ETEC op tijdstip $t = t_0$. Na dit tijdstip komen geen bacteriën meer de darmen binnen. Omdat het waarschijnlijk

is dat biggetjes in contact komen met uitwerpselen die bacteriën bevatten, onderzoeken we verder nog het geval dat een biggetje geherinfecteerd wordt met een fractie van de uitgescheiden bacteriën. Voor beide scenario's onderzoeken we of er een stabiele bacteriële distributie bestaat.

Acknowledgements

Several people have had a great influence on my life in the last four years, both on scientific as well as on private life.

On the scientific side, the first two people I wish to acknowledge are my supervisors Odo Diekmann and Marc Bonten. Odo, it has now been almost exactly five years since we met during a conference in Bedlewo. Your book 'Mathematical epidemiology of infectious diseases' and the discussions we had in Poland opened a whole new view on mathematics and biology to me and it is then that I decided to study mathematical biology and to come to the Netherlands. Thank you for giving me the freedom to work on such a variety of problems and for carefully reading all of my manuscripts. Marc, thanks for your time and expertise invested in Chapter 4.

I wish to thank the members of the reading committee, Rob de Boer, Gabriela Gomes, Mats Gyllenberg, Hans Heesterbeek and Hans Metz. In particular, I wish to thank Mats Gyllenberg for the many useful comments on Chapters 2 and 5 and Hans Metz for his valuable suggestions on Chapters 2 and 3.

During the four years of my PhD studies I attended several workshops and summer schools throughout Europe, three of which took place at the Instituto Gulbenkian de Ciência in Oeiras, Portugal, and were organised by Gabriela Gomes. Gabriela, it was a pleasure for me to take part in those summer schools. The atmosphere was always very friendly and stimulating and I had learned a lot there!

I wish to thank Hans Heesterbeek for his remarks on an early version of Chapter 3 that were made during one of our reading club meetings. Thanks also for co-writing the aforementioned book!

I am indebted also to Pieter Trapman; Pieter, thank you for your help on 'Samenvatting'!

Janez and my family have been of tremendous support to me during these last years: even though miles away you have always stayed close and I am very grate-

ful for all your love. Mom, thanks for Alice and for the beans (you know) - without them I would never have come this far. Eva and Lado, thank you for the beautiful cover design! And last, but most certainly not least: thank you, Janez, for your love, patience and your constant encouragements!

Curriculum Vitae

Barbara Boldin was born on the 5th of December 1976 in Jesenice, Slovenia, where she also grew up.

In 1995 she began to study mathematics at the University of Ljubljana, where she graduated in September 2000. She then continued with post-graduate studies of theoretical mathematics in Ljubljana, which she completed in July 2003 with the master's thesis entitled "Deterministic structured population epidemic models". In September 2003 she came to Utrecht to begin her PhD studies at the Department of Mathematics.

In the years 2003-2007, Barbara assisted in several math courses and attended workshops in Oberwolfach (Germany), Milano (Italy), Bath (England) and in Oeiras (Portugal). Four years of research, performed under supervision of prof. dr. Odo Diekmann and prof. dr. Marc Bonten, led to this thesis.