University of Ljubljana, Slovenia
FACULTY OF MATHEMATICS AND PHYSICS

Barbara Boldin

Deterministic structured population epidemic models

Master Thesis

Ljubljana, 2003

Contents

Preface

The aim of the thesis is to consider the invasion question for heterogeneous populations, that is, we are concerned with determining when will an infectious disease, after entering a heterogeneous population of hosts, cause an epidemic.
To do that we adopt two views of looking at the initial phase of a potential epidemic - the generation perspective and the real time growth perspective.

After the introductory part we first deal with the generation perspective.

We introduce the concept of a structured population and i-states and make a first stab at dealing with the i-state dynamics.

In the third Chapter we introduce the basic reproduction ratio R_0 and show a way to the mathematical definition of R_0. We review some of the known results regarding the basic reproduction ratio and the next generation operator and give some examples of R_0 calculation.

Chapter 4 is a mediator. On the one hand it fulfilles the promises made in the previous chapters and gives a detailed, time structured description of the next generation of infections, put in a most general setting where the distribution of individuals in the i-state space is described by a positive measure. On the other hand it prepares the ground for studying the (potential) epidemic outbreak in real time.

The latter is dealt with in a general setting in Chapter 5. We describe the population size and composition with a positive measure on the i-state space and characterise the real time growth rate by considering the Laplace - Stieltjes transform of the next generation operator. Several suffiicient mathematical conditions that guarantee the existence of the intrinsic growth rate r can be found - we are interested in those that offer biological interpretation and moreover pose as little restrictions in real situations as possible.

Chapter 1

Introduction

1.1 Mathematical epidemiology - what is it all about

Epidemiology is a study of infectious diseases, the causes of their occurence and their spread in space and time.
Mathematical epidemiology is about obtaining understanding of biological phenomena, translating assumptions regarding biological features to mathematical language, finding solutions of mathematical problems and last, but certainly not least, translating the results back to biology.
The main use of mathematics in epidemiology is to gain insight on epidemics, too see how the dynamics of an infectious disease depends on the basic parameters that characterise it.
Reality, however, is complex and even the most involved of the models are only sketches of it. To be able to describe the situation mathematically and to extrapolate on the basis of the mathematical results the modeler is first confronted with a challenge of making suitable simplifications.
But whichever the model, we must understand how the dynamics depends on the basic components of the model and how sensitive these parameters are. The latter is especially important when one makes conclusions concerning prediction and control.

1.2 How and when did it all begin

The first records of epidemic outbreaks and people's desire to comprehend the causes of epidemics go back as far as the ancient Greeks, e.g. the *Epidemics* of Hippocrates. The quest for clearer insight on epidemics, understanding the causes and the dynamics of infectious diseases had begun, but

the stage for (mathematical) epidemiology wasn't set until the 19th century. First of all, the mathematics itself was in the stage of development and secondly, but no less importantly, it was only in the 19th century that important discoveries in bacteriology were made.
There was, however, at least one important publication prior to those discoveries : in 1760 the Swiss mathematician Daniel Bernoulli studied the dynamics of an infectious disease, in particular, he considered the effect of cow-pox innoculation on the spread of smallpox. In his paper Bernoulli not only formulated and solved nonlinear ordinary differential equations but also translated mathematical results back to biological terms and this was probably the first time a mathematical model was used to evaluate the effectiveness of a vaccination programme.
By the end of the 19th century intense bacteriological research, mostly due to Pasteur and Koch, provided an insight into general mechanisms of epidemic spread and this, combined with mathematical theories allowed more elaborate studies in mathematical epidemiology, as opossed to purely empirical descriptions. In 1911 sir Ronald Ross, a medical doctor who is considered a founding father of modern epidemiology wrote :

> " ... As a matter of fact all epidemiology, considered as it is with variation of disease from time to time or from place to place, *must* be cosidered mathematically (...), if it is to be considered scientifically at all. (...) And the mathematical method of treatment is really nothing but the application af careful reasoning to the problems at hand. "

In his papers (1909 - 1917)[1] Ross studied various aspects of the transmission of malaria. He showed that there was a quantity, later to be named *the basic reproduction ratio*, which, being suppressed below one, guarantees the dissapearance of malaria from that area and that this quantitiy depends on the ratio of mosquito density to human density. Ross also developed a more general theory, which he called *a priori pathometry*.
A decade later, in 1927, Kermack and McKendrick [2] introduced an even greater degree of generality considering a large class of models. Their most outstanding result is the so called *threshold theorem*, according to which the introduction of a disease causing organism into a completely susceptible

[1] **Ross**, 1909, The Prevention of Malaria

[2] **Kermack, W.O., McKendrick, A.G.**, 1927, Contributions to the mathematical theory of epidemics, part I; 1932, Contributions to the mathematical theory of epidemics, part II (the problem of endemicity); 1933, Contributions to the mathematical theory of epidemics, part III (further studies of the problems of endemicity)

population will (under various assumptions) not give rise to an epidemic if the number of susceptibles is below a certain value. They also considered the problem of endemic diseases and related their findings to experimental epidemics in mice populations.
And the rest, they say, is history. Since then mathematical epidemiology has become a fast growing and well recognised subject and we don't even attempt to give a full description of it's developement, but rather reffer our reader to the back of this thesis for some references.

1.3 The basic questions of mathematical epidemiology

Let us consider a population of hosts. Suppose this population is *closed*, meaning that characteristic time scale of the disease is negligible compared to demographic turnover. Let us also suppose that the population is free from infection we are interested in, we call such population *virgin*, and that the disease causing organism is, in one way or another, introduced into the population. One might then be eager to know:

- does this introduction cause an epidemic ?

This is reffered to as *the invasion question*.
If the epidemic occurs,

- with what rate does the number of infecteds increase during the rise of the epidemic? When does this number reach maximum? How large will this maximum be?
- what proportion of the population will ultimately have experienced infection?

These were the questions that Kermack and McKendrick posed in their groundbreaking paper, and with some additional assumptions, they were also answered. Those assumptions were:

1. an infection triggers an autonomous process within the host; stating differently they considered microparasites,
2. contacts are made according to the mass action,
3. all individuals are equally susceptible,
4. the disease leads to either death or immunity,

5. the population is large enough to allow the deterministic description.

If we now release the closed population assumption and allow the inflow of new susceptibles it happens that, when the epidemic fades out, the population eventually becomes replenished by the inflow of new susceptibles. When the population is large enough again, another introduction of a disease causing agent will again cause an epidemic. If the disease is continually being transmitted it is called *endemic*. One might then ask:

- will there be a stable steady state? Or will there be oscillations?

Finally, there is the *regulation question* - how does the disease affect the growth rate of the population?

As mentioned in the Preface, we shall only be concerned with answering the invasion question for heterogeneous populations.

Chapter 2

What is structure ?

When an infectious disease enters a population of hosts, the individuals may differ from one another in characteristics that influence the spread of the disease. For example, think that for a certain disease the susceptibility to the infection changes with age, or, perhaps, that transmission probabilities vary for males and females,...
Thus, to make the model more realistic, we distinguish individuals according to those characteristics that influence the spread of the disease, including *contact pattern*, *susceptibility* and *infectivity*, and describe the way these characteristics change themselves.
So the first step in building a structured population model is to choose the characteristics that are relevant for the spread of the infection one is interested in, more accurately, to give precisely those characteristics that are needed to describe the future behaviour of the system. This is called *choosing the i-state*, where i stands for "individual." When choosing i-states one has to pay attention to infection transmission as well as population dynamics and we shall thus divide the i-state into two parts. Let us call the part of the i-state that describes the development of the infection within the individual the *d-state* (where d stands for "disease"). Stating more precisely, the d-state is the part of the i-state that describes the difference between infected and susceptible individuals.
The rest of the i-state reflects heterogeneity in the population and we shall refer to it as the individual's *h-state*.

2.1 d-states

The d-state determines the infectious output of an individual. Of course, response to an infection is a stochastic process that depends, among other things, on the individual's immune system. However, we shall avoid modelling this complex process and from our point of view there are but two kinds of d-states, *age of infection* and *degree of infection*.

- Age of infection

Once infection takes place, the infectious agent reproduces within the host with such a rate that further infections are no longer relevant. In other words, we treat the infection as a unique event. This includes viruses, bacteria and most protozoan diseases, for example measles, rabies, HIV,... We shall refer to such agents as *the microparasites*.
If there is no heterogeneity in the population, the only thing that determines the infectious output of an individual is the time t elapsed since the infection took place. We will call that time the individual's *infection age* and will denote the expected infectivity at time t with $A(t)$.
The moment the individual became infected is also called *the time of birth*, since, in that moment, the individual is born from the epidemiological point of view.
Not to be too wordy about it, let us now give an example.

Example 1

Assume that the population is closed with total size N, denote with S the number of susceptibles, with E the number of exposed (but not yet infectious),with I the number of infectious individuals and R the number of removed (meaning either immune, dead or in quarantine).
The *force of infection* is defined as a probability per unit of time for a susceptible to become infected. We shall assume that the force of infection is proportional to I and that the constant of proportionality is β. Moreover, assume that after infection an individual enters class E and that the time spent in E is exponentially distributed with parameter θ. After the latency period the individual enters compartment I and each infectious individual has a constant probabilty α per unit of time to become removed. Let us also assume that removed individuals are permanently immune.
Thus, the d-state variable has four possible values, S, E, I and R and the dynamics between the four compartments is described with the system of ordinary differential equations:

$$\frac{dS}{dt} = -\beta SI,$$

$$\begin{aligned}\frac{dE}{dt} &= \beta SI - \theta E,\\ \frac{dI}{dt} &= \theta E - \alpha I,\\ \frac{dR}{dt} &= \alpha I.\end{aligned}$$

If we also assume that every contact between an infected and a susceptible leads to a transmission with constant probability p and denote with $P(t, E)$ the probability to be in latency period at time t and with $P(t, I)$ the probability to be in class I at time t, then

$$A(t) = pP(t, I).$$

We assumed that the latency period is exponentialy distributed with parameter θ, therefore

$$P(t, E) = e^{-\theta t}.$$

In order to be in class I at time t, the individual must have entered I at some time $\tau \in (0, t)$ and stayed there in the interval (τ, t). Therefore we expect to get

$$P(t, I) = \int_0^t e^{-\alpha(t-\tau)}\theta P(\tau, E)d\tau$$

Indeed, as

$$\frac{dI}{dt} = \theta P(t, E) - \alpha P(t, I),$$

we get (after a short variaton-of-constants calculation) the desired result and with some more calculus

$$P(t, I) = \frac{\theta}{\alpha - \theta}(e^{-\theta t} - e^{-\alpha t}).$$

So

$$A(t) = p\frac{\theta}{\alpha - \theta}(e^{-\theta t} - e^{-\alpha t}).$$

The example is an $SEIR$ model, one of the so called *compartmental* models. Other examples include SIR, SIS, $SIRS$,... models, with each letter denoting the compartment one can reside in.

- Degree of infection

In this case the infection is a repeated process rather than a unique event. Diseases caused by helminths and other worm-like parasites fall into this category and we call such organisms *macroparasites*.

We shall restrict ourselves to microparasites, let us just remark that although age of infection and degree of infection seem like two distinct categories there are diseases, for example malaria, that belong to both.

2.2 h-states

Features on which an individual's h-state is based on can take *discrete values* (like sex, partnership status) or *continuous values* (for example age, spatial position). Furthermore, they can be *static* or *dynamic*.
We shall denote with Ω the state space of all possible values of the chosen set of characteristics and assume that it is a measurable space with a countably generated σ-algebra . Then Ω is called the *h-state space*.
In case when the d-state of an individual has a d-age representation and h-state can have more than one value we have to take into account that the probability of transmission may depend on both individuals involved in the contact, the infected and the susceptible.
Also, if the h-state is dynamic, or has dynamic components, we have to describe this dynamics.

The h-state has the following property: no matter what its dynamics are in the time interval $[0, t]$, if we want to know the h-state at time $t + s$, we only need to know the h-state at time t. In other words, the h-state at time t carries with it the relevant information of its dynamics in the interval $[0, t)$. Because of this feature, the theory of continuous time Markov chains comes naturally to describe h-state dynamics.
We shall deal with the h-state dynamics in the general setting in the following chapter. In this part we shall only consider one special example.

Let the h-state space be finite, say $\Omega = \{1, 2, ..., n\}$ and let $P(t) = (p_{ij}(t))_{i,j=1}^{n}$ denote the transition matrix,

$p_{ij}(t)$ = *the probability of being alive and having h-state i at time t, given one has h-state j at time* 0.

Then

$$P(0) = I \tag{2.1}$$

and the Chapman - Kolmogorov relation gives us

$$P(t+s) = P(t)P(s); \quad t, s \geq 0. \tag{2.2}$$

Thus, the set $\{P(t); t \geq 0\}$ forms a one parameter semigroup.

The often used modelling approach is to assume that the time spent in a certain h-state is exponentially distributed with some parameter. So let Σ describe jumps and let the diagonal matrix D contain (strictly positive) death rates. The system of ODE

$$\frac{d}{dt}P(t) = (\Sigma - D) \cdot P \tag{2.3}$$

then describes the changes in h-states and along with (2.1) we obtain

$$P(t) = e^{(\Sigma - D)t}.$$

Now let us, in order to make some additional interpretation concerning h-state dynamics, define *the spectral bound* of a matrix A as

$$s(A) = \max\{Re\lambda; \lambda \text{ is an eigenvalue of } A\}.$$

Since all the columns of Σ add to zero and the diagonal elements of D are strictly positive, the spectral bound of the matrix $\Sigma - D$ is negative and thus the integral

$$\int_0^\infty e^{(\Sigma - D)t} dt$$

converges. Using the Taylor series we find

$$(\Sigma - D)\int_0^t e^{(\Sigma - D)\tau} d\tau = \int_0^t e^{(\Sigma - D)\tau} d\tau (\Sigma - D) = e^{(\Sigma - D)t} - I.$$

Finally, taking the limit $t \to \infty$ we obtain the identity

$$\int_0^\infty e^{(\Sigma - D)t} dt = -(\Sigma - D)^{-1} \tag{2.4}$$

that offers some interpretation, that is, the left hand side of (2.4) tells us that $(-(\Sigma - D)^{-1})_{ij}$ equals the expected time an individual will spend in state i in the rest of its life after we have observed the individual in state j.

Now, the assumption of exponentially distributed times spent in h-states guarantees that

$$\lim_{t \to 0} p_{ij}(t) = \delta_{ij}, \tag{2.5}$$

which, along with (2.1) and (2.2), makes $\{P(t); t \geq 0\}$ a *strongly continuous* semigroup.
When we are dealing with a general h-state space, the semigroup property of $\{P(t); t \geq 0\}$ remains. The additional assumption of continuity in the h-state dynamics (for example (2.5)) guarantees the existence of an operator Q (which is $\Sigma - D$ in our notation), called the *infinitesimal generator*, such that

$$\frac{d}{dt}P(t) = Q \cdot P(t).$$

We shall not go into details here but rather refer to Rudin [1] and Pazy[2], which also verify the validity of (2.4).

Thus, when we have continuity in the h-state movement, the dynamics can be described in terms of an infinitesimal generator.
In situations where the individual's h-state changes in a discontinuous manner every attempt to describe its dynamics in terms of an infinitesimal generator fails and we need to describe h-state dynamics in some other way. One alternative approach is given in Chapter 4.

[1] **Rudin** W., Functional Analysis
[2] **Pazy** A., Semigroups of Linear Operator and Applications to Partial Differential Equations

Chapter 3

R_0

3.1 The basic reproduction ratio

Suppose an infectious agent enters a population of hosts that is in a demographic steady state. To answer the invasion question, that is, to find out whether this introduction leads to an epidemic, we consider the subpopulation of infected individuals from a generation perspective.
If there is no heterogeneity in the population we define the *basic reproduction ratio* R_0, as

$$R_0 := \textit{the expected number of new infections caused by one infected individual during its entire period of infectiousness.}$$

The basic reproduction ratio has threshold value 1 : the introduction leads to an epidemic if and only if $R_0 > 1$. When considering homogeneous populations one is often able to write an explicit expression for R_0 just by interpreting the biological problem. Indeed, if we return briefly to the Example 1 we find

$$R_0 = \frac{\beta}{\alpha} \cdot S,$$

with the following interpretation : each infected individual has an expected duration of infectious period $\frac{1}{\alpha}$. In that time, given a contact with the susceptible, there is a constant probability per unit of time β that the transmission is "successful". Therefore, the number of new cases per unit of time one infected produces is βS. So, clearly, the formula for R_0 follows.

In heterogeneous population, however, finding R_0 is not so straightforward. First of all, since we now distinguish individuals, the definition of the basic reproduction ratio now reads:

$R_0 :=$ *the expected number of new infections caused by a "typical" infected individual in its entire period of infectiousness.*

Of course, one can calculate the expected number of new cases with h-state $\xi \in \Omega$ caused by one individual that was born with h-state $\eta \in \Omega$. That way we obtain $R_0(\xi, \eta)$ for all $(\xi, \eta) \in \Omega \times \Omega$. The question is then, how to average all these numbers ?
The aim of this section is to provide the answer and to explain what we mean by "typical" in the definition of R_0. What we need is an average that would have almost the same meaning as in the homogeneous case: the introduction leads to an epidemic if and only if $R_0 > 1$.

Although the concept of R_0 dates back to 1909[1] and a lot of epidemic models have been made since, finding the right average for a general h-state space Ω was long an open problem in mathematical epidemiology and was solved in 1990[2]. Until then all "definitions" of R_0 were either wrong but gave the right result because of the special situations they were dealing with or were right mathematical definitions for special cases of Ω.

The essential assumptions that allow the answer to an invasion question in a general setting are

- in the initial phase of a (potential) epidemic the decreasing of the number of susceptibles can be neglected,
- as we also assume that populations are large, every contact an infected has is one with a susceptible.

Let us now present the idea that lead to the mathematical definition of R_0.

Finding the next generation of infected individuals means first studying the process at the individual level - how many infections does one infected individual with, say, h-state η cause ? How are these infections distributed in Ω ? And also, if the h-state is dynamic, how does the individual's h-state change with time? The next step would then be to add contributions of all individuals to get the next generation of infected individuals.
To get the idea we shall first assume that the h-state space is finite. Let $\{1, 2, ..., n\}$ be the set of all possible states. We shall postpone taking the h-state dynamics into account by assuming that the individual's h-state remains constant from the moment it becomes infected on. Let us denote

[1] **Ross**, 1909, The Prevention of Malaria

[2] **O. Diekmann, J.A.P. Heesterbeek, J.A.J. Metz**, 1990, On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations

k_{ij} : = *the expected number of new cases with h-state i caused by one individual that was itself infected while having h-state j, during its entire period of infectiousness.*

and consider generations of infected individuals. Let $\phi = (\phi_1, ..., \phi_n)$ be the vector describing the infecteds, that is, ϕ_i is the number of infected individuals with h-state i. In deterministic description we can then say that, if ϕ^0 describes the initial distribution of infecteds, the vector describing the first generation reads

$$\phi_i^1 = \sum_{j=1}^{n} k_{ij}\phi_j^0,$$

so, if $K = (k_{ij})_{i,j=1}^n$, then

$$\phi^1 = K\phi^0.$$

Continuing in this manner we get the n-th generation of infecteds as

$$\phi^n = K\phi^{n-1} = ... = K^n\phi^0.$$

The matrix K is called *the next generation operator*.
Now, how to average all k_{ij} ? Counting the number of infecteds, the first norm that comes to mind would be (note that all $\phi_i \geq 0$)

$$\|\phi\|_1 = \sum_{i=1}^{n} \phi_i$$

and the associated matrix norm,

$$||K||_1 = \sup_{\|\phi\|_1=1} \|K\phi\|_1.$$

As for any operator norm

$$\|K\phi\| \leq \|K\|\|\phi\|$$

we see that surely $\|K\|_1$ is the upper bound for the per-generation multiplication factor for the total number of cases, but as the next simple example shows, this norm is too coarse to be taken for the definiton of R_0.

Example 2

Let us consider a host-vector model (for example, think of a sexually transmitted disease in a heterosexual population with male and female as the h-states). Then

$$K = \begin{pmatrix} 0 & k_{12} \\ k_{21} & 0 \end{pmatrix}$$

and $\|K\|_1 = \max\{k_{12}, k_{21}\}$. So, if the infection is introduced in one individual with h-state 1, $\phi^0 = (1,0)$, then $\phi^1 = (0, k_{21})$, so the number of cases is multiplied by k_{21}. The next generation, however, $\phi^2 = (k_{12}k_{21}, 0)$. If, for example $k_{12} << k_{21}$, the definition $R_0 = \|K\|_1 = k_{21}$ would give us a much too pessimistic estimate of the number of new cases in the second generation as k_{21}^2.

Therefore, $\|.\|_1$ is not a proper way to average $\{k_{12}, k_{21}\}$. Nevertheless, the example shows the way to the right average - as the multiplication factor in two generations is $k_{12}k_{21}$, the average per-generation factor is $\sqrt{k_{12}k_{21}}$. So, we look at average multiplication factor after many generations, but per-generation,

$$\lim_{n\to\infty} \|K^n\|^{\frac{1}{n}}.$$

As this is the average per-generation multiplication factor in the long run, we take it to define R_0. Moreover, it is the well known relation for the spectral radius of K (which will be denoted with $spr(K)$), so

$$R_0 := \lim_{n\to\infty} \|K^n\|^{\frac{1}{n}} = spr(K) = \sup\{|\lambda|; \lambda \in \sigma(K)\}.$$

Since the next generation operator is a positive matrix, we can make use of the broad theory on non-negative matrices to arrive at additional information on the basic reproduction ratio and the spread of the infection.

All the relevant theorems on positive operators shall be collected in the Appendix, in this part we will only be concerned about the interpretation.

Indeed, since the next generation operator is positive, R_0 is in fact an eigevalue with a positive eigenvector (see Theorem 7 in the Appendix).

Let us now see whether R_0 really has the desired threshold property.
If $R_0 < 1$, then $K^n\phi^0 \to 0$ for an arbitrary ϕ^0 as $n \to \infty$. If $R_0 > 1$, however, $K^n\phi^0$ can be made arbitrarily large for a suitable choice of ϕ^0. Nevertheless, the fact that $R_0 > 1$ doesn't necessarily mean that the introduction will lead to an epidemic. If the infection is introduced into a subpopulation that has no contacts with other subgroups, the infection will die out. In mathematical jargon this reads: in order to guarantee that every initial distribution of infecteds leads to an epidemic in case when $R_0 > 1$, we have to assume *irreducibility* of K. Irreducibility is defined as follows:

Definition 1 *A square matrix $K \geq 0$ is* **irreducible***, if there exists no permutation matrix P such that*

$$P^{-1}AP = \begin{pmatrix} K_1 & 0 \\ L & K_2 \end{pmatrix}.$$

Note that, in the context of epidemic models, the 0 in the $P^{-1}AP$ means exactly that there exists a subgroup that produces no infections "outside" its own group.

Definition 2 *An irreducible matrix K is called* **primitive***, if $spr(K)$ is the only element of the peripheral spectrum of K.*

Using the Perron - Frobenius theorem[3] we make the following conclusions about the development of the epidemic - if the next generation operator is primitive, we can talk about a stable distribution. Indeed, write

$$\phi^0 = c(\phi^d)\phi^d + \phi,$$

where the vector ϕ lies in the subspace spanned by eigenvectors (and possibly generalised eigenvectors) corresponding to eigenvalues with modulus strictly less than R_0. Then

$$K^n(\phi^0) = c(\phi^d)R_0^n\phi^d + K^n(\phi) = R_0^n(c(\phi^d)\phi^d + \frac{1}{R_0^n}K^n(\phi))$$

and since R_0 is strictly dominant, $\frac{1}{R_0^n}K^n(\phi) \to 0$ as $n \to \infty$. However, if $c(\phi^d) = 0$, the number of cases would still converge to 0, even if $R_0 > 1$. But that can not be the case, since there would otherwise exist a nontrivial ideal invariant under K. In fact, we can calculate $c(\phi^d)$ the following way: let us normalize ϕ^{d*} in a way that $\phi^{d*} \cdot \phi^d = 1$. Since ϕ^{d*} is ortogonal to ϕ, we get

$$c(\phi^d) = \sum_{j=1}^{n} \phi_j^{d*}\phi_j^0,$$

and as ϕ^{d*} is strictly positive and $\phi^0 \geq 0$ it follows that $c(\phi^d) \neq 0$.

The normalised ϕ^0 can be interpreted as the probability distribution over Ω for the initial infections. Thus, no matter what this initial distribution ϕ^0 is, after many generations the distribution of infections over Ω is approximately ϕ^d (and this approximation improves with time) and the number of infections multiplies with R_0 in every generation.
The normalised eigenvector ϕ^d can be interpreted as the probability distribution for the birth states. Not only does the dynamics lead to this stable distribution, ϕ^d is also invariant - if the initial distribution of infecteds equals ϕ^d, then

$$K^n(\phi^0) = R_0^n\phi^0,$$

[3]Appendix

the distribution remains the same, the numbers are multiplied with R_0. Thus, by "typical" in the definition of R_0 we mean that h-states are distributed over Ω according to ϕ^d.

Further biological insight can be obtained when we write the following equivalent definition of irreducibilty: a positive matrix K is irreducible if and only if for every pair i, j there exists an $m = m(i, j) \in \mathbb{N}$ such that $(K^m)_{ij} > 0$. If there exists an $m \in \mathbb{N}$ independently of i, j, that is, $K^m > 0$, then K is called primitive.

Suppose the infection is introduced in one individual that has state j, $\phi^0 = e_j$. Since

$$(K^m)_{ij} = (K^m e_j)_i,$$

the $(i, j) - th$ element of K^m represents the number of infections with h-state i in the m-th generation caused by this initial infection. So, if K is irreducible, there will be, after a chain of infections with length $m(i, j)$, infections with h-state i. If K is primitive, we can choose m independently of i, j, meaning that, no matter what the h-state of the initial infection is, after a certain number of generations there will be infected individuals with h-state $i \in \{1, 2, ..., n\}$. Moreover, for all $k > m$, the number of infected individuals with h-state i is positive as well.

The idea of the next generation operator now extends to a situation where Ω has continuous elements, but still, we keep the assumption that the individual's h-state remains fixed the moment it becomes infected.
We introduce

$k(\xi, \eta)$: = *the expected number of new cases with h-state ξ caused by one individual that was itself infected while having h-state η, during its entire period of infectiousness.*

So, if $\phi^0(\eta)$ describes the initial distribution of infected individuals, the first generation should be of the form

$$\phi^1(\xi) = \int_\Omega k(\xi, \eta)\phi^0(\eta)d\eta.$$

Thus, to assure that the next generation operator is well defined we shall assume that $k(\xi, \eta)$ is a *kernel*, that is, if (Ω, Σ, μ) is the state space, we assume that $k(\xi, \eta)$ is a $\Sigma \times \Sigma$ measurable function on $\Omega \times \Omega$ such that for each distribution of infections ϕ the function

$$\eta \mapsto k(\xi, \eta)\phi(\eta)$$

is μ-integrable for almost all $\xi \in \Omega$ (μ).
The product of two kernel operators K and L with kernels k and l is a kernel operator with kernel

$$\int_\Omega k(\xi,\zeta)l(\zeta,\eta)d\zeta.$$

Thus, powers of K are kernel operators and next generations of infected individuals are obtained by repeatedly applying the next generation operator

$$K(\phi)(\xi) = \int_\Omega k(\xi,\eta)\phi(\eta)d\eta.$$

By definition then, $R_0 = spr(K)$.

As in the case of finite Ω we have to assume, in order to guarantee that every inital distribution leads to an epidemic when $R_0 > 1$, irreducibility of K, namely that for all subsets $X \subset \Omega$ of positive measure, such that $\Omega\backslash X$ also has positive measure we have

$$\int_{\Omega\backslash X} d\xi \int_X k(\xi,\eta)d\eta > 0.$$

Note that this definition has precisely the interpretation written in the case of finite Ω. For a given distribution of infected individuals ϕ that "lives" on a subset $X \subset \Omega$, the next generation $K(\phi)$ is strictly positive on some subset of positive measure disjoint from X, that is, ϕ produces new infections "outside" X.
In the case of finite Ω we have, when the next generation operator is primitive, convergence to a stable distibution. For a similar result in this case we would need that $spr(K) > 0$ and that it is a strictly dominant eigenvalue with a corresponding positive eigenvector. One sufficient condition that guarantees that this is indeed the case is the theorem of Jentzsch[4] from 1910. In our situation, the Jentzsch's theorem is used for $p = 1$ and to make sure that the spectral radius is a strictly positive eigenvalue with a corresponding positive eigenvector we need to assume, besides irreducibility, that some power of the next generation operator is compact.

When R_0 is the only element of the peripheral spectrum (as it is in the case when the kernel is strictly positive almost everywhere), then the iteration of the next generation operator leads to a stable distribution for the state at birth.

[4]Appendix

At times we cannot describe the distribution of individuals in Ω with a density function and another way of describing the next generation shows itself useful (or even necessary). That is, we take a measurable subset $\omega \subseteq \Omega$ and specify

$\Lambda(\eta, \omega) :=$ *the expected number of cases with h-state in* ω *caused by one individual that was itself infected while having h-state* η*, during its entire period of infectiousness.*

Let us denote with $\mathcal{M}_+(\Omega)$ the set of positive measures on Ω. The word *kernel* will now have the following definition:

Definition 3 *A function* $\Lambda : \Omega \to \mathcal{M}_+(\Omega)$ *is called a* **kernel** *if for any measurable set* $\omega \subseteq \Omega$ *the function*

$$\begin{aligned} \Omega &\to [0, \infty), \\ \eta &\mapsto \Lambda(\eta, \omega) \end{aligned}$$

is measurable.

We shall describe the distribution of infections in Ω with a positive measure m on Ω. The measure m assigns to every measurable subset $\omega \subseteq \Omega$ the number of cases with h-state in ω. The next generation of infections is then given by

$$(Km)(\omega) = \int_\Omega \Lambda(\eta, \omega) m(d\eta).$$

We shall return to this general setting at the end of the next Chapter, where we shall work out the details such as when does thus defined K give us a bounded linear operator on $M(\Omega)$ and find the conditions that guarantee the convergence to a stable distribution.

Right now we only remark that this formulation generalises the formulation in terms of a kernel k. More precisely, if the measures $\Lambda(\eta, .)$ and m are absolutely continuous, the corresponding density functions exist and we are back in the previous setting.

Before we present important special examples that allow the explicit expressions of R_0, let us make the following remark.

The definition of R_0 might strike as somewhat contradictory since on the one hand we linearise, that is, neglected the diminishing of the number of the susceptibles, saying that we consider the initial phase only, and, on the other hand, defined R_0 as the long term average multiplication factor. If it

takes many generations before one observes the stable distribution it might just be that the nonlinearity is already noticeable. However, looking at data in generation perspective is artificial anyhow and whether we can observe the real time growth rate r (which we shall introduce in the next chapter) depends on the difference between r and the next eigenvalue (ordered according to the real part) and "degrees" of irreducibility, such as the ratio of the maximum and minimum component of the dominant eigenvalue.
However, the threshold property remains:

$$\text{an epidemic occurs} \iff R_0 > 1.$$

3.2 On conditions that allow explicit expressions of R_0

All the information we need to know in order to answer the invasion question is given in the basic reproduction ratio. Since it is defined as a dominant eigenvalue of the next generation operator one might breathe a sigh of relief, thinking, to compute an eigenvalue of an operator is not an easy task in general, but there are various numerical techniques that give us accurate estimates of eigenvalues and since we only need to know whether $R_0 > 1$ this should be enough. However, explicit expressions of R_0 are of great importance as we strive not only for knowing its value, but also for understanding of its dependence on the various factors. When we understand how R_0 depends on basic parameters that govern the spread we can then influence those parameters to apply various control measures, to evaluate vaccination programmes,...
In some special cases the spectral radius of an operator is easy to compute. We now present some of those situations that are also relevant biologically.

3.2.1 One dimensional range

When the range of K is one dimensional there is but one non-zero eigenvalue. Biologically this corresponds to the situation when the probability distribution for the h-state of a newly infected individual is independent of the state of the individual that caused the infection. This condition is called *separable mixing*. Written in all three variants, we assume

$$k_{ij} = a_i b_j,$$

$$k(\xi, \eta) = a(\xi) b(\eta),$$

$$\Lambda(\eta, \omega) = a(\omega)b(\eta).$$

Let us write the eigenvalue problem for the last setting, as the first two follow immediately.
Since

$$(Km)(\omega) = a(\omega) \int_\Omega b(\eta)m(d\eta),$$

the range of K is spanned by the measure a (which is also the eigenvector corresponding to R_0) and

$$R_0 = \int_\Omega b(\eta)a(d\eta).$$

Thus, for the first two settings,

$$R_0 = \sum_{j=1}^{n} a_j b_j, \quad R_0 = \int_\Omega a(\eta)b(\eta)d\eta.$$

In case when a is proportional to b we speak of *proportionate mixing.*

3.2.2 Enhanced within group contacts

Sometimes one is able to derive a threshold quantity even though R_0 can not be explicitly calculated. In other words, one derives a quantity, say Q_0 such that

$$R_0 > 1 \iff Q_0 > 1.$$

Such is the case when individuals preferably have contacts with their own kind and practice separable mixing with other subgroups. That is, we assume that

$$k_{ij} = a_i b_j + \delta_{ij} c_j,$$
$$k(\xi, \eta) = a(\xi)b(\eta) + \delta_\eta(\xi)c(\eta),$$
$$\Lambda(\eta, \omega) = a(\omega)b(\eta) + \delta_\eta(\omega)c(\eta).$$

We shall assume that $\sup_\eta c(\eta) < 1$, since otherwise one infection automatically triggers an epidemic with just contacts in its own subgroup.
The eigenvalue problem now reads

$$a(\omega) \int_\Omega b(\eta)m(d\eta) + \int_\omega c(\eta)m(d\eta) = \lambda m(\omega).$$

Integrating the function $\xi \mapsto \frac{b(\xi)}{\lambda - c(\xi)}$ over Ω with respect to both sides we get

$$\int_\Omega \frac{b(\xi)}{\lambda - c(\xi)} a(d\xi) \int_\Omega b(\eta)m(d\eta) + \int_\Omega \frac{b(\xi)c(\xi)}{\lambda - c(\xi)} m(d\xi) = \lambda \int_\Omega \frac{b(\xi)}{\lambda - c(\xi)} m(d\xi)$$

and so

$$\int_\Omega \frac{b(\xi)}{\lambda - c(\xi)} a(d\xi) \int_\Omega b(\eta) m(d\eta) = \int_\Omega b(\eta) m(d\eta).$$

Thus,

$$\int_\Omega b(\eta) m(d\eta) = 0,$$

or

$$\int_\Omega \frac{b(\xi)}{\lambda - c(\xi)} a(d\xi) = 1.$$

Irreducibility of K guarantees that, since we are looking for the dominant eigenvalue, the first possibility is ruled out. Let us define

$$Q_0 = \int_\Omega \frac{b(\xi)}{1 - c(\xi)} a(d\xi).$$

The function $\lambda \mapsto \int_\Omega \frac{b(\xi)}{\lambda - c(\xi)} a(d\xi)$ is strictly decreasing on $(\sup_\xi c(\xi), \infty)$. Thus, since $\int_\Omega \frac{b(\xi)}{R_0 - c(\xi)} a(d\xi) = 1$, clearly $R_0 > 1 \iff Q_0 > 1$.
Now, for the remaining two settings we obtain

$$Q_0 = \sum_{j=1}^{n} \frac{a_j b_j}{1 - c_j}, \quad Q_0 = \int_\Omega \frac{a(\eta) b(\eta)}{1 - c(\eta)} d\eta.$$

Note that, compared to the separable mixing, every factor $a_j b_j$ (or $a(\eta)b(\eta)$) is replaced with $\frac{a_j b_j}{1 - c_j}$. This has the following interpretation: if we consider individuals with h-state j, then, if one of them is expected to infect c such individuals, the whole clan is expected to have size $1 + c + c^2 + ... = (1 - c)^{-1}$. Treating the whole clan as one individual, we are again in the separable mixing case and thus Q_0 is the expected number of new cases at the clan level and an epidemic occurs if and only if $Q_0 > 1$.

3.2.3 Finite dimensional range

When the next generation operator has a finite range, the problem of determining R_0 translates into computing the dominant eigenvalue of a matrix. However, in general, finite dimensional range of K doesn't have a biological interpretation and we shall restrict ourselves to one special example that does have an interpretation.
Let us assume that h-state has two components, a discrete one and a continuous one and denote with Ω_c the set over which the continuous variable ranges. The discrete variable can take values $1, 2, ..., n$, so $\Omega = \cup_{j=1}^{n} \{j\} \times \Omega_c$.

Moreover, denote for a measurable set $\omega \subseteq \Omega$, $\omega = (i, \omega_c)$ and $(j, \xi) \in \Omega$ with

$$\Lambda(\eta, \omega) = \Lambda_i(j, \xi)(\omega_c).$$

So $\Lambda_i(j, \xi)(\omega_c)$ denotes the expected number of cases with discrete component i and continuous component in ω_c caused by one individual with h-state (j, ξ) in its entire period of infectiousness. For example, think of modelling a sexually transmitted disease, taking gender for a discrete part of the h-state and sexual activity for the continuous one. We shall assume that, conditional on the discrete component being i, the probability distribution of the continuous component is a_i, independently of the h-state of the one that is responsible for the infection. The dependence in the discrete component is allowed and is described by b_{ij}. Thus, we have

$$\Lambda_i(j, \xi)(\omega_c) = a_i(\omega_c) b_{ij}(\xi).$$

The eigenvalue problem now reads

$$(Km)(\omega) = \sum_{j=1}^{n} \int_{\Omega_c} \Lambda_i(j, \xi)(\omega_c) m(j, d\xi) = a_i(\omega_c) \sum_{j=1}^{n} \int_{\Omega_c} b_{ij}(\xi) m(j, d\xi),$$

so the range of K is spanned by measures a_i on Ω_c. So, if $m(i, \omega_c) = c_i a_i(\omega_c)$ then $(Km)(i, \omega_c) = a_i(\omega_c) \sum_{j=1}^{n} \int_{\Omega_c} c_j b_{ij}(\xi) a_j(d\xi)$, meaning that the coefficients with respect to the basis $\{a_j\}_{j=1}^{n}$ transform according to the matrix $L = (l_{ij})$ where $l_{ij} = \int_{\Omega_c} b_{ij}(\xi) a_j(d\xi)$. Thus, R_0 is the dominant eigenvalue of the matrix L.
This condition is also called *multigroup separable mixing* and we shall give an example of such a situation in the next section.

3.3 Building up the kernel

So far we have given the next generation operator K in terms of a kernel, say

$$(K\phi)(\xi) = \int_{\Omega} k(\xi, \eta) \phi(\eta) d\eta.$$

When making a model we need to construct the kernel k from the basic parameters that are characteristic for the disease we are interested in. That is, we need to

- make submodels for the dynamics of h-states, contact process and specify transmission probabilities,

- put all the pieces together to obtain the specific form of the kernel.

We have already presented a submodel for the h-state dynamics in case when Ω is finite. Let that do for now, we shall deal with the h-state dynamics in the general setting later, first we want to present some examples.

Regarding contact structure there is not much that can be said in any generality. One way or another we need to specify

$c(\xi, \eta) :=$ *the probability per unit of time that an individual with current state η has contact with an individual with current state ξ.*

Let the probability of transmission be described by a function $p(t)$, where t is the time elapsed since the infection took place. Furthermore if we denote with $P(t, \omega, \eta)$ the probability that an individual originally with h-state η has h-state in $\omega \subseteq \Omega$, t units of time later, the kernel has the form

$$k(\xi, \eta) = \int_0^\infty p(t) \int_\Omega c(\xi, \zeta) P(t, d\zeta, \eta) dt.$$

Let us return again to the situation where the h-state space is finite. Assuming that the transmission probability is constant we obtain the identity

$$\begin{aligned} k_{ij} &= p \int_0^\infty \sum_{l=1}^n c_{il} \cdot P_{lj}(t) dt \\ &= p \cdot \sum_{l=1}^n c_{il} \int_0^\infty (e^{Qt})_{lj} dt \\ &= -p \cdot \sum_{l=1}^n c_{il} Q^{-1}_{lj} \end{aligned}$$

which has the following interpretation: k_{ij} is the expected number of new-cases with h-state i caused by one individual with state at birth j. The time the individual is expected to spend in state l equals $-Q^{-1}_{lj}$. In this time it contacts an individual with state i with probability c_{il} per unit of time and each such contact leads to a transmission with constant probability p. Summing over all possible states, we indeed get the expression for k_{ij}.

The matrix $-Q^{-1}$ describes the h-state dynamics. Putting the information about contact and transmission in the matrix $B = p(c_{ij})_{i,j=1}^n$, we have arrived at the expression for the next generation operator

$$K = -BQ^{-1}.$$

Let us now illustrate the construction of the next generation operator on two examples.

Example 3

Suppose that plants are grown in a field where cuttings are taken at some per capita rate γ. The cuttings are planted in a nursery, where they will mature. Plants in the nursery are replanted in the field with rate ζ. The grower maintains a fixed population size of plants in the field (denoted with N_1) and plants in the nursery (N_2). Suppose a fungus spreads in this host population and that this fungus can be directly transmitted between plants in the field and between plants in the nursery. We assume mass action with transmission rates β_1 for the plants in the field and β_2 for the plants in the nursery. In addition, there is a constant probability p that a cutting taken from an infected plant is itself infected. Let μ_1 and μ_2 denote the natural death rates of field and nursery plants, respectively. Finally, let ρ_1 and ρ_2 be the per capita infection - induced death rates of field and nursery plants respectively.

Let us construct the next generation operator and calculate R_0.

Since the plant can either be infected while standing in the field or standing in the nursery, there will be two states at birth, say $\Omega = \{1, 2\}$, where 1 denotes the plants in the field and 2 the plants in the nursery. Thus, R_0 will be the dominant eigenvalue of 2×2 matrix.
Since the plants don't move in the field and are moved with rate ζ from the nursery to the field we obtain the transition matrix Σ,

$$\Sigma = \begin{pmatrix} 0 & \zeta \\ 0 & -\zeta \end{pmatrix}.$$

Hence

$$\Sigma - D = \begin{pmatrix} -\mu_1 - \rho_1 & \zeta \\ 0 & -\mu_2 - \rho_2 - \zeta \end{pmatrix}$$

and

$$-(\Sigma - D)^{-1} = \begin{pmatrix} \frac{1}{\mu_1+\rho_1} & \frac{\zeta}{(\mu_1+\rho_1)(\mu_2+\rho_2+\zeta)} \\ 0 & \frac{1}{\mu_2+\rho_2+\zeta} \end{pmatrix}.$$

For the infectivity matrix B our assumptions lead to

$$B = \begin{pmatrix} \beta_1 N_1 & 0 \\ \gamma p & \beta_2 N_2 \end{pmatrix},$$

so

$$K = \begin{pmatrix} \frac{\beta_1 N_1}{\mu_1+\rho_1} & \frac{\zeta\beta_1 N_1}{(\mu_1+\rho_1)(\mu_2+\rho_2+\zeta)} \\ \frac{p\gamma}{\mu_1+\rho_1} & \frac{\beta_2 N_2 + p\gamma\zeta}{\mu_2+\rho_2+\zeta} \end{pmatrix}.$$

Hence

$$R_0 = \frac{1}{2}(k_{11} + k_{22}) + \frac{1}{2}\sqrt{k_{11}^2 - 2k_{11}k_{22} + k_{22}^2 + 4k_{12}k_{21}}.$$

All that remains is to insert the right elements of K to the expressions for R_0 but since this adds nothing to clarity, quite the contrary, we avoid it.

Example 4

It has been observed in, for example, Africa that certain genital ulcer diseases, particularly chancroid and syphilis, can increase the risk of HIV infection. The damage that these ulcerative sexually transmitted diseases (shortly USTDs) cause to the genital skin and membranes may facilitate both transmisson and aquisition. One might then study how much do USTDs that are endemic in the population affect the spread of HIV.
We shall thus study two diseases, d and D, having in mind that D is the major disease (incurable) and d the minor one.
First we shall give a simple model for the spread of D in the absence of d and then gradually introduce heterogeneity into the population and study how R_0 for the invasion of D to a population where d is endemic depends on the parameters that govern the spread of d.
So let us first assume that D enters a homogeneous population of D- susceptibles. Let N denote the population size, S the number of D-susceptibles, I the number of D-infected individuals. Moreover, let us assume that the life span of an individual is exponentially distributed with parameter μ and that the natural mortality rate is increased by σ due to disease D. The number of contacts per unit of time is assumed to be constant (denoted by ρ) and each contact between a D-infected and D-susceptible leads to a transmisson with constant probability p.
Then

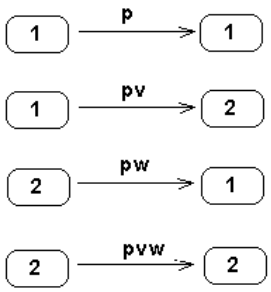
$$\frac{dI}{dt} = \frac{\rho p I S}{N} - (\mu + \sigma)I.$$

Denoting with R_0^D the basic reproduction ratio for the invasion of D we get

$$R_0^D = \frac{\rho p}{\mu + \sigma}.$$

The first step in introducing heterogeneity will be to distinguish individuals according to whether they are infected with d. Thus, we shall denote $\Omega = \{1, 2\}$, where 1 stands for d-free and 2 for d-infected individuals. We shall assume that d doesn't influence the contact rate but infectivity and susceptibility to D change - susceptibility to D is, for individuals having d,

enlarged by a factor v and a D-infectious individual is w times more infectious when also suffering from d. The transmission rates are given in the following figure.



Furthermore, assume that each d-infected individual is cured with rate γ per unit of time and that after recovery the susceptibility to d returns. Denoting with ζ the force of infection and μ the natural death and birth rate (and assuming that all newborns are d- susceptible), we now first calculate the steady state distribution of d-infected and d-free individuals in the absence of D. So let N_1 and N_2 describe numbers of d- free and d- infected individuals respectively. Then

$$\begin{aligned} \frac{dN_1}{dt} &= \mu N - \zeta N_1 - \mu N_1 + \gamma N_2, \\ \frac{dN_2}{dt} &= \zeta N_1 - \mu N_2 - \gamma N_2 \end{aligned}$$

and $N_1 + N_2 = N$. Together with this equality the steady state condition $\frac{dN_1}{dt} = \frac{dN_2}{dt} = 0$ gives us

$$N_1 = \frac{\gamma + \mu}{\gamma + \mu + \zeta} N, \qquad N_2 = \frac{\zeta}{\gamma + \mu + \zeta} N.$$

Now, the matrix B describing contacts and infectivity has the form

$$B = \frac{\rho p}{N} \begin{pmatrix} N_1 & wN_1 \\ vN_2 & vwN_2 \end{pmatrix},$$

since the expected number of contacts per unit time with 1- and 2- individuals equals $\frac{\rho N_1}{N}$ and $\frac{\rho N_2}{N}$ respectively; each of these contacts between d-free individuals is succesful with probability p and the probability is enlarged by a proper factor when any of the involved in a contact is suffering from d.

To obtain the matrix Q we need to restrict ourselves to a D- infected individual and that means leaving out the birth rate and increasing the death

rate to $\mu+\sigma$. We have

$$Q=\begin{pmatrix}-\zeta-\mu-\sigma & \gamma\\ \zeta & -\gamma-\mu-\sigma\end{pmatrix},$$

hence

$$-Q^{-1}=\frac{1}{(\mu+\sigma)(\mu+\sigma+\gamma+\zeta)}\begin{pmatrix}\mu+\sigma+\gamma & \gamma\\ \zeta & \mu+\sigma+\zeta\end{pmatrix}$$

and the next generation operator is given by $K=-BQ^{-1}$. Our assumption that the h-states of infectious and susceptible individuals have independent influence on the transmission of D is now reflected in the fact that the range of B (and thus the range of K) is one dimensional. Indeed, for $\phi=(\phi_1,\phi_2)$ we have

$$B\phi=\frac{\rho p}{N}(\phi_1+w\phi_2)\begin{pmatrix}N_1\\ vN_2\end{pmatrix}$$

and the range of B is spanned by $\phi^d=(N_1,vN_2)^T$. The same vector then spans the range of K and is thus the eigenvector corresponding to R_0, the dominant eigenvector. The birth states, "being d-free" and "being d-infected", are distributed in Ω in the ratio $N_1:vN_2$. Since

$$K\phi^d=-BQ^{-1}\phi^d=\frac{\rho p}{N}((1,w)Q^{-1}\phi^d)\phi^d,$$

it follows that

$$R_0=-\frac{\rho p}{N}(1,w)Q^{-1}\phi^d.$$

Now, we wish to compare R_0 with R_0^D, so we need to write down the explicit expression. Taking into account the steady state numbers N_1 and N_2 we obtain

$$R_0=\frac{\rho p(\gamma+\mu)(\mu+\sigma+\gamma+w\zeta)+vw\zeta(\frac{\gamma}{w}+\mu+\sigma+\zeta)}{(\gamma+\mu+\zeta)(\mu+\sigma)(\mu+\sigma+\gamma+\zeta)}.$$

Defining F by $R_0=R_0^DF$ we have

$$F=\frac{(\gamma+\mu)(\mu+\sigma+\gamma+w\zeta)+vw\zeta(\frac{\gamma}{w}+\mu+\sigma+\zeta)}{(\gamma+\mu+\zeta)(\mu+\sigma+\gamma+\zeta)}.$$

Rewriting F in the form

$$F \;=\; \frac{\gamma+\mu}{\gamma+\mu+\zeta}\frac{\mu+\sigma+\gamma+w\zeta}{\mu+\sigma+\gamma+\zeta}+\frac{\zeta}{\gamma+\mu+\zeta}\frac{v\gamma+vw(\mu+\sigma+\zeta)}{\mu+\sigma+\gamma+\zeta}$$

we observe that F is a convex combination of two numbers that are at least 1 (remember $v, w \geq 1$), so $F \geq 1$.

We can now study how these parameters influence F. We shall only make a couple of observations here. First of all we observe that for $v = w = 1$ we have $F = 1$ (and so $R_0^D = R_0$) which is to be expected - if d has no impact on the D transmission rates it is (from our point of view) as if it was not present.
We also expect that in the case when individuals recover quickly (γ large) or when the force of infection is small, the influence of d on R_0^D is negligible. And indeed,

$$\lim_{\gamma\to\infty} F(\gamma) = 1, \qquad \lim_{\zeta\to 0} F(\zeta) = 1.$$

For a relatively large force of infection hovewer, $F \approx vw$ as can be seen when we rewrite

$$F = \frac{(\mu+\gamma)(\frac{\mu}{\zeta}+\frac{\sigma}{\zeta}+\frac{\gamma}{\zeta}+w)}{\zeta(\frac{\mu}{\zeta}+\frac{\gamma}{\zeta}+1)(\frac{\mu}{\zeta}+\frac{\sigma}{\zeta}+\frac{\gamma}{\zeta}+1)} + \frac{vw(\frac{\gamma}{w\zeta}+\frac{\mu}{\zeta}+\frac{\sigma}{\zeta}+1)}{(\frac{\mu}{\zeta}+\frac{\gamma}{\mu}+1)(\frac{\mu}{\zeta}+\frac{\sigma}{\zeta}+\frac{\gamma}{\zeta}+1)}.$$

This example illustrates nicely how explicit expressions of R_0 are indeed of great importance. In order to understand how F depends on the parameters involved and to use this knowledge in attempts to limit the spread of D we could now continue to scrutinize F. For example, one easily checks that F is, as a function of γ, strictly decreasing (and thus one would want to increase γ with, for instance, better health care) and that it is, as a function of ζ strictly increasing (thereby we wish to lower ζ by, for example, campaigning against unprotected sex). In attempts to lower F, which one of the two options is preferable ? Under which conditions ?

Rather than digging further let us introduce another level of heterogeneity. As the transition probabilities may depend on whether the carrier is female or male we shall now, in addition to the $1, 2$ distinction, identify individuals by sex. So let f denote females, m males. The h-state now takes one of the four possible values: $\Omega = \{(f,1),(f,2),(m,1),(m,2)\}$ and since the part f, m of the state is static (well, for most of us anyway) the dynamics in Ω is described by

$$Q = \begin{pmatrix} Q_f & 0 \\ 0 & Q_m \end{pmatrix}.$$

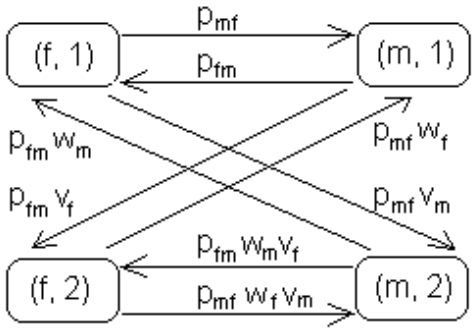
The matrices Q_f, Q_m have exactly the form of the previous Q but the parameters γ and ζ may differ for males and females. Then

$$-Q^{-1} = \begin{pmatrix} -Q_f^{-1} & 0 \\ 0 & -Q_m^{-1} \end{pmatrix}.$$

Concerning the contacts and transmission let us restrict ourselves to the heterosexual subpopulation, denote with ρ_f and ρ_m the contact rates for females and males respectively and let N_f, N_m denote the population sizes. Note that, since we only consider heterosexual contacts, we should have $\rho_f N_f = \rho_m N_m$. Now, dividing the populations of females and males further according to whether they are d-infected and denoting with $N_{f,1}, N_{f,2}, N_{m,1}$ and $N_{m,2}$ the appropriate subpopulation sizes, we arrive at

$$N_{i,1} = \frac{\gamma_i + \mu_i}{\gamma_i + \mu_i + \zeta_i} N_i, \qquad N_{i,2} = \frac{\zeta_i}{\gamma_i + \mu_i + \zeta_i} N_i,$$

where $i \in \{f, m\}$. To avoid further tedious enumerating of parameters let us give a diagram of transition probabilities,



Since we are only interested in heterosexual contacts, the matrix B has the form

$$B = \begin{pmatrix} 0 & B_{fm} \\ B_{mf} & 0 \end{pmatrix},$$

where

$$B_{fm} = \begin{pmatrix} \rho_m \frac{N_{f,1}}{N_f} p_{fm} & \rho_m \frac{N_{f,1}}{N_f} p_{fm} w_m \\ \rho_m \frac{N_{f,2}}{N_f} p_{fm} v_f & \rho_m \frac{N_{f,2}}{N_f} p_{fm} v_f w_m \end{pmatrix}$$

and

$$B_{mf} = \begin{pmatrix} \rho_f \frac{N_{m,1}}{N_m} p_{mf} & \rho_f \frac{N_{m,1}}{N_m} p_{mf} w_f \\ \rho_f \frac{N_{m,2}}{N_m} p_{mf} v_m & \rho_f \frac{N_{m,2}}{N_m} p_{mf} v_m w_f \end{pmatrix}.$$

Now, B_{fm} and B_{mf} have one dimensional range - the range of B_{fm} is spanned by $\psi_1 = (N_{f,1}, v_f N_{f,2})^T$, the range of B_{mf} is spanned by $\psi_2 = (N_{m,1}, v_m N_{m,2})^T$. Hence, taking into account the structure of B, the range

of B is spanned by the vectors

$$\phi_1 = \begin{pmatrix} N_{f,1} \\ v_f N_{f,2} \\ 0 \\ 0 \end{pmatrix}, \quad \phi_2 = \begin{pmatrix} 0 \\ 0 \\ N_{m,1} \\ v_m N_{m,2} \end{pmatrix},$$

which also span the range of K. The dominant eigenvector has the form $\phi^d = (c_1\phi_1, c_2\phi_2)^T$ and since

$$K\phi^d = \begin{pmatrix} -\frac{\rho_f}{N_m} p_{mf}((1, w_m)Q_m^{-1}\psi_2)\phi_1 \\ -\frac{\rho_m}{N_f} p_{fm}((1, w_f)Q_f^{-1}\psi_1)\phi_2 \end{pmatrix},$$

it follows that

$$K\phi = \begin{pmatrix} 0 & l_{fm} \\ l_{mf} & 0 \end{pmatrix} \phi,$$

where

$$l_{fm} = -\frac{\rho_f}{N_m} p_{mf} \begin{pmatrix} 1 \\ w_m \end{pmatrix} \cdot Q_m^{-1}\psi_2, \quad l_{mf} = -\frac{\rho_m}{N_f} p_{fm} \begin{pmatrix} 1 \\ w_f \end{pmatrix} \cdot Q_f^{-1}\psi_1.$$

Thus,

$$R_0 = \sqrt{l_{fm} l_{mf}}.$$

Another extension of the first model is to distinguish individuals according to their sexual activity. We shall let it take values in $\mathbb{N}$ and assume that it is a static trait. The h-state space thus consists of elements

$$\Omega = \{(i, j); i \in \mathbb{N}, j \in \{1, 2\}\}.$$

We shall assume that the force of infection is proportional to sexual activity, hence

$$\frac{dN_{i,2}}{dt} = i\zeta N_{i,1} - \gamma N_{i,2} - \mu N_{i,2}.$$

Denoting with N_i the numbers of individuals with sexual activity i we again have $N_i = N_{i,1} + N_{i,2}$ and

$$N_{i,1} = \frac{\gamma + \mu}{\gamma + \mu + i\zeta} N_i, \quad N_{i,2} = \frac{i\zeta}{\gamma + \mu + i\zeta} N_i.$$

The h-state dynamics is now described with the infinite matrix

$$Q = \begin{pmatrix} Q_1 & 0 & 0 & ... \\ 0 & Q_2 & 0 & ... \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix},$$

where the diagonal elements are of the form

$$Q_i = \begin{pmatrix} -i\zeta - \gamma - \sigma & \gamma \\ i\zeta & -\gamma - \mu - \sigma \end{pmatrix}.$$

Thus

$$Q^{-1} = \begin{pmatrix} Q_1^{-1} & 0 & 0 & ... \\ 0 & Q_2^{-1} & 0 & ... \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix},$$

If ρ denotes the contact rate for individuals with activity one, then the expected number of contacts of one $(j,\cdot)$ individual with $(i,\cdot)$ individuals equals $\frac{\rho i j N_i}{\sum_{k=1}^{\infty} k N_k}$. Hence, the matrix B has the form

$$B = \begin{pmatrix} B_{11} & B_{12} & ... \\ B_{21} & B_{22} & ... \\ \vdots & \vdots & \ddots \end{pmatrix},$$

where B_{ij} stands for

$$B_{ij} = \frac{\rho p i j}{\sum_{k=1}^{\infty} k N_k} \cdot \begin{pmatrix} N_{i,1} & w N_{i,1} \\ v N_{i,2} & v w N_{i,2} \end{pmatrix},$$

thus

$$K_{ij} = -B_{ij} Q_j^{-1}. \tag{3.1}$$

As a matter of fact, writing down B and multiplying it with Q^{-1} could be avoided by employing biological arguments instead. Indeed, if we take into account that the fraction $N_{i,2}/N_i$ of contacts with i individuals are with d-infected individuals, etc. and use proper enhanced transition probabilities, we arrive at (3.1).
Now, let us write $\phi = (\phi_1, \phi_2, ...)^T$, where

$$\phi_i = \begin{pmatrix} \phi_{i,1} \\ \phi_{i,2} \end{pmatrix}.$$

From the expression for K_{ij} we observe that the range of K has the i-th component spanned by $(N_{i,1}, v N_{i,2})$. The same goes for the i-th component of the dominant eigenvector ϕ^d and the corresponding coeficients are obtained by the inner product of $(1,w)^T$ with $\frac{\rho p}{\sum_{k=1}^{\infty} k N_k} \sum_{j=1}^{\infty} j Q_j^{-1} \phi_j^d$.

Since the coeficients of the dominant eigenvector represent the probability distribution for birth states we arrive at

$$R_0 = -\frac{\rho p}{\sum_{k=1}^{\infty} kN_k} \begin{pmatrix} 1 \\ w \end{pmatrix} \cdot \sum_{j=1}^{\infty} j^2 Q_j^{-1} \begin{pmatrix} N_{j,1} \\ vN_{j,2} \end{pmatrix}.$$

Now, combining both extensions of the initial model, that is, apart from $1, 2$ distinction, identifying individuals by both, sexual activity and gender, we have

$$\Omega = \{(i, j, k); i \in \mathbb{N}, j \in \{f, m\}, k \in \{1, 2\}\}$$

and again we obtain

$$Q = \begin{pmatrix} Q_1 & 0 & 0 & ... \\ 0 & Q_2 & 0 & ... \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix},$$

only now

$$Q_i = \begin{pmatrix} Q_{i,f} & 0 \\ 0 & Q_{i,m} \end{pmatrix},$$

where matrices $Q_{i,f}, Q_{i,m}$ have the same form as Q_i before with the exception that all the parameters now get index f or m.
Similarly, the matrices B_{ij} have the form

$$B_{ij} = \begin{pmatrix} 0 & B_{ij}^{fm} \\ B_{ij}^{mf} & 0 \end{pmatrix},$$

where B_{ij}^{fm} and B_{ij}^{mf} have the form of previous B_{ij} but with appropriate indexing of parameters.
Now, denote with $\phi = (\phi_1, \phi_2, ...)$ where ϕ_i are now four-vectors with

$$\phi_i = \begin{pmatrix} \phi_{i,f,1} \\ \phi_{i,f,2} \\ \phi_{i,m,1} \\ \phi_{i,m,2} \end{pmatrix}.$$

The i-th component of the range of K is now spanned by two vectors

$$\chi_i^1 = \begin{pmatrix} N_{i,f,1} \\ v_f N_{i,f,2} \\ 0 \\ 0 \end{pmatrix} \qquad \chi_i^2 = \begin{pmatrix} 0 \\ 0 \\ N_{i,m,1} \\ v_m N_{i,m,2} \end{pmatrix},$$

and the coeficients are obtained as

$$l_i^{fm} = -\sum_{j=1}^{\infty} \begin{pmatrix} 1 \\ w_m \end{pmatrix} Q_{j,m}^{-1} \psi_j^2, \quad l_i^{mf} = -\sum_{j=1}^{\infty} \begin{pmatrix} 1 \\ w_f \end{pmatrix} Q_{j,f}^{-1} \psi_j^1,$$

where

$$\psi_j^1 = \frac{\rho_m j}{\sum_{k=1}^{\infty} k N_k} p_{fm} (N_{j,f,1}, v_f N_{j,f,2})^T$$

and

$$\psi_j^2 = \frac{\rho_f j}{\sum_{k=1}^{\infty} k N_k} p_{mf} (N_{j,m,1}, v_m N_{j,m,2})^T.$$

Hence

$$R_0 = \sum_{j=1}^{\infty} j \sqrt{l_j^{fm} l_j^{mf}}.$$

3.4 Pair formation models

When constructing the next generation operator, we implicitly assumed that every contact an infected individual has, is one with a susceptible. This assumption seems to be far-fetched in certain situations, certainly when one considers sexually transmitted diseases, but this is not the only case - for example think of measles, influenza and schools, offices,... When modelling sexually transmitted diseases one has to, in order to make models more "realistic", take into account that individuals form partnerships for non-negligible periods of time. In that time they only have contacts with each other, assuming, of course, ideally faithful indviduals. Thus, from the point of view of the infective agent, all the contacts in the partnership are wasted once the infection is transmitted.

It was also assumed that the infectious output of an individual can be described with the age representation. However, in the pair formation case, there are two more parameters that come into play - the (partner's) survival and the rate of separation. Only when partners separate or one of the partners dies, becomes one available for other contacts and in this case the age representation is not always possible.

Let us see how these new aspects influence the basic reproduction ratio. Since the mentioned assumptions were crucial when constructing K, incorporating pairs into the model does not allow direct generalization of the next generation operator.

However, we shall present one situation that uses the ideas of the previous sections and allows the construction of the next generation operator.

We shall assume that there is a finite number of *disease states*, say $1, 2, ..., n$,. A newly infected individual starts its infected life in state one and from there on, disease states are passed through sequentially. Each disease state is characterised by its own infectivity and the rate of dying.
Concerning pairs we shall distinguish two kinds of pairs: the ones that are in the courtship period and those that are in the sexually active period. The courtship period is characterised by the absence of sexual contatcs. Thus, in that period, individuals are neither at risk of receiving nor of transmitting the infection. We shall assume that every pair starts with a courtship phase. As the probabilities of transmission may depend on the disease state of the infected partner we define the *partnership status* of an individual that takes values in $\{-1, 0, 1, 2, ..., n\}$. The label -1 means that the individual is single, 0 means that the individual is in a partnership with a susceptible. The status $j \in \{1, ..., n\}$ denotes that the individual's partner is in disease state j. The disease state and the partnership status together describe the *type* (i, j) of an infected individual. We shall indicate the type of individuals that are in the courtship period with $(i, j)_c$. Since we are only interested in R_0, we shall assume that every new partner of an infected individual is a susceptible and since each pair starts with the courtship phase, the set

$$\Lambda = \{(i, j); 1 \leq i \leq n, -1 \leq j \leq n\} \cup \{(i, 0)_c; 1 \leq i \leq n\}$$

describes all possible types.
In the construction of the next generation operator we shall need some sort of ordering on Λ. Since $|\Lambda| = n(n + 3)$, denote $\Sigma = \{1, 2, ..., n(n + 3)\}$ and let $L : \Lambda \to \Sigma$,

$$L(i, j) < L(i', j') \iff (i < i') \vee (i = i', j < j'),$$

$$L(i, -1) < L(i, 0)_c < L(i, 0)$$

describe the lexicographic ordering on Λ.
Before embarking on the construction of the next generation operator, we make the following assumptions:

- all individuals are equally susceptible,
- denote with μ_0 the death rate of susceptibles and with μ_j the death rate of individuals with state j
- the time spent in disease state j is, given that the individual survives, exponentially distributed with parameter θ_j. The disease state n is retained until death,

- the infectivity in state j is described by probability p_j, that is, each sexual contact with a susceptible leads to a transmition with probability p_j.

- every single individual has a constant probability ρ per unit of time to become a member of a courtship pair. The divorce rate is σ_c in the courtship phase and σ in the sexually active phase,

- a pair always starts with a courtship phase. The length of this phase is, conditional on the survival and no divorce of the partners, exponentially distributed with parameter α. By definition, the sexually active phase starts with one sexual contact,

- during the sexually active phase the partners have β contacts per unit of time.

The following diagram describes the possible changes of an individual of fixed disease state i,as long as it does not die,

Now, to obtain the next generation operator we need to specify the number of new infections with type $(1, i)$ caused by an individual that just became type $(1, j)$, during its entire remaining life. We shall denote that number with k_{ij}. Since the only two types that can cause new infections are types $(i, 0)_c$, these are first contacts, and $(i, 0)$, the next generation operator is a matrix $K \in \mathbb{R}^{n,n}$ with elements

$$k_{ij} = \textit{the sum of the expected number of type transitions } (i,0)_c \to (i,1) \textit{ and } (i,0) \to (i,1) \textit{ caused by an individual that just became type } (j,1), \textit{ during in its entire remaining life}$$

and R_0 is the dominant eigenvalue of K.

We shall consider the changes in types as a Markov process on state space Σ. Let $P(t)$ contain the following elements - let $P_{ij}(t)$ equal the probability of being alive and in state i at time t after starting in state j in time 0. We have

$$P(t) = e^{Qt},$$

where for $i \neq j$ the element Q_{ij} gives us the rate of leaving j to go to state i and $Q_{ii} = -\sum_{i \neq j} Q_{ji}$ equals the rate of dying of individuals in state i. In order that a newly infected individual with type $(1, j)$ causes an infection

of type $(i,1)$, the individual has to be alive at some time t after the infection and its state has to be either $(i,0)_c$ or $(i,0)$. Thus, for $1 \leq i,j \leq n$ we have

$$k_{ij} = p_i\alpha \int_0^\infty P_{L(i,0)_c L(1,j)}(t)dt + p_i\beta \int_0^\infty P_{L(i,0)L(1,j)}(t)dt,$$

or

$$k_{ij} = -p_i\alpha(Q^{-1})_{L(i,0)_c L(1,j)} - p_i\beta(Q^{-1})_{L(i,0)L(1,j)}.$$

So, all we need to calculate are the right elements of Q^{-1}. In fact, as we assumed that the disease states are passed through one by one, the matrix Q has the following structure

$$Q = \begin{pmatrix} A_1 & 0 & 0 & ... \\ D_1 & A_2 & 0 & ... \\ 0 & D_2 & A_3 & ... \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix},$$

where $D_i = diag(\theta_i) = \theta_i I_n$. The matrices A_i have the form

$$A_i = \begin{pmatrix} a_1(i) & \sigma_c + \mu_0 & \sigma + \mu_0 & \sigma + \mu_1 & \sigma + \mu_2 & \sigma + \mu_3 & ... & \sigma + \mu_n \\ \rho & a_2(i) & 0 & 0 & 0 & 0 & ... & 0 \\ 0 & (1-p_i)\alpha & a_3(i) & 0 & 0 & 0 & ... & 0 \\ 0 & p_i\alpha & p_i\beta & a_4(i) & 0 & 0 & ... & 0 \\ 0 & 0 & 0 & \theta_1 & a_5(i) & 0 & ... & 0 \\ 0 & 0 & 0 & 0 & \theta_2 & a_6(i) & ... & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & ... & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & ... & a_n(i) \end{pmatrix},$$

where

$$\begin{aligned} a_1(i) &= -\rho - \mu_i - \theta_i, \\ a_2(i) &= -\sigma_c - \mu_0 - \alpha - \mu_i - \theta_i, \\ a_3(i) &= -p_i\beta - \sigma - \mu_0 - \mu_i - \theta_i, \end{aligned}$$

and for $3 < m < n$

$$a_m(i) = -\sigma - \mu_i - \theta_i - \mu_{m-3} - \theta_{m-3}.$$

One easily checks that

$$Q^{-1} = \begin{pmatrix} A_1^{-1} & 0 & 0 & ... \\ B_{21} & A_2^{-1} & 0 & ... \\ B_{31} & B_{32} & A_3^{-1} & ... \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix},$$

where

$$B_{ij} = (-1)^{i+j} D_j D_{j+1} \cdot ... \cdot D_{i-1} A_i^{-1} A_{i-1}^{-1} \cdot ... \cdot A_j^{-1}.$$

Hence, we only need to calculate the inverse of every A_i.

Example 5

Let us look more closely at the situation where there is just one disease state. Then

$$Q = \begin{pmatrix} -\mu_1 - \rho & \mu_0 + \sigma_c & \mu_0 + \sigma & \mu_1 + \sigma \\ \rho & -\mu_0 - \mu_1 - \sigma_c - \alpha & 0 & 0 \\ 0 & (1-p)\alpha & -\mu_0 - \mu_1 - \sigma - p\beta & 0 \\ 0 & p\alpha & p\beta & -2\mu_1 - \sigma \end{pmatrix}$$

and R_0 is given by

$$p\alpha \int_0^\infty P_{L(1,0)_c L(1,1)}(t)dt + p\beta \int_0^\infty P_{L(1,0)L(1,1)}(t)dt = -p\alpha (Q^{-1})_{2,4} - p\beta (Q^{-1})_{3,4}.$$

After a bit of algebra we find

$$R_0 = \frac{p\alpha\beta(\mu_1+\sigma)(1-p)\rho + p\rho\alpha(\mu_1+\mu_0+\rho+p\beta)(\mu_1+\sigma)}{\mu_1(x+y)}, \qquad (3.2)$$

where

$$\begin{aligned} x &= (\mu_0+\mu_1+\sigma+\rho\beta)(2\mu_1+\sigma)(\mu_0+\mu_1+\rho+\sigma_c+\alpha), \\ y &= \rho\alpha(p\mu_0+(2-p)\mu_1+\sigma+p\beta). \end{aligned}$$

Taking various limits of (3.2) we can get some additional results that also give us the interpretation.

1. Let us first "collapse" the sexually active period to a point event, write $\beta = k\sigma + O(1)$ and let $\sigma \to \infty$. Then $1+k$ is the average number of contacts in one partnership and (3.2) simplifies to

$$R_0 = \frac{\rho\alpha p(1+k)}{\mu_1(1+pk)(\mu_0+\mu_1+\rho+\sigma_c+\alpha)} \qquad (3.3)$$

and this result has the following interpretation:

$$\frac{\rho}{\mu_1(\mu_0+\mu_1+\rho+\sigma_c+\alpha)}$$

is the expected time that an infected individual will have a susceptible partner, α is the rate of becoming sexually active and the "succes"-ratio per one partner is $p(1+k)/(1+pk)$.

2. In case when there are no pairs we would have $R_0 = p\rho/\mu_1$. We expect, when eliminating pair formation, the same result from (3.3). Indeed, let $\alpha \to \infty$ in (3.3) to obtain

$$R_0 = \frac{\rho p(1+k)}{\mu_1(1+pk)} \tag{3.4}$$

and put $k = 0$ (there are only first contacts) in (3.4).

3. If we let $\rho \to \infty$ we assume that there are no single individuals. Then

$$R_0 = \frac{\alpha p(1+k)}{\mu_1(1+pk)},$$

which has almost the same interpretation as the first case, with the only difference that now the newly infected individual is already in the courtship phase.

Introducing heterogenity

In the context of sexually transmitted diseases, the most important h-states that are incorporated into a model are gender, sexual activity, age and sexual orientation. We need to specifty both the h-state of the individual itself and that of its current partner, since the h-state of the partner can influence the death rate and in this way the rate with which the original individual becomes single. Let Ω denote the h-state space. The type of a sexually active individual is now described by

$$(i, j; \xi_i, \xi_j); \quad i \in \{1, ..., n\}, \ j \in \{0, 1, ..., n\}, \ \xi_i, \xi_j \in \Omega.$$

Furthermore, let $(i, -1, \xi_i)$ denote the type of an infected single individual and $(i, 0, \xi_i, \xi_0)$ the type of an infected individual in the courtship phase of a relationship.

Suppose we have an infected individual, say x, that was infected by an individual with disease state j and h-state ν. Suppose that x itself had h-state η at the moment of infection. Then x was born with type $(1, j; \eta, \nu)$. Now, during its infectious life, x might separate from its current partner and pair with a susceptible with, say, h-state ξ. In that time, x's disease state might have changed to, say θ. As in the case without heterogeneity, we need to specify the number of x's partner infections in it's entire remaining life, that is, we need to calculate the number of type transitions

$$(i, 0; \theta, \xi)_c \to (i, 1; \theta, \xi),$$

and

$$(i, 0; \theta, \xi) \to (i, 1; \theta, \xi).$$

We denote this number with $k_{ij}(\xi, \theta; \eta, \nu)$.
Let the current infectivity of x towards its partner be the probability $p_i(\xi, \theta)$, denote the rate of entering the sexually active phase by $\alpha(\xi, \theta)$ and let $\beta(\xi, \theta)$ describe the number of contacts per unit of time. Furthermore, for $i, j \in \Sigma$ let $P_{ij}(t, \xi, \theta; \eta, \nu)$ describe the probability that at time t after the individual was infected as type $(1, j; \eta, \nu)$, the individual is still alive and currently has type $(i, 0; \theta, \xi)$.
Then the biological interpretation again leads us to the expression

$$\begin{aligned} k_{ij}(\xi, \theta; \eta, \nu) &= p_i(\xi, \theta)\alpha(\xi, \theta) \int_0^\infty P_{L(i,0)_c L_(1,j)}(t, \xi, \theta; \eta, \nu)dt \\ &+ p_i(\xi, \theta)\beta(\xi, \theta) \int_0^\infty P_{L(i,0)L_(1,j)}(t, \xi, \theta; \eta, \nu)dt. \end{aligned}$$

Let $\phi = \phi(j; \eta, \nu)$ describe the distribution of newly infected individuals over the space $\{1, ..., n\} \times \Omega \times \Omega$. The next generation of infected individuals is now obtained by applying the next generation operator,

$$(K\phi)(i; \xi, \theta) = \sum_{j=1}^{n} \int_{\Omega\times\Omega} k_{ij}(\xi, \theta; \eta, \nu)\phi(j; \eta, \nu)d\eta d\nu$$

and R_0 is the spectral radius of the next generation operator.

Now, as mentioned before, the calculation of R_0 simplifies under additional assumptions. If we assume separable mixing,

$$k_{ij}(\xi, \theta; \eta, \nu) = a_i(\xi, \theta)b_j(\eta, \nu),$$

then

$$R_0 = \sum_{j=1}^{n} \int_{\Omega\times\Omega} a_j(\eta, \nu)b_j(\eta, \nu)d\eta d\nu.$$

Less restrictive is the assumption of multigroup separable mixing,

$$k_{ij}(\xi, \theta; \eta, \nu) = a_i(\xi, \theta)b_{ij}(\eta, \nu).$$

In this case, the eigenvalues of K are those of an $n \times n$ matrix $L = (l_{ij})_{i,j=1}^n$ with

$$l_{ij} = \int_{\Omega\times\Omega} a_j(\eta, \nu)b_{ij}(\eta, \nu)d\eta d\nu$$

and R_0 is the dominant eigenvalue of L.

Let us now give one example of incorporating heterogeneity.

Example 6

Let us distinguish individuals according to sex, so $\Omega = \{f, m\}$. If we take only heterosexual contacts into account, the next generation operator K is a $2n \times 2n$ matrix of the form

$$K = \begin{pmatrix} 0 & K_f \\ K_m & 0 \end{pmatrix},$$

where $K_f = (k_{ij}(f, m; m, f))_{i,j=1}^n$ and $K_m = (k_{ij}(m, f; f, m))_{i,j=1}^n$. As we only look at the heterosexual contacts, the matrices K_f and K_m can be obtained the same way as K before, with the only difference that (some) of the parameters may differ for males and females.
Since

$$K^2 = \begin{pmatrix} K_f K_m & 0 \\ 0 & K_m K_f \end{pmatrix}$$

and since for arbitrary matrices A, B,

$$r(A^2) = r(A)^2, \quad r(AB) = r(BA),$$

it follows that

$$R_0 = \sqrt{r(K_f K_m)}.$$

Let us assume, just like in Example 5, that there is only one disease state. Denote with $\mu_0, \mu_1, p, \beta, \sigma, \rho, \alpha$ the parameters for males and $\mu_0', \mu_1', p', \beta, \sigma', \rho', \alpha$ the parameters for females.
The transition matrix Q is given by

$$Q = \begin{pmatrix} Q_{fm} & 0 \\ 0 & Q_{mf} \end{pmatrix},$$

where Q_{fm}, Q_{mf} have the same form as the matrix Q in Example 5, but with appropriate setting of parameters. Thus

$$\begin{aligned} K_f &= k_{11}(f, m; m, f) = -p\alpha(Q_{mf}^{-1})_{2,4} - p\beta(Q_{mf}^{-1})_{3,4}, \\ K_m &= k_{11}(m, f; f, m) = -p'\alpha(Q_{fm}^{-1})_{2,4} - p'\beta(Q_{fm}^{-1})_{3,4}. \end{aligned}$$

and

$$R_0 = \sqrt{K_f K_m}.$$

Chapter 4

Incorporating the i-state dynamics and time structure

The purpose of this chapter is twofold. First we fulfill our promise made in the beginning of Chapter 2, that is, we find a way to describe the i-state dynamics in a general setting. We give a more a detailed description of the transmission, including information about the period between becoming inftected and infecting others. We do this to prepare the ground for the next Chapter where we shall study an epidemic outbreak in real time.

If the individual's i-state is dynamic, the specification of the next generation operator also includes the description of this dynamics. So far we have given the description of the state dynamics in the form of the *infinitesimal generator*. The main assumption that guaranteed its existence was that we have continuity in the i-state movement. More precisely, we assumed

$$\lim_{t \to 0} P(t, \eta, \xi) = \delta_\xi(\eta).$$

If the i-state of some individual moves in a discontinuos manner, the infinitesimal generator will not even exist and thus the i-state dynamics has to be described in some other way.

First we shall study the behaviour (meaning transmission process, survival and i-state dynamics) at the individual level (i-level). The distribution of i-states in Ω will be described with a (positive) measure on Ω. The next step will then be to add the individual contributions to describe the behaviour at the population level (p-level).
As in the previous chapter we shall only deal with the expected behaviour and assume that the environment is constant.

So let us, for $\eta \in \Omega$ and $\omega \subseteq \Omega$, define

$$\Lambda_t(\eta, \omega) = \textit{the expected number of new infections with i-state in } \omega, \textit{ produced in the time interval of length } t \textit{ by one individual with birth state } \eta.$$

The kernel Λ already introduced on page 20 is nothing but Λ_∞ in our present notation.

The *transmission kernel* Λ describes the first generation of infected individuals. To describe next generations we define

$$\Lambda_t^k(\eta, \omega) = \textit{the expected number of } k\textit{-th generation infections with i-state in } \omega, \textit{ produced in the time interval of length } t \textit{ by one individual with birth state } \eta.$$

The biological interpretation suggests the construction of the next generation kernels by induction as

$$\Lambda_t^{k+1}(\eta, \omega) = \int_{[0,t]\times\Omega} \Lambda_{t-s}(\xi, \omega)\Lambda_{ds}^k(\eta, d\xi), \tag{4.1}$$

so let us give the mathematical background that justifies it.

4.1 Kernels and the ⊗ - product

Let $M_+(\Omega)$ denote the set of non-negative measures on Ω, that is, measures on Ω with values in $\overline{\mathbb{R}}^+ = [0, \infty]$.

Definition 4 *A function $\Lambda : \mathbb{R}^+ \times \Omega \to M_+(\Omega)$ is called a* **kernel** *if for every measurable $\omega \subseteq \Omega$ the function*

$$(t, \eta) \mapsto \Lambda_t(\eta, \omega)$$

is measurable.

We introduce the following order relation on $M_+(\Omega)$ - for $m_1, m_2 \in M_+(\Omega)$ let,

$$m_1 \leq m_2 \overset{def}{\iff} m_1(\omega) \leq m_2(\omega) \text{ for every measurable } \omega \subseteq \Omega.$$

Thus, every monotonically increasing sequence of elements in $M_+(\Omega)$ has a limit in $M_+(\Omega)$.

We lift the order relation to the set of kernels by requiring that the inequality holds for every $(t, \eta) \in \mathbb{R}^+ \times \Omega$.
Since the pointwise limit of measurable functions is a measurable function, the set of kernels also inherits monotone sequential completeness. In particular, every series, whose terms are kernels, has a well defined sum, which is also a kernel.

We also introduce an algebraic structure on the set of kernels: since for every kernel Ψ and for every measurable set $\omega \subseteq \Omega$ the function $(t, s, \eta) \mapsto \Psi_{t-s}(\eta, \omega)$ is measurable,

$$(\Phi \otimes \Psi)_t(\eta, \omega) := \int_{[0,t]\times\Omega} \Phi_{t-s}(\xi, \omega)\Psi_{ds}(\eta, d\xi) \tag{4.2}$$

gives us a well defined product of kernels Φ and Ψ.
Let us check that $\Phi \otimes \Psi$ defined this way is a kernel.
Obviously, $\Phi \otimes \Psi$ is measure valued. We need to see that the function

$$(t, \eta) \mapsto \int_{[0,t]\times\Omega} \Phi_{t-s}(\xi, \omega)\Psi_{ds}(\eta, d\xi)$$

is measurable. The function $(t, s, \eta) \mapsto \Psi_{t-s}(\eta, \omega)$ is measurable and non-negative so it can be approximated with functions of the form $\sum_{i=1}^{n} \alpha_i \chi_{\mathcal{A}_i}$, where $\mathcal{A}_i$ are measurable subsets of $\mathbb{R}^+ \times \mathbb{R}^+ \times \Omega$. Moreover, it is sufficient to only look at the sets of the form $A \times B \times C$, where A and B are measurable subsets of $\mathbb{R}^+$ (equipped with the Borel σ-algebra) and C is a measurable subset of Ω. But since Ψ is a kernel the function

$$(t, \eta) \mapsto \int_{[0,t]\times\Omega} \chi_{A\times B\times C}\Psi_{ds}(\eta, d\xi) = \chi_A(t) \int_B \Psi_{ds}(\eta, C)$$

is measurable and so $\Phi \otimes \Psi$ is indeed a kernel.

Obviously, the product $\otimes$ is

1. distributive, that is, for kernels Θ, Φ, Ψ,

$$\Theta \otimes (\Phi + \Psi) = \Theta \otimes \Phi + \Theta \otimes \Psi,$$

$$(\Phi + \Psi) \otimes \Theta = \Phi \otimes \Theta + \Psi \otimes \Theta$$

   and

2. associative, that is

$$\Theta \otimes (\Phi \otimes \Psi) = (\Theta \otimes \Phi) \otimes \Psi$$

Also, let Ψ_∞ denote the pointwise limit of a non-decreasing sequence of kernels Ψ_n - we shall denote it as $\Psi_n \uparrow \Psi_\infty$. Then

$$\Psi_n \otimes \Phi \uparrow \Psi_\infty \otimes \Phi,$$

and

$$\Phi \otimes \Psi_n \uparrow \Phi \otimes \Psi_\infty.$$

Thus, the k-th power Φ^k of a kernel Φ is well defined and by monotone convergence the sum of these powers exists. Let us denote

$$\Psi^c = \sum_{k=1}^{\infty} \Psi^k$$

and put our findings in the next

Theorem 1 *Let Ψ be a kernel. The series*

$$\sum_{k=1}^{\infty} \Psi^k$$

converges in $\overline{\mathbb{R}}^+ = [0, \infty]$ for a given argument (t, η, ω). We denote the sum with Ψ^c and call it **the resolvent** *of Ψ with respect to $\otimes$. The resolvent is a kernel that satisfies the so called* **resolvent equation**

$$\Psi^c = \Psi + \Psi^c \otimes \Psi = \Psi + \Psi \otimes \Psi^c.$$

Another observation will show itself useful in the following sections: let $\Omega_1, \Omega_2, \Omega_3$ be measurable subsets of Ω. Furthermore let $\Psi : \mathbb{R}^+ \times \Omega_1 \to M_+(\Omega_2)$ and $\Phi : \mathbb{R}^+ \times \Omega_2 \to M_+(\Omega_3)$ be kernels, where we modify Definition 4 in an obvious way to know precisely what a kernel means. Then $\Phi \otimes \Psi : \mathbb{R}^+ \times \Omega_1 \to M_+(\Omega_3)$ defined as

$$(\Phi \otimes \Psi)_t(\eta, \omega) := \int_{[0,t]\times\Omega_2} \Phi_{t-s}(\xi, \omega)\Psi_{ds}(\eta, d\xi) \tag{4.3}$$

is a kernel.
We shall often specify the domains of the kernels writing such formulas as

$$(\Phi \otimes \Psi)|_{\Omega_3,\Omega_1} = \Phi|_{\Omega_3,\Omega_2} \otimes \Psi|_{\Omega_2,\Omega_1}.$$

4.2 The clan kernel and the renewal equation

We have now seen that (4.1) indeed gives us a well defined $(k+1)$-th generation of infections.
Also, the sum of all generation kernels

$$\Lambda^c = \sum_{k=1}^{\infty} \Lambda^k$$

exists and is a kernel. We shall call it the *clan kernel*. Moreover, the resolvent equations given in the previous section actually have a simple biological background. As new infections originating from one individual are either infections produced by this individual or infections caused by one of the secondary infections we arrive at

$$\Lambda^c = \Lambda + \Lambda^c \otimes \Lambda. \tag{4.4}$$

But, also, the clan infections are either the first generation infections or infections caused by clan members. So

$$\Lambda^c = \Lambda + \Lambda \otimes \Lambda^c. \tag{4.5}$$

This interpretation also reflects in the name *renewal equation* that is often used to denote equations (4.4) and (4.5).

Sometimes the biological background allows us to make certain reductions of the generation kernels and the clan kernel.
For example, if we are are looking at the size of the clan in some finite time interval of length t and there is a latency period, that is, there is a time gap between the time the individual itself becomes infected and the time it is infectious, then there will only be a finite number of nonzero terms in the Λ_t^c.
It is also possible that conceivable birth states take values in some set Ω_b that might be considerably smaller than Ω. When calculating next generations we can then limit ourselves to Ω_b. Also, when there is a latency period a further reduction is possible. In the following section we shall look closer at such a situation.

4.3 Reduction of the transmission kernel

Definition 5 *A measurable subset $\Omega_b \subset \Omega$ is a* **set representing the birth states** *if for all $t \in \mathbb{R}^+$ and for all $\eta \in \Omega$,*

$$\Lambda_t(\eta, \omega) = 0 \quad \textit{if } w \cap \Omega_b = \emptyset$$

In general, Ω_b is not unique. Of course, one would want Ω_b as small as possible, although that does not necessarily mean the easiest Ω_b to work with.

The *reduced transmission kernel* will be denoted with ${}_b\Lambda$ and is just Λ restricted to Ω_b, that is, ${}_b\Lambda_t(\eta, \omega) := \Lambda_t(\eta, \omega)$, the domain of η is Ω_b and the domain of ω are measurable subsets of Ω_b.
We shall write this shortly as

$${}_b\Lambda = \Lambda|_{\Omega_b, \Omega_b}. \tag{4.6}$$

We construct the next generation kernels by

$${}_b\Lambda_t^{k+1}(\eta, \omega) = \int_{[0,t]\times\Omega_b} {}_b\Lambda_{t-s}(\xi, \omega)_b\Lambda_{ds}^k(\eta, d\xi).$$

Now, to compute the first generation we need Λ, to compute the next generations we can restrict to ${}_b\Lambda$, so actually

$$\Lambda_t^{k+1}(\eta, \omega) = \int_{[0,t]\times\Omega_b} {}_b\Lambda_{t-s}^k(\xi, \omega)\Lambda_{ds}(\eta, d\xi). \tag{4.7}$$

In a situation when an infected individual is not infectious right away, birth states are separated from the states in which transmission is possible. We call the states in which an individual is infectious the *renewal states* and denote them with Ω_r. The point is that from the "counting of new infections" point of view the latency period contributes nothing and we consider a first entrance to the *renewal set* Ω_r as a kind of quasi birth.
To describe the situation mathematically we need to know the process of entering the renewal set.
Thus, we introduce a measure π on $\mathbb{R}^+ \times \Omega_r$ that describes when and where the individual with birth state $\eta \in \Omega_b$ enters Ω_r for the first time.
We shall assume that for every t and every measurable $\omega \subseteq \Omega_r$ the function

$$\begin{aligned} \Omega_b &\to \mathbb{R} \\ \eta &\mapsto \pi(\eta)(t, \omega) \end{aligned}$$

is measurable.
For every $t \in \mathbb{R}^+$, every $\eta \in \Omega_b$ and every measurable $\omega \subseteq \Omega_b$ the expected transmission, given the birth state, can be expressed in terms of the arrival to the renewal states and the transmission from the renewal states, i.e. we now require that

$${}_b\Lambda_t(\eta, \omega) = \int_{[0,t]\times\Omega_r} \Lambda_{t-s}(\xi, \omega)\pi(\eta)(ds, d\xi). \tag{4.8}$$

The transmission process actually begins when an individual first enters a renewal set. Also, the first generation of infection only becomes "active" when the newly infected enter Ω_r. Thus, we would like to give the expected transmission from those quasi - birth states, the renewal states.
So, for $\eta \in \Omega_r$ and a measurable $\omega \subseteq \Omega_r$ we define the *reduced renewal kernel* ${}_r\Lambda$ as

$$ {}_r\Lambda_t(\eta, \omega) := \int_{[0,t]\times\Omega_b} \pi(\xi)(t-s, \omega)\Lambda_{ds}(\eta, d\xi). \tag{4.9}$$

The next generations of infected individuals counted at the time of first hitting the renewal set can now be calculated with the powers of ${}_r\Lambda$ defined the same way we derived the powers of ${}_b\Lambda$.
Now, we would like to express the next generation kernels and the clan kernel in terms of powers of ${}_b\Lambda$ and ${}_r\Lambda$.
As the k-th generation of infections results from the infections caused when the $(k-1)$-th generation resides in the renewal set we obtain the following

Lemma 1 *For every $t \in \mathbb{R}^+$, $\eta \in \Omega_r$ and a measurable $\omega \subseteq \Omega_b$ we have*

$$ \Lambda^k_t(\eta, \omega) = \int_{[0,t]\times\Omega_r} \Lambda_{t-s}(\xi, \omega) {}_r\Lambda^{k-1}_{ds}(\eta, d\xi). \tag{4.10}$$

Proof We shall use the notation introduced in (4.6) and rewrite (4.8) – (4.10) as

$$ {}_b\Lambda = \Lambda|_{\Omega_b,\Omega_r} \otimes \pi|_{\Omega_r,\Omega_b},$$
$$ {}_r\Lambda = \pi|_{\Omega_r,\Omega_b} \otimes \Lambda|_{\Omega_b,\Omega_r}$$

and

$$ \Lambda^k|_{\Omega_b,\Omega_r} = \Lambda|_{\Omega_b,\Omega_r} \otimes \ {}_r\Lambda^{k-1}|_{\Omega_r,\Omega_r}.$$

Let us prove (4.10) by induction. Obviously, (4.10) holds for $k = 1$. Assume that it holds for $j = 1, 2, ..., k$. Then

$$\begin{aligned} \Lambda^{k+1}|_{\Omega_b,\Omega_r} &= \Lambda|_{\Omega_b,\Omega_b} \otimes \Lambda^k|_{\Omega_b,\Omega_r} \\ &= \Lambda|_{\Omega_b,\Omega_r} \otimes \pi|_{\Omega_r,\Omega_b} \otimes \Lambda|_{\Omega_b,\Omega_r} \otimes \ {}_r\Lambda^{k-1}|_{\Omega_r,\Omega_r} \\ &= \Lambda|_{\Omega_b,\Omega_r} \otimes \ {}_r\Lambda|_{\Omega_r,\Omega_r} \otimes \ {}_r\Lambda^{k-1}|_{\Omega_r,\Omega_r} \\ &= \Lambda|_{\Omega_b,\Omega_r} \otimes \ {}_r\Lambda^{k}|_{\Omega_r,\Omega_r}. \end{aligned}$$

And this completes the proof. ■

Also, the states at birth of the k-th generation are obtained by applying the k-th generation transmission kernel when an individual enters Ω_r. Indeed, let us prove

Lemma 2 *For every $t \in \mathbb{R}^+$, $\eta \in \Omega_b$ and every measurable $\omega \subseteq \Omega_b$ we have*

$$ {}_b\Lambda^k_t(\eta, \omega) = \int_{[0,t]\times\Omega_r} \Lambda^k_{t-s}(\xi, \omega)\pi(\eta)(ds, d\xi). \tag{4.11} $$

Proof We need to see that

$$ {}_b\Lambda^k = \Lambda^k|_{\Omega_b,\Omega_r} \otimes \pi|_{\Omega_r,\Omega_b}. $$

For $k = 1$ this equality is nothing but (4.8). Suppose that (4.11) holds for $j = 1, 2, ..., k$. Then

$$ \begin{aligned} {}_b\Lambda^{k+1} &= {}_b\Lambda|_{\Omega_b,\Omega_b} \otimes {}_b\Lambda^k|_{\Omega_b,\Omega_b} \\ &= {}_b\Lambda|_{\Omega_b,\Omega_b} \otimes {}_b\Lambda^k|_{\Omega_b,\Omega_r} \otimes \pi|_{\Omega_r,\Omega_b} \\ &= \Lambda^{k+1}|_{\Omega_b,\Omega_r} \otimes \pi|_{\Omega_r,\Omega_b}. \end{aligned} $$

■

Now we are ready to rewrite the expression for the clan kernel.

Theorem 2

1. *For every $t \in \mathbb{R}^+$, $\eta \in \Omega$ and every measurable $\omega \subseteq \Omega$*

$$ \Lambda^c_t(\eta, \omega) = \Lambda_t + \int_{[0,t]\times\Omega_b} {}_b\Lambda^c_{t-s}(\xi, \omega)\Lambda_{ds}(\eta, d\xi), $$

2. *If, for some renewal set Ω_r the kernel ${}_b\Lambda$ can be written in the form* (4.9) *then*

$$ {}_b\Lambda^c_t(\eta, \omega) = \int_{[0,t]\times\Omega_r} \Lambda^c_{t-s}(\xi, \omega)\pi(\eta)(ds, d\xi), $$

3.

$$ \Lambda^c_t(\eta, \omega) = \int_{[0,t]\times\Omega_r} \Lambda_{t-s}(\xi, \omega)_r\Lambda^c_{ds}(\eta, d\xi). $$

Proof The first equality is obtained from (4.7) by summing over k. Also with summation we get the second from (4.11) and the third from (4.10).

4.4 Survival and the i-state movement

In this section we shall describe the (possibly stochastic) i-state dynamics and later on incorporate it in the construction of the transmission kernel. Thus, we introduce

$u_t(\eta, \omega)$:= *the probability that an individual at some time, say s, has state η, is alive t units of time later and has state in $\omega \subseteq \Omega$.*

What we have in mind is that an individual's movement in Ω is a Markov process with death as an absorbing state.
We shall assume that u maps $\mathbb{R}^+ \times \Omega$ to the set $\mathcal{P}(\Omega)$ of probability measures on Ω, and for every $\omega \subseteq \Omega$ the function

$$\begin{aligned} \mathbb{R}^+ \times \Omega &\rightarrow [0,1] \\ (t, \eta) &\mapsto u_t(\eta, \omega) \end{aligned}$$

is measurable. We allow that for every $t \in \mathbb{R}^+$ and every $\eta \in \Omega$,

$$u_t(\eta, \Omega) \leq 1,$$

that is, $u_t(\eta, .)$ is a possibly defective probability measure.
Also, the definition of u suggests that the i-states are states in the Markovian sense, that is, for knowing the individual's state after some time t it does not matter whether we stop at some time, say s $(0 < s < t)$, remember it and look again at the state $t - s$ units of time later. In other words, we assume that the *Chapman - Kolmogorov* relation

$$u_t(\eta, \omega) = \int_\Omega u_{t-s}(\xi, \omega) u_{ds}(\eta, d\xi) \tag{4.12}$$

holds for every $t \in \mathbb{R}^+$, $0 \leq s \leq t$, $\eta \in \Omega$ and $\omega \subseteq \Omega$.
Written shortly,

$$u_t = u_{t-s} \otimes u_s. \tag{4.13}$$

Having the biological background in mind, we might want to make some additional requirements for u, for example the assumption that for every $\eta \in \Omega$

$$\lim_{t \to \infty} u_t(\eta, \Omega) = 0$$

would mean that all individuals are mortal.
Or, perhaps, that life expectancy is uniformly bounded,

$$\sup_{\eta \in \Omega} \int_0^\infty t u_{dt}(\eta, \Omega) < \infty.$$

Also, if we want to know the number of infected individuals that one infected individual caused in some time t it should not matter if we count the

infecteds at some earlier time, say $0 < s < t$ and follow the transmission from there on. Thus, we shall assume that another relation holds

$$\Lambda_t(\eta, \omega) = \Lambda_s(\eta, \omega) + \int_\Omega \Lambda_{t-s}(\xi, \omega) u_s(\eta, d\xi), \tag{4.14}$$

or in the shorthand notation

$$\Lambda_t = \Lambda_s + \Lambda_{t-s} \otimes u_s. \tag{4.15}$$

4.5 Putting the pieces together

Consider one infected individual that had at the time of infection state $\eta \in \Omega$. In the course of time this individual will infect other individuals and how many infections it will cause may depend on its current i-state. What we want now to calculate is the size of its clan and describe its distribution in Ω.
So we define

$$u_t^c(\eta, \omega) := u_t(\eta, \omega) + \int_{[0,t]\times\Omega} u_{t-s}(\xi, \omega) \Lambda_{ds}^c(\eta, d\xi). \tag{4.16}$$

Clearly, for any $\omega \subseteq \Omega$ the function

$$\begin{array}{rcl} \mathbb{R}^+ \times \Omega & \to & \mathbb{R} \\ (t, \eta) & \mapsto & u_t^c(\eta, \omega) \end{array}$$

is measurable.
The interpretation suggests that u^c should satisfy the Chapman - Kolmogorov relation. Let us check that that is indeed the case. First we shall prove a higher power analogue of (4.14). To do that we need to describe the distribution of the k-th generation, so we define

$$u_t^k(\eta, \omega) := \int_{[0,t]\times\Omega} u_{t-s}(\xi, \omega) \Lambda_{ds}^k(\eta, d\xi). \tag{4.17}$$

Let us prove the following

Lemma 3 *For any $t \in \mathbb{R}^+$, $\eta \in \Omega$ and $\omega \subseteq \Omega$ we have the relation*

$$\Lambda_t^k(\eta, \omega) = \Lambda_s^k(\eta, \omega) + \sum_{j=0}^{k-1} \int_\Omega \Lambda_{t-s}^{k-j}(\xi, \omega) u_s^j(\eta, d\xi). \tag{4.18}$$

Proof We shall prove the validity of this relation by induction. Now, for $k = 1$ this is nothing but (4.14) which holds by assumption. Assume that the relation holds for $i = 1, ..., k$. Now we want to calculate the $(k+1)$-th generation at some time t. If we stop at some time $0 < s < t$ and count the size of the $(k+1)$-th generation until that time we are then only left with the infections of the $(k+1)$-th generation that occur after time s. We shall divide the latter into two groups - those who result from the k-th generation before s and those which result from the k-th generation that also occured after s. Taking (4.1), (4.14) and (4.18) into account we write

$$\begin{aligned}
\Lambda_t^{k+1}(\eta,\omega) &= \Lambda_s^{k+1}(\eta,\omega) + \int_{[0,s]\times\Omega} \Lambda_{t-\tau}(\xi,\omega)\Lambda_{d\tau}^k(\eta,d\xi) \\
&+ \int_{(s,t]\times\Omega} \Lambda_{t-\tau}(\xi,\omega) \sum_{j=0}^{k-1} \int_\Omega \Lambda_{d(\tau-s)}^{k-j}(\zeta,d\xi) u_s^j(\eta,d\zeta) \\
&= \Lambda_s^{k+1}(\eta,\omega) + \int_{[0,s]\times\Omega} \int_\Omega \Lambda_{t-s}(d\zeta,\omega) u_{s-\tau}(\xi,d\zeta)\Lambda_{d\tau}^k(\eta,d\xi) \\
&+ \int_{(0,t-s]\times\Omega} \Lambda_{t-s-\sigma}(\xi,\omega) \sum_{j=0}^{k-1} \int_\Omega \Lambda_{d\sigma}^{k-j}(\zeta,d\xi) u_s^j(\eta,d\zeta) \\
&= \Lambda_s^{k+1}(\eta,\omega) + \int_\Omega \Lambda_{t-s}(d\zeta,\omega) u_s^k(\eta,d\zeta) \\
&+ \sum_{j=0}^{k-1} \int_\Omega \Lambda_{t-s}^{k+1-j}(\zeta,\omega) u_s^j(\eta,d\zeta) \\
&= \Lambda_s^{k+1}(\eta,\omega) + \sum_{j=0}^{k} \int_\Omega \Lambda_{t-s}^{k+1-j}(\zeta,\omega) u_s^j(\eta,d\zeta). \qquad \blacksquare
\end{aligned}$$

And then right away

Corollary 1 *For any $t \in \mathbb{R}^+$, $\eta \in \Omega$ and $\omega \subseteq \Omega$,*

$$\Lambda_t^c(\eta,\omega) = \Lambda_s^c(\eta,\omega) + \int_\Omega \Lambda_{t-s}^c(\xi,\omega) u_s^c(\eta,d\xi). \tag{4.19}$$

Proof The relation is obtained by summation of (4.18) over k. $\blacksquare$

We shall also need

Lemma 4 *For any $t \in \mathbb{R}^+$, $\eta \in \Omega$ and $\omega \subseteq \Omega$ we have the identity*

$$u_t^k(\eta,\omega) = \sum_{j=0}^{k} \int_\Omega u_{t-s}^{k-j}(\xi,\omega) u_s^j(\eta,d\xi) \tag{4.20}$$

Proof We take the definition (4.17) as a starting point and split it into two parts. For the first part we use the Chapman-Kolmogorov relation that was assumed to hold for u and for the second part we use (4.18) to arrive at

$$\begin{aligned}
u_t^k(\eta,\omega) &= \int_{[0,s]\times\Omega} u_{t-\tau}(\xi,\omega)\Lambda_{d\tau}^k(\eta,d\xi) + \int_{(s,t]\times\Omega} u_{t-\tau}(\xi,\omega)\Lambda_{d\tau}^k(\eta,d\xi) \\
&= \int_\Omega u_{t-s}(\zeta,\omega)\int_{[0,s]\times\Omega} u_{s-\tau}(\xi,d\zeta)\Lambda_{d\tau}^k(\eta,d\xi) \\
&+ \int_{(s,t]\times\Omega} u_{t-\tau}(\xi,\omega)\sum_{j=0}^{k-1}\int_\Omega \Lambda_{d(\tau-s)}^{k-j}(\zeta,d\xi)u_s^j(\eta,d\zeta) \\
&= \int_\Omega u_{t-s}(\zeta,\omega)u_s^k(\eta,d\zeta) \\
&+ \sum_{j=0}^{k-1}\int_\Omega\int_{(0,t-s]} u_{t-s-\sigma}(\xi,\omega)\Lambda_{d\sigma}^{k-j}(\zeta,d\xi)u_s^j(\eta,d\zeta) \\
&= \int_\Omega u_{t-s}(\zeta,\omega)u_s^k(\eta,d\zeta) + \sum_{j=0}^{k-1}\int_\Omega u_{t-s}^{k-j}(\zeta,\omega)u_s^j(\eta,d\zeta) \\
&= \sum_{j=0}^{k}\int_\Omega u_{t-s}^{k-j}(\zeta,\omega)u_s^j(\eta,d\zeta) \qquad \blacksquare
\end{aligned}$$

Now, the Chapman-Kolmogorov relation for u^c comes easily

Theorem 3 *For any $t \in \mathbb{R}^+$, $\eta \in \Omega$ and $\omega \subseteq \Omega$ we have*

$$u_t^c(\eta,\omega) = \int_\Omega u_{t-s}^c(\xi,\omega)u_s^c(\eta,d\xi). \tag{4.21}$$

Proof The relation is obtained by simply summing (4.20) with respect to k from 1 to ∞. $\blacksquare$

4.6 The population level

So far we have described the survival, the i-state movement and transmission for one infected individual. What remains is to add all the individual contributions to desribe the whole process at the population level.
We describe the size and the distribution of individuals in Ω with a positive measure $m \in M_+(\Omega)$. The size and the distribution of the next generation of infections caused by this initial distribution of infected individuals in some time interval of length t is then given via a linear operator

$\Lambda_t : M(\Omega) \to M(\Omega)$ defined by

$$(\Lambda_t m)(\omega) := \int_\Omega \Lambda_t(\eta, \omega) m(d\eta). \tag{4.22}$$

When

$$\sup_{\eta \in \Omega} \Lambda_t(\eta, \Omega) < \infty,$$

(4.22) defines a bounded linear operator on $M(\Omega)$, that leaves the cone $M_+(\Omega)$ invariant.

In the same spirit we define an operator that gives us the size of a clan produced in a time interval of length t,

$$(\Lambda_t^c m)(\omega) := \int_\Omega \Lambda_t^c(\eta, \omega) m(d\eta). \tag{4.23}$$

The operator U_t that describes the i-state dynamics in a time interval of length t is defined as

$$(U_t m)(\omega) := \int_\Omega u_t(\eta, \omega) m(d\eta). \tag{4.24}$$

and

$$(U_t^c m)(\omega) := \int_\Omega u_t^c(\eta, \omega) m(d\eta). \tag{4.25}$$

describes the i-structure of a clan after some time t.

Now, using the identities (4.12), (4.14), (4.20) and (4.21) we arrive at

Corollary 2 *For any $s, t \in \mathbb{R}^+$ the following identities hold*

$$\begin{aligned} \Lambda_{t+s} &= \Lambda_s + \Lambda_t \otimes U_s, \\ \Lambda_{t+s}^c &= \Lambda_s^c + \Lambda_t^c \otimes U_s^c, \end{aligned}$$

$$\begin{aligned} U_{t+s} &= U_t \otimes U_s, \\ U_{t+s}^c &= U_t^c \otimes U_s^c, \end{aligned}$$

so $\{U_t; t \geq 0\}$ and $\{U_t^c; t \geq 0\}$ form a one parameter semigroup.
Sometimes the following terminology is used: $\{U_t; t \geq 0\}$ (or $\{U_t^c; t \geq 0\}$) forms a **forward evolutionary system** *with a* **cumulative output family** *$\{\Lambda_t; t \geq 0\}$ (or $\{\Lambda_t^c; t \geq 0\}$).*

4.7 Epidemic or no epidemic ?

At the end of this chapter we are only left with writing the next generation operator in this general setting. To do that we denote for $\eta \in \Omega$ and $\omega \subseteq \Omega$

$\Lambda_\infty(\eta, \omega) :=$ *the expected number of cases with h-state in* ω *caused by one individual that was itself infected while having h-state* η*, during its entire period of infectiousness.*

Let the measure $m \in M_+(\Omega)$ again describe the size and the distribution of infected individuals. The linear space $M(\Omega)$ becomes, with total variation norm and the order relation defined in section 4.1., a Banach lattice and the operator $\Lambda_\infty : M(\Omega) \to M(\Omega)$ defined with

$$(\Lambda_\infty m)(\omega) := \int_\Omega \Lambda_\infty(\eta, \omega) m(d\eta) \tag{4.26}$$

is a linear operator on $M(\Omega)$. Since it gives us the next generation of infected individuals we call it the next generation operator.
When

$$\sup_{\eta \in \Omega} \Lambda_\infty(\eta, \Omega) < \infty,$$

(meaning loosely speaking that even the most infectious individual produces in all of its infectious period a finite number of infections), Λ_∞ is a bounded linear operator on $M(\Omega)$.

By definition then, the basic reproduction ratio R_0 is the spectral radius of Λ_∞. And again we have to, in order to guarantee that in the case of $R_0 > 1$ each initial distribution triggers an epidemic, assume irreducibility of Λ_∞.

Irreducibility of the kernel Λ_∞ is now defined as follows: Λ_∞ is **irreducible** if for every measurable subset $\omega \subseteq \Omega$ with positive measure such that $\Omega \backslash \omega$ also has positive measure,

$$\int_{\Omega \backslash \omega} \Lambda_\infty(\eta, \omega) d\eta > 0.$$

In order to obtain convergence to a stable distribution for states at birth, we again need to pose additional assumptions on the kernel.
Now, which conditions guarantee that the spectral radius of Λ_∞ is a strictly positive and a strictly dominant eigenvalue with a positive eigenvector ? In 1948 Krein and Rutman[1] proved that a positive, compact operator on a Banach Lattice with a strictly positive spectral radius possesses this property,

[1] Appendix

that is, the spectral radius is a strictly positive eigenvalue with a corresponding positive eigenvector. Some years later, in 1986 to be precise, de Pagter (we refer again to the Appendix) proved that each irreducible, compact nonzero positive operator on a Banach Lattice has a nonzero spectral radius. The strict dominance of the spectral radius is obtained with the assumption of strong irreducibility of Λ_∞.
Strong irreducibility is an analogue of primitivity (defined in the case when Ω is finite) and is defined as follows: a positive operator K on a Banach Lattice E is *strongly irreducible* if $K\phi$ is a weak order unit for every $0 < f \in E$. For kernel operators, however, there is an equivalent definition of strong irreducibilty[2], which offers some biological interpretation and goes as follows.

Let Σ denote the σ-algebra of measurable subsets of Ω and

$$\Delta = \{(\eta, \omega) \in \Omega \times \Sigma; \Lambda_\infty(\eta, \omega) = 0\} \subset \Omega \times \Sigma.$$

The kernel Λ_∞ is called **strongly irreducible** if Δ does not contain a subset $A \times \omega$, where A and ω are measurable subsets of Ω with positives measures.

In our context this means that, no matter what subsets of i-states with positive measures we choose, infected individuals with states in one set will eventually cause infections with i-states in the other set.

The assumptions of compactness of Λ_∞ (in fact, it is enough to assume that some power of Λ_∞ is compact) seems at first hard to check, but, actually we have, thanks to the fact that we are dealing with absolute kernel operators a condition that moreover offers interpretation. For details we refer to the book by Zaanen[3], here we shall only write down the condition and the interpretation.

The kernel Λ_∞ is **compact** if for every sequence ω_n of measurable subsets in Ω such that $\omega_n \downarrow 0$, the operator norms of the operators with kernels $\Lambda_\infty|_{\omega_n,\Omega}$ (for notation see (4.6)) tend to 0 as $n \to \infty$.
This assumption poses no real restriction on the kernel since it only means the following: we make a restriction on the "allowed" set of states at birth. When we make this set of allowed states smaller and smaller, the number of new cases with states in this allowed set will, no matter which initial distribution of individuals we take, go to zero as well. For example, think that age is the only i-state. If we make more and more restrictions on how old the individual is when it becomes infected, the number of newly infected with the required age will surely go to zero.

[2] **Zannen, A.C.**, Introduction to Operator Theory in Riesz Spaces
[3] **Zaanen, A.C.** Riesz Spaces II

Chapter 5

The intrinsic growth rate

Looking at the initial phase of a (potential) epidemic from a generation perspective has several advantages. First of all, it gives us a clear threshold: given that the next generation operator is irreducible (or, in other words: each subgroup of individuals necessarily causes infections outside its own group), the epidemic will occur if and only if $R_0 > 1$. Secondly, we often get explicit expressions for R_0 that give us further insight into the role of various factors.
However, where there are pros there are cons as well and one disadvantage is that, since the generations of infecteds may overlap each other, generations of infections are not what we actually observe.

So let us see whether there exists,in the context of a model, a number that corresponds to the real time growth rate. When such a number exists it will be denoted with r and called the *intrinsic growth rate*. We insist that growth in a generation perspective should correspond to growth in real time and vice versa, or stated differently, we insist that

$$R_0 > 1 \iff r > 0.$$

5.1 On the existence of the real time growth rate

Let us, to get a feeling for the problem, first consider the simplest example: let us assume that we are dealing with a homogeneous population where each individual has on average c contacts per unit of time. Each contact leads to transmission with constant probability p. Once an individual gets infected, there is a latency period of length $t_1 \geq 0$, in which an individual is infected, but not yet infectious. The latency period is followed by an

infectious period of length $t_2 - t_1$.
We define

I(t) := the expected number of infected individuals notified up to time t,

where time is measured from some convenient starting point on. The number $I(t)$ is also called the *prevalence*. The *incidence* $i(t)$ is defined as

i(t) := the expected number of new cases per unit of time, measured at time t

and will be proportional to $\frac{dI}{dt}$.
Since new cases at time t result from the infected individuals that were infected before that time and are still infectious at time t, the incidence satisfies the equation

$$i(t) = pc \int_{t_1}^{t_2} i(t-\tau) d\tau. \tag{5.1}$$

Now, with exponential increase (or decline) during the initial phase we have in mind that

$$I(t) \doteq Ce^{rt}$$

for some positive constant C and some real r. Thus, using (5.1) we find that r should satisfy the equality

$$1 = pc \int_{t_1}^{t_2} e^{-r\tau} d\tau. \tag{5.2}$$

And indeed, as the next argument shows, such an r exists and is unique.
Let $f(z) = pc \int_{t_1}^{t_2} e^{-z\tau} d\tau$. For $z = 0$ we obtain $f(0) = pc(t_2 - t_1) = R_0$. Moreover, as $z \to \infty$, $e^{-z\tau} \to 0$ uniformly on $[t_1, t_2]$, hence $f(z) \to 0$.
If $t_1 > 0$, the same argument applies when verifying that $f(z) \to \infty$ as $z \to -\infty$. If $t_1 = 0$ we restrict to some $\epsilon > 0$. Then again $e^{-z\tau} \to 0$ uniformly on $[\epsilon, t_2]$ and $f(z) \to \infty$ as $z \to -\infty$.

Thus, the solution of (5.2) exists and is unique so the intrinsic growth rate is well defined. Obviously,

$$R_0 > 1 \iff r > 0.$$

When dealing with the problem in a general setting, the existence of the real time growth rate is not always guaranteed. For some conditions that guarantee (in the context of the model) the existence of the intrinsic growth

we refer to the literature[1].
"R_0" deals with a situation where the distribution of individuals in Ω can be described with a density function, which is a positive element of $L^1(\Omega)$. However, sometimes it is advantageous (or even necessary) to describe the distribution of infected individuals in Ω with a (positive) measure on Ω. As this measure is not necessarily absolutely continuous, the corresponding density function may not exist. So let us now deal with the real time growth rate in this general setting.

Let Ω again denote the i-state space and let for $\eta \in \Omega$, $\omega \subseteq \Omega$

$$\Lambda_t(\eta, \omega) = \textit{the expected number of new infections with } i\textit{-state in } \omega, \textit{ produced in the time interval of length } t \textit{ by one individual with birth state } \eta.$$

By exponential increase (or decline) during the initial phase we now have in mind that

$$i(t, \omega) \doteq e^{rt}\Phi(\omega) \tag{5.3}$$

for some real r and some positive measure Φ on Ω.
Since new infections at time t with i-state in ω result from infected individuals that were infected before time t with some birth state $\eta \in \Omega$, we obtain

$$i(t, \omega) = \int_{\mathbb{R}^+ \times \Omega} \Lambda_{d\tau}(\eta, \omega) i(t - \tau, d\eta). \tag{5.4}$$

Using (5.3) and (5.4) we find that

$$\begin{aligned} \Phi(\omega) &= e^{-rt} i(t, \omega) \\ &= e^{-rt} \int_{\mathbb{R}^+ \times \Omega} \Lambda_{d\tau}(\eta, \omega) i(t - \tau, d\eta) \\ &= \int_{\mathbb{R}^+ \times \Omega} e^{-r\tau} \Lambda_{d\tau}(\eta, \omega) \Phi(d\eta), \end{aligned}$$

which means that Φ is an eigenvector corresponding to the eigenvalue 1 of an operator

$$(H_r m)(\omega) = \int_{\mathbb{R}^+ \times \Omega} e^{-r\tau} \Lambda_{d\tau}(\eta, \omega) m(d\eta). \tag{5.5}$$

So let $M(\Omega)$ denote the linear space of measures on Ω. Equipped with the total variation norm and with the following order relation: for $m_1, m_2 \in M(\Omega)$,

[1] Heesterbeek, J.A.P.: R_0, 1992

$$m_1 \leq m_2 \overset{def}{\iff} m_1(\omega) \leq m_2(\omega) \quad \text{for every } \omega \subseteq \Omega,$$

$M(\Omega)$ becomes a Banach lattice. The cone of all nonnegative measures will be denoted with $M_+(\Omega)$.
We shall describe the size and the distribution of infected individuals in Ω with a positive measure and consider the Laplace - Stieltjes transform of the next generation operator, that is, we define

$$(H_z m)(\omega) := \int_{\mathbb{R}^+ \times \Omega} e^{-z\tau} \Lambda_{d\tau}(\eta, \omega) m(d\eta), \tag{5.6}$$

for those $z \in \mathbb{C}$ for which (5.6) converges.
Returning briefly to the biological setting we observe that H_z counts infections while weighing a newly infected individual with factor e^{-zt} when t is the time elapsed between the infection and the time the individual responsible for the infection itself got infected.
So we are looking at a family of operators on a Banach lattice $M(\Omega)$ defined in (5.6) (note that these operators leave the cone of nonnegative measures invariant) and looking for a member that has an eigenvalue 1. In fact, under additional conditions on the operators H_z, the spectral radius (which will be denoted with $spr(.)$) is the only eigenvalue with a positive eigenvector and, since Φ should be positive we would be looking for a $z \in \mathbb{R}$ such that

$$spr(H_z) = 1. \tag{5.7}$$

As mentioned at the end of the previous Chapter, the assumptions of irreducibility of H_z and powers of operators H_z being compact are sufficient to guarantee that the spectral radius of H_z is a nonzero eigenvalue with a strictly positive eigenvector. Also (and this is important as well, as we would not wish, unless really necessary, to pose conditions that would severely reduce the set of kernels that come into play), none of the assumptions poses a real restriction on the kernels. Thus, our assumptions are

A_1: for every $z \in \mathbb{R}$ for which (5.6) converges, the operator H_z is irreducible

and

A_2: for every $z \in \mathbb{R}$ there exists a $n_z \in \mathbb{N}$ such that the operator $H_z^{n_z}$ is compact.

These two assumptions also guarantee (this Theorem can be found in Schaefer[2] and is also given in the Appendix) that

[2] **Schaefer, H.H.**, Banach Lattices and Positive Opertors

1. $spr(H_z)$ is a simple eigenvalue. The corresponding eigenvector will be denoted with Φ_z and normalized so that $\|\Phi_z\| = 1$.

2. Φ_z is a quasi - interior point of $M_+(\Omega)$, that is, $\langle F, \Phi_z \rangle > 0$ for every $F \in (M(\Omega))^*_+ \backslash \{0\}$,

3. the dual eigenvector $F_z \in (M(\Omega))^*$ corresponding to $spr(H_z)$ is a strictly positive functional.

Now,

$$R_0 = spr(H_0).$$

To prove the existence of r we need

a. $spr(H_z)$ depends continuously on z,

b. the function $z \mapsto spr(H_z)$ is strictly decreasing.

In case when $R_0 > 1$, the existence and uniqueness of the real time growth rate will be guaranteed if we find out

c. under which conditions there exists a $z_1 > 0$ with $spr(H_{z_1}) \leq 1$?

When $R_0 < 1$, the negative values of z come into play. If (5.6) is defined for $z > z_0 \in \mathbb{R}$ for some $z_0 \in \mathbb{R}$, the existence and uniqueness of r will be guaranteed if we find

d. under which conditions there exists a $z_2 \in (z_0, 0)$ with $spr(H_{z_2}) \geq 1$?

Let us first deal with **a.**

Definition 6 *The number s is a simple eigenvalue of a bounded linear operator $K : X \to X$ if there exist closed subspaces X_1, X_2 such that*

- $X = X_1 \oplus X_2$, $K(X_i) \subseteq X_i$, $i = 1, 2$,
- $K|_{X_2} = sI$, *dim* $X_2 = 1$,
- s *is in the resolvent set of* $K|_{X_1}$.

That means there exists an $F \in X^*$ such that $X_1 = \ker F$ and if $Kv = sv$, $v \neq 0$, $X_2 = \mathcal{L}(v)$ and $\langle F, v \rangle \neq 0$.

We shall need the following

Lemma 5 *Let $\{K_\lambda : X \to X;\ \lambda \in \mathbb{C}\}$ be a family of bounded, linear operators. Let*

$$\lim_{\lambda \to \mu} \|K_\lambda - K_\mu\| = 0.$$

Furthermore, let s be a simple eigenvalue of K_μ. Then for any $\epsilon > 0$ there exists $\delta > 0$ such that $|\mu - \lambda| < \delta$ implies that there exists $p \in \sigma(K_\lambda)$ such that $|p - s| < \epsilon$.

Proof Let F and v be as introduced above and define

$$g_\lambda(z) = \langle F, (z - K_\lambda)^{-1} v \rangle \ ; \qquad z \in \rho(K_\lambda).$$

Then

$$g_\mu(z) = \frac{1}{z - s} \langle F, v \rangle. \tag{5.8}$$

Let $\epsilon > 0$. Denote $B_\epsilon = \{z \in \mathbb{C}; |z - s| < \epsilon\}$. For a small enough ϵ we have $\partial B_\epsilon \subset \rho(K_\mu)$. Since $K_\lambda \to K_\mu$ the fact $z \in \partial B_\epsilon$ implies that for λ sufficiently close to μ we have $z \in \rho(K_\lambda)$ and also $(z - K_\lambda)^{-1} \to (z - K_\mu)^{-1}$ for $\lambda \to \mu$. By compactness of ∂B_ϵ the convergence is uniform. For λ sufficiently close to μ and $z \in \partial B_\epsilon$ we have

$$|\langle F, (z - K_\lambda)^{-1} v \rangle - \langle F, (z - K_\mu)^{-1} v \rangle| < \frac{|\langle F, v \rangle|}{2\epsilon},$$

hence using (5.8) we obtain,

$$|g_\lambda(z) - g_\mu(z)| < |g_\mu(z)|.$$

That means that, if g_λ is meromorphic at all, g_λ has the same index ($\sharp$ poles - $\sharp$ zeros) in B_ϵ as g_μ. But if $B_\epsilon \cap \sigma(K_\lambda)$ was empty, g_λ would be analytic on B_ϵ. ■

Now we are ready to prove

Theorem 4 *$spr(K_\lambda)$ depends continuously on λ.*

Proof Let

$$\mu = \lim_{j \to \infty} \lambda_j$$

and denote $spr_\infty = \limsup_j spr(K_{\lambda_j})$. Without any loss of generality we can assume

$$spr_\infty = \lim_{j \to \infty} spr(K_{\lambda_j}).$$

If $spr_\infty \notin \sigma(K_\mu)$, then $spr_\infty \in \rho(K_\lambda + spr_\infty - spr(K_\lambda))$, if only $\|(K_\lambda + spr_\infty - spr(K_\lambda)) - K_\mu\|$ is small enough. But that would mean that $spr(K_\lambda) \in$

$\rho(K_\lambda)$, which leads to contradiction. So $spr_\infty \in \sigma(K_\mu)$ and $spr(K_\mu) \geq spr_\infty$. Thus, $\limsup spr(K_\lambda) \leq spr(K_\mu)$.
Since $spr(K_\mu)$ is a simple eigenvalue it follows from the previous lemma that for any $\epsilon > 0$ and λ sufficiently close to μ

$$\sigma(K_\lambda) \cap \{z; |z - spr(K_\mu)| < \epsilon\} \neq \emptyset.$$

For such λ we have $spr(K_\lambda) \geq spr(K_\mu) - \epsilon$, so $\liminf spr(K_\lambda) \geq spr(K_\mu)$. Altogether now we have shown that

$$\lim_{j \to \infty} spr(K_\lambda) = spr(K_\mu).$$

That completes the proof. ■

Let us now see whether $z \mapsto spr(H_z)$ is strictly decreasing. That it is decreasing is clear, since for two non-negative operators A, B such that $A \leq B$ follows $spr(A) \leq spr(B)$. To show that it is in fact strictly decreasing will also guarantee that the solution of (5.7), if it exists, is unique.

Theorem 5 *The assumptions A_1 and A_2 guarantee that the function $z \mapsto r(H_z)$ is strictly decreasing.*

Proof Let $z, w \in \mathbb{R}$, $z > w$ and $\Phi \in M_+(\Omega)$. For those $\omega \subseteq \Omega$ for which $(H_z\Phi)(\omega) \neq 0$ we have

$$(H_w\Phi)(\omega) > (H_z\Phi)(\omega). \tag{5.9}$$

If we take $\Phi = \Phi_z$ we first observe that, since Φ_z is strictly positive and $spr(H_z) > 0$, (5.9) holds for every $\omega \subseteq \Omega$, so

$$H_w(\Phi_z) > spr(H_z)\Phi_z.$$

Denoting with F_w the dual eigenvector corresponding to $spr(H_w)$ we find that

$$\langle F_w, H_w(\Phi_z)\rangle \ > \ \langle F_w, spr(H_z)\Phi_z\rangle,$$

so

$$spr(H_w)\langle F_w, \Phi_z\rangle \ > \ spr(H_z)\langle F_w, \Phi_z\rangle$$

and since $\langle F_w, \Phi_z\rangle \ > 0$ we indeed get $spr(H_w) > spr(H_z)$. ■

To answer **c.** and **d.** let us write

Definition 7 *The Laplace - Stieltjes transform of the kernel Ψ is defined (for real values of z at first) as*

$$\widehat{\Psi}_z(\eta,\omega) := \int_{\mathbb{R}^+} e^{-zt}\Psi_{dt}(\eta,\omega).$$

We define

$$\|\Psi\|_z := \sup_{\eta\in\Omega} \widehat{\Psi}_z(\eta,\Omega).$$

A kernel Ψ is called **Laplace kernel** *if $\|\Psi\|_{z_0} < \infty$ for some $z_0 \in \mathbb{R}$.*

Now, let us get to **c.**

To find a sufficient condition that would guarantee that for some $z_0 > 0$ we have $spr(H_{z_0}) \leq 1$ we first need to consider what makes $spr(H_z)$ small. What makes the operator H_z small is the factor e^{-zt}. But this factor is only active for $t > 0$. If in the very short time interval after the infection took place the kernel Λ "blows up", the factor e^{-zt} has no chance of winning.

So let us assume that

A_3: for some $z_0 > 0$ we have $\|\Lambda\|_{z_0} \leq 1$.

Then

$$\begin{aligned} spr(H_{z_0}) &= \|H_{z_0}\Phi_{z_0}\| = (H_{z_0}\Phi_{z_0})(\Omega) \\ &= \int_{\mathbb{R}^+\times\Omega} e^{-z_0 t}\Lambda_{dt}(\eta,\Omega)\Phi_{z_0}(d\eta) \\ &\leq \|\Lambda\|_{z_0}\|\Phi_{z_0}\| \leq 1, \end{aligned}$$

Actually, the assumption A_3 can be replaced with a less restrictive one. Suppose that at the moment of infection an individual with state η produces two new infections with state η. Then, obviously, the kernel Λ "blows up". If, on the other hand it produces two new infections with different states, we still have a chance of suppressing $spr(H_z)$ by looking at further generations.

So let us first see what the product of Laplace - Stieltjes transforms is. Having in mind the biological background, we would expect it to equal the Laplace - Stieltjes transform of the convolution product since it should not matter whether we sample individuals of the second generation at time t and weigh them with factor e^{-zt} or stop at an arbitrary time $0 < s < t$, weigh

all the first generation infections at time s with e^{-zs} and then weigh the weighed first generation that at the remaining time $t-s$ causes infections of the second generation with $e^{-z(t-s)}$. And indeed, if Φ and Ψ are kernels and for some $z_0 \in \mathbb{R}$ Laplace - Stieltjes tranforms of Φ and Ψ exist then

$$\widehat{\Phi}_{z_0} \cdot \widehat{\Psi}_{z_0} = \widehat{(\Phi \otimes \Psi)}_{z_0},$$

where

$$(\Phi \otimes \Psi)_t(\eta, \omega) = \int_{[0,t]\times\Omega} \Phi_{t-s}(\xi, \omega)\Psi_{ds}(\eta, d\xi).$$

For details we refer to Widder.[3]

Thus, for every $k \in \mathbb{N}$

$$\widehat{\Lambda}^k_z(\eta, \omega) = \int_{\mathbb{R}^+} e^{-zu}\Lambda^k_{du}(\eta, \omega)$$

and to sufficiently lower the spectral radius it is enough to suppress the kernel of some k-th generation, that is, if we assume that

A'_3: for some $k \in \mathbb{N}$ and some $z_0 > 0$ we get $\|\Lambda^k\|_{z_0} \leq 1$

Then

$$\begin{aligned} spr^k(H_{z_0}) &= \|H^k_{z_0}\Phi_{z_0}\| = (H^k_{z_0}\Phi_{z_0})(\Omega) \\ &= \int_{\mathbb{R}^+\times\Omega} e^{-z_0 t}\Lambda^k_{dt}(\eta, \Omega)\Phi_{z_0}(d\eta) \\ &\leq \|\Lambda^k\|_{z_0}\|\Phi_{z_0}\| \leq 1, \end{aligned}$$

so $spr(H_{z_0}) \leq 1$.

All that remains is to answer **d.**
So let us assume that (5.6) exists for real values of z greater than some $z_0 \in \mathbb{R}$. One sufficient condition that guarantees that there will be a $z > z_0$ for which $spr(H_z) \geq 1$ is to assume that for some $z \in (z_0, 0)$

A_4 : $C_z := \inf_{\eta\in\Omega} \widehat{\Lambda}_z(\eta, \Omega) = \inf_{\eta\in\Omega} \int_{\mathbb{R}^+} e^{-zt}\Lambda_{dt}(\eta, \Omega) \geq 1.$

Then

$$spr(H_z) = \|H_z\Phi_z\| = \int_{\mathbb{R}^+\times\Omega} e^{-zt}\Lambda_{dt}(\eta, \Omega)\Phi_z(d\eta) \geq C_z\|\Phi_z\| \geq 1.$$

But, as in the previous case, it is enough that A_4 holds for the kernel of some k-th generation, that is if we assume

[3] **Widder**, The Laplace Transform

A_4': for some $k \in \mathbb{N}$ and $z \in (z_0, 0)$,

$$C_z^{(k)} := \inf_{\eta\in\Omega} \widehat{\Lambda^k}_z(\eta, \Omega) = \inf_{\eta\in\Omega} \int_{\mathbb{R}^+} e^{-zt}\Lambda_{dt}^k(\eta, \Omega) \geq 1,$$

then

$$spr^k(H_z) = \|H_z^k\Phi_z\| = \int_{\mathbb{R}^+\times\Omega} e^{-zt}\Lambda_{dt}^k(\eta, \Omega)\Phi_z(d\eta) \geq C_z^{(k)}\|\Phi_z\| \geq 1.$$

At the end of this section we make the following remark: in general there is no order corespondence between R_0 and r, that is, a high value of R_0 does not necessarily imply a high value of r. The point is that, if individuals produce many infections, but rather late after they themselves became infected, the real time growth may still be slow. The next simple argument shows that this is indeed the case.
Assume again that the population is homogeneous, that each individual has on average c contacts per unit of time and each of these contacts leads to transmission with constant probability p. Now, if an individual's infectious period starts t_1 units of time after it became infected and lasts until t_2 then $R_0 = pc(t_2 - t_1)$. But since R_0 doesn't depend on the actual values of t_1 and t_2 but only on their difference, we have $R_0 = pc(t_2 - t_1) = pc(t_2' - t_1') = R_0'$ if only $t_2 - t_1 = t_2' - t_1'$. However if $t_2 - t_2' = t_1 - t_1' = s > 0$ then

$$pc\int_{t_1}^{t_2} e^{-z(\tau+s)}d\tau = pc\int_{t_1}^{t_2} e^{-z\tau}e^{-rs}d\tau < pc\int_{t_1}^{t_2} e^{-z\tau}d\tau,$$

so $r' > r$. If s is large, r can be much smaller than r'. And since R_0 and r are both continuous functions of t_2 we can actually achieve, by making t_2 larger, that $R_0 > R_0'$ while still $r' > r$.

5.2 The computation of r

As mentioned in the introduction of this chapter, the main downside of looking at the initial phase of the spread of the infection in real time is the implicit characterisation of r.
However, as in the case of R_0, the computation of r simplifies when we pose aditional assumptions.
For example, if we assume *separabale mixing* the kernel is of the form

$$\Lambda_t(\eta, \omega) = \alpha(\omega)\beta(t, \eta).$$

Then

$$\begin{aligned}(H_z m)(\omega) &= \int_{\mathbb{R}^+\times\Omega} e^{-zt}\Lambda_{dt}(\eta,\omega)m(d\eta)\\ &= \int_{\mathbb{R}^+\times\Omega} e^{-zt}\alpha(\omega)\beta(dt,\eta)m(d\eta)\\ &= \alpha(\omega)\int_{\mathbb{R}^+\times\Omega} e^{-zt}\beta(dt,\eta)m(d\eta),\end{aligned}$$

so the range of H_z is spanned by a (positive) measure α which is then also an eigenvector corresponding to the spectral radius r, the only nonzero eigenvalue.
The intrinsic growth rate r is characterised by

$$\int_{\mathbb{R}^+\times\Omega} e^{-rt}\beta(dt,\eta)\alpha(d\eta) = 1.$$

The next situation that simplifies the computation of r is the *multigroup separable mixing* already mentioned in section 3.2.3.
There we assumed that the i-state has two components, a discrete one and a continuous one. The discrete one takes values in a set $\{1,2,...,n\}$ and the continuous one ranges in some set Ω_c, so $\Omega = \cup_{j=1}^n \{j\}\times\Omega_c$.
For measurable set $\omega\subseteq\Omega$, $\omega = (i,\omega_c)$ and $(j,\xi)\in\Omega$ we now write

$$\Lambda_t(\eta,\omega) = {}_i\Lambda_t(j,\xi)(\omega_c),$$

so ${}_i\Lambda_t(j,\xi)(\omega_c)$ denotes the expected number of new infections with a discrete component i and continuous component in ω_c caused in the time interval of length t by one individual with i-state (j,ξ).
We shall assume that, conditional on the discrete component being i, the probability distribution of the continuous component is a_i. The dependence in the discrete component will be allowed and described with b_{ij}.
We have

$${}_i\Lambda_t(j,\xi)(\omega_c) = a_i(\omega_c)b_{ij}(t,\xi).$$

Then

$$\begin{aligned}(H_z m)(\omega) &= \sum_{j=1}^n \int_{\mathbb{R}^+\times\Omega_c} e^{-zt}\ {}_i\Lambda_{dt}(j,\xi)(\omega_c)m(j,d\xi)\\ &= \sum_{j=1}^n \int_{\mathbb{R}^+\times\Omega_c} e^{-zt}a_i(\omega_c)b_{ij}(dt,\xi)m(j,d\xi)\\ &= a_i(\omega_c)\sum_{j=1}^n \int_{\mathbb{R}^+\times\Omega_c} e^{-zt}b_{ij}(dt,\xi)m(j,d\xi),\end{aligned}$$

so the range of H_z is spanned by measures a_i on Ω_c. Putting $m(i,\omega_c) = c_i a_i(\omega_c)$ we obtain

$$(H_z m)(i,\omega_c) = a_i(\omega) \sum_{j=1}^{n} \int_{\mathbb{R}^+ \times \Omega_c} e^{-zt} c_j b_{ij}(dt,\xi) a_j(d\xi).$$

That means the coeficients with respect to the basis $\{a_j\}_{j=1}^n$ transform according to the matrix $G = (g_{ij})_{i,j=1}^n$, where

$$g_{ij} = \int_{\mathbb{R}^+ \times \Omega_c} e^{-zt} b_{ij}(dt,\xi) a_j(d\xi).$$

The intrinsic growth rate r is then by definition the spectral radius of G,

$$r = spr(G).$$

At the start of this thesis we considered a special case where the i-state space is finite, say $\Omega = \{1,2,...,n\}$. We return now to this setting to present another characterisation of the intrinsic growth rate.

So let us again assume that the times spent in i-states are exponentially distributed and let Σ contain the transition rates. Moreover, let D denote the diagonal matrix of nonzero death rates. We put the information regarding contact and transmission in a matrix B, as in previous chapters. We saw that the next generation operator K is of the form

$$K = -B(\Sigma - D)^{-1}$$

and that R_0 equals the spectral radius of K. We shall now prove that the intrinsic growth rate r is in fact the spectral bound of the matrix $B + \Sigma - D$. To do that we define

Definition 8 *The* **spectral bound** *of a matrix A is defined as*

$$s(A) = \max\{Re\lambda; \lambda \text{ is an eigenvalue of } A\}.$$

and first prove the following

Lemma 6 *Let $A = (a_{ij})_{i,j=1}^n$ be a real matrix with nonnegative off - diagonal elements, that is*

$$a_{ij} \geq 0; \quad i \neq j.$$

Then e^{tA} is a positive matrix. Moreover

$$s(A) < 0 \iff (\det A \neq 0 \text{ and } -A^{-1} \geq 0).$$

Proof. Since the the off-diagonal elements of A are non-negative we have, for a sufficiently large θ, that $A + \theta I \geq 0$. It follows then from the Taylor series definition of a matrix exponential that

$$e^{t(A+\theta I)} \geq 0.$$

Hence

$$e^{tA} = e^{-t\theta} e^{t(A+\theta I)} \geq 0.$$

To prove the second part we first assume that $s(A) < 0$. Then $\det A \neq 0$, since $\det A = 0$ would mean that 0 is an eigenvalue of A and so $s(A) \geq 0$. Since $s(A) < 0$ the integral

$$\int_0^\infty e^{tA} dt$$

converges. Since $e^{tA} \geq 0$ we also have $\int_0^\infty e^{tA} dt \geq 0$. Using the Taylor series again we obtain

$$A \int_0^t e^{\tau A} d\tau = \int_0^t e^{\tau A} d\tau A = e^{tA} - I$$

and taking the limit $t \to \infty$ we have

$$\int_0^\infty e^{tA} dt = -A^{-1},$$

so $-A^{-1} \geq 0$.
To prove the converse we assume that A^{-1} exists and is negative. Then for sufficiently large θ the matrix $A + \theta I$ is positive and from the Perron - Frobenius theorem[4] we have that for some $\phi \geq 0$ and for some real r we have $A\phi = r\phi$ and $r = s(A)$. Applying A^{-1} to this identity we see that r has to be negative, so $s(A) < 0$. ■

We are now ready to prove

Theorem 6 *Let B be a positive matrix, Σ a positive off-diagonal matrix and D a positive diagonal matrix. Assume that the spectral bound $s(\Sigma - D)$ is negative. Let r denote the spectral bound $s(B + \Sigma - D)$ and let R_0 denote the dominant eigenvalue of the positive matrix $K = -B(\Sigma - D)^{-1}$. Then*

$$R_0 > 1 \iff r > 0.$$

[4] Appendix

Proof. Since $s(\Sigma - D) < 0$ the matrix $(\Sigma - D)^{-1}$ exists and since

$$(B + \Sigma - D)\phi = 0 \iff -B(\Sigma - D)^{-1}\psi = \psi$$

for $\psi = (\Sigma - D)\phi$ it follows

$$R_0 = 1 \iff r = 0.$$

Now assume that $r < 0$. Then by the previous lemma $(B + \Sigma - D)^{-1}$ exists and is negative, so

$$(\Sigma - D)^{-1}(-B(\Sigma - D)^{-1} - I)^{-1} \geq 0. \tag{5.10}$$

Let ψ be a positive eigenvector corresponding to the dominant eigenvalue R_0 of matrix K.
If $R_0 > 1$ we could rewrite the relation $K\psi = R_0\psi$ as

$$(-B(\Sigma - D)^{-1} - I)^{-1}\psi = (R_0 - 1)\psi.$$

If we apply a negative matrix $(\Sigma - D)^{-1}$ to both sides of the last equality we arrive at contradiction with (5.10). So $r < 0$ implies $R_0 < 1$.
To prove the converse assume that $R_0 < 1$. Then, applying lemma 6 to matrix $K - I$ we find that

$$(-B(\Sigma - D)^{-1} - I)^{-1} \leq 0$$

and so

$$(\Sigma - D)^{-1}(-B(\Sigma - D)^{-1} - I)^{-1} \geq 0.$$

Then also

$$-(B + \Sigma - D)^{-1} \geq 0.$$

Let ϕ be a positive eigenvector of $B+\Sigma-D$ corresponding to the eigenvalue r, $(B + \Sigma - D)\phi = r\phi$. If $r > 0$ we arrive, by applying the positive $-(B+\Sigma-D)^{-1}$ to both sides, at a contradiction. Thus $R_0 < 1$ implies $r < 0$ and this completes the proof. ■

Let us now do one example.

Example 7

In this example we return to the Example 3. We obtain

$$L = B + \Sigma - D = \begin{pmatrix} \beta_1 N_1 - \mu_1 - \rho_1 & \zeta \\ p\gamma & \beta_2 N_2 - \mu_2 - \rho_2 - \zeta \end{pmatrix}$$

and

$$r = \frac{1}{2}(l_{11} + l_{22}) + \frac{1}{2}\sqrt{l_{11}^2 - 2l_{11}l_{22} + l_{22}^2 + 4l_{12}l_{21}}.$$

Appendix A

On the spectrum of a positive operator

In this part we shall collect some of the known results regarding the spectrum of a positive operator that were needed in the preceding chapters. For the first part, the spectrum of positive matrices, we refer to Schaefer's book[1]. The rest of the theorems can be found in either Schaefer or one of the books by Zaanen[2],[3] and since we shall only give the theorems, we refer the inquiring reader to these two books for the proofs.

We shal denote the spectral radius by $spr(.)$ and first consider positive operators on finite dimensional vector spaces.

A.1 Positive matrices

Definition 9 *A square matrix $K = (k_{ij})_{i,j=1}^{n}$ is called* **positive**, *in symbols $K \geq 0$, if $k_{ij} \geq 0$ for all i, j. K is called* **strictly positive**, *in symbols $K > 0$, if $k_{ij} > 0$ for all i, j.*

Theorem 7 *If $K \geq 0$, the spectral radius is an eigenvalue (thus by definition the dominant eigenvalue). The dominant eigenvector ϕ^d can be chosen in such a way that $\phi^d \geq 0$. If $K \geq 0$ and $K\phi = \lambda\phi$ for some $\phi > 0$, then $\lambda = spr(K)$.*

[1] **Schaefer, H.H.**, Banach Lattices and Positive operators
[2] **Zaanen, A.C.**, Introduction to Operator Theory in Riesz Spaces
[3] **Zaanen, A.C.**, Riesz Spaces II

Definition 10 *A square matrix $K \geq 0$ is* **irreducible**, *if there exists no permutation matrix P such that*

$$P^{-1}AP = \begin{pmatrix} K_1 & 0 \\ L & K_2 \end{pmatrix}.$$

Theorem 8 (Perron - Frobenius, 1907) *Let $K \geq 0$ be irreducible. Then*

1. *the peripheral spectrum of K is fully cyclic, meaning that for some $l \in \mathbb{N}$,*

$$\{\lambda; \lambda \in \sigma(K), |\lambda| = spr(K)\} = \{spr(K)e^{\frac{2\pi i \cdot k}{l}}; k = 1, 2, ..., l\},$$

2. *$spr(K)$ is an algebraically simple eigenvalue,*
3. *ϕ^d and ϕ^{d*} (where ϕ^{d*} denotes the eigenvector of K^T corresponding to $spr(K)$) are strictly positive,*
4. *no other eigenvector of K is positive.*

Definition 11 *An irreducible matrix K is called* **primitive**, *if $spr(K)$ is the only element of the peripheral spectrum.*

A.2 Positive operators on Banach Lattices

Let E and F denote Banach lattices.

Definition 12 *A linear operator $K : E \to F$ is called* **positive** *(in symbols $K \geq 0$), if $K\phi \geq 0$ for all $\phi \geq 0$. K is called* **strictly positive** *(in symbols $K > 0$), if $K\phi > 0$ for all $\phi > 0$.*

The first generalization of the results of the first part is the next Theorem due to R. Jentzsch (1910). His results were later generalised, but we write this Theorem separately since it includes an important special example, the spaces $L^1(\mu)$.
Let us first write the definition of irreducibility of an integral operator on a space $L^p(\mu)$.

Definition 13 *Let (X, Σ, μ) be a σ-finite measure space and let $E = L^p(\mu)$ $(1 \leq p < \infty)$. Let $K \in \mathcal{L}(E)$ be a kernel operator given via the $(\Sigma \times \Sigma)$ measurable kernel $k \geq 0$ on $X \times X$, so that*

$$Kf(t) = \int_X k(s,t)f(s)d\mu(s)$$

holds a.e. (μ) for each $f \in E$. Then K is **irreducible** *if and only if*

$$\int_{X\backslash S} [\int_S k(s,t)d\mu(s)]d\mu(t) > 0$$

for each $S \in \Sigma$ such that $\mu(S) > 0$ and $\mu(X\backslash S) > 0$.
In particular, K is irreducible, whenever $k(s,t) > 0$ a.e. $(\mu \times \mu)$.

Theorem 9 (Jentzsch, 1910) *Let $E := L^p(\mu)$, where $1 \leq p \leq \infty$ and (Ω, Σ, μ) is a σ-finite measure space. Suppose $K \in \mathcal{L}(E)$ is an operator given by a $(\Sigma \times \Sigma)$ measurable kernel $k \geq 0$, satisfying these two assumptions:*

1. *there exists an $n \in \mathbb{N}$ such that K^n is compact,*

2. *K is irreducible,*

then $spr(K) > 0$ is an eigenvalue of K with a unique normalised eigenfunction ϕ satisfying $\phi > 0$ a.e. (μ). Moreover, if $k(s,t) > 0$ a.e. $(\mu \times \mu)$ then $spr(K)$ is strictly dominant.

The following theorem about positive compact operators on Banach lattices was published in 1948, the theorem is due to Krein and Rutman.

Theorem 10 (Krein - Rutman, 1948) *Let K be a positive and compact operator on E having a strictly positive spectral radius, $spr(K) > 0$. Then $spr(K)$ is an eigenvalue of K with a corresponding positive eigenelement.*

In the case of finite dimensional E irreducibility of K automatically guarantees that $spr(K) > 0$. This is also the case in general, though the fact wasn't proved in full generality until 1986. We define

Definition 14 *The positive operator K on E is called* **ideal irreducible** *if K does not leave any norm closed ideal invariant except E itself and $\{0\}$.*

We have

Theorem 11 (de Pagter, 1986) *If K is a non-zero positive, ideal irreducible and compact operator on E, then $spr(K) > 0$.*

Let E' now denote the order dual of Banach lattice E. For our purposes we may also use the following Theorem that can be found in Schaefer.

Theorem 12 *Let K be a positive irreducible operator with non-zero spectral radius, non-void peripheral point spectrum and possesses an invariant form $\phi = K'\phi$ ($0 < \phi \in E'_+$). The following is true*

1. *the fixed space F of K is one dimensional and spanned by a quasi-interior point point of E_+,*
2. *the peripheral point spectrum of K is a subgroup of the circle group,*
3. *each peripheral eigenvalue of K is simple,*
4. *$spr(K)$ is the unique eigenvalue of K with a positive eigenvector.*

Theorem 13 *If $K \in \mathcal{L}(E)$ is positive and irreducible, the remaining assumptions of Theorem 12 are automatically fulfilled when some power of K is compact.*

Bibliography

[1] Bailey, N.T.J *The Mathematical Theory of Infectious Diseases and its Applications*, Charles Griffin Company Limited, 1975

[2] *Classics of Theoretical Biology, part II*, Bulletin of Mathematical Biology, 1991

[3] Diekmann, O., Heesterbeek, J.A.P.:*Mathematical Epidemiology of Infectious Diseases, Model Building, Analysis and Interpretation*, Wiley, 2000

[4] Diekmann, O., Heesterbeek, J.A.P., Metz, J.A.J.: *On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations*, Journal of Mathematical Biology, 1990, 36, 349-388

[5] Diekmann, O., Dietz, O., Heesterbeek, J.A.P.: *The basic reproduction ratio for sexually transmitted diseases, Part I - Theoretical Considerations*, 1990

[6] Diekmann, O., Gyllenberg, M., Metz, J.A.J., Thieme, H.R.: *On the formulation and analysis of general deterministic population models, I. Linear Theory*, Journal of Mathematical Biology, 1998, 36, 349-388

[7] Gripenberg, G., Londen, S.O., Staffans, O.: *Volterra Integral and Functional Equations*, Cambridge University Press, 1990

[8] Heesterbeek, J.A.P., R_0, PhD Thesis, 1992

[9] Hirsch, M.W., Smale, S.: *Differential Equations, Dynamical Systems and Linear Algebra*, Academic Press, 1974

[10] Kato, T., *Perturbation Theory for Linear Operators*, Springer Verlag, 1966

[11] Mollison, D.: *Epidemic Models: Their Structure and Relation to Data*, Cambridge University Press, 1995

[12] Pazy, A.: *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer Verlag, 1983

[13] Rudin, W.: *Functional Analysis*, McGraw - Hill, 1991

[14] Rudin, W.: *Real and Complex Analysis*, McGraw - Hill, 1970

[15] Schaefer, H.H.: *Banach Lattices and Positive Operators*, Springer - Verlag, 1974

[16] Widder, D.V.: *The Laplace Transform*, Princeton University Press, 1946

[17] Zaanen, A.C.: *Introduction to Operator Theory in Riesz Spaces*, Springer - Verlag, 1997

[18] Zaanen, A.C.: *Riesz Spaces II*, North Holland Publishing Company, 1983