Mathematical methods to gain biological insight

Odo Diekmann & Bob Planqué

November 12, 2019

Contents

Chapter 1

Introduction

This is a master course for mathematics students about mathematical methods to gain insight in the mechanisms underlying biological phenomena.

The course consists of

- these lecture notes, which explain and illustrate the methods while referring to other sources for detailed accounts of the underlying mathematical theory;
- assignments which provide training in modelling and in the use of the methods.

Students work on assignments using both pen and paper and computer tools. Grades are to a large extent based on the handed in written texts and on oral presentations.

In the course, a lot of attention is paid to "translation": how do we get from biological information to a mathematical formulation of questions? And what do the mathematical results tell us about biological phenomena?

In addition, the course aims to introduce general physical ideas about temporal and spatial scales and how these can be used to great advantage when performing a mathematical analysis.

Prerequisites: basic knowledge about linear algebra, analysis, ODEs. (The key point, however, is the attitude: students should be willing to quickly fill in gaps in background knowledge.)

Chapter 2

Exploiting time scale differences: the quasi-steady-state approximation (QSSA)

The aim of this chapter is to illustrate, by way of examples, how one can take advantage of large differences in the rates of change of variables. The idea is simple: we first consider the fast variables for fixed values of the slow variables. If, in this setting, the fast variables go to a limit, a "steady" state, we next consider the *dynamics* of the slow variables while putting the fast variables at their "steady" state values. Note that the precise "steady" values of the fast variables depend on the values at which the slow variables were fixed, and so will vary slowly if we consider the dynamics of the slow variables. To reflect that phenomenon, we say *quasi*-steady-state. The key point is that we decompose a system into two lower-dimensional systems, which often substantially facilitates the analysis.

2.1 Michaelis-Menten enzyme kinetics

Enzymes are large protein molecules that catalyse reactions in the living cell. The simplest situation is that they transform substrate into product (by way of a conformational change) according to the reaction scheme

$$S + E \underset{k_-}{\overset{k_+}{\rightleftharpoons}} C \overset{k}{\to} P + E. \tag{2.1.1}$$

Note that we assume that the reaction $C \to P + E$ is irreversible. This makes the algebra simpler, and isn't too unrealistic for many reactions. The C stands for "complex", a molecule composed of two smaller ones. The meaning of the other symbols is, hopefully, evident.

The Law of Mass Action asserts that the rate at which a reaction proceeds is proportional to the product of the concentrations of the substances that are involved in the reaction. The k's are the constants of proportionality. If we denote the concentrations by small characters, the reaction scheme directly translates into the ODE system

$$\frac{ds}{dt} = -k_+ se + k_- c, \tag{2.1.2}$$

$$\frac{de}{dt} = -k_+ se + k_- c + kc, \tag{2.1.3}$$

$$\frac{dc}{dt} = k_+ se - k_- c - kc, \tag{2.1.4}$$

$$\frac{dp}{dt} = kc. \tag{2.1.5}$$

The atoms that constitute the enzyme also occur in C. The conservation of these atoms is reflected in $\frac{d}{dt}(e+c)=0$. So we can think of molecules that can jump back and forth between two states, the "free" state E and the "bound" state C. The rate (=probability per unit of time) at which a particle in the E state jumps to the C state is k_+s and the rate at which a particle in the C state jumps to the E state is $k_- + k$. If s would be constant, these rates would be constant too. For the *linear* system

$$\frac{d}{dt}\begin{pmatrix} e \\ c \end{pmatrix} = M \begin{pmatrix} e \\ c \end{pmatrix}, \tag{2.1.6}$$

with matrix

$$M = \begin{pmatrix} -k_+s & k_- + k \\ k_+s & -(k_- + k) \end{pmatrix} \tag{2.1.7}$$

we can interpret e and c in several ways:

- e is the probability that a particular particle is in state E and c is the complementary probability that it is in state C. We then normalise e and c such that $e+c=1$.
- e is the fraction of the particles that is in state E and c is the fraction of particles that is in state C. Again, we require that $e+c=1$.
- e and c are, as in (2.1.2), concentrations. The sum $e+c$ is constant in time, with the precise value determined by the initial condition.

To go from the third to the second, just consider $\frac{e}{e+c}$ and $\frac{c}{e+c}$. To go from the first to the second, just recall the Law of Large Numbers from probability theory. The point is that, by considering s as a given/prescribed quantity, we achieved that the particles are *independent* of one another, which is reflected in the fact that (2.1.6) is *linear*.

The matrix M has eigenvalue zero (as a direct consequence of the underlying conservation law). The second eigenvalue is $-k_+s - k_- - k$ (use that the sum of the two eigenvalues equals the trace of M and that one eigenvalue is zero). Since the second eigenvalue is negative, the solution of (2.1.7) converges for $t \to \infty$ to an eigenvector of M corresponding to eigenvalue zero. One such eigenvector is

$$\begin{pmatrix} e \\ c \end{pmatrix} = \begin{pmatrix} k_- + k \\ k_+s \end{pmatrix}, \tag{2.1.8}$$

and all others are a multiple of this one. We conclude that for a given constant value of s,

$$\begin{pmatrix} e(t) \\ c(t) \end{pmatrix} \overset{t\to\infty}{\longrightarrow} \frac{e_{\text{tot}}}{k_+s + k_- + k} \begin{pmatrix} k_- + k \\ k_+s \end{pmatrix}, \tag{2.1.9}$$

where $e_{\text{tot}} := e(0) + c(0)$. The difference between the left and the right hand side of (2.1.9) is bounded by a constant times

$$e^{-(k_+s+k_-+k)t}.$$

For that reason we say that the *time scale* of convergence is

$$(k_+s + k_- + k)^{-1}. \tag{2.1.10}$$

If we now substitute the right hand side of (2.1.9) for e and c in the equation for $\frac{ds}{dt}$ in (2.1.2), we obtain

$$\frac{ds}{dt} = -\frac{k_+ke_{\text{tot}}}{k_+s + k_- + k}s. \tag{2.1.11}$$

So as a consistency requirement for this approach we have that the time scale

$$\left(\frac{k_+ke_{\text{tot}}}{k_+s + k_- + k}\right)^{-1} \tag{2.1.12}$$

of changes in s according to (2.1.11), should be much longer than the time scale given by (2.1.10):

$$k_+ s + k_- + k \gg \frac{k_+ k e_{\text{tot}}}{k_+ s + k_- + k}. \tag{2.1.13}$$

In addition, we should provide (2.1.11) with an initial condition. We would like simply to put $s(0)$ equal to the true initial substrate concentration. Often an experiment is started by adding enzyme to substrate, so with no complex present at time zero. Then initially,

$$\frac{ds}{dt} \approx -k_+ e s,$$

and changes in s occur at the time scale

$$(k_+ e_{\text{tot}})^{-1}. \tag{2.1.14}$$

So if we require that

$$k_+ s(0) + k_- + k \gg k_+ e_{\text{tot}}, \tag{2.1.15}$$

then s hardly changes while e and c converge as described by (2.1.9). Next, note that (2.1.15) guarantees that (2.1.13) holds and that, accordingly, it suffices to require (2.1.15) or, equivalently,

$$\frac{k_+ e_{\text{tot}}}{k_+ s(0) + k_- + k} \ll 1. \tag{2.1.16}$$

HINT: To verify that (2.1.15) implies (2.1.13), first note that

$$0 < \frac{k}{k_+ s + k_- k} < 1$$

since all terms are positive. So if we multiply the right hand side of (2.1.15) by $k/(k_+ s + k_- + k)$ the left hand side is still much bigger.
See (Schnell and Maini, 2000), and the references given there, for justification as well as variants.

EXERCISE 2.1.1. The atoms that constitute the substrate also occur in C and in P. Formulate a conservation law and check that it is incorporated in (2.1.2)–(2.1.5).

EXERCISE 2.1.1. Check that in the approximation that leads to (2.1.11) we have

$$\frac{dp}{dt} = -\frac{ds}{dt}. \tag{2.1.17}$$

So the velocity at which substrate is transformed into product is initially given by

$$V = \frac{e_{\text{tot}} k k_+ s(0)}{k_+ s(0) + k_- + k}. \tag{2.1.18}$$

Rewrite this such that $\frac{1}{V}$ is a linear function of $\frac{1}{s(0)}$. Which reaction parameters can be estimated by measuring V as a function of the tunable initial substrate concentration?

This function V in (2.1.18) is the reason why one often uses

$$\frac{ds}{dt} = -\frac{V_m s}{K_m + s} \tag{2.1.19}$$

to model the change in concentration of substrate in enzyme kinetics. Equation (2.1.19) is often called the *Michaelis-Menten rate equation*. Here, $V_m = e_{\text{tot}} k$, and $K_m = (k_- + k)/k_+$, the *Michaelis constant*.

2.2 Scaling

Compared to the length of a human life, 100,000 milliseconds isn't very long, but 100,000 years is. Words like "small" or "large" are dangerous when we talk about quantities that carry a physical dimension, since the actual numbers depend on the choice of units. One way to justify a formal QSSA is to go through the systematic procedure of *scaling*, i.e., of reformulating the equations in terms of *non-dimensional* variables, and to identify a *small parameter* while doing so. We now illustrate this procedure in the context of (a reduced version of) system (2.1.2)–(2.1.5).

As noted above, changes in s at the start occur at the time scale given by (2.1.14), and this motivates us to choose

$$t = \frac{1}{k_+ e_0} t^*, \tag{2.2.1}$$

where e_0 is the initial condition for e. A natural scale for substrate is provided by the initial concentration, so we choose

$$s(t) = s(0) s^*(t^*). \tag{2.2.2}$$

As $c(t)$ is bounded by e_{tot}, we choose

$$c(t) = e_{\text{tot}} c^*(t^*), \tag{2.2.3}$$

and as a consequence

$$e(t) = e_{\text{tot}}(1 - c^*(t^*)). \tag{2.2.4}$$

Now, noting that

$$\frac{d}{dt} = \frac{dt^*}{dt}\frac{d}{dt^*} = k_+ e_0 \frac{d}{dt^*}, \tag{2.2.5}$$

we substitute all this into the first and the third equation of (2.1.2)–(2.1.5) and subsequently divide both sides of both equations by $k_+ s(0) e_{\text{tot}}$. The result is

$$\frac{ds^*}{dt^*} = -s^* + \left(\frac{k_-}{k_+ s(0)} + s^*\right) c^*, \tag{2.2.6}$$

$$\frac{dc^*}{dt^*} = \frac{s(0)}{e_{\text{tot}}} \left(s^* - \left(\frac{k_- + k}{k_+ s(0)} + s^*\right) c^*\right). \tag{2.2.7}$$

We next rename the three compound parameters that figure in these equations:

$$\varepsilon = \frac{e_{\text{tot}}}{s(0)}, \tag{2.2.8}$$

$$\kappa = \frac{k_- + k}{k_+ s(0)}, \tag{2.2.9}$$

$$\lambda = \frac{k}{k_+ s(0)}, \tag{2.2.10}$$

drop the stars and multiply the second equation by ε. This leads us to

$$\begin{aligned} \frac{ds}{dt} &= -s + (\kappa - \lambda + s)c, \\ \varepsilon \frac{dc}{dt} &= s - (\kappa + s)c. \end{aligned} \tag{2.2.11}$$

Note that ε is the initial ratio of enzyme and substrate molecules. Often an experiment consists of adding a little bit of enzyme to an excess of substrate, leading to a small value of ε. If ε is small and $c \not\approx \frac{s}{\kappa + s}$, then necessarily $\frac{dc}{dt}$ is very large, so c changes rapidly, i.e., at the time scale ε. For fixed s, the equation

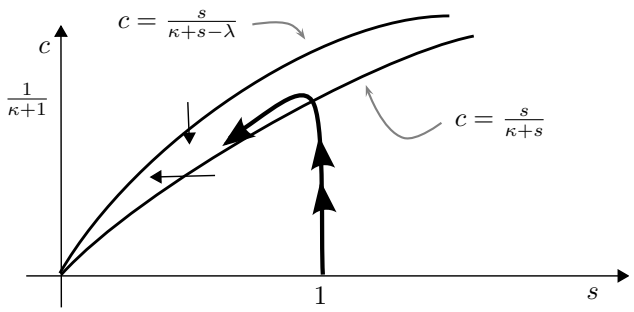


Figure 2.1. Dynamics of (2.2.11). The solution shoots up from the s-axis until it approaches $c = 1/(\kappa+1)$, after which it is trapped between the isoclines and converges towards the origin.

for c is linear and we see immediately that c converges to $\frac{s}{\kappa+s}$. If we substitute $c = \frac{s}{\kappa+s}$ in the equation for s we obtain

$$\frac{ds}{dt} = -\frac{\lambda s}{\kappa + s}, \tag{2.2.12}$$

as the equation describing the changes in s at the $\mathcal{O}(1)$ time scale. As initial condition we put $s(0) = 1$.

We can zoom in on the rapid change in c right at the start (when indeed $c \neq \frac{s}{\kappa+s}$), by introducing

$$t = \varepsilon t^*. \tag{2.2.13}$$

(The recycling of the $*$ should not confuse you! Note that t^* should be large to have a value of t that isn't small.) From

$$\frac{ds}{dt^*} = \varepsilon(-s + (\kappa - \lambda + s)c) \approx 0,$$

we deduce that $s(t^*) \approx s(0) = 1$. When we substitute this into

$$\frac{dc}{dt^*} = s - (\kappa + s)c,$$

we find

$$c(t^*) = \frac{1}{\kappa + 1}\left(1 - e^{-(\kappa+1)t^*}\right) \overset{t^*\to\infty}{\longrightarrow} \frac{1}{\kappa + 1}.$$

The sketchy phase plane portrait corresponding to (2.2.11) is depicted in Figure 2.1. A systematic method to analyse systems containing a small parameter is the method of *matched asymptotic expansions* (see, e.g., (Mishchenko and Rosov, 1980; Grasman, 1987; Kevorkian and Cole, 1996; Verhulst, 2005)). The key idea is to construct solutions in the form of a power series in ε, both for (2.2.11) and for the system

$$\frac{ds}{dt^*} = \varepsilon(-s + (\kappa - \lambda + s)c), \tag{2.2.14}$$

$$\frac{dc}{dt^*} = s - (\kappa + s)c. \tag{2.2.15}$$

(The first is called the *outer solution*, and the second the *inner solution*.) Unknown integration constants are determined by matching these expansions, i.e., by relating the limit $t \to 0$ in the outer solution to the limit $t^* \to \infty$ in the inner solution. In the present very simple case, this would amount to choosing $s(0) = 1$ as the initial condition for (2.2.12). See (Britton, 2003) for a fully worked treatment of the above problem.

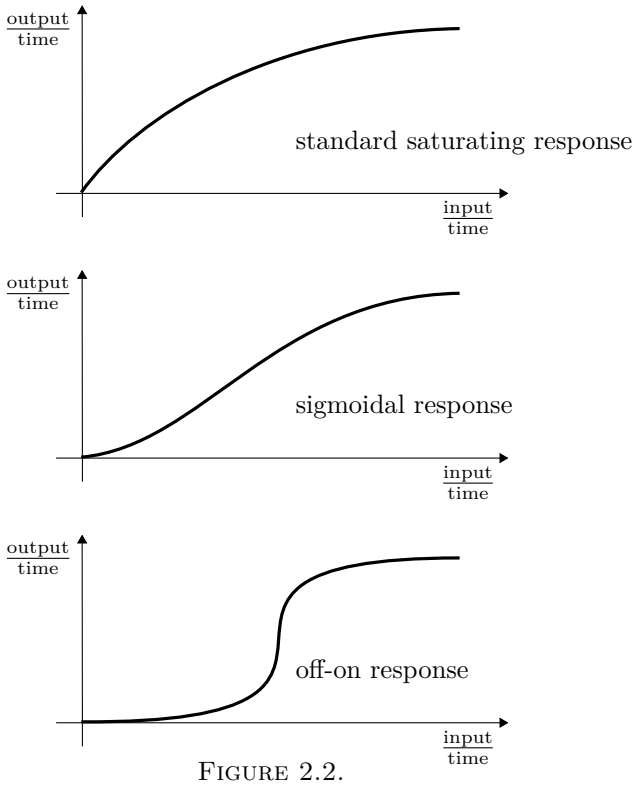


FIGURE 2.2.

2.3 Cooperative reactions and the sigmoidal response

In Section 2.1 our emphasis was on *processes* that were either slow or fast, while in Section 2.2 we focussed on *variables* that either change slowly or change fast. The first is the more mechanistic point of view, the second is more amenable to analysis. It may require ingenious transformations to translate the first into the second, see (Lee and Othmer, 2010).

In (2.1.11), we see that the velocity of the $S \to P$ transformation saturates for large substrate concentrations. This is a very general phenomenon: the output of a "factory" is proportional to input at low input, but is determined by "capacity" (and hence independent of input) for high input.

Sometimes one observes a remarkable special feature: the output accelerates in an intermediate input region. In more picturesque language: sometimes the response is s-shaped and is called sigmoidal. When the acceleration occurs in a narrow region, but is very strong, one has a "switch" between (almost) no production and production at (almost) full speed. The three profiles are sketched in Figure 2.2.

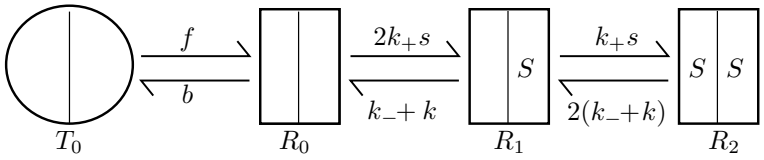
How can a sigmoidal response arise? What kind of biochemical mechanism can be responsible?

Some enzymes consist of several identical subunits (e.g., haemoglobin has four binding sites for O_2). They are call oligomers (cf. the word polymer to indicate that there are very many identical building blocks). Here we consider a dimer, consisting of two subunits. If a protein binds some smaller molecule to one of its subunits, that smaller molecule is often called a ligand (from *ligare* = to bind).

Now assume that each subunit can assume two different spatial conformations, i.e., shapes. And assume that the reactivity to bind a ligand depends on the conformation. Let's go to the extreme: in one of the two conformations the ligand cannot be bound at all.

Finally, let us assume that transitions between the two spatial conformations are concerted, i.e., all subunits switch simultaneously. Included in this assumption is that, if a ligand is bound to a reactive subunit, no switch to the inactive state can occur for any of the subunits. Now test your physical intuition: can you imagine how acceleration might arise? Whether or not you can, let's see what we can learn from a mathematical analysis.

Let us denote the inactive state by the symbol T (for "tight") and the reactive state by R. Let's use a subscript to denote the number of ligands bound. The considerations above can be summarised in the reaction scheme



and translated into the ODE system

$$\frac{dT_0}{dt} = -fT_0 + bR_0, \tag{2.3.1}$$

$$\frac{dR_0}{dt} = +fT_0 - bR_0 - 2k_+sR_0 + (k_- + k)R_1, \tag{2.3.2}$$

$$\frac{dR_1}{dt} = +2k_+sR_0 - (k_- + k)R_1 - k_+sR_1 + 2(k_- + k)R_2, \tag{2.3.3}$$

$$\frac{dR_2}{dt} = +k_+sR_1 - 2(k_- + k)R_2, \tag{2.3.4}$$

$$\frac{ds}{dt} = -2k_+sR_0 + k_-R_1 - k_+sR_1 + 2k_-R_2, \tag{2.3.5}$$

$$\frac{dp}{dt} = +kR_1 + 2kR_2. \tag{2.3.6}$$

Assume, for the time being, that s is constant. The enzyme molecule can be in four states. The four vector (T_0, R_0, R_1, R_2) satisfies a *linear* ODE with matrix M given by

$$M = \begin{pmatrix} -f & b & 0 & 0 \\ f & -b - 2k_+s & k_- + k & 0 \\ 0 & 2k_+s & -(k_- + k_+s + k) & +2(k_- + k) \\ 0 & 0 & k_+s & -2(k_- + k) \end{pmatrix}. \tag{2.3.7}$$

This matrix has an eigenvalue zero (how can we be so sure?). The corresponding eigenvector is easily computed by expressing the fourth component in the third, the third in the second, etcetera, while exploiting that terms at some level re-occur one level higher (this reflects the order structure in the reaction scheme: we can order the states such that only transitions to adjacent states are possible). If we normalise such that the sum of the components equals one, each component corresponds to the fraction of the enzyme molecules in the corresponding state:

$$\frac{1}{1 + \frac{f}{b}\left(1 + \frac{2k_+s}{k_- + k}\left(1 + \frac{k_+s}{2(k_- + k)}\right)\right)} \begin{pmatrix} 1 \\ \frac{f}{b} \\ \frac{f}{b}\frac{2k_+s}{k_- + k} \\ \frac{f}{b}\frac{2k_+s}{k_- + k}\frac{k_+s}{2(k_- + k)} \end{pmatrix}. \tag{2.3.8}$$

But are we sure that the solutions of the linear ODE converge to an eigenvector corresponding to eigenvalue zero? In other words, are we sure that all other eigenvalues have negative real part? Yes, we are! The point is that M defined in (2.3.7) is Positive-Off-Diagonal (POD), so that we can use the Perron-Frobenius theorem, discussed in more detail at the end of this section, to arrive at the desired conclusion (see also Graham (1987); Horn and Johnson (1990) for two general references). Equivalently, we can refer to the general theory of continuous time Markov Chains (Kemeny and Snell, 1963).

If we let Y denote the *fraction* of the subunits to which ligand is bound, then Y equals the sum of one half of the third component and the fourth component, so

$$Y = \frac{\frac{f}{b}\frac{k_+ s}{k_- + k}\left(1 + \frac{k_+ s}{k_- + k}\right)}{1 + \frac{f}{b}\left(1 + \frac{2k_+ s}{k_- + k}\left(1 + \frac{k_+ s}{2(k_- + k)}\right)\right)}. \tag{2.3.9}$$

Now, assuming an excess of substrate relative to the enzyme, let's look at changes in s and p at the slow time scale. In the QSSA the last two equations of (2.3.1) can be written as

$$\frac{dp}{dt} = 2kYe_{\text{tot}} = -\frac{ds}{dt} \tag{2.3.10}$$

(can you interpret the middle term? is this what you would expect?). We can conclude that the way in which the velocity of the $S \to P$ transformation depends on the substrate concentration can be read off from the formula (2.3.9) for Y.

In terms of

$$\tilde{s} := \frac{k_+ s}{k_- + k},$$

we have

$$Y = \frac{f}{b}\frac{\tilde{s}(1+\tilde{s})}{1 + \frac{f}{b}(1+\tilde{s})^2},$$

and one easily verifies the sigmoidal shape of this function. So as a final conclusion, we have that such a sigmoidal response may arise by the *cooperative* effect that the occupation of a first binding site prevents that other, as yet still free sites, switch back to the inactive state. Was this indeed what you anticipated when we asked you to test your physical intuition?

[This section is very much inspired by Chapter 4 of Segel (1984)]

2.3.1 Perron-Frobenius theory

Consider the *cone* of positive vectors v, i.e., $v_i \geq 0$ for all i. Let K be a positive square matrix, i.e., $K_{ij} \geq 0$. K leaves the cone invariant, since $Kv \geq 0$. The norm for our vectors is chosen to be $\|v\| = \sum_{i=1}^n |v_i|$; for K we choose the sup-norm,

$$\|K\| = \sup_{\|v\| \neq 0} \frac{\|Kv\|}{\|v\|} = \sup_{\|v\|=1} \|Kv\|.$$

Let the spectral radius $\rho(K)$ be defined by

$$\rho(K) = \lim_{m \to \infty} \|K^m\|^{1/m}.$$

Theorem 2.3.1*:* Let $K \geq 0$. Then

- $\rho(K)$ is an eigenvalue of K;
- For each eigenvalue λ we have $|\lambda| \leq \rho(K)$.

Corollary 2.3.2*:* $\rho(K)$ is the dominant eigenvalue. If $\rho(K)$ is strictly dominant (meaning $|\lambda| < \rho(K)$ if $\lambda \neq \rho(K)$), and $\rho(K)$ is algebraicly simple, then there exists a constant $c(v)$ depending on v such that

$$K^m v = c(v)(\rho(K))^m \psi + o(\rho(K)^m) \quad m \to \infty$$

where ψ is the eigenvector corresponding to $\rho(K)$.

Definition 2.3.3*:* A positive matrix K is called *irreducible* if for all i, j there exists an $m = m(i, j) \geq 1$ such that $(K^m)_{ij} > 0$.

If we consider a directed graph with n nodes, in which directed edges are drawn from point i to point j if and only if $K_{ij} > 0$, then irreducibility of K is equivalent to requiring that there exists a path from any i to and j in this graph.

Definition 2.3.4*:* A positive matrix K is called *primitive* or *aperiodic* there exists an m such that $(K^m)_{ij} > 0$ for i, j.

With these preliminaries, we can now state the Perron-Frobenius theorem.

Theorem 2.3.5*:* Let $K \geq 0$ be primitive, with $R_0 := \rho(K)$ as one of its eigenvalues. Then

- R_0 is strictly dominant;
- the left and right eigenvectors corresponding to R_0 have strictly positive components (hence $c(v) > 0$ if $v \geq 0$ and $v \neq 0$);
- R_0 is algebraicly simple;
- no other eigenvalue has a positive eigenvector.

For irreducible, but not primitive, matrices we still have the following.

Theorem 2.3.6*:* Let $K \geq 0$ be irreducible, but not primitive. Then all but the first point from the previous theorem still hold, but are eigenvalues λ with $|\lambda| = R_0$ forming roots of unity, i.e. $\lambda^k = 1$ for some $k \leq n$.

As an example of this latter case, consider

$$K = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Finally, let us briefly see in what precise way the Perron-Frobenius came to our rescue in the beginning of this paragraph. The matrix M in (2.3.7) is positive off-diagonal. By conservation of mass, the column vectors add up to zero, so that M has an eigenvalue zero, and we also immediately see that mass conservation is reflected in a left eigenvector $(1, 1, 1, 1)$ corresponding to eigenvalue 0. Since this left eigenvector is strictly positive, the 0 eigenvalue must be the dominant eigenvalue by the P-F theorem, and all other eigenvalues must have negative real part.

2.4 The force of infection in populations of variable size

If an infectious individual makes "contact" with a susceptible individual, the infectious agent is transmitted with a certain probability. If we simplify reality and assume that the just infected individual is infectious right away, we have the reaction scheme

$$S + I \xrightarrow{\beta} 2I,$$

where $\beta = cp$, with p the probability of transmission and c such that an infectious individual makes contact with susceptible individuals at rate cS. When we take the point of view of the susceptible, this is usually formulated as: the force of infection equals βI. Here the force of infection means "the probability per unit of time of becoming infected". Yet another common formulation is: the incidence (the number of new cases per unit time) equals βSI.

The underlying assumption here is that the number of contacts that an individual has per unit of time is proportional to population density. If you contrast commuter trains in the Tokyo district with pedestrians on a small rural road, this seems a fair assumption. But what if "contact" means sexual contact? Shouldn't, in

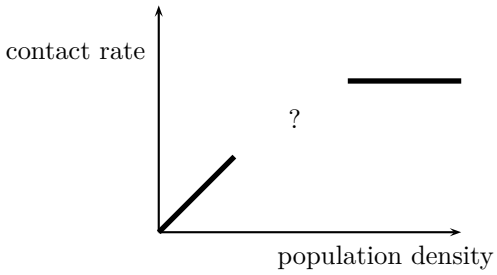


FIGURE 2.3.

this case, satisfaction lead to a saturating contact intensity? Or, if not satisfaction, then lack of time?

If we assume that individuals make contact at a fixed rate c (independent of population density), the force of infection equals $\beta\frac{I}{N}$, where N is the total population density (and so I/N can be interpreted as the chance that a contact is with an infectious individual; note that we assume throughout this section that the infection status has no influence whatsoever on the contact process).

When the total population density N is a constant, the difference between βI and $\beta\frac{I}{N}$ is just a difference in the interpretation of β. But what if N is variable, or if we want to compare various host populations with rather different values of N? So far, we covered the linear part of the contact response (βI), which probably works best at low densities, and the fully saturated part ($\beta\frac{I}{N}$), which approximates the process best at high densities (see Figure 2.3). Can we connect these two parts on the basis of a mechanistic submodel? The following positive answer is based on Heesterbeek and Metz (1993), where also variants appropriate for structured populations are considered.

The key idea is to introduce pair formation and dissolution and to restrict transmission to pairs. The average duration of partnerships then sets limits to contact intensity.

So let us distinguish *singles* (those individuals that at the time considered are not engaged in a pairwise contact) from *couples/pairs*. When the first have density $X(t)$, and the latter density $P(t)$, then

$$N(t) = X(t) + 2P(t), \tag{2.4.1}$$

since a pair consists of two individuals. If σ denotes the separation rate and if a single finds a partner at rate ρX (so at a rate proportional to the supply of candidates), then

$$\frac{dX}{dt} = -\rho X^2 + 2\sigma P, \tag{2.4.2}$$

$$\frac{dP}{dt} = \frac{1}{2}\rho X^2 - \sigma P. \tag{2.4.3}$$

Clearly this is a caricatural description of both pair formation and separation; remarkably, however, a description of the duration of relationships by an exponential distribution fits the data rather well.

EXERCISE 2.4.1. Verify that, with $\nu := \frac{\rho}{\sigma}$,

$$\bar{X} = \frac{\sqrt{1+4\nu N}-1}{2\nu}, \quad \bar{P} = \frac{1+2\nu N-\sqrt{1+4\nu N}}{4\nu}, \tag{2.4.4}$$

defines the unique steady state of (2.4.2) and that this steady state is globally asymptotically stable.

Next, on top of the distinction between singles and pairs, we introduce the distinction between susceptible and infectious individuals (for simplicity we shall assume that infected individuals do not lose their infectiousness). Let S_1 denote the density of susceptible singles, S_2 the density of pairs consisting of two susceptibles and let I_1 and I_2 be similarly defined. Let M (for "mixed") denote the density of pairs consisting of a susceptible and an infective.

Exercise 2.4.2. Interpret all terms in the ODE system

$$\frac{dS_1}{dt} = -\rho S_1(S_1 + I_1) + 2\sigma S_2 + \sigma M, \tag{2.4.5}$$

$$\frac{dI_1}{dt} = -\rho I_1(S_1 + I_1) + 2\sigma I_2 + \sigma M, \tag{2.4.6}$$

$$\frac{dS_2}{dt} = \frac{1}{2}\rho S_1^2 - \sigma S_2, \tag{2.4.7}$$

$$\frac{dM}{dt} = \rho S_1 I_1 - \sigma M - \beta M, \tag{2.4.8}$$

$$\frac{dI_2}{dt} = \frac{1}{2}\rho I_1^2 - \sigma I_2 + \beta M. \tag{2.4.9}$$

and check whether the pair formation and dissolution rules were correctly incorporated. What is the underlying assumption concerning transmission?

Exercise 2.4.3. The process of pair formation and dissolution proceeds *independently*, i.e., the S-I distinction has absolutely no influence at all on it. Shouldn't that allow us to recover (2.4.2) from (2.4.5)? Define X and P in terms of S_1, S_2, I_1, I_2, and M and check that (2.4.2) holds!

Exercise 2.4.4. The transmission process does *not* proceed independently of the single-pair distinction: on the contrary, it is restricted to mixed pairs. Define the density of susceptibles S and the density of infectives I in terms of S_1, S_2, I_1, I_2, and M, and show that

$$\frac{dS}{dt} = -\beta M, \tag{2.4.10}$$

$$\frac{dI}{dt} = \beta M, \tag{2.4.11}$$

exactly as one would expect (or wouldn't you?).

If β is very small relative to both ρ and σ, the variables S and I will change slowly relative to the speed at which the pair formation/dissolution process equilibrates. The idea of the time scale argument is now as follows: consider, for fixed S and I, the variant of (2.4.5)–(2.4.9) obtained by deleting the βM terms at the right hand side; compute how the limiting value of M depends on S and I; substitute the result into (2.4.10) to describe the slow changes in S and I. So the QSSA concerns (2.4.5)–(2.4.9) and involves omitting βM terms. Note that the system (2.4.2) is nonlinear, reflecting the dependence in the pair formation process. So we cannot rely on the spectral theory of POD matrices or on the theory of Markov processes. Yet the solution can be found by way of a simple combinatorial/probabilistic argument.

Exercise 2.4.5. Verify that the quasi-steady-state of (2.4.5)–(2.4.9) is given by

$$\begin{gathered} \bar{S}_1 = \tfrac{S}{N}\bar{X}, \qquad \bar{I}_1 = \tfrac{I}{N}\bar{X}, \\ \bar{S}_2 = \left(\tfrac{S}{N}\right)^2 \bar{P}, \quad \bar{M} = 2\tfrac{S}{N}\tfrac{I}{N}\bar{P}, \quad \bar{I}_2 = \left(\tfrac{I}{N}\right)^2 \bar{P}, \end{gathered} \tag{2.4.12}$$

and explain the logic behind these expressions. (One can check the stability by exploiting that S and I are constant and that X and P approach limits to reduce the five dimensional system to, essentially, one equation.)

EXERCISE 2.4.6. Write (2.4.10) in the form

$$\frac{dS}{dt} = -\beta C(N)\frac{SI}{N} = -\frac{dI}{dt}, \tag{2.4.13}$$

and check that

(i) $C(N) > 0$ for $N > 0$,
(ii) C is nondecreasing,
(iii) $C(N)$ is approximately linear in N for small N. More precisely, $C(N) = 2\nu N + \mathcal{O}(N^2)$ for $N \to 0$,
(iv) $C(N)$ is approximately constant for large N.

The function $C(N)$ studied in the last exercise thus provides us with a connection between the linear βI and saturated $\beta I/N$ responses with which we started, as illustrated in Figure 2.3. We refer to Section 10.2 of Diekmann and Heesterbeek (2000) for elaborations, remarks on generalisations and additional exercises.

2.5 Excitability

Cell membranes contain ion channels through which in particular potassium and sodium are pumped, enabling communication between neighbouring cells, and indeed signal transduction. These channels contain gates that open and close in response to voltage changes, giving rise to excitability, as we will see. Simply put, the current I through a membrane containing N channels may be specified by

$$I = Ng(V,t)\phi(V), \tag{2.5.1}$$

where $\phi(V)$ describes how the current through a single open channel depends on the voltage V, and $g(V,t)$ is the fraction of channels open at time t and voltage V. The simplest assumption for $\phi(V)$ is that the current depends linearly on the voltage, and is zero at a particular voltage, called the *Nernst potential*, V_N. So let us assume that

$$\phi(V) = V - V_N.$$

In general, V_N will be different for different ions. Let us consider channels with just one gate, that may be open or closed. The fraction of open channels g is assumed to change according to the 'reaction scheme'

$$\mathrm{C} \quad \underset{\beta(V)}{\overset{\alpha(V)}{\rightleftharpoons}} \quad \mathrm{O}$$

where C stands for the closed state of a channel, and O for the open state. This translates into an equation as follows,

$$\frac{dg}{dt} = \alpha(V(t))(1-g) - \beta(V(t))g. \tag{2.5.2}$$

In a voltage clamp experimental setup, the voltage is kept constant, and (2.5.2) simplifies to $\frac{dg}{dt} = \alpha(1-g) - \beta g$. We rewrite this as

$$\tau_g \frac{dg}{dt} = \frac{\alpha}{\alpha+\beta} - g$$

where $\tau_g = 1/(\alpha+\beta)$, to highlight that g converges to $\alpha/(\alpha+\beta)$ at the time scale τ_g. For channels with multiple subunits, we have to distinguish the different possibilities, analogous to our enzyme with multiple binding sites in Section 2.3. Let us consider a channel with two subunits, with possible states S_{00} (both subunits

closed), S_{10} (one of the two subunits open, so there are two possibilities here), and S_{11} (both gates closed), with dynamics specified by the scheme

$$S_{00} \underset{\beta}{\overset{2\alpha}{\rightleftharpoons}} S_{10} \underset{2\beta}{\overset{\alpha}{\rightleftharpoons}} S_{11}.$$

Denoting the fractions of channels with zero, one or two subunits open by x_0, x_1 and x_2 respectively, we have

$$\frac{dx_0}{dt} = \beta x_1 - 2\alpha x_0, \tag{2.5.3}$$

$$\frac{dx_2}{dt} = \alpha x_1 - 2\beta x_2. \tag{2.5.4}$$

The dynamics of x_1 is prescribed by the requirement $x_0 + x_1 + x_2 = 1$. If we denote the total fraction of open subunits by n, then we know that $x_0 = (1-n)^2$, $x_1 = 2n(1-n)$, $x_2 = n^2$ if we assume that subunits open and close independently from one another. This change of variables yields

$$\frac{dn}{dt} = \alpha(1-n) - \beta n, \tag{2.5.5}$$

so again in the form (2.5.2) for a channel with a single gate. This is true more generally: if there are k subunits in a channel, then $x_k = n^k$. In the Hodgkin-Huxley equations, it is assumed that a channel only pumps ions through its subunits if all the subunits are open, giving a term n^4 for a four subunit K^+ channel.

For Na^+ channels, there is a further complication, since these may be in either an active or inactive state. We model this using two kinds of subunits, three m units for the gate and one h unit for the activation/inactivation. The activation process is characterized by activation and inactivation rates γ and δ. We now have to keep track of double the number of channel types (open/closed, active/inactive). Transitions between the types follows the scheme

$$\begin{array}{ccccccc}
S_{00} & \underset{\beta}{\overset{3\alpha}{\rightleftharpoons}} & S_{10} & \underset{2\beta}{\overset{2\alpha}{\rightleftharpoons}} & S_{20} & \underset{\beta}{\overset{\alpha}{\rightleftharpoons}} & S_{30} \\
\delta \uparrow\downarrow \gamma & & \delta \uparrow\downarrow \gamma & & \delta \uparrow\downarrow \gamma & & \delta \uparrow\downarrow \gamma \\
S_{01} & \underset{\beta}{\overset{3\alpha}{\rightleftharpoons}} & S_{11} & \underset{2\beta}{\overset{2\alpha}{\rightleftharpoons}} & S_{21} & \underset{3\beta}{\overset{\alpha}{\rightleftharpoons}} & S_{31}
\end{array}$$

As before, we use variables x_{ij} for the fraction of channels in state S_{ij}. The system in x_{ij} may be simplified by changing variables as before, e.g. by setting $x_{00} = (1-m)^3(1-h)$ for the first variable. The result is two linear differential equations for m and h,

$$\frac{dm}{dt} = \alpha(1-m) - \beta m, \tag{2.5.6}$$

$$\frac{dh}{dt} = \gamma(1-h) - \delta h \tag{2.5.7}$$

A channel is now assumed to let through Na^+ ions only if it is both active and all of its channels are open, resulting in a term m^3h in the Hodgkin-Huxley equations.

To complete the Hodgkin-Huxley equations, we need a little electrostatics. The capacitance of an insulator (such as a membrane) is defined as the ratio of the charge across the capacitor to the voltage potential required to hold the charge, i.e.,

$$C_m = \frac{Q}{V}.$$

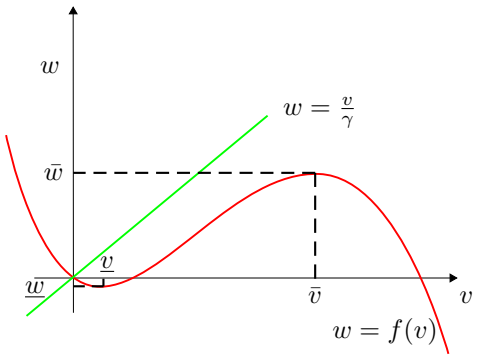


FIGURE 2.4.

Since the current I is equal to $\frac{dQ}{dt}$, we find that an ion current across a membrane gives rise to a voltage change according to

$$C_m \frac{dV}{dt} = I.$$

Using (2.5.1) and the assumptions on the potassium and sodium channels discussed above, we find

$$C_m \frac{dV}{dt} = -g_{\mathrm{K}} n^4 (V - V_{\mathrm{K}}) - g_{\mathrm{Na}} m^3 h (V - V_{\mathrm{Na}}) - g_L (V - V_L) + I_{\mathrm{app}}. \tag{2.5.8}$$

Here, the $g_L(V - V_L)$ is a term in which the other currents, for instance Cl^- ions, are lumped into a *leakage current*, and V_{K}, V_{Na} and V_L are the Nernst potentials for the respective (lumped) ions. Note the n^4 and m^3h terms in this equation. The complete Hodgkin-Huxley model is now given by (2.5.5), (2.5.6), (2.5.7) and (2.5.8). The only thing still missing in this model is the dependence of α and β on the voltage (which differs for each ion type). If you have become interested in the details, we refer to (Cronin, 1987; Keener and Sneyd, 2009) for a more thorough discussion of the Hodgkin-Huxley model. The original paper by (Hodgkin and Huxley, 1952) is also one of the most outstanding examples of experimental and theoretical work in mathematical physiology, and was awarded a Nobel Prize.

From this modelling discussion, it is clear that the system of equations that form the Hodgkin-Huxley model is too complicated to allow direct analysis. Consider instead the so-called *Fitzhugh-Nagumo* system

$$\varepsilon \frac{dv}{dt} = f(v) - w, \tag{2.5.9}$$

$$\frac{dw}{dt} = v - \gamma w, \tag{2.5.10}$$

with f a *cubic*, specified graphically in Figure 2.4. These equations were proposed as a somewhat caricatural simplification of the Hodgkin-Huxley equations, and is much more accessible.

Clearly, $(v, w) = (0, 0)$ is an equilibrium, which we call the rest state. The Jacobian matrix in this point has the sign pattern

$$\begin{pmatrix} - & - \\ + & - \end{pmatrix}$$

so the trace is negative and the determinant positive, implying that the rest state is *locally* asymptotically stable. We emphasise the word "locally", because, as we are going to show, the domain of *immediate* attraction is rather small. Yet the

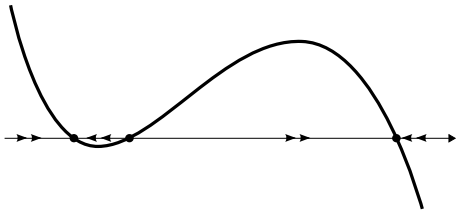


Figure 2.5.

domain of *ultimate* attraction is very large (presumably the entire plane). The key feature is that perturbations of the rest state may trigger a *large excursion* before the system settles back into the rest state. This is called *excitability* and it is a characteristic feature of the nerve pulse.

When $\varepsilon \ll 1$, the variable v will change fast as long as $f(v) \neq w$. Note that v will increase below the graph of $w = f(v)$ and decrease above that graph. Hence fixing w, we find the dynamics indicated in Figure 2.5. Slow changes in w are governed by

$$\dot{w} = f^{-1}(w) - \gamma w,$$

but one should keep in mind that f^{-1} is, on some of its domain, multivalued. In particular it follows that w increases to the right of the line $w = \frac{v}{\gamma}$ and decreases to the left of that line.

The function f is decreasing for $-\infty < v < \underline{v}$, where $\underline{v}$ is such that f assumes its local minimum $\underline{w}$ for $v = \underline{v}$, increasing for $\underline{v} < v < \bar{v}$, where $\bar{v}$ is such that f assumes its local maximum $\bar{w}$ for $v = \bar{v}$, and decreasing again for $\bar{v} < v < \infty$. Let us call the corresponding inverse functions f_-^{-1}, f_0^{-1}, and f_+^{-1}.

A key difference with the examples studied so far is that there are *two* phases of fast dynamics. Indeed, while w increases according to

$$\dot{w} = f_+^{-1}(w) - \gamma w,$$

it will come close to $\bar{w}$, where $v = f_+^{-1}(w)$ ceases to be an attractor for the fast v-dynamics (and in fact also ceases to exist). This then triggers a fast crossing to the branch $v = f_-^{-1}(w)$, after which the slow dynamics resumes with a gradual decay of w according to

$$\dot{w} = f_-^{-1}(w) - \gamma w,$$

towards the steady state $w = 0$.

Exercise 2.5.1. Verify that the orbit in the (v, w) plane leads to a graph of v as a function of time as depicted in Figure 2.6. When interpreting v as the voltage across the membrane surrounding the axon, this corresponds to a single nerve pulse being triggered.

Now let us introduce a positive parameter w_0 and change (2.5.9) into

$$\varepsilon \frac{dv}{dt} = f(v) - w + w_0, \tag{2.5.11}$$

$$\frac{dw}{dt} = v - \gamma w. \tag{2.5.12}$$

(Physiologically, this corresponds to forcing a current through the membrane.)

Once we have increased w_0 such that the steady state lies on the middle branch $v = f_0^{-1}(w)$, it is no longer an attractor. Simple graphical considerations suggest convergence towards a closed orbit in which slow and fast dynamics alternate (see Figure 2.7). Physiologically this corresponds to repetitive firing: the axon generates a never ending train of pulses in response to the applied current (as indeed it does

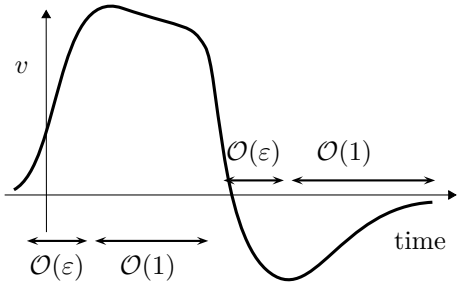


FIGURE 2.6.

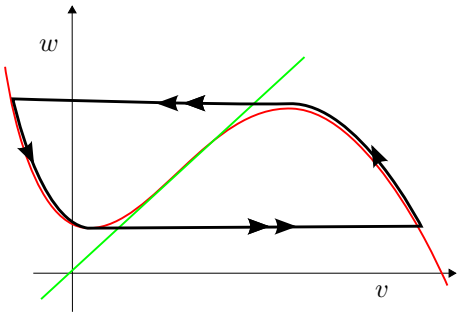


FIGURE 2.7.

in reality!). The mathematical jargon for such periodic behaviour with alternating slow and fast phases is "relaxation oscillation", see (Grasman, 1987).

EXERCISE 2.5.2. Use hand waiving to derive that

$$T = \int_{\underline{w}}^{\bar{w}} \left\{ \frac{1}{f_+^{-1}(w) - w} - \frac{1}{f_-^{-1}(w) - w} \right\} dw + \mathcal{O}(\varepsilon)$$

where T is the period of the oscillation.

With a substantial amount of work one can verify that the transition from a stable rest state to a stable relaxation oscillation with large amplitude is by way of a subcritical Hopf bifurcation. The abrupt major attractor change that occurs when w_0 passes the critical value is sometimes called a "hard" bifurcation (see Figure 2.8).

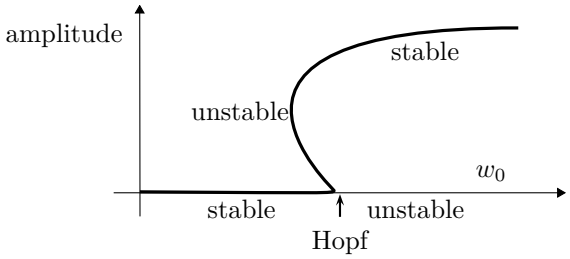


FIGURE 2.8.

2.6 Encore: using the Law of Mass Action to formulate differential equations and using a conservation law to check whether or not you made mistakes

Polymers are large molecules built from identical units, called monomers, in a chain-like manner. Let us denote a polymer consisting of k monomers by C_k. Assume, for simplicity, that changes in the size k can only occur by addition of a monomer or by fragmentation into two parts, one of size $k-1$ and one a monomer. Schematically,

$$C_{k-1} + C_1 \underset{\theta_-}{\overset{\theta_+}{\rightleftharpoons}} C_k$$

Assume, additionally, that $\theta_\pm$ are independent of k.

Exercise 2.6.1. (i) Use the Law of Mass Action to formulate a countable system of ODEs for the concentrations c_k

(ii) Formulate a conservation law and use this law to check your answer to (i).

Chapter 3

Phase plane analysis of prey-predator systems

3.1 The story of d'Ancona and Volterra

During World War I, Italian fishermen were forced to keep their boats ashore of the Adriatic sea, because of the danger of being sunk. With relief they resumed fishing after the war ended. Understandably, it vexed them to discover that the percentage of shark-like predatory fish (which didn't fetch an attractive price) in the haul had increased considerably compared to the pre-war period.

Hearing their complaints, the biologist d'Ancona wondered about an explanation. Unable to produce one, he posed the puzzle to his father in law, a famous mathematician named Volterra, who produced, first of all, the system of two differential equations

$$\frac{dv}{dt} = av - bvp, \tag{3.1.1}$$

$$\frac{dp}{dt} = dvp - cp. \tag{3.1.2}$$

Here, v stands for victim (the prey) and p for predator, and a, b, c and d are non-negative parameters.

Exercise 3.1.1. Explain the modelling assumptions underlying this system. In particular, explain the assumptions underlying the form of the terms.

Exercise 3.1.2. Draw the null-clines (i.e., zero isoclines). Indicate the steady-states by dots.

Exercise 3.1.3. Express the non-trivial steady state in terms of the parameters a, b, c and d.

Exercise 3.1.4. The effect of fishing can be captured by replacing a by $a - \mu$ and $-c$ by $-c - \mu$. Explain the rationale underlying this statement.

Exercise 3.1.5. Reproduce Volterra's explanation of the change in predator and prey fish before and after WOI.

Exercise 3.1.6. When spraying insecticides to protect a crop from herbivorous insects, one should verify first whether the herbivores are currently kept in check (perhaps, from the farmer's point of view, at an unacceptably high density) by a natural enemy which is sensitive to the insecticide too. Why?

3.2 The phase portrait of the Volterra-Lotka system

The zero-isoclines divide the positive quarter of the (v,p)-plane into four regions.

EXERCISE 3.2.1. Give a schematic picture of the direction of the flow by putting arrows on the zero-isoclines, and in each of the four regions (in the spirit of wind-directions: north-east, etc).

At a glance we see that orbits have a tendency to spin around the non-trivial steady state, but we can't see whether they spiral inward or outward. In an attempt to determine this very close to the steady state, we use linearisation.

EXERCISE 3.2.2. Compute the Jacobian matrix corresponding to the non-trivial steady state. Compute the eigenvalues of this matrix. What do you conclude?

In fact, the Volterra-Lotka (after the American mathematical biologist Lotka, who formulated exactly the same system of differential equations completely independently at about the same time) phase portrait is rather special: it consists of a collection of nested closed orbits. To demonstrate this, the function $L : \mathbb{R}^2_+ \to \mathbb{R}$ defined by

$$L(v,p) = dv - c\log v + bp - a\log p \tag{3.2.1}$$

is most helpful. The point is that L is a *constant of motion* (also called a *conserved quantity*), which means that the value of L does not change along an orbit of (3.1.1)–(3.1.2). In other words: orbits are contained in level sets of L.

We have to do three things:

(1) verify that indeed L is a constant of motion;
(2) explain how one derives L from (3.1.1)–(3.1.2) (this in order to explain how one could come to the idea of studying L in relation to (3.1.1)–(3.1.2));
(3) use L to show that the orbits of (3.1.1)–(3.1.2) are closed curves.

EXERCISE 3.2.3. Compute

$$\frac{d}{dt}L(v(t),p(t))$$

for an arbitrary solution $t \mapsto (v(t),p(t))$ of (3.1.1)–(3.1.2). Does it follow that L is a constant of motion?

EXERCISE 3.2.4. With (3.1.1)–(3.1.2) as a starting point, we compute

$$\frac{dp}{dv} = \frac{dp/dt}{dv/dt} = \frac{(-c+dv)p}{(a-bp)v} = \frac{-\frac{c}{v}+d}{\frac{a}{p}-b},$$

(while simply ignoring that the denominator does, occasionally, become zero). Use *separation of variables* to solve this differential equation. What do you conclude?

EXERCISE 3.2.5. Draw the graph of the function $v \mapsto dv - c\log v$. Conclude that the equation

$$dv - c\log v - c + c\log\frac{c}{d} = \varepsilon$$

has, for given $\varepsilon > 0$, exactly two solutions $v_\pm(\varepsilon)$. Repeat this analysis for $p \mapsto bp - a\log p$. Next, consider for given $\theta > 0$ the equation

$$L(v,p) - L\left(\frac{c}{d},\frac{a}{b}\right) = \theta.$$

Check that, for any ε with $0 < \varepsilon < \theta$ the four points $(v_\pm(\varepsilon), p_\pm(\theta-\varepsilon))$ are solutions. Thus, we obtain four paths, parametrised by ε, on which L takes the constant value $L\left(\frac{c}{d},\frac{a}{b}\right)+\theta$. Check that, by inclusion of the limiting points for $\varepsilon \to 0$ and $\varepsilon \to \theta$, we obtain a closed curve.

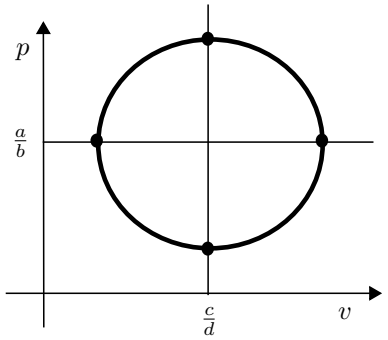


Figure 3.1.

So the Volterra-Lotka model predicts persisting prey-predator oscillations with an amplitude determined by the accidentalness of initial conditions. Yet, in section (3.1), Volterra's answer to d'Ancona was based on steady state values. Isn't there a discrepancy?

Exercise 3.2.6. Show that the average values of v and p over one period are equal to the steady state values, i.e.,

$$\frac{1}{T}\int_0^T v(\tau)\,d\tau = \frac{c}{d}, \quad \text{and} \quad \frac{1}{T}\int_0^T p(\tau)\,d\tau = \frac{a}{b},$$

where T is the period (corresponding to one full turn of (v,p) along a closed orbit).

Hint: divide the first equation of (3.1.1)–(3.1.2) by v and the second by p; use that $\frac{d}{dt}\log v(t) = \frac{1}{v(t)}\frac{dv}{dt}(t)$ and that $v(T) = v(0)$ and analogue identities for p.

We conclude with two side remarks, formulated as exercises.

Exercise 3.2.7. Show that the trivial steady state is a saddle point.

Exercise 3.2.8. Show that the number of parameters in (3.1.1)–(3.1.2) can be reduced from 4 to 1 by scaling of v, p and t.

3.3 The effect of limitations in prey growth

The Volterra-Lotka model gives an oversimplified description of prey-predator interaction. Yet it is a convenient reference point for the incorporation of additional mechanisms. In particular, the structural instability, as exemplified by the neutral stability of the steady state and the family of periodic orbits, allows us to to classify such mechanisms as either *stabilising* or *destabilising*, depending on whether the steady state in the modified model is (locally asymptotically) stable or unstable. In this and some of the following sections, we discuss various modifications in this spirit. Meanwhile, we enlarge our arsenal of phase plane techniques and train their use.

The so-called logistic equation

$$\frac{dv}{dt} = av - ev^2 = av\left(1 - \frac{e}{a}v\right) \tag{3.3.1}$$

describes, in a phenomenological manner, that, even in the absence of a predator, prey growth may be limited, with prey population size settling down at the so-called

carrying capacity

$$\tilde{v} = \frac{a}{e}. \tag{3.3.2}$$

The nonnegative parameter e describes the detrimental effects of crowding.

When $-c + d\tilde{v} < 0$, the predator population declines even at the maximal sustainable level of prey. We expect that the predator is doomed to go extinct.

EXERCISE 3.3.1. Determine the phase portrait of the system

$$\frac{dv}{dt} = v(a - ev - bp), \tag{3.3.3}$$

$$\frac{dp}{dt} = p(-c + dv), \tag{3.3.4}$$

for parameter combinations such that

$$-c + d\frac{a}{e} < 0. \tag{3.3.5}$$

In particular, draw the isoclines and a "wind-direction" vector field. Show by geometric arguments that the steady state $(\frac{a}{e}, 0)$ is globally asymptotically stable. Interpret this conclusion biologically.

When, on the other hand, we have the opposite inequality,

$$-c + d\frac{a}{e} > 0, \tag{3.3.6}$$

the predator population will start growing when introduced in an environment in which the prey is at its carrying capacity. We say that the predator can invade successfully.

EXERCISE 3.3.2. Assume (3.3.6). Show that

(i) $(\frac{a}{e}, 0)$ is a saddle point;

(ii) the nontrivial steady state

$$\left(\frac{c}{d}, \frac{ad - ec}{bd}\right)$$

is biologically meaningful, and that it is either a stable node or a stable spiral.

Next, draw the isoclines and the wind-direction field for this case.

In fact, the nontrivial steady state is *globally* asymptotically stable when (3.3.6) holds. In other words, orbits starting away from the steady state spiral in towards it. To demonstrate this, we use an auxiliary function (a slight modification of L introduced in the previous section), which is decreasing along orbits, and which achieves its minimum in the nontrivial steady state. Such functions are called *Lyapunov functions* and often denoted by the symbol V. There is no general method for their construction (but for some mechanical systems one can use the energy, for some chemical systems the entropy, and for some genetic systems the mean fitness), one has to rely on luck, intuition, experience and perseverance.

We define $V : \mathbb{R}^2_+ \to \mathbb{R}$ by

$$V(v, p) = d\left(v - \frac{c}{d}\log v\right) + b\left(p - \frac{ad - ec}{bd}\log p\right). \tag{3.3.7}$$

EXERCISE 3.3.3. Show that V decreases along orbits of (3.3.4). More precisely, show that

$$\frac{d}{dt}V(v(t), p(t)) \leq 0$$

for an arbitrary solution $t \mapsto (v(t), p(t))$ of (3.3.4), with equality if and only if $v(t) = \frac{c}{d}$.

From this information, one can deduce that every orbit approaches the nontrivial steady state. The precise argument is a bit technical but goes roughly like this: points in the ω-limit set[1] of an arbitrary orbit must lie in the same level set of V and all orbits starting in such ω-limit points must lie in that same level set too; only the nontrivial steady state satisfies these requirements (the appropriate version of Lyapunov's theorem goes by the name of LaSalle's Invariance Principle, see
`http://en.wikipedia.org/wiki/Krasovskii-LaSalle_principle`).

We conclude that density dependence in prey population growth is a *stabilising* mechanism for prey-predator interaction.

3.4 Deriving the Holling type II functional response by a time scale argument

The *functional response* is by definition the per predator per unit of time eaten number of prey, usually as a function of prey density. The *numerical response* is the per predator per unit of time produced number of offspring (as a result of eating prey). A frequent assumption (which ignores reproduction delays, etc.) is that the numerical response equals the functional response times a constant, the conversion efficiency.

Exercise 3.4.1. What is the functional response in the Volterra-Lotka model? What parameter combination corresponds to the conversion efficiency?

As the catching and "handling" of prey takes time and the digestive tract has only limited capacity, the functional response should saturate at high prey density. The so-called Holling type II functional response

$$v \mapsto \frac{bv}{1+b\beta v} \tag{3.4.1}$$

does indeed approach a limit $\frac{1}{\beta}$ for $v \to \infty$. The aim of this section is to derive this expression by a time-scale argument from a more complicated model in which we distinguish two types of predators, those searching for prey and those busy handling a prey caught earlier:

$$\frac{dv}{dt} = av - bvs, \tag{3.4.2}$$

$$\frac{ds}{dt} = -bvs + \frac{1}{\beta}(1+\eta)h - cs, \tag{3.4.3}$$

$$\frac{dh}{dt} = bvs - \frac{1}{\beta}h - ch. \tag{3.4.4}$$

Here s denotes the searching predators and h the handling predators, so $p = s + h$. Upon catching a prey, the predator turns from searching into handling. There is a probability per unit of time of $\frac{1}{\beta}$ that the handling is completed, whereupon the handling predator again turns into a searching one, but in addition produces $\eta = \frac{d}{b}$ offspring (which starts its life searching). Note that we may also say that the handling time is an exponentially distributed random variable with mean β and that the average energy content of a prey suffices to produce η predator.

[1] $(\bar{v}, \bar{p})$ belongs to the ω-limit set of $(v(0), p(0))$ iff there is a sequence $t_n \to \infty$ such that $(v(t_n), p(t_n)) \to (\bar{v}, \bar{p})$; clearly points on the same orbit have the same ω-limit set, so we also speak about the ω-limit set of an orbit; if one considers sequences $t_n \to -\infty$ one speaks about the α-limit set; note that α is the first and ω the last character in the Greek alphabet.

EXERCISE 3.4.2. Assume η, βa, βc and $\frac{p_0}{v_0}$ are all very small and of the same order of magnitude (this means that very little predator can be made out of one prey, that the handling time is negligible on the time scale of prey population growth and expected length of time of a predator's life, and that there are far less predators than prey). Meanwhile, assume that bp_0 is of the same order of magnitude as a (this means that predation causes changes in prey density at the same time scale as intrinsic prey population growth takes place). Perform a two-time-scale analysis of (3.4.4) in the spirit of Chapter 2. In particular, derive the system

$$\begin{aligned} \frac{dv}{dt} &= av - \frac{bvp}{1+b\beta v}, \\ \frac{dp}{dt} &= -cp + \frac{dvp}{1+b\beta v}, \end{aligned} \tag{3.4.5}$$

for changes in the prey and predator densities at the slow time scale.

3.5 The destabilising effect of a saturating functional response

EXERCISE 3.5.1.

(i) Draw the isoclines and the wind-direction-field for system (3.4.5), assuming that $\frac{d}{b\beta} > c$ (why?);
(ii) Show that the nontrivial steady state of (3.4.5) is unstable;
(iii) Formulate a conclusion in biological terms (taking the title of this section as a hint).

REMARK 3.5.1. It is a somewhat delicate task to determine where the orbits of (3.4.5) ultimately go. When prey density gets very large, the predator density satisfies approximately

$$\frac{dp}{dt} = \left(\frac{d}{b\beta} - c\right) p,$$

so p grows exponentially with exponent $\frac{d}{b\beta} - c$. When $\frac{d}{b\beta} - c < a$, the prey will escape from predator control, in the sense that both grow exponentially but the predator at a slower rate. When, on the other hand, $\frac{d}{b\beta} - c > a$, the predator overtakes the prey and causes its decline. So then cyclic alternations of growth and decline of each of the two species continue indefinitely.

3.6 The Rosenzweig-MacArthur model

Combining the two modifications of Volterra-Lotka discussed so far, we obtain the system

$$\begin{aligned} \frac{dv}{dt} &= v\left(a - ev - \frac{bp}{1+b\beta v}\right), \\ \frac{dp}{dt} &= p\left(-c + \frac{dv}{1+b\beta v}\right). \end{aligned} \tag{3.6.1}$$

Both this system, and the graphical stability criterion that we shall derive in the next exercise, carry the names of Rosenzweig and MacArthur. The system (3.6.1)

is of the form

$$\begin{aligned} \frac{dv}{dt} &= v\left(h(v) - p\phi(v)\right), \\ \frac{dp}{dt} &= p\left(-c + \psi(v)\right). \end{aligned} \tag{3.6.2}$$

with

(i) h decreasing and $h(K) = 0$ for some $K > 0$;
(ii) the functional response $v \mapsto v\phi(v)$ is increasing;
(iii) the numerical response $v \mapsto \psi(v)$ is increasing;

to which we add

(iv) $\psi^{-1}(c) < K$,

to avoid that the predator is doomed to go extinct. Systems of the form 3.6.2 are sometimes called Kolmogorov-type prey-predator systems (see (Freedman, 1980)).

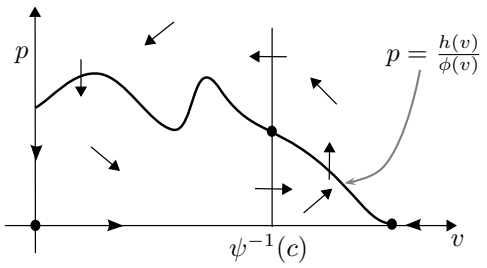


Figure 3.2.

Exercise 3.6.1.

(i) Demonstrate that (3.6.2) has exactly one steady state with both species present.
(ii) Compute the Jacobi matrix in this steady state and show that the determinant is positive.
(iii) Relate the sign of the trace of the Jacobi matrix to the sign of the derivative of $v \mapsto \frac{h(v)}{\phi(v)}$ in the steady state value of v. So, in other words, to whether the prey isocline $p = \frac{h(v)}{\phi(v)}$ is decreasing or increasing in the steady state.
(iv) Formulate carefully how the stability of the steady state can be read off from the isocline picture.
(v) Check the result of (iv) against your findings in the sections 3.3, system (3.3.4) and 3.5, system (3.4.5).
(vi) Use Poincaré-Bendixon theory (Amann, 1990; Perko, 2001; Hirsch et al., 2004) (and also the Appendix) to show that whenever
  – the internal steady state is unstable,
  – orbits stay bounded (i.e., cannot escape to infinity),
  there must exist at least one limit cycle.
(vii) Apply the results of (iv) and (vi) to system (3.6.1).
(viii) What can you say about Hopf bifurcation (see the Appendix) for system (3.6.1)?

For the following exercise, consult Rinaldi and Muratori (1992).

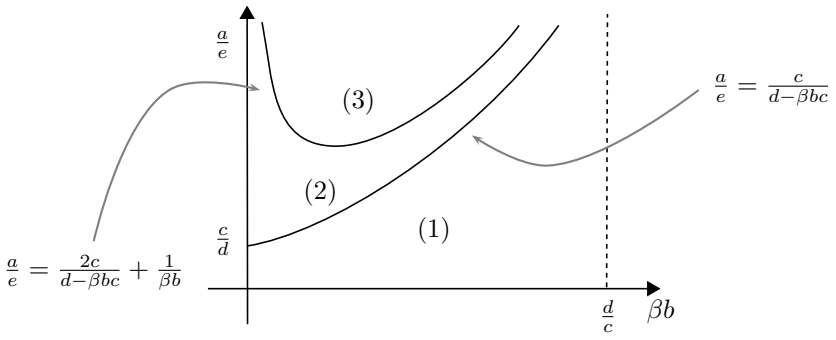


FIGURE 3.3.

EXERCISE 3.6.2. Show that under the scaling $\tau = ct$, $x = b\beta v$, $y = \frac{b}{a}p$, system (3.6.1) transforms into

$$\begin{aligned} \varepsilon \frac{dx}{d\tau} &= x\left(1 - \frac{x}{K} - \frac{y}{1+x}\right) \\ \frac{dy}{d\tau} &= y\left(-1 + \theta\frac{x}{1+x}\right), \end{aligned} \tag{3.6.3}$$

where $\varepsilon = \frac{c}{a}$, $\theta = \frac{d}{b\beta c}$, and $K = \frac{b\beta a}{e}$. Assume that $\theta > 1$ and $K > 1 + \frac{2}{\theta - 1}$.

Sketch the phase portrait for $\varepsilon \ll 1$ (this is quite subtle!). What does this assumption mean? Summarise your conclusions, i.e., interpret the phase portrait in biological terms.

Hints: the rescaling $\tau = \varepsilon t$ (where t differs from the original t) can be used to show that on the *fast time scale* t the y-component is more or less constant while x varies according to

$$\frac{dx}{dt} = x\left(1 - \frac{x}{K} - \frac{y}{1+x}\right) \tag{3.6.4}$$

where y is a parameter. By plotting the curve

$$y = (1+x)\left(1 - \frac{x}{K}\right) \tag{3.6.5}$$

in the positive (x, y)-quadrant, one can get an overview of the steady states of (3.6.4), and their stability, in dependence on y.

The *slow dynamics* are described by

$$\frac{dy}{d\tau} = y\left(-1 + \theta\frac{x(y)}{1+x(y)}\right), \tag{3.6.6}$$

with $x(y)$ a *stable* steady state of (3.6.4). All we really need to know is the sign of the right hand side of (3.6.6), in order to decide whether we move up (increasing y) or down (decreasing y) along a curve of *stable* steady states of (3.6.4). If along the curve the stability is lost, there is a switch back to the fast dynamics.

EXERCISE 3.6.3. Summarise (in a mixture of mathematical and biological terms) your knowledge about (3.6.1). You may use the information (which we now provide) that it has been proven that (3.6.1) admits at most one limit cycle (the proof is far from easy). A convenient way to sketch how the qualitative behaviour depends on the parameters is to draw the diagram as in Figure 3.3, and to describe the dynamical behaviour (in biological terms) for each of the three parameter domains.

Chapter 4

Movement in space

Our view of the world is structured by time and space and, we believe, this reflects reality: to interact, entities have to be at the same position at the same time. So far we concentrated on changes in time, but now we are going to incorporate spatial position. In the present chapter we only consider independent particles (molecules, bacteria, ...) but in the next we shall incorporate interaction.

4.1 Flux

The *density* of bacteria on an agar plate is, by definition, the number of bacteria per unit of area. Likewise, the *concentration* of a chemical substance in solution is the number of molecules per unit of volume. The density or concentration in a *point* is an idealisation, corresponding to the thought experiment of shrinking the area or volume to zero while focusing our attention on the point. We then write $u = u(t, x)$ and consider u as a smooth function of time t and position x. Note that we need to integrate $u(t, \cdot)$ over space to obtain an amount . If the total number is conserved, but the individual particles move, $u(t, \cdot)$ changes with time. How? How does redistribution over space manifest itself in changes in density/concentration?

Let us first consider a one-dimensional space (one might think of a river) and deterministic motion with prescribed velocity $c = c(x)$ (one might think of algae that float with the streaming water). The *flux* at x is the number of organisms that pass x, say from left to right, per unit of time. We denote the flux by $J = J(t, x)$. Clearly

$$J(t, x) = c(x)u(t, x), \tag{4.1.1}$$

as is indeed also suggested by the dimensional identity

$$\frac{\text{number}}{\text{time}} = \frac{\text{length}}{\text{time}} \cdot \frac{\text{number}}{\text{length}}. \tag{4.1.2}$$

Equally clearly,

$$\frac{d}{dt} \int_a^b u(t, x)\, dx = J(t, a) - J(t, b), \tag{4.1.3}$$

or, in words, if neither creation nor annihilation occurs, then the total number of organisms between a and b changes only by way of flux in at a and flux out at b (convince yourself that this terminology is appropriate when $a < b$ and $c > 0$ or when $a > b$ and $c < 0$, but should be adjusted otherwise). According to the fundamental theorem of calculus,

$$J(t, b) - J(t, a) = \int_a^b \frac{\partial J}{\partial x}(t, \xi)\, d\xi. \tag{4.1.4}$$

Hence,

$$\int_a^b \left(\frac{\partial u}{\partial t}(t,\xi) + \frac{\partial J}{\partial x}(t,\xi) \right) d\xi = 0, \tag{4.1.5}$$

and as this holds for arbitrary a and b, the integrand must be zero (see Lemma of DuBois-Reymond (Lin and Segel, 1998)), and so in combination with (4.1.1) we arrive at the *conservation law*

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(cu) = 0. \tag{4.1.6}$$

Two important variations on this theme are

(i) in *higher space dimension* the flux J is a *vector* and the conservation law takes the form

$$\frac{\partial u}{\partial t} + \nabla \cdot J = 0, \tag{4.1.7}$$

with the *divergence* of the flux $\nabla \cdot J$ defined by

$$\nabla \cdot J = \sum_{i=1}^{n} \frac{\partial J_i}{\partial x_i} \tag{4.1.8}$$

(more explanation below).

(ii) The motion of pollen that the botanist Brown observed under his microscope was very irregular. So much so that it became the prototype for *random* motion. A phenomenological description takes *Fick's law*

$$J = -d\nabla u, \tag{4.1.9}$$

as the constitutive relation that links the flux J to the density u by requiring that J is proportional to the *gradient* ∇u, with d a constant of proportionality called the *diffusion constant*, since when we substitute (4.1.9) into (4.1.7) we obtain the *diffusion equation*

$$\frac{\partial u}{\partial t} = d\Delta u, \tag{4.1.10}$$

where $\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$ is the *Laplacian*. Note that d has dimension $(\text{length})^2/\text{time}$.

In the next subsection we shall provide various derivations that yield a quasi-mechanistic underpinning of Fick's Law. We conclude this subsection with a few observations on the notion of flux in higher dimensions.

Consider a point in two-space. If we want to talk about the traffic of particles in that point, we need to specify a direction. This we do by choosing a unit vector m. The flux J at the point is a vector such that, whatever choice of m, the number of particles crossing per unit of time a straight line L perpendicular to m in an interval of length h centred at the focus point equals

$$J \cdot m \; h + o(h) \quad \text{as } h \to 0.$$

For deterministic motion we have just as in the one-dimensional case that the flux is the product of the velocity, which is now a vector, and the density (so in particular, the traffic is maximal in the direction of the velocity and zero in the direction perpendicular to the velocity).

In the present context, the analogue of the fundamental theorem of calculus is the *Divergence Theorem*

$$\int_\Omega \nabla \cdot F \, dA = \int_{\partial\Omega} F \cdot n \, ds, \tag{4.1.11}$$

where n is the outward pointing unit vector (outward normal) perpendicular to the boundary $\partial\Omega$ of the domain Ω.

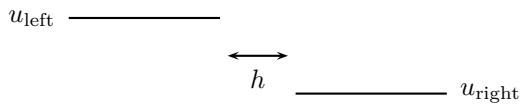


FIGURE 4.1.

In three dimensions, one replaces L by a plane and the interval of length h by a subset of this plane of area h. The Divergence Theorem now reads

$$\int_\Omega \nabla \cdot F \, dV = \int_{\partial\Omega} F \cdot n \, dS, \tag{4.1.12}$$

and is usually called *Gauss' Theorem*.

Note that if we substitute Fick's Law (4.1.9) into the identity of the Divergence or Gauss' Theorem, we obtain

$$\int_\Omega \Delta u = \int_{\partial\Omega} \frac{\partial u}{\partial n}. \tag{4.1.13}$$

4.2 Various ways to motivate Fick's Law

Derivation 1 Consider one-dimensional space and suppose that at every point particles move at speed c, but half of them to the left and half to the right. Consider a point where the density jumps over a gap of length h from a constant density u_{left} to a constant density u_{right} (see Fig. 4.1). In a time interval of length Δt the net transport to the right equals

$$\frac{1}{2}(u_{\text{left}} - u_{\text{right}})c\Delta t.$$

So per unit of time $\frac{1}{2}(u_{\text{left}} - u_{\text{right}})c$ is transported, which we write as

$$\frac{u_{\text{left}} - u_{\text{right}}}{h} \frac{1}{2} ch.$$

When we now take the limit $h \to 0$ while assuming that

$$\frac{1}{2}hc \to d,$$

we obtain

$$\text{flux} = -d\frac{\partial u}{\partial x}.$$

The key point of this very debatable "derivation" is that it clearly shows that in the limit we should have $c \to \infty$. So, in a sense, we consider particles that move infinitely fast but never can make up their mind about the direction in which they go.

Derivation 2 Imagine a particle moving on a one-dimensional lattice that we represent by $\mathbb{Z}$. We take time discrete and at every time step the particle moves to the left with probability $\frac{1}{2}$ and to the right with probability $\frac{1}{2}$. If the particle is at position zero at time zero then the probability $p_i(n)$ that it is at position i at time n is given explicitly by

$$p_i(n) = \begin{cases} \binom{n}{\frac{1}{2}(n+i)} \left(\frac{1}{2}\right)^n & \text{for n+i even and } -n \le i \le n, \\ 0 & \text{otherwise.} \end{cases}$$

(Indeed, if the particle makes k steps to the right then it makes $n-k$ steps to the left, and to end up at i we should have $k-(n-k)=i$. Hence $k=\frac{1}{2}(n+i)$. The probability that k out of n steps are to the right equals $\binom{n}{k}(\frac{1}{2})^k(\frac{1}{2})^{n-k}=\binom{n}{k}(\frac{1}{2})^n$.)

If we now take $i=\frac{x}{\lambda}$ and $n=\frac{t}{\tau}$, let both λ and τ approach zero but in such a way that $\lambda^2/2\tau$ converges to d, then the binomial distribution converges to the normal distribution (see e.g. Section 7.3 in (Chung, 1974), or better still, verify this yourself)

$$p(t,x)=\frac{1}{2\sqrt{\pi dt}}e^{-\frac{x^2}{4dt}}.$$

This, as we shall see later on, is the fundamental solution of the one-dimensional diffusion equation. Note that one can interpret λ/τ as the speed and that this speed grows beyond any bound.

Alternatively, we can shorten the distance between the lattice points as well as the time intervals between steps. Then, by performing a formal Taylor expansion for p we derive the diffusion equation directly from the random walk assumptions by taking a limit. The next derivation is essentially of this type, but considers right away both space and time as continuous variables.

Note once again that for independently moving particles we need not make a distinction between the density of many particles and the probability density for one particle.

Derivation 3 We postulate that

$$u(t+\tau,x)=\int_{-\infty}^{\infty}u(t,x-y)\frac{1}{\varepsilon}\phi\left(\frac{y}{\varepsilon}\right)\,dy \tag{4.2.1}$$

for a function ϕ satisfying $\phi\geq 0$, $\int_{-\infty}^{\infty}\phi(y)\,dy=1$, and $\phi(-y)=\phi(y)$. Then, in particular, $\int_{-\infty}^{\infty}y\phi(y)\,dy=0$. The identity (4.2.1) states that between times t and $t+\tau$ particles are moved over a distance y with probability density $\frac{1}{\varepsilon}\phi\left(\frac{y}{\varepsilon}\right)$ and the symmetry guarantees that there is no preferred direction. A formal Taylor expansion yields

$$u(t+\tau,x)=u(t,x)+\tau\frac{\partial u}{\partial t}(t,x)+\cdots, \tag{4.2.2}$$

$$u(t,x-y)=u(t,x)-y\frac{\partial u}{\partial x}(t,x)+\frac{1}{2}y^2\frac{\partial^2 u}{\partial x^2}(t,x)+\cdots. \tag{4.2.3}$$

Substituting these expressions in (4.2.1) we find

$$\tau\frac{\partial u}{\partial t}(t,x)=\frac{1}{2}\int_{-\infty}^{\infty}y^2\frac{1}{\varepsilon}\phi\left(\frac{y}{\varepsilon}\right)ydy\frac{\partial^2 u}{\partial x^2}(t,x)+\cdots, \tag{4.2.4}$$

$$=\frac{\varepsilon^2}{2}\int_{-\infty}^{\infty}z^2\phi(z)dz\frac{\partial^2 u}{\partial x^2}(t,x)+\cdots \tag{4.2.5}$$

If we now let both τ and ε converge to zero but in such a manner that

$$\frac{\varepsilon^2}{2\tau}\int_{-\infty}^{\infty}z^2\phi(z)\,dz\to d,$$

we arrive at

$$\frac{\partial u}{\partial t}=d\frac{\partial^2 u}{\partial x^2}. \tag{4.2.6}$$

EXERCISE 4.2.1. Let $u(t,x)$ satisfy (4.2.6). Show that u as a function of t decreases where u as a function of x has a maximum, and that u as a function of t increases where u as a function of x has a minimum. Conclude that the diffusion equation has an equalising effect. Do you agree that this is already embodied in Fick's Law?

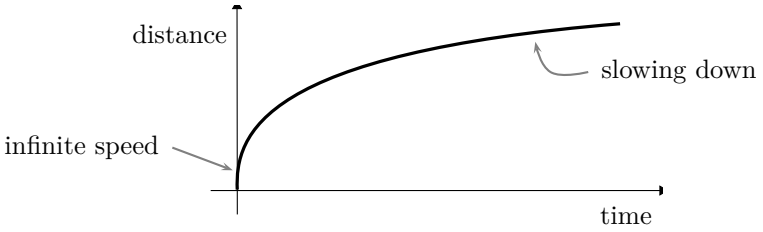


Figure 4.2.

4.3 Transport by diffusion

The two observations

(1) dim $d = \frac{(\text{length})^2}{\text{time}}$,

(2) the diffusion equation (4.2.6) is *invariant* under a scaling

$$t^* = \varepsilon^2 t, \quad x^* = \varepsilon x,$$

both motivate the following statements:

- the average distance over which diffusion transports particles in a given time interval of length t is proportional to $\sqrt{dt}$
- the average time it takes to diffuse over a distance h is proportional to h^2/d.

Please contrast Figure 4.2 with the deterministic straight line distance = velocity $\cdot$ time. It appears that the efficiency of diffusion as a transport mechanism depends very much on the distance to be travelled! We need the circulatory blood system for active transport of, among other things, oxygen. But the very last bit of transport to the muscle tissue is by diffusion! See (Vogel, 1988, Chapter 8) for some general considerations.

4.4 How to measure the diffusion coefficient

A capillary tube is inserted into a suspension of bacteria of known concentration (see Fig. 4.3). After a prescribed period of time, the tube is extracted and the number of bacteria that have entered is counted. Assume that the bacteria can be described in terms of a concentration u, that they move randomly, that the concentration at the mouth of the tube is always a constant, u_0 say, that there are no bacteria in the tube at the beginning of the experiment, and that the concentration in the tube varies only in the length direction and not in the radial direction. A mathematical formulation of these assumptions reads

$$\frac{\partial u}{\partial t} = d\frac{\partial^2 u}{\partial x^2} \quad 0 < x < \infty,\ t > 0 \tag{4.4.1}$$

$$u(t,0) = u_0 \quad t > 0 \tag{4.4.2}$$

$$u(0,x) = 0 \quad x > 0, \tag{4.4.3}$$

where x measures the distance down the (infinitely long, by debatable assumption) tube.

Exercise 4.4.1. Derive the expression

$$d = \frac{\pi N^2}{4u_0^2 A^2 T} \tag{4.4.4}$$

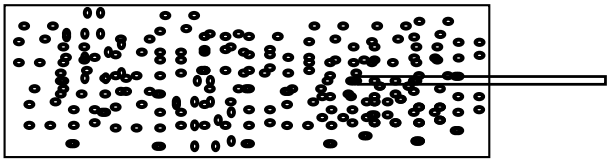


FIGURE 4.3.

where N denotes the number of bacteria in the tube at time T and A the cross-sectional area of the tube.

The way (4.4.4) is used is: for several choices of u_0 and T the experiment is performed and N is determined. The right hand side is then computed and if, within reasonable accuracy, the value is the same for all u_0, T and N combinations, then we have confidence in the model and, in addition, the value serves as an estimate for d. A typical value is 0.2 cm^2/hr.

Hints and remarks: take as a starting point the *fundamental solution*

$$\frac{1}{2\sqrt{\pi dt}} e^{-\frac{x^2}{4dt}},$$

which we will derive later on in Section 5.1. The fundamental solution serves as a building block: since the equation is linear, the *superposition principle* applies. The fundamental solution is the solution of the diffusion equation with initial data $u_0(x) = \delta(x)$, the Dirac delta function. For instance, if we replace this initial condition by the general condition $u(0, x) = g(x)$, for $x \in \mathbb{R}$, then

$$u(t, x) = \frac{1}{2\sqrt{\pi dt}} \int_{\mathbb{R}} e^{-\frac{(x-y)^2}{4dt}} g(y)\, dy.$$

To make this formula applicable to (4.4.3), we need the trick of extending the domain and the initial condition to $(-\infty, \infty)$ in such a way that the boundary condition automatically holds (essentially this is based on symmetry). The right choice is

$$u(0, x) = 2u_0, \quad x < 0,$$

so that the value for $x = 0$ is (for $t = 0$, but in fact also for $t > 0$) exactly the average of the value to the left and the value on the right. You should now arrive at

$$\frac{u(t, x)}{u_0} = \frac{2}{\sqrt{\pi}} \int_{x/\sqrt{4dt}}^{\infty} e^{-\xi^2}\, d\xi. \tag{4.4.5}$$

To be clear: deriving (4.4.5) forms part of the Exercise! To derive (4.4.4), you may want to use integration by parts.

4.5 About sojourn times

Suppose particles enter a compartment at a rate F. Let N denote the total number of particles in the compartment. To find a relation between N and F we need to know how long particles stay in the compartment. We assume that this so-called sojourn time is a stochastic variable T with a continuous probability distribution.

EXERCISE 4.5.1. Assume that both F and N depend on time t. Explain in words the bookkeeping considerations underlying the identity.

$$N(t) = \int_0^{\infty} F(t - \sigma) P(T \geq \sigma)\, d\sigma. \tag{4.5.1}$$

Exercise 4.5.2. Now assume that both N and F are constant. Let f denote the probability density of T, so, in particular,

$$P(T \geq \sigma) = \int_\sigma^\infty f(s)ds, \quad \text{and} \quad \int_0^\infty f(s)\,ds = 1.$$

Let

$$\tau = \int_0^\infty \sigma f(\sigma)\,d\sigma$$

denote the mean of T. Show that $N = F\tau$. Are you surprised? Finally, reflect a moment on the possibility that $\tau = \infty$. How would you interpret the result $N = F\tau$ in that case ?

4.6 How long does it take?

Suppose particles are released at $x = L$ and removed upon arrival at $x = 0$. We want to check that the rule of thumb formulated in Section 4.3 applies. To do so, we use a trick: we consider a steady situation with continuous release and removal rather than following an individual particle (the point being that in this manner we let the equation take care of the statistics; this works since we are satisfied with the average, the expected, time and do not aim to derive the full probability distribution).

Exercise 4.6.1. Why should we supplement the steady state equation

$$d\frac{\partial^2 u}{\partial x^2} = 0$$

with the *boundary conditions*

$$u(L) = u_0 \quad \text{and} \quad u(0) = 0.$$

Compute the steady particle density, i.e., find a function u that satisfies the equation as well as the boundary conditions. Express the influx J_{in}, i.e., the number of particles that enter at $x = L$ per unit of time, in terms of u_0, d and L. Next, compute the total number N of particles that are present.
How are J_{in} and N related (recall Section 4.2)? Compute the average sojourn time. Check in particular that it does not depend on u_0 (did you already anticipate this?) and that it confirms nicely to the rule of thumb.

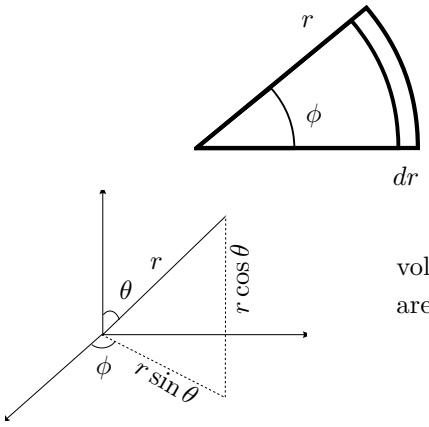
The efficiency of diffusion as a transport mechanism depends not only on size but also on *shape*, in particular on the dimension (1, 2 or 3) of the domain. We now want to demonstrate that it has advantages for a cell to arrange the chemical "factories" along a two-dimensional membrane (incidentally, recent findings indicate that a cell is partly an assembly-belt and that the traditional picture of a freely floating 3D chemical soup is fundamentally flawed). In this connection it is also good to realise that diffusion can only "work" if there is an excess of particles, as any one of them may go the wrong direction and/or take ages before reaching the target (if at all).

Exercise 4.6.2. Consider a *radially symmetric* two-dimensional setting. Show that the conservation equation takes the form

$$r\frac{\partial u}{\partial t} = -\frac{\partial}{\partial r}(rJ),$$

that Fick's law amounts to

$$J = -d\frac{\partial u}{\partial r},$$



area $\sim \phi r dr$ $\quad x_1 = r\cos\phi$
length $\sim \phi r$ $\quad x_2 = r\sin\phi$

volume $\sim r^2\phi\theta dr$ $\quad x_1 = r\sin\theta\cos\phi$
area $\sim r^2\phi\theta$ $\quad x_2 = r\sin\theta\sin\phi$
$\quad x_3 = r\cos\theta$

FIGURE 4.4.

and that accordingly the diffusion equation reads

$$\frac{\partial u}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(dr\frac{\partial u}{\partial r}\right).$$

Next show that in three dimensions one obtains

$$\frac{\partial u}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(dr^2\frac{\partial u}{\partial r}\right).$$

For both these exercises, a hint: see Figure 4.4.

Now suppose the particle density is held at $u_0 > 0$ on a circle/ball of radius L and at zero at a circle/ball of radius $a < L$. Derive that the average sojourn time is given by, respectively

$$\frac{1}{2d}\left\{L^2\left(\log\frac{L}{a} - \frac{1}{2}\right) + \frac{1}{2}a^2\right\},$$

and

$$\frac{1}{ad}\left\{\frac{1}{3}L^3 - \frac{a}{2}L^2 + \frac{1}{6}a^3\right\}.$$

Reflect on the difference for large L.

As a final note along this theme, let us consider phytoplankton cells drifing in an ocean of depth L, undergoing random motion due to turbulent eddy diffusivity in the ocean's mixed layer, and about to be devoured by clams waiting at the ocean floor. How long does a phytoplankton cell drift on average? Does it drift long enough to able to grow, absorp light and take up nutrients, and divide? Gravity is less of a problem than random motion by turbulent eddy diffusivity. The key observation: let $T(x)$ be the *expected* time till absorption (i.e., until it is eaten by the big monster at the boundary). Then

$$T(x) = t + \int_{-\infty}^{\infty} \Phi(t, \xi - x)T(\xi)d\xi + \text{small correction}$$

for t small and x not very near to the boundary [This requires some additional explanation]. Differentiate with respect to t, and use that Φ satisfies the diffusion equation to find

$$0 = 1 + d\int_{-\infty}^{\infty} \Big(\Phi(t, \xi - x)\Big)_{xx} T(\xi)d\xi + \cdots$$

Integrate by parts twice and let $t \to 0$ to find

$$0 = 1 + T''(x).$$

The boundary condition are $T(0) = 0$ (the monster!) and $T'(L) = 0$ (no flux, reflecting). Then the explicit solution is given by

$$T(x) = -\frac{1}{2}\frac{x^2}{d} + \frac{L}{d}x = \frac{x}{d}\Big(L - \frac{x}{2}\Big). < /0,$$

T reaches its maximum $L^2/2d$ at $x = L$, and for $x = L/2$ we have $T = \frac{3L^2}{8d}$. The vertical eddy diffusivity in the ocean's mixed layer is approximately 10^{-4} $\mathrm{m^2 s^{-1}}$, so if the ocean is about 10 meters deep, it takes about 4 days to reach the bottom and be devoured by the clams.

4.7 A remark on boundary conditions

The idea behind the boundary condition $u(L) = u_0$ in Exercise 4.6.1 is that to the right of $x = L$ there is a reservoir of particles which is held at a constant density. Alternatively, we might imagine a pumping device that somehow manages to generate a constant influx. In that case we should put as boundary conditions

$$d\frac{\partial u}{\partial x}(L) = \text{ prescribed influx } \sim \frac{\text{number}}{\text{time}}$$

(note that, as we saw in Section 4.2, Derivation 1, the flux equals $-d\frac{\partial u}{\partial x}$ if our orientation is from left to right; but the domain is to the left of $x = L$, i.e., the inward normal points to the left). One can redo Exercise 4.6.1 with the alternative boundary condition and arrive, of course, at the same answer.

If we model animals that can move freely in some domain Ω, but cannot (for whatever reason) leave Ω, we should put *no-flux boundary conditions*

$$\frac{\partial u}{\partial n}\Big|_{\partial\Omega} = 0.$$

These are also called *(zero-)Neumann conditions* and we omitted, as usual, the factor d since when we put zero at the right-hand side it has no influence (but be aware of this factor when the flux isn't required to reduce to zero!). $\partial\Omega$ is called a *reflecting* boundary.

If u is the density of plants and the diffusion term is used to describe the dispersal of seeds, it may be that the complement of Ω is simply unsuitable habitat in which no plant can grow. We may then impose *(zero-)Dirichlet conditions*

$$u\Big|_{\partial\Omega} = 0,$$

but should realise that such a form of heterogeneity of the world as a whole has a strong impact on pattern formation (we shall return to this point in the next chapter).

Mixed boundary conditions

$$\Big[-(1-\theta)d\frac{\partial u}{\partial n} + \theta u\Big]_{\partial\Omega} = 0$$

are, from a mathematical point of view, a one-parameter family that forms a homotopy between no-flux and zero-Dirichlet and describe, from a biological point of view, a partially reflecting boundary to a completely hostile exterior. Their relevance in a biological modelling context is not clear at all.

In some diffusion problems arising in population genetics, the spatial variable x is actually a fraction of the population carrying a certain allele. In such problems

d depends on x and declines to zero when x approaches the boundary points $x = 0$ and $x = 1$. The classification of the mechanistic effect of the boundary is far more subtle in such a situation, see (Feller, 1952, 1954, 1955).

Boundary conditions should be chosen on the basis of modelling considerations, even though this is far less straightforward than one is tempted to believe. Much mathematical work on biology inspired equations is wasted for the simple reason that boundary conditions, in particular zero-Dirichlet conditions, are chosen out of habit and without a critical reflection on their meaning and effect.

Chapter 5

Linear diffusion

5.1 The fundamental solution

The aim of this section is to derive the fundamental solution to the diffusion equation

$$\frac{\partial u}{\partial t} = d\frac{\partial^2 u}{\partial x^2}, \quad x \in \mathbb{R}, \ \ t \geq 0, \tag{5.1.1}$$

subject to far out boundary conditions

$$u(t, \pm\infty) = 0, \quad t \geq 0, \tag{5.1.2}$$

using *dimensional analysis*. This technique often reveals the basic structure of solutions to partial differential equations, by simply asking which (combination) of the variables actually determine the dependent variable we want to study.

Let us model the concentration of some species living on the real line, dispersing according to (5.1.1). Assume that at time $t = 0$, all individuals are in one particular location $x = 0$. Since the number of individuals remains constant in time, we know that for each $t > 0$,

$$\int_{-\infty}^{\infty} u(t, x)dx = 1. \tag{5.1.3}$$

(Exploiting the linearity of the diffusion equation, we have just taken the liberty of scaling u such that (5.1.3) holds.) A solution u is now completely determined by all other quantities involved, so we are looking for a function f such that

$$u = f(t, x, d). \tag{5.1.4}$$

We have already seen in Section 4.3 that (5.1.1) is invariant under the scaling $t^* = \varepsilon^2 t$, $x^* = \varepsilon x$. This suggests that we could write f as a function of $x/\sqrt{t}$. However, x^2/t is not dimensionless, and we therefore cannot expect solutions to be dependent on $x/\sqrt{t}$ *only*. Observe, however, from (5.1.1) that the diffusion constant d has dimension (length)^2/time. So the combination $x/\sqrt{dt}$ *is* dimensionless.

On the other hand, u has dimension 1/length, so we at least need f to be of the form $w\phi(x/\sqrt{dt})$ for some function w with dimension 1/length and a dimension-less function ϕ. The conservation equation (5.1.3) now yields

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} u(t, x)dx \\ &= \int_{-\infty}^{\infty} w\phi\left(\frac{x}{\sqrt{dt}}\right)dx \\ &= \int_{-\infty}^{\infty} w\sqrt{dt}\phi(\xi)d\xi, \end{aligned}$$

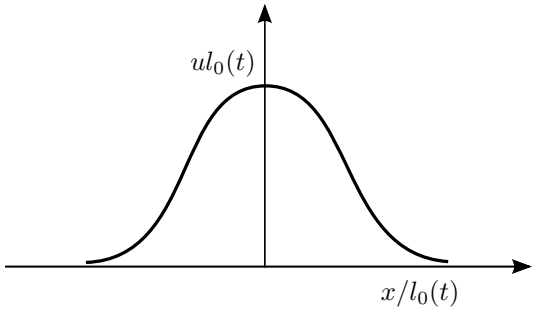


FIGURE 5.1. In the right time-dependent variables ($x/l_0(t)$ and $ul_0(t)$ the fundamental solution of the diffusion equation has a unique profile for all time.

where $\xi = x/\sqrt{dt}$. So $w = 1/\sqrt{dt}$ seems the obvious candidate to make all dimensions fit. In all, we look for a u of the form

$$u = \frac{1}{\sqrt{dt}}\phi\left(\frac{x}{\sqrt{dt}}\right).$$

This strategy of finding the structure of solutions by considering dimensions is applicable much more generally (Barenblatt, 1996).

Note that if we define the time-dependent length scale $l_0(t) := \sqrt{dt}$, then

$$u = \frac{1}{l_0(t)}\phi\left(\frac{x}{l_0(t)}\right),$$

so if we plot $ul_0(t)$ versus $x/l_0(t)$, we find *one* curve for all time. See Figure 5.1. This shows that this solution possesses the property of *self-similarity*: when scaling both the spatial variable and the (population) density in an appropriate time-dependent manner, nothing changes at all. In fact one can also find the form of the solution by, from the very beginning, searching for a solution such that

$$u(t,x) = \lambda^\alpha u(\lambda t, \lambda^\beta x) \quad \text{for all } \lambda > 0,$$

and constants α and β to be chosen suitably. The choice $\lambda = t^{-1}$ then reveals that we are looking for a function of one variable.

The great advantage of having to find $\phi(\xi)$ instead of $f(t,x,d)$ is that the (partial differential) diffusion equation (5.1.1) reduces to an *ordinary* differential equation in which, moreover, neither the independent nor the dependent variable carries a physical dimension.

EXERCISE 5.1.1. Show that, in the new variable ξ, (5.1.1) becomes

$$\frac{d^2\phi}{d\xi^2} + \frac{\xi}{2}\frac{d\phi}{d\xi} + \frac{\phi}{2} = 0. \tag{5.1.5}$$

Integrating once, we find that

$$\frac{d\phi}{d\xi} + \frac{\xi}{2}\phi = \text{constant}. \tag{5.1.6}$$

Since u is symmetric with respect to reflection in 0, ϕ should be symmetric around $\xi = 0$, and therefore $\frac{d\phi}{d\xi} = 0$ at $\xi = 0$. The constant on the right hand side is therefore zero. Using for instance integrating factors to solve (5.1.6), we conclude that

$$\phi(\xi) = Ae^{-\xi^2/4}, \tag{5.1.7}$$

for some constant A.

Using the well-known integral identity

$$\int_{\mathbb{R}} e^{-\xi^2} d\xi = \sqrt{\pi}.$$

we finally arrive at a solution of the diffusion equation, which we denote by Φ and which is explicitly given by the formula

$$\Phi = \frac{1}{2\sqrt{\pi dt}} e^{-\frac{x^2}{4dt}}.$$

We call Φ the *fundamental solution*. Note that $\Phi > 0$ for arbitrary x, no matter how small we choose $t > 0$. This is yet another manifestation of the infinite speed of propagation that is embodied in the diffusion equation. Also note that Φ is a Gauss distribution with mean zero and variance $2dt$. In particular, Φ is astronomically small for large $|x|$. So it is not so clear how we should interpret the positivity of Φ. We return to the question of the speed of propagation in Section 5.3 below. Finally note that the variance goes to zero for $t \downarrow 0$. So the distribution at $t = 0$ corresponds to a unit (recall (5.1.3)) mass concentrated at $x = 0$. In the mathematically precise sense of distributions, $u(t, .)$ converges to the Dirac delta for $t \downarrow 0$.

The reason Φ is called the *fundamental solution* is that by linearity of the diffusion equation we may apply *superposition*: given initial data $u(0, x) = u_0(x)$, the solution of the diffusion problem can be expressed as a convolution of the initial data and the fundamental solution:

$$u(t, x) = \int_{-\infty}^{\infty} \Phi(t, x - y) u_0(y) dy.$$

This section is greatly inspired by (Barenblatt, 1996, Section 2.1).

5.2 Separation of variables and spectral theory

If $\frac{du}{dt} = ru$ we know that u grows exponentially when $r > 0$, while it decays exponentially if $r < 0$. Now suppose that, additionally, u diffuses in a spatial domain. Is the conclusion still true? Does u develop any spatial pattern? What is the influence of boundary conditions? For simplicity we restrict our attention to a one-dimensional spatial domain. To begin with we provide the *diffusion equation*

$$\frac{\partial u}{\partial t} = d\frac{\partial^2 u}{\partial x^2} + ru \tag{5.2.1}$$

with so-called no-flux boundary conditions

$$\frac{\partial u}{\partial x}(t, 0) = 0 = \frac{\partial u}{\partial x}(t, L) \tag{5.2.2}$$

at the endpoints $x = 0$ and $x = L$ of the interval we consider.

Exercise 5.2.1. Explain why we can, without any loss of generality, either take $d = 1$ or $L = 1$. Also explain why, for $r > 0$, it is no loss of generality to take $r = 1$.

We choose to take $d = 1$, but to keep r as it is (so we scale the spatial variable but we do not scale time).

Exercise 5.2.2. Show that, in order for

$$u(t, x) = a(t)\phi(x)$$

to be a solution, we must necessarily have that, for some λ,

$$a(t) = \text{constant} \cdot e^{\lambda t},$$

and

$$\phi''(x) = (\lambda - r)\phi(x) \tag{5.2.3}$$

$$\phi'(0) = 0 = \phi'(L). \tag{5.2.4}$$

EXERCISE 5.2.3. Show that

(i) both $\phi(x) = \cos \mu x$ and $\phi(x) = \sin \mu x$ satisfy the differential equation $\phi'' = (\lambda - r)\phi$, provided $\lambda - r = -\mu^2$.

(ii) only $\phi(x) = \cos \mu x$ satisfies, in addition, the left boundary condition $\phi'(0) = 0$.

(iii) in order for $\phi(x) = \cos \mu x$ to also satisfy the right boundary condition $\phi'(L) = 0$, we should have

$$\mu = \frac{k\pi}{L} \quad \text{for some integer } k \geq 0.$$

(iv) finally, verify that (5.2.3)–(5.2.4) does *not* have a solution if $\lambda - r > 0$.

EXERCISE 5.2.4.

(i) Verify that, while making the preceding two exercises, you have deduced that the following statement is true: for $k = 0, 1, 2, \ldots$,

$$u(t, x) = e^{rt} e^{-\left(\frac{k\pi}{L}\right)^2 t} \cos\left(\frac{k\pi}{L} x\right) \tag{5.2.5}$$

is a solution of (5.2.1)—(5.2.2).

(ii) A very simple argument shows that of all these solutions the one with $k = 0$ has the fastest growth (or the least decay, when $r < 0$) for $t \to \infty$. Formulate this argument.

(iii) An even simpler argument shows that the solution with $k = 0$ is "flat", i.e., has no spatial structure. Provide also this argument.

The spatial solutions found in (5.2.5) can be used as building blocks for a representation of the general solution. By "general solution" we mean that we add to (5.2.1)—(5.2.2) an *initial condition*

$$u(0, x) = u_0(x), \tag{5.2.6}$$

where u_0 is a rather arbitrary function defined on $[0, L]$. Suppose that we can find coefficients $\{b_k\}_{k=0}^{\infty}$ such that

$$u_0(x) = \sum_{k=0}^{\infty} b_k \cos\left(\frac{k\pi}{L} x\right). \tag{5.2.7}$$

Then, by the superposition principle, which holds since (5.2.1)—(5.2.2) is a linear problem, and (5.2.5) we have

$$u(t, x) = e^{rt} \sum_{k=0}^{\infty} b_k e^{-\left(\frac{k\pi}{L}\right)^2 t} \cos\left(\frac{k\pi}{L} x\right). \tag{5.2.8}$$

The justification of (5.2.7) is the subject of Fourier analysis.

EXERCISE 5.2.5.

(i) In the above introduction of Section 5.2 we formulated three questions. Provide answers to the first two of these.

(ii) Alternatively to the no-flux boundary conditions (5.2.2), we can consider the situation in which the concentration is held zero at the boundary (imagine a big monster at the boundary that eats everything that gets there):

$$u(t, 0) = 0 = u(t, L). \tag{5.2.9}$$

It follows that now

$$u(t,x) = e^{rt} \sum_{k=1}^{\infty} a_k e^{-\left(\frac{k\pi}{L}\right)^2 t} \sin\left(\frac{k\pi}{L}x\right). \tag{5.2.10}$$

Answer the first two questions in the introduction of this section for this situation. Hint: note that the term with $k = 0$ is now missing, as $\sin 0 = 0$.

(iii) Give a (partial) answer to the third question in the introduction.

Exercise 5.2.6. Consider a rectangular domain Ω with sides of length L_1 and L_2. Determine the eigenvalues and eigenvectors of the diffusion problem with no-flux boundary conditions. Conclude that the modes are naturally numbered by a *pair* of integers. If one orders the eigenvalues according to μ_{k_1,k_2}, one obtains an ordering of these pairs. Investigate the influence of the ratio L_1/L_2 on this ordering of pairs.

5.2.1 A digression on general theory

For general bounded open subsets Ω of $\mathbb{R}^n$, the eigenfunctions of the Laplace operator provided with zero Dirichlet boundary conditions form an orthonormal basis for $L^2(\Omega)$, i.e., every element f of $L^2(\Omega)$ can be written as

$$f = \sum_{i=1}^{\infty} \langle f, v_i \rangle v_i,$$

with $\Delta v_i = \lambda_i v_i$ for $i = 1, 2, \ldots$ The eigenvalues λ_i are real and negative and $\lambda_i \to -\infty$ as $i \to \infty$. The eigenvalue λ_1 is simple and the corresponding eigenfunction v_1 is positive (if we make the right choice; note that $-v_1$ is also an eigenfunction, so "is of one sign" is a slightly more accurate formulation). In fact, this positivity characterizes v_1: if $\lambda_i \neq \lambda_1$ then v_i cannot be chosen to be positive!

A cautionary note: often results are states for $-\Delta$ and then the eigenvalues are positive and converge to $+\infty$ for $i \to \infty$.

One can prove this result by first showing that a Green's function exists, and next using this function to convert the boundary value problem for the differential equation into an integral equation. Then general spectral theory of compact self-adjoint operators can be used. The positivity follows via the Krein-Rutman theorem, which is the infinite-dimensional version of Perron-Frobenius.

The idea of a *principal eigenvalue* with corresponding positive eigenfunction extends to operators of the form $Lu = \Delta u + ru$ where r is a function of x, rather than a scalar. To determine the sign of the principle eigenvalue (in order to decide about growth or decline) is a nontrivial task.

In the case of a one-dimensional spatial variable, this is part of the so-called Sturm-Liouville theory. The no-flux boundary condition is treated just as easily as the zero Dirichlet case (just compare the Exercises 5.2.4 and 5.2.5).

For higher dimensional Ω, one needs a bit of regularity of $\partial\Omega$ when dealing with no-flux boundary conditions. It is remarkably hard to find a precisely formulated result for the case of a no-flux boundary condition in the literature. After extensive searching we found Chapter 11, §A in Smoller (1983).

Finally, note that there also exist variational characterizations of the eigenvalues and eigenfunctions and these are particularly useful for dealing with the principal eigenvalue.

5.3 The asymptotic speed of propagation

This exercise is, in a way, a continuation of the preceding one. But now we consider a biological population living in a very large domain. In fact, the domain

is so large that we use the plane $\mathbb{R}^2$ as an idealised mathematical description. So consider

$$\frac{\partial u}{\partial t} = d\Delta u + ru, \quad r > 0, \tag{5.3.1}$$

where $\Delta u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}$ and $x = (x_1, x_2) \in \mathbb{R}^2$. There are many situations in which one wants to know how fast the area occupied by the population expands. We shall derive the answer in two quite different ways. The first consists of analysing the fundamental solution

$$u(t,x) = \frac{1}{4\pi dt} e^{rt - \frac{|x|^2}{4dt}}, \quad \text{where } |x|^2 = x_1^2 + x_2^2. \tag{5.3.2}$$

describing the effect of a release at $t = 0$ at $x = 0$. The second relies on a search for travelling plane wave solutions, i.e., solutions of the form

$$u(t,x) = w(x \cdot \nu - ct), \tag{5.3.3}$$

where $w : \mathbb{R} \to \mathbb{R}$ defines the *profile*, the unit vector $\nu \in \mathbb{R}^2$ the *direction* and the real number c the *speed*.

EXERCISE 5.3.1. With u given by (5.3.2), for fixed x we have $\lim_{t\to\infty} u(t,x) = \infty$, while for fixed t we have $\lim_{|x|\to\infty} u(t,x) = 0$. To find out where, roughly, the transition from 0 to ∞ is located, we can consider $\lim_{t\to\infty} u(t,x)$ under various assumptions about how fast $|x(t)| \to \infty$ as $t \to \infty$.

(i) Show that this limit equals zero if $|x(t)|^2 > (4dr + \varepsilon)t^2$ for some $\varepsilon > 0$.
(ii) Show that, on the other hand, this limit equals ∞ if $|x(t)|^2 < (4dr - \varepsilon)t^2$ for some $\varepsilon > 0$.
(iii) Give arguments in favour of the assertion: "the population expands with speed $2\sqrt{dr}$".
(iv) Substitute (5.3.3) into (5.3.1) and derive an equation for w in which c figures as an additional (to d and r) parameter. Why did the ν drop out?
(v) Try for w an exponential function. Express the exponent in terms of c, r and d.
(vi) The biological interpretation requires w to be positive. This condition leads to a lower bound for the wave speed c. Which bound is this?
(vii) Comparing answers to (iii) and (vi), you find that the minimal speed of plane wave solutions coincides with the population expansion speed as derived from (5.3.2). Can you give an intuitive argument why this is to be expected? Hint: think in terms of fireworks that are ignited by fuses that we make as long or short as we want but that can also, via connections, be ignited by nearest neighbours.
(viii) Consider the plane wave solution with minimal speed. Check that at a fixed position the population grows like e^{2rt}, whereas a uniform (i.e., position independent) population grows as e^{rt}. Can you explain where the difference stems from?

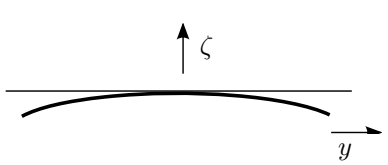


FIGURE 5.2.

We conclude this section with a more general look at the speed of propagation, without using the travelling wave Ansatz. Let us try to zoom in at the transition

region by choosing a fixed direction ζ of unit length, and write $x = \alpha(t)\zeta + y$, with $\zeta \cdot y = 0$, where α is a "local" one- dimensional coordinate corresponding to the ζ direction (Figure 5.2). With these assumptions, $|x|^2 = \alpha^2 + |y|^2$, and hence

$$u(t,x) = \frac{1}{4\pi dt} e^{rt\left(1-\frac{\alpha^2}{4rdt^2}\right)} e^{-\frac{|y|^2}{4dt}}.$$

For y in a bounded set, the last factor converges to 1 uniformly as $t \to \infty$. We would like to know at what speed $\alpha(t)$ has to progress such that the limit will be different from both zero and infinity. Call this limit ψ. Put $\frac{1}{4\pi dt} e^{rt\left(1-\frac{\alpha^2}{4rdt^2}\right)} = \psi$. Then solving for α^2, we find

$$\alpha^2 = 4drt^2\left(1 - \frac{\log 4\pi dt}{rt} - \frac{\log \psi}{rt}\right),$$

and hence

$$\alpha = 2\sqrt{dr}t - \sqrt{\frac{d}{r}}\left(\log(4\pi dt) - \log(\psi)\right) + \mathcal{O}\left(\frac{\log^2(t)}{t}\right).$$

We write this as $\alpha = m(t) + \theta + \mathcal{O}\left(\frac{\log^2(t)}{t}\right)$, with $\theta = -\sqrt{\frac{d}{k}}\log(\psi)$. The new function $m(t)$ satisfies the differential equation

$$\dot{m}(t) = 2\sqrt{dr} - \sqrt{\frac{d}{r}}\frac{1}{t}.$$

Thus we see that the speed at which α needs to proceed converges algebraically to $2\sqrt{dr}$. Note that $\theta = -\sqrt{\frac{d}{r}}\log(\psi) \iff \psi = e^{-\theta\sqrt{\frac{r}{d}}}$.

Since ζ is arbitrary, we conclude that the fundamental solution u decomposes into plane waves travelling in all directions with speed $2\sqrt{dr}$, and that these waves describe the transition from inside the critically growing ball ($\psi \to \infty$, $\theta \to -\infty$), to outside ($\psi \to 0$, $\theta \to \infty$).

Travelling waves derive from the combination of a homogeneous medium and time translation invariance. The waves (in particular, their speeds) are independent of the direction ζ since the medium is isotropic.

On finite but large domains we still can use *self-similar solutions* (here travelling waves) to describe the *intermediate asymptotics* when the details of the initial condition do not matter anymore while boundary conditions do not yet influence the dynamics in a substantial way. For "self-similar", see Grindrod, box E (Grindrod, 1991), but also the book by Barenblatt devoted to self-similarity and intermediate asymptotics (Barenblatt, 1996). For each c the equation is invariant with respect to a group of transformations

$$T_\varepsilon^c = \begin{cases} x & \to & x + \varepsilon c, \\ t & \to & t + \varepsilon, \\ u & \to & u. \end{cases}$$

Hence, given a solution we can generate other (possibly, but not necessarily, different) solutions by applying T_ε^c. A similarity solution is one for which the group orbit $T_\varepsilon^c u$ consists of only one point.

Chapter 6

Reaction-diffusion equations

6.1 Introduction

It happens rarely that the changes in time of some quantity are due to just one mechanism: as a rule several mechanisms are involved. In a finite time interval the contributions of the various mechanisms are entangled. The great success of differential equations as a modelling tool derives to a large extent from the fact that in infinitesimal time intervals (by which we just mean that we let the length of the considered time interval go to zero) the contributions to the *rate* of change become independent and can simply be added. So in the modelling phase we can concentrate on one mechanism at a time and derive the corresponding term for the ultimate differential equation. The *solutions* of the differential equation then take into account the joint, intertwined influence of all mechanisms.

In the present chapter we consider the system of equations

$$\frac{\partial u}{\partial t} = d\Delta u + f(u) \tag{6.1.1}$$

where

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{pmatrix} \tag{6.1.2}$$

is a vector of, say, the concentrations of k different chemical substances that are subject to both diffusion and reaction. The function $f : \mathbb{R}^k \to \mathbb{R}^k$ describes the velocities and the stoichiometry of the various reactions. The Laplacian acts componentwise and d is a diagonal matrix with elements $d_1, \ldots, d_k$, so $d\Delta u$ is the vector

$$\begin{pmatrix} d_1\Delta u_1 \\ d_2\Delta u_2 \\ \vdots \\ d_k\Delta u_k \end{pmatrix} \tag{6.1.3}$$

If the space variable

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$$

has m components, then

$$\Delta u_i = \frac{\partial^2 u_i}{\partial x_1^2} + \frac{\partial^2 u_i}{\partial x_2^2} + \cdots + \frac{\partial^2 u_i}{\partial x_m^2} \tag{6.1.4}$$

In Chapters 2 and 3 we concentrated on the reaction part and ignored spatial dependence. In Chapter 4 we concentrated on the diffusion part and ignored reactions. At the level of *formulation* we do not need to invest further work: (6.1.1) is obtained by adding the contributions to the rate of change of u. But at the level of model *analysis*, things are not at all clear. What new phenomena can we expect? And are the values of k and m important in this respect? And what about boundary conditions?

Concerning the last point, we shall give most attention to two cases in which the boundary conditions per se do not have any "forcing" influence:

- a bounded domain Ω with no-flux boundary conditions at $\partial\Omega$
- $\Omega = \mathbb{R}^m$ (with very mild boundedness conditions which are usually not even mentioned)

(But for $k = 1$, $m = 1$, $\Omega =$ interval, we shall also study the case of zero Dirichlet boundary conditions, to get at least a feel for the difference between no-flux and zero Dirichlet boundary conditions.)

The new phenomena that we shall encounter are:

- pattern formation
- growth (or decay) by way of travelling waves, i.e., growth by way of spatial expansion

But for mathematicians a very important point is also that we need to enlarge our toolbox, as we enter into the realm of *infinite* dimensional dynamical systems (the infinity aspect of a partial differential equation stems from the fact that $x \in \mathbb{R}$, so for each fixed x we in fact have a differential equation, coupled to all the others; another way of looking at it is that, for fixed time t, solutions to partial differential equations are *functions* $u(t, \cdot)$, which depend on infinitely many $x \in \mathbb{R}$). Due to the smoothing effect of diffusion, not that much changes though. If we consider, for instance, linearised stability, then, in a sense, the main difference is that we have to analyse countably many matrices rather than just one.

6.2 Stability criteria for uniform steady states

A solution of (6.1.1) that is independent of time is called a *steady state*. If that solution is also independent of spatial position, we speak about a *uniform steady state*. If we denote both such a solution and the value it takes in $\mathbb{R}^k$ by $\bar{u}$, then we should have

$$f(\bar{u}) = 0 \tag{6.2.1}$$

For $k = 1$, we can find solutions of (6.2.1) by a graphical analysis and, in one go, also determine their stability with respect to the reaction dynamics with spatial dependence ignored. See Figure 6.1. The analytical criterion is

$$Df(\bar{u}) < 0 \implies \bar{u} \text{ is stable}$$
$$Df(\bar{u}) > 0 \implies \bar{u} \text{ is unstable}$$

For $k > 1$, (6.2.1) is short hand for k equations in as many unknowns and in the absence of space dependence the stability can be determined from the eigenvalues of the Jacobian matrix $Df(\bar{u})$:

- if $\Re\lambda < 0$ for *all* eigenvalues λ of $Df(\bar{u})$, then $\bar{u}$ is (locally asymptotically; in fact exponentially) stable

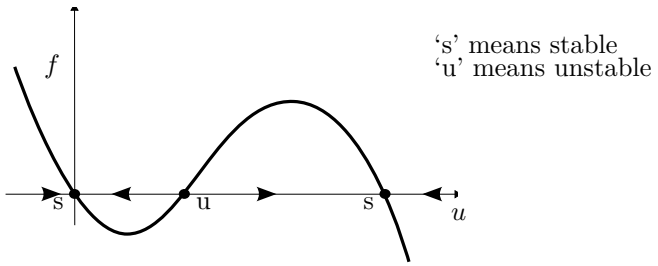


Figure 6.1.

- if $\Re\lambda > 0$ for *some* eigenvalue λ of $Df(\bar{u})$, then $\bar{u}$ is unstable

The proof is based on linearisation, i.e., on an analysis of the linearised equation

$$\frac{dv}{dt} = Df(\bar{u})v \tag{6.2.2}$$

combined with estimates for higher order terms (note that the proof of stability is much easier than the proof of instability, since in general an unstable steady state is a saddle point and there are solutions that actually do approach the steady state for $t \to \infty$). The connection between (6.2.2) and the eigenvalue problem

$$Df(\bar{u})\bar{v} = \lambda\bar{v} \tag{6.2.3}$$

is separation of variables (with variables t and the index indicating the component): if we substitute

$$v(t) = \psi(t)\bar{v} \tag{6.2.4}$$

with $\psi(t) \in \mathbb{R}$ and $\bar{v} \in \mathbb{R}^k$ into (6.2.2) we find that

$$\frac{\psi'}{\psi}\bar{v} = Df(\bar{u})\bar{v}$$

which can hold only if ψ'/ψ is a constant (i.e., independent of t), say λ. Hence $\psi(t) = \psi(0)e^{\lambda t}$, and (6.2.3) must hold.

Now, let us include the diffusion term and investigate its impact. If we supply (6.1.1) with no-flux boundary conditions

$$\frac{\partial u}{\partial n}\Big|_{\partial\Omega} = 0 \tag{6.2.5}$$

then solutions of (6.2.1) yield uniform steady states. The linearised equation now reads

$$\frac{\partial v}{\partial t} = d\Delta v + Df(\bar{u})v \tag{6.2.6}$$

$$\frac{\partial v}{\partial n}\Big|_{\partial\Omega} = 0 \tag{6.2.7}$$

We apply, as before, separation of variables, but this time there are *three* variables: t, x, and the index that indicates the component. So we substitute

$$v(t,x) = \psi(t)\phi(x)\bar{v} \tag{6.2.8}$$

with $\psi(t)$, $\phi(x) \in \mathbb{R}$, and $\bar{v} \in \mathbb{R}^k$. After division by $\psi(t)\phi(x)$ we obtain

$$\frac{\psi'(t)}{\psi(t)}\bar{v} = \frac{\Delta\phi(x)}{\phi(x)}d\bar{v} + Df(\bar{u})\bar{v} \tag{6.2.9}$$

which requires that ψ'/ψ and $\Delta\phi/\phi$ are constant. Hence, as before, $\psi(t) = \psi(0)e^{\lambda t}$. With foresight (inspired by Section 5.2) we call the constant value that $\Delta\phi/\phi$ takes

$-\mu^2$, so require that

$$\Delta\phi = -\mu^2\phi \tag{6.2.10}$$

$$\frac{\partial\phi}{\partial n}\Big|_{\partial\Omega} = 0 \tag{6.2.11}$$

and, in addition,

$$[Df(\bar{u}) - \mu^2 d]\bar{v} = \lambda\bar{v} \tag{6.2.12}$$

Note the order: we can determine the relevant values of μ first by studying (6.2.10), and only after that determine, for each relevant μ, the values of λ that satisfy

$$\det\left(Df(\bar{u}) - \mu^2 d - \lambda I\right) = 0 \tag{6.2.13}$$

with I the $k \times k$ identity matrix. Equation (6.2.13) is often called a *dispersion relation*, as it links a characteristic of the time dependence, λ, to a characteristic of the space dependence, μ.

We now first concentrate on (6.2.10)–(6.2.11) for bounded Ω. In Section 5.2 we dealt with the case $m = 1$, $d = 1$, $\Omega = [0, L]$, and found that μ should be of the form

$$\mu = \frac{k\pi}{L}, \quad k = 0, 1, 2, \ldots \tag{6.2.14}$$

This generalises in the sense that for

$$\mu_0 = 0 \tag{6.2.15}$$

we have a solution $\phi =$ constant and that there exist μ_1, μ_2, ..., with $\mu_{i+1} > \mu_i$ for $i = 0, 1, 2, \ldots$, with corresponding λ_i determined by the dispersion relation (6.2.13), for which (6.2.10)–(6.2.11) has a nontrivial solution while there is no such solution for any other value of μ. The mathematical background of this result has various facets (see e.g., (Renardy and Rogers, 1993)):

- elliptic differential equations
- self-adjoint operators with compact resolvent
- positivity

Unless Ω is symmetric, for instance a rectangle in $\mathbb{R}^2$, it is not feasible to determine the μ_i for $i \geq 1$ explicitly. But the fact that we know that $\mu_0 = 0$ is the smallest of all μ_i is often very helpful!

In the special case $k = 1$, i.e., a *scalar* equation, (6.2.13) reads

$$\lambda = Df(\bar{u}) - d\mu^2$$

and it follows that

- for every μ there is exactly one λ, which is real
- the λ's are *ordered* exactly as the $-\mu^2$
- $\lambda = Df(\bar{u})$ is the largest

The so-called *Principle of Linearised Stability*, see the "Theorem" below, now implies that $\bar{u}$ is exponentially stable if $Df(\bar{u}) < 0$ and unstable if $Df(\bar{u}) > 0$ (we often formulate this as: $\bar{u}$ is linearly stable iff $Df(\bar{u}) < 0$).

"THEOREM" (Principle of Linearised Stability)
(i) if for every *eigenvalue* $-\mu^2$ *of the diffusion problem with no-flux boundary conditions,* every *eigenvalue* λ *of* $Df(\bar{u}) - \mu^2 d$ *has negative real part, then* $\bar{u}$ *is a (locally exponentially)* stable *steady state.*
(ii) if for some *eigenvalue* $-\mu^2$ *of the diffusion problem with no-flux boundary conditions,* some *eigenvalue* λ *of* $Df(\bar{u}) - \mu^2 d$ *has positive real part, then* $\bar{u}$ *is an* unstable *steady state.*

Several questions now come to mind:

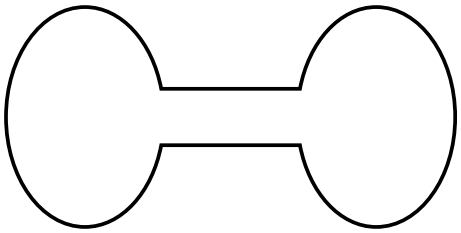


Figure 6.2.

- if $k > 1$ and all eigenvalues of $Df(\bar{u})$ have negative real part, does it follow that $\Re\lambda < 0$ for all λ that satisfy (6.2.13) for some $\mu = \mu_k$? The, perhaps surprising, answer is: NO. It was Alan Turing's great idea that diffusion driven instability is possible in the case of *systems* of equations provided the diffusion constants of the various components are sufficiently different. In Section 6.6 you shall demonstrate this in detail. The bottom line is that *pattern formation* takes place when reaction-stable uniform steady states turn unstable due to differences in the diffusion constants of the various components
- for $k = 1$ and no-flux boundary conditions, is it possible to have a *stable non-uniform* steady state? The short answer is: no, unless you force it by combining bistable dynamics with a special domain shape involving almost disconnected components. In more detail:
  (i) if $m = 1$, $\Omega = [0, L]$, then no, see Section 6.3
  (ii) if $m > 1$, and Ω is convex, then no, see (Kishimoto and Weinberger, 1985)
  (iii) if $m = 2$ and Ω is a "halter" domain (see Figure 6.2) and f has at least two stable steady states, then yes, provided the connecting pipe line is thin enough (the non-uniform steady state is close to different reaction-stable steady state values in the two ball-like parts of the domain; see (Matano, 1979))

6.3 Scalar Reaction-Diffusion equations: global bifurcation theory based on phase plane analysis and symmetry arguments

One reason to focus on steady states and their stability is that the corresponding analysis is relatively easy. But for scalar equations there is a better reason: solutions do converge to steady states. To demonstrate this, we introduce the Lyapunov function

$$V(\phi) = \frac{1}{2}\int_0^L (\phi'(x))^2\, dx - \int_0^L F(\phi(x))dx \tag{6.3.1}$$

where

$$F(w) := \int_0^w f(\sigma)d\sigma \tag{6.3.2}$$

Let $u = u(t, x)$ be a solution of

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u) \tag{6.3.3}$$

$$\frac{\partial u}{\partial x}(t, 0) = 0 = \frac{\partial u}{\partial x}(t, L) \tag{6.3.4}$$

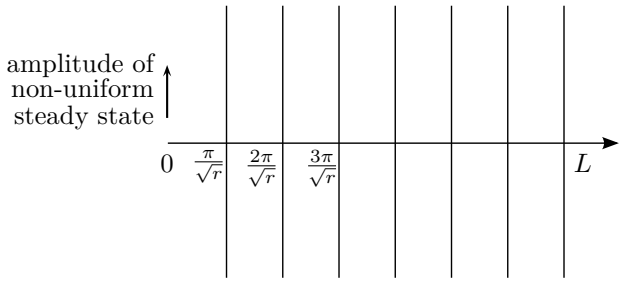


FIGURE 6.3.

(Note that we scaled the spatial variable such that the diffusion constant d equals 1.) Then, performing a partial integration, and using the boundary conditions,

$$\begin{aligned}\frac{d}{dt}V(u(t,\cdot)) &= \int_0^L \left\{ \frac{\partial u}{\partial x}(t,x)\frac{\partial^2 u}{\partial t \partial x}(t,x) - f(u(t,x))\frac{\partial u}{\partial t}(t,x) \right\} dx \\ &= -\int_0^L \left(\frac{\partial^2 u}{\partial x^2}(t,x) + f(u(t,x)) \right) \frac{\partial u}{\partial t}(t,x) dx \\ &= -\int_0^L \left(\frac{\partial u}{\partial t}(t,x) \right)^2 dx \\ &\leq 0\end{aligned}$$

This rules out the possibility of time-periodic solutions or any other kind of persistent behaviour for which $\frac{\partial u}{\partial t}$ is not identically zero. So only steady states are potentially attractors. Note that local minima of V are *stable* steady states. It appears that non-constant steady states are saddle points of V.

The function V is also a Lyapunov function when we impose Dirichlet rather than no-flux boundary conditions. Moreover, using one of Green's formulas, the proof is easily extended to the case of higher space dimension, i.e., $m > 1$.

We now know that non-uniform steady states cannot be stable when $k = 1$, $m = 1$, $\Omega = [0, L]$ and we impose no-flux boundary conditions. But do they exist?

In Section 5.2 we found that in the *linear* case the diagram in Figure 6.3 summarizes the situation: if we consider the growth rate r as fixed and the length of the domain L as a parameter, then a non-uniform steady state only exists if

$$L = L_k = \frac{k\pi}{\sqrt{r}}, \quad k = 1, 2, \ldots$$

and it is then, modulo a multiplicative constant, given by

$$\bar{u}(x) = \bar{u}_k(x) = \cos\left(\frac{k\pi}{L}x\right)$$

The physicists jargon is that $\bar{u}_k$ is the k-th *spatial mode* and that this mode turns unstable if L is increased beyond L_k.

In this section we replace the linear function $u \mapsto ru$ by the nonlinear function $u \mapsto f(u)$ and next investigate how Figure 6.3 deforms as a result of the nonlinearity. We shall find Figure 6.4, which we call a *bifurcation* diagram (cf. Appendix) with bifurcation parameter L. The choice of L as the key parameter is somewhat arbitrary: it is easy to translate Figure 6.4 into a bifurcation diagram with parameter either the diffusion constant d or the derivative $f'(0)$. Indeed, by scaling we can transform the equation

$$u_t = u_{xx} + f(u) \quad x \in [0, L],$$

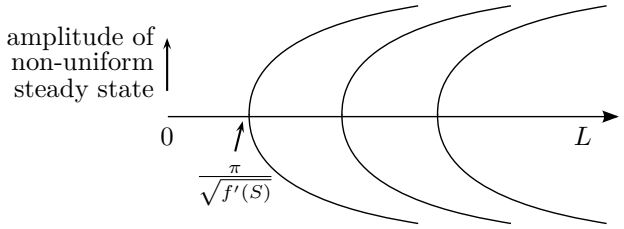


Figure 6.4.

into

$$u_t = \frac{1}{L^2} u_{\xi\xi} + f(u) \quad (\text{taking } \xi = x/L,\ \xi \in [0,1])$$

and next into

$$u_\tau = u_{\xi\xi} + L^2 f(u) \quad (\text{taking } \tau = t/L^2,\ \xi \in [0,1])$$

To investigate the steady state problem, we first rewrite the second order equation, $u_{xx} + f(u) = 0$ as a first-order system of ODEs:

$$u_x = v \tag{6.3.5}$$

$$v_x = -f(u) \tag{6.3.6}$$

The point is that we can analyse this first order system by phase plane methods (so we are going to look at (6.3.5)–(6.3.6) from a dynamical systems point of view, but note that this is just an auxiliary tool and that it has *nothing* to do with the infinite dimensional dynamical system generated by the diffusion equation (6.1.1) with appropriate boundary conditions!)

With, as defined in (6.3.2),

$$F(u) := \int_0^u f(\sigma) d\sigma$$

we find that this first order system has a conserved quantity,

$$H(u,v) := \frac{1}{2} v^2 + F(u), \tag{6.3.7}$$

also called a Hamiltonian. Since $v^2 = (-v)^2$, the phase portrait is symmetric with respect to reflection in the u-axis. Orbits are mapped by $(u,v) \mapsto (u,-v)$ to orbits, which, however, are traversed in the opposite direction. If f happens to be antisymmetric in u (i.e., $f(-u) = -f(u)$), then $F(-u) = F(u)$, and we also have symmetry of the phase plane with respect to reflection in the v-axis. Orbits are now also mapped by $(u,v) \mapsto (-u,v)$ to orbits.

We now first show that there are no bifurcations from a *stable* (with respect to well-stirred dynamics) uniform steady state. Assume that $f(0) = 0$, and $f'(0) < 0$. Then F has a maximum in $u = 0$, and consequently the origin in phase space is a saddle point (Figure 6.5) in the sense of dynamical systems (and also a saddle point as a critical point of the function H).

Therefore, there are locally near the origin neither connections from the v-axis to the v-axis nor connections from the u-axis to the u-axis. So for both the boundary value problem with no-flux conditions and for the zero Dirichlet boundary value problem we can conclude that bifurcations from such a steady state are impossible.

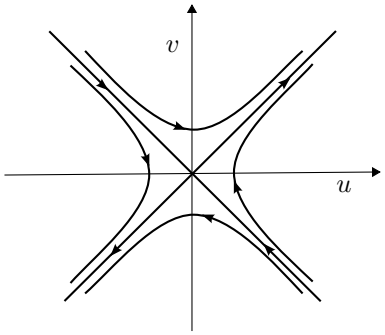


FIGURE 6.5.

EXERCISE 6.3.1. Use the principle of linearised stability, to show that the stable uniform state 0 remains stable if we add diffusion and add either no-flux BCs or compatible (i.e., zero in this case) Dirichlet BCs. In particular, show that the eigenvalues are just the eigenvalues of the Laplacian shifted over $f'(0)$, so in the negative direction, making the new solution even more stable.

Our next aim is to use phase plane analysis to derive the bifurcation diagram depicted in Figure 6.4 for the no-flux nonlinear boundary value problem

$$u_{xx} + f(u) = 0$$
$$u_x(0) = 0 = u_x(L)$$

with bifurcation parameter L. We assume that the graphs of f and F (recall (6.3.2)) have the form shown in Figure 6.6.

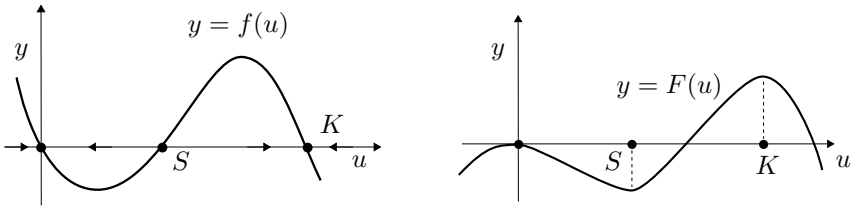


FIGURE 6.6.

Note that $u = 0$ and $u = K$ are stable as steady states for the dynamical systems generated by the ODE $\dot{u} = f(u)$, while the unstable steady state $u = S$ separates their domains of attraction. Also note that we assumed that $F(K) > 0$ (in other words, that the area below the u-axis and above the graph of f is less than the area above the u-axis and below that graph of f, if we consider the interval $[0, K]$). The consequence is that the phase portrait is as depicted in Figure 6.7.

In particular: the family of closed orbits surrounding $(S, 0)$ "ends" in a homoclinic loop issuing from the saddle $(0, 0)$. In an ecological context, u represents the density of a population subject to an Allee effect (meaning that it is bound to go extinct for low densities but that, due to positive density dependence, it grows to the carrying capacity K if abundant enough; the underlying mechanism may be sexual reproduction, so the difficulty of finding a mate when the species is rare in the considered area).

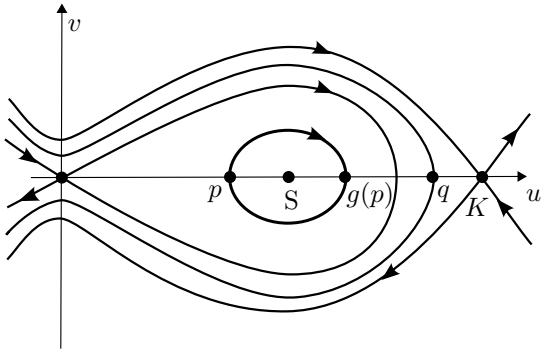


Figure 6.7.

We parameterise the family of closed orbits by the minimum value that u takes and we shall denote this minimum value by p. Note that the corresponding value of H equals $F(p)$. We denote

- the corresponding maximum value of u by $g(p)$
- the corresponding period by $T(p)$

Note that

$$T(p) = 2\int_p^{g(p)} \frac{du}{\sqrt{2(F(p) - F(u))}} \tag{6.3.8}$$

Exercise 6.3.2. (i) Derive this identity
(ii) Show that $\lim_{p\downarrow 0} T(p) = \infty$
(iii) Show that $\lim_{p\uparrow S} T(p) = \frac{2\pi}{\sqrt{f'(S)}}$

The question whether or not T is a monotone function of p is, in general, not easy to answer (see e.g. (Chicone, 1987)). But in any case, we know that the range of T includes the interval $(\frac{2\pi}{\sqrt{f'(S)}}, \infty)$.

Now suppose that $2L$ belongs to the range of T. Because of the symmetry of the phase portrait with respect to reflection in the u-axis, we know that it takes "as long" for u to increase from p to $g(p)$ as it takes u to decrease subsequently again from $g(p)$ to p. So if $T(p) = 2L$, each of these changes happens during a stretch L of the independent variable. The corresponding solutions are indicated by N_1 and they are sketched in Figure 6.8. The index specifies the number for intervals on which the solution is monotone. If we denote one solution by u_+ and the other by u_-, then

$$u_+(x) = u_-(L - x)$$

or, in other words, one solution is obtained from the other by a reflection in the midpoint of the interval.

The solutions indicated by N_2 correspond to p such that $T(p) = L$, so they correspond to a full turn. The midpoint of the x-interval is reached after half a turn, so these solutions are themselves symmetric with respect to reflection in the midpoint. Both solutions have exactly one interval of increase and one of decrease, but if we first decrease and then increase the maximum is at the boundary and the minimum in the interior, while with the other order it is the other way around (see Figure 6.8).

The solutions indicated by N_3 correspond to p such that $\frac{3}{2}T(p) = L$, so to $1\frac{1}{2}$ turns. In general we consider p such that $\frac{k}{2}T(p) = L$ and solutions that make

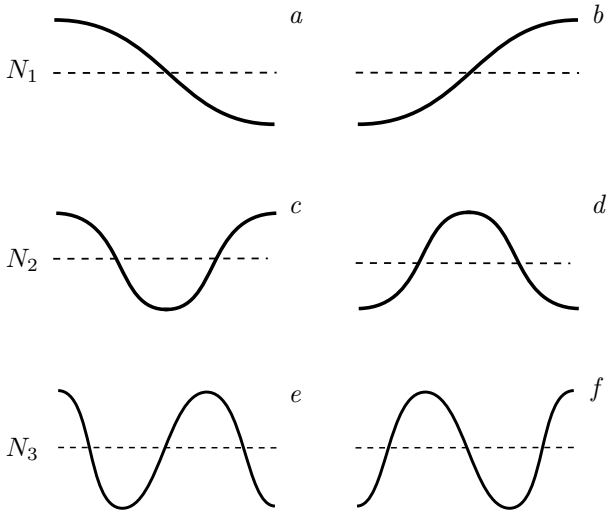


FIGURE 6.8.

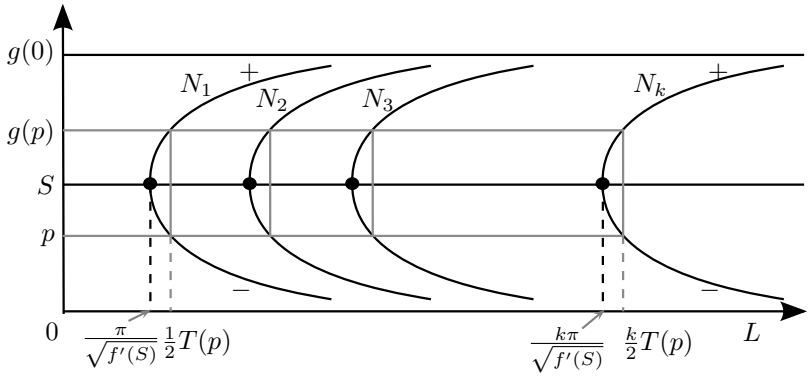


FIGURE 6.9.

$\frac{k}{2}$ turns.[1] For k even these are symmetric, while for k odd the two solutions are related to each other by a symmetry.

Assuming that $p \mapsto T(p)$ is monotone we can now draw a more detailed version of Figure 6.4, shown in Figure 6.9.

If T is *not* monotone, there are wiggles in these branches.

How should we interpret this diagram in the context of the infinite-dimensional dynamical system generated by (6.3.3)–(6.3.4)? The constant steady state $u \equiv S$ is now a *saddle point* (recall (6.3.1) and note that the term involving F has the opposite sign as the term involving F in the formula (6.3.7) for the Hamiltonian H). For small L it has a one-dimensional unstable manifold, but as L increases the dimension of this manifold increases (or, in more physical jargon, more and more modes turn unstable). Presumably, the domains of attraction of the stable constant (i.e., spatially uniform) steady states $u \equiv 0$ and $u \equiv K$ are separated by the union of the stable manifolds of $u \equiv S$ and all the existing non-constant steady

[1]Above we used k to denote the number of components of a system of equations. Below we shall use the symbol k to specify a mode number, i.e., to indicate the number of minima and maxima of a steady state solution. The aim of this warning is to avoid that you get confused.

states "around" it. So the bifurcations change the structure of the flow within this separatrix.

Consider once more the zero-flux nonlinear boundary value problem

$$u_{xx} + f(u) = 0, \tag{6.3.9}$$

$$u_x(0) = 0 = u_x(L), \tag{6.3.10}$$

with parameter L. We now want to derive relations between solutions by using extension and symmetry arguments instead of phase plane analysis. (The key advantage of such arguments is that they also work for systems.)

Exercise 6.3.3. Show that if u is a solution of (6.3.9)–(6.3.10), so is w, defined by $w(x) = u(L - x)$. We call u *symmetric* if $w = u$. What symmetry does this amount to?

Exercise 6.3.4. Assume $f(0) = 0$, $f'(0) > 0$. Show that then bifurcations occur if $L = \frac{k\pi}{\sqrt{f'(0)}}$. Consider $k = 1$. Give arguments in favour of the claim that the bifurcating solutions are *not* symmetric. Show that, consequently, the bifurcation must be a pitchfork.

Exercise 6.3.5. Whenever u is a solution, extend it to a $2L$-periodic function by

$$\begin{aligned} &u(-x) := u(x), && 0 \le x \le L, \\ &u(x + 2L) := u(x). \end{aligned}$$

Show that the extension is a solution for parameter value kL, $k = 1, 2, 3, \ldots$. Conclude that the branch bifurcating for $k = 1$ repeats itself for every higher value of k.

Exercise 6.3.6. Show that u is symmetric if and only if the extension has period L. Show that all branches corresponding to even k consist of symmetric solutions. How are in that case the two solutions (one for each subbranch) related to each other?

Exercise 6.3.7. Show that *some* branches for the problem with *Dirichlet* boundary conditions $u(0) = S = u(L)$ can be obtained from extended solutions of the zero flux boundary value problem.

Next, let us look at the situation where there is a big monster at the boundary, so where we replace the no-flux boundary conditions by the zero-Dirichlet conditions

$$u(0) = 0 = u(L) \tag{6.3.11}$$

In terms of the phase portrait Figure 6.7 this means that, in search for steady states, we look for pieces of orbits that start and end at the v-axis.

A glance at Figure 6.7 shows that we can parameterise candidate orbits by the maximum q of u, with

$$g(0) < q < K$$

and that the corresponding steady state solutions of the boundary value problem are symmetric with respect to reflection in the midpoint $x = \frac{L}{2}$ where the maximum q is assumed. Let us denote by $\tilde{T}(q)$ the "time" it takes to arrive at the negative v-axis when starting at the positive v-axis. Then

$$\tilde{T}(q) = 2 \int_0^q \frac{du}{\sqrt{2(F(q) - F(u))}}$$

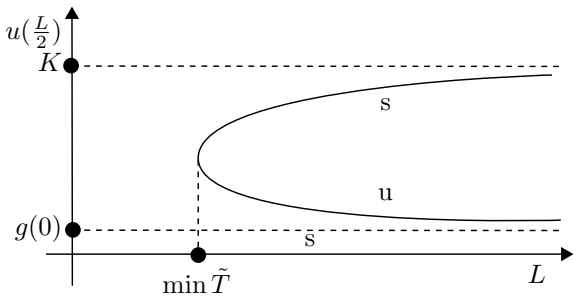


FIGURE 6.10.

and, as movement slows down near saddle points, necessarily

$$\lim_{q \downarrow g(0)} \tilde{T}(q) = \infty$$

$$\lim_{q \uparrow K} \tilde{T}(q) = \infty$$

So $\tilde{T}$ assumes a minimum and the equation $\tilde{T}(q) = L$ has for

$$L < \min \tilde{T} \quad \text{no solution}$$

$$L > \min \tilde{T} \quad \text{at least two solutions}$$

If there are *exactly* two solutions of $\tilde{T}(q) = L$ for L larger than the minimum of $\tilde{T}$, then the bifurcation diagram has the form shown in Figure 6.10.

So there is a saddle-node bifurcation for $L = \min \tilde{T}$ at which a stable steady state and an unstable (saddle) steady state are born. Presumably the stability character of these two steady states does not change when L is further increased. The stable manifold of the saddle steady state serves again as a separatrix (note that $u \equiv 0$ is a stable steady state for all L). By using Maximum Principle arguments, cf. (Aronson and Weinberger, 1975, 1978; Ludwig et al., 1979), or more precisely, by constructing sub- and supersolutions, one can get some partial information about the initial conditions for (6.3.3) and (6.3.11) that yield solutions converging to either $u \equiv 0$ or to the stable non-uniform steady state. Note that for very large L the values that the latter takes are very close to K on most of the interval.

Our main conclusion is that the population can persist, despite the big monster at the boundary, *provided* the domain is large enough.

6.4 The non-existence of patterns for scalar equations

We will now show that non-trivial equilibrium solutions ('patterns') to scalar reaction-diffusion equations subject to Neumann boundary conditions can never be stable. This is often loosely summarised with the phrase in the title of this section.

Consider

$$\begin{cases} u_t = u_{xx} + f(u) & \text{on } [0, L] \\ u_x(t, 0) = u_x(t, L) = 0 \end{cases} \tag{6.4.1}$$

Let $U(x)$ be a non-trivial equilibrium solution, i.e., $U(x) \not\equiv U_0$. The eigenvalue problem for V, which describes the linear stability of U, is given by

$$\begin{cases} V_{xx} + [\frac{\partial f}{\partial u}(U(x)) - \lambda]V = 0 & \text{on } [0, L] \\ V_x(0) = V_x(L) = 0 \end{cases} \tag{6.4.2}$$

This is a standard Sturm-Liouville problem. The eigenvalues, denoted with λ_i^N to signify that they belong to the Neumann problem, are again all real and simple,

$-\infty < \ldots < \lambda_2^N < \lambda_1^N < \lambda_0^N$, and the eigenfunction corresponding with λ_i has i zeroes. We need to show that $\lambda_0^N > 0$, thus showing that $U(x)$ is unstable. Note that a similar ordering of eigenvalues exists for the Dirichlet problem. Also now the i-th Dirichlet eigenvalue, λ_i^D has an eigenfunction with i zeroes.

Now observe that, by differentiating the steady state equation for U with respect to x, we know that $W = U_x$ solves

$$\begin{cases} W_{xx} + \frac{\partial f}{\partial u}(U(x))W = 0 & \text{on } [0, L], \\ W(0) = W(L) = 0. \end{cases}$$

Therefore, problem (6.4.2) with homogeneous Dirichlet boundary conditions instead of homogeneous Neumann boundary conditions has $\lambda = 0$ as eigenvalue. Denoting the eigenvalues of the Dirichlet problem by λ_i^D, $i = 0, 1, \ldots$, we thus know that $\lambda_i^D = 0$ for a certain i.

We can now use the following Lemma, stating that we can order the eigenvalues of two solutions, if we can order the solutions in some manner.

Lemma 6.4.1*:* Let $g(x)$ be given. Let Φ and Ψ satisfy

$$V_{xx} + (g(x) - \lambda)\, V = 0$$

with eigenvalues λ and μ respectively. Assume $\Phi(0) = \Phi(L) = 0$, and $\Phi(x) > 0$ on $(0, L)$, and $\Psi(x) > 0$ on $[0, L]$. Then $\lambda < \mu$.

Proof. Multiplying the eigenvalue equation for Φ by Ψ and vice versa, and subtracting these two, we find

$$\Phi\Psi'' - \Phi''\Psi + (\mu - \lambda)\Phi\Psi = 0.$$

Hence, integrating over $[0, L]$, and using partial integration yields

$$\Psi\Phi'\Big|_0^L - \Phi\Psi'\Big|_0^L + (\mu - \lambda) \int_0^L \Phi\Psi = 0.$$

The last integral is strictly positive by the assumptions on Φ and Ψ. Since Φ is a Dirichlet solution, the second term vanishes. The first term is a priori only non-positive, which would yield $\lambda \leq \mu$. Note, however, that if $\Phi'(0) = 0$ or $\Phi'(L) = 0$, then $\Phi \equiv 0$, and hence by uniqueness of the boundary value problem we find $\Phi'(0) > 0$ and $\Phi'(L) < 0$. Therefore $\lambda < \mu$. □

Recall that the smallest eigenvalue λ_0^N, which is the one we are actually interested in, corresponds to an eigenfunction without zeroes. Therefore, this function solves (6.4.2) and has the properties of Ψ in the lemma. So from the lemma we conclude that $\lambda_0^N > \lambda_0^D \geq \lambda_i^D = 0$, and that indeed $U(x)$ is unstable.

This argument also works in higher space dimensions, provided that the domain Ω is convex. As we discussed in Section 6.2, stable non-uniform steady states do exist if we choose a bistable function f on a halter-shaped domain (recall Figure 6.2).

Now we turn to the scalar Dirichlet problem

$$\begin{cases} u_t = u_{xx} + f(u) & \text{on } [0, L] \\ u(0) = u(L) = 0 \end{cases} \tag{6.4.3}$$

Let again $U(x)$ be an equilibrium solution and assume that $U(x)$ changes sign on $(0, L)$. Then $U(x)$ is again unstable!

To see this, we again study $W = U_x$. Since U changes sign, there exist $0 < x_1 < x_2 < L$ such that $W(x_1) = W(x_2) = 0$. So W solves

$$\begin{cases} w_{xx} + f'(U)w = 0 & \text{on } [0, L], \\ w(x_1) = w(x_2) = 0. \end{cases}$$

Applying Lemma 6.4.1 again to $\Phi = \pm W$ (choose the sign so that $\Phi > 0$), with $\lambda = 0$, and $\Psi = \Phi_0^D$ with $\mu = \lambda_0^D$, restricted to $[x_1, x_2]$, we conclude $\mu = \lambda_0^D > \lambda = 0$, and that U is unstable.

6.5 Travelling waves for mono- and bistable scalar Reaction-Diffusion

Diffusion, as a mechanism to generate signals used in for instance development, is a very slow process, and doesn't work efficiently over large distances. As we already saw in Section 5.3, augmenting diffusion with some kind of reaction, be it multiplication of a species or the interaction between different species or chemicals, can lead to patterns that travel much faster viz. with constant speed. It may therefore come as no surprise that reaction-diffusion is a mechanism that abounds in all kinds of biological areas. In this section we are going to study the existence of travelling wave profiles for nonlinear scalar reaction-diffusion equations, determine at what speeds these can travel, and in what direction.

Two prototype equations will be studied: the monostable and the bistable case. The first is also known as the Fisher-Kolmogorov equation, and is given by

$$u_t = du_{xx} + ku(1-u), \quad x \in \mathbb{R} \tag{6.5.1}$$

which, using the rescaled variables $t^* = kt$ and $x^* = x\sqrt{k/d}$ becomes

$$u_t = u_{xx} + u(1-u). \tag{6.5.2}$$

Here the stars have immediately been dropped. It is a model equation which was originally devised by Fisher to model the spread of a favourable gene in a population (Fisher, 1937). It was simultaneously (and presumably independently) studied by, as Aronson (1985) put it, the famous troika of Kolmogorov, Petrovskii and Piscunov (Kolmogorov et al., 1937). The bistable equation in non-dimensional form is given by

$$u_t = u_{xx} + u(u-a)(1-u), \quad x \in \mathbb{R} \tag{6.5.3}$$

where $0 < a < 1$. As we will see, both these equations admit travelling wave profiles, but the range of speeds with which such waves progress is quite different.

Let us first consider the monostable case. If we ignore space for the moment, the equation reduces to

$$u' = u(1-u).$$

This is the standard model for logistic growth, having $u = 0$ and $u = 1$ as steady states, the first of which is unstable, and the second stable. This suggests it might be possible to find travelling wave profiles $w(z) = u(x - ct)$ that connect 0 and 1 and which travel at speed c. Let us try to find these. Substituting the travelling wave Ansatz, taking also into consideration the choice of behaviour at $\pm\infty$, we obtain

$$-cw' = w'' + w(1-w) \tag{6.5.4}$$

$$\lim_{z\to-\infty} w(z) = 1, \quad \lim_{z\to\infty} w(z) = 0 \tag{6.5.5}$$

Writing (6.5.4) in phase plane form,

$$w' = v \tag{6.5.6}$$

$$v' = -cv - w(1-w) \tag{6.5.7}$$

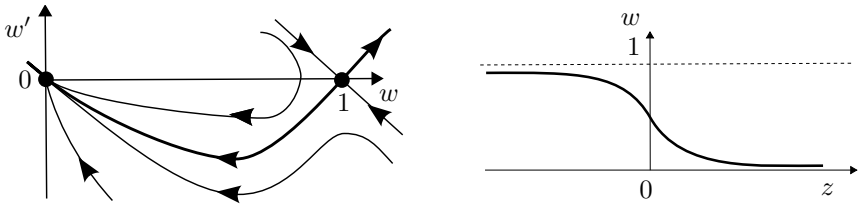


Figure 6.11.

we again find two equilibria, $(w, v) = (0, 0)$ and $(1, 0)$. The Jacobian of this system is

$$\begin{pmatrix} 0 & 1 \\ -1 + 2w & -c \end{pmatrix} \tag{6.5.8}$$

and hence we find the following eigenvalues for the two steady states. For $(0, 0)$,

$$\lambda_{\pm} = -\frac{c}{2} \pm \frac{1}{2}\sqrt{c^2 - 4}$$

whereas for $(1, 0)$,

$$\lambda_{\pm} = -\frac{c}{2} \pm \frac{1}{2}\sqrt{c^2 + 4}$$

Hence, if $c^2 > 4$, the origin is a stable node, and if $c^2 < 4$, it is a stable spiral, giving rise to physically unrealistic solutions since the solutions then become negative. The other steady state, $(1, 0)$, is always a saddle. The phase plane for (6.5.6) for the case $c^2 > 4$, see Figure 6.11, now strongly suggests that the relevant part of the unstable manifold $(1, 0)$ always has to connect to the stable origin, and thus form a heteroclinic orbit connecting $w = 1$ and $w = 0$. Since we can perform this construction for any $c^2 > 4$, we find a continuum of possible wave speeds for travelling waves of the Fisher-Kolmogorov equation (Kolmogorov et al., 1937; Hadeler and Rothe, 1975).

Having argued that travelling waves do exist for a continuum of wave speeds, we may wonder for which initial conditions the solution in time tends to a travelling wave (or a combination of two travelling waves, one moving to the left, and one to the right)? Kolmogorov, Petrovskii and Piscunov (Kolmogorov et al., 1937) already showed in 1937 that initial conditions of the form

$$\begin{cases} w \equiv 1 \text{ for } x < p \\ w \equiv 0 \text{ for } x > p \end{cases}$$

for some $p \in \mathbb{R}$ do indeed tend to travelling wave profiles which travel with minimum speed $c = 2$. But also a localized initial condition, relevant for instance in models of introduced species, grows out to form an expanding block with two fronts, one travelling to the left and the other to the right (see Figure 6.12).

Let us now turn to the bistable equation and repeat the linear stability analysis. For a general scalar reaction diffusion equation

$$u_t = u_{xx} + f(u)$$

the equation for the travelling wave profile $w(z) = u(x - ct)$ is

$$w'' + cw' + f(w) = 0.$$

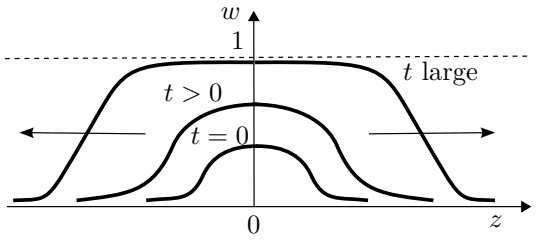


FIGURE 6.12.

Writing this as a two-dimensional system

$$w' = v \tag{6.5.9}$$

$$v' = -cv - f(w) \tag{6.5.10}$$

we have the Jacobian

$$\begin{pmatrix} 0 & 1 \\ -f'(w) & -c \end{pmatrix} \tag{6.5.11}$$

with eigenvalues

$$\lambda_{\pm} = \frac{c}{2} \pm \frac{1}{2}\sqrt{c^2 - 4f'(w)}.$$

For our particular problem, $f(u) = u(u-a)(1-u)$, so there are three equilibria, which for the 2D system are written as $(w, v) = (0, 0)$, $(a, 0)$, or $(1, 0)$. Computing the eigenvalues at each of these steady states, we find that $(0, 0)$ and $(1, 0)$ are always saddle points, and $(a, 0)$ is a stable node if $c^2 > 4f'(a)$ and a stable spiral if $c^2 < 4f'(a)$. Finding travelling wave solutions w with speed c connecting a stable and unstable equilibrium, such as from $w = 0$ to $w = a$ or from $w = a$ to $w = 1$ is possible for many choices of c, essentially by the reasoning outlined above for the monostable case. Can we also find heteroclinic orbits connecting the unstable manifold of $w = 0$ to the stable manifold of $w = 1$? It is not likely that this is going to be possible: for most speeds c the solution coming from the unstable manifold will converge to the stable node or spiral $(a, 0)$, or will overshoot (the unstable direction at $(1, 0)$ then converges to $(a, 0)$) . However, the phase planes for small and for large c (see Figure 6.13, top row), together with continuity arguments, do suggest that for some exceptional intermediate c^* a heteroclinic from 1 to 0 exists (see Figure 6.13, bottom). This can indeed be made rigorous.

There is a more general rule: if a reaction term $f(u)$ has a number of roots, they generically come alternatingly as saddle points and stable nodes or spirals. Most (in terms of c) heteroclinic orbits connect stable and saddle steady states, and only a few connect two saddle equilibria.

Let us now turn to the question of the *direction* of the wave. After all, within biological contexts it may be all important to know if a wave of some disease or introduced species retreats or advances. A simple argument gives the direction of the speed, i.e., the sign of c. Recall that travelling wave profiles $w(z)$ solve

$$w'' + cw' + f(w) = 0. \tag{6.5.12}$$

Equation (6.5.12) is invariant under $z \mapsto -z$, $c \mapsto -c$, so we need to choose the behaviour of solutions for $z \to \pm\infty$ to be able to determine the true direction of travelling wave solutions. We focus on solutions w which tend to 1 for $z \to -\infty$ and to 0 for $z \to \infty$. In particular, then also $w'(z) \to 0$ for $z \to \pm\infty$.

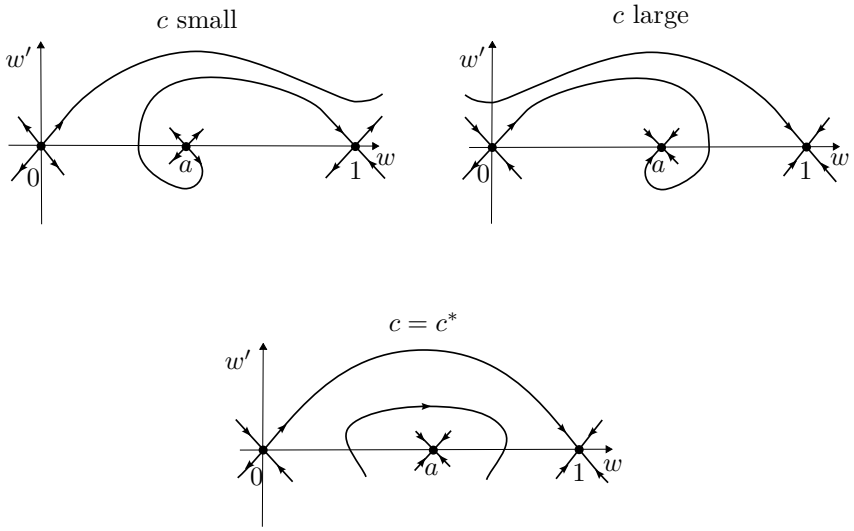


Figure 6.13.

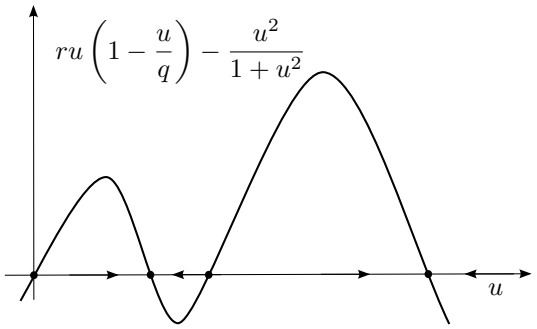


Figure 6.14.

Multiplying (6.5.12) by w' and integrating over $\mathbb{R}$, we find

$$0 = \int_{-\infty}^{\infty} [w'' + cw' + f(w)]w' dz \tag{6.5.13}$$

$$= 0 + c\int_{-\infty}^{\infty} w'^2\, dz + \int_{-\infty}^{\infty} f(w)w' dz \tag{6.5.14}$$

$$= c\int_{-\infty}^{\infty} w'^2\, dz + \int_{1}^{0} f(w) dw \tag{6.5.15}$$

where we have used partial integration and the above limits. We thus conclude that the sign of c is given by $\int_0^1 f(w)dw$, since

$$c = \frac{\int_0^1 f(w)dw}{\int_{-\infty}^{\infty} w'^2\, dz}$$

In the monostable case, $\int_0^1 u(1-u)du = 1/6$, so the wave is always describing an increase of u (for x fixed, $u(x-ct)$ increases from 0 to 1). In the bistable case, $\int_0^1 u(u-a)(1-u)du = \frac{1}{12}(1-2a)$. So for $0 < a < \frac{1}{2}$, the unique wave speed c^* is positive, for $a = \frac{1}{2}$ we find a standing wave ($c^* = 0$), while for $\frac{1}{2} < a < 1$, the wave speed is negative.

Let us consider a well-known example from scalar reaction-diffusion equations, the spread of the spruce budworm, which is a pest in North American forests. The equation, in nondimensional form, reads

$$u_t = u_{xx} + ru\left(1 - \frac{u}{q}\right) - \frac{u^2}{1+u^2}$$

The reaction term is sketched in Figure 6.14. Depending on parameter values, there are up to four equilibria, and if we ignore space dependence two of these are stable, and two unstable. As before, including space again and looking for travelling wave profiles, the time-stable steady states become saddle points, and the time-unstable steady states become stable nodes or spirals. There exist travelling wave profiles connecting the two saddles, which have a certain unique speed. By adjusting the (scaled) carrying capacity q (for instance, by limiting the amount of food available to the budworms), the direction of the wave may be controlled, and thus the outbreak of these pests may be contained. See (Murray, 2002) for a detailed discussion of this much-studied problem.

6.6 Pattern formation: The Turing instability

A key aim of developmental biology is to understand morphogenesis: how can, starting from a uniform state, spatial structure, i.e., pattern, develop? Localised differentiation of cells is certainly an essential component. But how do cells know which differentiation pathway to follow? If this hinges on positional information, then how do these cells know "where" they are? Genetic information needs physico-chemical mechanisms to be expressed, to be translated into form.

Earlier we observed that for a *scalar quantity* that diffuses and reacts, spatial structure disappears (rather than originates), unless we force it upon the system by the boundary conditions or the shape of the domain (recall the halter from Figure 6.2). What if there are *several* quantities that interact and diffuse?

Here we focus on the system of reaction-diffusion equations 6.4.1 for $k = 2$ (two components) and $m = 1$ (one-dimensional spatial domain) and establish conditions such that a uniform steady state, that is stable as a steady state of the purely kinetic system, turns unstable if we allow both components to diffuse, but with rather different diffusion constants. So differences in the time scale of spatial transport of the various components can interfere with the interaction and localised instability can manifest itself as spontaneous pattern formation (to show this in mathematical detail one needs to go beyond linearised instability and apply bifurcation methods to construct non-uniform steady states).

It is most efficient to first consider

$$\frac{\partial u_1}{\partial t} = d_1 \frac{\partial^2 u_1}{\partial x^2} + f_1(u_1, u_2) \tag{6.6.1}$$

$$\frac{\partial u_2}{\partial t} = d_2 \frac{\partial^2 u_2}{\partial x^2} + f_2(u_1, u_2) \tag{6.6.2}$$

for $x \in \mathbb{R}$ and only later consider the effect of no-flux boundary conditions on a bounded domain.

Let

$$\bar{u} = \begin{pmatrix} \bar{u}_1 \\ \bar{u}_2 \end{pmatrix}$$

be such that $f(\bar{u}) = 0$, and assume that all eigenvalues of the Jacobian matrix

$$M = Df(\bar{u}) = \begin{pmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u_2} \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2} \end{pmatrix}_{u=\bar{u}}$$

with entries m_{ij} have negative real part.

Exercise 6.6.1. Repeat the steps leading to equation (6.2.13) and show that, written out in detail, this equation reads

$$\lambda^2 + \lambda(-m_{22} - m_{11} + d_2\mu^2 + d_1\mu^2) + \{(m_{11} - d_1\mu^2)(m_{22} - d_2\mu^2) - m_{12}m_{21}\} = 0 \tag{6.6.3}$$

Exercise 6.6.2. By assumption the inequalities

$$\begin{aligned} m_{11} + m_{22} &< 0, \\ m_{11}m_{22} - m_{12}m_{21} &> 0 \end{aligned} \tag{6.6.4}$$

hold (why?). Verify that if we write (6.6.3) as

$$\lambda^2 + \theta_1\lambda + \theta_2 = 0 \tag{6.6.5}$$

then $\theta_1 > 0$. Explain why we may conclude from this that destabilization is never by way of Hopf bifurcation.

Exercise 6.6.3. A transcritical bifurcation occurs when $\lambda = 0$ is a root of (6.6.5) (or, more precisely, if a real root of (6.6.5) changes sign when parameters are varied). Evidently, this requires

$$0 = \theta_2 := d_1d_2(\mu^2)^2 - (d_1m_{22} + d_2m_{11})\mu^2 + m_{11}m_{22} - m_{12}m_{21} \tag{6.6.6}$$

Check that, as a function of μ^2, θ_2 describes a parabola with a minimum at

$$\mu^2 = \frac{1}{2}\left(\frac{m_{11}}{d_1} + \frac{m_{22}}{d_2}\right) \tag{6.6.7}$$

Compute that the minimum value equals

$$\theta_2^{\min} = m_{11}m_{22} - m_{12}m_{21} - \frac{1}{4}\frac{(d_1m_{22} + d_2m_{11})^2}{d_1d_2} \tag{6.6.8}$$

Show that $\theta_2^{\min} < 0$ iff

$$m_{11}d_2 + m_{22}d_1 > 2\sqrt{d_1d_2(m_{11}m_{22} - m_{12}m_{21})} > 0 \tag{6.6.9}$$

Check that the right hand side of (6.6.7) is positive if (6.6.9) holds (why is this important?). Show that under our assumptions, (6.6.9) cannot hold if $d_1 = d_2$. Show that (6.6.9) requires (under our assumptions) that m_{11} and m_{22} have opposite signs. Show that then also m_{12} and m_{21} should have opposite signs.

Exercise 6.6.4. If the sign structure of M is

$$\begin{pmatrix} + & - \\ + & - \end{pmatrix}$$

we call species/substance 1 an *activator* and species/substance 2 an *inhibitor*. Explain the rationale of this terminology.

Exercise 6.6.5. Without loss of generality we may assume that the sign structure is as assumed in the preceding exercise. Substantiate this claim.

Hint: the other possibilities are

$$\begin{pmatrix} + & + \\ - & - \end{pmatrix}, \quad \begin{pmatrix} - & + \\ - & + \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} - & - \\ + & + \end{pmatrix}$$

It is, of course, rather arbitrary how we number the species. In addition one might do the bookkeeping in terms of $-u_2$ rather than u_2 (but note carefully that often the interpretation requires quantities to be *positive*; yet *deviations* from a strictly positive steady state value may assume both signs. The message is that "without loss of generality" is a subtle notion when the interpretation leads to constraints on mathematical transformations).

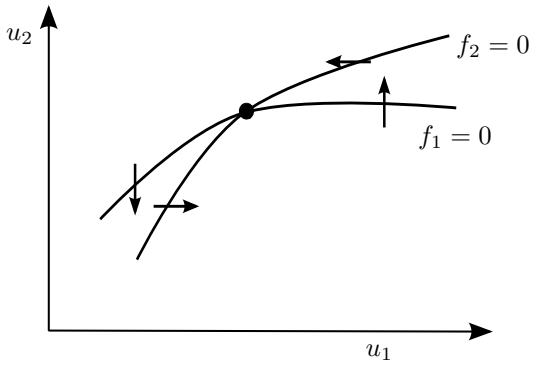


FIGURE 6.15. Local phase portrait of the kinetic system in Exercise 6.6.8.

.

EXERCISE 6.6.6. One often encounters statements like "Diffusive instability requires long range inhibition and short range activation". With the sign structure of M as in Exercise 6.6.4 we can rewrite (6.6.9), with the middle part omitted, as

$$\tau_1 d_1 < \tau_2 d_2 \tag{6.6.10}$$

with $\tau_1 := m_{11}^{-1}$ and $\tau_2 = |m_{22}|^{-1}$. Explain the relation between this inequality and the statement between the quotation marks.

EXERCISE 6.6.7. Assume that $M - \mu^2 d$ has eigenvalue zero and that M has activator-inhibitor sign structure, cf. Exercise 6.6.4. Let $\bar{v}$ be the eigenvector corresponding to this eigenvalue zero. Show that $\operatorname{sign} \bar{v}_1 = \operatorname{sign} \bar{v}_2$. Explain in a hand-waiving manner that accordingly the two components of a bifurcating non-uniform steady state are in-phase, meaning that one increases as a function of x if and only the other does too. What changes if the sign pattern of M is $\begin{pmatrix} + & + \\ - & - \end{pmatrix}$?

EXERCISE 6.6.8. Assume that M has activator-inhibitor sign structure and that M has a positive determinant. Show that the local phase portrait of the kinetic system is as shown in Figure 6.15. Hint: solve $f_i = 0$ for u_2 as a function u_1 by way of the Implicit Function Theorem. How does the phase portrait look if $\begin{pmatrix} + & + \\ - & - \end{pmatrix}$?

EXERCISE 6.6.9. Show that by *scaling* of the spatial variable we may arrive at a *ratio* of diffusion coefficients and that this, for a particular choice of scaling, amounts to replacing μ^2 by μ^2/D_1 so that (6.6.6) transforms into

$$0 = \theta_2 = \frac{d_2}{d_1}(\mu^2)^2 - \left(\frac{d_2}{d_1} m_{11} + m_{22}\right) \mu^2 + m_{11} m_{22} - m_{12} m_{21} \tag{6.6.11}$$

or, if we solve for d_2/d_1 as a function of μ^2,

$$\frac{d_2}{d_1} = \frac{m_{22}\mu^2 - m_{11} m_{22} + m_{12} m_{21}}{\mu^2(\mu^2 - m_{11})} \tag{6.6.12}$$

Show that under the conditions (6.6.4) and (6.6.9) the graph of the right hand side, as a function of μ^2, is as depicted in Figure 6.16.

Thus we have determined the stability boundary in the two parameter plane formed by the ratio d_2/d_1 of the diffusion constants and the mode parameter μ^2 (which is a continuous quantity when the spatial domain is the line $-\infty < x < \infty$). This graph is an essential ingredient for the stability analysis for finite intervals with no-flux boundary conditions, as we shall see in the next exercise.

EXERCISE 6.6.10. Now restrict the spatial domain to

$$0 \le x \le L$$

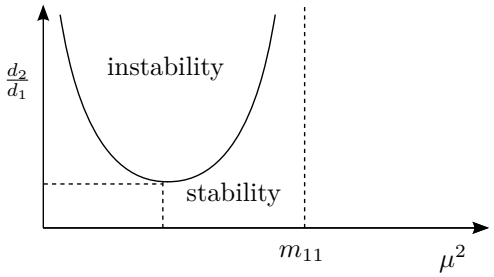


Figure 6.16. The graph of the right hand side of (6.6.12) as a function of μ^2 in Exercise 6.6.9.

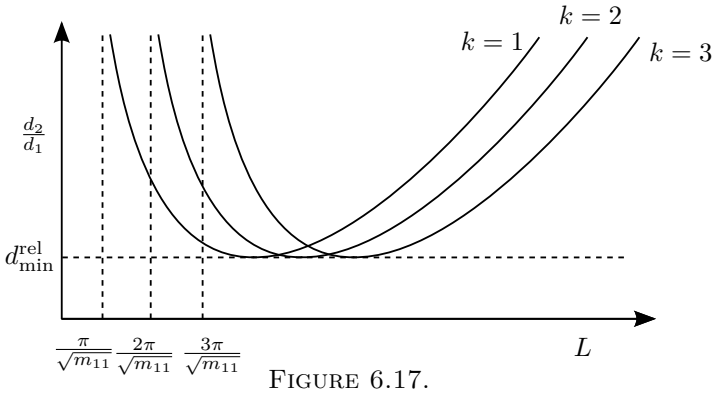


Figure 6.17.

and impose the no-flux boundary conditions

$$u_x(0) = 0 = u_x(L)$$

Derive that necessarily

$$\mu \in \left\{ \frac{k\pi}{L} \ : \ k = 1, 2, \ldots \right\}$$

By considering L as a parameter we regain continuity (i.e., we eliminate to a certain extent the imposed discreteness), but we obtain denumerably many curves, one for each mode (see Figure 6.17). Describe the instability domain in $(\frac{d_2}{d_1}, L)$-space and its boundary. Describe what happens when we consider L as a free parameter for d_2/d_1 fixed at a value just slightly above $d_{\min}^{\text{rel}}$ (you may find it useful to think in terms of "resonance" between the "natural" wave length associated with the instability on the one hand, and the size of the domain on the other).

Exercise 6.6.11. If we cross the curve corresponding to a particular k, the nonlinear system undergoes a pitchfork bifurcation. How can we be so sure about "pitchfork" without doing any calculations? Also, formulate that there is a certain arbitrariness in the pattern that arises in a real experiment.

Exercise 6.6.12. (builds on the Appendix on Bifurcation Methods) Derive a formula for the direction of the pitchfork bifurcation, i.e., determine whether it is is subcritical or supercritical. Formulate the Principle of Exchange of Stability for this particular case.

Exercise 6.6.13. Reflect on the patterns that one expects to see on a two-dimensional rectangular spatial domain, depending on the ratio of the two lengths (so when we enlarge the domain but keep this ratio fixed). Next, go to the zoo and look for spotted bodies and striped tails!

We now consider a concrete example which is rather debatable from a modelling point of view (in particular because of the H in the denominator) but which has the great advantage that the calculations are not too cumbersome. It is often called the Gierer-Meinhardt model, see (Meinhardt, 1982).

The system of reaction diffusion equations

$$\frac{\partial A}{\partial t} = 1 + R\frac{A^2}{H} - A + \frac{\partial^2 A}{\partial x^2}, \tag{6.6.13}$$

$$\frac{\partial H}{\partial t} = Q(A^2 - H) + P\frac{\partial^2 H}{\partial x^2} \tag{6.6.14}$$

provided with no-flux boundary conditions

$$\frac{\partial A}{\partial x}(t,0) = 0 = \frac{\partial A}{\partial x}(t,L), \tag{6.6.15}$$

$$\frac{\partial H}{\partial x}(t,0) = 0 = \frac{\partial H}{\partial x}(t,L) \tag{6.6.16}$$

describes the interaction between an autocatalytic activator A and an inhibitor H in a one-dimensional spatial domain, the x-interval $[0, L]$. The system (6.6.14) is already in a scaled form and only three parameters, R, Q, and P (all assumed positive) remain. In particular, the spatial variable x has been scaled to make the diffusion constant of A equal to 1. In the term RA^2/H, the denominator H is an approximation to (constant + H) and accordingly, predictions based on (6.6.14) should not be trusted when they involve small values of H. The greatest advantage of this approximation is that it makes the calculations below far simpler, which is why it is made.

EXERCISE 6.6.14.

(i) Find the uniform (in x) steady (in t) state.

(ii) Compute the Jacobian matrix of the reaction part.

(iii) Show that the constant steady state is *stable* with respect to homogeneous (i.e., x-independent) perturbations, provided the parameter inequality

$$\frac{R-1}{R+1} < Q \tag{6.6.17}$$

holds.

EXERCISE 6.6.15.

(i) Show that the k-th mode, characterised by dependence on x of the form $\cos\left(\frac{k\pi x}{L}\right)$, is stable when

$$P\left(\frac{k\pi}{L}\right)^4 + \left(Q - P\frac{R-1}{R+1}\right)\left(\frac{k\pi}{L}\right)^2 + Q > 0 \tag{6.6.18}$$

but unstable when the reverse inequality holds. In your answer, do not start with equation (6.6.6), but go through the arguments starting with the dispersion relation (6.2.13) which leads to (6.6.3) in Exercise 6.6.1.

(ii) Deduce from (6.6.18) that a necessary condition for instability is

$$Q < P\frac{R-1}{R+1} \tag{6.6.19}$$

(iii) Deduce, by comparing (6.6.17) and (6.6.19), that a necessary condition for pattern forming instability is

$$P > 1$$

and interpret this condition.

Exercise 6.6.16. Show that the quadratic polynomial in α,

$$P\alpha^2 + \left(Q - P\frac{R-1}{R+1}\right)\alpha + Q$$

has minimum value

$$-\frac{1}{4}P\left(\frac{R-1}{R+1} - \frac{Q}{P}\right)^2 + Q$$

and check that it is attained for a *positive* value of α when (6.6.19) holds.

Exercise 6.6.17. (Re)formulate the results in biological terms and draw conclusions. In particular, think about how diffusion facilitates pattern formation, what kind of activator-inhibitor system this is, and which modes are potentially unstable for given parameter values.

6.7 Discrete-space version of spectral decoupling as in Turing

Let X be a $k \times n$ matrix with components x_{ij}, to be interpreted as the density of the i-th species in the j-th patch. So for $j = 1, \dots, n$, the column k-vector $x_{ij} = (x_{1j}, \dots, x_{kj})^T$ describes the desities of the various species in the j-th patch, while for $i = 1, \dots, k$, the row n-vector $(x_{i1}, \dots, x_{in})$ describes the densities of species i in the various patches.

Let now $f : \mathbb{R}^k \to \mathbb{R}^k$ describe given (nonlinear) local interactions; $C : \mathbb{R}^n \to \mathbb{R}^n$ a linear map describing redistribution, so that c_{ij} describes movement from i to j; and $M : \mathbb{R}^k \to \mathbb{R}^k$ a linear and diagonal map describing the tendency to migrate with components m_i, $i = 1, \dots, k$.

We now lift all three of these maps to $\mathbb{R}^{k\times n}$ by the definitions

$$\begin{aligned} f(X) &= \Big(f(x_{\cdot 1}, \dots, x_{\cdot n})\Big) && \text{local interaction} \\ XC &= \Big(x_{1\cdot}C, \dots, x_{k\cdot}C)\Big)^T && \text{species independent redistribution} \\ MX &= (Mx_{\cdot 1}, \dots, Mx_{\cdot n}) && \text{patch independent migration tendency} \end{aligned}$$

The problems to be investigated are

$$\dot{X} = f(X) + MXC \quad \text{continuous time,} \tag{6.7.1}$$

and

$$X' = f(X) + Mf(X)C \quad \text{discrete time.} \tag{6.7.2}$$

Compared with the continuous-space problem

$$\frac{\partial u}{\partial t} = D\Delta u + f(u),$$

matrix M has the role of D, and C the role of the Laplacian Δ.

If $a \in \mathbb{R}^k$ and $b \in \mathbb{R}^n$ we shall write $X = ab$ as a shorthand for $X_{ij} = a_i b_j$. Let $\mathbb{1}_n = (1, \dots, 1)$, and *assume* that $\mathbb{1}_n C = 0$. Let y be a $\mathbb{R}^k$-valued function of time such that

$$\dot{y} = f(y) \quad \text{or} \quad y' = f(y)$$

so that y is a solution of the 1-patch problem. Then $X = y\mathbb{1}_n$ is a solution of the n-patch problem (note that $f(X) = f(y)\mathbb{1}_n$ when verifying the discrete time case). Conversely, if $X = y\mathbb{1}_n$ is a solution of the n-patch problem, then necessarily y is a solution of the 1-patch problem. We call these the *flat* solutions.

Let $X = y\mathbb{1}_n$ be a flat solution. The Jacobi $k\times k$-matrix $Df(y)$ is lifted to $\mathbb{R}^{k\times n}$ exactly as M. Then the linearised problems of (6.7.1) and (6.7.2) are respectively

$$\dot{\xi} = Df(y)\xi + M\xi C,$$
$$\xi' = Df(y)\xi + MDf(y)\xi C.$$

Next *assume* that C has n linearly independent eigenvectors ψ^i, $i = 1,\dots,n$, corresponding to eigenvalues λ_i, i.e.,

$$\psi^i C = \lambda_i \psi^i, \quad i = 1,\dots,n.$$

The result, that any n-vector can be written as a linear combination of the ψ^i, can be lifted to $\mathbb{R}^{k\times n}$:

$$\xi = \sum_{i=1}^{n} a^i \psi^i \quad \text{for some } a_i \in \mathbb{R}^k, i = 1,\dots,n.$$

With this respresentation for ξ we have

$$Df(y)\xi = \sum_{i=1}^{n} Df(y)a^i\psi^i$$

(where the right hand side $Df(y)$ is again the $k \times k$ matrix), and

$$M\xi C = \sum_{i=1}^{n} Ma^i\psi^i C = \sum_{i=1}^{n} \lambda_i Ma^i\psi^i.$$

The independence of the ψ^i then implies that the $k \times n$-dimensional problems decouple into n problems of dimension k,

$$\dot{a}^i = Df(y)a^i + \lambda Ma^i,$$

or

$$(a^i)' = Df(y)a^i + \lambda MDf(y)a^i.$$

And each of these can be analysed by spectral methods, i.e., the decay or growth of the solutions is completely determined by the eigenvalues of

$$Df(y) + \lambda_i M \quad \text{position relative to imaginary axis,}$$

or

$$Df(y) + \lambda_i MDf(y) \quad \text{position relative to unit circle.}$$

The analysis above is *completely the same* as the standard Turing instability analysis of

$$\frac{\partial u}{\partial t} = D\Delta u + f(u).$$

To conclude, we make some remarks about C and, in particular, the interpretation of the assumption $\mathbb{1}_n C = 0$.

The matrix C_n is positive-off-diagonal (i.e., only diagonal elements can be negative). If migrants cannot die in the process, we should have the conservation relation

$$\sum_{j=1}^{n} c_{ij} = 0.$$

If redistribution is only governed by relative distances, we should have the symmetry

$$c_{ji} = c_{ij},$$

and consequently

$$\sum_{i=1}^{n} c_{ij} = 0.$$

The latter relation expresses, when written in the form

$$\sum_{i=1,\ i\neq j}^{n} c_{ij} = -c_{jj},$$

that immigration into a patch is matched by emigration from a patch at the per capita level. This could be called a "no accumulation" condition. As a consequence of the conservation relation, the column n-vector $\mathbb{1}_n^T$ is a right eigenvector of C_n corresponding to eigenvalue zero. Under the no-accumulation condition $\mathbb{1}_n$ is a left eigenvector corresponding to eigenvalue zero.

It helps to think of λ in

$$Df(y) + \lambda M$$

as a continuous variable (taking only negative values).

Chapter 7

Chemotaxis

7.1 Introduction

Organisms often direct their movement by external cues, a process called taxis. Depending on the cue in question, we may term such directed movement thermotaxis (warmth), phototaxis (light), chemotaxis (chemical substances), and so on, and this may be both an attracting or repelling movement. Several kinds of common bacteria, such as *E. coli*, *Salmonella*, or the slime mould *Dictyostelium* (*Dicty* in short) have been shown to form intricate patterns when grown in semi-solid or liquid media in the laboratory. An example pattern showing spirals of the chemoattractant cyclic AMP (cAMP) in *Dictyostelium* is presented in Figure 7.1.

The reason for aggregation or repulsion may be many. Bacteria use scents to find food; immune cells chemotactically find enemies such as bacteria; insects use pheromones to find each other, either for reproduction as in moths, or to hunt collectively, as in army ants; reproduction is also the main drive behind Dicty aggregation.

Because of the simplicity of this organism, and the richness of its collective behaviour, *Dicty* has been used as a model organism for many years, especially to understand how signal transduction of cAMP induces aggregation and subsequent development (Othmer and Schaap, 1998).

Aggregation in *Dicty* is basically done through a feedback loop, in which the individual cells release cAMP into the environment while reacting to the cAMP levels they perceive. When general food levels become too low to sustain the slime mold cells individually, the cells start to produce cAMP in order to aggregate. They first form centra, which then concentrate to become 'slugs'. These slugs move about for a while until a suitable place is found. There, the multicellular slug undergoes cell differentiation to form a (pre)stalk on which a (pre)spore is formed. The stalk contains nonreproducing *Dicty* cells, i.e., there is cooperation among the cells so that few may reproduce. This is indeed an intricate evolutionary question, and to find the answer you should read the Selfish Gene and The Extended Phenotype by Richard Dawkins! The fruiting body on top finally contains spores which are dispersed by the wind. Some excellent videos of this remarkable progression may be found on the internet, including those made by the person who started much of this field, John Bonner, whilst still an undergraduate.

Mathematically, the description of the evolution of concentrations of organisms in some domain Ω begins with a general conservation law. This states that the total amount of organisms in Ω at time $t + \delta t$ must be equal to the total amount at time t plus the net concentration of particles which either flows out of Ω or is created inside Ω within the timespan δt. If we denote this net flow out of Ω by a

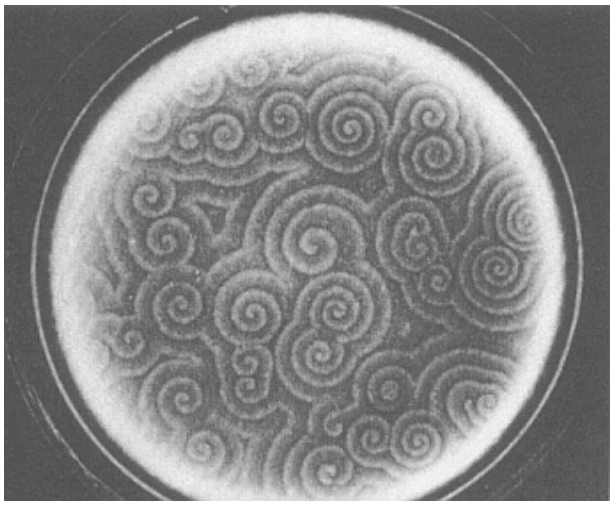


FIGURE 7.1. Spiraling waves of cyclic AMP, the chemoattractant used by colonies of *Dictyostelium* bacteria

flux $J(t,x)$ then we can write

$$\int_\Omega u(t+\delta t,x)dx = \int_\Omega u(t,x)dx - \delta t \int_{\partial\Omega} J(t,x)dS \tag{7.1.1}$$

Since this argument holds for any domain Ω, and $\int_{\partial\Omega} J(t,x)dS = \int_\Omega \nabla \cdot J dx$, we have the general evolution equation

$$u_t + \nabla \cdot J = 0. \tag{7.1.2}$$

This flux J may be due to different kinds of motion, such as diffusion or taxis. We could also easily take creation of particles in Ω into account. If we let $f(t,x)$ denote the creation of organisms at time t and position x, then the above equation becomes simply

$$u_t + \nabla \cdot J - f = 0$$

Phenomenologically, we may pose the following equation for organisms whose movements are described as stochastic random walks with a bias towards chemoattractant concentrations:

$$u_t + \nabla \cdot (d\nabla u - u\chi(s)\nabla s) \tag{7.1.3}$$

known as a Keller-Segel equation (Keller and Segel, 1971). Here the chemotactic flux J_c due to attraction by the chemical s is

$$J_c = u\chi(s)\nabla s$$

where $\chi(s)$ is termed the *chemotactic sensitivity*. The Keller-Segel equation (7.1.3) is often coupled to an equation for the chemoattractant s, which usually diffuses and is produced either by the cells itself or by an external source, e.g.

$$s_t = \Delta s + u - s.$$

Note that we have conveniently scaled the diffusion constant for s to 1 by rescaling space, the production rate by u to 1 by rescaling u and the degredation rate of s to 1 by rescaling time. We supply no-flux BCs to the equations for u and S to complete the Keller-Segel model.

The plan for the rest of the chapter is as follows. First we will study the initiation of pattern formation for a Keller-Segel model in one spatial dimension, and briefly discuss whether solutions exist for all time or may blow up in finite

time. Lastly, we will try to derive macroscopic Keller-Segel-like equations from microscopic behaviour of the individual cells, such as run-and-tumble movement in bacteria, signal transduction, and so on.

7.2 Initiation of pattern formation

Let us focus on the following set of equations on a domain $[0, L]$, with no-flux BCs,

$$\frac{\partial u}{\partial t} = d\frac{\partial^2 u}{\partial x^2} - \chi\frac{\partial u}{\partial x}\frac{\partial s}{\partial x} - \chi u\frac{\partial^2 s}{\partial x^2}, \tag{7.2.1}$$

$$\frac{\partial s}{\partial t} = \frac{\partial^2 s}{\partial x^2} + u - s. \tag{7.2.2}$$

The last term in the u equation is due to cross diffusion. Note that there is conservation of mass for u, since $\frac{d}{dt}\int u\,dx = 0$.

There exists a uniform steady state $u = \bar{u}$, $s = \bar{s}$, as long as $\bar{u} = \bar{s}$, and we can thus treat $\bar{u}$ as a parameter. Introducing $u = \bar{u} + U$ and $s = \bar{s} + S$. Linearizing around $(\bar{u}, \bar{s})$ gives

$$\frac{\partial U}{\partial t} = d\frac{\partial^2 U}{\partial x^2} - \chi\bar{u}\frac{\partial^2 S}{\partial x^2}, \tag{7.2.3}$$

$$\frac{\partial S}{\partial t} = \frac{\partial^2 S}{\partial x^2} + U - S. \tag{7.2.4}$$

Making the usual separation of variables Ansatz

$$\begin{pmatrix} U \\ S \end{pmatrix}(t, x) = e^{\lambda t}\cos\mu x \begin{pmatrix} U_0 \\ S_0 \end{pmatrix}$$

gives

$$\lambda U_0 = \mu^2 d U_0 + \chi\bar{u}\mu^2 S_0,$$
$$\lambda S_0 = -\mu^2 S_0 + U_0 - S_0,$$

which in matrix form reads

$$\begin{pmatrix} \lambda + \mu^2 d & -\chi\bar{u}\mu^2 \\ -1 & \lambda + \mu^2 + 1 \end{pmatrix}\begin{pmatrix} U_0 \\ S_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The condition for a nontrivial solution is thus given by requiring that this matrix has determinant zero, i.e.,

$$(\lambda + \mu^2 d)(\lambda + \mu^2 + 1) - \chi\bar{u}\mu^2 = 0,$$

which is written more transparently as

$$\lambda^2 + (\mu^2(d+1) + 1)\lambda + \mu^2(d(\mu^2 + 1) - \chi\bar{u}) = 0.$$

Denoting the first and zeroth order coefficients of this polynomial in λ by B and C, we know that

$$\lambda_\pm = \frac{-B \pm \sqrt{B^2 - 4C}}{2}.$$

Observe that if $\Im\lambda \neq 0$ then $\Re\lambda < 0$, and that if $\Im\lambda = 0$ then $\lambda_- < 0$ and $\lambda_+ > 0$ if and only if $C < 0$.

We conclude that the uniform steady state is *unstable* iff

$$C < 0 \quad \Longleftrightarrow \quad d(\mu^2 + 1) < \chi\bar{u}. \tag{7.2.5}$$

We may offer directly the following biological interpretation of this inequality. Instability, and thus the initiation of pattern formation, is promoted by a high initial concentration of cells, a high chemotactic sensitivity χ, or a low random motility d.

Additionally, we do well to remember that several quantities were hidden inside the nondimensional equations by the rescaling of variables we had assumed. Making these explicit again revels that a high rate of cAMP production and low degredation rate of cAMP also promote pattern formation. Finally, we focus on the spatial mode $\mu = k\pi/L$ of the perturbation we have analyzed. Inequality (7.2.5) is more easily satisfied for small μ, i.e. for long waves, or equivalently, on large intervals $[0, L]$.

Overall, we may conclude that the feedback loop of involving signal production and moving towards stronger signals may lead to growing peaks provided the "equalizing" influence of diffusion is not too strong.

Finally we make some remarks on the *asymptotic behaviour* for large t. If the spatial dimension is 1, then solutions stay bounded. (Note that $\int u\,dx$ remains constant, so there is never blow up in the L^1-norm. Also $\int S \to \int u$ as $t \to \infty$.) If the spatial dimension is two, then if $\int u$ is large enough, a Dirac may form in finite time. If the spatial dimension is 3, then Dirac's may form in infinite time. Keller-Segel chemotaxis in the plane may thus account for aggregation in one spot. (There are are some ways to extend the model in order to capture the subsequent movement of Dirac's.)

The literature on Keller-Segel-like models has grown enormously since the early 1970s. The interested reader may consult (Horstmann, 2003) or (Perthame, 2007).

7.3 Derivation of chemotaxis models

One of the main obstacles to use (7.1.3) directly is that one has to specify $\chi(S)$. There is no general theory which allows us to translate the bacteria's perception of the chemoattractant and their subsequent change of behaviour (moving towards higher chemoattractant concentrations) to a macroscopic chemotactic sensitivity function.

In this section we will show how one can obtain Keller-Segel equations, or other evolution equations for chemotactic bacteria, using the dynamics at a *mesoscopic* scale as starting point. The main point is that it is often easier to describe dynamics on a level at which pattern is *not* observed, and then lift these equations to the level at which it *is* observed. In the current context, it is easy to specify how individual particles change their direction due to external cues or random motion. This gives us evolution equations for a density $u(t, x, v)$, say, which thus depends on velocities v. The mathematical goal is then to derive an evolution equation for a function n, say, which does not depend on v anymore, but only on time and space (which is the quantity one observes when one describes the bacterial patterns such as in Figure 7.1).

The simplest example in which we can derive a parabolic equation from a mesoscopic one is in one space dimension. Let particles move according to a so-called velocity-jump process. In this process, the particles move at a certain speed (here assumed to be a constant s), and reorient at random instants in time according to a Poisson process with intensity λ. In one space dimension, the particles can only move in two directions. Let $u^{\pm}(t, x)$ be the density of particles at (t, x) and moving to the right (+) or left (-) respectively. Then $u^{\pm}$ satisfy the hyperbolic equations

$$\frac{\partial u^+}{\partial t} + s\frac{\partial u^+}{\partial x} = -\lambda u^+ + \lambda u^- \tag{7.3.1}$$

$$\frac{\partial u^-}{\partial t} - s\frac{\partial u^-}{\partial x} = -\lambda u^- + \lambda u^+ \tag{7.3.2}$$

The density of particles at (t,x), $u(t,x)$, is the sum of $u^+(t,x)$ and $u^-(t,x)$, and the particle flux j equals $s(u^+ - u^-)$. These satisfy

$$\frac{\partial u}{\partial t} + \frac{\partial j}{\partial x} = 0 \tag{7.3.3}$$

$$\frac{\partial j}{\partial t} + 2\lambda j = -s^2 \frac{\partial u}{\partial x} \tag{7.3.4}$$

Differentiating the first of these equations with respect to t and the second to x and combining both equations leads to the telegraph equation

$$\frac{\partial^2 u}{\partial t^2} + 2\lambda \frac{\partial u}{\partial t} = s^2 \frac{\partial^2 u}{\partial x^2}$$

The diffusion equation follows formally in the limit $\lambda \to \infty$, $s \to \infty$, while keeping

$$\frac{s^2}{\lambda} =: d \tag{7.3.5}$$

constant.

To understand pattern formation in bacteria or other chemotactic organisms such as ants, we need to derive continuum models in higher dimensions. Now particles can travel in an infinity of directions, and we denote by $p(t,x,v)$ the density of particles at time t and position $x \in \mathbb{R}^n$ moving in the direction $v \in V := sS^{n-1}$ (still with constant speed s). This density now satisfies the transport equation

$$\frac{\partial u}{\partial t} + \nabla \cdot (vu) = -\lambda u + \lambda \int_V T(v,v')u(t,x,v')dv'$$

which resembles the Boltzmann equation, but it is linear in u rather than quadratic. It is nothing but a conservation equation like (7.1.2), but now over the domain $\mathbb{R}^n \times V$ rather than a domain $\Omega \subset \mathbb{R}^n$. Within the current context, $T(v,v')$ is a turning kernel, and signifies the probability of changing direction from v' to v if a switch is made, which happens with probability λ. It has a number of obvious properties. Most importantly, $T \geq 0$ and

$$\int_V T(v,v')dv' = 1$$

The main goal here is to find an evolution equation for $n(t,x) := \int u(t,x,v)dv$ such as a diffusion equation or a Keller-Segel equation. The method makes crucial use of a small parameter which has to be identified in the mesoscopic model. Within the current context, this parameter, ε, is usually the ratio sL/T, where L is a typical length scale of the pattern and T a measure of the typical time scale. The method now consists of three steps

(1) introduce a scaling which reflects the type of model you want to find using a small parameter
(2) write u as an asymptotic expansion in this small parameter
(3) find out what equations the different parts of this expansion have to satisfy in order for the problem to be solvable. In many cases, these solvability conditions give rise to the evolution equations you are after

In our case, we will study a parabolic scaling, i.e., $\xi = \varepsilon x$ and $\tau = \varepsilon^2 t$. Hyperbolic scalings are also often useful, and give rise to chemotaxis models with quite different properties. The rescaled transport equation now becomes

$$\varepsilon^2 \frac{\partial u}{\partial \tau} + \varepsilon \nabla_\xi \cdot (vu) = -\lambda u + \lambda \int_V T(v,v')u(\tau,\xi,v')dv' \tag{7.3.6}$$

The second step is two write u as an asymptotic expansion

$$u(\tau,\xi,v) = \sum_{i=0}^{k} \varepsilon^i u_i(\tau,\xi,v) + \mathcal{O}(\varepsilon^{k+1}) \tag{7.3.7}$$

Let us denote

$$\mathcal{L}\phi(v) = -\lambda\phi(v) + \lambda \int_V T(v,v')\phi(\tau,\xi,v')dv'$$

for functions $\phi \in L^2(V)$. Note that the natural choice of function space here would be $L^1(V)$, but choosing L^2 makes the exposition more straightforward, since the dual of L^2 is again L^2.

Plugging this expansion (7.3.7) into (7.3.6) and grouping the terms in the resulting equation by orders of ε, we find

$$\mathcal{O}(\varepsilon^0): \qquad \mathcal{L}u_0 = 0 \tag{7.3.8}$$

$$\mathcal{O}(\varepsilon^1): \qquad \mathcal{L}u_1 = v\cdot\nabla u_0 \tag{7.3.9}$$

$$\mathcal{O}(\varepsilon^2): \qquad \mathcal{L}u_2 = \frac{\partial u_0}{\partial t} + v\cdot\nabla u_1 \tag{7.3.10}$$

$$\vdots$$

$$\mathcal{O}(\varepsilon^i): \qquad \mathcal{L}u_i = \frac{\partial u_{i-2}}{\partial t} + v\cdot\nabla(u_{i-1}) \quad 3 \le i \le k \tag{7.3.11}$$

The properties of T imply that 0 is a simple eigenvalue of $\mathcal{L}u = -\lambda u + \lambda\int_V T(.,v')u(v')dv'$ with eigenfunction $u \equiv 1$. We can hence conclude that, since $\mathcal{L}u_0 = 0$, u_0 does not depend on v! This means that u_0 only depends on τ and ξ and is the dependent variable for which we are trying to derive an evolution equation.

Were $\mathcal{L}$ invertible, we could simply proceed by first setting

$$u_1 = \mathcal{L}^{-1}(v\cdot\nabla u_0)$$

and then

$$u_2 = \mathcal{L}^{-1}\frac{\partial u_0}{\partial t} + v\cdot\nabla(\mathcal{L}^{-1}(v\cdot\nabla u_0))$$

and so on. But $\mathcal{L}$ is singular, and is only invertible on the orthogonal complement in $L^2(V)$ of the eigenspace at eigenvalue 0, $\langle 1\rangle^\perp$. This is nothing but those functions $\phi \in L^2$ such that

$$\int_V \phi(v)dv = 0$$

Hence, to be able to express u_1 in u_0, we have to make sure that the right hand side of (7.3.9) satisfies this orthogonality condition, which reads

$$\int_V (v\cdot\nabla u_0)dv = 0. \tag{7.3.12}$$

Then $u_1 = \mathcal{F}(v\cdot\nabla u_0)$, where $\mathcal{F}$ is the pseudoinverse of $\mathcal{L}$ (i.e., the inverse of $\mathcal{L}$ where it is well-defined).

To solve (7.3.10) we similarly have to require that

$$\int_V \frac{\partial u_0}{\partial t} + v\cdot\nabla u_1 dv = 0$$

Using $u_1 = \mathcal{F}(v\cdot\nabla u_0)$ this becomes

$$\int_V \frac{\partial u_0}{\partial t} + v\cdot\nabla(\mathcal{F}(v\cdot\nabla u_0))dv = 0$$

Since u_0 is independent of v, the integrand vanishes, and we find the desired evolution equation for u_0,

$$\frac{\partial u_0}{\partial t} + v \cdot \nabla(\mathcal{F}(v \cdot \nabla u_0)) = 0 \tag{7.3.13}$$

which can be written in the more familiar form

$$\frac{\partial u_0}{\partial t} - \nabla \cdot (d \nabla u_0) = 0 \tag{7.3.14}$$

where

$$d = -\frac{1}{|S^{n-1}|} \int_V v \mathcal{F} v dv$$

Fortunately, in many cases this pseudoinverse $\mathcal{F}$ can be computed explicitly. In the simplest case, when $T(v, v') = 1/|S^{n-1}|$ and $V = sS^{n-1}$, we find

$$d = \frac{s^2}{\lambda n}$$

This is a straighforward generalisation of the diffusion constant (7.3.5) we found in the one-dimensional telegraph equation.

This technique of scaling, substituting an asymptotic expansion and finding an evolution equation as a solvability condition, is a very general one and occurs in many applied mathematics problems. Let us here extend this idea to incorporate sensing of a chemoattractant, with a Keller-Segel model as the final result.

The main ingredient we need to add to include chemotaxis is to change the turning kernel T. We suppose that this is now a function of an external signal $S(t, x)$. Intuitively, if a bacterium senses the signal it should swim in the direction of highest concentration. The probability of choosing a new direction should thus depend on the concentration around the position x, in other words on ∇S. We will make this assumption at the very end.

We continue the above technique at the rescaled transport equation, which now reads

$$\varepsilon^2 \frac{\partial u}{\partial \tau} + \varepsilon \nabla_\xi \cdot (vu) = -\lambda u + \lambda \int_V T(v, v', S) u(\tau, \xi, v') dv'$$

Next to the asymptotic expansion (7.3.7) of u, we also introduce an expansion for T. Let us assume that the influence of S only occurs in the order ε term:

$$T(v, v', S) = T_0(v, v') + \varepsilon T_1(v, v', S) \tag{7.3.15}$$

Substituting this gives

$$\varepsilon^2 \frac{\partial u}{\partial \tau} + \varepsilon \nabla_\xi \cdot (vu) = \mathcal{L}_0 u + \varepsilon \lambda \int_V T_1(v, v', S) u(\tau, \xi, v') dv'$$

where

$$\mathcal{L}_0 u = -\lambda u + \lambda \int_V T_0(., v') u(v') dv'$$

We continue by substituting the expansion for u. The u_i again satisfy a coupled set of equations analogous to (7.3.8)–(7.3.11), which can only be solved under certain solvability conditions. The lowest order contribution u_0 is still independent of v, and the solvability condition for u_0 at order ε^2 is now

$$\int_V \left(\frac{\partial u_0}{\partial t} + (v \cdot \nabla) \mathcal{F}_0 (v \cdot \nabla u_0) - \lambda (v \cdot \nabla) \mathcal{F}_0 \left(\int_V T_1(v, v') dv' u_0 \right) \right) dv$$
$$- \lambda_0 \int_V \int_V T_1(v, v', S) u_1(v') dv' dv = 0 \tag{7.3.16}$$

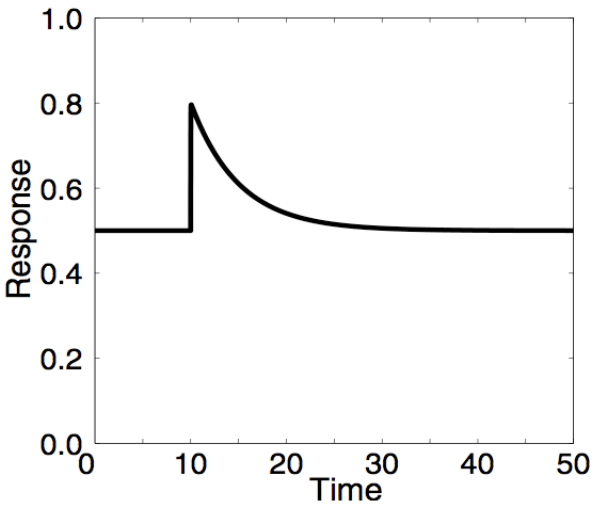


FIGURE 7.2. Typical dynamics of the internal excitation variable y_1 as the bacterium passes a sudden increase of chemoattractant. Note that the bacterium is first more excitable, but that this excitability slowly decreases back to a rest state. Adapted from Erban and Othmer (2004).

Here $\mathcal{F}_0$ is the pseudoinverse of $\mathcal{L}_0$. If we define

$$v_c = -\frac{\lambda}{|S^{n-1}|}\int_V\int_V v\mathcal{F}_0 T_1(v,v',S)dv'dv$$

as the chemotactic velocity, then u_0 satisfies an equation which starts to resemble a Keller-Segel equation

$$\frac{\partial p_0}{\partial \tau} = \nabla\cdot(d\nabla u_0 - v_c u_0)$$

If we moreover make the same simplifying assumptions as before, $T_0 = 1/|S^{n-1}|$, then

$$d = \frac{s^2}{\lambda n}, \quad v_c = \frac{1}{|S^{n-1}|}\int_V\int_V vT_1(v,v',S)dvdv'$$

Finally, to obtain the classical Keller-Segel model, we make the additional assumption that T_1 depends linearly on ∇S to which we hinted at the beginning of this derivation. Then v_c is of the form $\chi(S)\nabla S$, with

$$\chi(S) = \frac{\lambda k(S)}{|S^{n-1}|}d$$

Note, however, that we are effectively not very much further. Rather than having to choose an arbitrary function $\chi(S)$, we now have to choose an equally arbitrary $k(S)$.

There are many variations and extensions on this theme. We may also introduce dependence of the turning rate λ on the chemoattractant, and again find that the this dependence has to be of order ε to give us a Keller-Segel equation. Starting either with a turning kernel T or a turning rate λ in which this S-dependence is already present in the $\mathcal{O}(1)$ term (T_0 or λ_0) does not result in a Keller-Segel equation, but reduces the evolution equation to simple diffusion.

One of the most exciting extensions in this field has been the additional modelling of the signal transduction pathways by which bacteria sense the chemoattractant. Rather than modelling directly how the turning angle depends

on S (which resulted in our having to choose $k(S)$ in the chemotactic sensitivity $\chi(S)$), we let it depend on some internal state of the bacterium. This pathway is known to be very complex indeed, and mathematical models of its reaction dynamics often contain 30 or more dependent variables. Fortunately, it has been shown convincingly that this system may be approximated well using two phenomenologically chosen variables, a fast excitation variable y_1 changing at time scale τ_e, and a slow adaptation variable y_2 varying at time scale τ_a. See Figure 7.2 for the kind of dynamics this creates. These two ingredients of excitation and adaptation are very commonly found in many sensory systems following an external concentration. We now introduce an internal state $y = (y_1, y_2)$ for each individual particle, evolving according to

$$\frac{dy}{dt} = h(y, S) \tag{7.3.17}$$

or more specifically,

$$\tau_e \frac{dy_1}{dt} = g(S(x)) - (y_1 + y_2) \tag{7.3.18}$$

$$\tau_a \frac{dy_2}{dt} = g(S(x)) - y_2 \tag{7.3.19}$$

We may resort to the conservation equation (7.1.2) again, but now the flux is not in space or in velocity space, but in internal state space. The evolution of particle densities and their internal states can be given by

$$\frac{\partial u}{\partial t} + \nabla_y \cdot (hu) = 0$$

So rather than treating y as an dependent variable, we can treat it as an independent variable. This is entirely analogous to the situation in which $dx/dt = v$ and u solves $u_t + \nabla \cdot (vu) = 0$.

Putting this new ingredient into the transport equation is straighforward. Now $u(t, x, y, v)$ satisfies

$$\frac{\partial u}{\partial t} + \nabla_x \cdot (vu) + \nabla_y \cdot (hu) = -\lambda(y)u + \lambda(y) \int_V T(v, v', y)u(t, x, v', y)dv' \tag{7.3.20}$$

The three-step technique works also for this more elaborate example. Assuming that the turning rate depends linearly on the excitation variable (which detects the signal and thus influences when to change direction) we write $\lambda(y) = \lambda_0 - by_1$ for $\lambda_0 > 0$, $b > 0$. A macroscopic evolution equation for $U(x, t) = \int \int u(t, x, y, v)dydv$, can now be derived, and is indeed a classical Keller-Segel equation for large times

$$\frac{\partial U}{\partial t} = \nabla \cdot \Big(\frac{s^2}{\lambda_0} \nabla U - \Big[\frac{bs^2 \tau_a g'(S(x))}{\lambda_0 (1 + 2\lambda_0 \tau_a)(1 + 2\lambda_0 \tau_e)} \Big] U \nabla S \Big). \tag{7.3.21}$$

Not only have we understood under what circumstances mesoscopic dynamics give rise to Keller-Segel equations, we have also obtained insight how the parameters specifying the individual bacteria's behaviour influence the diffusion or aggregation parts of the final macroscopic equation (7.3.21).

This is but one way to derive Keller-Segel models from lower-level dynamics. One may also start from a stochastic description, see e.g. Stroock (1974); Stevens (1995). For more information and much detail on recent developments on velocity-jump processes and their relation to chemotaxis models, see Alt (1980); Othmer et al. (1988); Othmer and Hillen (2002); Erban and Othmer (2004, 2007); Xue and Othmer (2009). The analysis of the resulting chemotaxis models has become a large field. The interested reader may consult Horstmann Horstmann (2003) for a detailed review of many mathematical aspects of chemotaxis.

Chapter 8

Physiologically structured populations

Often we are not only interested in just the spatial or temporal dynamics of a population, but also in the *internal* structure of the individuals in this population. Examples include the demography of a population (how many individuals of what age), or its size distribution. The first of these, age structure, is considerably easier than the latter, since age simply passes linearly with time, while size may depend on the environment (how much food was consumed, for instance). This chapter thus focuses on the simpler case of age structure. This does not preclude interesting behaviour, as we will see when discussing the consumption of one's own kind: cannibalism of adults of juveniles.

In any model description of physiologically structured populations, a distinction must be made between the state of an individual or *i-state*, and how it changes, and the population state or *p-state* and its change. Individuals are subject to the rest of the population and the environment, and the whole population in turn changes as a result of individual births and deaths (and potential other causes, such as immigration). This distinction thus immediately means more involved bookkeeping and a greater proliferation of processes and constants than in more macroscopic model approaches such as reaction-diffusion equation. Nevertheless, much of the modelling we will encounter boils down to mere bookkeeping and well-chosen notation.

8.1 Age structure

Let us first introduce the function $n(t, a)$ which denotes the density of individuals of age a at time t. The total number of individuals at time t will be denoted by $A(t)$, and these together give birth to offspring at a rate $b(t)$. We further denote the per capita birth rate at age a by $\beta(a)$. The population birth rate $b(t)$ is given by

$$b(t) = \int_0^\infty \beta(a) n(t, a) da. \tag{8.1.1}$$

Individuals are assumed to survive up to age a with probability $\mathcal{F}(a)$, and die at a per capita rate $\mu(a)$. The number of individuals of age a at time t is then those individuals that were born a time ago and are still alive,

$$n(t, a) = b(t - a)\mathcal{F}(a). \tag{8.1.2}$$

The relation between $\mathcal{F}$ and μ is given by

$$\frac{d\mathcal{F}}{da} = -\mu(a)\mathcal{F}. \tag{8.1.3}$$

which can be solved using an initial condition $\mathcal{F}(0) = 1$,

$$\mathcal{F}(a) = e^{-\int_0^a \mu(\alpha)d\alpha}. \tag{8.1.4}$$

Equivalently, we can express μ in $\mathcal{F}$,

$$\mu(a) = -\frac{d}{da}\log \mathcal{F}(a).$$

Usually, age-specific mortality rates can not be estimated reliably from demographic data, but $\mathcal{F}(a)$ *can* be, and is called the cohort-survival probability.

If we start an experiment at time $t = 0$ with some known initial age density $n(0, a)$, then (8.1.2) only holds for $0 \leq a \leq t$. For $a > t$ we instead have the conditional survival probability

$$n(t, a) = \frac{\mathcal{F}(a)}{\mathcal{F}(a-t)} n(0, a-t).$$

After all, these individuals already had a positive age at time zero, and had to have survived from birth to the initial time of the experiment. Now, (8.1.1) becomes

$$b(t) = \int_0^t \beta(a)\mathcal{F}(a)b(t-a)da + f(t) \tag{8.1.5}$$

with

$$f(t) = \int_t^\infty \frac{\mathcal{F}(a)}{\mathcal{F}(a-t)}\beta(a)n(0, a-t)da. \tag{8.1.6}$$

Note that $f(t)$ is a known quantity since it involves the initial condition $n(0, t)$ which we assume to be known. An equivalent way to prescribe an initial condition would be to supply a function $b(a)$ for $a \leq 0$: if we take $t = 0$ in (8.1.2), we find

$$n(0, a) = b(-a)\mathcal{F}(a), \quad a \geq 0,$$

which allows us to find $n(0, a)$ from $b(a)$ and vice versa.

Another useful quantity known from data is the average number of offspring produced by age a, $L(a)$, given by

$$L(a) = \int_0^a \beta(a)\mathcal{F}(a)da, \tag{8.1.7}$$

which is equivalent to

$$\beta(a) = \frac{1}{\mathcal{F}(a)}\frac{dL}{da}(a).$$

Combining (8.1.1) and (8.1.2) we find the *renewal equation*

$$b(t) = \int_0^\infty \beta(a)\mathcal{F}(a)b(t-a)da, \tag{8.1.8}$$

a Volterra integral equation of convolution type, which is to be solved with the initial condition $b(a)$, $a < 0$. Both (8.1.5) and (8.1.8) are called *renewal equations*.

An alternative way to describe the dynamics of age structure is to keep in mind that age increases at rate 1 as time progresses, so that $n(t, a)$ satisfies the first-order partial differential equation

$$\frac{\partial n}{\partial t} + \frac{\partial n}{\partial a} = -\mu n, \tag{8.1.9}$$

first studied by McKendrick (1926). As boundary condition we supply a birth rate, $n(t, 0) = b(t)$ by (8.1.1). The left hand side of (8.1.9) is just a version of the general bookkeeping rule (4.1.7) from Chapter 4, i.e., (8.1.9) is a transport equation. Here the flux is $\frac{\partial n}{\partial a}\frac{da}{dt} = \frac{\partial n}{\partial a}$ since $\frac{da}{dt} = 1$.

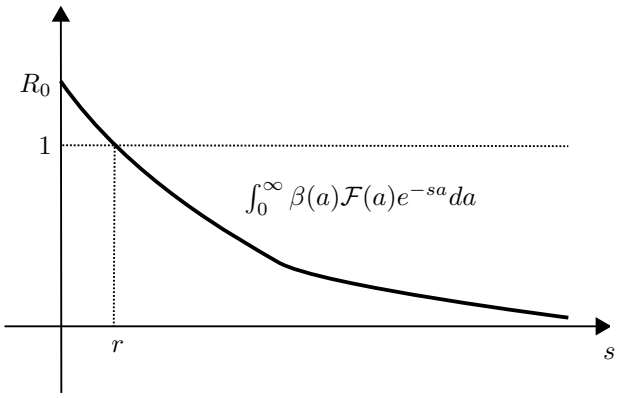


FIGURE 8.1. For $R_0 > 1$, the characteristic equation (8.1.10) has a real positive solution r.

8.1.1 Growth or decline

Equation (8.1.5) is more complicated than (8.1.8), but is more useful for existence and uniqueness questions. However, as time progresses the extra term $f(t)$ vanishes since all the individuals that existed at time 0 will have died. Therefore, (8.1.5) becomes identical to (8.1.8). Since this equation is linear in b we can try to find possible solutions by setting $b(t) = b_0 e^{rt}$. Such a solution satisfies (8.1.8) if

$$1 = \int_0^\infty \beta(a)\mathcal{F}(a)e^{-ra}da. \tag{8.1.10}$$

This is the *Euler-Lotka equation* for the population growth rate r. The *basic reproduction number*

$$R_0 := \int_0^\infty \beta(a)\mathcal{F}(a)da = L(\infty) \tag{8.1.11}$$

gives the population growth on a *generation basis*, while the *Malthusian parameter* r measures the real time growth of the population. It is also called the intrinsic rate of increase. The most important relation between R_0 and r is that

$$R_0 > 1 \iff r > 0. \tag{8.1.12}$$

More precisely, equation (8.1.10) has a strictly positive real root r if and only if $R_0 > 1$, which is conveniently illustrated in Figure 8.1. Intuitively, it seems obvious that a population can only grow in real time if and only if it can grow on a per-generation basis. The Euler-Lotka equation states that the discounting rate r should be adjusted to make the discounted basic reproduction number exactly 1.

Cautionary example. If we assume that cells give birth to two offspring exactly at time 1, then L jumps from zero to two at age 1 (ignoring the possibility of dying), and the difference equation $b(t) = 2b(t-1)$ leads to the characteristic equation $1 = 2e^{-z}$ with roots $z = \log 2 + 2\pi ki$, $k \in \mathbb{Z}$. This is analogous to the situation with discrete age classes, modelled using Leslie matrices, in which the matrix is nonprimitive (see Section 2.3.1).

The large-time behaviour of the renewal equation for the birth rate, equation (8.1.5), is best studied using the Laplace transform, which is defined by

$$\bar{h}(s) = \int_0^\infty h(t)e^{-st}dt.$$

Its most useful feature in the present context is that the Laplace transform converts convolutions into products, so that (8.1.5) is equivalent to

$$\bar{b}(s) = \overline{\beta\mathcal{F}}(s)\bar{b}(s) + \bar{f}(s),$$

which we can solve for $\bar{b}(s)$ to find

$$\bar{b}(s) = \frac{\bar{f}(s)}{1 - \overline{\beta\mathcal{F}}(s)}. \tag{8.1.13}$$

Note that the Euler-Lotka equation (8.1.10) also has the form of the Laplace transform of $\overline{\beta\mathcal{F}}(s)$. Indeed,

$$\overline{\beta\mathcal{F}}(s) = \int_0^\infty \beta(a)\mathcal{F}(a)e^{-sa}da = 1$$

when s equals the intrinsic growth rate r. All other roots have real part smaller than r. We may transform $\bar{b}(s)$ back to $b(t)$ using the Inverse Laplace transform

$$h(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \bar{h}(s)e^{st}ds,$$

where the c is greater than the real part of all singularities of $\bar{h}(s)$. In the case that the roots of the Euler-Lotka equation are simple, denoted by r, s_1, s_2, ..., the original birth rate, backtransformed from (8.1.13), has a particularly transparent form,

$$b(t) = \frac{f(r)}{(\overline{\beta\mathcal{F}})'(r)}e^{rt} + \sum_{i=1}^{\infty}\frac{f(s_i)}{(\overline{\beta\mathcal{F}})'(s_i)}e^{s_i t}.$$

This shows that the large-time behaviour of the birth rate is dominated by the first term, which indeed becomes the asymptotic growth rate for the total population.

One can also show that $n(t,a)$ for almost all initial conditions $n(0,a) = n_0(a)$, converges towards a stable age distribution $v(a)$ that grows exponentially at rate r (Feller, 1941; Gripenberg et al., 1990; Diekmann et al., 1995), where

$$v(a) = \frac{e^{-ra}\mathcal{F}(a)}{\int_0^\infty e^{-r\alpha}\mathcal{F}(\alpha)d\alpha}.$$

The age density $v(a)$ is found by interpretation: it is the density of individuals of age a, so that

$$\begin{aligned} v(a) &= \lim_{t\to\infty}\frac{n(t,a)}{\int_0^\infty n(t,\alpha)d\alpha} \\ &= \lim_{t\to\infty}\frac{b(t-a)\mathcal{F}(a)}{\int_0^\infty b(t-\alpha)\mathcal{F}(\alpha)d\alpha} \\ &\approx \frac{e^{r(t-a)}\mathcal{F}(a)}{\int_0^\infty e^{r(t-\alpha)}\mathcal{F}(\alpha)d\alpha} \text{ for } t \text{ large} \\ &= \frac{e^{-ra}\mathcal{F}(a)}{\int_0^\infty e^{-r\alpha}\mathcal{F}(\alpha)d\alpha}. \end{aligned}$$

Note that the age density is steeply declining when the population is growing fast. In other words, a slowly growing population has many elderly people at steady state, which reflects the current situation in many western countries.

A much more general discussion of Laplace transforms to solve renewal equations, and indeed of delay equations in general, may be found in (Diekmann et al., 1995).

8.2 Cannibalism

Things start to become more interesting in age structure models if the old consume the young. Under what conditions can cannibalism have a positive effect

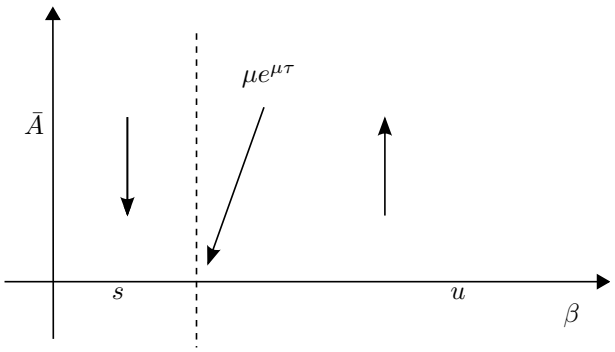


FIGURE 8.2. Without taking cannibalism into account we find that for small per capita birth rates β, populations (indicated by size of the adult population $\bar{A}$) crash (s for stable trivial solution), while for large ones they grow unbounded (u for unstable trivial solution.

on total (or just adult) population sizes? To start, let us divide the population into adults A, and juveniles, which mature to adulthood at a fixed age τ.

8.2.1 Focusing on the victims: higher juvenile death rates

The condition for long term survival of the population is given by $R_0 > 1$. With the introduction of adults and juveniles, the long term survival of the population is given by

$$\beta \frac{1}{\mu} \mathcal{F}(\tau) > 1,$$

Simply put, the expected number of offspring from one newborn individual is the product of the birth rate, the life expectancy of an individual and the probability to reach adulthood, and this should exceed one. To keep things simple, we assume that both juveniles and adults have a fixed death rate μ (due to all other causes than cannibalism), so that

$$\mathcal{F}(\tau) = e^{-\mu\tau}.$$

From the previous discussion we know that the adult population either collapses if $R_0 < 1$, or grows unbounded when $R_0 > 1$. When these inequalities are written from the perspective of birth rate β, we find Figure 8.2.

We now build up the model by first considering the fate of juveniles, by augmenting the juvenile death rate but keeping the birth rate fixed, i.e., independent of cannibalism. The death rate μ is assumed to become

$$\mu + h(a)A$$

where $h(a)$ is a vulnerability function that depends on the age a of the juvenile. Note that we have assumed independence of individuals here, to invoke the Law of Mass Action (the cannibalism rate is assumed to be proportional to the product of juvenile and adult population sizes). We assume that h has a graph more or less as in Figure 8.3: its support is confined to $(0, \tau)$.

Now we introduce an important piece of notation. Let ϕ_t denote the *history of a function* ϕ, in the sense that

$$\phi_t(\theta) := \phi(t + \theta), \quad \theta \leq 0.$$

In this notation, individuals born at time $t - a$ survive to age a at time t with a probability that depends on the size of the adult population during the time period

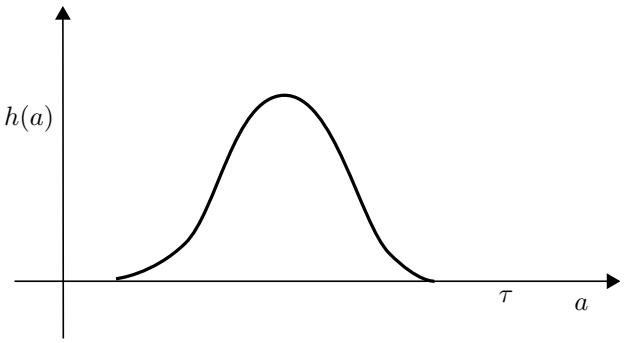


FIGURE 8.3. The vulnerability function $h(a)$ as function of age a, with compact support in $(0, \tau)$.

$[t-a, t]$,

$$\mathcal{F}(a, A_t) = e^{-\mu a - \int_0^a h(\alpha) A(t-a+\alpha) d\alpha},$$

As we said at the start, assume that cannibalism does not augment the birth rate of adults, and keep β constant. The renewal equation for b now becomes

$$b(t) = \beta \int_\tau^\infty \mathcal{F}(a, A_t) b(t-a) da. \tag{8.2.1}$$

By definition, the adults in the population comprise those individuals that were born more than τ time ago and survived to reach adulthood,

$$A(t) = \int_\tau^\infty \mathcal{F}(a, A_t) b(t-a) da. \tag{8.2.2}$$

By our assumptions on h, we can evaluate this integral in steps, which is called the "method of steps" in the theory of delay equations. We can take the derivative of $A(t)$ using (8.2.2) to arrive at

$$\frac{dA}{dt}(t) = b(t-\tau)\mathcal{F}(\tau, A_t) - \mu A(t),$$

which may be interpreted in biological terms in a clear way: the number of adults increases due to births τ time ago, and surviving their youth in the face of cannibalism, and decreases due to natural mortality. If we use that

$$b(t) = \beta A(t)$$

in dA/dt, we find

$$\frac{dA}{dt}(t) = \beta A(t-\tau)\mathcal{F}(\tau, A_t) - \mu A(t), \tag{8.2.3}$$

which nicely summarises this model in one delay differential equation.

Let us now focus on steady states and see if Figure 8.2 changes already. A prerequisite for a steady state is that the population growth rate over generations, R_0, is exactly one. The adult population size is of course a constant A. Recall that we have

$$R_0 = L(\infty) = \beta \int_0^\infty \mathcal{F}(a, A) da,$$

so that

$$R_0 = \beta \frac{1}{\mu} e^{-\mu\tau - HA},$$

where

$$H = \int_0^\tau h(\sigma) d\sigma.$$

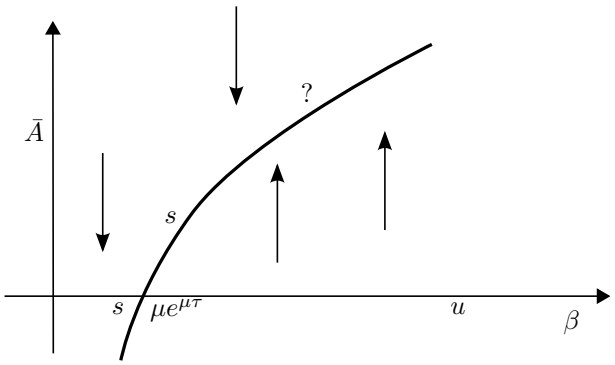


FIGURE 8.4. A supercrticical transcritical bifurcation. Stability changes at $\beta = \mu e^{\mu\tau}$, but we don't yet know whether this stability is maintained for larger β, as indicated by the question mark.

The steady state equation $R_0(A) = 1$ has solution

$$A = \bar{A} = \frac{1}{H} \log\Big(\frac{\beta}{\mu} e^{-\mu\tau}\Big).$$

For biologically relevant steady states, we need $\bar{A} \geq 0$, which means that

$$\beta > \mu e^{\mu\tau},$$

the same value as before. Plotting the steady state $\bar{A}$ as a function of β, we find Figure 8.4, in constrast to Figure 8.2. From this analysis we only find that the steady state is stable near the bifurcation point (See the Appendix for a brief introduction to bifurcation theory). However, as the question mark in the figure indicates, it is not yet clear whether the steady state remains stable for larger β. This is not so easy to establish.

To summarise, the point where populations become viable remains the same, but we now find that populations do not grow unbounded, but that they are kept in check by cannibalism.

8.2.2 Focusing on the perpetrators: including higher birth rates

So far, we have assumed that cannibalism was uncommon enough (with respect to additional resources available to the adults) that birth rates were not affected by it. Now we do include a higher birth rate due to consumption of juveniles, and assume that β becomes $\beta + B(t)$ where

$$B(t) = \int_0^\tau E(\alpha) h(\alpha) \mathcal{F}(\alpha, A_t) b(t-\alpha) d\alpha. \tag{8.2.4}$$

In words, cannibalised juveniles that contribute to extra births are born at time $t - \alpha$, survive to the age of being eaten by conspecifics with probability $\mathcal{F}(\alpha, A_t)$ (which thus depends on the historic population size of adults, are predated upon with rate $h(\alpha)$ and converted into offspring according to some conversion function $E(\alpha)$). The renewal equation for the birth rate now becomes

$$b(t) = (\beta + B(t)) \int_\tau^\infty \mathcal{F}(a, A_t) b(t-a) da. \tag{8.2.5}$$

The complete model is now given by equation (8.2.5) for b, (8.2.4) for B, and (8.2.3) for A.

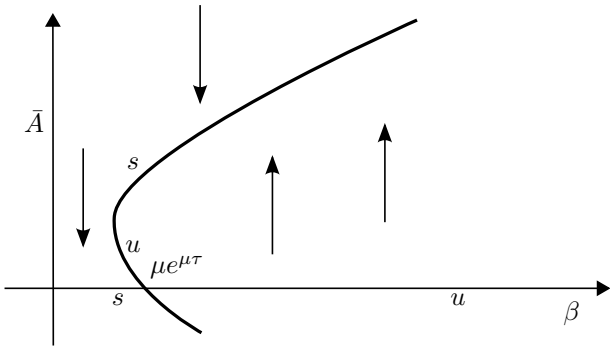


FIGURE 8.5. A subcritical transcritical bifurcation

The steady states are given by solving the following three equations for the three constants, b, A and B,

$$R_0 = 1 \iff 1 = (\beta + B)\frac{1}{\mu}e^{-\mu\tau - HA}, \tag{8.2.6}$$

$$A = \beta \int_\tau^\infty \mathcal{F}(a, A)da = \frac{\beta}{\mu}e^{-\mu\tau - HA}, \tag{8.2.7}$$

$$B = b\int_0^\tau E(\alpha)h(\alpha)e^{-\mu\alpha - \int_0^\alpha h(\sigma)d\sigma}d\alpha. \tag{8.2.8}$$

Solving (8.2.7) for b, substituting this into (8.2.8), then solving for B and finally substituting this into (8.2.6), we find

$$\beta = \mu e^{\mu\tau + HA}\Big(1 - A\int_0^\tau E(\alpha)h(\alpha)e^{-\mu\alpha - A\int_0^\alpha h(\sigma)d\sigma}d\alpha\Big). \tag{8.2.9}$$

This is one equation in A, with solution $\bar{A}$. It of course gives β as a function of $\bar{A}$, so we can at least make a sketch. Figure 8.5 is one such sketch, where a subcritical transcritical bifurcation is shown. The criterion for this to be the case is given by the sign of $\beta'(0)$,

$$\frac{d\beta}{dA}\Big|_{A=0} = He^{\mu\tau}\Big(1 - \frac{1}{H}\int_0^\tau E(\alpha)h(\alpha)e^{-\mu a}da\Big). \tag{8.2.10}$$

So we find a supercritical bifurcation such as in Figure 8.4 when

$$\frac{1}{H}\int_0^\tau E(\alpha)h(\alpha)e^{-\mu a}da > 1, \tag{8.2.11}$$

and a subcritical bifurcation, see Figure 8.5, when

$$\frac{1}{H}\int_0^\tau E(\alpha)h(\alpha)e^{-\mu a}da < 1. \tag{8.2.12}$$

This condition has an interesting biological interpretation, revolving around the positive and negative gains from cannibalism. To derive this interpretation, we introduce $\Phi(a)$, the expected number of newborns produced as a result of cannibalism, per newborn, up to reaching age a, at constant adult population size A. Note that this quantity is not the complete benefit of cannibalism, since these newborns themselves give rise to $\Phi(a)$ new offspring through cannibalism, making the total number of newborns of these first two generations $\Phi(a)+\Phi^2(a)$. Summing

over all generations in this way, we find that the total benefit is thus

$$\sum_{k=1}^{\infty} \Phi^k. \tag{8.2.13}$$

This variable Φ satisfies

$$\begin{cases} \frac{d\Phi}{da}(a) & = AE(a)h(a)e^{-\mu a - A\int_0^a h(\sigma)d\sigma}, \\ \Phi(0) & = 0, \end{cases}$$

so that

$$\Phi(\tau) = A\int_0^{\tau} E(\alpha)h(\alpha)e^{-\mu\alpha - A\int_0^{\alpha} h(\sigma)d\sigma}d\alpha. \tag{8.2.14}$$

For small A, $\Phi(\tau) < 1$, and hence our focal quantity (8.2.13) equals

$$\frac{1}{1-\Phi(\tau)}.$$

Now we turn to the negative effect of cannibalism, which is much simpler: it is the term e^{-HA} in the probability to survive cannibalism till adulthood. The net gain from cannibalism is thus positive if the overall growth rate exceeds one, i.e.,

$$\frac{e^{-HA}}{1-\Phi(\tau)} > 1,$$

which we rewrite as

$$e^{-HA} > 1 - \Phi(\tau)$$

Close to $A = 0$, e^{-HA} is well approximated by $1 - HA$, so that the inequality above becomes $1 - HA > 1 - \Phi(\tau)$, i.e.,

$$\frac{\Phi(\tau)}{HA}\Big|_{A=0} > 1.$$

Using expression (8.2.14) for Φ, we find precisely condition (8.2.11) which we sought to derive.

Finally, to get some feeling for the behaviour of steady states away from the bifurcation point $A = 0$, assume that $h(a)$ is concentrated at a point, $h(a) = H\delta(a - \hat{\tau})$. Then the birth rate β becomes

$$\beta = \mu e^{\mu\tau + HA}\Big(1 - E(\hat{\tau})e^{-\mu\hat{\tau}}(1 - e^{-HA})\Big).$$

Therefore,

$$\frac{d\beta}{dA}(A) = H\mu e^{\mu\tau + HA}(1 - E(\hat{\tau})e^{-\mu\hat{\tau}}).$$

The last factor was found as the condition for *local* stability of $A = 0$, but since the above argument is valid all along the $\beta(A)$ curve, it is now also the *global* condition.

One remarkable point to note is that if we assume that the birth rate due to normal (non-cannibalistic feeding) β is zero, then the steady state $\bar{A}$ solves

$$\bar{A} = -\frac{1}{H}\log\left(1 - e^{\mu\tau}(E(\hat{\tau}))^{-1}\right),$$

which is potentially positive! In that case, the "bend" in Figure 8.5 extends beyond the $\bar{A}$-axis. The adult population manages to feed exclusively on juvenile conspecifics, and the only energy uptake from outside the population is from juveniles feeding on other food. Indeed, there are lakes in which there is only one, predatory, fish species, whose juveniles feed exclusively on zooplankton, and all adults are obligate cannibalists. See (van den Bosch et al., 1988) for a full account of this remarkable story.

Chapter 9

Infectious Diseases

There was a time (the middle of the last century) when the World Health Organization declared impending victory in the battle against infectious diseases and expressed that medicine should now focus on cardiovascular diseases and cancer. Today we are aware that bugs hit back hard and fast (evolution is an ally of anybody and does not care about human interest in particular), and that new bugs emerge. On top of that, both livestock and crops are vulnerable to a multitude of bacteria, viruses and fungi. So any tool that enhances our understanding of infectious disease dynamics is most welcome (since understanding is a first step towards control). The present chapter is a short introduction to mathematical modeling tools in this direction. The monograph by Diekmann et al. (2012) provides a more encompassing treatment of the subject.

9.1 The force of infection

A very special feature of population dynamics of micro-parasites is that the parasites are represented indirectly in the form of infected hosts (for macro-parasites, in particular worms, hosts are often characterized in terms of parasite load, giving a somewhat more direct representation). A manifestation of this indirectness is that the analogue of the birth rate is the *incidence*, i.e., the number of new cases per unit of time. Ignoring, for convenience, the practical complication of delayed onset of symptoms and delayed reporting, and pretending that a newly infected host is immediately recognizable as such, the incidence is

$$i(t) = \frac{\#\text{ new cases}}{\text{time}} \quad \text{at time } t. \tag{9.1.1}$$

As the analogue of the age of an individual we introduce

$$\tau = \text{time on the clock that starts when host gets infected.} \tag{9.1.2}$$

With some delay (whence the introduction of τ) infected individuals produce propagules that have the potential to infect other hosts. We define

$$S(t) = \text{size of the population of susceptible individuals,} \tag{9.1.3}$$

where "size" may refer to spatial density, to absolute numbers or to the fraction of the total population, depending on the context. We postulate

$$i(t) = F(t)S(t), \tag{9.1.4}$$

and say

$$\begin{aligned} F(t) &= \text{force of infection at time } t \\ &= \text{probability per unit time for a susceptible to become infected.} \end{aligned} \tag{9.1.5}$$

The constitutive equation

$$F(t) = \int_0^\infty i(t-\tau)A(\tau)\,d\tau \tag{9.1.6}$$

closes the feedback loop. It expresses that the force of infection is made up from contributions by individuals that themselves were infected some time τ ago. Indeed

$$A(\tau) = \text{expected contribution to the force of infection at time } \tau \text{ after becoming infected.} \tag{9.1.7}$$

So $A(\tau)$ is proportional to the expected production of propagules at time τ after becoming infected. The constant of proportionality incorporates the relevant quantification of the contact process that provides suitable opportunities for propagules to jump over from infected hosts to susceptible hosts. The function $A(\tau)$ is the key modeling ingredient.

By combining (9.1.4) and (9.1.6) we obtain the *renewal equation*

$$i(t) = S(t)\int_0^\infty i(t-\tau)A(\tau)\,d\tau \tag{9.1.8}$$

for the incidence i. To complete the model formulation, we need to specify the dynamics of S.

9.2 The introductory phase

Susceptibles are the substrate for the pathogen. Susceptibles are needed for reproduction, but are "consumed" in the process, in particular when immunity results. If no new fuel is provided, a fire will go extinct. If we light a pile of wet wood, extinction is immediate. If we light a pile of very dry wood, a blazing fire may result and yet, ultimately, the fire goes extinct. Even though the end result is the same, the two situations are clearly rather different and we want to distinguish them from each other. To do so, we ignore that the fire itself diminishes the amount of combustible material. This denial of a fact allows us to easily and clearly make the distinction between the introductory phase of the two cases. Once the distinction is made, we should realize that in the case of the dry wood our denial is exposed, that we need to take the fact into account once the fire becomes substantial.

The "denial of a fact" alluded to above, is in mathematics called "linearization". The usual procedure is to formulate a complete nonlinear model, compute some derivatives and then use these to define the linearized system. When discussing the introduction of a species, there is a shortcut! One has to think in terms of environmental condition and feedback. In the present setting, and from the point of view of the parasites, the environmental condition is fully determined by $S(t)$. Feedback entails that S decreases when hosts get infected. Linearization simply means that we ignore the feedback. Indeed, if we consider (9.1.8) and assume that $i(t)$ is small, then small deviations from a given $S(t)$ lead to quadratically small terms that are ignored in the linearization.

Finally, a relatively easy situation of great relevance is when the environmental condition is constant in time. The key point is that, when the environment is constant, it does not matter *when* you are "born" (i.e., infected), and hence it makes perfect sense to study the (implicit discrete time) dynamics of generations. We will use this heavily in Section 9.4 below.

So let us take $S(t)$ to be constant in time, S^0. The linear renewal equation

$$i(t) = S^0\int_0^\infty i(t-\tau)A(\tau)\,d\tau \tag{9.2.1}$$

is, apart from notation, identical to (8.1.8). So motivated by (8.1.11) we call

$$R_0 = R_0(S^0) = S^0 \int_0^\infty A(\tau)\, d\tau \tag{9.2.2}$$

the *basic reproduction number*. In the current context

$$R_0 = \text{expected number of secondary cases per primary case.} \tag{9.2.3}$$

The corresponding Euler-Lotka equation

$$1 = S^0 \int_0^\infty e^{-\lambda\tau} A(\tau)\, d\tau \tag{9.2.4}$$

has a unique real root $\lambda = r$ and

$$\operatorname{sign} r = \operatorname{sign}(R_0 - 1). \tag{9.2.5}$$

If $r > 0$ and the incidence grows like e^{rt}, we say that a *major outbreak* is initiated. If $r < 0$ we say that only *minor outbreaks* are possible (cf. the wet wood). We say

$$R_0 \text{ has threshold value one} \tag{9.2.6}$$

to express that for a major outbreak to be possible, R_0 should exceed the value one. Since R_0 is proportional to S^0 we can, in particular when S^0 is a spatial density, also say that the population density has to exceed a critical density before a large outbreak is possible. And indeed, the burden of infectious disease for mankind increased a lot when, in the Middle Ages, population density in walled towns grew substantially.

9.3 The final size in a closed population

If the population is demographically closed (no birth/death, no emigration/immigration) and the disease leads to permanent immunity, the change in $S(t)$ is entirely due to new infections. Or, in equations

$$\frac{dS}{dt}(t) = -i(t) = -F(t)S(t) \tag{9.3.1}$$

and, upon integration

$$S(t) = S^0 e^{-\int_{-\infty}^t F(\sigma)\, d\sigma}, \tag{9.3.2}$$

where S^0 is the size of the susceptible population way back in the past, before the disease entered stage. Combining (9.1.4), (9.3.2) and (9.1.6) we obtain the nonlinear renewal equation

$$F(t) = S^0 \int_0^\infty F(t-\tau) e^{-\int_{-\infty}^{t-\tau} F(\sigma)\, d\sigma} A(\tau)\, d\tau \tag{9.3.3}$$

Integrating this identity from $-\infty$ to t, while using

$$F(t-\tau) e^{-\int_{-\infty}^{t-\tau} F(\sigma)\, d\sigma} = -\frac{d}{dt} e^{-\int_{-\infty}^{t-\tau} F(\sigma)\, d\sigma},$$

we deduce

$$\int_{-\infty}^t F(\sigma)\, d\sigma = S^0 \int_0^\infty \Big[1 - e^{-\int_{-\infty}^{t-\tau} F(\sigma)\, d\sigma}\Big] A(\tau)\, d\tau. \tag{9.3.4}$$

This is a convolution type nonlinear renewal equation for

$$\int_{-\infty}^t F(\sigma)\, d\sigma = -\ln \frac{S(t)}{S^0} \tag{9.3.5}$$

(cf. (9.3.2)). Taking the limit $t \to \infty$ we find

$$-\ln \frac{S(\infty)}{S^0} = \Big(1 - \frac{S(\infty)}{S^0}\Big) R_0. \tag{9.3.6}$$

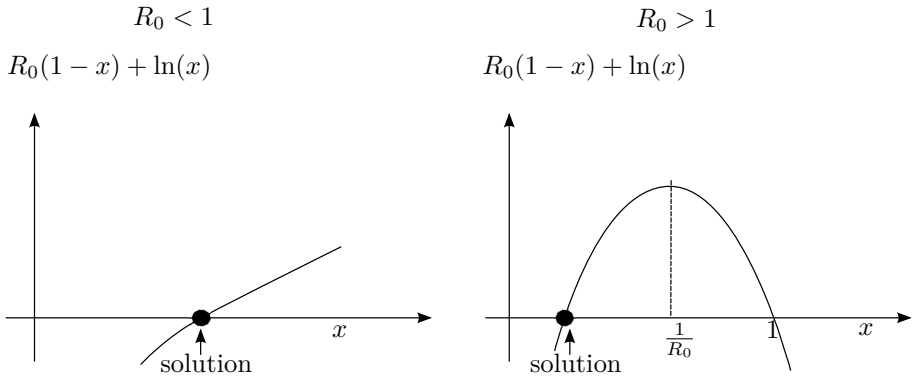


FIGURE 9.1. Solutions of (9.3.6) for $R_0 < 1$ (left) and $R_0 > 1$ (right).

(A rigorous proof combines monotonicity with respect to t with integrability of A.) A simple graphical argument establishes that (9.3.6) has a unique solution with

$$0 < \frac{S(\infty)}{S^0} < 1$$

when $R_0 > 1$, and no such solution (meaning that $S(\infty)/S^0 = 1$ is the relevant solution) when $R_0 < 1$, see Figure 9.1.

So once again we conclude that a major outbreak requires $R_0 > 1$ but now we found the additional information that the *size* of such a major outbreak, expressed by $1-S(\infty)/S^0$ as the fraction of the population that falls victim to the disease, can be determined by finding the appropriate root of (9.3.6). In particular we conclude that some fraction escapes (note, however, that $S(\infty)/S^0 \sim \exp(-R_0)$ for $R_0 \gg 1$).

Cautionary remark: the final size equation depends on the assumptions about the contact process, see (Diekmann et al., 2012, Section 1.3.3).

9.4 The probability of a minor outbreak (an interlude on branching processes)

Some of the formulations above were a bit awkward or cryptic, since we wanted to avoid suggesting that $R_0 > 1$ guarantees that a major outbreak occurs. It *does* once a small *fraction* of the large (to justify our deterministic description) population is infected. But it *does not* if only a small *number* of individuals is infected. The point is that the potential to generate a lot of secondary cases is computed as an average, an expected value, but that as long as numbers are small we need to take into account that the actual realizations may not at all realize this potential. Bad luck can have decisive impact when numbers are small. Indeed, the pathogen may go extinct after triggering only a minor outbreak.

Branching processes are stochastic models that can capture such effects very well. We now return to the setting of Section 9.2, i.e., we again ignore the feedback and assume that the availability of substrate, of susceptibles, is constant in time. Recall that in a constant environment we can adopt a generation perspective.

A stochastic formulation of "ignore the feedback" reads "assume that individuals behave (in our case: reproduce) independently of each other". The specification of a model then amounts to providing numbers

$$q_k \geq 0, \quad k = 0, 1, 2, 3, \ldots \tag{9.4.1}$$

such that

$$\sum_{k=0}^{\infty} q_k = 1 \tag{9.4.2}$$

and with the interpretation that q_k is the probability that an individual produces k offspring during its entire life. In more technical jargon: the total number of offspring of an individual is a random variable with probability distribution $\{q_k\}_{k=0}^{\infty}$. One also says: offspring is an independent and identically distributed (abbreviated to i.i.d.) random variable. Note that in the context of infectious diseases one should interpret "offspring" as "host infected by the individual under consideration".

To be able to characterize the probability of extinction, it is useful to introduce the *generating function*

$$g(z) := \sum_{k=0}^{\infty} q_k z^k \tag{9.4.3}$$

and to verify that g has the following properties

$$\begin{aligned} &i)\ g(0) = q_0, \\ &ii)\ g(1) = 1, \\ &iii)\ g'(1) = \sum_{k=1}^{\infty} k q_k, \\ &iv)\ g'(z) > 0,\ \text{for } z \geq 0, \\ &v)\ g''(z) > 0 \text{ for } z \geq 0. \end{aligned} \tag{9.4.4}$$

Now assume that $q_0 > 0$, which means that with positive probability an individual will beget no offspring at all. If we start with exactly one individual, then q_0 is the probability that the first generation consists of zero individuals. More generally, let us define

$$\begin{array}{rcl} z_n & := & \text{probability that the } n\text{th generation consists of} \\ & & \text{zero individuals, given that the zeroth generation} \\ & & \text{consists of one individual} \end{array}$$

Then the above observation can be reformulated as $z_1 = q_0$.

If the first generation consists of k individuals, then, in order that the second generation consists of zero individuals, each of these should die childless. The independence implies that this event has probability q_0^k. Thus we find

$$z_2 = q_0 + \sum_{k=1}^{\infty} q_k q_0^k = q_0 + \sum_{k=1}^{\infty} q_k z_1^k,$$

which is the case $n = 1$ of the general recursion relation

$$z_{n+1} = q_0 + \sum_{k=1}^{\infty} q_k z_n^k = g(z_n). \tag{9.4.5}$$

The interpretation of (9.4.5) is as follows: the $(n+1)$-generation consists of zero individuals if already the first generation consisted of zero individuals, or if the first generation consisted of k individuals and each of these seeds a clan that is extinct by generation n.

Using the relations (9.4.4) we can easily analyse the recursion (9.4.5) graphically by making "staircase plots", see Figure 9.2.

If $g'(1) \leq 1$ then $\lim_{n\to\infty} z_n = 1$, or in words, the population goes extinct with probability one.

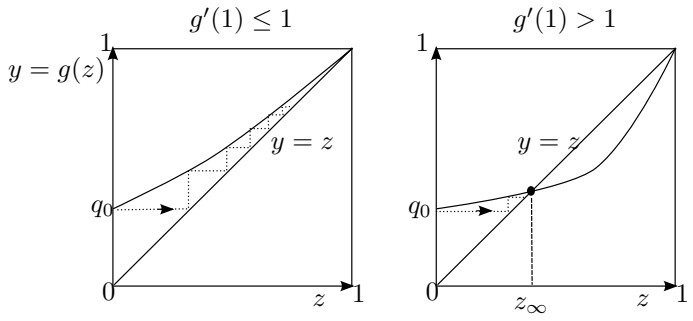


FIGURE 9.2. Staircase plots of $z_{n+1} = g(z_n)$ for the case $g'(1) \leq 1$ (left) and $g'(1) > 1$ (right).

If $g'(1) > 1$, then $\lim_{n\to\infty} z_n = z_\infty$, where z_∞ is the unique root in $(0, 1)$ of the equation

$$z = g(z) \tag{9.4.6}$$

So in that case the population goes extinct with probability z_∞ and with the complementary probability $1 - z_\infty$ it persists indefinitely.

In the epidemic context we interpret this as follows: with probability z_∞ the introduction of one case leads to a minor outbreak, where "minor" means that the total number of cases is, expressed as a fraction of the very large total population of hosts, negligible; with probability $1 - z_\infty$ the introduction of one case leads to a major outbreak, where "major" means that a non-negligible fraction of the very large population of hosts becomes infected (such that our assumption that feedback can be ignored becomes untenable).

We end this section with the crucial observation that R_0 is just the mean number of offspring from a single individual, which in terms of the probability distribution $\{q_k\}_{k=0}^{\infty}$ means

$$R_0 = \text{expected number of offspring} = \sum_{k=1}^{\infty} k q_k = g'(1). \tag{9.4.7}$$

For more in-depth treatments on the use of branching processes in biology, consult (Kimmel and Axelrod, 2002; Haccou et al., 2005).

9.5 Model ingredients (special submodels)

In Section 9.2 we introduced S^0 and $A(\tau)$ as model ingredients and in Section 9.4 we worked with $\{q_k\}$. A legitimate question is: how are such ingredients related to each other? By combining the respective relations for the basic reproduction number R_0, (9.2.2) and (9.4.7), we have as a consistency condition

$$S^0 \int_0^{\infty} A(\tau)\, d\tau = \sum_{k=1}^{\infty} k q_k \tag{9.5.1}$$

but one would like to come up with less coarse relations. In order to derive these, we need to go "down" and provide a more detailed model specification from which both $A(\tau)$ and $\{q_k\}$ can be deduced. We shall do this for a very special and very simple submodel that enjoys great popularity.

Assume that an individual can be either Susceptible or Infectious or Removed (meaning recovered and immune, or deceased). We shall use the labels S, I and R to denote these i-states. (Note that we shall soon use the same labels to denote the *number* of individuals (or the fraction of the total population) in that state.) Assume that the length of the period that an infected individual stays infectious is exponentially distributed with parameter α. Assume also that, while infectious, an individual has a constant contribution β to the force of infection. Then

$$A(\tau) = \beta e^{-\alpha\tau}. \tag{9.5.2}$$

At the i-level these assumptions correspond to

$$\begin{aligned} \frac{dp_S}{dt} &= -Fp_S, \\ \frac{dp_I}{dt} &= Fp_S - \alpha p_I, \\ \frac{dp_R}{dt} &= \alpha p_I, \end{aligned} \tag{9.5.3}$$

where $p_{\text{label}}(t)$ is the probability that an individual is in the state corresponding to the label at time t , and where F denotes the force of infection (we need p-level data to specify F). Of course, we need to provide an initial condition for (9.5.3). If we consider an individual that has just entered state I at time t, the initial condition is $p_S(t) = 0$, $p_I(t) = 1$, $p_R(t) = 0$.

If we denote the time on the clock that starts at t by τ, then

$$p_I(t + \tau) = e^{-\alpha\tau}.$$

Thus we derived/explained one of the two factors in (9.5.2).

At the p-level the assumptions correspond to

$$\begin{aligned} \frac{dS}{dt} &= -FS, \\ \frac{dI}{dt} &= FS - \alpha I, \\ \frac{dR}{dt} &= \alpha I, \end{aligned} \tag{9.5.4}$$

supplemented by

$$F = \beta I. \tag{9.5.5}$$

Exercise 9.5.1. Demonstrate that (9.1.8), (9.3.1) and (9.5.2) are equivalent to (9.5.4), (9.5.5).

In other words (and as already demonstrated by Kermack and McKendrick in 1927 (and republished in (Kermack and McKendrick, 1991a,b)), the familiar ODE formulation of the SIR-model,

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI, \\ \frac{dI}{dt} &= \beta SI - \alpha I, \\ \frac{dR}{dt} &= \alpha I, \end{aligned} \tag{9.5.6}$$

is but a very special case in a very particular guise of the general class of deterministic models that have S^0 and $A(\tau)$ as ingredients.

But let us now address the task of finding expressions for the q_k in this particular situation. To do so, we need to provide a more detailed specification of the contact process at the i-level. The obvious option is to assume that an individual makes contact according to a Poisson process with an intensity

that scales with the population size. In addition, we need to specify the probability of transmission, given a contact between an infectious and a susceptible individual. The parameter β is the product of the intensity-scaling-parameter and the probability of transmission.

The upshot is that, in the initial phase where feedback can be ignored, an individual that has an infectious period of length T generates k secondary cases with probability

$$P(k,T) = \frac{(\beta S^0 T)^k}{k!} e^{-\beta S^0 T}.$$

It follows that

$$q_k = \alpha \int_0^\infty P(k,t) e^{-\alpha t} dt = \alpha \int_0^\infty \frac{(\beta S^0 t)^k}{k!} e^{-\beta S^0 t} e^{-\alpha t}\, dt, \tag{9.5.7}$$

and, using that the generating function for a Poisson distribution with intensity parameter μ is $g(z) = e^{\mu(z-1)}$, consequently

$$g(z) = \alpha \int_0^\infty e^{\beta S^0 t(z-1) - \alpha t}\, dt = \frac{\alpha}{\alpha - \beta S^0 (z-1)}. \tag{9.5.8}$$

The equation

$$z = g(z)$$

is a quadratic equation in this case, and since we know that $z = 1$ is a solution, we do not even have to apply the general formula to find the other solution. We find, for $R_0 = \beta S^0/\alpha > 1$, as the relevant solution

$$z_\infty = \frac{\alpha}{\beta S^0} = \frac{1}{R_0}. \tag{9.5.9}$$

So for this particular submodel, the probability of a minor outbreak (upon introducing one newly infected individual) is simply the inverse of R_0 when we are in the supercritical situation $R_0 > 1$.

Side remark: See (Diekmann et al., 2012, Section 1.3.4) for a Markov chain description of an epidemic in a *finite* population, using the same submodel for infectiousness as above.

9.6 Endemicity

In order to include demography, we may change (9.5.6) into

$$\begin{aligned} \frac{dS}{dt} &= B - \mu S - \beta SI, \\ \frac{dI}{dt} &= -\mu I + \beta SI - \alpha I, \\ \frac{dR}{dt} &= -\mu R + \alpha I. \end{aligned} \tag{9.6.1}$$

Here B is a population level constant birth rate, and μ a per capita constant death rate (so the population size equals B/μ and if the variables S, I and R are actually fractions we should choose $B = \mu$). According to this description, the survival probability of an individual to at least age a equals

$$\mathcal{F}(a) = e^{-\mu a}, \tag{9.6.2}$$

which is at variance with observed survival probabilities in present-day developed countries. So let us try to find an alternative for (9.6.1) incorporating a general survival probability, i.e., a stable age distribution of the general form

$$B\mathcal{F}(a), \tag{9.6.3}$$

where $B > 0$ and $\mathcal{F}$ is a monotone non-increasing positive function with $\mathcal{F}(0) = 1$ (and for all practical purposes, compact support). We assume

- the disease leads to permanent immunity
- age has no influence on contact (i.e., the probability that two individuals have contact does not depend on their ages)

The second assumption is clearly as much at variance with known facts as an exponential survival function (see (Mossong et al., 2008)). But our aim here is to introduce some improvement in the formalism for infectious disease modelling and not to go all the way to realistic models of specific diseases.

The key point of the second assumption is that the force of infection is *not* age specific. Realising that an individual with age a at time t must have been born at time $t - a$, we write

$$S(t, a) = B\mathcal{F}(a)e^{-\int_0^a F(t-a+\sigma)\,d\sigma}, \tag{9.6.4}$$

where, as before, S refers to "susceptible". Equation (9.1.4) is replaced by

$$i(t, a) = F(t)S(t, a) \tag{9.6.5}$$

and equation (9.1.6) by

$$F(t) = \int_0^\infty \int_0^\infty i(t - \tau, a)\frac{\mathcal{F}(a + \tau)}{\mathcal{F}(a)}\,da\,A(\tau)\,d\tau. \tag{9.6.6}$$

Here, the conditional survival probability incorporates that, in principle, an individual that was infected at age a at time $t - \tau$ may not be alive anymore at time t. (Note that, when the support of A makes that only τ values that do not exceed on or two weeks are relevant, while individuals live on average for 80 years, the difference between $\mathcal{F}(a + \tau)/\mathcal{F}(a)$ and 1 is negligible except, possibly, for very high values of a.) Combining (9.6.4)–(9.6.6) we obtain a *scalar* nonlinear renewal equation

$$F(t) = B\int_0^\infty F(t - \tau)\int_0^\infty e^{-\int_0^a F(t-\tau-a+\sigma)\,d\sigma}\mathcal{F}(a + \tau)\,da\,A(\tau)\,d\tau \tag{9.6.7}$$

for the force of infection F. Clearly, $F(t) \equiv 0$ is a constant solution of (9.6.7), usually called the "disease free steady state". Linearisation around this steady state amounts to replacing $\exp(-\int_0^a F(t - \tau - a + \sigma)\,d\sigma)$ by 1. We now briefly list some observations, inviting the reader to embark upon the exercises in Section 9.9 for background, motivation, understanding.

The linearisation of the disease free steady state leads to

$$R_0 = B\int_0^\infty \int_0^\infty \mathcal{F}(a + \tau)\,da\,A(\tau)\,d\tau \tag{9.6.8}$$

and the Euler-Lotka equation

$$1 = B\int_0^\infty e^{-\lambda\tau}\int_0^\infty \mathcal{F}(a + \tau)\,da\,A(\tau)\,d\tau \tag{9.6.9}$$

Whenever $R_0 > 1$, (9.6.7) has an endemic steady state $F > 0$ characterized by the equation

$$1 = B\int_0^\infty \int_0^\infty e^{-aF}F(a + \tau)\,da\,A(\tau)\,d\tau \tag{9.6.10}$$

With a bit of effort one can linearise about the endemic steady state and next deduce the characteristic equation

$$1 = B\int_0^\infty \int_0^\infty F(a + \tau)e^{-aF}\left(1 - \frac{F}{\lambda}(1 - e^{-a\lambda})\right)\,da\,e^{-\lambda\tau}A(\tau)\,d\tau \tag{9.6.11}$$

It can be shown that for the special case $\mathcal{F}(a) = e^{-\mu a}$ all roots of (9.6.11) belong the left half plane of $\mathbb{C}$. It is unknown whether or not this also holds for general survival functions $\mathcal{F}$.

9.7 Heterogeneity

Recall the description of (9.1.7) as an expectation: there are many possibilities for how much an infected individual does contribute to the force of infection and all of these are incorporated by averaging. In Chapter 2 of (Diekmann et al., 2012), called "Heterogeneity: the art of averaging", it is explained that differences in susceptibility are less easily handled. Part II of that book provides the appropriate methodology. Here we restrict to an example.

We use a label i to distinguish individuals from each other according to relevant properties (for concreteness, think of i as an indicator of sexual activity in the context of an STD (Sexually Transmitted Disease) in a homosexual population; so note that i specifies the "type" of an individual, and not the individual itself; in particular, we work with a deterministic model and the subpopulation of individuals of type i should be large for all i). Let i range from 1 to m. We consider a closed population, as in Section 9.3. We replace (9.3.1) by

$$\frac{dS_i}{dt}(t) = -F_i(t)S_i(t) \tag{9.7.1}$$

and

$$F_i(t) = \sum_{j=1}^{m} c_{ij} \int_0^\infty F_j(t-\tau)S_j(t-\tau)A_j(\tau)\,d\tau, \tag{9.7.2}$$

where we interpret c_{ij} as the number of contacts per unit time in an ordered pair consisting of an i-individual and a j-individual. So whenever contacts are symmetric, the order is irrelevant and we should have $c_{ij} = c_{ji}$ (note that blood-transfusion provides an example of asymmetric contact). By integrating (9.7.2), using also (9.7.1), we obtain

$$\int_{-\infty}^{t} F_i(\sigma)\,d\sigma = \sum_{j=1}^{m} c_{ij}S_j^0 \int_0^\infty \left[1 - e^{-\int_{-\infty}^{t} F_j(\sigma)\,d\sigma}\right] A_j(\tau)\,d\tau, \tag{9.7.3}$$

as the analogue of (9.3.4).

Now we define the $m \times m$ matrix K by

$$K_{ij} = c_{ij}S_j^0 \int_0^\infty A_j(\tau)\,d\tau, \tag{9.7.4}$$

and next the basic reproduction number by

$$R_0 = \text{ spectral radius of } K, \tag{9.7.5}$$

and finally the Malthusian parameter r as a real root of the characteristic equation

$$\text{spectral radius}\left(c_{ij}S_j^0 \int_0^\infty e^{-\lambda\tau}A_j(\tau)\,d\tau\right) = 1.$$

EXERCISE 9.7.1. Whenever $K_{ij} = a_i b_j$ we speak about "separable mixing". Analyse this very special case, with due attention for the interpretation.

9.8 Spatial epidemic spread

The general idea is the replace i by a continuous variable x and c_{ij} by a function of two variables x and ξ that is in fact only a function of the distance $|x-\xi|$. One obtains greater generality by incorporating this dependence on two spatial positions in the kernel A. If we think of a fungal disease spreading via dispersal of spores in an agricultural crop, we may describe the introductory phase by the linear equation

$$F(t,x) = S^0 \int_0^\infty \int_{-\infty}^\infty F(t-\tau,\xi)A(\tau,x-\xi)\,d\xi d\tau \tag{9.8.1}$$

when the density of susceptible plants S^0 is uniform, i.e., independent of the spatial position. In the spirit of Section 6.5 we can now look for travelling wave solutions of the form

$$F(t,x) = e^{-\lambda(x-ct)} \tag{9.8.2}$$

In order for (9.8.2) to provide a solution of (9.8.1) we should have that

$$L(c,\lambda) = 1 \tag{9.8.3}$$

where

$$L(c,\lambda) = S^0 \int_0^\infty \int_{-\infty}^\infty e^{-\lambda(c\tau-\xi)}A(\tau,\xi)\,d\xi d\tau \tag{9.8.4}$$

Under appropriate conditions on A one can define c_0 as the minimal value of c for which (9.8.3) has a real solution λ. With considerable effort one can subsequently prove that c_0 is the asymptotic speed of propagation, see (Radcliffe and Rass, 2003). Also see (Metz et al., 2000).

9.9 Infectious diseases with demographic turnover taken into account

We consider the following variables

- t time
- S number of susceptibles
- i incidence, i.e., the number of new cases per unit of time
- F force of infection, i.e., the probability per unit time that a susceptible is infected
- B population birth rate
- μ per capita death rate

We assume that the infection leads to permanent immunity and that the infection does *not* lead to a higher death rate. These assumptions lead to the following equation

$$\frac{dS}{dt}(t) = B - \mu S(t) - i(t), \tag{9.9.1}$$

$$i(t) = F(t)S(t), \tag{9.9.2}$$

$$F(t) = \int_0^\infty i(t-\tau)A(\tau)e^{-\mu\tau}d\tau. \tag{9.9.3}$$

Exercise 9.9.1. Interpret equation (9.9.3). In particular, describe in words the meaning of $A(\tau)$.

Exercise 9.9.2. Show that the definition

$$S(t) = B\int_0^\infty e^{-\mu a - \int_0^a F(t-a+\sigma)d\sigma}da \tag{9.9.4}$$

is compatible with (9.9.1)–(9.9.2), i.e., show that S defined by (9.9.4) satisfies (9.9.1) with $i(t)$ defined by (9.9.2). Hint: rewrite (9.9.4) as

$$S(t) = B \int_{-\infty}^{t} e^{-\mu(t-\eta)-\int_{\eta}^{t} F(\sigma)d\sigma} d\eta.$$

EXERCISE 9.9.3. Describe the meaning of (9.9.4) when a is interpreted as the age of an individual.

EXERCISE 9.9.4. By substituting (9.9.4) into (9.9.2) and next (9.9.2) into (9.9.3), we obtain the nonlinear renewal equation

$$F(t) = \int_{0}^{\infty} F(t-\tau) B \int_{0}^{\infty} e^{-\mu a-\int_{0}^{a} F(t-\tau-a+\sigma)d\sigma} da A(\tau) e^{-\mu\tau} d\tau. \tag{9.9.5}$$

Determine the steady states of this equation. Hint: do not try to take a time derivative to set to zero.

EXERCISE 9.9.5. What is the condition for the (in)stability of the trivial steady state $F = 0$? What is the ecological interpretation of this condition? How does this condition relate to the (non-)existence of a biologically meaningful nontrivial steady state?

EXERCISE 9.9.6. Linearise the renewal equation (9.9.5) around the nontrivial steady state $\bar{F}$.

EXERCISE 9.9.7. Derive the corresponding characteristic equation

$$1 = \frac{\lambda+\mu}{\lambda + B\bar{A}(\mu)} \frac{\bar{A}(\lambda+\mu)}{\bar{A}(\mu)}, \tag{9.9.6}$$

where

$$\bar{A}(s) := \int_{0}^{\infty} e^{-st} A(t) dt$$

is the Laplace transform of $A(t)$.

EXERCISE 9.9.8. Assume that $B\bar{A}(\mu) > \mu$ (recall Exercise 9.9.5 to interpret this condition). Show that (9.9.6) cannot have a root with $\Re\lambda \geq 0$. Hint: take absolute values at both sides of (9.9.6).

Corollary 9.9.1*:* The endemic steady state is locally asymptotically stable whenever it exists.

EXERCISE 9.9.9. As an encore, consider the special case

$$A(\tau) = \beta e^{-\alpha\tau}.$$

Verify that in that case (9.9.2), (9.9.3) can be replaced by the familiar ODE

$$\frac{dI}{dt} = \beta IS - (\alpha+\mu)I, \tag{9.9.7}$$

where $I = \frac{1}{\beta}F$ represents the number of infectious individuals.

Chapter 10

Adaptive Dynamics

10.1 Introduction

How does natural selection shape the life history of a species? A popular, but naive, view is that survival of the fittest leads to optimal adaptation to the environment. The complication is that often part of the environmental conditions (in particular food availability and predation pressure) are, in turn, influenced by the population itself. And if indeed the ecological feedback loop mediates the selection, we cannot pretend that the environment is unaffected while evolution takes its course.

Adaptive Dynamics is a theory/approach that focuses on two processes:

- ecological interaction via environmental variables
- mutation

While doing so, it ignores that phenotypes are determined by genotypes and that sexual reproduction shakes up genotypes. In other words: no genes and no sex. The phenotypes are characterised by "traits" (which are sometimes also called "strategies") and in principle clonal reproduction leads to offspring with exactly the same trait. But occasionally offspring may carry a slightly mutated trait, i.e., variation is created by rare and small mutations. Can we predict how the interplay of selection and mutation leads to the evolutionary dynamics of the trait?

The aim of this chapter is

- to explain concepts, like unbeatable strategy (often called ESS, for Evolutionarily Stable Strategy), invasion exponent, selection gradient and trait substitution sequence
- to describe results, like the pessimization principle, the principle of indifference and the classification of singular points (introducing in particular branching points, where populations turn dimorphic)

We shall do so by way of the example of consumer-resource dynamics. In fact, we first consider competition for one resource and then continue to investigate how the repertoire of possibilities increases when the environmental condition (from the point of view of the consumer) is two-dimensional, simply since there are two (substitutable) resources. The trait concerns the up-take rate. When dealing with one resource, we shall introduce a trade-off that relates the up-take rate to the conversion efficiency. When dealing with two resources, the trait determines the relative up-take rate and hence we can analyse the (dis)advantages of being a specialist or a generalist.

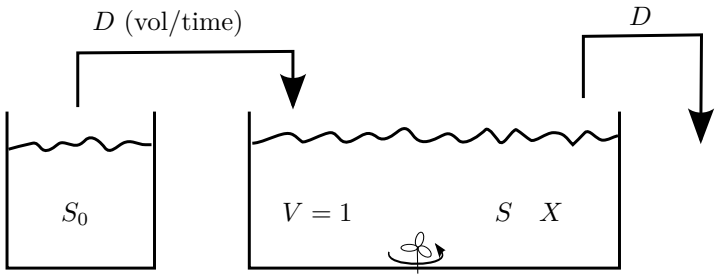


FIGURE 10.1. The chemostat

10.2 The pessimization principle

The chemostat (see Figure 10.1) is a laboratory device to culture micro-organisms like algae or bacteria. In a vessel of volume $V = 1$ all they need to grow is provided in excess, except for one substance (e.g., phosphate) which accordingly is called the limiting substrate (or, resource). We refer to Smith and Waltman (1995) for a systematic exposition of the insights about chemostat dynamics that can be obtained by modelling, with due attention for the mathematical methods that are used to derive these insights.

We denote by S and X the concentrations of, respectively, the substrate and the micro-organisms. The substrate in the fresh medium is denoted by S_0 and is assumed to be a given constant. By choosing the time unit appropriately, we can achieve that the rate D, at which fluid is pumped in and out of the vessel, equals 1. As the vessel is continuously stirred, the concentrations in the outflowing medium are exactly what they are inside the vessel.

We assume that up-take of substrate by the micro-organisms is governed by the Law of Mass Action. The proportionality constant we call u and it this u that we consider as the trait. In other words, our aim is to understand the evolutionary dynamics of u due to the combined effect of selection and mutation. The idea is that by fine-tuning their biochemical machinery, the micro-organisms may improve their efficiency for binding the substrate molecules to the cell surface and next swallow them. But of course this may come at a physiological cost. Let η denote the conversion efficiency, i.e., to produce one unit of micro-organism η^{-1} units of substrate are needed. Then one can imagine that to build chemical pathways that increase u, the micro-organism needs to sacrifice chemical pathways involved in conversion, so that η decreases. At first we shall neglect such a constraint, but later on we will investigate the trade-off that it generates.

The consumer-resource dynamics are described by

$$\begin{aligned} \frac{dS}{dt} &= S_0 \quad -S - uSX \\ \frac{dX}{dt} &= \quad\; -X + \eta uSX = (-1 + \eta uS)X \end{aligned} \tag{10.2.1}$$

We require that $\eta u S_0 > 1$, which amounts to assuming that the growth rate $-1+\eta uS$ is positive when micro-organisms are introduced in the virgin environment characterised by $S = S_0$. By consumption the growing population of micro-organisms reduces the substrate concentration to the steady state level

$$\bar{S} = \frac{1}{\eta u} \tag{10.2.2}$$

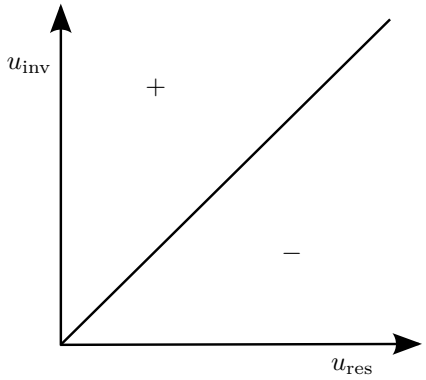


FIGURE 10.2. PIP showing the sign of the invasion exponent. Note that there is necessarily neutrality at the diagonal.

EXERCISE 10.2.1. Show that (10.2.1) has a globally stable steady state. Hint: try to reduce the dimension by a conservation law for S_{total}, the free substrate plus the substrate incorporated in X.

Note that, of course, the growth rate $-1+\eta u S$ equals zero in the steady state.

Now imagine that in this steady situation, by mutation, a micro-organism arises that has a slightly different up-take coefficient, i.e., a slightly different trait. Of course, it may be unlucky and be washed out right away, in which case nothing happens. In other words, strictly speaking we should deal with the demographic stochasticity pertinent to low numbers and, for instance, adopt a description in terms of branching processes. But if we only want to know whether the mutant has any chance at all of surviving in an environment set by the resident, we might as well adopt a deterministic description and pretend that the mutant constitutes a small fraction of the population of micro-organisms. So we want to know whether it has a positive growth rate.

As there are now two types of micro-organisms, we need labels to distinguish them from one another. We shall use u_{res} to denote the trait of the resident and u_{inv} to denote the trait of the mutant, where inv stands for invader. The *invasion exponent* is, by definition, the population growth rate of the invader in the environmental conditions as set by the resident. So in this particular case,

$$\text{invasion exponent} = -1 + \eta u_{\text{inv}} \frac{1}{\eta u_{\text{res}}} = -1 + \frac{u_{\text{inv}}}{u_{\text{res}}}$$

(More generally, it makes sense to define "fitness" as the long term population growth rate as a function of two variables, the trait and the environmental condition, see Metz et al. (1992).)

Clearly the invasion exponent is positive if $u_{\text{inv}} > u_{\text{res}}$, exactly as common sense predicts. We can embody this graphically in a PIP, a Pairwise Invasibility Plot, see Figure 10.2. Next we observe that for this trait and this ecological interaction, there is never mutual invasibility: if u_2 can invade successfully (meaning that its growth rate is positive) in the environment set by u_1, then u_1 *cannot* invade in the environment set by u_2. It is a folk theorem that this implies that a successful invader outcompetes the resident, i.e., drives it to extinction (and by doing so becomes the new resident). We then say that a *trait substitution* took place. For small mutation the folk theorem has been verified, see Geritz (2005).

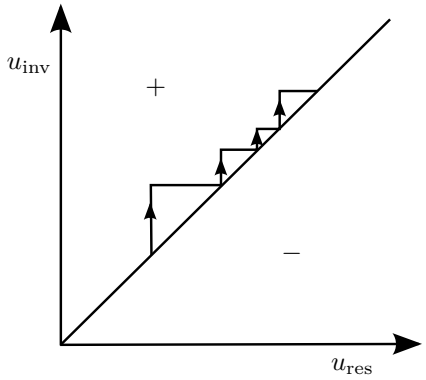


FIGURE 10.3. A trait substitution sequence: note that the height of the jumps is a stochastic facet of mutation, but that the direction of evolutionary change (i.e. the fact that u increases) is a consequence of natural selection.

To verify the folk theorem for the present situation, we need to study the three-dimensional system

$$\begin{aligned} \frac{dS}{dt} &= S_0 - S - u_{\text{res}}SX \quad - u_{\text{inv}}SY \\ \frac{dX}{dt} &= \quad -X + \eta u_{\text{res}}SX \\ \frac{dY}{dt} &= \quad -Y \qquad\qquad + \eta u_{\text{inv}}SY \end{aligned} \tag{10.2.3}$$

EXERCISE 10.2.2. Use reduction to a two-dimensional system and next Poincaré-Bendixson theory (see Appendix) to prove the folk theorem for system (10.2.3).

Note that in our analysis we exclude the possibility of a new mutation during the period of ecological interaction between resident and invader. That is, we assume *time scale separation* in the sense that the time scale of ecological interaction is very short relative to the time scale at which mutation yields candidates for evolutionary change. The upshot is that the combined effect of mutation and selection manifests itself as a *trait substitution sequence*, see Figure 10.3.

If we write

$$\textit{invasion exponent} = s_{u_{\text{res}}}(u_{\text{inv}})$$

we can determine the direction in which *small* mutational steps and selection will drive the trait, by computing the *selection gradient*

$$\frac{\partial}{\partial v} s_u(v)\Big|_{v=u}$$

In the present case the selection gradient equals $\frac{1}{u}$ and, in particular, it is positive for every u. We conclude once more that, given our assumptions, evolution leads to an ever increasing up-take rate u.

Realizing that nothing comes for free, we change the assumptions.

Indeed, let us assume that η is a linearly decreasing function of u. More precisely, we assume that all variables are already dimensionless and that

$$\eta = \eta(u) = 1 - u \tag{10.2.4}$$

and that $S_0 > 4$. The function $u \mapsto \eta(u)u$ assumes its maximum $\frac{1}{4}$ for $u = \frac{1}{2}$. There exists an interval $(u_{\min}, u_{\max})$ such that $\eta(u)uS_0 > 1$ for u in this interval

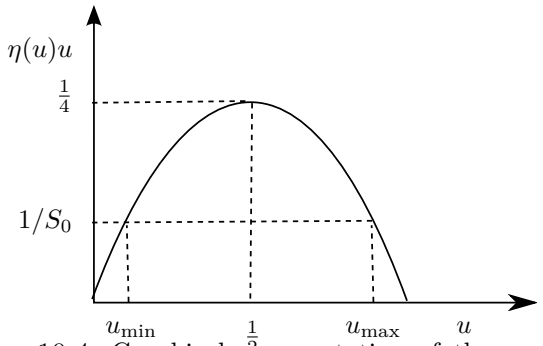


FIGURE 10.4. Graphical representation of the way to find the interval where $\eta(u)uS_0 > 1$.

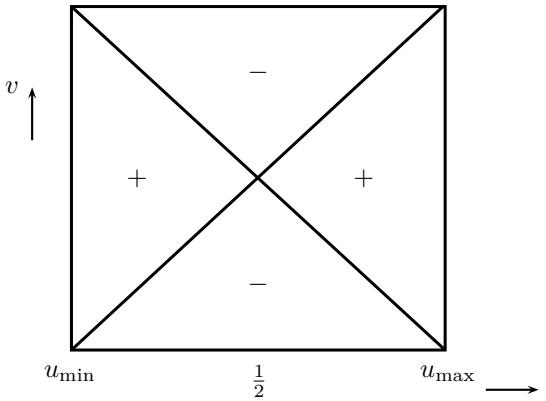


FIGURE 10.5. The PIP when η is given by (10.2.4).

(see Figure 10.4). The invasion exponent is given by

$$s_u(v) = -1 + \frac{(1-v)v}{(1-u)u} = \frac{(v-u)(1-v-u)}{(1-u)u} \tag{10.2.5}$$

and we see that, in addition to the diagonal neutrality curve, we have a neutrality curve $v = 1-u$ (Figure 10.5). The two neutrality curves intersect at the *singular point* $u = \frac{1}{2}$ where the selection gradient $\frac{1-2u}{(1-u)u}$ vanishes. The PIP (see Figure 10.5) clearly shows that small mutations lead to increasing u when $u_{\text{res}} < \frac{1}{2}$ but to decreasing u when $u_{\text{res}} > \frac{1}{2}$. In other words, evolutionary change brings u_{res} ever closer to $\frac{1}{2}$. One says that the singular point $u = \frac{1}{2}$ is *convergence stable* . Moreover, the vertical line through $u = \frac{1}{2}$ lies entirely in the $-$ region. So $u = \frac{1}{2}$ is *unbeatable* or, as it is often called, *evolutionarily stable* (the word "stable" is actually a misnomer; it would be better to replace it by "steady", but the standard meaning of ESS is probably itself evolutionarily stable in the sense that attempts to change it are doomed to fail). A singular point that is both ESS and convergence stable is often called a CSS for Continuously Stable Strategy. It is a (local) *attractor* with respect to the adaptive dynamics.

Now recall 10.2.2: the steady state substrate level is given by

$$\bar{S} = \frac{1}{\eta u}$$

Since $\eta(u)u$ is maximal for $u = \frac{1}{2}$, we conclude that $\bar{S}$ is minimal for $u = \frac{1}{2}$. This is the *pessimization principle*: if the population growth rate is a monotone increasing

function of a one-dimensional variable that fully describes the environmental condition, then the evolutionary winner is the trait that minimises this variable.

EXERCISE 10.2.3. Suppose that a population of individuals with trait x grows under constant environmental conditions characterised by a variable I like

$$e^{r(x,I)t}.$$

Assume that both x and I are real numbers. Assume that r is monotone in I and that the equation $r(x, I) = 0$ has, for given x, a unique solution $\bar{I}(x)$. Assume that population dynamics always leads to steady state and that the folk theorem applies. Show that adaptive dynamics leads to a trait substitution sequence for which $\bar{I}(x)$ is monotone. Show that if the sequence converges, then $\bar{I}$ necessarily has an extremum at the limit and interpret this result in terms of local evolutionary stability (i.e., unbeatability).

We conclude that adaptive dynamics is rather predictable if the ecological feedback loop that drives the selection just involves a one-dimensional environmental condition. Will the repertoire become richer if we change to a higher dimensional environmental condition?

10.3 The principle of indifference

Assume that the fresh medium contains not one, but two resources, in concentrations S_{10} and S_{20} respectively. We assume that these are substitutable, meaning that they both provide the one substance that is not available in excess. We also assume (in order to make the algebra as simple as possible) that the conversion efficiency η for both equals one. So we consider the system

$$\begin{aligned}\frac{dS_1}{dt} &= S_{10} - S_1 - u_1 S_1 X \\ \frac{dS_2}{dt} &= S_{20} - S_2 - u_2 S_2 X \\ \frac{dX}{dt} &= -X + u_1 S_1 X + u_2 S_2 X = (-1 + u_1 S_1 + u_2 S_2) X\end{aligned} \tag{10.3.1}$$

Provided $u_1 S_{10} + u_2 S_{20} > 1$ there exists a unique steady state. It is found by first choosing $\bar{X}$ as the unique solution of

$$-1 + \frac{u_1 S_{10}}{1 + u_1 X} + \frac{u_2 S_{20}}{1 + u_2 X} = 0 \tag{10.3.2}$$

and next defining $\bar{S}_1$ and $\bar{S}_2$ by the explicit formulas

$$\bar{S}_1 = \frac{S_{10}}{1 + u_1 \bar{X}}, \quad \bar{S}_2 = \frac{S_{20}}{1 + u_2 \bar{X}} \tag{10.3.3}$$

EXERCISE 10.3.1. Show that this steady state is globally asymptotically stable. Hint: first show that $S_1 + S_2 + X \to S_{10} + S_{20}$ for $t \to \infty$, next perform a standard phase plane analysis of the two dimensional system obtained by putting $X = S_{10} + S_{20} - S_1 - S_2$ in the equations for S_1 and S_2.

Now that we understand the population dynamics, we can superimpose the adaptive dynamics. We introduce a one-dimensional trait x and assume that both u_1 and u_2 depend on x. More specifically, we assume that u_1 is an increasing function of x and u_2 a decreasing function. The idea is the same as before: the micro-organism can improve the chemical pathways involved in uptake of S_1, but only at the expense of the chemical pathways involved in uptake of S_2.

The invasion exponent is now given by

$$s_x(y) = -1 + u_1(y)\bar{S}_1(x) + u_2(y)\bar{S}_2(x) \tag{10.3.4}$$

and consequently,

$$\text{selection gradient } = u_1'(x)\bar{S}_1(x) + u_2'(x)\bar{S}_2(x) \tag{10.3.5}$$

Singular points are, by definition, characterised by a vanishing selection gradient. The condition

$$u_1'(x)\bar{S}_1(x) + u_2'(x)\bar{S}_2(x) = 0 \tag{10.3.6}$$

can be interpreted as follows: an infinitesimal change in x leads to an infinitesimal gain or loss in uptake of S_1 and to an infinitesimal loss or gain in uptake of S_2, and singular points are exactly those for which the two cancel, i.e., for which there is no net change.

This is the marginal value theorem of economic optimisation under constraints and it is the mechanism behind the ideal free distribution in population ecology. We like to call it the Principle of Indifference: a population of individuals tunes the environmental conditions such that the various options that are open become equally attractive, so that infinitesimal re-allocations don't make any difference.

Let us consider the special case that $S_{10} > 1$, $S_{20} > 1$ and

$$u_1(x) = x, \quad u_2(x) = 1 - x \tag{10.3.7}$$

in some detail. In this case (10.3.6) reduces to

$$\bar{S}_1 = \bar{S}_2. \tag{10.3.8}$$

Let us denote the common value by $\bar{S}$. Since in steady state we should have (note that $u_1 + u_2 = 1$ for all x!)

$$0 = -1 + u_1\bar{S} + u_2\bar{S} = -1 + \bar{S}$$

it follows that necessarily $\bar{S} = 1$. Next, if we solve for $\bar{X}$ in both identities in (10.3.3) we find

$$\frac{S_{10} - \bar{S}}{u_1\bar{S}} = \frac{S_{20} - \bar{S}}{u_2\bar{S}}$$

which, using (10.3.7) and $\bar{S} = 1$, after a few manipulations leads to the expression

$$x_{\text{singular}} = \frac{S_{10} - 1}{S_{10} + S_{20} - 2} \tag{10.3.9}$$

Much more important than this explicit expression, however, is the following observation. From (10.3.4), (10.3.8) and (10.3.7) we see that

$$s_{x_{\text{singular}}}(y) = 0 \quad \text{for all } y! \tag{10.3.10}$$

or, in words, the second neutrality curve, that intersects the diagonal at the singular point, is a vertical line (see Figure 10.6). Clearly small mutation will bring the resident trait to a small neighbourhood of the singular point. But once we are very near, an overshoot may occur. The key point of the overshoot is that, in the present situation, it brings us in a region of *mutual invasibility* (implying that successful invasion is *not* followed by extinction of the former resident—instead we find *coexistence*: the monomorphism is replaced by a dimorphism).

Exercise 10.3.2. Check that a dimorphic population with traits at opposite sides of x_{singular} sets the environmental condition at $\bar{S}_1 = \bar{S}_2 = 1$, after which there is complete neutrality.

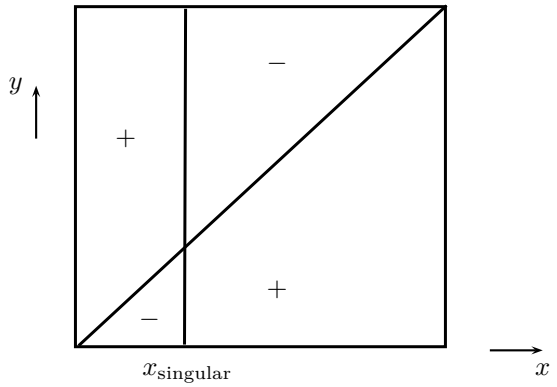
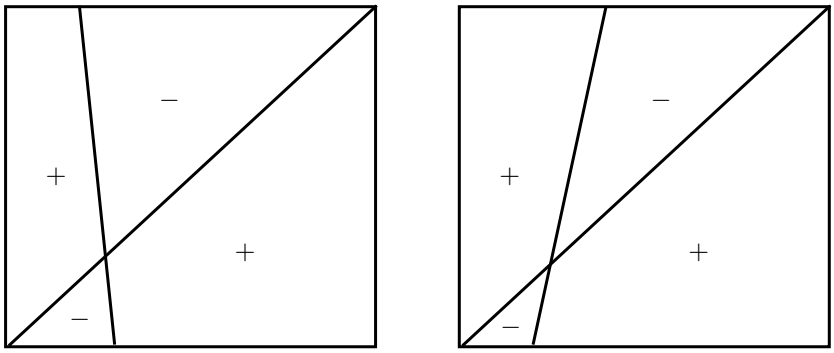


FIGURE 10.6. The PIP for the special case (10.3.7).



Tilted to the left:
convergence stable and ESS

Tilted to the right:
convergence stable but *not* ESS

FIGURE 10.7. The two generic perturbations of a vertical neutrality curve (with the key local features extended globally, eventhough that need not be the case).

The key point about (10.3.7) is that it makes the invasion exponent $s_x(y)$ linear in y, such that the local indifference condition that characterises the singular point has a global effect. Indeed, the vertical neutrality curve is a very strong manifestation of the Principle of Indifference and as such it is not robust (the local characterisation of the singular point persists under perturbation, but the global features of the second neutrality curve do not). We will now change (10.3.7) and investigate the effects. But before doing any algebra, we consider the geometry. From Figure 10.7 it appears that under perturbation the singular point may or may not turn into a (local) ESS. In the first case, overshoots may, by mutual invasibility, still lead to dimorphisms but one can show that these are *converging*, meaning that a mutant in between the two resident traits can successfully invade, resulting in extinction of one of the two resident traits. So after the trait substitution the two resident traits are closer together, whence the name *converging dimorphism*. In the second case, a *diverging dimorphism* arises. So then the dimorphic population structure is lasting (maybe not everlasting, as later on (i.e., far away from the singular point) other things may happen) rather than transient.

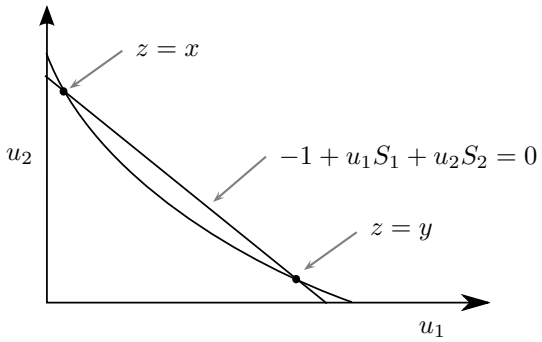


FIGURE 10.8. Plot of the curve $z \mapsto (u_1(z), u_2(z))$ and of the straight line $-1 + u_1S_1 + u_2S_2 = 0$. Intersections correspond to the dimorphism. The invasibility condition $s_{x,y}(z) > 0$ means that the point on the curve corresponding to trait z should be above the straight line. So in the depicted situation we have a diverging dimorphism.

So let us look at dimorphic populations. These are described by the system

$$\begin{aligned}
\frac{dS_1}{dt} &= S_{10} - S_1 - u_1(x)S_1X_1 - u_1(y)S_1X_2 \\
\frac{dS_2}{dt} &= S_{20} - S_2 - u_2(x)S_2X_1 - u_2(y)S_2X_2 \\
\frac{dX_1}{dt} &= \quad -X_1 + u_1(x)S_1X_1 + u_2(x)S_2X_1 \\
\frac{dX_2}{dt} &= \quad -X_2 + u_1(y)S_1X_2 + u_2(y)S_2X_2
\end{aligned} \tag{10.3.11}$$

EXERCISE 10.3.3. Show that the corresponding steady states are given by

$$\begin{pmatrix} S_1 \\ S_2 \end{pmatrix} = \frac{1}{u_1(x)u_2(y) - u_1(y)u_2(x)} \begin{pmatrix} u_2(y) - u_2(x) \\ u_1(x) - u_1(y) \end{pmatrix} \tag{10.3.12}$$

and

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \frac{u_2(y)S_{10}}{u_2(y)-u_2(x)} - \frac{u_1(y)S_{20}}{u_1(x)-u_1(y)} - \frac{u_1(y)-u_2(y)}{u_1(x)u_2(y)-u_1(y)u_2(x)} \\ \frac{-u_2(x)S_{10}}{u_2(y)-u_2(x)} + \frac{u_1(x)S_{20}}{u_1(x)-u_1(y)} + \frac{u_2(x)-u_1(x)}{u_1(x)u_2(y)-u_1(y)u_2(x)} \end{pmatrix} \tag{10.3.13}$$

Show that whenever the expressions at the right hand side of (10.3.13) are positive (as they should be in order to be meaningful) then $0 < S_i < S_{i0}$ for $i = 1, 2$ with S_i defined by (10.3.12). The global stability of this steady state is established in Ballyk et al. (2005).

The invasion exponent in such a dimorphism is given by

$$s_{x,y}(z) = -1 + u_1(z)S_1 + u_2(z)S_2 \tag{10.3.14}$$

with S_i given by (10.3.12). Choosing rather general functions $u_1(z)$ and $u_2(z)$ once more, the situation of a diverging dimorphism is depicted in Figure 10.8.

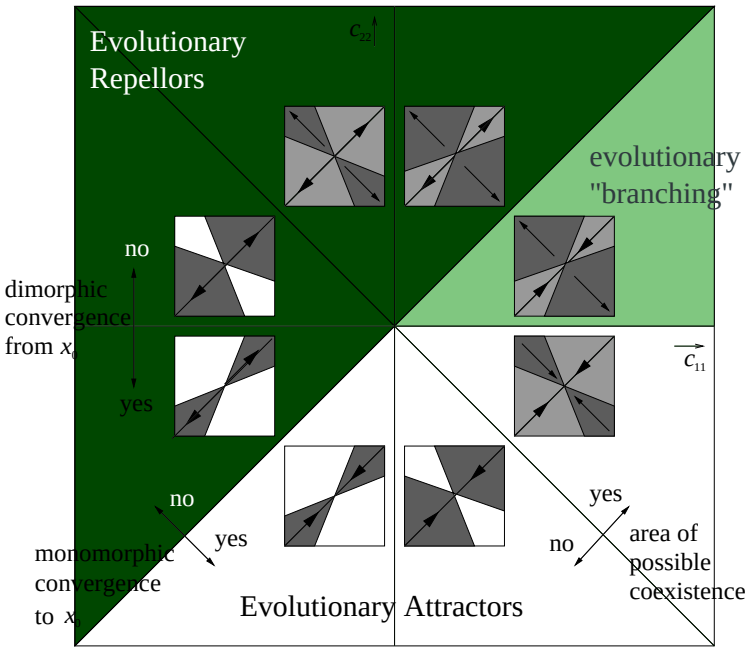


FIGURE 10.9. Classification of singular points.

Chapter 11

Stochastic approaches to dynamical problems

This book deals mostly with deterministic approaches to biological situations. This is partly our own shortcoming, since stochastic models are also found in abundance in the literature. The reasons for a probabilistic look at dynamics are several. First of all, one can ask new questions that determinstic models cannot answer. For example, in epidemiology, one of the main questions one is interested in is the probability that an outbreak of a disease will occur. Such a question is beyond the analyses that differential equations of continuous variables have to offer. A second good reason to study probabilistic models is that these models give a host of new variables to track and measure. Examples here are the mean, variance and higher moments of the probability density function of the process under consideration. The first of these may have a direct counterpart in the deterministic description (and may indeed be identical), but the time evolution of the variance is also often interesting in applications. We might also be interested in the first time a particular event occurs.

When setting up a stochastic model, the first thing one might try to is to discretize all the independent variables, vz. time *and* state space. This gets you into the realm of discrete Markov chains, a topic covered in many courses in probability and not (yet) covered here. Instead, we discretize only state space, and keep time continuous, which takes us into the field of *Master Equations*.

If we discrete neither time nor space, one can still develop stochastic models, using stochastic differential equations. This is technically by far the hardest, and requires a separate master's course. It is therefore left out here.

11.1 Master Equations

Let us consider a random variable which takes values in $\mathbb{Z}$. In many instances, these values are naturally restricted to a subset, e.g., $\mathbb{N}$ or a finite interval within $\mathbb{N}$, but we leave such restricted processes for later.

We assume that in small enough time intervals of length Δt, transitions may occur so that the random variable changes state from n, say, to either $n-1$ or $n+1$. In other words, we only consider state transitions between neighbouring states. (The field of Master Equations also deals with the more general case when arbitrary state transitions occur, but we restrict ourselves to the case above, which applies naturally in many biological settings.) Let the probability that a transition $n \to n+1$ occurs within a time interval of length Δt be given by $\lambda_n \Delta t + \mathcal{O}((\Delta t)^2)$, and let analogously the transition $n \to n-1$ occur with probability $\mu_n \Delta t + \mathcal{O}((\Delta t)^2)$.

Note that these functions λ_n and μ_n may indeed depend on n, and often do as we will see. The point is that the probability for an event to occur depends on the time span which one considers, and that the above probability only make sense for *small* Δt.

Let $P_{n,m}(t)$ be the probability that the random variable is in state n at time t when initially starting at state m at time $t = 0$. Clearly, $P_{n,m}(t)$ satisfies the following equation up to order $(\Delta t)^2$,

$$P_{n,m}(t + \Delta t) = \lambda_{n-1}\Delta t P_{n-1,m}(t) + \mu_{n+1}\Delta t P_{n+1,m}(t) + (1 - (\lambda_n + \mu_n))\Delta t P_{n,m}(t). \tag{11.1.1}$$

It considers the fate of the variable each time starting out at the same position and arriving at state n at time t and is therefore called a *forward equation*. (There also exists a backward equation, but it is used less often.) Taking the limit $\Delta t \to 0$, we find the forward master equation for $P_{n,m}(t)$,

$$\frac{d}{dt}P_{n,m}(t) = \lambda_{n-1}P_{n-1,m}(t) + \mu_{n+1}P_{n+1,m}(t) - (\lambda_n + \mu_n)P_{n,m}(t), \tag{11.1.2}$$

$$\tag{11.1.3}$$

If we first solve this equation for $P_{n,m}(t)$ with initial conditions

$$P_{n,m}(0) = \delta_{n,m},$$

then the complete solution satisfying the general initial condition

$$P_{n,m}(0) = p(m),$$

is given by superposition

$$P_n(t) = \sum_m P_{n,m}(t)p(m).$$

In general, it is too much to ask for a full description of $P_{n,m}(t)$. Instead, the first few moments of $P_{n,m}(t)$ are enough to get a good feel for the behaviour of solutions. So let us introduce the moments of $P_{n,m}(t)$ by

$$\begin{aligned}
\langle n \rangle &:= \sum_n nP_{n,m}(t), \quad \text{the } \textit{mean} \text{ of } P_{n,m}(t),\\
\langle n^2 \rangle &:= \sum_n n^2 P_{n,m}(t),\\
&\cdots\\
\langle n^k \rangle &:= \sum_n n^k P_{n,m}(t).
\end{aligned}$$

The *variance* of $P_{n,m}(t)$ is then given by $\langle n^2 \rangle - \langle n \rangle^2$. These moments satisfy their own differential equations, which may be derived from the forward (and backward) master equation by multiplying by n^k and summing over n. For instance, the mean $\langle n \rangle$ satisfies the equation

$$\begin{aligned}
\frac{d}{dn}\langle n \rangle = \frac{d}{dt}\sum_n nP_{n,m}(t) = \Big(\sum_n n\lambda_{n-1}P_{n-1,m} - \sum_n n\lambda_n P_{n,m}\Big)\\
+ \Big(\sum_n n\mu_{n+1}P_{n+1,m} - \sum_n n\mu_n P_{n,m}\Big).
\end{aligned} \tag{11.1.4}$$

Note that

$$\sum n\lambda_{n-1}P_{n-1,m} = \sum (n+1)\lambda_n P_{n,m},$$

so that the first term on the right hand side of (11.1.4) becomes $\sum \lambda_n P_{n,m}$. Similarly, the second term reduces to $\sum \mu_n P_{n,m}$. We thus find that

$$\frac{d}{dt}\langle n\rangle = \sum \lambda_n P_{n,m} - \sum \mu_n P_{n,m} =: \langle \lambda_n\rangle - \langle \mu_n\rangle .$$

Along the same lines, the variance $\mathrm{var}(n)$ satisfies

$$\frac{d}{dt}\,\mathrm{var}(n) = 2(\langle n\lambda_n\rangle - \langle n\mu_n\rangle) + (1-2\langle n\rangle)\langle \lambda_n\rangle + (1+2\langle n\rangle)\langle \mu_n\rangle .$$

Initial conditions $P_{n,m}(0) = \delta_{n,m}$ carry over to initial conditions for the moments,

$$\langle n^k\rangle(0) = m^k, \quad k = 1, 2, \ldots,$$

and in particular, $\mathrm{var}(n)(0) = 0$. These equations for the moments are not directly solvable without an explicit description of λ_n and μ_n. If λ_n and μ_n are *linear* functions of n (or constants), then the equation for $\langle n^k\rangle$ only depends on $\langle n^k\rangle$ and lower moments, making this a closed system of equations. E.g., if we consider

$$\lambda_n = an, \quad \mu_n = bn,$$

for real constants a and b, then the differential equation for the mean $\langle n\rangle$ becomes simply

$$\frac{d}{dt}\langle n\rangle = (a-b)\langle n\rangle ,$$

so that $\langle n\rangle = me^{(a-b)t}$, just as in the deterministic case.

If, however, λ_n and μ_n are polynomials of degree higher than 1, then the equation for $\langle n^k\rangle$ depends on higher moments, $\langle n^{k+1}\rangle$, for each k, and we would never be able to close this without making extra assumptions. To make headway in such situations, *moment closure techniques* are invoked which prescribe how higher order moments may be written in terms of lower order moments to close the equations. This is quite a field in itself, particularly in epidemiology.

11.2 Generating functions for master equations

Let

$$G_m(t,z) := \sum_n P_{n,m}(t) z^n$$

be the *probability generation function* for $P_{n,m}(t)$. $G_m(t,z)$ is often also called the *z-transform* of $P_{n,m}(t)$. Evidently, finding $G_m(t,z)$ is equivalent to finding the complete solution $P_{n,m}(t)$. Having found the generating function, it is easy to compute the moments of $P_{n,m}(t)$. First, introduce $y = \log z$, so that $z = e^y$, and note that

$$\frac{\partial^k}{\partial y^k} G(t,e^y) = \sum_n n^k P_{n,m}(e^y)^n.$$

Hence

$$\frac{\partial^k}{\partial y^k} G(t,e^y)\Big|_{y=0} = \frac{\partial^k}{\partial \log z^k} G(t,z)\Big|_{z=1} = \langle n^k\rangle_m(t).$$

So let us find out how to translate the master equations for $P_{n,m}(t)$ into a partial differential equation for $G_m(t,z)$. We start with the time derivative of $G_m(t,z)$,

and compute, using the original master equation (11.1.2) for $P_{n,m}(t)$,

$$\frac{\partial}{\partial t}G_m(t,z) = \sum_n z^n \frac{d}{dt}P_{n,m}(t) \tag{11.2.1}$$

$$= \sum_n \left[\lambda_{n-1}P_{n-1,m}(t) + \mu_{n+1}P_{n+1,m}(t) - (\lambda_n + \mu_n)P_{n,m}(t)\right] z^n. \tag{11.2.2}$$

Now observe that

$$\sum_n \lambda_{n-1}P_{n-1,m}(t)z^n = \sum_n \lambda_n P_{n,m}(t)z^{n+1}$$

and similarly for the $\mu_{n+1}P_{n+1,m}(t)$ sum, so that (11.2.1) simplifies to

$$\frac{\partial}{\partial t}G_m(t,z) = \sum_n \left(\lambda_n(z-1) + \mu_n\left(\frac{1}{z}-1\right)\right) P_{n,m}(t)z^n$$

$$= \left(\lambda_n(z-1) + \mu_n\left(\frac{1}{z}-1\right)\right) G_m(t,z). \tag{11.2.3}$$

This is not yet a partial differential equation for $G(t,z)$, and for general λ_n and μ_n it cannot be reduced to one. But for polynomial λ_n, μ_n it can. Note that if we introduce $f_1(n) = \lambda_n$ and $f_{-1}(n) = \mu_n$, then

$$f_1\Big(z\frac{\partial}{\partial z}\Big)z^n = f_1(n)z^n,$$

since $z\frac{\partial}{\partial z}z^n = znz^{n-1} = nz^n$. For instance, if $\lambda_n = n^2$, then

$$f_1\Big(z\frac{\partial}{\partial z}\Big)z^n = z\frac{\partial}{\partial z}\Big(z\frac{\partial}{\partial z}(z^n)\Big) = n^2z^n = f_1(n)z^n.$$

Equation (11.2.3) is now

$$\frac{\partial}{\partial t}G_m(t,z) = \Big((z-1)f_1\Big(z\frac{\partial}{\partial z}\Big) + \Big(\frac{1}{z}-1\Big)f_{-1}\Big(z\frac{\partial}{\partial z}\Big)\Big)G_m(t,z), \tag{11.2.4}$$

a PDE for $G_m(t,z)$.

11.3 Examples of known solutions

Luckily, the PDE for the generating function can be explicitly solved for a number of concrete cases that arise in applications. In each of the cases below, we assume that a process starts at position m at time 0. If we now for instance consider a Poisson process with the properties $\lambda_n = \nu \in \mathbb{R}$, $\mu_n = 0$, then

$$G(t,z) = z^m e^{\nu t(z-1)},$$

from which we compute that $\langle n\rangle = m + \nu t$, and $\mathrm{var}(n) = \nu t$. Similarly, setting $\mu_n = \rho$, $\lambda_n = 0$, we find

$$G(t,z) = z^m e^{\rho t(\frac{1}{z}-1)},$$

and $\langle n\rangle = m - \nu t$, $\mathrm{var}(n) = \rho t$. A general Poisson process with $\lambda_n = \nu$, $\mu_n = \rho$ yields

$$G(t,z) = z^m e^{-(\nu+\rho)t + (\nu z + \rho/z)t},$$

so that $\langle n\rangle = m + (\nu - \rho)t$, and $\mathrm{var}(n) = (\nu + \rho)t$.

Another important example is when we consider simple birth and death processes. In such processes, the rates at which events occur scale with the population size and are thus density dependent. The simplest and most common choice is for λ_n and μ_n to depend linearly on n. The choice $\lambda_n = \lambda n$, $\mu_n = 0$ (for

example, a unidirectional chemical reaction $A \to B$, which we will consider later), gives

$$G(t,z) = \left(1 - e^{\lambda t}\left(1 - \frac{1}{z}\right)\right).$$

The mean is found to be $\langle n \rangle = me^{\lambda t}$, just as we would get from the deterministic model $u' = \lambda u$, but we now we also know how the variance changes in time, $\mathrm{var}(n) = me^{\lambda t}(e^{\lambda t} - 1)$. For the general birth and death process, with $\lambda_n = \lambda n$, $\mu_n = \mu n$, and assuming $\lambda \neq \mu$, we introduce auxially variables

$$\sigma = e^{(\lambda - \mu)t}, \gamma = \frac{\lambda}{\mu}.$$

With these, the generating function is found to be

$$G(t,z) = \left[\frac{(\sigma - 1) + (\gamma - \sigma)z}{(\gamma\sigma - 1) + \gamma(1 - \sigma)z}\right]^m.$$

The mean $\langle n \rangle$ is simply $m\sigma$, and the variance is

$$\mathrm{var}(n) = \frac{m\sigma(\gamma + 1)(\sigma - 1)}{\gamma - 1}.$$

There are many more explicitly known solutions, and also the $P_{n,m}(t)$ are known in each case (but are more cumbersome). For variations, we may consider so-called restricted processes, which have $\mathbb{N}$ as state space instead of $\mathbb{Z}$, with either absorbing boundaries (hitting 0 means the end of the process) or reflecting boundaries (hitting 0 means the random variable will 'bounce back' to 1 in the next time step). Just as for regular PDE theory, we then need to change the boundary conditions, or extend the initial conditions to obtain the master equations on such restricted state spaces. Much detail may be found in the classic reference by Goel and Richter-Dyn (1974). This book also contains a wealth of information about explicitly known expressions for dwell times (the time before an event occurs), and related quantities.

11.4 Species diversity on islands

A classical field in ecology is that of biogeography, the distribution of species over time and space. Current species distributions are often the result of a mix of processes occurring over shorter (on the order of thousands of years, e.g., ice ages) and much longer time spans (millions of years). The relative contributions of each are often hard to unravel. Perhaps the easiest situation to consider is one in which numbers of individuals (and thus also species) is relatively small and turnover of individuals is large. Such a situation may be found on islands which are closer or farther away from a mainland continent. On the continent, the evolutionary dynamics will contribute vitally to the species composition, but especially on young islands, this cannot be the case. In that case, the species dynamics should be chiefly caused by immigration and emigration of new individuals from the mainland.

This was precisely the case considered in a famous monograph by MacArthur and Wilson (1967), a theory still taught in introductory courses in ecology. The idea is very simple. We consider individuals coming from a fixed mainland. Initially, we assume that there are no species at all on the island. Hence, new individuals will almost invariably be new species, but the immigration rate of new species will decline as diversity increases. The established species on the island are assumed to be in some sort of direct or indirect competition, so that sometimes species go extinct. If the probability that any species on the island goes extinct is fixed, then the probability that any one goes extinct scales with species numbers.

The simplest concrete implementation of the above processes is $\lambda_n = \lambda(K-n)$, $\mu_n = \mu n$. After a simple change of variables for λ_n, this is a simple birth and death process discussed in the previous section. The generating function $G_m(t,z)$ is thus explicitly known, and so are the $P_{n,m}(t)$. For this context, the most interesting quantity is

$$P_{n,0}(t) = \binom{K}{n}\Big(\frac{1+\gamma\sigma}{1+\gamma}\Big)^{K-n}\Big(\frac{\gamma(1-\sigma)}{1+\gamma}\Big)^{n},$$

where $\gamma = \lambda/\mu$, $\sigma = e^{-(\lambda+\mu)t}$. For $t \to \infty$, the species composition approaches a binomial distribution,

$$P_{n,0}(t) \to \binom{K}{n}\gamma^n(1+\gamma)^{-K}, \quad , n = 1, \ldots, K.$$

The mean number of species over time is given by

$$\langle n\rangle = \frac{K\lambda}{\lambda+\mu}\Big(1 - e^{-(\lambda+\mu)t}\Big) \to \frac{K\lambda}{\lambda+\mu}.$$

The variance is found to be similar but somewhat more complicated, with equilibrium value

$$\frac{K\lambda\mu}{(\lambda+\mu)^2}.$$

This process thus gives a wealth of information to test in real island systems, and for a comprehensive overview of such result we refer to the monograph by MacArthur and Wilson. For a more modern theory of biodiversity which also takes evolutionary time scales into account, and which has caused quite a stir in the ecological community, read (Hubbell, 2001).

11.5 Simple (bio)chemical reactions

We now turn to systems of chemical reactions, one of the main fields of applications, next to population genetics, of master equations. We start simple, with a unimolecular reaction

$$\mathrm{A} \xrightarrow{\ k\ } \mathrm{B}.$$

The random variable we track is the number of A molecules at time t. The master equation for the probability of finding n molecules of type A when initially starting out with m such molecules is given by

$$\frac{d}{dt}P_{n,m}(t) = \mu_{n+1}P_{n+1,m}(t) - \mu_n P_{n,m}(t). \tag{11.5.1}$$

After all, there are two ways by which the probability of finding n molecules may change: either there were $n+1$ molecules, and one molecule of A was converted into B, or there were n molecules, and after a single reaction there are now $n-1$. Chemical reactions clearly fall in the class of birth and death processes, since the probability that any one of the n molecules is converted through reaction (11.5) is proportional to the number of molecules. Hence, we choose $\mu_n = kn$.

From Section 11.3, we know that the mean satisfies the deterministic equation,

$$\frac{d}{dt}\langle n\rangle = -k\langle n\rangle,$$

so that $\langle n\rangle = me^{-kt}$. This nice coincidence between the mean of the stochastic process and the deterministic version does not generally hold true, but is restricted to unimolecular reactions.

In the same vain, we may write down the master equations for Michaelis-Menten kinetics of the enzymatic conversion of a substrate. Recall equations (2.1.2)–(2.1.5) from Section (2.1),

$$\begin{aligned}
\frac{ds}{dt} &= -k_+ se + k_- c, \\
\frac{de}{dt} &= -k_+ se + k_- c + kc, \\
\frac{dc}{dt} &= \ \ k_+ se - k_- c - kc, \\
\frac{dp}{dt} &= \qquad\qquad\quad + kc.
\end{aligned}$$

Recall that $e(t) + c(t) = e_0$, the total amount of enzyme in the reaction vessel. The product P does not appear in the reaction equations, and can thus be ignored, leaving us with two variables, e and s. We now make e and s discrete variables. The master equations are written for the variable $P(e, s; t)$, the probability of finding e molecules of type E, and s molecules of type S at time t. From the above reaction scheme we deduce that

$$\begin{aligned}
\frac{d}{dt}P(e,s;t) = {} & k_+(e+1)(s+1)P(e+1,s+1;t) \\
& + k_-(e_0-(e-1))P(e-1,s-1;t) \\
& + k(e_0-(e-1))P(e-1,s;t) \\
& - k_+ es + (k_- + k)(e_0 - e))P(e,s;t).
\end{aligned}$$

For instance, if the reaction

$$\mathrm{C} \xrightarrow{\ k\ } \mathrm{P} + \mathrm{E}$$

takes place, then $e - 1 \to e$, $s \to s$, and $c = e_0 - (e - 1) \to e_0 - e$, so that we get a contribution

$$k(e_0 - (e-1))P(e-1, s; t)$$

to the master equations above.

Certainly from a simulation perspective, but also from one of biological insight, it would be preferable to reduce the above system of equations for $P(e, s; t)$. Computationally, this system has a number of equations totalling the product of s_0, the initial substrate concentration, and e_0. But we also know that "not all chemical species are equal", i.e., that the enzyme is often bound in equilibrium to the substrate, and that the substrate changes but slowly compared to the enzyme kinetics. We would like to perform some kind of QSSA reduction, just as in Section (2.1.19) resembling the Michaelis-Menten equation

$$\frac{d\bar{s}}{d\tau} = -\frac{V_{\max}\bar{s}}{K_m + \bar{s}}.$$

Now the hopeful end result would be of the form of unimolecular reaction (11.5.1), with master equations

$$\frac{d}{dt}P(\bar{s};t) = a_{k+1}P(\overline{s+1};t) - a_k P(\bar{s},t). \tag{11.5.2}$$

This is the subject of the next section.

11.6 The QSSA approximation for general chemical reactions

Until the 1970s, the main focus in master equations research was to derive full descriptions of $P_{n,m}(t)$, its generating function, or its moments, akin to trying to find explicit solutions for ODE or PDE problems. This has severe limitations, and it is often more useful to try to find *structure* in equations, or uncover *qualitative* properties of solutions. The QSSA approximation falls in the first category. By reducing the number of equations, they become more manageable, but more interestingly, the new equations show how the long-term dynamics may be captured in a succinct form.

A general reaction scheme has master equations of the form

$$\frac{d}{dt}P(x,t) = \sum_{k=1}^{m} a_k(x - v_k)P(x - v_k; t) - a_k(x)P(x; t). \tag{11.6.1}$$

Here, x is a vector denoting the numbers of molecules of each species x_i, there are m reactions involved, $a_k(x)$ measures the propensity of the k-th reaction, and v_k denotes the number of molecules of the various species involved in the k-th reaction (the *stoichiometry* of the reaction). If a QSSA reduction is to be expected, then we should be able to divide the chemical species into *primary* species and *intermediary* species. If we would apply this to the familiar Michaelis-Menten context, the former would play the role of the substrate and would have slow dynamics, and the latter would have the role of enzyme and complex and have fast dynamics. In other words, the intermediary species will have zero net flux in the slow time scale. Let us thus split our vector x in these two classes, and write $x = (y, z)$, where y form the (numbers of molecules of) primary species, and z the intermediary species. Similarly, the stoichiometry vectors v_k will be split into their v_k^y and v_k^z counterparts. Then the master equations (11.6.1) read

$$\frac{d}{dt}P(y,z) = \sum_{k=1}^{m}[a_k(y - v_k^y, z - v_k^z)P(y - v_k^y, z - v_k^z) - a_k(y,z)P(y,z)]. \tag{11.6.2}$$

The time dependence of $P(y, z; t)$ is suppressed in this notation to keep the formulas a bit more transparent. We want to separate out $P(y; t)$ to find the master equations that it satisfies, and to do so we use conditional probabilities for finding z given y,

$$P(y, z; t) = P(z|y; t)P(y; t).$$

By the product rule, (11.6.2) is equivalent to

$$P(y)\frac{d}{dt}P(z|y) + P(z|y)\frac{d}{dt}P(y) = \sum_{k=1}^{m}[a_k(y - v_k^y, z - v_k^z)P(z - v_k^z|y - v_k^y)P(y - v_k^y) - a_k(y,z)P(z|y)P(y)].$$

The main point is that z should hardly change when y is fixed, i.e.,

$$\frac{d}{dt}P(z|y) \approx 0.$$

Assuming that z conditional on y is Markovian, we may approximate $\frac{d}{dt}P(z|y)$ by

$$\frac{d}{dt}P(z|y) = \sum_{k=1}^{m}[a_k(y - v_k^y, z - v_k^z)P(z - v_k^z|y) - a_k(y,z)P(z|y)$$

Setting this to zero and solving for $P(z|y)$ (now truly time independent), we find

$$P(z|y)\frac{d}{dt}P(y) \approx \sum_{k=1}^{m}[a_k(y-v_k^y, z-v_k^z)P(z-v_k^z|y-v_k^y)P(y-v_k^y) - a_k(y,z)P(z|y)P(y)].$$

Since $\sum_z P(z|y;t) = 1$, we may find the marginal distribution $P(y;t)$ by summing the above equation over z. Then

$$\frac{d}{dt}P(y;t) = \sum_z P(z|y)\frac{d}{dt}P(y;t) \approx \sum_{k=1}^{m}[b_k(y-v_k^y)P(y-v_k^y) - b_k(y)P(y), \quad (11.6.3)$$

where

$$b_k(y) := \sum_z a_k(y,z)P(z|y).$$

Equation is a problem only for y, and is our QSSA approximation of (11.6.1).

Let us apply this to our favourite benchmark, Michaelis-Menten kinetics. As in Section 2.1, we need to rescale the variables so that a small parameter appears in the system. Vector x consists only of two chemical species s and e, the first of which is primary, the second intermediate. Set

$$\bar{s} = \frac{s}{s_0}, \quad \bar{e} = \frac{e}{e_0}, \quad \epsilon = \frac{e_0}{s_0}, \quad \tau = e_0^2 t.$$

In this scaling, $\bar{e}$ and $\bar{s}$ are no longer integers, but change with increments $d = \frac{1}{e_0}$. We find master equations of the form

$$\epsilon\frac{d}{d\tau}P(\bar{e},\bar{s};\tau) = f(P(\bar{e},\bar{s};\tau), P(\bar{e}-d,s;\tau), P(\bar{e}+d,s;\tau), P(\bar{e}+d,s+\epsilon d;\tau)),$$

for a suitable function f. Setting $\epsilon = 0$, we obtain an algebraic relation, which may be inserted into the master equation for $P(e,s;t)$ for unscaled e and s in original time t. In short,

$$\frac{d}{dt}P(e,s;t) \approx P(e|s;t)\frac{d}{dt}P(s;t),$$

and taking the marginal density over e, we find

$$\frac{d}{dt}P(s;t) = k_2\mathbb{E}(e|s+1)P(s+1;t) - k_2\mathbb{E}(e|s)P(s;t),$$

where the expectation is of the form

$$\mathbb{E}(e|s) = \frac{e_0 s}{K_m + s}.$$

To sum up, we have indeed recaptured the Michaelis-Menten rate equation

$$\frac{d}{dt}P(s;t) = \frac{V_{\max}(s+1)}{K_m+s+1}P(s+1;t) - \frac{V_{\max}s}{K_m+s}P(s;t)$$

which has the form (11.5.2) we sought to find. See (Rao and Arkin, 2003) for more details.

Chapter 12

Pitfalls in modelling biological phenomena

In this chapter, we review a number of common modelling mistakes, made by beginners and experts alike.

12.1 Theme 1: continuous versus discrete time

First of all, the principle of mass action applies to rates, not to maps. Consider for instance everyone's favourite starting point for a model of population dynamics of one species, the (continuous) logistic equation,

$$\frac{dN}{dt} = rN\Big(1 - \frac{N}{K}\Big). \tag{12.1.1}$$

Naively, we might write down its direct discrete counterpart, which is maybe even more famous among mathematicians for its relation with period-doubling and chaos, the (discrete) logistic map

$$N_{t+1} = rN_t\Big(1 - \frac{N_t}{K}\Big). \tag{12.1.2}$$

The maximum of the latter right hand side is $rK/4$, achieved for $N = K/2$. This maximum exceeds K if $r > 4$ and then N_t can become negative, even if N_0 is positive. In other words, the positive cone $N \geq 0$ is not invariant under the dynamics of the logistic map. There is evidently something wrong with this map as a discrete time population model.

The appropriate discrete time version of the logistic equation (12.1.1) is the *Beverton-Holt map*,

$$N_{t+1} = \frac{r}{1 + \frac{r-1}{K}N_t}N_t. \tag{12.1.3}$$

Note that the multiplication factor

$$\frac{r}{1 + \frac{r-1}{K}N}$$

is a decreasing function of N, while

$$N \mapsto \frac{rN}{1 + \frac{r-1}{K}N}$$

is increasing, leading to montone orbits, just as for the logistic differential equation.

For the *Ricker map*,

$$N_{t+1} = re^{-\frac{\log r}{K}N_t}N_t, \tag{12.1.4}$$

the multiplication factor

$$re^{-\frac{\log r}{K}N}$$

is decreasing as well, but

$$N \mapsto re^{-\frac{\log r}{K}N}N$$

has a humped graph, creating the possibility of oscillations (one speaks of overcompensatory density dependence). Ricker introduced this map in order to incorporate that adults might eat the eggs and larvae that they produced (so, in other words, to incorporate cannibalistic activity). In a certain year, the N adults produce $\tilde{r}N$ eggs. During a certain time window they eat, perhaps accidentally, aE eggs per unit of time if egg density equals E. Let τ denote time during this time window and let τ be varying between 0 and $\tau_{\max}$. Then

$$\frac{dE}{d\tau} = -aEN,$$
$$E(0) = \tilde{r}N,$$

and consequently

$$E(\tau_{\max}) = \tilde{r}Ne^{-a\tau_{\max}N}.$$

Now assume that the adult density in the next year is proportional to $E(\tau_{\max})$ with, say, constant of proportionality p. Then we obtain the Ricker map by choosing

$$r = \tilde{r}p, \quad K = \frac{\log r}{a\tau_{\max}}.$$

It is, quite generally, much better to derive discrete time models by using continuous time building block as above (see (Geritz and Kisdi, 2004)). Reproduction is often concentrated in a very short period of the year, but interactions (predation, competition for food, transmission of an infectious disease) often happen during the entire year. Hybrid models are built by combining continuous time differential equations describing interaction with a reproduction event at a specified period sequence of time points. Sometimes, one can, like we did above when deriving the Ricker map, explicitly solve the differential equations and obtain an explicit map. In general, one obtains a map that is implicitly defined in terms of solutions of differential equations.

EXERCISE 12.1.1. What is wrong with the discrete time formulation

$$S_{t+1} = S_t - \beta S_t I_t,$$
$$I_{t+1} = (1-\alpha)I_t + \beta S_t I_t,$$

of a model for the epidemic spread of an infectious disease? Reformulate the equations such that they do make sense, and derive how one arrives from these at the formulation above.

EXAMPLE 12.1.1. Let us consider a metapopulation model, i.e., populations living in patches and coupled by migration. The two patch variant of the logistic equation reads

$$\frac{dN}{dt} = r_1 N\Big(1 - \frac{N}{K_1}\Big) - d_1 N + d_2\theta_{12}M, \tag{12.1.5}$$
$$\frac{dM}{dt} = r_2 M\Big(1 - \frac{M}{K_2}\Big) + \theta_{21}d_1 N - d_2 M. \tag{12.1.6}$$

Here N denotes the population density in patch 1 and M the population density in patch 2, and d_i is the rate of emigration out of patch i. The constants θ_{ij} incorporate two effects,

- a migrating individual may not survive the trip;
- if the patches are not of equal size, one needs to convert the change of density in one patch to the change of density in the other (the safe way of doing this is by computing the *number* of emigrants).

Let us make our life easy and assume that $r_1 = r_2 = r$, $K_1 = K_2 = K$, $d_1 = d_2 = d$, the patches are of equal size and no death occurs during travel from one patch to the other. Note, incidentally, that the formulation ignores that the travelling may take time.

Due to seasonality, a discrete time model is more often appropriate (in particular for plants that flower once per year during a short period). Suppose the one patch formulation is

$$N_{t+1} = f(N_t),$$

then what is the two-patch formulation? Is it

$$N_{t+1} = f(N_t) - dN_t + dM_t,$$
$$M_{t+1} = f(M_t) + dN_t - dM_t?$$

No! In fact, one has to go into the derivation of f from a description of the life cycle and then incorporate dispersion of seeds (or whatever is the appropriate migration mechanism).

For instance, consider an annual plant and assume that the variables correspond to the density of flowering plants. Let $p(x)$ be the probability that, given seed density x, a seed survives the winter, wins the local competition with other seeds and develops into a flowering plant the next year. Define

$$f(x) = xp(x).$$

Assume that a flowering plant produces on average r seeds. Then in the one-patch situation we obtain

$$N_{t+1} = f(rN_t),$$

but in the two-patch situation

$$N_{t+1} = f((1-d)rN_t drM_t),$$
$$M_{t+1} = f(drN_t M_t(1-d)rM_t).$$

Here, we hope it is clear how to put the indices and insert the θ when patches differ (in size, quality, or both) and dispersing seeds may land in unsuitable habitat.

As a final remark, note that the two-patch model for plants would become quite different potentially if the moment of census would have been chosen differently. Then one would for instance count ungerminated seeds, rather then plants, which have quite different dynamics. This is clearly a silly thing to do in the above example, but when modelling parasitoids with complex life cycles, there are often several obvious choices to be made of what to do the bookkeeping on, with different final results.

EXERCISE 12.1.2. What if the population census is based on seed density at the beginning of the winter?

More realistic models often take a seed bank into account.

12.2 Theme 2: the danger of delay

The otherwise quite respectable ecologist G. E. Hutchinson published in 1948 the paper *Circular causal systems in ecology* and introduced in that paper the delayed logistic equation

$$\dot{u}(t) = ru(t)\Big(1 - \frac{u(t-1)}{K}\Big),$$

which became a favourite prototype of a scalar nonlinear differential equation with delay for quite a few applied mathematicians. The problem with this equation is that is very hard, if not impossible, to interpret the term

$$-ru(t)\frac{u(t-1)}{K}.$$

If we think of *maturation delay*, more appropriate versions are

$$\dot{u}(t) = \beta u(t-1) - \mu u(t) - \sigma u(t)u(t),$$

(where the last term reflects increased per capita mortality at higher densities) and

$$\dot{u}(t) = \beta u(t-1)e^{-u(t-1)} - \mu u(t),$$

(where the first term incorporates egg cannibalism or other effects, see the paper by Gurney et al. (1980)). Another, more phenomenological, option is

$$\dot{u}(t) = \beta(K - u(t-1))_+ + u(t-1) - \mu u(t),$$

where $x_+ := \max\{0, x\}$ and where the first term reflects reduced fertility at higher densities. Now recall the equations (**??**) from Chapter **??**

$$\frac{dv}{dt} = v\Big(1 - ev - \frac{bp}{1+b\beta v}\Big), \tag{12.2.1}$$

$$\frac{dp}{dt} = p\Big(-c + \frac{dv}{1+b\beta v}\Big), \tag{12.2.2}$$

corresponding to a prey/resource (v) - predator/consumer (p) model. A persistent mistake made by applied mathematicians that love delay equations, but are not willing to devote a small fraction (say 5 or 10 percent) of their research time to thinking about the model motivation, assumptions, meaning, etc., is to give all functions argument t except for the rightmost term in the second equations, which they take as

$$\frac{dv(t-1)}{1+b\beta v(t-1)}.$$

If again we think of maturation delay and if we assume that juvenile predators have another source of food, we should write the second equation as

$$\frac{dp}{dt}(t) = -cp(t) + \frac{dp(t-1)v(t-1)}{1+b\beta v(t-1)}.$$

Next, recall equation (6.1.1) from Chapter **??**,

$$\frac{\partial u}{\partial t} = d\Delta u + f(u),$$

with for instance $f(u) = ru(1 - \frac{u}{K})$. How do we incorporate maturation delay in the setting of spatially distributed populations? When juveniles do *not* move, we can replace $f(u)$ by, for instance

$$\beta u(t-1, x)e^{-u(t-1,x)} - \mu u(t, x).$$

So this would correspond to a situation where eggs are dropped and stay put until they hatch (with the mother moving on). A debatable model, for sure, but at least without inconsistency.

If juveniles do move, we have to realise that those who are at position x at time t have different histories in the sense that they were at different positions in the recent past. It is now safer to write the equation in the form

$$\frac{\partial u}{\partial t} = d\Delta u + h - \mu u,$$

and to specify separately how the source term h depends on the history of u as a function of the spatial variable x. If movement and survival of juveniles is *not*

influenced by population density, we may introduce as a higher level modelling ingredient

$$p(1,x,\xi) := \text{probability that a juvenile that is at } \xi \text{ at some time is still alive one time unit later and is then located at } x$$

and next write

$$h(t,x) = \int_\Omega p(1,x,\xi)\beta u(t-1,\xi)e^{-u(t-1,\xi)}d\xi.$$

If $\Omega = \mathbb{R}^n$, juveniles diffuse with diffusion constant $\tilde{d}$ and are subject to a per capita mortality rate $\tilde{\mu}$, we may compute p by solving

$$\frac{\partial p}{\partial t} = \tilde{d}\Delta p - \tilde{\mu}p,$$
$$p(0,x,\xi) = \delta_\xi(x).$$

and evaluate for $t = 1$. For bounded domains, we need to take boundary conditions into account. See (Gourley et al., 2004) for further details. The moral is

- distinguish carefully between death terms and birth terms when writing bookkeeping equations
- think in terms of i-state, i.e., the state of individuals, and mechanisms (how does density dependence come about? is death affected by density or birth or movement or maturation or all of these?)
- do *not* write vague sentences like "what happens now, manifests itself somewhat later" or "all processes take some time to complete" as a motivation for changing somewhere in an equation an argument t by $t-1$ (or, the unscaled variant $t-\tau$)

Exercise 12.2.1. Starting from the Lotka-Volterra prey-predator model described by (3.1.1)–(3.1.2) on page 25, formulate a variant where both prey and predator diffuse and prey eaten is converted into predator offspring with gestation delay of one unit of time.

12.3 Theme 3: going astray while bookkeeping

The PDE formulation of age-dependent population growth is given by

$$\frac{\partial n}{\partial t} + \frac{\partial n}{\partial a} = -\mu n$$
$$n(t,0) = \int_0^\infty \beta(\alpha)nt,\alpha)d\alpha.$$

The differential equation captures aging and survival, while the boundary condition expresses that the influx of newborns equals the rate at which offspring is produced.

Next, suppose we distinguish individuals from each other by their *size* x (rather than their *age* a). Assume that newborn individuals have size x_b. The differential equation now reads

$$\frac{\partial n}{\partial t} + \frac{\partial}{\partial x}(gn) = -\mu n, \quad x > x_b,$$

where $g = g(x)$ is the growth rate of an individual of size x (i.e., a new modelling ingredient that we need to specify). You should recognize gn as the *flux*, i.e., the density times velocity, recall (4.1.6), page 34. Every now and then it happens that a mathematician who is not really interested in modelling writes the boundary condition as

$$n(t,x_b) = \int_{x_b}^\infty \beta(\xi)n(t,\xi)d\xi.$$

A first way to see that this must be wrong is to consider the physical dimension: at the left side we have a number per unit of size, while at the right we have a number per unit of time. Another way is to integrate the differential equation with respect to x from x_b to infinity, and to interpret the boundary term as the inflow of newborns. Equivalently, just recall the definition of "flux". Any of these considerations leads to the correct boundary condition

$$g(x_b)n(t,x_b) = \int_{x_b}^{\infty} \beta(\xi)n(t,\xi)d\xi.$$

Now return to the age-structured population and recall from equation (8.1.2) the formula

$$n(t,a) = b(t-a)\mathcal{F}(a),$$

where $b(t)$ is the population birth rate and $\mathcal{F}(a)$ the survival probability (so $\mathcal{F}(a) = \exp(-\int_0^a \mu(\alpha)d\alpha)$ and $b(t) = \int_0^{\infty} \beta(\alpha)n(t,\alpha)d\alpha$ if one specifies the model in terms of a per capita death rate μ and a per capita fertility rate β). In some models it is assumed that individuals die instantaneously upon reaching some maximal age $a_{\max}$. In other words, $\mathcal{F}(a) = 0$ for $a > a_{\max}$. Next it is sometimes assumed that $a_{\max}$ is *dynamic*, say determined by environmental conditions like resource abundance, etc. Now note that if one allows the possibility that

$$\frac{da_{\max}}{dt} > 1$$

and if one does not modify the formula $n(t,a) = b(t-a)\mathcal{F}(a)$, it may happen that already dead individuals are resurrected!

In case of a size-structured population, a frequent assumption is that individuals become adult (i.e., start to reproduce) upon reaching a certain size x_A. The individual growth rate g is often a function of both the i-state x and the environmental condition S, where S is the resource concentration. To facilitate the exposition, assume that g is independent of x.

In order to know at what time the individuals were born that mature (i.e., become adult) at time t, we need to solve the equation

$$x_A = x_b + \int_{t-\tau}^{t} g(S(\eta)d\eta$$

for τ, given the food history as described by the function S. By differentiation with respect to t we deduce that

$$0 = g(S(t)) - g(S(t-\tau))\Big(1 - \frac{d\tau}{dt}(t)\Big)$$

and hence that

$$\frac{d\tau}{dt}(t) = 1 - \frac{g(S(t))}{g(S(t-\tau))}.$$

Note that, under reasonable assumption concerning S and g, this implies that $\frac{d\tau}{dt} < 1$, or in words, rejuvenation is impossible!

Suppose that juveniles die at a per capita rate μ_1, adults at a per capita rate μ_2 and that adults produce offspring at a per capita rate β. Let A denote the size of the adult population. It is tempting to describe the dynamics of A by the equation

$$\frac{dA}{dt}(t) = \beta e^{-\mu_1\tau(t)}A(t-\tau(t)) - \mu_2 A(t),$$

but this is incorrect. The safe way of deriving the equation is to write first

$$A(t) = \int_{\tau(t)}^{\infty} b(t-a)\mathcal{F}(a)da = \int_{-\infty}^{t-\tau(t)} b(\sigma)\mathcal{F}(t-\sigma)d\sigma,$$

(where, admittedly, our notation for the somewhat complicated survival probability arising from a possible difference between μ_1 and μ_2 is rather sloppy). Differentiation then leads to

$$\frac{dA}{dt}(t) = -\mu A(t) + b(t-\tau(t))\mathcal{F}(\tau(t))\Big(1 - \frac{d\tau}{dt}(t)\Big).$$

If we now insert $b(t) = \beta A(t)$, $\mathcal{F}(\tau(t)) = e^{-\mu_1 \tau(t))}$ and the expression for $\frac{d\tau}{dt}(t)$, we arrive at

$$\frac{dA}{dt}(t) = -\mu A(t) + \beta A(t-\tau(t))e^{-\mu_1 \tau(t)} \frac{g(S(t))}{g(S(t-\tau(t))},$$

where the last factor accounts for the fact that, modulo changes due to death, the *flux* is conserved, and not the density!

Chapter 13

Appendix

13.1 Bifurcation theory

13.1.1 Structural (in)stability by way of example

The phase portrait of the Volterra-Lotka system (3.1.1)–(3.1.2) is *structurally unstable*. By this we mean that arbitrarily small perturbations of the vector field yield phase portraits which are truly different. Indeed, the family of closed orbits can either develop into inward spiralling orbits (Section 3.3) or into outward spiralling orbits (Section 3.4), depending on how we perturb.

For the Rosenzweig-MacArthur system (3.6.1) there are two relations among parameters that mark transitions in the essential features of the phase portrait. When

$$\frac{a}{e} = \frac{c}{d - bc\beta} \tag{13.1.1}$$

the predator isocline hits the prey-only steady state on the v-axis. One can think of this relation as defining a 5-dimensional surface in the 6-dimensional parameter space. On the side of the surface where

$$\frac{a}{e} < \frac{c}{d - bc\beta}$$

the predator goes extinct and all orbits converge to $(v, p) = (\frac{a}{e}, 0)$. On the other side of the surface, the system has a positive (non-trivial) steady state, which is stable for parameter values near to the surface. In fact, the second relation

$$\frac{a}{e} = \frac{2c}{d - bc\beta} + \frac{1}{b\beta} \tag{13.1.2}$$

characterises the loss of stability of the positive steady state (indeed, when (13.1.2) holds the predator isocline hits the top of the parabola forming the prey isocline).

There are both technical and physiological/psychological reasons to prefer two-dimensional pictures. So we shall consider four of the six parameters as fixed and only two as variable. The relations (13.1.1) and (13.1.2) then define curves in the plane, which we can sketch.

We choose the dimensionless compound parameters $\frac{a}{e}$ and $b\beta$ as our variable parameters. The first is the carrying capacity of the prey, the second a measure for predator efficiency. Considering $\frac{d}{c}$ as fixed, we then represent (13.1.1), (13.1.2) and the results of Section 3.6 graphically in Figure 13.1.

Each of the regions 1, 2 and 3 is characterised by a certain phase portrait. If we pick a parameter point inside a region, the phase portrait is *structurally stable* (in the sense that small changes of the vector field do not change it). In contrast, the regions are separated from one another by boundary curves corresponding to

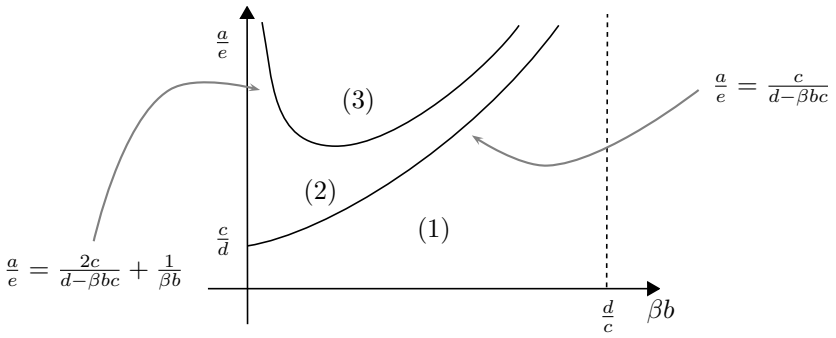


FIGURE 13.1. The phase portrait of the Rosenzweig-MacArthur system.

parameter values for which the phase portrait is structurally unstable. The curve defined by (13.1.2) is called the *stability boundary* in the parameter space, as it separates region 2, where the positive steady state is stable, from region 3, where this steady state is unstable.

13.1.2 Topological equivalence and structural (in)stability

What exactly do we mean when we say that two phase portraits (or two dynamical systems) are the same? How do we translate the intuitive idea of two pictures that look the same into a precise formal definition?

A *homeomorphism* is a one-to-one map that is continuous and has a continuous inverse. Two dynamical systems are called *topologically equivalent* if there exists a homeomorphism between their state spaces that maps orbits onto orbits, while preserving the order in which orbits are traversed in the course of time. One then also speaks (as we did above) about the topological equivalence of the two phase portraits.

The aim of the qualitative theory of dynamical systems is to give a catalogue of all equivalence classes and, in addition, to be able to derive from information about the generating vector field what is the phase portrait in a particular case. This is too ambitious. A more pragmatic approach is to concentrate on special orbits (points, circles) or, more generally, invariant subsets (tori, chaotic attractors) and describe the structure of the phase portrait in a neighbourhood (the *local* phase portrait) as well as the behaviour within the invariant subset.

The (local) phase portrait is called *structurally stable* if the topological equivalence class does not change when we perturb the vector field a little. To make this precise, we need to specify how we measure perturbations of the vector field. We do this by means of the C^1-topology, which basically means that the values of both the function itself and of its derivative should be close in order to call the perturbation small.

We emphasise that often it makes sense to require that the perturbations preserve a certain structure (like invariance of the v- and p-axes for prey-predator systems, or Hamiltonian structure for mechanical systems).

Structurally unstable systems might be called degenerate, atypical, non-generic. Degeneracy, however, occurs in degrees. The relations (13.1.1) and (13.1.2) describe degeneracies that occur naturally in one-parameter families of vector fields, in the sense that they cannot be eliminated by a small perturbation of the one-parameter family (if you visualise the one-parameter family as a curve crossing the two curves corresponding to (13.1.1) and (13.1.2), you see that small perturbation in the curve cannot eliminate the crossings). Higher degeneracies require more parameters to be a natural phenomenon.

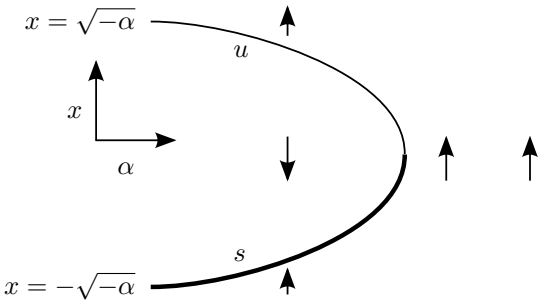


Figure 13.2. Bifurcation diagram for a fold bifurcation corresponding to equation (13.1.3). Note that there are no solutions for $\alpha > 0$. Bold lines signify stable (s) solution branches, thin lines unstable (u) ones.

A *bifurcation diagram* is a (local) partitioning of the parameter space in regions of (local) topological equivalence of the dynamical system, together with a representative picture of the flow for each region (and each boundary). A boundary point is called a *bifurcation point* (and the *codimension* is the difference between the dimension of the parameter space and the dimension of the boundary; equivalently, the codimension is defined as the number of independent conditions determining the bifurcation; in other words, the degree of degeneracy is measured by the codimension and the two relations (13.1.1) and (13.1.2) each determine a codimension one bifurcation; incidentally, the bifurcation corresponding to (13.1.2) is called a *Hopf bifurcation* and is characterised by the appearance of a limit cycle when a steady state changes stability character due to a pair of complex conjugated eigenvalues crossing the imaginary axis, see the next section).

13.1.3 Bifurcations associated with steady states

The local phase portrait near a steady state is structurally stable when none of the eigenvalues of the linearisation is *critical*. The meaning of "critical" depends on the structure of the time variable: in the discrete time case it means "having modulus 1", and in the continuous time case it means "having real part zero".

So bifurcations are characterised by critical eigenvalues. We distinguish three cases:

(i) eigenvalue 1 (discrete time) or 0 (continuous time)
(ii) two complex conjugate roots on the unit circle (discrete time) or on the imaginary axis (continuous time)
(iii) eigenvalue -1 in the discrete time case

For each case, we describe a simple example displaying the typical bifurcation diagram. For case (i) we describe several additional examples that show the effect of special structure. A much more detailed discussion on bifurcation theory can be found in Kuznetsov (2004).

(ia) *fold* (also called saddle-node, limit point, turning point): Let the state variable be $x \in \mathbb{R}$ and the parameter $\alpha \in \mathbb{R}$. With the differential equation

$$\dot{x} = \alpha + x^2 \tag{13.1.3}$$

we associate the bifurcation diagram in Figure 13.2.

Exercise 13.1.1. Construct the diagram for $\dot{x} = \alpha - x^2$.

Exercise 13.1.2. Analyse the discrete time system

$$x(n+1) = \alpha + x(n) \pm (x(n))^2.$$

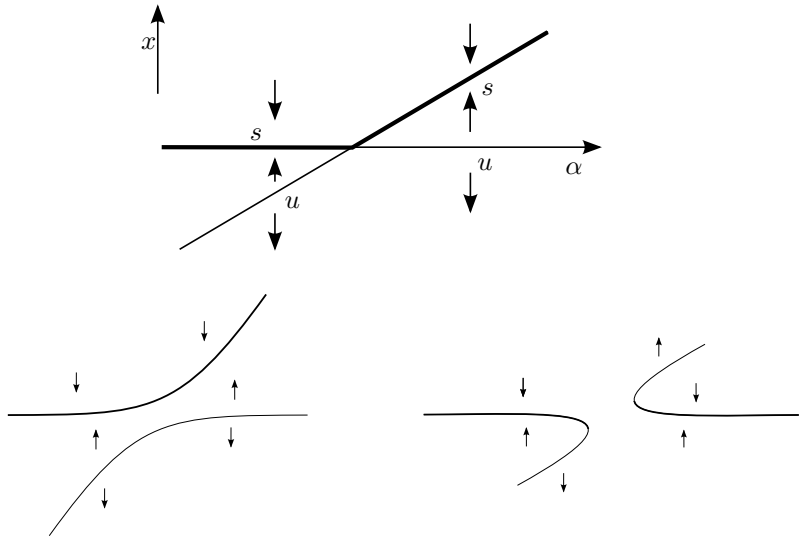


FIGURE 13.3. Bifurcation diagram for a transcritical bifurcation (top). Bifurcation diagrams of perturbed systems can be seen below it. In the left case there are no bifurcations, in the right there are two folds. Bold lines signify stable (s) solution branches, thin lines unstable (u) ones.

(ib) *transcritical bifurcation*: The top part of Figure 13.3 corresponds to

$$\dot{x} = x(\alpha - x),$$

and illustrates the *Principle of Exchange of Stability*. It is typical for systems of the form $\dot{x} = xg(x, \alpha)$ but if one allows perturbations that destroy this structure one can get the bifurcation diagrams shown in the bottom part of Figure 13.3.

(ic) *pitchfork bifurcation* (from which the name "bifurcation" originated):
For

$$\dot{x} = x(\alpha - x^2) \tag{13.1.4}$$

we have the diagram shown in the left part of Figure 13.4. The right part shows the bifurcation structure of

$$\dot{x} = x(\alpha + x^2). \tag{13.1.5}$$

This form of bifurcation is typical for systems with reflection symmetry.

In the $-$ case we speak about a *supercritical* bifurcation since the nontrivial steady states exist for parameter values for which the trivial steady state is no longer stable. The nontrivial steady states inherit the stability character of the trivial steady state and for parameter values near the critical parameter value they are close to the trivial steady state. When we increase the parameter α we notice a change when we pass the critical value $\alpha = 0$, but not a dramatic change. The bifurcation is called *soft*. In contrast, the *subcritical* $+$ case exhibits a *hard/catastrophic* bifurcation since, when the trivial steady state loses stability, no nearby attractor takes over (for the present caricatural system, orbits approach infinity).

EXERCISE 13.1.3. When perturbations destroy the symmetry, the diagram may break into two components. Draw the various components.

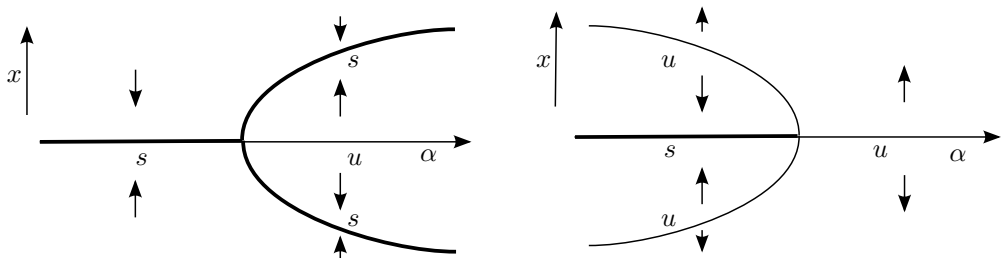


FIGURE 13.4. Bifurcation diagrams for a pitchfork bifurcation. The left figure corresponds to equation (13.1.4), the right to (13.1.5). Bold lines signify stable (s) solution branches, thin lines unstable (u) ones.

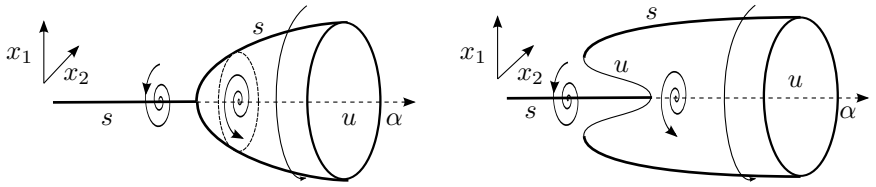


FIGURE 13.5. Supercritical (left) and subcritical Hopf bifurcation diagrams. Bold lines signify stable (s) solution branches, thin lines unstable (u) ones.

(ii) *Hopf bifurcation*: The system

$$\dot{x_1} = \alpha x_1 - x_2 \pm x_1(x_1^2 + x_2^2) \tag{13.1.6}$$

$$\dot{x_2} = x_1 - \alpha x_2 \pm x_2(x_1^2 + x_2^2) \tag{13.1.7}$$

takes in polar coordinates the form

$$\dot{r} = r(\alpha \pm r^2) \tag{13.1.8}$$

$$\dot{\phi} = 1 \tag{13.1.9}$$

There are two different bifurcation diagrams, for the + and − cases, shown in Figure 13.5. The − case is called a supercritical Hopf bifurcation, and the + case a subcritical Hopf bifurcation.

The discrete time version of the Hopf bifurcation is often called the *Neimark-Sacker bifurcation*. It is characterised, like the continuous time case, by the origination of an invariant circle. But it is somewhat more subtle than the continuous time case in that there are various possibilities for the dynamics on this circle: there may or may not be an attracting periodic orbit.

(iii) *flip (or period doubling) bifurcation*: The typical example is

$$x(n+1) = -(1+\alpha)x(n) \pm (x(n))^3.$$

What happens is that at $\alpha = 0$ the fixed point $\bar{x} = 0$ loses its stability by an eigenvalue that leaves the unit circle at -1, and that two points of period two originate. There is again a subcritical and a supercritical case.

EXERCISE 13.1.4. The discrete time logistic equation

$$x(n+1) = \alpha x(n)(1 - x(n))$$

has a nontrivial fixed point

$$\bar{x} = 1 - \frac{1}{\alpha}.$$

Show that at $\alpha = 3$ a supercritical period doubling occurs. Hint: define $f(x) = \alpha x(1-x)$ and compute $f^{(2)}(x) = f(f(x))$. The equation $x = f^{(2)}(x)$ is a fourth order polynomial equation, but we know the solutions $x = 0$ and $x = 1 - \frac{1}{\alpha}$, so we can reduce it to a quadratic equation.

Determine the stability of the period two solution. Hint: consider a periodic point of period two as a fixed point of the iteration $x(n+1) = f^{(2)}(x(n))$.

Show that l'histoire se répète: at $\alpha = 1 + \sqrt{6}$ the period two solution loses stability (and a period four solution arises; it is not part of the exercise to prove this last part of the statement).

Hysteresis refers to the phenomenon that the state in which we find a system may depend on the history, viz. how parameters attained their current values. A dynamical system may have several attractors (at the same parameter values), in which case we speak about *bistability* or *multistability*. The domain of attraction of an attractor may enlarge or shrink as parameters are varied. For fixed parameter values the system may be brought from (the domain of attraction of) one attractor to another by a sufficiently large disturbance. The same transition may alternatively be achieved by parameter variations that are sufficiently large, in which case we may come across hysteresis. Continuous changes in parameters may not be reversible. Coexistence of attractors implies that both chance and necessity play a role: we can predict on the basis of deterministic relations between cause and effect, yet initial conditions and parameter histories may have a decisive influence on what will actually happen.

Figure 13.6 shows a typical bifurcation plot exhibiting hysteresis. As a parameter is changed from low to high, the system first follows the lower branch of stable steady states, until a saddle-node bifurcation occurs. It then jumps to the higher branch. By lowering the parameter again, the system now first follows the higher branch, and after another saddle-node bifurcation drops down to the lower branch again.

EXERCISE 13.1.5. Consider the consumer-resource interaction in the chemostat as described in Section (10.1), but now in the form

$$\frac{dS}{dt} = DS_0 \quad -DS - g(S)X \tag{13.1.10}$$

$$\frac{dX}{dt} = \qquad -DX + \eta g(S)X \tag{13.1.11}$$

with up-take function g that we don't know and that we like to determine experimentallly. The dilution rate D can be easily tuned by the experimenter by just turning the tap (of course S_0 can also be adjusted, but this requires more work, so here we assume S_0 is fixed once and for all). For any particular value of D, the steady state values $\bar{S}$ and $\bar{X}$ can be measured in the outflowing fluid. As a first check on the model, we can compute $\bar{X}(S_0 - \bar{S})^{-1}$ for all measured

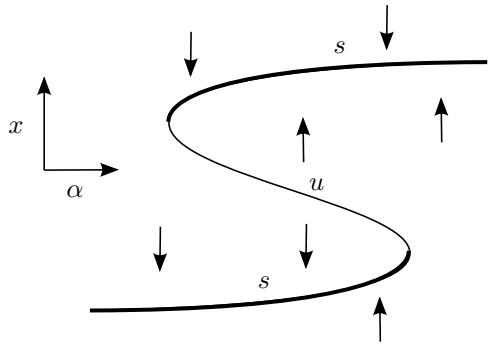


Figure 13.6. A typical hysteresis diagram, with two saddle-node bifurcations at the turning points. Bold lines signify stable (s) solution branches, thin lines unstable (u) ones.

$(\bar{S}, \bar{X})$ combinates with $\bar{S} < S_0$, to see whether this yields (approximately) the same number. Explain the rationale underlying this test.

Next we can plot the points

$$\left(\bar{S}, \frac{D(S_0 - \bar{S})}{\bar{X}}\right)$$

in the plane. Convince yourself that the result should look more or less like Figure **??** when g is of the Michaelis-Menten/Holling type. The dashed lines indicate unobserved unstable (and possibly unphysical) steady states. How would you classify the bifurcation in the terminology introduced in this section?

Some substrates are toxic at high concentrations. In such a case the graph of g may look like Figure **??**. Assume that S_0 exceeds the value of S for which g is maximal. Following the experimental procedure described above, what do you observe when increasing D? (The answer should be a picture!) And what do you observe if, subsequently, you repeat the experiment while decreasing D? (You may assume that the algae are never washed out completely, or that every now and then re-seeding with a tiny amount of algae is carried out.) Combine the two pictures now into one, add unobserved untable steady states and interpret the result in terms of hysteresis.

13.1.4 Poincaré-Bendixson theory

For dynamical systems on the plane $\mathbb{R}^2$, much information on the existence of steady states or of periodic orbits may be obtained using topological arguments. So let us consider the system

$$\dot{u} = f(u), \quad u \in \mathbb{R}^2 \tag{13.1.12}$$

where continuously differentiable f will suffice for our purposes. In Section 3.3 we briefly encountered the notion of the ω-limit set of an orbit of system (13.1.12): a point $p \in \mathbb{R}^2$ belongs to the ω-limit set of an orbit $u(t)$ if there exists a sequence $n \to \infty$ such that $u(t_n) \to p$ as $t_n \to \infty$. Different points on the same orbit have the same ω-limit set, so we can speak of the ω-limit set of an orbit too. The ω-limit set of a positive orbit γ (taking only $t \geq 0$) will be denoted by $\omega(\gamma)$.

The Poincaré-Bendixson Theorem now states that bounded orbits either converge to periodic orbits (or are periodic already), or converge towards a steady state.

Theorem 13.1.1 (Poincaré-Bendixson)*:* Consider a bounded positive orbit γ of (13.1.12) with ω-limit set $\omega(\gamma)$. Then one of the following possibilities holds:

(1) $\omega(\gamma)$ is an equilibrium

(2) $\omega(\gamma)$ is a perdiodic orbit
(3) $\omega(\gamma)$ consists of equilibria and orbits having these equilibria as their α- and ω-limit sets (heteroclinic orbits or one homoclinic orbit).

Of course, analogous results hold for bounded negative orbits (taking $t \leq 0$ only) too. These either come from steady states or from limit cycles, or are steady states or periodic themselves. For a proof of this important result, see e.g. (Coddington and Levinson, 1955; Verhulst, 1996).

13.2 Lyapunov-Schmidt reduction

One of the main problems in nonlinear partial differential equations is to understand the structure of the set of steady state solutions as parameters are varied. In most cases, the Implicit Function Theorem (IFT) ensures that in the neighbourhood of a steady state (u_0, λ_0) at some given parameter λ_0 the steady state solutions can be parametrized by λ. In other words, we find a one-parameter family of steady states $\bar{u}(\lambda)$ such that $\bar{u}(\lambda_0) = u_0$. When the IFT cannot be invoked, we are at a *bifurcation point*. Of course, these points are our main points of interest, since there the number of steady state solutions changes. The idea of Lyapunov-Schmidt reduction is to split the problem into to parts, in such a way that "known" solution can be divided out and the IFT still applies. The method is fully general for finite or infinite systems of nonlinear equations. To get a feel for the method, let us stick to a finite dimensional setting first. The extension to infinite dimensions (and thus, PDEs or integral equations) is then mostly a matter of some technical requirements on the operators involved.

13.2.1 Finite dimensional Lyapunov-Schmidt reduction

Consider a smooth nonlinear map $H : \mathbb{R}^n \times \mathbb{R}^{k+1} \to \mathbb{R}^n$, and the problem of finding the *variables* $x_1, \ldots, x_n$ and *parameters* $\lambda_0, \lambda_1, \ldots, \lambda_k$ such that H vanishes,

$$H_i(x, \lambda) = 0, \quad i = 1, \ldots, n. \tag{13.2.1}$$

Let $DH(0,0)$ be the Jacobian matrix $(\frac{\partial H_i}{\partial x_j})$ evaluated at the origin $(x, \lambda) = (0, 0)$. If $DH(0,0)$ has full rank n, then we may apply the IFT, and know that (13.2.1) may be solved uniquely for x as a function of λ. This is the case in which there is no bifurcation. Consider now the minimally degenerate case, when $\operatorname{rank} DH(0,0) = n-1$. Write $L = DH(0,0)$ for convenience. The Lyapunov-Schmidt reduction now consists of five steps.

Step 1: Choose vector space complements of $\ker L$ and $\operatorname{range} L$ respectively,

$$\mathbb{R}^n = \ker L \oplus M,$$
$$\mathbb{R}^n = N \oplus \operatorname{range} L.$$

Step 2: Let P be the projection of $\mathbb{R}^n$ onto $\operatorname{range} L$, with $\ker P = N$. The complementary projection $(I-P)$ has $\operatorname{range}(I-P) = N$, and $\ker(I-P) = \operatorname{range} L$.

Step 3: Project the original equations (13.2.1) and use the IFT. Since $H(x, \lambda) = 0$, we know that

$$PH(x, \lambda) = 0, \tag{13.2.2}$$
$$(I - P)H(x, \lambda) = 0. \tag{13.2.3}$$

The basic idea of Lyapunov-Schmidt reduction is that (13.2.2) may be solved using the IFT restricted to $\operatorname{range} L$, i.e., in $n-1$ of the components of x. Equation (13.2.3) then gives an equation for the remaining unknown if we substitute the solution from (13.2.2). Why does this work? Decompose the vector $x \in \mathbb{R}^n$ into $v + w$,

with $v \in \ker L$ and $w \in M$. Eq. (13.2.2) is now $PH(v+w,\lambda) = 0$. Note that the derivative with respect to w of $PH(v+w,\lambda)$ at the origin is

$$P(DH(0,0)) = PL = L, \tag{13.2.4}$$

since P is precisely the projection of $\mathbb{R}^n$ onto the range of L. Most importantly, $L : M \to \operatorname{range} L$ is invertible, so we can use the IFT to solve (13.2.2). Denote the solution $w = W(v,\lambda)$, so $W : \ker L \times \mathbb{R}^{k+1} \to M$. The argument of W is just the collection of all remaining variables and parameters for which we cannot solve (13.2.1). With this solution, (13.2.2) reads

$$PH(v + W(v,\lambda),\lambda) = 0, \quad \text{where } W(0,0) = 0. \tag{13.2.5}$$

Step 4: Now substitute this solution into the other equation (13.2.3) still to be solved,

$$(I-P)H(v + W(v,\lambda),\lambda) =: h(v,\lambda). \tag{13.2.6}$$

The zeros of $h(v,\lambda)$ are in one-to-one correspondence with the zeros of $H(x,\lambda)$,

$$h(v,\lambda) = 0 \iff H(v + W(v,\lambda),\lambda) = 0.$$

Step 5: To actually compute things, we have to choose explicit coordinates on $\ker L$ and $\mathbb{R}^{k+1}$ on which h is defined, to get a map $g : \mathbb{R} \times \mathbb{R}^{k+1} \to \mathbb{R}$ instead of h. The usual choice is to choose vectors $v_i \in \ker L$, $v_i^* \in (\operatorname{range} L)^\perp$ (in our case just one each, v_0 and v_0^*, say, because $\dim \ker L = \dim(\operatorname{range} L)^\perp = 1$). Every vector $v \in \ker L$ can be written as tv_0 for $t \in \mathbb{R}$. Now define

$$g(t,\lambda) = \langle v_0^*, h(tv_0,\lambda)\rangle.$$

Since $h(tv_0,\lambda) \in N$, $g(t,\lambda) = 0$ if and only if $h(tv_0,\lambda) = 0$. Hence, the zeros of g are also in one-to-one correspondence with the zeros of $H(x,\lambda)$. Writing out g in full, the projection $I - P$ occurring in the definition of $h(v,\lambda)$ drops out and we find

$$g(t,\lambda) = \langle v_0^*, H(tv_0 + W(tv_0,\lambda),\lambda)\rangle. \tag{13.2.7}$$

This is the final *bifurcation equation* that gives a complete picture of the steady state solutions of $H(x,\lambda) = 0$ we were interested in, in the neighbourhood of the bifurcation point $(x,\lambda) = (0,0)$. The reason that the projection $I - P$ disappears in (13.2.7) is that $v_0^* \in (\operatorname{range} L)^\perp$, and that $Pv \in \operatorname{range} L$ for any vector v, so that $\langle v_0^*, Pv\rangle = 0$. Hence $\langle v_0^*, v\rangle = \langle v_0^*, (I-P)v\rangle$ for any $v \in \mathbb{R}^n$.

The character of the bifurcation is determined by the derivatives of $g(t,\lambda)$. In principle, this just requires taking derivatives of (13.2.5) and (13.2.6) and using the chain rule. Without spelling out all the details of these calculations (which can for instance be found in (Golubitsky and Schaeffer, 1985)), the end result is

$$g_t = 0, \tag{13.2.8}$$

$$g_{tt} = \left\langle v_0^*, D^2H(v_0,v_0)\right\rangle, \tag{13.2.9}$$

$$g_{ttt} = \left\langle v_0^*, D^3H(v_0,v_0,v_0) - 3D^2H(v_0, L^{-1}PD^2H(v_0,v_0)\right\rangle, \tag{13.2.10}$$

$$g_\lambda = \left\langle v_0^*, H_\lambda\right\rangle, \tag{13.2.11}$$

$$g_{\lambda t} = \left\langle v_0^*, DH_\lambda \cdot v_0 - D^2H(v_0, L^{-1}PH_\lambda\right\rangle. \tag{13.2.12}$$

Here, all derivatives are evaluated at $(t,\lambda) = (0,0)$, and

$$D^2H(v,w) = \sum_{i,j=1}^{n} \frac{\partial^2 H}{\partial y_i \partial y_j}(t,\lambda) v_i w_j.$$

As an example, the bifurcation point $(0,0)$ exhibits a pitchfork if the following conditions are met

$$g = g_t = g_{tt} = g_\lambda = 0, \quad g_{ttt}g_{\lambda t} < 0.$$

13.2.2 Infinite dimensional Lyapunov-Schmidt reduction

All we really need to require from some operator L defined on a function space X for the reduction method to work is that L has a finite dimensional kernel and finite dimensional corange. Such operators are called *Fredholm* operators.

Definition 13.2.1*:* Let $L : X \to Y$ be a bounded linear operator between Banach spaces. L is called *Fredholm* if $\ker L \subset X$ is finite dimensional and $\operatorname{range} L \subset Y$ is closed and has finite dimensional codimension. The *index* of L is defined by $\dim \ker L - \operatorname{codim} \operatorname{range} L$.

To perform Step 1 in the reduction, we need to choose closed subspaces M and N such that $X = \ker L \oplus M$ and $Y = N \oplus \operatorname{range} L$. Precisely for Fredholm operators, such closed subspaces M and N can be found.

Often, the index of L is zero, which ensures that if $\ker L = 0$, then L is invertible. If L is a differential operator, X and Y are typically subspaces of $L^2(\Omega)$ for some bounded domain Ω, and L^2 is equipped with inner product

$$\langle u, v \rangle = \int_\Omega u(x)v(x)dx,$$

The context of a particular problem often gives C^k functions for X and Y. We would like to set $M = (\ker L)^\perp$ and $N = (\operatorname{range} L)^\perp$, but it is generally not guaranteed that X is fully spanned by $\ker L$ and its complement, or Y by $\operatorname{range} L$ and its complement. However, in the important case that L is an *elliptic differential operator* this choice for M and N *is* justified. Of course, reaction-diffusion problems are always of this type.

The one additional technical requirement for the Implicit Function Theorem in Banach spaces is that $\operatorname{range} L$ is closed. Note that Fredholm operators have this property by definition.

The upshot is that Lyapunov-Schmidt reduction works for Fredholm operators of index zero.

13.2.3 A concrete example

Let us consider the following problem

$$H(u, \lambda) := Lu + \lambda f(u) = 0, \tag{13.2.13}$$

for $u \in C^2([0, 1])$, $u(0) = 0 = u(1)$, $Lu = u''$ and f a smooth function. Suppose we have a solution (u_0, λ_0). We are interested which other functions solve (13.2.13) in the neighbourhood of λ_0. Linearizing around u_0, we find the derivative in the direction v to be

$$Av := Lv + \lambda Df(u_0)v.$$

L has eigenvalues $-m^2\pi^2$, for $m = 1, 2, \ldots$ and corresponding eigenfunction $\phi_m = \sin m\pi x$. The kernel of A is nonempty if $\lambda Df(u_0) = n^2\pi^2$ for some $n \geq 1$. If this is not the case, then A has full rank and we can use the IFT to find a unique solution $u(\lambda)$ parametrized by λ with the property that $u(\lambda_0) = u_0$. Suppose that $n = 1$, however, and let $A = L + \pi^2$. The kernel of A is all functions v of the form $t\phi_1$, for some $t \in \mathbb{R}$. To see this, write v in Fourier series, $v = \sum_{n=1}^\infty c_n\phi_n$, so that

$$Av = \sum_{n=1}^\infty (-n^2\pi^2 + \pi^2)\phi_n = \sum_{n=2}^\infty (-n^2\pi^2 + \pi^2)\phi_n,$$

and $Av = 0$ if and only if $c_n = 0$, $n \geq 2$. The range of A is all functions w such that $Av = w$. Using Fourier series for w, it is easily seen that w is in the range if and only if w does not contain a multiple of ϕ_1. Hence, $\operatorname{range} A = (\ker A)^\perp$. Let

P be the projection onto the range of A, so that $I - P$ is the projection onto the kernel of A,

$$(I - P)v(x) = 2\phi_1(x) \int_0^1 \phi_1(y)v(y)dy.$$

We can now split $L^2([0,1])$ (or its C^2 subspace?) as $\ker A \oplus \text{range}\, A$. A function u is split accordingly into $u_1 + u_2$.

To perform the third and most important step in the reduction, project the original equation (13.2.13),

$$PH(u,\lambda) = 0, \tag{13.2.14}$$

$$(I - P)H(u,\lambda) = 0. \tag{13.2.15}$$

The first of these, $PH(u,\lambda) = P(Lu + \lambda f(u)) = 0$ is written as

$$0 = P(Lu + \pi^2 u + (\lambda f(u) - \pi^2 u)) = P(Au + h(u,\lambda)),$$

where $h(u,\lambda) := \lambda f(u) - \pi^2 u$ are the higher order terms split off from Au. Now we can properly project with P. Note that

$$PAu = PA(u_1 + u_2) = PAu_2 = Au_2$$

by construction, as discussed in (13.2.4). Au_2 is thus the derivative of $PH(u,\lambda)$ at (u_0,λ_0), and is invertible on range A. Therefore, we apply the IFT and solve for u_2 as a function of u_1 and λ, with solution $U_2(u_1,\lambda)$. Note that $u_1 = t\phi_1$ since $u_1 \in \ker A$, so we can write $U_2(u_1,\lambda)$ as $\tilde{U}_2(t,\lambda)$.

Finally, in step 4, this solution is inserted into the remaining equation to be solved, (13.2.15). Since

$$(I - P)H(u,\lambda) = (I - P)(Au + h(u,\lambda)) = (I - P)h(u,\lambda)$$

equation (13.2.15) becomes

$$g(t,\lambda) := (I - P)h(t\phi_1 + \tilde{U}_2(t,\lambda),\lambda) = 0,$$

which by definition $I - P$ is

$$\left\langle \phi_1, h(t\phi_1 + \tilde{U}_2(t,\lambda),\lambda) \right\rangle = 0.$$

As noted before, this can be rewritten to

$$\left\langle \phi_1, H(t\phi_1 + \tilde{U}_2(t,\lambda),\lambda) \right\rangle = 0.$$

This, then, is the (finite dimensional!) bifurcation equation for this problem.

Two example papers that specifically deal with Lyapunov-Schmidt reduction in systems of reaction-diffusion equations are (Mimura et al., 1979; Mimura and Nishiura, 1979).

Bibliography

Alt, W. (1980). Biased random walk models for chemotaxis and related diffusion approximations. *J. Math. Biology*, 9:147–177.

Amann, H. (1990). *Ordinary Differential Equations: An Introduction to Nonlinear Analysis*. Walter de Gruyter.

Aronson, D. G. (1985). The role of diffusion in mathematical population biology: Skellam revisited. In *Mathematics in biology and medicine (Bari, 1983)*, volume 57 of *Lecture Notes in Biomath.*, pages 2–6. Springer, Berlin.

Aronson, D. G. and Weinberger, H. F. (1975). Nonlinear diffusion in population genetics, combustion, and nerve pulse propagation. In *Partial differential equations and related topics (Program, Tulane Univ., New Orleans, La., 1974)*, pages 5–49. Lecture Notes in Math., Vol. 446. Springer, Berlin.

Aronson, D. G. and Weinberger, H. F. (1978). Multidimensional nonlinear diffusion arising in population genetics. *Adv. in Math.*, 30(1):33–76.

Ballyk, M., McCluskey, C., and Wolkowicz, G. (2005). Global analysis of competitiong for perfectly substituable resources with linear response. *J. Math. Biology*, 51:458–490.

Barenblatt, G. I. (1996). *Scaling, self-similarity, and intermediate asymptotics*. Cambridge texts in applied mathematics. Cambridge University Press.

Britton, N. F. (2003). *Essential Mathematical Biology*. Springer-Verlag, London.

Chicone, C. (1987). The monotonicity of the period function for planar Hamiltonian vector fields. *J. Diff. Eq.*, 69:310–321.

Chung, K. (1974). *Elemntary probability theory with stochastic processes*. Springer Verlag.

Coddington, E. A. and Levinson, N. (1955). *Theory of ordinary differential equations*. McGraw-Hill.

Cronin, J. (1987). *Mathematical aspects of Hodgkin-Huxley neural theory*. Cambridge University Press.

Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press.

Dawkins, R. (1982). *The Extended Phenotype*. W.H. Freeman, Oxford.

Diekmann, O., Heesterbeek, H., and Britton, T. (2012). *Mathematical Tools for Understanding Infectious Disease Dynamics*. Princeton University Press, Princeton, NJ.

Diekmann, O. and Heesterbeek, J. (2000). *Mathematical Epidemiology of Infectious Diseases*. Wiley.

Diekmann, O., van Gils, S. A., Lunel, S. M. V., and Walther, H.-O. (1995). *Delay Equations. Functional, Complex, and Nonlinear Analysis*. Springer Verlag, New York.

Erban, R. and Othmer, H. G. (2004). From individual to collective behavior in bacterial chemotaxis. *SIAM J. Appl. Math.*, 65(2):361–391.

Erban, R. and Othmer, H. G. (2007). Taxis equations for amoeboid cells. *J. Math. Biology*, 54(6):847–885.

Feller, W. (1941). On the intergral equation of renewal theory. *Ann. Math. Stat.*, 12:243–267.

Feller, W. (1952). The parabolic differential equation and the associated semi-group of transformations. *Annals of Math.*, 55:468–519.

Feller, W. (1954). Diffusion processes in one dimension. *Trans. Am. Math. Soc.*, 77:1–31.

Feller, W. (1955). On differential operators and boundary conditions. *Comm. Pure Applied Math.*, 8:203–216.

Fisher, R. A. (1937). The advance of advantageous genes. *Ann. of Eugenics*, 7:355–369.

Freedman, H. (1980). *Deterministics Mathematical Models in Population Ecology.* Marcel Dekker.

Geritz, S. A. H. (2005). Resident-invader dynamics and the coexistence of similar strategies. *J. Math. Biology*, 50:67–82.

Geritz, S. A. H. and Kisdi, E. (2004). On the mechanistic underpinning of discrete-time population models with complex dynamics. *J. Theor. Biology*, 228:261–269.

Goel, N. S. and Richter-Dyn, N. (1974). *Stochastic Models in Biology.* Blackburn Press.

Golubitsky, M. and Schaeffer, D. G. (1985). *Singularities and Groups in Bifurcation Theory, Volume I.* Springer-Verlag, New York.

Gourley, S. A., So, J. W.-H., and Wu, J. H. (2004). Nonlocality of reaction-diffusion equation induced by delay: biological modeling and nonlinear dynamics. *J. Math. Sci.*, 124:5119–5153.

Graham, A. (1987). *Nonnegative matrices and applicable topics in linear algebra.* John Wiley & Sons.

Grasman, J. (1987). *Asymptotic Methods for Relaxation Oscillations and Applications.* Applied Mathematics Series. Springer-Verlag.

Grindrod, P. (1991). *Patterns and Waves in Reaction-Diffusion: Techniques and Applications.* Oxford University Press.

Gripenberg, G., Londen, S.-O., and Staffens, O. (1990). *Volterra Integral and Functional Equations.* Cambridge University Press, Cambridge, UK.

Gurney, W. S. C., Blythe, S. P., and Nisbet, R. M. (1980). Nicholson's blowflies revisited. *Nature*, 287:17–21.

Haccou, P., Jagers, P., and Vatutin, V. A. (2005). *Branching Processes: Variation, Growth, and Extinction of Populations.* Cambridge University Press, Cambridge, MA.

Hadeler, K. P. and Rothe, F. (1975). Travelling fronts in nonlinear diffusion equations. *J. Math. Biology*, 2:251–263.

Heesterbeek, J. A. P. and Metz, J. A. J. (1993). The saturating contact rate in marriage en epidemic models. *J. Math. Biology*, 31:529–539.

Hirsch, M. W., Smale, S., and Devaney, R. L. (2004). *Differential Equations, Dynamical Systems, and an Introduction to Chaos.* Academic Press.

Hodgkin, A. L. and Huxley, F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiology*, 117:500–544.

Horn, R. A. and Johnson, C. R. (1990). *Matrix Analysis.* Cambridge University Press.

Horstmann, D. (2003). From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. I. *Jahresber. Deutsch. Math.-Verein.*, 105(3):103–165.

Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography.* Princeton University Press.

Hutchinson, G. E. (1948). Circular causal systems in ecology. *Ann. New York Acad. Sciences*, 50:221–246.

Keener, J. and Sneyd, J. (2009). *Mathematical Physiology Vol. I and II*. Springer-Verlag.

Keller, E. F. and Segel, L. A. (1971). Travelling bands of chemotactic bacteria: a theoretical analysis. *J. Theor. Biology*, 30:235–248.

Kemeny, J. G. and Snell, K. L. (1963). *Finite Markov chains*. D. Van Nostrand Company, Princeton, NJ.

Kermack, W. and McKendrick, A. (1991a). Contributions to the mathematical theory of epidemics—i. *Bull. Math. Biol.*, 53(1-2):33–55.

Kermack, W. and McKendrick, A. (1991b). Contributions to the mathematical theory of epidemics—II. The problem of endemicity. *Bull. Math. Biol.*, 53(1-2):57–87.

Kevorkian, J. and Cole, J. D. (1996). *Multiple Scale and Singular Perturbation Methods*. Springer-Verlag.

Kimmel, M. and Axelrod, D. E. (2002). *Branching Processes in Biology*. Springer-Verlag, New York.

Kishimoto, K. and Weinberger, H. (1985). The spatial homogeneity of stable equilibria of some reaction diffusion systems on convex domains. *J. Diff. Equations*, 58:15–21.

Kolmogorov, A. N., Petrovskii, I., and Piscunov, N. (1937). Étude de léquations de la diffusion avec croissance de la quantité de matière et son application a un problème biologique. *Bull. Univ. Moscow, Ser. Internat., Sec. A*, 1:1–25.

Kuznetsov, Y. A. (2004). *Elements of applied bifurcation theory*. Springer-Verlag.

Lee, C. H. and Othmer, H. G. (2010). A multi-time-scale analysis of chemical reaction networks: I. Deterministic systems. *J. Math. Biology*, 60:387–450.

Lin, C. C. and Segel, L. A. (1998). *Mathematics Applied to Deterministic Problems in the Natural Sciences*. SIAM: Society for Industrial and Applied Mathematics.

Ludwig, D., Aronson, D. G., and Weinberger, H. F. (1979). Spatial patterning of the spruce budworm. *J. Math. Biology*, 8(3):217–258.

MacArthur, R. H. and Wilson, E. O. (1967). *The Theory of Island Biogeography*. Princeton University Press.

Matano, H. (1979). Asymptotic behavior and stability of solutions of semilinear diffusion equations. *Publ. Res. Inst. Math. Sci.*, 15:401–458.

McKendrick, A. G. (1926). Applications of mathematics to medical problems. *Proc. Edinburgh Math. Soc.*, 44:98–130.

Meinhardt, H. (1982). *Models of Biological Pattern Dormation*. Academic Press, available online as pdf at `http://www.eb.tuebingen.mpg.de/departments/former-departments/h-meinhardt/82-book/Bur82.htm`.

Metz, J. A. J., Mollison, D., and van den Bosch, F. (2000). The dynamics of invasion waves. In Dieckmann, U., Law, R., and Metz, J. A. J., editors, *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*, pages 482–512. Cambridge University Press, Cambridge, MA.

Metz, J. A. J., Nisbet, R. M., and Geritz, S. A. H. (1992). How should we define "fitness" for general ecological scenarios? *Trends Ecol. Evol.*, 7:198–202.

Mimura, M. and Nishiura, Y. (1979). Spatial patterns for an interaction-diffusion equation in morphogenesis. *J. Math. Biology*, 7:243–263.

Mimura, M., Nishiura, Y., and Yamaguti, M. (1979). Some diffusive prey and predator systems and their bifurcation problems. *Annals New York Academy of Sciences*, 316:490–510.

Mishchenko, E. F. and Rosov, N. K. (1980). *Differential Equations with Small Parameters and Relaxation Oscillations*. Plenum, New York.

Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., Massari, M., Salmaso, S., Tomba, G. S., Wallinga, J., Heijne, J., Sadkowska-Todys, M., Rosinska, M., and Edmunds, W. J. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.*, 5(e74).

Murray, J. D. (2002). *Mathematical Biology I: An Introduction*. Springer Verlag.

Othmer, H. G., Dunbar, S. R., and Alt, W. (1988). Models of dispersal in biological systems. *J. Math. Biology*, 26:263–398.

Othmer, H. G. and Hillen, T. (2002). The diffusion limit of transport equations. II. Chemotaxis equations. *SIAM J. Appl. Math.*, 62(4):1222–1250.

Othmer, H. G. and Schaap, P. (1998). Oscillatory cAMP signalling in the development of *Dictyostelium discoideum*. *Comments Theor. Biol.*, 5:175–282.

Perko, L. (2001). *Differential Equations and Dynamical Systems*. Springer-Verlag.

Perthame, B. (2007). *Transport Equations in Biology*. Birkhaüser Basel.

Radcliffe, J. and Rass, L. (2003). *Spatial Deterministic Epidemics*, volume 102 of *Math. Surveys and Monographs*. American Mathematical Society, Providence, RI.

Rao, C. V. and Arkin, A. P. (2003). Stochastic chemical kinetics and the quasi-steady-state assumption: application to the Gillespie algorithm. *J. Chem. Phys.*, 118(11):4999–5010.

Renardy, M. and Rogers, R. C. (1993). *An introduction to partial differential equations*. Springer-Verlag.

Rinaldi, S. and Muratori, S. (1992). Slow-fast limit cycle in predator-prey models. *Ecological modeling*, 61:287–308.

Schnell, S. and Maini, P. (2000). Enzyme kinetics at high enzyme concentration. *Bull. Math. Biol.*, pages 483–499.

Segel, L. A. (1984). *Modeling dynamic phenomena in molecular and cellular biology*. Cambridge University Press.

Smith, H. L. and Waltman, P. (1995). *The theory of the chemostat*. Cambridge University Press.

Smoller, J. (1983). *Shock Waves and Reaction-diffusion Equations*. Springer Verlag.

Stevens, A. (1995). Trail following and aggregation of myxobacteria. *J. Biol. Sys.*, 3:1059–1068.

Stroock, D. W. (1974). Some stochastic processes which arise from a model of the motion of a bacterium. *Z. Wahrscheinlichkeitstheorie*, 28:305–315.

van den Bosch, F., de Roos, A. M., and Gabriel, W. (1988). Cannibalism as a life boat mechanism. *J. Math. Biology*, 26:614–633.

Verhulst, F. (1996). *Nonlinear differential equations and dynamical systems*. Springer-Verlag.

Verhulst, F. (2005). *Methods and Applications of Singular Perturbations*. Springer-Verlag.

Vogel, S. (1988). *Life's devices: the physical world of animals and plants*. Princeton University Press.

Xue, C. and Othmer, H. G. (2009). Multiscale models of taxis-driven patterning in bacterial populations. *SIAM J. Appl. Math.*, 70:133–169.