

Mathematical methods to gain biological insight

Odo Diekmann & Bob Planqué
February 20, 2012

Contents

Chapter 1. Introduction	5
Chapter 2. Exploiting time scale differences: the quasi-steady-state approximation (QSSA)	6
2.1 Michaelis-Menten enzyme kinetics	6
2.2 Scaling	9
2.3 Cooperative reactions and the sigmoidal response	11
2.4 The force of infection in populations of variable size	13
2.5 Excitability	16
2.6 Encore: Enzyme kinetics	19
Chapter 3. Phase plane analysis of prey-predator systems	20
3.1 The story of d'Ancona and Volterra	20
3.2 The phase portrait of the Volterra-Lotka system	21
3.3 The effect of limitations in prey growth	22
3.4 Deriving the Holling type II functional response	24
3.5 The destabilising effect of a saturating functional response	25
3.6 The Rosenzweig-MacArthur model	26
Chapter 4. Movement in space	29
4.1 Flux	29
4.2 Various ways to motivate Fick's Law	31
4.3 Transport by diffusion	33
4.4 How to measure the diffusion coefficient	33
4.5 About sojourn times	35
4.6 How long does it take?	35
4.7 A remark on boundary conditions	37
Chapter 5. Linear diffusion	38
5.1 The fundamental solution	38
5.2 Separation of variables and spectral theory	40
5.3 The asymptotic speed of propagation	43
Chapter 6. Reaction-diffusion equations	46
6.1 Introduction	46
6.2 Stability criteria for uniform steady states	47
6.3 Scalar Reaction-Diffusion equations	50
6.4 Travelling waves for Reaction-Diffusion	59
6.5 The non-existence of patterns for scalar equations	63
6.6 Pattern formation: The Turing instability	64

Chapter 7. Chemotaxis	70
7.1 Introduction	70
7.2 Derivation of chemotaxis models	71
7.3 Understanding pattern formation in <i>E. coli</i> bacteria	77
Chapter 8. Adaptive Dynamics	85
8.1 Introduction	85
8.2 The pessimization principle	86
8.3 The principle of indifference	90
Chapter 9. Appendix	95
9.1 Bifurcation theory	95
Bibliography	103

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, if they wish in couples, 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



Note that we assume that the reaction $C \rightarrow P + E$ 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

$$\frac{ds}{dt} = -k_+se + k_-c, \quad (2.1.2)$$

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

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

$$\frac{dp}{dt} = +kc. \quad (2.1.5)$$

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

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

with matrix

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

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 $t \rightarrow \infty$ to an eigenvector of M corresponding to eigenvalue zero. One such eigenvector is

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

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

$$\begin{pmatrix} e(t) \\ c(t) \end{pmatrix} \xrightarrow{t \rightarrow \infty} \frac{e_{\text{tot}}}{k_+s + k_- + k} \begin{pmatrix} k_- + k \\ k_+s \end{pmatrix}, \quad (2.1.9)$$

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

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

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

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

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

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

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

$$k_+s + k_- + k \gg \frac{k_+ke_{\text{tot}}}{k_+s + k_- + k}. \quad (2.1.13)$$

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,

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

and changes in s occur at the time scale

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

So if we require that

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

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,

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

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 righthand side of (2.1.15) by $k/(k_+s + k_- + k)$ the left hand side is still much bigger.

See [52], 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

$$\frac{dp}{dt} = -\frac{ds}{dt}. \quad (2.1.17)$$

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

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

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

$$\frac{ds}{dt} = -\frac{V_m s}{K_m + s} \quad (2.1.19)$$

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

$$t = \frac{1}{k_+ e_0} t^*, \quad (2.2.1)$$

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

$$s(t) = s(0) s^*(t^*). \quad (2.2.2)$$

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

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

and as a consequence

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

Now, noting that

$$\frac{d}{dt} = \frac{dt^*}{dt} \frac{d}{dt^*} = k_+ e_0 \frac{d}{dt^*}, \quad (2.2.5)$$

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

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

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

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

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

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

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

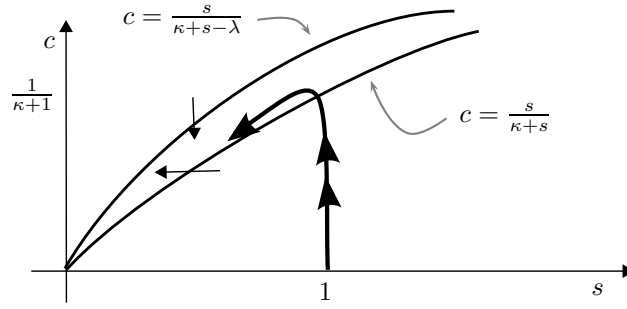


FIGURE 2.1. Dynamics of (2.2.11). The solution shoots up from the s -axis until it approaches $c = 1/(\kappa + 1)$, after which it is trapped between the isoclines and converges towards the origin.

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

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

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 $c \approx \frac{s}{\kappa + s}$, then necessarily $\frac{dc}{dt}$ is very large, so c changes rapidly, i.e., at the time scale ε . For fixed s , the equation 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

$$\frac{ds}{dt} = -\frac{\lambda s}{\kappa + s},\quad (2.2.12)$$

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

$$t = \varepsilon t^*.\quad (2.2.13)$$

(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

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

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., [44, 24, 34, 60]). 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

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

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

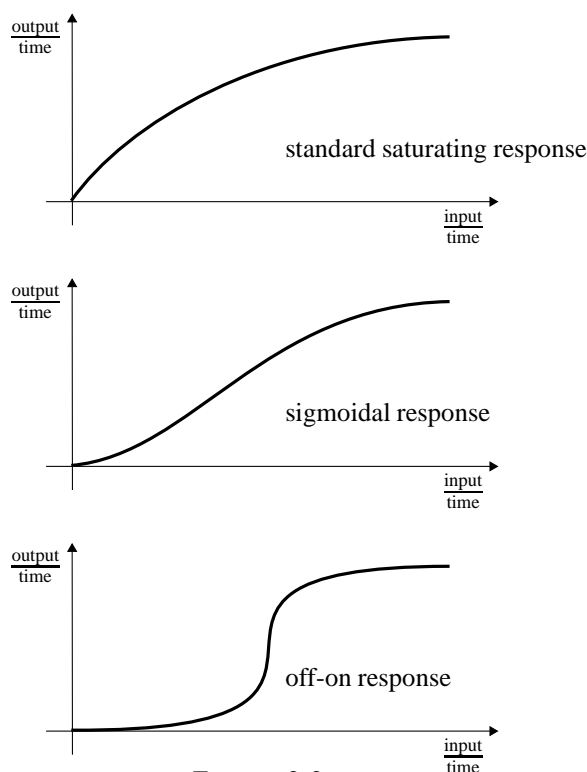


FIGURE 2.2.

(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 $t \rightarrow 0$ in the outer solution to the limit $t^* \rightarrow \infty$ 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 [8] for a fully worked treatment of the above problem.

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 [38].

In (2.1.11), we see that the velocity of the $S \rightarrow P$ 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.

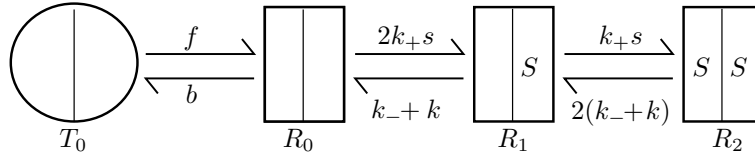
How can a sigmoidal response arise? What kind of biochemical mechanism can be responsible?

Some enzymes consist of several identical subunits (e.g., haemoglobin has four binding sites for O_2). They are called 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's 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

$$\frac{dT_0}{dt} = -fT_0 + bR_0, \quad (2.3.1)$$

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

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

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

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

$$\frac{dp}{dt} = +kR_1 + 2kR_2. \quad (2.3.6)$$

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

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

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:

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

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 [23, 29] to arrive at the desired conclusion. Equivalently, we can refer to the general theory of continuous time Markov Chains [33].

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

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

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

$$\frac{dp}{dt} = 2kY e_{\text{tot}} = -\frac{ds}{dt} \quad (2.3.10)$$

(can you interpret the middle term? is this what you would expect?). We can conclude that the way in which the velocity of the $S \rightarrow P$ 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 [53]]

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

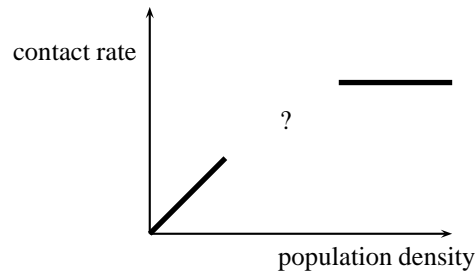
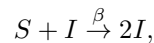


FIGURE 2.3.

just infected individual is infectious right away, we have the reaction scheme



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 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 [27], 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

$$N(t) = X(t) + 2P(t), \quad (2.4.1)$$

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

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

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

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 $\nu := \frac{\rho}{\sigma}$,

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

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

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

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

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

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

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

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

$$\frac{dS}{dt} = -\beta M, \quad (2.4.10)$$

$$\frac{dI}{dt} = \beta M, \quad (2.4.11)$$

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) 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) is given by

$$\begin{aligned} \bar{S}_1 &= \frac{S}{N} \bar{X}, & \bar{I}_1 &= \frac{I}{N} \bar{X}, \\ \bar{S}_2 &= \left(\frac{S}{N}\right)^2 \bar{P}, & \bar{M} &= 2\frac{S}{N}\frac{I}{N}\bar{P}, & \bar{I}_2 &= \left(\frac{I}{N}\right)^2 \bar{P}, \end{aligned} \quad (2.4.12)$$

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

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

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 $N \rightarrow 0$,
- (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 [14] for elaborations, remarks on generalisations and additional exercises.

2.5 Excitability

Consider the system

$$\varepsilon \frac{dv}{dt} = f(v) - w, \quad (2.5.1)$$

$$\frac{dw}{dt} = v - \gamma w, \quad (2.5.2)$$

$$(2.5.3)$$

with f a *cubic*, specified graphically in Figure 2.4.

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

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

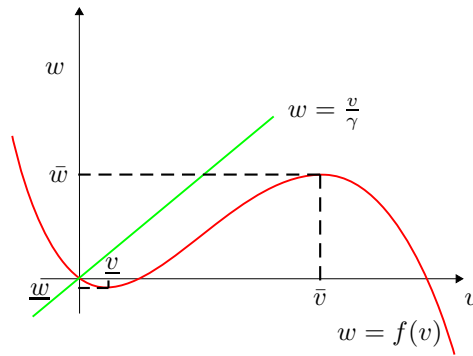


FIGURE 2.4.

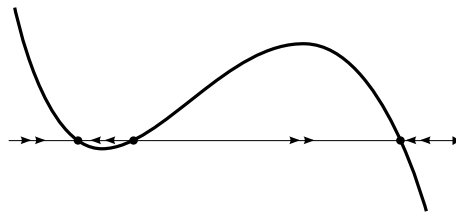


FIGURE 2.5.

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 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. Indeed, (2.5.1) is called the Fitzhugh-Nagumo system and it was derived as a somewhat caricatural simplification of the Hodgkin-Huxley equations, in which the space variable has been neglected. This corresponds to the space clamped version of the famous Hodgkin-Huxley experiments (see [13, 31]).

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

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

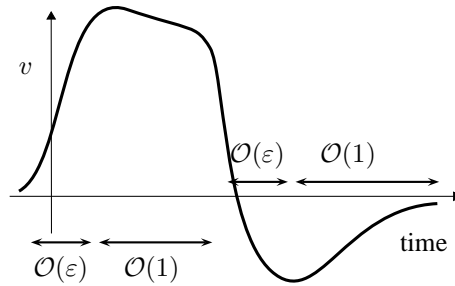


FIGURE 2.6.

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 (2.5.1) into

$$\varepsilon \frac{dv}{dt} = f(v) - w + w_0, \quad (2.5.4)$$

$$\frac{dw}{dt} = v - \gamma w. \quad (2.5.5)$$

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

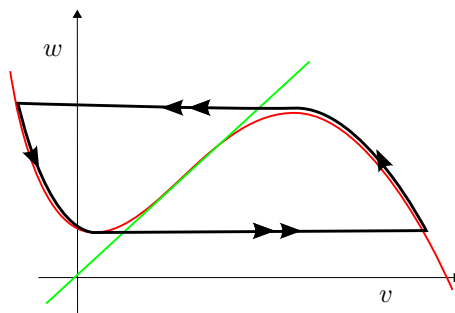


FIGURE 2.7.

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 in reality!). The mathematical jargon for such periodic behaviour with alternating slow and fast phases is “relaxation oscillation”, see [24].

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).

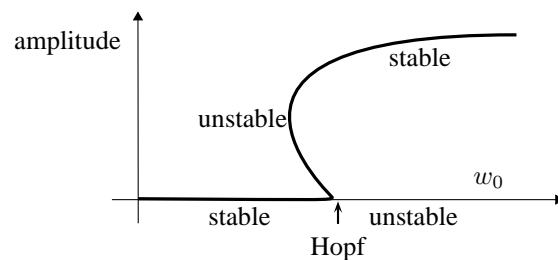
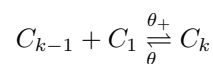


FIGURE 2.8.

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,



Assume, additionally, that θ_{\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

$$\frac{dv}{dt} = av - bvp, \quad (3.1.1)$$

$$\frac{dp}{dt} = dvp - cp. \quad (3.1.2)$$

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.

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.

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 Jacobi 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 $L : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ defined by

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

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

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

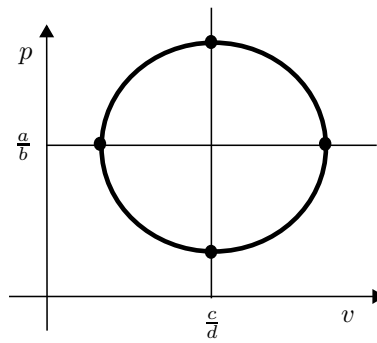


FIGURE 3.1.

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 $\varepsilon \rightarrow 0$ and $\varepsilon \rightarrow \theta$, we obtain a closed curve.

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 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

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

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*

$$\tilde{v} = \frac{a}{e}. \quad (3.3.2)$$

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

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

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

for parameter combinations such that

$$-c + d\frac{a}{e} < 0. \quad (3.3.5)$$

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,

$$-c + d\frac{a}{e} > 0, \quad (3.3.6)$$

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 $V : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ by

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

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 set¹ 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

$$v \mapsto \frac{bv}{1 + b\beta v} \quad (3.4.1)$$

does indeed approach a limit $\frac{1}{\beta}$ for $v \rightarrow \infty$. 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

¹ (\bar{v}, \bar{p}) belongs to the ω -limit set of $(v(0), p(0))$ iff there is a sequence $t_n \rightarrow \infty$ such that $(v(t_n), p(t_n)) \rightarrow (\bar{v}, \bar{p})$; 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 $t_n \rightarrow -\infty$ one speaks about the α -limit set; note that α is the first and ω the last character in the Greek alphabet.

types of predators, those searching for prey and those busy handling a prey caught earlier:

$$\frac{dv}{dt} = av - bvs, \quad (3.4.2)$$

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

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

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.

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

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

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

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

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

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

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 [21]).

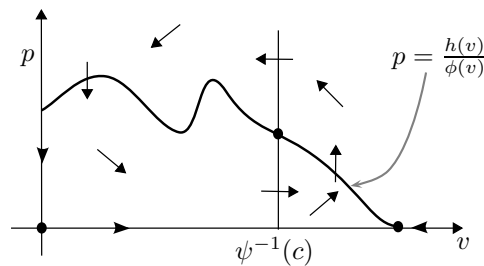


FIGURE 3.2.

Exercise 3.6.1.

- (i) Demonstrate that there is 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 [2, 49, 28] (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 [51].

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

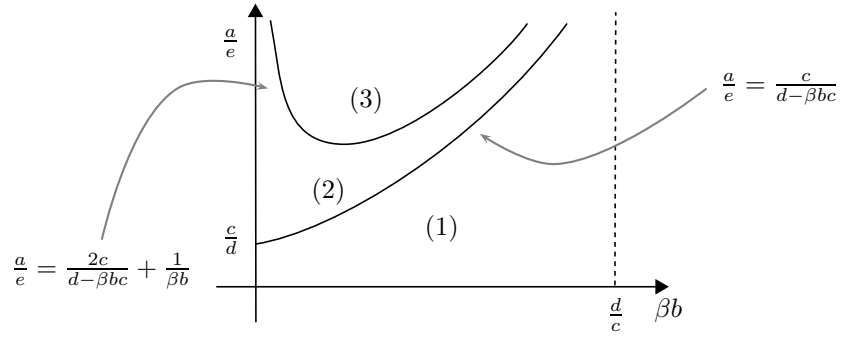


FIGURE 3.3.

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

$$J(t, x) = c(x)u(t, x), \quad (4.1.1)$$

as is indeed also suggested by the dimensional identity

$$\frac{\text{number}}{\text{time}} = \frac{\text{length}}{\text{time}} \cdot \frac{\text{number}}{\text{length}}. \quad (4.1.2)$$

Equally clearly,

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

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,

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

Hence,

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

and as this holds for arbitrary a and b , the integrand must be zero (see Lemma of DuBois-Reymond [39]), and so in combination with (4.1.1) we arrive at the *conservation law*

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

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

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

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

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

(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*

$$J = -d\nabla u, \quad (4.1.9)$$

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*

$$\frac{\partial u}{\partial t} = d\Delta u, \quad (4.1.10)$$

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

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

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).

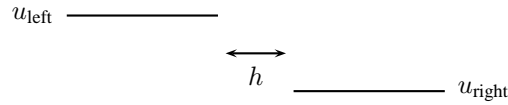


FIGURE 4.1.

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

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

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

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

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

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

$$\int_{\Omega} \Delta u = \int_{\partial\Omega} \frac{\partial u}{\partial n}. \quad (4.1.13)$$

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 $h \rightarrow 0$ while assuming that

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

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 $c \rightarrow \infty$. 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

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

(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 $\binom{n}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k} = \binom{n}{k} \left(\frac{1}{2}\right)^n$.)

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 [11], 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

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

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

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

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

Substituting these expressions in (4.2.1) we find

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

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

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

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

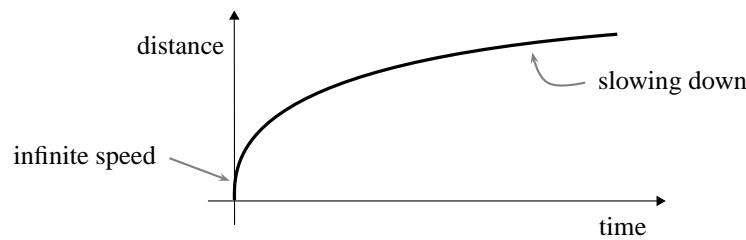


FIGURE 4.2.

we arrive at

$$\frac{\partial u}{\partial t} = d \frac{\partial^2 u}{\partial x^2}. \quad (4.2.6)$$

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?

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 · 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 [61, 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

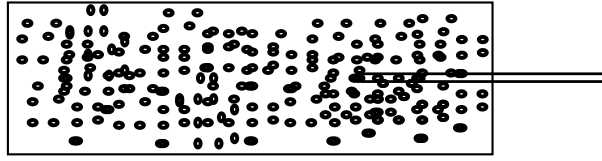


FIGURE 4.3.

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

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

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

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

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

Exercise 4.4.1. Derive the expression

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

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 $0.2 \text{ cm}^2/\text{hr}$.

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

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

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 book-keeping considerations underlying the identity.

$$N(t) = \int_0^\infty F(t - \sigma)P(T \geq \sigma) d\sigma. \quad (4.5.1)$$

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

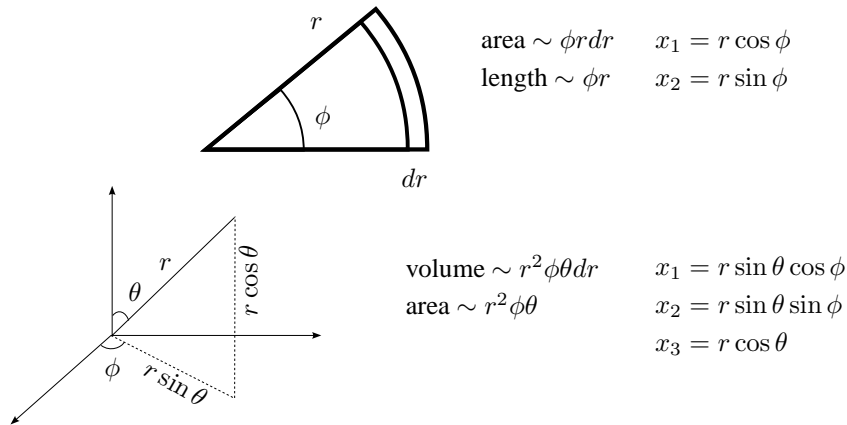


FIGURE 4.4.

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

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

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 .

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*

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

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

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

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 [17, 18, 19].

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

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

subject to far out boundary conditions

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

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 $t > 0$,

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

(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

$$u = f(t, x, d). \quad (5.1.4)$$

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 $(\text{length})^2/\text{time}$. So the combination x/\sqrt{dt} is dimensionless.

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

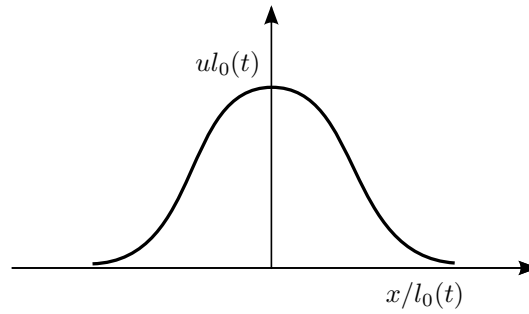


FIGURE 5.1. In the right time-dependent variables $(x/l_0(t))$ and $ul_0(t)$ the fundamental solution of the diffusion equation has a unique profile for all time.

ϕ . The conservation equation (5.1.3) now yields

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

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 [7].

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

$$\frac{d^2 \phi}{d\xi^2} + \frac{\xi}{2} \frac{d\phi}{d\xi} + \frac{\phi}{2} = 0. \quad (5.1.5)$$

Integrating once, we find that

$$\frac{d\phi}{d\xi} + \frac{\xi}{2}\phi = \text{constant}. \quad (5.1.6)$$

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

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

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, $u(t, \cdot)$ 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 [7, 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*

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

with so-called no-flux boundary conditions

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

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

$$\phi''(x) = (\lambda - r)\phi(x) \quad (5.2.3)$$

$$\phi'(0) = 0 = \phi'(L). \quad (5.2.4)$$

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 $k = 0, 1, 2, \dots$,

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

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 $t \rightarrow \infty$. 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*

$$u(0, x) = u_0(x), \quad (5.2.6)$$

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

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

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

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

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):

$$u(t, 0) = 0 = u(t, L). \quad (5.2.9)$$

It follows that now

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

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.

5.3 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

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

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

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

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

$$u(t, x) = w(x \cdot \nu - ct), \quad (5.3.3)$$

where $w : \mathbb{R} \rightarrow \mathbb{R}$ 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 $\lim_{t \rightarrow \infty} u(t, x) = \infty$, while for fixed t we have $\lim_{|x| \rightarrow \infty} u(t, x) = 0$. To find out where, roughly, the transition from 0 to ∞ is located, we can consider $\lim_{t \rightarrow \infty} u(t, x)$ under various assumptions about how fast $|x(t)| \rightarrow \infty$ as $t \rightarrow \infty$.

- (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?

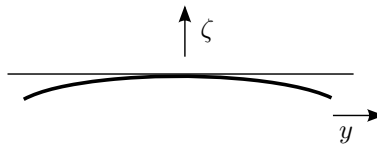


FIGURE 5.2.

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 $t \rightarrow \infty$. 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

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

We write this as $\alpha = m(t) + \theta + \mathcal{O}\left(\frac{\log^2(t)}{t}\right)$, with $\theta = -\sqrt{\frac{d}{r}} \log(\psi)$. 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 ($\psi \rightarrow \infty$, $\theta \rightarrow -\infty$), to outside ($\psi \rightarrow 0$, $\theta \rightarrow \infty$).

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 [25], but also the book by Barenblatt devoted to self-similarity and intermediate asymptotics [7]. For each c the equation is invariant with respect to a group of transformations

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

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

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

where

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{pmatrix} \quad (6.1.2)$$

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

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

If the space variable

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

has m components, then

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

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 $k = 1$, $m = 1$, $\Omega = \text{interval}$, 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

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

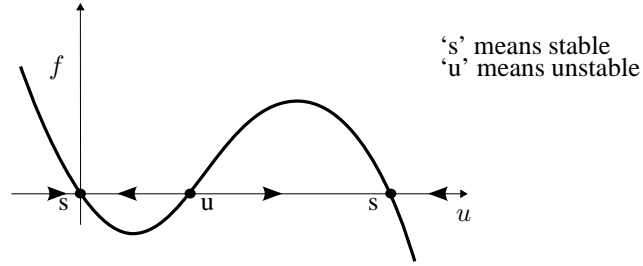


FIGURE 6.1.

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
- 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

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

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 $t \rightarrow \infty$). The connection between (6.2.2) and the eigenvalue problem

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

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

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

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

$$\left. \frac{\partial u}{\partial n} \right|_{\partial\Omega} = 0 \quad (6.2.5)$$

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

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

$$\left. \frac{\partial v}{\partial n} \right|_{\partial\Omega} = 0 \quad (6.2.7)$$

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

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

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

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

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

$$\Delta\phi = -\mu^2\phi \quad (6.2.10)$$

$$\left. \frac{\partial\phi}{\partial n} \right|_{\partial\Omega} = 0 \quad (6.2.11)$$

and, in addition,

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

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

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

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

$$\mu = \frac{k\pi}{L}, \quad k = 0, 1, 2, \dots \quad (6.2.14)$$

This generalises in the sense that for

$$\mu_0 = 0 \quad (6.2.15)$$

we have a solution $\phi = \text{constant}$ and that there exist μ_1, μ_2, \dots , with $\mu_{i+1} > \mu_i$ for $i = 0, 1, 2, \dots$, 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., [50]):

- 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

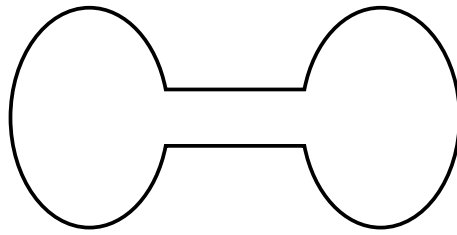


FIGURE 6.2.

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:

- 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 [35]
 - (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 [41])

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

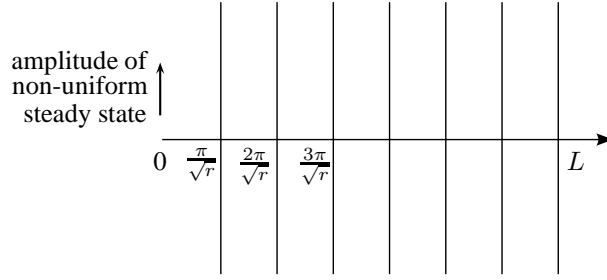


FIGURE 6.3.

steady states. To demonstrate this, we introduce the Lyapunov function

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

where

$$F(w) := \int_0^w f(\sigma) d\sigma \quad (6.3.2)$$

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

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

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

(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

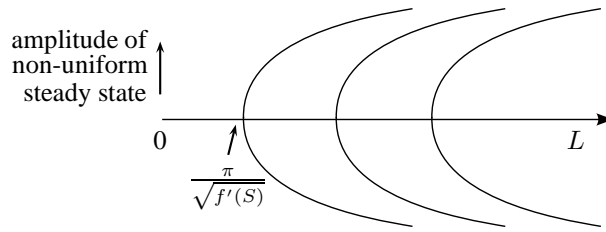


FIGURE 6.4.

parameter, then a non-uniform steady state only exists if

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

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 *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

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

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).

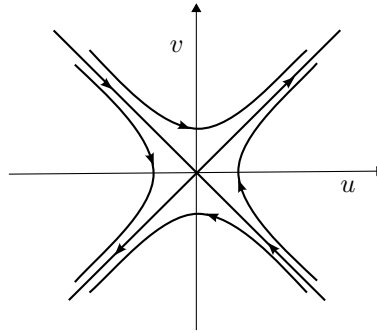


FIGURE 6.5.

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.

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

$$\begin{aligned} u_{xxx} + f(u) &= 0 \\ u_x(0) = 0 &= u_x(L) \end{aligned}$$

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.

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

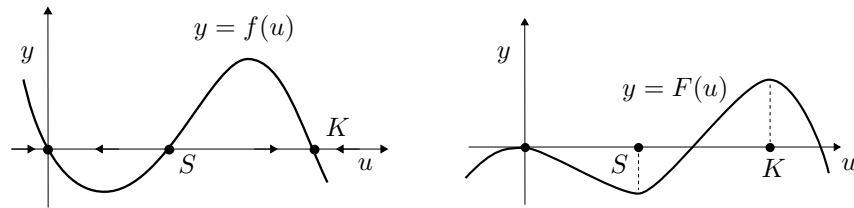


FIGURE 6.6.

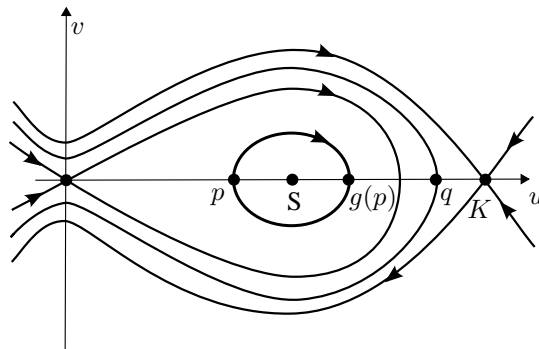


FIGURE 6.7.

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).

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

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

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. [10]). 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

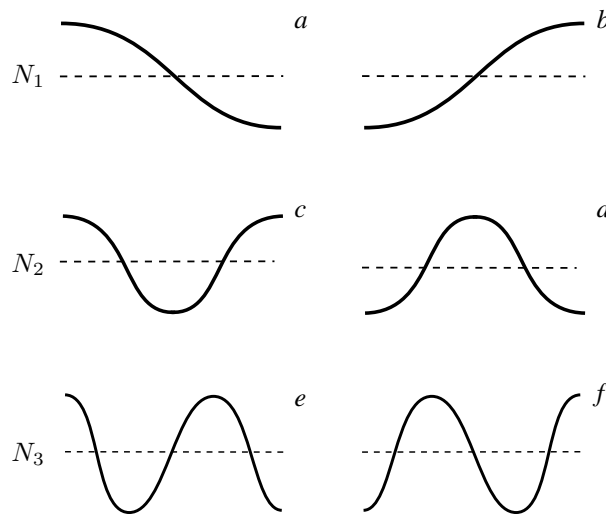


FIGURE 6.8.

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 $\frac{k}{2}$ 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

¹Above 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.

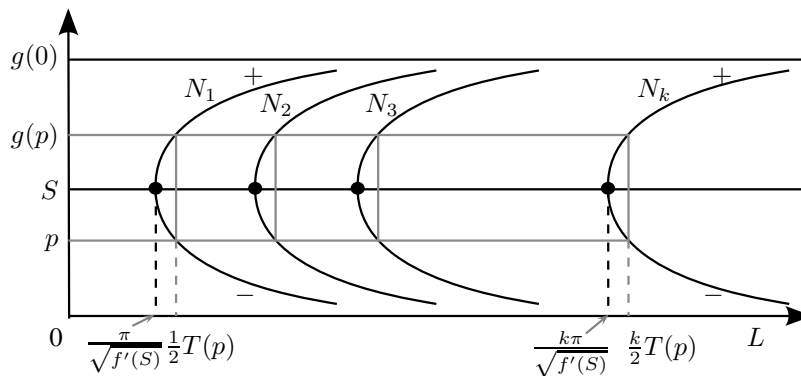


FIGURE 6.9.

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 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

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

$$u_x(0) = 0 = u_x(L), \quad (6.3.10)$$

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

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

Show that the extension is a solution for parameter value kL , $k = 1, 2, 3, \dots$. 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.

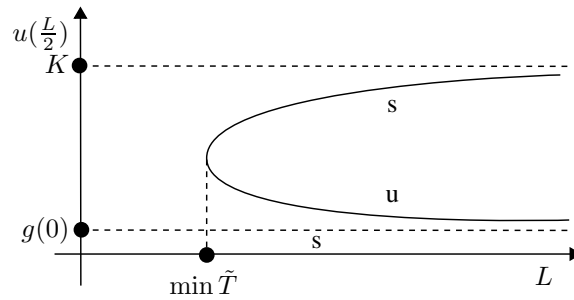


FIGURE 6.10.

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

$$u(0) = 0 = u(L) \quad (6.3.11)$$

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

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

and, as movement slows down near saddle points, necessarily

$$\begin{aligned} \lim_{q \downarrow g(0)} \tilde{T}(q) &= \infty \\ \lim_{q \uparrow K} \tilde{T}(q) &= \infty \end{aligned}$$

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

$$\begin{aligned} L < \min \tilde{T} & \text{ no solution} \\ L > \min \tilde{T} & \text{ at least two solutions} \end{aligned}$$

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. [4, 5, 40], 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 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

$$u_t = du_{xx} + ku(1 - u), \quad x \in \mathbb{R} \quad (6.4.1)$$

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

$$u_t = u_{xx} + u(1 - u). \quad (6.4.2)$$

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 [20]. It was simultaneously (and presumably independently) studied by, as Aronson [3] put it, the famous troika of Kolmogorov, Petrovskii and Piscunov [36]. The bistable equation in non-dimensional form is given by

$$u_t = u_{xx} + u(u - a)(1 - u), \quad x \in \mathbb{R} \quad (6.4.3)$$

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

$$-cw' = w'' + w(1 - w) \quad (6.4.4)$$

$$\lim_{z \rightarrow -\infty} w(z) = 1, \quad \lim_{z \rightarrow \infty} w(z) = 0 \quad (6.4.5)$$

Writing (6.4.4) in phase plane form,

$$w' = v \quad (6.4.6)$$

$$v' = -cv - w(1 - w) \quad (6.4.7)$$

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

$$\begin{pmatrix} 0 & 1 \\ -1 + 2w & -c \end{pmatrix} \quad (6.4.8)$$

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

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

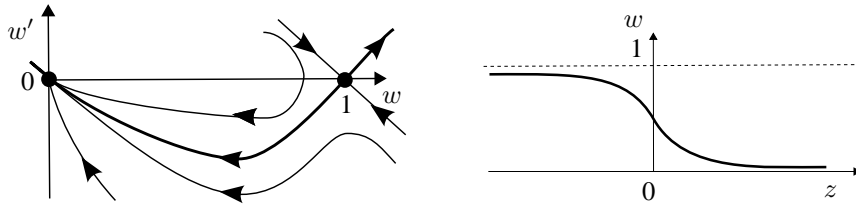


FIGURE 6.11.

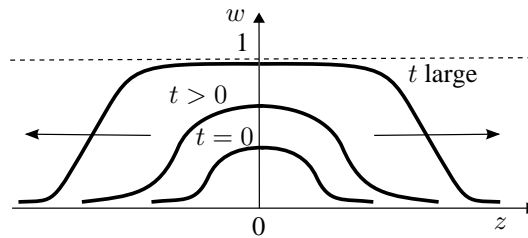


FIGURE 6.12.

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 (6.4.6) 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 [36, 26].

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 [36] 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

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

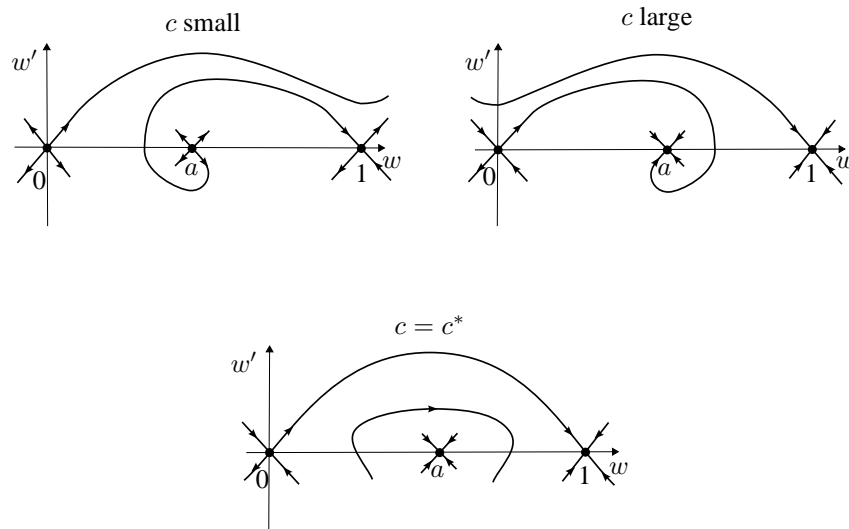


FIGURE 6.13.

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

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

Writing this as a two-dimensional system

$$w' = v \tag{6.4.9}$$

$$v' = -cv + f(w) \tag{6.4.10}$$

we have the Jacobian

$$\begin{pmatrix} 0 & 1 \\ -f'(w) & -c \end{pmatrix} \tag{6.4.11}$$

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.

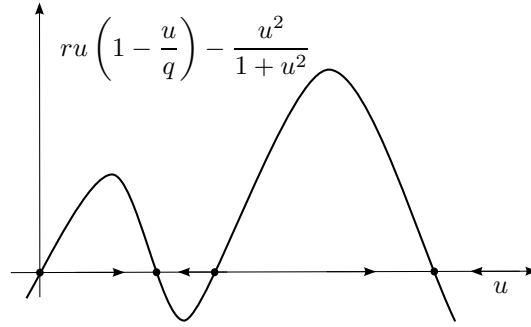


FIGURE 6.14.

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

$$w'' + cw' + f(w) = 0. \quad (6.4.12)$$

Equation (6.4.12) is invariant under $z \mapsto -z$, $c \mapsto -c$, so we need to choose the behaviour of solutions for $z \rightarrow \pm\infty$ to be able to determine the true direction of travelling wave solutions. We focus on solutions w which tend to 1 for $z \rightarrow -\infty$ and to 0 for $z \rightarrow \infty$. In particular, then also $w'(z) \rightarrow 0$ for $z \rightarrow \pm\infty$. Multiplying (6.4.12) by w' and integrating over \mathbb{R} , we find

$$0 = \int_{-\infty}^{\infty} [w'' + cw' + f(w)]w' dz \quad (6.4.13)$$

$$= 0 + c \int_{-\infty}^{\infty} w'^2 dz + \int_{-\infty}^{\infty} f(w)w' dz \quad (6.4.14)$$

$$= c \int_{-\infty}^{\infty} w'^2 dz + \int_1^0 f(w)dw \quad (6.4.15)$$

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 [45] for a detailed discussion of this much-studied problem.

6.5 The non-existence of patterns for scalar equations

We will now show that non-trivial equilibrium solutions ('patterns') to scalar RD 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

$$\begin{cases} u_t = u_{xx} + f(u) & \text{on } [0, L] \\ u_x(t, 0) = u_x(t, L) = 0 \end{cases} \quad (6.5.1)$$

Let $U(x)$ be a non-trivial equilibrium solution, i.e., $U(x) \neq U_0$. The eigenvalue problem for V , which describes the linear stability of U , is given by

$$\begin{cases} V_{xx} + [\frac{\partial f}{\partial u}(U(x)) - \lambda]V = 0 & \text{on } [0, L] \\ V_x(0) = V_x(L) = 0 \end{cases} \quad (6.5.2)$$

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, $-\infty < \dots < \lambda_2^N < \lambda_1^N < \lambda_0^N$, 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 (6.5.2) 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 , $i = 0, 1, \dots$, 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 6.5.1. *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 $\lambda < \mu$. \square

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 (6.5.2) 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

$$\begin{cases} u_t = u_{xx} + f(u) & \text{on } [0, L] \\ u(0) = u(L) = 0 \end{cases} \quad (6.5.3)$$

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 6.5.1 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.

6.6 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 6.5.1 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

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

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

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 Jacobi 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

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

Exercise 6.6.2. By assumption the inequalities

$$m_{11} + m_{22} < 0, \quad (6.6.4)$$

$$m_{11}m_{22} - m_{12}m_{21} > 0$$

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

$$\lambda^2 + \theta_1\lambda + \theta_2 = 0 \quad (6.6.5)$$

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

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

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

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

Compute that the minimum value equals

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

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

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

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).

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 $\text{sign } \bar{v}_1 = \text{sign } \bar{v}_2$. 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 ,

$$\frac{d_2}{d_1} = \frac{m_{22}\mu^2 - m_{11}m_{22} + m_{12}m_{21}}{\mu^2(\mu^2 - m_{11})} \tag{6.6.12}$$

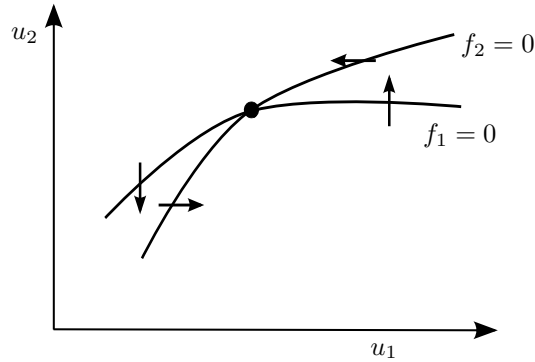
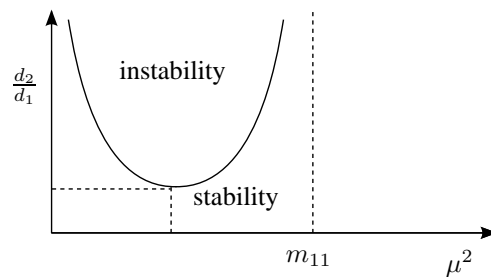


FIGURE 6.15. Local phase portrait of the kinetic system in Exercise 6.6.8.

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.

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

$$0 \leq x \leq L$$

and impose the no-flux boundary conditions

$$u_x(0) = 0 = u_x(L)$$

Derive that necessarily

$$\mu \in \left\{ \frac{k\pi}{L} : k = 1, 2, \dots \right\}$$

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}^{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).

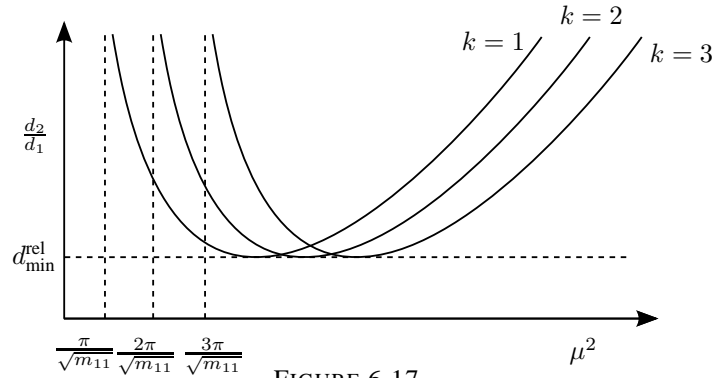


FIGURE 6.17.

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 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 [42].

The system of reaction diffusion equations

$$\frac{\partial A}{\partial t} = 1 + R \frac{A^2}{H} - A + \frac{\partial^2 A}{\partial x^2}, \quad (6.6.13)$$

$$\frac{\partial H}{\partial t} = Q(A^2 - H) + P \frac{\partial^2 H}{\partial x^2} \quad (6.6.14)$$

provided with no-flux boundary conditions

$$\frac{\partial A}{\partial x}(t, 0) = 0 = \frac{\partial A}{\partial x}(t, L), \quad (6.6.15)$$

$$\frac{\partial H}{\partial x}(t, 0) = 0 = \frac{\partial H}{\partial x}(t, L) \quad (6.6.16)$$

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 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 $(\text{constant} + H)$ 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 Jacobi 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

$$\frac{R-1}{R+1} < Q \quad (6.6.17)$$

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

$$p\left(\frac{k\pi}{L}\right)^4 + \left(Q - P\frac{R-1}{R+1}\right)\left(\frac{k\pi}{L}\right)^2 + Q > 0 \quad (6.6.18)$$

but unstable when the reverse inequality holds.

- (ii) Deduce from (6.6.18) that a necessary condition for instability is

$$Q < P\frac{R-1}{R+1} \quad (6.6.19)$$

- (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.

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* 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 in *Dictyostelium* is presented in Figure 7.1.

Some of the ingredients at play are clear:

- the cells move in a biased random walk due to both random movement and moving up (or down) the gradient of a chemoattractant
- the chemoattractant is produced, with two different cases to consider. Either an external source produces chemical attractants, such as when leukocytes track the signal molecules produced by bacteria that have entered our bodies, or the cells sensitive to the chemoattractant produce it themselves.
- the bacteria consume nutrients, such as succinate.

It is not at all clear, however, how these different parts work together to induce patterns, or how we may be able to predict the time and length scales that characterize these patterns. The sheer variety of these patterns also indicates that we may need to look for different kinds of explanations in different situations. Experiments on these model species, such as the one shown in Figure 7.1, have yielded some insights into why patterns are formed, but it has proved difficult to get good intuition that explains why one sees the particular observed pattern. This is thus a clear case in which mathematical modelling may help to understand what causes pattern formation.

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



FIGURE 7.1. Spiraling waves of cyclic AMP, the chemoattractant used by colonies of *Dictyostelium* bacteria

timespan δt . If we denote this net flow out of Ω by a flux $J(t, x)$ then we can write

$$\int_{\Omega} u(t + \delta t, x) dx = \int_{\Omega} u(t, x) dx - \delta t \int_{\partial\Omega} J(t, x) dS \quad (7.1.1)$$

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

$$u_t + \nabla \cdot J = 0. \quad (7.1.2)$$

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:

$$u_t + \nabla \cdot (d\nabla u - u\chi(S)\nabla S) \quad (7.1.3)$$

known as a Keller-Segel equation [32]. Here the chemotactic flux J_c due to attraction by the chemical 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. 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.

7.2 Derivation of chemotaxis models

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

$$\frac{\partial u^+}{\partial t} + s \frac{\partial u^+}{\partial x} = -\lambda u^+ + \lambda u^- \quad (7.2.1)$$

$$\frac{\partial u^-}{\partial t} - s \frac{\partial u^-}{\partial x} = -\lambda u^- + \lambda u^+ \quad (7.2.2)$$

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

$$\frac{\partial u}{\partial t} + \frac{\partial j}{\partial x} = 0 \quad (7.2.3)$$

$$\frac{\partial j}{\partial t} + 2\lambda j = -s^2 \frac{\partial u}{\partial x} \quad (7.2.4)$$

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 $\lambda \rightarrow \infty$, $s \rightarrow \infty$, while keeping

$$\frac{s^2}{\lambda} =: d \quad (7.2.5)$$

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

$$\varepsilon^2 \frac{\partial u}{\partial \tau} + \varepsilon \nabla_{\xi} \cdot (vu) = -\lambda u + \lambda \int_V T(v, v') u(\tau, \xi, v') dv' \quad (7.2.6)$$

The second step is to write u as an asymptotic expansion

$$u(\tau, \xi, v) = \sum_{i=0}^k \varepsilon^i u_i(\tau, \xi, v) + \mathcal{O}(\varepsilon^{k+1}) \quad (7.2.7)$$

Let us denote

$$\mathcal{L}\phi(v) = -\lambda\phi(v) + \lambda \int_V T(v, v') \phi(\tau, \xi, v') dv'$$

for functions $v \in L^2(V)$. 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 (7.2.7) into (7.2.6) and grouping the terms in the resulting equation by orders of ε , we find

$$\mathcal{O}(\varepsilon^0) : \quad \mathcal{L}u_0 = 0 \quad (7.2.8)$$

$$\mathcal{O}(\varepsilon^1) : \quad \mathcal{L}u_1 = v \cdot \nabla u_0 \quad (7.2.9)$$

$$\mathcal{O}(\varepsilon^2) : \quad \mathcal{L}u_2 = \frac{\partial u_0}{\partial t} + v \cdot \nabla u_1 \quad (7.2.10)$$

⋮

$$\mathcal{O}(\varepsilon^i) : \quad \mathcal{L}u_i = \frac{\partial u_{i-2}}{\partial t} + v \cdot \nabla(u_{i-1}) \quad 3 \leq i \leq k \quad (7.2.11)$$

The properties of T imply that 0 is a simple eigenvalue of $\mathcal{L}u = -\lambda u + \lambda \int_V T(\cdot, v') u(v') dv'$ 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 (7.2.9) satisfies this orthogonality condition, which reads

$$\int_V (v \cdot \nabla u_0) dv = 0. \quad (7.2.12)$$

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 (7.2.10) 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 ,

$$\frac{\partial u_0}{\partial t} + v \cdot \nabla (\mathcal{F}(v \cdot \nabla u_0)) = 0 \quad (7.2.13)$$

which can be written in the more familiar form

$$\frac{\partial u_0}{\partial t} - \nabla \cdot (d \nabla u_0) = 0 \quad (7.2.14)$$

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 straightforward generalisation of the diffusion constant (7.2.5) 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 (7.2.7) of u , we also introduce an expansion for T . Let us assume that the influence of S only occurs in the order ε term:

$$T(v, v', S) = T_0(v, v') + \varepsilon T_1(v, v', S) \quad (7.2.15)$$

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

$$\mathcal{L}_0 u = -\lambda u + \lambda \int_V T_0(\cdot, v') u(v') dv'$$

We continue by substituting the expansion for u . The u_i again satisfy a coupled set of equations analogous to (7.2.8)–(7.2.11), 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

$$\int_V \left(\frac{\partial u_0}{\partial t} + (v \cdot \nabla) \mathcal{F}_0(v \cdot \nabla u_0) - \lambda (v \cdot \nabla) \mathcal{F}_0 \left(\int_V T_1(v, v') dv' u_0 \right) \right) dv - \lambda_0 \int_V \int_V T_1(v, v', S) u_1(v') dv' dv = 0 \quad (7.2.16)$$

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

$$d = \frac{s^2}{\lambda n}, \quad v_c = \frac{1}{|S^{n-1}|} \int_V \int_V v T_1(v, v', S) dv dv'$$

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 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

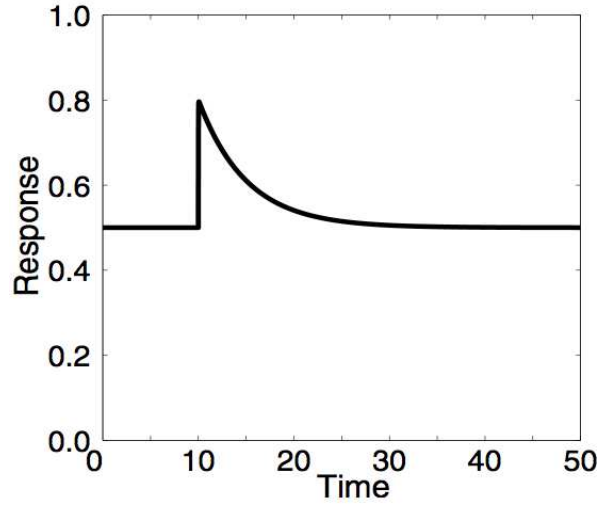


FIGURE 7.2. Typical dynamics of the internal excitation variable y_1 as the bacterium passes a sudden increase of chemoattractant. Note that the bacterium is first more excitable, but that this excitability slowly decreases back to a rest state. Adapted from [15].

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

$$\frac{dy}{dt} = h(y, S) \quad (7.2.17)$$

or more specifically,

$$\tau_e \frac{dy_1}{dt} = g(S(x)) - (y_1 + y_2) \quad (7.2.18)$$

$$\tau_a \frac{dy_2}{dt} = g(S(x)) - y_2 \quad (7.2.19)$$

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 a 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 straightforward. Now $u(t, x, y, v)$ satisfies

$$\frac{\partial u}{\partial t} + \nabla_x \cdot (vu) + \nabla_y \cdot (hu) = -\lambda(y)u + \lambda(y) \int_V T(v, v', y)u(t, x, v', y)dv' \quad (7.2.20)$$

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 $U(x, t) = \int \int u(t, x, y, v) dy dv$, can now be derived, and is indeed a classical Keller-Segel equation for large times

$$\frac{\partial U}{\partial t} = \nabla \cdot \left(\frac{s^2}{\lambda_0} \nabla U - \left[\frac{bs^2\tau_a g'(S(x))}{\lambda_0(1 + 2\lambda_0\tau_a)(1 + 2\lambda_0\tau_e)} \right] U \nabla S \right). \quad (7.2.21)$$

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 (7.2.21).

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. [56, 55]. For more information and much detail on recent developments on velocity-jump processes and their relation to chemotaxis models, see [1, 47, 48, 15, 16, 62]. The analysis of the resulting chemotaxis models has become a large field. The interested reader may consult Horstmann [30] for a detailed review of many mathematical aspects of chemotaxis.

7.3 Understanding pattern formation in *E. coli* bacteria

Bacteria are by far the best studied type of organism which deploys chemotaxis. A schematic model that takes into account cell densities $u(t, x)$, chemoattractant $S(t, x)$ and a nutrient $n(t, x)$ could include the following. The cell density diffuses, uses chemotaxis and proliferates; the chemoattractant diffuses, is produced and/or taken up by the bacteria; the nutrient might diffuse and is taken up by bacteria.

Apart from modelling the chemotactic sensitivity we discussed in the previous section, quantifying the other contributions to the evolution equations for u , S and n is a real challenge. We here introduce one concrete model from [46, 57], and outline the kind of analysis one can perform to understand pattern formation.

For the chemotactic sensitivity we take a commonly employed choice,

$$u\chi(S) = u \frac{k_1}{(k_2 + S)^2}$$

For cell growth and death, we choose a kind of logistic growth dependent on the amount of nutrient

$$k_3 u \left(k_4 \frac{n^2}{k_9 + n^2} - u \right)$$

For the production of chemoattractant, we choose a Holling type III response,

$$k_5 n \frac{u^2}{k_6 + u^2}$$

Chemoattractant consumption is assumed to be just linear in both u and S , but for nutrient consumption we choose again a Holling type III functional response,

$$k_8 u \frac{n^2}{k_9 + n^2}$$

All parameters k_1, \dots, k_9 need to be measured experimentally. In all, we consider

$$u_t = d_u \Delta u - \nabla \cdot \left(\frac{k_1 u}{(k_2 + S)^2} \nabla S \right) + k_3 u \left(\frac{k_4 n^2}{k_9 + n^2} - u \right) \quad (7.3.1)$$

$$S_t = d_S \Delta S + k_5 n \frac{u^2}{k_6 + u^2} - k_7 u S \quad (7.3.2)$$

$$n_t = d_n \Delta n - k_8 u \frac{n^2}{k_9 + n^2} \quad (7.3.3)$$

These equations are to be solved on a finite domain Ω , corresponding to the finite dish in which the bacteria grow and move. At time $t = 0$ we specify initial concentrations $u_0(x)$, $S_0(x)$ and $n_0(x)$.

Let us try to see what kind of analysis is required to understand pattern formation (7.3.1)–(7.3.3). If we include both diffusion and consumption of the nutrient n , the only long term behaviour of this nutrient concentration is inevitably to vanish everywhere. But then growth of cells must halt, and death of cells prevail. We conclude that the full model only has one steady state: the trivial equilibrium. This fact alone makes it clear that linear stability analysis of this steady state will not give us true insight in pattern formation. An analysis of intermediate-time solutions is necessary, which is much more involved, as this is a fully nonlinear problem. However, the linear stability analysis does give us an idea which perturbations are linearly unstable, and how the stability of certain modes depends on the parameters in the system.

7.3.1 Experiments in liquid media

Experiments on bacterial chemotaxis are often done in two different kinds of media, each with their own kinds of patterns. In liquid media, dynamics are comparatively fast. This allows us safely to neglect cell proliferation, and the nutrient n is not considered the main food source. The model above now simplifies to

$$u_t = d_u \Delta u - \nabla \cdot \left(\frac{k_1 u}{(k_2 + S)^2} \nabla c \right) \quad (7.3.4)$$

$$S_t = d_S \Delta S + k_5 n \frac{u^2}{k_6 + u^2} \quad (7.3.5)$$

$$n_t = d_n \Delta n \quad (7.3.6)$$

In this more simplified liquid model (7.3.4)–(7.3.6), the nutrient does approach a nontrivial uniform steady state \bar{n} as $t \rightarrow \infty$, since Ω is finite. By inspection we also find a nontrivial steady state $(u_0, 0, 0)$. Perturbing slightly from this steady state, the chemoattractant will grow without bound, which saturates the chemotactic response. Therefore, the movement of bacteria will eventually be dominated by diffusion, and we again find a uniform equilibrium to which the system converges. So also in this system, it is clear that linear stability analysis will not give much information on pattern formation. Instead we need to look at what the solutions look like after some time has elapsed, but not enough for the influence of chemotaxis to have waned entirely.

The equation for the nutrient $n(t, x)$ is now decoupled. We therefore assume that n has spread out to become uniformly constant n_0 , say. We assume furthermore that Ω is one-dimensional. We nondimensionalise the two remaining equations by setting

$$v = \frac{u}{u_0}, \quad w = \frac{S}{k_2}, \quad t^* = \frac{k_5 n_0}{k_2} t, \quad \xi^* = \sqrt{\frac{k_5 n_0}{d_S k_2}} x, \quad \delta = \frac{d_u}{d_S}, \quad \alpha = \frac{k_1}{d_S k_2}, \quad \mu = \frac{k_6}{u_0^2} \quad (7.3.7)$$

Note that α contains the original parameters involved in the chemotactic sensitivity $\chi(S)$. Using that $n/n_0 = 1$ is the new rescaled nutrient concentration, we obtain

$$v_t = \delta v_{xx} - \alpha \frac{\partial}{\partial x} \left(\frac{v}{(1+w)^2} \frac{\partial w}{\partial x} \right) \quad (7.3.8)$$

$$w_t = w_{xx} + \frac{v^2}{\mu + v^2} \quad (7.3.9)$$

where we have dropped the stars on the rescaled time and space variables for clarity of presentation. In this scaling, the steady state $(u, S) = (u_0, 0)$ is $(v, w) = (1, 0)$.

Which solutions $v(t, x)$, $w(t, x)$ that are heterogeneous in space initially grow and then decay in time? Let us introduce order $\mathcal{O}(\varepsilon)$ perturbations of the steady state $(1, 0)$,

$$v(t, x) = 1 + \varepsilon f_k(t) \sum a_k e^{ikx} \quad (7.3.10)$$

$$w(t, x) = \frac{1}{\mu + 1} t + \varepsilon g_k(t) \sum b_k e^{ikx} \quad (7.3.11)$$

The choice of the factor $1/(1 + \mu)$ for w stems from the fact that initially $v(t, x)$ is near 1 for all x , so w_t is close to $1/(1 + \mu)$ for all x . Combining this with $w(0, x) = 0$ everywhere, $w(t, x)$ is hence approximated well by $1/(1 + \mu)t$ for all x .

The domain Ω is finite and one-dimensional, so as before we only find integer modes m , related to k by

$$k^2 = \frac{m^2 \pi^2}{l^2}$$

where l is the dimensionless length of the domain. Substituting the Ansatz (7.3.10)–(7.3.11), we need to avoid coupling of modes to keep the problem tractable. These are all of $\mathcal{O}(\varepsilon^2)$, so we need to consider only the contributions up to $\mathcal{O}(\varepsilon)$. For each k , we get

$$\frac{dF(\tau)}{d\tau} = -\delta k^2 F(\tau) + \alpha(\mu + 1)^2 \frac{k^2}{\tau^2} G(\tau) \quad (7.3.12)$$

$$\frac{dG(\tau)}{d\tau} = -k^2 G(\tau) + \frac{2\mu}{(\mu + 1)^2} F(\tau) \quad (7.3.13)$$

where $F(\tau) := f(t)$, $G(\tau) := g(t)$ and $\tau = t + \mu + 1$. From these equations it is immediately clear that if $t \rightarrow \infty$, then $\tau \rightarrow \infty$ and $k^2/\tau^2 \rightarrow 0$ for fixed k . Hence $F'(\tau)$ approaches $-\delta k^2 F(\tau)$, so for large times F is a decaying exponential. But if $F \rightarrow 0$, then also $G'(\tau) \rightarrow -k^2 G$ and G also decays exponentially.

The second derivative of F satisfies

$$\frac{d^2 F}{d\tau^2} = -\delta k^2 \frac{dF}{d\tau} + \alpha(\mu + 1)^2 k^2 \frac{d}{d\tau} \left(\frac{G(\tau)}{\tau^2} \right)$$

Hence, using that G may be written as a function of F and F' , we find a second order equation with non-constant coefficients,

$$F'' + \left(k^2(\delta + 1) + \frac{2}{\tau} \right) F' + k^2 \left(\delta k^2 + \frac{2\delta}{\tau} - \frac{2\alpha\mu}{\tau^2} \right) F = 0 \quad (7.3.14)$$

It is possible to find a solution in terms of so-called confluent hypergeometric functions, but this teaches us little about the solution's behaviour. Some headway may be made by

assuming that these coefficients do not change as quickly in time as F and its derivatives do. We write the last equation in the form

$$F'' + D(\tau)F' + N(\tau)F = 0$$

with

$$D(\tau) = k^2(\delta + 1) + \frac{2}{\tau}, \quad N(\tau) = k^2 \left(\delta k^2 + \frac{2\delta}{\tau} - \frac{2\alpha\mu}{\tau^2} \right)$$

The influence of the dimensionless chemotactic parameter α is confined to $N(\tau)$. Note that $D(\tau) > 0$ for all $\tau > 0$, and that $N(\tau) > 0$ for τ sufficiently large. If we for the moment assume that D and N are constant, then F is given by

$$F(\tau) = c_1 e^{\lambda_+ \tau} + c_2 e^{\lambda_- \tau},$$

where $\lambda_{\pm} = -\frac{D(\tau)}{2} \pm \frac{1}{2} \sqrt{D^2(\tau) - 4N(\tau)}$. The characteristic equation is given by $\lambda^2 + D\lambda + N = 0$, so we know that λ_+ has positive real part if $N < 0$. The tipping point is exactly when $N = 0$.

Let us now focus how chemotaxis changes the sign of $N(\tau)$ through the influence of α . As α increases, N decreases for fixed τ , and $N(\tau) < 0$ for τ small. As τ then increases, N changes sign and the eigenvalues both have negative real parts. Which modes m or wavenumbers k are unstable? The form of $N(\tau)$ suggests that the lowest frequencies are the most unstable. At $t = 0$, $\tau = \tau_0 := 1 + \mu$. Solving $N(\tau_0) = 0$ for k , we get

$$\frac{2}{\delta(1 + \mu)} \left(\frac{\alpha\mu}{1 + \mu} - \delta \right) =: K^2$$

So for $k^2 > K^2$ modes are stable for all $\tau \geq \tau_0$.

If we now choose a fixed mode and compute solutions over small τ intervals in which D and N change little, we can piece together an approximation of a the true solution to (7.3.14), and we expect this to be an increasing function of τ while λ_+ remains positive, and a decreasing function thereafter. In other words, F increases until $N(\tau)$ changes from negative to positive, after which the solution decays to zero. See Figure 7.3 for an example solution $F(\tau)$. If we compare different solutions $F(\tau)$ computed for different wavenumbers k (see again Figure 7.3), we notice that first higher modes appear, which also decay fast. There is a wavenumber which has highest amplitude, but this also disappears quite soon. After this has reached its peak, modes with smaller wavenumber are dominant. This gives a nice indication of the behaviour of smaller aggregates joining into larger ones over time, as observed experimentally.

7.3.2 Experiments in semi-solid media

In experiments with *Salmonella* or *E. coli* bacteria growing in a semi-solid medium, so-called swarm rings are often observed. These are rings emanating from the inoculated center. Over time new concentric rings are formed. The inner rings slowly break up into small regions. See Figure 7.4 for some striking examples. In these experiments food is often available in excess, so we can assume that food consumption is negligible in our model. This reduces the model (7.3.1)–(7.3.3), and this new model does have a nontrivial steady state, and we perform linear stability analysis to study it.

Experimental observation indicates that after the inoculation event in the center of the petri dish, the bacteria first produce a homogeneous distribution over the entire dish, after which the pattern is formed on top of this distribution.

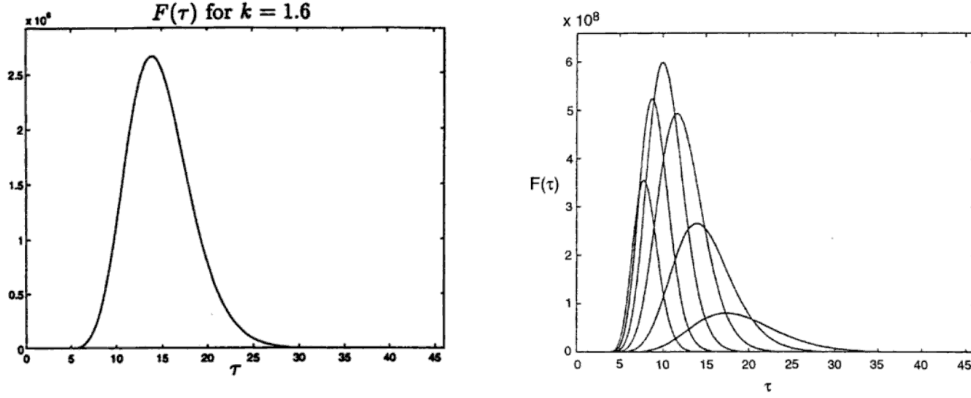


FIGURE 7.3. Left: Example solution $F(\tau)$ for a given mode $k = 1.6$. Right: comparison of solutions $F(\tau)$ for various k . The functions that decay the fastest have highest wavenumber, and vice versa. This illustrates the coalescence of smaller aggregates over time. See the text for more detail. From [58].

First we nondimensionalise the full model again. Using a slightly different scaling than for the liquid model, we now find

$$v_t = \delta \Delta v - \alpha \nabla \cdot \left(\frac{v}{(1+w)^2} \nabla w \right) + \rho v \left(\sigma \frac{z^2}{1+z^2} - v \right) \quad (7.3.15)$$

$$w_t = \Delta w + \beta z \frac{v^2}{\mu + v^2} - v z \quad (7.3.16)$$

$$z_t = d_z \Delta z - \kappa v \frac{z^2}{1+z^2} \quad (7.3.17)$$

We neglect food consumption, so we set $\kappa = 0$ and the nondimensional nutrient z to be uniformly constant. The reduced model now becomes

$$v_t = \delta \Delta v - \alpha \nabla \cdot \left(\frac{v}{(1+w)^2} \nabla w \right) + \rho v \left(\sigma \frac{z}{1+z^2} - v \right) \quad (7.3.18)$$

$$w_t = \Delta w + \beta z \frac{v^2}{\mu + v^2} - v z \quad (7.3.19)$$

which has a steady state

$$(v^*, w^*) = \left(\sigma \frac{z^2}{1+z^2}, \beta z \frac{v^*}{\mu + v^*} \right)$$

We write this reduced model in the following more general form

$$v_t = \delta \Delta v - \alpha \nabla \cdot (v \chi(w) \nabla w) + f(v, w) \quad (7.3.20)$$

$$w_t = \Delta w + g(v, w) \quad (7.3.21)$$

We linearise by setting

$$v = v^* + \varepsilon v_1, \quad w = w^* + \varepsilon w_1$$

and find

$$\frac{\partial v_1}{\partial t} = \delta \Delta v_1 - \alpha v^* \chi(w^*) \Delta w_1 + f_v^* v_1 + f_w^* w_1 \quad (7.3.22)$$

$$\frac{\partial w_1}{\partial t} = \Delta w_1 + g_v^* v_1 + g_w^* w_1 \quad (7.3.23)$$

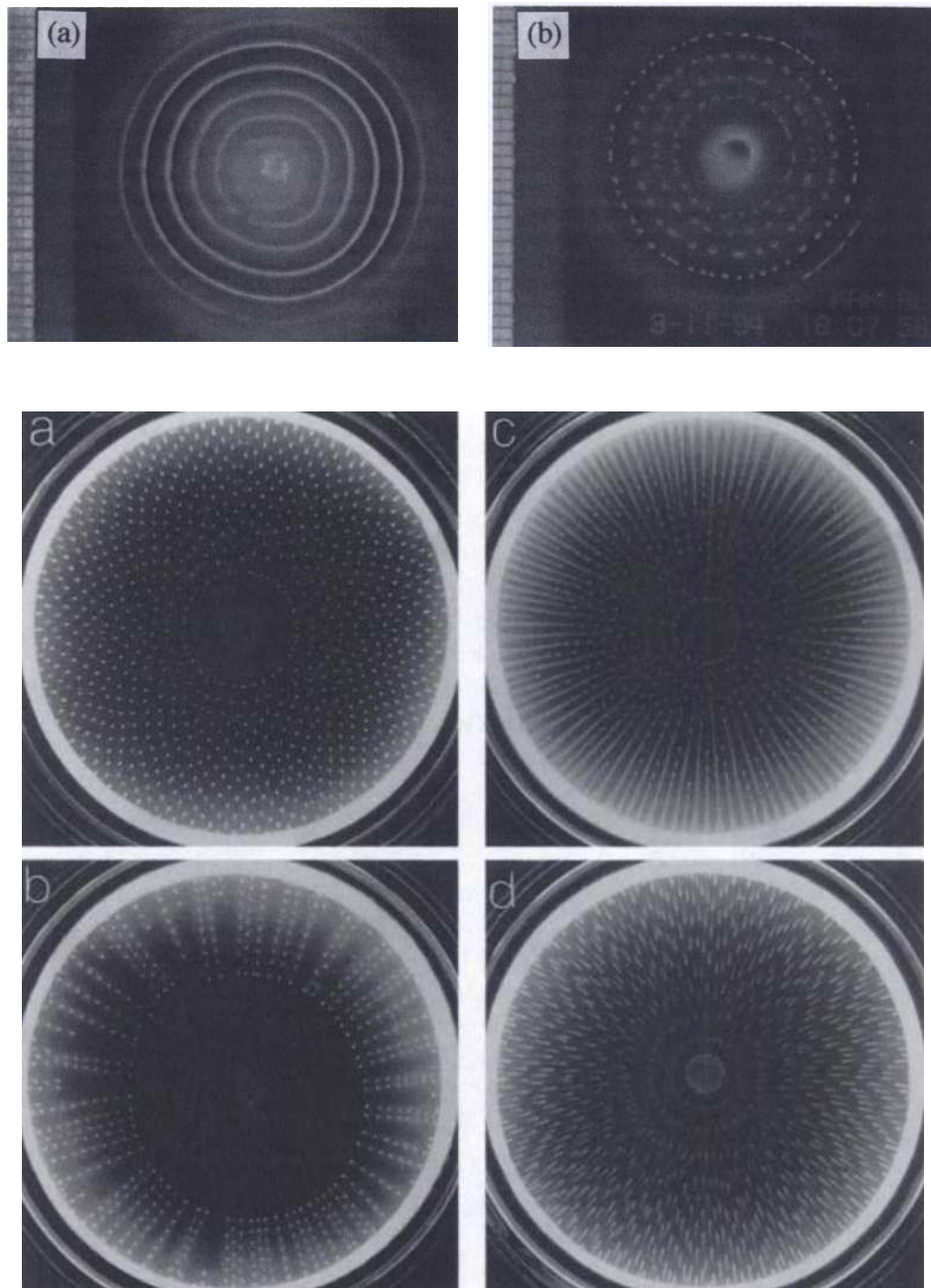


FIGURE 7.4. Top row: typical *Salmonella* chemotaxis patterns in semi-solid medium. Taken from [9]. After an initial inoculation with about 10^4 bacteria in the center, the left figure shows solid rings after 48h, the right broken rings after 70h. The bottom four images are typical *E. coli* patterns in semi-solid medium, also from [9].

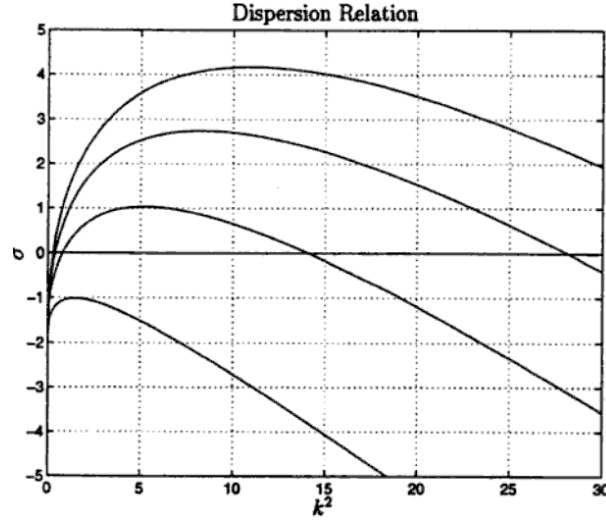


FIGURE 7.5. The dispersion relation $\lambda(k^2)$ (here denoted by σ on the y -axis) for varying μ . Higher μ gives higher curves, and thus a wider range of λ 's where the dispersion relation is positive. From [58].

where f_v^* denotes $\frac{\partial f}{\partial v}(v^*, w^*)$, etc. In short,

$$\frac{\partial}{\partial t} \begin{pmatrix} v_1 \\ w_1 \end{pmatrix} = (a + D\Delta) \begin{pmatrix} v_1 \\ w_1 \end{pmatrix} \quad (7.3.24)$$

where

$$A = \begin{pmatrix} f_v^* & f_w^* \\ g_v^* & g_w^* \end{pmatrix}, \quad D = \begin{pmatrix} d_v & -\alpha v^* \chi(w^*) \\ 0 & 1 \end{pmatrix}$$

Note that D is non-diagonal, so this is a cross-diffusion problem. We look for solutions by setting

$$\begin{pmatrix} v_1 \\ w_1 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} e^{\lambda t + i\mathbf{k}x}$$

Substituting this into (7.3.24) leads to

$$[\lambda I + D|\mathbf{k}|^2 - A] \begin{pmatrix} v_1 \\ w_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

So we study when the determinant of $\lambda I + D|\mathbf{k}|^2 - A$ becomes zero. Setting $k^2 = |\mathbf{k}|^2$ we get

$$\lambda^2 + b(k^2)\lambda + c(k^2) = 0,$$

and

$$\lambda_{\pm} = \frac{1}{2}(-b(k^2) \pm \sqrt{b^2(k^2) - 4c(k^2)})$$

for suitable functions $b(k^2)$ and $c(k^2)$. We are as usual only interested in the largest eigenvalue, λ_+ and its behaviour for varying k^2 .

To study pattern formation, we need to require that the homogeneous steady state is stable. In other words, if $\mathbf{k} = (0, 0)$, then v_1 and w_1 are exponential functions only of time, and we require these to be decreasing functions. So we demand that $\lambda_+(0) < 0$, which translates into

$$b(0) = f_v^* + g_w^* < 0 \quad (7.3.25)$$

$$c(0) = \det A > 0 \quad (7.3.26)$$

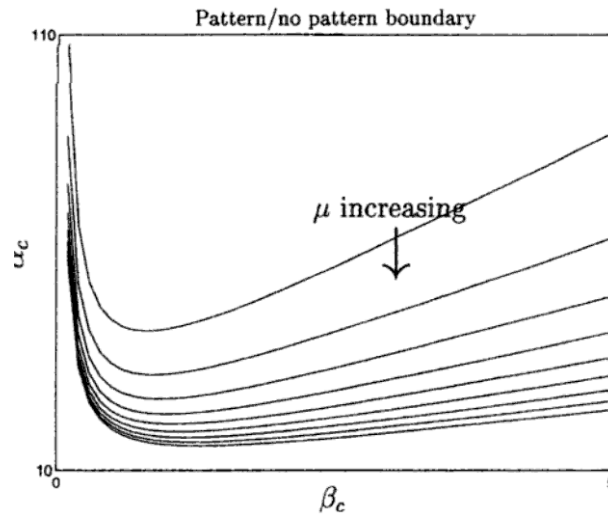


FIGURE 7.6. The boundary between the stable and unstable perturbations in the (k_c, α) -plane, for various values of μ . (The parameter k_c is here called β_c .) Points above such a boundary give rise to patterns, points below do not. As μ increases, the region where patterns are possible increases. From [58].

In the end we find that $\lambda_+ > 0$ only in some range $0 \leq k_1^2 < k^2 < k_2^2$, and increasing μ makes the interval of k values where λ_+ is positive larger (see Figure 7.5). The first mode k_c which becomes unstable has the property that k_c is the only real root of $\lambda_+(k)$. Analytically, this allows us to find a condition on the nondimensional chemotaxis parameter α ,

$$\alpha = -\frac{(\delta_v g_w^* + f_v^*) + 2\sqrt{\delta_u \det A}}{v^* \chi(w^*) g_v^*}, \quad k_c^2 = \sqrt{\frac{\det A}{\delta_u}}$$

This allows us to draw the boundary between stable and unstable perturbations in the (k_c, α) -plane, see Figure 7.6.

This analysis does not tell us what patterns are formed, only which perturbations are unstable. The full nonlinear problem is much more involved, and requires the study of the Ginzburg-Landau equations.

Chapter 8

Adaptive Dynamics

8.1 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

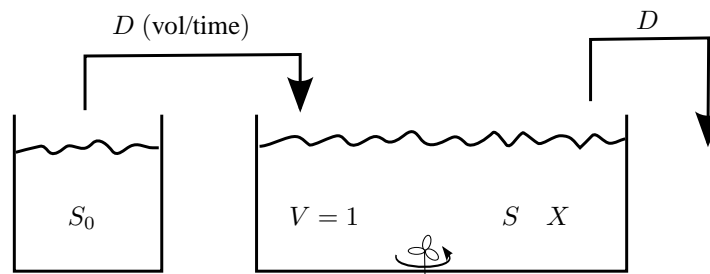


FIGURE 8.1. The chemostat

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.

8.2 The pessimization principle

The chemostat (see Figure 8.1) 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 [54] 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 is 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

$$\begin{aligned}\frac{dS}{dt} &= S_0 - S - uSX \\ \frac{dX}{dt} &= -X + \eta uSX = (-1 + \eta uS)X\end{aligned}\tag{8.2.1}$$

We require that $\eta uS_0 > 1$, 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

$$\bar{S} = \frac{1}{\eta u}\tag{8.2.2}$$

Exercise 8.2.1. Show that (8.2.1) 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 $-1 + \eta uS$ 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 [43].)

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 8.2. 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 [22].

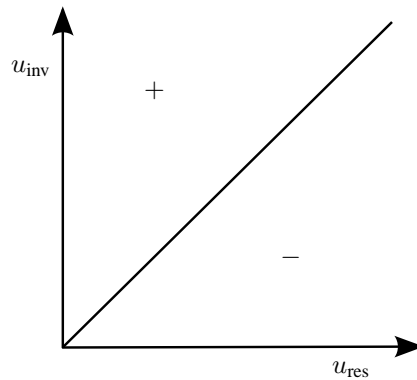


FIGURE 8.2. PIP showing the sign of the invasion exponent. Note that there is necessarily neutrality at the diagonal.

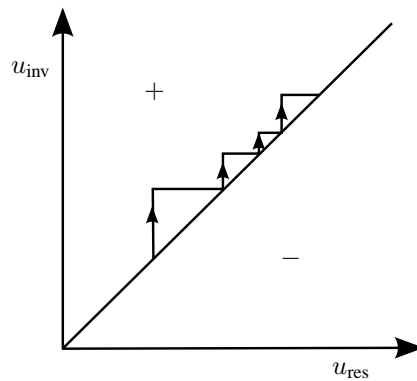


FIGURE 8.3. A trait substitution sequence: note that the height of the jumps is a stochastic facet of mutation, but that the direction of evolutionary change (i.e. the fact that u increases) is a consequence of natural selection.

To verify the folk theorem for the present situation, we need to study the three-dimensional system

$$\begin{aligned} \frac{dS}{dt} &= S_0 - S - u_{\text{res}}SX - u_{\text{inv}}SY \\ \frac{dX}{dt} &= -X + \eta u_{\text{res}}SX \\ \frac{dY}{dt} &= -Y + \eta u_{\text{inv}}SY \end{aligned} \quad (8.2.3)$$

Exercise 8.2.2. Use reduction to a two-dimensional system and next Poincaré-Bendixson theory (see Appendix) to prove the folk theorem for system (8.2.3).

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 8.3.

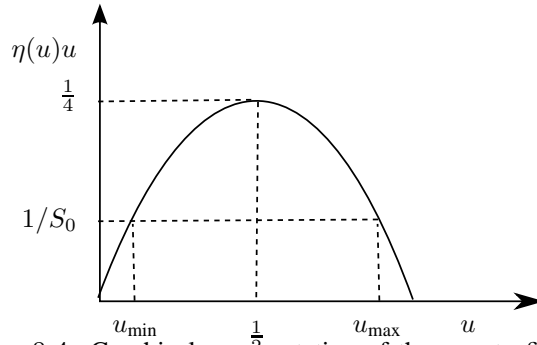


FIGURE 8.4. Graphical representation of the way to find the interval where $\eta(u)uS_0 > 1$.

If we write

$$\text{invasion exponent} = s_{u_{\text{res}}}(u_{\text{inv}})$$

we can determine the direction in which *small* mutational steps and selection will drive the trait, by computing the *selection gradient*

$$\left. \frac{\partial}{\partial v} s_u(v) \right|_{v=u}$$

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

$$\eta = \eta(u) = 1 - u \quad (8.2.4)$$

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 (see Figure 8.4). The invasion exponent is given by

$$s_u(v) = -1 + \frac{(1-v)v}{(1-u)u} = \frac{(v-u)(1-v-u)}{(1-u)u} \quad (8.2.5)$$

and we see that, in addition to the diagonal neutrality curve, we have a neutrality curve $v = 1 - u$ (Figure 8.5). 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 8.5) 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 8.2.2: the steady state substrate level is given by

$$\bar{S} = \frac{1}{\eta u}$$

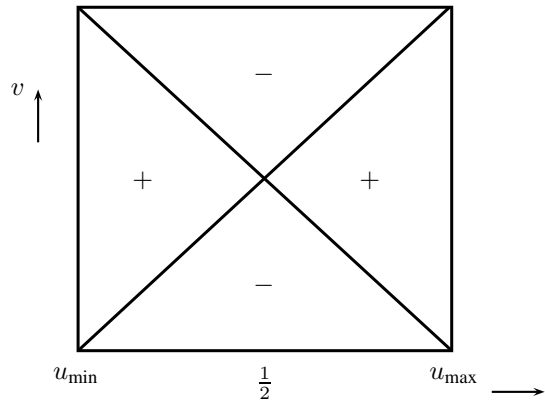


FIGURE 8.5. The PIP when η is given by (8.2.4).

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 8.2.3. 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?

8.3 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

$$\begin{aligned} \frac{dS_1}{dt} &= S_{10} - S_1 - u_1 S_1 X \\ \frac{dS_2}{dt} &= S_{20} - S_2 - u_2 S_2 X \\ \frac{dX}{dt} &= -X + u_1 S_1 X + u_2 S_2 X = (-1 + u_1 S_1 + u_2 S_2) X \end{aligned} \quad (8.3.1)$$

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

$$-1 + \frac{u_1 S_{10}}{1 + u_1 \bar{X}} + \frac{u_2 S_{20}}{1 + u_2 \bar{X}} = 0 \quad (8.3.2)$$

and next defining \bar{S}_1 and \bar{S}_2 by the explicit formulas

$$\bar{S}_1 = \frac{S_{10}}{1 + u_1 \bar{X}}, \quad \bar{S}_2 = \frac{S_{20}}{1 + u_2 \bar{X}} \quad (8.3.3)$$

Exercise 8.3.1. Show that this steady state is globally asymptotically stable. Hint: first show that $S_1 + S_2 + X \rightarrow S_{10} + S_{20}$ for $t \rightarrow \infty$, 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

$$s_x(y) = -1 + u_1(y)\bar{S}_1(x) + u_2(y)\bar{S}_2(x) \quad (8.3.4)$$

and consequently,

$$\text{selection gradient} = u_1'(x)\bar{S}_1(x) + u_2'(x)\bar{S}_2(x) \quad (8.3.5)$$

Singular points are, by definition, characterised by a vanishing selection gradient. The condition

$$u_1'(x)\bar{S}_1(x) + u_2'(x)\bar{S}_2(x) = 0 \quad (8.3.6)$$

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

$$u_1(x) = x, \quad u_2(x) = 1 - x \quad (8.3.7)$$

in some detail. In this case (8.3.6) reduces to

$$\bar{S}_1 = \bar{S}_2. \quad (8.3.8)$$

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 (8.3.3) we find

$$\frac{S_{10} - \bar{S}}{u_1 \bar{S}} = \frac{S_{20} - \bar{S}}{u_2 \bar{S}}$$

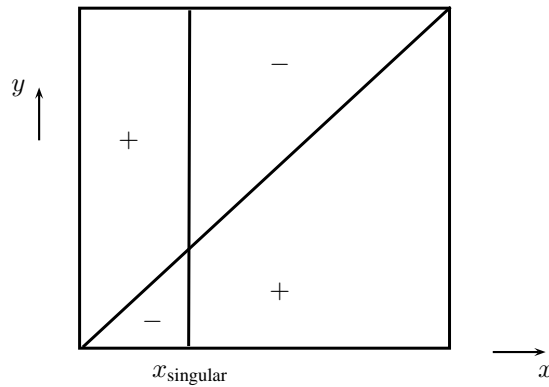


FIGURE 8.6. The PIP for the special case (8.3.7).

which, using (8.3.7) and $\bar{S} = 1$, after a few manipulations leads to the expression

$$x_{\text{singular}} = \frac{S_{10} - 1}{S_{10} + S_{20} - 2} \quad (8.3.9)$$

Much more important than this explicit expression, however, is the following observation. From (8.3.4), (8.3.8) and (8.3.7) we see that

$$s_{x_{\text{singular}}}(y) = 0 \quad \text{for all } y! \quad (8.3.10)$$

or, in words, the second neutrality curve, that intersects the diagonal at the singular point, is a vertical line (see Figure 8.6). 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 8.3.2. 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.

The key point about (8.3.7) 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 (8.3.7) and investigate the effects. But before doing any algebra, we consider the geometry. From Figure 8.7 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.

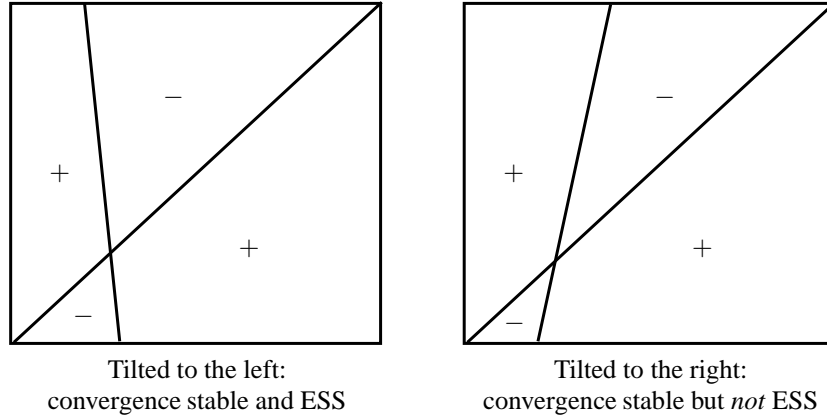


FIGURE 8.7. The two generic perturbations of a vertical neutrality curve (with the key local features extended globally, even though that need not be the case).

So let us look at dimorphic populations. These are described by the system

$$\begin{aligned}
 \frac{dS_1}{dt} &= S_{10} - S_1 - u_1(x)S_1X_1 - u_1(y)S_1X_2 \\
 \frac{dS_2}{dt} &= S_{20} - S_2 - u_2(x)S_2X_1 - u_2(y)S_2X_2 \\
 \frac{dX_1}{dt} &= -X_1 + u_1(x)S_1X_1 + u_2(x)S_2X_1 \\
 \frac{dX_2}{dt} &= -X_2 + u_1(y)S_1X_2 + u_2(y)S_2X_2
 \end{aligned} \tag{8.3.11}$$

Exercise 8.3.3. Show that the corresponding steady states are given by

$$\begin{pmatrix} S_1 \\ S_2 \end{pmatrix} = \frac{1}{u_1(x)u_2(y) - u_1(y)u_2(x)} \begin{pmatrix} u_2(y) - u_2(x) \\ u_1(x) - u_1(y) \end{pmatrix} \tag{8.3.12}$$

and

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \frac{u_2(y)S_{10}}{u_2(y) - u_2(x)} - \frac{u_1(y)S_{20}}{u_1(x) - u_1(y)} - \frac{u_1(y) - u_2(y)}{u_1(x)u_2(y) - u_1(y)u_2(x)} \\ \frac{-u_2(x)S_{10}}{u_2(y) - u_2(x)} + \frac{u_1(x)S_{20}}{u_1(x) - u_1(y)} + \frac{u_2(x) - u_1(x)}{u_1(x)u_2(y) - u_1(y)u_2(x)} \end{pmatrix} \tag{8.3.13}$$

Show that whenever the expressions at the right hand side of (8.3.13) 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 (8.3.12). The global stability of this steady state is established in [6].

The invasion exponent in such a dimorphism is given by

$$s_{x,y}(z) = -1 + u_1(z)S_1 + u_2(z)S_2 \tag{8.3.14}$$

with S_i given by (8.3.12). Choosing rather general functions $u_1(z)$ and $u_2(z)$ once more, the situation of a diverging dimorphism is depicted in Figure 8.8.

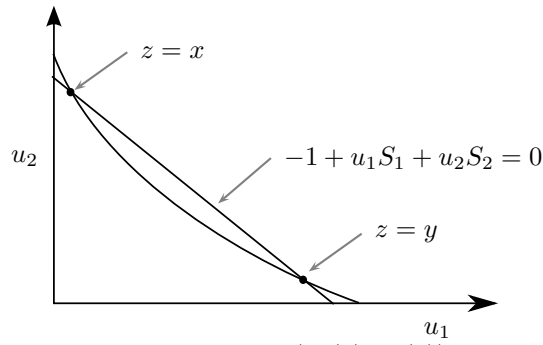


FIGURE 8.8. Plot of the curve $z \mapsto (u_1(z), u_2(z))$ and of the straight line $-1 + u_1S_1 + u_2S_2 = 0$. Intersections correspond to the dimorphism. The invasibility condition $s_{x,y}(z) > 0$ means that the point on the curve corresponding to trait z should be above the straight line. So in the depicted situation we have a diverging dimorphism.

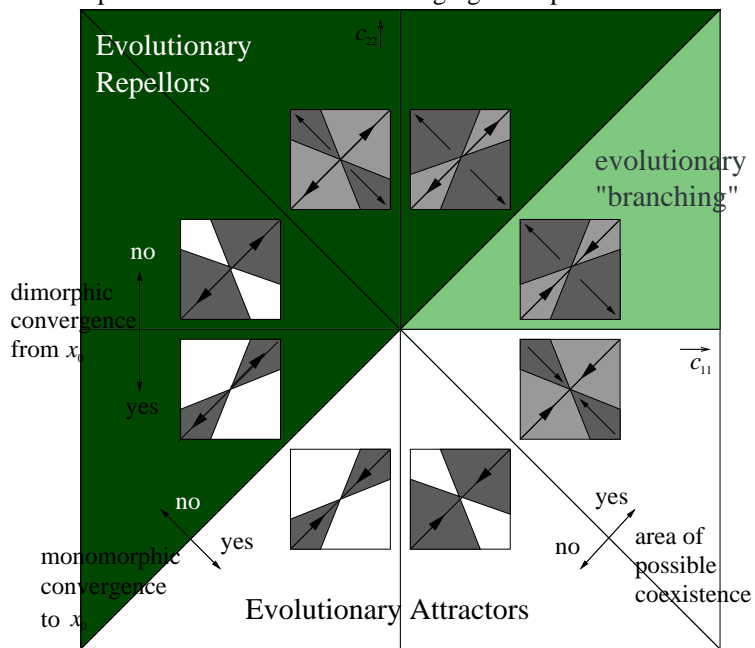


FIGURE 8.9. Classification of singular points...

Chapter 9

Appendix

9.1 Bifurcation theory

9.1.1 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

$$\frac{a}{e} = \frac{c}{d - bc\beta} \quad (9.1.1)$$

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

$$\frac{a}{e} = \frac{2c}{d - bc\beta} + \frac{1}{b\beta} \quad (9.1.2)$$

characterises the loss of stability of the positive steady state (indeed, when (9.1.2) 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 (9.1.1) and (9.1.2) 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.

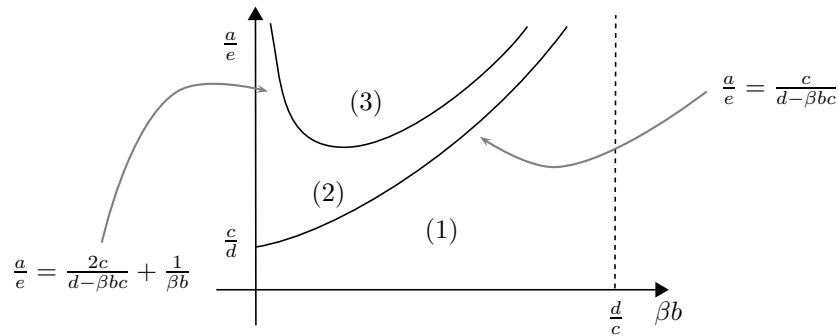


FIGURE 9.1. The phase portrait of the Rosenzweig-MacArthur system.

Considering $\frac{d}{c}$ as fixed, we then represent (9.1.1), (9.1.2) and the results of Section 3.6 graphically in Figure 9.1.

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 parameter values for which the phase portrait is structurally unstable. The curve defined by (9.1.2) 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.

9.1.2 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 (9.1.1) and (9.1.2) 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 (9.1.1) and (9.1.2), you see that small perturbation in the curve cannot eliminate the crossings). Higher degeneracies require more parameters to be a natural phenomenon.

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 (9.1.1) and (9.1.2) each determine a codimension one bifurcation; incidentally, the bifurcation corresponding to (9.1.2) 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).

9.1.3 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 [37].

- (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

$$\dot{x} = \alpha + x^2 \tag{9.1.3}$$

we associate the bifurcation diagram in Figure 9.2.

Exercise 9.1.1. Construct the diagram for $\dot{x} = \alpha - x^2$.

Exercise 9.1.2. Analyse the discrete time system

$$x(n+1) = \alpha + x(n) \pm (x(n))^2.$$

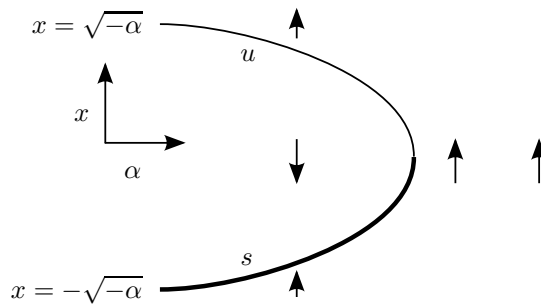


FIGURE 9.2. Bifurcation diagram for a fold bifurcation corresponding to equation (9.1.3). Note that there are no solutions for $\alpha > 0$. Bold lines signify stable (*s*) solution branches, thin lines unstable (*u*) ones.

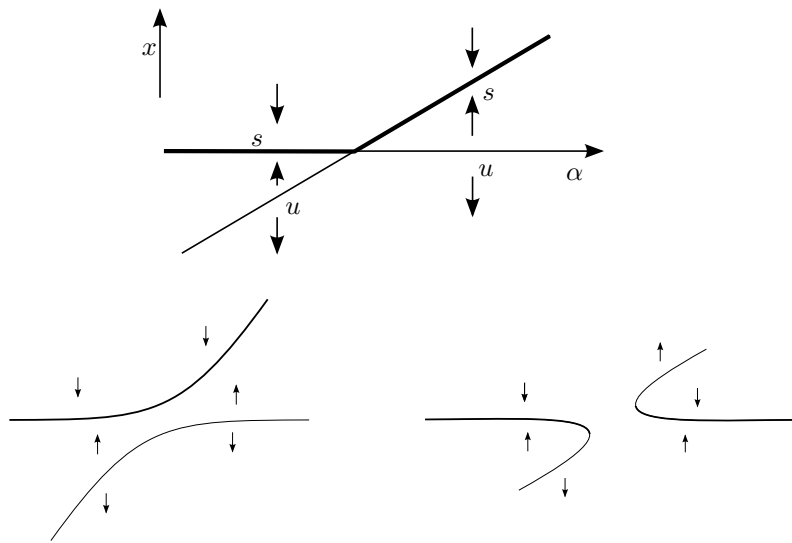


FIGURE 9.3. Bifurcation diagram for a transcritical bifurcation (top). Bifurcation diagrams of perturbed systems can be seen below it. In the left case there are no bifurcations, in the right there are two folds. Bold lines signify stable (*s*) solution branches, thin lines unstable (*u*) ones.

(ib) *transcritical bifurcation*: The top part of Figure 9.3 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 9.3.

(ic) *pitchfork bifurcation* (from which the name “bifurcation” originated):

For

$$\dot{x} = x(\alpha - x^2) \tag{9.1.4}$$

we have the diagram shown in the left part of Figure 9.4. The right part shows the bifurcation structure of

$$\dot{x} = x(\alpha + x^2). \tag{9.1.5}$$

This form of bifurcation is typical for systems with reflection symmetry.

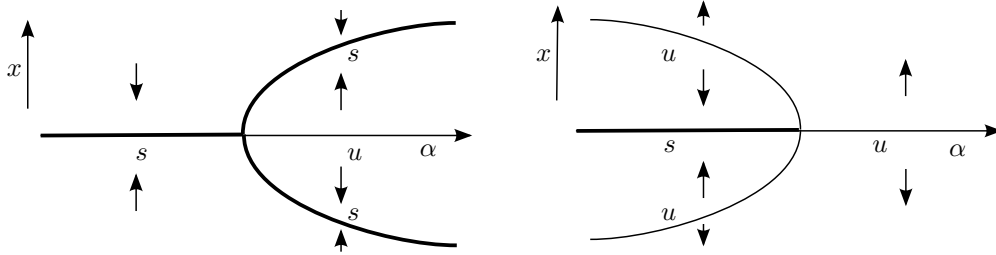


FIGURE 9.4. Bifurcation diagrams for a pitchfork bifurcation. The left figure corresponds to equation (9.1.4), the right to (9.1.5). Bold lines signify stable (s) solution branches, thin lines unstable (u) ones.

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 9.1.3. When perturbations destroy the symmetry, the diagram may break into two components. Draw the various components.

(ii) *Hopf bifurcation:* The system

$$\dot{x}_1 = \alpha x_1 - x_2 \pm x_1(x_1^2 + x_2^2) \quad (9.1.6)$$

$$\dot{x}_2 = x_1 - \alpha x_2 \pm x_2(x_1^2 + x_2^2) \quad (9.1.7)$$

takes in polar coordinates the form

$$\dot{r} = r(\alpha \pm r^2) \quad (9.1.8)$$

$$\dot{\phi} = 1 \quad (9.1.9)$$

There are two different bifurcation diagrams, for the $+$ and $-$ cases, shown in Figure 9.5. 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 9.1.4. The discrete time logistic equation

$$x(n+1) = \alpha x(n)(1-x(n))$$

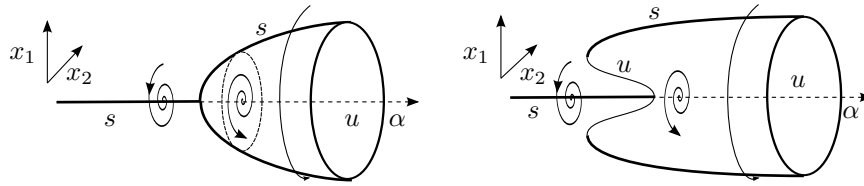


FIGURE 9.5. Supercritical (left) and subcritical Hopf bifurcation diagrams. Bold lines signify stable (s) solution branches, thin lines unstable (u) ones.

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 9.6 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 9.1.5. Consider the consumer-resource interaction in the chemostat as described in Section (8.1), but now in the form

$$\frac{dS}{dt} = DS_0 - DS - g(S)X \quad (9.1.10)$$

$$\frac{dX}{dt} = -DX + \eta g(S)X \quad (9.1.11)$$

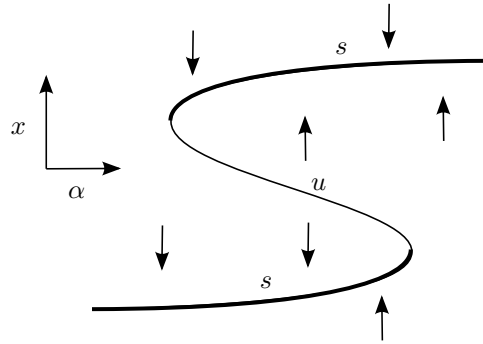


FIGURE 9.6. A typical hysteresis diagram, with two saddle-node bifurcations at the turning points. Bold lines signify stable (s) solution branches, thin lines unstable (u) ones.

with up-take function g that we don't know and that we like to determine experimentally. 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 (\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 unstable steady states and interpret the result in terms of hysteresis.

9.1.4 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

$$\dot{u} = f(u), \quad u \in \mathbb{R}^2 \quad (9.1.12)$$

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 (9.1.12): a point $p \in \mathbb{R}^2$ belongs to the ω -limit set of an orbit $u(t)$ if there exists a sequence $n \rightarrow \infty$ such that $u(t_n) \rightarrow p$ as $t_n \rightarrow \infty$. 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 9.1.1 (Poincaré-Bendixson). *Consider a bounded positive orbit γ of (9.1.12) with ω -limit set $\omega(\gamma)$. Then one of the following possibilities holds:*

- (1) $\omega(\gamma)$ is an equilibrium
- (2) $\omega(\gamma)$ is a periodic 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. [12, 59].

Bibliography

- [1] W. Alt, *Biased random walk models for chemotaxis and related diffusion approximations*, J. Math. Biology **9** (1980), 147–177.
- [2] H. Amann, *Ordinary differential equations: An introduction to nonlinear analysis*, Walter de Gruyter, 1990.
- [3] D. G. Aronson, *The role of diffusion in mathematical population biology: Skellam revisited*, Mathematics in biology and medicine (Bari, 1983), Lecture Notes in Biomath., vol. 57, Springer, Berlin, 1985, pp. 2–6. MR MR812871
- [4] D. G. Aronson and H. F. Weinberger, *Nonlinear diffusion in population genetics, combustion, and nerve pulse propagation*, Partial differential equations and related topics (Program, Tulane Univ., New Orleans, La., 1974), Springer, Berlin, 1975, pp. 5–49. Lecture Notes in Math., Vol. 446. MR MR0427837 (55 #867)
- [5] ———, *Multidimensional nonlinear diffusion arising in population genetics*, Adv. in Math. **30** (1978), no. 1, 33–76. MR MR511740 (80a:35013)
- [6] M.M. Ballyk, C.C. McCluskey, and G.S.K. Wolkowicz, *Global analysis of competition for perfectly substitutable resources with linear response*, J. Math. Biology **51** (2005), 458–490.
- [7] G. I. Barenblatt, *Scaling, self-similarity, and intermediate asymptotics*, Cambridge texts in applied mathematics, Cambridge University Press, 1996.
- [8] N. F. Britton, *Essential mathematical biology*, Springer-Verlag, London, 2003.
- [9] E. O. Budrene and H. C. Berg, *Complex patterns formed by motile cells of Escherichia coli*, Nature **376** (1991), 49–53.
- [10] C. Chicone, *The monotonicity of the period function for planar Hamiltonian vector fields*, J. Diff. Eq. **69** (1987), 310–321.
- [11] K.L. Chung, *Elementary probability theory with stochastic processes*, Springer Verlag, 1974.
- [12] E. A. Coddington and N. Levinson, *Theory of ordinary differential equations*, McGraw-Hill, 1955.
- [13] J. Cronin, *Mathematical aspects of Hodgkin-Huxley neural theory*, Cambridge University Press, 1987.
- [14] O. Diekmann and J.A.P. Heesterbeek, *Mathematical epidemiology of infectious diseases*, Wiley, 2000.
- [15] Radek Erban and Hans G. Othmer, *From individual to collective behavior in bacterial chemotaxis*, SIAM J. Appl. Math. **65** (2004/05), no. 2, 361–391. MR MR2123062 (2005j:35220)
- [16] ———, *Taxis equations for amoeboid cells*, J. Math. Biol. **54** (2007), no. 6, 847–885. MR MR2309023 (2008b:92011)
- [17] W. Feller, *The parabolic differential equation and the associated semi-group of transformations*, Annals of Math. **55** (1952), 468–519.
- [18] ———, *Diffusion processes in one dimension*, Trans. Am. Math. Soc. **77** (1954), 1–31.
- [19] ———, *On differential operators and boundary conditions*, Comm. Pure Applied Math. **8** (1955), 203–216.
- [20] R. A. Fisher, *The advance of advantageous genes*, Ann. of Eugenics **7** (1937), 355–369.
- [21] H. Freedman, *Deterministic mathematical models in population ecology*, Marcel Dekker, 1980.
- [22] S. A. H. Geritz, *Resident-invader dynamics and the coexistence of similar strategies*, J. Math. Biol. **50** (2005), 67–82.
- [23] A. Graham, *Nonnegative matrices and applicable topics in linear algebra*, John Wiley & Sons, 1987.
- [24] J. Grasman, *Asymptotic methods for relaxation oscillations and applications*, Applied Mathematics Series, Springer-Verlag, 1987.
- [25] P. Grindrod, *Patterns and waves in reaction-diffusion: Techniques and applications*, Oxford University Press, 1991.
- [26] K. P. Hadeler and F. Rothe, *Travelling fronts in nonlinear diffusion equations*, J. Math. Biology **2** (1975), 251–263.
- [27] J. A. P. Heesterbeek and J. A. J. Metz, *The saturating contact rate in marriage en epidemic models*, J. Math. Biol. **31** (1993), 529–539.
- [28] M. W. Hirsch, S. Smale, and R. L. Devaney, *Differential equations, dynamical systems, and an introduction to chaos*, Academic Press, 2004.
- [29] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, 1990.

- [30] D. Horstmann, *From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. I*, Jahresber. Deutsch. Math.-Verein. **105** (2003), no. 3, 103–165. MR MR2013508 (2005f:35163)
- [31] J. Keener and J. Sneyd, *Mathematical physiology*, Springer-Verlag, 1998.
- [32] E. F. Keller and L. A. Segel, *Travelling bands of chemotactic bacteria: a theoretical analysis*, J. Theor. Biol. **30** (1971), 235–248.
- [33] J. G. Kemeny and K. L. Snell, *Finite Markov chains*, D. Van Nostrand Company, Princeton, NJ, 1963.
- [34] J. Kevorkian and J. D. Cole, *Multiple scale and singular perturbation methods*, Springer-Verlag, 1996.
- [35] K. Kishimoto and H. Weinberger, *The spatial homogeneity of stable equilibria of some reaction diffusion systems on convex domains*, J. Diff. Equations **58** (1985), 15–21.
- [36] A. N. Kolmogorov, I. Petrovskii, and N. Piscunov, *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** (1937), 1–25.
- [37] Y. A. Kuznetsov, *Elements of applied bifurcation theory*, Springer-Verlag, 1995.
- [38] C. H. Lee and H. G. Othmer, *A multi-time-scale analysis of chemical reaction networks: I. Deterministic systems*, J. Math. Biol. **60** (2010), 387–450.
- [39] C. C. Lin and L. A. Segel, *Mathematics applied to deterministic problems in the natural sciences*, SIAM: Society for Industrial and Applied Mathematics, 1998.
- [40] D. Ludwig, D. G. Aronson, and H. F. Weinberger, *Spatial patterning of the spruce budworm*, J. Math. Biol. **8** (1979), no. 3, 217–258. MR MR657283 (83j:92061)
- [41] H. Matano, *Asymptotic behavior and stability of solutions of semilinear diffusion equations*, Publ. Res. Inst. Math. Sci. **15** (1979), 401–458.
- [42] H. Meinhardt, *Models of biological pattern formation*, Academic Press, available online as pdf at <http://www.eb.tuebingen.mpg.de/departments/former-departments/h-meinhardt/82-book/Bur82.htm>, 1982.
- [43] J. A. J. Metz, R. M. Nisbet, and S. A. H. Geritz, *How should we define “fitness” for general ecological scenarios?*, Trends Ecol. Evol. **7** (1992), 198–202.
- [44] E. F. Mishchenko and N. Kh. Rosov, *Differential equations with small parameters and relaxation oscillations*, Plenum, New York, 1980.
- [45] J. D. Murray, *Mathematical biology i: An introduction*, Springer Verlag, 2002.
- [46] J. D. Murray, D. Manoussaki and S. R. Lubkin, and R. B. Vernon, *A mechanical theory of in vitro vascular network formation*, Vascular Morphogenesis: In Vivo, In Vitro, and In Mente (C. D. Little, V. Mironov, and E. H. Sage, eds.), Birkhäuser, Boston, 1998, pp. 173–198.
- [47] H. G. Othmer, S. R. Dunbar, and W. Alt, *Models of dispersal in biological systems*, J. Math. Biol. **26** (1988), 263–398.
- [48] Hans G. Othmer and Thomas Hillen, *The diffusion limit of transport equations. II. Chemotaxis equations*, SIAM J. Appl. Math. **62** (2002), no. 4, 1222–1250. MR MR1898520 (2003c:35154)
- [49] L. Perko, *Differential equations and dynamical systems*, Springer-Verlag, 2001.
- [50] M. Renardy and R. C. Rogers, *An introduction to partial differential equations*, Springer-Verlag, 1993.
- [51] S. Rinaldi and S. Muratori, *Slow-fast limit cycle in predator-prey models*, Ecological modeling **61** (1992), 287–308.
- [52] S. Schnell and P.K. Maini, *Enzyme kinetics at high enzyme concentration*, Bull. Math. Biol. (2000), 483–499.
- [53] L. A. Segel, *Modeling dynamic phenomena in molecular and cellular biology*, Cambridge University Press, 1984.
- [54] H. L. Smith and P. Waltman, *The theory of the chemostat*, Cambridge University Press, 1995.
- [55] A. Stevens, *Trail following and aggregation of myxobacteria*, J. Biol. Sys. **3** (1995), 1059–1068.
- [56] D. W. Stroock, *Some stochastic processes which arise from a model of the motion of a bacterium*, Z. Wahrscheinlichkeitstheorie **28** (1974), 305–315.
- [57] R. Tyson, S. R. Lubkin, and J. D. Murray, *Model and analysis of chemotactic bacterial patterns in a liquid medium*, J. Math. Biol. **38** (1999), no. 4, 359–375. MR MR1687391
- [58] R. C. Tyson, *Pattern formation by E. coli—mathematical and numerical investigation of a biological phenomenon*, Ph.D. thesis, Department of Applied Mathematics, University of Washington, Seattle, 1996.
- [59] F. Verhulst, *Nonlinear differential equations and dynamical systems*, Springer-Verlag, 1996.
- [60] ———, *Methods and applications of singular perturbations*, Springer-Verlag, 2005.
- [61] S. Vogel, *Life’s devices: the physical world of animals and plants*, Princeton University Press, 1988.
- [62] C. Xue and H. G. Othmer, *Multiscale models of taxis-driven patterning in bacterial populations*, SIAM J. Appl. Math. **70** (2009), 133–169.