Further Mathematical Biology
Lecture Notes

Prof Helen M Byrne
Michaelmas Term 2018

Contents

Chapter 1

Introduction

An outline for this course.

- We will observe that many phenomena in ecology, biology and biochemistry can be modelled mathematically.
- We will focus initially on systems where spatial variation is either absent or, at least, not important. In such cases only temporal evolution needs to be described, typically via ordinary differential equations.
- We are inevitably confronted with systems of ordinary differential equations, and thus we will study analytical techniques for extracting information from such equations.
- We will then consider systems where there is explicit spatial variation. The resulting models must also incorporate spatial effects.
- In ecological and biological applications the main physical phenomenon governing spatial variation is typically, but not exclusively, diffusion. Thus we invariably consider parabolic partial differential equations. Mathematical techniques will be developed to study such systems.
- These studies will be in the context of ecological, biological and biochemical applications. In particular we will draw examples from:
    1. enzyme-substrate dynamics and other biochemical reactions;
    2. Trans-membrane ion channels and nerve pulses;
    3. epidemics;
    4. the propagation of an advantageous gene through a population;
    5. biological pattern formation mechanisms;
    6. chemotaxis;
    7. tumour growth.

Acknowledgements: these lecture notes build on material originally developed by Professors Philip Maini, Ruth Baker, Eamonn Gaffney, Jon Chapman and Andrew Fowler. I am extremely grateful to them for allowing me to re-use and extend their notes.

1.1 References

The main references for this lecture course will be:

- J. D. Murray, Mathematical Biology, 3rd edition, Volume I [**?**].
- J. D. Murray, Mathematical Biology, 3rd edition, Volume II [**?**].

Other useful references include (but are no means compulsory):

- J. P. Keener and J. Sneyd, Mathematical Physiology [**?**].
- L. Edelstein-Keshet, Mathematical Models in Biology [**?**].
- N. F. Britton, Essential Mathematical Biology [**?**].

Note: If you have not taken the Part A short option *Modelling in Mathematical Biology*, you are encouraged to work through the lecture notes. They are made available on the course web-site.

Chapter 2

Enzyme kinetics

2.1 Introduction

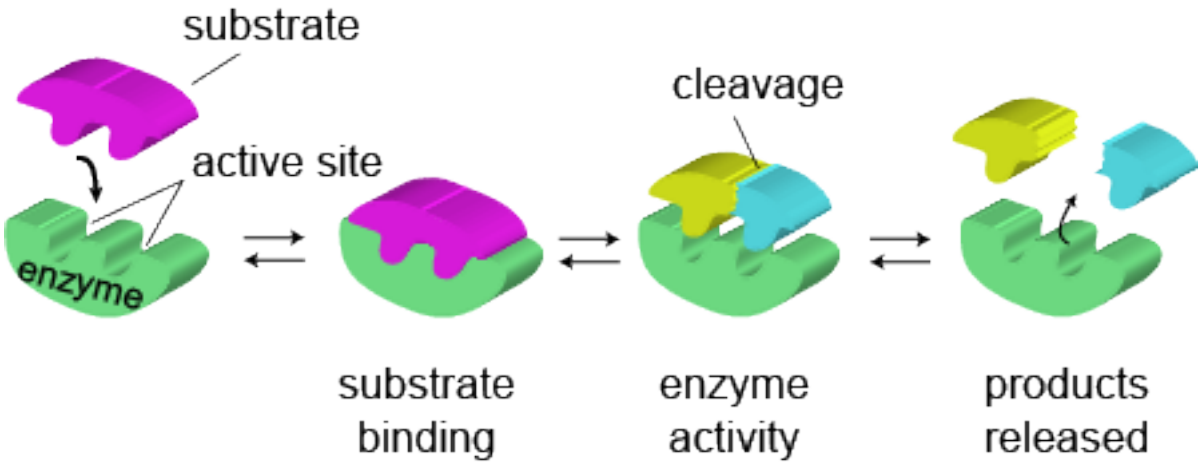


Figure 2.1: How do enzymes work? An enzyme has an active site where the substrates and enzyme fit together so that the substrates react. After the reaction, the products are released and the enzyme assumes its original shape.

Biochemical reactions are extremely important for correct biological function. For example, biochemical reactions are involved in:

- Metabolism and its control;
- Immunological responses;
- Cell-signalling processes.

Biochemical processes are often controlled by enzymes. Enzymes are proteins that catalyse biochemical reactions by lowering activation energy. Even when present in very small amounts, enzymes can have a dramatic effect on a system (see Figure 2.3). Table 2.1 illustrates how effective enzymes can be at accelerating reactions in biochemical systems.

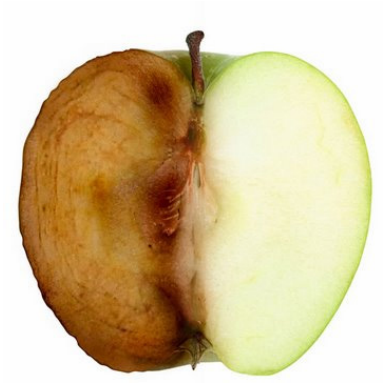


Figure 2.2: The enzyme catecholase catalyses a reaction between the molecule catechol and oxygen. The product of this reaction is polyphenol, the brown substance that accumulate when apples are exposed to air.

Enzyme	Substrate	Product	Rate without enzyme	Rate with enzyme	Accel. due to enzyme
Hexokinase	Glucose	Glucose 6-Phosphate	< 0.0000001	1300	> 13 billion
Phosphorylase	–	–	< 0.000000005	1600	> 320 billion
Alcohol Dehydrogenase	Ethanol	Acetaldehyde	< 0.000006	2700	> 450 million
Creatine Kinase	Creatine	Creatine Phosphate	< 0.003	40	$> 13,000$

Table 2.1: Examples illustrating the impact that enzymes can have on reaction rates.

In this chapter we will focus on enzyme kinetics. These can be thought of as a special case of an interacting species model (for details, see lecture notes on short course 'Mathematical Modelling in Biology'). In all cases we will neglect spatial variation.

References.

- J. D. Murray, Mathematical Biology, 3rd edition, Volume I, Chapter 6 [**?**].
- J. P. Keener and J. Sneyd, Mathematical Physiology, Chapter 1 [**?**].

2.2 The Law of Mass Action

Throughout this chapter, we will consider reactions involving m chemical species $C_1, \ldots, C_m$.

- The concentration of C_i, denoted c_i, is defined to be the number of molecules of C_i per unit volume.

- A standard unit of concentration is moles m^{-3}, often abbreviated to mol m^{-3}. Recall that 1 mole $= 6.023 \times 10^{23}$ molecules.

We will use **The Law of Mass Action** to construct the reaction rates. In words, the Law of Mass Action states:

The Law of Mass Action. A chemical reaction proceeds at a rate proportional to the concentrations of the participating reactants. The constant of proportionality is called the rate constant.

Suppose $C_1, \ldots, C_m$ undergo the reaction

$$\lambda_1 C_1 + \lambda_2 C_2 + \ldots + \lambda_m C_m \underset{k_b}{\overset{k_f}{\rightleftharpoons}} \nu_1 C_1 + \nu_2 C_2 + \ldots + \nu_m C_m. \tag{2.1}$$

The Law of Mass Action states that the forward reaction proceeds at rate

$$k_f c_1^{\lambda_1} c_2^{\lambda_2} \ldots c_m^{\lambda_m}, \tag{2.2}$$

while the backward (or reverse) reaction proceeds at the rate

$$k_b c_1^{\nu_1} c_2^{\nu_2} \ldots c_m^{\nu_m}, \tag{2.3}$$

where k_f and k_b are (non-negative) dimensional constants that must be determined empirically.

Note 1. Strictly speaking, to treat k_f, k_b as constant, we must assume that the temperature is constant. This is a good approximation for most biochemical reactions occurring in, for example, physiological systems. However, if one wanted to model reactions that produce significant amounts of heat for example, burning petrol, one must include temperature dependence in k_f and k_b and, subsequently, keep track of how the temperature of the system changes as the reaction proceeds. This typically makes the modelling significantly more difficult. Below we assume that we are dealing with systems where the temperature is approximately constant as the reaction proceeds.

Note 2. The Law of Mass Action for chemical reactions can be derived from statistical mechanics under quite general conditions (see for example L. E. Riechl, A Modern Course in Statistical Physics [**?**]).

Note 3. The Law of Mass Action is used in a variety of biological scenarios. For example, we use it to write down equations describing interactions between people infected with, and people susceptible to, a pathogen during an epidemic. In such circumstances the validity of the Law of Mass Action must be taken as a modelling assumption; in such scenarios one cannot rely on thermodynamic/statistical mechanical arguments to justify the Law of Mass Action.

Example: Stoichiometry. Suppose m molecules of A react reversibly with n molecules of B to create C:

$$mA + nB \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C$$

Then the Law of Mass Action takes the form

$$\left.\begin{aligned} \tfrac{da}{dt} &= -mk_1 a^m b^n + mk_{-1}c \\ \tfrac{db}{dt} &= -nk_1 a^m b^n + nk_{-1}c \\ \tfrac{dc}{dt} &= k_1 a^m b^n - k_{-1}c \end{aligned}\right\}$$

as m molecules of A and n molecules of B must collide to produce one molecule of C.

Note: Mass conservation supplies

$$a + mc = \text{constant}, \quad b + nc = \text{constant}.$$

2.3 Michaelis-Menten kinetics

Michaelis-Menten kinetics approximately describe the dynamics of a number of enzyme systems. The reactions are

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C, \tag{2.4}$$

$$C \xrightarrow{k_2} P + E, \tag{2.5}$$

where C represents the complex SE, and s, e, p and c denote the concentrations of S, E, P and C respectively. From the Law of Mass Action, we can derive the following ordinary differential equations for s, e, p and c:

$$\frac{\mathrm{d}s}{\mathrm{d}t} = -k_1 se + k_{-1}c, \tag{2.6}$$

$$\frac{\mathrm{d}c}{\mathrm{d}t} = k_1 se - k_{-1}c - k_2 c, \tag{2.7}$$

$$\frac{\mathrm{d}e}{\mathrm{d}t} = -k_1 se + k_{-1}c + k_2 c, \tag{2.8}$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = k_2 c. \tag{2.9}$$

Note: the equation for p decouples and, hence, we can neglect it (at least initially).

The initial conditions are:

$$s(0) = s_0, \quad e(0) = e_0 \ll s_0, \quad c(0) = 0, \quad p(0) = 0. \tag{2.10}$$

Key Point. In systems described by the Law of Mass Action, linear combinations of the variables are often conserved. In this example we have

$$\frac{\mathrm{d}}{\mathrm{d}t}(e+c)=0 \quad \Rightarrow \quad e=e_0-c, \tag{2.11}$$

and, hence, the equations simplify to:

$$\frac{\mathrm{d}s}{\mathrm{d}t} = -k_1(e_0-c)s+k_{-1}c, \tag{2.12}$$

$$\frac{\mathrm{d}c}{\mathrm{d}t} = k_1(e_0-c)s-(k_{-1}+k_2)c, \tag{2.13}$$

with the dynamics of p and e readily achievable once the dynamics of s and c are known.

2.3.1 Non-dimensionalisation

We non-dimensionalise as follows:

$$\tau=k_1e_0t, \quad u=\frac{s}{s_0}, \quad v=\frac{c}{e_0}, \quad \lambda=\frac{k_2}{k_1s_0}, \quad \epsilon \stackrel{def}{=} \frac{e_0}{s_0} \ll 1, \quad K \stackrel{def}{=} \frac{k_{-1}+k_2}{k_1s_0}, \tag{2.14}$$

which yields

$$u' = -u+(u+K-\lambda)v, \tag{2.15}$$

$$\epsilon v' = u-(u+K)v, \tag{2.16}$$

where $u(0)=1$, $v(0)=0$ and $\epsilon \ll 1$. Normally $\epsilon \sim 10^{-6}$. Setting $\epsilon=0$ yields

$$v=\frac{u}{u+K}, \tag{2.17}$$

which is inconsistent with the initial conditions! We have a singular perturbation problem; there must be a (boundary) region with respect to the time variable around $t=0$ where $v' \nsim \mathcal{O}(1)$. Indeed, for the stated initial conditions we find $v'(0) \sim \mathcal{O}(1/\epsilon)$, with $u(0)$, $v(0) \leq \mathcal{O}(1)$. This gives us the scaling we need to perform a singular perturbation analysis.

2.3.2 Singular perturbation investigation

We consider

$$\sigma=\frac{\tau}{\epsilon}, \tag{2.18}$$

with

$$u(\tau,\epsilon) = \tilde{u}(\sigma,\epsilon)=\tilde{u}_0(\sigma)+\epsilon\tilde{u}_1(\sigma)+\ldots, \tag{2.19}$$

$$v(\tau,\epsilon) = \tilde{v}(\sigma,\epsilon)=\tilde{v}_0(\sigma)+\epsilon\tilde{v}_1(\sigma)+\ldots. \tag{2.20}$$

Proceeding in the usual way, we find that $\tilde{u}_0$, $\tilde{v}_0$ satisfy

$$\frac{\mathrm{d}\tilde{u}_0}{\mathrm{d}\sigma}=0 \quad \Rightarrow \quad \tilde{u}_0=\text{constant}=1, \tag{2.21}$$

and

$$\frac{\mathrm{d}\tilde{v}_0}{\mathrm{d}\sigma}=\tilde{u}_0-(1+K)\tilde{v}_0=1-(1+K)\tilde{v}_0 \quad \Rightarrow \quad \tilde{v}_0=\frac{1-e^{-(1+K)\sigma}}{1+K}, \tag{2.22}$$

which gives us the *inner* solution.

To find the *outer* solution we expand

$$u(\tau,\epsilon) = u_0(\tau) + \epsilon u_1(\tau) + \ldots, \tag{2.23}$$
$$v(\tau,\epsilon) = v_0(\tau) + \epsilon v_1(\tau) + \ldots, \tag{2.24}$$

within the equations

$$u' = -u + (u + K - \lambda)v, \tag{2.25}$$
$$\epsilon v' = u - (u + K)v, \tag{2.26}$$

to find that

$$\frac{\mathrm{d}u_0}{\mathrm{d}\tau} = -u_0 + (u_0 + K - \lambda)v_0, \tag{2.27}$$

and

$$0 = u_0 - (u_0 + K)v_0. \tag{2.28}$$

This gives

$$v_0 = \frac{u_0}{u_0 + K} \quad \text{and} \quad \frac{\mathrm{d}u_0}{\mathrm{d}\tau} = -\frac{\lambda u_0}{u_0 + K}. \tag{2.29}$$

In order to match the solutions as $\sigma \to \infty$ and $\tau \to 0$ we require

$$\lim_{\sigma\to\infty} \tilde{u}_0 = \lim_{\tau\to 0} u_0 = 1 \quad \text{and} \quad \lim_{\sigma\to\infty} \tilde{v}_0 = \lim_{\tau\to 0} v_0 = \frac{1}{1+K}. \tag{2.30}$$

The resulting solution looks like that shown in Figure 2.3.

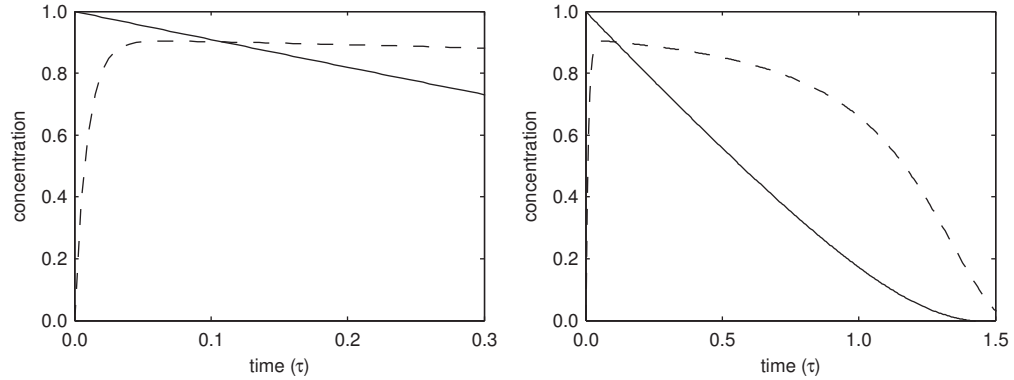


Figure 2.3: Numerical solution of the non-dimensional Michaelis-Menten equations clearly illustrating the two different time scales. The u dynamics are indicated by the solid line and the v dynamics by the dashed line. Parameters are $\epsilon = 0.01$, $K = 0.1$ and $\lambda = 1.0$.

Often the initial, fast, transient is not seen or modelled: we consider only the outer equations, with a suitably adjusted initial condition (ultimately determined from consistency/matching with the inner solution). In particular, we often use *Michaelis-Menten kinetics* where the equations are simply:

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -\frac{\lambda u}{u+K} \quad \text{with} \quad u(0) = 1 \quad \text{and} \quad v = \frac{u}{u+K}. \tag{2.31}$$

Definition. When the time derivative is fast, *i.e.* of the form

$$\epsilon \frac{\mathrm{d}v}{\mathrm{d}\tau} = g(u, v), \tag{2.32}$$

where $\epsilon \ll 1$ and $g(u, v) \sim \mathcal{O}(1)$, taking the temporal dynamics to be trivial,

$$\frac{\mathrm{d}v}{\mathrm{d}\tau} \simeq 0, \tag{2.33}$$

is known as the *pseudo-steady state hypothesis*. This is a common assumption in the literature. We have seen its validity for enzyme kinetics, at least away from the inner region.

Note. While the Michaelis-Menten kinetics derived above are a useful approximation, they hinge on the validity of the Law of Mass Action. Even in simple biological systems the Law of Mass Action may break down. One (of many) reasons, and one that is potentially relevant at the sub-cellular level, is that the system in question has too few reactant molecules to justify the statistical mechanical assumptions underlying the Lass of Mass Action. Another reason is that the reactants are not well-mixed, but vary spatially as well as temporally. We will see what happens in this case later in the course.

2.4 More complex systems

Here we consider a number of other simple systems involving enzymatic reactions. In each case the Law of Mass Action is used to write down a system of ordinary differential equations describing the dynamics of the various reactants (See J. Keener and J. Sneyd, Mathematical Physiology, for more details).

2.4.1 Several enzyme reactions and the pseudo-steady state hypothesis

We can have multiple enzymes. In general the system of equations reduces to

$$u' = f(u, v_1, \ldots, v_n), \tag{2.34}$$

$$\epsilon_i v_i' = g_i(u, v_1, \ldots, v_n), \tag{2.35}$$

for $i \in \{1, \ldots, n\}$, while the pseudo-steady state hypothesis gives a single ordinary differential equation

$$u' = f(u, v_1(u), \ldots, v_n(u)), \tag{2.36}$$

where $v_1(u), \ldots, v_n(u)$ are the appropriate roots of the equations

$$g_i(u, v_1, \ldots, v_n) = 0, \quad i \in \{1, \ldots, n\}. \tag{2.37}$$

Exercise. Consider an enzymatic reaction in which an enzyme can be activated or inactivated by a chemical X as follows:

$$E + X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E_1, \qquad E_1 + X \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E_2, \qquad E_1 + S \overset{k_3}{\to} P + Q + E.$$

Suppose further that X is supplied at a constant rate, and removed at a rate proportional to its concentration.

1. Write down differential equations for the evolution of E, E_1, E_2, X and S.
2. Show that $E + E_1 + E_2$ is a conserved quantity, E^* say.
3. Nondimensionalise the system, scaling E, E_1 and E_2 with E^*, X and S with $X_0 = X(0)$, and time with $1/(k_1 E^*)$. Assuming that $\delta = E^*/X_0 \ll 1$, use the resulting "quasi-steady" equations for the dimensionless quantities e, e_1, e_2 to solve for these variables in terms of x and s, and hence obtain the following system of two ODEs for x and s only:

   $$\frac{dx}{d\tau} = \alpha_0 - \nu_4 x - \frac{\kappa_3 x s}{\mu_1 + \kappa_3 s + x + \kappa_2 x^2/\mu_2}, \quad \frac{ds}{d\tau} = -\frac{\kappa_3 x s}{\mu_1 + \kappa_3 s + x + \kappa_2 x^2/\mu_2}.$$

   Identify all parameters and variables in these equations.

2.4.2 Allosteric enzymes

Here the binding of one substrate molecule at one site affects the binding of another substrate molecules at other sites. A typical reaction scheme is:

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C_1 \xrightarrow{k_2} P + E \tag{2.38}$$

$$S + C_1 \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} C_2 \xrightarrow{k_4} C_1 + P. \tag{2.39}$$

Further details on the investigation of such systems can be found in J. D. Murray, Mathematical Biology Volume I, and J. P. Keener and J. Sneyd, Mathematical Physiology.

2.4.3 Autocatalysis and activator-inhibitor systems

Here a molecule catalyses its own production. The simplest example is the reaction scheme

$$A + B \overset{k}{\to} 2B, \tag{2.40}$$

although the positive feedback in autocatalysis is usually ameliorated by inhibition from another molecule. This leads to an example of an activator-inhibitor system which can have a very rich behaviour. Other examples of these systems are given below.

Example 1

This model qualitatively incorporates activation and inhibition:

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{a}{b+v} - cu, \tag{2.41}$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = du - ev. \tag{2.42}$$

Example 2

This model is commonly referred to as the Gierer-Meinhardt model and was proposed in 1972:

$$\frac{\mathrm{d}u}{\mathrm{d}t} = a - bu + \frac{u^2}{v}, \tag{2.43}$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = u^2 - v. \tag{2.44}$$

Example 3

This model is commonly referred to as the Thomas model. Proposed in 1975, it is an empirical model based on a specific reaction involving uric acid and oxygen:

$$\frac{\mathrm{d}u}{\mathrm{d}t} = a - u - \rho R(u,v), \tag{2.45}$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \alpha(b-v) - \rho R(u,v), \tag{2.46}$$

where

$$R(u,v) = \frac{uv}{1+u+Ku^2}, \tag{2.47}$$

represents the interactive uptake.

Chapter 3

Ions and Excitable Systems

3.1 Introduction

3.1.1 Background

The cell membrane is a *phospholipid bilayer* separating the cell interior (the *cytoplasm*) from the extracellular environment. The membrane contains numerous proteins, and is approximately 7.5nm thick. The most important property of the cell membrane is its selective permeability: it allows the passage of some molecules but restricts the passage of others, thereby regulating the passage of materials into and out of the cell. Many substances penetrate the cell membrane at rates reflected by their diffusive behaviour in a pure phospholipid bilayer. However, certain molecules and ions such as glucose, amino acids and Na^+ pass through cell membranes much more rapidly, indicating that the membrane proteins selectively facilitate transport.

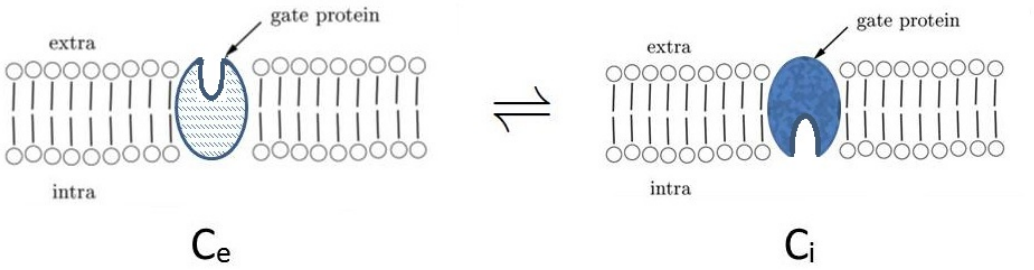


Figure 3.1: A schematic of the phospholipid membrane double layer, with a gating protein in one of two configurations, C_e and C_i, spanning the membrane, as part of a passive, carrier-mediated transport system.

The membrane contains water-filled pores with diameters of about 0.8nm, and protein-lined pores, called *channels* or *gates*, which allow the passage of specific molecules. Both the intracellular and extracellular environments comprise (among other things) a dilute aqueous solution of dissolved salts, mainly NaCl and KCl, which dissociate into Na^+, K^+ and Cl^- ions. The cell membrane acts as a barrier to the free flow of these ions and to the flow of water.

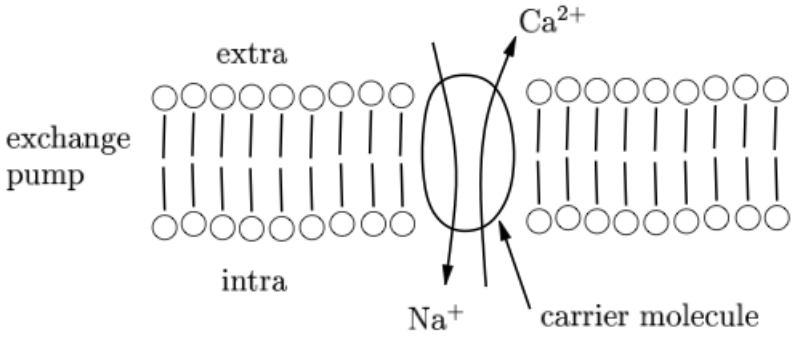


Figure 3.2: Representation of an exchange pump which actively transports across the cell membrane.

The mechanisms that facilitate transport across the cellular membrane can be divided into *active* and *passive* processes. Active processes requires energy expenditure, while passive processes result solely from the random motion of molecules, for example, diffusion.

Action potentials, or 'nerve impulses', are brief changes in the membrane potential of a cell produced by the flow of ionic current across the cell membrane. They enable communication by many cell types, including neurons, cardiac and muscle cells. In section 3.3 we will study the Fitzhugh Nagumo equations which describe nerve impulses in axons. First, we outline the basic physical concepts needed to study ion channels and nerve impulses, and introduce basic mathematical models for ionic currents and channel gating.

Note: Much of the material covered in this course is extended and studied in greater detail in the Part C Course *Mathematical Physiology*.

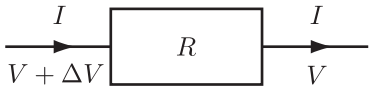
3.1.2 Basic Concepts

First, we note that:

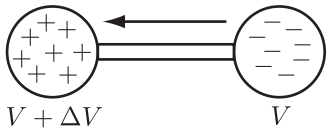
- numerous fundamental particles, ions and molecules have an electric charge, *e.g.* the electron, e^-, and the sodium ion, Na^+;
- it is an empirical fact that total charge is conserved;
- electric charges exert electrical forces on one another such that like charges repel and unlike charges attract. The electric potential, denoted V, is the potential energy of a unit of charge due to such forces and is measured in volts;
- a concentration of positive particles has a large *positive* potential, while a concentration of negative particles has a large, *but negative* potential;
- electric current is defined to be the rate of flow of electric charge, measured in Amps.

Resistance

Ohms Law, $\Delta V = IR$, holds in most situations, where ΔV is the change in potential, I is the current flowing and R, which may depend on material properties and geometries *but not on I nor V* is the resistance.



Key point. Suppose one uses a wire of low resistance to connect a region with a concentration of positive charges to a region with a concentration of negative charges. The charges will, very quickly, flow onto/off the wire until the potential is constant and there is no further flow of charge.



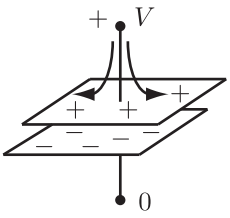
Ionic Currents and Capacitance

The flow of ions across the cell membrane due to concentration differences leads to a build up of charge near the cell membrane and a potential difference across the cell membrane. The cell membrane acts as a capacitor. The voltage (or potential difference) across any capacitor is related to the charge stored Q by

$$V = \frac{Q}{C},$$

where C is the capacitance.

Capacitance. A simple example of a capacitor is two conducting plates, separated by an insulator (e.g. an air gap).

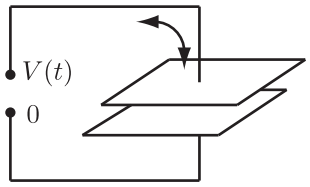


Connecting a battery to the plates, as illustrated, using wires of low resistance causes charge to flow onto/off the plates. It will equilibrate very quickly. Let $\pm Q_{eqm}$ denote the

charges stored on the two plates. The capacitance of the plates, C, is defined to be

$$C = \frac{Q_{eqm}}{V} > 0, \tag{3.1}$$

where C is a constant, independent of V. Thus the higher the capacitance, the better the plates are at storing charge, for a given potential.



If I is the ionic current out of the cell (the rate of flow of positive charges outwards),then the stored charge changes according to

$$I = -\frac{dQ}{dt}. \tag{3.2}$$

Thus, assuming the capacitance is constant

$$C\frac{dV}{dt} + I = 0. \tag{3.3}$$

This equation is the basis for much theoretical electrophysiology. The difference between various models relates to the expression used for the ionic current I.

The simplest models assume linear dependent of I on V (as in Ohms law). For a single ion S, with Nernst potential V_S, this gives an ionic current

$$I_S = g_S(V - V_S),$$

where the constant g_S is the ion-specific membrane conductance, since the current must vanish when $V = V_S$. If more than one ion is present, then the currents from the different ions are added together to produce the total ionic current.

3.2 Channel gating

3.2.1 Simple Gates

In practice, g_S is not constant: it depends on V and time t. One explanation for this is that the channels are not always open (they may be open or closed), and the transition rates between open and closed states depends on the potential difference, V. The membrane conductance may then be written as ng_S, where g_S is the constant conductance that would result if all channels were open, and n is the proportion of open channels.

For a generic ion, let n be the proportion of *open* ion channels. Denoting the open channels by O and the closed channels by C, the reaction scheme is simply

$$\mathrm{C} \underset{\beta(V)}{\overset{\alpha(V)}{\rightleftharpoons}} \mathrm{O} \tag{3.4}$$

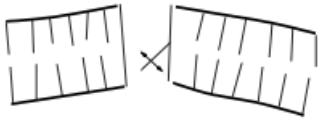


Figure 3.3: Schematic diagram of channel gating.

where $\alpha(V)$ and $\beta(V)$ represent voltage dependent rates of switching between the closed and open states. Using the law of mass action we obtain

$$\frac{dn}{dt} = \alpha(V)(1-n) - \beta(V)n, \tag{3.5}$$

or, equivalently,

$$\tau_n(V)\frac{dn}{dt} = n_\infty(V) - n, \tag{3.6}$$

where $n_\infty(V) = \alpha/(\alpha+\beta)$ is the equilibrium value of n and $\tau_n(V) = 1/(\alpha+\beta)$ is the timescale for approach to this equilibrium (Note: both n_∞ and τ_n can be determined experimentally).

3.2.2 Multiple gates

The simple model presented above can be generalised to channels with multiple identical subunits, each of which can be in either the open or closed state.

We assume that the channel consists of two "gates", both of which may exist in open or closed states. The ion channel is open only if *both* "gates" are open; the ion channel is closed if any one gate within the ion channel is closed. If we denote by S_i ($i \in \{0,1,2\}$) the proportion of channels with exactly i gates open, then

$$S_0 + S_1 + S_2 = 1 \tag{3.7}$$

and the reaction scheme

$$S_0 \underset{\beta(V)}{\overset{2\alpha(V)}{\rightleftharpoons}} S_1 \underset{2\beta(V)}{\overset{\alpha(V)}{\rightleftharpoons}} S_2. \tag{3.8}$$

The 2s arise because there are two possible states with one gate open and one gate closed. Since each gate is identical we lump these two states into one variable S_1. Using mass action kinetics gives

$$\frac{dS_0}{dt} = \beta(V)(1-S_0-S_2) - 2\alpha(V)S_0, \tag{3.9}$$

$$\frac{dS_2}{dt} = \alpha(V)(1-S_0-S_2) - 2\beta(V)S_2, \tag{3.10}$$

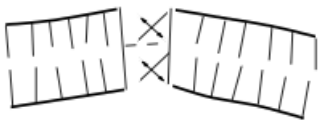


Figure 3.4: Schematic diagram of two identical gate units.

where, in general, V is a function of time (and possibly space too). We could also write down an equation for S_1, but it is superfluous since S_1 can be determined from (3.7).

Let n denote the *proportion of open gates;* then

$$\frac{dn}{dt} = \alpha(V)(1-n) - \beta(V)n. \tag{3.11}$$

Simple substitution shows that (3.7) , (3.9) and (3.10) are satisfied by

$$S_0 = (1-n)^2, \quad S_1 = 2n(1-n), \quad S_2 = n^2. \tag{3.12}$$

In fact, it is possible to derive a stronger result. Suppose

$$S_0 = (1-n)^2 + y_0, \quad S_2 = n^2 + y_2, \tag{3.13}$$

so that $S_1 = 2n(1-n) - y_0 - y_2$. It then follows that

$$\frac{dy_0}{dt} = -2\alpha y_0 - \beta(y_0 + y_2), \quad \frac{dy_2}{dt} = -\alpha(y_0 + y_2) - 2\beta y_2. \tag{3.14}$$

This linear system has eigenvalues $-(\alpha+\beta)$, $-2(\alpha+\beta)$, and so y_0, y_2 decay exponentially to zero. Thus, regardless of the initial conditions, the solution will still approach exponentially that given by (3.11) and (3.12) (*i.e.* (3.11) and (3.12) define a stable invariant manifold of the full system (3.7), (3.9) and (3.10)).

The analysis of a two-gated channel generalises easily to channels containing three or more gates. In the case of k identical gates the fraction of open channels is n^k, where n again satisfies (3.11). We will find that a model with 4 gates agree with empirical observations of K^+ channels, and will be used below.

3.2.3 Non-identical gates

Often channels are controlled by multiple proteisn. Each protein may control a set of identical gates, but the gates of each protein are assumed to be different and independent. Consider, for example, a channel with two types of gate, m and h say, each of which may be open or closed. For illustrative purposes, we will assume that the channel has two m subunits and one h subunit. With S_{ij} the proportion of channels with $i \in \{0,1,2\}$ of the m-gates open and $j \in \{0,1\}$ of the h-gates open, the reaction scheme is

Simple substitution shows that the corresponding law of mass action equations are satisfied by

$$S_{00} = (1-m)^2(1-h), \quad S_{10} = 2m(1-m)(1-h), \quad S_{20} = m^2(1-h), \tag{3.15}$$

$$S_{01} = (1-m)^2 h, \quad S_{11} = 2m(1-m)h, \quad S_{21} = m^2 h, \tag{3.16}$$

so that the proportion of open channels is m^2h, provided

$$dm/dt = \alpha(V)(1-m) - \beta(V)m \tag{3.17}$$

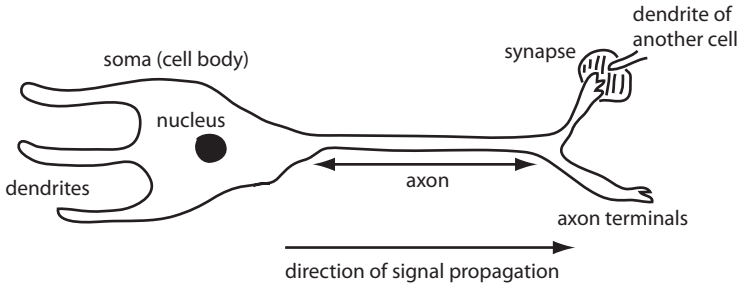
$$dh/dt = \gamma(V)(1-h) - \delta(V)h. \tag{3.18}$$

As before, such solutions form a stable invariant manifold.

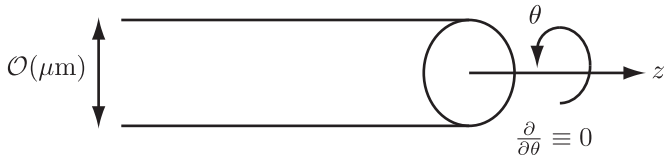
$$\begin{array}{ccccc} S_{00} & \underset{\beta}{\overset{2\alpha}{\rightleftharpoons}} & S_{10} & \underset{2\beta}{\overset{\alpha}{\rightleftharpoons}} & S_{20} \\ \gamma \downarrow\uparrow \delta & & \gamma \downarrow\uparrow \delta & & \gamma \downarrow\uparrow \delta \\ S_{01} & \underset{\beta}{\overset{2\alpha}{\rightleftharpoons}}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

3.3 The Fitzhugh Nagumo equations

An axon is a part of a nerve cell. The nerve signal along the axon is, in essence, a propagating pulse in the potential difference across the plasma (*i.e.* outer) membrane of the axon. This potential difference, V, arises due to the preferential permeability of the axon plasma membrane which allows potassium and sodium ions, K^+ and Na^+, to pass through the membrane at rates which differ between the two ions and vary with V. In the rest state, $V = V_{rest} \simeq -70\text{mV}$ (millivolts); in a nerve signal pulse in V rises to a peak of $\sim 15\text{mV}$. We wish to model this pulse.

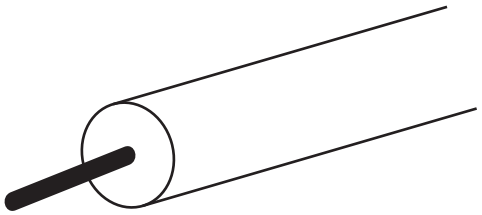


The axon can be viewed as a cylindrical tube. An axon is axisymmetric, so we may ignore θ dependence in our models of the axon.

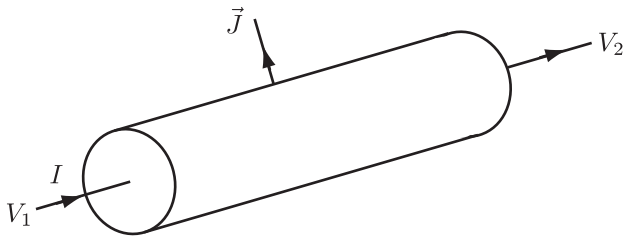


3.3.1 Deducing the FitzHugh Nagumo Equations

A common, simplifying, experimental scenario is to *space-clamp* the axon, *i.e.* to place a conducting wire along the axon's axis of symmetry.



- The interior of the axon will quickly equilibrate, and there will be no spatial variation in the potential difference, nor any current, along the inside of the axon.



- Thus, by conservation of charge, the *total current flowing across the axon membrane must be zero.*

- **Note:** any changes in the interior due, for example to the axon allowing K^+ and Na^+ to pass through its membrane, will occur on a much slower timescale and, hence, one has that the interior of the space-clamped axon has no spatial variation in its potential difference, no current flowing along the inside of the axon, and, most importantly, the *total* current flowing through the axon membrane is zero.

The basic model for the *space-clamped* axon plasma membrane potential is given by

$$\begin{aligned} 0 &= \text{total transmembrane current per unit area,} \\ &= c\frac{\mathrm{d}V}{\mathrm{d}t} + I_{Na} + I_K + I_0 + I_{applied}(t), \end{aligned} \tag{3.19}$$

where

- $I_{applied}(t)$ is the applied current, *i.e.* the current injected through the axon plasma membrane in the experiment, which is only function of time in most experimental set-ups. We will take $I_{applied}(t) = 0$ below.

- $c\,\mathrm{d}V/\mathrm{d}t$ is the capacitance current through a unit area of the membrane. *Recall:*

  Rate of flow on/off capacitor $= C\,\mathrm{d}V/\mathrm{d}t$.

  Therefore rate of flow of charge per unit area of membrane $= c\,\mathrm{d}V/\mathrm{d}t$ where c is the membrane capacitance per unit area.

- I_{Na}, I_K are the voltage dependent Na^+ and K^+ currents. I_0 is a voltage dependent background current.

  These currents actually take complicated forms, involving numerous other variables which satisfy complex equations, that can be simplified, if somewhat crudely. An excellent account is given in T. F. Weiss, Cellular Biophysics.

The resulting equations may be written in terms of the non-dimensional variables $v = (V - V_{rest})/|V_{rest}|$ and $\tau = t/T$ where $T = 6$ ms, the time scale of a typical nerve pulse. These equations state

$$\epsilon\frac{\mathrm{d}v}{\mathrm{d}\tau} = Av(\delta - v)(v - 1) - n, \tag{3.20}$$

$$\frac{\mathrm{d}n}{\mathrm{d}\tau} = -\gamma n + v, \tag{3.21}$$

where A, γ, ϵ, δ are positive parameters such that $A, \gamma \sim \mathcal{O}(1)$, $0 < \epsilon \ll \delta \ll 1$.

Key point. Equations, (3.20)-(3.21) are known as the **Fitzhugh Nagumo equations**. They model (in approximate terms!) the spatially independent behaviour of a *space-clamped* axon.

3.3.2 A brief analysis of the Fitzhugh Nagumo equations

We have

$$\epsilon\frac{\mathrm{d}v}{\mathrm{d}\tau} = Av(\delta - v)(v - 1) - n, \tag{3.22}$$

$$\frac{\mathrm{d}n}{\mathrm{d}\tau} = -\gamma n + v, \tag{3.23}$$

where A, γ, ϵ, δ are positive parameters such that $A, \gamma \sim \mathcal{O}(1)$, $0 < \epsilon \ll \delta \ll 1$.

The (n, v) phase plane

The *nullclines* of equations (3.20)-(3.21) are the lines where $\dot{v} = 0$ and $\dot{n} = 0$. A plot of the nullclines separates the (v, n) phase plane into four regions, as shown in Figure 3.5.

There are several things to note about the dynamics.

- There is one stationary point which is a stable focus.

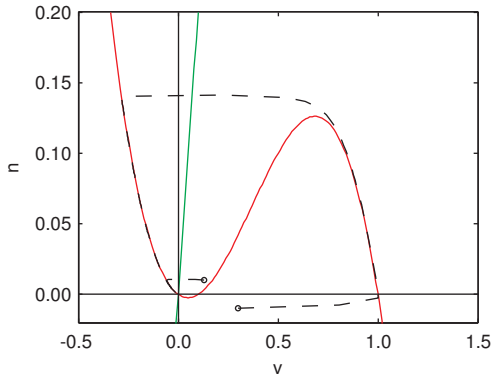


Figure 3.5: The phase plane for the Fitzhugh-Nagumo equations with the v nullcline shown in red and the n nullcline in green. The trajectories for two different initial perturbations from the steady state are shown as dashed lines. Parameters are as follows: $A = 1$, $\gamma = 0.5$, $\delta = 0.1$ and $\epsilon = 0.001$.

- Thus, with initial conditions sufficiently close to the stationary point, the system evolves to the stationary point in a simple manner.
- Consider initial conditions with $n \sim 0$, but v increased sufficiently. The system does not simply relax back to the equilibrium. However, one can understand the qualitative behaviour of the system by considering the phase plane.
- We anticipate that $v = (V - V_{rest})/|V_{rest}|$ behaves in the manner shown in Figure 3.6 for a sufficiently large perturbation in v.

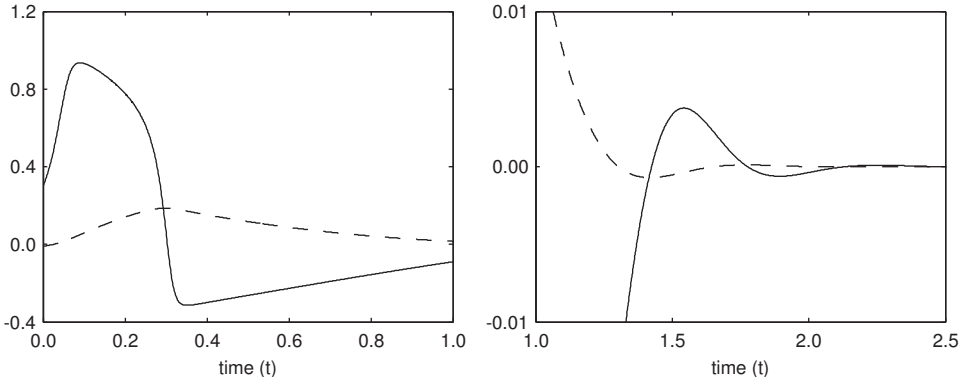


Figure 3.6: Solutions of the Fitzhugh Nagumo equations with v dynamics indicated by the solid line and n dynamics by the dashed line. The right-hand figure shows the oscillations that arise for large t. Parameters are as follows: $A = 1$, $\gamma = 0.5$, $\delta = 0.1$ and $\epsilon = 0.01$.

- This is essentially a nerve pulse (although because of the space clamping all of the nerve axon is firing at once).

Definition. A system which, for a sufficiently large perturbation from a stationary point, undergoes a large change before eventually returning to the same stationary point is referred to as *excitable*.

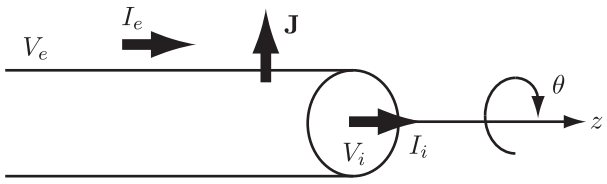
3.4 Modelling nerve signal propagation - NON-EXAMINABLE

In the following, we generalise the ideas we have seen for modelling the plasma membrane potential of an axon to scenarios where this potential can vary along the axon. This material is not examinable and will be better understood after we have covered spatial models. It is included here for completeness only (and to whet your appetite for the Mathematical Physiology Course at Part C).

3.4.1 The cable model

In the model we are about to develop we make following assumptions.

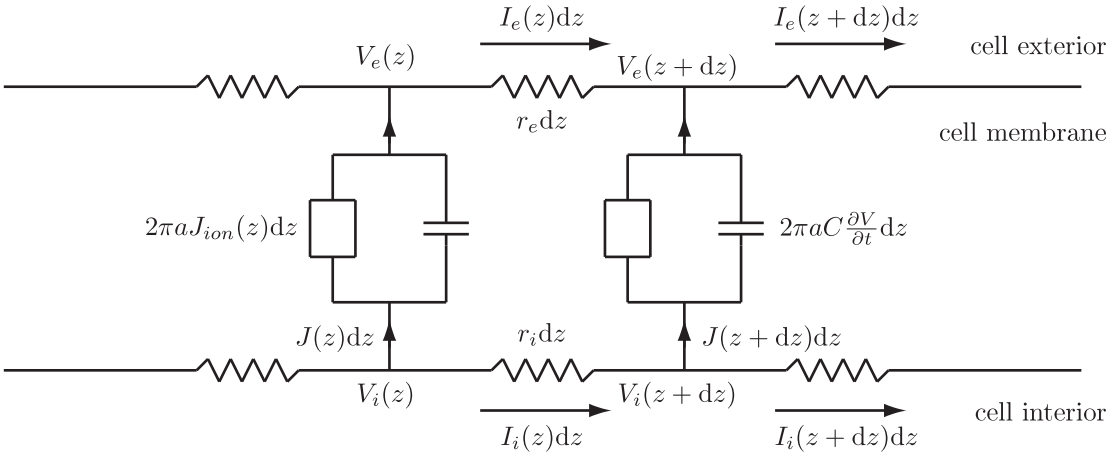
- The cell membrane is a cylindrical membrane separating two conductors of electric current, namely the extracellular and intracellular mediums. These are assumed to be homogeneous and to obey Ohm's law.
- The model has no θ dependence.
- A circuit theory description of current and voltages is adequate, *i.e.* quasi-static terms of Maxwell's equations are adequate; for example, electromagnetic radiation effects are totally negligible.
- Currents flow through the membrane in the radial direction only.
- Currents flow through the extracellular medium in the axial direction only and the potential in the extracellular medium is a function of z only. Similarly for the potential in the intracellular medium.



These assumptions are appropriate for unmyelinated nerve axons. Deriving the model requires considering the following variables:

- $I_e(z,t)$ – external current;
- $I_i(z,t)$ – internal current;
- $J(z,t)$ – total current through the membrance per unit length;

- $J_{ion}(z,t)$ – total ion current through the membrance per unit area;
- $V(z,t) = V_i(z,t) - V_e(z,t)$ – transmembrane potential;
- r_i – internal resistance per unit length;
- r_e – external resistance per unit length;
- C – membrane capacitance per unit area.



Consider the axial current in the extracellular medium, which has resistance r_e per unit length. We have

$$V_e(z+\mathrm{d}z) - V_e(z) = -r_e I_e(z)\mathrm{d}z \quad \Rightarrow \quad r_e I_e(z) = -\frac{\partial V_e}{\partial z}, \tag{3.24}$$

where the minus sign appears because of the convention that positive current is a flow of positive charges in the direction of increasing z. Hence, if $V_e(z+\mathrm{d}z) > V_e(z)$ then positive charges flow in the direction of decreasing z giving a negative current. Similarly,

$$r_i I_i(z) = -\frac{\partial V_i}{\partial z}. \tag{3.25}$$

Using conservation of current, we have

$$I_e(z+\mathrm{d}z,t) - I_e(z,t) = J(z,t)\mathrm{d}z = I_i(z,t) - I_i(z+\mathrm{d}z,t), \tag{3.26}$$

which gives

$$J(z,t) = -\frac{\partial I_i}{\partial z} = \frac{\partial I_e}{\partial z}. \tag{3.27}$$

Hence

$$J = \frac{1}{r_i}\frac{\partial^2 V_i}{\partial z^2} = -\frac{1}{r_e}\frac{\partial^2 V_e}{\partial z^2}, \tag{3.28}$$

and so

$$\frac{\partial^2 V}{\partial z^2} = (r_i + r_e)J. \tag{3.29}$$

Putting this all together gives

$$0 = -\frac{\partial(I_i + I_e)}{\partial z} = \frac{\partial}{\partial z}\left(\frac{1}{r_e}\frac{\partial V_e}{\partial z} + \frac{1}{r_i}\frac{\partial V_i}{\partial z}\right) = \left(\frac{r_e + r_i}{r_e r_i}\right)\frac{\partial^2 V_e}{\partial z^2} + \frac{1}{r_i}\frac{\partial^2 V}{\partial z^2}, \tag{3.30}$$

and so

$$0 = \frac{1}{r_i}\frac{\partial^2 V}{\partial z^2} - \left(\frac{r_e + r_i}{r_i}\right)\frac{\partial I_e}{\partial z} = \frac{1}{r_i}\left(\frac{\partial^2 V}{\partial z^2} + (r_e + r_i)J(z,t)\right). \tag{3.31}$$

We also have that

$$J(z,t) = 2\pi a\left(J_{ion}(V,z,t) + C\frac{\partial V}{\partial t}\right), \tag{3.32}$$

and finally, therefore,

$$\frac{1}{2\pi a(r_i + r_e)}\frac{\partial^2 V}{\partial z^2} = C\frac{\partial V}{\partial t} + J_{ion}(V,z,t). \tag{3.33}$$

This gives an equation relating the cell plasma membrane potential, V, to the currents across the cell plasma membrane due to the flow of ions, $J_{ion}(V,z,t)$.

Note 1. Note even though, physically, there is no diffusion, we still have a parabolic partial differential equation, so the techniques we have previously studied are readily applicable.

Note 2. From the above equation one can model cell plasma membrane potentials given suitable initial and boundary conditions, and a suitable expression for $J_{ion}(z,t)$.

We use the same expression for $J_{ion}(z,t)$, *i.e.* the expression for $I_{Na} + I_K + I_0$ as in the Fitzhugh Nagumo model of a space-clamped axon.

Thus with $v = (V - V_{rest})/|V_{rest}|$ and $x = Kz$, where K is a constant, we have

$$\epsilon\frac{\partial v}{\partial \tau} = \epsilon^2\frac{\partial^2 v}{\partial x^2} + Av(\delta - v)(v - 1) - n, \tag{3.34}$$

$$\frac{\mathrm{d}n}{\mathrm{d}\tau} = -\gamma n + v, \tag{3.35}$$

where $0 < A, \gamma \sim \mathcal{O}(1)$, $0 < \epsilon \ll \delta \ll 1$.

Note that K has been chosen so that the coefficient infront of the v_{xx} term is ϵ^2. This means, with respect to such variables, the front of the nerve pulse is extremely sharp. Hence, for such a scaling to exist, the extent of the nerve pulse must be less than ϵL, where L is the length of the axon; this constraint holds true for typical parameter estimates. The reason for the choice of this scaling is simply mathematical convenience in a travelling wave analysis.

We are interested in *nerve pulses*, so we take the boundary conditions to be n, $v \to 0$ as $x \to \pm\infty$.

We thus again have a system of parabolic partial differential equations to solve, and we are particularly interested in travelling pulse solutions. This entails that a travelling wave

analysis would be most insightful. With the travelling wave coordinate $y = x - c\tau$ and $v(y) = v(x, \tau)$, $n(y) = n(x, \tau)$, we obtain

$$\epsilon^2 \frac{\mathrm{d}^2 v}{\mathrm{d}y^2} + \epsilon c \frac{\mathrm{d}v}{\mathrm{d}y} + Av(\delta - v)(v - 1) - n = 0, \tag{3.36}$$

$$c\frac{\mathrm{d}n}{\mathrm{d}y} - \gamma n + v = 0. \tag{3.37}$$

We have $0 < A \sim \mathcal{O}(1)$, $0 < \gamma^{-1}$, $\delta, \epsilon \ll 1$. One can readily investigate these ordinary differential equations to find that the travelling wave speed is *unique*, giving a unique prediction for the speed of a nerve pulse in terms of biophysical parameters.

Chapter 4

Introduction to spatial variation

We have considered biological, biochemical and ecological phenomena for which spatial effects are not important. This is, however, often not the case. Consider a biochemical reaction as an example. Suppose this reaction involves solutes in a relatively large, *unstirred* solution. Then the system dynamics are governed not only by the rates at which the biochemicals react, but also by possible spatial variation in solute concentrations; in such cases, diffusion of the reactants can occur. Modelling such systems requires that we account for both reaction and diffusion.

A similar problem arises in population and ecological models when we wish to describe the tendency of a species to spread into a region it has not previously populated. Notable examples include ecological invasions, where one species invades another's territory (as with grey and red squirrels in the UK [**?**]), or the spread of disease. When developing mathematical descriptions of some, though by no means all, of these ecological and disease-spread systems, the appropriate transport mechanism is again diffusion; when modelling such systems we must include both reaction and diffusion.

In addition, motile cells can move in response to external influences, such as chemical concentrations, light, mechanical stress and electric fields, among others. Of particular interest is modelling when motile cells respond to gradients in chemical concentrations, a process known as **chemotaxis**; we will also consider this scenario.

In the following chapters, we will learn how to model such phenomena and how (when possible) to solve the resulting partial differential equations, for a range of models drawn from biology, biochemistry and ecology.

Most of the PDEs that we will study can be written in the general form

$$\begin{pmatrix} \text{rate of change} \\ \text{of species} \end{pmatrix} = \begin{pmatrix} \text{net movement/flux} \\ \text{of species} \end{pmatrix} + \begin{pmatrix} \text{net rate of production} \\ \text{of species} \end{pmatrix}$$

This is the **Principle of Mass Balance**.

As an example, consider Fisher's equation for growth and spread of a population in one spatial dimension:

$$\frac{\partial N}{\partial t} = \underbrace{D\frac{\partial^2 N}{\partial x^2}}_{\text{diffusion}} + \underbrace{rN\left(1-\frac{N}{K}\right)}_{\text{logistic growth}}$$

References.

- J. D. Murray, Mathematical Biology Volume I, Chapter 11 [**?**].
- N. F. Britton, Essential Mathematical Biology, Chapter 5 [**?**].

4.1 Derivation of the reaction-diffusion equations

Let $i \in \{1, \ldots, m\}$. Suppose the chemical species C_i, of concentration c_i, is undergoing a reaction such that, in the absence of diffusion, one has

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} = R_i(c_1, c_2, \ldots, c_m). \tag{4.1}$$

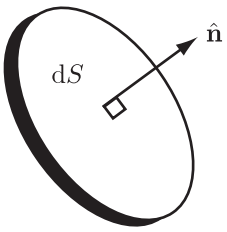
In equation (4.1, $R_i(c_1, c_2, \ldots, c_m)$ is the total rate of production/destruction of C_i *per unit volume*, *i.e.* it is the rate of change of the concentration c_i.

Let t denote time, and $\boldsymbol{x}$ denote the position vector of a point in space. We define

- $c(\boldsymbol{x}, t)$ to be the concentration of a chemical (typically measured in mol m^{-3}).
- $\boldsymbol{q}(\boldsymbol{x}, t)$ to be the flux of the same chemical (typically measured in mol m^{-2} s^{-1}).

Now the flux of a chemical is defined to be such that, for a given infinitesimal surface element, of area $\mathrm{d}S$ and unit normal $\hat{\mathbf{n}}$, the amount of chemical flowing through the surface element in an infinitesimal time interval, of duration $\mathrm{d}t$, is given by

$$\hat{\mathbf{n}} \cdot \boldsymbol{q}\ \mathrm{d}S\mathrm{d}t. \tag{4.2}$$



Definition. *Fick's Law of Diffusion* relates the flux $\boldsymbol{q}$ to the gradient of c via

$$\boldsymbol{q} = -D\nabla c, \tag{4.3}$$

where D, the diffusion coefficient, is independent of c and ∇c.

Using the Principle of Mass Balance, we have, for any closed volume V (fixed in time and space), with bounding surface ∂V,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V c_i \,\mathrm{d}V = -\int_{\partial V} \boldsymbol{q}\cdot\mathbf{n}\,\mathrm{d}S + \int_V R_i(c_1, c_2, \ldots, c_m)\,\mathrm{d}V, \quad i \in \{1, \ldots, m\}. \tag{4.4}$$

Hence

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_V c_i \,\mathrm{d}V &= -\int_V \nabla\cdot\boldsymbol{q}\,\mathrm{d}V + \int_V R_i(c_1, c_2, \ldots, c_m)\,\mathrm{d}V &(4.5)\\ &= \int_V \{\nabla\cdot(D_i\nabla c_i) + R_i(c_1, c_2, \ldots, c_m)\}\,\mathrm{d}V, &(4.6)\end{aligned}$$

and thus for any closed volume, V, with surface ∂V, we have

$$\int_V \left\{\frac{\partial c_i}{\partial t} - \nabla\cdot(D_i\nabla c_i) - R_i\right\}\mathrm{d}V = 0, \quad i \in \{1, ..., m\}. \tag{4.7}$$

Hence

$$\frac{\partial c_i}{\partial t} = \nabla\cdot(D_i\nabla c_i) + R_i, \quad x \in \mathcal{D}, \tag{4.8}$$

which constitutes a system of reaction-diffusion equations for the m chemical species in the finite domain $\mathcal{D}$. Such equations must be supplemented with initial and boundary conditions for each of the m chemicals.

Warning. Given, for example, that

$$\int_0^{2\pi} \cos\theta\,\mathrm{d}\theta = 0 \quad \not\Rightarrow \quad \cos\theta = 0, \quad \theta \in [0, 2\pi], \tag{4.9}$$

are you sure one can deduce equation (4.8)?

Suppose

$$\frac{\partial c_i}{\partial t} - \nabla\cdot(D_i\nabla c_i) - R_i \neq 0, \tag{4.10}$$

at some $\boldsymbol{x} = \boldsymbol{x}^*$. Without loss of generality, we can assume the above expression is positive *i.e.* the left-hand side of equation (4.10) is positive.

Then $\exists\,\epsilon > 0$ such that

$$\frac{\partial c_i}{\partial t} - \nabla\cdot(D_i\nabla c_i) - R_i > 0, \tag{4.11}$$

for all $\boldsymbol{x} \in \mathcal{B}_\epsilon(\boldsymbol{x}^*)$.

In this case

$$\int_{\mathcal{B}_\epsilon(\boldsymbol{x}^*)} \left[\frac{\partial c_i}{\partial t} - \nabla\cdot(D_i\nabla c_i) - R_i\right]\mathrm{d}V > 0, \tag{4.12}$$

contradicting our original assumption, equation (4.7).

Hence our initial supposition is false and equation (4.8) holds for $\boldsymbol{x} \in \mathcal{D}$.

Remark. With one species, with a constant diffusion coefficient, in the absence of reactions, we have the diffusion equation. In one dimension this reduces to

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}. \tag{4.13}$$

For a given length scale, L, and diffusion coefficient, D, the timescale of the system is $T = L^2/D$. For a cell, $L \sim 10^{-5}\text{m} = 10^{-3}\text{cm}$, and for a typical protein $D \sim 10^{-7}\text{cm}^2\text{s}^{-1}$ would not be unreasonable. Thus the timescale for diffusion to homogenise spatial gradients of this protein within a cell is

$$T \sim \frac{L^2}{D} \sim \frac{10^{-6}\ \text{cm}^2}{10^{-7}\ \text{cm}^2\ \text{s}^{-1}} \sim 10\ \text{s}, \tag{4.14}$$

therefore we can often neglect diffusion in a cell. However, as the scale doubles the time scale squares *e.g.* $L \to L \times 10 \Rightarrow T \to T \times 10^2$ and $L \to L \times 10^2 \Rightarrow T \to T \times 10^4$.

Note. The above derivation generalises to situations more general than modelling chemical or biochemical diffusion. For example, let $I(x, y, t)$ denote the number of infected people per unit area. Assume the infectives, on average, spread out via a random walk and interact with susceptibles, as described by the Law of Mass Action (see Section (5.2.1)). Then the flux of infectives, $\boldsymbol{q}_I$, is given by

$$\boldsymbol{q}_I = -D_I \nabla I, \tag{4.15}$$

where D_I is a constant, with dimensions of (length)^2 (time)$^{-1}$. Thus, via precisely the same ideas and arguments as above, we have that

$$\frac{\partial I}{\partial t} = \nabla \cdot (D_I \nabla I) + rIS - aI, \tag{4.16}$$

where $S(x, y, t)$ is the number of susceptibles per unit area, and r is the rate at which susceptibles become infected on contact with infecteds, and a is the rate at which infecteds recover from the disease (see Section 5.2.1 for more details).

Fisher's Equation. A common example is the combination of logistic growth and diffusion which, in one spatial dimension, gives rise to Fisher's Equation:

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + ru\left(1 - \frac{u}{K}\right). \tag{4.17}$$

This equation was first proposed to model the spread of an advantageous gene through a population. See Section 5.1 for more details.

4.2 Chemotaxis

As briefly mentioned earlier, motile cells can move in response to spatial gradients in chemical concentrations, a process known as chemotaxis. This leads to slightly more complicated transport equations, as we shall see.

The diffusive flux for the population density of the cells, n, is as previously: $\boldsymbol{J}_D = -D_n \nabla n$. The flux due to chemotaxis (assuming it is an attractant rather than a repellent) takes the form:

$$\boldsymbol{J}_C = n\chi(a)\nabla a = n\nabla\Phi(a), \tag{4.18}$$

where a is the chemical concentration and $\Phi(a)$ increases monotonically with a. Clearly $\chi(a) = \Phi'(a)$; the cells move in response to a gradient of the chemical in the direction in which the function $\Phi(a)$ is increasing at the fastest rate.

Thus the total flux $\boldsymbol{J}$ is

$$\boldsymbol{J} = \boldsymbol{J}_D + \boldsymbol{J}_C = -D_n \nabla n + n\chi(a)\nabla a. \tag{4.19}$$

If we assume that the behaviour of the cells is dominated by their diffusive and chemotactic transport, and their rate of reproduction and/or death, then we can use the principle of mass balance to derive a PDE that describes how their distribution changes over time. We need an additional (reaction-diffusion) equation for the chemical: we assume it diffuses and, typically, is secreted and degrades. In this way, we arrive at the following equations for the cells n and the cell-derived chemical a:

$$\frac{\partial n}{\partial t} = \nabla \cdot (D_n \nabla n) - \nabla \cdot (n\chi(a)\nabla a) + f(n, a), \tag{4.20}$$

$$\frac{\partial a}{\partial t} = \nabla \cdot (D_a \nabla a) + \lambda n - \mu a. \tag{4.21}$$

In the above the above $f(n, a)$ is often taken to be a logistic growth term while the function $\chi(a)$ describing chemotaxis has many forms, including

$$\chi(a) = \frac{\chi_0}{a}, \tag{4.22}$$

$$\chi(a) = \frac{\chi_0}{(k + a)^2}, \tag{4.23}$$

where the latter represents a receptor law, with $\Phi(a)$ taking a Michaelis-Menten form [**?**].

4.3 Positional Information and Pattern Formation

Patterns are ubiquitous in biology. Consider, for example, animal coat markins on tigers, leopards and tropical fish. Consider, also, the well-defined pattern of bones and digits (fingers, thumbs and toes) and teeth that appear during human development. There are two main theories about how such patterns arise:

- Alan Turing's concept of **diffusion-driven instability** which we will study in chapter 6 (see, also, Turing, 1952);
- Lewis Wolpert's theory of **positional information** which is also known as the French Flag Model (see: Wolpert, 1969 and Wolpert 1971); we study this theory below.

The French Flag Model

Consider a one-dimensional chain of cells that occupies the region $0 \leq x \leq L$. Suppose that a morphogen, $m(x,t)$, or signalling molecule enters the domain through $x = 0$, diffuses across the domain (with diffusion coefficient D), and is removed at $x = L$. If we assume that initially there is no mophogen in the domain, then the distribution of $m(x,t)$ can be described by the following equations

$$\frac{\partial m}{\partial t} = D\frac{\partial^2 m}{\partial x^2}, \tag{4.24}$$

$$\text{with} \quad m(0,t) = m_0, \quad m(L,t) = 0, \quad m(x,0) = 0, \tag{4.25}$$

where the positive constant m_0 defines the morphogen concentration at $x = 0$.

We assume that the morphogen rapidly establishes a fixed spatial profile which we determine by setting $\frac{\partial}{\partial t} = 0$ in equation (4.24):

$$0 = \frac{d^2 m}{dx^2} \quad \Rightarrow m(x) = m_0\left(1 - \frac{x}{L}\right) \tag{4.26}$$

since $m(x = 0) = m_0$ and $m(x = L) = 0$. Cells on the left (near $x = 0$) sense high morphogen levels and respond in some way (e.g. they turn blue). Cells in the centre and on the right sense intermediate and low levels of morphogen respectively and response in different ways (e.g. they turn white and red: see Figure 4.1).

To determine the widths of the red, white and blue regions, we introduce the positive constants $0 < m_W < m_B < m_0$ and define the spatial locations $0 < x_B < x_W < L$ such that

$$m(x = x_B) = m_B, \quad m(x = x_W) = m_W.$$

Since $m(x) = m_0(1 - x/L)$, it is straightforward to show:

$$\text{width of blue region} = x_B = \left(1 - \frac{m_B}{m_0}\right) L$$

$$\text{width of white region} = x_W - x_B = \left(\frac{m_B}{m_0} - \frac{m_W}{m_0}\right) L$$

$$\text{width of red region} = L - x_W = \left(\frac{m_W}{m_0}\right) L$$

Remarks:

- The sizes of the red, white and blue regions are independent of the morphogen diffusion coefficient: do you think this is realistic?
- How do the widths of the different change as the domain size L and m_0 are varied? How do they depend on the threshold morphogen levels m_B and m_W?
- More complex models for positional information can be developed, to account for

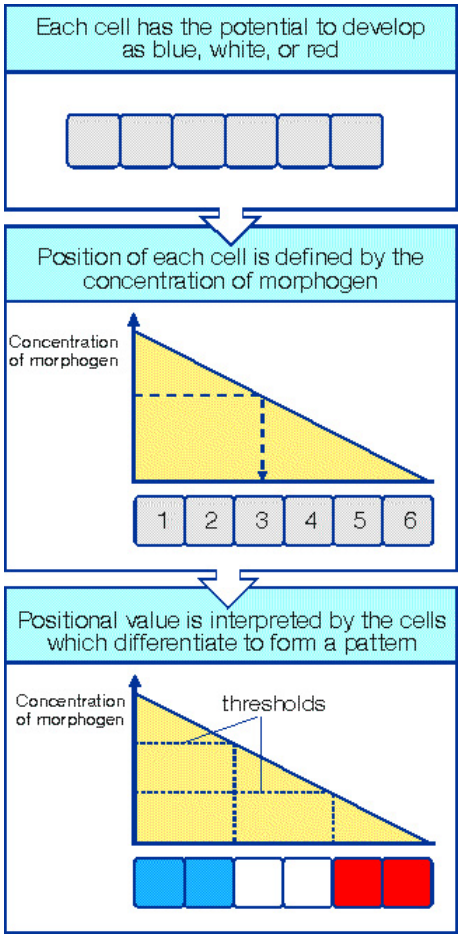


Figure 4.1: Schematic diagram illustrating the principles of positional information.

  - multiple morphogens,
  - different boundary conditions,
  - the decay of morphogens as they diffuse across the domain.

- In other biological applications (e.g. the intestinal crypt), positional information may determine whether cells proliferate, mature and/or die and, in this way, specify tissue size. In chapter 6, we will study problems of this type, where the domain size depends on the distribution of a morphogen.

Chapter 5

Fisher's Equation and travelling waves

Certain types of models exhibit wave-type behaviour. Here we will be interested in travelling waves, *i.e.* those that move at constant speed, without change in shape.

References.

- J. D. Murray, Mathematical Biology Volume I, Chapter 13.
- J. D. Murray, Mathematical Biology Volume II, Chapter 1.
- N. F. Britton, Essential Mathematical Biology, Chapter 5.

5.1 Fisher's equation: an investigation

Fisher's equation, after suitable non-dimensionalisation, can be written as

$$\frac{\partial \beta}{\partial t} = \frac{\partial^2 \beta}{\partial z^2} + \beta(1-\beta), \tag{5.1}$$

where β, z, t are dimensionless variables.

Clearly the solution of these equations will depend on the initial and boundary conditions we impose. For the time being, we state these conditions as

$$\beta(z,t) \to \beta_{\pm\infty} \quad \text{as} \quad z \to \pm\infty \quad \text{and} \quad \beta(z, t=0) = \beta_0(z), \tag{5.2}$$

where $\beta_{\pm\infty}$, β_0, are constants.

5.1.1 Key points

- We will investigate whether a solution exists for the above equation which propagates *without a change in shape* and at a constant (but as yet unknown) speed v. Such wave solutions are defined to be *travelling wave solutions*.

- Our investigation of the existence of a travelling wave solution will be substantially easier if we first transform to the moving coordinate frame $y = z - v\tau$ as, by the definition of a travelling wave, the wave profile will be independent of time in a frame moving at speed v.

- Using the chain rule and noting that we seek a solution that is time independent with respect to the y variable, we have

$$\frac{\partial \beta}{\partial t} = \frac{\partial \beta}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial \beta}{\partial \tau}\frac{\partial \tau}{\partial t} \quad \text{and} \quad \frac{\partial \beta}{\partial z} = \frac{\partial \beta}{\partial y}\frac{\partial y}{\partial z} + \frac{\partial \beta}{\partial \tau}\frac{\partial \tau}{\partial z}, \tag{5.3}$$

  i.e.

$$\frac{\partial}{\partial t} \mapsto -v\frac{\partial}{\partial y} + \frac{\partial}{\partial \tau} \quad \text{and} \quad \frac{\partial}{\partial z} \mapsto \frac{\partial}{\partial y}. \tag{5.4}$$

  Assuming $\beta = \beta(y)$ so that $\partial\beta/\partial\tau = 0$ the partial differential equation, (5.1), transforms to

$$\beta'' + v\beta' + \beta(1-\beta) = 0 \quad \text{where} \quad {}' = \frac{\mathrm{d}}{\mathrm{d}y}. \tag{5.5}$$

- One must choose appropriate boundary conditions at $\pm\infty$ for the travelling wave equations. These are the same as the boundary conditions for the full partial differential equation (but rewritten in terms of y), *i.e.*

$$\beta(y) \to \beta_{\pm\infty} \quad \text{as} \quad y \to \pm\infty, \tag{5.6}$$

  where the constants $\beta_{\pm\infty}$ are identical to those specified in equation (5.2).

- We require that $\beta_{\pm\infty}$ only take the values zero or unity:

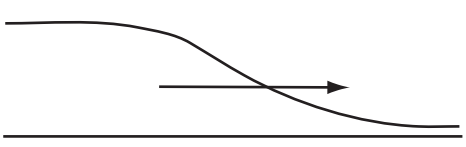
$$\int_{-\infty}^{\infty} \left[\beta'' + v\beta' + \beta(1-\beta)\right] \mathrm{d}y = 0, \tag{5.7}$$

  gives

$$\left[\beta' + v\beta\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \beta(1-\beta)\mathrm{d}y = 0. \tag{5.8}$$

  If we want $\beta \to$ constant as $y \to \pm\infty$ and β, β' finite for $\forall y$ we must have either $\beta \to 0$ or $\beta \to 1$ as $y \to \infty$ and similarly for $y \to -\infty$.

- With the boundary conditions $\beta(-\infty) = 1$ and $\beta(\infty)) = 0$, we anticipate $v > 0$.



  Indeed there are no solutions of the Fisher travelling wave equations with these boundary conditions and $v \leq 0$.

- Solutions to equations (5.1) and (5.2) are unique. The proof is an exercise in the theory of partial differential equations.

- The solutions of the travelling wave equations are not unique. One may have solutions for different values of the unknown v. Also, if $\beta(y)$ solves (5.5) for any fixed value of v then, for the same value of v, so does $\beta(y + A)$, where A is any constant. For both v and A fixed the solution of the travelling wave equations are normally unique.

- Note that the solutions of the travelling wave equations, (5.5), can only be solutions of the full partial differential equation, when considered on an infinite domain. [Realistically one requires that the length scale of variation of the system in question is much less than the length scale of the physical domain for a travelling wave to (have the potential to) be an excellent approximation to the reaction-diffusion wave solutions on the physical, *i.e.* finite, domain].

- One "loses" the initial conditions associated with equations (5.1) and (5.2). The solution of the travelling wave equations given above for β are only a solution of the full PDE, (5.1), for all time if the travelling wave solution is consistent with the initial conditions specified in (5.2).

- However, often (or rather usually!), one finds that for a particular choice of v the solutions of the full PDE system, (5.1) and (5.2), tend, as $t \to \infty$, to a solution of the travelling wave equations (5.5), with fixed v and A, for a very large class of initial conditions.

- The Russian mathematician Kolmogorov proved that solutions of the full partial differential equation system, (5.1) and (5.2), do indeed tend, as $t \to \infty$, to a solution of the travelling wave equations for $v = 2$ for a large class of initial conditions.

5.1.2 Existence and the phase plane

We will investigate the existence of solutions of Fisher's equation, equation (5.5), with the boundary conditions $\beta(-\infty) = 1$ and $\beta(\infty) = 0$ and $v > 0$, via a series of exercises involving the phase plane (β, β').

Consider the travelling wave equation

$$\frac{\mathrm{d}^2\beta}{\mathrm{d}y^2} + v\frac{\mathrm{d}\beta}{\mathrm{d}y} + \beta(1-\beta) = 0, \tag{5.9}$$

with $v > 0$ and the boundary conditions $(\beta(-\infty), \beta(\infty)) = (1, 0)$.

Exercise 1. Show that the stationary point at $(\beta, \beta') = (1, 0)$ is always a saddle point and the stationary point at $(\beta, \beta') = (0, 0)$ is a stable node for $v \geq 2$ and a stable focus for $v < 2$.
Solution. Writing $\beta' = \gamma$ gives

$$\frac{\mathrm{d}}{\mathrm{d}y}\begin{pmatrix} \beta \\ \gamma \end{pmatrix} = \frac{\mathrm{d}}{\mathrm{d}y}\begin{pmatrix} \beta \\ \beta' \end{pmatrix} = \begin{pmatrix} \gamma \\ -v\gamma - \beta(1-\beta) \end{pmatrix} = \begin{pmatrix} f(\beta,\gamma) \\ g(\beta,\gamma) \end{pmatrix}. \tag{5.10}$$

The Jacobian, $\boldsymbol{J}$, is given by

$$\boldsymbol{J} = \begin{pmatrix} \frac{\partial f}{\partial \beta} & \frac{\partial f}{\partial \gamma} \\ \frac{\partial g}{\partial \beta} & \frac{\partial g}{\partial \gamma} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1+2\beta & -v \end{pmatrix}. \tag{5.11}$$

At $(0,0)$ we have

$$\det(\boldsymbol{J} - \lambda \boldsymbol{I}) = \det \begin{pmatrix} -\lambda & 1 \\ -1 & -v-\lambda \end{pmatrix} \quad \Rightarrow \quad \lambda^2 + v\lambda + 1 = 0, \tag{5.12}$$

and hence

$$\lambda = \frac{-v \pm \sqrt{v^2-4}}{2}. \tag{5.13}$$

Therefore:

- if $v < 2$ we have $\lambda = -v/2 \pm i\mu$ and hence a stable spiral;
- if $v > 2$ we have $\lambda = -v/2 \pm \mu$ and hence a stable node;
- if $v = 2$ we have $\lambda = -1$ and hence a stable node.

At $(1,0)$ we have

$$\det(\boldsymbol{J} - \lambda \boldsymbol{I}) = \det \begin{pmatrix} -\lambda & 1 \\ 1 & -v-\lambda \end{pmatrix} \quad \Rightarrow \quad \lambda^2 + v\lambda - 1 = 0, \tag{5.14}$$

and hence

$$\lambda = \frac{-v \pm \sqrt{v^2+4}}{2}. \tag{5.15}$$

Therefore $(1,0)$ is a saddle point.

Exercise 2. Explain why solutions of Fisher's travelling wave equations must tend to phase plane stationary points as $y \to \pm\infty$. Hence explain why solutions of (5.9) with $v < 2$ are unphysical.

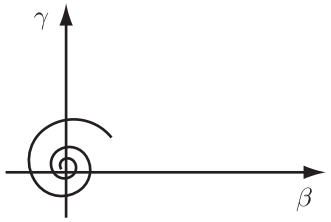
Solution. $(\beta, \gamma) \equiv (\beta, \beta')$ will change as y increases, unless at a stationary point. Therefore they will keep moving along a phase space trajectory as $y \to \infty$ unless the $y \to \infty$ limit evolves to a stationary point.

To satisfy $\lim_{y\to 0} \beta(y) = 0$, we need to be on a phase space trajectory which "stops" at $\beta = 0$. Therefore we must be on a trajectory which tends to a stationary point with $\beta = 0$ as $y \to \infty$.

Hence (β, β') must tend to $(0,0)$ as $y \to \infty$ to satisfy $\lim_{y\to\infty} \beta(y) = 0$ as $y \to \infty$.

An analogous argument holds as $y \to -\infty$.

If $v < 2$ then $\beta < 0$ at some point on the trajectory: this is unphysical:



Exercise 3. Show that the gradient of the unstable manifold at $(\beta, \beta') = (1, 0)$ is given by

$$\frac{1}{2}\left(-v + \sqrt{v^2+4}\right). \tag{5.16}$$

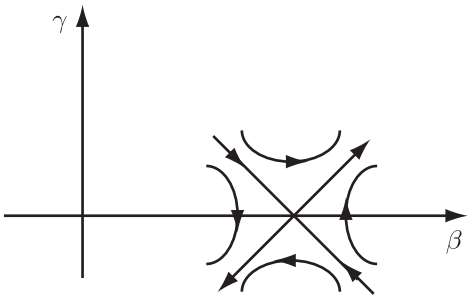
Sketch the qualitative form of the phase plane trajectories near to the stationary points for $v \geq 2$.

Solution. We require the eigenvectors of the Jacobian at $(1, 0)$:

$$\begin{pmatrix} 0 & 1 \\ 1 & -v \end{pmatrix}\begin{pmatrix} 1 \\ q_\pm \end{pmatrix} = \lambda_\pm \begin{pmatrix} 1 \\ q_\pm \end{pmatrix} \quad \Rightarrow \quad q_\pm = \lambda_\pm \quad \text{and} \quad 1 - vq_\pm = \lambda_\pm q_\pm. \tag{5.17}$$

Hence

$$\boldsymbol{v}_\pm = \begin{pmatrix} 1 \\ \frac{1}{2}\left[-v \pm \sqrt{v^2+4}\right] \end{pmatrix}. \tag{5.18}$$



Exercise 4. Explain why any physically relevant phase plane trajectory must leave $(\beta, \beta') = (1, 0)$ on the unstable manifold pointing in the direction of decreasing β.

Solution. Recall that, close to the stationary point,

$$\begin{pmatrix} \beta \\ \gamma \end{pmatrix} - \begin{pmatrix} \beta^* \\ \gamma^* \end{pmatrix} = a_- e^{\lambda_- y}\boldsymbol{v}_- + a_+ e^{\lambda_+ y}\boldsymbol{v}_+. \tag{5.19}$$

The solution moves away from the saddle along the unstable manifold, which corresponds to a_-.

Exercise 5. Consider $v \geq 2$. With $\gamma \stackrel{def}{=} \beta'$ show that for $\beta \in (0,1]$ the phase plane trajectories, with gradient

$$\frac{\mathrm{d}\gamma}{\mathrm{d}\beta} = -v - \frac{\beta(1-\beta)}{\gamma}, \tag{5.20}$$

satisfy the constraint $\mathrm{d}\gamma/\mathrm{d}\beta < -1$ whenever $\gamma = -\beta$.

Solution.

$$\left.\frac{\mathrm{d}\gamma}{\mathrm{d}\beta}\right|_{\beta=-\gamma} = -v + (1-\beta) \leq (-v+2) - (1+\beta) < -1. \tag{5.21}$$

Exercise 6. Hence show that with $v \geq 2$ the unstable manifold leaving $(\beta, \beta') = (1,0)$ and entering the region $\beta' < 0$, $\beta < 1$ enters, and can never leave, the region

$$\mathcal{R} \stackrel{def}{=} \{(\beta,\gamma) \mid \gamma \leq 0,\ \beta \in [0,1],\ \gamma \geq \beta\}. \tag{5.22}$$

Solution. Along $\mathcal{L}_1 = \{(\beta,\gamma) \mid \gamma = 0,\ \beta \in (0,1)\}$ the trajectories point vertically into $\mathcal{R}$ as

$$\left|\frac{\mathrm{d}\gamma}{\mathrm{d}\beta}\right| \to \infty \quad \text{as we approach } \mathcal{L}_1 \text{ and } \gamma' = -\beta(1-\beta) < 0. \tag{5.23}$$

Along $\mathcal{L}_2 = \{(\beta,\gamma) \mid \beta = 1,\ \gamma \in (-1,0)\}$ we have

$$\left.\frac{\mathrm{d}\gamma}{\mathrm{d}\beta}\right|_{\mathcal{L}_2} = -v - \frac{\beta(1-\beta)}{\gamma} = -v < 0. \tag{5.24}$$

Hence trajectories that enter $\mathcal{R}$ cannot leave. There any trajectory must end at a stationary point and trajectories are forced to the point $(\beta,\gamma) = (0,0)$.

Exercise 7. Thus prove that that there exists a monotonic solution, with $\beta \geq 0$, to equation (5.9) for every value of $v \geq 2$ and, with $v \geq 2$ fixed, the phase space trajectory is unique.

Solution. The above analysis is valid for $v \geq 2$. For v fixed a trajectory enters the region $\mathcal{R}$ along the unstable manifold (only one unstable manifold enters $\mathcal{R}$). The solution is monotonic as $\gamma < 0$ throughout $\mathcal{R}$.

Figure 5.1 shows the results of numerical simulation of the Fisher equation (5.1) with initial and boundary conditions given by (5.2) at a series of time points.

5.1.3 Relation between the travelling wave speed and initial conditions

We have seen that, for v fixed, the phase space trajectory of Fisher's travelling wave equation is unique. The non-uniqueness associated with the fact that if $\beta(y)$ solves Fisher's travelling wave equation then so does $\beta(y+A)$ for A constant simply corresponds to a shift along the phase space trajectory. This, in turn, corresponds simply to translation of the travelling wave.

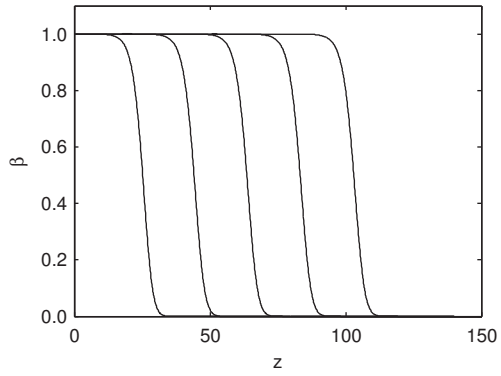


Figure 5.1: Solution of the Fisher equation (5.1) with initial and boundary conditions given by (5.2) at times $t = 10, 20, 30, 40, 50$.

Key question. We showed above that if Fisher's equation possesses a travelling wave (TW) solution then its wavespeed satisfies $v^2 \geq 4$. When are such TW solutions realised, and with what wavespeeds? A rough estimate can be obtained by studying Fisher's equation near the leading front of the TW where β is small. Linearising about $\beta = 0$ yields the following linear PDE:

$$\frac{\partial \beta}{\partial t} = \frac{\partial^2 \beta}{\partial x^2} + \beta.$$

We will assume that

$$\beta(x, 0) \sim Be^{-ax} \quad x \to \infty \quad (B, a > 0),$$

and seek TW solutions of the linearised Fisher equation of the form

$$\beta(x, t) \sim Be^{-a(x-vt)}.$$

The linearised PDE yields the following *dispersion relation* which defines the wavespeed v in terms of the decay rate of the initial data

$$av = a^2 + 1 \quad \Rightarrow v = a + \frac{1}{a} \geq 2,$$

with equality iff $a = 1$. We will consider separately the cases $a > 1$ and $a < 1$.

Case 1: $a < 1$**.**

$$\begin{aligned} a < 1 \quad & \Rightarrow e^{-ax} > e^{-x} \\ & \Rightarrow \text{I.C.s decay less rapidly than TW with min wavespeed} \\ & \Rightarrow \text{behaviour dominated by I.C.s} \\ & \Rightarrow v = a + a^{-1}. \end{aligned}$$

Case 2: $a > 1$**.**

$$\begin{aligned} a > 1 \quad & \Rightarrow e^{-ax} < e^{-x} \\ & \Rightarrow \text{I.C.s decay more rapidly than TW with min wavespeed} \\ & \Rightarrow \text{behaviour dominated by TW with min wavespeed} \\ & \Rightarrow v = 2. \end{aligned}$$

Non-Examinable: initial conditions of compact support

Kolmogorov considered the equation

$$\frac{\partial \psi}{\partial \tau} = \frac{\partial^2 \psi}{\partial z^2} + \psi(1-\psi), \tag{5.25}$$

with the boundary conditions

$$\psi(z,\tau) \to 1 \quad \text{as} \quad z \to -\infty \quad \text{and} \quad \psi(z,\tau) \to 0 \quad \text{as} \quad z \to \infty, \tag{5.26}$$

and non-negative initial conditions satisfying the following: there is a K, with $0 < K < \infty$, such that

$$\psi(z,\tau=0) = 0 \quad \text{for} \quad z > K \quad \text{and} \quad \psi(z,\tau=0) = 1 \quad \text{for} \quad z < -K. \tag{5.27}$$

He proved that $\psi(z,\tau)$ tends to a Fisher travelling wave solution with $v = 2$ as $t \to \infty$.

This can be applied to equations (5.1) and (5.2) providing the initial conditions are non-negative and the initial condition for β satisfies the above constraint, *i.e.* there is a K, with $0 < K < \infty$, such that

$$\beta(z,\tau=0) = 0 \quad \text{for} \quad z > K \quad \text{and} \quad \beta(z,\tau=0) = 1 \quad \text{for} \quad z < -K. \tag{5.28}$$

Under such constraints β also tends to a Fisher travelling wave solution with $v = 2$.

5.2 Models of epidemics

The study of infectious diseases has a long history and there are numerous detailed models of a variety of epidemics and epizootics (*i.e.* animal epidemics). Here we will only scratch the surface. In what follows we consider a basic model and show how it can be used to make general comments about epidemics and, in fact, approximately describe some specific epidemics.

5.2.1 The SIR model (revision from Short Course)

Consider a disease for which the population can be placed into three compartments:

- the susceptible compartment, S, who can catch the disease;
- the infective compartment, I, who have and transmit the disease;
- the removed compartment, R, who have been isolated, or who have recovered and are immune to the disease, or have died due to the disease during the course of the epidemic.

Assumptions

- The epidemic is of short duration course so that the population is constant (counting those who have died due to the disease during the course of the epidemic).
- The disease has a negligible incubation period.
- If a person contracts the disease and recovers, they are immune (and hence remain in the removed compartment).
- The numbers involved are sufficiently large to justify a continuum approximation.
- The 'dynamics' of the disease can be described by applying the Law of Mass Action to:

$$S + I \xrightarrow{r} 2I, \quad I \xrightarrow{a} R. \tag{5.29}$$

The model

Then the equations describing the time evolution of numbers in the susceptible, infective and removed compartments are given by

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -rIS, \tag{5.30}$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = rIS - aI, \tag{5.31}$$

$$\frac{\mathrm{d}R}{\mathrm{d}t} = aI, \tag{5.32}$$

subject to

$$S(t=0) = S_0, \quad I(t=0) = I_0, \quad R(t=0) = 0. \tag{5.33}$$

Note that

$$\frac{\mathrm{d}}{\mathrm{d}t}(S + I + R) = 0 \quad \Longrightarrow \quad S + I + R = S_0 + I_0. \tag{5.34}$$

Key questions in an epidemic situation are, given r, a, S_0 and I_0,

1. Will the disease spread, *i.e.* will the number of infectives increase, at least in the short-term?

   Solution.

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -rIS \quad \Rightarrow \quad S \text{ is decreasing and therefore } S \leq S_0. \tag{5.35}$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = I(rS - a) < I(rS_0 - a). \tag{5.36}$$

   Therefore, if $S_0 < a/r$ the infectives never increase, at least initially.

2. If the disease spreads, what will be the maximum number of infectives at any given time?

   Solution.
   $$\frac{\mathrm{d}I}{\mathrm{d}S} = -\frac{(rS-a)}{rS} = -1 + \frac{\rho}{S} \quad \text{where} \quad \rho \stackrel{def}{=} \frac{a}{r}. \tag{5.37}$$
   Integrating gives
   $$I + S - \rho \ln S = I_0 + S_0 - \rho \ln S_0, \tag{5.38}$$
   and so, noting that $\mathrm{d}I/\mathrm{d}S = 0$ for $S = \rho$, the maximum number of infectives is given by
   $$I_{max} = \begin{cases} I_0 & S_0 \leq \rho \\ I_0 + S_0 - \rho \ln S_0 - \rho \ln \rho - \rho & S_0 > \rho \end{cases}. \tag{5.39}$$

3. How many people in total catch the disease?

   Solution. From 1, $I \to 0$ as $t \to \infty$. Therefore the total number who catch the disease is
   $$R(\infty) = N_0 - S(\infty) - I(\infty) = N_0 - S(\infty), \tag{5.40}$$
   where $S(\infty) < S_0$ is the root of
   $$S_\infty - \rho \ln S_\infty = N_0 - \rho \ln S_0, \tag{5.41}$$
   obtained by setting $S = S_\infty$ and $N_0 = I_0 + S_0$ in equation (5.38).

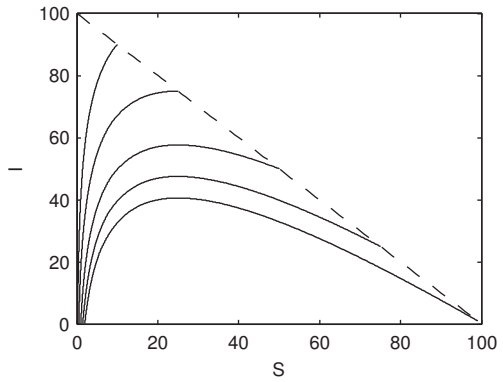


Figure 5.2: Numerical solution of the SIR model, equations (5.30)-(5.32), where the solid lines indicate the phase trajectories and the dashed line $S + I = S_0 + I_0$. Parameters are as follows: $r = 0.01$ and $a = 0.25$.

5.2.2 An SIR model with spatial heterogeneity

We consider an application to fox rabies. We will make the same assumptions as for the standard SIR model, plus:

- healthy, *i.e.* susceptible, foxes are territorial and, on average, do not move from their territories;

- rabid, *i.e.* infective, foxes undergo behavioural changes and migrate randomly, with an effective, constant, diffusion coefficient D;
- rabies is fatal, so that infected foxes do not return to the susceptible compartment but die, and hence the removed compartment does not migrate.

Taking into account rabid foxes' random motion, the SIR equations become

$$\frac{\partial S}{\partial t} = -rIS, \tag{5.42}$$
$$\frac{\partial I}{\partial t} = D\nabla^2 I + rIS - aI, \tag{5.43}$$
$$\frac{\partial R}{\partial t} = aI. \tag{5.44}$$

The I and S equations decouple, and we consider these in more detail. We assume a one-dimensional spatial domain $x \in (-\infty, \infty)$ and apply the following scalings/non-dimensionalisations,

$$I_* = \frac{I}{S_0}, \quad S_* = \frac{S}{S_0}, \quad x_* = \sqrt{\frac{rS_0}{D}}x, \quad t_* = rS_0 t, \quad \lambda = \frac{a}{rS_0}, \tag{5.45}$$

where S_0 is the population density in the absence of rabies, to obtain

$$\frac{\partial S}{\partial t} = -IS, \tag{5.46}$$
$$\frac{\partial I}{\partial t} = \nabla^2 I + I(S - \lambda), \tag{5.47}$$

where asterisks have been dropped for convenience in the final expression.

Travelling waves

We seek travelling wave solutions with

$$S(x,t) = S(y), \quad I(x,t) = I(y), \quad y = x - ct, \quad c > 0, \tag{5.48}$$

which results in the system

$$0 = cS' - IS, \tag{5.49}$$
$$0 = I'' + cI' + I(S - \lambda), \tag{5.50}$$

where $' = \mathrm{d}/\mathrm{d}y$.

We assume $\lambda = a/(rS_0) < 1$ below. This is equivalent to the condition for disease spread in the earlier SIR model.

Boundary conditions

We assume a healthy population as $y \to \infty$:

$$S \to 1 \quad \text{and} \quad I \to 0, \tag{5.51}$$

and as $y \to -\infty$ we require

$$I \to 0. \tag{5.52}$$

Bound on travelling wave speed

We write $S = 1 - P$ and linearise about the wavefront:

$$-cP' - I = 0 \quad \text{and} \quad I'' + cI' + I(1 - \lambda). \tag{5.53}$$

The I equation decouples and analysis of this equation gives a stable focus at $(I, I') = (0, 0)$ if the eigenvalues

$$\mu = \frac{-c \pm \sqrt{c^2 - 4(1 - \lambda)}}{2}, \tag{5.54}$$

are complex. This requires

$$c \geq 2\sqrt{1 - \lambda}. \tag{5.55}$$

Severity of epidemic

$S(-\infty)$ is a measure of the severity of the epidemic. We have $I = cS'/S$ and therefore

$$\frac{\mathrm{d}}{\mathrm{d}y}(I' + cI) + cS'\left(\frac{S - \lambda}{S}\right) = 0. \tag{5.56}$$

Therefore

$$(I' + cI) + c(S - \lambda \ln S) = \text{constant} = c, \tag{5.57}$$

by evaluating the equation as $y \to \infty$.

In this case

$$S(-\infty) - \lambda \ln S(-\infty) = 1, \quad \text{where} \quad S(-\infty) < 1, \tag{5.58}$$

gives the severity of the epidemic.

Further comments on travelling wave speed

Typically, the wave evolves to have minimum wave speed:

$$c \simeq c_{min}. \tag{5.59}$$

Chapter 6

Pattern formation

Examples of spatial pattern and structure can be seen just about everywhere in the natural world. Here we will be concerned with (i) building and analysing models which generate patterns, and (ii) understanding how self-organising principles may generate shape and form.

References.

- J. D. Murray, Mathematical Biology Volume II, Chapters 2 and 3.
- N. F. Britton, Essential Mathematical Biology, Chapter 7.

6.1 Minimum domains for spatial structure

Consider the following dimensionless model for budworm which spread by diffusion on a one-dimensional domain, $0 \leq x \leq L$:

$$u_t = Du_{xx} + f(u), \quad \text{where} \quad f(u) = ru\left(1 - \frac{u}{q}\right) - \frac{u^2}{1+u^2}. \tag{6.1}$$

We suppose that exterior to the domain conditions are extremely hostile to budworm so that we have the boundary conditions

$$u(0,t) = 0, \quad u(L,t) = 0. \tag{6.2}$$

Note that $f'(0) = r > 0$.

Question. Clearly $u = 0$ is a solution. However, if we start with a small initial distribution of budworm, will the budworm die out, or will there be an outbreak of budworm? In particular, how does what happens depend on the domain size?

Solution. For initial conditions with $0 \leq u(x, t = 0) \ll 1$, sufficiently small, we can approximate $f(u)$ by $f'(0)u$ at least while $u(x,t)$ remains small. Then our equations are, approximately,

$$u_t = Du_{xx} + f'(0)u, \quad u(0,t) = 0, \quad u(L,t) = 0. \tag{6.3}$$

We look for a solution of the form (invoking completeness of Fourier series):

$$u(x,t) = \sum_{n=1}^{\infty} a_n(t) \sin\left(\frac{n\pi x}{L}\right). \tag{6.4}$$

This gives that the time-dependent coefficients satisfy

$$\frac{\mathrm{d}a_n}{\mathrm{d}t} = -\frac{Dn^2\pi^2}{L^2}a_n + f'(0)a_n = \sigma_n a_n, \tag{6.5}$$

and hence

$$u(x,t) = \sum_{n=1}^{\infty} a_n(0) \exp\left[\left(f'(0) - \frac{Dn^2\pi^2}{L^2}\right)t\right] \sin\left(\frac{n\pi x}{L}\right). \tag{6.6}$$

For the solution to decay to zero, we require that all Fourier modes decay to zero as $t \to \infty$. Hence, we require that

$$\sigma_n < 0 \quad \forall n \quad \Rightarrow \quad f'(0) - \frac{Dn^2\pi^2}{L^2} < 0 \quad \forall n, \tag{6.7}$$

or, equivalently,

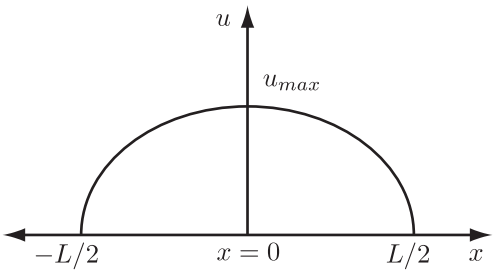
$$f'(0) < \frac{Dn^2\pi^2}{L^2} \quad \Rightarrow \quad L \leq \sqrt{\frac{D\pi^2}{f'(0)}} \stackrel{def}{=} L_{crit}. \tag{6.8}$$

Hence there is a critical lengthscale, L_{crit}, beyond which an outburst of budworm is possible in a spatially distributed system.

6.1.1 Domain size

On first inspection it is perhaps surprising that L_{crit} increases with the diffusion coefficient, *i.e.* diffusion is destabilising the zero steady state.

We can further investigate how the nature of a steady state pattern depends on the diffusion coefficient. Suppose $L > L_{crit}$ and that the steady state pattern is of the form:



We therefore have

$$0 = Du_{xx} + f(u). \tag{6.9}$$

Multiplying by u_x and integrating with respect to x, we have

$$0 = \int Du_x u_{xx}\,\mathrm{d}x + \int u_x f(u)\,\mathrm{d}x. \tag{6.10}$$

Thus we have

$$\frac{1}{2}Du_x^2 + F(u) = \text{constant} = F(u_{max}) \quad \text{where} \quad F'(u) = f(u). \tag{6.11}$$

We can therefore find a relation between L, D, integrals of

$$F(u) \stackrel{def}{=} \int_0^u f(y)\,\mathrm{d}y, \tag{6.12}$$

and $\max(u)$, the size of the outbreak, as follows:

$$u_x = -\left(\frac{2}{D}\right)^{\frac{1}{2}} \sqrt{F(u_{max}) - F(u)} \quad \text{since } x > 0 \text{ and therefore } u_x < 0. \tag{6.13}$$

Integrating, gives

$$2\int_0^{L/2} \mathrm{d}x = -(2D)^{\frac{1}{2}} \int_{u_{max}}^0 \frac{1}{\sqrt{F(u_{max}) - F(\bar{u})}}\,\mathrm{d}\bar{u}, \tag{6.14}$$

and hence

$$L = (2D)^{\frac{1}{2}} \int_0^{u_{max}} \frac{1}{\sqrt{F(u_{max}) - F(\bar{u})}}\,\mathrm{d}\bar{u}. \tag{6.15}$$

Therefore u_{max} is a function of $L/\sqrt{2D}$ and the root of equation (6.15).

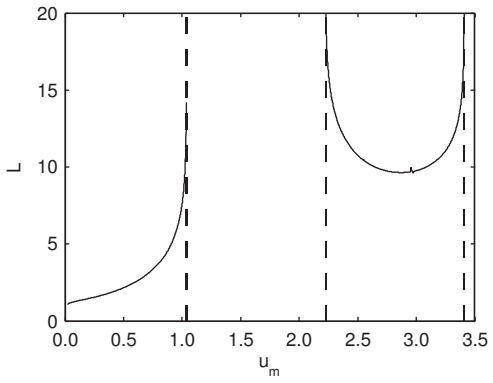


Figure 6.1: Numerical simulation of the u_m-L space, equation (6.15) with $r = 0.6$, $q = 6.2$ and $D = 0.1$.

6.2 Diffusion-driven instability

Consider a two component system

$$u_t = D_u\nabla^2 u + f(u, v), \tag{6.16}$$

$$v_t = D_v\nabla^2 v + g(u, v), \tag{6.17}$$

for $\boldsymbol{x} \in \Omega$, $t \in [0, \infty)$ and Ω bounded.

The initial conditions are

$$u(x, 0) = u_0(x), \quad v(x, 0) = v_0(x), \tag{6.18}$$

and the boundary conditions are either Dirichlet, *i.e.*

$$u = u_B, \quad v = v_B, \quad \boldsymbol{x} \in \partial\Omega, \tag{6.19}$$

or homogeneous Neumann, *i.e.*

$$\boldsymbol{n} \cdot \nabla u = 0, \quad \boldsymbol{n} \cdot \nabla v = 0, \quad \text{for} \quad \boldsymbol{x} \in \partial\Omega, \tag{6.20}$$

where $\boldsymbol{n}$ is the outward pointing normal on $\partial\Omega$.

Definition. *Patterns* are stable, time-independent, spatially heterogeneous solutions of equations (6.16)-(6.17).

Definition. A *diffusion-driven instability*, also referred to as a *Turing instability*, occurs when a steady state that is stable in the absence of diffusion becomes unstable when diffusion is present.

Remark. Diffusion-driven instabilities, in particular, can drive pattern formation in chemical systems and there is significant, but not necessarily conclusive, evidence that they can drive pattern formation in a variety of biological systems. A key point is that this mechanism can drive the system from close to a homogeneous steady state to a state with spatial pattern and structure. The fact that diffusion is responsible for this is initially quite surprising. Diffusion, in isolation, disperses a pattern; yet diffusion, combined with kinetic terms, can often drive a system towards a state with spatial structure.

6.2.1 Linear analysis

We wish to understand when a diffusion-driven instability occurs. Using vector and matrix notation we define

$$\boldsymbol{u} = \begin{pmatrix} u \\ v \end{pmatrix}, \quad \boldsymbol{F}(\boldsymbol{u}) = \begin{pmatrix} f(u, v) \\ g(u, v) \end{pmatrix}, \quad \boldsymbol{D} = \begin{pmatrix} D_u & 0 \\ 0 & D_v \end{pmatrix}, \tag{6.21}$$

and write the problem with homogeneous Neumann boundary conditions as follows:

$$\boldsymbol{u}_t = \boldsymbol{D}\nabla^2\boldsymbol{u} + \boldsymbol{F}(\boldsymbol{u}), \tag{6.22}$$

i.e.

$$\frac{\partial}{\partial t}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} D_u & 0 \\ 0 & D_v \end{pmatrix}\nabla^2\begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} f(u, v) \\ g(u, v) \end{pmatrix}, \tag{6.23}$$

with

$$\boldsymbol{n} \cdot \nabla\boldsymbol{u} = 0, \quad \boldsymbol{x} \in \partial\Omega, \tag{6.24}$$

i.e.

$$\boldsymbol{n}.\nabla u = 0 = \boldsymbol{n}.\nabla v \quad \boldsymbol{x} \in \partial\Omega. \tag{6.25}$$

Let $\boldsymbol{u}^*$ be such that $\boldsymbol{F}(\boldsymbol{u}^*) = \boldsymbol{0}$. Implicit in this definition is the assumption that $\boldsymbol{u}^*$ is a constant vector.

Let $\boldsymbol{w} = \boldsymbol{u} - \boldsymbol{u}^*$ with $|\boldsymbol{w}| \ll 1$. Then we have

$$\frac{\partial \boldsymbol{w}}{\partial t} = D\nabla^2 \boldsymbol{w} + \boldsymbol{F}(\boldsymbol{u}^*) + \boldsymbol{J}\boldsymbol{w} + \text{higher order terms}, \tag{6.26}$$

where

$$\boldsymbol{J} = \left(\begin{array}{cc} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \end{array} \right)\Bigg|_{\boldsymbol{u}=\boldsymbol{u}^*}, \tag{6.27}$$

is the Jacobian of $\boldsymbol{F}$ evaluated at $\boldsymbol{u} = \boldsymbol{u}^*$. Note that $\boldsymbol{J}$ is a *constant* matrix.

Neglecting higher order terms in $|\boldsymbol{w}|$, we have

$$\boldsymbol{w}_t = \boldsymbol{D}\nabla^2 \boldsymbol{w} + \boldsymbol{J}\boldsymbol{w}, \quad \boldsymbol{n} \cdot \nabla \boldsymbol{w} = 0, \quad \boldsymbol{x} \in \partial\Omega. \tag{6.28}$$

This is a linear equation and so we look for a solution in the form of a linear sum of separable solutions. To do this, we must first consider a general separable solution given by

$$\boldsymbol{w}(\boldsymbol{x}, t) = A(t)\boldsymbol{p}(\boldsymbol{x}), \tag{6.29}$$

where $A(t)$ is a scalar function of time. Substituting from equation (6.29) into equation (6.28) yields

$$\frac{1}{A}\frac{\mathrm{d}A}{\mathrm{d}t}\,\boldsymbol{p} = \boldsymbol{D}\nabla^2 \boldsymbol{p} + \boldsymbol{J}\boldsymbol{p}. \tag{6.30}$$

Clearly to proceed, with $\boldsymbol{p}$ dependent on $\boldsymbol{x}$ only, we require $\dot{A}/A$ to be time independent. It must also be independent of $\boldsymbol{x}$ as A is a function of time only. Thus $\dot{A}/A$ is constant.

We take $\dot{A} = \lambda A$, where λ is an, as yet, undetermined constant. Thus

$$A = A_0 \exp(\lambda t), \tag{6.31}$$

for $A_0 \neq 0$ constant. Hence we require that our separable solution is such that

$$\left[\lambda \boldsymbol{p} - \boldsymbol{J}\boldsymbol{p} - \boldsymbol{D}\nabla^2 \boldsymbol{p}\right] = 0. \tag{6.32}$$

Suppose $\boldsymbol{p}$ satisfies the equation

$$\nabla^2 \boldsymbol{p} + k^2 \boldsymbol{p} = 0, \quad \boldsymbol{n} \cdot \nabla \boldsymbol{p} = 0, \quad \boldsymbol{x} \in \partial\Omega, \tag{6.33}$$

where $k \in \mathbb{R}$. This is motivated by the fact in one-dimensional on a bounded domain, we have $p'' + k^2 p = 0$; the solutions are trigonometric functions which means one immediately has a Fourier series when writing the sum of separable solutions.

Then we have

$$\left[\lambda \boldsymbol{p} - \boldsymbol{J}\boldsymbol{p} + \boldsymbol{D}k^2\boldsymbol{p}\right] = 0, \tag{6.34}$$

and thus

$$\left[\lambda \boldsymbol{I} - \boldsymbol{J} + \boldsymbol{D}k^2\right]\boldsymbol{p} = 0, \tag{6.35}$$

with $|\boldsymbol{p}|$ not identically zero. Hence

$$\det\left[\lambda \boldsymbol{I} - \boldsymbol{J} + k^2\boldsymbol{D}\right] = 0. \tag{6.36}$$

This can be rewritten as

$$\det\begin{pmatrix} \lambda - f_u + D_u k^2 & -f_v \\ -g_u & \lambda - g_v + D_v k^2 \end{pmatrix} = 0, \tag{6.37}$$

which gives the following quadratic in λ:

$$\lambda^2 + \left[(D_u + D_v)k^2 - (f_u + g_v)\right]\lambda + h(k^2) = 0, \tag{6.38}$$

where

$$h(k^2) = D_u D_v k^4 - (D_v f_u + D_u g_v)k^2 + (f_u g_v - g_u f_v). \tag{6.39}$$

Note 1. Fixing model parameters and functions (*i.e.* fixing D_u, D_v, f, g), we have an equation which gives λ as a function of k^2.

Note 2. For any k^2 such that equation (6.33) possesses a solution, denoted $\boldsymbol{p}_k(\boldsymbol{x})$ below, we can find a $\lambda = \lambda(k^2)$ and, hence, a general separable solution of the form

$$A_0 e^{\lambda(k^2)t}\boldsymbol{p}_k(\boldsymbol{x}). \tag{6.40}$$

The most general solution formed by the sum of separable solutions is therefore

$$\sum_{k^2} A_0(k^2) e^{\lambda(k^2)t}\boldsymbol{p}_k(\boldsymbol{x}), \tag{6.41}$$

if there are countable k^2 for which equation (6.33) possesses a solution. Otherwise the general solution formed by the sum of separable solutions is of the form

$$\int A_0(k^2) e^{\lambda(k^2)t}\boldsymbol{p}_{k^2}(\boldsymbol{x})\,\mathrm{d}k^2, \tag{6.42}$$

where k^2 is the integration variable.

Unstable points

If, for any k^2 such that equation (6.33) possesses a solution, we find $Re(\lambda(k^2)) > 0$ then:

- $\boldsymbol{u}^*$ is (linearly) unstable and perturbations from the stationary state will grow;
- while the perturbations are small, the linear analysis remains valid; thus the perturbations keep growing until the linear analysis is invalid and the full non-linear dynamics comes into play;

- a small perturbation from the steady state develops into a growing spatially heterogeneous solution which subsequently seeds spatially heterogeneous behaviour of the full non-linear model;
- a spatially heterogeneous pattern can emerge from the system from a starting point which is homogeneous to a very good approximation.

Stable points

If, for all k^2 such that equation (6.33) possesses a solution, we find $Re(\lambda(k^2)) < 0$ then:

- $\boldsymbol{u}^*$ is (linearly) stable and perturbations from the stationary state do not grow;
- patterning will not emerge from perturbing the homogeneous steady state solution $\boldsymbol{u}^*$;
- the solution will decay back to the homogeneous solution[1].

6.3 Detailed study of the conditions for a Turing instability

For a Turing instability we require the homogeneous steady state to be **stable without diffusion** and **unstable with diffusion present**. Here we analyse the requirements for each of these conditions to be satisfied.

6.3.1 Stability without diffusion

First, we require that in the absence of diffusion the system is stable. This is equivalent to

$$Re(\lambda(0)) < 0, \tag{6.43}$$

for *all* solutions of $\lambda(0)$, as setting $k^2 = 0$ removes the diffusion-driven term in equation (6.36) and the preceding equations.

We have that $\lambda(0)$ satisfies

$$\lambda(0)^2 - [f_u + g_v]\,\lambda(0) + [f_u g_v - f_v g_u] = 0. \tag{6.44}$$

Insisting that $Re(\lambda(0) < 0)$ gives us the conditions

$$f_u + g_v < 0 \tag{6.45}$$

$$f_u g_v - f_v g_u > 0. \tag{6.46}$$

[1]Technical point: strictly, this conclusion requires completeness of the separable solutions. This can be readily shown in 1D on bounded domains. (Solutions of $p'' + k^2 p = 0$ on bounded domains with Neumann conditions are trigonometric functions and completeness is inherited from the completeness of Fourier series). Even if completeness of the separable solutions is not clear, numerical simulations of the full equations are highly indicative and do not, for the models typically encountered, contradict the results of the linear analysis. With enough effort and neglecting any biological constraints on model parameters and functions, one may well be able to find $D_u,\ D_v,\ f,\ g$ where there was such a discrepancy but that is not the point of biological modelling.

The simplest way of deducing (6.45) and (6.46) is by brute force.

The roots of the quadratic are given by

$$\lambda(0)_{\pm} = \frac{(f_u + g_v) \pm \sqrt{(f_u + g_v)^2 - 4(f_u g_v - f_v g_u)}}{2}. \tag{6.47}$$

6.3.2 Instability with diffusion

Now consider the effects of diffusion. In addition to $Re(\lambda(0)) < 0$, we must show, for diffusion-driven instability, that there exists k^2 such that

$$Re(\lambda(k^2)) > 0, \tag{6.48}$$

so that diffusion does indeed drive an instability.

We have that $\lambda(k^2)$ satisfies

$$\lambda^2 + \left[(D_u + D_v)k^2 - (f_u + g_v)\right]\lambda + h(k^2) = 0, \tag{6.49}$$

where

$$h(k^2) = D_u D_v k^4 - (D_v f_u + D_u g_v)k^2 + (f_u g_v - g_u f_v), \tag{6.50}$$

and

$$\alpha = (f_u + g_v) - (D_u + D_v)k^2 < 0. \tag{6.51}$$

Thus $Re(\lambda(k^2)) > 0$ requires that

$$Re\left(\alpha \pm \sqrt{\alpha^2 - 4h(k^2)}\right) > 0 \quad \Rightarrow \quad h(k^2) < 0. \tag{6.52}$$

Hence we must find k^2 such that

$$h(k^2) = D_u D_v k^4 - (D_v f_u + D_u g_v)k^2 + (f_u g_v - g_u f_v) < 0, \tag{6.53}$$

so that we have $k^2 \in [k_-^2, k_+^2]$ where $h(k_\pm^2) = 0$. Figure 6.2 shows a plot of a caricature $h(k^2)$.

We conclude that we have instability whenever

$$k^2 \in \left[\frac{A - \sqrt{A^2 - B}}{2D_u D_v}, \frac{A + \sqrt{A^2 - B}}{2D_u D_v}\right] = \left[k_-^2, k_+^2\right], \tag{6.54}$$

where

$$A = D_v f_u + D_u g_v \quad \text{and} \quad B = 4D_u D_v (f_u g_v - g_u f_v) > 0, \tag{6.55}$$

and there exists a solution of the following equation

$$\nabla^2 \boldsymbol{p} + k^2 \boldsymbol{p} = 0, \quad \boldsymbol{n} \cdot \nabla \boldsymbol{p} = 0, \quad \boldsymbol{x} \in \partial\Omega, \tag{6.56}$$

for k^2 in the above range.

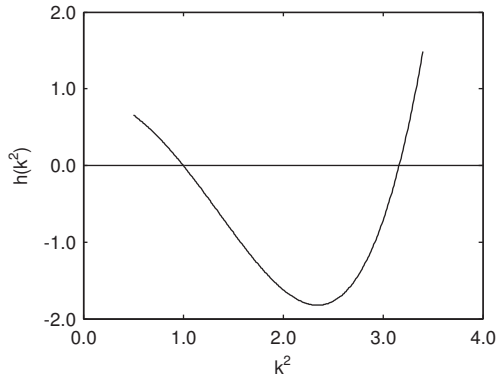


Figure 6.2: A plot of a caricature $h(k^2)$.

Insisting that k is real and non-zero (we have considered the $k = 0$ case above) we have

$$A > 0 \quad \text{and} \quad A^2 - B > 0, \tag{6.57}$$

which gives us that when $Re(\lambda(k^2)) > 0$, the following conditions hold:

$$A > 0 : \qquad D_v f_u + D_u g_v > 0, \tag{6.58}$$

$$A^2 - B > 0 : \quad (D_v f_u + D_u g_v) > 2\sqrt{D_u D_v (f_u g_v - f_v g_u)}. \tag{6.59}$$

6.3.3 Summary

We have found that diffusion-driven instability can occur when conditions (6.45), (6.46), (6.58), (6.59) hold. Then the instability is driven by separable solutions which solve equation (6.33) with k^2 in the range (6.54).

Key point 1. Note that constraints (6.45) and (6.58) immediately gives us that $D_u \neq D_v$. Thus one cannot have a diffusion-driven instability with *identical* diffusion coefficients.

Key point 2. From constraints (6.45), (6.46), (6.58) the signs of f_u, g_v must be such that $\boldsymbol{J}$ takes the form

$$\boldsymbol{J} = \begin{pmatrix} + & + \\ - & - \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} + & - \\ + & - \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} - & - \\ + & + \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} - & + \\ - & + \end{pmatrix}. \tag{6.60}$$

Key point 3. A Turing instability typically occurs via *long-range inhibition and short-range activation.* In more detail, suppose

$$\boldsymbol{J} = \begin{pmatrix} + & - \\ + & - \end{pmatrix}. \tag{6.61}$$

Then we have $f_u > 0$ and $g_v < 0$ by the signs of $\boldsymbol{J}$. In this case $D_v f_u + D_u g_v > 0 \Rightarrow D_v > D_u$. Hence the activator has a lower diffusion coefficient and spreads less quickly than the inhibitor.

6.3.4 The threshold of a Turing instability.

The threshold is defined such that equation (6.39), *i.e.*

$$D_u D_v k_c^4 - (D_v f_u + D_u g_v) k_c^2 + (f_u g_v - g_u f_v) = 0, \tag{6.62}$$

has a single root, k_c^2.

Thus we additionally require

$$A^2 = B \quad i.e. \quad (D_v f_u + D_u g_v)^2 = 4 D_u D_v (f_u g_v - g_u f_v) > 0, \tag{6.63}$$

whereupon

$$k_c^2 = \frac{A}{2 D_u D_v} = \frac{D_v f_u + D_u g_v}{2 D_u D_v}. \tag{6.64}$$

Strictly one also requires that a solution exists for

$$\nabla^2 \mathbf{p} + k^2 \boldsymbol{p} = 0, \quad \boldsymbol{n} \cdot \nabla \boldsymbol{p} = 0, \quad \boldsymbol{x} \in \partial\Omega, \tag{6.65}$$

when $k^2 = k_c^2$. However, the above value of k_c^2 is typically an excellent approximation.

6.4 Extended example 1

Consider the one-dimensional case:

$$u_t = D_u u_{xx} + f(u, v), \tag{6.66}$$

$$v_t = D_v v_{xx} + g(u, v), \tag{6.67}$$

for $x \in [0, L]$, $t \in [0, \infty)$ and zero flux boundary conditions at $x = 0$ and $x = L$.

The analogue of

$$\nabla^2 \boldsymbol{p} + k^2 \boldsymbol{p} = 0, \quad \boldsymbol{n} \cdot \nabla \boldsymbol{p} = 0, \quad \boldsymbol{x} \in \partial\Omega, \tag{6.68}$$

is

$$p_{xx} + k^2 p = 0, \quad p'(0) = p'(L) = 0, \tag{6.69}$$

which gives us that

$$p_k(x) = A_k \cos(kx), \quad k = \frac{n\pi}{L}, \quad n \in \{1, 2, \ldots\}, \tag{6.70}$$

where A_k is k-dependent in general but independent of t and x.

Thus the separable solution is of the form

$$\sum_k A_k e^{\lambda(k^2)t} \cos(kx), \tag{6.71}$$

where the sum is over the allowed values of k *i.e.*

$$k = \frac{n\pi}{L}, \quad n \in \{1, 2, \ldots\}. \tag{6.72}$$

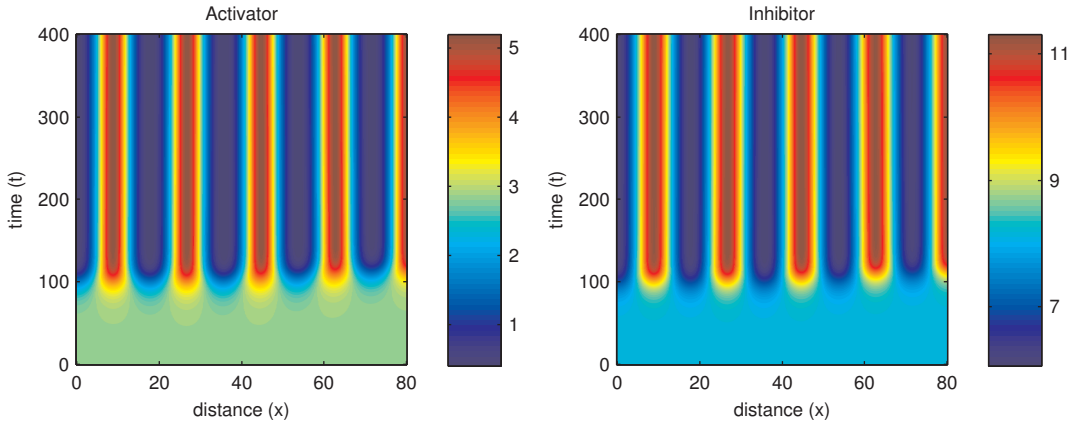


Figure 6.3: Numerical simulation of the Gierer-Meinhardt model for pattern formation.

6.4.1 The influence of domain size

If the smallest allowed value of $k^2 = \pi^2/L^2$ is such that

$$k^2 = \frac{\pi^2}{L^2} > \frac{A + \sqrt{A^2 - B}}{2D_u D_v} = k_+^2, \tag{6.73}$$

then we cannot have a Turing instability.

Thus for very small domains there is *no* pattern formation via a Turing mechanism. However, if one slowly increases the size of the domain, then L increases and the above constraint eventually breaks down and the homogeneous steady state destabilises leading to spatial heterogeneity.

This pattern formation mechanism has been observed in chemical systems. It is regularly hypothesised to be present in biological systems (*e.g.* animal coat markings, fish markings, the interaction of gene products at a cellular level, the formation of ecological patchiness) though the evidence is not conclusive at the moment.

6.5 Extended example 2

Consider the two-dimensional case with spatial coordinates $\boldsymbol{x} = (x, y)^T$, $x \in [0, a]$, $y \in [0, b]$, and zero flux boundary conditions. We find that the allowed values of k^2 are

$$k_{m,n}^2 = \left[\frac{m^2\pi^2}{a^2} + \frac{n^2\pi^2}{b^2}\right], \tag{6.74}$$

with

$$p_{m,n}(\boldsymbol{x}) = A_{m,n} \cos\left(\frac{m\pi x}{a}\right) \cos\left(\frac{n\pi y}{b}\right), \quad n, m \in \{0, 1, 2, \ldots\}, \tag{6.75}$$

excluding the case where n, m are both zero.

Suppose the domain is long and thin, $b \ll a$. We may have a Turing instability if

$$k_{m,n}^2 = \left[\frac{m^2\pi^2}{a^2} + \frac{n^2\pi^2}{b^2}\right] \in \left[k_-^2, k_+^2\right] \quad \text{where} \quad h(k_\pm^2) = 0. \tag{6.76}$$

For b sufficiently small, this requires $n = 0$ and therefore no spatial variation in the y direction.

This means we have that the seed for pattern formation predicted by the linear analysis is a separable solution which is "stripes"; this typically invokes a striped pattern once the non-linear dynamics sets in.

For a large rectangular domain, $b \sim a$ sufficiently large, it is clear that a Turing instability can be initiated with n, $m > 0$. This means we have that the seed for pattern formation predicted by the linear analysis is a separable solution which is "spots". This typically invokes a spotted pattern once the non-linear dynamics sets in.

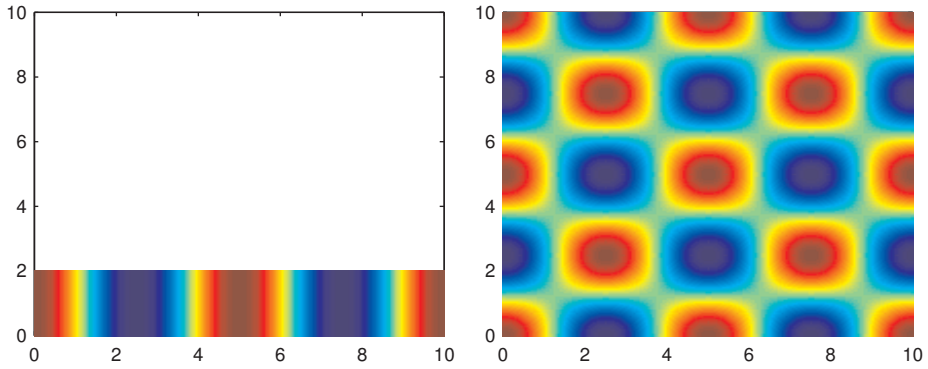
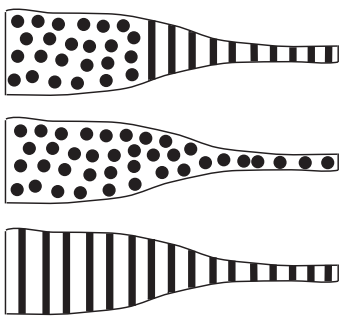


Figure 6.4: Changes in patterning as the domain shape changes.

Figure 6.4 shows how domain size may affect the patterns formed. On the left-hand side the domain is long and thin and only a striped pattern results, whilst the on the right-hand side the domain is large enough to admit patterning in both directions.

Suppose we have a domain which changes its aspect ratio from rectangular to long and thin. Then we have the following possibilities:



This leads to an interesting prediction, in the context of animal coat markings, that if it is indeed driven by a Turing instability, then one should not expect to see an animal with a striped body and a spotted tail.

Figure 6.5: Animal coat markings which are consistent with the predictions of pattern formation by a Turing instability.

Common observation is consistent with such a prediction (see Figure 6.5) but one should not expect universal laws in the realms of biology as one does in physics (see Figure 6.6). More generally, this analysis has applications in modelling numerous chemical and biochemical reactions, in vibrating plate theory, and studies of patchiness in ecology and modelling gene interactions.

Figure 6.6: Animal coat markings which are inconsistent with the predictions of pattern formation by a Turing instability.

Chapter 7

Domain Growth in Biology

In many biological systems, growth is accompanied by size changes. For example, children typically increase in height as they grow. Mathematical models that can predict such growth are very useful and have many applications. For example, domain growth can affect the animal coat patterns that arise in reaction-diffusion systems (see Figure 6.1). Alternatively, in tissue engineering, experimentalists aim to create new tissues (e.g. skin, bone, cartilage) to replace damaged ones.

When testing new drugs and treatments in the laboratory, experimentalists need a system which is reliable, safe and reproducible. Common methods for doing this involve culturing (for example) tumour cells, as 2D monolayers or as 3D multicellular spheroids. The tumour cells cluster together and undergo diffusion limited growth: vital nutrients (eg oxygen, glucose) diffuse through the cellular mass, enabling the cells to stay alive and proliferate (see Figure 6.2).

An additional reason for studying the growth of multicellular spheroids is that they mimic the early stages of tumour growth, before the tumour has developed a blood supply. During this growth phase, tumour cells aggregate to form a mass which increases in size as the cells proliferate, their growth depending on local levels of nutrients that diffuse through the tissue. In order to model the tumour's growth we need to be able to predict the nutrient concentration within the tumour (this determines how the tumour grows and will ultimately enable us to determine the position of the outer tumour boundary). Note

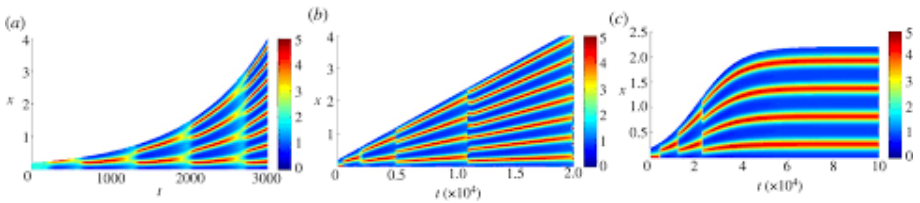


Figure 7.1: Series of images showing how the rate of domain growth can affect the spatial patterns generated by a reactio-diffusion system. For further information, see EJ Crampin, EA Gaffney and PK Maini (1999). Bull Math Biol. 61: 1093-1120.

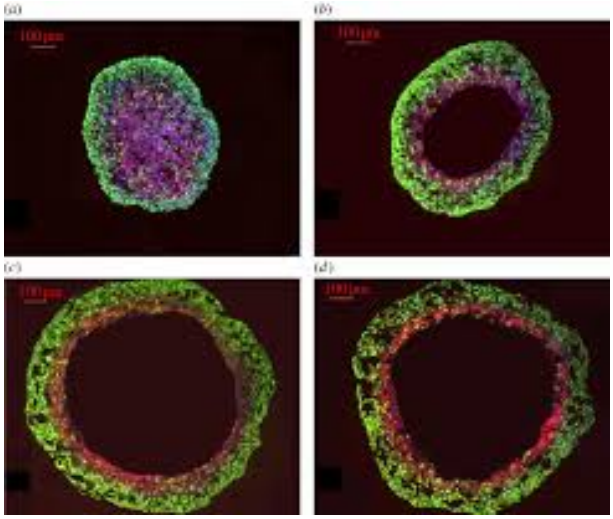


Figure 7.2: Series of images showing how the size and structure of multicellular tumour spheroids change over time

that the domain on which we solve for the nutrient concentration grows as the tumour grows: this is an example of a **moving boundary problem**.

References (for interest).

- H.P. Greenspan (1972) Models for the growth of a solid tumour by diffusion. *Stud. Appl. Math.* 52: 317-344.
- E.J. Crampin, E.A. Gaffney and P.K. Maini (1999). Reaction and diffusion on growing domains: scenarios for robust pattern formation. *Bull. Math. Biol.* 61: 1093-1120.
- D.S. Jones and B.D. Sleeman (2004). Differential equations and mathematical biology. Taylor and Francis.

7.1 A simple model of 1D tumour growth

Consider a 3D slab of tumour cells, which are uniform in the y and z directions. Let $x = R(t)$ denote the position of the outer tissue boundary. We need an equation for the **nutrient concentration** $C(x,t)$**,** and **suitable boundary conditions.** For **moving boundary problems,** since the position of the boundary is unknown, we will need an extra condition to determine it. This will come from assumptions about how the tumour grows in response to its consumption of nutrients.

Nutrient concentration, $C(x,t)$**.** We assume that the nutrient diffuses and is consumed by tumour cells at a constant rate. Thus, $C(x,t)$ satisfies

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} - \lambda \qquad \text{for } |x| < R(t) \quad \text{(within the tumour)}, \tag{7.1}$$

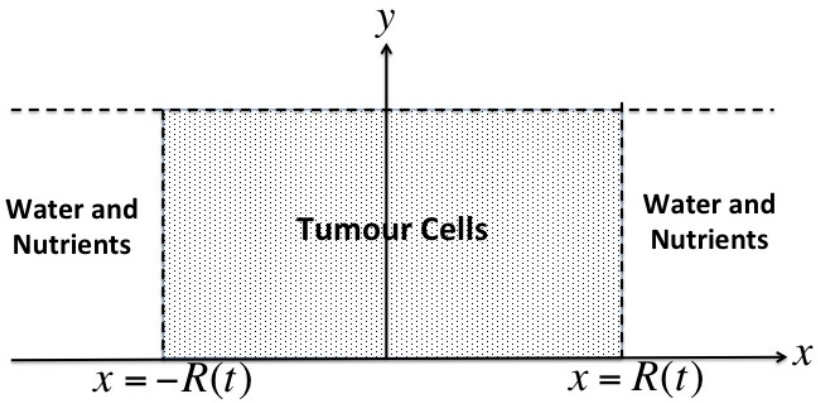


Figure 7.3: Schematic diagram of 1D tumour

where $D > 0$ is the diffusion constant and $lambda > 0$ the constant rate at which tumour cells consume nutrient. We assume that outside the tumour the nutrient is constantly replenished, so its concentration there is maintained at a constant value:

$$C(x,t) \equiv C^* \qquad \text{for } |x| > R(t).$$

We assume that the tumour is symmetric about $x = 0$, so we only need solve on $0 \leq x \leq R(t)$, with boundary conditions

$$\frac{\partial C}{\partial x} = 0 \quad \text{at } x = 0, \tag{7.2}$$

$$C(R(t),t) = C^*. \tag{7.3}$$

We need an additional condition (or equation) that describes how tumour growth (represented by $R(t)$) depends on nutrient levels.

Tumour boundary, $x = R(t)$. The following equation defines how the growth (represented by $x = R(t)$) depends on the nutrient concentration:

$$\frac{dR}{dt} = \int_0^{R(t)} P(C)dx, \tag{7.4}$$

where $P(C)$ represents the **local proliferation rate** at a given point within the tumour (it depends only on the availability of nutrient at that point). In general, we expect $P(C)$ to be an *increasing function* of C. To close the model, we prescribe the *initial* position of the tumour boundary:

$$R(t=0) = R_0. \tag{7.5}$$

Note. Strictly speaking, we should impose initial conditions for the nutrient: $C(x,0) = C_{in}(x)$, say. In practice, as we explain below, it is not necessary to impose initial conditions for $C(x,t)$.

7.2 Model reduction: nondimensionalise

Our model of tumour growth consists of equations (7.1)-(7.5). Before solving these equations, it is helpful to nondimensionalise and simplify them. We nondimensionalise as follows:

$$x = R_0\xi, \quad R(t) = R_0 r(\tau), \quad C(x,t) = C^* c(\xi,\tau), \quad P(C) = P_0 p(c), \quad t = \frac{\tau}{P_0}, \tag{7.6}$$

where P_0 is a typical tumour proliferation rate (e.g. when $C \equiv C^*$). Equation (7.1) then becomes

$$\left(\frac{R_0^2 P_0}{D}\right)\frac{\partial c}{\partial \tau} = \frac{\partial^2 c}{\partial \xi^2} - \mu, \qquad \text{where } \mu = \frac{\lambda R_0^2}{C^* D}. \tag{7.7}$$

The coefficient $R_0^2 P_0/D$ is the ratio of the **diffusion timescale,** R_0^2/D**,** to a typical timescale for **tumour proliferation,** $1/P_0$**.** Typically (e.g. from experimental data), the diffusion timescale is *much shorter* (minutes, for a small tumour) than the proliferation timescale (tumour growth occurs over weeks) and, thus,

$$\frac{R_0^2 P_0}{D} \ll 1.$$

We can therefore **neglect the time derivative** in the diffusion equation and, to leading order, we obtain the following dimensionless model:

Dimensionless model equations:

$$0 = \frac{\partial^2 c}{\partial \xi^2} - \mu \quad \text{for } 0 \le \xi \le r(\tau), \tag{7.8}$$

$$c(\xi,\tau) \equiv 1 \quad \text{for } \xi > r(\tau), \tag{7.9}$$

$$\frac{\partial c}{\partial \xi}(0,\tau) = 0, \tag{7.10}$$

$$c(r(\tau),\tau) = 1, \tag{7.11}$$

$$\frac{dr}{d\tau} = \int_0^{r(\tau)} p(c)\,d\xi, \tag{7.12}$$

$$r(0) = 1. \tag{7.13}$$

Solving (7.8) gives

$$c = \frac{\mu\xi^2}{2} + A(\tau)\xi + B(\tau).$$

Although we have removed the explicit τ-dependence from the c-equation, $A = A(\tau)$ and $B = B(\tau)$ still depend on τ via the **boundary conditions.** Imposing conditions (7.10) and (7.11) on c gives

$$c(\xi,\tau) = 1 - \frac{\mu}{2}(r^2(\tau) - \xi^2). \tag{7.14}$$

Points to note:

- c attains its **minimum value** at $\xi = 0$ where $c(0,\tau) = 1 - \mu r^2(\tau)/2$.
- As the tumour grows and $r(\tau)$ increases, the nutrient concentration at the tumour centre decreases.
- For physically-realistic solutions, we require (at least) $c(0,\tau) > 0$.
- More on this later!

Question: How does $r(\tau)$ evolve? Armed with solution (7.14) for c, we can determine $r(\tau)$ via equation (7.12). We consider the simplest case for which p is a linear function of c so that $p(c) = c$ (nondimensionalisation of P chosen to make coefficient of c unity). Then (7.12) becomes

$$\begin{aligned} \frac{dr}{d\tau} &= \int_0^{r(\tau)} c(\xi,\tau)d\xi \qquad (7.15) \\ &= \int_0^{r(\tau)} \left(1 - \frac{\mu}{2}(r^2(\tau) - \xi^2)\right) d\xi \\ \Rightarrow \frac{dr}{d\tau} &= r(\tau)\left(1 - \frac{\mu r^2(\tau)}{3}\right) = f(r) \qquad (7.16) \end{aligned}$$

which is similar to logistic growth (indeed, $y = r^2$ undergoes logistic growth). We can integrate to find $r(\tau)$ explicitly, or we can use phase-plane/stability analysis.

Phase plane analysis.

- In the physically-relevant regime, $r \geq 0$, and there are two steady states: $r_1^* = 0$ and $r_2^* = \sqrt{3/\mu}$.
- Their stability is determined by the sign of $f'(r^*)$.
- Near $r_1^* = 0$ f is increasing and positive $\Rightarrow$ small perturbations from $r_1^* = 0$ continue to increase, i.e. this tumour-free steady state is unstable.
- Near $r_2^* = \sqrt{3/\mu}$, if $r < r_2^*$ then f is positive, so the solution increases back towards r_2^*; if $r > r_2^*$ then f is negative, and the solution decreases towards r_2^*. i.e. this steady state is stable.
- **Exercise:** demonstrate this by linearising about the critical points.

All solutions to (7.16) will approach the unique stable steady state

$$r^* = \sqrt{3/\mu}.$$

Below we consider whether this steady state is physically realistic.

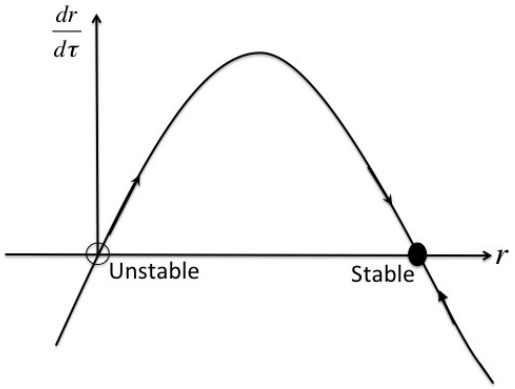


Figure 7.4: Phase-plot of $dr/d\tau$ versus $r(\tau)$, illustrating the steady states and their stability.

7.3 Cell death at low nutrient concentration

As nutrient levels fall a cell's ability to proliferate and remain alive diminishes. We model these effects by assuming that there exist a **threshold** nutrient concentration, $c_N \in (0, 1)$, such that

- $\mathbf{c > c_N} \Rightarrow$ **proliferation** of **live** cells;
- $\mathbf{c < c_N} \Rightarrow$ cell **death and degradation** ("necrosis").

With these new features we predict tumour evolution to be as follows.

1. At $\tau = 0$, the tumour has unit radius ($r(0) = 1$), and nutrient concentration given by (7.14), where we assume that
$$c(0, 0) = 1 - \frac{\mu}{2} > c_N$$
so that initially the tumour is well-nourished.
2. The tumour boundary **grows** according to (7.16) until **either** the steady state is attained, **or** until the minimum nutrient concentration $c(0, \tau) = c_N$, and **necrosis** sets in at centre of the tumour.
3. Which occurs first? Necrosis or a steady state?

We determine whether necrosis occurs before the tumour attains a steady state by determining the tumour size $r = r_1$ at which necrosis first occurs. We do this by setting $c(0, \tau) = c_N$ in equation (7.14):

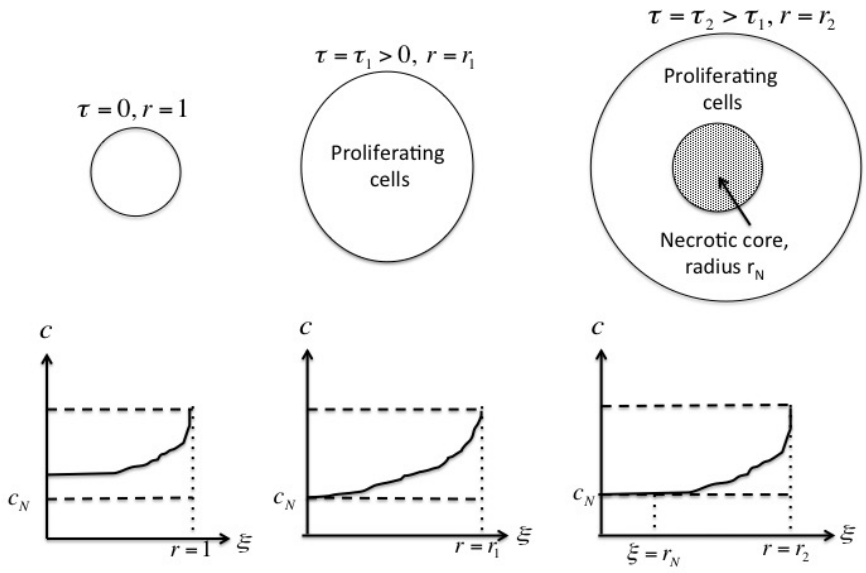


Figure 7.5: Series of diagrams illustrating how necrosis is initiated.

$$c_N = 1 - \frac{\mu r_1^2}{2} \quad \Rightarrow \quad r_1 = \left(\frac{2(1-c_N)}{\mu}\right)^{1/2} < r_* = \sqrt{3/\mu}.$$

Thus r_1 is always attained **before** the steady state is reached: the steady state is **never physically feasible.**

To calculate the time $\tau = \tau_1$ of necrosis onset, we separate variables and integrate equation (7.16), between the appropriate limits:

$$\int_1^{r_1} \frac{dr}{r(1-\mu r^2/3)} = \int_0^{\tau_1} d\tau = \tau_1.$$

In order to describe tumour evolution after the onset of necrosis (*i.e.* for $\tau > \tau_1$), we must **modify** the model.

7.4 Revised model: proliferation and necrosis

In the original model, λ in (7.1) (or μ in (7.8)) represents **uniform nutrient uptake**. In reality only **live** cells will absorb nutrient. Thus we replace (7.8) by

$$\frac{\partial^2 c}{\partial \xi^2} = \mu H(c - c_N) \quad 0 \le \xi \le r(\tau), \quad \text{where} \quad H(c-c_N) = \begin{cases} 1 & \text{if } c > c_N, \\ 0 & \text{if } c \le c_N. \end{cases}$$

Equations (7.9), (7.10) and (7.11) are unchanged but the growth equation (7.12) must be modified since only **live** cells proliferate and contribute to growth, while **dead** cells **degrade** and effectively **remove mass** from the system. Thus, again assuming proliferation (where it occurs) to be linear in c, we take

$$\frac{dr}{d\tau} = \int_0^{r(\tau)} p(c)d\xi = \int_{r_N(\tau)}^{r(\tau)} cd\xi - \int_0^{r_N(\tau)} \delta\, d\xi.$$

That is, the net proliferation function $p(c)$ has the form

$$p(c) = \begin{cases} c & \text{if } c > c_N \\ -\delta & \text{if } c \le c_N, \end{cases}$$

where $\delta > 0$ is the death rate when nutrient levels are too low ($c \le c_N$) to sustain viable cells.

We note that the edge of the necrotic region, $\xi = r_N(\tau)$, say, will **move in time**: it is located where the nutrient concentration first dips below the threshold value $c(r_N(\tau), \tau) = c_N$. We now have **two** moving boundaries to determine.

Model summary (for $\tau > \tau_1$):

$$\frac{\partial^2 c}{\partial \xi^2} = \mu H(c - c_N) \quad 0 \le \xi \le r(\tau), \tag{7.17}$$

$$c(\xi, \tau) \equiv 1 \quad \xi > r(\tau), \tag{7.18}$$

$$\frac{\partial c}{\partial \xi}(0, \tau) = 0, \tag{7.19}$$

$$c(r(\tau), \tau) = 1, \tag{7.20}$$

$$c(r_N(\tau), \tau) = c_N, \quad \text{with } c \text{ continuous across } \xi = r_N(\tau), \tag{7.21}$$

$$\frac{\partial c}{\partial \xi} \quad \text{continuous across } \xi = r_N(\tau), \tag{7.22}$$

$$\frac{dr}{d\tau} = \int_0^{r(\tau)} p(c)d\xi = \int_{r_N(\tau)}^{r(\tau)} cd\xi - \int_0^{r_N(\tau)} \delta\, d\xi, \tag{7.23}$$

$$r(\tau_1) = r_1 = \left(\frac{2(1 - c_N)}{\mu}\right)^{1/2}. \tag{7.24}$$

Note: for $\tau > \tau_1$, the outer moving boundary $r(\tau)$ is determined by (7.23), while the boundary of the necrotic region $r_N(\tau)$ is determined **implicitly** by the condition $c(r_N(\tau), \tau) = c_N$.

Model solution. We solve for c in each of the two regions, noting that (7.17) is equivalent to

$$\frac{\partial^2 c}{\partial \xi^2} = \begin{cases} 0 & 0 \le \xi \le r_N(\tau), \\ \mu & r_N(\tau) \le \xi \le r(\tau). \end{cases}$$

$$\Rightarrow \quad c(\xi, \tau) = \begin{cases} A_1\xi + B_1 & 0 \le \xi \le r_N(\tau) \\ \frac{\mu\xi^2}{2} + A_2\xi + B_2 & r_N(\tau) \le \xi \le r(\tau). \end{cases}$$

We fix $A_1(\tau)$ and $B_1(\tau)$ by imposing conditions (7.19) and (7.21), which give

$$A_1 = 0, \quad B_1 = c_N.$$

Conditions (7.20), (7.21) and (7.22) supply:

$$\begin{aligned} \frac{\mu r^2(\tau)}{2} + A_2 r(\tau) + B_2 &= 1, \\ \frac{\mu r_N^2(\tau)}{2} + A_2 r_N(\tau) + B_2 &= c_N, \\ \mu r_N(\tau) + A_2 &= 0. \end{aligned}$$

In this way we obtain $A_2(\tau)$ and $B_2(\tau)$ in terms of $r_N(\tau)$, and an equation relating $r_N(\tau)$ to $r(\tau)$:

$$A_2 = -\mu r_N(\tau), \quad B_2 = c_N + \frac{\mu r_N^2(\tau)}{2}, \tag{7.25}$$

$$1 - c_N = \frac{\mu}{2}(r(\tau) - r_N(\tau))^2. \tag{7.26}$$

Equation (7.26) predicts that the **width of the tumour's proliferating rim** remains fixed after the onset of necrosis:

$$r(\tau) - r_N(\tau) = \sqrt{\frac{2}{\mu}(1 - c_N)} = \alpha. \tag{7.27}$$

The nutrient concentration in each region is given by

$$c(\xi, \tau) = \begin{cases} c_N & 0 \le \xi < r_N(\tau), \\ c_N + \frac{\mu}{2}(\xi - r_N(\tau))^2 & r_N(\tau) \le \xi \le r(\tau), \\ 1 & \xi > r(\tau), \end{cases}$$

with equation (7.26) relating $r_N(\tau)$ and $r(\tau)$ and ensuring continuity of c at $r(\tau)$.

It remains to solve equation (7.23) subject to the "*initial*" condition (7.24):

$$\begin{aligned} \frac{dr(\tau)}{d\tau} &= \int_{r_N(\tau)}^{r(\tau)} c\, d\xi - \int_0^{r_N(\tau)} \delta\, d\xi \\ &= \int_{r_N(\tau)}^{r(\tau)} \left[\frac{\mu}{2}(\xi - r_N(\tau))^2 + c_N\right] d\xi - \int_0^{r_N(\tau)} \delta\, d\xi \\ &= (r(\tau) - r_N(\tau)) \left[\frac{\mu}{6}(r(\tau) - r_N(\tau))^2 + c_N\right] - \delta r_N(\tau) \\ \Rightarrow \frac{dr(\tau)}{d\tau} &= -\delta r(\tau) + \frac{\alpha}{3}(1 + 2c_N + 3\delta), \quad \text{with} \quad r(\tau_1) = r_1, \end{aligned}$$

where equation (7.27) was used to eliminate $r_N(\tau)$. This ODE for $r(\tau)$ is readily solved:

$$r(\tau) = \left(r_1 - \frac{\beta}{\delta}\right) e^{-\delta(\tau - \tau_1)} + \frac{\beta}{\delta}, \qquad \beta = \frac{\alpha}{3}(1 + 2c_N + 3\delta),$$

when $\delta \neq 0$. In this case, the tumour evolves to a final dimensionless radius of β/δ.

If $\delta = 0$ (dead cells do not degrade) then we have **constant growth,**

$$\frac{dr(\tau)}{d\tau} = \beta, \qquad r(\tau_1) = r_1.$$

In this case, the tumour radius does not attain a steady state: it **grows linearly with** τ.

7.5 Summary

- Growth processes can give rise to **moving boundaries** in mathematical models
- We assumed that domain growth is driven entirely by **nutrient consumption** and, hence, that we must account for nutrient transport and uptake within the growing regions
- As the timescale for **tumour growth** is typically much longer than the timescale for **nutrient diffusion** we can consider the **quasi-steady** limit in which the diffusion equation is **quasi-steady**
- For a simple 1D model, with symmetry about ξ-axis, uniform nutrient concentration outside the tumour, and a cell proliferation rate **linear in** c, we obtained an ODE for tumour boundary $r(\tau)$
- We showed that this model is breaks down for large times because the nutrient concentration at the centre of the tumour becomes negative (before the steady state can be attained)
- We modified the model to allow cells to die when insufficient nutrient $c < c_N$ (necrotic core)
- This modification lead to a model with **2 moving boundaries**: edge of tumour (fixed by proliferation condition) and edge of necrotic core (fixed by condition $c = c_N$)
- If dead cells **degrade** then the new model leads to **steady state** for the tumour; otherwise **linear growth** ensues.
- There are many ways in which we could make the model more realistic. For example, we could solve in 3D geometry, and/or we could incorporate the effects of an externally-supplied drug (chemotherapy).

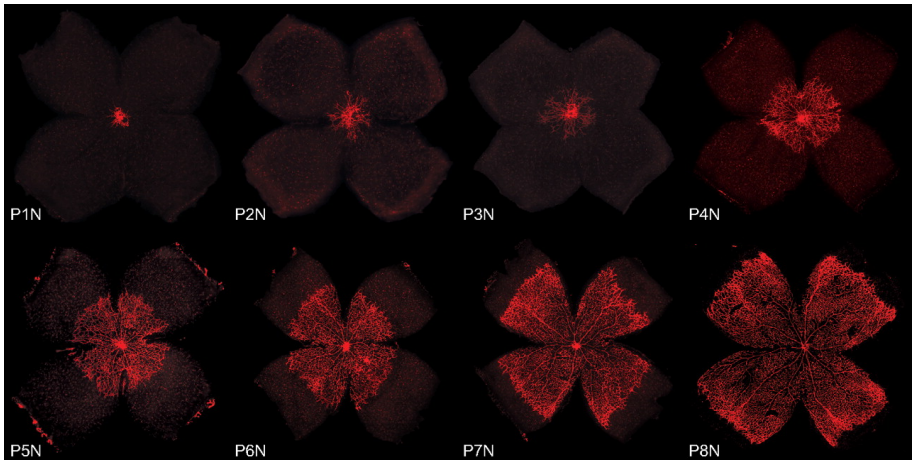


Figure 7.6: Series of images showing how the blood supply to the retina evolves during development.

Chapter 8

From Discrete to Continuum Models

8.1 Introduction

The PDE systems we have studied thus far typically view species as densities or concentrations which vary continuously with position and time. In practice, cells (and other biological quantities, eg animals, people) are individual agents (see Figure 8.1). This raises the question of when it is valid to represent a population of discrete cells as a continuum. In this section, we will investigate this issue by focussing on two case studies:

- One-dimensional, biased, random walks;
- Age-structured populations.

The approaches that we use to derive continuum descriptions for the case studies are explored in more detail in the course *B5.1 Stochastic modelling of biological systems.*

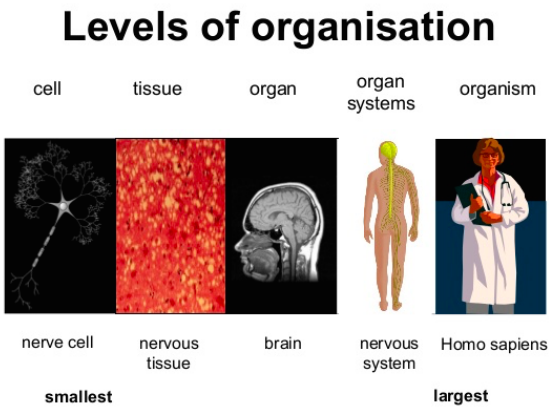


Figure 8.1: A schematic diagram illustrating the different levels of organisation within a human.

References (optional).

- AM Turing (1952 and 1990). The chemical basis of morphogenesis. Bull. Math Biol. 52(1): 153-197.
- T Hillen and KJ Painter (2009). A user's guide to PDE models for chemotaxis. J. Math. Biol. 58: 183-217.

8.2 Biased Random Walks and Advection-Diffusion Equations

Consider a population of cells moving along the real line. We decompose the real line into a series of boxes or compartments, all of width Δx. We denote by $n_i(t)$ the density of cells in the $i-$th box, so that $n_i(t)\Delta x =$ total number of cells in the $i-$th box (see Figure 8.2).

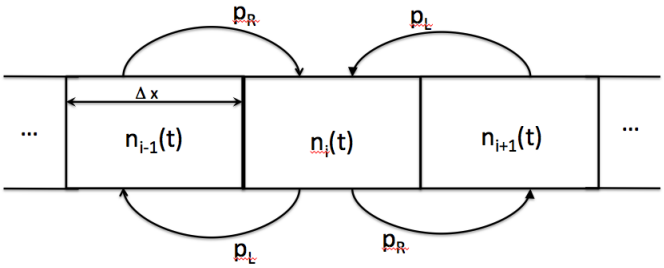


Figure 8.2: Schematic diagram of biased random walk. On each time step, a cell in box i moves left into box $(i-1)$ with probability $0 \leq p_L \leq 1$ or right into box $(i+1)$ with probability $p_R = 1 - p_L$.

We assume that on each time step $\Delta t > 0$, each cell moves left or right with probabilities $0 \leq p_R \leq 1$ and $p_L = 1 - p_R$ respectively (*i.e.* all cells move to a different box during the time step Δt). We use the principle of mass balance to derive **discrete conservation equations (DCEs)** for $n_i(t + \Delta t)\Delta x$, the number of cells in box i at time $t + \Delta t$:

$$n_i(t + \Delta t) \cdot \Delta x = p_R \cdot n_{i-1}(t) \cdot \Delta x + p_L \cdot n_{i+1}(t) \cdot \Delta x$$

$$\Rightarrow n_i(t + \Delta t) - n_i(t) = p_R \cdot n_{i-1}(t) + p_L \cdot n_{i+1}(t) - n_i(t). \tag{8.1}$$

If the box size Δx is sufficiently small then we can identify a continuous density $N(i\Delta x, t) = N(x, t)$ with the discrete density $n_i(t)$, and rewrite equation (8.1) in terms of $N(x, t)$ as follows:

$$N(x, t + \Delta t) - N(x, t) = p_R(N(x - \Delta x, t) - N(x, t)) + p_L(N(x + \Delta x, t) - N(x, t)).$$

We perform Taylor series expansions in x and t to obtain

$$\frac{\partial N}{\partial t} = (p_L - p_R)\left(\frac{\Delta x}{\Delta t}\right)\frac{\partial N}{\partial x} + \frac{(\Delta x)^2}{2\Delta t}\frac{\partial^2 N}{\partial x^2} + O(\Delta t) + O(\Delta x^3/\Delta t).$$

Taking the limit as $\Delta t, \Delta x \to 0$ we deduce

$$\frac{\partial N}{\partial t} = D\frac{\partial^2 N}{\partial x^2} - c\frac{\partial N}{\partial x}, \tag{8.2}$$

where

$$c = \lim_{\Delta t, \Delta x \to 0} (p_R - p_L)\frac{\Delta x}{\Delta t} \qquad \text{and} \qquad D = \lim_{\Delta t, \Delta x \to 0} \frac{(\Delta x)^2}{2\Delta t}.$$

Notes:

- If $p_R = p_L = 1/2$ then $c = 0$ and

  $$\frac{\partial N}{\partial t} = D\frac{\partial^2 N}{\partial x^2},$$

  so that the **diffusion equation** represents the continuum limit of an **unbiased random walk**.

- When deriving the diffusion equation, we assume that the cells move a distance Δx in time Δt. Then their speed is $\Delta x/\Delta t$. If Δx, $\Delta t \to 0$ while $\frac{(\Delta x)^2}{\Delta t} = 2D$, then $\Delta x/\Delta t \Rightarrow \infty$ *i.e.*, the cells move with infinite speed, which is not physically realistic!

- If $\Delta x/\Delta t = O(1)$, then $D \to 0$ as $\Delta x, \Delta t \to 0$ and

  $$\frac{\partial N}{\partial t} + c\frac{\partial N}{\partial x} = 0, \tag{8.3}$$

  where $c > 0$ if $p_R > p_L$. Thus, if the distance travelled is similar in magnitude to the time step of interest then we recover a **rightward-travelling wave**, with $N(x,t) = N_{TW}(x - ct)$, for a population of cells undergoing a rightward-biased random walk.

8.2.1 Boundary Conditions

In practice, cells (or particles) move on a finite domain ($0 \le x \le S$, say). We derive boundary conditions to close equation (8.2) by considering how the boundaries affect cell movement. There are several cases to consider.

Absorbing boundary: a cell that crosses $x = S$ disappears.

$$N(S, t + \Delta t) = p_R N(S - \Delta x, t)$$

$$\Rightarrow 0 = (1 - p_R)N(S,t) + O(\Delta t) + O(\Delta x)$$

$$\Rightarrow N(S,t) = 0 \quad \text{provided} \quad p_R \neq 1.$$

Note: if $p_R = 1$ then we cannot have $(p_R - p_L) \to 0$ as $\Delta t, \Delta x \to 0$, *i.e.* we cannot recover an advection-diffusion equation in this limit.

Reflecting boundary: a cell that attempts to move beyond $x = S$ bounces back.

$$N(S, t + \Delta t) = p_R N(S - \Delta x, t) + p_R N(S, t)$$

$$\Rightarrow N(S,t) + \Delta t \cdot \frac{\partial N}{\partial t}(St) = p_R \left(N(S,t) - \Delta x \frac{\partial N}{\partial x}(S,t) \right) + p_R N(S,t)$$

$$\Rightarrow 0 = (p_R - p_L)N(S,t) - p_R \Delta x \frac{\partial N}{\partial x}(S,t) - \Delta t \frac{\partial N}{\partial t}(S,t)$$

Multiply by $(\Delta x/\Delta t)$, recalling the definitions of c and D, to deduce that

$$0 = cN(S,t) - 2p_R D \frac{\partial N}{\partial x}(S,t) - \Delta x \frac{\partial N}{\partial t}(S,t)$$

so that, in the limit as $\Delta t, \Delta x \to 0$, we have

$$\Rightarrow \mu \frac{\partial N}{\partial x}(S,t) = \frac{c}{2D} N(S,t),$$

where $\mu = p_R$.

8.2.2 Model Extensions

There are many ways to extend the basic framework used above to derive an advection-diffusion equation for cell movement. The key point is to include all the relevant (source and sink) terms when using the principle of mass balance to derive the discrete conservation equations. Other processes that could be included are:

- Chemotaxis: here, the probabilities of moving left and right from box i will depend on the concentration of a diffusible chemoattractant in boxes $i \pm 1$ and box i (see Question Sheet 4);
- Proliferation and death;
- Competition for space.

Exercise: Suppose that cells at position (x, t)

- move left/right with probabilities p_L, p_R,
- stay still with probability p_S,
- die with probability p_D.

Suppose further that $p_L + p_R + p_S + p_D = 1$. Derive a continuum PDE (reaction-diffusion-advection) equation for the cell population. State clearly how the probabilities p_L, p_R, p_S and p_D should scale with Δt and Δx in order to recover your PDE.

Exercise: A population of cells moves on a uniform, 2D grid (*i.e.* $\Delta x = \Delta y$). The probabilities of moving up, down, left and right are p_N, p_S, p_W and p_E respectively. The sum of these probabilities is one. Derive the corresponding diffusion equation for the cell population, stating clearly how p_N, p_S, p_W and p_E must scale with Δt and Δx.

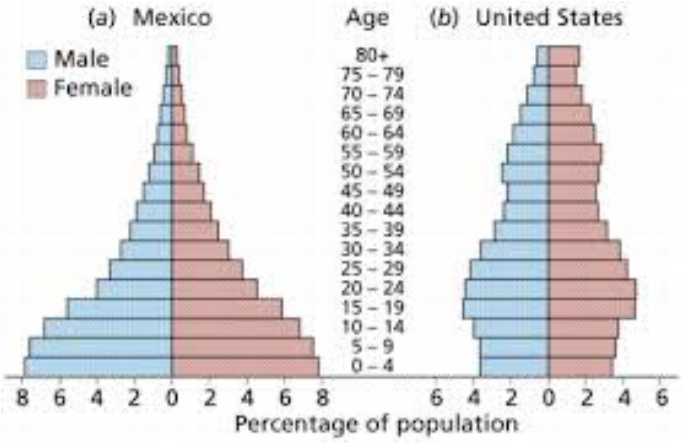


Figure 8.3: Histograms highlighting differences in the age distribution of males and females in Mexico and the USA

Notes:

- Turing's original model of pattern formation was formulated for a discrete chain of cells [see: AM Turing (1952). The chemical basis of morphogenesis. Bull. Math. Biol. 52(1): 153-197].
- In the above examples, we assume that, at a given time, cells move 'right' and 'left' (and possibly 'up' and 'down') independently of what they were doing at the previous time point. In practice, cells cannot instantaneously change direction. Models that account for this will be introduced in the course *Stochastic modelling of biological systems'*.

8.3 Age-Structured Populations

Simple models of population dynamics treat all individuals as identical whereas sophisticated models may consider distinct subgroups. For example, SIR models of diseases such as measles, chicken pox and HIV/AIDS typically decompose the population into 3 subgroups (Susceptibles, Infectives and Recovered), and time-dependent ODEs describe their evolution. In practice, many processes involved in population growth and/or disease spread depend on age: young and old individuals may have higher mortality rates and be more susceptible to a disease; only individuals of a certain age may be able to produce offspring; school children who frequently come into contact with each other may play a stronger role in disease transmission than other age-groups. In this section we will use the approach outlined in the previous section to develop age-structured population models and then show how they can be used to study population dynamics, epidemics and cell-cycle dynamics.

References

- AG McKendrick (1926). Applications of mathematics to medical problems. *Proc. Edinb. Math. Soc.* 44: 98-130.
- GF Webb (1985).Theory of age dependent population dynamics. Marcel Dekker, New York.
- H von Foerster (1959). Some remarks on changing populations. In: *The Kinetics of Cellular Proliferation* (F. Stahlman Jr., ed.). New York, Grune and Stattan, pp. 383-407.
- F Hoppensteadt (1975). Mathematical theories of populations: demographics, genetics and epidemics. *SIAM*, Philadelpia, Penn.

Model development (von Foerster's Equation)

Consider a population (of humans, cells, animals, ...) structured by age. Let $N_i(t)$ denote the number of individuals of age i at time t ($i = 0, 1, 2, \ldots$). Over one time step Δt (= 1 year, 1 day, ...), individuals may die, age and reproduce. As before, we use the principle of mass balance to derive discrete conservation equations for $N_i(t)$. For $i = 1, 2, 3, \ldots$, we have

$$\left(\begin{array}{c}\text{change in number of}\\ \text{individuals of age } i\\ \text{during time step}\end{array}\right) = \left(\begin{array}{c}\text{loss due to}\\ \text{cell death}\end{array}\right) + \left(\begin{array}{c}\text{change due}\\ \text{to ageing}\end{array}\right)$$

$$\Rightarrow N_i(t+\Delta t) - N_i(t) = -\mu N_i(t)\Delta t + N_{i-1}(t) - N_i(t).$$

For $i = 0$ (new-borns), we have

$$N_0(t+\Delta t) - N_0(t) = \underbrace{\Sigma_{i=0}^{\infty} b_i N_i(t)\Delta t}_{\text{reproduction}} - \underbrace{N_0(t)}_{\text{ageing}},$$

or, equivalently,

$$N_0(t+\Delta t) = \Sigma_{i=0}^{\infty} b_i N_i(t)\Delta t.$$

We identify with $N_i(t)$ a cell density $n(a,t)$ where

$$N_i(t) = n(i\Delta a, t)\Delta a.$$

Typically $\Delta a, \Delta t$ are small and, so, we can perform Taylor series expansions to switch from a discrete to a continuous description. In more detail,

$$n(a, t+\Delta t) - n(a,t) = -\mu n(a,t)\Delta t + n(a-\Delta a, t) - n(a,t)$$

$$\Rightarrow \frac{\partial n}{\partial t}\Delta t + O(\Delta t^2) = -\mu n(a,t)\Delta t - \frac{\partial n}{\partial a}\Delta a + O(\Delta a^2).$$

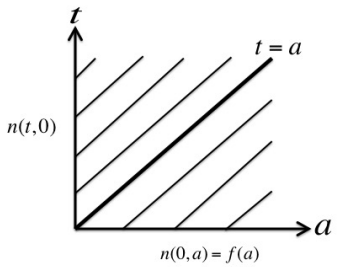


Figure 8.4: Schematic showing the characteristic curves associated with von Foerster's equation.

Dividing by Δt and taking the limit as $\Delta t, \Delta a \to 0$, with $(\Delta a/\Delta t = 1)$ since a is chronological age, we deduce that $n(a,t)$ satisfies the following linear PDE:

von Foerster's Equation:

$$\frac{\partial n}{\partial t} + \frac{\partial n}{\partial a} = -\mu(a)n. \tag{8.4}$$

Boundary conditions for von Foerster's Equation

$$n(0,t) = \int_0^\infty b(a)n(a,t)da, \quad \text{birth rate of population.} \tag{8.5}$$

where $b(a)$ = birth rate of individuals of age a.

Initial conditions for von Foerster's Equation

$$n(a,0) = f(a), \quad \text{initial age-distribution of population,} \tag{8.6}$$

We use the method of characteristics to solve equations (8.4)-(8.6). The characteristic curves are straight lines along which $\frac{da}{dt} = 1$ and, hence,

$$a = \begin{cases} t + a_0, & \text{for } a > t \\ t - t_0, & \text{for } a < t \end{cases}$$

Here a_0 represents the *initial age* of an individual who has age $a > t$ at time t, and t_0 represents the time at which an individual of age $0 < a < t$ was born.

From Figure 8.4, it is clear that

- characteristic curves for which $0 < t < a$ emanate from the initial data (where $t = 0$)

- characteristics for which $0 < a < t$ emanate from the boundary conditions (at $a = 0$).

Since information from the boundary and initial conditions propagates along characteristics, we deduce that the solution will have different forms in $0 < t < a$ and $0 < a < t$. We construct these solutions below.

Region 1 ($0 < t < a$). In this region

$$\frac{\partial n}{\partial t} + \frac{\partial n}{\partial a} = -\mu(a)n, \quad n(0,a) = f(a),$$

$$\Rightarrow a = t + a_0, \quad \frac{dn}{dt} = -\mu n, \quad n(a_0, 0) = f(a_0).$$

$$\Rightarrow n(a,t) = f(a_0)e^{-\int_{a_0}^{a}\mu(\theta)d\theta}.$$

Since $a_0 = a - t$ along characteristic curves, we have

$$\boxed{n(a,t) = f(a-t)e^{-\int_{a-t}^{a}\mu(\theta)d\theta}, \qquad \text{for } 0 < t < a. \tag{8.7}}$$

Region 2 ($0 < a < t$). In this region

$$\frac{\partial n}{\partial t} + \frac{\partial n}{\partial a} = -\mu(a)n, \quad n(0,t) = \int_0^\infty b(a)n(a,t)da,$$

$$\Rightarrow a = t - t_0, \quad \frac{dn}{dt} = -\mu n, \quad n(0,t_0) = \int_0^\infty b(a)n(a,t_0)da \text{ when } t = 0.$$

$$\boxed{\Rightarrow n(a,t) = n(0,t-a)e^{-\int_{0}^{a}\mu(\theta)d\theta}, \qquad \text{for } 0 < a < t. \tag{8.8}}$$

Using Equations (8.7) and (8.8), we deduce

$$n(0,t) = \int_0^\infty b(a)n(a,t)da = \underbrace{\int_0^t b(a)n(a,t)da}_{(0<a<t)} + \underbrace{\int_t^\infty b(a)n(a,t)da}_{(t<a<\infty)},$$

and, hence,

$$\boxed{n(0,t) \equiv \int_0^t \left(b(a)n(0,t-a)e^{-\int_0^a \mu(\theta)d\theta}\right)da + \int_t^\infty \left(b(a)f(a-t)e^{-\int_{a-t}^a \mu(\theta)d\theta}\right)da, \tag{8.9}}$$

Equation (8.9) defines $n(0,t)$ in terms of known functions and $n(0,\tau)$ $(0 \le \tau < t)$. Although this is a linear equation for $n(0,t)$, it is usually difficult to solve.

Worked Example

Suppose $\mu(a) = \mu$, constant and $b(a) = bH(a_R - a)$ where $H(.)$ is the Heaviside step function ($H(x) = 1$ if $x > 0$ and $H(x) = 0$ otherwise). Then $n(a,t)$ satisfies

$$\frac{\partial n}{\partial t} + \frac{\partial n}{\partial a} = -\mu n,$$

with $n(0,a) = 1$ and $n(0,t) = \int_0^\infty b(a)n(a,t)da$, where

$$b(a) = \begin{cases} b, & \text{for } 0 < a < a_R, \\ 0, & \text{otherwise.} \end{cases}$$

Region 1 ($0 < t < a$). With $\mu(a) = \mu$ and $f(a) = 1$, equation (8.7) supplies

$$n(a,t) = e^{-\mu t}. \tag{8.10}$$

Region 2 ($0 < a < t$). In this region, equations (8.8) and (8.9) supply

$$n(a,t) = n(0,t-a)e^{-\mu a} \qquad \text{where} \qquad n(0,t) = b\int_0^{a_R} n(a,t)da. \tag{8.11}$$

When solving equation (8.11), we consider separately cases for which $0 < t < a_R$ and $a_R < t < \infty$.

Case 1. If $0 < t < a_R$, then

$$n(0,t) = b\int_0^{a_R} n(a,t)da = b\int_0^t n(a,t)da + b\int_t^{a_R} n(a,t)da$$

$$\Rightarrow n(0,t) = b\int_0^t n(0,t-a)e^{-\mu a}da + b\int_t^{a_R} e^{-\mu t}da, \tag{8.12}$$

where we have exploited the fact that $n(a,t) = e^{-\mu t}$ for $0 < t < a$. We rewrite equation (8.12) in terms of $N(t) = n(0,t)$:

$$\begin{aligned} N(t) &= b\int_0^t N(t-a)e^{-\mu a}da + b(a_R - t)e^{-\mu t} \\ &= b\int_0^t N(\tau)e^{-\mu(t-\tau)}d\tau + b(a_R - t)e^{-\mu t}. \end{aligned}$$

Differentiating this expression for $N(t)$ with respect to t we obtain the following ODE for $N(t)$:

$$\frac{dN}{dt} = (b-\mu)N - be^{-\mu t}$$

$$\Rightarrow N(t) = \hat{N}e^{(b-\mu)t} + e^{-\mu t}$$

where $\hat{N}$ is a constant of integration. Substituting with $N(t) = n(0,t)$ in equation (8.11) (and recalling that $n(a,t) = e^{-\mu t}$ for $0 < t < a$), we deduce that for $0 < a < t < a_R$

$$n(a,t) = \begin{cases} N(t-a)e^{-\mu a} = e^{-\mu t}\left(\hat{N}e^{b(t-a)} + 1\right) & \text{for } 0 < a < t \\ e^{-\mu t} & \text{for } 0 < t < a. \end{cases} \tag{8.13}$$

Case 2. For $a_R < t$, and with $N(t) = n(0,t)$, equations (8.11) supply

$$N(t) = b\int_0^{a_R} N(t-a)e^{-\mu a}da = b\int_{t-a_R}^{t} N(\tau)e^{-\mu(t-\tau)}d\tau,$$

where, as before, $N(t) = n(0,t)$. If we differentiate this expression for $N(t)$ with respect to t we obtain a **delay differential equation** for $N(t)$:

$$\frac{dN}{dt} = (b-\mu)N(t) - bN(t-a_R)e^{-\mu a_R}.$$

We seek solutions of the form $N(t) = \tilde{N}e^{\omega t}$ where

$$\omega = (b-\mu) - be^{-(\omega+\mu)a_R}. \tag{8.14}$$

With $N(t) = \tilde{N}e^{\omega t}$, equation (8.13) supplies

$$n(a,t) = N(t-a)e^{-\mu a} = \tilde{N}e^{\omega(t-a)}e^{-\mu a}.$$

If $\Re e(\omega) > 0$ then our age-dependent population increases over time; if $\Re e(\omega) < 0$ then it decays. For a stable, age-structured population, we require $\omega = 0$. From equation (8.14) we deduce that the population will be stable if the birth rate b, the death rate μ and the parameter a_R satisfy

$$b = \frac{\mu}{1 - e^{-\mu a_R}}. \tag{8.15}$$

In this case (when $\omega = 0$),

$$n(a,t) = \begin{cases} e^{-\mu a}, & \text{for } 0 < a < t \\ e^{-\mu t}, & \text{for } a_R < t < a. \end{cases}$$

Remarks: Equation (8.15) states how the birth and death rates should be related in order to achieve a stable, age-structured population (i.e. one which neither explodes nor dies out). We can use this equation to draw the following conclusions:

- $b > \mu$: since only individuals of age $0 < a < a_R$ reproduce, the birth rate b must exceed the death rate to achieve a stable population;
- If $a_R \to \infty$, then $b \to \mu$: if all individuals reproduce, then a stable population will be achieved if the birth and death rates balance;
- If $a_R \to 0$, then $b \to 1/a_R \gg 1$: as the reproductive lifespan of the population decreases, their birth rate must increase in order to maintain a stable population.

Separable Solutions for Age-Structured Models

Guided by the solution for the worked example, we now seek separable solutions to von Foerster's equation of the form:

$$n(a,t) = e^{\gamma t}F(a),$$

i.e. we assume that the age distribution is altered by a time-dependent factor which decays or grows depending on whether $\Re e(\gamma) < 0$ or $\Re e(\gamma) > 0$. Substituting this separable form into equation (8.4) supplies the following ODE for $F(a)$:

$$\frac{dF}{da} = -(\mu(a) + \gamma)F \quad \Rightarrow \quad F(a) = F(0)\exp\left\{-\gamma a - \int_0^a \mu(\theta)d\theta\right\}.$$

Imposing boundary conditions (8.6) we deduce

$$\begin{aligned} n(0,t) &= e^{\gamma t}F(0) = \int_0^\infty b(a)e^{\gamma t}F(a)da \\ &= e^{\gamma t}F(0)\int_0^\infty b(a)\exp\left\{-\gamma a - \int_0^a \mu(\theta)d\theta\right\}da. \end{aligned}$$

Cancelling by a nonzero factor $e^{\gamma t}F(0)$, we deduce

$$1 = \int_0^\infty b(a)\exp\left\{-\gamma a - \int_0^a \mu(\theta)d\theta\right\}da \;\equiv\; \Phi(\gamma). \tag{8.16}$$

Since $\Phi(\gamma)$ is a monotonic decreasing function of γ, we deduce that equation (8.16) admits a unique solution for γ.

In general, a separable solution will not satisfy the initial conditions $n(0,a) = f(a)$. However, in the limit as $t \to \infty$, equation (8.9) supplies

$$n(0,t) \sim \int_0^t b(a)n(0,t-a)\exp\left\{-\int_0^a \mu(\theta)d\theta\right\}da.$$

If we seek solutions to this equation of the form $n(a,t) \sim e^{\gamma t}F(a)$, then equation (8.16) is recovered.

Exercise 8.2: By seeking **separable solutions** to the worked example, verify that equation (8.16) is a necessary condition for obtaining a stable, age-structured population.

Exercise 8.3: Suppose that $\mu(a) = \mu$, constant, $n(0,a) = 1$ and $b(a)$ is given by

$$b(a) = \begin{cases} b, & \text{for } a_L < a < a_R, \\ 0, & \text{otherwise.} \end{cases}$$

By seeking separable solutions to von Foerster's equation, show that

$$b = \mu/(e^{-\mu a_L} - e^{-\mu a_R})$$

is a necessary condition for obtaining a stable, age-structured population. Explain what happens in the limit at $(a_R - a_L) \to 0$.

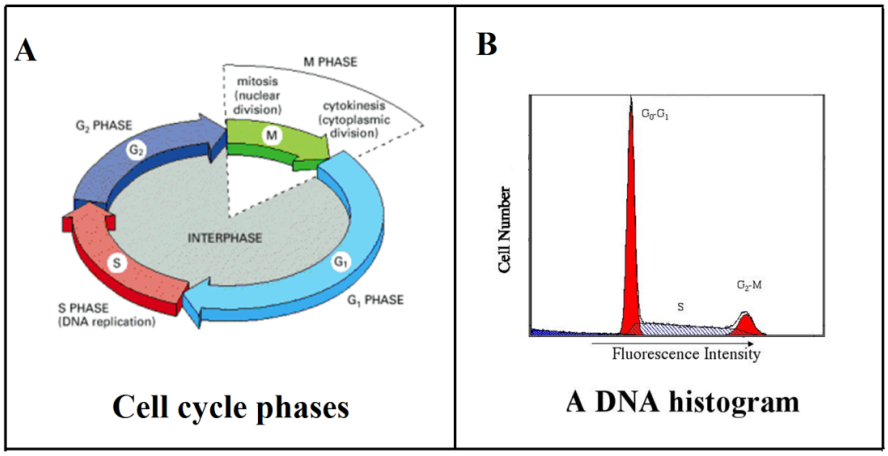


Figure 8.5: Schematic showing the different phases of the cycle cycle and how DNA content varies with cell cycle phase.

8.3.1 Structured Models for Proliferating Cells

Cells reproduce by duplicating their contents and then dividing in two (see Figure 8.6). The duration of the cell cycle varies widely: from 8 minutes in fly embryos to more than a year for mammalian liver cells. In this subsection we will adapt the age-structured models developed above to study cell cycle dynamics.

Worked Example

A tissue contains two types of cells:

- $p(t,s)$ = number of cycling cells at position $0 \leq s < T$ of the cell cycle at time t;
- $q(t,s)$ = number of quiescent (or non-cycling) cells arrested at position $0 \leq s < t$ at time t.

The evolution of the cycling population is governed by a nonlinear PDE similar to equation (8.4); as quiescent cells do not progress around the cell cycle, their evolution is governed by a time-dependent ODE. In particular

$$\frac{\partial p}{\partial t} + \frac{\partial p}{\partial s} = -\underbrace{\mu N p}_{\text{cell death}} - \underbrace{\lambda N p}_{\text{exit cell cycle}} + \underbrace{\frac{\gamma q}{N_0 + N}}_{\text{re-enter cell cycle}}, \tag{8.17}$$

$$\frac{\partial q}{\partial t} = -\mu N q + \lambda N p - \frac{\gamma q}{N_0 + N} \tag{8.18}$$

where $N(t) = \int_0^T (p(t,s)+q(t,s))ds =$ total number of cells at time t and μ, λ, γ and N_0 are positive constants. In addition, we impose the following boundary and initial conditions:

$$p(0,s) = p_0(s), \quad q(0,s) = q_0(s), \quad p(t,0) = 2p(t,T).$$

The final condition states that **at the end of the cell cycle (when $s = T$) a cell produces 2 cells of age $s = 0$.**

We seek separable solutions of the form $p(t,s) = e^{\theta t}P(s)$ and $q(t,s) = e^{\theta t}Q(s)$ with $\theta = 0$. Then $N(t) = \int_0^T (p(t,s) + q(t,s))ds = N$, constant, and

$$\frac{dP}{ds} = -(\mu+\lambda)NP + \frac{\gamma Q}{N_0+N} \quad \text{and} \quad 0 = \lambda NP - \left(\mu N + \frac{\gamma}{N_0+N}\right) Q$$

$$\Rightarrow \frac{dP}{ds} = -\mu N(P+Q) \quad \text{and} \quad Q = \left(\frac{\lambda N(N_0+N)}{\gamma+\mu N(N_0+N)}\right) P.$$

Eliminating Q we deduce

$$\frac{1}{P}\frac{dP}{ds} = -\mu N\left(1 + \frac{\lambda N(N_0+N)}{\gamma+\mu N(N_0+N)}\right) \equiv -\omega, \text{ say.}$$

$$\Rightarrow P(s) = P_\infty e^{-\omega s} \quad \text{and} \quad Q(s) = Q_\infty e^{-\omega s} = \left(\frac{\lambda N(N_0+N)}{\gamma+\mu N(N_0+N)}\right) P_\infty e^{-\omega s}.$$

Now $P(s=0) = 2P(s=T) \Rightarrow 1 = 2e^{-\omega T}$

$$\Rightarrow \quad \frac{\ln 2}{T} = \omega = \mu N\left(1 + \frac{\lambda N(N_0+N)}{\gamma+\mu N(N_0+N)}\right).$$

This equation defines the total population size N in terms of the cell cycle length T and other parameters.

We determine the proportion of cycling cells by noting that

$$N = \int_0^T [P(s)+Q(s)]ds = \left(\frac{1-e^{-\omega T}}{\omega}\right)(P_\infty + Q_\infty),$$

$$\Rightarrow 2\omega N = \frac{\omega}{\mu N}P_\infty \quad \text{or} \quad P_\infty = 2\mu N^2,$$

$$\Rightarrow P(s) = 2\mu N^2 e^{-\omega s} \quad \text{and} \quad Q(s) = 2\mu N^2\left(\frac{\lambda N(N_0+N)}{\gamma+\mu N(N_0+N)}\right)e^{-\omega s}.$$

Exercise 8.4:

- Suppose $\gamma \to 0$: how are N, P and Q defined? Interpret your results.
- Suppose $\lambda \to 0$: how are N, P and Q defined? Interpret your results.

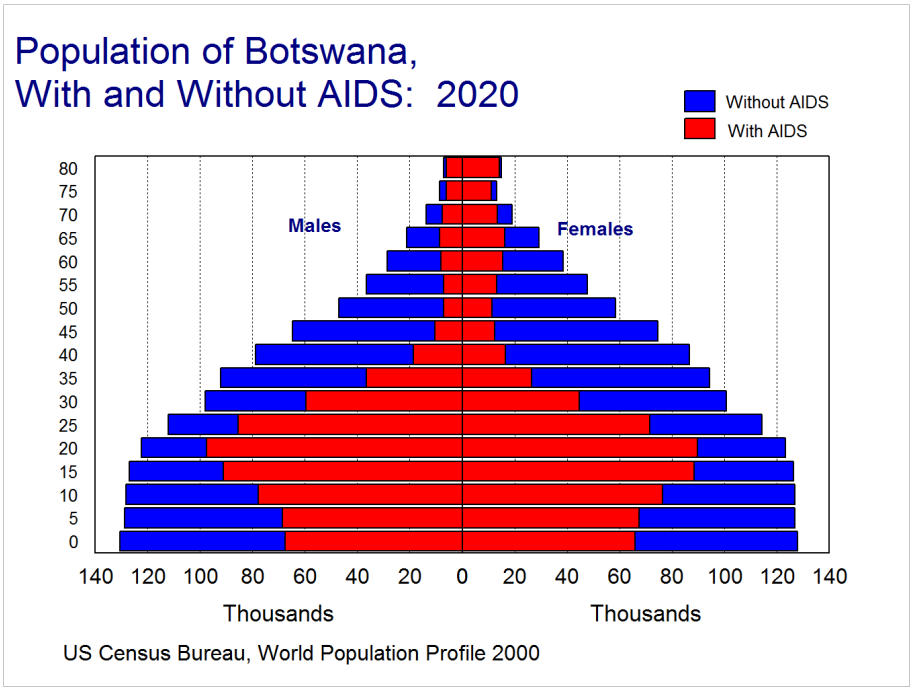


Figure 8.6: Schematic showing the age distribution of people with AIDS in Botswana.

8.3.2 Age-dependent epidemic models (not examinable)

One of the main reasons for developing age-structured models is to study the spread of diseases for which age is an important factor for susceptibility and infectiousness. For example, vulnerability to AIDS/HIV decreases dramatically with age (see Figure 8.5). In this section, we will extend our previous age-structured model of population growth to describe the spread of a disease. In order to do this, we divide the population into two, age-structured sub-populations:

$$\textbf{Susceptibles}, S(t,a) \text{ and } \textbf{Infectives}, I(t,a).$$

Applying the arguments used to derive equation (8.4), we assume that $S(t,a)$ and $I(t,a)$ satisfy the following PDEs:

$$\frac{\partial S}{\partial t} + \frac{\partial S}{\partial a} = \underbrace{-\left(\int_0^\infty r(\alpha) I(t,\alpha) d\alpha\right) S(t,a)}_{\text{infection}} - \underbrace{\mu S}_{\text{death}}, \tag{8.19}$$

$$\frac{\partial I}{\partial t} + \frac{\partial I}{\partial a} = \underbrace{\left(\int_0^\infty r(\alpha) I(t,\alpha) d\alpha\right) S(t,a)}_{\text{infection}} - \underbrace{\mu I}_{\text{death}}, \tag{8.20}$$

where, for simplicity, we assume that susceptibles and infectives die at the same, constant rate μ.

We close equations (8.19) - (8.20) by prescribing the following initial and boundary conditions:

$$S(0,a) = S_0(a), \;\; I(0,a) = I_0(a), \;\; S(t,0) = \int_0^\infty b(a)S(t,a)da, \;\; I(t,0) = 0,$$

i.e. all newborns are susceptible.

In general solutions for $S(t,a)$ and $I(t,a)$ require numerical approaches. Here, we **seek separable solutions** of the form

$$S(t,a) = e^{\gamma t}S(a), \quad I(t,a) = e^{\gamma t}I(a),$$

with $\gamma = 0$, i.e. **time-independent solutions**. Then equations (8.19) and (8.20) supply:

$$\frac{dS}{da} = -\left(\int_0^\infty r(\alpha)I(\alpha)d\alpha\right)S(a) - \mu S,$$

$$\frac{dI}{da} = \left(\int_0^\infty r(\alpha)I(\alpha)d\alpha\right)S(a) - \mu I,$$

$$\Rightarrow \frac{d}{da}(S+I) = -\mu(S+I), \qquad \Rightarrow S + I = \Lambda e^{-\mu a},$$

where Λ is a constant of integration. If $r(a) = r$, constant, then

$$\frac{dI}{da} = r\underbrace{\int_0^\infty I(\alpha)d\alpha}_{\equiv I_{TOT}}S(a) - \mu I = rI_{TOT}\Lambda e^{-\mu a} - (\mu + rI_{TOT})I,$$

$$\Rightarrow I(a) = \Lambda e^{-\mu a} + Ae^{-(\mu + rI_{TOT})a}.$$

where

$$I_{TOT} = \int_0^\infty I(a)da = \frac{A}{\mu + rI_{TOT}} + \frac{\Lambda}{\mu},$$

$$\Rightarrow I(a) = \Lambda e^{-\mu a} - \left(\frac{\Lambda}{\mu} - I_{TOT}\right)(\mu + rI_{TOT})\, e^{-(\mu + rI_{TOT})a}.$$

Now

$$S(a) = \Lambda e^{-\mu a} - I(a) = \left(\frac{\Lambda}{\mu} - I_{TOT}\right)(\mu + rI_{TOT})e^{-(\mu + rI_{TOT})a}$$

where $S(0) = \int_0^\infty b(a)S(a)da$. Substituting for $S(0)$ and $S(a)$ we deduce that I_{TOT} satisfies

$$1 = \int_0^\infty b(a)e^{-(\mu + rI_{TOT})}da.$$

If $b(a) = b^{-\theta a}$ (i.e. an individual's birth rate decreases with age), then

$$I_{TOT} = (b - \mu - \theta)/r.$$

Exercise 8.5: Derive an expression for I_{TOT} when $b(a) = bae^{-\theta a}$.

Appendix A

The phase plane

Throughout this appendix we will be concerned with systems of two coupled, first-order, autonomous, non-linear ordinary differential equations.

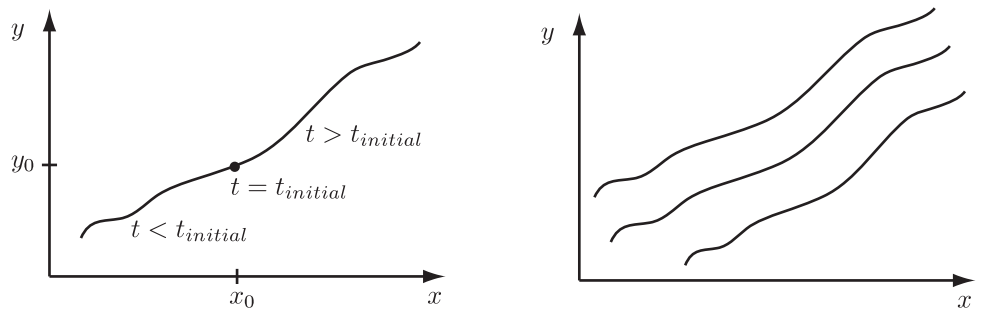
Disclaimer. This material should have been covered elsewhere (for example in your course on differential equations) and hence below is intended to review, rather than introduce and lecture this topic.

We can represent solutions to the equations

$$\frac{\mathrm{d}x}{\mathrm{d}t} = X(x,y), \tag{A.1}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = Y(x,y), \tag{A.2}$$

as trajectories (or "integral paths") in the phase plane, that is the (x,y) plane. Suppose, for the initial condition $x(t = t_{initial}) = x_0$, $y(t = t_{initial}) = y_0$ we plot, in the (x,y) plane, the solution of (A.1):



We can do exactly the same for all the values of $\{t_{initial}, x_{initial}, y_{initial}\}$, to build-up a graphical representation of the solutions to the equations (A.1) and (A.2) for many initial conditions. This plot is referred to as the *"phase plane portrait"*.

A.1 Properties of the phase plane portrait

The gradient of the integral path through the point (x_0, y_0) is given by

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}t} \Big/ \frac{\mathrm{d}x}{\mathrm{d}t} = \left(\frac{Y(x,y)}{X(x,y)}\right)\Bigg|_{(x_0,y_0)} = \frac{Y(x_0,y_0)}{X(x_0,y_0)}. \tag{A.3}$$

Key point 1. Note that if $Y(x_0, y_0) = 0$ and $X(x_0, y_0) \neq 0$ then

$$\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)\Bigg|_{(x_0,y_0)} = 0, \tag{A.4}$$

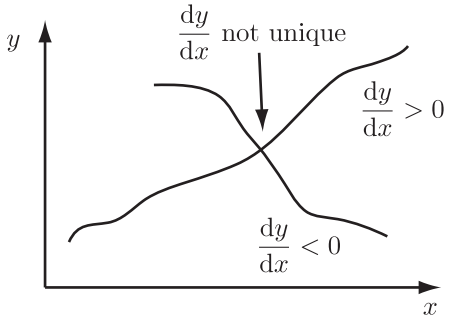
which corresponds to a horizontal line segments in the phase plane.

Key point 2. If $Y(x_0, y_0) \neq 0$ and $X(x_0, y_0) = 0$ then

$$\left|\frac{\mathrm{d}y}{\mathrm{d}x}\right| \to \infty \quad \text{as} \quad (x, y) \to (x_0, y_0), \tag{A.5}$$

which corresponds to a vertical line segment in the phase plane.

Key point 3. Assuming that either $X(x_0, y_0) \neq 0$ or $Y(x_0, y_0) \neq 0$, then two path integral curves do not cross at the point (x_0, y_0). This is because under these circumstances $\mathrm{d}y/\mathrm{d}x$ takes a unique value, *i.e.* the following is *not* possible:



A.2 Equilibrium points

Definition. A point in the phase plane where $X(x_0, y_0) = Y(x_0, y_0) = 0$ is defined to be an *equilibrium point*, or equivalently, *a stationary point.*

The reason for the above definition is because if $(x, y) = (x_0, y_0)$ then both $\mathrm{d}x/\mathrm{d}t$ and $\mathrm{d}y/\mathrm{d}t$ are zero, and hence (x, y) do not change as t increases; hence $x(t)$, $y(t)$ remain at (x_0, y_0) for all time.

Key point 1. Integral curves cannot cross at points which are not equilibrium points.

Key point 2. If an integral path ends it must end on a stationary point.

Key point 3. As we shall see below, equilibrium points are only approached as $t \to \infty$ or $t \to -\infty$.

However, what about the gradient of integral paths at (x_0, y_0)? We informally have

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{0}{0}, \tag{A.6}$$

which is not uniquely defined—the value ultimately depends on the details of how quickly $X(x, y)$ and $Y(x, y)$ approach zero as $(x, y) \to (x_0, y_0)$, and this generally depends on the direction upon which (x, y) approaches (x_0, y_0).

A.2.1 Equilibrium points: further properties

Suppose the equations (A.1) and (A.2) have an equilibrium point at (x_0, y_0). Thus $X(x_0, y_0) = Y(x_0, y_0) = 0$. To determine the behaviour of integral paths close to the equilibrium point we write

$$x = x_0 + \bar{x}, \qquad y = y_0 + \bar{y}, \tag{A.7}$$

where it is assumed that $\bar{x}$, $\bar{y}$ are sufficiently small to allow the approximations that we will make below.

By Taylor expansion, we have

$$\begin{aligned} X(x, y) &= X(x_0 + \bar{x}, y_0 + \bar{y}) = X(x_0, y_0) + \bar{x}\frac{\partial X}{\partial x}(x_0, y_0) + \bar{y}\frac{\partial X}{\partial y}(x_0, y_0) + \text{h.o.t.} \\ &= \bar{x}\frac{\partial X}{\partial x}(x_0, y_0) + \bar{y}\frac{\partial X}{\partial y}(x_0, y_0) + \text{h.o.t.}, \end{aligned} \tag{A.8}$$

using the fact $X(x_0, y_0) = 0$. Similarly, we have

$$\begin{aligned} Y(x, y) &= Y(x_0 + \bar{x}, y_0 + \bar{y}) = Y(x_0, y_0) + \bar{x}\frac{\partial Y}{\partial x}(x_0, y_0) + \bar{y}\frac{\partial Y}{\partial y}(x_0, y_0) + \text{h.o.t.}, \\ &= \bar{x}\frac{\partial X}{\partial x}(x_0, y_0) + \bar{y}\frac{\partial X}{\partial y}(x_0, y_0) + \text{h.o.t.} \end{aligned} \tag{A.9}$$

Note that x_0 and y_0 are constant, and hence have zero time derivative. Hence, by use of Taylor expansions and neglecting higher orders (*i.e.* taking $\bar{x}$, $\bar{y}$ sufficiently small), we can neglect terms of the order $\mathcal{O}\left(\bar{x}\bar{y}, \bar{x}^2, \bar{y}^2\right)$ and hence we can write equations (A.1) and (A.2) in the form

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = \begin{pmatrix} \frac{\partial X}{\partial x}(x_0, y_0) & \frac{\partial X}{\partial y}(x_0, y_0) \\ \frac{\partial Y}{\partial x}(x_0, y_0) & \frac{\partial Y}{\partial y}(x_0, y_0) \end{pmatrix} \boldsymbol{u} \stackrel{def}{=} \boldsymbol{J}\boldsymbol{u} \quad \text{where} \quad \boldsymbol{u} \stackrel{def}{=} \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix}. \tag{A.10}$$

Definition. The matrix

$$\boldsymbol{J} = \begin{pmatrix} \frac{\partial X}{\partial x}(x_0, y_0) & \frac{\partial X}{\partial y}(x_0, y_0) \\ \frac{\partial Y}{\partial x}(x_0, y_0) & \frac{\partial Y}{\partial y}(x_0, y_0) \end{pmatrix}, \tag{A.11}$$

is defined to be the *Jacobian* matrix at the equilibrium point (x_0, y_0).

A.3 Summary

The key points thus far are as follows.

1. We have taken the full *non-linear* equation system, (A.1) and (A.2), and expanded about one of its (possibly many) equilibrium points taken to be located at (x_0, y_0), using Taylor expansions of $X(x, y)$, $Y(x, y)$.

2. We assume that we are sufficiently close to (x_0, y_0) to enable us to only consider linear terms of the order of $(x - x_0)$, $(y - y_0)$.

3. In this way, we obtain a set of two coupled, *linear*, autonomous ordinary differential equations, *i.e.* equation (A.10) above, which in principle we can solve!

4. This procedure is sometimes referred to as "a *linearisation* of equations (A.1) and (A.2) about the point (x_0, y_0)".

5. In *virtually all* cases the behaviour of the linearised system is the same as the behaviour of the full non-linear equations sufficiently close to the point (x_0, y_0). In this respect one should note that the statement immediately above can be formulated more rigorously and proved for all the types of stationary points except:

   - centre type equilibrium points, *i.e.* case [3c] below;
   - the degenerate cases where $\lambda_1 = 0$ and/or $\lambda_2 = 0$, which are briefly mentioned in item 2 on page (94). These stationary points can be considered non-examinable.

   The relevant theorem is "Hartmann's theorem", as discussed further in P. Glendinning, Stability, Instability and Chaos [**?**].

6. However, one should also note that the solution of the linearised equations may behave substantially differently from the solutions of the full *non-linear* equations, (A.1) and (A.2), sufficiently far from (x_0, y_0).

A.4 Investigating solutions of the linearised equations

We now have a set of two coupled, linear, autonomous ordinary differential equations, (A.10). It is useful to look for a solution of the form

$$\boldsymbol{u} = \boldsymbol{u}_0 e^{\lambda t}, \tag{A.12}$$

for some constant, λ. Substituting this into equation (A.10) we obtain

$$\lambda \boldsymbol{u}_0 e^{\lambda t} = \boldsymbol{J}\boldsymbol{u}_0 e^{\lambda t} \quad i.e. \quad (\boldsymbol{J} - \lambda \boldsymbol{I})\,\boldsymbol{u}_0 = 0. \tag{A.13}$$

For a non-zero solution, we must have $\boldsymbol{u}_0 \neq (0,0)$ and hence we require

$$\det(\boldsymbol{J} - \lambda \boldsymbol{I}) = 0, \tag{A.14}$$

where $\boldsymbol{I}$ is the 2×2 identity matrix.

This quadratic equation has two roots for λ, denoted λ_1, λ_2, which are possibly equal and possibly complex; these are, of course, the eigenvalues of $\boldsymbol{J}$ evaluated at the point (x_0, y_0).

A.4.1 Case I

λ_1, λ_2 real, with $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, $\lambda_1 \neq \lambda_2$. *Without loss of generality* we take $\lambda_2 > \lambda_1$ below.

We have two distinct, real eigenvalues. Let the corresponding eigenvectors be denoted by $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$. We thus have

$$\boldsymbol{J}\boldsymbol{e}_1 = \lambda_1 \boldsymbol{e}_1, \quad \boldsymbol{J}\boldsymbol{e}_2 = \lambda_2 \boldsymbol{e}_2. \tag{A.15}$$

We seek a solution of the form

$$\boldsymbol{u} = A_1 \boldsymbol{e}_1 + A_2 \boldsymbol{e}_2. \tag{A.16}$$

Substituting this into equation (A.10), we find, by comparing coefficients of $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, that

$$\frac{\mathrm{d}A_1}{\mathrm{d}t} = \lambda_1 A_1, \quad \frac{\mathrm{d}A_2}{\mathrm{d}t} = \lambda_2 A_2, \tag{A.17}$$

and hence

$$A_1 = A_1(t=0)e^{\lambda_1 t}, \quad A_2 = A_2(t=0)e^{\lambda_2 t}. \tag{A.18}$$

Thus we have

$$\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} \overset{def}{=} \boldsymbol{u} = A_1(t=0)e^{\lambda_1 t}\boldsymbol{e}_1 + A_2(t=0)e^{\lambda_2 t}\boldsymbol{e}_2, \tag{A.19}$$

which gives us a representation of the solution of (A.10) for general initial conditions. This information is best displayed graphically, and we do so below according to the values of λ_2, λ_1.

Note. The equilibrium point *i.e.* $(\bar{x}, \bar{y}) = (0,0)$ can only be reached either as $t \to \infty$ or $t \to -\infty$.

1. $\lambda_1 < \lambda_2 < 0$. The phase plot of the linearised equations in the $(\bar{x},\ \bar{y})$ plane looks like one of the two possibilities in Figure A.1.

Definition. An equilibrium point which results in this case is called a *stable node*, with the word "stable" referring to the fact that integral paths *enter* the node, *i.e.* the equilibrium point at $(0,0)$.

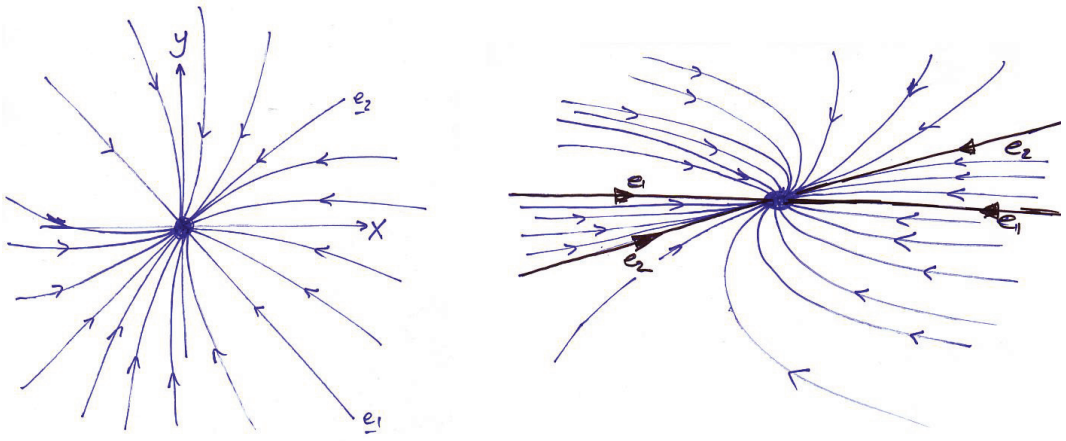


Figure A.1: Possible phase portraits of a stable node. The equilibrium point in each case is denoted by the large dot.

2. $\lambda_2 > \lambda_1 > 0$. We still have

$$\boldsymbol{u} = A_1(t=0)e^{\lambda_1 t}\boldsymbol{e}_1 + A_2(t=0)e^{\lambda_2 t}\boldsymbol{e}_2. \tag{A.20}$$

However, the direction of the arrows is reversed as the signs of λ_1, λ_2 are changed. The phase plane portraits are the same as in Figure A.1 except the direction of the arrows is reversed.

Definition. An equilibrium point which results in this case is called an *unstable node*, with the word "unstable" referring to the fact that integral paths *leave* the node, *i.e.* the equilibrium point at $(0, 0)$.

3. $\lambda_2 > 0 > \lambda_1$. Once more, we still have

$$\boldsymbol{u} = A_1(t=0)e^{\lambda_1 t}\boldsymbol{e}_1 + A_2(t=0)e^{\lambda_2 t}\boldsymbol{e}_2, \tag{A.21}$$

but again the phase plane portrait is slightly different—see Figure A.2.

Definition. An equilibrium point which results in this case is called a *saddle point*.

Definition. The two integral paths originating from the saddle point are sometimes referred to as the *unstable manifolds* of the saddle point. Conversely, the integral paths tending to the saddle point are sometimes referred to as the *stable manifolds* of the saddle point. This forms part of a nomenclature system commonly used in more advanced dynamical systems theory; see P. Glendinning, Stability, Instability and Chaos [**?**].

A.4.2 Case II

λ_2, λ_1 real. One, or more, of the following also holds:

$$\lambda_2 = \lambda_1, \qquad \lambda_1 = 0, \qquad \lambda_2 = 0. \tag{A.22}$$

We typically will not encounter these degenerate cases in this course. We briefly note that behaviour of the full equations, (A.1), can be highly nontrivial when the linearisation reduces to these degenerate cases. Further details of such cases can be found in P. Glendinning, Stability, Instability and Chaos [**?**], which is on the reading list for this course. When $\lambda_1, \lambda_2 = 0$, Hartmann's theorem doesn't hold.

A.4.3 Case III

λ_2, λ_1 complex. The complex eigenvalues of a real matrix always occur in complex conjugate pairs. Thus we take, without loss of generality,

$$\lambda_1 = a - ib = \lambda_2^*, \qquad \lambda_2 = a + ib = \lambda_1^*, \tag{A.23}$$

where a, b real, $b \neq 0$, and * denotes the complex conjugate.

We also have two associated complex eigenvectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, satisfying

$$\boldsymbol{J}\boldsymbol{e}_1 = \lambda_1 \boldsymbol{e}_1, \quad \boldsymbol{J}\boldsymbol{e}_2 = \lambda_2 \boldsymbol{e}_2, \tag{A.24}$$

which are complex conjugates of each other, *i.e.* $\boldsymbol{e}_1 = \boldsymbol{e}_2^*$.

Using the same idea as in Case I above, we have

$$\boldsymbol{u} = A_1(t=0)e^{\lambda_1 t}\boldsymbol{e}_1 + A_2(t=0)e^{\lambda_2 t}\boldsymbol{e}_2, \tag{A.25}$$

though now, in general, $A_1(t=0)$, λ_1, $\boldsymbol{e}_1$, $A_2(t=0)$, λ_2, $\boldsymbol{e}_2$ are complex, and hence so is $\boldsymbol{u}$.

Restricting $\boldsymbol{u}$ to be real gives

$$\boldsymbol{u} = A_1(t=0)e^{\lambda_1 t}\boldsymbol{e}_1 + A_1^*(t=0)e^{\lambda_2 t}\boldsymbol{e}_2 = A_1(t=0)e^{\lambda_1 t}\boldsymbol{e}_1 + (A_1(t=0)e^{\lambda_1 t}\boldsymbol{e}_1)^*, \tag{A.26}$$

and this is real, as for any complex number z, we have $z + z^* \in \mathbb{R}$.

After some algebra this reduces to

$$\boldsymbol{u} = e^{at}\left[\boldsymbol{M}\cos(bt) + \boldsymbol{K}\sin(bt)\right] = e^{at}\left[\begin{pmatrix} M_1 \\ M_2 \end{pmatrix}\cos(bt) + \begin{pmatrix} K_1 \\ K_2 \end{pmatrix}\sin(bt)\right], \tag{A.27}$$

where $\boldsymbol{M} = (M_1, M_2)^T$, $\boldsymbol{K} = (K_1, K_2)^T$ are real, constant vectors, which can be expressed in terms of $A_1(t=0)$, $A_2(t=0)$ and the components of the eigenvectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$. Equivalently, we have

$$\bar{x} = e^{at}\left[\cos(bt)M_1 + \sin(bt)K_1\right], \quad \bar{y} = e^{at}\left[\cos(bt)M_2 + \sin(bt)K_2\right], \tag{A.28}$$

where M_1, M_2, K_1, K_2 are real constants.

1. $a > 0$. We have $\bar{x}$, $\bar{y}$ are, overall, increasing exponentially but are oscillating too. For, example, with $K_1 = 0$, $M_1 = 1$ we have $\bar{x} = e^{at}\cos(bt)$, which looks like:

   Note that the overall growth of $\bar{x}$ is exponential at rate a. Thus, in general, the phase plane portrait looks like one of the examples shows in Figure A.3.

   Note. The sense of the rotation, clockwise or anti-clockwise, is easily determined by calculating $\mathrm{d}\bar{y}/\mathrm{d}t$ when $\bar{y} = 0$ or $\mathrm{d}\bar{x}/\mathrm{d}t$ when $\bar{x} = 0$.

   Definition. An equilibrium point which results in the above, is called an *unstable spiral* or, equivalently, an *unstable focus*. The word "unstable" refers to the fact that integral paths *leave* the equilibrium point.

2. $a < 0$. This is the same as 1. except now the phase plane portrait arrows point towards the equilibrium point as $\bar{x}$ and $\bar{y}$ are exponentially decaying as time increases rather than exponentially growing.

   Definition. An equilibrium point which results in this case is called a *stable spiral* or, equivalently, a *stable focus*. The word "stable" refers to the fact that integral paths *enter* the equilibrium point.

3. $a = 0$. Thus we have $\lambda_2 = -ib = -\lambda_1$, $b \neq 0$, b real, and

$$\bar{x} = [\cos(bt)M_1 + \sin(bt)K_1], \quad \bar{y} = [\cos(bt)M_2 + \sin(bt)K_2], \tag{A.29}$$

   where M_1, M_2, K_1, K_2 are constants. Note that

$$K_2\bar{x} - K_1\bar{y} = L\cos(bt), \quad -M_2\bar{x} + M_1\bar{y} = L\sin(bt), \tag{A.30}$$

   where $L = K_2M_1 - K_1M_2$. Letting $x^* = K_2\bar{x} - K_1\bar{y}$ and $y^* = -M_2\bar{x} + M_1\bar{y}$, we have

$$(x^*)^2 + (y^*)^2 = L^2, \tag{A.31}$$

   i.e. a circle in the (x^*, y^*) plane, enclosing the origin, which is equivalent to, in general, a closed ellipse, in the $(\bar{x}, \bar{y})$ plane enclosing the origin.

   Note. As with 3. above, the sense of the rotation, clockwise or anti-clockwise, is easily determined by calculating $\mathrm{d}\bar{y}/\mathrm{d}t$ when $\bar{y} = 0$ or $\mathrm{d}\bar{x}/\mathrm{d}t$ when $\bar{x} = 0$.

   Definition. An equilibrium point which results in this case, is called a *centre*. A centre is an example of a limit cycle.

Definition. A *limit cycle* is an integral path which is closed (and which does not have any equilibrium points).

A.5 Linear stability

Definition. An equilibrium point is *linearly stable* if the real parts of both eigenvalues λ_1, λ_2 are negative.

From the expressions for $\boldsymbol{u}$ above, for example

$$\boldsymbol{u} = A_1(t=0)e^{\lambda_1 t}\boldsymbol{e}_1 + A_2(t=0)e^{\lambda_2 t}\boldsymbol{e}_2, \tag{A.32}$$

when λ_1, λ_2 real, we see that any perturbation away from the equilibrium decays back to the equilibrium point.

Definition. An equilibrium point is *linearly unstable* if the real parts of at least one of the eigenvalues λ_1, λ_2 is positive (and the other is non-zero).

Other situations are in general governed by the non-linear behaviour of the full equations and we do not need to consider them here.

A.5.1 Technical point

The behaviour of the linearised equations and the behaviour of the non-linear equations sufficiently close to the equilibrium point are guaranteed to be the same for any of the equilibrium points [I 1-3], [III 1,2] or [II] with $\lambda_1 = \lambda_2 \neq 0$. All these equilibrium points are such that $Re(\lambda_1)\ Re(\lambda_2) \neq 0$. This is the essence of Hartmann's theorem. This guarantee does not hold for centres, or the equilibrium points described in [II] with $\lambda_1\lambda_2 = 0$.

The underlying reasons for this are as follows.

- First, note from the above that integral paths which meet the equilibrium point can either grow/decay at exponential rate $Re(\lambda_1)$, or exponential rate $Re(\lambda_2)$, or consist of the sum of two such terms. Second, note that in the above we took a Taylor expansion. Including higher order terms in this Taylor expansion can lead to a small correction for the rate of exponential decay towards or growth away from the stationary point exhibited by the integral paths. These corrections tend to zero as one approaches the equilibrium point.

- Consider the centre equilibrium point, which has $Re(\lambda_1) = Re(\lambda_2) = 0$, and an exponential growth/decay of zero. If the corrections arising from the Taylor series are always positive, the exponential growth/decay rate of all integral paths sufficiently near the stationary point is always (slightly) positive. Hence these integral paths grow exponentially away from the stationary point. However, b is non-zero, so $\bar{x}$ and $\bar{y}$ are still oscillating. Hence one has the non-linear equations behave like a stable focus.

- If $Re(\lambda_1), Re(\lambda_2) \neq 0$, then for all integral paths reaching the stationary point, the above mentioned corrections, sufficiently close to the equilibrium point, are negligible, *e.g.* they cannot change exponential growth into exponential decay or *vice-versa*.

This allows one to show that stationary points with $Re(\lambda_1)\ Re(\lambda_2) \neq 0$ are guaranteed to have the same behaviour for the linearised and the non-linear equations sufficiently close to the equilibrium point.

A.6 Summary

We will typically only encounter stationary points [I 1-3], [III 1-3]. Of these stationary points, all but centres exhibit the same behaviour for the linearised and the non-linear equations sufficiently close to the equilibrium point as plotted above and in D. W. Jordan and P. Smith, Mathematical Techniques [**?**].

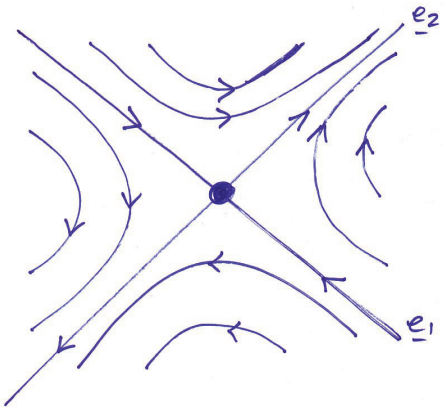
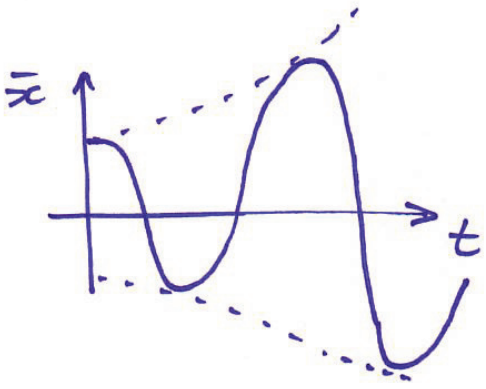


Figure A.2: The phase portrait of a saddle point. The equilibrium point is denoted by the large dot.



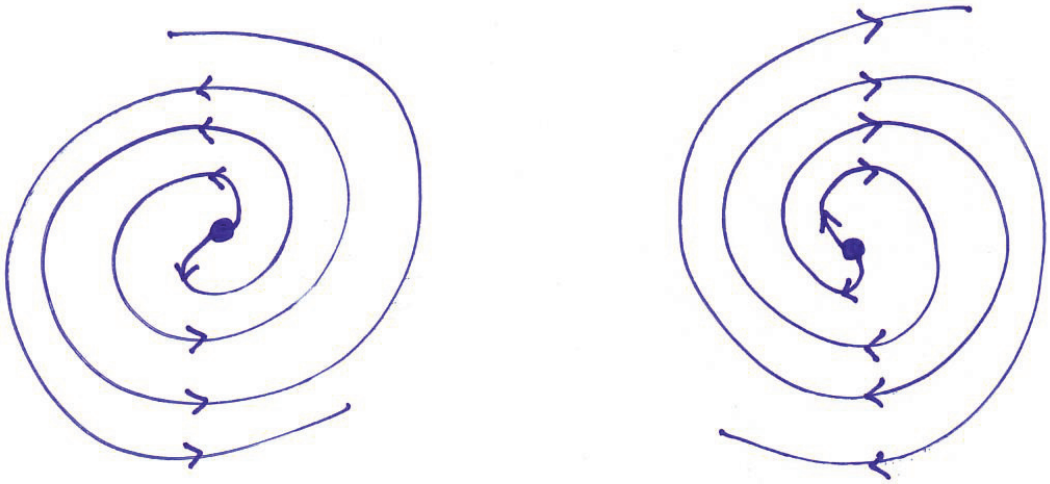


Figure A.3: Possible phase portraits of a focus. The equilibrium point in each case is denoted by the large dot.

Bibliography

[1] N.F. Britton (2005). *Essential Mathematical Biology.* Springer Undergraduate Mathematics Series. Springer.

[2] L. Edelstein-Keshet (2005). *Mathematical Models in Biology.* SIAM Classics in Applied Mathematics.

[3] A. Gierer and H. Meinhardt (1972). A theory of biological pattern formation. *Kybernetik* 12: 30-39.

[4] P. Glendinning (1999). *Stability, Instability and Chaos: An Introduction to the Theory of Nonlinear Differential Equations.* Cambridge Texts in Applied Mathematics.

[5] T. Hillen and K. Painter (2009). A user's guide to PDE models for chemotaxis. *J Math Biol* 58 (1): 183-217.

[6] D.W. Jordan and P. Smith (2002). *Mathematical Techniques: An Introduction for the Engineering, Physical and Mathematical Sciences.* Oxford University Press, 3rd edition.

[7] J.P. Keener and J. Sneyd (1998). *Mathematical Physiology,* volume 8 of *Interdisciplinary Applied Mathematics.* Springer, New York, 1st edition.

[8] J.D. Murray (2003). *Mathematical Biology I: An Introduction,* volume I. Springer-Verlag, 3rd edition.

[9] J.D. Murray (2003). *Mathematical Biology II: Spatial Models and Biochemical Applications,* volume II. Springer-Verlag, 3rd edition.

[10] A. Okubo, P.K. Maini, M.H. Williamson and J.D. Murray (1989). On the spread of the grey squirrel in Great Britain. *Proc Roy Soc Lond B* 238 (1291): 113-125.

[11] L.E. Reichl (2009). *A modern course in statistical physics.* Wiley-VCH, 3rd edition.

[12] A.M. Turing (1952). The chemical basis of morphogenesis. *Roy Soc Lond Phil Trans B* 237: 37-72.

[13] L. Wolpert (2001). Positional information and the spatial pattern of cellular differentiation. *J Theor Biol* 1: 1-47.