Mathematical Biology

University of Cambridge Part II Mathematical Tripos

David Tong

Department of Applied Mathematics and Theoretical Physics,
Centre for Mathematical Sciences,
Wilberforce Road,
Cambridge, CB3 OBA, UK

http://www.damtp.cam.ac.uk/user/tong/mathbio.html
d.tong@damtp.cam.ac.uk

Recommended Books and Resources

Here are a bunch of books that I've found useful while preparing for this course:

- J.D. Murray, *Mathematical Biology, Volumes 1 and 2*

This is the default reference for practitioners in the field. It's a remarkably easy read and explains things with clarity without holding your hand. The material that we need for the first part of the course can be found in volume 1, but you'll need to open volume 2 when we get to spatially organised systems.

- Dick Neal, *Introduction to Population Biology*

The mathematics is straightforward and, at times, might come across a little laboured. But the explanations of the underlying biology are well written and the book is peppered with many interesting examples.

- Steven Strogatz, *Nonlinear Dynamics and Chaos*

Not a biology book per se, but a spectacularly good introduction to the dynamical systems that underlie much of this course. It has a number of biologically leaning examples.

In addition, there are many online lecture notes, including ones by past lecturers of this course that I have freely taken from. You can find links to these on the course webpage.

Contents

"I write about biology from the point of view of a physicist. Some physicists are arrogant and some are humble. I prefer to be humble. Arrogant physicists say that biology needs better concepts; since physicists are good at concepts, our job is to tell biologists how to think. Humble physicists say that biology needs better hardware; since physicists are good at hardware, our job is to invent new tools for biologists to use. With the exception of Max Delbruck and Francis Crick and a few other pioneers in the heroic age of molecular biology, physicists who tried to teach biologists how to think have failed dismally."

Freeman Dyson (being very Freeman Dyson).

Acknowledgements

These lecture notes are far from original. They closely follow the notes of Julia Gog, Ray Goldstein, and Peter Haynes, all of whom gave previous versions of this course. Huge thanks to Julia, Ray, and Peter for putting so much preparation into this course and for helping me with my dumb questions. I'd also like to thank Andrew Gonzalez at McGill University, who knows a thing or two about ecology and was extremely generous in sharing it with me.

0 Introduction

> "This process of "model building", essentially that of discarding all but the essentials and focusing on a model simple enough to do the job but not too hard to see all the way through, is possibly the least understood – and often the most dangerous – of all the functions of a theoretical physicist."
>
> *Philip Anderson*

Imagine that you possess the following superpower. When presented with a new object, you pick up a pen and write down a description of it. Maybe you just jot down a few words. Maybe the muse takes hold and you become expansive, filling many paragraphs.

Now comes the superpower. When your description is finished, you stare at the page and have a little think before rearranging the words in a different order, following arcane rules known only to those who possess the superpower. You shuffle an adjective here and a verb there until a feeling of calm comes over you and the words form a pattern that, incredibly, when read, reveals something new about the original object, something that you didn't know before, something that must have been hiding there all along yet only emerges after you play this game. For those imbued with real skill, this superpower can be used to reveal things that no one on the planet previously knew.

This superpower sounds like magic. And yet it's a power that each of you can wield and the purpose of these lectures is merely to hone it. The trick, of course, is that you must describe the original object in the language of mathematics. The existence of such a superpower is the reason why mathematics is special. It is why mathematics is a greater form of expression than, say, poetry: partly because it's more useful, but mostly because it's more magical.

There are many stories about the world written in the language of mathematics. One of the mysteries of this enterprise is that, at the most fundamental level – at the level of quantum field theory and general relativity – the laws of physics are fully described in terms of the most simple equations. But that's not our interest here. Instead, we want to turn to topics in biology: population dynamics and the spread of disease and the interactions of enzymes and many others besides. These topics, like everything in biology, tend to be complicated. And this gives us an immediate headache because, for the kind of systems that we're interested in, no equation with fewer than, say, 10^{23} terms is likely to capture the full complexity of the situation. What to do?

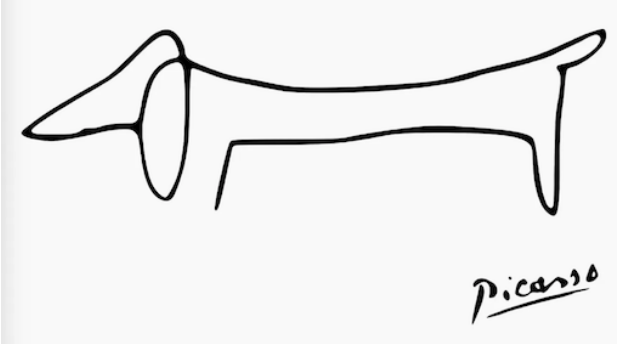


Figure 1. A good model.

One option is to aim for realism. We could continue to add more and more terms to our equations, hoping to match the intricate details of the system, perhaps inspired by results elsewhere in physics where theoretical calculations agree with experiment to many decimal places. This is not the approach we will take here.

Instead, we will strive to create mathematical caricatures of biological systems. We will strip away the complexity and focus on the key principles that lie underneath. The result will be equations that are akin to Picasso's line drawings of animals: they capture the spirit of the beast, but they may not be particularly useful if you're a vet learning how to do surgery. The idea is that, with a few strokes of the pen, we can construct equations that describe the essence of a thing and then solve these equations to learn more about what it does and how it does it.

Throughout these lectures we will solve many such equations. However, one of the things that we will *not* do is to explain how to construct such models in the first place. This is an important skill, but it's typically a skill that first requires a deep knowledge of a particular topic – say, the migratory patterns of monarch butterflies, or the human respiratory system – before you then try to extract from the wealth of information the important features that can be profitably distilled into equations. We won't go through this long and complicated process in what follows, focussing instead on the end product: some simple examples of mathematical models. We will look at many such examples. The hope is that by exploring a wide range of mathematical models, and learning what sorts of behaviour you might expect to emerge, you will be better placed to construct your own models when you finally learn all there is to know about those monarch butterflies.

1 Population Dynamics and Other Stories

People are born. People die. People move on. This, in a nutshell, is population dynamics. Our goal in this section is to drape equations around these words to understand how the population of various species changes over time.

The general approach that we will take will be mirrored throughout these lectures: first we try to isolate the relevant dynamical degree of freedom; then we introduce models of increasing complexity to capture some basic idea about life or death or movement. Each of the complications will usually (but not always) change the form of our equations in some way and a large part of this chapter will be devoted to describing the techniques needed to solve the resulting equations.

The basic question that we have to address when building any mathematical model of the world is: what are the dynamical degrees of freedom? In other words, what are the right variables to use? For our immediate purposes, the answer would seem to be blindingly obvious: we want to understand the size of the population and how it changes with time. We denote this as $N(t)$.

Even here, however, there are subtleties and we may well need to further refine our dynamical variable or add more. Does it matter if we treat $N(t)$ as a continuous variable or should we insist that it's something discrete? Does the age profile of the population matter, in which case we should work with $n(a,t)$, the number of people with a particular age a. Does the way the population is distributed spatially matter, in which case we should work with a population density $n(\mathbf{x},t)$. Does it matter if the population is interacting with some other species? Are there other things that we've just completely, perhaps even unintentionally, ignored?

The answer to all of these question is: it depends. And even if you devote your life to the study of some particular population, it still may not be obvious. The art of mathematical modelling is largely in realising what you should include and what you can safely ignore. The way to proceed is to start simple and then introduce each of the potential complications above to see what qualitative and quantitative effect they have on the dynamics. Indeed, it's often only by comparing the data to the mathematical models that we can start to understand which complexities are important and which can be discarded.

As this section progresses, we will be increasingly flexible about what we think of as a "population". Later, we will look at "populations" of viruses, chemicals and electrical signals in the brain. The thing that links these different phenomena is that

they can all be described by coupled, first order differential equations. Indeed, the real purpose of this whole section is to see a few things that can be modelled by these simple dynamical systems. Unlike in other courses, we will not develop the mathematics systematically but, instead, work our way through various examples, most of which will exhibit different mathematical features. As we proceed, we will build up a toolkit of results that allow us to examine these kinds of systems more generally.

1.1 First You're Born, Then You Die

We're going to start by writing down the simplest models for population growth. These describe just the single variable $N(t)$ and we will take both the population N and time t to be continuous. This seems reasonable in the case of time, less so for the population which is, in reality, an integer. But we can justify this approximation if we're dealing with situation in which N is suitably large so that the difference between N and $N+1$ doesn't make any material difference to the situation.

In what follows, we will ignore both immigration and emigration. This means that population change is determined by birth and death rates alone.

1.1.1 Exponential Growth

For our first attempt at writing down a model, we will assume that the population has some fixed birth and death rates. Each person is the population has some probability of giving birth and some probability of dying. We're not distinguishing population by age and so this probability is the same for everyone.

Suppose that in some time Δt, the number of births is $bN\Delta t$ and the number of deaths is $dN\Delta t$ with b and d constant. We call b and d the birth and death rates, respectively. More precisely, they are the number of births/deaths per individual (or *per capita*) per unit time.

In the time Δt, the change in the population will be

$$N(t+\Delta t) \approx N(t) + (b-d)N(t)\Delta t \ . \tag{1.1}$$

Dividing by Δt, and taking the limit $\Delta t \to 0$, we get the first order differential equation

$$\frac{dN}{dt} = (b-d)N \ . \tag{1.2}$$

This equation only depends on the difference of b and d, not their individual values. (As we will see in Section 4, this conclusion will change when we look at random fluctuations.) The equation is easily solved and we have

$$N = N_0\, e^{rt} \quad \text{with} \quad r = b - d \ . \tag{1.3}$$

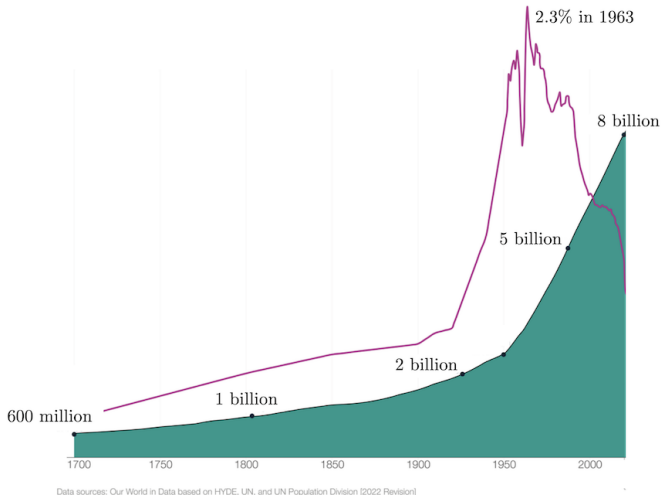


Figure 2. The growth of the world's population, taken from the Our World in Data website. The purple line shows the percentage increase each year, so a constant purple line would correspond to exponential growth.

Here N_0 is the initial population at time $t = 0$. We learn that populations will grow exponentially if the birth rate is higher than the death rate, and shrink exponentially if it's the other way around.

The idea of exponential population growth is often attributed to an essay by Thomas Malthus in 1798. He writes

> "*This* [increase of population] *constantly tends to subject the lower classes of the society to distress and to prevent any great permanent amelioration of their condition*"

He was not the last person to deduce wide-sweeping, moralising conclusions about the state of society based on flimsy mathematical underpinnings.

So does this exponential growth hold up in practice? Well, like many things in this course, the answer is: to an extent. But often you have to seek out specific examples where it works. An obvious place to look is the global population of humans. The data is shown in Figure 2. For much of the past 300 years, the population growth has been super-exponential, with the exponent r growing over time. (This exponent is roughly what is plotted on the purple line.) But, since 1963, the exponent r has been decreasing. Current projections suggest that r will become negative before the year 2100, with the world's population peaking at around 10 billion.

1.1.2 The Logistic Equation

In an attempt to be more realistic, we could envisage that the birth and death rates themselves depend on the size of the population. That leads us to the general class of models

$$\frac{dN}{dt} = [b(N) - d(N)]N \tag{1.4}$$

with general functions $b(N)$ and $d(N)$ that encode whatever features of population growth you think are important.

As we're treading slowly, an obvious guess for the right-hand side is to replace the linear function that led to exponential growth with a quadratic function. This suggests that we examine solutions to the equation

$$\frac{dN}{dt} = rN - \lambda N^2 \ . \tag{1.5}$$

This is a famous differential equation, known as the *logistic equation*. It was first written down in 1838 by the Belgian mathematician Pierre Francois Velhurst. The idea of the extra term is that growth rates become smaller as resources become scarce, perhaps due to lack of food or to overcrowding. The logistic equation comes with a natural maximum population that can be supported. This is most easily seen if we rewrite the equation as

$$\frac{dN}{dt} = r\left(1 - \frac{N}{K}\right)N \quad \text{with} \quad K = \frac{r}{\lambda} \ . \tag{1.6}$$

Here K is referred to as the *carrying capacity*.

If $r, \lambda > 0$, then small populations with $N < K$ will have an increasing population with $dN/dt > 0$. Meanwhile, larger populations with $N > K$ will have a decreasing population with $dN/dt < 0$. In this way, the logistic equation captures the idea that there is a natural ceiling to the size of a population.

It's straightforward to solve the logistic equation. We have

$$\begin{aligned} \int dN \ \frac{K}{(K-N)N} &= \int dN \left(\frac{1}{N} + \frac{1}{K-N}\right) \\ &= \log\left|\frac{N}{K-N}\right| - \log\left|\frac{N_0}{K-N_0}\right| = rt \ . \end{aligned} \tag{1.7}$$

Here we've introduced the integration constant N_0 which is designed so that $N(t = 0) = N_0$. We've taken the modulus sign in the logarithms because we don't know if

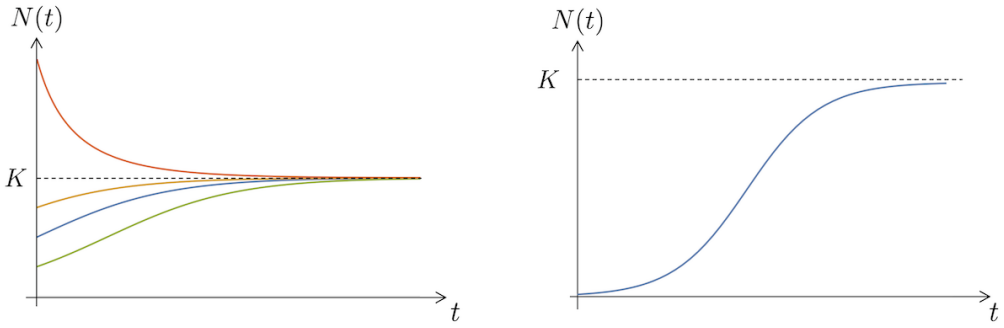


Figure 3. Solutions to the logistic equation. On the right, the curves are plotted with (from top to bottom) $N_0 = 2K$, $N_0 = \frac{3}{4}K$, $N_0 = \frac{1}{2}K$ and $N_0 = \frac{1}{2}K$. On the right, the curve is plotted with $N = \frac{1}{100}K$ where it exhibits the kind of characteristic tanh-like squashed s-shape that is sometimes called a *sigmoidal*.

$N > K$ or $N < K$. But if $N(t) > K$ at one time, then it remains so at all times (and similar for $N < K$.) This means that once we combine the two logs we can remove the modulus signs and write

$$\log\left(\frac{N}{K-N}\frac{K-N_0}{N_0}\right) = rt \ . \tag{1.8}$$

Rearranging, we have

$$N = \frac{N_0 K e^{rt}}{K + N_0(e^{rt}-1)} \ . \tag{1.9}$$

This general form $f(x) = a/(b + e^{-x})$ is known as the *logistic function*. The function is plotted in Figure 3 for various values of N_0. Regardless of the initial value of the population N_0, the curves converge on $N \to K$ at late times. The logistic function is well approximated by our earlier exponential function for times such that $N_0(e^{rt}-1) \ll K$

There are other ways to write the solutions to the logistic equation that highlight the difference between the two cases $N > K$ and $N < K$. To see this, it's perhaps simplest to return to the integral solution (1.7) and write it as

$$\log\left|\frac{N}{K-N}\right| = r(t-t_0) \ . \tag{1.10}$$

Now we have the integration constant t_0 and we retain the modulus signs. Rearranging with $N < K$, we find

$$N = \frac{Ke^{r(t-t_0)}}{1+e^{r(t-t_0)}} = \frac{K}{2}\left[1 + \tanh\left(\frac{1}{2}r(t-t_0)\right)\right] \ . \tag{1.11}$$

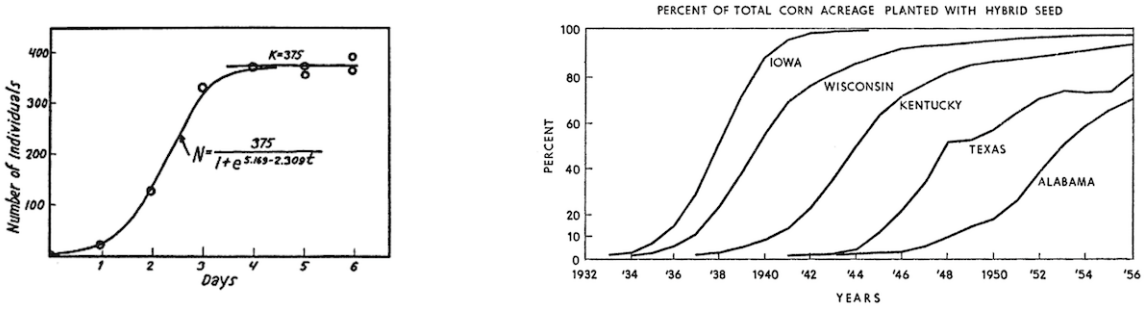


Figure 4. On the left: the population of a single celled organism over time. This picture is taken from the classic, and wonderfully named, 1934 book *A Struggle for Existence* by G.F. Gause. (In addition, here is a more recent paper showing how to fit Gause's data statistically.) On the right: the percentage of the total corn planted with hybrid seed in different US states, between 1932 and 1956, taken from the 1957 paper Hybrid Corn: An Exploration in the Economics of Technological Change by Griliches.

Meanwhile, for $N > K$, the same manipulations give

$$N = \frac{Ke^{r(t-t_0)}}{e^{r(t-t_0)} - 1} = \frac{K}{2}\left[1 + \coth\left(\frac{1}{2}r(t-t_0)\right)\right] \quad . \tag{1.12}$$

Again, we can ask: how well does the logistic curve do in modelling real world populations? It by no means a universal curve, but it certainly is more ubiquitous than exponential growth and you can find the characteristic s-shape curve appearing in many different places. One example is shown on the left of Figure 4 where the time evolution of the population of single celled organisms known as "paramecium caudatum" is plotted. (Unsurprisingly, when comparing data to equations, things tend to look better when the population in question lives in a bottle in controlled conditions, rather than in the wild.)

The Spread of Beneficial Mutation

There are many other situations where data seems to fit the logistic equation[1], notably when there is a natural ceiling in place. In particular, the equation arises when describing the fraction of a population that has adapted in some advantageous way. One

[1]You can find claims that the growth of children's vocabulary can be well fitted by the logistic function. You can see the data here and make up your own mind. It's worth mentioning that the logistic function (1.9) coincides with the Fermi-Dirac distribution in theoretical physics, although I don't know of a way to view that distribution as a solution to the logistic equation. (You can read more about the way this distribution arises in the lectures on Statistical Physics.)

famous example is the percentage of farmers who adopted a particular breed of hybrid corn, as shown on the right of Figure 4.

For example, the fraction $p(t)$ of a population which carry a gene with a beneficial mutation is described by the logistic equation. To derive this, let $N(t)$ be the total population. A mutation is beneficial if it increases the rate at which individuals have offspring. Suppose that, in the absence of the mutation, the population grows at a rate r but that, with the mutation present, this is increased to $r + s$. At time t,

$$N_{\rm mutant}(t) = p(t)N(t) \quad \text{and} \quad N_{\rm normal}(t) = (1 - p(t))N(t) \ . \tag{1.13}$$

At time $t + \delta t$, these numbers are

$$\begin{aligned} N_{\rm mutant}(t + \delta t) &= \big(1 + (r + s)\delta t\big)\, pN \\ N_{\rm normal}(t + \delta t) &= (1 + r\delta t)\,(1 - p)N \end{aligned} \tag{1.14}$$

where both p and N on the right-hand side are evaluated at time t and, here and below, we're dropping terms of order $\mathcal{O}(\delta t^2)$. This means that the total population at time $t + \delta t$ is

$$\begin{aligned} N(t + \delta t) &= N_{\rm mutant}(t + \delta t) + N_{\rm normal}(t + \delta t) \\ &= (1 + r\,\delta t + sp(t)\,\delta t)N(t) \ . \end{aligned} \tag{1.15}$$

Meanwhile, the fraction of mutants at time $t + \delta t$ is

$$\begin{aligned} p(t + \delta t) &= \frac{N_{\rm mutant}(t + \delta t)}{N(t + \delta t)} \\ &= \frac{(1 + (r + s)\delta t)}{1 + r\,\delta t + sp(t)\,\delta t} p(t) \ . \end{aligned} \tag{1.16}$$

We now Taylor expand the denominator, again throwing away terms of order $\mathcal{O}(\delta t^2)$, to get

$$\begin{aligned} p(t + \delta t) &= (1 + (r + s)\delta t)\,(1 - r\,\delta t - sp(t)\,\delta t)\,p(t) \\ &= (1 + s(1 - p(t))\delta t)\,p(t) \ . \end{aligned} \tag{1.17}$$

We then have

$$\frac{dp}{dt} = \lim_{\delta t \to 0} \frac{p(t + \delta t) - p(t)}{\delta t} = sp(t)(1 - p(t)) \tag{1.18}$$

which is the promised logistic equation.

Nondimensionalisation

Applied mathematicians have invented one of the ugliest words in the English language and then they go around saying it like it's completely normal. That word is *nondimensionalisation*.

The idea is to work with dimensionless variables. To achieve this, we absorb some of the constants in the equations into the dynamical variables. For example, for the logistic equation (1.6), we define

$$x = \frac{N}{K} \quad \text{and} \quad \tau = rt \tag{1.19}$$

so that the equation becomes

$$\frac{dx}{d\tau} = x(1-x) \ . \tag{1.20}$$

Correspondingly, the solution (1.9) is

$$x = \frac{x_0 e^\tau}{1 + x_0(e^\tau - 1)} \ . \tag{1.21}$$

Nondimensionalisation has the advantage that it allows us to see the wood for the trees, stripping away anything that is inconsequential for the analysis. Of course, it's straightforward to put the constants back in by reverting to the original variables.

You will often see the practice of rescaling certain variables to make them nondimensional, but then retaining their original names. We will also be guilty of this at times, but will flag when we're doing it.

1.1.3 Fixed Points

In general, the class of population models that we're considering take the nondimensionalised form

$$\frac{dx}{dt} = f(x) \tag{1.22}$$

for some function $f(x)$. One particularly simple question that we can ask is: what are the steady state solutions, such that $dx/dt = 0$?

Clearly, the steady state solutions are the roots of the function, meaning those values $x = x_\star$ such that $f(x_\star) = 0$. Given such a fixed point, we can further ask: is it stable? If the population deviates a little from $x = x_\star$, does it move towards $x_\star$ or away? The answer to this follows from some simple analysis. We write

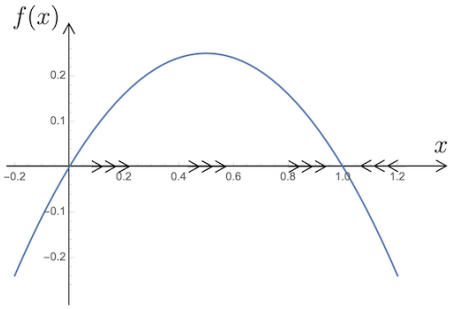
$$x(t) = x_\star + \epsilon(t) \tag{1.23}$$

with $\epsilon(t) \ll 1$. Taylor expanding, we then have

$$\frac{dx}{dt} = f'(x_\star)\epsilon(t) + \mathcal{O}(\epsilon^2) \ . \tag{1.24}$$

We see that the fixed point $x = x_\star$ is stable if $f'(x_\star) < 0$ and is unstable if $f'(x_\star) > 0$.

In practice, we can just plot the function $f(x)$ and see whether the graph crosses the axis with positive or negative slope. For example, for the logistic equation we have $f(x) = x(1-x)$ which is plotted on the right. The fixed point at $x_\star = 0$ has $f'(0) > 0$ and is unstable, while the other fixed point at $x_\star = 1$ has $f'(1) < 0$ and is stable.



An Example: Adding Predation

Here's a more intricate example. Suppose that we add an extra term to the logistic equation, so that it reads

$$\frac{dN}{dt} = r\left(1 - \frac{N}{K}\right)N - \frac{BN^2}{A^2 + N^2} \ . \tag{1.25}$$

This additional term represents the effects of predators, which increases monotonically as the population increases but reaches a maximum value as $N \to \infty$.

The are different ways to nondimensionalise this equation. We could define $x = N/K$ as before. But we will instead make the different choice

$$x = \frac{N}{A} \ , \quad \tau = \frac{Bt}{A} \quad \text{and} \quad \alpha = \frac{Ar}{B} \ , \quad \beta = \frac{K}{A} \ . \tag{1.26}$$

Here α and β are dimensionless parameters. With this rescaling, the equation becomes

$$\frac{dx}{d\tau} = f(x) = \alpha\left(1 - \frac{x}{\beta}\right)x - \frac{x^2}{1+x^2} \ . \tag{1.27}$$

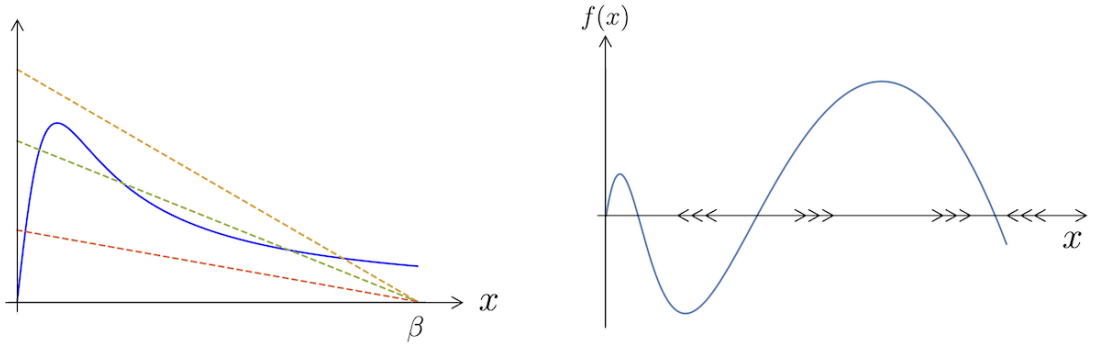


Figure 5. One the left: graphic solutions to equation (1.28). the dashed lines are plotted for a fixed value of β and varying α. On the right, the function $f(x)$ defined in (1.27) for values of the parameters where there are four roots.

There's one obvious fixed point at $x = 0$. The others are more complicated as we need to solve

$$\alpha\left(1 - \frac{x}{\beta}\right) = \frac{x^2}{1+x^2} \ . \tag{1.28}$$

Multiplying this out gives us an unwieldy cubic. It has either one root or three depending on the parameters. To see this, we can look for solutions graphically. On the left of Figure 5, the function $x^2/(1+x^2)$ is plotted as a solid line, while the linear function $\alpha(1-x/\beta)$ is plotted for a fixed value of β and varying α. We see that for α large and small, there is just a single intersection point, while there is a window of values of α where there are three intersections.

If there is just a single additional root of $f(x)$, other than $x = 0$, then it is necessarily the stable point. If there are three additional roots, then two of them are stable and one is unstable. Indeed, a plot of $f(x)$ for a choice of α and β where there there are four roots is shown on the right of Figure 5. You can see that, for any function $f(x)$, the fixed points have to alternate between unstable and stable. The solutions to (1.22) always sit between two fixed points and flow, asymptotically, to the stable one.

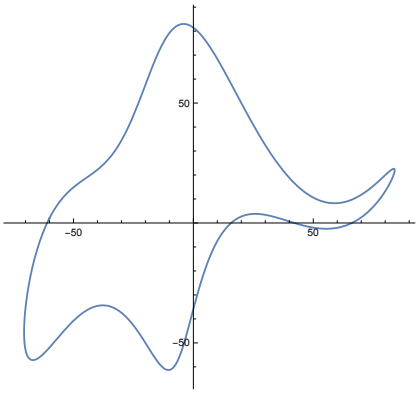
The equation (1.27), is thought to be a decent model for a class of tree-eating bugs, with the predation term arising because birds like bugs. These kind of models can inform the strategy that you take to reduce the infection by varying any parameters in the model that may be under your control. For example, you could spray the trees with something unpleasant to reduce the carrying capacity K of the bug. Or you could deliberately increase the population of predators, increasing the value of B.

1.1.4 In Praise of Parsimony

Throughout these lectures, we will typically start with some very simple model to describe a particular phenomenon. It will capture some aspect of what's going on but, when compared to data, we will usually find that things are more complicated. The obvious thing to do is to return to our original equations and improve them by, say, adding an extra term. That's what we did above when including the effect of predation.

It might be tempting to continue to add extra terms and parameters until the mathematical model agrees more closely with the data. We're going to resist this temptation. Mathematical biology doesn't, for the most part, have the precision of more fundamental areas of physics. We're never going to find the biological analog of, say, the blackbody radiation curve that we met in the Statistical Physics lectures and matches the light from the Big Bang to an accuracy of 10^{-5}. That's because humans and rabbits and bugs and viruses are all more complicated than atoms and photons. And including more and more terms in an attempt to match what you see runs the risk that you aren't adding any kind of insight beyond data fitting.

Instead our goal is more modest in scope. We would like to understand the qualitative features of a biological system, distilling its essence into a few simple equations.



For those who would prefer to see a closer match between predictions and data, it's worth recalling von Neumann's famous words to Fermi (all the more appropriate for a course on mathematical biology)

> " With four parameters I can fit an elephant, and with five I can make him wiggle his trunk."

An almost-realisation of this can be achieved by the parametric plot[2]

$$\begin{aligned} x(t) &= -60\cos t + 30\sin t - 8\sin 2t + 10\sin 3t \\ y(t) &= 50\sin t + 18\sin 2t - 12\cos 3t + 14\cos 5t \ . \end{aligned}$$

[2]This is taken from the paper "Drawing an Elephant with Four Complex Parameters" by Jürgen Mayer, Khaled Khairy, and Jonathon Howard.

The result is shown in the figure. If you count only the amplitudes as parameters then you've got 8. The authors, somewhat cheekily, paired the parameters together in a random way and declared success with 4 complex numbers. (Presumably they could have equally argued that they did it with just 2 quaternions.)

1.2 Delay

The class of models (1.22) includes an arbitrary function $f(x)$, designed to capture the intricacies of population interactions. But the kind of solutions that we get are always going to be the same: the population will evolve monotonically towards a fixed point. This means that any population that exhibits different behaviour – say oscillations – must have something else going on.

In this section, we explore the mathematics of a different class of models. These have the property that the change in the population depends on its size evaluated at a previous time. This might be argued to be more realistic. For example, it takes time for eggs to hatch. And it takes time for babies to grow into fully functioning, and reproducing, members of society.

We will look at the *delayed logistic equation*, also known as the *Hutchinson-Wright equation*. In nondimensionalised variables, it is

$$\frac{dx(t)}{dt} = \alpha\, x(t)\big[1 - x(t-T)\big] \ . \tag{1.29}$$

As promised, the change in population depends on the current population both at the current time t and at a previous time $t - T$.

As we will see, the delayed logistic equation (1.29) exhibits interesting behaviour and is used widely to model various biological phenomena. But the equation doesn't match the words that we've used to motivate it: the delay isn't in the birth term, but instead it is in the carrying capacity term. This means that it is the resources available at the earlier time that affect the growth rate[3]. For example, you might think of animals

[3]This is the insight of Hutchinson in his 1948 paper " Circular Causal Systems in Ecology". (He has a cute footnote saying that he turned to his friend, Lars Onsager, for help in solving the equation. Onsager has a formidable reputation as one of the greatest, and least comprehensible, physicists of the mid-20$^{\text{th}}$ century.) In contrast, Wright's 1955 paper "A non-linear difference-differential equation" discusses the equation where the delay is in the birth rate

$$\frac{dx(t)}{dt} = \alpha\, x(t-T)\big[1 - x(t)\big] \ .$$

We'll look at an equation in a similar spirit below when we discusses delay models applied to blowfly populations.

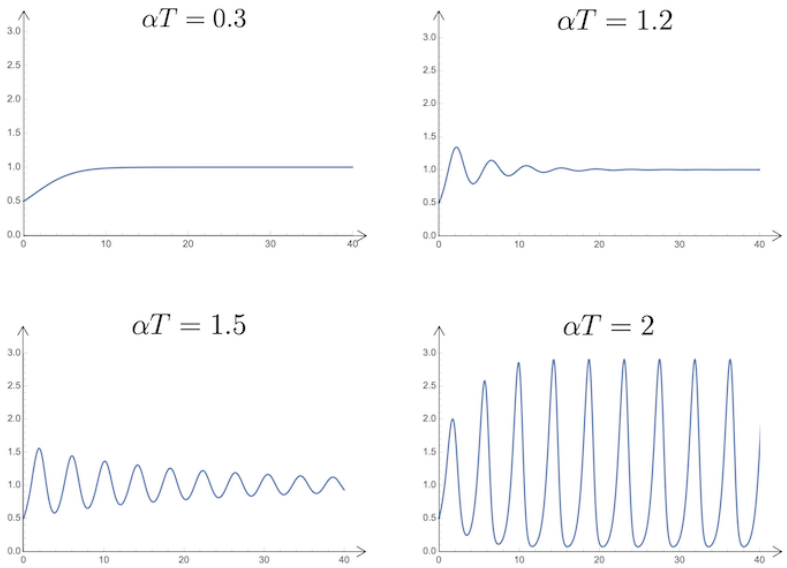


Figure 6. Numerical solutions to the delayed logistic equation, with $x(t)$ plotted vertically and t plotted horizontally. These were produced with $T = 1$ and α given by the value of αT specified in the graph. The initial history was taken to be $x(t) = \frac{1}{2}$ for $-1 \leq t \leq 0$.

storing nutrients, or hoarding resources, like squirrels and nuts, or humans and toilet paper.

The delayed logistic equation is an example of a general class of *delay differential equations*. These kind of equations are rare in theoretical physics, where things tend to happen locally in time, so are likely unfamiliar. Usually a first order differential equation needs just a single initial condition, say $x(t = 0)$. Here we need a functions worth of initial conditions, telling us the history $x(t)$ for $-T \leq t \leq 0$, before we can find a solution.

Even before we try to solve (1.29), we can get a feel for what will happen. In the original logistic equation, the population is capped above at $x = 1$ (in dimensionless variables). But the delay means that the population at time t can exceed $x = 1$ and continue to grow provided that the past population $x(t - T)$ is less than one. The population then stops growing only at a time T after it first hits the would-be ceiling at $x = 1$. Then the population will decrease. This story then repeats, resulting in oscillations around the $x = 1$ fixed point.

Numerical Results

Delay differential equations like (1.29) typically don't have closed form solutions. We will make some analytic progress shortly, but first we can study the solutions numerically. The equation has two parameters: α and T. But, after rescaling time, there is really just a single dimensionless parameter αT. We want to understand how solutions depend on this parameter.

The results are shown in Figure 6, plotted for various values of αT. We do indeed see the observed oscillations when αT is suitably large. But when αT is small, they die away. And by the time that $\alpha T \lesssim 0.5$, the oscillations are pretty much invisible. Our goal is to understand this behaviour.

1.2.1 The Linear Delay Differential Equation

The solution $x = 1$ is still a fixed point of the delayed logistic equation (1.29). We will make progress by looking at small deviations away from this fixed point. To this end, we write

$$x(t) = 1 + \epsilon(t) \ . \tag{1.30}$$

Substituting this into delayed logistic equation and dropping terms of order ϵ^2, we find

$$\frac{d\epsilon(t)}{dt} = -\alpha\, \epsilon(t - T) \ . \tag{1.31}$$

This is still a delayed differential equation, but it's one of the few that we can solve analytically. These solutions take the form

$$\epsilon(t) = \epsilon_0 e^{st} \tag{1.32}$$

for some constant s. There's a standard trick that we use when solving linear equations of this kind: we take $s \in \mathbb{C}$. At first this looks slightly odd because ϵ is a (rescaled) measure of the population and when you're driving past those road signs that state the population of a town, they very rarely give a complex number. But because (1.31) is linear, if we've got a complex solution then we can simply take the real and imaginary parts to find real solutions.

At heart, the trick of taking $s \in \mathbb{C}$ is simply so that we can write things like e^{it} rather than $\cos t$ and $\sin t$. This means that any imaginary part of s is telling us that the solution oscillates about the fixed point. Meanwhile, a real part of s is telling us that the solution converges to the fixed point (if $\mathrm{Re}(s) < 0$) or moves away from the fixed point (if $\mathrm{Re}(s) > 0$).

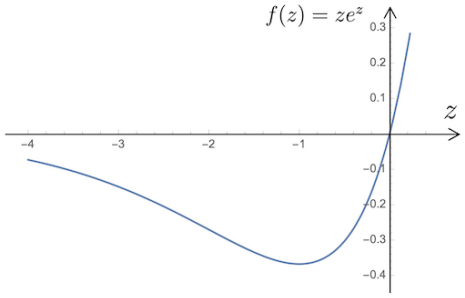
The ansatz (1.32) solves our delay differential equation (1.31) provided that

$$se^{sT} = -\alpha \ . \tag{1.33}$$

If we rescale and define the variable $z = sT$, this becomes

$$ze^z = -\alpha T \ . \tag{1.34}$$

The solutions to this equation are somewhat subtle. We will proceed slowly.



First, we can look for purely real solutions. This is straightforward. We plot the curve $f(z) = ze^z$ in the figure. The right-hand side of (1.34) is negative so we see immediately that solutions must have $z < 0$. In other words, these are stable. The function has a minimum at $z = -1$ where $f(-1) = -e^{-1}$. This means that there are only purely real solutions provided that the decay time is suitably short,

$$\alpha T \leq \frac{1}{e} \approx 0.37 \ . \tag{1.35}$$

This agrees with what we saw numerically: for low values of αT, there are no oscillations. If $\alpha T \ll 1$, then we can zoom into the origin of the graph where $f(z) = ze^z \approx z$. We see that the solution for very small αT is just $s = -\alpha$, which coincides with the solution to the logistic equation near the fixed point.

When $\alpha T > e^{-1}$, the solutions to (1.34) become complex. The general class of solutions are known as *Lambert W functions*. Once we allow complex values of z, there isn't a unique solution because of the branch cut inherent in taking the log of a complex number. Instead, the different branches of solutions are labelled by $k \in \mathbb{Z}$ and are written as

$$z = W_k(-\alpha T) \ . \tag{1.36}$$

Here $W_k(x)$ are the Lambert W functions. There are no closed form expressions, but their properties are very well studied. (In particular, there is an inbuilt Mathematica function `LambertW`$[k, x]$ that will do the job for you.)

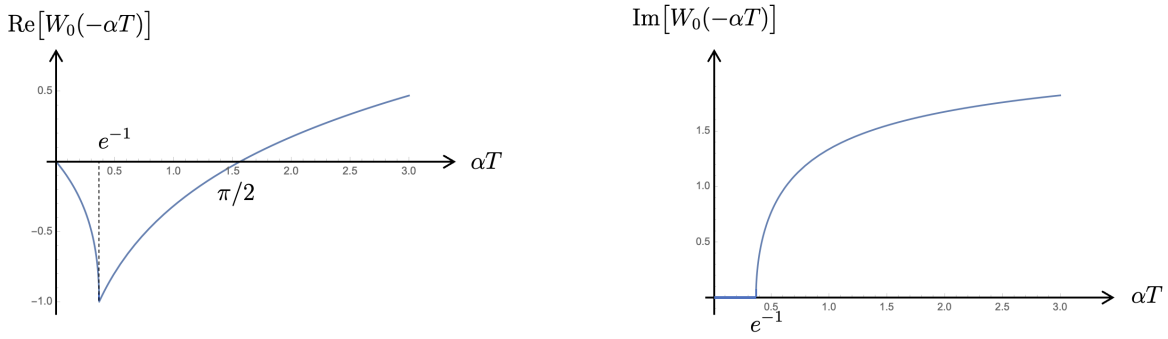


Figure 7. The real and imaginary parts of the Lambert W function $W_0(-x)$.

The purely real solutions that we have found above are part of the Lambert W function $W_0(x)$. The real and imaginary parts of this solution are plotted in Figure 7. We see that the imaginary part vanishes for $\alpha T < 1/e$, agreeing with what we found above. Meanwhile, the real part has a sharp kink at that point. For our immediate purposes, the important thing is that the real part is negative for $\alpha T < \pi/2$, and then becomes positive.

It's straightforward to reproduce these two key features with a little thought. First note that when we take $z = \pi i/2$, we have

$$z = \frac{\pi i}{2} \quad \Longrightarrow \quad z e^z = -\frac{\pi}{2} \tag{1.37}$$

So this is indeed a purely imaginary solution when $\alpha T = \pi/2$.

Next, write $z = \sigma + i\omega$. If we rewrite (1.34) as $z = -\alpha T e^{-z}$, we have

$$\sigma + i\omega = -\alpha T e^{-\sigma}(\cos\omega - i\sin\omega) \quad \Longrightarrow \quad \begin{cases} \sigma = -\alpha T e^{-\sigma}\cos\omega \\ \omega = +\alpha T e^{-\sigma}\sin\omega \end{cases} \tag{1.38}$$

where we have simply decomposed the first equation into real and imaginary parts. When $\alpha T = \pi/2$, we know that these equations are solved when we have $\sigma = 0$ and $\omega = \pi/2$. We want to show that when $\alpha T < \pi/2$, we necessarily have $\sigma < 0$.

To see this, it's simplest to consider two separate cases. First, if $|\omega| > \alpha T$ then the second equation in (1.38) tells us that $\omega = \alpha T e^{-\sigma}\sin\omega < \alpha T e^{-\sigma}$ so we must have $\sigma < 0$. Second, if $|\omega| < \alpha T < \pi/2$ then $\cos\omega > 0$ and so the first equation in (1.38) tells immediately that $\sigma < 0$.

The upshot of this argument is there are three distinct regimes, in which solutions to the delayed logistic equation have the following behaviour close to the stable fixed point at $x = 1$:

- For $0 < \alpha T < 1/e$, solutions monotonically approach the stable fixed point.
- For $1/e < \alpha T < \pi/2$, solutions oscillate about the stable, with the oscillations decaying exponentially.
- For $\alpha T > \pi/2$, the unstable fixed point is unstable. This is a more general lesson: too much delay causes instability. Our linear analysis isn't sufficient to tell us what actually happens but, from last of the numerical solution shown in Figure 6, with $\alpha T = 2$, we see that the solution oscillates about $x = 1$ without decay.

  We can, however, use our analysis to estimate the frequency of oscillations. For αT slightly greater than $\pi/2$, we have $\mathrm{Im}(z) \approx \pi/2$ so the oscillations should be close to $e^{i\pi t/2T}$, meaning that they have approximate period $2T/\pi$. That's indeed what is seen numerically.

The discussion above relates only to the first branch of solutions, with $z = W_0(-\alpha T)$. There are also other branches with $z = W_k(-\alpha T)$. For example, it's simple to check that there exists purely imaginary solutions whenever

$$z = (2k+1)\frac{\pi i}{2} \quad \Longrightarrow \quad ze^z = -(-1)^k(2k+1)\frac{\pi}{2} \ . \tag{1.39}$$

This is where $\mathrm{Re}\big[W_k(-\alpha T)\big] = 0$ and this particular branch turns from stable to unstable. Importantly, however, higher branches only become unstable at larger values of αT. This ensures that, for $0 < \alpha T < \pi/2$, the fixed point $x = 1$ is stable.

You might then wonder whether it's possible to find stable solutions for larger values of αT, which use only these higher branches with $k \neq 0$. The answer is "yes", but such solutions are finely tuned. In general, the solution to the linearised equation (1.32) will be

$$x(t) = \sum_{k\in\mathbb{Z}} A_k e^{z_k t/T} \quad \text{for} \quad -T \leq t \leq 0 \ . \tag{1.40}$$

Here $z_k = W_k(-\alpha T)$ are the exponents from the different branches of the Lambert W function. The coefficients A_k are determined by the initial conditions. Recall, in particular, that for delay equations of this type we don't have a single initial condition, but a whole history's worth $x(t)$ for $-T \leq t \leq 0$: that's why there are an infinite number of coefficients A_k. As we have seen, the $k = 0$ solution is unstable for $\alpha T > \pi/2$. If

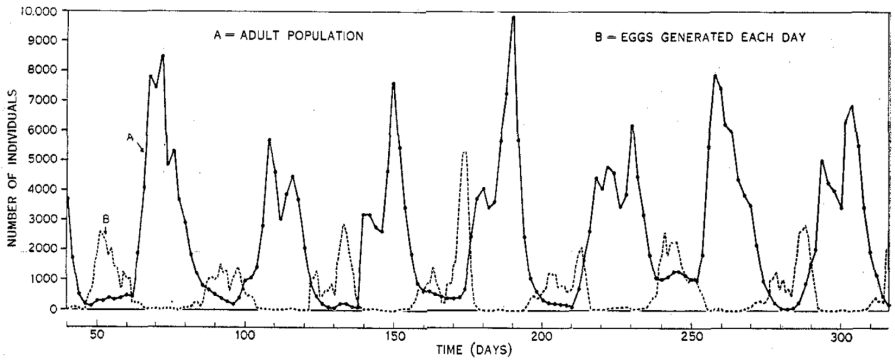


Figure 8. Data of blowfly population. The solid line is the adult population, the dotted line the eggs.

the initial conditions mean that $A_0 = 0$ then we can discard this solution and extend the regime of stability further, with the $k = 1$ solution the first to turn unstable at $\alpha T = 5\pi/2$.

1.2.2 Blowflies

If you're a sheep farmer, then you really care about blowflies (little bastards). The flies lay their eggs in the fleece, often attracted by open wounds. This doesn't do the sheep much good: mortality rates are around 2 to 3%.

In a famous experiment in the 1950s, the British-Australian zoologist Nicholson kept blowfly as pets in the lab, regulating their supply of food for both the adult and larval population. Nicholson observed wild fluctuations in the blowfly population. His results are reproduced in Figure 8. These oscillations were subsequently explained by the kind of delay differential equation that we're considering here[4].

This time we will construct a model in which the birth rate is delayed. We assume that the egg production per capita is given by $P_0 e^{-N/N_0}$. The total egg production is then

$$P(N) = P_0 N e^{-N/N_0} \ . \tag{1.41}$$

We take a constant death rate μ. Our delay differential equation is then

$$\frac{dN(t)}{dt} = P(t-T) - \mu N(t) \ . \tag{1.42}$$

[4]Nicholson's classical study was An Outline of the Dynamics of Animal Population. The model below was first propose by Gurney, Blythe, and Nisbet in Nicholson's blowflies revisited.

We have four constant parameters in this equation: P_0, N_0, μ and the delay time T. As a first step, we nondimensionalise our variables and parameters. We write

$$x = \frac{N}{N_0} \ , \quad \tau = \mu t \ , \quad \text{and} \quad a = \mu T \ , \quad b = P_0 T \tag{1.43}$$

This leaves us with the following equation that depends on just two dimensionless parameters, a and b, with a appearing as the delay time,

$$\frac{dx}{d\tau} = \frac{b}{a}\, x(\tau - a)\, e^{-x(\tau-a)} - x(\tau) \ . \tag{1.44}$$

First, we look at the equilibrium point $x = x_\star$, obeying

$$\frac{b}{a} x_\star e^{-x_\star} = x_\star \quad \Longrightarrow \quad x_\star = \log \frac{b}{a} \ . \tag{1.45}$$

This is a physical equilibrium point with $x_\star > 0$ only if $b > a$. We will assume this in what follows.

Next, we look at small perturbations about the fixed point. We write

$$x(\tau) = \log \frac{b}{a} + \epsilon(\tau) \ . \tag{1.46}$$

Substituting into our delay equation (1.44), we have

$$\begin{aligned} \frac{d\epsilon(\tau)}{d\tau} &= \frac{b}{a}\big(x_\star + \epsilon(\tau - a)\big) e^{-x_\star} e^{-\epsilon(\tau-a)} - x_\star - \epsilon(\tau) \\ &= \big(x_\star + \epsilon(\tau - a)\big)\big(1 - \epsilon(\tau - a)\big) - x_\star - \epsilon(\tau) \\ &= (1 - x_\star)\epsilon(\tau - a) - \epsilon(\tau) \ . \end{aligned} \tag{1.47}$$

Now we follow our previous analysis. We look for solutions of the form

$$\epsilon(\tau) = \epsilon_0 e^{s\tau} \ . \tag{1.48}$$

These obey (1.47) if

$$s = (1 - x_\star) e^{-sa} - 1 \ . \tag{1.49}$$

Again, we can look for (typically) complex solutions to this equation. This time the solutions are harder to come by. But it's rather straightforward to prove the following result:

Claim: The fixed point $x_\star = \log(b/a)$ is stable if $b/a < e^2$.

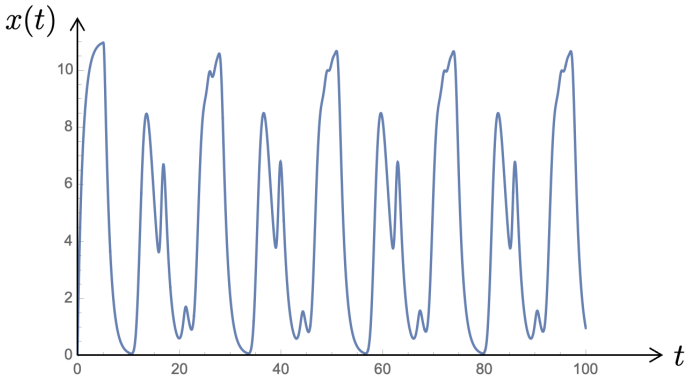


Figure 9. A typical oscillatory solution from the delay differential equation with $b \gg a$.

Proof: To see this, we decompose the equation (1.49) into real and imaginary parts by writing $s = \sigma + i\omega$. We have

$$\sigma = (1 - x_\star)e^{-\sigma a}\cos(\omega a) - 1 \quad \text{and} \quad \omega = -(1 - x_\star)e^{-\sigma a}\sin(\omega a)\ . \tag{1.50}$$

If $b/a < e^2$ then $|1 - x_\star| < 1$. Suppose that the system is unstable, so that $\sigma > 0$. Then the right we have $|(1 - x_\star)e^{-\sigma a}\cos(\omega a)| < 1$ and hence the right-hand side of the first equation above is necessarily negative. However the left-hand side is just σ which, by assumption, is positive. Hence we must have $\sigma < 0$ in this regime. □

Note that we haven't proven that the system is unstable for $b/a > e^2$ and, indeed, this isn't always the case. But it is true that, for suitably large b/a, the system is unstable and again exhibits oscillations. A typical numerical solution with $b \gg a$, with pretty funky oscillations, is shown in Figure 9.

1.2.3 And... Breathe

While our primary focus in this section is to understand the evolution of populations, the techniques that we're introducing are useful in many other contexts. Here we take a slight detour to explain how delay models can be used to model breathing.

There are two reasons that you breathe: one is to bring oxygen into your body, the other is to send carbon dioxide out. Because your body is good at keeping you alive, there is a feedback mechanism at play, and the volume V of the breath that you take depends on the concentration C of CO_2 in your blood. The relation between the two can be modelled by an equation of the form

$$V(C) = V_{\max}\frac{C^m}{A^m + C^m} \tag{1.51}$$

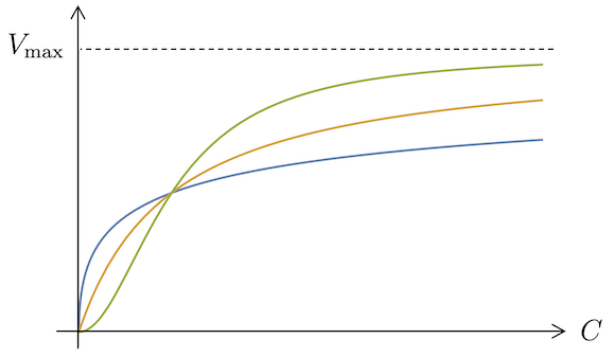


Figure 10. The Hill function plotted for $m = \frac{1}{2}$, for $m = 1$, and for $m = 2$.

for some constants A and m. Curves of this kind are sometimes called *Hill functions*. A variety of such curves, for fixed A and varying m, are plotted in Figure 10.

The concentration of CO_2 depends on time, so we have $C(t)$. Suppose that you take in CO_2 at a constant rate α. The amount that you exhale is proportional to both $C(t)$ and to the volume of your breath. But the feedback mechanism isn't entirely efficient, and the breath that you take at time t depends on the concentration at an earlier time $t - T$ for some delay T. The upshot is that we have a delay differential equation,

$$\begin{aligned} \frac{dC(t)}{dt} &= \alpha - b\,C(t)\,V(t) \\ &= \alpha - b\,C(t)\,V(C(t-T)) \\ &= \alpha - bV_{\max}C(t)\frac{C(t-T)^m}{A^m + C(t-T)^m} \ . \end{aligned} \tag{1.52}$$

Here b is a constant that specifies how good your breathing is at removing CO_2.

We can do our usual nondimensionalisation. We define

$$C' = \frac{C}{A} \ , \quad t' = \frac{\alpha t}{A} \quad \text{and} \quad \beta = \frac{AbV_{\max}}{\alpha} \ , \quad T' = \frac{\alpha T}{A} \ . \tag{1.53}$$

This gives us a delay equation for $C'(t')$ but, to keep our equations cleaner, we then just relabel $C' \to C$ and $t' \to t$ and also $T' \to T$. The upshot is that we get the delay differential equation

$$\frac{dC(t)}{dt} = 1 - \beta C(t)\frac{C(t-T)^m}{1 + C(t-T)^m} \ . \tag{1.54}$$

First, we look for the equilibrium solution $C_\star$, which sits at

$$\frac{C_\star^{m+1}}{1+C_\star^m} = \frac{1}{\beta} \ . \tag{1.55}$$

There is a unique solution $C_\star$ to this equation. This corresponds to a regular breathing pattern. That is generally regarded as good. The question we would like to ask is: when is this breathing pattern stable?

The idea here is that something might change to alter your breathing pattern. In particular there is a famous irregularity in breathing known as *Cheynes-Stokes respiration*. If you get it, it's most likely that you have only days to live or you're on the top of a mountain (or possibly both). While it might not be much comfort, it's thought that this can be traced to some complicated physiology which results in a change of the coefficient m in the Hill function (1.51).

We look at small perturbations of form

$$C(t) = C_\star + \epsilon(t) \ . \tag{1.56}$$

By now, the path should be familiar. We substitute this into (1.54) to derive the linearised delay differential equation. A little algebra gives

$$\begin{aligned} \frac{d\epsilon(t)}{dt} &= -\frac{1}{C_\star}\epsilon(t) - \frac{m}{C_\star}\left(1 - \frac{C_\star^m}{1+C_\star^m}\right)\epsilon(t-T) \\ &= -p\epsilon(t) - q\epsilon(t-T) \ . \end{aligned} \tag{1.57}$$

In the second line we've defined the positive constants $p = 1/C_\star$ and $q = m/C_\star(1+C_\star^m)$. To solve this linear equation, we make the usual ansatz

$$\epsilon(t) = \epsilon_0 e^{st} \ . \tag{1.58}$$

This leaves us with the algebraic equation

$$s = -p - qe^{-sT} \ . \tag{1.59}$$

We decompose this into real and imaginary parts by writing $s = \sigma + i\omega$ to get

$$\begin{aligned} \sigma &= -p - qe^{-\sigma T}\cos(\omega T) \\ \omega &= qe^{-\sigma T}\sin(\omega T) \ . \end{aligned} \tag{1.60}$$

We can immediately read off some results.

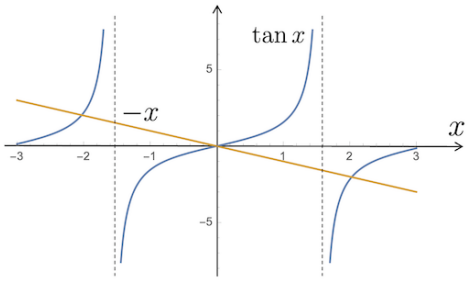
First, when $T = 0$ so there is no delay, we have $\sigma = -(p+q) < 0$ so the system is stable.

Second, $\omega = 0$ is always a solution but there will be a second solution with $\omega \neq 0$ whenever $qe^{-\sigma T} > 1$

Third, the system is stable whenever $p > q$. To see this, note that we have $(\sigma+p)^2 + \omega^2 = q^2 e^{-2\sigma T}$, so $qe^{-\sigma T} > |\sigma + p|$. Suppose that the system is unstable, so that $\sigma > 0$. In this case, $q > qe^{-\sigma T} > |\sigma + p| > p$. So instability necessarily implies $q > p$. That isn't to say that all systems with $q > p$ are necessarily unstable; some may be stable. But all those with $q < p$ are definitely stable.

This suggests that, if we want to look for when instability occurs, we should start with $p > q$, fix p and then increase q. The point of marginal stability occurs when $\sigma = 0$ and we have

$$\begin{aligned} p &= -q\cos(\omega T) \\ \omega &= q\sin(\omega T) \end{aligned} \quad \Longrightarrow \quad \begin{aligned} p\tan(\omega T) &= -\omega \\ p^2 + \omega^2 &= q^2 \end{aligned} \; .$$



We can always find solutions to the first of these equations in the region $\omega T \in (\pi/2, \pi)$, as shown in the figure on the right where we plot $\tan x$ and $-x$. Call this intersection point $\omega_\star(p, T)$. The second equation above then gives a necessary relationship between p, q and T for marginal stability:

$$q^2 = p^2 + \omega_\star^2 \; . \tag{1.61}$$

For q less than this value, breathing is stable. For q greater than this value, you're in trouble.

There's a lot of objects above that we've only defined implicitly – like $C_\star$ and $\omega_\star$ – but, given particular values of the parameters you could easily find these numerically. Nonetheless, the implicit nature of these constants makes it a little tricky to determine, say, the critical value of m in the function (1.51) in terms of T and β (or equivalently $C_\star$). We can make some minor progress if we substitute $p = 1/C_\star$ and $q = m/C_\star(1+C_\star^m)$ into (1.61). Using the fact that $\pi/2T < \omega_\star < \pi/T$, we get that the critical value of m lies between

$$1 + \frac{C_\star \pi^2}{4T^2} < \frac{m^2}{(1+C_\star^m)^2} < 1 + \frac{C_\star^2 \pi^2}{T^2} \; . \tag{1.62}$$

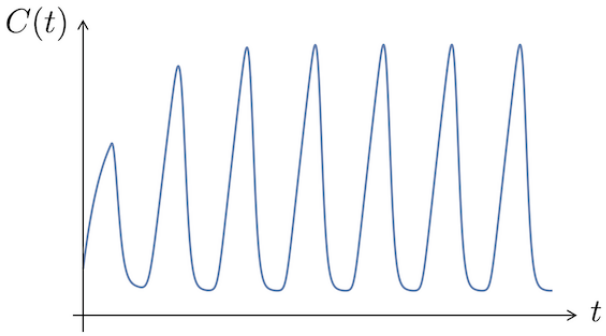


Figure 11. A numerical solution corresponding to an unstable breathing pattern.

The left-hand inequality immediately tells us that instability only kicks for some value of $m > 1$. A numerical solution to the original delay differential equation (1.52) in the unstable regime is shown in Figure 11.

1.3 Age Concern

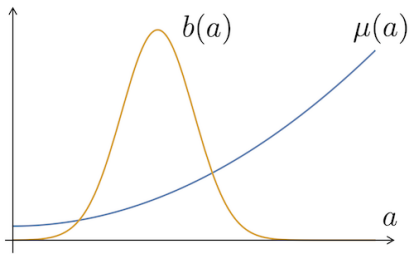
Until now, we've treated the population N as just a single number. But, for many questions, we may want to be more discerning. Here we want to take into account one particularly important variable: age.

To this end, we consider the population density $n(a, t)$. Here t is time, as before, while a is the age of individuals. Roughly speaking, $n(a, t)$ counts the number of individuals of age a. More precisely, we should think of $n(a, t)\, da$ as the number of individual with age between a and $a + da$. The total population at time t is

$$N(t) = \int_0^\infty da\; n(a, t) \; . \tag{1.63}$$

For mortals, we can reduce the upper limit of this integral to the lifespan of the species: a few weeks for blowfly, somewhat longer for humans.

It's reasonable to expect both the birth rate $b(a)$ and the death rate $\mu(a)$ to depend on the age of the population: the young are fertile, the old vulnerable. We might expect typical functions to look something like those shown on the right.



We now want to think about how the population ages. Suppose that your age is a at time t. It is a sad fact of life that at a short time δt later, you have only two options open to you: either your age is $a + \delta t$, or you're dead. In equations, this translates to

$$n(a + \delta t, t + \delta t) = n(a, t) - \mu(a) n(a, t) \delta t \tag{1.64}$$

where we're dropping terms of order δt^2. We can Taylor expand the left-hand side and divide through by δt to get the partial differential equation

$$\frac{\partial n}{\partial t} + \frac{\partial n}{\partial a} = -\mu(a)\, n(a, t) \ . \tag{1.65}$$

This is sometimes called the *von Foerster equation*.

So far this equation only captures death. To inject a level of joy into the proceedings, we also need to talk about births. It's hard to be born at any age other than $a = 0$ and so the birth rates arise as a boundary condition on the function $n(a, t)$ at $a = 0$:

$$n(0, t) = \int_0^\infty da\ b(a)\, n(a, t) \ . \tag{1.66}$$

In any realistic situation, we don't have any problem with convergence: both the population and birth rate will vanish for some $a > a_{\text{ancient}}$.

A Cute Analogy with Fluid Mechanics

For a one-dimensional fluid with velocity u, we usually define the *material derivative*,

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} \ . \tag{1.67}$$

This tells us how any quantity changes as we drift along with the fluid. For our population with age structure, we have the corresponding material derivative

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \frac{\partial}{\partial a} \ . \tag{1.68}$$

The analog of the fluid velocity in this equation is just $u = 1$, corresponding to the fact that we get older by one year per year. In this case, the material derivative tells us how things change if we're aging with the population (which we are!).

The fluid analogy is nice enough, but it breaks down when it comes to the initial boundary condition (1.66) dictated by the birth rates. This is because, from the perspective of fluid mechanics, this is a "non-local" boundary condition: what happens at $a = 0$ depends on the value of the dynamical field $n(a, t)$ at all values of a.

1.3.1 Separable Solutions

Inspired by the Malthusian growth that we saw earlier, we will look for separable solutions of (1.65) of the form

$$n(a,t) = \tilde{n}(a)e^{rt} \tag{1.69}$$

for some function $\tilde{n}(a)$ and some $r \in \mathbb{R}$. This means that some age profile $\tilde{n}(a)$ remains unchanged, growing (or shrinking) exponentially. We substitute this into (1.65) to get

$$\frac{\partial \tilde{n}}{\partial a} = -\big[r + \mu(a)\big]\,\tilde{n}(a) \quad \Longrightarrow \quad \tilde{n}(a) = n_0 e^{-ra} \exp\left(-\int_0^a ds\ \mu(s)\right) . \tag{1.70}$$

We see that the age profile is determined only by the death rate. That exponentiated integral $e^{-\int_0^a ds\,\mu(s)}$ has the interpretation of the probability of surviving to age a.

We still have to impose the birthing condition (1.66) which leaves us with the following, slightly unusual expression

$$\phi(r) := \int_0^\infty da\ b(a)e^{-ra} \exp\left(-\int_0^a ds\ \mu(s)\right) = 1 . \tag{1.71}$$

Here that double integral defines the function $\phi(r)$ which, you can see by inspection, is a monotonically decreasing function of r. The birthing condition requires $\phi(r) = 1$ which we view as determining the exponent r.

The population may grow or shrink depending on the sign of r. To get a handle on r, we can look at $\phi(0)$. Because $\phi(r)$ is monotonically decreasing, we can immediately say that

- If $\phi(0) > 1$ then $r > 0$ and the population grows in size.
- If $\phi(0) < 1$ then $r < 0$ and the population shrinks.

In this sense, $\phi(0)$ can be viewed as the average number of offspring per individual.

Note that adding age structure hasn't changed the overall story of exponential growth. In particular, it doesn't open up the possibility for oscillations in the way that delay equations did.

As a sanity check, we can look at what this model gives us in the case that birth and death rates are constant, so $b(a) = b$ and $\mu(a) = d$. Then the function $\phi(r)$ defined in (1.71) becomes

$$\phi(r) = \int_0^\infty da\ be^{-(r+d)a} = \frac{b}{r+d} . \tag{1.72}$$

Evaluated at $r = 0$, we have $\phi(0) = b/d$, reproducing our old Malthusian result: if $b > d$ then the population increases exponentially, while if $b < d$ it shrinks exponentially.

1.3.2 More General Solutions

There is a more general class of solutions to the equation (1.65). To motivate this, note that the left-hand-side takes the form of a wave equation. If we could somehow banish death, so that $\mu(a) = 0$, then the solutions would be

$$\frac{\partial n}{\partial a} + \frac{\partial n}{\partial t} = 0 \quad \Longrightarrow \quad n(a,t) = f(a-t) \tag{1.73}$$

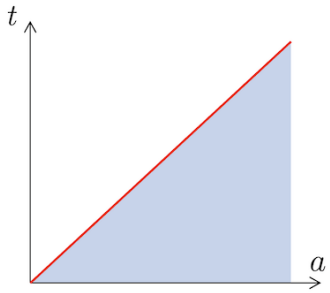
for any function $f(a-t)$. This is telling us that there is some initial population profile $f(a)$ and it just propagates forward in the "age"-direction as people get older. Sadly, death is unavoidable but, at least in this context, is easily dealt with and (1.65) has the solution

$$n(a,t) = f(a-t)\, \exp\left(- \int_0^a ds\ \mu(s) \right) \ . \tag{1.74}$$

This is the same age profile propagating forward, but now you lose a few as you go.

If we start with some initial age distribution $n(a,0) = n_0(a)$, then that's enough to tell us what happens for all $a > t$. We have

$$\begin{aligned} n_0(a) &= f(a)\, \exp\left(- \int_0^a ds\ \mu(s) \right) \\ \Longrightarrow \quad n(a,t) &= n_0(a-t)\, \exp\left(- \int_{a-t}^a ds\ \mu(s) \right) . \end{aligned}$$



But this form of the solution holds only when $a > t$, which is the shaded region in the diagram. This is the region where it's sufficient to know the initial population profile. For the other region, with $t > a$, we need to take the births into account. For this, we need the boundary condition that specifies the birth rate (1.66). Substituting the wave-like solution (1.74) into this boundary condition gives

$$n(0,t) = f(-t) = \int_0^\infty da\ b(a)\, n(a,t) \ . \tag{1.75}$$

We have the solution for $n(a,t)$ when $a > t$, but not when $a < t$. We can make use of this by splitting the integral above into two pieces and writing

$$\begin{aligned} f(-t) &= \int_0^t da\ b(a)\, n(a,t) + \int_t^\infty da\ b(a)\, n(a,t) \\ &= \int_0^t da\ b(a)\, f(a-t) e^{-\int_0^a ds\, \mu(s)} + \int_t^\infty da\ b(a)\, n_0(a-t) e^{-\int_{a-t}^a ds\, \mu(s)} \ . \end{aligned} \tag{1.76}$$

This is a really horrible equation! In principle, it should be used to determine the function $f(-t)$ for $t > 0$ by integrating $f(-t')$ from $t' = 0$ to $t' = t$. We then use this expression for $f(-t)$ in (1.74) to determine the population at times $t > a$. In practice, this is easier said that done.

There is one minor simplification that we can make. The second term in (1.76) gives the contribution from the offspring of the original population $n_0(a)$. The exponential suppression is telling us that must they survive from age $a-t$ to age a. At late times, we expect that the number of these offspring will tend to zero as the old fail to reproduce, and the resulting population profile will be independent of the initial condition. Indeed, the separable solution didn't give us the opportunity to impose an initial condition. In this case, we have to solve the integral equation

$$f(-t) \approx \int_0^t da\; b(a)\, f(a-t)\, \exp\left(-\int_0^a ds\; \mu(s)\right) \quad . \tag{1.77}$$

It's still not a pleasant equation to solve.

An Example

We can illustrate these ideas with a simple example. We take a constant death rate, $\mu(a) = d$, and a birth rate that is a step function, where you can only give birth if you're younger than some fixed age A,

$$b(a) = \begin{cases} b & 0 < a < A \\ 0 & a > A \end{cases} \quad . \tag{1.78}$$

To keep things simple, we'll take our initial population to be $n_0(a) = 1$ for all ages a. Admittedly, this isn't particularly realistic (even Methuselah didn't make it past 1000) but, as we saw above, we expect that the initial population will soon be unimportant. We now need to solve for the population in two different regimes: $a > t$ and $a < t$. For $a > t$, we have simply

$$n(a,t) = e^{-dt} \quad \text{for } a > t \; . \tag{1.79}$$

The other regime $a < t$ is the trickier one. Here we must treat the regimes $t < A$ and $t > A$ separately. For $t < A$, the integral equation (1.76) becomes

$$\begin{aligned} f(-t) &= b\int_0^t da\; f(a-t)e^{-ad} + b\int_t^A da\; e^{-dt} \\ &= b\int_0^t d\tau\; \left(f(-\tau)e^{-d(t-\tau)}\right) + b(A-t)e^{-dt} \quad \text{for} \quad a < t < A \end{aligned} \tag{1.80}$$

where, in the second line, we changed the integration coordinate to $\tau = t - a$. This is actually an equation we can solve. To do this, it's simplest to define $\tilde{n}(t) = n(0,t) = f(-t)$. We have

$$\tilde{n}(t) = b\int_0^t d\tau\ \left(\tilde{n}(\tau)e^{-d(t-\tau)}\right) + b(A-t)e^{-dt} \quad \text{for} \quad a < t < A\ . \tag{1.81}$$

We can turn this into a more familiar differential equation simply by differentiating with respect to t. We have

$$\begin{aligned}\frac{d\tilde{n}}{dt} &= b\tilde{n}(t) - bd\int_0^t d\tau\ \tilde{n}(\tau)e^{-d(t-\tau)} - b(1+d(A-t))e^{-dt} \\ &= (b-d)\tilde{n}(t) - be^{-dt} \quad \text{for} \quad a < t < A\end{aligned} \tag{1.82}$$

where, to get to the second line, we've substituted in the integral expression in (1.81). But this equation is easily solved. We have

$$\tilde{n}(t) = \hat{N}e^{(b-d)t} + e^{-dt} \quad \text{for} \quad a < t < A\ . \tag{1.83}$$

Putting this together, we get the following solution

$$n(a,t) = \begin{cases} e^{-dt} & \text{for } t < a \\ e^{-dt}\left[\hat{N}e^{b(t-a)} + 1\right] & \text{for } a < t < A \end{cases}\ . \tag{1.84}$$

A non-vanishing $\hat{N}$ can be viewed as an injection of births and leads to a discontinuity in $n(a,t)$ along the line $a = t$.

We're left with understanding what happens for $A > 0$. Here we can follow the same steps and derive an integral equation for $\tilde{n}(t) = n(0,t) = f(-t)$.

$$\tilde{n}(t) = b\int_{t-A}^t d\tau\ \tilde{n}(\tau)e^{-d(t-\tau)} \quad \text{for } t > a, A\ . \tag{1.85}$$

This time, when we differentiate with respect to time, this becomes a delay differential equation

$$\frac{d\tilde{n}(t)}{dt} = (b-d)\tilde{n}(t) - \tilde{n}(t-A)e^{-dA}\ . \tag{1.86}$$

We saw how to deal with linear delay differential equations of this kind in Section 1.2. We can look for solutions of the form $\tilde{n}(t) \sim e^{st}$ where the exponent s must obey

$$s = (b-d) - be^{-(d+s)A}\ . \tag{1.87}$$

As usual, the population grows if $\mathrm{Re}(s) > 0$ and shrinks if $\mathrm{Re}(s) < 0$. The population is stable if $s = 0$, which, from (1.87), requires a fine tuning between birth and death rates, now given by

$$b = \frac{d}{1 - e^{-dA}} \ . \tag{1.88}$$

We see that we must have $b > d$ to maintain a stable population. This is because, in this model, only the young in the population can reproduce while anyone can die.

1.4 Interacting Species

So far we've discussed the evolution of just a single population. At times there was a different species lurking in the background, like when we included terms describing predation, but this other species was very much an NPC in the story.

Things become significantly more interesting when there are two or more populations that interact with each other. This allows for a much richer collection of dynamics.

For much of this section, we will restrict ourselves to just two independent species. But the basis mathematical formalism is just as easy to state regardless of the number. Suppose that we have n different dynamical variables, $u_i(t)$ with $i = 1, \ldots, n$. We will study a class of first order dynamical systems that takes the form

$$\frac{du_i}{dt} = f_i(u_1, \ldots, u_n) \ . \tag{1.89}$$

A fixed point of this system is a steady state solution $u_i = u_i^\star$ obeying

$$f_i(u_1^\star, \ldots, u_n^\star) = 0 \quad \text{for each } i = 1, \ldots, n \ . \tag{1.90}$$

Given a fixed point, we would like to know whether it is stable or unstable. It's quite possible that it will be stable in some directions and unstable in others. To determine this, we expand

$$u_i(t) = u_i^\star + \xi_i(t) \tag{1.91}$$

where $\xi_i(t)$ is taken to be a small perturbation. We then expand (1.89) to linear order in ξ:

$$\frac{d\xi_i}{dt} = f_i(u^\star) + \left.\frac{\partial f_i}{\partial u_j}\right|_{u^\star} \xi_j + \ldots \tag{1.92}$$

The constant term vanishes precisely because we're at a fixed point. We're left with

$$\frac{d\xi_i}{dt} = J_{ij}\,\xi_j \quad \text{with} \quad J_{ij} = \left.\frac{\partial f_i}{\partial u_j}\right|_{u^\star} \ . \tag{1.93}$$

The matrix J is called the *Jacobian*. The stability of a fixed point is determined by the eigenvalues of this matrix. Suppose that we have an eigenvalue λ with corresponding eigenvector $\mathbf{x}$,

$$J\mathbf{x} = \lambda\mathbf{x} \ . \tag{1.94}$$

In general, λ could be real or complex. If $\mathrm{Re}(\lambda) < 0$, then the fixed point is stable in the $\mathbf{x}$ direction, while if $\mathrm{Re}(\lambda) > 0$ the dynamics will be unstable. If λ includes a complex part, then the trajectories will typically spiral towards or away from the fixed point.

In general, we will also want to plot the trajectories in the space $\mathbb{R}^n$ parameterised by u_i. We will gain plenty of experience in doing this as we go through various examples.

In much of what follows, we will consider situations with $n = 2$ species. In this case J is a 2×2 matrix with eigenvalues λ_1 and λ_2 and it's usually easiest to compute them by considering

$$\mathrm{Tr}\, J = \lambda_1 + \lambda_2 \quad \text{and} \quad \det J = \lambda_1\lambda_2 \ . \tag{1.95}$$

In particular, if $\mathrm{Tr}\, J < 0$ and $\det J > 0$ then both eigenvalues must be negative and the fixed point is stable.

1.4.1 Predator-Prey: The Lotka-Volterra Equations

The poster child for mathematical biology is the *Lotka-Volterra model*. It's a simple and instructive model for two species interacting. But, as we will see, the results it gives are far from generic.

Here's the setting. There are two species. The first is the prey and has population $N(t)$. These are the cute things that you are most likely rooting for. The other is the predator with population $P(t)$. These are the villains of the piece. Their dynamics is governed by the set of equations

$$\frac{dN}{dt} = N(a - bP) \quad \text{and} \quad \frac{dP}{dt} = P(dN - c) \tag{1.96}$$

with a, b, c and d are all positive numbers. These are the *Lotka-Volterra equations*

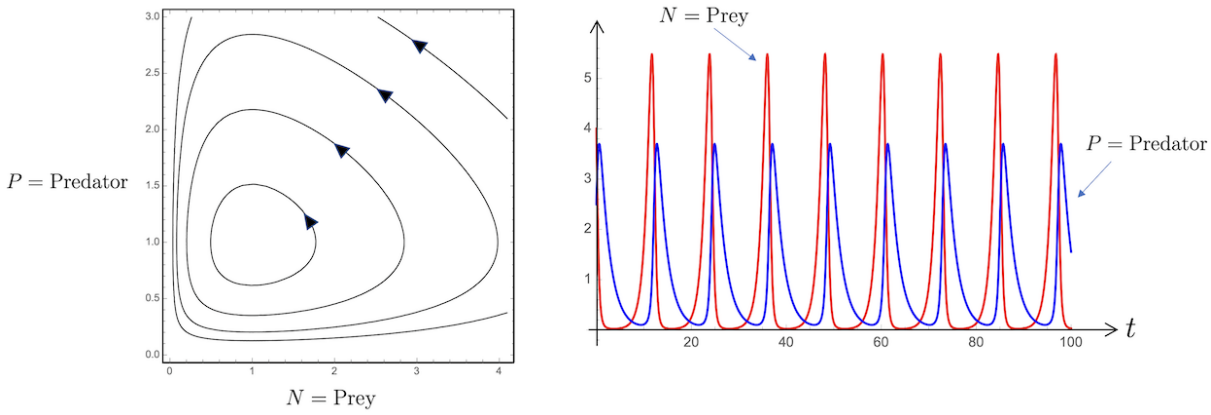


Figure 12. Solutions to the Lotka-Volterra equations plotted in the phase plane on the left, and as a function of time on the right. Both of these were plotted with $\alpha = 0.5$.

The physics behind the various terms and minus signs is straightforward. The $+aN$ term is telling us that, in the absence of predators, the prey prospers (because it eats grass, or because it preys on something even more helpless that doesn't even get a mention in the equations). Meanwhile, the $-cP$ term tells us that, in the absence of prey, the predators die out. Each equation also contains an NP term that captures what happens when the two populations interact. The $\pm$ signs tell us that the result is good for one, less good for the other.

As usual, we can do some rescaling. We write $u = dN/c$ and $v = bP/a$ and $t' = at$ and then, annoyingly, relabel t' as t. The end result is the set of non-dimensionalised equations

$$\frac{du}{dt} = u(1-v) \quad \text{and} \quad \frac{dv}{dt} = \alpha v(u-1) \ . \tag{1.97}$$

These equations depend on the single dimensionless parameter $\alpha = c/a > 0$.

Rather unusually, it's straightforward to find exact solutions to these equations. We can think of $P = P(N)$ or, equivalently, $v = v(u)$ to get a direct relationship between the two populations. Dividing the two equations in (1.97) gives

$$\frac{dv}{du} = \frac{\alpha v(u-1)}{u(1-v)} \ . \tag{1.98}$$

This differential equation is separable and we have

$$\int dv \ \frac{1-v}{v} = \alpha \int du \ \frac{u-1}{u} \quad \Longrightarrow \quad \log v - v + \alpha(\log u - u) = \text{constant} \ . \tag{1.99}$$

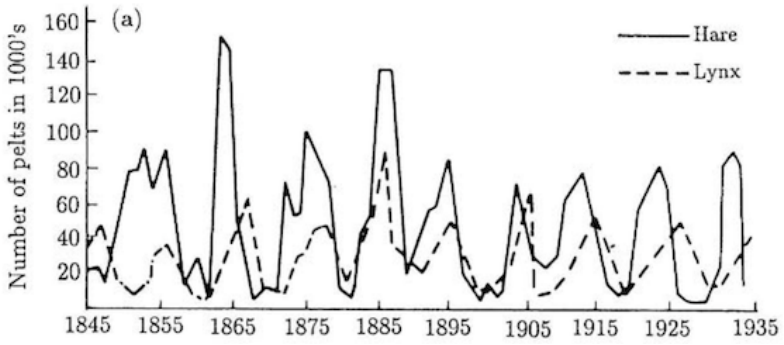


Figure 13. Sales of hare and lynx pelts, taken from Murray's book on Mathematical Biology.

We can plot these orbits in (u, v)-plane which, in this context, is also known as the *phase plane*. The resulting orbits are shown on the left of Figure 12. The trajectories all exhibit the same essential behaviour: they orbit the fixed point $(u, v) = (1, 1)$.

This is telling us that the populations of both species oscillate in time. It's more difficult to extract information about the time dependence, and this typically needs to be done numerically. An example is shown on the right of Figure 12.

We can also compute the average population $\langle u \rangle$ and $\langle v \rangle$. We take, for example, the first equation in (1.97) and write it as $\dot{u}/u = 1 - v$. Integrating the left-hand side over a single period T gives

$$\int_0^T dt\ \frac{\dot{u}}{u} = \int_{u_{\rm start}}^{u_{\rm finish}} du\ \frac{1}{u} = \Big[\log u\Big]_{u_{\rm start}}^{u_{\rm finish}} = 0\ . \tag{1.100}$$

This vanishes because the orbits are closed, so $u_{\rm start} = u_{\rm finish}$. Integrating the right-hand side over a single period must similarly vanish: we have

$$\int_0^T dt\ (1 - v) = T - T\langle v \rangle = 0 \quad \Longrightarrow \quad \langle v \rangle = 1\ . \tag{1.101}$$

We see that the average sits at the fixed point which, back in our original variables, is $\langle P \rangle = a/b$. A similar argument shows that $\langle u \rangle = 1$.

From either of the plots in Figure 12, we can extract a story about the underlying ecology. First, the number of prey increases. After a short time, this results in a corresponding increase in the number of predators. This decreases the prey, resulting in less to eat and a decrease in predators which then allows the prey to thrive and so the cycle of life repeats.

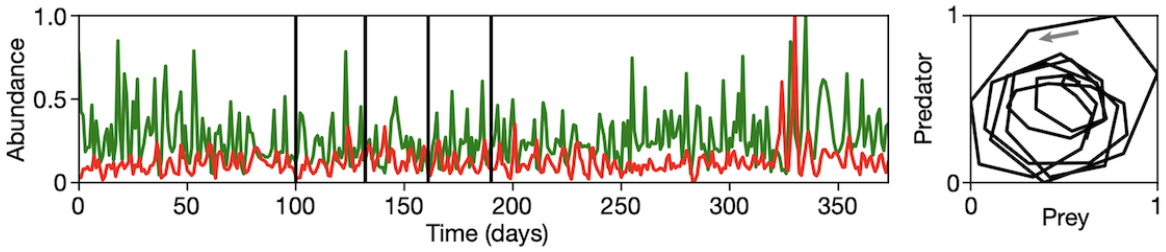


Figure 14. The population of prey plankton (in green) and predator algae (in red), both normalised to one, together with the trajectories in the phase plane. This data is taken from the paper Long-term cyclic persistence in an experimental predator-prey system by Blasius et al.

An obvious question is: does this match what is seen in the wild? The answer is: sometimes but not often. As we will soon see, the mathematical structure of the Lotka-Volterra equations is rather special and doesn't survive most perturbations. For that reason, we should be suspicious about its applicability. Nonetheless, there is a famous and rather wonderful story that seems to give support to this model. For many decades, the Hudson Bay Trading Company kept records of the numbers of pelts they sold. These include pelts of hares which are natural prey, and lynx which are natural predators. The results are shown in Figure 13 and are closely resemble the Lotka-Volterra oscillations. Clearly there are many other factors at play here, and so it's unsurprising that the data is rather messy. One puzzle is that, around 1885, it looks like the lynx population rises before the hare population, which prompted many theories, the best of which can be viewed here.

Cleaner data can be found for populations in bottles, rather than in the wild. The population of rotifer, which is a kind of plankton, and their algae prey are shown in Figure 14, clearly exhibiting predator-prey cycles.

The Hamiltonian Structure of Lotka-Volterra

We see from (1.99) that there is a conserved quantity in the Lotka-Volterra equations,

$$H = v - \log v + \alpha(u - \log u) \ . \tag{1.102}$$

The value of H determines the chosen orbit. It takes its minimum value at the fixed point, where $H_{\rm min} = (1 + \alpha)$. Orbits that are further out have larger values of H.

The existence of a conserved quantity is reminiscent of the conserved energy in classical mechanics and you might wonder if there's a deeper connection. It turns out that

the Lotka-Volterra system is an example of a Hamiltonian system of the kind that we met in the course on Classical Dynamics. Specifically, we can view the coordinates $q = \log v$ and $p = \log u$ as position and momenta, and the Lotka-Volterra equations can then be written a

$$\frac{dq}{dt} = \frac{\partial H}{\partial p} \quad \text{and} \quad \frac{dp}{dt} = -\frac{\partial H}{\partial q} \quad \text{with} \quad H = e^q - q + \alpha(e^p - p) \ . \tag{1.103}$$

This Hamiltonian structure makes the Lotka-Volterra equations rather special and, to some extent, unrealistic. In physics, the fundamental laws are all Hamiltonian, a fact that can ultimately be traced to the quantum nature of reality. In mathematical biology, there is no such reason that the underlying laws should be pretty. Indeed, we will soon see more realistic generalisations of the Lotka-Volterra equations that do not preserve this Hamiltonian structure.

Stability Analysis

While we can understand the solutions to Lotka-Volterra equations by direct integration, this won't be true of the generalisations that we look at. To prepare ourselves, it will be useful to look again at the Lotka-Volterra equations from other perspectives which, although they are less powerful, will easily generalise to the more complicated situations.

The first method is to do a stability analysis of the fixed point. In fact, there are two fixed points at $(u, v) = (0, 0)$ and $(u, v) = (1, 1)$. The Jacobian matrix (1.93) takes the general form

$$J = \begin{pmatrix} 1 - v & -u \\ \alpha v & \alpha(u - 1) \end{pmatrix} \ . \tag{1.104}$$

To understand the stability, we need to compute the eigenvalues λ of this matrix at each of the fixed points. For the trivial fixed point we have

$$(u, v) = (0, 0) \quad \Longrightarrow \quad \lambda = 1, -\alpha \ . \tag{1.105}$$

This means that the origin is a saddle point. This is fortunate, but unsurprising. It's fortunate because it means that our two species will not naturally be driven to extinction. It's unsurprising because we set things up such that the prey flourish in the absence of predators, while the predators suffer in the absence of prey.

The second fixed point is more interesting. We have

$$(u, v) = (1, 1) \quad \Longrightarrow \quad \lambda = \pm i\sqrt{\alpha} \ . \tag{1.106}$$

So in this case, the eigenvalues are pure imaginary. Taken at face value, this tells us that the trajectories in the phase plane orbit the fixed point. To see this, note that the eigenvectors of the Jacobian are $\mathbf{x}_\pm = (\pm i, \sqrt{\alpha})$. The most general linearised solution is then

$$(u, v) = (1, 1) + A_+ \mathbf{x}_+ e^{i\sqrt{\alpha}t} + A_- \mathbf{x}_- e^{-i\sqrt{\alpha}t} \tag{1.107}$$

where A_+ and A_- can be viewed as initial conditions and should be chosen so that (u, v) is real. For example, we could take $A_+ = A_- = \frac{1}{2}$, in which case we have the trajectory

$$(u, v) = (1, 1) + \big(-\sin(\sqrt{\alpha}t), \sqrt{\alpha}\cos(\sqrt{\alpha}t)\big) \ . \tag{1.108}$$

As we have seen, the trajectories in the Lotka-Volterra model do, in fact, orbit the fixed point but the linear analysis shown here isn't enough to demonstrate it. That's because, in general, a fixed point with purely imaginary eigenvalues could be rendered stable or unstable from higher order effects.

Nullclines

The second method to analyse dynamical systems of this type is to search for the *nullclines*. These are simply the curves

$$\begin{aligned} \dot{u} = 0 \quad &\Longrightarrow \quad u = 0 \ \text{ or } \ v = 1 \\ \dot{v} = 0 \quad &\Longrightarrow \quad v = 0 \ \text{ or } \ u = 1 \ . \end{aligned} \tag{1.109}$$

These are plotted on the left of Figure 15. On the red nullclines, we have $\dot{u} = 0$, and hence the flows are either left or right. On the blue nullclines we have have $\dot{v} = 0$ and hence the flows are up or down. We necessarily have a fixed point whenever different nullclines meet because both $\dot{u} = \dot{v} = 0$. We see that the nullclines split the phase plane into quadrants where flows point north-east, north-west, south-west, or south-east. From this, we can surmise the general topology of the flows. The full flow structure is shown on the right of Figure 15.

Gone Fishing

We will see several generalisations of the Lotka-Volterra equations shortly. But there is one that is very straightforward and yet comes with a counterintuitive punchline.

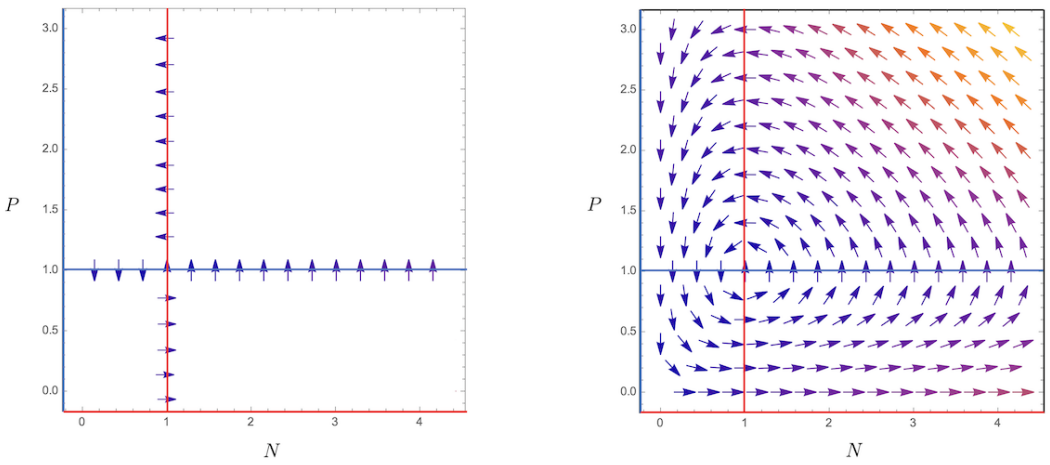


Figure 15. On the left, the nullclines divide the phase plane into four quadrants, where the flows are up/down and left/right. The fixed point sits at the point where the different nullclines meet. On the right, the full flow in the phase plane.

Suppose that there is an additional effect at play, where the one or both of the species is hunted. This is usually phrased as "fishing" in the literature, as a nod to Volterra's initial motivation for writing down this system of equations. We could, for example, add an additional term to capture the effect of hunting the prey. In this case, the original equations (1.96) could be changed to

$$\frac{dN}{dt} = N(a - bP) - hN \quad \text{and} \quad \frac{dP}{dt} = P(dN - c) \ . \tag{1.110}$$

The additional hN term captures the effect of hunting. This is straightforward because it doesn't change the structure of the equations at all, at least if $h < a$. Everything that we said above still holds, but the fixed point shifts to

$$(N, P) = \left(\frac{c}{d}, \frac{a-h}{b}\right) \ . \tag{1.111}$$

The surprise is that hunting the prey hasn't changed the average prey population at all. Instead, counterintuitively, it has reduced the average predator population!

This cute effect is the reason why biologists first studied the Lotka-Volterra equations. (Lotka was the first to introduce the equation but that was in the context of chemical reactions.) The Italian biologist Umberto D'Ancona noted that the proportion of sharks and skates and other slightly dangerous predatory things dramatically increased in the Adriatic during the First World War when fishing was largely curtailed. He spent a long

time puzzling over this until finally doing the sensible thing and asking a mathematician for help. Fortunately, his father-in-law was Vito Volterra, one of Italy's great 20^{th} century mathematicians. This was the motivation for Volterra to write these equations and show that, as observed, in the absence of hunting, the predator population should increase.

1.4.2 Predator-Prey: A Logistic Twist

No set of equations is ever the last word in mathematical biology. There are always ways to finesse the model, to include some extra factor that may, or may not, change the qualitative behaviour. In this section and the next, we look at two such generalisations of the Lotka-Volterra equations.

We already introduced the logistic equation back in Section 1.1: it includes an additional term for a single species that limits the ability to reproduce as the population grows. We can easily incorporate such terms into our predator-prey model, so that they Lotka-Volterra equations (1.96) become

$$\frac{dN}{dt} = N\left(a - bP - \frac{N}{K_1}\right) \quad \text{and} \quad \frac{dP}{dt} = P\left(dN - c - \frac{P}{K_2}\right) \ . \tag{1.112}$$

Here we've introduced two carrying capacities, K_1 and K_2, for the two species. After rescaling, the equations can be written as

$$\frac{du}{dt} = u(1 - v - \mu_1 u) \quad \text{and} \quad \frac{dv}{dt} = \alpha v(u - 1 - \mu_2 v) \tag{1.113}$$

where the (inverse) carrying capacities are to be found in the positive constants μ_1 and μ_2. We will take $\mu_i < 1$.

These equations have three fixed points, provided that $\mu_1 < 1$. There is the trivial fixed point $(u, v) = (0, 0)$ and a new fixed point in which the predators are extinct while the prey reach their logistic equilibrium: $(u, v) = (1/\mu_1, 0)$. The coexistence fixed point is now

$$(u^\star, v^\star) = \frac{1}{1 + \mu_1\mu_2}(1 + \mu_2, 1 - \mu_1) \ . \tag{1.114}$$

Note that we have $v^\star > 0$ when $\mu_1 < 1$. Note also that $u^\star > 1$ and $v^\star < 1$, so the additional terms have the effect of reducing the predator population while increasing the prey population at the fixed point.

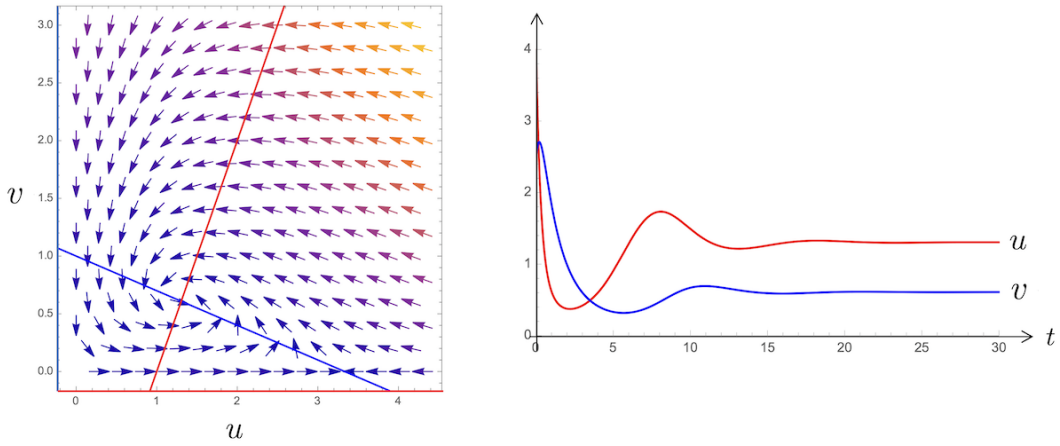


Figure 16. On the left: the flow in phase plane. The nullclines are shown in red and blue and determine where the flow is horizontal or vertical respectively. On the right, the time dependence of $u(t)$ and $v(t)$. Both plots were made with $\alpha = 0.5$, $\mu_1 = 0.3$ and $\mu_2 = 0.5$.

We can perform the usual stability analysis. About the fixed point (1.114), the Jacobian is

$$J = \begin{pmatrix} 1 - v^\star - 2\mu_1 u^\star & -u^\star \\ \alpha v^\star & \alpha(u^\star - 1 - 2\mu_2 v^\star) \end{pmatrix} = \begin{pmatrix} -\mu_1 u^\star & -u^\star \\ \alpha v^\star & -\alpha\mu_2 v^\star \end{pmatrix} \tag{1.115}$$

where, to get to the second expression, we've used the fact that, for example, $1 - v^\star - \mu_1 u^\star = 0$. At this point, we don't need to use the explicit expressions in (1.114) for the fixed point: it's sufficient to stare at the signs of the terms. Recall that the determinant of a 2×2 matrix is the product of eigenvalues while the trace is the sum. We have $\det J > 0$ and $\mathrm{Tr} J < 0$ which means that both eigenvalues must be negative. We learn that the fixed point is now stable.

The resulting flows in phase plane are shown in Figure 16, together with the nullclines which now sit at an angle. Note that there is an additional fixed point on the u-axis, where two different nullclines meet. On the right of Figure 16, the time-dependent solutions are plotted. This makes the dynamics clear: after a number of mild oscillations, the two populations settle down to an equilibrium with more (non-dimensionalised) prey than predators.

A Lyapunov Function

There is another way of seeing that all trajectories spiral into the fixed point. This comes from looking at the "Hamiltonian" that was constant on Lotka-Volterra orbits

$$H = v - \log v + \alpha(u - \log u) \ . \tag{1.116}$$

For our new model, it will prove useful to introduce the slightly different function

$$\tilde{H} = v - v^\star \log v + \alpha(u - u^\star \log u) \ . \tag{1.117}$$

This coincides with the Hamiltonian H when $u^\star = v^\star = 1$ for the Lotka-Volterra model. The importance of this new function can be seen by taking its time derivative

$$\begin{aligned} \frac{d\tilde{H}}{dt} &= \left(1 - \frac{v^\star}{v}\right)\dot{v} + \alpha\left(1 - \frac{u^\star}{u}\right)\dot{u} \\ &= \alpha\,(v - v^\star)\,(u - 1 - \mu_2 v) + \alpha\,(u - u^\star)\,(1 - v - \mu_1 u) \ . \end{aligned} \tag{1.118}$$

Now recall that $1 = v^\star + \mu_1 u^\star$ and $1 = u^\star - \mu_2 v^\star$. We use this to replace the 1's in the expression above to get

$$\begin{aligned} \frac{d\tilde{H}}{dt} &= \alpha\big[(v - v^\star)(u - u^\star - \mu_2(v - v^\star)) - (u - u^\star)(v - v^\star + \mu_1(u - u^\star)\big] \\ &= -\alpha\big[\mu_2(v - v^\star)^2 + \mu_1(u - u_\star)^2\big] \\ &\leq 0 \ . \end{aligned} \tag{1.119}$$

That's rather nice. We've managed to construct a monotonically decreasing function $\tilde{H}(u, v)$ along the flow. Indeed, it stops changing only when we hit the fixed point $(u^\star, v^\star)$. This, again shows that all flows necessarily end up at the fixed point. This is an example of a *Lyapunov function*.

1.4.3 Predator-Prey: I'm Full Now

In both the original Lotka-Volterra model, and the logistic generalisation above, there is no end to the predator's appetite: they continue to voraciously work their way through the prey, like pac-man munching those dots.

A more realistic model might be to allow the predators to become sated at some point, so that the amount of predation saturates as $u \to \infty$, rather than continuously increasing. Here is an example of a model that has this property,

$$\frac{du}{dt} = u\left(1 - u - \frac{bv}{u + d}\right) \quad \text{and} \quad \frac{dv}{dt} = \alpha v\left(1 - \frac{v}{u}\right) \ . \tag{1.120}$$

We see that there is novelty in both equations. For the prey population $u(t)$, the final predation term has the promised effect that, while proportional to the predators, the coefficient plateaus as $u \to \infty$. Meanwhile, the evolution of predators is different from what we had before. Now the reproduction term is $\alpha > 0$, but their carrying capacity is equal to the prey population.

We can use the techniques that we developed above to see the behaviour of this system. The non-trivial fixed point has $v^\star = u^\star$ and

$$1 - u^\star = \frac{bu^\star}{u^\star + d} \quad \Longrightarrow \quad u^\star = \frac{1}{2}\left[1 - b - d + \sqrt{(1-b-d)^2 + 4d}\right] \tag{1.121}$$

where we've picked the root that is positive. We can now check the stability. Here there's a short calculation to do in evaluating the Jacobian J at the fixed point. It's straightfortward to show that

$$\det J = \alpha u^\star \left[1 + \frac{db}{(u^\star + d)^2}\right] > 0 \ . \tag{1.122}$$

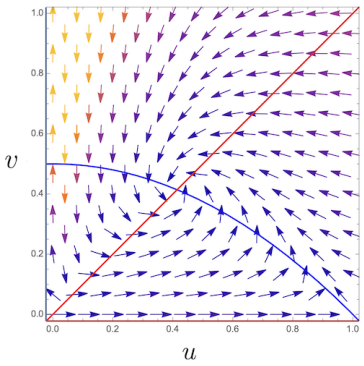
This tells us that the eigenvalues λ are either both positive, both negative or, come in complex conjugate pairs. Stability is determined by whether $\mathrm{Re}\,\lambda$ is positive or negative and we can see this by looking at the trace,

$$\mathrm{Tr}\, J = u^\star \left[\frac{bu^\star}{(u^\star + d)^2} - 1\right] - \alpha \ . \tag{1.123}$$

One way to read this is that for fixed b and d (and hence fixed $u^\star$), there is a critical value of α given by

$$\alpha_{\rm crit} = u^\star \left[\frac{bu^\star}{(u^\star + d)^2} - 1\right] \ . \tag{1.124}$$

For $\alpha > \alpha_{\rm crit}$ the fixed point is stable and for $\alpha < \alpha_{\rm crit}$ the fixed point is unstable.



The space of parameters is somewhat more interesting because for certain values of b and d, we have $\alpha_{\rm crit} < 0$ so the system is stable for all $\alpha > 0$. An example of the resulting flows, together with the nullclines, is shown in the figure to the right (plotted with $b = 2$ and $d = 1$ and $\alpha = 0.5$).

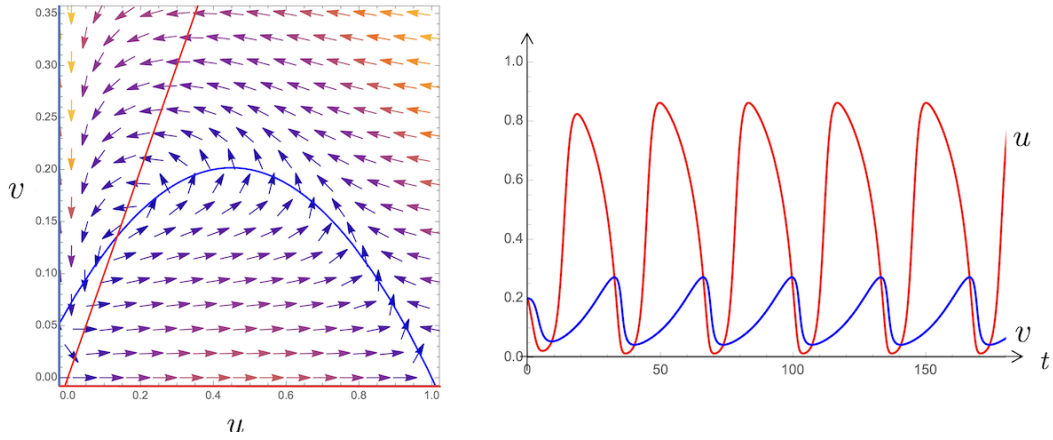


Figure 17. On the left: flows in the phase plane with an unstable fixed point, together with the nullclines. On the right, a numerical solution to the equations of motion with initial conditions that start close to the fixed point. Both of these plots were made with $b = 1.5$ and $d = 0.1$ which gives $\alpha_{\text{crit}} \approx 0.36$. We have then chosen $\alpha = 0.1$ to be in the unstable regime.

Things are more interesting when the fixed point is unstable. If we don't end up at the fixed point, then where do we end up? The answer is that the dynamics converges towards a *limit cycle*, meaning a closed trajectory in the phase plane that attracts nearby trajectories. The flows in the phase plane, together with a numerical solution to the equations of motion are shown in Figure 17. We see that this model again gives rise to the oscillations in populations that was characteristic of the original Lotka-Volterra model, but with one important difference: all initial conditions converge to the same cyclic behaviour.

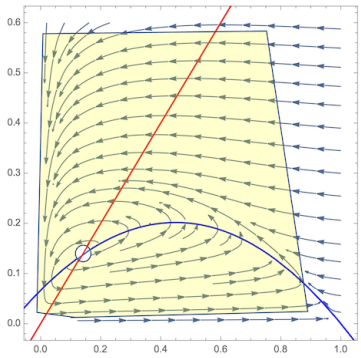
The Poincaré-Bendixson Theorem

It's natural to ask: is there a way to demonstrate the existence of a limit cycle without resorting to numerics? There is a standard technique that can be applied to do this which we sketch here.

The general idea is to find a closed, bounded region $S \subset \mathbb{R}^2$ for which all flows go *into* S, and none come out. More precisely, we pick an outward pointing normal $\mathbf{n}$ on the boundary ∂S of S and require that

$$\mathbf{n} \cdot \begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} < 0 \quad \text{everywhere on } \partial S \ . \tag{1.125}$$

This is telling us that, once inside the region S, you're trapped there forever. The *Poincaré-Bendixson theorem* then states (roughly) that if there are no fixed points in S then there will necessarily be a limit cycle.



Given this theorem, we can prove the existence of a limit cycle in our model if we can find such a region S. For our particular model, a sketch of the region S is shown shaded in the figure. It has two boundaries. An inner boundary consists of a small circle that excludes the fixed point and the flows are necessarily into S because the fixed point is unstable. The outer boundary is constructed so that (1.125) holds. This is straight forward to show for three of the four sides because the nullclines dictate the direction of the flow. The slightly tricky one is the left-hand boundary and one has to work a little harder to show that this too can be made to obey (1.125).

1.4.4 Competition

For our next example, we turn to a slightly different scenario, albeit one that is modelled by the same kind of equations. Rather than considering a populations of predators and prey, we will try to even the playing field a little. Instead, we consider two species competing for the same resources.

We call the populations N_1 and N_2 and describe their dynamics by the coupled logisticesque equations

$$\begin{aligned}\frac{dN_1}{dt} &= r_1 N_1\left(1-\frac{N_1}{K_1}-b_1\frac{N_2}{K_2}\right)\\ \frac{dN_2}{dt} &= r_2 N_2\left(1-\frac{N_2}{K_2}-b_2\frac{N_1}{K_1}\right)\end{aligned}\quad . \tag{1.126}$$

We take $r_1, r_2, b_1, b_1 > 0$.

There are various rescalings that we can do to simplify this equation. An obvious one is to remove the carrying capacities by defining $N_i' = N_i/K_i$. We do this, then rename $N_i' \to N_i$. There are further rescalings that we can do by absorbing one of the reproduction rates r_i into time, but they destroy the symmetry of the equations so we

choose not to. We then have

$$\begin{aligned}\frac{dN_1}{dt} &= r_1 N_1 \left(1 - N_1 - b_1 N_2\right) \\ \frac{dN_2}{dt} &= r_2 N_2 \left(1 - N_2 - b_2 N_1\right) \ .\end{aligned} \tag{1.127}$$

If we set either $N_1 = 0$ or $N_2 = 0$, then the dynamics of the other is described by the familiar logistic equation that we discussed in Section 1.1. Here, we would like to understand how the two species interact with each other.

As we'll see, this model is rather straightforward, certainly compared to some of the predator-prey generalisations that we discussed above. To start, we can look at the fixed points. There are four. One is $(N_1, N_2) = 0$ and is devoid of life. Two others have one of the species extinct,

$$(N_1, N_2) = (1, 0) \ , \quad (N_1, N_2) = (0, 1) \tag{1.128}$$

while the third equilibrium point has the two species coexisting:

$$(N_1, N_2) = \frac{1}{1 - b_1 b_2}(1 - b_1 \ , \ 1 - b_2) \ . \tag{1.129}$$

Populations have to be positive, so the last of these is a viable fixed point if either both $b_1, b_2 < 1$ or if both $b_1, b_2 > 1$. But if one of these coefficients is greater than one and the other less than one, then there is no coexistence.

What about stability? The Jacobian matrix takes the form

$$J = \begin{pmatrix} r_1(1 - 2N_1 - b_1 N_2) & -r_1 b_1 N_1 \\ -r_2 b_2 N_2 & r_2(1 - 2N_2 - b_2 N_1) \end{pmatrix} \ . \tag{1.130}$$

We need to evaluate this on each of the four fixed points (assuming that all four exist) and compute the eigenvalues λ. This is a straightforward exercise. We have:

$$(N_1, N_2) = (0, 0) \quad \Longrightarrow \quad \lambda = r_1, r_2 \ . \tag{1.131}$$

So this fixed point is unstable. The next two fixed points we get

$$\begin{aligned}(N_1, N_2) = (1, 0) \quad &\Longrightarrow \quad \lambda = -r_1 \ , \ r_2(1 - b_2) \\ (N_1, N_2) = (0, 1) \quad &\Longrightarrow \quad \lambda = -r_2 \ , \ r_1(1 - b_1) \ .\end{aligned} \tag{1.132}$$

Here we see a more interesting pattern: the first fixed point is stable if $b_2 > 1$ and the second is stable if $b_1 > 1$. If either of these coefficients is less than one, then the respective fixed point becomes a saddle. The flow in phase plane for two the cases $b_1 < 1 < b_2$ and $b_2 < 1 < b_1$ are shown in Figure 18.

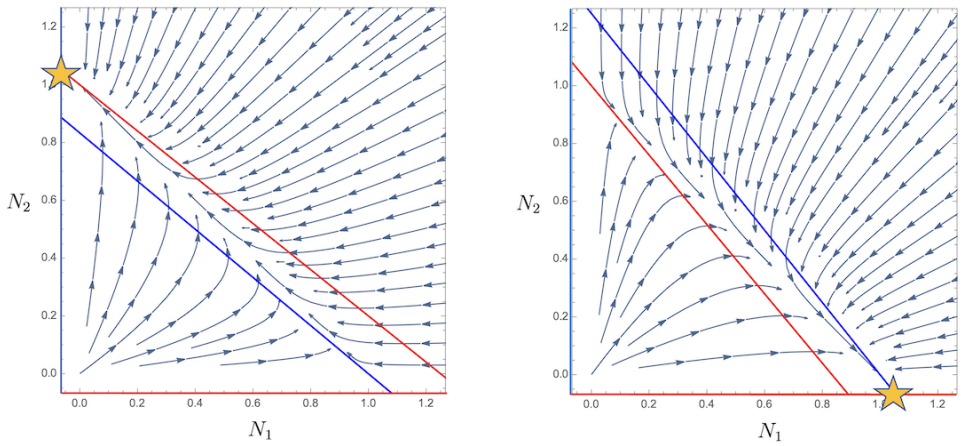


Figure 18. The phase plane flows for two cases $b_2 < 1 < b_1$ (on the left) and $b_1 < 1 < b_2$ (on the right) with the nullclines also shown. The fixed points are where a red and blue nullcline meet. The gold star denotes the stable fixed point.

Finally, for the fixed point corresponding to coexistence, the eigenvalues are more complicated. A little algebra gives

$$\begin{aligned}(N_1, N_2) &= \frac{1}{1-b_1 b_2}(1-b_1\,,\,1-b_2)\\ \implies \quad \lambda &= \frac{1}{2(1-b_1 b_2)}\Big[r_1(b_1-1)+r_2(b_2-1) \\ &\qquad \pm \sqrt{(r_1(b_1-1)+r_2(b_2-1))^2+4(b_1-1)(b_2-1)(b_1 b_2-1)}\Big]\,.\end{aligned} \tag{1.133}$$

You can check that the number under the square root is always positive. Moreover, both eigenvalues are negative provided that $b_1, b_2 < 1$, while one is negative and the other positive if $b_1, b_2 > 1$.

The flows in the phase plane are plotted for these two cases in Figure 19. We can now piece the full story together. The coefficient b_1 is a measure of the disruption that the second species has on the first. Similarly, b_2 measures the disruption of the first species on the second. If $b_1, b_2 < 1$, then there is a stable coexistence fixed point and the two species share the resources, with the population of each lower due to the presence of the other species.

Meanwhile if, say, $b_1 > 1$ and $b_2 < 1$ then this means that the second species is the more disruptive and it's game over for the first: the fixed point has $N_1 = 0$.

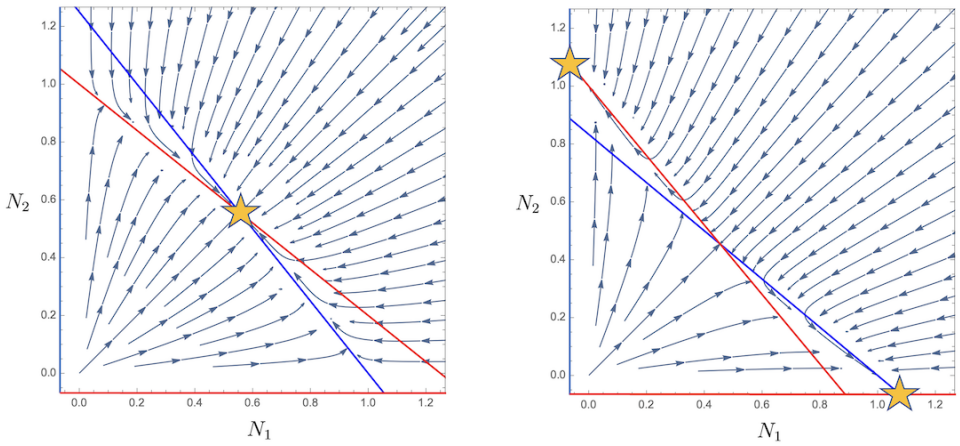


Figure 19. The phase plane flows for two two cases $b_1, b_2 < 1$ (on the left) and $b_1, b_2 > 1$ (on the right) with the nullclines also shown. The fixed points are where a red and blue nullcline meet and the stable fixed point(s) denoted by a gold star.

Finally, if both $b_1 > 1$ and $b_2 > 1$ then there are stable fixed points with one or the other species extinct. Which fixed point you end up in depends on the initial conditions. Increasing, say, b_1 increases the basin of attraction for the second species to become the winner.

There is a well known dictum in ecology known as the *principle of competitive exclusion*. It says that two species which compete for the same limited resources cannot coexist as one will have an advantage and will ultimately win out. This is indeed what the simple model above predicts whenever $b_1 > 1$ or $b_2 > 1$. But, as we've seen, the two species can certainly live in happy coexistence in this model when $b_1, b_2 < 1$. I'm not an ecologist but it seems strange to me to take something that is *not* predicted by equations and then elevate it to a "principle".

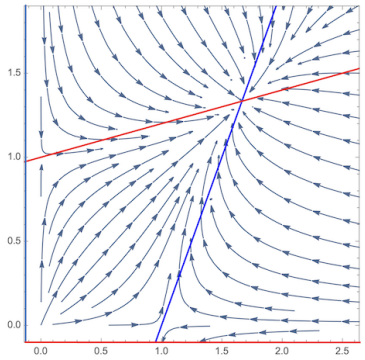
Be Kind

It may be that our two species get along in some symbiotic way. In this case, we can continue to describe their interactions through the equations (1.127), but we take $b_1, b_2 < 0$.

As before, there are four fixed points: the trivial one, the two with one species extinct (1.128), and the coexisting fixed point (1.129),

$$(N_1, N_2) = \frac{1}{1 - b_1 b_2}(1 - b_1\,,\, 1 - b_2)\ .$$

We see that this fixed point only exists if $b_1 b_2 < 1$. This fixed point is stable, and the resulting flows are shown in the figure. This time, the presence of each species enhances the population of the other.



As $b_1 b_2 \to 1$ the two species are *too* beneficial, and the populations run away to infinity! Indeed, for any $b_1 b_2 > 1$, the population runs off to infinity.

1.4.5 Dengue Fever

In this section we will develop a slightly different competition model, this one designed to describe attempts to mitigate a particularly unpleasant disease.

Dengue fever is transmitted by mosquitos in the tropics. It's pretty unpleasant and you don't want to catch it. Happily, help is at hand in the form of a bacteria known as *Wolbachia*. When mosquitos are infected with this bacteria, it blocks transmission of the dengue virus. This brings hope that by introducing Wolbachia-infected mosquitos into the wild, they may help reduce or eradicate the dengue virus.

We will assume that all mosquitos carry dengue, but those infected with Wolbachia are harmless. At this point there's something of a story to tell. Mosquitos can only pass Wolbachia onto their offspring. (Just kissing is not enough.) And we have the following complications:

- If a female is infected, all her eggs will be infected.
- An infected female lays fewer eggs than the uninfected.
- An infected female can mate with an infected male. But if an uninfected female mates with an infected male, there will be no viable eggs.
- Infected mosquitos don't live as long.

Our task is to translate these facts into equations. Indeed, the art of mathematical modelling is constructing equations from words. We introduce the following variables

$$\begin{aligned} x &= \text{number of uninfected females} \\ y &= \text{number of infected females} \ . \end{aligned} \tag{1.134}$$

We assume that the number of (un)infected males simply tracks the female population. Then we model the facts above by the following equations,

$$\begin{aligned}\frac{dx}{dt} &= x\left(r\frac{x}{x+y} - d - \epsilon(x+y)\right)\\ \frac{dy}{dt} &= y\big(\lambda r - \mu d - \epsilon(x+y)\big)\ .\end{aligned} \tag{1.135}$$

If we set either $x = 0$ or $y = 0$ then the other equation reduces to the logistic equation. The various terms in the equations have the following interpretation:

- The proportion of uninfected males is $x/(x+y)$. An uninfected female mating can only mate with an uninfected male and this increases the uninfected proportional to r.
- An infected female can mate with any male. This increases the rate of infected mosquitos by λr, with $\lambda < 1$ because infected females lay fewer eggs.
- Uninfected mosquitos die off at a rate proportional to d. Infected mosquitos die off at the quicker rate of μd with $\mu > 1$.
- There is a logistic-like competition for resources, giving rise to the $\epsilon(x+y)$ terms. This is the same for infected and uninfected.

With these equations in hand, we can now rescale to remove unnecessary constants, with $t \to rt$ and $x \to \epsilon x/r$ and $y \to \epsilon y/r$. We're left with the pair of equations,

$$\frac{dx}{dt} = x\left(\frac{x}{x+y} - \frac{d}{r} - (x+y)\right) \quad \text{and} \quad \frac{dy}{dt} = y\left(\lambda - \frac{\mu d}{r} - (x+y)\right)\ . \tag{1.136}$$

It's useful to collect the remaining constants together so we write

$$\frac{dx}{dt} = x\left(x_0 - \frac{y}{x+y} - (x+y)\right) \quad \text{and} \quad \frac{dy}{dt} = y\left(y_0 - (x+y)\right)\ . \tag{1.137}$$

Here

$$x_0 = 1 - \frac{d}{r} \quad \text{and} \quad y_0 = \lambda - \frac{\mu d}{r}\ . \tag{1.138}$$

(Note that the extra 1 in x_0 is what turns the $x/(x+y)$ in (1.136) into $y/(x+y)$ in (1.137)). The uninfected mosquitos breed more than they die (sadly), so $r > d$ and, correspondingly, $0 < x_0 < 1$. There is no a priori bound on y_0 but we will assume that $y_0 > 0$ so that populations of infected mosquitos also grow otherwise our goal of eradicating dengue is hopeless. This means that we have $0 < y_0 < x_0 < 1$.

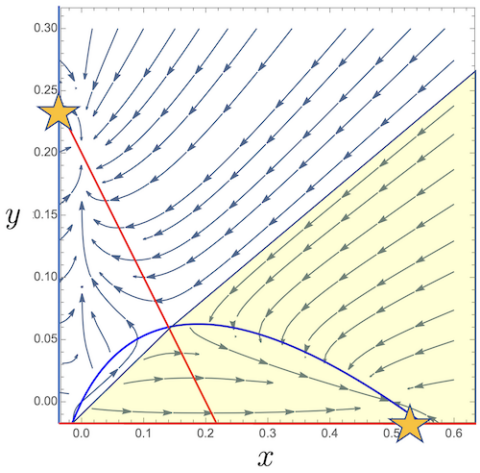


Figure 20. Flows in the phase plane have two stable fixed points shown with gold stars. These correspond to all mosquitos infected (on the y-axis, or all mosquitos uninfected (on the x-axis). Where you end up depends on the initial conditions. The basis of attraction to have uninfected mosquitos is shown shaded. This is plotted with $x_0 = 0.5$ and $y_0 = 0.2$.

Now we're in business. There are four fixed points: the trivial one $(x, y) = (0, 0)$, two where one of the populations is extinct, $(x, y) = (x_0, 0)$ and $(x, y) = (0, y_0)$ and one where both infected and uninfected coexist $(x^\star, y^\star)$, with

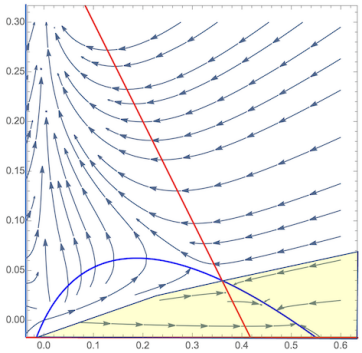
$$x^\star = y_0(1 - x_0 + y_0) \quad \text{and} \quad y^\star = y_0(x_0 - y_0) \ . \tag{1.139}$$

You can check that this is a saddle.

A plot of the flows in the phase plane, together with the nullclines, are shown in Figure 20. The two stable fixed points have either all mosquitos infected, or all uninfected. If you want to end up solely with infected mosquitos, then you need to introduce a sufficient number so that you sit in the basin of attraction of the fixed point $(x, y) = (0, y_0)$. This is the unshaded region in Figure 20.

Suppose that your goal is to eradicate dengue fever. You might wonder if it's possible to change the parameters in our equations to work to our benefit. We see that the end result ultimately depends on just two parameters: x_0 and y_0. There's not much that we can do about x_0, which depends on the birth and death rates of the uninfected mosquitos. In contrast, as shown in (1.138), y_0 depends on λ, which is the drop in egg production for infected mosquitos, and on μ which captures the reduced lifespan of infected mosquitos, and you might optimistically hope that these can be changed by altering the strain of Wolbachia. We could either try to increase λ, or to decrease μ.

From (1.138), we see that μ multiplies $d/r \ll 1$, suggesting that it might be more profitable to attempt to increase λ so that the number of eggs produced by infected mosquitos is closer to those produced by the uninfected. Either way, if we increase y_0 then the nullcline moves. An example is shown in the right, where the phase plane flow is plotted for $x_0 = 0.5$ and $y_0 = 0.4$ (as opposed to $y_0 = 0.2$ in Figure 20). Again the basin of attraction for uninfected mosquitos is shaded and is visibly smaller than that shown in Figure 20.



1.4.6 The Large Diversity Limit

Many systems in nature involve more than two interacting species. At the extreme end, the human gut contains somewhere close to 1000 different kinds of microbes. Some marine ecosystems contain a similar number of species. Can we fruitfully model such complicated systems?

To start, we can proceed as before. If we have M different species, each with population $N_i(t)$ with $i = 1, \ldots, M$, then we can describe the dynamics by the set of equations

$$\frac{dN_i}{dt} = f_i(N) \ . \tag{1.140}$$

In general, this dynamics is likely to be complicated. But there's one question that has a rather nice answer: is it possible for multiple species to coexist?

Rephrasing this, is it possible to find a stable fixed point $f(N^\star) = 0$ when the number of interacting species is very large, so $M \gg 1$? Expanding about the fixed point, the dynamics is

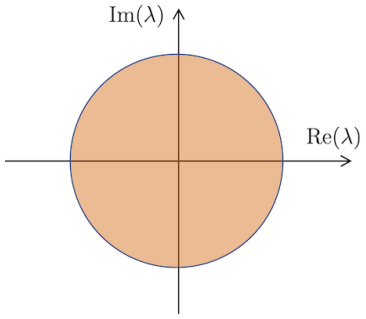
$$\frac{dN_i}{dt} = J_{ij}(N_j - N_j^\star) \quad \text{with} \quad J_{ij} = \frac{\partial f_i(N^\star)}{\partial N_j} \ . \tag{1.141}$$

This fixed point is stable if all the eigenvalues of J_{ij} have negative real part.

Now, it is of course always *possible* to find functions f_i such that all the eigenvalues of J_{ij} all have negative real part. But is it likely?

There is an important branch of mathematics, known as *random matrix theory*, that allows us to answer questions of this kind. Suppose that you have a matrix where each element is chosen randomly, and independently, from some probability distribution. We then ask: what is the probability distribution of the eigenvalues? Rather wonderfully, it turns out that many features of the eigenvalue distribution are independent of the original choice of probability distribution that you choose for the individual elements.

We won't derive the key properties of random matrices here, but instead just state the key result. Suppose that you sample the individual elements of the matrix from a Gaussian distribution with zero mean and variance σ^2. Then, in the limit $M \to \infty$, the eigenvalues λ are uniformly distribution on the complex plane in a disc of radius $r = \sqrt{M\sigma^2}$, as shown in the figure.



How can we use this for our question of population stability? Here's a simplified model. We will take the Jacobian matrix in (1.141) to have the form

$$J = -\mathbb{1} + J_{\text{random}} \ . \tag{1.142}$$

That is: in the absence of interactions, the fixed point is stable and all eigenvalues have been scaled to $\lambda = -1$. But we then add to this random interactions with other species, captured by the additional term J_{random}. Invoking the result from random matrix theory, we see that, in the $M \to \infty$ limit, the eigenvalue with the largest real part has

$$\text{Re}(\lambda_{\text{max}}) = -1 + \sqrt{M\sigma^2} \ . \tag{1.143}$$

In other words, a stable ecosystem where the interactions between different species have variance σ^2 can support at most

$$M_{\text{max}} = \frac{1}{\sigma^2} \tag{1.144}$$

different species. This is known as *May's stability criterion*.

The model above involves many simplifying assumptions but, nonetheless, the idea that there is a trade-off between diversity and stability has been influential among ecologists.

1.5 Epidemiology

2020 was a weird year. For many of us, it's a blur of Zoom calls and government mandated daily walks and lots of talk about the number R_0 and why it's too big. The purpose of this section is to re-live this experience, this time with equations.

1.5.1 The SIR Model

The classic epidemic model is named after its three variables,

$$\begin{aligned} S &= \text{number of people who are susceptible to the disease} \\ I &= \text{number of people who are infected} \\ R &= \text{number who have recovered or are dead.} \end{aligned}$$

One of the lessons that came out of the 2020 pandemic is that there is a tension between the goals of scientists and those of politicians. This model highlights the tension pretty clearly. If you're a mathematical modeller then, at least at this basic level, it doesn't matter if someone recovers or dies from the disease: either way, they no longer contribute to its spread. But I've been told that there are some politicians who appreciate the distinction between these two outcomes. If you want to phrase things more delicately, you could say $R =$"removed".

The equations that model the spread of the disease are:

$$\frac{dS}{dt} = -\beta IS \ , \quad \frac{dI}{dt} = +\beta IS - \nu I \ , \quad \frac{dR}{dt} = \nu I \tag{1.145}$$

with positive constants $\beta, \nu > 0$.

These equations have an intuitive underpinning. The susceptible turn into infected at a rate that is proportional to IS. The idea is that this product captures the interaction between the two groups, in way that is analogous to the NP terms that we met in the predator-prey equations. Meanwhile, the infected either recover or die at a constant rate. Whichever path they take, they are removed from the process and no longer contribute to the dynamics.

The first thing to note is that the total number of people $N = S+I+R$ is a constant. That's tautologically true if you still count people who die! Moreover, we don't really care about the recovered/dead at all as they don't feed back into the other two. That means that we can focus just on the first two equations in (1.145) and then reconstruct $R(t) = N - S(t) - I(t)$.

The first question to ask is: are we in trouble? Is the number of infected people going to increase? From the second equation in (1.145), we see that the answer doesn't depend on how many are already infected, as long as $I \neq 0$: it depends only on how many susceptible there are:

$$\dot{I}(0) > 0 \quad \Longleftrightarrow \quad \beta S(0) > \nu \ . \tag{1.146}$$

Here ν is the rate at which individuals recover, and β is the rate at which each susceptible is infected. The epidemic starts if the rate at which total population is infected is faster than the rate at which they recover.

Typically, a disease will start with just a few infected among a large population and we can take $S(0) \approx N$. This motivates us to define the *reproductive ratio*, pronounced "R-naught" or "R-zero",

$$R_0 = \frac{\beta N}{\nu} \ . \tag{1.147}$$

Here $1/\nu$ is the transmission period; β is the transmission rate per contact; and N is the initial number of contacts. The epidemic grows if $R_0 > 1$. For the original strain of COVID-19, R_0 was somewhere between 2 and 5. For polio, $R_0 \approx 4 - 6$; for mumps $R_0 \approx 10 - 12$. For measles, $R_0 \approx 16 - 18$.

Suppose that we are well prepared and vaccinate a fraction p of the population before the disease hits. Then the number of susceptibles is reduced to $(1-p)N$, lowering the effective reproductive ratio to $(1-p)R_0$. Or, said differently, if the unvaccinated population has a given R_0, then we need to vaccinate a fraction $p > (R_0 - 1)/R_0$ to stop the disease spreading. Crucially, we don't need to vaccinate everyone, just enough to reach herd immunity which protects the whole community.

This is a good point to pause and comment on a more philosophical aspect of mathematical modelling. All the mathematical models that we write down in this course are, to put it bluntly, wrong. They are all, at best, caricatures of the underlying reality. Most likely they omit many important details. All of which begs the question: why should we trust them? This is particularly important when it comes to putting in place mechanisms to counter disease. One answer to this question is that you can search for things in the models that are robust: things that don't change when you add extra layers to the model and include further details. Crucially, it turns out that in epidemiology, one such robust quantity is the threshold for vaccination that we have computed above. And, indeed, it works well in practice.

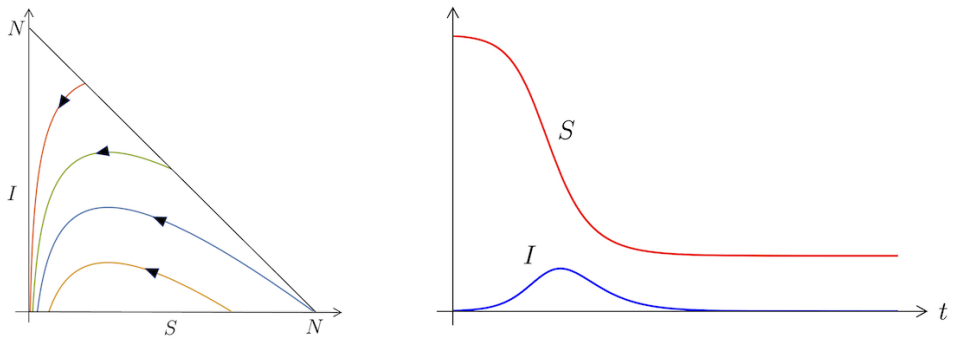


Figure 21. On the left, flows of the SIR model in the phase plane. On the right, a typical time evolution of the susceptible and infected population.

We can solve for the trajectories in the phase plane in much the same way as the Lotka-Volterra model. We view $I = I(S)$ and, dividing the first two equations in (1.145), we're left with

$$\frac{dI}{dS} = \frac{\nu}{\beta S} - 1 = \frac{N}{R_0 S} - 1 \ . \tag{1.148}$$

This is easily integrated to give

$$I(S) = \frac{N}{R_0} \log S - S + c \tag{1.149}$$

with c the integration constant. Curves for different choices of c are plotted on the left of Figure 21, adorned with arrows which show that the number of susceptibles are always decreasing. The curves are plotted in the triangular region to reflect the fact that $S + R < N$, the total population. An epidemic that starts with just a few infected is described by the blue curve which intersects the corner of the triangle.

To get the time data, you need to solve the equations numerically. A typical example is shown on the right of Figure 21. The number of infected grows and then shrinks to zero, while the number of susceptibles falls to a constant value S_∞.

From both graphs, we see that the number of infected reaches a maximum value. From (1.148), we can easily see that the maximum number of infected occurs when $S = N/R_0$. An intuitive way to think about this comes from defining the *effective reproductive ratio* $R_{\text{eff}} = \beta S/\nu$. The epidemic starts to recede when $R_{\text{eff}} = 1$.

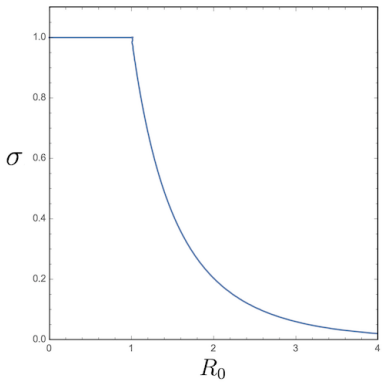


Figure 22. The lucky ones, as a function of R_0.

There's one piece of good news hiding in this analysis: we're not all going to die. Or even become infected. The epidemic burns itself out and some number S_∞ never catch the disease. This occurs when $I(S) = 0$. From the left plot of Figure 21, we see that there are two values of S when this happens: we call these S_0 and S_∞. Setting $I(S) = 0$ in (1.149), we have

$$\frac{N}{R_0}\log S_0 - S_0 = \frac{N}{R_0}\log S_\infty - S_\infty \ . \tag{1.150}$$

It's natural to assume that we start with very few infected and lots of susceptibles, so $I_0 \ll N$ and $S_0 \approx N$. We will denote the lucky fraction of the population as σ, so that

$$S_\infty = \sigma N \ . \tag{1.151}$$

From (1.150), this lucky fraction satisfies

$$\log\sigma = R_0(\sigma - 1) \ . \tag{1.152}$$

The result is plotted in Figure 22. For $R_0 < 1$ there's no need to panic: essentially 100% of the population escapes unscathed. But we can see that things get more worrisome as R_0 increases. In particular, for R_0 large we have $\sigma \approx e^{-R_0}$. So by the time you get to $R_0 \approx 10$, that's more or less everyone that's going to get infected.

A Quick History of the SIR Equations

The SIR equations are usually attributed to a famous and influential 1927 paper by William Kermack and Anderson McKendrick, titled "*A Contribution to the Mathematical Theory of Epidemics*". This is viewed as the beginning of the study of epidemiology.

However, they were not the first. That honour belongs to Ronald Ross and Hilda Hudson who constructed the equations ten years before[5]. Ross is famous in medical circles for understanding the role of mosquitos in the transmission or malaria, research for which he won the Nobel prize for Medicine in 1902. As he got older, he became increasingly aware that the kind of questions he wanted to address required mathematics, and he gave many talks boring doctors with equations they couldn't understand. The relevant variables were introduced by Ross alone in a 1916 paper with the title "*A Priori Pathometry*", his preferred term for what we now call epidemiology. However, Ross soon realised that going further was beyond his mathematical abilities and so he turned to the Royal Society, requesting "a government grant for a lady mathematician to assist" him, the "lady" added because in 1916 male mathematicians were elsewhere. He joined forces with Hilda Hudson, a geometer who had been an undergraduate at Newnham College, Cambridge (at a time when, famously, women could not be awarded degrees), and this resulted in two further papers, the second of which contains the SIR equations.

One of the difficulties in writing a paper that starts a new field is that there are very few experts around who can referee it. The papers of Ross and Hudson were sent to the eminent astronomer, Sir Arthur Eddington. Referee reports were rather more blunt back then. Eddington writes "the theory is, of course, rather dull, and the mathematics elementary, but it is desirable that some such treatment of statistics be made." It's interesting that even this very first referee report appreciated the point of the subject: the mathematics isn't deep, the value lies in the application.

1.5.2 Just When You Thought It Was Safe...

I caught covid last month. So it's certainly not gone away. Of course, what happened is that my immunity wore off and I moved from the recovered R group back into the susceptible S group.

It's straightforward to incorporate this, or other variations, into the SIR model. Here we consider a different feature: people die (of something other than the disease in question); and people are born. And all of those new born babies sit straight in the susceptible camp. To capture this, we add various terms to our original SIR equations

[5]You can hear more of this story in the engaging YouTube talk by June Barrow-Green.

(1.145), which now becomes

$$\begin{aligned}\frac{dS}{dt} &= -\beta IS + bN - \mu S \\ \frac{dI}{dt} &= \beta IS - \nu I - \mu I \\ \frac{dR}{dt} &= \nu I - \mu R \ .\end{aligned} \tag{1.153}$$

Here $N = S + I + R$ as before; b is the birth rate, and μ the death rate. We've assumed that this disease doesn't kill anyone, an assumption that manifests itself in the fact that members of all groups contribute to the births. Moreover, we assume that the death rate is constant regardless of whether you've had the disease or not. We see that $dN/dt = (b - \mu)N$ like the Malthus model of Section 1.1.

To make things simple, let's assume (perhaps unrealistically) that the population is stable, with $b = \mu$ so that N is constant. Our reproductive ratio again comes from looking at the $\dot{I}$ equation when $S = N$; it is

$$R_0 = \frac{\beta N}{\nu + \mu} \ . \tag{1.154}$$

At first glance, it looks like that's good news: dying reduces R_0! But some thought suggests that the effect is minimal: for any disease in which you recover over a period of time that's much shorter than your lifetime, we necessarily have $\nu \gg \mu$, and so $R_0 \approx \beta N/\nu$ as before.

Since N is constant, the first two equations again decouple from the third. This time, the influx of births means that there is a fixed point $(S^\star, I^\star)$ with

$$S^\star = \frac{\nu + \mu}{\beta} = \frac{N}{R_0} \quad \text{and} \quad I^\star = \mu \frac{N - S^\star}{\beta S^\star} = \frac{\mu}{\beta}(R_0 - 1) \ . \tag{1.155}$$

By now, we know what we're doing. We determine the stability by computing the Jacobian at the fixed point

$$J = \begin{pmatrix} -\mu R_0 & -(\nu + \mu) \\ \mu(R_0 - 1) & 0 \end{pmatrix} \ . \tag{1.156}$$

The eigenvalues are complex and given by

$$\lambda = -\frac{\mu R_0}{2} \pm \frac{1}{2}\sqrt{\mu^2(R_0 - 2)^2 - 4\mu\nu(R_0 - 1)} \ . \tag{1.157}$$

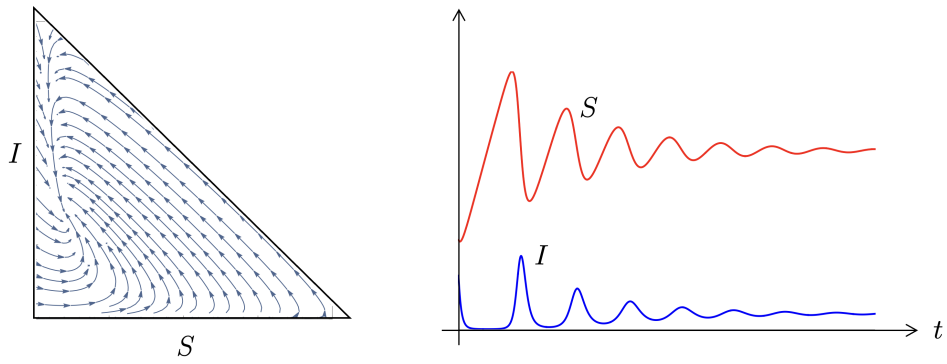


Figure 23. On the left, flows of the modified SIR model in the phase plane. On the right, a typical time evolution of the susceptible and infected population.

If we're in the realistic situation where recovery is not measured in a timescale less than decades then $\mu \ll \nu$ and we have

$$\lambda \approx -\frac{\mu R_0}{2} \pm i\omega \quad \text{with} \quad \omega = \sqrt{\mu\nu(R_0 - 1)} \; . \tag{1.158}$$

These eigenvalues tell us that we have a stable fixed point, which trajectories spiral towards with a period The resulting dynamics is plotted, both in the phase plane, and as a function of time, in Figure 23

This suggests that we will observe transient oscillations on a time period that is the geometric mean of the lifetime and infection recovery time.

$$T = \frac{2\pi}{\omega} = \frac{2\pi}{\sqrt{\mu\nu(R_0 - 1)}} \approx \frac{1}{\sqrt{\mu\nu}} \; . \tag{1.159}$$

For example, measles has $R_0 \approx 20$ with a recovery rate of about 12 days. So putting $1/\nu \approx 12$ days and setting a human lifetime at $1/\mu \approx 70$ years, we expect to see oscillations with a period of $T \approx 2.2$ years. In fact, there's an extra complication here, namely schools. There is a delta-function injection of a susceptible population every September when kids start school. This is sufficient to keep the measles outbreaks occurring, but with an oscillation that is roughly 2 years rather than 1 year. Data of measles outbreaks in the UK over the past decade is shown in Figure 24.

1.5.3 Zombie Apocalypse

The mathematics that we've developed in this section is crucially important to model and, ideally, contain different diseases. Here we apply our newfound skill to address one of the most concerning issues facing humanity: a zombie apocalypse.

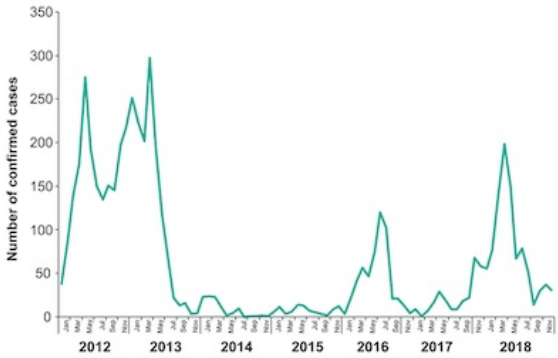


Figure 24. Data from the UK government on measles outbreaks over the past decade.

We model the zombie apocalypse by a variant of the SIR model, known as the SZR model[6]. Once again, the model is named after its three variables

$$S = \text{number of humans}$$
$$Z = \text{number of zombies}$$
$$R = \text{number of dead who are not yet reanimated}$$

The relevant equations are

$$\frac{dS}{dt} = -\beta SZ \ , \quad \frac{dZ}{dt} = \beta SZ + \zeta R - \alpha SZ \ , \quad \frac{dR}{dt} = \alpha SZ - \zeta R \ . \tag{1.160}$$

The biological interpretation of the various terms is as follows. When humans and zombies meet, one of two things can happen. The first option is that the zombie bites the human, turning him or her into a companion zombie. This happens with rate β. The second is that the human succeeds in killing the zombie . Perhaps this requires a beheading, or a stake through the heart – just pick your favourite movie scenario. Either way, the zombie removal happens with rate α.

However, the dead don't stay dead for long. They are resurrected as zombies with rate ζ. This, as we shall see, gives something of an unfair advantage to the zombies.

As with the SIR model, the total number $N = S + Z + R$ is constant. This means that we can eliminate the equation for R, leaving us with

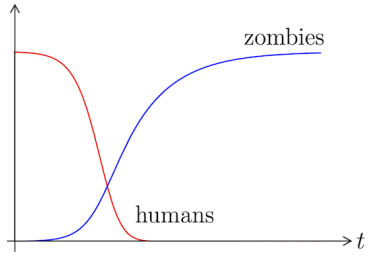
$$\frac{dS}{dt} = -\beta SZ \quad \text{and} \quad \frac{dZ}{dt} = (\beta - \alpha)SZ + \zeta(N - S - Z) \ . \tag{1.161}$$

[6]This was first introduced in the paper "When zombies attack!: Mathematical modelling of an outbreak of zombie infection" by P. Munz, I. Hudea Joe Imad, and R.J. Smith?, published in Infectious Disease Modelling Research Progress, 2009.

These equations have two fixed points. The first is $S = N$ and $Z = 0$. That's where we want to be. The other is $Z = N$ and $S = 0$. That's a less appealing outcome. To figure out where we end up, we can compute the Jacobian at the fixed point

$$J = \begin{pmatrix} -\beta Z & -\beta S \\ (\beta - \alpha)Z - \zeta & (\beta - \alpha_S - \zeta \end{pmatrix} . \tag{1.162}$$

You can check that at the fixed point with $S = N$, we have $\det J = 0 - \zeta\beta N < 0$. This immediately tells us that this fixed point is a saddle. Meanwhile, the fixed point with $Z = N$ where the zombies triumph has $\det J = \beta N J > 0$ and $\mathrm{Tr}\, J = -\beta N - \zeta < 0$ and so is stable. That's not good news.



A numerical plot of the outcome is shown in the figure to the right. Zombies triumph. Humans lose. This happens regardless of the choice of parameters α, β, and ζ. You might think there could be a some small chance that you could battle it out against the odds. There isn't[7].

1.6 Chemical Reactions

The kinds of dynamical system that we've been developing so far have many applications that removed from population dynamics. Here we give an example in which we treat the population of chemicals in reactions.

1.6.1 The Law of Mass Action

The simplest chemical reaction is that two chemicals, A and B, form a product C,

$$A + B \xrightarrow{\mathrm{k}} C . \tag{1.163}$$

Here k is the rate constant of the reaction. It's straightforward to model this in terms of equations. We let A, B and C denote the concentration of the chemical and write the equations

$$\frac{dA}{dt} = -kAB , \quad \frac{dB}{dt} = -kAB , \quad \frac{dC}{dt} = +kAB . \tag{1.164}$$

[7]In 2025, I set this problem as an exam question. But in Cambridge all exam questions must be vetted by a committee and they rejected my question on the grounds that it wasn't biologically realistic. I suggested that the same set of equations could be used to describe the discussion that takes place among Cambridge examiners, with $S =$ "those with a sense of humour", $Z =$ "those with no sense of humour", and $R =$ "those who don't know". This too was rejected.

The fact that the time derivatives are proportional to the product of the concentrations $A \times B$ is known as the *law of mass action* in chemistry. It's the same conceptual idea that we've invoked in both the predator-prey models and the epidemiology models but without giving it a name. However, the law of mass action is on much firmer footing in the context of chemical reactions where it can be derived using the kind of technology that we introduced in the lectures on Statistical Physics. In contrast, in population dynamics and epidemiology, the idea that we get the right phenomenology by multiplying together the interacting variables is an assumption.

It's straightforward to solve these equations. First we note that $A + C = A_0$ is constant, as is $B + C = B_0$. We can then use this to write the last equation in (1.164) purely in terms of C,

$$\frac{dC}{dt} = k(A_0 - C)(B_0 - C) \ . \tag{1.165}$$

This has the solution

$$C(t) = A_0 B_0 \frac{1 - e^{(A_0 - B_0)kt}}{B_0 - A_0 e^{(A_0 - B_0)kt}} \tag{1.166}$$

where we've chosen the integration constant so that $C(0) = 0$. This expression has the nice property that $\lim_{t\to\infty} C(t) = \min(A_0, B_0)$, reflecting the fact that the reaction stops whenever one of the chemicals is exhausted.

What is done can sometimes be undone. It may be that the reaction can also go in the opposite direction, albeit at a different rate:

$$A + B \underset{\mathrm{k_2}}{\overset{\mathrm{k_1}}{\rightleftharpoons}} C \ . \tag{1.167}$$

Chemists often use the notation k_{-1} for the inverse reaction rate k_2. In this case, the third equation in (1.164) becomes

$$\frac{dC}{dt} = k_1 AB - k_2 C \tag{1.168}$$

with k_1 and k_2 the two reactions rates. We still have $A + C = A_0$ constant and $B + C = B_0$ constant, so this equation becomes

$$\frac{dC}{dt} = k_1(A_0 - C)(B_0 - C) - k_2 C \tag{1.169}$$

Now there is a fixed point, when the right-hand side vanishes. We define

$$\kappa = \frac{k_2}{k_1} \ . \tag{1.170}$$

From (1.168), we see that, despite the similarity in their names, k_2 and k_1 actually have different dimensions, so that κ has the dimensions of concentration. Indeed, the fixed point obeys the quadratic

$$C^2 - (A_0 + B_0 + \kappa)C + A_0 B_0 = 0 \tag{1.171}$$

together with the additional requirement that $0 < C < \min(A_0, B_0)$.

Lotka-Volterra Revisited

Here's an interesting class of reactions: we have four different chemicals, A, B, C and D which react as

$$A + B \xrightarrow{k_1} 2B \quad \text{and} \quad B + C \xrightarrow{k_2} 2C \quad \text{and} \quad C \xrightarrow{k_3} D \tag{1.172}$$

We take the reaction rates to be k_1, k_2, and k_3 respectively. The dynamics is then given by

$$\begin{aligned} \frac{dA}{dt} &= -k_1 AB \\ \frac{dB}{dt} &= +k_1 AB - k_2 BC \qquad (1.173) \\ \frac{dC}{dt} &= +k_2 BC - k_3 C \qquad (1.174) \\ \frac{dD}{dt} &= +k_3 C \ . \end{aligned}$$

Suppose that we now intervene and ensure that there is a constant concentration of the chemical A. Then we can replace this variable with its constant values in (1.173) and this equation, together with (1.174), reduces to the Lotka-Volterra equations that we studied in detail in Section 1.4. Our earlier results tell us that the concentration of the chemicals B and C will oscillate over time. Indeed, this was the context in which Alfred Lotka first wrote down these equations.

Stoichiometry

Consider a reaction in which m molecules of type A combine with n molecules of type B to produce a molecule of type C. If the reaction is reversible, then we write

$$mA + nB \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C \ . \tag{1.175}$$

This is modelled by the equations

$$\begin{aligned}\frac{dA}{dt} &= -mk_1A^mB^n + mk_2C \\ \frac{dB}{dt} &= -nk_1A^mB^n + nk_2C \\ \frac{dC}{dt} &= k_1A^mB^n - k_2C \ .\end{aligned} \tag{1.176}$$

Here the powers A^m and B^n are a consequence of the law of mass action. Meanwhile, the pre-factors of m and n are designed to ensure that $A + mC$ and $B + nC$ are both constant, as they should be.

1.6.2 Michaelis-Menten Enzyme Kinetics

Some biochemical reactions can be yawningly slow. They get a helping hand from enzymes. These are proteins which give the chemicals a small hug, increasing the reaction rate.

The original set of chemicals are called the *substrate* S. The end result is called the *product* P. The reaction is helped on its way by a population of enzymes E. When the enzymes bind with the substrate, they form a combination known as a *complex* C. The reaction takes place through the *Michaelis-Menten reaction*:

$$S + E \underset{\mathrm{k_2}}{\overset{\mathrm{k_1}}{\rightleftharpoons}} C \xrightarrow{\mathrm{k_3}} E + P \ . \tag{1.177}$$

Note that the first reaction, with the enzymes binding to the substrate, can go both ways. But the second, where the enzyme releases the final product, goes only in one direction. The equations governing this reaction are

$$\begin{aligned}\frac{dS}{dt} &= -k_1SE + k_2C \\ \frac{dE}{dt} &= -k_1SE + (k_2 + k_3)C \\ \frac{dC}{dt} &= +k_1SE - (k_2 + k_3)C \\ \frac{dP}{dt} &= +k_3C \ .\end{aligned} \tag{1.178}$$

We will assume that $S(0) = S_0$ and $E(0) = E_0$, while $C(0) = P(0) = 0$ as initial conditions. We usually assume that $S_0 \gg E_0$, meaning that the limit on the reaction rate is set by the number of enzymes rather than the lack of substrate.

There are two conservation laws within these equations. These are

$$E + C = E_0 \quad \text{and} \quad S + C + P = S_0 \ . \tag{1.179}$$

The first reflects the fact that the enzymes aren't used up: they are either empty and denoted as E or full and denoted as C. The second tells us that the end result is to change substrate into product without losing any.

We can use the first of these to eliminate the enzyme concentration E, and focus on equations just for S and C. (Note that the product P is just a dumping ground and doesn't affect the other variables.) We have

$$\begin{aligned} \frac{dS}{dt} &= -k_1 E_0 S + (k_1 S + k_2) C \\ \frac{dC}{dt} &= k_1 E_0 S - (k_1 S + k_2 + k_3) C \ . \end{aligned} \tag{1.180}$$

We can clean these up a little by rescaling variables. We write $s = S/S_0$ and $c = C/E_0$. This is the proportion of substrate and the proportion of occupied enzymes respectively, with $s(0) = 1$ and $c(0) = 0$. If we also rescale time by defining $\tau = k_1 E_0 t$ then the equations become

$$\begin{aligned} \frac{ds}{d\tau} &= -s + (s + \mu - \lambda) c \\ \frac{dc}{d\tau} &= \frac{1}{\epsilon}\big(s - (s + \mu) c\big) \end{aligned} \tag{1.181}$$

where we've now got just three constants

$$\lambda = \frac{k_3}{k_1 S_0} \quad \text{and} \quad \mu = \frac{k_2 + k_3}{k_1 S_0} \quad \text{and} \quad \epsilon = \frac{E_0}{S_0}. \tag{1.182}$$

It's straightforward to solve these equations numerically and the result is shown in Figure 25 where, crucially, we have taken $\epsilon \ll 1$. The result is rather striking: all trajectories head immediately to a common curve, which they then follow down to the origin.

The key to understanding this feature is to appreciate that, with $\epsilon \ll 1$, the enzyme dynamics $c(t)$ happens on a much faster time scale than the substrate dynamics $s(t)$. The system therefore relaxes quickly to the $\dot{c} = 0$ curve, given by

$$c = \frac{s}{s + \mu} \ . \tag{1.183}$$

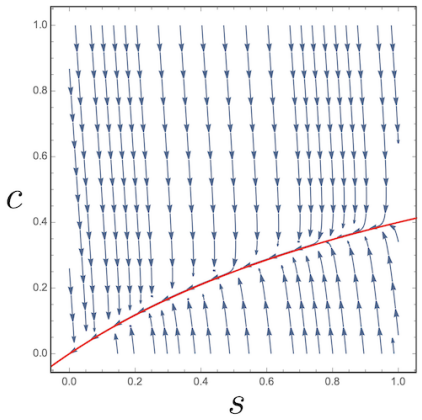


Figure 25. The flows in the (s, c) phase plane, plotted with $\epsilon = 0.1$. We have also taken $\mu = 1.5$ and $\lambda = 0.5$.

That's the curve shown in red in Figure 25. There is then a much slower precession along this curve which is captured by substituting (1.183) into the equation for $\dot{s}$,

$$\frac{ds}{d\tau} = -\frac{\lambda s}{s+\mu} \ . \tag{1.184}$$

This tells us how the substrate is depleted towards $s = 0$. As the amount of substrate changes, so too does the concentration of complexes, quickly adapting to remain in the equilibrium given by (1.183).

Biochemists are often interested in the reaction velocity R, given by

$$R = \frac{dP}{dt} = k_2 C = k_1 k_2 E_0 \, \frac{S}{k_1 S + k_2 + k_3} \ . \tag{1.185}$$

This is the *Michaelis-Menten equation*, often written in terms of the maximum reaction velocity $V = k_2 E_0$ and the so called *Michaelis-Menten constant* $\kappa = (k_2 + k_3)/k_1$. You can extract these constants by plotting the linear graph of R^{-1} against S^{-1}.

1.6.3 Asymptotic Behaviour

It's possible to make analytic progress by solving the equation perturbatively in the small parameter ϵ. However, naively setting $\epsilon = 0$ changes the character of the differential equation so we need to tread a little carefully.

Let's start by being naive. If we set $\epsilon = 0$ in (1.181) then we get the differential equation (1.184). We'll call the function that solves this $\tilde{s}(\tau)$: it is easy to check that

this is given implicitly by

$$\tilde{s} + \mu \log \tilde{s} = -\lambda\tau + \text{constant} \ . \tag{1.186}$$

We can fix the constant by requiring the initial condition $\tilde{s}(\tau = 0) = 1$, which tells us that

$$\tilde{s} + \mu \log \tilde{s} = -\lambda\tau + 1 \ . \tag{1.187}$$

This should be viewed as an approximation to the late time behaviour of the system: it holds after the fast dynamics has happened and we can impose the constraint (1.183).

We can also get an approximation for the short time behaviour. To do this, we rescale the time coordinate and write

$$\tau = \epsilon T \ . \tag{1.188}$$

The two equations of motion (1.181) become

$$\frac{ds}{dT} = \epsilon\big(-s + (s + \mu - \lambda)c\big) \quad \text{and} \quad \frac{dc}{dT} = s - (s + \mu)c \ . \tag{1.189}$$

Now the ϵ is sitting on the right-hand side of the equation, rather than the left-hand side, and we can do a standard perturbative expansion, writing

$$s(T) = s_0(T) + \epsilon s_1(T) + \ldots \quad \text{and} \quad c(T) = c_0(T) + \epsilon c_1(T) + \ldots \ . \tag{1.190}$$

This is a short time expansion. To leading order in ϵ, we have

$$\frac{ds_0}{dT} = 0 \quad \text{and} \quad \frac{dc_0}{dT} = s_0(T) - (s_0(T) + \mu)c_0(T) \tag{1.191}$$

subject to the initial conditions $s(0) = 1$ and $c(0) = 0$. The first of these equations tells us that $s_0(T) = 1$, while the second is then solved by

$$c_0(T) = \frac{1}{1+\mu}\left[1 - e^{-(1+\mu)T}\right] \quad \Longrightarrow \quad c_0(\tau) = \frac{1}{1+\mu}\left[1 - e^{-(1+\mu)\tau/\epsilon}\right] \ . \tag{1.192}$$

We see that $c_0(\tau)$ has the characteristic $e^{-\tau/\epsilon}$ behaviour at short times. Famously, this kind of function vanishes faster than any polynomial as $\epsilon \to 0$.

For both $s(\tau)$ and $c(\tau)$, we now have two approximations to the solutions. The short time solutions are $s_0(\tau)$ (which, admittedly, is trivially constant) and $c_0(\tau)$. The long time solution, after the fast dynamics is exhausted, is $\tilde{s}(\tau)$ given implicitly by (1.187) and, from (1.183), $\tilde{c}(\tau) = \tilde{s}(\tau)/(\tilde{s}(\tau) + \mu)$. In the world of matched asymptotic expansions, $s_0(\tau)$ and $c_0(\tau)$ are called *inner solutions* while $\tilde{s}(\tau)$ and $\tilde{c}(\tau)$ are called *outer solutions*. Our next task is to patch these together.

This is straightforward for $s(\tau)$ because the short time solution is just a constant, $s_0(\tau) = 1$ and, indeed, this matches the short time behaviour $\lim_{\tau\to 0} \tilde{s}(\tau) = 1$. You can check that $\tilde{s}(\tau)$ is a good approximation to the true solution when $\epsilon \ll 1$.

For $c(\tau)$, we have to be a little more careful. We have

$$\lim_{\tau\to\infty} c_0(\tau) = \lim_{\tau\to 0} \tilde{c}(\tau) = \frac{1}{1+\mu} \ . \tag{1.193}$$

This is telling us that they will patch nicely together. We do this by adding the two solutions and subtracting their common piece, a process known as the *uniform approximation*,

$$c_{\rm uniform}(\tau) = c_0(\tau) + \tilde{c}(\tau) - \frac{1}{1+\mu} \ . \tag{1.194}$$

This function is still known only implicitly because $\tilde{s}(\tau)$ is given by the relation (1.187). Nonetheless, you can check that $c_{\rm uniform}(\tau)$ gives good agreement to the numerical solution for $c(\tau)$ when $\epsilon \ll 1$.

1.7 Neuron Excitations

One of the great advances of mathematical biophysics was a system of equations, first written down in 1952 by Hodgkin and Huxley, to describe the way that neurons fire in the brain. Neurons have a long sticky-out bit called an axon, which conducts pulses of electricity known as *action potentials* and these mediate signalling from one neuron to the next. These action potentials form from some external chemical prompt, but in a way such that a reasonably small prompt is amplified to an unambiguous pulse of electricity. The question is: how does this happen?

Rather than looking at the propagation of the signal along the axon, Huxley and Hodgkin instead considered the easier situation in which the signal is constant in space and focussed on how the signal fires in time. (We will rectify this omission and examine how the signal propagates in space in Section 3.2.3.) They wrote down a rather complicated set of equations, in which there are three kinds of electric current, carried by potassium ions, sodium ions, and other stuff that they called the "leakage current". These currents have to pass through gates and the probability that these gates are open or closed depends in some complicated way on the voltage. We won't go into the detail of the Hodgkin-Huxley model, but the end result is four equations in four variables which they were able to solve numerically.

The solutions matched well the experimental data from studies of squid giant axons. (As opposed to giant squid axons; the squid in question was rather small by squid standards, the axon rather large by axon standards.) The giant axon is the part of the squid's nervous system that is responsible for escaping danger by squirting out water like a jet propulsion system; the large size of the axon means that it has lower resistivity and so signals propagate faster, allowing quicker reaction times. It also means that it's big enough to stick an electrode down there and measure the voltage. You can read about the Huxley-Hodgkin model in many places, including a fairly decent summary on Wikipedia.

As an aside: Hodgkin and Huxley were successive masters of Trinity College Cambridge in the 70s and 80s. Hodgkin was long before my time, but I did meet Andrew Huxley a couple of years before he died. I sat down next to him at lunch, embarrassingly ignorant of who he was even though his portrait was hanging on the wall in front of me. He asked me what I did and, when I told him I was a physicist, there was a lovely pause before he said: "I met J.J. Thomson once". (This is the same J.J. Thomson who discovered the electron in 1897.)

The key feature in the Hodgkin-Huxley model is that a small input gives rise to a big, spiked response. There are many other biological situations where similar behaviour is seen, including heart muscle cells and the blooms of plankton. The general class of models are called *excitable systems* and in this section we explore the phenomenology of the simplest.

1.7.1 FitzHugh-Nagumo Model

The FitzHugh-Nagumo model is the following pair of coupled differential equations,

$$\begin{aligned} \frac{du}{dt} &= \frac{1}{\epsilon}\left(u + v - \frac{1}{3}u^3 - z(t)\right) \\ \frac{dv}{dt} &= -(u - a + bv) \ . \end{aligned} \tag{1.195}$$

In the context of signal propagation by neurons, $u(t)$ represents the voltage difference across a membrane, while $v(t)$ represents the ease with which sodium and potassium ions can cross the membrane.

The model depends on three positive constants, a, b and ϵ. The first pair sit in the regions

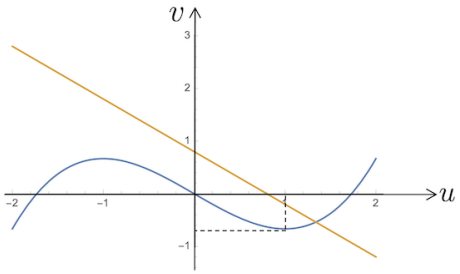
$$0 < b \leq 1 \ , \quad 1 - \frac{2b}{3} < a < 1 \ . \tag{1.196}$$

The remaining constant ϵ is arbitrary but, as the name suggests, we will soon take $\epsilon \ll 1$ which ensures that the dynamics of $u(t)$ is fast, while that of $v(t)$ is slow. In addition, there is an *input function* $z(t)$ which we get to specify and it acts as a forcing term in the first equation.

We start by setting $z(t) = 0$ and look at the dynamics of the equations. There is a single fixed point $(u^\star, v^\star)$ which is the intersection of the two nullclines

$$v = u\left(\frac{1}{3}u^2 - 1\right) = \frac{a-u}{b} \ . \tag{1.197}$$

We have to solve a cubic which is a little awkward, but we can extract the essential features if we plot the two graphs together, as shown in the figure. First note that the slope of the cubic at the origin is -1, while the slope of the line is $-1/b$ which, by the first inequality in (1.196), is necessarily steeper. This tells us that there is only one intersection point.



Second, the cubic has roots at $u = 0$ and $u = \pm\sqrt{3}$. The minimum is at $u = 1$ where we have $v = -2/3$. The slightly weird inequality in (1.196) ensures that the linear graph sits above the cubic at its minimum, which means that they must meet later. In other words, we have

$$u^\star > 1 \ . \tag{1.198}$$

Happily, this is all the information we need to determine the stability of the fixed point. The Jacobian is

$$J = \begin{pmatrix} (1-u^{\star\,2})/\epsilon & 1/\epsilon \\ -1 & -b \end{pmatrix} \ . \tag{1.199}$$

We have $\mathrm{Tr}\, J < 0$ and $\det J > 0$, ensuring that the real part of both eigenvalues is negative and the fixed point is stable.

Now we assume that $\epsilon \ll 1$ so that the dynamics separates into a fast piece and a slow piece. The $u(t)$ variable is the fast piece, quickly tending almost horizontally towards the cubic $v = u^3/3 - u$. (In the context of neurons, this happens on a time scale of about 1 ms.) Then the $v(t)$ dynamics is slower. The result is plotted numerically in Figure 26. Note that the flow is unusual. If the flow hits the cubic to the right of the

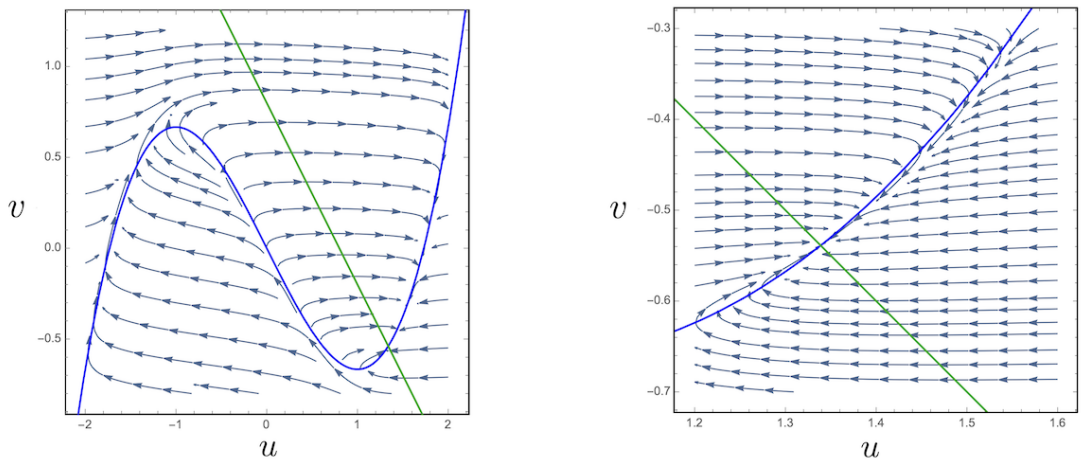


Figure 26. The phase plane dynamics of the FitzHugh-Nagumo model, plotted with $a = 0.8$, $b = 1$ and $\epsilon = 0.2$. Because ϵ is reasonably small, the dynamics separates into a fast mode, which heads towards the cubic, and a much slower mode which traces the curve of the cubic towards the fixed point where the nullclines meet. A close up of the flow near the fixed point is shown on the right.

local minimum at $u = 1$, then the dynamics simply follows the cubic down towards the fixed point. If, in contrast, it hits the cubic to the left of that local minimum then it takes the long way round to the fixed point, moving up, before sweeping right as shown in Figure 26. The fact that some paths take these long deviations will be important in what follows.

Response to an External Input

Now we explore the role of the forcing term $z(t)$ in the equations (1.195). Suppose that by time $t = 0$, the system has happily settled down at the fixed point. Now we perturb the system by turning on $z(t)$, with the simple profile

$$z(t) = \begin{cases} 0 & \text{for } t < 0 \\ V_0 & \text{for } t > 0 \end{cases} \tag{1.200}$$

with V_0 constant.

The effect of the perturbation $V_0 > 0$ is to shift the cubic nullcline upwards, while leaving the linear nullcline alone. This means that the fixed point also shifts up and to the left. What happens next depends crucially on whether the new fixed point is stable or unstable. We saw previously that the fixed point is stable provided that $u^\star > 1$

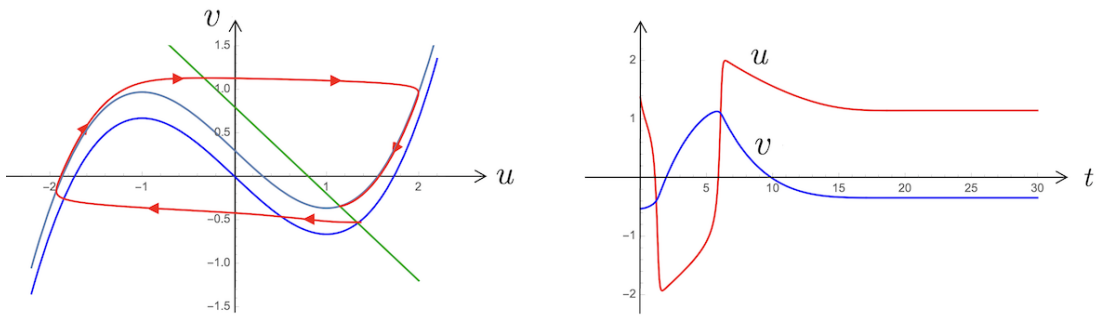


Figure 27. On the left: the phase plane motion, shown in red, after the perturbation. The system starts at the original fixed point, where the lower cubic curve intersects the straight line. It then goes on a long detour in phase space before ending up at the new fixed point, where the upper cubic curve intersects the straight line. On the right: the resulting time dependence of $u(t)$ and $v(t)$. These plots are made with the same values as Figure 26, together with $V_0 = 0.3$.

which ensures that it sits to the right of the local minimum in the cubic. It's simple to check that this persists for our perturbed system provided that $V_0 < V_{\rm crit}$ with

$$V_{\rm crit} = \frac{2}{3} + \frac{a-1}{b} \ . \tag{1.201}$$

We start by looking at the situation with $V_0 < V_{\rm crit}$. If the perturbation V_0 is very small, so that the old fixed point sits above the minimum of the new cubic, then nothing dramatic happens: the system simply flows directly to the new fixed point.

But a slightly larger V_0, still with $V_0 < V_{\rm crit}$, has a much more dramatic effect. If the old fixed point sits below the minimum of the new cubic then the system takes a large diversion to reach the new fixed point. This is shown on the left of Figure 27 which depicts both the original and the new nullclines, together with the trajectory in the phase space shown in red. This trajectory starts at the original fixed point, where the lower cubic curve intersects the line, and ends at the new fixed point at the intersection of the upper cubic curve. Although the two fixed points are close to each in phase space, the motion takes a huge detour to get from one to the other. The result is that a small perturbation of the original system can lead to a large pulse in the variables $u(t)$ and $v(t)$, before they settle down to equilibrium values that are close to their original values. This is shown on the right of Figure 27. In the context of neurons, this pulse in the voltage is what causes the neuron to fire.

The situation for $V_0 > V_{\rm crit}$ is shown in Figure 28. Once again, the system takes a long detour to get from the old fixed point to the new. But now, once it arrives at the

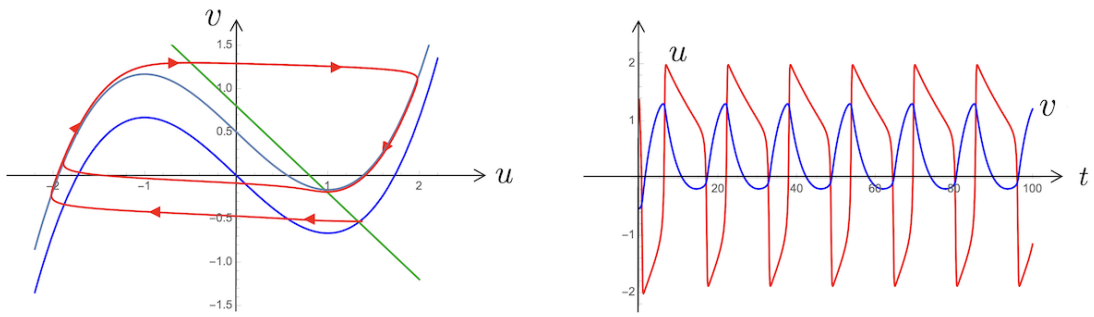


Figure 28. On the left: the phase plane motion, shown in red, after the perturbation. Now the new fixed point is unstable and the system performs indefinite loops. On the right: the resulting time dependence of $u(t)$ and $v(t)$. These plots are made with the same values as Figure 26, together with $V_0 = 0.5$.

new fixed point, it finds that it's not particularly welcome as the fixed point is unstable and it gets pushed away. It is now condemned to loop forever on the limit cycle. The resulting dynamics for $u(t)$ and $v(t)$ are shown on the right of Figure 28 and exhibit an infinite series of pulses.

For both $V_0 < V_{\rm crit}$ and for $V_0 > V_{\rm crit}$, we see how a small change in the equation can lead to a large change in the resulting dynamics, either to initiate a single, transient pulse or to send the system into a new limit cycle.

2 Discrete Time

Sometimes it is more natural to model populations and other biological quantities using discrete, rather than continuous, time. A species may give birth only in one particular season. Or you may be interested in how some mutation evolves from one generation to the next. Alternatively, it may be that you really do want to think of time as continuous but you're solving an equation numerically and want to know the pitfalls that could arise when making time discrete.

Whatever the motivation, in this section we will study the behaviour of systems of a single variable x_n where the index $n \in \mathbb{Z}$ plays the role of time. In the simplest case, the variable evolves through some function

$$x_{n+1} = f(x_n) \ . \tag{2.1}$$

We will learn that one dimensional systems of this kind can be significantly richer than their differential equation counterparts.

2.1 Linear Examples

We start our discussion by looking at a couple of simple examples. These won't bring any substantially new dynamics to the table beyond what we've seen in continuous systems. But they will allow us to get used to some features of discrete time dynamics.

2.1.1 Hello Poppy

Poppies are annual flowers, living for just one year. Their seeds sit in the ground over winter. Some fraction of them germinate the following year, some the year after that, and some not at all. We would like to model this mathematically.

We'll introduce the following variable

- Let x_n be the number of plants in season n.
- Let γ be the number of seeds produced by each plant.
- Let σ be the probability that a seed germinates after one year.
- Let τ be the probability that, having failed to germinate the first year, a seed is successful the next.

Now we're in business and can write down the equation that describes the poppy population: it is

$$x_n = \sigma\gamma x_{n-1} + \tau(1-\sigma)\gamma x_{n-2} \ . \tag{2.2}$$

What kind of solutions should we be looking for? We could look for a steady state, but generically the only one is $x^\star = 0$. And, because the equation is linear, any small perturbation around this just brings us back to (2.2).

Instead, motivated by the form of the equation (2.2), we look for solutions of the form

$$x_n = p^n \tag{2.3}$$

for some p. If we substitute this into (2.2), we get a quadratic

$$p^2 = \sigma\gamma p + \tau(1-\sigma)\gamma \ . \tag{2.4}$$

This has roots

$$p_\pm = \frac{\sigma\gamma}{2} \pm \frac{1}{2}\sqrt{\sigma^2\gamma^2 + 4\tau(1-\sigma)\gamma} \ . \tag{2.5}$$

We have $p_- < 0 < p_+$ and $|p_+| > |p_-|$. The general solution takes the form

$$x_n = Ap_-^n + Bp_+^n \ . \tag{2.6}$$

Because $|p_+| > |p_-|$, the second term will dominate at large n. The question is: does the population grow or shrink?

This depends on the size of p_+. If $p_+ > 1$ then the population grows over time; if $p_+ < 1$ then it shrinks. The critical case is

$$p_+ = 1 \quad \Longrightarrow \quad \gamma[\sigma + \tau(1-\sigma)] = 1 \ . \tag{2.7}$$

This makes sense: the quantity $[\sigma + \tau(1-\sigma)]$ is the probability that a seed germinates, either in the first year or the second. So the quantity $\gamma[\sigma + \tau(1-\sigma)]$ is the average number of offspring that a given plant produces. If this number is greater than one, then $p_+ > 1$ and the poppies flourish. If this number is less than one, then $p_+ < 1$ and it's goodbye poppy.

2.1.2 Breathe Again

In Section 1.2.3, we introduced a model of breathing in which the volume of the breath, V, depends on the concentration of CO_2 in your blood. Because breaths are things you can count, it makes sense to construct such model using discrete time.

We previously introduced the breathing model to illustrate delay equations. We'll keep this feature, with the volume of the n^{th} breath V_n determined by the concentration of CO_2 k breaths previously,

$$V_{n+1} = \alpha C_{n-k} \ . \tag{2.8}$$

We will then model the change in the CO_2 level by the equation

$$C_{n+1} - C_n = M - \beta V_{n+1} \ . \tag{2.9}$$

Here $M, \alpha, \beta > 0$. Note that this model is not a straightforward discretisation of our previous differential equation (1.52). Indeed, here we have a rather simple linear system, contrasting with the more complicated non-linear delay differential equation that we previously studied.

We can eliminate the volume V_n entirely, and focus just on the CO_2 concentration,

$$C_{n+1} = M + C_n - \alpha\beta C_{n-k} \ . \tag{2.10}$$

There is a steady-state solution given by $C_n = C^\star$ with

$$C^\star = \frac{M}{\alpha\beta} \ . \tag{2.11}$$

Now we can look at perturbations away from this steady state. We will be particularly interested in how the qualitative behaviour of the solutions depends on the delay parameter k.

$\underline{k = 0:}$ To kick things off, let's analyse (2.10) when $k = 0$ so there is no delay and the volume of breath depends on the present concentration of CO_2. We perturb around the fixed point and write $C_n = C^\star + \epsilon_n$ where, as the name suggests $\epsilon_n \ll 1$. Then we have

$$\epsilon_{n+1} = (1 - \alpha\beta)\epsilon_n \ . \tag{2.12}$$

We see that the steady state is stable provided that $\alpha\beta < 2$ and is unstable for $\alpha\beta > 2$.

$\underline{k = 1:}$ Now what happens if we introduce the smallest possible delay $k = 1$? Perturbing around the fixed point, the equation (2.10) becomes

$$\epsilon_{n+1} = \epsilon_n - \alpha\beta\epsilon_{n-1} \ . \tag{2.13}$$

This time we're going to look for solutions of the form $\epsilon_n = p^n$ for some p. The equation above then becomes a quadratic:

$$p^2 - p + \alpha\beta = 0 \quad \Longrightarrow \quad p = p_\pm = \frac{1}{2}\left(-1 \pm \sqrt{1 - 4\alpha\beta}\right) \ . \tag{2.14}$$

For $\alpha\beta < \frac{1}{4}$, both p_+ and p_- are real and, moreover $|p_\pm| < 1$. Then we have the general solution

$$\epsilon_n = Ap_+^n + Bp_-^n \tag{2.15}$$

which decays as $\epsilon \to 0$ as $n \to 0$. We see that, once again, the fixed point $C^\star$ is stable for $\alpha\beta$ small enough.

For $\alpha\beta > \frac{1}{4}$, the roots $p_\pm$ become complex. This means that the system now oscillates about the fixed point but doesn't otherwise change our approach. We still want to know if $|p_\pm|$ is less than one, and hence the fixed point is stable, or greater than one and hence unstable. We have

$$|p_\pm|^2 = \frac{1}{4} + \left(\alpha\beta - \frac{1}{4}\right) = \alpha\beta \ . \tag{2.16}$$

So we learn that, complex oscillations aside, the system is stable if $\alpha\beta < 1$ and unstable if $\alpha\beta > 1$. The upshot of this is that the delay reduces the range of $\alpha\beta$ over which the system is stable.

We can look more closely what happens at $\alpha\beta = 1$ and the system becomes unstable. Here we have

$$p_\pm = \frac{1}{2}\left(-1 \pm \sqrt{3}i\right) = e^{\pm\pi i/3} \ . \tag{2.17}$$

We see that the system has 6-fold periodicity at this point, with $p_\pm^6 = 1$.

2.2 The Logistic Map

The fun with discrete maps really gets going when we look at non-linear maps. A great deal of all that's interesting and surprising about these maps can be found lurking inside the deceptively simple example

$$x_{n+1} = f(x_n) = rx_n(1 - x_n) \ . \tag{2.18}$$

This is the *logistic map*. Understanding the mysteries of the logistic map will occupy us for the rest of this section.

We will take our parameter to lie in the range $x_n \in [0, 1]$. The logistic map keep us within this range provided that the parameter r is bounded by

$$0 \leq r \leq 4 \ . \tag{2.19}$$

If $r = 4$, the maximum value of the logistic map (at $x_n = 1/2$) gives $x_{n+1} = 1$. Much of our interest will be focussed on how the dynamics of the logistic map changes as we very r.

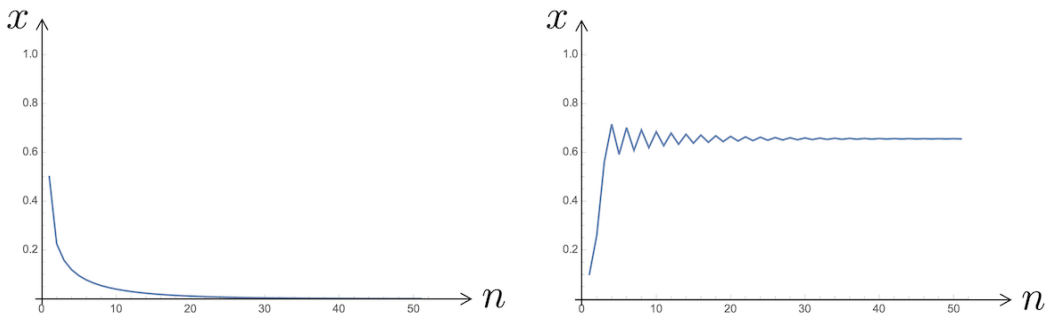


Figure 29. Two plots of x_n. On the left, we have taken $r = 0.9$ and $x_0 = 0.5$. We see that the map quickly tends to the origin. On the right, we have taken $r = 2.9$, close to the end of the window of stability, and $x_0 = 0.1$. The map now tends to $x^\star$, oscillating about it as it goes.

The logistic map a discrete version of the logistic equation that we studied in section 1.1. However, after nondimensionalisation, the logistic equation has *no* free parameters. That's not true of the logistic map (2.18), which crucially depends on the parameter r which can't be absorbed into rescaling time because that's now a discrete variable. Indeed, if you discretise the logistic equation and rescale then you will end up with the logistic map (2.18) with $r = 1 + \epsilon^2$ where $\epsilon \ll 1$ is related to the small time interval that you choose. We will soon see that, for this value of r, the logistic map does indeed recover the qualitative behaviour of the logistic equation. But, for other values of r, wildly different things can happen.

2.2.1 The Fixed Points

The logistic map has two fixed points: one at $x = 0$ and the other at

$$x^\star = 1 - \frac{1}{r} \ . \tag{2.20}$$

This is only a fixed point for $r > 1$.

What is the stability of these two fixed points? First we can look near the origin where we write $x_n = \epsilon_n \ll 1$. We then have

$$x_{\rm n+1} = r\epsilon_n(1 - \epsilon_n) = r\epsilon_n + \mathcal{O}(\epsilon_n^2) \ . \tag{2.21}$$

So the origin is a stable fixed point for $r < 1$, as each successive iteration takes us closer to it. It is unstable for $r > 1$.

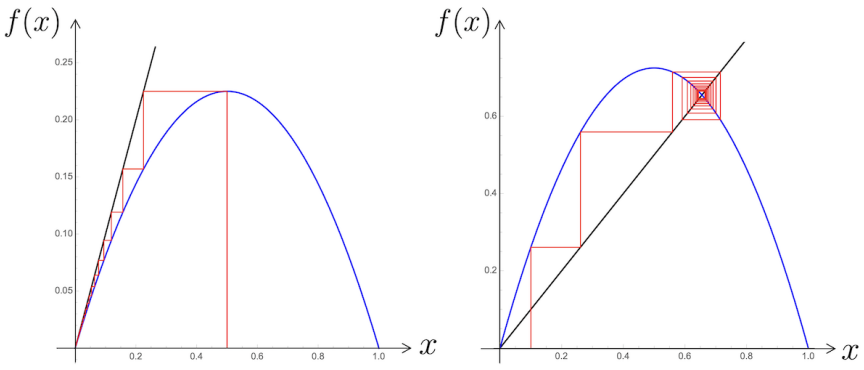


Figure 30. Cobweb diagrams, made with the same parameters as Figure 29. On the left, $r < 1$ and on the right $1 < r < 3$.

Meanwhile, for the non-trivial fixed point (2.20), we write $x_n = x^\star + \epsilon_n$ and look at

$$x_{n+1} = x^\star + f'(x^\star)\epsilon_n + \mathcal{O}(\epsilon_n^2) \quad \text{with} \quad f'(x^\star) = 2 - r \ . \tag{2.22}$$

This fixed point is stable for $1 < r < 3$ and is unstable for $r > 3$. Moreover, in the stable regime $2 < r < 3$, we see that $f'(x^\star)$ is negative and so successive terms will jump either side of $x^\star$, while honing in to the fixed point.

We can see this analysis bearing out by simply plotting successive iterations of the logistic map. This is shown in Figure 29 for $r < 1$ and for $2 < r < 3$, where we see that the results converge to the origin and to $x^\star$ respectively.

Cobweb Diagrams

There's a nice graphical way to see how the map behaves. We plot $y = f(x)$ together with the line $y = x$. Start at some value of x and $y = 0$ and move up until you hit the graph $f(x)$. Then it's simple to convince yourself that successive iterations of the map are implemented by bouncing at right angles between the graph and the line.

For $r < 1$, the line sits above the graph $f(x)$ and the bouncing takes you down to the origin. This is shown in on the left of Figure 30. For $r > 1$, the line leaves the origin below the graph $f(x)$ and intersects it again at $x^\star$. For $1 < r < 3$, the bouncing zooms in to the fixed point as shown on the right of Figure 30.

2.2.2 Bifurcation

We've understood the behaviour of the logistic map for $r < 3$. But what happens for $r > 3$ when the fixed point $x^\star$ is unstable? We can easily answer this by looking at

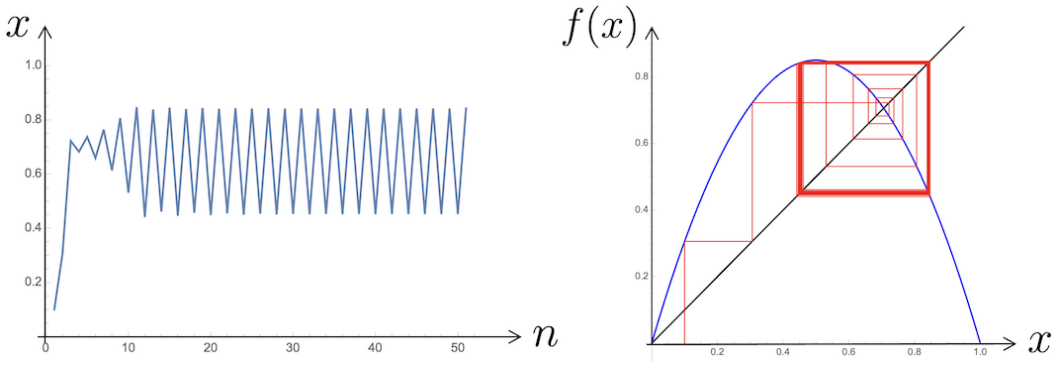


Figure 31. The logistic map for $r = 3.4$ rapidly reaches a stable 2-cycle, bouncing between two different points.

some numerics. Both the behaviour of x_n and the cobweb diagram are shown in Figure 31 for $r = 3.4$. We see that the map starts by honing in on the fixed point $x^\star$, but quickly realises that this is unstable and settles down to a periodic pattern, bouncing between two different points. On the cobweb diagram, the trajectory repeatedly traces out the rectangle as shown in the figure.

We will call a trajectory that bounces between p different points a *p-cycle*. What we have on our hands in Figure 31 is a 2-cycle.

We can understand this behaviour analytically. We look at the map

$$\begin{aligned} f^2(x) &= f(f(x)) \\ &= rf(x)(1-f(x)) \\ &= r^2x(1-x)(x-rx(1-x)) \ . \end{aligned} \tag{2.23}$$

This has fixed points

$$x = f^2(x) \quad \Longrightarrow \quad x(1-r+rx)(1+r-r(1+r)x+r^2x^2) = 0. \tag{2.24}$$

The first two factors give us our previous fixed points, $x = 0$ and $x = x^\star$. Now, however, we see that f_2 has two further fixed points, where the second factor vanishes. These are given by

$$x_\pm = \frac{1}{2r}\left(1+r \pm \sqrt{(r-3)(r+1)}\right) \ . \tag{2.25}$$

We see that these fixed points are only real when $r > 3$, which is what we wanted. These are the two points that the system bounces between, as evident in the numerical solution of Figure 31.

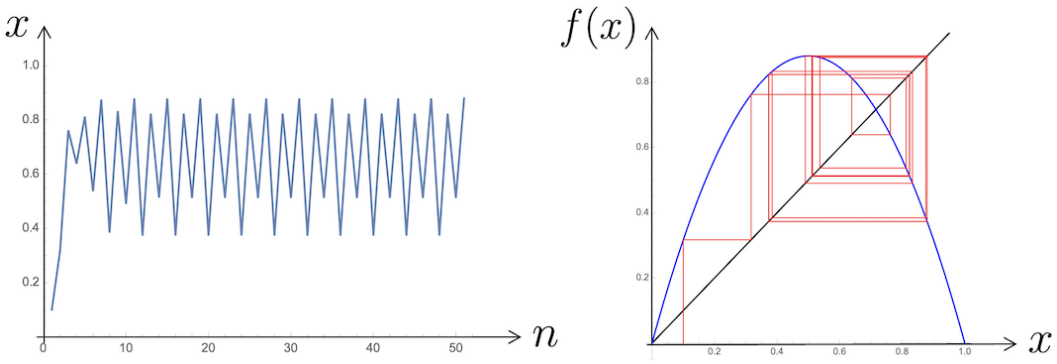


Figure 32. The logistic map for $r = 3.52$ reaches a stable 4-cycle, bouncing between four different points.

Next we should ask: are these fixed points of $f^2(x)$ stable? Here, the standard perturbation analysis prompts us to look at the derivative $f^{2\prime}(x_\pm)$ and see if its modulus is bigger than one (in which case the fixed points are unstable) or less than one (in which case they are stable). We have

$$\frac{df^2}{dx} = \frac{d}{dx} f(f(x)) = f'(f(x))\, f'(x) \ . \tag{2.26}$$

If we put $x = x_+$ then $f(x_+) = x_-$ and so, rather cutely, we have

$$\left.\frac{df^2}{dx}\right|_{x_+} = f'(x_-)f'(x_+) = r^2(1-2x_-)(1-2x_+) = -r^2 + 2r + 4 \ . \tag{2.27}$$

You can check that, for r slightly greater than 3, the modulus of the right-hand side is less than one and so the fixed points of $f^2(x)$ are stable. This is the behaviour that we see in the figure. But the fixed points turn unstable when

$$-r^2 + 2r + 4 < -1 \quad \Longrightarrow r > 1 + \sqrt{6} \approx 3.449 \tag{2.28}$$

where we've taken the positive root of the quadratic.

The net result is that the 2-cycle behaviour seen in Figure 31 only holds for the regime $3 < r \lesssim 3.45$. Which begs the question: what happens for greater values of r?

Again, we can plot things numerically to get an idea. This is shown in Figure 32 for $r = 3.52$ where we see that now the value of x bounces periodically between four fixed points. Again, we could get an analytic handle on this by studying the fixed points of $f^4(x)$ although this will now be a polynomial of order 8 and somewhat harder to analyse. (Actually, after factoring out the fixed points of $f(x)$ and $f^2(x)$, you're left with a quartic to deal with.)

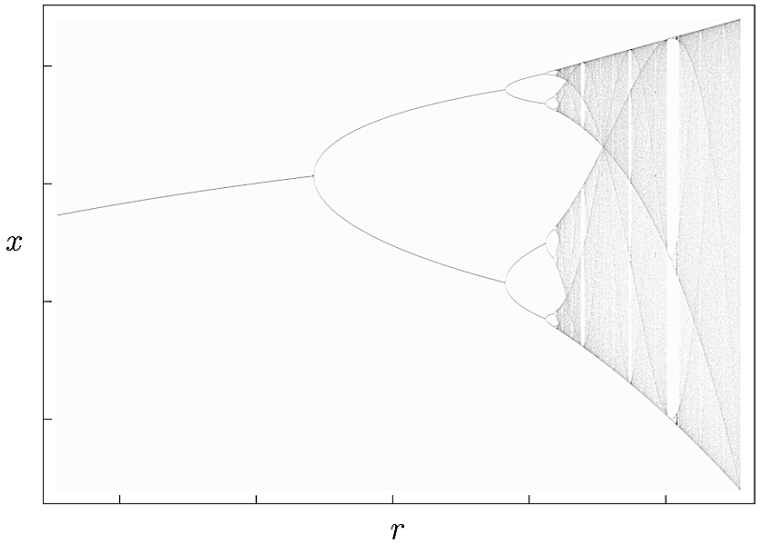


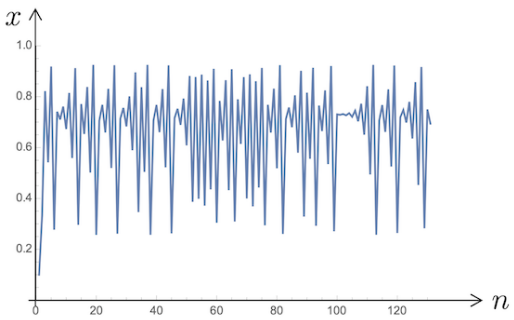
Figure 33. The long term behaviour of the logistic map as a function of r. We see the single stable fixed point for $1 < r < 3$ turning into the stable 2-cycle, then the stable 4-cycle, then the stable 2^n-cycle. For $r > r_\infty$, there is (typically, but not always) no stable cycles. But occasionally, out of the murk, windows of order appear.

This story now repeats. The 4-cycle seen in Figure 32 does not persist forever but, it turns out, becomes unstable for $r \gtrsim 3.55$, at which point we find ourselves with a stable 8-cycle. And so on. As r increases, we find ourselves with the period of the cycle constantly doubling in size.

2.2.3 And Then... Chaos

The period of cycles keep doubling and the range over which this happens gets shorter and shorter. Until we get to approximately

$$r > r_\infty = 3.5699 \ . \tag{2.29}$$



At this point, all hell breaks loose. For most (but not all!) values of r above this value, there is no discernible pattern in the long term behaviour, which bounces around seemingly at random. The plot for $r = 3.7$ is shown in the figure to the right.

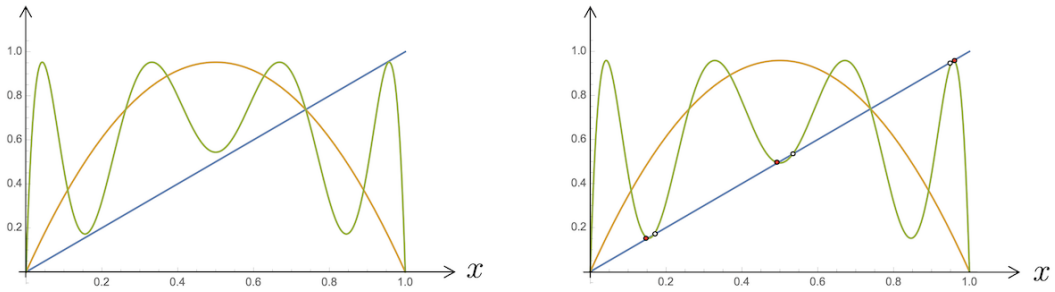


Figure 34. Plots of $y = x$, $y = f(x)$ and $y = f^3(x)$. On the left, $r = 3.8$ and on, the right, $r = 3.835$. In the latter case, there are solutions to $x = f^3(x)$, reflecting the emergence of a stable 3-cycle.

We can get more insight into this by staring (for a long time!) at the iconic plot shown in Figure 33. This depicts the long-time behaviour of the logistic map for different values of r. For $1 < r < 3$, there is the single fixed point that we found above. At $r \approx 3.45$, this bifurcates into the period 2-cycle and this later bifurcates into the period 4-cycle and so on. For $r > r_\infty$, we see that a seeming continuum of values of x appear. This is the regime of chaos.

But, perhaps most surprisingly, the chaos does not persist for all values of $r > r_\infty$. Occasionally, we find windows of order. These are the white stripes that are apparent in Figure 33 where we again find cycles of some order.

The most prominent of these is the white stripe around $r = 3.88$. Towards the left of this stripe, you can see three points in the diagram. This reflects the fact that, in a small window, there is a stable 3-cycle. This is something new, as all previous p-cycles had p a power of 2.

We could try to repeat the story above to understand how this 3-cycle comes about, looking for fixed points of $f^3(x)$. But, as for all higher powers, this is tricky because it's a higher order polynomial and we don't have analytic expressions for its values. But we can plot the functions $y = x$, $y = f(x)$ and $y = f^3(x)$ to see where the fixed points lie. These functions are plotted in Figure 34 for $r = 3.8$ (on the left) where there is no 3-cycle, and for $r = 3.835$ (on the right) which is in the window with the stable 3-cycle. We can see that, as r changes, the dips in $f^3(x)$ come down sufficiently to intersect the straight line, revealing the fixed points. In fact, the straight line intersects the curve $f^3(x)$ twice each time: these are denoted with red and white dots. The stable 3-cycle

corresponds to intersections where $f^3(x)$ is shallower and these are coloured red. The white dots are then an accompanying unstable 3-cycle.

The period doubling that we saw previously also takes place within these windows of order. For example, the 3-cycle turns into a 6-cycle, which then turns into a 12-cycle, and so on.

2.2.4 The Logistic Map in Ecology

The logistic map is certainly very pretty. But, so far, we haven't really explained why it might be useful in modelling population dynamics.

The potential utility of the logistic map, with its panoply of different behaviours, was advertised in a famous paper by the mathematical ecologist Bob May[8]. The first compelling evidence that period bifurcations and the ensuing chaos are at play in the world of population dynamics was presented in a study from the 1990s on flour beetles[9]. The population at discrete time n consists of larvae L_n, pupae P_n and adults A_n and can be modelled as

$$\begin{aligned} L_{n+1} &= bA_n e^{-c_1 A_n - c_2 L_n} \\ P_{n+1} &= r_L L_n \\ A_{n+1} &= P_n e^{-c_3 A_n} + r_A A_n \ . \end{aligned} \tag{2.30}$$

Here the constants b, r_L and r_A capture reproductive and death rates while the constants c_1, c_2 and c_3 are more gruesome, describing the cannibalistic tendencies of flour beetles. This set of equations exhibits many of the features of the logistic map, including a series of period bifurcations before descending into chaos as the parameters are varied. The rates r_L and r_A were artificially varied in the experiment by simply removing individuals from the population. There is then a feedback effect where the constants c_i, with $i = 1, 2, 3$, depend on r_L and r_A. In an experiment that took place over many years, the authors observed both period doubling and chaotic behaviour in the population.

The current consensus is that chaos is possible, but not common, in nature. However, this has been challenged recently in work that suggests chaos may not be so rare after

[8]The paper is "Simple mathematical models with very complicated dynamics", Nature 261 (1976) and makes for an easy and fun read.

[9]The original paper is R.F. Constantino et al, "Chaotic Dynamics in an Insect Population", Science, vol 275 (1997).

n	0	1	2	3	4	5	6
r_n	3	3.44948974	3.54409035	3.56440726	3.56875941	3.56969160	3.56989125
δ_n		4.751	4.656	4.668	4.668	4.669	4.669

Table 1. Numerical values for r_n and δ_n

all[10]. It's hard to know for sure because it is challenging to collect the long time series needed to estimate Lyapunov exponents and other tell tale signs of chaos.

There is something a little unnerving in finding chaos in population dynamics. One might, quite reasonably, think that wild swings in population size are due to some extreme event, say weather or famine. Moreover, if we can learn to control such events then we could restore order to the universe. But chaotic systems exhibit wild swings for no underlying reason other than the inherent dynamics itself. Chaos is a control freak's worst nightmare.

2.3 Universality

There's something magical lurking in the discussion above. As we increase r, we get a series of period doublings. We can ask: for what values of r does the period double?

We calculated the first two of these above. The single fixed point becomes unstable and bifurcates into a 2-cycle at $r = r_0$ with

$$r_0 = 3 \ . \tag{2.31}$$

The 2-cycle then bifurcates further into a 4-cycle at

$$r_1 = 1 + \sqrt{6} \approx 3.44948974 \ . \tag{2.32}$$

The value of r_n for which the 2^n cycle bifurcates into a 2^{n+1}-cycle can be calculated numerically and is shown in Table 2.3.

With these values in hand, we see that there is a pattern. We define the ratio of differences

$$\delta_n = \frac{r_n - r_{n-1}}{r_{n+1} - r_n} \ . \tag{2.33}$$

[10]See T. Rogers, B. Johnson, and S. Munch, "Chaos is not rare in natural ecosystems", Nature Ecology & Evolution, 6, (2022).

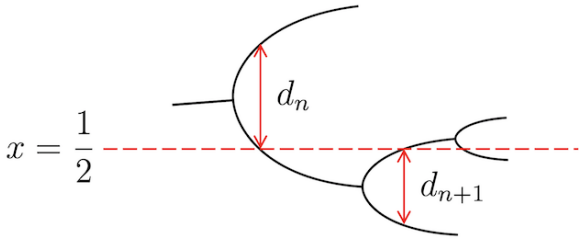


Figure 35. The heights of successive bifurcations, measured relative to the line $x = 1/2$, are d_n.

These too are given in Table 2.3 where they are clearly converging to

$$\delta = \lim_{n\to\infty} \delta_n = 4.669\ldots \ . \tag{2.34}$$

The magic is that this number appears in many other maps. For example, if you consider the map

$$f(x) = r\sin(\pi x) \tag{2.35}$$

then, as you vary $r \in [0, 1]$, you will again see period doubling at a rate that converges towards the same value of δ. What we're seeing here is that, hidden within the logistic map, is a new mathematical constant, δ. This is known as the *Feigenbaum constant*. Or, more precisely, it is one of two Feigenbaum constants.

The other Feigenbaum constant comes from noting that the heights of successive bifurcations get smaller in Figure 33. We would like to find a way to characterise their height. We do this, by measuring the height relative to the line $x = 1/2$ (corresponding to the maximum of the function $f(x)$). We call successive heights d_n, as shown in the Figure 35. (As an aside, the kind of bifurcating diagrams shown in Figure 35 are affectionately known as fig tree diagrams, not because they look particularly like fig trees but because this is a direct translation of the German word "Feigenbaum".)

The second Feigenbaum constant comes from noting that the ratio of heights also converges to

$$\alpha = \lim_{n\to\infty} \frac{d_n}{d_{n+1}} = -2.5029\ldots \tag{2.36}$$

Here the minus sign reflects the fact that, as shown above, the heights are measured alternatey above and below the line $x = 1/2$. The value of α is another universal constant, in the sense that the same number emerges for many different maps.

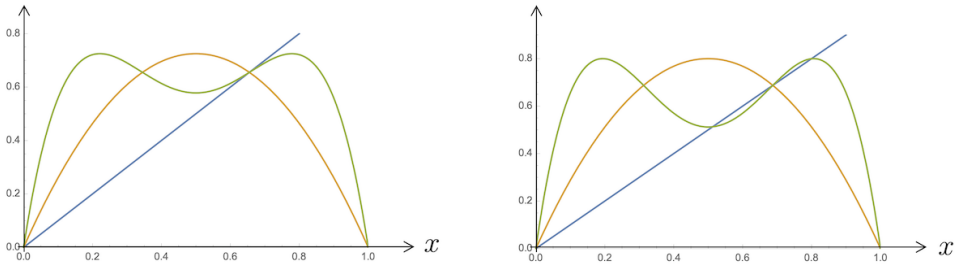


Figure 36. Plots of $y = x$, $f(x)$ and $f^2(x)$ for $r = 2.9$ (on the left) and for $r = 3.2$ (on the right).

The Feigenbaum constants α and δ arise for many maps, but not for all maps. The maps $f(x)$ in question should be smooth and "unimodal", which means that they go up, then down with a single maximum. Moreover, the maximum should be quadratic. Any map with these features will exhibit period doubling with the Feigenbaum constants α and δ.

Of course, that begs the question: what about other maps? Here too there is an interesting story. Suppose, for example, that we consider maps that have a quartic, rather than quadratic, maximum. Then you'll find the same kind of period doubling but with different constants, $\alpha \approx -1.7$ and $\delta \approx 7.3$.

The Feigenbaum constants are reminiscent of other mathematical constants, known as *critical exponents*, that arise in the theory of phase transitions. You can read more about them in the lectures on Statistical Field Theory. Indeed, we will borrow some ideas of "renormalisation" from the theory of phase transitions below when we describe how to calculate the Feigenbaum constants.

2.3.1 Zooming in on Bifurcations

Let's first get a sense for why this universality might be happening. We can go to the very first bifurcation, in which the stable fixed point changes to the 2-cycle.

We can get some understanding of what's going on by staring at Figure 36 where we plot $y = x$, $y = f(x)$ and $y = f^2(x)$. On the left, these are plotted for $r = 2.9$ where the only solutions to $x = f^2(x)$ are also solutions to $x = f(x)$. On the right, we plot these same functions for $r = 3.2$. Now we see that there are additional solutions to $x = f^2(x)$ that are *not* solutions to $x = f(x)$. This transition happens at $r = 3$, and this is where the single stable fixed point of $f(x)$ becomes unstable and is replaced by the stable 2-cycle, corresponding to the new fixed points of $f^2(x)$.

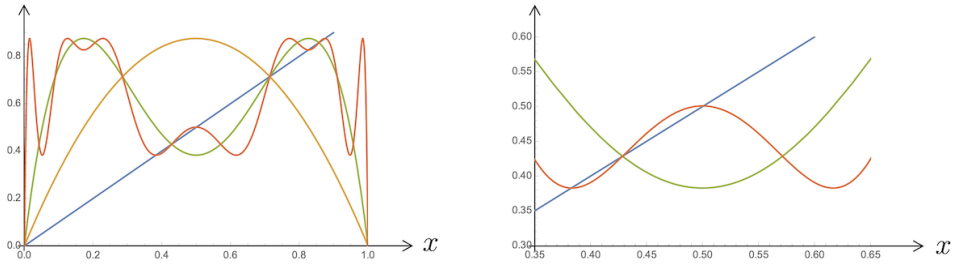
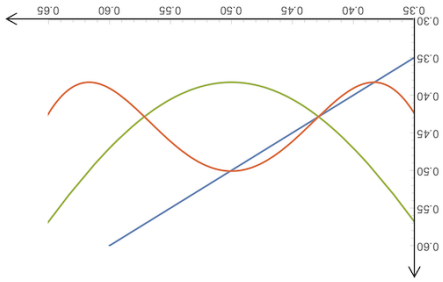


Figure 37. Plots of $y = x$, $f(x)$, $f^2(x)$ and, $f^4(x)$. This is plotted for $r = 3.5$, after the 2-cycle becomes unstable and there are new fixed points of $f^4(x)$. These can be shown in the plot on the left and, more clearly, in the plot on the right where we've zoomed in to the relevant piece of the graph.

So far, so good. Now what happens when we increase r further so that this 2-cycle turns into a 4-cycle? This time, the fixed points of $f^2(x)$ are becoming unstable. We can see this happening in the same graphical manner as before, this time also plotting $f^4(x)$ to see how its fixed points emerge at some value of r. This is shown on the left of Figure 37 where it's all a little cluttered. But we can zoom in to the relevant part, as shown on the right of 37.

Now comes the key observation. Take the zoomed in plot in Figure 37. Flip it upside down and reflect it. It looks like this:



But that looks very much like the right-hand plot of Figure 36!

Now we can repeat this, looking at how the fixed points of $f^4(x)$ become unstable as new fixed points of $f^8(x)$ appear, and then how these become unstable as new fixed points of $f^{16}(x)$ appear, and so on. At each stage, we zoom in and flip and what we're left with is always a figure that looks like the one above. The existence of the two universal Feigenbaum constants suggests that, as we do this procedure over and over again, we might converge on some universal function. Our task is to try to understand some properties of this function and to extract the Feigenbaum constants from it.

2.3.2 Renormalisation

We want to put the ideas above on a firmer footing. How do we implement the iterated "taking $f^2(x)$, zooming and flipping" procedure described above in more concrete terms?

There are a number of steps that we need to take. First, we will describe the class of functions that we care about. It's convenient to rescale things slightly. We will consider regular, unimodal functions such that

- $x_{n+1} = f(x_n)$ with $x \in [-1, +1]$.
- The map is symmetric about the maximum at $x = 0$, with $f(0) = 1$
- The maximum is quadratic.

For example, after some rescaling the logistic map can be rewritten as

$$x_{n+1} = 1 - rx_n^2 \ . \tag{2.37}$$

This map exhibits all the universal features described above, including the two Feigenbaum constants α and δ. But everything that we say below holds for any map obeying the three criteria above.

Part of our iteration in going from $f(x)$ to $f^2(x)$ involves increasing the parameter r, so that we get to the point where $f^{2n}(x)$ develops fixed points. For example, Figure 36 is plotted with $r = 2.9$ while Figure 37 is plotted with $r = 3.5$. Let's first look more carefully at how we do this.

A map $x \to f(x)$ has a fixed point $x^\star = f(x^\star)$. This fixed point is stable if $f'(x^\star) < 1$ and is unstable if $f'(x^\star) > 1$. The map fixed point is said to be *superstable* if

$$f'(x^\star) = 0 \ . \tag{2.38}$$

This is the most stable that a fixed point can be. In this case, the convergence towards the fixed point is typically exponential rather than power-law.

It's clear that a fixed point of our class of maps is superstable if $x^\star$ coincides with the maximum of the map $x_{\rm max}$ which, by construction, we've taken to be $x_{\rm max} = 0$. Similarly, for the higher maps $f^2(x)$ with two new fixed points x_+ and x_-, we saw in (2.27) that $df^2/dx(x_+) = f'(x_+)f'(x_-)$, so this fixed point is superstable if either x_+ or x_- coincides with $x_{\rm max}$.

We will denote the value of r at which the fixed point of $f^{n-1}(x)$ is superstable as R_n. These are shown in Figure 38. Note that $r_n < R_n < r_{n+1}$. It turns out that the superstable points R_n converge in the same manner as the bifurcation points r_n,

$$\delta = \lim_{n\to\infty} \frac{R_n - R_{n-1}}{R_{n+1} - R_n} = 4.669\ldots \tag{2.39}$$

In what follows, we will phrase everything in terms of maps evaluated at the superstable points R_n.

The Renormalisation Map

Now we can start to put our iteration process in place. We start with a map $f(x; R_0)$ which has a superstable fixed point. We then want to turn this into the appropriately zoomed and flipped map $f^2(x; R_1)$ which, importantly, also has a superstable fixed point. Moreover, we want the map $f^2(x)$ to fall into our general class of maps, obeying the various criteria listed above.

To achieve, this suppose that

$$f(1) = -a\ . \tag{2.40}$$

Then, because $f(0) = 1$, we have $f^2(0) = f(1) = -a$. So to keep ourselves within the class of maps with $f(0) = 1$, we should rescale f^2. This is the zooming described above. We should also rescale x to ensure that the domain remains in $x \in [-1, +1]$, with a minus sign to give us the necessary reflection. The upshot is that the zooming and flipping procedure is described by the following action on a map

$$f(x; R_0) \mapsto -\frac{1}{a} f^2\left(-ax; R_1\right)\ . \tag{2.41}$$

Both the original map $f(x)$ and the new map $f^2(x)$ are evaluated at the appropriate point $r = R_n$ where they have a superstable fixed point. We say that the map $f(x; R_0)$ has been *renormalised*, terminology stolen from quantum field theory.

The renormalisation map (2.41) can itself be viewed as a dynamical system, but not one acting on a single variable x but now acting on a class of functions $f(x)$. Said differently, this is a dynamical system with an infinite number of degrees of freedom. Nonetheless, we could push on and think of it like any other dynamical system. We could, for example, reiterate the process n times

$$f(x, R_0) \mapsto \left(-\frac{1}{a}\right)^n f^{2n}\left((-a)^n x; R_n\right)\ . \tag{2.42}$$

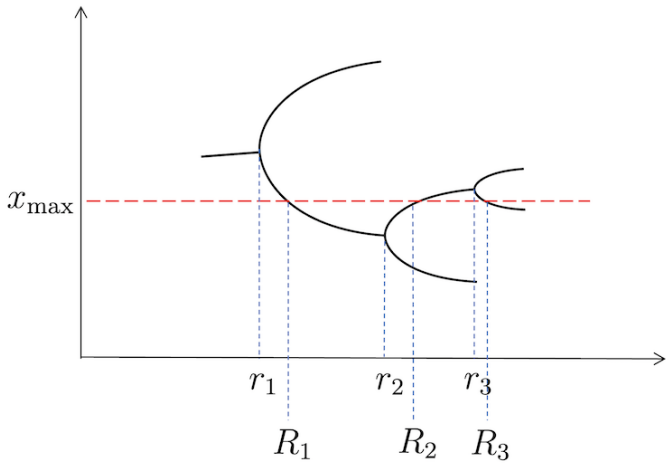


Figure 38. The bifurcations happen at points $r = r_n$; the fixed points are superstable at $r = R_n$.

Then we can ask: does this process converge? In other words, is there some universal function defined by

$$g(x) = \lim_{n\to\infty} \left(-\frac{1}{a}\right)^2 f^{2n}\left((-a)^n x; R_n\right) \ ? \tag{2.43}$$

It's not at all obvious that such a function $g(x)$ exists. But the universality observed in the Feigenbaum constants suggests that, under the right circumstances, it might. But what are these circumstances?

2.3.3 The Feigenbaum constant α

To get a sense for when universal function $g(x)$ in (2.43) might exist, we need to think more carefully about the meaning of that rescaling factor a. It is rescaling the x coordinate by a, but we expect from our previous discussion that this is what the Feigenbaum constant α is doing. This means that we should identify

$$\alpha = -\frac{1}{a} \ . \tag{2.44}$$

Moreover, it suggests that the limit (2.43) should only exist if we take a very specific value of α. We just need to compute this value.

Now we're on the home straight. If there's a universal function $g(x)$ that is the limit of the renormalisation map then, when we put it in the renormalisation map, nothing

should happen. In other words, $g(x)$ should be a fixed point of the renormalisation map and satisfy

$$g(x) = \alpha\, g^2\left(\frac{x}{\alpha}\right) \ . \tag{2.45}$$

This slightly strange, self-referential equation defines both the function $g(x)$ and the constant α. Indeed, as we've chosen $g(0) = 1$, we have

$$1 = \alpha g(1) \quad \Longrightarrow \quad \alpha = \frac{1}{g(1)} \ . \tag{2.46}$$

To make progress, we can simply Taylor expand $g(x)$ around the origin. We know that it is a symmetric function with a quadratic maximum, so we can Taylor expand

$$g(x) = 1 + \sum_{n=1}^{N} c_{2n} x^{2n} \ . \tag{2.47}$$

Substitute this into (2.45) and compare various terms. The constant terms give

$$1 = \alpha(1 + c_2 + c_4 + \ldots) \tag{2.48}$$

which just reiterates the result (2.46). The x^2 terms give

$$\alpha = 2c_2 + 4c_4 + \ldots \ .. \tag{2.49}$$

The x^{2n} terms are polynomials of degree $2n$ in the variables α and c_{2n}. The upshot is that we have $N+1$ equations in $N+1$ variables which we can solve numerically. (Because these are higher order polynomials, there are several solutions and you have to make sure that you get the right one.) The higher the value of N, the better the accuracy. For low values we have:

$$\begin{aligned} N = 2 \quad &\Longrightarrow \quad c_2 \approx -1.52 \ , \ c_4 \approx 0.13 \ , \ \alpha \approx -2.53 \\ N = 3 \quad &\Longrightarrow \quad c_2 \approx -1.52 \ , \ c_4 \approx 0.073 \ , \ c_6 \approx 0.046 \ , \ \alpha \approx -2.479 \ . \end{aligned} \tag{2.50}$$

By the time you get to $N = 6$, you have $\alpha \approx -2.502897$, which is correct to 1 part in 10^6.

You can see that our ansatz (2.47) assumes that the maximum is quadratic. You could repeat the calculation setting $c_2 = 0$ so that the maximum is quartic. Then you get the different universal constant $\alpha \approx -1.7$ appropriate for such quartic maps.

2.3.4 The Feigenbaum Constant δ

Computing the other Feigenbaum constant δ is a little more involved. We will be somewhat heuristic in what follows, but still put together enough of the story to allow us to compute δ.

To start, we define the renormalisation map to act on any function $\phi(x)$ as

$$\text{Ren}[\phi(x)] = \alpha\phi^2\left(\frac{x}{\alpha}\right) \ . \tag{2.51}$$

This is the same as our previous map (2.41), except we're not changing the variable r in any way: just iterating the map and rescaling. We take the constant α to be the Feigenbaum constant.

Now consider the one-parameter family of functions $f(x;r)$ that defines our original map. We know that there are special values of $r = R_n$ where this map has a superstable 2^n-cycle. The claim of universality is that they converge as

$$R_n = R_\infty - \frac{A}{\delta^n} \tag{2.52}$$

for some (non-universal) constant A and with δ the Feigenbaum constant (2.34). This is what we would like to show. We will first need to develop some machinery to do this.

Expanding About the Universal Function

The renormalisation map $\text{Ren}[\phi(x)]$ has a "fixed point", or more precisely a "fixed function", which is our universal function $g(x)$,

$$\text{Ren}[g(x)] = \alpha g^2\left(\frac{x}{\alpha}\right) = g(x) \ . \tag{2.53}$$

Usually, when presented with a dynamical system with a fixed point, our first inclination is to linearise around the fixed point to see what happens in its vicinity. The same is true here. We look at functions $\phi(x)$ that are close to the universal function, with

$$\phi(x) = g(x) + \epsilon\,\theta(x) \tag{2.54}$$

with $\epsilon \ll 1$ and $\theta(x)$ some other arbitrary function. Now we act with the renormalisation map of $f(x;r)$. The map isn't linear so we have to tread slowly. We have

$$\begin{aligned}
\text{Ren}[\phi(x)] &= \alpha\phi\left(\phi\left(\frac{x}{\alpha}\right)\right) \\
&= \alpha\phi\left(g\left(\frac{x}{\alpha}\right) + \epsilon\theta\left(\frac{x}{\alpha}\right)\right) \\
&= \alpha\left[\phi\left(g\left(\frac{x}{\alpha}\right)\right) + \epsilon\phi'\left(g\left(\frac{x}{\alpha}\right)\right)\theta\left(\frac{x}{\alpha}\right)\right] \\
&= \alpha g\left(g\left(\frac{x}{\alpha}\right)\right) + \epsilon\alpha\left[\theta\left(g\left(\frac{x}{\alpha}\right)\right) + g'\left(g\left(\frac{x}{\alpha}\right)\right)\,\theta\left(\frac{x}{\alpha}\right)\right]
\end{aligned} \tag{2.55}$$

where, in the third and fourth lines we've Taylor expanded and dropped terms of order ϵ^2.

When you Taylor expand a function, the first correction is the derivative. Here we've Taylor expanded a *functional* $\text{Ren}[\phi(x)]$, which is a function of a function. The term multiplying the ϵ parameter should be thought of as the functional derivative of $\text{Ren}[\phi(x)]$. In fancy maths words, it's called the Fréchet derivative. We define

$$\text{DRen}_g[\theta(x)] = \alpha\left[\theta\left(g\left(\frac{x}{\alpha}\right)\right) + g'\left(g\left(\frac{x}{\alpha}\right)\right)\,\theta\left(\frac{x}{\alpha}\right)\right] \ . \tag{2.56}$$

We can then write

$$\text{Ren}[\phi(x)] = g(x) + \epsilon\,\text{DRen}_g[\theta(x)] \ . \tag{2.57}$$

where we've used the fact that $g(x)$ is a fixed point to get the first term.

How should we think of this? If this were a dynamical system with a finite number of degrees of freedom then, when we linearise around a fixed point, we get a matrix. And to understand the behaviour of the fixed point we have to look at the eigenvectors and eigenvalues of that matrix. The same is true here, except that we should look for eigenfunctions, $\theta_i(x)$ of the weird operator DRen. These obey

$$\text{DRen}_g[\theta_i(x)] = \lambda_i\theta_i(x) \ . \tag{2.58}$$

Here λ_i is the corresponding eigenvalue and i is an (infinite) index that labels the different eigenthings.

Our next step would be to solve for the different eigenvalues using (2.58). And we will, eventually, do this. The trouble is that this is a hard equation to solve and it's useful to have some motivation for doing so! That's where we're going next. We will argue that, out of the infinite number of eigenvalues, there is just single one λ_0 that has $|\lambda_0| > 1$. And, rather wonderfully, this eigenvalue coincides with the Feigenbaum constant δ.

Expanding About the Edge of Chaos

To extract the Feigenbaum constant, we need to look more carefully at the original map $f(x;r)$ and, in particular, this map evaluated at the superstable points $r = R_n$.

To start, we look at the map when it sits at the edge of chaos at $r = R_\infty$. We define

$$F(x) = f(x; R_\infty) \ . \tag{2.59}$$

At this point, we need to make an assumption: we assume that if you act with successive iterations of the map Ren on the function $F(x)$, then you will quickly converge towards the universal function $g(x)$ defined in (2.45),

$$\text{Ren}^n[F](x) \approx g(x) \quad \text{for } n \text{ suitable large.} \tag{2.60}$$

Said in more sophisticated language, we assume that $F(x)$ lies on the stable manifold of the fixed point $g(x)$. This assumption is, it turns out, true, but we will not prove it here. It seems plausible because the renormalisation map Ren differs from our original renormalisation by not changing the value of r, but we've already tuned the value of r to its final resting place R_∞ when considering $F(x)$.

Next we ask: what if we act with the renormalisation map on $f(x; r)$ with r close to R_∞? We write

$$f(x; r) = F(x) + \epsilon \left.\frac{\partial f(x; r)}{\partial r}\right|_{r=R_n} \quad \text{with} \quad \epsilon = r - R_\infty \ . \tag{2.61}$$

Here we've dropped terms of order ϵ^2 and higher. Acting with the renormalisation map on $f(x; r)$ is just a matter of repeating the calculation (2.55), with $g(x)$ replaced by $F(x)$ and $\theta(x)$ replaced by $\partial f/\partial r$. We can express the result in terms of our operator DRen defined in (2.56),

$$\text{Ren}[f(x; r)] = \alpha F\left(F\left(\frac{x}{\alpha}\right)\right) + \epsilon\, \text{DRen}_F\left[\frac{\partial f(x; r)}{\partial r}\right] \ . \tag{2.62}$$

Now we act with successive renormalisation maps on this function. We know that $F(x)$ tends towards the universal function $g(x)$, as in (2.60). Acting successively on the function $f(x; r)$, with r close to R_∞ then gives

$$\text{Ren}^n[f(x; r)] \approx g(x) + (r - R_\infty)\, \text{DRen}_g^n\left[\frac{\partial f(x; r)}{\partial r}\right] \quad \text{for suitable large } n. \tag{2.63}$$

Admittedly, that $\approx$ sign is doing some heavy lifting here. We've taken n iterations of the map DRen_g, each of then evaluated around the function $g(x)$ rather than the function $F(x)$ or some intermediate. That should really be justified. But we're not going to.

To proceed, we expand the function $\partial f/\partial r$ in terms of eigenfunctions of the operator DRen. We write

$$\frac{\partial f(x;r)}{\partial r} = \sum_i a_i \theta_i(x) \ . \tag{2.64}$$

Clearly there's yet another assumption here that the $\theta_i(x)$ form a complete basis of functions. You may have guessed by now that it's not an assumption we're going to justify. If we substitute this ansatz into the result (2.63), we have

$$\text{Ren}^n[f(x;r)] \approx g(x) + (r - R_\infty) \sum_i \lambda_i^n a_i \theta_i(x) \ . \tag{2.65}$$

Now we're in good shape to make the final argument.

The Feigenbaum Constant is an Eigenvalue

We make the following claim:

Claim: If there is just a single eigenvalue λ_0 with $|\lambda_0| > 1$, then

$$\lambda_0 = \delta \tag{2.66}$$

with δ the Feigenbaum constant.

Proof: The key idea is to get different expressions for the renormalisation $\text{Ren}^n[f(x;r)]$. First, if the assumption is correct, and all eigenvalues other than λ_0 have modulus $|\lambda_i| < 1$, then the iterations in (2.65) quickly kill all but the $\theta_0(x)$ eigenfunction,

$$\text{Ren}^n[f(x;r)] \approx g(x) + (r - R_\infty)\lambda_0^n \, a_0 \, \theta_0(x) \ . \tag{2.67}$$

Now we think about this result applied to the case with $r = R_n$. The nice thing about the function $f(x; R_n)$ is that it has a superstable n-cycle and, moreover, we know that $x = 0$ is one of the points on this n cycle. In other words, if we act with $f(x; R_n)$ a total of 2^n times, starting at $x = 0$, then we get back to $x = 0$,

$$f^{2n}(x = 0; R_n) = 0 \ . \tag{2.68}$$

But this means that the left-hand side of (2.67) vanishes, with the extra scaling by α in the renormalisation group map unimportant because $0/\alpha = 0\alpha = 0$. So we have

$$R_n - R_\infty = -\frac{g(0)}{a_0 \, \theta_0(0)} \frac{1}{\lambda_0^n} = \frac{\text{constant}}{\lambda_0^n} \ . \tag{2.69}$$

But this is precisely the geometric progression (2.52) seen in the bifurcations, with $\lambda_0 = \delta$. □

It remains to find the eigenvalue $\lambda_0 = \delta$. For this, we need to solve the eigenfunction equation (2.58)

$$\mathrm{DRen}_g[\theta(x)] = \alpha\left[\theta\left(g\left(\frac{x}{\alpha}\right)\right) + g'\left(g\left(\frac{x}{\alpha}\right)\right)\,\theta\left(\frac{x}{\alpha}\right)\right] = \delta\,\theta(x)\ . \tag{2.70}$$

This too can be found using a power series ansatz for $\theta(x)$, together with our previous expansion for the universal function $g(x)$. Expanding to order $N = 6$ is sufficient to give $\delta \approx 4.66914$, accurate to one part in 10^5.

3 Spatial Variations

In the past two sections, we considered the dynamics of functions that depend only on time. But there are other dimensions in our universe and these too can be important in biology. If, for example, you're a rabbit then the location of a fox is important to you. The purpose of this section is to include the effects of spatial localisation in our models.

Mathematically, this means that our variables depend both on time t and on space $\mathbf{x}$. For example, instead of working with a total population $N(t)$, we instead have a *population density* $n(\mathbf{x}, t)$. If you integrate this density over some region V then it tells you the total population inside that region

$$N(t) = \int_V d^3x \; n(\mathbf{x}, t) \; . \tag{3.1}$$

The fact that we're dealing with functions of space and time means that our system will no longer be described by a system of ordinary differential equations. Instead, we must embrace partial differential equations and all they have to offer.

The Continuity Equation

Many of the variables of interest in mathematical biology are counting things. And these things are, by and large, conserved.

That sentence may seem strange given that everything we've done so far is devoted to understanding the time evolution of these variables. If they were truly conserved then they wouldn't change! But, as a starting point, the conservation of things is important. For example, it's true that you were born and you will die but, if you're lucky, there's a good 70 to 80 years in between in which neither of these things happens. This means that the population is approximately conserved. We can then start to look at how birth and death rates change this conclusion.

Crucially, when things are conserved in physics (and, indeed, in biology) they are conserved *locally*. The amount of conserved quantity can change in one region of space, but only because it moves to a neighbouring region. There is an important and ubiquitous equation that captures this fact: the density of some stuff $n(\mathbf{x}, t)$ is conserved if there exists a vector function $\mathbf{J}(\mathbf{x}, t)$, known as a *current density* or *flux*, that obeys

$$\frac{\partial n}{\partial t} + \nabla \cdot \mathbf{J} = 0 \; . \tag{3.2}$$

This is the *continuity equation*. We've met it previously in courses on Electromagnetism (where electric charge is conserved), Fluid Mechanics (where mass is conserved), and Quantum Mechanics (where probability is conserved).

To see why the continuity equation (3.2) implies conservation, we integrate both sides over a region of space V with boundary $S = \partial V$ and then invoke Gauss' divergence theorem,

$$\frac{dN}{dt} = \int_V d^3x \, \frac{\partial n}{\partial t} = -\int_V d^3x \, \nabla \cdot \mathbf{J} = -\int_S d^2\mathbf{S} \cdot \mathbf{J} \ . \tag{3.3}$$

We learn that the total population $N(t)$ in some region V can change with time, but only if there is a flux of the current $\mathbf{J}$ out of the boundary of the region. Often we will be interested in some closed region V from which there is no escape.In this case we have $\mathbf{J} = 0$ on the boundary S and, correspondingly, N is constant. For convenience, we may sometimes, unrealistically in the context of biology, take $V = \mathbb{R}^3$ and require that $\mathbf{J} \to 0$ suitably fast asymptotically.

Anything that is conserved obeys the continuity equation (3.2). But, as we have stressed, our populations and other beasts are typically not fully conserved. In this case, it's straightforward to amend the continuity equation: we just include an additional term on the right-hand side

$$\frac{\partial n}{\partial t} + \nabla \cdot \mathbf{J} = F(\mathbf{x}, t) \ . \tag{3.4}$$

This function $F(\mathbf{x}, t)$ captures any loss or creation of the quantity of interest. For example, if $n(\mathbf{x}, t)$ is the population density then $F(\mathbf{x}, t)$ may describe the birth and death rates, now allowed to vary in both space and time. The function F may itself depend on n or (less commonly) $\mathbf{J}$.

Local conservation means that it's not enough to talk only about the density $n(\mathbf{x}, t)$: we also need to introduce the current density $\mathbf{J}(\mathbf{x}, t)$. In general, this could be an independent variable. But there are two situations that arise most commonly:

- Diffusion: When the underlying density is subject to constant, random fluctuations, the result is that the current is given by

  $$\mathbf{J}(\mathbf{x}, t) = -D\nabla n(\mathbf{x}, t) \ . \tag{3.5}$$

  This is known as *Fick's law* and D is a constant known as the *diffusivity*. Fick's law is telling us that there is a current from high density regions to low density

regions. We will get more intuition for this result in Section 4 when we discuss the effects of randomness and, in particular, when we derive the Fokker-Planck equation. For now, we will take this result as an assumption and see where it takes us. With a current of this form, the continuity equation becomes

$$\frac{\partial n}{\partial t} = D\nabla^2 n \ . \tag{3.6}$$

This is the *heat equation*. Processes governed by this equation are said to undergo *diffusion*. We will devote Section 3.1 to solving this equation. For now, note that this equation makes it clear that the dimension of the diffusivity is $[D] = L^2/T$.

- Advection: An alternative scenario is that the stuff we care about is sitting in some moving, background medium and just going with the flow. In this case, the current is given by

$$\mathbf{J}(\mathbf{x}, t) = n(\mathbf{x}, t)\, \mathbf{u}(\mathbf{x}, t) \tag{3.7}$$

where $\mathbf{u}(\mathbf{x}, t)$ is the underlying velocity field. This kind of current typically arises in fluid mechanics and is known as *advection*. In this case, you need to separately specify the form of the velocity field or, if it's a dynamical variable, introduce more equations (such as the Euler or Navier-Stokes equations) that govern its properties.

In general, it's quite possible that the current has both a diffusive piece and an advective piece.

In both the examples above, there can be no current $\mathbf{J}(\mathbf{x}, t)$ if $n(\mathbf{x}, t) = 0$. That's in contrast to what happens in electromagnetism where it's quite possible to have $\mathbf{J} =\neq 0$ even if the charge density $n = 0$. That's because electromagnetism comes with both positive and negative charges which cancel out. But if the negative charges move, while the positive charges stay still – which is what happens in a wire conducting electricity – then the current is non-zero. In contrast, the diffusion and advection described above typically happens in situations where $n(\mathbf{x}, t) > 0$. This, of course, is the situation for populations.

There are other ways to generalise the ideas above, some of which we will meet later. For example, the diffusivity D in (3.5) could depend on the function $n(\mathbf{x}, t)$. The same is true of the forcing function F in (3.4). If we have standard diffusion, together with a field-dependent forcing $F(n)$, we are left with a class of equations that take the form

$$\frac{\partial n}{\partial t} - D\nabla^2 n = F(n) \ . \tag{3.8}$$

These are known as *reaction-diffusion equations*, with the $D\nabla^2 n$ term capturing diffusion and the $F(n)$ term said to be the "reaction". This is the class of equations that we will mostly focus on in this section.

3.1 Diffusion

We first study the solutions to the heat equation (3.6) in various situations. To keep things simple, we will restrict ourselves to the one-dimensional case,

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} \ . \tag{3.9}$$

As we proceed, we'll also look at some generalisations of this equation.

3.1.1 Diffusion on a Finite Interval

We start by placing our system on a finite interval $x \in [0, L]$. We will impose boundary conditions on both ends

$$n(0,t) = n_0 \quad \text{and} \quad n(L,t) = n_1 \ . \tag{3.10}$$

with n_0 and n_1 both constant. In the context of thermodynamics, we might think of $n(x,t)$ as the temperature along a one-dimensional rod with the two ends sitting in some heat reservoir, held at fixed temperatures. In the context of biology, we could think of $n(x,t)$ as the population of something small (say, an ant or a bacterium) which can move along a narrow tube. The ends of the tube are connected to two population reservoirs, each held at constant population density. The heat equation (3.9) then tells us how the temperature/population varies along the tube.

To start, we can look for a steady state solution with no time dependence. This is straightforward. We have

$$\frac{\partial^2 n}{\partial x^2} = 0 \quad \Longrightarrow \quad n(x,t) = n^\star(x) = n_0 + (n_1 - n_0)\frac{x}{L} \tag{3.11}$$

where we've implemented the boundary conditions (3.10). We've called this solution $n^\star(x)$ because it's analogous to the fixed points that we found in the dynamical systems of Section 1.

For this steady state, the flux is $J = -D\,\partial n/\partial x = D(n_0 - n_1)/L$ is constant. If $n_0 > n_1$ then $J > 0$. If $n_0 < n_1$ then $J < 0$. In both cases, there is a net flux from the high density population to the low density population.

As we mentioned briefly above (and will see more in Section 4), diffusion typically happens where there is some underlying randomness in the situation. Having a net flux $J > 0$ doesn't mean that the ants are marching in lockstep from left to right. There may be some ants wandering in one direction and some in the other. But there's an overall preference for them to travel left to right. Our diffusion model doesn't capture these microscopic fluctuations: only the overall trend of the flow.

Time Dependence

What happens when we deviate from the steady state? Now we need to specify the initial value of the field at $t = 0$. The resulting solution takes the form

$$n(x, 0) = n^\star(x) + c(x, t) \ . \tag{3.12}$$

where $c(x, t)$ also solves the heat equation

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} \tag{3.13}$$

now with boundary conditions $c(0, t) = c(L, t) = 0$.

We can look for separable solutions of the form

$$c(x, t) = f(x)\, g(t) \ . \tag{3.14}$$

Substituting this into the heat equation, we see that the two functions must obey

$$f\dot{g} = Dgf'' \ . \tag{3.15}$$

Dividing through by fg, we have

$$\frac{\dot{g}}{g} = D\frac{f''}{f} \ . \tag{3.16}$$

But the left-hand side is a function only of t, and the right-hand side is a function only of x, which means that actually both sides must be constant. The solution for $f(x)$ is constrained by the boundary conditions $f(0) = f(L) = 0$ which tells us that solutions must be of the form

$$f''(x) \propto f(x) \quad \Longrightarrow \quad f(x) = \sin\left(\frac{\pi n x}{L}\right) \quad \text{with} \quad n = 1, 2 \ldots \tag{3.17}$$

Now (3.16) fixes the form of $g(t)$,

$$\dot{g} = -\lambda_n g \quad \Longrightarrow \quad g = e^{-\lambda_n t} \quad \text{with} \quad \lambda_n = \frac{D\pi^2 n^2}{L^2} \ . \tag{3.18}$$

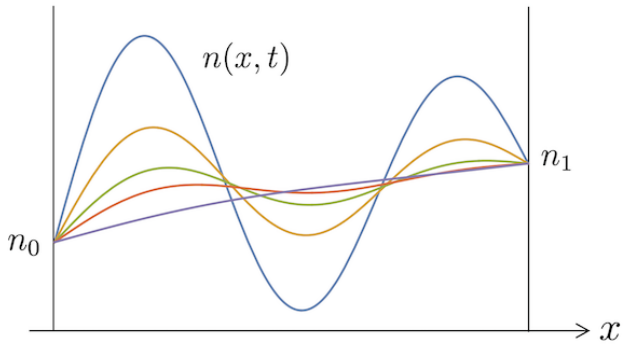


Figure 39. Diffusion of an initial wiggly profile quickly settles down to the linear, steady state.

Because the heat equation is linear, we can simply add together separable solutions for different n. Moreover, the most general solution can be constructed in this way and takes the form

$$c(x,t) = \sum_{n=1}^{\infty} c_n e^{-\lambda_n t} \sin\left(\frac{\pi n x}{L}\right) \ . \tag{3.19}$$

Here the c_n are determined by the initial conditions where they are essentially Fourier components of the initial profile at time $t = 0$. We now see the key feature of the heat equations: all the higher Fourier modes die off exponentially quickly, tending towards the steady state solution. The higher the Fourier mode, so the more wiggly the profile, the faster it decays away. This is the characteristic behaviour of the heat equation: it smooths things out. An example of the evolution of $n(x,t)$, plotted for increasing values of t is shown in Figure 39.

No Flux Boundary Condition

We can look at generalisations of this set-up. For example, instead of fixing the value of $n(x,t)$ on both ends, we could instead require that, say, $n(0,t) = n_0$ on the far left, but

$$J = -D\frac{\partial n}{\partial x} = 0 \ . \tag{3.20}$$

on the far right at $x = L$. Mathematically we say that we are imposing Dirichlet boundary conditions on the left, and Neumann boundary conditions on the right. Physically, we might have a tube connected to a population of ants but, having travelled all the way down it, the ants are disappointed to find that the end is closed off, ensuring that there is no flux.

Now the steady state solution is simply $n(x,t) = n_0$ and we can again write the most general solution as $n(x,t) = n_0 + c(x,t)$. We can proceed largely as before, looking first for separable solutions of the form $n(x,t) = f(x)\,g(t)$ which must obey (3.16). The novelty now comes in the boundary conditions that are imposed on $f(x)$ which hold if we take

$$f(x) = \sin\left(\frac{(2n-1)\pi x}{2L}\right) \quad . \quad \text{with} \quad n = 1, 2, \ldots \tag{3.21}$$

The most general solution is then

$$n(x,t) = n_0 + \sum_{n=1}^{\infty} c_n e^{-\lambda_n t} \sin\left(\frac{(2n-1)\pi x}{2L}\right) \quad . \quad \text{with} \quad \lambda_n = \frac{D(2n-1)^2\pi^2}{4L^2} \quad . \tag{3.22}$$

Again, we see the key feature: the faster the wiggle, the faster they die out.

3.1.2 How to Cook a Turkey

If you buy a turkey, the instructions will typically tell you to cook it for 20 minutes per kg, and then another 70 minutes for a 2-4 kg turkey, or another 90 minutes for a 4-10 kg turkey. These slightly convoluted rules arise because the relationship between the cooking time and the weight is not linear. The correct relationship was suggested by the particle physicist Pief Panofsky who pointed out that the cooking time τ scales with the mass M by the relation

$$\tau \sim M^{2/3} \; . \tag{3.23}$$

We can derive this formula using the ideas of diffusion described above[11].

The temperature $T(\mathbf{x}, t)$ of the turkey is described by the heat equation, now in 3d

$$\frac{\partial T}{\partial t} = D\nabla^2 T \; . \tag{3.24}$$

In fine tradition, we will assume that our turkey is spherical. (The analysis below also holds for spherical cows.) Then, further assuming that the oven is also spherically symmetric we can think of $T = T(r,t)$ with r the radial coordinate and the heat equation takes the form

$$\frac{\partial T}{\partial t} = \frac{D}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right) \quad . \tag{3.25}$$

[11]This is taken from the paper Physics in Turkey Cooking by Jin, Wang and Wang. I'm grateful to Ray Goldstein for pointing me to this paper.

At this point, we use a trick and write $V(r,t) = rT(r,t)$. (We used a similar trick in the lectures on Quantum Mechanics when solving the Schrödinger equation in 3d.) Then we have $T' = V'/r - V/r^2$ and, rather wonderfully, the heat equation becomes

$$\frac{\partial V}{\partial t} = D\frac{\partial^2 V}{\partial r^2} \ . \tag{3.26}$$

We see that, in this new variable V, we're back solving the 1d diffusion equation. And we know how to do that! A separable solution takes the form

$$V(r,t) = e^{-\lambda Dt}\big[A\cos(\sqrt{\lambda}r) + B\sin(\sqrt{\lambda}r)\big] \ . \tag{3.27}$$

for some $\lambda > 0$. The temperature T is given by $T = V/r$ so if we want to avoid a divergence at $r = 0$ then we need to set $A = 0$. Our solution will involve only $\sin(\sqrt{\lambda}r)$.

We can determine the allowed values of λ, together with the constants A and B, by looking at the boundary conditions. If the turkey has radius R, then we have the boundary condition

$$T(r,t) = T_{\rm hot} \quad \text{for all } t \text{ and } r \geq R \tag{3.28}$$

Here $T_{\rm hot}$ is the temperature of the oven. This is telling us that we can deviate from the uniform temperature only inside the turkey, $r < R$. We do this by taking $\sqrt{\lambda} = n\pi/R$ with $n \in \mathbb{Z}^+$ and writing down the general solution

$$T(r,t) = T_{\rm hot} + \frac{1}{r}\sum_{n=1}^{\infty}\left[V_n \sin\left(\frac{n\pi r}{R}\right) e^{-n^2\pi^2 Dt/R^2}\right] \ . \tag{3.29}$$

The coefficients V_n are set by the initial conditions. We'll take this to be

$$T(r,0) = T_0 \ll T_{\rm hot} \quad \text{for} \quad 0 \leq r < R \tag{3.30}$$

This initial data is discontinuous at $r = R$ where the temperature jumps from T_0 to $T_{\rm hot}$, but it's straightforward to implement this. We just need to pick the coefficients V_n so that that it gives the Fourier decomposition of a linear function r, cancelling the $1/r$ in the denominator. It's simple to check that this is achieved by the solution

$$T(r,t) = T_{\rm hot} - \frac{2R}{\pi}\frac{(T_{\rm hot} - T_0)}{r}\sum_{n=1}^{\infty}\left[\frac{(-1)^n}{n}\sin\left(\frac{n\pi r}{R}\right) e^{-n^2\pi^2 Dt/R^2}\right] \ . \tag{3.31}$$

These initial conditions decay away in characteristic time

$$\tau = \frac{R^2}{n^2\pi^2 D} \ . \tag{3.32}$$

When the lowest $n = 1$ mode has decayed away, the turkey is cooked, approaching the steady state solution $T(r,t) = T_{\rm hot}$. Importantly, $\tau \sim R^2$. This is the origin of the Panofsky turkey rule (3.23), since the mass is proportional to volume $M \sim R^3$.

To put some numbers of this, we need to know the diffusivity for heat in a turkey. That can be easily measured to be $D \approx 2 \times 10^{-3}$ cm^{2}s^{-1}. Suppose that our turkey has radius $R \approx 10$ cm, then we find $\tau \approx 5000$ seconds, or about 80 minutes. You might want to wait for, say $2 \times \tau$, to be convinced that you're not going to get salmonella, so pop it in for three hours and voilà. Don't let anyone tell you that maths isn't useful.

3.1.3 A First Look at Diffusion With Growth

In preparation for more interesting things to come (noticeably, the Turing instability) we can look at what happens if we deform the heat equation. We will take our substance to diffuse, but now with a linear growth term on the right-hand side

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} + \lambda n \tag{3.33}$$

with λ constant.

We'll again take the system on an interval $x \in [0, L]$ and we'll take the simple situation where we impose boundary condition $n(0, t) = n(0, L) = 0$. If there was no growth term, then the system would settle down to the empty state $n = 0$. Conversely, if there was no diffusion, so $D = 0$, then we know that the population n will grow exponentially quickly. In combining the two terms we have introduced what story tellers call narrative tension. Something interesting should now happen.

There is a trick to solving this equation. We define a new variable

$$\tilde{n}(x, t) = e^{-\lambda t} n(x, t) \quad \Longrightarrow \quad \frac{\partial \tilde{n}}{\partial t} = D\frac{\partial^2 \tilde{n}}{\partial t^2} \ . \tag{3.34}$$

We see that this new variable solves our original heat equation. We can just import our previous solution (3.19) to find

$$n(x, t) = \sum_{n=1}^{\infty} c_n e^{(\lambda - \lambda_n)t} \sin\left(\frac{\pi n x}{L}\right) \quad \text{with} \quad \lambda_n = \frac{D\pi^2 n^2}{L^2} \ . \tag{3.35}$$

Although the maths was straightforward, the resulting physics is novel. There is a critical length of the interval

$$L_c = \sqrt{\frac{D\pi^2}{\lambda}} \ . \tag{3.36}$$

For $L < L_c$, the system settles down to the boring steady state $n = 0$ where everything diffuses out the end points. But for suitably long intervals, $L > L_c$, the system becomes unstable with the lowest $n = 1$ mode the first to start growing. As we make L yet longer, successive modes also become unstable.

The idea that adding an additional term to the diffusion equation can lead to spatial instability is something that we will see again shortly.

3.1.4 Diffusion on the Line

The boundary conditions played a crucial role in constructing the solutions above. What happens if we want to solve the heat equation

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} \tag{3.37}$$

on an infinite line?

We will insist that our density is localised somewhere (say, near the origin) and, moreover that $J \sim \partial n/\partial x \to 0$ as $x \to \pm\infty$. This then ensures that the total amount of stuff

$$N = \int_{-\infty}^{+\infty} dx\ n(x,t) \tag{3.38}$$

is constant, with

$$\frac{dN}{dt} = \int_{-\infty}^{+\infty} dx\ \frac{\partial n}{\partial t} = D\int_{-\infty}^{+\infty} dx\ \frac{\partial^2 n}{\partial x^2} = D\left[\frac{\partial n}{\partial x}\right]_{-\infty}^{+\infty} = 0\ . \tag{3.39}$$

We won't give the most general solution to the heat equation. Instead, we will find a particular solution that is "self-similar", meaning that after scaling space and time in a certain way, it looks the same.

The essence of these self-similar solutions is that we can replace the partial differential equation (3.37) with an appropriate ordinary differential equation, where the relevant variable is a suitable combination of x and t. To figure out what linear combination works, we do a little dimensional analysis.

We have two variables x and t and two constants with dimension $[D] = L^2T^{-1}$ and $[N] = L$. (Here we're assuming that $n(x,t)$ itself is dimensionless; you could assign it a dimension of "people density" or "bugs density" or whatever, but this won't change the conclusions below). We then introduce the dimensionless combination

$$\xi = \frac{x}{\sqrt{Dt}}\ . \tag{3.40}$$

Furthermore, we look for solutions of the form

$$n(\mathbf{x},t) = \frac{N}{\sqrt{Dt}}\, f(\xi)\ . \tag{3.41}$$

The idea here is that the constant N sets the overall scale of the solution and the factor of $(Dt)^{-1/2}$ ensures that the function $f(\xi)$ is dimensionless. At this point we have to roll up our sleeves and figure out what the heat equation looks like when written in terms of ξ. We have

$$\frac{\partial \xi}{\partial t} = -\frac{1}{2}\frac{\xi}{t} \quad \text{and} \quad \frac{\partial \xi}{\partial x} = \frac{1}{\sqrt{Dt}} = \frac{\xi}{x} \ . \tag{3.42}$$

The time derivative of $n(x,t)$ is then

$$\begin{aligned} \frac{\partial n}{\partial t} &= -\frac{1}{2t}\frac{N}{\sqrt{Dt}} f + \frac{N}{\sqrt{Dt}} f'(\xi)\frac{\partial \xi}{\partial t} \\ &= -\frac{1}{2t}\frac{N}{\sqrt{Dt}}(f + \xi f') \\ &= -\frac{1}{2t}\frac{N}{\sqrt{Dt}}\frac{d}{d\xi}(\xi f) \ . \end{aligned} \tag{3.43}$$

Meanwhile, the spatial derivatives are

$$\frac{\partial}{\partial x} = \frac{1}{\sqrt{Dt}}\frac{\partial}{\partial \xi} \quad \text{and} \quad \frac{\partial^2}{\partial x^2} = \frac{1}{Dt}\frac{\partial^2}{\partial \xi^2} \ . \tag{3.44}$$

Putting this together, the heat equation (3.37) becomes the ordinary differential equation

$$\frac{d^2 f}{d\xi^2} + \frac{1}{2}\frac{d}{d\xi}(\xi f) \ . \tag{3.45}$$

It's simple to integrate this once:

$$\frac{df}{d\xi} + \frac{1}{2}\xi f = \text{constant} \ . \tag{3.46}$$

If we want a localised solution, with $f, f' \to 0$ as $\xi \to \infty$ then this constant must vanish. We learn that we must solve

$$\frac{df}{d\xi} = -\frac{1}{2}\xi f \quad \Longrightarrow \quad f(\xi) = Ae^{-\xi^2/4} \ . \tag{3.47}$$

The normalisation condition (3.38) translates to the requirement

$$\int_{-\infty}^{+\infty} d\xi \ f(\xi) = 1 \quad \Longrightarrow \quad A = \frac{1}{\sqrt{4\pi}} \ . \tag{3.48}$$

The upshot of this analysis is that we have a self-similar solution to the heat equation given by

$$n(x,t) = \frac{N}{\sqrt{4\pi Dt}} e^{-x^2/4Dt} \ . \tag{3.49}$$

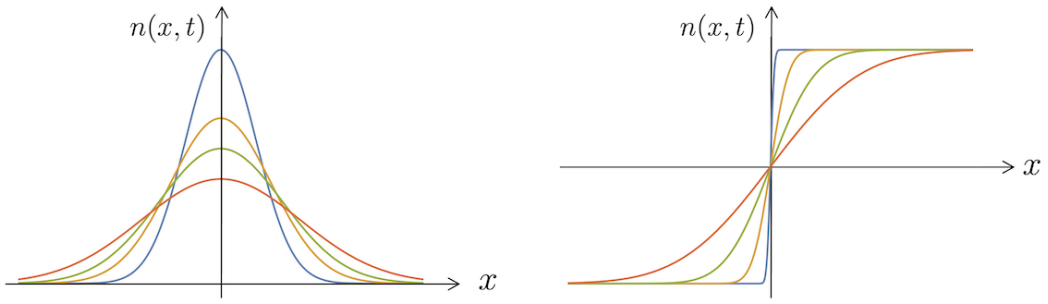


Figure 40. On the left: diffusion of a Gaussian wavepacket. On the right: diffusion of the error function. In both cases, diffusion takes the edge off.

This is a Gaussian of ever-spreading width. If we trace it back to $t \to 0^-$, it becomes a delta-function localised at the origin. Again, we see the tendency of the heat equation to take a solution and spread it out. The resulting profile for various values of t is shown on the left of Figure 40.

Changing Boundary Conditions at Infinity

We can get solutions with different boundary conditions using a slight variation of this argument. Suppose that we want a solution to the heat equation such that

$$n(x,t) \to \begin{cases} +1 & x \to +\infty \\ -1 & x \to -\infty \end{cases} . \tag{3.50}$$

Now there's no analog of the conserved quantity N because the spatial integral over $n(x,t)$ diverges. But, inspired by the approach above, we could look for solutions of the form

$$n(x,t) = t^\alpha g(\xi) \quad \text{with} \quad \xi = \frac{x}{\sqrt{Dt}} \tag{3.51}$$

and some constant α that we need to determine. The two sides of the heat equation then become

$$\frac{\partial n}{\partial t} = t^{\alpha-1}\left(\alpha - \frac{1}{2}\xi g'\right) \quad \text{and} \quad \frac{\partial^2 n}{\partial x^2} = \frac{t^\alpha}{Dt} g'' . \tag{3.52}$$

This time, the factors of t work out on both sides. But if we want to impose the boundary conditions $n(x,t) \to \pm 1$ on both sides, then we had better take $\alpha = 0$. The heat equation becomes

$$g'' + \frac{1}{2}\xi g' = 0 . \tag{3.53}$$

Again, we can integrate to get

$$g'(\xi) = Ae^{-\xi^2/4} \quad \Longrightarrow \quad g(\xi) = B + A\int_0^\xi d\eta\; e^{-\eta^2/4} \tag{3.54}$$

with A and B both integration constants. This definite integral defines the so-called *error function*

$$\text{Erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x dy\; e^{-y^2}\;. \tag{3.55}$$

It has the property that $\text{Erf}(x) \approx 2x/\sqrt{\pi}$ for $|x| \ll 1$ and $\text{Erf}(x) \to \pm 1$ as $x \to \pm\infty$. The integration constants A and B are then fixed by the boundary conditions (3.50), and we have the solution

$$n(x,t) = \text{Erf}\left(\frac{x}{\sqrt{4Dt}}\right)\;. \tag{3.56}$$

The evolution of this function with t is shown on the right-hand side of Figure 40.

Growth Revisited

We saw in Section 3.1.3 that interesting things happen if we add linear growth to the heat equation on the interval, so we have

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} + \lambda n\;. \tag{3.57}$$

Previously our interest in this came largely from the fact that there was a transition, with the growth winning when $L > L_c$ and dissipation winning when $L < L_c$. Now with L effectively infinite, you might imagine that this crossover no longer happens and growth always wins. And you would be right. For example, the dissipating Gaussian wavepacket (3.49) now solves (3.57) with

$$n(x,t) = \frac{N}{\sqrt{4\pi Dt}}e^{\lambda t}e^{-x^2/4Dt}\;. \tag{3.58}$$

For any fixed value of x, the height of the wavepacket grows exponentially in time for $\lambda > 0$. It shrinks for $\lambda < 0$.

Note that our solution (3.58) doesn't take the form $n(x,t) = t^\alpha g(x/t^\beta)$ that we would look for in a similarity solution. Indeed, you can check that no such solution of this form exists. Repeating the steps that we took above, we would end up having to solve

$$t^{\alpha-1}(\alpha g - \beta\xi g') = t^{\alpha-2\beta}Dg'' + \lambda t^\alpha g \tag{3.59}$$

and there's no way to pick α and β so that this holds for all t. This, it turns out, is rather typical: similarity solutions rarely exist when we try to solve equations with three of more terms.

3.1.5 Non-Linear Diffusion

In all the examples above, we have taken the diffusivity D to be constant. But that's not necessarily the case. In general, we could have a current given by

$$\mathbf{J} = -D(n; \mathbf{x}, t)\nabla n \tag{3.60}$$

where, as shown, D can vary over space and time or even depend on $n(\mathbf{x}, t)$ itself. Here we look at the latter situation. We will consider a diffusivity given by

$$D = k\, n(\mathbf{x}, t) \tag{3.61}$$

with $k > 0$ constant. This means that diffusion is greater when the population density is greater. It's as if the individuals are keen to get away from each other.

We again restrict ourselves to one spatial dimension. The continuity equation now gives the non-linear diffusion equation

$$\frac{\partial n}{\partial t} = -\frac{\partial J}{\partial x} = k\frac{\partial}{\partial x}\left(n\frac{\partial n}{\partial x}\right) \ . \tag{3.62}$$

As before, we can look for self-similar solutions that depend only on a single dimensionless combination of x and t. This time, however, the dimensions of our constants are different. In addition to length L and time T, we'll also need the dimension of $n(x, t)$ which, quite reasonably, we'll denote as n.

The two constants in the game are k and, provided that $J \to 0$ as $x \to \pm\infty$, the total population

$$N = \int_{-\infty}^{+\infty} n(x, t) \ . \tag{3.63}$$

These have dimensions

$$[k] = L^2T^{-1}n^{-1} \quad \text{and} \quad [N] = nL \ . \tag{3.64}$$

From these, we see that to construct a dimensionless variable proportional to x, we must take

$$\xi = \frac{x}{(Nkt)^{1/3}} \ . \tag{3.65}$$

We then take the ansatz

$$n(x, t) = \frac{N}{(Nkt)^{1/3}} f(\xi) \tag{3.66}$$

where the overall factor is designed so that $f(\xi)$ is a dimensionless function.

By now, the path should be a familiar one. We have

$$\frac{\partial \xi}{\partial t} = -\frac{1}{3}\frac{\xi}{t} \quad \text{and} \quad \frac{\partial \xi}{\partial x} = \frac{1}{(Nkt)^{1/3}} = \frac{\xi}{x} \ . \tag{3.67}$$

The time derivative of $n(x,t)$ is then

$$\begin{aligned} \frac{\partial n}{\partial t} &= -\frac{1}{3t}\frac{N}{(Nkt)^{1/3}} f + \frac{N}{(Nkt)^{1/3}} f'(\xi)\frac{\partial \xi}{\partial t} \\ &= -\frac{1}{3t}\frac{N}{(Nkt)^{1/3}}\frac{d}{d\xi}(\xi f) \ . \end{aligned} \tag{3.68}$$

Meanwhile, the first spatial derivative is

$$\frac{\partial n}{\partial x} = \frac{N}{(Nkt)^{2/3}} f'(\xi) \tag{3.69}$$

and so the combination in (3.62) becomes

$$\frac{\partial}{\partial x}\left(n\frac{\partial n}{\partial x}\right) = \frac{1}{(Nkt)^{1/3}}\frac{\partial}{\partial \xi}\left(\frac{Nf}{kt}\frac{df}{d\xi}\right) = \frac{1}{kt}\frac{N}{(Nkt)^{1/3}}\frac{d}{d\xi}(ff') \ . \tag{3.70}$$

Putting these together, the non-linear diffusion equation (3.62) becomes

$$\frac{d}{d\xi}(ff') = -\frac{1}{3}\frac{d}{d\xi}(\xi f) \quad \Longrightarrow \quad ff' = -\frac{1}{3}\xi f \tag{3.71}$$

where we've eliminated an integration constant by requiring that $f, f' \to 0$ as $x \pm \infty$.

We see that we have two different solutions. The first is $f = 0$, which is rather boring. The second is

$$f' = -\frac{\xi}{3} \quad \Longrightarrow \quad f = -\frac{\xi^2}{6} + \text{constant} \ . \tag{3.72}$$

That looks more interesting but, sadly, it doesn't satisfy our boundary conditions $f \to 0$ as $\xi \to \pm\infty$.

What's going on?! We've got one nice quadratic solution that doesn't satisfy the boundary conditions, and one boring solution $f = 0$ that doesn't satisfy the fact that total population is a constant N. Indeed, it's simple to check that the integral (3.63) translates to an integral of $f(\xi)$,

$$\int_{-\infty}^{+\infty} d\xi \ f(\xi) = 1 \ . \tag{3.73}$$

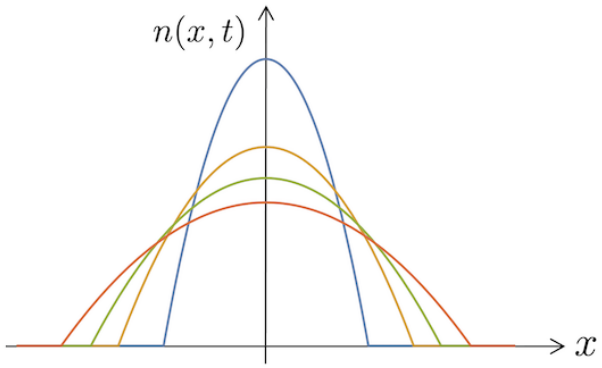


Figure 41. The non-linear diffusion equation results in this weird portion of a parabola expanding outwards.

We can make progress by splicing together these two solutions. We take

$$f(\xi) = \begin{cases} A - \xi^2/6 & |\xi| < \xi_0 = \sqrt{6A} \\ 0 & |\xi| \geq \xi_0 \end{cases} \tag{3.74}$$

The crossover ξ_0 is chosen so that the function is continuous (and, moreover, so that $f(\xi)$ is everywhere non-negative). The derivative of the $f(\xi)$ is discontinuous at ξ_0 and we should really work a little harder to show that this kind of splicing is allowed. The reason it's acceptable can be traced to the condition $f(f' - \xi/3) = 0$ in (3.71) with, roughly speaking, the $f = 0$ beating the fact that f' isn't well-defined at the splice. In particular, the current $J \sim ff'$ vanishes at the point $\xi = \pm\xi_0$. In the mathematics literature, these are sometimes called *weak solutions*, which means that they can be shown to satisfy the original equation in some well-defined sense.

The constant A is fixed by the normalisation condition (3.73),

$$1 = \int_{-\xi_0}^{\xi_0} d\xi \; f(\xi) = 2A\xi_0 - \frac{1}{9}\xi_0^3 = \frac{4\sqrt{6}}{3}A^{3/2} \; . \tag{3.75}$$

This gives $A = (3/32)^{1/3}$ and $\xi_0 = \sqrt{6A} = (9/2)^{1/3}$. This then gives us our final result: going back to the x and t variables, the density $n(x,t)$ takes the shape of finite piece of parabola, spreading out over time

$$n(x,t) = \frac{1}{6}\frac{N^{2/3}}{(kt)^{1/3}}\left[\left(\frac{9}{2}\right)^{2/3} - \frac{x^2}{(Nkt)^{2/3}}\right] \quad \text{for} \quad x < \left(\frac{9}{2}Nkt\right)^{1/3} \; . \tag{3.76}$$

with $n(x,t) = 0$ outside this region. The result is shown in Figure 41.

The end result is slightly odd, not least because we have come to expect that sharp edges are washed out by diffusion, but here the corner at $n(x,t) = 0$ persists for all time. This is a novelty that comes from the non-linear aspect of diffusion. In particular, we have $D \sim n$ so it's not possible for the system to diffuse when $n = 0$. Instead, the population piles up near the edges and, as it grows, diffuses faster.

3.2 Travelling Waves

Once we have both diffusion and some forcing term, interesting things can happen. Before we get to the interesting things, let's look at some boring things.

The one-dimensional reaction-diffusion equation takes the form

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} + F(n) \tag{3.77}$$

for some "reaction" forcing function $F(n)$. An obvious way to proceed is to look for spatially homogeneous solutions, with $n(x,t) = n(t)$. Then the equation becomes

$$\frac{\partial n}{\partial t} = F(n) \ . \tag{3.78}$$

But this is precisely the kind of equation that we started exploring in Section 1. And with just a single variable $n(t)$, there's not a great deal that can happen. We look for fixed points $n^\star$ that obey

$$F(n^\star) = 0 \ . \tag{3.79}$$

Now we can look at perturbations around this fixed point. The novelty is that these perturbations need not be spatially homogeneous: we write

$$n(x,t) = n^\star + \epsilon(x,t) \ . \tag{3.80}$$

Substituting this into (3.77), we get a reaction-diffusion equation for $\epsilon(x,t)$,

$$\frac{\partial \epsilon}{\partial t} = D\frac{\partial^2 \epsilon}{\partial x^2} + \lambda\epsilon \quad \text{with} \quad \lambda = \frac{\partial F}{\partial n}(n^\star) \ . \tag{3.81}$$

But this is the diffusion with linear growth that we already studied in Section 3.1. If we're studying the equation on the domain $x \in \mathbb{R}$ then things are particularly straightforward: the perturbation grows if $\lambda > 0$ and decays if $\lambda < 0$. (If we're instead working on an interval then, as we saw in Section 3.1.3, there is a phase transition in the behaviour as we vary the length of the interval.)

The real interest occurs when the system is unstable because the perturbative analysis above quickly breaks down and fails to tell us what really happens. In this section, we will explore some important examples of reaction-diffusion equations and see some of the novel things that can occur. A recurring lesson will be that non-linear PDEs like (3.77) can offer a much richer experience than boring linear PDEs like (3.81).

3.2.1 The KPP-Fisher Equation

The first non-linear dynamical system that we explored in Section 1 was the logistic equation. That too will be our first non-linear PDE. We will call the dimensionless dynamical variable $p(x,t)$ (rather than $n(x,t)$) and consider the 1d reaction-diffusion equation

$$\frac{\partial p}{\partial t} = \frac{\partial^2 p}{\partial x^2} + p(1-p) \ . \tag{3.82}$$

This is the *KPP-Fisher equation*, with the initials reflecting the important work done by Kolmogorov, Petrovsky and Piskunov.

Fisher originally introduced this equation in 1937 to describe the spread of advantageous genes, with $p(x,t)$ the percentage of the population that carries the gene[12]. We already saw in Section 1.1.2 that the logistic equation describes the spread of a beneficial mutation. The novelty here is the diffusion term which captures how this beneficial mutation evolves in space.

The logistic equation has two fixed points: $p = 0$ and $p = 1$. The latter is stable, the former unstable. The question that we would like to ask is: suppose that we start at $p = 0$ and then perturb the system with a spatially localised disturbance, like the alpha variation of COVID-19 kicking off in Kent. How does it subsequently propagate?

Here is a guess. Suppose that we start with a small disturbance, localised in a region of size L around $x = 0$ at time $t = 0$. From what we've learned about diffusion and growth, we might expect that this perturbation will grow in both height and width, with the top plateauing at the fixed point $p = 1$. If we're sitting at some distance $x \ll L$ from the initial perturbation, we have to wait some time until this perturbation hits us. A cartoon of this dynamics is shown in Figure 42.

[12]The paper is "The wave of advance of advantageous genes", published in the queasily named Annals of Eugenics.

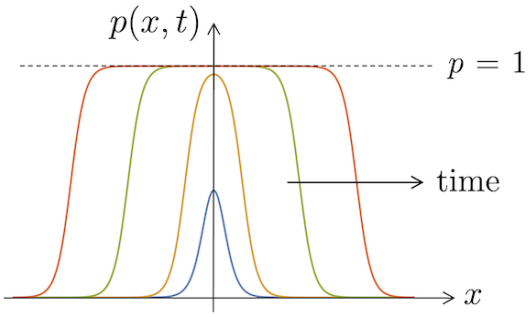


Figure 42. A cartoon of the evolution of a small perturbation in the Fisher equation, spreading out over time as a wavefront.

This suggests that we might look for wave-like solutions to the KPP-Fisher equation. It's worth pointing out that we're not guaranteed that such solutions exist. Indeed, the basic diffusion equation does not support wave-like solutions. But the addition of a reaction term changes the story and, as we'll now show, such waves typically do exist.

We don't know how fast such a wave travels so we'll leave this as arbitrary for now and call it c. We will then look for solutions of the form

$$p(x,t) = f(\xi) \quad \text{with} \quad \xi = x - ct \tag{3.83}$$

with $c > 0$ the as yet unknown wave speed. We don't know if such a solution exists, but it seems like a reasonable place to look. If we substitute this into the Fisher equation, we get an ordinary differential equation

$$-cf' = f'' + f(1-f) \ . \tag{3.84}$$

Our task is to analyse solutions to this equation. Here we offer a number of ways to do this.

Phase Plane Analysis

To start, we can turn our second order differential equation into a pair of first order differential equations,

$$f' = g \quad \text{and} \quad g' = -cg - f(1-f) \ . \tag{3.85}$$

This is the kind of system that we became adept at solving in Section 1 and we know the drill by now. First we look for fixed points. There are two $(f,g) = (0,0)$ and $(f,g) = (1,0)$.

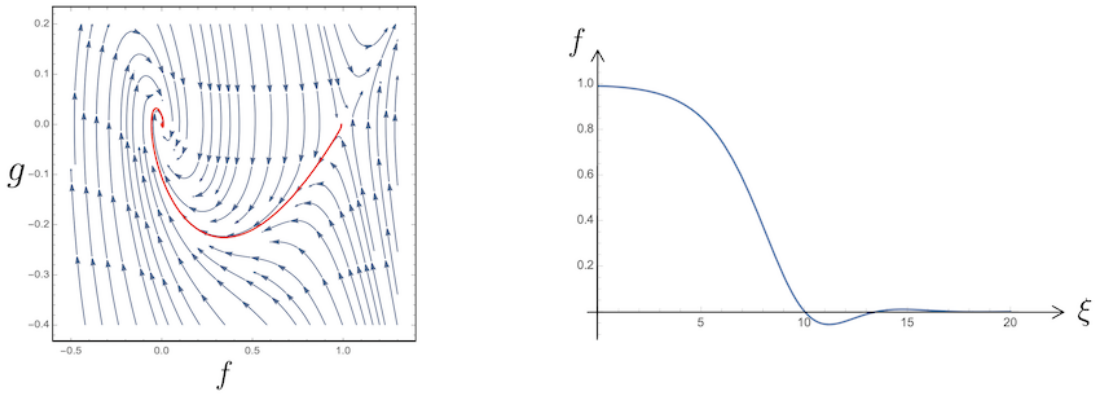


Figure 43. Numerical solutions with $c = 1$, both in the phase plane (on the left) and the evolution of $f(\xi)$. The red line shows the trajectory of a solution that starts close to the unstable fixed point $(f, g) = (1, 0)$.

Next we look at stability. The Jacobian is

$$J = \begin{pmatrix} 0 & 1 \\ -1 + 2f & -c \end{pmatrix} . \tag{3.86}$$

For $(f, g) = (1, 0)$, we have $\det J = -1$ and $\mathrm{Tr}\, J = -c$ so this fixed point is necessarily a saddle.

The other fixed point at $(f, g) = (0, 0)$ is more interesting. The eigenvalues λ of the Jacobian are

$$\lambda^2 + c\lambda + 1 = 0 \quad \Longrightarrow \quad \lambda = -\frac{c}{2} \pm \frac{1}{2}\sqrt{c^2 - 4} \, . \tag{3.87}$$

For $c < 2$, the eigenvalues are complex, with negative real part, so the fixed point is a stable focus, with trajectories spiralling in. For $c \geq 2$, the eigenvalues are real and negative (strictly, one vanishes when $c = 2$) and so the fixed point is stable.

The fact that the flows in the phase plane have qualitatively different behaviour for $c < 2$ and $c > 2$ is important. In particular, we can look at the kind of solutions we get with $c < 2$. These are plotted numerically in Figure 43. While these are fine formal solutions to the Fisher equation, because they spiral into the origin they necessarily have a region of ξ for which $f < 0$. But if we're thinking of $f(\xi)$ as the fraction of a population then we want $f \geq 0$. This means that, for our present purpose, we discard the solutions with $c < 2$.

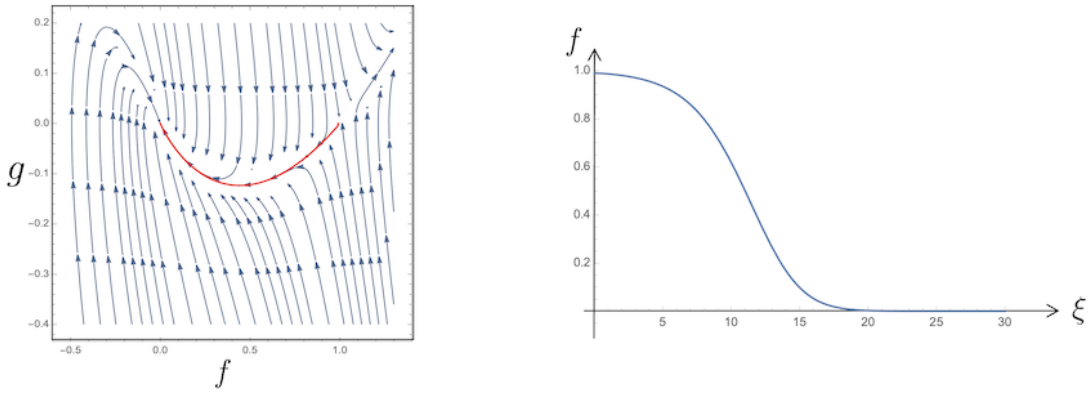


Figure 44. Numerical solutions with $c = 2$, both in the phase plane (on the left) and the evolution of $f(\xi)$. Again, the red line shows the trajectory of a solution that starts close to the unstable fixed point $(f, g) = (1, 0)$.

That leaves us with $c \geq 2$. Here there is no such concern. A numerical plot of this solution (shown with $c = 2$) is depicted in Figure 44. Now there is a solution that starts near the unstable fixed point and heads directly towards the stable fixed point. These are the class of solutions that we will be interested in.

It's worth pointing out that something a little strange has happened here. We wanted to find solutions where we start at $f = 0$ and then perturb slightly to see what happens. Instead, our phase space analysis has resulted in solutions that seem to go the opposite way, with

$$f(\xi) \to 1 \quad \text{and} \quad \xi \to -\infty \quad \text{and} \quad f(\xi) \to 0 \quad \text{as} \quad \xi \to +\infty \ . \tag{3.88}$$

In fact, this is just because the ξ coordinate is defined as $\xi = x - ct$ and that minus sign is the cause for the strange behaviour. For fixed x, these same solutions obey $f(t) \to 0$ as $t \to -\infty$ and $f(t) \to 1$ as $t \to +\infty$.

A Mechanical Analogy

There's a way to translate the story above into something familiar. The equation (3.84) is the kind of thing that we studied in our course on Dynamics and Relativity. If we write it as

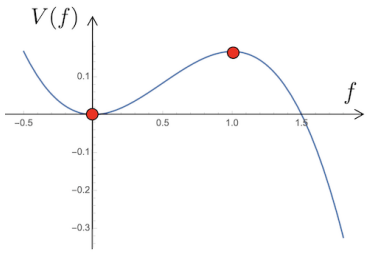
$$f'' = f(f-1) - cf' = -\frac{dV}{df} - cf' \tag{3.89}$$

with

$$V(f) = \frac{1}{2}f^2 - \frac{1}{3}f^3 \ . \tag{3.90}$$

then it looks like the equation of motion for a particle moving in potential V with a friction term $-cf'$.

The potential is plotted in the figure. It has two critical points, at $f = 0$ and at $f = 1$. In this analogy, as in the dynamical system, the point $f = 1$ is unstable and the point $f = 0$ is stable. That's compatible with what we saw above: in the $\xi = x - ct$ coordinate, we go flow from $f = 1$ to $f = 0$ rather than the other way around.



We've seen above that we get qualitatively different behaviour for $c < 2$ and for $c > 2$. It's simple to see why in this mechanical analogy, where c dictates the strength of the friction force. For $c < 2$, the system is "underdamped", meaning that, as it rolls down the hill, it overshoots the minimum at $f = 0$, oscillating back and forth before settling down. This is the behaviour seen in Figure 43.

In contrast, for $c > 2$ the system is overdamped, slowing enough so that it stops when it ultimately reaches the minimum at $f = 0$. The phase plane analysis tells us that the crossover between these two behaviours happens at $c = 2$ when the system has critical damping.

Linearised Analysis

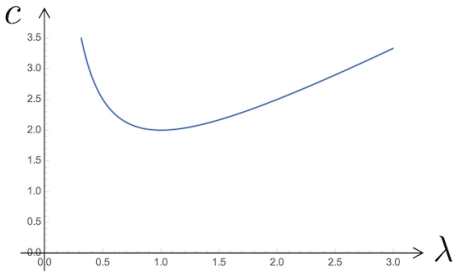
We can learn more about the travelling wave by looking at the leading edge of the wave, where $f \approx 0$. This means that we're looking at the region of the $f(\xi)$ graph in Figure 44 where $f(\xi)$ is approaching the ξ-axis. Here it's appropriate to linearise the equation (3.84) and work with

$$-cf' = f'' + f \ . \tag{3.91}$$

We make the obvious ansatz $f(\xi) = e^{-\lambda\xi}$, with $\lambda > 0$ so that this solution decays towards $f \to 0$ as ξ increases. We see that this solves the equation if

$$\lambda^2 - c\lambda + 1 = 0 \quad \Longrightarrow \quad c = \lambda + \frac{1}{\lambda} \ . \tag{3.92}$$

This is plotted in the figure. We can think of $1/\lambda$ as the width of the wavefront. We learn that the speed and shape of the wave are linked.



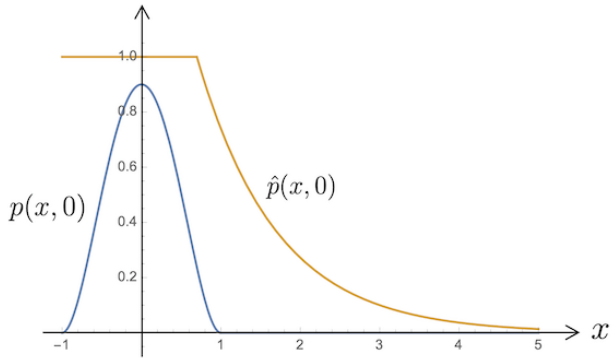


Figure 45. The initial profile $p(x,0)$ that evolves through the full non-linear equation is bounded above by the exponentially decaying profile $\hat{p}(x,0)$ that will evolve through the linearised Fisher equation.

The Speed of the Non-Linear Wave

So far, nothing has told us what speed c the wave travels at if we start with a given initial, localised perturbation $p(x,0)$. We only know that this speed must be $c \geq 2$.

The full non-linear analysis is complicated but the final result, proven by Kolomgorov, is beautifully simple. If we start with some initial conditions that vanish outside of some interval, i.e. $p(x,0) = 0$ for all $|x| > x_0$, the the system will ultimately settle down a wave that travels with speed $c = 2$. In other words, the non-linear system travels at the *slowest* possible speed of the linearised system.

We won't prove this result here, but we can motivate it with the following argument. First note that the non-linear speed must be *one* of the allowed linear speeds $c \geq 2$ just because the linearised analysis is valid at the wavefront. But, with suitably localised initial conditions, we can show that the non-linear speed must be less than (or equal to) that of a linear wave.

To see this, let's take initial conditions that are strictly localised in some region

$$p(x,0) = 0 \quad \text{for all } |x| \geq x_0 \ . \tag{3.93}$$

We'll evolve this with the full non-linear Fisher equation (3.82).

Our strategy is to set this profile in a race against a wavefront with initial profile

$$\hat{p}(x,0) = Ae^{-\lambda x} \ . \tag{3.94}$$

We will pick A so that $p(x,0) < \hat{p}(x,0)$. It's simple to check that this is always possible to construct such a bounding profile for any choice of $\lambda > 0$ simply by picking a suitable A. Crucially, we evolve the profile $\hat{p}(x,t)$ through the *linearised* Fisher equation

$$\frac{\partial \hat{p}}{\partial t} = \frac{\partial^2 \hat{p}}{\partial x^2} + \hat{p} \ . \tag{3.95}$$

Now we set these two profiles off. We will show that it's not possible for the non-linear $p(x,t)$ to overtake the linearised $\hat{p}(x,t)$. At heart, this follows because the missing term in the linearised equation is $-p^2$ and, with the minus sign, this only serves to delay the non-linear evolution.

We can put some meat on this argument by defining $g(x,t) = \hat{p}(x,t) - p(x,t)$. By construction, we have $g(x,0) > 0$. Watching this function evolve in time, we have

$$\frac{\partial g}{\partial t} = \frac{\partial \hat{p}}{\partial t} - \frac{\partial p}{\partial t} = \frac{\partial^2 g}{\partial x^2} + g + p^2 \geq \frac{\partial^2 g}{\partial x^2} + g \ . \tag{3.96}$$

We learn that the evolution of $g(x,t)$ is at least as fast as diffusion with linear growth. And with $g(x,0)$ positive, the function $g(x,t)$ can never go negative. This is telling us that $p(x,t)$ is bounded above by $\hat{p}(x,t)$ for all time. In other words, the non-linear wave can never overtake the linear wave.

But the analysis above holds for any choice of λ and, in particular, for $\lambda = 1$ which travels at the slowest speed $c = 2$. It tells us that the non-linear wave can travel no faster than $c = 2$. But, as we've seen previously, the wave ansatz only makes sense for $c \geq 2$. Hence the non-linear wave must travel at the slowest possible speed $c = 2$.

3.2.2 Front Propagation in Bistable Systems

Here is a simple modification of the Fisher equation

$$\frac{\partial p}{\partial t} = \frac{\partial^2 p}{\partial x^2} - p(p-r)(p-1) \ . \tag{3.97}$$

We'll take $0 < r < 1$.

Now, in the homogeneous system, both $p = 0$ and $p = 1$ are stable fixed points. The intermediate value of $p = r$ is unstable. For this reason, systems of this type are referred to as *bistable*.

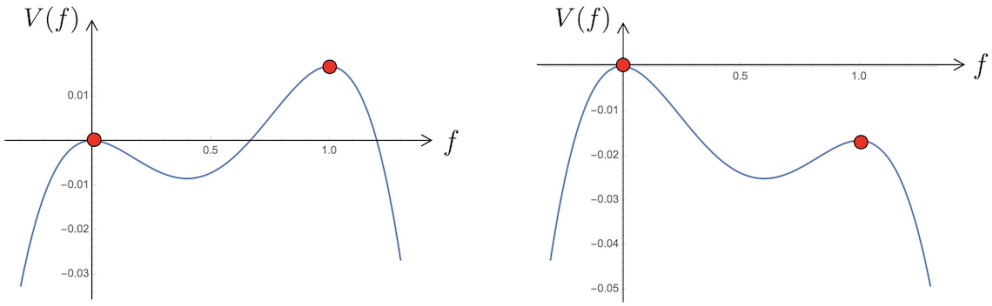
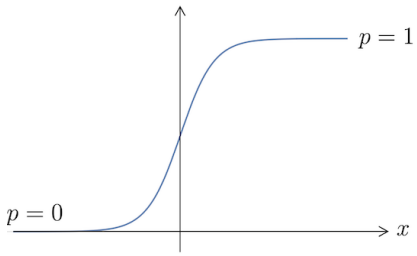


Figure 46. On the left: the potential $V(f)$ plotted with $r = 0.4$. On the right: the potential plotted with $r = 0.6$.



Suppose that we start in a system with $p = 0$ to the left, as $x \to -\infty$, and $p = 1$ to the right, as $x \to +\infty$. In between there has to be a transition, which we will call a "front". It looks something like the plot shown in the figure. The question that we would like to ask is: what happens next? Does the front advance to the left, or to the right? Since both $p = 0$ and $p = 1$ are stable fixed points, it's not immediately obvious which will win. We could think of this as a model for how diseases, mutations, or chemicals spread or die out.

We can use the kind of analysis that we developed for the Fisher equation to answer this. We will again look for a travelling wave solution, with

$$p(x,t) = f(\xi) \quad \text{with} \quad \xi = x - ct \tag{3.98}$$

for some velocity c. With this ansatz, the reaction-diffusion equation (3.97) becomes

$$-cf' = f'' - f(f-r)(f-1) \ . \tag{3.99}$$

This time we can get all the information we need from the mechanical analogy. We write

$$f'' = -\frac{dV}{df} - cf' \quad \text{with} \quad V(f) = -\frac{1}{4}f^4 + \frac{1}{3}(1+r)f^3 - \frac{1}{2}rf^2 \ . \tag{3.100}$$

This potential has two maxima, at $f = 0$ and $f = 1$. But, crucially, the shape of the potential depends on the value of r. For $r < 0.5$, the maximum at $f = 1$ is higher; for $r > 0.5$, the maximum at $f = 0$ is higher. Two representative examples are shown in Figure 46.

When $r < 0.5$, the ball rolls down from $f = 1$. If the friction term, captured by $-cf'$, is large, then the ball will ultimately come to rest at the local minimum at $f = r$. if the friction is low, then the ball will sail past the local maximum at $f = 0$ and into oblivion. In both cases, there is a formal solution to the reaction-diffusion equation but not one with the boundary conditions that we want. However, there is a special value of c such that the friction is just right and the ball rolls down from $f = 1$ where it sat at $\xi \to -\infty$, coming to rest at $f = 0$ at $\xi \to +\infty$. This is the velocity c that we want.

When $r > 0.5$, the heights of the two maxima are inverted, and now there is a critical velocity where the ball rolls the other way: from $f = 0$ to $f = 1$.

But, as in the previous section, we have to remember that the mechanical analogy reverses what actually happens because $\xi = x - ct$. This means that

- For $r < 0.5$ we have, for fixed x, $p(x,t) \to 1$ as $t \to \infty$ and so the front moves to the left.
- For $r > 0.5$ we have, for fixed x, $p(x,t) \to 0$ as $t \to \infty$ and so the front moves to the right.

In contrast to the Fisher equation, the mechanical analogy tells us that the front must travel at a very specific velocity c. But what is it? We can make progress by computing the analog of the work done in our mechanical system. We can take the equation of motion (3.100), multiply by f', and integrate. We have

$$\int_{-\infty}^{+\infty} d\xi \; f'f'' = \int_{-\infty}^{+\infty} d\xi \; \left(-f' \frac{dV}{df} - cf'^2 \right) = \int_{-\infty}^{+\infty} d\xi \; \left(-\frac{dV}{d\xi} - cf'^2 \right) \; . \tag{3.101}$$

The left-hand side is a total derivative of $\frac{1}{2}f'^2$ but our boundary conditions mean that $f' \to 0$ as $\xi \to \pm\infty$. This leaves us with an expression for the velocity

$$c = -\frac{\Delta V}{\int d\xi \; f'^2} \; . \tag{3.102}$$

Here ΔV is the difference in energy between the two maxima,

$$\Delta V = V(f=1) - V(f=0) = \frac{1}{6}\left(\frac{1}{2} - r \right) \; . \tag{3.103}$$

Note that if $r < 0.5$ then $\Delta V > 0$ and so $c < 0$ and the front travels to the left. Meanwhile, if $r < 0.5$ then $c > 0$ and the front travels to the right. This agrees with what we saw above.

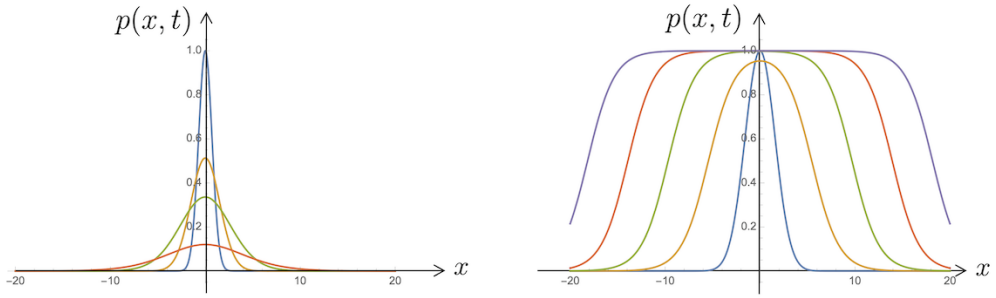


Figure 47. Numerical solutions for $r = 0.2$. The plot on the left starts with the initial condition $p(x, 0) = e^{-x^2}$. The one on the right starts with the wider initial condition $p(x, 0) = e^{-x^2/5}$. We see that the first flounders while the second flourishes.

To determine the actual speed c, we still need to do the integral $\int d\xi\ f'^2$. And, for this, we typically need to first figure out the solution $f(\xi)$. That's not so easy, but there is one value of r for which we can solve the equation of motion exactly. This is the value $r = 0.5$, when the local maxima of the potential have the same height and, correspondingly, the front is static with $c = 0$ and doesn't move. In this case, you can check that the function $f(\xi)$ is given by

$$f = \frac{1}{2}\left(1 - \tanh\left(\frac{\xi}{2\sqrt{2}}\right)\right) \ . \tag{3.104}$$

If we are cheeky and use this as a proxy for the function $f(\xi)$ in the integral, then we get a speed

$$c = 2\left(r - \frac{1}{2}\right) \ . \tag{3.105}$$

Strictly, this calculation holds only when $r = 0.5$ where the speed vanishes! But we can view this as a good approximation to the speed for $r \approx 0.5$. Indeed, if we were more careful we could set up a perturbation expansion in $r - 0.5$, with this the leading order term.

Localised Perturbations

There's a closely related question that we can ask of this system? Suppose that we start in the $p = 0$ state but then introduce a small region of $p = 1$ state, localised around the origin. What then happens? Does this $p = 1$ state expand to take over the world? Or does the $p = 0$ state fight back and shrink it?

The analysis that we've done above goes part way towards answering this. If $r > 0.5$ then the $p = 0$ state will win and the $p = 1$ insurgents will be crushed. Meanwhile, $r < 0.5$ then there's an opportunity for the $p = 1$ state to expand. But it's not guaranteed. That's because our previous analysis assumed that the system settled down to some wavelike behaviour with the fronts propagating at some constant speed. But the question of when this happens depends on how wide the initial localised perturbation is.

We won't give any detailed analysis of this behaviour here, but instead exhibit some simple numerical solutions. Figure 47 shows the time evolution of two different initial perturbations, the one on the left narrower than the one on the right. You can see clearly that the one on the left shrinks over time, while the one on the right thrives, ultimately forming wavefronts that are described by the analysis we did previously.

Dengue Revisited

We can connect the analysis above to one of the stories that we met earlier in these lectures. In Section 1.4.5, we described how Wolbachia bacteria could be introduced to mosquitos to inhibit the spread of dengue fever. Our model of choice was the two-dimensional dynamical system (1.137)

$$\frac{du}{dt} = u\left(u_0 - \frac{v}{u+v} - (u+v)\right) \quad \text{and} \quad \frac{dv}{dt} = v\left(v_0 - (u+v)\right) \ . \tag{3.106}$$

Here u is the population of uninfected mosquitoes and v the infected. The two constants are restricted to lie in $0 < v_0 < u_0 < 1$. (We've switched variables from x and y in (1.137) to u and v above to avoid confusion with the spatial coordinate x that we will soon introduce.)

This is a two-dimensional dynamical system, but it has hidden within it a one-dimensional system that governs the fraction of infected mosquitos

$$p = \frac{v}{u+v} \ . \tag{3.107}$$

You can check that the dynamics of p is governed by

$$\frac{dp}{dt} = -p(p-r)(p-1) \tag{3.108}$$

with $r = u_0 - v_0 \in (0, 1)$. This is precisely the homogeneous system that we discussed above. We can upgrade it to a system that includes spatial localisations by promoting $p(t)$ to a function $p(x, t)$. If we further assume that the population of infected mosquitos diffuses then we are back to solving the equation (3.97). We learn that a localised population of infected mosquitos will only spread if the parameter $r = x_0 - y_0 < 0.5$.

Admittedly this analysis also assumed, perhaps unrealistically, that the mosquitos fly only along a one-dimensional line. You might want to upgrade the system further to allow for 2d diffusion and redo the analysis before you contact the WHO.

Waves in Higher Dimensions

It's reasonably straightforward to extend the analysis above to waves in higher dimensions. Indeed, for plane waves, the story is identical. But for axially symmetric (in 2d) or spherically symmetric (in 3d) there's an additional term that we have to deal with. For example, a 2d axially symmetric system, the Fisher equation reads

$$\frac{\partial p}{\partial t} = \frac{\partial^2 p}{\partial r^2} + \frac{1}{r}\frac{\partial p}{\partial r} + p(1-p) \ . \tag{3.109}$$

Now we get that extra $\partial p/\partial r$ term on the right-hand side. For waves with $c > 0$, on the wavefront, $\partial p/\partial r < 0$. You can see this in the plots in Figure 47 so this additional term acts to slow the wave down. However, it also comes with a $1/r$ which means that it plays less of a role as the spreading circle gets big and the boundary looks more like a plane wave. The upshot is that the wave speed approaches $c \to 2$ as $r \to \infty$.

3.2.3 Wave Propagation in Neurons

We met a model for how neurons are excited in Section 1.7. Our slimmed-down model of choice was due to FitzHugh and Nagumo and was described by the following pair of ordinary differential equations (1.195)

$$\begin{aligned} \epsilon \frac{du}{dt} &= u + v - \frac{1}{3}u^3 \\ \frac{dv}{dt} &= -(u - a + bv) \ . \end{aligned} \tag{3.110}$$

Here $u(t)$ represents the voltage difference across a membrane, while $v(t)$ represents the ease with which various ions can cross the membrane. The constants a and b sit in the region

$$0 < b \leq 1 \ , \quad 1 - \frac{2b}{3} < a < 1 \ . \tag{3.111}$$

Meanwhile, we take $\epsilon \ll 1$ to ensure that the dynamics of $u(t)$ is fast, while that of $v(t)$ is much slower. This separation into two times scales is, in part, responsible for the interesting property of these equation: namely that a small stimulus gives rise to a large response which, in the context of neurons, is known as the action potential.

Now we extend (3.110) to include spatial propagation. First, we take both variables to depend on space and time, so we have $u(t,x)$ and $v(t,x)$. The voltage difference u diffuses, but v does not. Correspondingly, we extend the equations (3.110) to

$$\begin{aligned} \epsilon \frac{du}{dt} &= u + v - \frac{1}{3}u^3 + \epsilon^2 \frac{\partial^2 u}{\partial x^2} \\ \frac{dv}{dt} &= -(u - a + bv) \ . \end{aligned} \tag{3.112}$$

The diffusion term comes with a ϵ^2 to ensure that the spatial and temporal variations are both fast.

We will again search for travelling wave solutions of the form

$$u(x,t) = u(\xi) \quad \text{and} \quad v(x,t) = v(\xi) \quad \text{with} \quad \xi = x - ct \ . \tag{3.113}$$

Substituting this ansatz into the equations (3.112), we arrive at the pair of ordinary differential equations

$$\begin{aligned} -\epsilon c u' &= u + v - \frac{1}{3}u^3 + \epsilon^2 u'' \\ -cv' &= -(u + a + bv) \ . \end{aligned} \tag{3.114}$$

We could eliminate $v(\xi)$ from this pair of equations, but only at the expense of having a third order differential equation for $u(\xi)$. That makes the problem somewhat harder to solve than the analogous equation (3.84) for the Fisher equation.

To proceed, we're going to make use of the two different time scales in the problem, encapsulated in the small $\epsilon \ll 1$ factors in the equations. We know from our previous analysis that the spatially homogeneous system undergoes a cycle as shown in Figure 48 (which is a repeat of Figure 27). We split this cycle into four distinct pieces, labelled AB, BC, CD, and DA. (The beginning and final point are both labelled A even though they are slightly separated.) From the figure, we can see that the transitions AB and CD happed rapidly, while BC and DA are more leisurely. This gives us the motivation to look at each in turn.

Fast from A to B: For the fast motion, it is appropriate to rescale the spacetime coordinate and write

$$\xi = \epsilon\zeta \ . \tag{3.115}$$

With this rescaling, the equations (3.114) become

$$\begin{aligned} -cu' &= u + v - \frac{1}{3}u^3 + u'' \\ -cv' &= -\epsilon(u + a + bv) \end{aligned} \tag{3.116}$$

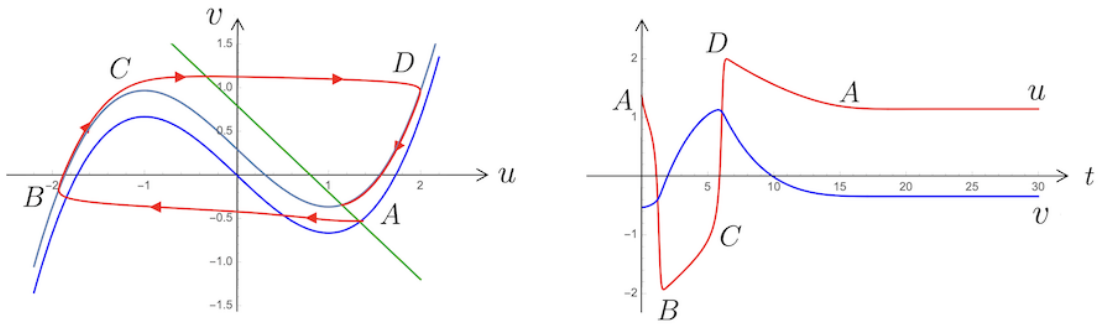


Figure 48. The phase plane motion (on the left, in red) and time dependence (on the right) for the spatially homogeneous model. These are now split into four distinct segments, labelled AB, BC, CD and DA.

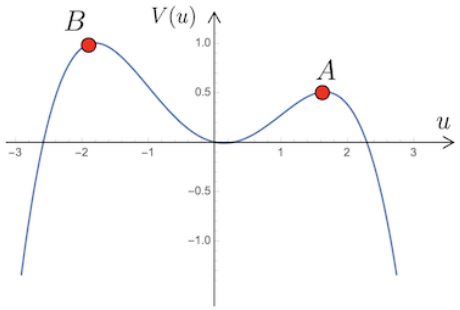
where the primes now denote differentiation with respect to ζ, rather than ξ. The $\epsilon \ll 1$ has now shifted to the right-hand side of the second equation, telling us that v is approximately constant, and equal to its fixed point value $v^\star$, on this segment. Indeed, this can also be seen in the plots of Figure 48 which confirms that the fixed point sits in the region $-1 < v^\star < 0$. This means that the dynamics is captured by the single, second order differential equation

$$-cu' = u + v^\star - \frac{1}{3}u^3 + u'' \ . \tag{3.117}$$

Now we're in business: this is the same kind of equation that we met previously and we have various ways to think about it. For example, if we embrace the mechanical analogy, we write

$$u'' = -\frac{dV}{d\zeta} - cu' \quad \text{with} \quad V(u) = -\frac{1}{12}u^4 + \frac{1}{2}u^2 + v^\star u. \tag{3.118}$$

The potential is plotted in the figure to the right. We see that this is the same kind of bistable system that we met in Section 3.2.2. In the mechanical analogy, the ball rolls down from the global maximum at B towards the local maximum at A. (Recall that things run from B to A, rather than A to B, in the mechanical analogy because our "time" is $\xi = x - ct$ with that minus sign in front of ct.) The is a critical value of the friction coefficient c for



which the ball miraculously ends up perched at the point A. This determines the speed c of the propagating wave. An expression for this critical velocity can be derived using (3.102).[13]

Slow from B to C: The next phase of the cycle, from B to C. For this, we should return to our original unscaled variable ξ and drop the terms proportional to $\epsilon \ll 1$ in (3.114). We're left with

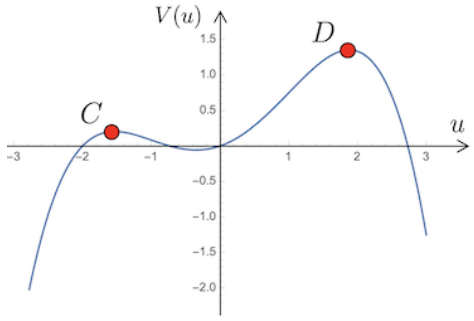
$$v = \frac{1}{3}u^3 - u \quad \text{and} \quad -cv' = -(u + a + bv) \ . \tag{3.119}$$

The first of these equations is telling us that the dynamics lies on the slow manifold, which is the cubic shown on the left of Figure 48. The second equation is just a first order system. We learn that the value of v slowly increases.

The question that we would like to ask is: at what value of v does the system exit this slow motion? Let's call this value $v^{\star\star}$. You might naively think that this is the same value of $v^{\star\star}$ that appears in the spatially homogeneous system shown in Figure 48. But that turns out *not* to be the case. To see why, we need to turn to the next phase of the cycle.

Fast from C to D: The transition from C to D is again fast and we can use the same analysis as we saw from A to B. We're again left with the equation (3.117), but now with the fixed point $v^\star$ replaced by the new value $v^{\star\star}$

$$-cu' = u + v^{\star\star} - \frac{1}{3}u^3 + u'' \ . \tag{3.120}$$



We will have $v^{\star\star} > 0$ (which contrasts with $v^\star < 0$. This ensures that the effective potential in (3.118) has the positions of the local and global maxima swapped, so that the dynamics goes the other way. An example of such a potential is shown on the right. But what is the value of $v^{\star\star}$?

At this point, we use a variant of our previous argument. In the mechanical analogy, we want to roll from point D to point C with a critical value of the "friction" c so

[13] You can watch an experiment measuring the velocity of these action potentials in earthworms in this Youtube video. (Be warned: it does involve sticking pins into an anesthetised earthworm, although the author of the video promised that the worm was then released back into the garden where it lived a long and rewarding life.)

that we end up perched precisely at the local maximum C. Usually we would use this argument to determine c. But, in the present case, we already know the value of c: it was determined by our analysis of the AB part of the cycle. So now we replay this argument, but now varying the value of $v^{\star\star}$, so that the effective potential changes until the two maxima are exactly the right height to allow a trajectory from D to C that works for the chosen value of c. In that we, we find the value of $v^{\star\star}$ that allows the entire pulse – from A to B to C to D to A again – to all propagate at the same speed c.

Slow from D to A: The final part of the story, from D to A, is again a gentle relaxation along the slow manifold. This is described by the pair of equations (3.119). This concludes the cycle. The end result is that all parts of the pulse propagate at the speed c that was determined by the $A \to B$ part of the analysis.

3.3 Turing Instability

As we have seen, diffusion encourages things to spread out, damping any wild spatial variations. This makes it rather surprising that, in the right circumstances, diffusion can render a system unstable giving rise to spatial variations. This is known as the Turing instability.

For this to occur, we need multiple variables. We'll consider the simplest such system with just two variables. $u(\mathbf{x}, t)$ and $v(\mathbf{x}, t)$, subject to the reaction-diffusion equations

$$\frac{\partial u}{\partial t} = f(u, v) + D_1 \nabla^2 u \quad \text{and} \quad \frac{\partial v}{\partial t} = g(u, v) + D_2 \nabla^2 v \ . \tag{3.121}$$

We have two diffusion constants, D_1 and D_2, and two functions $f(u, v)$ and $g(u, v)$ that determine the dynamics.

We will assume that, in the absence of diffusion, the dynamics admits a stable fixed point $(u^\star, v^\star)$ such that

$$f(u^\star, v^\star) = g(u^\star, v^\star) = 0 \ . \tag{3.122}$$

The Jacobian is

$$J = \begin{pmatrix} f_u & f_v \\ g_u & g_v \end{pmatrix}_{(u^\star, v^\star)} \tag{3.123}$$

where we're using the notation $f_u = \partial f / \partial u$ and so on. The statement that the fixed point is stable means that we have

$$\begin{aligned} \text{Tr}\, J &= f_u + g_v < 0 \\ \det J &= f_u g_v - f_v g_u > 0 \ . \end{aligned} \tag{3.124}$$

Now we consider perturbations that are not spatially homogeneous. We will write

$$u(\mathbf{x},t) = u^\star + \hat{u}(t)e^{i\mathbf{k}\cdot\mathbf{x}} \quad \text{and} \quad v(\mathbf{x},t) = v^\star + \hat{v}(t)e^{i\mathbf{k}\cdot\mathbf{x}} \ . \tag{3.125}$$

This notation deserves some explanation because suddenly, without warning, our variables $u(\mathbf{x},t)$ and $v(\mathbf{x},t)$ appear to be complex! This is just a trick. We will work with the form only when linearising about the fixed point, which means that we're always able to take real and imaginary parts at will. We could just as easily write $\cos(\mathbf{k}\cdot\mathbf{x})$ instead of $e^{i\mathbf{k}\cdot\mathbf{x}}$. We prefer the latter because it's typically easier to work with exponentials than trigonometric functions.

What happens with such a perturbation? If we substitute this ansatz into (3.121) and keep only terms linear in $\hat{u}$ and $\hat{v}$, then we get the equations

$$\begin{aligned} \frac{d\hat{u}}{dt} &= f_u\hat{u} + f_v\hat{v} - D_1k^2\hat{u} \\ \frac{d\hat{v}}{dt} &= g_u\hat{u} + g_v\hat{v} - D_2k^2\hat{v} \end{aligned} \tag{3.126}$$

with $k^2 = \mathbf{k}\cdot\mathbf{k}$ and where f_u and other partial derivatives are all evaluated at the fixed point.

Your first inclination is to think that the diffusion terms only make the system more stable. After all, they both appear with negative signs on the right-hand side! But that's not the way these things work! To check stability, we have to look at the modified Jacobian,

$$J_{\text{new}} = \begin{pmatrix} f_u - D_1k^2 & f_v \\ g_u & g_v - D_2k^2 \end{pmatrix} \ . \tag{3.127}$$

This has

$$\text{Tr}\, J_{\text{new}} = \text{Tr}\, J - (D_1 + D_2)k^2 < 0 \ . \tag{3.128}$$

So the trace is indeed more negative. This is what our previous intuition was telling us. However, the determinant is not so straightforward. It is

$$\begin{aligned} \det J_{\text{new}} &= (f_u - D_1k^2)(g_v - D_2k^2) - f_vg_u \\ &= \det J - k^2(D_1g_v + D_2f_u) + D_1D_2k^4 \ . \end{aligned} \tag{3.129}$$

The system is stable only if $\det J_{\text{new}} > 0$. But we see that this is not guaranteed! There's that middle term which could possibly turn the whole thing negative.

To see if this happens, we write

$$\det J_{\rm new} = Ak^4 - Bk^2 + C \ . \tag{3.130}$$

This dips below zero for some region of $k^2 > 0$ only if $B > 0$ and the quadratic has real roots. This, in turn, requires that the discriminant $B^2 - 4AC$ is positive. Translated back to the expression (3.129), we learn that the determinant can be negative for certain values of k provided that

$$D_1 g_v + D_2 f_u > \sqrt{4D_1 D_2 (f_u g_v - f_v g_u)} \ . \tag{3.131}$$

This is the condition for *Turing Instability*.

Note that the system is stable to both long wavelength modes (with $k \ll 1$) and short wavelength modes (with $k \gg 1$). The long wavelength modes are close to the homogeneous system which is known to be stable. The short wavelength modes are eliminated quickly by diffusion. The instability occurs only in some intermediate regime,

$$k_\star^2 - \Delta < k^2 < k_\star^2 + \Delta \tag{3.132}$$

with

$$k_\star = \sqrt{\frac{B}{2A}} \quad \text{and} \quad \Delta = \frac{\sqrt{B^2 - 4AC}}{2A} \ . \tag{3.133}$$

You could imagine starting with a system that does not exhibit the Turing instability, and then slowly varying parameters until you reach the phase transition at $B^2 = 4AC$. At this point, the modes with wavenumber

$$k_\star = \left(\frac{C}{A}\right)^{1/4} = \left(\frac{\det J}{D_1 D_2}\right)^{1/4} \tag{3.134}$$

are the first to become unstable.

Instability Requires Different Diffusivities

The Turing instability feels counterintuitive. At heart, the idea is that the matrix J has eigenvalues with negative real parts, but the matrix

$$J_{\rm new} = J - k^2 \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix} \tag{3.135}$$

does not. It feels like subtracting things off the diagonal should only decrease both eigenvalues. But, as we have seen, it's possible to decrease one and increase the other. That's the algebraic crutch on which the instability relies

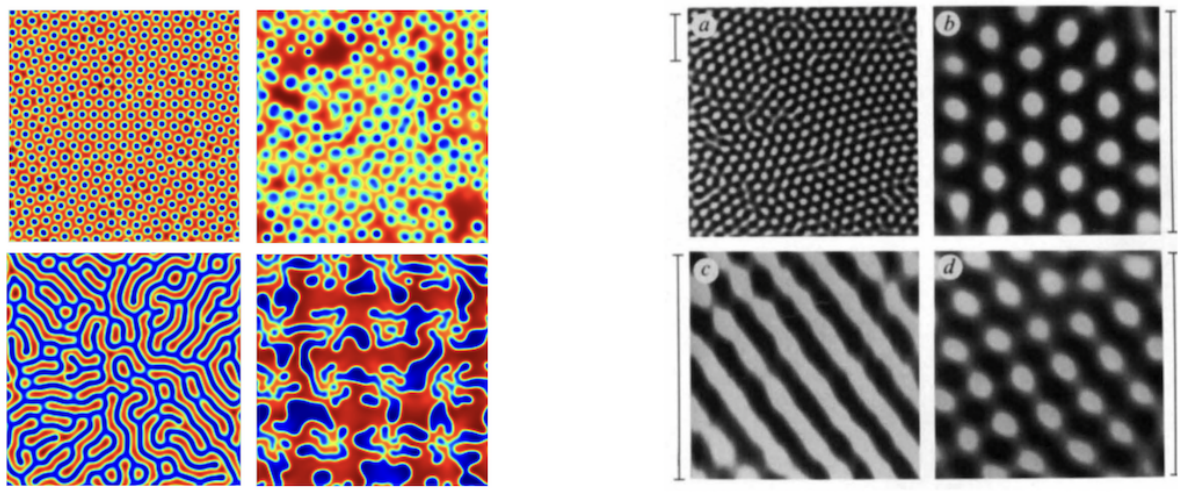


Figure 49. On the left: numerical simulations of various reaction-diffusion systems with a Turing instability. On the right: Turing instabilities seen in chemical reactions.

For this to happen, however, it's important that $D_1 \neq D_2$. Indeed, if $D_1 = D_2 = D$ then we have $J_{\rm new} = J - k^2 D\mathbb{1}$. If the eigenvalues of J are λ_i then the eigenvalues of $J_{\rm new}$ are $\lambda_i - k^2 D$ and both decrease.

Another way of seeing this is to define the ratio of diffusivities

$$d = \frac{D_1}{D_2} \ . \tag{3.136}$$

Then the condition for instability (3.131) becomes

$$f_u + dg_v > 2\sqrt{d \det J} \ . \tag{3.137}$$

The right-hand side is clearly greater than zero so we must have

$$f_u + dg_v > 0 \ . \tag{3.138}$$

Meanwhile, we know that

$$\mathrm{Tr}\, J = f_u + g_v < 0 \ . \tag{3.139}$$

So clearly $d = 1$ does not do the job. For an instability to kick, the diffusivities must be suitably different. Moreover, this simple algebra highlights what's actually going on. For the instability to occur, one of f_u and g_v must be positive and the other negative. Suppose that we have $f_u > 0$ and $g_v < 0$. Then the variable u is called an *activator*

Figure 50. This is Eddie, possibly Turing unstable, definitely not Turing complete.

because it wants to increase things while the variable v is called the *inhibitor* because it wants to reduce things. The result (3.139) is telling us that, in the homogeneous situation, the inhibitor wins. But if we take $d \ll 1$, so the inhibitor diffuses faster than the activator, then the result (3.137) is telling us that, ultimately the activator can triumph.

3.3.1 Pattern Formation

Our analysis above only finds the instability. Obviously the next question is: what becomes of it? We know that the system will necessarily become inhomogeneous and we might expect that this happens at a characteristic wavelength

$$\lambda_\star \approx \frac{2\pi}{k_\star} \tag{3.140}$$

To understand the resulting pattern, we need to study the full non-linear behaviour of a system. And this typically means doing numerics. Some examples of patterns from various numerical simulations, together with experimental results seen in certain chemical reactions are shown in Figure 49.[14]

It is thought that other pattens seen in nature, including animals coats, can be traced to a Turing diffusion instability of the kind described above.

[14]The numerical plots are taken from this blogpost which includes the script used to create them. The experimental data is taken from the paper Transition from a uniform state to hexagonal and striped Turing patterns, by Q. Ouyang and H. Swinney, Nature, 1991.

Effects of a Boundary

In the above discussion, we took our perturbations (3.125) proportional to $e^{i\mathbf{k}\cdot\mathbf{x}}$. This is appropriate for an infinite domain. But it may be that we want to solve the reaction-diffusion equation on some finite domain. Typically we would then impose Neumann boundary conditions $\mathbf{n}\cdot\nabla u = \mathbf{n}\cdot\nabla v = 0$ on the boundary ∂V, where $\mathbf{n}$ is the vector normal to the boundary. We should now take the perturbations to be eigenfunctions of the Laplacian on this domain, rather than $e^{i\mathbf{k}\cdot\mathbf{x}}$.

In a finite domain, the eigenvalues k^2 will typically be discrete and we get an instability if one of them lies in the window (3.132). The lowest eigenvalue k_0 is set by the size L of the domain,

$$k_0 \approx \frac{2\pi}{L} \ . \tag{3.141}$$

If the system is suitably small, the we will have $k_0^2 > k_\star^2 + \Delta$ and no instability will occur. This mimics what we saw in Section 3.1.3, where linear growth does not lead to an instability for small system sizes.

Suppose, for example, that the domain consists of a rectangle with $x \in [0, L_1]$ and $y \in [0, L_2]$. Then the eigenvalues of the Laplacian are $e^{i\mathbf{k}\cdot\mathbf{x}}$ but with $\mathbf{k}$ quantised as

$$k_1 = \frac{m\pi}{L_1} \quad \text{and} \quad k_2 = \frac{n\pi}{L_2} \quad m, n \in \mathbb{Z} \ . \tag{3.142}$$

If the rectangle is reasonably large and square, then one would expect to find k_1 and k_2 sitting within the window of instability (3.132). As we've stressed, it's far from clear what the resulting pattern would be but suppose that we end up with spots. Now consider a narrower rectangle, say with $L_2 \ll L_1$. Then we could have a situation in which there is no instability in the y-direction, but only in the x-direction, resulting in stripes rather than spots. This suggests that if the Turing instability is responsible for animal patterns then we might expect to see animals with spotted coats, but striped tails.

3.3.2 An Example

Our discussion above was rather abstract. It's useful to have one example where we can see this in practice. We take

$$\frac{\partial u}{\partial t} = -u + u^2 v + \nabla^2 u \quad \text{and} \quad \frac{\partial v}{\partial t} = b - u^2 v + d\nabla^2 v \ . \tag{3.143}$$

The model depends on just two parameters, b and the ratio of diffusivities d. You can think of this as modelling a chemical reaction, with two molecules u and v combining

as $2u + v \to 3u$. In addition, the u molecule decays at a constant rate (hence the $-u$ term) and the v molecule is produced at a constant rate (the $+b$ term). This is variant of the so-called Gray-Scott model.

First we look at the homogeneous situation. There is a single fixed point $(u^\star, v^\star) = (b, 1/b)$ with Jacobian

$$J = \begin{pmatrix} -1 + 2u^\star v^\star & u^{\star\,2} \\ -2u^\star v^\star & -u^{\star\,2} \end{pmatrix} = \begin{pmatrix} 1 & b^2 \\ -2 & -b^2 \end{pmatrix} \ . \tag{3.144}$$

We have $\mathrm{Tr}\, J = 1 - b^2$. And we have $\det J = b^2 > 0$. So the fixed point is stable for $b > 1$.

Now we include the diffusion term. Our modified Jacobian becomes

$$J_{\rm new} = \begin{pmatrix} 1 - k^2 & b^2 \\ -2 & -b^2 - dk^2 \end{pmatrix} \ . \tag{3.145}$$

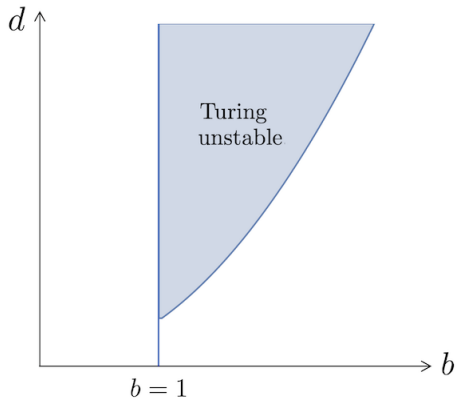
The determinant is

$$\det J_{\rm new} = dk^4 - (d - b^2)k^2 + b^2 \ . \tag{3.146}$$

We require $d > b^2 > 1$ for the roots to be positive. Indeed, we knew from our general discussion that $d = 1$ wouldn't do the job because the diffusivities had to be different. This quadratic (in k^2) has roots provided that

$$d - b^2 \geq 2b\sqrt{d} \ . \tag{3.147}$$

A little bit of algebra shows that this holds provided $d > (3 + 2\sqrt{2})b^2$ (using the fact that $b > 1$). The region of parameter space that exhibits the Turing instability is shown in the figure.



3.4 Chemotaxis

Bacteria and other single-celled organisms may be small, but they know what they like. They will happily swim towards nutrients, or away from poison, following the gradients of the chemical. This process is known as *chemotaxis*.

We will model the situation with two variables: $n(\mathbf{x}, t)$ is the density of bacteria and $c(\mathbf{x}, t)$ is the density of some chemical that they have developed a fondness for. For the chemical, we will have the usual kind of reaction-diffusion equation

$$\frac{\partial c}{\partial t} - D_c \nabla^2 c = G(n, c) \tag{3.148}$$

where $G(n, c)$ is some function that specifies the dynamics. However, for the bacteria we will include an additional term in the equation that captures the fact they swim towards the chemical. We have the general form

$$\frac{\partial n}{\partial t} + \nabla \cdot \mathbf{J} = F(n, c) \tag{3.149}$$

where $F(n, c)$ is a function that we will specify shortly, while the current is given by the now-familiar diffusion term, plus something else:

$$\mathbf{J} = -D_n \nabla n + n\mathbf{v} \ . \tag{3.150}$$

Here $\mathbf{v}$ is the velocity of the bacteria. They swim in the direction in which the chemical concentration is greatest, so we take

$$\mathbf{v} = \chi \nabla c \tag{3.151}$$

with $\chi > 0$ a constant.

Although the mathematical expression (3.151) is natural to write down, it's worth pausing to ask how it comes about physically. After all, bacteria are around 10^{-6} m long. It seems unlikely that they come equipped with sensitive equipment to allow them to detect the gradient of chemicals across such a small distance and then swim towards the place it's greatest! So what's the physics that leads to (3.151)?

The answer is pretty cool. Single-celled organisms with flagella have two modes of transport, called "runs" and "tumbling". On a run, they point themselves in one direction and motor along at a speed of around 10 - 50 μm per second. But every second or so, they stop and do a tumble, reorienting themselves in a random direction, before they head off again. The net effect is a random walk. The way they swim in the direction of greater chemical concentration is by reducing the frequency of tumbling when the concentration is higher, ensuring that they spend a longer time where the climate is nice.

The upshot is that our set of chemotaxis equations read

$$\begin{aligned}\frac{\partial n}{\partial t} &= D_n\nabla^2 n - \chi\nabla\cdot(n\nabla c) + F(n,c)\\ \frac{\partial c}{\partial t} &= D_c\nabla^2 c + G(n,c)\ .\end{aligned} \tag{3.152}$$

The novelty is, of course, the chemotaxis term proportional to χ. Our goal is to understand what qualitative effect this has on the dynamics., We will see that it provides another avenue for the system to become unstable.

3.4.1 An Example

We will explore the chemotaxis equations by looking at an example. We take

$$F(n,c) = \gamma - \delta n \quad \text{and} \quad G(n,c) = \alpha n - \beta c \tag{3.153}$$

with $\alpha, \beta, \gamma, \delta > 0$. The idea here is the bacteria are being constantly injected into the system, and then die at some rate δ. Meanwhile the bacteria are extreting their favourite chemical at a rate α, which is subsequently decaying at a rate β.

To start, we again assume that the system is homogenous and look for a fixed point $(n^\star, c^\star)$ of the local dynamics such that $F(n^\star, c^\star) = G(n^\star, c^\star) = 0$. There's a unique fixed point

$$(n^\star, c^\star) = \left(\frac{\gamma}{\delta}, \frac{\alpha\gamma}{\beta\delta}\right)\ . \tag{3.154}$$

The Jacobian at this fixed point is

$$J = \begin{pmatrix} -\delta & 0 \\ \alpha & -\beta \end{pmatrix}\ . \tag{3.155}$$

We see immediately that the two eigenvalues are $-\delta$ and $-\beta$ and both are negative. So our fixed point is stable.

Now we look at the effect of the gradient terms and, in particular, the chemotaxis term. We will perturb about the fixed point but, as in the previous section, allow our perturbations to vary in both space and time. We write

$$n(\mathbf{x},t) = n^\star + u(t)e^{i\mathbf{k}\cdot\mathbf{x}} \quad \text{and} \quad c(\mathbf{x},t) = c^\star + v(t)e^{i\mathbf{k}\cdot\mathbf{x}}\ . \tag{3.156}$$

We take both $u(t)$ and $v(t)$ to be small which means that, when substituting into the chemotaxis equations (3.152), we drop all terms quadratic or higher in these variables. We're then left with the following pair of equations

$$\begin{aligned}\frac{\partial u}{\partial t} &= -D_n k^2 u + \chi n^\star k^2 v - \delta u \\ \frac{\partial v}{\partial t} &= -D_c k^2 + \alpha u - \beta v \ .\end{aligned} \tag{3.157}$$

Now we have a slightly different dynamical system, with a modified Jacobian matrix

$$J_{\rm new} = J - k^2 \begin{pmatrix} D_n & -\chi\gamma/\delta \\ 0 & D_c \end{pmatrix} \ . \tag{3.158}$$

If the system is to be stable against spatially varying perturbations, then this new Jacobian $J_{\rm new}$ must also have negative eigenvalues. But does it?

First note that, in the absence of chemotaxis, with $\chi = 0$, the eigenvalues are again negative. So this system does not exhibit a Turing instability. But what happens when $\chi \neq 0$?

To compute the eigenvalues, we can look at

$$\mathrm{Tr}\, J_{\rm new} = -(\beta + \delta) - k^2(D_n + D_c) < 0 \ . \tag{3.159}$$

So there's certainly no problem there. But the determinant gives

$$\det J_{\rm new} = (\beta + k^2 D_n)(\delta + k^2 D_c) - \frac{\alpha\gamma\chi}{\delta} k^2 \ . \tag{3.160}$$

The system is only stable if $\det J_{\rm new} > 0$. But we see that the chemotaxis term contributes with a minus sign. If you make χ big enough, then there is guaranteed to a window of k values for which the system goes unstable. Indeed, the mode with wavenumber k is unstable if

$$\frac{\alpha\gamma\chi}{\beta} > \frac{(\beta + k^2 D_n)(\delta + k^2 D_c)}{k^2} \ . \tag{3.161}$$

If we start with χ small and then slowly increase it, the first mode to go unstable is $k_\star$ which is the minimum of the function on the right-hand side above. A short calculation shows that this is given by

$$k_\star = \left(\frac{\beta\delta}{D_n D_c}\right)^{1/4} \ . \tag{3.162}$$

If we substitute this into (3.161), we find that the minimum value before the instability kicks in is

$$\chi_{\min} = \frac{\delta}{\alpha\beta}\left(\sqrt{\delta D_c} + \sqrt{\beta D_n}\right)^2 \ . \tag{3.163}$$

For $\chi > \chi_{\min}$, the system will again settle down to some spatially inhomogeneous configuration.

4 Random Variations

The differential equations that we've worked with so far in these lectures are deterministic. You set the initial conditions and what then follows is set in stone. I know some people whose lives are like that, but most things in the biological world are not. Instead, important features of our lives are dictated by randomness, the kind of event that, to quote the classics, blindsides you at 4pm on an idle Tuesday.

The purpose of this section is to learn how to incorporate such random fluctuations into our equations. We will do this by studying the evolution of probability distributions over the space of outcomes. Throughout, we will take time to be continuous but the outcomes themselves may be either discrete or continuous. In the latter case, the probability distribution will be governed by the famous Fokker-Planck equation.

4.1 Discrete Outcomes

We start by considering the situation where the possible outcomes are discrete. We will build up slowly, first considering just two possible outcomes, then ∞, then ∞^2.

4.1.1 Two Outcomes

Suppose that there are just two possible states in our system, A and B. We would like to understand the probability $P(A,t)$ to be in state A and the related probability $P(B,t) = 1 - P(A,t)$ to be in state B.

For this, we need to stipulate the underlying dynamics which tells us how the system evolves between A and B. This too will be probabilistic. We will assume that we are dealing with a *Markov process*, meaning that the probability to transition from one state to the other depends only on the current state. For simplicity, we consider the following rules.

- If in state A, the system has a probability per unit time λ to transition to state B.
- Once in state B, the system stays there.

The real purpose of this warm-up example is to understand what we mean by "probability per unit time". In a short time δt, the probability that we jump from A to B is $\lambda\,\delta t$. Equivalently, in the same short time δt, the probability that we remain in state A is $1 - \lambda\,\delta t$.

From this information, we can write down a differential equation that governs the probability. If we know $P(A,t)$ at time t then, at time $t+\delta t$, the probability is

$$P(A,t+\delta t) = P(A,t)(1-\lambda\,\delta t)\ . \tag{4.1}$$

If we now Taylor expand the left-hand side, we have

$$P(A,t) + \frac{dP(A,t)}{dt}\delta t = P(A,t)(1-\lambda\,\delta t) \quad \Longrightarrow \quad \frac{dP(A,t)}{dt} = -\lambda P(A,t)\ . \tag{4.2}$$

This is easily solved. If we start off most definitely in state A, so $P(A,0)=1$, then we have

$$P(A,t) = e^{-\lambda t}\ . \tag{4.3}$$

We see that a constant probability per unit time to jump from A to B means an exponential depletion of A.

There are further questions that we can ask of this simple system. We could, for example, ask for the probability distribution $f(t)$ for the time t that we make the jump from A to B. To get this, we first consider the probability that we made the jump at some time $t<T$,

$$\text{Prob}[t<T] = \int_0^T f(t)\ dt\ . \tag{4.4}$$

But this can be identified with the probability that we're in state B at time T,

$$\text{Prob}[t<T] = P(B,T) = 1-e^{-\lambda T}\ . \tag{4.5}$$

Equating these two expressions and differentiating (and, perhaps confusingly, replacing the dummy variable T with t), gives

$$f(t) = \frac{dP(B,t)}{dt} = \lambda e^{-\lambda t}\ . \tag{4.6}$$

This is the probability distribution. It obeys

$$\int_0^\infty f(t)\ dt = 1 \tag{4.7}$$

as probability distributions should. From this, we can easily round up the usual statistical suspects. The expected time to make the jump is

$$\langle t\rangle = \int_0^\infty t f(t)\ dt = \frac{1}{\lambda}\ . \tag{4.8}$$

Meanwhile, the variance is given by

$$\mathrm{Var}(t) = \langle t^2 \rangle - \langle t \rangle^2 = \int_0^\infty t^2 f(t)\, dt - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \ . \tag{4.9}$$

We see that the standard deviation $\sigma(t) = \sqrt{\mathrm{Var}(t)} = 1/\lambda$ is the same as the mean. This is telling us that fluctuations are important in this system. If we knew only about the average time to jump, this wouldn't agree particularly well with observations in any given case.

4.1.2 Discrete Population Size

With this simple example under our belts, let's now turn to a situation where the outcomes are labelled by $n \in \mathbb{N} = \{0, 1, 2, \ldots\}$. (As an aside: mathematicians can't make up their minds whether or not zero is a natural number. Here I have decided for them.) We can think of n as labelling the population size.

Again, we need to specify the dynamics of the system. Here we take constant probability per unit time λ to jump from n to $n+1$. This is known as a *Poisson process*.

(You might reasonably argue that it is unrealistic for a population to jump from $n = 0$ to $n = 1$. You might, for that matter, argue that it's equally unrealistic for most populations to jump from $n = 1$ to $n = 2$. If you're worried, think "immigration" rather than "birth".)

Again, we can translate this statement into a differential equation. We have, for $n \geq 1$,

$$P(n, t+\delta t) = (1 - \lambda\, \delta t) P(n,t) + \lambda \delta t\, P(n-1, t) \ . \tag{4.10}$$

Here the first term captures the probability that we remain in state n, while the second captures the probability that we jump up from state $n-1$. Taylor expanding the left-hand side then gives us the differential equation

$$\frac{dP(n,t)}{dt} = \lambda \big[P(n-1,t) - P(n,t) \big] \ . \tag{4.11}$$

Equations like (4.11) (or (4.2)) that govern the evolution of a probability distribution are called, rather pompously, *master equations*. In the present case, it is a differential equation in t and a difference equation in n.

The Generating Function

The most systematic way to solve the master equation (4.11) is to introduce the *generating function*

$$\phi(s,t) = \sum_{n=0}^{\infty} s^n P(n,t) = \langle s^n \rangle \tag{4.12}$$

where we think of $s \in [0,1]$, ensuring convergence. The generating function is a lovely object that cleanly captures many of the things we most care about in the distribution. For example, the average population size is

$$\langle n(t) \rangle = \sum_{n=0}^{\infty} nP(n,t) = \left. \frac{\partial \phi(s,t)}{\partial s} \right|_{s=1} . \tag{4.13}$$

Similarly,

$$\begin{aligned} \langle n^2(t) \rangle &= \sum_{n=0}^{\infty} n^2 P(n,t) \\ &= \sum_{n=0}^{\infty} n(n-1)P(n,t) + \sum_{n=0}^{\infty} nP(n,t) \\ &= \left. \frac{\partial^2 \phi(s,t)}{\partial s^2} \right|_{s=1} + \left. \frac{\partial \phi(s,t)}{\partial s} \right|_{s=1} . \end{aligned} \tag{4.14}$$

Combining these, we can extract the standard deviation.

If we know the generating function then we can easily reconstruct the probability distribution by differentiating

$$P(n,t) = \left. \frac{1}{n!} \frac{\partial^n}{\partial s^n} \phi(s,t) \right|_{s=0} . \tag{4.15}$$

Finally, the generating function satisfies a boundary condition at $s = 1$ that comes from the observation that

$$\phi(1,t) = \sum_{n=0}^{\infty} P(n,t) = 1 \ . \tag{4.16}$$

There may be an additional boundary condition at $t = 0$ coming from an initial condition on the probability distribution.

With this in mind, let's now return to our master equation (4.11) and use it to construct a differential equation for the generating function. We have

$$\begin{aligned}\frac{\partial\phi(s,t)}{\partial t} &= \sum_{n=0}^{\infty} s^n \frac{\partial P(n,t)}{\partial t} \\ &= \lambda \sum_{n=0}^{\infty} s^n \big[P(n-1,t) - P(n,t)\big] \\ &= \lambda \sum_{n=0}^{\infty} \big[s^{n+1}P(n,t) - s^n P(n,t)\big] \\ &= \lambda(s-1)\phi(s,t) \ . \end{aligned} \tag{4.17}$$

where, in the second line, we've used $P(-1,t) = 0$. This is a differential equation in t. We can simply integrate it, treating s as a constant to get

$$\phi(s,t) = \phi(s,0)e^{\lambda(s-1)t} \ . \tag{4.18}$$

The function $\phi(s,0)$ is fixed by the initial probability distribution at time $t = 0$. We will take this to be $P(n,0) = \delta_{n,0}$, meaning that everything kicks off at $n = 0$ and $P(0,0) = 1$. This gives $\phi(s,0) = 1$ and we have

$$\phi(s,t) = e^{\lambda(s-1)t} \ . \tag{4.19}$$

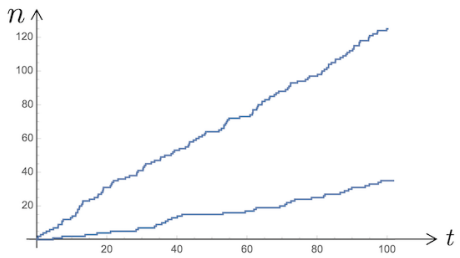
The associated probability density is then given by, using (4.15), by

$$P(n,t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \ . \tag{4.20}$$

This is the *Poisson distribution*. The expectation and standard deviation can be computed from (4.13) and (4.14) and are given by

$$\langle n(t)\rangle = \lambda t \quad \text{and} \quad \sigma = \sqrt{\lambda t} \ . \tag{4.21}$$

We see that $\sigma/\langle n\rangle = 1/\sqrt{\langle n\rangle}$, meaning that fluctuations get less important over time as the population grows. Two Poisson processes, one with $\lambda = 1.3$ and the other with $\lambda = 0.3$, are shown in the figure.



A plot of the probability distribution for different times t is shown in Figure 51. We can get an analytic handle on the evolution of the probability distribution if we invoke Stirling's approximation,

$$n! \approx \sqrt{2\pi n}\, n^n e^{-n} \ . \tag{4.22}$$

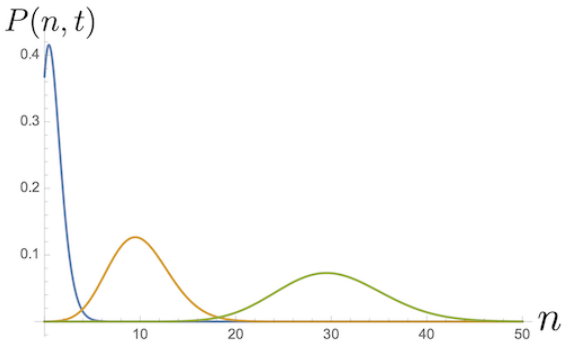


Figure 51. The march of probability, plotted here (for $\lambda = 1$) for times $t = 1$ (in blue), $t = 10$ (in orange), and $t = 30$ (in green).

This is proved, for example, in the lectures on Statistical Physics. With this approximation, the probability distribution (4.21) can be written as

$$P(n,t) \approx \frac{1}{\sqrt{2\pi}} e^{-\lambda t} e^{g(n,t)} \quad \text{with} \quad g(n,t) = n + n\log(\lambda t) - n\log n - \frac{1}{2}\log n \ . \tag{4.23}$$

The function $g(n,t)$ has a maximum at $\partial g/\partial n = 0$ which, you can check, is given at large n by $n = n^\star \approx \lambda t$. Expanding about this maximum gives us an approximate expression for the exponent

$$\begin{aligned} g(n,t) &\approx g(n^\star) + \frac{1}{2}(n-n^\star)^2 \frac{\partial^2 g}{\partial n^2} + \ldots \\ &= \lambda t - \frac{(n-\lambda t)^2}{2\lambda t} - \frac{1}{2}\log(\lambda t) \ . \end{aligned} \tag{4.24}$$

This then translates into a late time, large n, expression for the probability distribution:

$$P(n,t) \approx \frac{1}{\sqrt{2\pi\lambda t}} e^{-(n-\lambda t)^2/2\lambda t} \ . \tag{4.25}$$

We see that, at late times, the probability distribution settles down to a Gaussian distribution, marching forwards with linear growth $n \approx \lambda t$.

4.1.3 Birth and Death Again

We find ourselves turning once again, like the great Russian novelists, to the grand questions of life and death. This time, with a probabilistic slant.

We'll stick with our population model with states given by $n \in \mathbb{N}$. Now the dynamics includes the possibility for birth (or immigration) which increases n by one, and death which decreases n by one. The probability per unit time for these is:

- For $n \to n+1$, we take a constant rate, λ.
- For $n \to n-1$, we take a constant rate *per capita*, so that the rate is βn.

To avoid continually writing out these words, we summarise this in a reaction-like diagram

$$n-1 \underset{\beta \mathrm{n}}{\overset{\lambda}{\rightleftarrows}} n \underset{\beta(\mathrm{n}+1)}{\overset{\lambda}{\rightleftarrows}} n+1 \ . \tag{4.26}$$

We will skip the step of writing $P(n, t+\delta t)$ and just jump immediately to the master equation for the probability distribution which is

$$\frac{dP(n,t)}{dt} = -(\lambda+\beta n)P(n,t) + \lambda P(n-1,t) + \beta(n+1)P(n+1,t) \ . \tag{4.27}$$

You can trace the origin of each of these terms to the diagram above. We'll take this equation to hold for $n \geq 0$, with the proviso that $P(n=-1,t)=0$.

The generating function $\phi(s,t)$ is again defined by (4.12). It obeys

$$\begin{aligned}
\frac{\partial \phi(s,t)}{\partial t} &= \sum_{n=0}^{\infty} s^n \frac{\partial P(n,t)}{\partial t} \\
&= \sum_{n=0}^{\infty} s^n \big[-\lambda P(n,t) + \lambda P(n-1,t) - \beta n P(n,t) + \beta(n+1)P(n+1,t) \big] \\
&= \sum_{n=0}^{\infty} \big[-s^n \lambda + s^{n+1}\lambda - s^n \beta n + s^{n-1}\beta n \big] P(n,t)
\end{aligned} \tag{4.28}$$

where, in the final line, we've shifted the summation variable to gather all terms of the form $P(n,t)$. It's simple to write the first two terms using the generating function: they are proportional to $\phi(s,t)$ and $s\phi(s,t)$ respectively. For the second two terms, we have an extra factor of n in the sum. This arises by differentiating the generating function

$$\frac{\partial \phi(s,t)}{\partial s} = \sum_{n=0}^{\infty} n s^{n-1} P(n,t) \ . \tag{4.29}$$

In this way, having transition rates that are proportional to n, like the β rates above, leads to a partial differential equation for the generating function,

$$\frac{\partial\phi(s,t)}{\partial t} = (s-1)\left(\lambda\phi(s,t) - \beta\frac{\partial\phi(s,t)}{\partial s}\right) \ . \tag{4.30}$$

Note that we have an overall factor of $s-1$. This should be expected because, as shown in (4.16), we have $\phi(1,t) = 1$ for all t, so $\partial\phi/\partial t$ should vanish at $s = 1$.

We're left with (4.30) to solve. To do this, we make the (not immediately obvious) ansatz

$$\phi(s,t) = \exp\big((s-1)f(t)\big) \tag{4.31}$$

for some to-be-determined function $f(t)$. Our ansatz automatically obeys the constraint $\phi(1,t) = 1$. Substituting into (4.30), we see that all s-dependence happily drops out and we are left with the a differential equation only for $f(t)$:

$$\frac{df}{dt} = \lambda - \beta f(t) \ . \tag{4.32}$$

We take the initial condition $P(0,0) = 1$ or, equivalently, $\phi(s,0) = 1$. This requires $f(0) = 1$ and the equation above has solution

$$f(t) = \frac{\lambda}{\beta}(1 - e^{-\beta t)}) \ . \tag{4.33}$$

So our generating function takes the double-exponential form

$$\phi(s,t) = \exp\left(\frac{\lambda}{\beta}(s-1)(1-e^{-\beta t})\right) \ . \tag{4.34}$$

With this in hand, we can now compute various expectation values. The average population size is

$$\langle n(t)\rangle = \left.\frac{\partial\phi}{\partial s}\right|_{s=1} = \frac{\lambda}{\beta}\left(1 - e^{-\beta t}\right) \ . \tag{4.35}$$

The variance can be computed from (4.14) to be

$$\begin{aligned}\sigma^2(t) &= \langle n^2\rangle - \langle n\rangle^2 \\ &= \left.\frac{\partial^2\phi}{\partial s^2}\right|_{s=1} + \left.\frac{\partial\phi}{\partial s}\right|_{s=1} - \left(\left.\frac{\partial\phi}{\partial s}\right|_{s=1}\right)^2 = \frac{\lambda}{\beta}\left(1 - e^{-\beta t}\right) \ .\end{aligned} \tag{4.36}$$

So, again we have $\sigma/\langle n\rangle = 1/\sqrt{\langle n\rangle}$.

In the limit $t \to \infty$, the system settles down to a steady state. In the present case, we can extract this straightforwardly from the generating function (4.34). However, it is rare that we can find an exact expression for the generating function. Nonetheless, it's often possible to get the steady state by returning to the differential equation that governs the generating function, in this case (4.30). In the steady state, we have $\partial\phi/\partial t = 0$ and so

$$\frac{\partial \phi}{\partial s} = \frac{\lambda}{\beta}\phi \quad \Longrightarrow \quad \phi(s) = \exp\left(\frac{\lambda}{\beta}(s-1)\right) \tag{4.37}$$

where we've used the boundary condition $\phi(s=1) = 1$ to fix the overall normalisation. Translated to a steady-state probability distribution, this is

$$P(n) = \frac{1}{n!}\left(\frac{\lambda}{\beta}\right)^n e^{-\lambda/\beta} \ . \tag{4.38}$$

This is again a Poisson distribution. Note that for these kinds of stochastic models, the steady state means that we have a constant probability distribution, rather than a constant n.

More Offspring Means More Variation

We can make a simple change to the model above, and suppose that a birth results in M new individuals. In this case, the master equation (4.27) is replaced by

$$\frac{dP(n,t)}{dt} = -(\lambda + \beta n)P(n,t) + \lambda P(n-M,t) + \beta(n+1)P(n+1,t) \ . \tag{4.39}$$

You can rerun the steps above to find the new equation governing the generating function,

$$\frac{\partial \phi(s,t)}{\partial t} = \lambda(s^M - 1)\phi(s,t) - \beta(s-1)\frac{\partial \phi(s,t)}{\partial s} \ . \tag{4.40}$$

Now this equation is harder to solve. We could restrict ourselves to look for long-time steady state solutions with $\partial\phi/\partial t = 0$, so that we have to solve

$$\frac{\partial \phi}{\partial s} = \frac{\lambda}{\beta}\frac{s^M - 1}{s-1}\phi(s) \ . \tag{4.41}$$

This is somewhat easier to solve. For example, if $M = 2$, then we have

$$\phi(s) = \exp\left(\frac{\lambda}{\beta}\left(s + \frac{s^2}{2} - \frac{3}{2}\right)\right) \ . \tag{4.42}$$

And, from this, we can then reconstruct the steady state probability distribution.

Alternatively, we could just jump immediately to what we're most interested in: the expectation $\langle n(t)\rangle$ and the variance, which follows from $\langle n^2(t)\rangle$. We can derive equations that govern both of these quantities. For the expectation, we have

$$\begin{aligned}
\frac{d\langle n(t)\rangle}{dt} &= \sum_{n=0}^{\infty} n\frac{dP(n,t)}{dt} \\
&= \sum_{n=0}^{\infty} n\big[-(\lambda+\beta n)P(n,t) + \lambda P(n-M,t) + \beta(n+1)P(n+1,t)\big] \\
&= \sum_{n=0}^{\infty} \big[-\lambda n - \beta n^2 + \lambda(n+M) + \beta(n-1)n\big]P(n,t) \\
&= \sum_{n=0}^{\infty} \big[\lambda M - \beta n\big]P(n,t) \\
&= \lambda M - \beta\langle n(t)\rangle \ . \qquad (4.43)
\end{aligned}$$

We see that we get a simple differential equation for $\langle n(t)\rangle$ which we can now just solve. Using the initial condition $\langle n(0)\rangle = 0$, the solution is

$$\langle n(t)\rangle = \frac{\lambda M}{\beta}\left(1 - e^{-\beta t}\right) \ . \qquad (4.44)$$

This takes the same functional form as our previous result (4.35), but with the birth rate λ now increased to λM. That makes sense. However, there's more to be seen if we look at the variance. This too obeys its own differential equation,

$$\begin{aligned}
\frac{d\langle n^2(t)\rangle}{dt} &= \sum_{n=0}^{\infty} n^2\frac{dP(n,t)}{dt} \\
&= \sum_{n=0}^{\infty} n^2\big[-(\lambda+\beta n)P(n,t) + \lambda P(n-M,t) + \beta(n+1)P(n+1,t)\big] \\
&= \sum_{n=0}^{\infty} \big[-\lambda n^2 - \beta n^3 + \lambda(n+M)^2 + \beta(n-1)^2 n\big]P(n,t) \\
&= \sum_{n=0}^{\infty} \big[\lambda M^2 + (2\lambda M + \beta)n - 2\beta n^2\big]P(n,t) \\
&= \lambda M^2 + (2\lambda M + \beta)\langle n(t)\rangle - 2\beta\langle n^2(t)\rangle \ . \qquad (4.45)
\end{aligned}$$

We already have an expression for $\langle n(t)\rangle$, so this is a differential equation for $\langle n^2(t)\rangle$. Things are easier if we look at the steady state distribution. Here we have $\langle n\rangle = \lambda M/\beta$ and so

$$\langle n^2\rangle = \frac{1}{2\beta}\left(\lambda M^2 + \frac{\lambda M}{\beta}(2\lambda M + \beta)\right) \ . \qquad (4.46)$$

Note that there is an M^2 term, as well as a term linear in M. For $M \gg 1$, this quadratic term dominates and we have variance

$$\text{var}(n) \approx \frac{\lambda}{2\beta} M^2 \ . \tag{4.47}$$

While the average population scales as λM, the variation scales as λM^2. This too makes sense: a birth now gives a jump of M in the population, rather than just one, and so the jumps around the mean value are larger.

Non-Linear Growth Rates

For the model above, we have birth and death rates that were either constant or proportional to n. And this was reflected in the nice differential equation (4.43) that we derived for the expectation $\langle n(t) \rangle$. We might wonder if the stochastic growth rates that we enter into the master equation always arise in the equation for the expectation value in this way. The answer, sadly, is no. Things are less pleasant when the rates depend non-linearly on n.

For example, we might try to cook up something akin to the logistic equation by taking a birth rate proportional to λn and a death rate proportional to βn^2. It's straightforward to write down the corresponding master equation,

$$\frac{dP(n,t)}{dt} = -(\lambda n + \beta n^2)P(n,t) + \lambda(n-1)P(n-1,t) + \beta(n+1)^2 P(n+1,t) \ . \tag{4.48}$$

We can then retrace our steps that led to (4.43). This time we have

$$\begin{aligned} \frac{d\langle n(t)\rangle}{dt} &= \sum_{n=0}^{\infty} n \frac{dP(n,t)}{dt} \\ &= \sum_{n=0}^{\infty} \big[-\lambda n^2 - \beta n^3 + \lambda n(n+1) + \beta(n-1)n^2 \big] P(n,t) \\ &= \sum_{n=0}^{\infty} \big[\lambda n - \beta n^2 \big] P(n,t) \\ &= \lambda \langle n(t) \rangle - \beta \langle n^2(t) \rangle \ . \end{aligned} \tag{4.49}$$

But that's not so useful: the equation for $\langle n \rangle$ requires us to know something about $\langle n^2 \rangle$. And the equation for $\langle n^2 \rangle$ will need us to know about $\langle n^3 \rangle$ and so on. The set of equations doesn't close and to make progress we need to make some approximation about these higher order moments, or turn to numerical simulation. Nonetheless, although we can't solve such models completely, as we now show, there are some things that we can say.

4.1.4 Extinction

"*Do one calculation every day that scares you.*"

Eleanor Roosevelt.

In stochastic models, with populations fluctuating up and down, it's quite possible that the population fluctuates to zero and stays there. We would like to know the likelihood of such an extinction event.

Extinction isn't possible in the simple model above since, even if the population does die down to zero, it can still grow again. (We suggested at the time that this may be due to immigration rather than laziness when designing the model.) But we can make things more realistic by considering a model where the birth and death rates depend on the population size n. We write these as b_n and d_n respectively, and then take $b_0 = 0$, meaning that if the population hits $n = 0$ it stays there. That's extinction. This is summarised in the following reaction diagram:

$$0 \underset{\mathrm{d}_1}{\overset{0}{\rightleftarrows}} 1 \underset{\mathrm{d}_2}{\overset{\mathrm{b}_1}{\rightleftarrows}} 2 \quad \ldots \quad n-1 \underset{\mathrm{d}_n}{\overset{\mathrm{b}_{n-1}}{\rightleftarrows}} n \underset{\mathrm{d}_{n+1}}{\overset{\mathrm{b}_n}{\rightleftarrows}} n+1 \quad \ldots \tag{4.50}$$

Now we ask: what's the probability of extinction? Or, more precisely: suppose that the population sits at some healthy number n. Whats the probability Q_n that it will eventually become extinct? We will see that, under one further reasonable assumption, this probability is necessarily one.

Our strategy is to set up a recurrence relation for Q_n. The probability of extinction for a population n can be related to

$$\begin{aligned} Q_n &= \text{Prob(birth next)}\, Q_{n+1} + \text{Prob(death next)}\, Q_{n-1} \\ &= \frac{b_n}{b_n + d_n} Q_{n+1} + \frac{d_n}{b_n + d_n} Q_{n-1} \ . \end{aligned} \tag{4.51}$$

Rearranging, gives the recurrence relation

$$Q_{n+1} - Q_n = \frac{d_n}{b_n}(Q_n - Q_{n-1}) = \left(\prod_{i=1}^{n} \frac{d_i}{b_i} \right) (Q_1 - Q_0) \ . \tag{4.52}$$

All the n dependence on the left-hand side sits in that product. The next question that we want to ask is: does the product converge for large n? The answer, in any realistic model, is no! Our requirement for realism is that as the population swells, the death

rate exceeds the birth rate. Specifically, we require that there exist an integer N and a number $R > 1$ such that

$$\frac{d_n}{b_n} \geq R \quad \text{for all } n > N \ . \tag{4.53}$$

In this case, the product in (4.52) can get arbitrarily large as n gets large. But the left-hand side is a difference of probabilities, so $(Q_{n+1} - Q_n) \in [-1, +1]$. This means that the only way (4.52) can be satisfied for very large n is if

$$Q_{n+1} - Q_n = Q_1 - Q_0 = 0 \tag{4.54}$$

for all n. In particular, we must have $Q_n = Q_0$ but if the population is at $n = 0$ then it's already extinct and so $Q_0 = 1$. We learn that $Q_n = 1$ for all n. Closed systems go extinct under reasonable assumptions. All men must die.

There is one glimmer of light in this calculation: we didn't yet compute how long we've got left! Happily, it turns out that the expected lifetime of a species can be very large.

4.1.5 Multiple Populations: Wildebeest and Flies

We can extend the ideas above to multiple populations. Here we describe a simple system which, to add some colour, we will think of as a population m of wildebeest and a population n of flies. The birth and death rates are taken to be

- The wildebeest have a birth rate λ_1 and death rate $\beta_1 m$.
- The flies have a birth rate (or, said differently, an import rate) of $\lambda_2 m$ and a death rate of $\beta_2 n$. Note that the "birth rate" is proportional to the number of wildebeest, which might sound slightly weird, but we should interpret this as wildebeest attracting flies from elsewhere into the system.

We can capture this in two reaction diagrams (or, alternatively in one 2d reaction diagram). The growth of the wildebeest population is described by

$$(m-1, n) \underset{\beta_1\,\mathrm{m}}{\overset{\lambda_1}{\rightleftarrows}} (m, n) \underset{\beta_1(\mathrm{m}+1)}{\overset{\lambda_1}{\rightleftarrows}} (m+1, n) \ . \tag{4.55}$$

Meanwhile, the growth of the flies is described by

$$(m, n-1) \underset{\beta_2\,\mathrm{n}}{\overset{\lambda_2\,\mathrm{m}}{\rightleftarrows}} (m, n) \underset{\beta_2(\mathrm{n_1}^+))}{\overset{\lambda_2\,\mathrm{m}}{\rightleftarrows}} (m, n+1) \ . \tag{4.56}$$

The system is slightly unusual in that the wildebeest population is unaffected by the flies, but the flies care about the wildebeest. Note, also, that everything is linear which will make the system tractable.

To keep our equations looking vaguely reasonable, we will write the probability that we are in the state (m, n) as $p_{m,n}(t)$. We can read off the master equation from the reaction diagrams above:

$$\begin{aligned}\frac{dp_{m,n}}{dt} &= \lambda_1\big[p_{m-1,n} - p_{m,n}\big] + \beta_1\big[(m+1)p_{m+1,n} - mp_{m,n}\big] \\ &\quad + \lambda_2\big[mp_{m,n-1} - mp_{m,n}\big] + \beta_2\big[(n+1)p_{m,n+1} - np_{m,n}\big] \ . \end{aligned} \tag{4.57}$$

From this, we can compute the evolution of the average population size. For wildebeest, we have

$$\begin{aligned}\frac{d\langle m\rangle}{dt} &= \sum_{m,n} mp_{m,n} \\ &= \lambda_1\big[\langle m+1\rangle - \langle m\rangle\big] + \beta_1\big[\langle (m-1)m\rangle - \langle m^2\rangle\big] \\ &\quad +\lambda_2\big[\langle m^2\rangle - \langle m^2\rangle\big] + \beta_2\big[\langle mn\rangle - \langle mn\rangle\big] \\ &= \lambda_1 - \beta_1\langle m\rangle \ . \end{aligned} \tag{4.58}$$

For the flies, we have

$$\begin{aligned}\frac{d\langle n\rangle}{dt} &= \sum_{m,n} np_{m,n} \\ &= \lambda_1\big[\langle n\rangle - \langle n\rangle\big] + \beta_1\big[\langle mn\rangle - \langle mn\rangle\big] \\ &\quad +\lambda_2\big[\langle m(n+1)\rangle - \langle mn\rangle\big] + \beta_2\big[\langle (n-1)n\rangle - \langle n^2\rangle\big] \\ &= \lambda_2\langle m\rangle - \beta_2\langle n\rangle \ . \end{aligned} \tag{4.59}$$

We see again that the wildebeest population (4.58) doesn't depend on the flies, while the converse is not true. The steady state is given by

$$\langle m\rangle = \frac{\lambda_1}{\beta_1} \quad \text{and} \quad \langle n\rangle = \frac{\lambda_2}{\beta_2}\langle m\rangle = \frac{\lambda_1\lambda_2}{\beta_1\beta_2} \ . \tag{4.60}$$

We can also look at the fluctuations, starting by computing quadratic expectations. We've already computed $\langle m^2\rangle$ in our previous birth/death model. (It's given by (4.45) after setting $M = 1$.) We have

$$\frac{d\langle m^2\rangle}{dt} = \lambda_1 + (2\lambda_1 + 1)\langle m\rangle - 2\beta_1\langle m^2\rangle \ . \tag{4.61}$$

We can similarly compute the evolution of $\langle n^2\rangle$ and $\langle mn\rangle$. They are given by

$$\begin{aligned}\frac{d\langle n^2\rangle}{dt} &= \lambda_2\langle m\rangle + \beta_2\langle n\rangle + 2\lambda_2\langle mn\rangle - 2\beta_2\langle n^2\rangle \\ \frac{d\langle mn\rangle}{dt} &= \lambda_1\langle n\rangle + \lambda_2\langle m^2\rangle - (\beta_1 + \beta_2)\langle mn\rangle \ . \end{aligned} \tag{4.62}$$

In the steady state (4.60), variance of the flies is then given by

$$\begin{aligned}\text{var}(n) &= \langle n^2\rangle - \langle n\rangle^2 \\ &= \frac{1}{2}\langle n\rangle + \frac{\lambda_2}{2\beta_2}\langle m\rangle + \frac{\lambda_2}{\beta_2}\langle mn\rangle - \langle n\rangle^2 \\ &= \langle n\rangle + \frac{\lambda_2}{\beta_2}\big[\langle mn\rangle - \langle n\rangle\langle m\rangle\big] \ . \end{aligned} \tag{4.63}$$

The term in square brackets is the *covariance* between the variables m and n,

$$\text{cov}(m,n) = \langle mn\rangle - \langle m\rangle\langle n\rangle \ . \tag{4.64}$$

We learn that the fluctuations of the flies has two terms: an intrinsic fluctuation in the birth and death rates of the flies, proportional to $\langle n\rangle$, and an additional fluctuation proportional to $\text{cov}(m,n)$ that tracks the fluctuations in wildebeest.

4.2 Meet the Fokker-Planck Equation

In the previous section, we studied various examples of the "master equation", which governs how a probability distribution over a discrete set of outcomes evolves. In this section, we would like to generalise this idea to describe a probability distribution over a continuous set of outcomes. The simplest example is a probability distribution $P(\mathbf{x},t)$ of some substance distributed spread over some spatial coordinate $\mathbf{x}$. The resulting equation is called the *Fokker-Planck equation*.

We're going to derive the Fokker-Planck equation starting from our discrete master equation. If the width of the probability distribution is much broader than then size between the spacing, then it makes sense to approximate the discrete variable with a continuous variable.

For all our examples above, we thought of the discrete variable $n \in \mathbb{N}$ as the population size. There will be times when we want to keep that interpretation, but we might also want to think of n as labelling the position of some object that is restricted to lie on a lattice. (For example, such a set-up arises in Solid State Physics when we think of an electron moving in a solid.) In this case, we could relabel $n = x$ to denote position. We will adopt this notation below.

Suppose that the hopping rate to jump from site n to site $n+r$ is given by $W(n,r)$. Here $r \in \mathbb{Z}$ can be positive or negative. Then the master equation for the probability distribution over sites $x = n$ is given by

$$\frac{\partial P(x,t)}{\partial t} = \sum_{r\in\mathbb{Z}} \Big[W(x-r,r)P(x-r,t) - W(x,r)P(x,t)\Big] \ . \tag{4.65}$$

Here the first term captures the fact that the particle could hop from any site to x, while the second term captures the fact that it could hop away to any site. Typically, this hopping rate will be "short range", meaning that $W(x,r)$ drops off quickly as r gets large.

Now, the term in the square brackets looks like $f(x-r) - f(x)$ where $f(x) = W(x,r)P(x,t)$. For r small, we Taylor expand

$$f(x-r) = f(x) - r\frac{df}{dx} + \frac{r^2}{2}\frac{d^2f}{dx^2} + \ldots \tag{4.66}$$

We apply this Taylor expansion to the master equation (4.65) and drop the $\ldots$ terms. The decision to truncate the Taylor expansion after the second derivative is important and will have consequence below. We're left with

$$\begin{aligned}\frac{\partial P(x,t)}{\partial t} &= \sum_{r\in\mathbb{Z}}\Big[-r\frac{\partial}{\partial x}\big(W(x,r)\,P(x,t)\big) + \frac{r^2}{2}\frac{\partial^2}{\partial x^2}\big(W(x,r)\,P(x,t)\big)\Big] \\ &= -\frac{\partial}{\partial x}\big(u(x)\,P(x,t)\big) + \frac{\partial^2}{\partial x^2}\big(D(x)\,P(x,t)\big)\ .\end{aligned} \tag{4.67}$$

This is the *Fokker-Planck equation*[15]. It involves two functions, $u(x)$ and $D(x)$, given by

$$u(x) = \sum_{r\in\mathbb{Z}} r\,W(x,r) \quad \text{and} \quad D(x) = \frac{1}{2}\sum_{r\in\mathbb{Z}} r^2\,W(x,r)\ . \tag{4.69}$$

We assume that $W(x,r)$ drops off quickly enough at large r so that both of these sums converge. You can read more about the Fokker-Planck equation, viewed from a slightly different perspective, in the lectures on Kinetic Theory.

The total probability is necessarily conserved, with $\int dx\ P(x,t) = 1$ for all time. Things that are conserved obey a continuity equation, and probability is no exception. We can recast the Fokker-Planck equation in this form, writing

$$\frac{\partial P}{\partial t} + \frac{\partial J}{\partial x} = 0 \quad \text{with} \quad J = uP - \frac{\partial}{\partial x}(DP)\ . \tag{4.70}$$

[15]It's not uncommon to see the Fokker-Planck equation written as

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}(A\,P) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(BP) \tag{4.68}$$

with the obvious relation $A(x) = u(x)$ and $B(x) = 2D(x)$. But, as we'll see soon, the variables u and D are more evocative than A and B.

Written in this way, we see that the current J takes the form that we anticipated in Section 3. The first term corresponds to advection, with $u(x)$ some background flow that carries the probability with it. The second term gives rise to diffusion. We'll see these interpretations borne out in what follows.

From the definitions (4.69), we see that for positive rates $W(x,r) > 0$, it might be possible for the advection $u(x)$ to vanish, but it will never be possible for the diffusion $D(x)$ to vanish. In this sense, diffusion is really the essential element that arises from stochastic processes.

Evolving Moment by Moment

We can look at how various expectation values change with the Fokker-Planck equation. The average position $\langle x(t)\rangle$ evolves as

$$\frac{d\langle x\rangle}{dt} = \int dx\ x\,\frac{\partial P(x,t)}{\partial t} = -\int dx\ x\frac{\partial(uP)}{\partial x} + \int dx\ x\frac{\partial^2(DP)}{\partial x^2}\ . \tag{4.71}$$

We integrate by parts, using the fact that any normalised probability distribution must vanish asymptotically. For the second term, we can integrate by parts twice to get zero. So only the first term contributes, giving

$$\frac{d\langle x\rangle}{dt} = \int dx\ uP = \langle u(x)\rangle\ . \tag{4.72}$$

We see that the time evolution of the mean depends on the average of $u(x)$, and not on the function $D(x)$. This confirms what we said above: the function $u(x)$ acts like advection, governing the overall drift of the probability distribution.

For the variance, we first look at

$$\frac{d\langle x^2\rangle}{dt} = -\int dx\ x^2\frac{\partial(uP)}{\partial x} + \int dx\ x^2\frac{\partial^2(DP)}{\partial x^2}\ . \tag{4.73}$$

Now both terms survive integration by parts. We have

$$\frac{d\langle x^2\rangle}{dt} = 2\langle xu(x)\rangle + 2\langle D(x)\rangle\ . \tag{4.74}$$

The variation is, as usual, $\mathrm{var}(x) = \langle x^2\rangle - \langle x\rangle^2$ and obeys the equation

$$\frac{d(\mathrm{var}(x))}{dt} = \frac{d\langle x^2\rangle}{dt} - 2\langle x\rangle\frac{d\langle x\rangle}{dt} = 2\langle D(x)\rangle + 2\,\mathrm{cov}(x,u(x))\ . \tag{4.75}$$

We see the same kind of behaviour as for our wildebeest problem, with two terms contributing to the variance. The first is the expectation value of $D(x)$, the second a covariance between x and $u(x)$ given by $\mathrm{cov}(x,u(x)) = \langle xu(x)\rangle - \langle x\rangle\langle u(x)\rangle$.

4.2.1 Constant Drift and Diffusion

To illustrate these ideas, let's return to the particularly simple model that marches tentatively forward at a constant rate λ. In the discrete case, this gave rise to the Poisson process, governed by the master equation (4.11)

$$\frac{dP(n,t)}{dt} = \lambda\big[P(n-1,t) - P(n,t)\big] \ . \tag{4.76}$$

Comparing to (4.65), we have $W(n,1) = \lambda$ for all n, with $W(n,r) = 0$ for $r \neq 1$. The corresponding Fokker-Planck equation is

$$\frac{\partial P}{\partial t} = -\lambda\frac{\partial P}{\partial x} + \frac{\lambda}{2}\frac{\partial^2 P}{\partial x^2} \ . \tag{4.77}$$

We can use our results above, with $u = \lambda$ and $D = \frac{1}{2}\lambda$, to compute how the moments evolve. We have

$$\frac{d\langle x\rangle}{dt} = \lambda \quad \Longrightarrow \quad \langle x\rangle = \lambda t \tag{4.78}$$

and

$$\frac{d\langle x^2\rangle}{dt} = 2\lambda\langle x\rangle + \lambda \quad \Longrightarrow \quad \langle x^2\rangle = \lambda t + (\lambda t)^2 \ . \tag{4.79}$$

It's worth pointing out that both of these agree with the corresponding discrete model, where we also had

$$\langle n\rangle = \lambda t \quad \text{and} \quad \langle n^2\rangle = \lambda t + (\lambda t)^2 \ . \tag{4.80}$$

It's natural to ask: does the Fokker-Planck equation coincide with the discrete Poisson process? The answer is no: the first two moments coincide, but not higher moments. You can check, for example, that

$$\begin{aligned} \langle x^3\rangle &= (\lambda t)^3 + 3(\lambda t)^2 \\ \langle n^3\rangle &= (\lambda t)^3 + 3(\lambda t)^2 + \lambda t \ . \end{aligned} \tag{4.81}$$

This is a typical feature of the Fokker-Planck equation when compared to a discrete master equation. The fact that the two agree for the first two moments, and then disagree, can be traced to our truncation of the Taylor expansion at the second derivative in (4.67). Note, however, that the two agree for large time, which reflects the fact that $\langle x\rangle = \lambda t$ and so $\langle x\rangle \gg r$ for large time.

In this simple case, it's not difficult to solve the Fokker-Planck equation. Motivated by the fact that we have a constant drift $u = \lambda$, we introduce the variable $\xi = x - \lambda t$ and consider the ansatz

$$P(x,t) = G(\xi,t) \ . \tag{4.82}$$

Then the Fokker-Planck equation (4.77) becomes the diffusion equation

$$\frac{\partial G}{\partial t} = \frac{\lambda}{2}\frac{\partial^2 G}{\partial \xi^2} \ . \tag{4.83}$$

We've already seen solutions to this equation in Section 3.1. If we start with a delta-function initial condition, then the probability distribution is given by an ever-spreading Gaussian, now with an overall drift set by $u = \lambda$,

$$P(x,t) = \frac{1}{\sqrt{2\pi\lambda t}} e^{-(x-\lambda t)^2/2\lambda t} \ . \tag{4.84}$$

Diffusion Revisited

It's straightforward to cook up a situation in which the drift vanishes, but diffusion remains. Suppose that we have a particle that lives on a line, with position $n \in \mathbb{Z}$. This time, it bounces back and forth at the same rate λ, so we have $W(n,1) = W(n,-1) = \lambda$. Now we have $A = 0$ and $D = \lambda$ and the Fokker-Planck equation coincides with the heat equation that we studied in Section 3,

$$\frac{\partial P}{\partial t} = D\frac{\partial^2 P}{\partial x^2} \ . \tag{4.85}$$

This illustrates how diffusion arises from underlying randomness. This is what happens, for example, in Brownian motion in which small particles, suspended in a liquid, move in an erratic motion as they are constantly bombarded by surrounding molecules.

4.2.2 Birth and Death Once More

We can also look at the Fokker-Planck equation for our birth and death model with master equation (4.27)

$$\frac{dP(n,t)}{dt} = -(\lambda + \beta n)P(n,t) + \lambda P(n-1,t) + \beta(n+1)P(n+1,t) \ . \tag{4.86}$$

We compare this to (4.65) to find $W(n,1) = \lambda$ and $W(n,-1) = \beta n$. We replace the discrete n with the continuous x which, in this context, still measure the population. The Fokker-Planck equation for $P(x,t)$ then becomes

$$\frac{\partial P}{\partial t} = -\frac{\partial (uP)}{\partial x} + \frac{\partial^2 (DP)}{\partial x^2} \tag{4.87}$$

with

$$u(x) = W(n,1) - W(n,-1) = \lambda - \beta x$$
$$D(x) = \frac{1}{2}(W(n,1) + W(n,-1)) = \frac{1}{2}(\lambda + \beta x) \ . \tag{4.88}$$

This is one of the simplest examples of the Fokker-Planck equation. Following our expectations above, we have (4.72),

$$\frac{d\langle x\rangle}{dt} = \lambda - \beta\langle x\rangle \quad \Longrightarrow \quad \langle x\rangle = \frac{\lambda}{\beta}\left(1 - e^{-\beta t}\right) \ . \tag{4.89}$$

This is identical to the discrete birth and death model that we met previously. (See, for example, (4.44) with $M = 1$.) Similarly, we have from (4.74)

$$\frac{d\langle x^2\rangle}{dt} = \lambda + (2\lambda + \beta)\langle x\rangle - 2\beta\langle x^2\rangle \ . \tag{4.90}$$

This too agrees with the differential equation (4.45) that governs the discrete model. The variance then obeys

$$\frac{d(\mathrm{var}(x))}{dt} = \lambda + \beta\langle x\rangle - 2\beta\mathrm{var}(x) \quad \Longrightarrow \quad \mathrm{var}(x) = \frac{\lambda}{\beta}\left(1 - e^{-\beta t}\right) \ . \tag{4.91}$$

where we've used the expression for $\langle x\rangle$ in (4.89) and implemented the initial condition $\mathrm{var}(x) = 0$ when $t = 0$.

4.2.3 Fokker-Planck With More Variables

It's straightforward to generalise the Fokker-Planck equation to include more variables so that we work with the vector $\mathbf{x} \in \mathbb{R}^d$. In the context of physics, $\mathbf{x}$ is a spatial coordinate; in the context of ecology, $\mathbf{x} = \mathbf{n}$ is a variable that describes the population of d different species.

We can follow our earlier definition, starting with (4.65) which, with multiple variables, reads

$$\frac{\partial P(\mathbf{x},t)}{\partial t} = \sum_{\mathbf{r}\in\mathbb{Z}^d}\Big[W(\mathbf{x}-\mathbf{r},\mathbf{r})P(\mathbf{x}-\mathbf{r},t) - W(\mathbf{x},\mathbf{r})P(\mathbf{x},t)\Big] \ . \tag{4.92}$$

Again we Taylor expand $f(\mathbf{x}) = W(\mathbf{x},\mathbf{r})P(\mathbf{x},t)$ and write

$$f(\mathbf{x}-\mathbf{r}) = f(\mathbf{x}) - r_i\frac{\partial f}{\partial x_i} + \frac{r_i r_j}{2}\frac{\partial^2 f}{\partial x_i \partial x_j} + \ldots \tag{4.93}$$

Dropping the ... terms leaves us with the multi-dimensional Fokker-Planck equation

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x_i}\big(u_i(\mathbf{x})\,P(\mathbf{x},t)\big) + \frac{\partial^2}{\partial x_i x_j}\big(D_{ij}(\mathbf{x})\,P(\mathbf{x},t)\big)\ . \tag{4.94}$$

with our two functions now given by

$$u_i(\mathbf{x}) = \sum_{\mathbf{r}\in\mathbb{Z}^d} r_i\,W(\mathbf{x},\mathbf{r}) \quad \text{and} \quad D_{ij}(\mathbf{x}) = \frac{1}{2}\sum_{\mathbf{r}\in\mathbb{Z}^d} r_i r_j\,W(\mathbf{x},\mathbf{r})\ . \tag{4.95}$$

We see that the advection term now involves a vector function $\mathbf{u}(\mathbf{x})$, reflecting its interpretation it plays as a background velocity field. Meanwhile the diffusion term now involves a symmetric matrix $D(\mathbf{x})$.

We can again see the meaning of the advection term by computing

$$\begin{aligned}\frac{d}{dt}\langle x_i\rangle &= \int d^dx\ x_i\frac{\partial P(\mathbf{x},t)}{\partial t}\\ &= \int d^dx\ x_i\left(-\frac{\partial(u_jP)}{\partial x_j} + \frac{\partial(D_{jk}P)}{\partial x_j\partial x_k}\right) = \int d^dx\ u_iP\end{aligned} \tag{4.96}$$

where, in the final equality, we've integrated by parts. We see that we have the obvious generalisation

$$\frac{d}{dt}\langle\mathbf{x}\rangle = \langle\mathbf{u}\rangle\ . \tag{4.97}$$

Meanwhile, the time derivative of the fluctuations is captured by

$$\begin{aligned}\frac{d}{dt}\langle x_ix_j\rangle &= \int d^dx\ x_ix_j\left(-\frac{\partial(u_kP)}{\partial x_k} + \frac{\partial(D_{kl}P)}{\partial x_k\partial x_l}\right)\\ &= \langle x_iu_j\rangle + \langle u_ix_j\rangle + 2\langle D_{ij}\rangle\ .\end{aligned} \tag{4.98}$$

It's useful to define the symmetric *covariance matrix*

$$C_{ij} = \mathrm{cov}(x_i,x_j) = \langle x_ix_j\rangle - \langle x_i\rangle\langle x_j\rangle\ . \tag{4.99}$$

This contains the variance as its diagonal terms, $\mathrm{var}(x_i) = C_{ii}$, with the off-diagonal terms telling us about correlations between different variables. You can check that

$$\frac{dC_{ij}}{dt} = \mathrm{cov}(x_i,u_j) + \mathrm{cov}(x_j,u_i) + 2\langle D_{ij}\rangle\ . \tag{4.100}$$

Often, we would like to understand the steady state of a distribution, which means that want to find solutions where the right-hand side of this equation vanishes. But that's not so straightforward because the right-hand side of this equation depends on $\langle x_iu_j\rangle$ and we don't necessarily have a good handle on this. To illustrate how to proceed, we turn to an example.

4.2.4 Wildebeest and Flies Again

We will revisit the story of the Wildebeest and flies from Section 4.1.5. This is a two-dimensional system, described by $\mathbf{x} = (m, n)$ where m is the population of wildebeest and n the population of flies. We previously derived the master equation (4.57) when treating the population as discrete,

$$\begin{aligned}\frac{dp_{m,n}}{dt} &= \lambda_1\big[p_{m-1,n} - p_{m,n}\big] + \beta_1\big[(m+1)p_{m+1,n} - mp_{m,n}\big] \\ &\quad + \lambda_2\big[mp_{m,n-1} - mp_{m,n}\big] + \beta_2\big[(n+1)p_{m,n+1} - np_{m,n}\big] \ . \end{aligned} \tag{4.101}$$

Comparing to (4.92), we can read off the non-vanishing values of $W(\mathbf{x}, \mathbf{r})$. They are:

- $W(\mathbf{x}, \mathbf{r}) = \lambda_1$ when $\mathbf{r} = (1, 0)$.
- $W(\mathbf{x}, \mathbf{r}) = \beta_1 m$ when $\mathbf{r} = (-1, 0)$.
- $W(\mathbf{x}, \mathbf{r}) = \lambda_2 m$ when $\mathbf{r} = (0, 1)$.
- $W(\mathbf{x}, \mathbf{r}) = \beta_2 n$ when $\mathbf{r} = (0, -1)$.

From this, we can read off the functions in the Fokker-Planck equation. The advection velocity is

$$\begin{aligned}\mathbf{u} &= \sum_{\mathbf{r}} \mathbf{r}\, W(\mathbf{x}, \mathbf{r}) \\ &= \lambda_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta_1 m \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \lambda_2 m \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \beta_2 n \begin{pmatrix} 0 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 - \beta_1 m \\ \lambda_2 m - \beta_2 n \end{pmatrix} \ . \end{aligned} \tag{4.102}$$

Meanwhile, the diffusion matrix is

$$\begin{aligned} D &= \frac{1}{2}\left[\lambda_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \beta_1 m \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \lambda_2 m \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \beta_2 n \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right] \\ &= \frac{1}{2}\begin{pmatrix} \lambda_1 + \beta_1 m & 0 \\ 0 & \lambda_2 m + \beta_2 n \end{pmatrix} . \end{aligned} \tag{4.103}$$

Now that we have explicit expressions for $\mathbf{u}$ and D, we can return to the question: what does it mean to have a steady state probability distribution for this model?

It's simple to find the steady state for expectation values $\langle \mathbf{x} \rangle$ because, as we've seen in (4.97), this is given by $\langle \mathbf{u} \rangle = 0$, so

$$\langle m \rangle = \frac{\lambda_1}{\beta_1} \quad \text{and} \quad \langle n \rangle = \frac{\lambda_1 \lambda_2}{\beta_1 \beta_2} \ . \tag{4.104}$$

This agrees with the results from the discrete model (4.60). But now we want to extend this to think about fluctuations, as captured in the covariance matrix C_{ij}. In the steady state, we want to find solutions to $dC_{ij}/dt = 0$ and that means that we need to compute $\text{cov}(x_i, u_j)$ and $\langle D_{ij} \rangle$. We see that $\mathbf{u}$ is linear in $\mathbf{x} = (m, n)$ and so we can write

$$\mathbf{u} = \boldsymbol{\lambda} + a\mathbf{x} \ \text{ with } \ \boldsymbol{\lambda} = \begin{pmatrix} \lambda_1 \\ 0 \end{pmatrix} \ \text{ and } \ a = \begin{pmatrix} -\beta_1 & 0 \\ \lambda_2 & -\beta_2 \end{pmatrix} \ . \tag{4.105}$$

This gives

$$\text{cov}(x_i, u_j) = a_{jk} \text{cov}(x_i, x_k) = a_{jk} C_{ik} \ . \tag{4.106}$$

We also have

$$\langle D_{ij} \rangle = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \lambda_2 / \beta_1 \end{pmatrix} \ . \tag{4.107}$$

Now we can look for steady state solutions for the covariance matrix C_{ij}. In steady state, our evolution equation (4.100) becomes

$$\frac{dC}{dt} = aC + Ca^T + 2\langle D \rangle = 0 \ . \tag{4.108}$$

This is a matrix equation, with a, C and $\langle D \rangle$ all 2×2 matrices. It is an example of a *Lyapunov equation*. The equation is easily solved by writing it out in components and doing some linear algebra. We find (recalling that $C_{12} = C_{21}$.

$$\begin{aligned} -2\beta_1 C_{11} + 2\lambda_1 &= 0 \\ \lambda_2 C_{11} - \beta_2 C_{12} - \beta_1 C_{12} &= 0 \\ 2\lambda_2 C_{12} - 2\beta_2 C_{22} + 2\frac{\lambda_1 \lambda_2}{\beta_1} &= 0 \ . \end{aligned} \tag{4.109}$$

Rearranging, then gives the variances

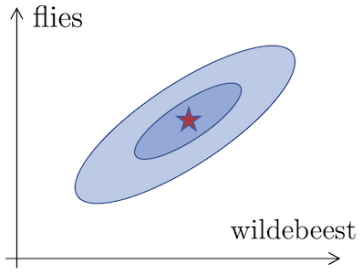
$$\text{var}(m) = C_{11} = \frac{\lambda_1}{\beta_1} \quad \text{and} \quad \text{var}(n) = C_{22} = \frac{\lambda_1 \lambda_2}{\beta_1 \beta_2} \left(1 + \frac{\lambda_2}{\beta_1 + \beta_2} \right) \tag{4.110}$$

while the covariance is

$$\mathrm{cov}(m,n) = C_{12} = \frac{\lambda_1\lambda_2}{\beta_1(\beta_1+\beta_2)} \ . \tag{4.111}$$

We see that $C_{12} > 0$ so the wildebeest and flies covary positively: if there is more of one, then there is likely to be more of the other. That's to be expected given our initial assumptions which were that wildebeest attract flies.

With this information, we can plot the range in which we expect to find populations of wildebeest and flies. This is shown schematically in the figure where the mean is shown as a star. Around that, we draw ellipses whose semi-axes are determined by the eigenvectors and eigenvalues of the covariance matrix C. In the figure, we've sketched ellipses corresponding to one standard deviation and, outside, 95% confidence levels.



Deriving a General Lyapunov Equation

For our example above, we were lucky because both $\mathbf{u}(\mathbf{x})$ and $D(\mathbf{x})$ were linear in the variables $\mathbf{x}$. That meant that the steady state condition $dC_{ij}/dt = 0$ could be expressed entirely in terms of the covariant matrix C_{ij} and some constant matrices. But that won't always be the case.

To find the steady state for more general, non-linear systems, we typically have to make a (not always justified) approximation. We approximate the advection velocity to be linear and the diffusion matrix to be constant,

$$\mathbf{u} = \boldsymbol{\lambda} + a\mathbf{x} \quad \text{and} \quad D = b \tag{4.112}$$

where both a and b are constant matrices. In steady state, the covariance matrix then satisfies the Lyapunov equation

$$aC + Ca^T + b = 0 \ . \tag{4.113}$$

The solution can again be found using some linear algebra.

4.3 An Invitation to Fluctuation and Dissipation

To finish, we can make contact with some basic ideas from elsewhere in physics, notably the lectures on Statistical Physics. We will start by considering something very basic: a particle of mass m, with Newton force law

$$m\ddot{\mathbf{x}} = -\gamma\dot{\mathbf{x}} - \nabla V + \mathbf{f} \ . \tag{4.114}$$

The first two terms on the right-hand side are very familiar: the first is a friction term, with the strength of friction dictated by the coefficient γ; the second is a conservative force arising from a potential $V(\mathbf{x})$. The novelty is the third term, consisting of the additional force $\mathbf{f}$. This we take to be a random force. You can think of this as arising because the particle is suspended in some liquid, and is being constantly bombarded by the underlying molecules, causing it to bounce back and forth in some random way. This is a famous process known as *Brownian motion*.

There are various ways of dealing with equations like (4.114). The most systematic way is to think of the random force $\mathbf{f}$ as coming from some probability distribution, and then figuring out how to translate that into a corresponding probability distribution $P(\mathbf{x})$ for the position of the particle. In this context, (4.114) is known as the *Langevin equation*. You can read more about this in the lectures on Kinetic Theory. Here, instead, we will make direct contact with the Fokker-Planck equation. Our goal is to write down a Fokker-Planck equation for the probability distribution $P(\mathbf{x}, t)$.

For this, we should be in the limit where the motion is friction dominated and the acceleration term in (4.114) can be ignored. In this case, we take the average of (4.114) and use the fact that $\langle \mathbf{f} \rangle = 0$ because the random force is just as likely to hit from any direction. This then tell us that the average velocity of the particle is dictated by the potential

$$\langle \dot{\mathbf{x}} \rangle = -\frac{1}{\gamma} \langle \nabla V \rangle \ . \tag{4.115}$$

But this is the same kind of equation that we get from the Fokker-Planck equation (4.97) if we set $\mathbf{u} = -\nabla V/\gamma$. This suggests that the probability distribution of the particle is governed by a Fokker-Planck equation that takes the form

$$\frac{\partial P}{\partial t} = \frac{1}{\gamma} \nabla \cdot (P \nabla V) + D \nabla^2 P \tag{4.116}$$

for some diffusion constant D. We have made the additional assumptions here that the matrix $D_{ij} = D\delta_{ij}$ is diagonal, on grounds of rotational invariance, and, moreover, that D is independent of x on grounds of translational invariance. It remains to determine the diffusion constant D in terms of the variables in the original set-up.

This is where idea from statistical mechanics come in. First, we look at the equilibrium probability distribution, obeying

$$\nabla \cdot \left(\frac{1}{\gamma} P \nabla V + D \nabla P \right) = 0 \ . \tag{4.117}$$

We can view this as a differential equation for $P(\mathbf{x})$, one that is solved by

$$P(\mathbf{x}) \sim \exp\left(-\frac{1}{\gamma D} V(\mathbf{x}) \right) \tag{4.118}$$

up to an overall normalisation that we've ignored. Now suppose that the random force $\mathbf{f}$ arises because the particle sits in a fluid at temperature T. Then we know that the probability distribution must take the usual Boltzmann form

$$P(\mathbf{x}) \sim \exp\left(-\frac{1}{k_B T} V(\mathbf{x}) \right) \tag{4.119}$$

with k_B the Boltzmann constant. (There is no kinetic term in this expression because we're in a friction-dominated environment where we can ignore the $m\ddot{x}$ term in the original equation of motion.) Equating these two expressions, we learn that the diffusion constant must be given by

$$D = \frac{k_B T}{\gamma} \ . \tag{4.120}$$

This is the *Einstein relation*. It is the key result in the fourth of his famous collection of 1905 papers. (The one that didn't introduce special relativity or pioneer the idea of the quantum!) The relation is rather surprising: the diffusion constant tells us how much the particle is kicked around by the environment, while the friction term tells us how difficult it is for the particle to plough through the same environment. Remarkably, the two are related.

The Einstein relation is an example of a more general idea known as the *fluctuation-dissipation theorem*, which relates the fluctuations experienced by a system to the dissipation (i.e. friction) experienced by the system. It's an important idea in many areas of physics.