Illustrative examples of the uses of mathematical modelling

Alan D. Rendall

1 Introduction

In these notes I present a number of examples which illustrate how mathematical models can be applied to improve understanding of scientific questions. The focus is on topics from biology and medicine, although models arising in chemistry and physics will also be mentioned. The examples have been ordered roughly by the mathematical complexity of the models, starting with the simplest. For each example I will begin by introducing some of the scientific background before going on to describe the mathematical modelling. These notes were originally prepared as background for a lecture in Karlstad and so I included Swedish translations of some animal names.

2 The circulation of the blood

The mathematics in this example is just simple arithmetic. It illustrates how the use of even the most elementary mathematical ideas can influence the progress of science. From a more general point of view it is an early example of the successful application of the scientific method. The idea of the circulation of the blood, which we take for granted today, gained wide acceptance due the work of the English physician William Harvey. Perhaps the central event in this development was the publication of his book 'De motu corporis' in 1628. This overturned an incorrect theory which had been believed in and applied by doctors for almost 1500 years since it was introduced by Galen.

The heart consists of two halves which are separated by a sheet of muscle called the septum. Each half is attached to a system of blood vessels. In Galen's theory blood is produced in the liver and consumed in the tissues, providing them with nourishment. Some of the blood passes through the septum. This blood carries 'vital spirits' from the lungs to the tissues. Harvey studied medicine with Fabricius in Padua, at that time a centre of medical learning. It is interesting to note that the professor of mathematics in Padua at that time was Galileo Galilei, another pioneer of the scientific method. Fabricius observed that the veins contained valves which restricted the blood to flow in one direction. This was already in conflict with Galen's theory. Harvey collected different pieces of evidence for the blood circulation but the most striking was

based on a quantitative argument. Harvey did an experiment on a dog where he measured how much blood left the heart with each heartbeat. Multiplying this by the number of heartbeats in a given time allowed him to compute how much blood passed through the heart in that time and thus, according to Galen's theory, would have had to be produced and consumed in that time. By a reasonable scaling, this could give an estimate for the corresponding quantity in the case of a human. The answer was that the amount of blood involved was three times the weight of the body each hour. This was so implausible that it was an extremely effective argument for Harvey. There was enormous resistance among physicians and because of public criticism Harvey lost many of the patients from his private practice. Nevertheless, with the help of this argument, the truth eventually prevailed. It should be pointed out that in Harvey's theory the blood has to pass through the tissues from the arteries to the veins and the way in which it could do so was no more known that the hypothetical pores of the septum in Galen's theory. The capillaries, which make the necessary connection, were later discovered using the microscope by Malpighi in 1661, four years after Harvey's death. Further information on Harvey can be found in [11].

3 Multidrug therapy for HIV-AIDS

The human immunodeficiency virus (HIV) is one of the biggest health problems of the modern world. A major change in this problem was brought about by the introduction of multidrug therapy in the mid 1990s. Before that the best that therapy could do was to hold back the progress of the disease a little. After that it became possible to avoid death due to the virus essentially indefinitely by administering suitable drugs. This situation is not without difficulties. The drugs are very expensive and have a lot of unpleasant side effects. It is necessary to take them steadily in order to prevent the disease returning. Now AIDS has become primarily a disease of the third world. There people often cannot afford the drugs and even if they are provided the other problems in these countries prevent people from having sufficient motivation to take them regularly on a long term basis. Nevertheless this therapy has had a huge positive effect on the situation of those infected with the virus, particularly in the developed world. Both of the fundamental papers on this subject which were published in 1995 had a mathematician on the team to do necessary modelling. One of these papers is [1].

When someone becomes infected with HIV they get flu-like symptoms for a few days. These soon go away and no obvious symptoms remain. Only after a long time (about ten years) do the symptoms of AIDS appear. The virus infects certain white blood cells known as T-helper or $CD4^+$ T cells. During the long period between infection and the outbreak of AIDS the number of these cells decreases slowly but steadily. Nobody understands the mechanism of this decrease and this is perhaps the central theoretical open question in this subject. When the number of these cells has dropped enough the immune system ceases to work and the patient suffers from opportunistic infections. Normally our

immune system is fighting many pathogens which we do not even notice. When this defence is removed all sorts of diseases can break out. In an HIV infected individual the definition of AIDS is that there are less than 200 $CD4^+$ T cells per microlitre of blood, compared with 1000 in a healthy individual. For a long time nobody suspected this evolutionary process. It was thought that during the long period of apparent health the virus was in some sense dormant and not doing anything. This was completely wrong. By doing measurements of the number of virus particles at short time intervals and using mathematical modelling the turnover rate of the virus can be computed and it is very large. This results in a high mutation rate and the production of many different mutant versions of the virus (quasispecies). The quasispecies undergo strong natural selection so that they can rapidly become resistant to any drug. Knowing the reproduction rate of the virus it can be computed how many drugs with independent mechanisms are required to avoid the development of resistance on a certain timescale. The answer is that three drugs are needed to ensure long-term protection. It was the use of three-drug therapy that produced the breakthrough.

How is the rate of viral turnover computed? The starting point is the following system of ODE (basic model of virus dynamics).

$$\dot{x} = \lambda - dx - \beta xv \tag{1}$$

$$\dot{y} = \beta xv - ay \tag{2}$$

$$\dot{v} = ky - uv \tag{3}$$

Here x, y and v are the numbers of uninfected cells, infected cells and virus particles, respectively. The other quantities in the equations are positive constants. Notice that in this system the immune system plays no role. There is free interaction between the virus and the host cells. This model is closely related to models commonly used in population dynamics. In modelling HIV the role of the animals in the population dynamics models is taken over by the cells and the virions. There is a parameter $R_0 = \frac{\beta\lambda k}{adu}$ which is diagnostic for the nature of the dynamics. It is called the basic reproductive ratio. If it is less than one all solutions converge to a stationary solution without virus. If it is greater than one then the stationary solution without virus becomes unstable and all solutions converge to an endemic stationary solution with a non-zero amount of virus. In the literature it is often assumed without further comment that solutions will always converge to some stationary solution. That this is actually the case was proved by Korobeinikov [3] using a Lyapunov function. Interestingly this function was found using the mathematical analogy between these models and models of the spread of infection through a population of animals or humans which had previously been analysed.

The turnover rate of the virus can be determined using one of the drugs against HIV. Suppose that a patient is treated with a reverse transcriptase inhibitor which prevents infected cells from producing new virus. Suppose for simplicity that this drug is 100% effective. In terms of the model this corresponds to setting $\beta = 0$. Then the equations for y and v decouple from that for x and are linear. They can be solved explicitly. The number of virus par-

ticles is a sum of two decaying exponential functions. In one of these the rate is determined by the death rate of the cells while in the other it is determined by the rate of removal of virus. Under the assumption that the second rate is faster the second rate can be used to estimate the turnover rate of the virus. This second rate was determined by frequent measurements of the amount of virus in the blood of patients and this demonstrates the high turnover of virus. The half-life of an infected cell is about two days. More information on these subjects can be found in [5].

4 A heteroclinic lizard and the big bang

The common side-blotched lizard *Uta stansburiana* is found in the western United States. Males of this species come in three forms with different throat colours, orange, blue and yellow. When competing for mates they have the following cyclic pattern of dominance: orange > blue > yellow > orange. It is possible to set up a simple ODE model for the population dynamics of these animals. The unknowns are the proportions of the three different forms and by definition these must add up to one. The state space of the model is a simplex in $\mathbf{R}^3$. The corners of the simplex are stationary solutions. the edges are heteroclinic orbits. The solution tends to one corner as $t \to -\infty$ and another as $t \to +\infty$. Concatenating these gives a heteroclinic cycle. It turns out that all solutions in the interior converge to the boundary as $t \to \infty$ while moving around the triangle for ever. The heteroclinic cycle is globally stable. The different types of males dominate in a cyclic manner. More information about this can be found in [2]

A similar mathematical situation is found in the study of the big bang, the initial singularity of our universe. The basic equations are the Einstein equations of general relativity and it is of interest to study solutions which depend only on a suitable time coordinate. This leads to a system of four ordinary differential equations. Particles may move away from each other and towards each other in a direction-dependent manner. In one simple but important model it is enough to consider the motion in three orthogonal directions. The distances between the particles in these three directions are denoted by $a(t)$, $b(t)$ and $c(t)$. Define quantities k_i by $\dot{a}/a$, $\dot{b}/b$ and $\dot{c}/c$. In the approach to the big bang the quantity $H = \dot{a}/a + \dot{b}/b + \dot{c}/c$ is never zero and we can introduce the normalized quantities $p_i = k_i/H$. They satisfy $\sum_i p_i = 1$. There are stationary solutions which satisfy the equation $\sum_i p_i^2 = 1$ and form a circle in the state space. There are numerous heteroclinic orbits which link one of these stationary solutions to another. Concatenating these gives many heteroclinic chains which can be very complicated. The simplest one is a heteroclinic cycle which is a triangle. It can be shown that there is a codimension one submanifold which has the property that any solution which starts on that manifold converges to the triangle as the big bang is approached. Thus the three directions take turns being the one which expands towards the past just as the three forms of the lizard alternate. In the idealized model this happens infinitely often as

the singularity is approached. There is a forty-year old picture of the dynamics of solutions of the Einstein equations (BKL picture) which suggests that this kind of oscillatory behaviour (approach to a heteroclinic chain) is typical but a mathematical proof of this is still lacking. The case of the triangle was treated in [4]

Although the areas of application of mathematics involved in these two cases are so different there is a structural relationship between them. When studying dynamical systems with a view to applications it is common to concentrate on those which are structurally stable, i.e. those where the qualitative behaviour of the solutions is not changed by small changes in the coefficients in the system. This is natural since the fact that measurements are always subject to error means that we can never know the values of the parameters to be used in a model exactly. Heteroclinic cycles can always be destroyed by a small change in the coefficients and so the phenomena I have just been discussing do not have the desired property of structural stability. This is related to a point made by Karl Sigmund in a plenary talk he gave at the International Congress of Mathematicians in Berlin in 1998 and this was one of the first things which sparked off my interest in mathematical biology. He remarked that a lot of what can be found in textbooks on dynamical systems is not helpful for the problems which come up in mathematical biology. The reason he gave was the occurrence of absolute elements which cannot be perturbed. Independently of the details of the way populations are modelled one thing is clear: if the population of a species is zero at one time it will remain zero. This means that there is an invariant manifold which cannot be perturbed away. We should look at systems which are structurally stable modulo an invariant manifold and this does permit heteroclinic chains. In the case of the big bang we have a similar situation where the vanishing of one of the scale factors a, b, c, which ends the evolution, is an absolute element.

5 Liesegang rings

On 6th May 2008 I attended a lecture by Arnd Scheel at the Free University in Berlin about Liesegang patterns. On the overhead projector used the name Liesegang was written. Is this a coincidence? Not entirely, as I will explain later. Liesegang patterns are widely known due to the work of Raphael Eduard Liesegang in 1896 although he was probably not the first one to make such a observation. An experimental set-up which can be used is the following. A Petri dish is covered with a layer of a gel containing potassium dichromate. Then a drop of silver nitrate solution is deposited at the centre of the dish. It is observed that over a period of hours coloured rings appear which are centred on the point where the drop of solution was. It turns out that similar observations can be made in many other chemical systems. Another typical set-up is to take a test tube filled with gel containing one chemical and put a solution of another chemical on top of the gel. In this case horizontal bands are produced. Part of the fascination of the Liesegang phenomenon is the striking visual patterns

which accompany it. Of course it should not be forgotten that not everything which looks the same must have the same underlying mechanism. This kind of phenomenon appears to be widespread in chemistry and has also been invoked in connection with certain biological phenomena.

It is tempting to try and find a mathematical explanation of the Liesegang phenomenon. It has been observed that the position of the rings follows a geometric progression, that the time of their appearance goes as the square root of the distance and empirical laws for their thickness have also been stated. Thus there are some definite things which a mathematical model could try to reproduce. One type of model involves systems of reaction diffusion equations. These have the form

$$\frac{\partial u_i}{\partial t} = D_i \Delta u_i + f(u). \tag{4}$$

Here u_i are real-valued functions of time t and some spatial variables and they are denoted collectively by u. They represent the concentrations of chemicals. Δ is the Laplacian with respect to the spatial variables and models diffusion of the chemicals. The D_i are diffusion coefficients. The function f describes how the chemicals react. In the past this phenomenon has been modelled using non-smooth functions f, something which is at the very least aesthetically unpleasant. The aim of Scheel's work was to find an alternative model where f is smooth. Suppose we restrict to one spatial dimension, which is appropriate for describing bands rather than rings. Rather than describing the process of formation of the rings it is easier to concentrate on the final steady state. In the case of one space dimension this is modelled by a solution of a system of ODE. In fact it turns out to be important to find a homoclinic orbit of this dynamical system, a solution which converges to a stationary solution in the past and to the same stationary solution in the future. The bands are then described by solutions which asymptotically approach the homoclinic solution.

Homoclinic orbits, like heteroclinic cycles, are not structurally stable. How can this be consistent with the frequent appearance of this phenomenon in chemical and biological systems? This is the same problem as we saw before with the lizards and it has the same type of solution. There must be some absolute element which restricts relevant perturbations to those which preserve that restriction. This is the explanation proposed by Scheel. In this case the absolute element comes from the fact that the rings are formed by precipitation which is an irreversible reaction. Pertubations which would make this reaction 'slightly reversible' are not permitted. In terms of the ODE system for steady states this irreversibility leads to a partial decoupling and perturbations should not restore the coupling. It then turns out that there are smooth nonlinearities leading to a homoclinic orbit which is stable under the restricted class of perturbations. More information can be found in [8].

What about the overhead projector? Liesegang, born in 1869, was not a typical academic - he began studying chemistry but did not graduate. Instead he became one of the directors of the company founded by his father. This company was involved in photography and the experiments with the patterns were related to that. The company was also involved in producing optical

instruments and this makes the connection to the overhead projector. It went bankrupt in 2002. Later in life Liesegang worked as a scientist at the Kaiser Wilhelm Institute for Biophysics. In 1937 he became director of the Institute for Colloid Research.

6 Kestrels impose a Dirichlet boundary condition

In the north of England there is an area called the Kielder Forest with a lake in the middle and the region around the lake is inhabited by a population of the field vole *Microtus agrestis* (Swedish name *åkersork*). It is well known that populations of voles undergo large fluctuations in time. What is less known is what the spatial dependence is like. There are two alternative scenarios. In the first the population density of voles oscillates in a way which is uniform in space. In the seond it is a travelling wave of the form $U(x - ct)$. In that case the population at a fixed point in space oscillates in time but the phase of the oscillations is different at different spatial points. There is relatively little observational data on this. A dedicated observer collected data on the population of voles in the Kielder Forest which provides information on both the spatial and temporal variation of the population density. This data is the basis for the modelling which I will now describe.

The main predators of the voles are weasels *Mustela nivalis* (Swedish *vessla*). It is possible set up a model where the unknowns are the population densities of voles and weasels. Their interaction is modelled in a simple way common in predator-prey models. Their spatial motion is described by a diffusion term. In this way a system of reaction diffusion equations is obtained. These are parabolic equations and so they are non-local in space. The unknowns are defined on a region with boundary, the complement of the lake. Because of this we need not only initial values to determine a solution but also boundary conditions. How should they be chosen? In the area around the lake there live certain birds of prey (kestrel, Swedish *tornfalk*). They hunt voles from the air. In most of the area being considered there is very thick vegetation and the voles can easily hide from the kestrels. Thus the direct influence of the kestrels on the vole population is negligible and the kestrels do not need to be included in the reaction-diffusion system. They do, however, have a striking indirect effect. On the edge of the lake there is a narrow strip with little vegetation and any vole which ventures into that area is in great danger of being caught be a kestrel. This means that the kestrels essentially enforce the vanishing of the population density of voles at the edge of the lake. In other words they impose homogeneous Dirichlet boundary conditions on one of the unknowns at the boundary. Note that this is incompatible with spatially uniform oscillations. On the boundary oscillations are ruled out by the Dirichlet condition. When the PDE are solved numerically what is seen that the shore of the lake generates a train of travelling waves which propagate away from it. Information on the original work by Jonathan

Sherratt and collaborators can be found in [9] and [10].

7 Chemical reaction network theory and signal transduction in T cells

T cells and their population dynamics were already encoutered in the section on HIV. Another application of mathematical modelling to T cells concerns the processes which take place inside them. Like in other living cells an important role is played by signalling networks, networks of chemical reactions which transfer information. The task of a T cell within the immune system is to recognize a foreign substance (its antigen) and then become activated and react in a particular way. The main event causing T cell activation is when the antigen binds to a molecule on the surface of the cell called the T cell receptor. The information then has to be passed to transcription factors in the nucleus of the cell. These bind to the DNA with the result that certain proteins are produced in greater or lesser amounts. This constitutes a change in the behaviour of the cell. In T cell activation there are three important transcription factors involved in activation called NFAT, NFκB and AP-1. Here only the first of these will be considered. There is a signalling pathway which starts at the T cell receptor and ends at NFAT.

Part of the NFAT pathway is as follows. Calcium ions flow into the cytosol (the part of the cell outside the nucleus) and caused a protein called calcineurin to become active. In the resting state of the cell there are thirteen phosphate groups bound to the NFAT molecule which is situated in the cytosol. Calcineurin catalyses the removal of these and when this has happened the NFAT can move into the nucleus and bind to the DNA. It is found experimentally that the activation of the NFAT pathway is digital, i.e. in a single cell either almost no NFAT is in the nucleus or almost all of it is. This process is like a switch and there are essentially no cells where the pathway is 'half on'. A mathematical model explaining this digital behaviour in a way which depends on the large number of phosphate groups has been set up by Salazar and Höfer [7]. Their analysis was based on explicitly determining certain stationary solutions of (a slightly simplified version of) the system of ODE they derived. They did not address the issue of whether these stationary solutions are of relevance for the dynamics. Later I showed that all solutions converge to the stationary solutions they found as $t \to \infty$ [6].

The system of Salazar and Höfer is large (56 equations and 134 parameters). This is typical of many systems of chemical reactions arising in biology. How can we hope to control the global dynamics of general solutions of such a big system? The obvious thing which comes to mind is to find a Lyapunov function but where should that come from? I was able to use a theory called chemical reaction network theory (CRNT) to accomplish this for the Salazar-Höfer system. When this theory applies it gives very strong conclusions. There remains the question of how often it does apply in examples arising in applications. In the case of

the Salazar-Höfer model one of the simplest results of CRNT, the Deficiency Zero Theorem, applies. Looking at the proof of this theorem reveals that one of the main tools used is the construction of Lyapunov function, confirming what was previously indicated. (The other main tool is a result from linear algebra, the Perron-Frobenius theorem.) This Lyapunov function in turn originates in thermodynamics. In special cases it agrees with the Helmholtz free energy.

References

[1] Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M., Markowitz, M. 1995 Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature 373, 123–126.

[2] Hofbauer, J. and Sigmund, K. 1998 Evolutionary games and population dynamics. Cambridge University Press, Cambridge.

[3] Korobeinikov, A. 2004 Global properties of basic virus dynamics models. Bull. Math. Biol. 66, 879–883.

[4] Liebscher, S., Härterich. J., Webster, K. and Georgi, M. 2011 Ancient dynamics in Bianchi models: approach to periodic cycles. Commun. Math. Phys. 305, 59–83.

[5] Nowak, M. A. and May, R. M. 2000 Virus dynamics: mathematical principles of immunology and virology. Oxford University Press: Oxford.

[6] Rendall, A. D. 2012 Mathematics of the NFAT signalling pathway. SIAM J. Appl. Dyn. Sys. 11, 988-1006.

[7] Salazar, C. and Höfer, T. 2003 Allosteric regulation of the transcription factor NFAT1 by multiple phosphorylation sites: a mathematical analysis. J. Mol. Biol. 327, 31–45.

[8] Scheel, A. 2009 Robustness of Liesegang patterns. Nonlinearity 22, 457–483.

[9] Sherratt, J. A., Lambin, X., Thomas, J. A. and Sherratt, T. N. 2002 Generation of periodic waves by landscape features in cyclic predator-prey systems. Proc. R. Soc. Lond. B 269, 327–334.

[10] Sherratt, J. A. 2003 Periodic travelling wave selection by Dirichlet boundary conditions in oscillatory reaction-diffusion systems. SIAM J. Appl. Math. 63, 1520–1538.

[11] Strathern, Paul, 2005, A brief history of medicine. Robinson, London.