

Mathematising biology

Smallbone, Kieran and Wang, Yunjiao

2015

MIMS EPrint: **2015.21**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: <http://eprints.maths.manchester.ac.uk/>

And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

Mathematising biology*

Kieran Smallbone[†] Yunjiao Wang[‡]

6 February 2015

Dave Broomhead was editor of *Mathematics Today* for ten years; his lively editorials could approach current mathematical affairs from unexpected directions. In the October 2003 edition, he described the 16th-century engraving *Melencolia I* by Albrecht Dürer [1]. It shows an angel, looking pretty fed up, clutching a pair of dividers and staring off into space. Amongst the confusion and clutter surrounding the angel are various mathematical objects, including an unusual polyhedron (now known as Dürer’s solid) and a magic square giving the date of the work as 1514. Dave could empathise with the state of mind of the angel – the solitary artist (or mathematician), lost in thought as they struggle to bring order to the surrounding chaos before creating something new. He also felt that the disorder in which the angel sits could easily be his office!

The idea of mathematical research as a solitary activity is at odds with funding bodies’ drive for “team science”, whereby different disciplines must mingle to merge their languages and cultures to pursue a common goal. Biology, in particular, has no well-developed, mathematical basis. Dave writes [1]:

In the history of science, many of the great advances have been made possible by the introduction of appropriate, innovative mathematical ideas. [...] Understanding the structural and dynamical complexity of biological systems is arguably the greatest intellectual challenge before us.

The first steps toward mathematising biology were taken in 1913, by German biochemist Leonor Michaelis and Canadian doctor Maud Menten [2]. They proposed a mathematical model of an enzymatic reaction mechanism, invertase. It involves an enzyme E reversibly binding to a substrate S to form a complex ES, which in turn is converted into a product P and the enzyme (see Fig. 2(a)). Applying the law of mass action results in a system of four non-linear ordinary differential equations. However, using a simplifying assumption – for example, that the enzyme concentration is much less than the substrate concentration – the rate of reaction from S to P reduces to

$$R = V_{\max} \frac{S}{S + K_m}, \quad (1)$$

*A shorter version of this article – omitting the appendix – will appear in a special issue of *Mathematics Today* in memory of Dave Broomhead.

[†]Manchester Centre for Integrative Systems Biology, University of Manchester.

[‡]Department of Mathematics, Texas Southern University.

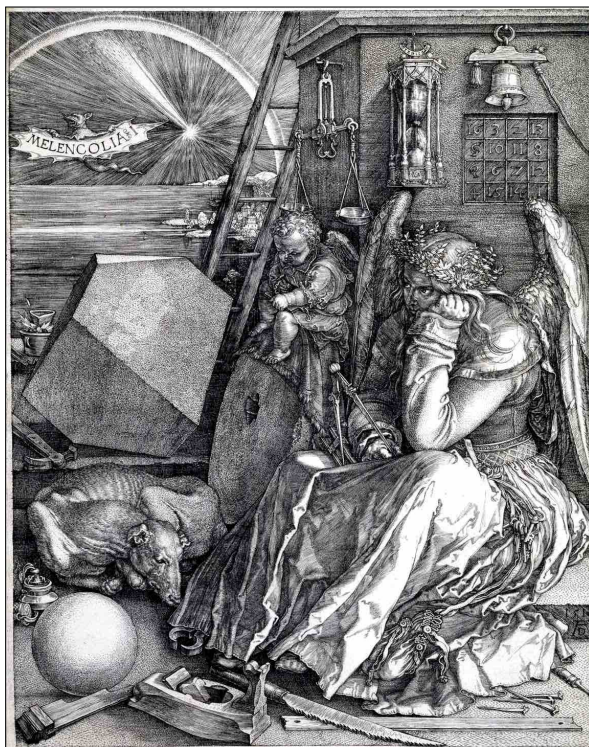


Figure 1: Dave Broomhead in his office. Or, *Melencolia I*. [Image in the public domain.]

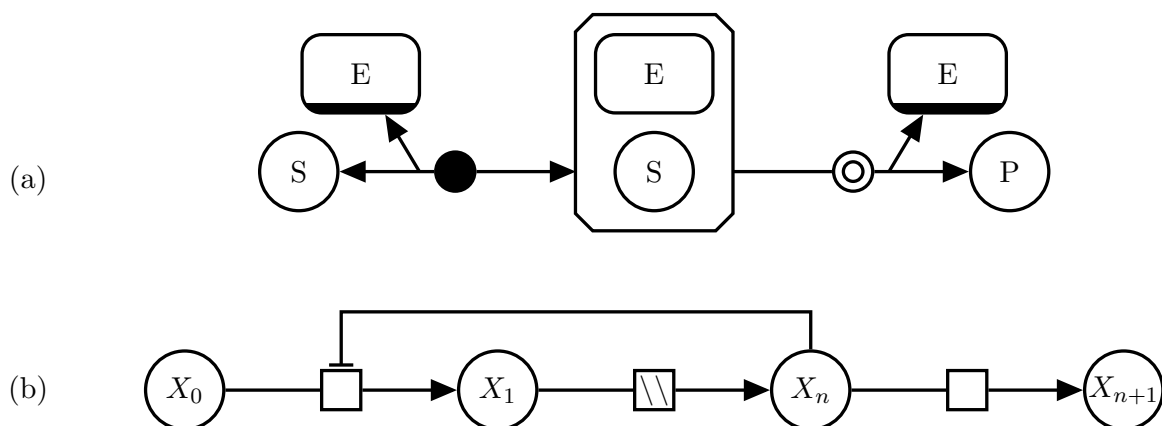


Figure 2: (a) Michaelis–Menten model of enzymatic reactions. (b) A biochemical pathway created through connection of $n + 1$ Michaelis–Menten reactions. The chain may oscillate either due to oscillatory input, or due to the presence of a negative feedback loop. Note that these diagrams are drawn in standardised SBGN format [6].

where V_{\max} is the maximum reaction rate, attained as $S \rightarrow \infty$, and K_m is the substrate concentration at which the reaction rate is at half-maximum, known as the Michaelis–Menten constant.

One hundred years later, the model of Michaelis and Menten remains central to biochemical dogma. But we have also come to realise that, inside a cell, reactions do not exist in isolation. Metabolic networks contain thousands of reactions whose substrates interact with one another in non-linear ways. Systems biology characterises complex biological systems – such as metabolic networks – through mathematical modelling, and aims to discover emergent properties that cannot be found through reductionist analyses.

We may put together a simple systems biology model of a biochemical pathway by connecting a series of enzymatic reactions through a single chain. The model has a fixed input substrate concentration X_0 and a fixed output product concentration X_{n+1} separated via n variables X_i , and $n + 1$ reactions R_i (see Fig. 2(b)). We can model these reactions using Michaelis–Menten kinetics:

$$R_i = V_i \frac{X_i}{X_i + K_i}, \quad (2)$$

so the equations associated with the pathway have the form $X'_i = R_{i-1} - R_i$.

The system has a unique steady state, but this is reached only if $V_i > R_0$ for all i . If this doesn't hold, one of the intermediate metabolite concentrations will increase unboundedly – a biologically unreasonable prediction. The maximum rates V_i are biological parameters that can be measured in the lab and, within the cell, are distributed over a range of values. The probability that any $V_i < R_0$ is some $p > 0$, and so the probability that our model “works” is $(1 - p)^n \rightarrow 0$ as $n \rightarrow \infty$. Despite starting with a very good model at the single reaction level, once we consider multiple reactions simultaneously, the model is almost certain to fail once it reaches a certain size. Biology is complex, and building systems biology models is hard.

In 2005, Marty Golubitsky gave a seminar at the University of Manchester [3]. He presented a result on the dynamics of a three-node feed-forward chain, where the evolution of each node is dependent only on itself and the preceding node, as in our example above. Dave asked what would happen if the first node of the chain is forced to have a periodic signal. Rhythmicity is common in biology; we experience the sleep–wake cycle, a beating heart, and breathing. At the molecular level, the p53 signalling pathway oscillates in response to DNA damage; the NF- κ B signalling pathway – which Dave studied extensively (e.g. [4, 5]) – shows damped oscillations in response to a wide variety of stimuli.

We can subject our pathway to periodic forcing, by defining

$$X_0 = \bar{X}_0(1 + \alpha \sin \omega t), \quad (3)$$

with $\alpha \leq 1$. Without this signal, ($\alpha = 0$), the system has a stable equilibrium. With the forcing ($\alpha > 0$ and $\omega \neq 0$), the downstream nodes oscillate with the same period as the signal (see Fig. 3(a)).

Sustained oscillations may also occur through the introduction of a negative feedback loop, as shown in Fig. 2 (b). We allow the final variable to feedback negatively to the first reaction,

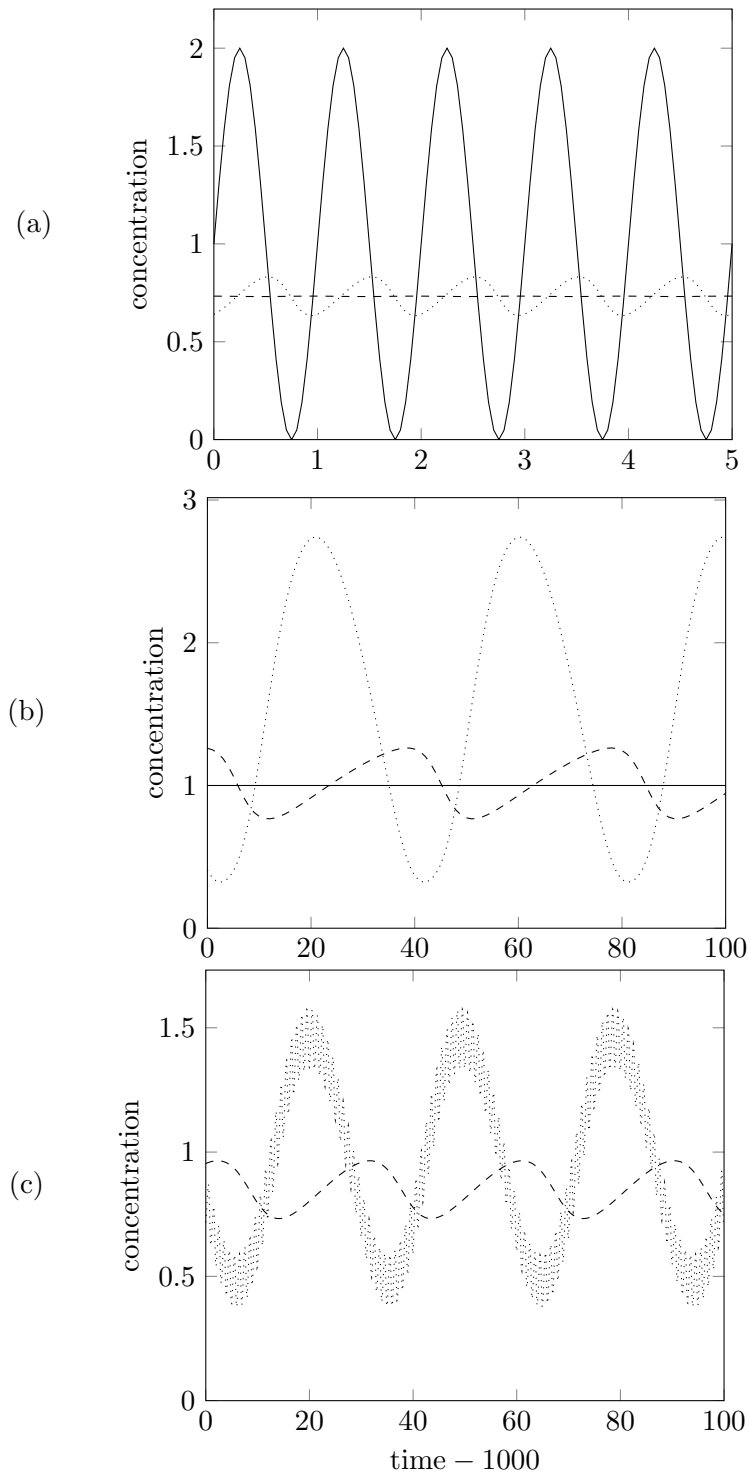


Figure 3: Late-time simulations of the pathway. (a) With periodic forcing of the first node. (b) With oscillations driven by a negative feedback loop. (c) The combined effect of forced and inherent oscillations. Shown are changes in X_0 (solid), X_1 (dotted) and X_n (dashed) with time.

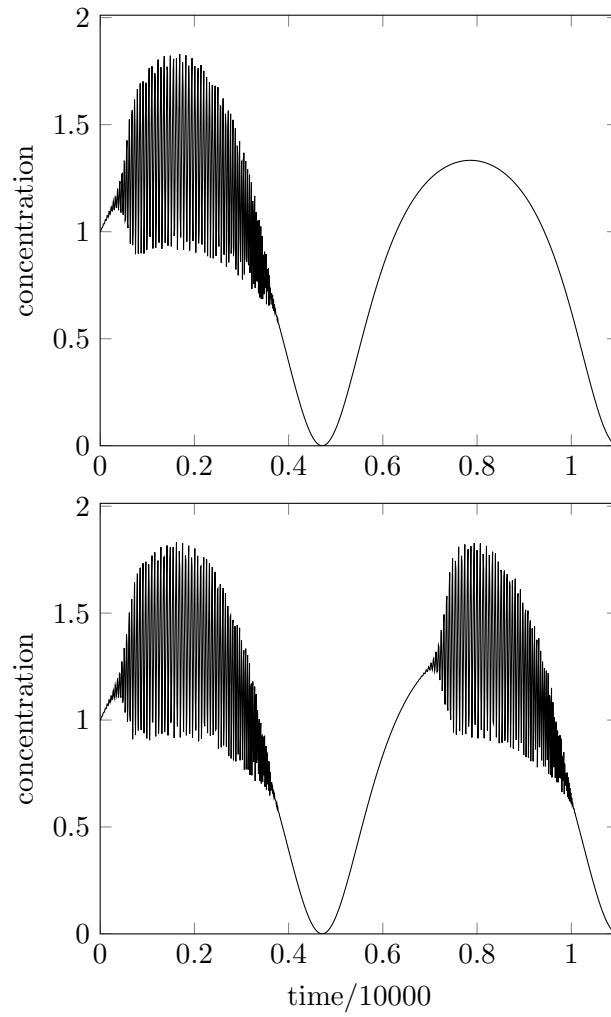


Figure 4: Chaotic behaviour observed in the combined system. A tiny change in initial conditions between the top and bottom panels leads to a large change in system behaviour. Shown are changes in X_n with time.

setting

$$R_0 = V_0 \frac{X_0}{X_0 + K_0} \frac{K_I}{X_n + K_I}. \quad (4)$$

As V_0 increases to a sufficiently large value, we pass through a Hopf bifurcation, causing a transition from the stable steady state dynamics of the feed-forward pathway, to the sustained oscillatory behaviour of the negative feedback loop. In Fig. 3(b) we simulate this model, which has slower oscillations than with periodic forcing.

The natural step is to combine these two models, inducing pathway oscillations both via periodic forcing and negative feedback. X_0 varies as in Eq. (3), and R_0 is defined as in Eq. (4). The results the simulation are in Fig. 3(c); we see a combination of oscillations, at two time-scales.

In addition to oscillations, quasi-periodic and chaotic behaviour may occur for some parameter values. For example, reducing the frequency ω of the input signal – so that the negative feedback becomes the fast time-scale, and the forced oscillation the slow time-scale – leads to the results in Fig. 4. The top and bottom panels simulate from near-identical initial conditions, but have very different outcomes.

Such sensitivity to initial conditions is one of the characteristics of chaotic dynamics. But initial data is important in determining the behaviour of all models, a fact Dave emphasised when writing about NF- κ B [5]:

From the point of view of a mathematician this is such an obvious point that it might seem hardly worth the emphasis, but the initial state of a cell is an extremely difficult idea for the experimentalist. [...] Even with a perfect model—but without knowledge of the initial state—there are likely to be distinct qualitative differences between the behaviour of model and experiment.

In this article, we have started with a very simple biological system – a pathway with negative feedback – and exposed it to external rhythmicity – as occurs naturally. The intuitive, verbal reasoning approaches favoured by many biologists are clearly insufficient to describe the resulting chaotic dynamics. This is not the only type of chaos encountered by mathematical biologists. Returning to *Melencolia*, Dave writes [1]:

[T]he vastness of the clutter that surrounds mathematical biologists would tax even Dürer's considerable representational talents. For the past decade biologists have been given inordinate quantities of money to generate data in astronomical amounts. So huge is this morass that a whole new discipline has emerged which effectively treats the data itself as an experimental system to be probed and trawled in the hope [of] retrieving something useful.

It is over ten years since Dave wrote this editorial and there remains a great disparity between the amount spent on data collection and the amount spent on developing a mathematical underpinning for biology. Experience from other areas of science has taught us that quantitative

methods are needed to develop comprehensive theoretical models for interpretation, organisation and integration of this data. Once thought of as too simplistic to describe complex biological phenomena, we now see that mathematical models, continuously revised by new information, can be used to guide experimental design and interpretation. We hope that funding bodies' continued interest in interdisciplinary science presages a time when serious investment is made into the effort to mathematise biology.

Appendix

Model 1: Michaelis–Menten kinetics

Applying the law of mass action, which states that the rate of a reaction is proportional to the product of the concentrations of the reactants, gives a system of four non-linear ordinary differential equations that define the rate of change of reactants with time

$$\begin{aligned} S' &= -k_f ES + k_r C \\ E' &= -k_f ES + k_r C + k_{cat} C \\ C' &= k_f ES - k_r C - k_{cat} C \\ P' &= k_{cat} C \end{aligned} \tag{A-1}$$

where C denotes the concentration of the enzyme-substrate complex. We take as initial conditions $S(0) = S_0$, $E(0) = E_0$ and $C(0) = P(0) = 0$. Analysis performed by British botanist George Briggs and geneticist John Haldane in 1925 assumed that the concentration of the complex does not change on the time-scale of product formation [7] – known today as the quasi-steady state approximation – which leads to reaction rate

$$P' = -S' = V_{\max} \frac{S}{S + K_m}, \tag{A-2}$$

where $V_{\max} = k_{cat} E_0$ and $K_m = (k_{cat} + k_r)/k_f$. Note that Michaelis and Menten’s original analysis resulted in the same final formula, albeit via a different assumption. The quasi-steady state approximation can be shown to be valid if $E_0 \ll S_0 + K_m$ (see, for example, [8]); the catalytic effectiveness of enzymes is reflected in the small concentrations needed in their reactions as compared with the concentrations of the substrates involved, and as such this constraint normally holds (see Fig. A-1).

Model 2: Biochemical pathway

The system’s unique steady state is given by

$$X_i = \frac{R_0 K_i}{V_i - R_0}. \tag{A-3}$$

As concentrations may only take positive values, this steady state exists if and only if $V_i > R_0$ for all i .

We simulate the system with $n = 3$ (see Fig. A-2). When $V_n < R_0$, we see that the system does not reach a steady state, and $X_n \rightarrow \infty$. The eigenvalues at steady state are given by

$$\lambda_i = -\frac{(V_i - R_0)^2}{V_i K_i} = -\frac{R_0 K_i}{X_i (X_i + K_i)} < 0, \tag{A-4}$$

and so, if the steady state exists, it is stable. Note that Michaelis–Menten constants are often of similar magnitude to substrate concentrations [9], allowing us to simplify $K_i \approx X_i$ and $\lambda_i \approx -R_0/2K_i$. Thus the stiffness ratio of the system may be approximated as

$$\phi = \frac{\max_i(|\operatorname{Re}(\lambda_i)|)}{\min_j(|\operatorname{Re}(\lambda_j)|)} \approx \frac{K_{\max}}{K_{\min}}. \quad (\text{A-5})$$

Michaelis–Menten constants are distributed over a range of values, and thus the stiffness of this simple model will increase with its size, and in turn simulation will prove more difficult.

Model 3: Periodic forcing

In Fig. 3(a) we simulate this system, using parameter values $n = 3$, $V_i = 2$, $K_i = 1$, $\bar{X}_0 = 1$, $\alpha = 1$ and $\omega = 2\pi$. Notice that, although the mean value of the signal is the same as the fixed point $X_i = 1$, the mean value of subsequent variables is lower, and the amplitude of their response lower than the signal.

Model 4: Negative feedback

If we run a simulation with $V_0 = 4$ and $K_I = 1$, the system is very similar to the original, with stable fixed point $X_i = R_i = 1$. However, as we increase V_0 , we encounter a Hopf bifurcation at $V_0 = 32$. For example, at $V_0 = 40$, the system displays sustained oscillations about the unstable fixed point $X_i = 10$, $R_i = 20/11$. Rescaling so that this fixed point is given by $X_i = 1$ and $R_i = 1$, as before, leads to parameter values $V_0 = 22$, $K_0 = 1$, $K_I = 0.1$ and $V_i = 1.1$ and $K_i = 0.1$ for $i \geq 1$. This system is simulated in Fig. 3(b). Its period of oscillation is around 39 time units.

Model 5: Combined oscillations

In Fig. 3(c) we simulate the combined model with parameter values as previously. In Fig. 4 we reduce the frequency ω from 2π to 10^{-3} , making the signal become the slow dynamics of the system. The top panel shows results from initial conditions $X_i = 1$, whilst in the bottom panel we make the tiny change $X_n = 1 - 10^{-3}$.

The models described here are available in standardised SBML format [10] from the BioModels database [11] with accession numbers

1. MODEL1503180002
2. MODEL1503180003
3. MODEL1503180004
4. MODEL1503180005
5. MODEL1503180006

References

- [1] Broomhead D.S. (2003) Editorial. *Mathematics Today*, vol. 39, pp. 136-137.
- [2] Michaelis L. and Menten M.L. (1913) Die Kinetik der Invertinwirkung. *Biochemische Zeitschrift*, vol. 49, pp. 333–369. Translation at doi:10.1021/bi201284u.
- [3] Golubitsky M. Coupled systems: theory & examples. Seminar at the University of Manchester, 26 October 2005.
- [4] Nelson D.E., et al. (2008) Oscillations in NF- κ B signaling control the dynamics of gene expression. *Science*, vol. 306, pp. 704–708. doi:10.1126/science.1099962 .
- [5] Broomhead D.S., et al. (2008) Oscillations in the NF- κ B Signaling Pathway. MIMS EPrint 2009.105.
- [6] Le Novère N., et al. (2009) The Systems Biology Graphical Notation. *Nature Biotechnology*, vol. 27, pp. 735–741.
- [7] Briggs, G.E. and Haldane, J.B.S. (1925) A note on the kinematics of enzyme action. *Biochemical Journal*, vol. 19, pp. 338–339.
- [8] Murray, J.D. (2002) *Mathematical Biology: I. An Introduction*. Springer. ISBN 978-0-387-95223-9
- [9] Smallbone K. and Mendes P. (2013) Large-Scale Metabolic Models: From Reconstruction to Differential Equations. *Industrial Biotechnology*, vol. 9, pp. 179–184. doi:10.1089/ind.2013.0003
- [10] Hucka M., et al. (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, vol. 19, pp. 524–531. doi:10.1093/bioinformatics/btg015
- [11] Li C., et al. (2010) BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Syst Biol* 4:92. doi:10.1186/1752-0509-4-92

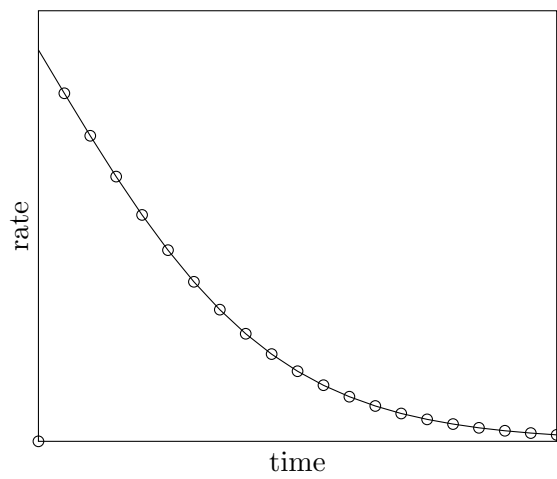


Figure A-1: After its initial transient, the Michaelis–Menten rate law (line) provides an excellent approximation to the full system of ODEs (circles). Parameter values used are $k_f = k_r = k_{cat} = 1$, $S_0 = 1$ and $E_0 = 10^{-3} \ll S_0$.

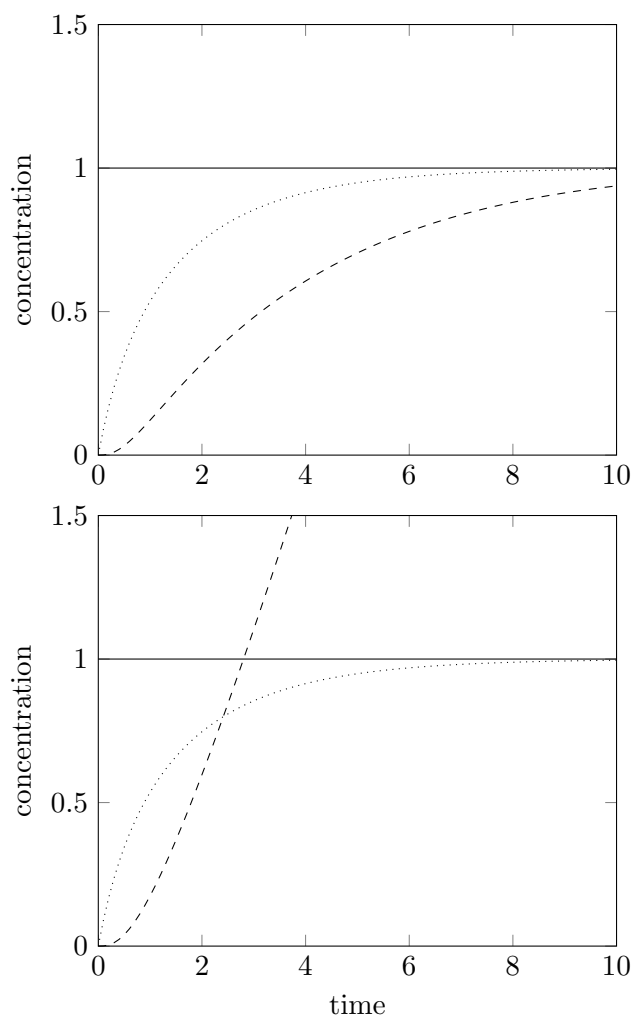


Figure A-2: Simulations of the simple biochemical pathway with $n = 3$, parameters $V_i = 2$, $K_i = 1$ and $X_0 = 1$, and initial conditions $X_i = 0$. Shown are changes in X_0 (solid), X_1 (dotted) and X_n (dashed) with time. The network in the upper panel reaches a steady state given by $X_i = 1$ and $R_i = 1$. In the lower panel, $V_n = \frac{1}{2} < R_0$ and no steady state is reached.