PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL PHYSICAL & ENGINEERING SCIENCES

Emerging areas of mathematical modelling

J. R. King

Phil. Trans. R. Soc. Lond. A 2000 **358**, 3-19
doi: 10.1098/rsta.2000.0516

THE ROYAL SOCIETY

Emerging areas of mathematical modelling

By J. R. King

Division of Theoretical Mechanics, School of Mathematical Sciences, University of Nottingham, Nottingham NG7 2RD, UK
(john.king@nottingham.ac.uk)

Through the use of specific examples, we illustrate growing areas of mathematical modelling that have arisen more recently than traditionally dominant subjects such as fluid mechanics. As high-level mathematics finds application in an expanding range of disciplines, new methods are needed to investigate the resulting models, and some of these are also illustrated. A number of recurring themes are identified, suggesting likely trends for subsequent developments.

Keywords: industrial mathematics; mathematical medicine; high-order systems; pattern formation

1. Introduction

The increasing influence of mathematical modelling, both within its traditional territory in the physical sciences and beyond (notably in biology, medicine and finance), can be attributed not only to the growing need for quantitative understanding but also to improvements in mathematical capability. The latter have resulted from recent rapid advances both in computer power, making possible the solution of problems of greater complexity, and in the sophistication of mathematical methods, particularly as applied to nonlinear problems. The subject of mathematical modelling is extremely broad in its scope and variety, and models can be classified in a number of different ways, such as deterministic against stochastic and continuous against discrete; a distinction particularly worth emphasizing here is that between models designed to represent the key features of a particular situation and those intended to study generic nonlinear phenomena. The second of these types has received by far the greater publicity, to a large degree due to popular interest in topics such as chaos and fractals; our emphasis here is very much on the former.

Our approach will be to outline (in §§ 2–5) four representative topics, each being an area of growing significance; we do this primarily through concrete examples intended to illustrate more general principles. We conclude in § 6 by drawing together some of the common themes and spelling out some of their implications for likely future developments.

2. Industrial mathematics: impurity diffusion

A considerable amount of activity in industrial mathematics ultimately has its root in the Oxford (now European) Study Groups with Industry, which were initiated in 1968 and have subsequently spawned similar meetings in a large number of countries.

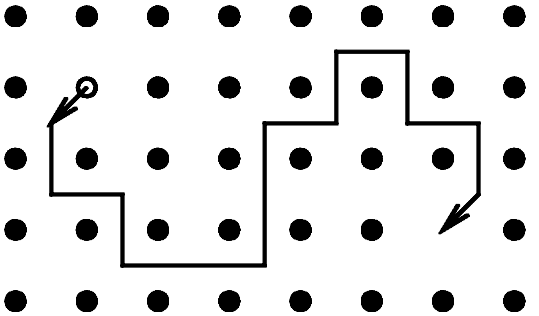

Figure 1. Schematic of dissociative mechanism (○, impurity atom; ●, semiconductor atom).

They are week-long workshops in which academic mathematicians address problems presented by the industrial participants on the first day, which often have yet to be formulated in mathematical terms. The academics present are thus confronted with issues they would not normally have considered otherwise; the consequences have included a great deal of novel mathematics (some indication of the scope of which is provided by Tayler (1986)), greatly enhancing worldwide activity in such areas as moving-boundary problems (cf. § 5 below), and strong stimuli to avoid the narrowly focused research, often in well-trodden areas, that characterizes some academic activity. One of the areas that has been a rich source of new problems, and from which we draw our example, is that of semiconductor-device fabrication. The example is intended to be indicative of some of the ways in which a mathematical model can be formulated.

Impurities are introduced in a controlled way into semiconductor wafers in order to modify their electrical properties in the fashion required. Diffused impurity profiles very often exhibit strongly nonlinear effects, being far from the profiles (such as complementary error functions) that characterize linear diffusion. A number of mechanisms can be responsible for such behaviour, the one we consider here being a substitutional–interstitial diffusion process known as the dissociative mechanism. A schematic clarifying the terminology is given in figure 1, where an impurity atom is shown leaving a substitutional (lattice) site, executing a random walk (i.e. diffusing) between interstitial sites, and then returning to the substitutional state upon encountering a vacancy (an unoccupied lattice site). Reference to much of the earlier literature is given in Tuck (1988); here we concentrate on some two-dimensional effects identified in Meere *et al*. (1995).

Model formulation can proceed from the consideration of atomic jumps of the type illustrated in figure 1; this leads to discrete formulations (details of which are given in King *et al*. (1995) for a different diffusion process, namely the interstitialcy mechanism) that have some novel features. However, diffusion lengths are usually, in practice, much larger than the atomic spacing, and the appropriate model can then be obtained either directly by continuum considerations, for which the details of the crystal lattice may be unimportant, or by taking the continuum limit of the discrete formulation. In any case, for the simplest version of the dissociative mechanism, the governing equations can be deduced as follows. Denoting by s, i and V, respectively, a substitutional impurity atom, an interstitial impurity atom and a vacancy (the concentrations of which are denoted by c_{s}, c_{i} and c_{V}), the dissociative process can

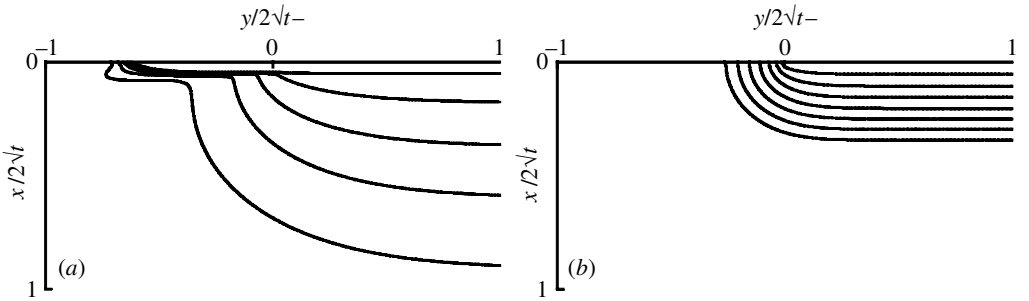

Figure 2. Contours of total impurity concentration for diffusion under a mask edge; the semiconductor surface $x = 0$ is covered by an impervious mask in $y < 0$ and by a source of impurity in $y > 0$.

be treated as a reversible chemical reaction, namely

$$\mathrm{s} \rightleftharpoons \mathrm{i} + \mathrm{V}. \tag{2.1}$$

This reaction progresses in the forward direction (illustrated by the left-hand arrow in figure 1) at a rate $k_1 c_\mathrm{s}$, this being proportional to the number of substitutional atoms available to become interstitial; it operates in the reverse direction (see the right-hand arrow in figure 1) at a rate, $k_2 c_\mathrm{i} c_\mathrm{V}$, determined by the probability that an interstitial encounters a vacancy. The probability that a vacancy occupies a given lattice site is, thus, treated as being independent of whether or not there is a neighbouring interstitial. Vacancies and interstitials are taken to be able to execute random walks over lattice and interstitial sites, respectively, with diffusivities D_V and D_i, and the substitutionals are taken to be unable to diffuse. Accounting for these diffusive effects and for the reaction (2.1) leads to the reaction–diffusion system:

$$\left.\begin{aligned} \frac{\partial c_\mathrm{s}}{\partial t} &= k_2 c_\mathrm{i} c_\mathrm{V} - k_1 c_\mathrm{s}, \\ \frac{\partial c_\mathrm{i}}{\partial t} &= D_\mathrm{i} \nabla^2 c_\mathrm{i} - k_2 c_\mathrm{i} c_\mathrm{V} + k_1 c_\mathrm{s}, \\ \frac{\partial c_\mathrm{V}}{\partial t} &= D_\mathrm{V} \nabla^2 c_\mathrm{V} - k_2 c_\mathrm{i} c_\mathrm{V} + k_1 c_\mathrm{s}. \end{aligned}\right\} \tag{2.2}$$

In practice, the reaction typically proceeds much more rapidly than diffusive processes and (2.2) simplifies for large k_1 and k_2 to the diffusion problem:

$$k_1 c_\mathrm{s} = k_2 c_\mathrm{i} c_\mathrm{V}, \qquad \frac{\partial}{\partial t}(c_\mathrm{s} + c_\mathrm{i}) = D_\mathrm{i} \nabla^2 c_\mathrm{i}, \qquad \frac{\partial}{\partial t}(c_\mathrm{s} + c_\mathrm{V}) = D_\mathrm{V} \nabla^2 c_\mathrm{V}. \tag{2.3}$$

One-dimensional solutions for impurity in-diffusion from a source at the surface typically exhibit strongly nonlinear behaviour, with the impurity profile showing a region of rapid variation near the surface and with a pronounced depletion of vacancies. The latter results from impurity atoms occupying vacancies as they diffuse into the semiconductor and is responsible for the former. Two-dimensional solutions exhibit additional features of interest. Figure 2 illustrates impurity diffusion under a mask (the model used is in fact slightly more sophisticated than (2.2) in that it accounts for the charges carried by the diffusing species and the electric fields that these charges induce; the qualitative behaviour of the models is similar, however).

Figure 2*b* shows the behaviour when c_s is small, so that c_V is almost constant and the first two equations of (2.3) reduce to the linear diffusion equation. For the simulation in figure 2*a*, however, vacancy concentrations are strongly depleted; this figure illustrates the greatly enhanced diffusion that can occur beneath the mask. The effect results from the surface acting as a source for vacancies, whose concentration is, therefore, little depleted immediately beneath the mask. This enables an interstitial impurity to become substitutional there very readily, providing a rapid diffusion pathway. This phenomenon of enhanced lateral diffusion is well known experimentally and is highly undesirable in practice; as suggested by figure 2*b*, it does not occur for simpler diffusion processes. Its successful prediction by (2.2) illustrates the efficacy of models that aim to provide concise mathematical descriptions of the specific physical processes involved.

3. Mathematical biology: multicell tumour spheroid growth

Applications to biology and medicine are among the areas of fastest growth in the scope of mathematical modelling studies. One of the approaches that has found most widespread application involves the use of reaction–diffusion systems of a type not unlike the example outlined in the previous section. Indeed, many workers in the field concentrate almost exclusively on models of reaction–diffusion type, partly motivated no doubt by the success of such systems in generating a wide variety of interesting patterns of the sort observed in animal coat patterns and elsewhere; the field has been very strongly influenced by the seminal early studies of Turing (1952). While it seems clear that reaction–diffusion phenomena are crucial in many biological and medical processes, it is very often the case that other effects (notably those involving material deformation) are equally important. In many biomedical applications there is, thus, a very great need for the more effective combination of concepts from reaction–diffusion theory with those from continuum mechanics; the model outlined below provides a simple example of such an amalgam. Continuum mechanics, which describes the deformation of solids and the flow of liquids, together with the stresses that cause or result from them, is an extremely well-developed subject. Nevertheless, its application to mathematical biology brings a number of relatively novel challenges, notably those that result from tissue growth and deformation occurring simultaneously.

Multicell spheroids grown *in vitro* are widely studied experimentally, partly in view of their possible relevance to the early (avascular) stages of tumour growth *in vivo* (before the tumour has acquired its own blood supply). A schematic of the phases of multicell spheroid growth is shown in figure 3. When a spheroid is sufficiently small, ample externally supplied nutrient is able to penetrate right through to its centre; all the cells are thus able to proliferate, resulting in exponential growth. As the tumour grows, nutrient consumption by the cells near to the surface results in those at the centre receiving an insufficient supply, causing them to become quiescent (alive but dormant) and then necrotic (dead and decomposed).

Important early modelling work includes that of Greenspan (1972). The model we describe here is given in Ward & King (1997) and treats the whole tumour on the same footing, the spatial structure indicated in figure 3 being predicted from the model rather than being enforced when it is formulated. As well as reproducing the observed spatial heterogeneity, the model also successfully predicts the first two phases of tumour growth (exponential and linear). The dependent variables involved

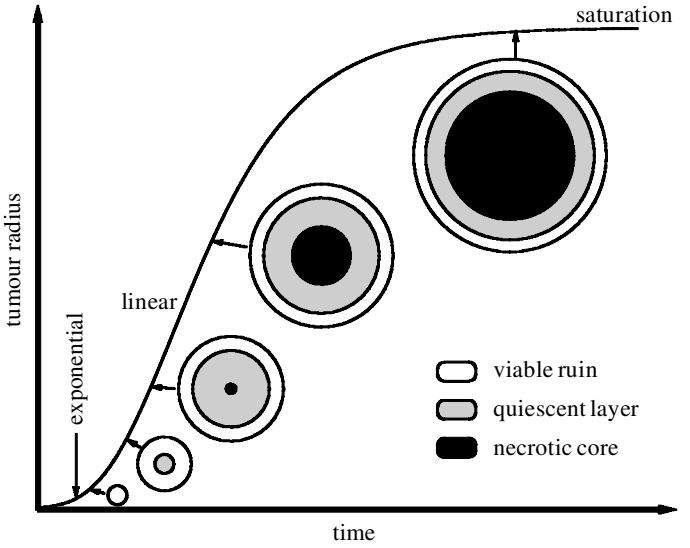

Figure 3. Schematic of multicell spheroid growth and heterogeneity.

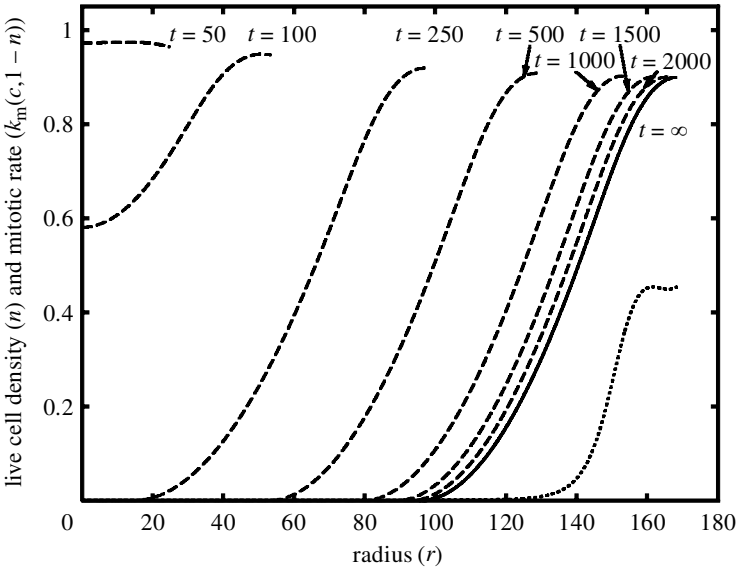

Figure 4. Evolution of live cell density in time (dashed lines) towards the steady-state (saturation) profile (solid line). The steady-state mitotic distribution is also shown (dotted line); the region in which this is very small, but the live cell density is not, is one of cell quiescence (cf. the schematic in figure 3).

are $n(\mathbf{x},t)$ and $m(\mathbf{x},t)$, the concentrations of live and dead cells, $\mathbf{v}(\mathbf{x},t)$, the local velocity within the tumour (driven by volume creation through cell birth and loss by cell death), and $c(\mathbf{x},t)$, the nutrient concentration. We have

$$\frac{\partial n}{\partial t} + \nabla \cdot (\mathbf{v}n) = (k_{\mathrm{m}}(c) - k_{\mathrm{d}}(c))n, \qquad \frac{\partial m}{\partial t} + \nabla \cdot (\mathbf{v}m) = k_{\mathrm{d}}(c)m, \tag{3.1}$$

where k_{m} is the rate of mitosis (cell division) and k_{d} is the rate of cell death. The first equation of (3.1) states that $\partial n/\partial t$, the rate of change of the live cell density, is determined by the net birth rate of cells, $(k_{\mathrm{m}} - k_{\mathrm{d}})c$, and by their transport due to the local velocity field, $\nabla \cdot (\mathbf{v}n)$. Taking the materials to be incompressible and free of voids, we have $V_{\mathrm{L}}n + V_{\mathrm{D}}m = 1$ within the tumour, where V_{L} and V_{D} are, respectively, live and dead cell volumes. It then follows from (3.1) that

$$\nabla \cdot \mathbf{v} = k_{\mathrm{m}}(c)nV_{\mathrm{L}} - k_{\mathrm{d}}(c)n(V_{\mathrm{L}} - V_{\mathrm{D}}),$$

which makes explicit how cell birth and death drive the velocity field. Finally, the nutrient concentration satisfies the reaction–convection–diffusion equation

$$\frac{\partial c}{\partial t} + \nabla \cdot (\mathbf{v}c) = D\nabla^2 c - \beta k_{\mathrm{m}}(c)n, \tag{3.2}$$

the terms on the right-hand side of which represent nutrient diffusion (diffusivity D) and consumption (parameter β). Taking the tumour to be spherically symmetric, equations (3.1)–(3.2) (together with $m = (1 - V_{\mathrm{L}}n)/V_{\mathrm{D}}$ and appropriate boundary and initial conditions) form a complete model, numerical and asymptotic solutions of which are given in Ward & King (1997). Since the model is unable to predict the final phase of growth saturation, rather than reproducing such results here, we instead give, in figure 4, a numerical simulation of the model of Ward & King (1999), which generalizes and refines the earlier model by incorporating the consumption and egress of necrotic material. Two large-time outcomes (the states to which the solution is attracted, no matter whence it starts) are then possible, namely growth saturation (illustrated in figure 4) and continued linear growth (represented mathematically by a travelling wave). The bifurcation analysis in Ward & King (1999) identifies the transition between these outcomes in terms of the values of the model parameters. *In vitro*, the growth saturation regime is always observed; that of linear growth may, however, be of relevance to behaviour *in vivo*.

A spherically symmetric tumour may become unstable (see, for example, Byrne & Chaplain 1997), developing 'hot spots' on the surface at which growth is more rapid, leading to fingering of the type described in a different context in § 5 below. Such instabilities may have important consequences for the invasiveness of a tumour *in vivo*. If the tumour lacks radial symmetry, an extra ingredient is needed in the modelling, namely a constitutive law relating deformations and stresses, and a number of such laws have been investigated. Enhanced pressures within a tumour have undesirable implications for the delivery of chemotherapeutic drugs, providing an illustration of the importance of the interplay between continuum mechanics and reaction–diffusion effect to which we have already alluded.

The adequate modelling of (vascular) tumour growth *in vivo* requires the incorporation of numerous extra aspects, of which we note the tumour's acquisition of its own blood supply through the process of angiogenesis (which has been widely modelled; see Anderson & Chaplain (1998) and references therein) and the interaction of a tumour with the surrounding (normal) tissue (which has received little modelling attention; see Perumpanani *et al.* (1997), however). In common with many other topics in mathematical biology, a vast range of important open problems remains in this area, providing continued impetus for the growth of the field.

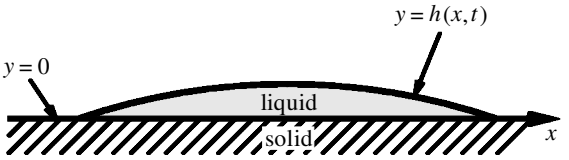

Figure 5. Schematic of spreading fluid droplet.

4. High-order diffusion

Our intention in this section is to highlight an area of growing mathematical interest with important physical applications, and to illustrate some of its difficulties.

Evolution equations of the form

$$\frac{\partial h}{\partial t} = \frac{\partial}{\partial x}\left(h^n \frac{\partial p}{\partial x}\right) \tag{4.1}$$

(where n is a positive constant) arise in describing the slow spreading of a thin fluid film over a horizontal surface $y = 0$ (see the schematic in figure 5), with $y = h(x,t)$ denoting the height of the droplet and $p(x,t)$ the pressure with the fluid. For $n = 3$, (3.1) is Reynolds's equation from lubrication theory, to which the exponent $n = 2$ is also relevant in circumstances in which the fluid is able to 'slip' over the solid. Equation (3.1) with $n = 1$ is the corresponding model for flow in a porous medium or a Hele–Shaw cell. To close the problem, the force driving the spreading of the droplet needs to be prescribed and the cases of interest here are when, suitably normalized, $p = h$ (gravity driven), $p = -\partial^2 h/\partial x^2$ (surface tension driven), and $p = \partial^4 h/\partial x^4$ (when an elastic plate covers the droplet surface), giving, respectively,

$$\frac{\partial h}{\partial t} = \frac{\partial}{\partial x}\left(h^n \frac{\partial h}{\partial x}\right), \tag{4.2 a}$$

$$\frac{\partial h}{\partial t} = -\frac{\partial}{\partial x}\left(h^n \frac{\partial^3 h}{\partial x^3}\right), \tag{4.2 b}$$

$$\frac{\partial h}{\partial t} = \frac{\partial}{\partial x}\left(h^n \frac{\partial^5 h}{\partial x^5}\right), \tag{4.2 c}$$

the first three members of a hierarchy of diffusion equations. The vanishing of the 'diffusivity', h^n, at $h = 0$ results in an important finite speed of propagation property for (4.2), whereby an interface separating a dry region (in which $h = 0$) from a wet one ($h > 0$) moves at a finite rate, rather than complete wetting occurring immediately. Equation (4.2 *a*) arises in numerous other contexts; it and its close relatives have been the subjects of a vast amount of study over the last 50 years (reviews include Kalashnikov (1987)) and they are now very well understood. These developments have formed an important branch of applied analysis, (4.2 *a*) having provided a proving ground for the development of a number of powerful methods. The focus of much attention is now turning to (4.2 *b*), and to other high-order systems, with formulations of this type occurring across a wide range of disciplines (see, for example, Bernis 1995). Our goal in what follows is to indicate some of the mathematical difficulties (often reflecting physical phenomena) associated with (4.2 *b*) that go beyond

those of (4.2 *a*); a considerable number of open questions remain, with (4.2 *b*) proving a very useful test-bed for the development of methods for analysing high-order problems.

Much of the analysis of (4.2 *a*) is based on comparison theorems, whereby particular solutions can be used to give information about much more general ones. A simple illustration is that, because $h = 0$ satisfies (4.2 *a*), it follows that if a solution is initially positive it remains so. By contrast, no comparison theorem exists for (4.2 *b*), which creates a great deal of extra mathematical difficulty and means, in particular, that the film thickness h may be able to drop to zero, implying film rupture and the possible development of dry spots.

A second major difficulty in the analysis of (4.2 *b*) concerns the lack of uniqueness of the solution; this leads to the issue of how a physically relevant solution is to be prescribed. We can highlight this difficulty, and also give information about whether or not a droplet is predicted to spread, by considering travelling-wave solutions $h = h(\zeta)$, $\zeta = x - qt$, where q is a constant. Taking $h = 0$ for $\zeta \geqslant 0$ and prescribing conservation of fluid at $\zeta = 0$, we have, for (4.2 *a*), that

$$-q = h^{n-1}\frac{\mathrm{d}h}{\mathrm{d}\zeta}, \quad \text{giving } h = (nq(-\zeta)_+)^{1/n}, \tag{4.3}$$

where $(-\zeta)_+ = \max(0, -\zeta)$, the qualitative behaviour of which is insensitive to the value of n. For (4.2 *b*) we instead have (cf. Boatto *et al.* 1993)

$$q = h^{n-1}\frac{\mathrm{d}^3 h}{\mathrm{d}\zeta^3}, \tag{4.4}$$

one solution of which is

$$h = \left(\frac{n^3 q}{3(3-n)(2n-3)}(-\zeta)_+^3\right)^{1/n}, \quad n \neq \tfrac{3}{2}, 3. \tag{4.5}$$

It would not be appropriate to go into all the consequences of such expressions here, but we note the following points.

(i) The solution (4.5) is only of the same type as (4.3), in the sense that $q > 0$ is required (implying that the droplet spreads outwards), when $\frac{3}{2} < n < 3$. Unlike (4.2 *a*), equation (4.2 *b*) exhibits a number of critical values of the exponent n at which the qualitative behaviour of solutions changes significantly. One of the most important of these is $n = 3$, since it turns out that for $n \geqslant 3$ no solutions exist in which the droplet expands. This result relates to well-known difficulties associated with how a point at which the free surface $y = h$ meets the solid (known as a contact line) moves; the appropriate formulation of the physical behaviour near to such a point remains a topic of debate.

(ii) For $n < \frac{3}{2}$, (4.4) has solutions with $h \sim B(-\zeta)^2$ as $\zeta \to 0^-$ for constant B. Like (4.5), this expression has zero contact angle for $n < 3$ (i.e. $\mathrm{d}h/\mathrm{d}\zeta = 0$ at $\zeta = 0$), but the non-uniqueness property can be illustrated by noting the existence of other solutions with finite contact angle, for which

$$h \sim A(-\zeta) + B(-\zeta)^2 - \frac{q}{(4-n)(3-n)(2-n)A^{n-1}}(-\zeta)^{4-n}, \quad \text{as } \zeta \to 0^-, \tag{4.6}$$

for arbitrary constants A and B, a minor modification of (4.6) being needed if $n = 2$. The local behaviour indicated by (4.6) generalizes from travelling waves to other solutions (with A, B and q becoming time dependent) and leads to an infinite family of possible solutions; this is in sharp contrast to the second-order case (4.2 *a*) for which the local behaviour (4.3) specifies the solution uniquely. Uniqueness for (4.2 *b*) can, however, be recovered if, for example, the contact angle $\partial h/\partial x$ or the pressure $-\partial^2 h/\partial x^2$ is specified at the contact line; this is again a reflection of singular behaviour of such a point from the physical point of view.

It is hoped that the above remarks illustrate some of the difficulties associated with high-order systems, even those as apparently straightforward as (4.2 *b*), together with the scope for further developments. Equation (4.2 *c*), which is relevant in semiconductor fabrication and to surface crusting on spreading melts, for example, leads to an additional set of extra complications. The area is likely to remain one of intense study for many years to come.

We conclude this section by noting an example of a related fourth-order system, namely the Cahn–Hilliard equation used in the study of phase separation in alloys and of interest partly for its pattern-forming properties. For a degenerate mobility version of this equation, of the form

$$\frac{\partial u}{\partial t} = \nabla \cdot ((1-u^2)\nabla(-\gamma\nabla^2 u + \quad (u)))$$

(u being the difference in the mass fraction of the two components of the alloy), it has recently been shown by Cahn *et al*. (1996) that, in a limit case, the velocity at which the contour $u = 0$ propagates in the normal direction is proportional to the Laplacian of its curvature. We shall use a related interfacial dynamics law as an illustration in the next section (see equation (5.8)).

5. Moving-boundary problems

(*a*) *Interfacial dynamics*

In many physical applications, the appropriate form of the mathematical model contains an interface whose evolution must be determined in the course of solving the problem. The coupling between the location of this moving boundary and the other unknowns leads to nonlinearities that cause significant complications. Moving-boundary problems originally arose in the modelling of melting and freezing, in which the interface separates the solid and fluid phases; they are now very widespread, occurring in, for example, corrosion, oil extraction, option pricing and, as already implied, tumour growth and free surface flows (see, for example, Elliott & Ockendon 1982). The simplest class of such problems, an example of which we shall use to illustrate a variety of relevant phenomena, is that in which the moving-boundary location is the only unknown, being governed by an evolution equation that involves quantities evaluated only on the moving boundary itself. Such models are widely used to mimic crystal growth, for example; moreover, they can be derived systematically as limit cases of reaction–diffusion problems and (as already indicated) of models of phase separation.

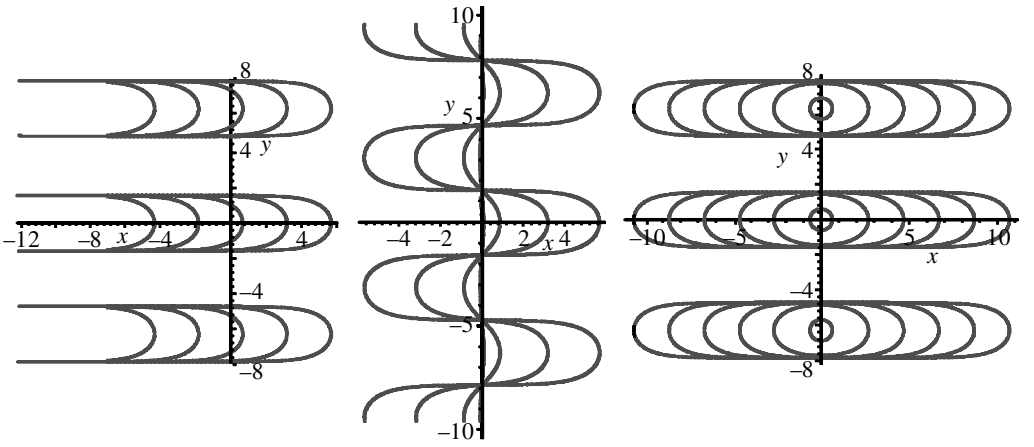

Figure 6. Finger development in an ill-posed moving-boundary problem. (*a*) Steadily propagating fingers, (5.5); (*b*) the growth of fingers from an almost planar interface, (5.6); and (*c*) the growth of 'bubbles', (5.7).

One of the simplest models of the class in question takes the form

$$v_{\mathrm{n}} = \kappa, \tag{5.1}$$

where v_{n} is the normal velocity and κ the mean curvature of the moving boundary; we consider curves in two dimensions, each point of which thus moves in the perpendicular direction with speed given by the local value of the curvature.

Equation (5.1) is well-posed, tending, in particular, to smooth out protuberances on a curve. However, in many applications, ill-posed models are of more interest, leading to phenomena such as fingering and dendrite growth. The time reversal of (5.1), namely $v_{\mathrm{n}} = -\kappa$, is such an ill-posed model. Writing the curve in the form $y = f(x,t)$ yields

$$\frac{\partial f}{\partial t} = -\frac{1}{(1+(\partial f/\partial x)^2)}\frac{\partial^2 f}{\partial x^2}, \tag{5.2}$$

which is one of the nonlinear diffusion equations identified by Akhatov *et al.* (1987) and Bluman *et al.* (1988) as special cases from the symmetry point of view; such symmetry approaches are now widely used in the analysis of nonlinear differential equations, much progress having recently been made in the area. The results of the symmetry analysis suggest re-expressing (5.2) in terms of the variables

$$z = x + \mathrm{i}f, \qquad g(z,t) = x - \mathrm{i}f, \tag{5.3}$$

which enables (5.2) to be written in its simplest form, namely

$$\frac{\partial g}{\partial t} = -\frac{1}{(\partial g/\partial z)}\frac{\partial^2 g}{\partial z^2}. \tag{5.4}$$

That this formulation arises from symmetry considerations is a noteworthy result because (5.3) expresses the moving boundary in the form $\bar{z} = g(z,t)$ (where $z = x+\mathrm{i}y$, $\bar{z} = x-\mathrm{i}y$ are complex conjugates); such an approach is known for other reasons to be

of value in many moving-boundary problems, g being termed the Schwarz function (see Howison 1992). New solutions to (5.4) were constructed in King (1993); from these, we have—in addition to the well-known travelling wave solution shown in figure 6*a* and given by

$$\cos(qf) = \tfrac{1}{2}\mathrm{e}^{q(x-qt)}, \qquad g - qt = -\ln(1 - \mathrm{e}^{-q(z-qt)})/q, \tag{5.5}$$

where q is a constant—two generalizations, namely

$$\cos(qf) = \mathrm{e}^{-q^2t}\sinh(qx), \qquad \tanh(\tfrac{1}{2}qg) = \frac{1 - \mathrm{i}\mathrm{e}^{-q^2t}}{1 + \mathrm{i}\mathrm{e}^{-q^2t}}\tanh(\tfrac{1}{2}qz), \tag{5.6}$$

shown in figure 6*b*, which illustrates how an originally nearly planar interface leads to fingering, with the fingers eventually being of the form (5.5), and

$$\cos(qf) = \mathrm{e}^{-q^2t}\cosh(qx), \qquad \tanh(\tfrac{1}{2}qg)\tanh(\tfrac{1}{2}qz) = \tanh(\tfrac{1}{2}q^2t), \tag{5.7}$$

corresponding to the 'bubbles' illustrated in figure 6*c* and again evolving for large time to (5.5) and its mirror image. Analogous solutions to the more complicated Hele–Shaw moving-boundary problem have been widely studied.

Because (5.2) is ill-posed, for most initial conditions a solution will cease to exist in finite time. The usual way to remedy this is to regularize the problem by introducing an additional stabilizing term (surface tension provides a commonly adopted physical regularization); a typical such generalization is (see, for example, Brower *et al.* 1984)

$$v_\mathrm{n} = -\kappa - \varepsilon^2\frac{\partial^2\kappa}{\partial s^2}, \tag{5.8}$$

where ε is a small constant and s is the arc length along the curve. This again takes a particularly convenient form when written in terms of the Schwarz function, namely

$$\frac{\partial g}{\partial t} = -\frac{1}{(\partial g/\partial z)}\frac{\partial^2 g}{\partial z^2} + 2\varepsilon^2\frac{\partial}{\partial z}\left(\frac{1}{(\partial g/\partial z)^{1/2}}\frac{\partial^2}{\partial z^2}\left(\frac{1}{(\partial g/\partial z)^{1/2}}\right)\right).$$

It is natural to speculate whether the finger solution (5.5) to the unregularized problem plays any role in the large-time behaviour of (5.8); we expand on this below, similar issues arising in many pattern-forming moving-boundary problems. We note that the Schwarz function, $g(z,t)$, has singularities in the complex plane at $z = qt + 2n\pi\mathrm{i}/q$ for (5.5), $z = qt + 2n\pi\mathrm{i}/q$ and $z = -qt + (2n+1)\pi\mathrm{i}/q$ for (5.6), and at $z = \pm qt + 2n\pi\mathrm{i}/q$ for (5.7), where n is an arbitrary integer. Such apparently non-physical singularities (in the sense that they are not present in the physical domain, which, in the current case, is simply the curve $y = f(x,t)$) are widely believed to play an important role in determining the physical behaviour. The simple explicit solutions above give some indication of how such singularities can evolve. In particular, for (5.6) we have a singularity at $(x,y) = (qt,0)$, while on the curve $y = f$ we have $y = 0$ at $x = \sinh^{-1}(\mathrm{e}^{q^2t})/q$; the latter satisfies $x \sim \mathrm{e}^{q^2t/q}$ as $t \to -\infty$ and $x \sim qt + \ln 2$ as $t \to +\infty$. The singularities of the Schwarz function thus move closer to the curve as t increases (this being a common feature of ill-posed problems), approaching a spacing of $\ln 2$ as $t \to +\infty$. For initial data for which the solution to (5.4) ceases to exist in finite time, one or more singularities will impinge on the curve at that finite time; such behaviour is generic for (5.4) because of its ill-posedness and it would be of interest to know the full class of solutions for which it does not occur, this class obviously including (5.5)–(5.7).

(*b*) *Asymptotics beyond all orders*

Asymptotic problems are those involving a small parameter (ε in the case of (5.8), with $0 < \varepsilon \ll 1$) that can be exploited to simplify their analysis. The final issue we wish to illustrate involves an area of asymptotics in which there has recently been dramatic progress, with important further developments likely in the near future. The methods in question emphasize the crucial role played by complex-plane singularities in dictating the real behaviour.

We have already alluded to the possible role of (5.5) in governing the large-time behaviour of (5.8) for small ε. An issue that arises immediately is the selection of the wavespeed q, which is arbitrary in (5.5). As with the Hele–Shaw problem, a natural proposal is that it is the regularization that provides the physical mechanism responsible for selecting the speed, even though the regularizing term is small. In a similar manner to Kruskal & Segur (1991), we formulate the travelling wave problem for (5.5) in terms of $\theta(s)$, the angle the moving boundary makes with the vertical. For fingers propagating steadily in the x-direction at speed q, we then have $v_{\rm n} = q\cos\theta$, $\kappa = -\mathrm{d}\theta/\mathrm{d}s$, so that (5.1) becomes

$$\varepsilon^2 \frac{\mathrm{d}^3\theta}{\mathrm{d}s^3} + \frac{\mathrm{d}\theta}{\mathrm{d}s} = q\cos\theta;$$

suitably rescaling s and redefining ε (now with $0 \leqslant \varepsilon \leqslant 1$), this is equivalent to

$$\varepsilon^2 \frac{\mathrm{d}^3\theta}{\mathrm{d}s^3} + (1-\varepsilon^2)\frac{\mathrm{d}\theta}{\mathrm{d}s} = \cos\theta, \tag{5.9}$$

which will be the most convenient form for us to work with. We should like to solve (5.9) subject to

$$\theta \to \pm\tfrac{1}{2}\pi, \quad \text{as } s \to \pm\infty, \tag{5.10}$$

to give fingers of the type shown in figure 6*a* for $\varepsilon = 0$. In fact, no such solutions exist for $\varepsilon > 0$, and our goal is to illustrate some of the subtleties involved in the behaviour for ε small but non-zero.

Imposing the boundary condition as $s \to +\infty$, we find that

$$\theta \sim \tfrac{1}{2}\pi - A\mathrm{e}^{-s}, \quad \text{as } s \to +\infty, \tag{5.11}$$

the arbitrary constant A being the only degree of freedom in this limit. This implies that (5.10) represents four boundary conditions in all, making the problem overspecified (a difficulty that is compounded by the translation invariance of (5.9)–(5.10), which enables us to specify the value of A to give an initial-value problem); this suggests why there is no solution that satisfies (5.11) for $\varepsilon > 0$. The periodicity property $\theta(s + \pi\mathrm{i}) = \pi - \theta(s)$ follows from (5.9) and (5.11).

For small ε, it is natural to expand the solution to (5.9) in the form

$$\theta \sim \sum_{n=0}^{\infty} \varepsilon^{2n}\theta_n(s), \tag{5.12}$$

giving, in particular,

$$\frac{\mathrm{d}\theta_0}{\mathrm{d}s} = \cos\theta_0, \qquad \frac{\mathrm{d}\theta_1}{\mathrm{d}s} = -\theta_1\sin\theta_0 + \frac{\mathrm{d}\theta_0}{\mathrm{d}s} - \frac{\mathrm{d}^3\theta_0}{\mathrm{d}s^3},$$

and, hence, $\theta_0 = \sin^{-1}(\tanh s)$, $\theta_1 = 2\,\mathrm{sech}\, s \tanh s$; since

$$\theta_n = \sum_{k=1}^{n} a_{n,k}\, \mathrm{sech}^{2k-1} s \tanh s, \quad n \geqslant 1, \tag{5.13}$$

for constant $a_{n,k}$, it is straightforward to calculate θ_n up to any order; it is easily seen that a solution that satisfies (5.10) exists up to any power of ε, so the expansion (5.12) apparently fails to reproduce the expected non-existence result. However, two features of the asymptotic solution indicate that such appearances may be misleading. The first is that the series (5.12) is divergent; care must therefore be exercised in giving it meaning with the tail of the expansion (the behaviour of θ_n for large n) containing the information we require (see Chapman *et al.* 1998). Secondly, it is clear from (5.13) that θ_n possesses complex-plane singularities at $s = \frac{1}{2}(2n+1)\pi\mathrm{i}$ (reflecting those of the Schwarz function mentioned above) and that these singularities become more severe as n increases; this indicates the need for a separate analysis in the complex plane close to these singularities (see Kruskal & Segur 1991). Stokes lines are lines in the complex plane across which an exponentially small quantity is switched on, this quantity being of size $\mathrm{e}^{-\pi/2\varepsilon}/\varepsilon^{1/2}$ here ('beyond all orders' of the algebraic expansion (5.12)). They are initiated at the complex plane singularities and, in the current example, run along the imaginary axis. Such Stokes lines, hidden beyond all orders of the asymptotic expansion, can conveniently be made manifest by optimal truncation (Chapman *et al.* 1998). The exponentially small quantity that is turned on is of the form

$$\Theta \sim \lambda \mathrm{e}^{-\pi/2\varepsilon} \cosh^{1/2} s \cos(s/\varepsilon)/\varepsilon^{1/2}, \tag{5.14}$$

where $\lambda \approx 8.48$, this being an asymptotic solution of the homogeneous remainder equation

$$\varepsilon^2 \frac{\mathrm{d}^3\Theta}{\mathrm{d}s^3} + (1-\varepsilon^2)\frac{\mathrm{d}\Theta}{\mathrm{d}s} = -\Theta \sin\theta_0.$$

The solution θ that satisfies (5.11) contains the exponentially small term Θ to the left of the Stokes line (i.e. in $\mathrm{Im}\, s < 0$) but not to the right. A crucial point is that the Stokes line crosses the real axis and, thus, influences the solution on the real line, the term (5.14) causing the boundary condition as $s \to -\infty$ to be violated, thus implying the non-existence of a travelling wave.

Since no steadily propagating fingers exist, the question of how (5.8) ultimately evolves remains unanswered; the application of the methods of asymptotics beyond all orders to such time-dependent problems remains in its infancy. One possibility is that fingers do develop but that they throw off sidebranches of the type observed in dendritic (snowflake-like) growth; these may, in turn, produce their own sidebranches, leading to the possibility of very complicated (ultimately fractal) morphologies. Such issues indicate the scope for further developments in the area.

6. Discussion

The examples given above have enabled us to illustrate some fast-developing fields in mathematical modelling and to touch upon a variety of powerful techniques that are the subject of current advances, such as symmetry methods, asymptotics beyond

all orders, rigorous analysis, and complex-variable methods. They also suggest areas that will be the subject of intensive study in the future.

One such area was hinted at by the example of § 2, namely the link between microscopic (cellular, molecular, etc.) and macroscopic behaviour. Understanding the interactions between widely different scales remains a great challenge, which will grow in significance as, among other issues, semiconductor devices shrink and comprehensive understanding of biological systems is sought. Further relevant examples are provided by the key role played by vortices in determining the large-scale behaviour of fluids and superconductors and of dislocations in determining that of crystals. One way of viewing vortices and dislocations is as line singularities in partial differential equations (cf. Chapman *et al.* 1997), illustrating another recurring theme; local singularities also play a central role in the study of contact line behaviour (cf. § 4), while complex plane singularities control the evolution of many nonlinear systems (cf. § 5).

The coupling between material deformation and reaction–diffusion effects is another topic that deserves much greater study, given its significance not only in biology but also in many other disciplines, examples being provided by the complicated interactions between stress and diffusion that can arise in materials modelling. One of the difficulties in mathematical biology is the huge range of coupled effects that may be present, so a skill required for successful modelling is the ability to formulate tractable models that make possible the investigation of which effects are of most significance. The scope for developments in mathematical biology and medicine is particularly large, the hope and expectation being that ideas of significant clinical value will result.

Despite the growing power of computers, analytical methods will continue to play a central role in mathematical modelling. Simple explicit solutions such as those in § 4 provide very valuable insight, singularity development cannot be investigated solely by numerical means and asymptotic methods can cope very successfully with problems that are extremely demanding (or indeed intractable) numerically. Computer algebra provides an effective tool for making analytical approaches viable in a wider range of problems. Rigorous analysis is also likely to become more central to modelling studies as its power grows and as the increasing complexity (or smaller scale) of the phenomena being studied makes physical intuition less reliable. One role of a particular tool of analysis, namely regularization, has been outlined in § 5. It is of much more widespread value, however: it can, for example, be used to address the issues of non-uniqueness mentioned in § 4. High-order systems will remain a focus of analysis for a considerable period; some of the reasons were noted in § 4, and the significance of such systems is also implicit in § 2, the 'bird's beak' profiles of figure 2*a* arising naturally for certain reaction–diffusion systems but not from single equations, for which the profiles of figure 2*b* are more representative.

Mathematical modelling is a vast (and fast-growing) field; the above selection of topics reflects the author's own interests and is far from comprehensive (as, of necessity, is the list of references). There are numerous other areas in which mathematical modelling is growing in importance; to illustrate the extent of these, we mention applications involving materials modelling (smart materials, etc.), food and agricultural sciences (in quantifying issues of food safety, for example), chemistry, quantum devices and optics. The scope of such applications requires a wide variety

of models; to offset the emphasis on diffusive systems in §§ 2–5, the very widespread importance of other classes of partial differential equations (namely, hyperbolic and elliptic) should be mentioned, as should that class of models involving quite different mathematical formulations (integral equations, discrete systems, stochastic models and so on). Each such description presents its own mathematical challenges, as recent progress in the understanding of nonlinear and hyperasymptotic phenomena, in particular, confirms. 'Black box' numerical packages, such as computational fluid dynamics codes, are now widely used, particularly in industrial and engineering contexts. While they clearly have an important role to play, uncritical reliance upon them accordingly has its dangers.

Successful mathematical modelling is very much a multidisciplinary endeavour. While we have chosen to concentrate here on the mathematical demands that it makes, the importance to its health and vitality of genuine collaborations with scientists and engineers across a broad spectrum of disciplines cannot be overstated. Providing mathematicians rise to this challenge, the subject has an exceptionally bright future.

I am very grateful to Mark Bowen, Martin Meere and John Ward for providing the figures, and to the Leverhulme Trust for financial support.

References

Akhatov, I. Sh., Gazizov, R. K. & Ibragimov, N. Kh. 1987 Group classification of nonlinear filtration equations. *Sov. Math. Dokl.* **35**, 384–386.

Anderson, A. R. A. & Chaplain, M. A. J. 1998 Continuous and discrete models of tumor-induced angiogenesis. *Bull. Math. Biol.* **60**, 857–900.

Bernis, F. 1995 Viscous flows, fourth order nonlinear degenerate parabolic equations and singular elliptic problems. In *Free boundary problems: theory and applications* (ed. J. I. Diaz, M. A. Herrero, A. Liñan & J. L. Vazquez). London: Pitman.

Bluman, G. W., Reid, G. J. & Kumei, S. 1988 New classes of symmetries for partial differential equations. *J. Math. Phys.* **29**, 806–811.

Boatto, S., Kadanoff, L. P. & Olla, P. 1993 Traveling-wave solutions to thin-film equations. *Phys. Rev.* E **48**, 4423–4431.

Brower, R. C., Kessler, D. A., Koplik, J. & Levine, H. 1984 Geometrical models for interface evolution. *Phys. Rev.* A **29**, 1335–1342.

Byrne, H. M. & Chaplain, M. A. J. 1997 Free boundary value problems associated with the growth and development of multicellular spheroids. *Eur. J. Appl. Math.* **8**, 639–658.

Cahn, J. W., Elliott, C. M. & Novick-Cohen, A. 1996 The Cahn–Hilliard equation with a concentration dependent mobility: motion by minus the Laplacian of the mean curvature. *Eur. J. Appl. Math.* **7**, 287–301.

Chapman, S. J., Elliott, C. M., Head, A. K., Howison, S. D., Leslie, F. M. & Ockendon, J. R. (eds) 1997 Vortices, dislocations and line singularities in partial differential equations. Preface. *Phil. Trans. R. Soc. Lond.* A **355**, 1947.

Chapman, S. J., King, J. R. & Adams, K. L. 1998 Exponential asymptotics and Stokes lines in nonlinear ordinary differential equations. *Proc. R. Soc. Lond.* A **454**, 2733–2755.

Elliott, C. M. & Ockendon, J. R. 1982 *Weak and variational methods for moving boundary problems*. London: Pitman.

Greenspan, H. P. 1972 Models for the growth of a solid tumour by diffusion. *Stud. Appl. Math.* **51**, 317–340.

Howison, S. D. 1992 Complex variable methods in Hele–Shaw moving boundary problems. *Eur. J. Appl. Math.* **3**, 209–224.

Kalashnikov, A. S. 1987 Some problems of the qualitative theory of nonlinear degenerate 2nd order parabolic equations. *Russ. Math. Surveys* **42**, 169–222.

King, J. R. 1993 Exact polynomial solutions to some nonlinear diffusion equations. *Physica* D **64**, 35–65.

King, J. R., Sharp, T. E., Tuck, B. & Rogers, T. G. 1995 Mathematical modelling of the interstitialcy diffusion mechanism. *Proc. R. Soc. Lond.* A **450**, 623–649.

Kruskal, M. D. & Segur, H. 1991 Asymptotics beyond all orders in a model of crystal growth. *Stud. Appl. Math.* **85**, 129–181.

Meere, M. G., King, J. R. & Rogers, T. G. 1995 Asymptotic analysis of some two-dimensional diffusion problems. *Proc. R. Soc. Lond.* A **448**, 213–236.

Perumpanani, A. J., Sherratt, J. A. & Norbury, J. 1997 Mathematical modelling of capsule formation and multinodularity in benign tumour growth. *Nonlinearity* **10**, 1599–1614.

Tayler, A. B. 1986 *Mathematical models in applied mechanics.* Oxford: Clarendon.

Tuck, B. 1988 *Atomic diffusion in III–V semiconductors.* Bristol: Adam Hilger.

Turing, A. M. 1952 The chemical basis of morphogenesis. *Phil. Trans. R. Soc. Lond.* B **237**, 37–79.

Ward, J. P. & King, J. R. 1997 Mathematical modelling of avascular-tumour growth. *IMA J. Math. Appl. Med. Biol.* **14**, 39–69.

Ward, J. P. & King, J. R. 1999 Mathematical modelling of avascular-tumour growth. II. Modelling growth saturation. *IMA J. Math. Appl. Med. Biol.* **16**, 171–211.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

AUTHOR PROFILE

J. R. King

John King graduated in mathematics from the University of Cambridge and undertook postgraduate research in industrial applied mathematics, supervised by Alan Tayler, at the University of Oxford. He then spent two years as a Postdoctoral Research Associate in the Department of Mathematical Sciences at Rensselaer Polytechnic Institute, working for Julian Cole. He returned to England to take up a Lectureship in Theoretical Mechanics at the University of Nottingham, where he was subsequently promoted to a Readership in Nonlinear Systems and then appointed to his current post, the Chair in Theoretical Mechanics previously held by Tryfan Rogers. Now aged 38, his research interests range across a number of areas of mathematical modelling.