

# Techniques of applied mathematics

A. C. Fowler  
Mathematical Institute, Oxford University

Version 2.0

October 11, 2005

# Contents

Preface . . . . .	iv
<b>1 Introduction</b>	<b>1</b>
1.1 Conservation and constitutive laws . . . . .	3
1.2 The law of mass action . . . . .	4
1.3 The $S-I-R$ model . . . . .	5
1.4 Saturation and cooperativity . . . . .	7
1.4.1 Cooperativity . . . . .	7
1.5 Notes and references . . . . .	9
1.5.1 $SIR$ model . . . . .	9
1.5.2 Spruce budworm model . . . . .	9
Exercises . . . . .	9
<b>2 Non-dimensionalisation and approximation</b>	<b>12</b>
2.1 Non-dimensionalisation . . . . .	12
2.1.1 The wave equation . . . . .	13
2.1.2 The heat equation . . . . .	14
2.2 Scaling . . . . .	16
2.2.1 The wave equation, again . . . . .	16
2.3 Simple approximation methods . . . . .	17
2.4 Perturbation methods . . . . .	22
2.5 Notes and references . . . . .	23
Exercises . . . . .	23
<b>3 Perturbation methods</b>	<b>27</b>
3.1 Regular perturbation theory . . . . .	28
3.2 Singular perturbation theory: boundary layers . . . . .	29
3.2.1 Internal layers . . . . .	33
3.3 Multiple scales and averaging . . . . .	34
3.3.1 Poincaré-Lindstedt method . . . . .	36
3.3.2 Van der Pol oscillator . . . . .	37
3.3.3 Averaging . . . . .	39
3.4 Notes and references . . . . .	40
Exercises . . . . .	40

<b>4</b>	<b>Stability and oscillations</b>	<b>43</b>
4.1	Linear stability . . . . .	43
4.2	Linear systems . . . . .	46
4.2.1	The fundamental solution . . . . .	47
4.2.2	Floquet theory . . . . .	47
4.2.3	Mathieu equation . . . . .	49
4.3	Nonlinear stability . . . . .	49
4.4	Phase plane analysis . . . . .	49
4.4.1	An example of Murray . . . . .	51
4.4.2	Travelling waves on the Martian north polar ice cap . . . . .	55
4.5	Relaxation oscillations . . . . .	56
4.5.1	The Van der Pol oscillator . . . . .	58
4.6	Belousov-Zhabotinskii reaction . . . . .	60
4.7	Notes and references . . . . .	60
	Exercises . . . . .	60
<b>5</b>	<b>Oscillation and comparison methods</b>	<b>64</b>
5.1	Derivation of Sturm-Liouville systems . . . . .	64
5.1.1	Waves in a bucket . . . . .	65
5.1.2	Laplace's equation . . . . .	66
5.1.3	Instability in thermally activated shear flow . . . . .	67
5.2	Sturm-Liouville theory . . . . .	70
5.2.1	Eigenvalues and eigenfunctions . . . . .	71
5.2.2	Singular Sturm-Liouville systems . . . . .	75
5.3	Comparison methods . . . . .	75
5.4	Notes and references . . . . .	76
	Exercises . . . . .	77
<b>6</b>	<b>Integral equations and eigenfunctions</b>	<b>79</b>
6.1	Integral equations . . . . .	79
6.1.1	Volterra integral equations . . . . .	79
6.1.2	Fredholm integral equations . . . . .	81
6.1.3	The delta function . . . . .	82
6.1.4	The Fredholm alternative . . . . .	82
6.2	Eigenfunction expansions . . . . .	83
6.2.1	Neumann series . . . . .	83
6.2.2	Inner products and norms . . . . .	84
6.2.3	Spectral theory . . . . .	86
6.3	Variational methods . . . . .	86
6.3.1	A variational principle . . . . .	86
6.3.2	Fourier series . . . . .	88
6.3.3	The Fredholm alternative revisited . . . . .	89
6.3.4	Rayleigh-Ritz method . . . . .	89
6.4	Notes and references . . . . .	89
	Exercises . . . . .	89

<b>7</b>	<b>Waves and shocks</b>	<b>92</b>
7.1	Traffic flow . . . . .	99
7.2	Burgers' equation . . . . .	100
7.3	The Fisher equation . . . . .	103
7.3.1	Stability . . . . .	104
7.4	Solitons . . . . .	104
7.5	Snow melting . . . . .	105
7.6	Notes and references . . . . .	109
	Exercises . . . . .	109
<b>8</b>	<b>Similarity solutions</b>	<b>112</b>
8.1	The heat equation . . . . .	112
8.1.1	Error function profile . . . . .	115
8.2	The porous medium equation . . . . .	116
8.3	Snow melting revisited . . . . .	118
8.4	Notes and references . . . . .	119
	Exercises . . . . .	119
<b>9</b>	<b>Nonlinear diffusion</b>	<b>122</b>
9.1	The viscous droplet . . . . .	122
9.2	Advance and retreat: waiting times . . . . .	124
9.3	Blow-up . . . . .	125
9.4	Notes and references . . . . .	131
	Exercises . . . . .	131
	<b>References</b>	<b>134</b>

# Preface

These notes accompany the B5 course, ‘Techniques of applied mathematics’. There are some parts which are not written yet, you will notice these by the sections with no content! Modifications and additions will be made occasionally, and the most up to date version will be kept on the web at

`http://www.maths.ox.ac.uk/~fowler/courses/tech/tech.html`

Edition numbers  $a.b$  mean the following:  $a$  refers to the edition available at the Mathematical Institute reception, and  $b$  refers to the web-based updates. Institute hard copies will thus always be 1.0, 2.0, etc. Amendments in updated versions will be indicated *here*, in this preface, but these lists of amendments will disappear at each new hardcopy edition.

My thanks to John Norbury for a rapid critical reading of the rapidly written chapters 5 and 6.

A. C. Fowler  
October 11, 2005

## Version number

2.0 Michaelmas Term 2005 release.

# Chapter 1

## Introduction

Mathematical modelling is that abstruse subject which forms the connecting tissue between the problems of the real world which we wish to solve, and the quantitative analysis which we undertake to do so. Almost any problem which requires a quantitative answer, whether it be in industry, medicine, economics, biology or geophysics (for example) involves the formulation of the problem as a *mathematical model*, and it is this formulation, and the techniques which one uses to solve the model, which form the subject of these notes.

There are, perhaps, three kinds of model: statistical, discrete and continuous. We can illustrate their difference with a simple (and topical) example. Suppose fox hunting in England were to be banned. One of the arguments against this is that hunting provides a control on the number of foxes. As with many such assertions, this is one which might be true, or might not. It is a quantitative assertion, and the only way to find out whether it is true is to examine it scientifically. We do this by proposing a model, and examining its validity<sup>1</sup>. The statistical method examines the probability that one of two alternative hypotheses is true, by using actual measured data. For example, if hunting were to be banned, one might examine fox population numbers during the five years preceding and following such a ban. As with all models, it is in the interpretation of the results that the skill lies. This is particularly true of statistics, where the aim is to eliminate possibilities, rather than propose constituent mechanisms.

Statistical models have to deal with the issue of predictability. Suppose fox numbers double following a ban; then it would seem likely that there is a connection. But there are other factors which are involved in determining fox populations, and it is always arguable whether one particular factor has a deciding control.

Statistical models are *diagnostic*: they try to interpret process from measured data. Discrete and continuous models are examples of *prognostic* models: they propose a descriptive model of a phenomenon, and then predict what will happen in the future. Such models require validation. This consists initially of matching observation to theory, and often will suggest experiments which can be done to confirm the theory.

A discrete model will propose a difference equation for the variables of interest.

---

<sup>1</sup>This is in fact the scientific method: the models that we propose are no more than hypotheses, and science should never present itself as purveying absolute truth.

In the case of foxes, this might be an estimate for the fox population density  $u_n$  at a particular time during the  $n$ -th year, and a simple such discrete model is the logistic model with harvesting:

$$u_{n+1} = ru_n \left(1 - \frac{u_n}{K}\right) - hu_n, \quad (1.1)$$

where  $r$  is the specific reproductive rate, and the term  $hu_n$  represents harvesting (via hunting); the coefficient  $h$  represents the effort involved, which one might suppose would be proportional to the number of hunts. This model contains the simple idea that excess populations become limited by competition for resources (the nonlinear term involving  $K$  implies decreasing growth rate at larger values of  $u_n$ ). A discrete model such as (1.1) might be appropriate for fox populations, which have an annual rut, so that the reproductive cycle is repeated annually.

Suppose, for the sake of argument, that (1.1) is a reasonable model. Note that  $r > 1$  is a pure number, and so also is  $h$ , while  $K$  has the same units as  $u_n$ . In the absence of hunting, there is a steady population  $K(r - 1)/r$ , and this is reduced if  $h \neq 0$  by a factor  $\left[1 - \frac{h}{r - 1}\right]$ . So the efficacy of hunting in this model depends on this particular quantity; hunting is effective if  $h/(r - 1)$  is significant. This provides a predicted outcome, provided the parameters  $h$  and  $r$  can be reasonably assessed.

Continuous models are used in the same way, but describe processes using differential equations. It is this kind of model which forms the focus of these notes. The rationale for continuous models is the *continuum hypothesis*, which states that the actual behaviour of a discrete variable, such as a population, can be accurately represented by the evolution of a continuous (and usually differentiable) variable. The basis for this assumption is the ‘fine-grained’ nature of the variable: in a large population, a change by one individual is a small relative change, and can be viewed as being a finite difference approximation to the infinitesimal changes of the calculus.

A continuous version of (1.1) might be the ordinary differential equation

$$\frac{du}{dt} = \rho u \left(1 - \frac{u}{C}\right) - \mu u, \quad (1.2)$$

whose behaviour can be analysed in a similar way to its discrete counterpart.

There are two particular points of view which we can bring to bear on the mathematical models which describe the phenomena which concern us: these are the dynamical systems approach, or equivalently the bifurcation theory approach; and the perturbation theory approach. Each has its place in different contexts, and sometimes they overlap.

The bifurcation theory approach most usually (but not always) is brought to bear on problems which have some kind of complicated time-dependent behaviour. The idea is that we seek to understand the observations through the understanding of a number of simpler problems, which arise successively through bifurcations in the mathematical model, as some critical parameter is changed. A classic example of this approach is in the study of the origin of chaos in the Lorenz equations, or the onset of complicated forms of thermal convection in fluids.

In its simplest form (e.g., in weakly nonlinear stability theory) the perturbative approach is similar in method to the bifurcational one; however, the ethos is rather different. Rather than try and approach the desired solution behaviour through a sequence of simpler behaviours, we try and break down the solution by making approximations, which (with luck) are in fact realistic. In real problems, such approximations are readily available, and part of the art of the applied mathematician is having the facility of being able to judge how to make the right approximations. In these notes, we follow the perturbative approach. It has the disadvantage of being harder, but it is able to get closer to a description of how realistic systems may actually behave.

## 1.1 Conservation and constitutive laws

The basic building blocks of continuous mathematical models are conservation laws. The continuum assumption adopts the view that the physical medium of concern may be considered continuous, whether it be a porous medium (for example, sand on a beach) or a fluid flow. The continuum hypothesis works whenever the length or time scales of interest are (much) larger than the corresponding microscale. For example, the formation of dunes in a desert (length scale hundreds of metres) can be modelled as a continuous process, since the microscale (sand grain size) is much smaller. Equally, the modelling of large animal populations or of snow avalanches treats the corresponding media as continuous.

Conservation laws arise as mathematical equations which represent the idea that certain quantities are conserved — for example, mass, momentum (via Newton’s law) and energy. More generally, a conservation law refers to an equation which relates the increase or decrease of a quantity to terms representing supply or destruction.

In a continuous medium, the typical form of a conservation law is as follows:

$$\frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{f} = S. \quad (1.3)$$

In this equation,  $\phi$  is the quantity being ‘conserved’ (expressed as amount per unit volume of medium, i.e. as a density;  $\mathbf{f}$  is the ‘flux’, representing transport of  $\phi$  within the medium, and  $S$  represents source ( $S > 0$ ) or sink ( $S < 0$ ) terms. Derivation of the point form (1.3) follows from the integral statement

$$\frac{d}{dt} \int_V \phi dV = - \int_{\partial V} \mathbf{f} \cdot \mathbf{n} dS + \int_V S dV, \quad (1.4)$$

after application of the divergence theorem (which requires  $\mathbf{f}$  to be continuously differentiable), and by then equating integrands, on the basis that they are continuous and  $V$  is arbitrary. Derivation of (1.3) thus requires  $\phi$  and  $\mathbf{f}$  to be continuously differentiable, and  $S$  to be continuous.

Two basic types of transport are advection (the medium moves at velocity  $\mathbf{u}$ , so there is an advective flux  $\phi \mathbf{u}$ ) and diffusion, or other gradient-driven transport (such as chemotaxis). One can thus write

$$\mathbf{f} = \phi \mathbf{u} + \mathbf{J}, \quad (1.5)$$



where  $\mathbf{J}$  might represent diffusive transport, for example. The very simplest conservation law is that of conservation of mass, where the conserved quantity is the density  $\rho$ , and the mass flux is entirely due to advection:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0. \quad (1.6)$$

Invariably, conservation laws give more terms than equations. In (1.5), for example, we have one scalar equation for  $\phi$ , but other quantities  $\mathbf{J}$  and  $\mathbf{S}$  are present as well, and equations for these must be provided. Typically, these take the form of constitutive laws, and are based squarely on experimental measurement. For example, diffusive transport is represented by the assumption

$$\mathbf{J} = -D \nabla \phi, \quad (1.7)$$

where  $D$  is a diffusion coefficient. In the heat equation, this is known as Fourier's law, and the heat equation itself takes the familiar form

$$\frac{\partial}{\partial t}(\rho c_p T) + \nabla \cdot [\rho c_p T \mathbf{u}] = \nabla \cdot [k \nabla T] + Q, \quad (1.8)$$

where  $Q$  represents any internal heat source or sink.

## 1.2 The law of mass action

Apart from conservation laws, which typically involve transport terms in the form of divergences, the other type of term which frequently arises in models is an algebraic 'reaction' term. The source term  $Q$  in (1.8) is a simple example of this. Slightly more complicated is the Newtonian cooling law

$$\frac{dT}{dt} = h(T_0 - T), \quad (1.9)$$

which represents cooling (or warming) of a material of temperature  $T$  in surroundings of ambient temperature  $T_0$ . Here,  $hT_0$  represents heat gain from the surroundings, and  $hT$  is the corresponding heat loss.

A particularly common and interesting way in which source terms arise is in the nonlinear interaction between different constituent populations. These may represent individuals of a plant or animal species, or they may represent different types of cells in the body, or concentrations of different reacting chemical species. The basis for describing such interactions is called the law of mass action, and is most simply understood in the reaction between two substances, which we denote as  $A$  and  $B$ . The reaction is written as



indicating that a product  $P$  is formed at a rate  $r$ , which may typically have dimensions  $\text{mol l}^{-1} \text{s}^{-1}$ ; 'mol' indicates moles, and 'l' indicates litres. The law of mass action says that

$$r \propto AB, \quad (1.11)$$

where  $A$  and  $B$  indicate the concentrations of the respective reactants. This law represents the idea that the rate of a chemical reaction is proportional to the rate at which the reacting molecules run into each other, and in a homogeneous, well-mixed environment, this is simply proportional to the size of each population. Insofar as the same basic assumption applies to any population, the law of mass action will apply there too.

### 1.3 The $S-I-R$ model

A simple illustration of the law of mass action is the  $SIR$  model for epidemics in a population of fixed size. The population is divided into three classes, susceptible (those who have yet to catch the disease), infected, and removed (or recovered). The removed class may represent those who have recovered from the disease and are no longer susceptible (as, for example, in a flu epidemic) or those who have died from the disease. Three equations which describe the interactions of these three sub-populations are

$$\begin{aligned}\dot{S} &= -kSI, \\ \dot{I} &= kSI - rI, \\ \dot{R} &= rI.\end{aligned}\tag{1.12}$$

The nonlinear interaction term  $kSI$  represents the rate of infection, and assumes free dispersal of the infected class amongst the general population. Also of note is the first order decay term  $-rI$ , which describes ‘first order’ kinetics, and is indicative of the law of mass action applied to a single population. However, in the case of an epidemic, this term is of dubious provenance, as it implies that there is an exponential distribution of infection times, whereas in reality diseases tend to last for a reasonably fixed period. Consideration of this leads us to consider an age-structured model.

Let us suppose that  $i(t, a)$  represents the density (with respect to  $a$ ) of infected individuals at time  $t$  who have been infected for a time  $a$  (i.e., their ‘age’ of infection is  $a$ ). We suppose the rate of removal  $r(a)$  is a function of age, so that

$$\frac{di}{dt} = -r(a)i,\tag{1.13}$$

and the evolution of age with time is simply

$$\frac{da}{dt} = 1.\tag{1.14}$$

These are the characteristic equations for the age-structured evolution equation

$$\frac{\partial i}{\partial t} + \frac{\partial i}{\partial a} = -r(a)i.\tag{1.15}$$

The total infected class is

$$I = \int_0^\infty i \, da,\tag{1.16}$$

which on integrating (1.15) implies

$$\frac{dI}{dt} = - \int_0^\infty r(a) i da + i|_{a=0}. \quad (1.17)$$

Evidently  $i|_{a=0}$  is the rate of addition to the infected class, thus we take

$$i|_{a=0} = kSI, \quad (1.18)$$

and this provides the initial condition for  $i$  on  $a = 0$ .

If  $r$  is constant we recover (1.12). For  $r = r(a)$ , we solve (1.15) using the method of characteristics, i. e., we solve the ordinary differential equations (1.13) and (1.14) subject to the initial conditions on  $a = 0$  which can be parameterised as

$$t = \eta, \quad i = i_0(\eta) \equiv kS(\eta)I(\eta), \quad a = 0, \quad (1.19)$$

valid for  $\eta > 0$ . (The other part of the initial condition is applied at  $t = 0$ ,  $a > 0$ , but this part of the initial condition only contributes to a transient part of the solution for a finite time.)

The solution of (1.13) and (1.14) is then found to be

$$i = i_0(t - a) \exp \left[ - \int_0^a r(s) ds \right], \quad (1.20)$$

valid for  $t > a$ . For  $t < a$ , the solution depends on the initial condition at  $t = 0$ , and if, for example,  $i = 0$  initially, then  $i = 0$  for  $t < a$ .

Now let us consider the particular case where there is a fixed period of infection,  $\tau$ . This means essentially that  $r = 0$  for  $a < \tau$ , but becomes rapidly infinite<sup>2</sup> as  $a$  increases through  $\tau$ . As a consequence, we have that

$$\int_0^a r(s) ds = \begin{cases} 0, & a < \tau, \\ \infty, & a > \tau, \end{cases} \quad (1.21)$$

and thus

$$\begin{aligned} i &= i_0(t - a), & a < \tau, \\ i &= 0, & a > \tau. \end{aligned} \quad (1.22)$$

Finally we have that (for  $t > \tau$ )

$$I = \int_{t-\tau}^t i_0(s) ds. \quad (1.23)$$

In view of the definition of  $i_0$ , we can deduce from this that

$$I = -S(t) + S(t - \tau), \quad (1.24)$$

and therefore that  $S$  satisfies the differential delay equation

$$\dot{S} = kS[S - S_\tau], \quad (1.25)$$

where  $S_\tau = S(t - \tau)$ .

---

<sup>2</sup>  $r$  is actually then what is called a generalised function, or distribution.

## 1.4 Saturation and cooperativity

The simple quadratic interaction terms of the law of mass action become more complicated either when there are more reactions, or when the reaction rates are themselves modified by external factors. An example of the latter sort of term arises in the simple model for spruce budworm infestations given by Murray (2002). The model describes the evolution of spruce budworm moth density, denoted by  $B$ , as

$$\dot{B} = r_B B \left(1 - \frac{B}{K_B}\right) - \frac{\beta B^2}{\alpha^2 + B^2}. \quad (1.26)$$

The spruce budworm larvae develop in pine forests in Northern Canada, where they feed on the foliage. The model describes the growth of the larval density as a term proportional to population,  $r_B B$  (representing the year to year growth of larval populations through the successive stages of pupation, moth development and egg formation). The specific growth rate  $r_B$  is modified by a term which represents the effects of crowding on the population. Thus we have the logistic growth model

$$\dot{B} = r_B B \left(1 - \frac{B}{K_B}\right). \quad (1.27)$$

The quantity  $K_B$  represents a saturation limit beyond which  $B$  cannot grow, and this represents one way in which the effects of saturation may be modelled. It is not the only way, and not necessarily the most realistic.

The second term on the right hand side represents predation of the larvae by birds, for whom the larvae form part (but certainly not all) of the staple diet. The form of this term is of some interest. It is apparently independent of the size of the bird population. This implicitly assumes that there are so many birds that there is always one available to eat available larvae. Alternatively, it assumes a constant bird population, whose size is not controlled by larval density, but by other factors such as climate.

At high values of  $B$ , the predation rate saturates. Presumably this represents the fact that the appetite of birds and other predators is finite. Saturation is often modelled by a logistic term  $\propto \frac{B}{B + B_0}$ , which is a linear at small  $B$ , corresponding to first order kinetics. When the dependence is quadratic, as in (1.26), the rate is said to display the *Allee effect*. There are (at least) two motivations for such a quadratic dependence. The first is that when  $B$  is very small, one can imagine birds finding other prey more easily, and making less effort to locate larvae. A quadratic dependence would indicate that the effort or time spent searching for larvae is proportional to their density. The second is that at higher densities, the availability of larvae may attract birds. If bird density is proportional to larval density, then a quadratic dependence also ensues.

### 1.4.1 Cooperativity

Cooperativity refers to a biochemical reaction process in which an enzyme reacts with a chemical species, often referred to as a substrate. The very simplest version (which

is not actually cooperative!) is the Michaelis-Menten reaction scheme



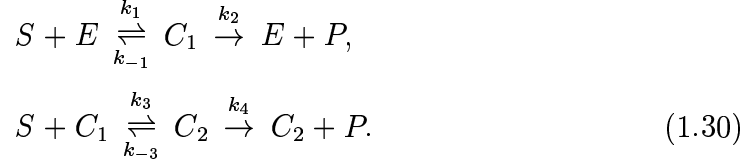
which represents the reaction of a substrate ( $S$ ) with an enzyme ( $E$ ) to form a complex ( $C$ ) which produces a product ( $P$ ), together with the original enzyme. The enzyme thus catalyses the reaction (i.e., it is not absorbed by the reaction).

Although the law of mass action is applied to these reactions separately, it is usually the case that the first reaction (the formation of the complex) is very fast, essentially because there is normally not much enzyme present. As a consequence of this, the forward rate  $k_1SE$  is approximately equal to the backward rate  $k_{-1}C$ . In addition, the total enzyme is conserved, either as free enzyme  $E$  or in the bound form  $C$ , so that  $E + C = E_0$  is constant. It follows from these observations that the product rate formation, i.e., the rate of reaction, is

$$r = k_2C \approx \frac{k_2E_0S}{S + K_1}, \quad (1.29)$$

where  $K_1 = k_{-1}/k_1$ . This provides a reaction-based explanation for saturating rate kinetics.

The actual phenomenon of cooperativity refers to the ability of some enzymes to have multiple binding sites for certain substrates. For example, with two binding sites we might consider the reaction scheme



It is easy to analyse this in the same way, assuming the binding reactions are fast. We find that the overall production rate is

$$r = \frac{(k_2K_1S + k_4K_1K_3S^2)E_0}{1 + K_1S + K_1K_3S^2}, \quad (1.31)$$

where  $K_1 = k_{-1}/k_1$  and  $K_3 = k_{-3}/k_3$ . The predation rate in (1.26) is a particular case of this, if  $K_3$  is large enough,

$$r \approx \frac{k_4K_1K_3E_0S^2}{1 + K_1K_3S^2}. \quad (1.32)$$

It is easy to extend these ideas to enzymes with  $n$  binding sites, and we then find that the reaction rates are expressible as ratios of  $n$ -th degree polynomials. In particular, such models can produce the rates described by Hill functions, in the form

$$r \approx \frac{kS^n}{S^n + S_0^n}; \quad (1.33)$$

such models are often used as empirical rate terms, for example in respiratory control modelling.

## 1.5 Notes and references

There are quite a lot of books about mathematical modelling, but they do vary in range. What you certainly don't want are books on simple methods: how to do the Laplace transform, and so on. There are a number of relatively recent books on the principles of mathematical modelling, for example Tayler (1986), Haberman (1998), Fowkes and Mahony (1994), Howison (2005); these are aimed at undergraduates, as is the present text. Fowler (1997) aims at a more graduate level. The classic of the type is the book by Lin and Segel (1974), now very dated but certainly an absolute model of clarity.

There are obviously many books which describe the technical material which is used here, but not so many which are written from the viewpoint of the practising applied mathematician. Of several fitting this latter description, we mention Keener (2000), Ockendon *et al.* 2003, and Carrier and Pearson (1976).

### 1.5.1 *SIR* model

The *SIR* model was introduced to epidemiology by Kermack and McKendrick (1927). It is described in the book by Murray (2002), where various extensions to the model are described. Recovery of the susceptible population (for example by population growth) can lead to oscillations, indicative of repetitive outbreaks. If these occur in a spatially distributed domain, then population migration (modelled as a diffusion term) can cause an epidemic to propagate as a travelling wave. Apparently this is how the Black Death spread in Europe in the fourteenth century.

The short book by Hoppensteadt (1975) describes age dependent population models, as well as the *SIR* model.

### 1.5.2 Spruce budworm model

Murray's (2002) book also describes the simplest version of the spruce budworm model. This itself is a simplification of a three variable continuous model described by Ludwig *et al.* (1979), which itself is an applied mathematician's attempt to propose a simpler version of the complex simulation model of Jones (1979). The Ludwig model is analysed in greater depth by Fowler (1997), and the Jones 'budworm site model' has been analysed by Hassell *et al.* (1999). Royama (1992) discusses the ecological background of the problem.

## Exercises

1.1 Consider the discrete population model (1.1):

$$u_{n+1} = ru_n \left(1 - \frac{u_n}{K}\right) - hu_n.$$

By writing  $u_n = Lw_n$  for some suitable choice of  $L$ , show that the model takes the form

$$w_{n+1} = \lambda w_n(1 - w_n),$$

and determine  $\lambda$ . What is the effect of increasing  $h$  on the behaviour of the population? What happens if  $h > r - 1$ ?

1.2 Consider the continuous population model (1.2):

$$\frac{du}{dt} = \rho u \left(1 - \frac{u}{C}\right) - \mu u.$$

By writing  $u = Mw$  for some suitable choice of  $M$ , show that the model takes the form

$$\frac{dw}{dt} = \alpha w(1 - w),$$

and determine  $\alpha$ . What is the effective of increasing  $\mu$  on the population? What happens if  $\rho < \mu$ ?

1.3 Starting from an integral conservation law, derive the heat equation in the form

$$\frac{\partial}{\partial t}(\rho c_p T) + \nabla \cdot [\rho c_p T \mathbf{u}] = \nabla \cdot [k \nabla T] + Q,$$

and the mass conservation equation in the form

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0.$$

1.4 The *SIR* model of an epidemic is

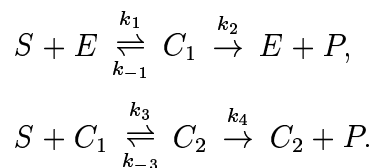
$$\begin{aligned}\dot{S} &= -kSI, \\ \dot{I} &= kSI - rI, \\ \dot{R} &= rI,\end{aligned}$$

where  $S$  represents the susceptible population,  $I$  the infected population, and  $R$  the recovered (or removed) population. Suppose that at  $t = 0$ ,  $S = S_0$ ,  $I$  is very small (perhaps a single individual), and  $R = 0$ . Show that an epidemic will occur if the *reproductive rate*

$$R = \frac{kS_0}{r}$$

is greater than one, and find  $I$  as a function of  $S$ .

1.5 A model for the reaction of a substrate with an enzyme with two binding sites is given by



If the reversible reaction rates  $k_{\pm 1}$  and  $k_{\pm 3}$  are very fast, show that the rate of formation of product  $P$  is approximately

$$r = \frac{(k_2 K_1 S + k_4 K_1 K_3 S^2) E_0}{1 + K_1 S + K_1 K_3 S^2},$$

where  $K_1 = k_{-1}/k_1$  and  $K_3 = k_{-3}/k_3$ .



# Chapter 2

## Non-dimensionalisation and approximation

Once we have a model, we have to try and solve it. There are two kinds of solutions: exact, analytical solutions, and approximate solutions. Exact solutions are explicit formulae; for example we can exactly solve quadratic equations, and certain differential equations, such as that describing simple harmonic motion. We also consider that solutions such as Taylor series constitute analytic solutions: they can be computed to arbitrary accuracy. The same applies to quadratures, such as the solution of

$$\frac{du}{dt} = f(u), \quad u(0) = u_0, \quad (2.1)$$

which has an implicitly defined solution

$$\int_0^u \frac{d\xi}{f(\xi)} = t. \quad (2.2)$$

Approximate solutions are those where one solves an approximate equation, or an approximating sequence of equations. Approximate methods are best applied when the approximation is based on the size of certain terms. In this chapter we will illustrate the use of such methods, firstly on simple algebraic equations, and then on some differential equations. The whole subject of perturbation theory is extensive, and a thorough discussion is beyond the scope of these notes.

### 2.1 Non-dimensionalisation

In order to approximate a solution, we need to be able to neglect terms which are small. This raises a concept of fundamental importance, which is that ‘small’ and ‘large’ are adjectives which can only apply quantitatively when a comparison is made between quantities of the same dimension. An equivalent assertion is that we can only make approximations based on the small size of parameters if these parameters are dimensionless. It makes no intrinsic sense to say a quantity is small if it still has dimensions. A speed of  $1 \text{ cm s}^{-1}$  is small if you are a rocket, but large if you are a

giant African land snail. An ice sheet is thick if you are a human being, but thin if you are a planet. So we always associate large or small quantities with dimensionless numbers, that is, parameters which measure quantities with respect to some reference value. The process of isolating these parameters is called *non-dimensionalisation*.

### 2.1.1 The wave equation

Putting a mathematical model into non-dimensional form is fundamental. Although technically trivial, there is a certain art to the process of non-dimensionalisation, and the associated concept of scaling, and the only real way to learn how to do it is by example. Let us begin with a simple model, the wave equation, which one learns how to derive in first year applied mathematics courses. We suppose a string, for example a guitar string, is wound between two points a distance  $l$  apart. If the tension in the string is  $T$  and its density (per unit length) is  $\rho$ , then an application of Newton's second law to an infinitesimal segment of the string leads to the equation

$$\rho \frac{\partial^2 y}{\partial t^2} = T \frac{\partial^2 y}{\partial x^2}, \quad y = 0 \quad \text{on} \quad x = 0, l, \quad (2.3)$$

where  $x$  is distance along the string, and  $y$  is its transverse displacement.

The main assumption that is usually stated in deriving this equation is that the displacement  $y$  is small. However, there are at least two other implicit assumptions which are made. One is that gravity is unimportant; the other is that there is no longitudinal displacement.

For a guitar string, these seem to be reasonable assumptions, but why? We expect the effect of gravity to be a deviation of the displacement from the vertical, and this is evidently valid for the guitar string. It is not valid for the hanging chain, or for the wire between telegraph poles. Why? I would say, for the chain, the density is too large; for the telegraph wire, the distance  $l$  is too large; while the guitar string is straight because it is tight:  $T$  is large. These facts suggest that the 'size' of the gravitational acceleration  $g$  may in fact be measured by the dimensionless parameter  $\rho gl/T$ , which appears to be the only independent dimensionless parameter which can be formed from those in the model if we include gravity.

How can this suspicion be confirmed? From first principles, we derive the wave equation, including gravity, in the form

$$\rho \frac{\partial^2 y}{\partial t^2} = T \frac{\partial^2 y}{\partial x^2} - \rho g, \quad y = 0 \quad \text{on} \quad x = 0, l. \quad (2.4)$$

Next we write the model in dimensionless form. We do this by *non-dimensionalising* the variables so that they are numerically of order one (written  $O(1)$ ). Specifically, we write

$$x = lx^*, \quad y = y_0 y^*, \quad t = (l/c)t^*, \quad (2.5)$$

where

$$c = \sqrt{\frac{T}{\rho}} \quad (2.6)$$

is the wave speed, and  $y_0$  is a measure of the displacement: for example, it could be the maximum initial displacement. The dimensionless model is then obtained in the form

$$\frac{\partial^2 y^*}{\partial t^{*2}} = \frac{\partial^2 y^*}{\partial x^{*2}} - \beta, \quad (2.7)$$

where

$$\beta = \frac{\rho g l^2}{T y_0}, \quad (2.8)$$

with  $y^* = O(1)$  initially, and  $y^* = 0$  on  $x = 0, l$ . All of the terms in the equations and in the initial and boundary conditions are dimensionless, and all the coefficients which appear (such as  $\beta$ ) are dimensionless.

It is conventional at this dimensionless stage to dispense with the asterisks in writing the variables, and this we now do. The process of choosing particular scales for the variables, or *scaling*, is motivated by the following ideas. Firstly, there is a natural length scale to the problem,  $l$ , which is the dimension of the geometric domain on which the problem is to be solved. Further, there is a natural length scale for the displacement,  $y_0$ , which is present in the initial conditions. Next, of the three terms in (2.4), we anticipate (at least for the guitar string) that the gravity term will be ‘small’. It follows that the other two should be the same size, and we choose the time scale so that these two terms ‘balance’, that is their dimensionless scales (here  $T y_0^2 / l^2$ ) are the same. When the model is written in the dimensionless form (2.7), we then have equal dimensionless coefficients multiplying these two terms: here they are both equal to one.

The final, essential idea is that, in general, a dimensionless function  $u(x, t)$  which varies by  $O(1)$  over an  $x$  range of  $O(1)$  will have derivatives of  $O(1)$ . This is true, for example, for  $\sin x$ ,  $e^{-x}$ , and  $x^2$ : it is *not* true for the function  $e^{-10x}$ , which varies rapidly over a distance of order  $x \sim 0.1$  near  $x = 0$ . With this assumption, the derivative terms  $\partial^2 y / \partial t^2$  and  $\partial^2 y / \partial x^2$  are  $O(1)$ , and it follows that the relative size of the gravity term is given by  $\beta$ . Thus gravity is negligible if  $\beta \ll 1$ , and indeed this means large tension, small density or short length, as we surmised.

### 2.1.2 The heat equation

Next we consider a form of the heat equation, (1.8). We write it in the form (assuming density  $\rho$  and specific heat  $c_p$  are constant)

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T = \kappa \nabla^2 T + H, \quad (2.9)$$

where  $H = Q / \rho c_p$ . We have assumed  $\nabla \cdot \mathbf{u} = 0$ , which follows from the conservation of mass in the form

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \quad (2.10)$$

together with  $\rho = \text{constant}$ .

Suppose we are to solve (2.9) in a domain  $D$  of linear dimension  $l$ , on the boundary of which we prescribe

$$T = T_B \quad \text{on} \quad \partial D, \quad (2.11)$$

where  $T_B$  is constant. We also have an initial condition

$$T = T_0(\mathbf{x}) \quad \text{in } D, \quad t = 0, \quad (2.12)$$

and we suppose  $\mathbf{u}$  is given, of order  $U$ .

We can make the variables dimensionless in the following way:

$$\mathbf{x} = l\mathbf{x}^*, \quad t = [t]t^*, \quad T = T_B + (\Delta T)T^*. \quad (2.13)$$

Again, we do this in order that both dependent and independent variables be of numerical order one,  $O(1)$ . If we can do this, then we would suppose *a priori* that derivatives such as  $\nabla^* T^*$  ( $\nabla = l^{-1} \nabla^*$ ) will also be of numerical  $O(1)$ , and the size of various terms will be reflected in the dimensionless parameters which occur.

In writing (2.13), it is clear that  $l$  is a suitable length scale, as it is the size of  $D$ . For example, if  $D$  was a sphere we might take  $l$  as its radius or diameter. We also suppose that the origin is in  $D$ ; if not, we could write  $\mathbf{x} = \mathbf{x}_0 + l\mathbf{x}^*$ , where  $\mathbf{x}_0 \in D$ : evidently  $\mathbf{x}^* = O(1)$  in  $D$ .

A similar motivation underlies the choice of an ‘origin shift’ for  $T$ . In the absence of a heat source, the temperature will tend to the uniform state  $T \equiv T_B$  as  $t \rightarrow \infty$ . If  $H \neq 0$ , the final state will be raised above  $T_B$  (if  $H > 0$ ) by an amount dependent on  $H$ . We take  $\Delta T$  to represent this amount, but we do not know what it is in advance — we will choose it by *scaling*. The subtraction of  $T_B$  from  $T$  before non-dimensionalisation is because the model for  $T$  contains only derivatives of  $T$ , so that it is really the variation of  $T$  about  $T_B$  which we wish to scale.

In a similar way, the time scale  $[t]$  is not prescribed in advance, and we will choose it also by scaling, in due course.

With the substitutions in (2.13), the heat equation (2.9) can be written in the form

$$\left( \frac{l^2}{\kappa[t]} \right) \frac{\partial T^*}{\partial t^*} + \left( \frac{Ul}{\kappa} \right) \mathbf{u}^* \cdot \nabla^* T^* = \nabla^{*2} T^* + \left( \frac{Hl^2}{\kappa \Delta T} \right), \quad (2.14)$$

where we have written  $\mathbf{u} = U\mathbf{u}^*$ , so that  $\mathbf{u}^* = O(1)$ . This equation is dimensionless, and the bracketed parameters are dimensionless. They are somewhat arbitrary, since  $[t]$  and  $\Delta T$  have not yet been chosen: we now do so by scaling.

The solution of the equation can depend only on the dimensionless parameters. It is thus convenient to *choose*  $[t]$  and  $\Delta T$  so that two of these are set to some convenient value. There is no unique way to do this.

The temperature scale  $\Delta T$  appears only in the source term. Since it is this which determines the temperature rise, it is natural to *choose*

$$\Delta T = \frac{Hl^2}{\kappa}. \quad (2.15)$$

It is also customary to choose the time scale so that the two terms of the advective derivative on the left of (2.14) are the same size, and this gives the convective time scale

$$[t] = \frac{l}{U}. \quad (2.16)$$

Finally, we remove the asterisks. When this is done, the dimensionless equation takes the form

$$Pe \left[ \frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T \right] = \nabla^2 T + 1, \quad (2.17)$$

where the Péclet number is

$$Pe = \frac{Ul}{\kappa}, \quad (2.18)$$

and the solution of the model depends only on this parameter (as well as the initial condition). The boundary condition is

$$T = 0 \quad \text{on} \quad \partial D, \quad (2.19)$$

and the initial condition is

$$T = \theta(\mathbf{x}) \quad \text{at} \quad t = 0, \quad (2.20)$$

where

$$\theta(\mathbf{x}) = \frac{T_0(l\mathbf{x}) - T_B}{\Delta T}. \quad (2.21)$$

## 2.2 Scaling

A well-scaled problem generally refers to a model in which the dimensionless parameters are  $O(1)$  or less. Evidently, this can be ensured simply by dividing through by the largest parameter in any equation. More importantly, if parameters are numerically small, then (as we discuss below) approximate solutions can be obtained by neglecting them. The problem is well-scaled if the resulting approximation makes sense. For example, (2.17) is well-scaled for any value of  $Pe$ . However, the problem  $\varepsilon T_t = \varepsilon \nabla^2 T + 1$ , with  $\varepsilon \ll 1$ , is not well scaled. One makes a problem well-scaled in this situation by *rescaling* the variables, and we now consider the wave equation (2.7) again in this light.

### 2.2.1 The wave equation, again

In the statement of the wave equation with gravity, there are in fact two dimensionless parameters:

$$B = \frac{\rho g l}{T}, \quad \varepsilon = \frac{y_0}{l}. \quad (2.22)$$

The parameter  $\varepsilon$  is a measure of the amplitude of the motion, and it is on the basis that  $\varepsilon \ll 1$  that we derive the linear wave equation in the first place. The parameter  $\beta = B/\varepsilon$ , so the assumption of negligible gravity is equivalent to the assumption that  $B \ll \varepsilon \ll 1$ .

If  $B \sim \varepsilon$ , then  $\beta = O(1)$ . The problem is sensibly scaled, but gravity is no longer negligible. There is a steady state  $y = \frac{1}{2}\beta x(l - x)$  (the hanging chain), and, because the equation is linear, the string simply oscillates about this steady state.

Now suppose that  $\beta \gg 1$ . The model is now not correctly scaled (because the limit  $\beta \rightarrow \infty$  gives no sensible approximation). In fact the model suggests that a

steady state will have  $y \sim \beta$ , and that the string will oscillate about this steady state with a similar amplitude. In order to obtain a sensibly scaled problem, we simply have to *rescale* the displacement by writing  $y = \beta Y$ , and the discussion can proceed as before, except that the initial condition has  $Y \ll 1$ .

It seems that all is now well; we have discussed the cases  $\beta \ll 1$ ,  $\beta = O(1)$ , and  $\beta \gg 1$ . However, there is more concern when  $\beta$  becomes as large as  $O(1/\varepsilon)$ . In this case, the model suggests oscillations about a steady state of order  $\beta \sim 1/\varepsilon$ , and dimensionally of order  $l$ . Because of this, the basis of the original derivation is suspect, and the model must be re-derived: in fact this can be done (see question 2.1), so that the model equation (2.7) remains valid. However, the initial value scale  $y_0$  is *not* now an appropriate scale for  $y$ ; the appropriate (dimensional) scale is  $y \sim l$ , and this *a posteriori* adjustment is the rescaling alluded to above. In dimensionless terms, we rescale the model by writing  $y = (l/y_0)Y = Y/\varepsilon$ , and we find

$$\frac{\partial^2 Y}{\partial t^2} = \frac{\partial^2 Y}{\partial x^2} - B, \quad (2.23)$$

and this version of the model is appropriate if  $B = O(1)$ . If  $B \gg 1$ , a further rescaling simply takes  $Y \sim B$ .

## 2.3 Simple approximation methods

Suppose we wish to solve the equation

$$x^3 - \varepsilon x - 1 = 0, \quad (2.24)$$

where  $\varepsilon$  is relatively small: for example if  $\varepsilon = 0.1$ . Formulae do in fact exist for writing solutions of cubic (and also quartic) equations, but they are fairly unpleasant and are rarely used. A much better way is to use an approximation method, based on the idea that the *parameter*  $\varepsilon$  in equation (2.24) is small.

Graphically, it is clear (see figure 2.1) that when  $\varepsilon$  is small, (2.24) will have a unique, positive real root, and in fact it will be close to  $x = 1$  (since  $1 + \varepsilon x \approx 1$ ). (In passing, note that there will be exactly one root for  $\varepsilon < \varepsilon_c$ , where  $\varepsilon_c$  is the value corresponding to tangency of the line with the curve; at this value  $x^3 = 1 + \varepsilon x$  and  $3x^2 = \varepsilon$ , so that  $\varepsilon_c = 3/2^{2/3} \approx 1.89$ .)

A simple iterative method to solve (2.24) is

$$x_{n+1} = (1 + \varepsilon x_n)^{1/3}. \quad (2.25)$$

If we choose  $\varepsilon = 0.1$  and  $x_0 = 1$ , then successively

$$\begin{aligned} x_1 &= 1.0322801, \\ x_2 &= 1.0332889, \\ x_3 &= 1.0333204, \\ x_4 &= 1.0333214, \end{aligned} \quad (2.26)$$

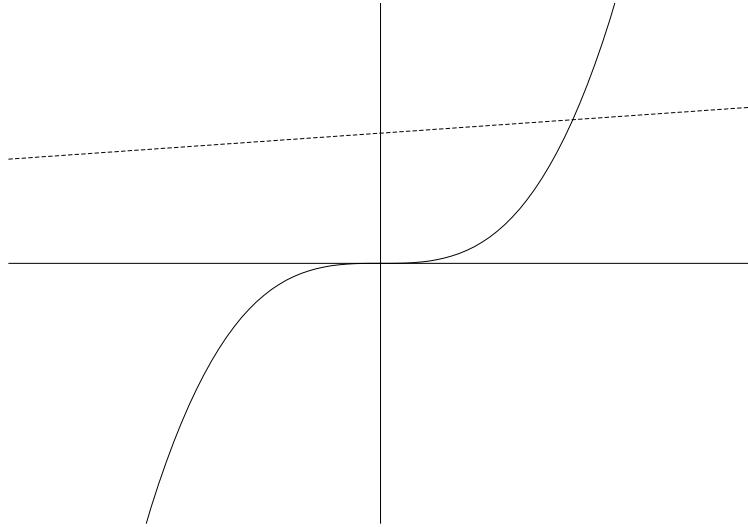


Figure 2.1: The graphs of  $x^3$  and  $1 + \varepsilon x$ , with  $\varepsilon = 0.1$ .

and this last value is the root. It is of course trivial to compute the positive root for a range of  $\varepsilon$ , but it would be convenient to have an *analytic* (as opposed to numerical) approximation. We construct this by using a *perturbation method*.

We can use the binomial expansion to write (2.24) in the form

$$x = (1 + \varepsilon x)^{1/3} = 1 + \frac{1}{3}\varepsilon x - \frac{1}{9}\varepsilon^2 x^2 + \frac{5}{81}\varepsilon^3 x^3 \dots \quad (2.27)$$

Since  $\varepsilon \ll 1$ , we see that  $x \approx 1 + O(\varepsilon)$ : that is, terms of *order*  $\varepsilon$ , i.e. of size about  $\varepsilon$ . A better estimation is then

$$x \approx 1 + \frac{1}{3}\varepsilon x \approx 1 + \frac{1}{3}\varepsilon, \quad (2.28)$$

and we can see that this relatively crude approximation is in fact accurate to four decimal places when  $\varepsilon = 0.1$ ! Repeating this idea, we would have

$$\begin{aligned} x &\approx 1 + \frac{1}{3}\varepsilon x - \frac{1}{9}\varepsilon^2 x^2 \\ &\approx 1 + \frac{1}{3}\varepsilon(1 + \frac{1}{3}\varepsilon) - \frac{1}{9}\varepsilon^2(1 + \dots)^2 \approx 1 + \frac{1}{3}\varepsilon \end{aligned} \quad (2.29)$$

(there is no  $O(\varepsilon^2)$  term), but a more methodical procedure is to anticipate (by inspection) that the root can be written in the form of a series

$$x = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \dots; \quad (2.30)$$

we substitute this into (2.24), expand in powers of  $\varepsilon$ , and *equate coefficients of powers of*  $\varepsilon$ : a little thought indicates why this procedure is necessary. For (2.24), we thus have

$$(x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \dots)^3 - \varepsilon(x_0 + \varepsilon x_1 + \dots) - 1 = 0, \quad (2.31)$$

whence

$$\begin{aligned} (x_0^3 - 1) &+ \varepsilon(3x_0^2x_1 - x_0) + \varepsilon^2(3x_0^2x_2 + 3x_0x_1^2 - x_1) \\ &+ \varepsilon^3(3x_0^2x_3 + 3x_0x_1x_2 + x_1^3 - x_2) + \dots = 0, \end{aligned} \quad (2.32)$$

so that, sequentially,

$$\begin{aligned} x_0^3 - 1 &= 0, \\ 3x_0^2x_1 - x_0 &= 0, \\ 3x_0^2x_2 + 3x_0x_1^2 - x_1 &= 0, \\ 3x_0^2x_3 + 3x_0x_1x_2 + x_1^3 - x_2 &= 0, \\ &\dots, \end{aligned} \quad (2.33)$$

from which we obtain

$$x_0 = 1, \quad x_1 = \frac{1}{3}, \quad x_2 = 0, \quad x_3 = -\frac{1}{81}, \quad (2.34)$$

and hence the root is

$$x \approx 1 + \frac{1}{3}\varepsilon - \frac{1}{81}\varepsilon^3 + O(\varepsilon^4). \quad (2.35)$$

With  $\varepsilon = 0.1$ , we have  $x \approx 1.033321$ : practically, the exact result. Even, for  $\varepsilon = 1$ , the approximate result is 1.321, while the exact root is 1.3247.

## Singular approximations

The approximation above is called a *regular* approximation, because the limit when  $\varepsilon = 0$  gives an approximation to the root. Now consider the cubic

$$\varepsilon x^3 - x - 1 = 0, \quad \varepsilon \ll 1. \quad (2.36)$$

Graphically (figure 2.2), there are clearly three real roots. One of these is near  $-1$  and can be recovered by a regular approximation. The others are at large values of  $|x|$ , and are determined by *balancing*  $\varepsilon x^3$  with  $x$ ; that is,  $\varepsilon x^3 \sim x$  when  $x \sim \varepsilon^{-1/2}$ , so we first write

$$x = \varepsilon^{-1/2}X, \quad (2.37)$$

and then

$$X^3 - X - \varepsilon^{1/2} = 0, \quad (2.38)$$

with approximate roots  $X \approx 0, \pm 1$ . The root  $X \approx 0$  corresponds to  $x \approx -1$  and is determined by the regular approximation for  $x$ . The larger roots are determined by a regular approximation of (2.38), as a power series in  $\varepsilon^{1/2}$ . Thus

$$X = X_0 + \varepsilon^{1/2}X_1 + \varepsilon X_2 + \dots, \quad (2.39)$$

and substituting this into (2.38) and equating powers of  $\varepsilon^{1/2}$ , leads to the approximate solutions (written in terms of  $x$ )

$$x \approx \pm \frac{1}{\sqrt{\varepsilon}} + \frac{1}{2} \mp \frac{3}{8}\sqrt{\varepsilon} + \frac{7}{32}\varepsilon + \dots \quad (2.40)$$



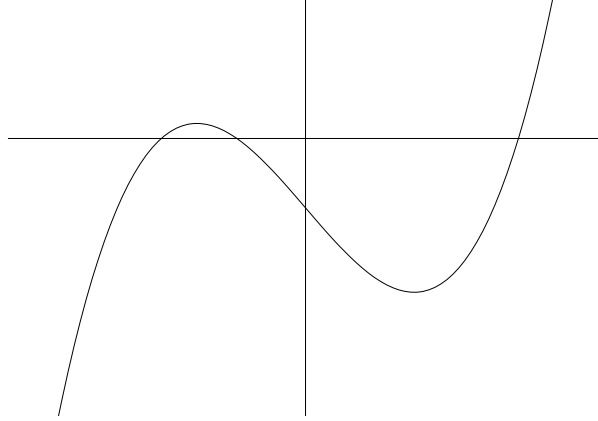


Figure 2.2: The graph of  $\varepsilon x^3 - x - 1$ , for  $\varepsilon = 0.1$ .

The upper sign gives the positive root, the lower the negative root. For example if  $\varepsilon = 0.1$ , the approximate positive root from (2.40) is 3.565567, while the exact root is 3.5770894. For  $\varepsilon = 1$ , the approximate root is 1.1468753, while as we have seen, the exact root is 1.3247. This is less good, but can be improved by taking further terms in the series (2.40). We see that approximation methods can provide a very useful way of solving algebraic equations.

Now suppose we wish to solve

$$\tan x = \tanh x. \quad (2.41)$$

Each function is odd, so  $x = 0$  is a solution, and graphically (figure 2.3) it is clear that there is a sequence of positive roots  $x_1^*, x_2^*, \dots$  (and thus also negative roots  $-x_1^*, -x_2^*, \dots$ ), and that as  $n \rightarrow \infty$ ,

$$x_n^* \approx (n + \frac{1}{4})\pi. \quad (2.42)$$

Suppose we put

$$x_n^* = (n + \frac{1}{4})\pi + \theta; \quad (2.43)$$

then

$$\begin{aligned} \sin x_n^* &= \frac{(-1)^n}{\sqrt{2}}(\cos \theta + \sin \theta), \\ \cos x_n^* &= \frac{(-1)^n}{\sqrt{2}}(\cos \theta - \sin \theta), \end{aligned} \quad (2.44)$$

so (2.41) is

$$\tan x_n^* = \frac{\cos \theta + \sin \theta}{\cos \theta - \sin \theta} = \frac{1 - e^{-2x_n^*}}{1 + e^{-2x_n^*}}. \quad (2.45)$$

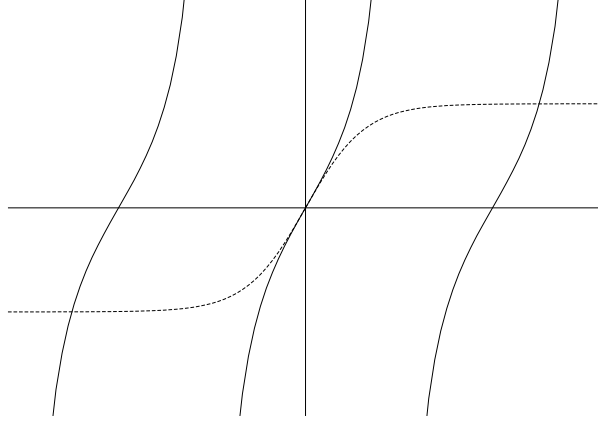


Figure 2.3: The graphs of  $\tan x$  and  $\tanh x$ .

We expect  $\theta \ll 1$ , and also  $x_n^* \gg 1$ , and thus, since

$$\begin{aligned}\cos \theta &\approx 1 - \frac{1}{2}\theta^2, \\ \sin \theta &\approx \theta,\end{aligned}\tag{2.46}$$

we have, with use of the binomial expansion for  $(1 + y)^{-1}$ ,

$$\begin{aligned}\frac{1 - \frac{1}{2}\theta^2 \dots + \theta \dots}{1 - \frac{1}{2}\theta^2 \dots - \theta \dots} &= (1 - e^{-2x_n^*})(1 - e^{-2x_n^*} + \dots) \\ \Rightarrow (1 + \theta - \frac{1}{2}\theta^2 \dots)(1 + \{\theta + \frac{1}{2}\theta^2 + \dots\} + \{\theta^2 + \dots\} \dots) &= 1 - 2e^{-2x_n^*} \dots \\ \Rightarrow 1 + 2\theta + 2\theta^2 + \dots &= 1 - 2e^{-(2n+\frac{1}{2})\pi - 2\theta} \dots \\ \Rightarrow \theta + \theta^2 + \dots &= -e^{-(2n+\frac{1}{2})\pi}(1 - 2\theta \dots),\end{aligned}\tag{2.47}$$

so that, finally,

$$\theta \approx -e^{-(2n+\frac{1}{2})\pi},\tag{2.48}$$

and

$$x_n^* \approx (n + \frac{1}{4})\pi - e^{-(2n+\frac{1}{2})\pi}.\tag{2.49}$$

## Numerical approximation

Iterative numerical methods are widely used to solve algebraic equations. A general iterative method to find a root of  $f(x) = 0$  is to define a sequence  $x = x_0, x_1, \dots$ , satisfying  $x_{n+1} = g(x_n)$ , where the function  $g$  is chosen so that  $x = g(x)$  if  $f(x) = 0$ : for example,  $g(x) = f(x) + x$ . A simple iterative method to solve

$$L(x) = R(x)\tag{2.50}$$

is as follows: define

$$L(x_{r+1}) = R(x_r),\tag{2.51}$$

i.e.  $x_{r+1} = L^{-1} \circ R(x_r)$ . The sequence will converge if  $|(L^{-1} \circ R)'| < 1$  at a root, and using the chain rule (via  $f = L^{-1} \circ R$  if  $L[f(x)] = R(x)$ , so  $L'f' = R'$ , and at a root  $f(x) = x$ ,  $f' = R'/L'$ ) we find that this is  $|L'| > |R'|$ . Consulting figure 2.3, we see that the iteration

$$x_{r+1} = \tan^{-1}[\tanh x_r] + n\pi \quad (2.52)$$

will converge to  $x_n^*$  (since  $\tan^{-1}$  is defined to be less than  $\pi/2$  on a calculator).

The lowest root ( $n = 1$ ) is approximated by

$$x_1^* \approx \frac{5}{4}\pi - e^{-5\pi/2} = 3.9266026 \quad (2.53)$$

compared with the exact value 3.9266024. Here an approximation based on the limit  $n \rightarrow \infty$  is accurate even when  $n = 1$ .

## 2.4 Perturbation methods

Let us consider (2.17) with (2.19) and (2.20), and suppose that  $\theta_0 \leq O(1)$ . If  $Pe \ll 1$ , we obtain an approximation by putting  $Pe = 0$ :  $\nabla^2 T + 1 \approx 0$ . Evidently, we cannot satisfy the initial condition, and this suggests that we rescale  $t$ : put  $t = Pe\tau$ , so that (approximately)

$$\frac{\partial T}{\partial \tau} = \nabla^2 T + 1; \quad (2.54)$$

now we can satisfy the initial condition (at  $\tau = 0$ ) too. Often one abbreviates the rescaling by simply saying, ‘rescale  $t \sim Pe$ , so that  $T_t \approx \nabla^2 T + 1$ ’.

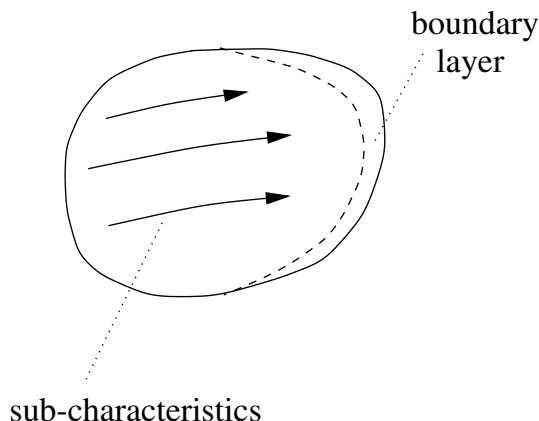


Figure 2.4: Sub-characteristics and boundary layer for the equation (2.17). The sub-characteristics are the flow lines  $d\mathbf{x}/dt = \mathbf{u}$ , and the boundary layer (of thickness  $O(1/Pe)$ ) is on the part of the boundary where the flow lines terminate.

On the other hand, if  $Pe \gg 1$ , then  $T_t + \mathbf{u} \cdot \nabla T \approx 0$ , and we can satisfy the initial condition but not the boundary condition on all of  $\partial D$ , since the approximating equation is hyperbolic (its characteristics are called ‘sub-characteristics’). To remedy

this, one has to rescale  $\mathbf{x}$  near the part of the boundary where the boundary condition is not satisfied. This gives a spatially thin region, called (evidently) a boundary layer, of thickness  $1/Pe$  (see figure 2.4).

The other possibility is if  $\theta \gg 1$ , say  $\theta \sim \theta_0 \gg 1$ . We discuss only the case  $Pe \gg 1$  (see also exercise 1.2). Since  $T \sim \theta_0$  initially, we need to rescale  $T$ , say  $T = \theta_0 \tilde{T}$ . then  $Pe[\tilde{T}_t + \mathbf{u} \cdot \nabla \tilde{T}] = \nabla^2 \tilde{T} + \theta_0^{-1}$ , and with  $\tilde{T} = O(1)$ , we have  $\tilde{T}_t + \mathbf{u} \cdot \nabla \tilde{T} \approx 0$  for  $Pe \gg 1$ . The initial function is simply advected along the flow lines (sub-characteristics), and the boundary condition  $\tilde{T} = 0$  is advected across  $D$ . In a time of  $O(1)$ , the initial condition is ‘washed out’ of the domain. Following this, we revert to  $T$ , thus  $T_t + \mathbf{u} \cdot \nabla T = Pe^{-1}(\nabla^2 T + 1)$ . Evidently  $T$  will remain  $\approx 0$  in most of  $D$ , with a boundary layer near the boundary as shown in figure 2.4. If  $n$  is the coordinate normal to  $\partial D$  in this layer, then  $\mathbf{u} \cdot \nabla T \sim u_n \partial T / \partial n \sim Pe T$ ,  $Pe^{-1} \nabla^2 T \sim Pe^{-1} \partial^2 T / \partial n^2 \sim Pe T$ , and in the steady state, these must balance the source term  $Pe^{-1}$ : thus in fact, the final state has the rescaled  $T \sim Pe^{-2}$ , and this applies also for  $\theta_0 \leq O(1)$ .

These ideas of perturbation methods are very powerful, but a full exposition is beyond the scope of these notes. Nevertheless, they will relentlessly inform our discussion. While it is possible to use formal perturbation expansions, it is sufficient in many cases to give more heuristic forms of argument, and this will typically be the style we choose.

## 2.5 Notes and references

Non-dimensionalisation is technically trivial, analytically essential, and difficult to get right. You won’t find many books that dwell on it, because of its apparent triviality. But, it is also clear that you can get it very wrong. The clearest precise exposition for the uninitiated is in the book by Lin and Segel (1974), and other than that it is largely a matter of example and practice. What is clear is that you can easily get it wrong!

## Exercises

- 2.1 Derive the wave equation describing oscillations of a string of length  $l$  from first principles, when gravity is included, assuming displacements are small and of order  $y_0$ . Show how to non-dimensionalise the equation to obtain the form

$$\frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2} - \beta,$$

and define  $\beta$ .

Now suppose that  $y_0 \sim l$ . Suppose that in the unstretched state where the displacement  $y = 0$ , the density  $\rho_0$  is constant. By careful consideration of the application of Newton’s second law to an infinitesimal segment of length

$ds$  (stretched from its original length  $dx$ ), show that the (dimensional) wave equation can be derived in the form

$$\rho_0 \frac{\partial^2 y}{\partial t^2} = T_0 \frac{\partial^2 y}{\partial x^2} - \rho_0 g,$$

assuming only that displacements are purely vertical, where  $T_0$  is the horizontal component of the tension exerted at the fixed end points at  $x = 0$ ,  $x = l$ .

Non-dimensionalise the model in this case and describe the form of the resulting oscillation.

2.2 A population of size  $N$  is subject to immigration at rate  $I$ , and mutual pair destruction at a rate  $kN^2$ , so that  $\dot{N} = I - kN^2$ . By appropriate scaling of the variables, show that the model can be written in the form  $\dot{x} = 1 - x^2$ .

2.3 A model describing the interactions of populations of a herbivore ( $H$ ) and plankton ( $P$ ) is

$$\begin{aligned} \frac{dP}{dt} &= rP \left[ (K - P) - \frac{BH}{C + P} \right], \\ \frac{dH}{dt} &= DH \left[ \frac{P}{C + P} - AH \right]. \end{aligned}$$

Explain what the terms in the equations mean, and suggest ecological reasons for their form.

Show how to non-dimensionalise the model to obtain the dimensionless pair of equations

$$\begin{aligned} \dot{x} &= x \left[ k - x - \frac{y}{1+x} \right], \\ \dot{y} &= \alpha y \left[ \frac{x}{1+x} - ay \right], \end{aligned}$$

and define the dimensionless parameters  $k$ ,  $a$  and  $\alpha$ . By identifying the units of the physical constants in the model, show that these parameters are indeed dimensionless.

2.4 Suppose

$$Pe \left[ \frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T \right] = \nabla^2 T + 1 \quad \text{in } D,$$

with

$$\begin{aligned} T &= 0 \quad \text{on } \partial D, \\ T &= \theta_0 \Theta(\mathbf{x}) \quad \text{in } D \quad \text{at } t = 0, \end{aligned}$$

and  $\Theta = O(1)$ ,  $\theta_0 \gg 1$ ,  $Pe \ll 1$ . Discuss appropriate scales for the various phases of the solution.

2.5 Explain why the iterative method

$$x_{n+1} = (1 + \varepsilon x_n)^{1/3}$$

used to solve (2.24) will converge to its solution. Does this depend on the value of  $\varepsilon$ ?

2.6 (i) Find approximations to the solution of

$$\varepsilon x^3 - x - 1 = 0, \quad \varepsilon \ll 1,$$

which is close to  $x = -1$ . Compare with the numerical solution when  $\varepsilon = 0.1$ ;  $\varepsilon = 0.01$ .

(ii) Use perturbation methods to find approximate roots to the equation

$$xe^{-x} = \varepsilon, \quad 0 < \varepsilon \ll 1.$$

(Use graphical methods to find the location of the roots. For the larger root, take logs and note that if  $x \gg 1$ , then  $x \gg \ln x$ .)

2.7 Each of the equations

$$z^5 - \varepsilon z - 1 = 0,$$

$$\varepsilon z^5 - z - 1 = 0,$$

has five (possibly complex) roots. Find approximations to these if  $\varepsilon \ll 1$ . Can you refine the approximations?

2.8 Suppose that  $m$  satisfies the polynomial

$$\frac{\varepsilon\alpha}{\gamma}m^3 + \varepsilon(\lambda + B)m^2 - (\lambda - A)m - 2\gamma(\lambda + B) = 0, \quad (*)$$

where  $\alpha, \gamma, A, B, \lambda$  are  $O(1)$  and  $\varepsilon \ll 1$ . Show that the roots are approximately given by

$$m_0 \approx \frac{2\gamma(\lambda + B)}{A - \lambda}, \quad m_{\pm} \approx \pm i \left( \frac{\gamma(A - \lambda)}{\varepsilon\alpha} \right)^{1/2}.$$

Now suppose that  $\gamma = g/\varepsilon$  and  $m = M/\varepsilon$ . Show that  $M$  satisfies

$$\frac{\alpha}{g}M^3 + (\lambda + B)M^2 + (A - \lambda)M - 2g(\lambda + B) = 0. \quad (**)$$

Suppose that  $g$  is large. Show that the roots are approximately given in this case by

$$M_{\pm} \approx \pm \sqrt{2g}, \quad M_0 \approx -\frac{(\lambda + B)g}{\alpha}.$$

By consideration of the roots of (\*\*) as  $g$  decreases, can you ascertain the path of the roots in the complex plane, and how  $M_{\pm}$  and  $M_0$  are transformed to the roots  $m_{\pm}$  and  $m_0$  of (\*)?

2.9 The Lorenz equations are

$$\begin{aligned}\dot{x} &= -\sigma x + \sigma y, \\ \dot{y} &= (r - z)x - y, \\ \dot{z} &= xy - bz.\end{aligned}$$

These equations are already dimensionless, and  $r$ ,  $b$  and  $\sigma$  are dimensionless parameters. Is it possible to rescale the equations in order to reduce the number of independent parameters? If so, how? If not, why not?

Suppose the parameter  $r$  is large. Show how to rescale the equations so that the damping terms in the equations are small (of size  $\varepsilon = 1/\sqrt{r}$ ), and show that if the terms in  $\varepsilon$  are put to zero, there are two first integrals of the equations. What does this imply about the solutions when  $\varepsilon = 0$ ?

2.10 The *SIR* model of an epidemic of a fatal disease is

$$\begin{aligned}\dot{S} &= -kSI, \\ \dot{I} &= kSI - rI, \\ \dot{R} &= rI,\end{aligned}$$

where  $S$  represents the susceptible population,  $I$  the infected population, and  $R$  the removed (dead) population. Suppose that at  $t = 0$ ,  $S = S_0$ ,  $I$  is very small (perhaps a single individual), and  $R = 0$ .

Non-dimensionalise the model to obtain the form

$$\begin{aligned}\dot{s} &= -sf, \\ \dot{f} &= sf - \frac{f}{R}, \\ \dot{\rho} &= \frac{f}{R},\end{aligned}$$

where the dimensionless parameter

$$R = \frac{kS_0}{r}$$

is called the reproductive rate.

Find a first integral for  $f$  in terms of  $s$ , and show that an epidemic will occur if  $R > 1$ .

Now suppose that the epidemic is weak, in the sense that  $R - 1 = \varepsilon \ll 1$ . Show that a suitable approximation of the model can be made by defining

$$s = 1 - \varepsilon\sigma, \quad f = \varepsilon^2 F, \quad t = \frac{T}{\varepsilon},$$

and show that the (scaled) death rate  $F$  of the resultant approximation is approximately given by

$$F = \frac{1}{2} \operatorname{sech}^2 \frac{1}{2} T.$$

# Chapter 3

## Perturbation methods

In chapter 2 we were introduced to the basic ideas of approximation techniques, which are central to the art of the practical applied mathematician. In this chapter we introduce some of the formal analytic methods which have been devised to provide approximate solutions to problems which we will suppose are framed as differential equations. We begin with some generalities, but the subject is best taught and learned by way of example, because by its nature there is no general theory which will provides a method of solution for all problems.

Suppose we have an ordinary differential equation for a function  $u(t)$ , which we write in the formal way

$$N(D, \varepsilon; u) = 0; \quad (3.1)$$

here  $N$  denotes a nonlinear differential operator of the ordinary derivative  $D = \frac{d}{dt}$ . Often  $N$  will be a polynomial function of both  $D$  and  $\varepsilon$ . If we suppose that  $N$  is analytic in  $\varepsilon$ , then we can expand the equation as

$$N(D, \varepsilon; u) = N_0(D; u) + \varepsilon N_1(D; u) + \dots = 0, \quad (3.2)$$

and this suggests that we seek a perturbation series for the solution of the form

$$u = \sum_{s=0}^{\infty} \varepsilon^s u_s, \quad (3.3)$$

and if we substitute this into (3.2), then we find that

$$\sum_{r,s} \varepsilon^{r+s} N_{r,s}(D; u_s, u_{s-1}, \dots) = 0. \quad (3.4)$$

Here we have expanded each operator term as

$$N_r \left( D; \sum_{s=0}^{\infty} \varepsilon^s u_s \right) = \sum_{s=0}^{\infty} \varepsilon^s N_{r,s}(D; u_s, u_{s-1}, \dots). \quad (3.5)$$

Note that each operator  $N_{r,s}$  depends on functions  $u_k$  with indices up to  $s$ . By changing the indices of summation, we can write (3.4) in the form

$$\sum_{j=0}^{\infty} \varepsilon^j \hat{N}_j(D; u_j, u_{j-1}, \dots) = 0, \quad (3.6)$$



where

$$\hat{N}_j(D; u_j, u_{j-1}, \dots) = \sum_{s=0}^j N_{j-s,s}(D; u_s, u_{s-1}, \dots). \quad (3.7)$$

The formal procedure is now to equate powers of  $\varepsilon$  to zero, yielding a sequence of equations

$$\hat{N}_j(D; u_j, u_{j-1}, \dots) = 0. \quad (3.8)$$

It is clear that we can, at least in principle, solve these sequentially for  $u_j$  to find

$$u_j = F(u_{j-1}, u_{j-2}, \dots). \quad (3.9)$$

This sequential procedure then yields our solution.

Apart from the fact that there are differential operators involved, what we have described above is simply an iterative procedure for evaluating a Taylor series solution for  $u$  in terms of  $\varepsilon$ . What makes perturbation methods interesting is that the series that is so constructed is rarely convergent. The more usual divergent power series is called an asymptotic series, or an asymptotic expansion, and is written

$$u \sim \sum_{s=0}^{\infty} \varepsilon^s u_s. \quad (3.10)$$

The symbol ‘ $\sim$ ’ denotes asymptotic equivalence, and in general we write  $u \sim v$  (as  $\varepsilon \rightarrow 0$ , for example) if the ratio  $u/v$  tends to a fixed limit as  $\varepsilon \rightarrow 0$ ; we also write  $u = O(v)$  in this case. Other notation is to write  $u = o(v)$  (or  $u \ll v$ ) if  $u/v \rightarrow 0$  as  $\varepsilon \rightarrow 0$ .

Despite its divergence, the series (3.10) can still yield a good approximation to the solution when  $\varepsilon$  is small. The reason for this is that the terms will then initially decrease, and so the finite sum  $\sum_0^N u_s$  will yield an increasingly accurate solution if  $N$  is fixed and  $\varepsilon$  is decreased.

Why should the series for  $u$  diverge if the original operator  $N$  is analytic in  $u$  and  $\varepsilon$ ? The divergence of an asymptotic series such as (3.10) indicates that the function  $u(t, \varepsilon)$  has a singularity at  $\varepsilon = 0$ . This must certainly be the case if  $u$  does not approach the solution  $u_0$  of the unperturbed equation

$$N_0(D; u_0) = 0, \quad (3.11)$$

and this is often the case. It is, for example, generally true for singularly perturbed problems, and these are precisely the types of problem which are of most interest (because they are the most difficult).

### 3.1 Regular perturbation theory

We illustrate some of the above abstract discussion by consideration of some specific examples. Let us consider the boundary value problem

$$u'' - u = \varepsilon u^2, \quad u(0) = 0, \quad u(1) = 1, \quad (3.12)$$

where  $\varepsilon \ll 1$ , and we will denote the independent variable as  $x$ , since boundary value problems usually arise in a spatial context.

Following our description above, we seek a solution in the form (3.10), and substituting this into (3.12) and equating powers of  $\varepsilon$ , we find the sequence of equations

$$\begin{aligned} u_0'' - u_0 &= 0, \\ u_1'' - u_1 &= u_0^2, \\ u_2'' - u_2 &= 2u_0u_1, \end{aligned} \tag{3.13}$$

and so on. The boundary conditions are also expanded in this way, so that we have

$$\begin{aligned} u_0 = u_1 = u_2 = 0 &\quad \text{at } x = 0, \\ u_0 = 1, \quad u_1 = u_2 = 0 &\quad \text{at } x = 1. \end{aligned} \tag{3.14}$$

It is easy, if messy, to solve this sequence of equations. We have

$$\begin{aligned} u_0 &= \frac{\sinh x}{\sinh 1}, \\ u_1 &= \frac{-\frac{1}{2} + \frac{1}{6} \cosh 2x + \frac{1}{3} \cosh x}{\sinh^2 1} + \frac{\left(\frac{1}{2} - \frac{1}{6} \cosh 2 - \frac{1}{3} \cosh 1\right) \sinh x}{\sinh^3 1}, \end{aligned} \tag{3.15}$$

for example. We can see that  $u_0$  contains terms in  $e^x$  and  $e^{-x}$ , while  $u_1$  contains terms from  $e^{2x}$  down to  $e^{-2x}$ . It is not difficult to show by induction that  $u_r$  contains terms from  $e^{(r+1)x}$  down to  $e^{-(r+1)x}$ . This suggests that the perturbation series will be convergent for sufficiently small  $\varepsilon$ ; see also question 3.2. Note that this will not be true for an infinite interval of integration, for example if we prescribe two initial conditions at  $x = 0$  and require solutions valid in  $0 < x < \infty$ . We discuss this case further in section 3.3 below.

## 3.2 Singular perturbation theory: boundary layers

Next we consider an equation which is fairly similar to (3.12), but with the small coefficient  $\varepsilon$  multiplying the highest derivative:

$$\varepsilon u'' - u' = u^2, \quad u(0) = 0, \quad u(1) = 1, \tag{3.16}$$

and  $\varepsilon \ll 1$ .

We use the same procedure as before, writing

$$u \sim u_0(x) + \varepsilon u_1(x) + \dots, \tag{3.17}$$

and then equating terms of the same order to obtain

$$\begin{aligned} u_0' + u_0^2 &= 0, \\ u_1' + 2u_0u_1 &= u_0'', \\ u_2' + 2u_0u_2 &= u_1'' - u_1^2, \end{aligned} \tag{3.18}$$

and so on. The associated boundary conditions are

$$\begin{aligned} u_0 = u_1 = u_2 = 0 & \quad \text{at } x = 0, \\ u_0 = 1, \quad u_1 = u_2 = 0 & \quad \text{at } x = 1. \end{aligned} \quad (3.19)$$

The essential distinguishing feature about this problem is that the small coefficient multiplies the highest derivative of the equation. Because of this, the leading order problem (for  $u_0$ ) is of lower degree than the full equation, and in general this means that it is not possible to satisfy all of the boundary conditions.

An issue of choice then arises, because this leading order ‘outer’ solution may be chosen differently depending on which boundary conditions are satisfied. In addition, nonlinearity of the equation may allow the possibility of genuinely multiple solutions.

In the present case, we have at leading order

$$u'_0 = -u_0^2. \quad (3.20)$$

Let us suppose, apparently arbitrarily, that we apply the left hand boundary condition

$$u_0 = 0 \quad \text{at } x = 0, \quad (3.21)$$

and in general we will then aim to satisfy the boundary conditions on  $u_n$  at  $x = 0$ . As we shall see, the issue of choice is generally resolved by requiring consistency of the solution procedure.

Picard’s theorem tells us that

$$u_0 \equiv 0 \quad (3.22)$$

is the only solution of (3.20) with (3.21). Since  $u_0$  is constant, it is not difficult to see that the successive terms  $u_1$ ,  $u_2$  and the rest will also all be zero, so that this outer solution is given by  $u = 0$  to all orders of  $\varepsilon$ . Evidently,  $u = 0$  cannot be the exact solution (though it exactly solves the equation) because it does not satisfy the boundary condition at  $x = 1$ . Here we have an early glimpse of the notion of *exponential asymptotics*: the outer solution for  $u$  is smaller than any algebraic power of  $\varepsilon$ , and this suggests that it may be exponentially small, or  $u = \exp \left[ -O \left( \frac{1}{\varepsilon} \right) \right]$ .

In any event, we have an approximation to the solution which does not satisfy the boundary condition at  $x = 1$ . This is associated with the loss of the highest derivative, and such problems are called singular perturbation problems, presumably because the resulting perturbation solution is singular (having, in this case, a discontinuity at  $x = 1$ ). In order to resolve the failure of the method, we need to be able to bring back the highest derivative term into the approximation scheme, and the way in which we can formally do this is by adjusting the length scale. We do this near  $x = 1$ , since that is where we have the problem. We put

$$x = 1 - \delta X, \quad (3.23)$$

and suppose that  $\delta$ , as yet unidentified, is small. Changing the independent variable in (3.16), we find that  $u$ , now taken as a function of  $X$ , satisfies the equation

$$\frac{\varepsilon}{\delta^2} u'' + \frac{1}{\delta} u' = u^2, \quad (3.24)$$

where now  $u' = du/dX$ .

This scaling of  $x$  places us in a *boundary layer* next to the boundary at  $x = 1$ . Since the object of the rescaling is to bring back the highest derivative term at leading order, we see that we must choose

$$\delta = \varepsilon, \quad (3.25)$$

and with this the equation can be written

$$u'' + u' = \varepsilon u^2. \quad (3.26)$$

The issue of appropriate boundary conditions for (3.26) is crucial to the procedure. One is fairly clear, that we specify

$$u = 1 \quad \text{at} \quad X = 0; \quad (3.27)$$

this is just the right hand boundary condition.

The other condition is less clear. We already have an outer expansion (3.17) which satisfies the left hand boundary condition, so the other condition can not be that. Nevertheless, we press on, aiming to find an *inner* expansion for (3.26) of the form

$$u \sim U_0(X) + \varepsilon U_1(X) + \dots, \quad (3.28)$$

and then, equating powers of  $\varepsilon$ , we have

$$\begin{aligned} U_0'' + U_0' &= 0, \\ U_1'' + U_1' &= U_0^2, \end{aligned} \quad (3.29)$$

and so on.

The leading order solution for  $U_0$  is

$$U_0 = e^{-X} + A(1 - e^{-X}), \quad (3.30)$$

where  $A$  is a constant, and further constants arise in the determination of higher order terms. The issue is how we choose the constant  $A$  (and the other constants which arise at higher order). The idea is this: we have two expansions, an inner and outer expansion, each representing an approximation to the solution in different regions; one where  $1 - x \sim O(1)$ , the other where  $1 - x \sim O(\varepsilon)$ . We suppose that each expansion should blend into the other in a fuzzy intermediate zone where  $\varepsilon \ll 1 - x \ll 1$ . This region is called a matching region, and the method is thus called the method of *matched asymptotic expansions*.

There is a methodical procedure for matching the expansions, which is beyond the scope of the present discussion, but is necessary if matching is to be carried out for the higher order terms. However, its implementation for the leading order terms is usually fairly straightforward. The simplest situation is when the outer solution tends to a constant at the boundary adjoining the boundary layer, while the inner solution tends to a constant as the inner variable becomes large. This is exactly the situation that applies here. In this case the matching condition is to ensure that these two constants are the same. Comparing (3.22) and (3.30), we see that the two solutions match if  $A = 0$ . Thus we have

$$\begin{aligned} u &\sim 0, & 1 - x &\sim O(1), \\ u &\sim \exp[-(1 - x)/\varepsilon], & 1 - x &\sim O(\varepsilon). \end{aligned} \quad (3.31)$$

## Uniform solutions

Although the inner and outer expansions formally blend together, actual graphical illustration of them generally indicates discontinuity, and it is attractive to write a solution which is uniformly valid. It is possible to do this. The trick is to realise that, since both inner and outer expansions are supposed to be valid in the matching region, we can get a uniform solution by adding the inner and outer solutions and then subtracting that part of the solution which each represents in the matching region. In the present example this is trivial, since the common part of the outer and inner solutions is zero, and in fact the inner solution is uniformly valid.<sup>1</sup>

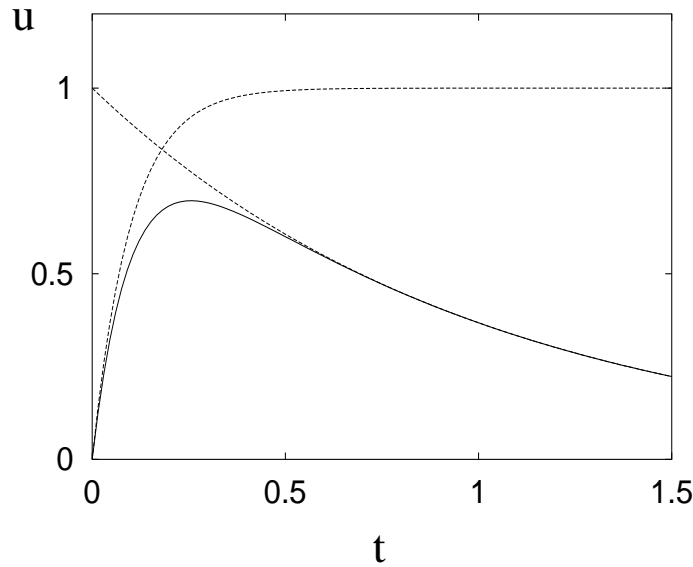


Figure 3.1: Inner, outer and uniform solutions to (3.33) when  $\varepsilon = 0.1$ .

An easy illustration of this idea is the function

$$u(t) = e^{-t} - e^{-t/\varepsilon}, \quad (3.32)$$

which is a uniform first order expansion of the solution to the initial value problem

$$\varepsilon \ddot{u} + \dot{u} + u = 0, \quad u(0) = 0, \quad \dot{u}(0) = \frac{1}{\varepsilon}, \quad \varepsilon \ll 1. \quad (3.33)$$

The first order outer solution of this problem is  $u = Ae^{-t}$ , while the first order inner solution is  $u = 1 - e^{-T}$ , where  $t = \varepsilon T$ . Matching requires  $A = 1$ , and the common part of both approximations in the intermediate matching region is then just their common limit, which is one. Thus the uniform first order solution adds the inner and outer solutions and subtracts one. Figure 3.1 shows that both inner and outer approximations are well approximated in their regions of validity by the uniform expansion.

---

<sup>1</sup>This lends substance to our earlier comment concerning exponential asymptotics.

### 3.2.1 Internal layers

The question of which boundary or boundaries one should attach boundary layers to is a matter of possibilities. Because the rapidly varying solution in a boundary layer must blend into a slowly varying outer solution, it needs to decay in the matching region. Very often, this decay is exponential, and the choice of exponential decay rather than exponential growth determines the boundary layer location. For example, we found that (3.16) allowed a boundary layer at a right hand boundary, but (3.33) will only allow boundary layers on the left.

More generally, the equation

$$\varepsilon y'' + a(x)y' + b(x)y = 0 \quad (3.34)$$

will allow left hand boundary layers where  $a > 0$  and right hand boundary layers where  $a < 0$  (see also question 3.6). This raises a potential problem, which we can illustrate by consideration of the nonlinear problem

$$\varepsilon y'' + 2yy' + y(1 - y^2) = 0, \quad (3.35)$$

with  $0 < \varepsilon \ll 1$  and

$$y(-1) = -1, \quad y(1) = 1. \quad (3.36)$$

As we expect both the derivative terms to be important in any boundary layer (why?), we see that the sign of  $y$  at the boundaries precludes boundary layers there, and any region of rapid change must occur in the interior of the domain. Such regions are called interior layers, or in partial differential equation models they are called shocks (there will be more on these in chapter 7).

(3.35) has outer solutions satisfying both boundary conditions given by

$$y = \pm 1 \quad \text{for} \quad x \lesseqgtr x_0, \quad (3.37)$$

where  $x_0$  is indeterminate. There is also a possible outer solution  $y = 0$ , which we will ignore in this discussion. If the solution is unique, then it is odd (since  $-y(-x)$  is also a solution), and then  $x_0 = 0$ . But because the problem is nonlinear, uniqueness (or even existence) is not guaranteed.

In the transition region, we put

$$x = x_0 + \varepsilon X; \quad (3.38)$$

then to leading order  $y$  satisfies

$$y'' + 2yy' = 0, \quad y(\pm\infty) = \pm 1, \quad (3.39)$$

and thus

$$y = \tanh X, \quad (3.40)$$

where a constant of integration which fixes where  $y = 0$  can be absorbed into the definition of  $x_0$ .

This solution describes the interior layer, but there is nothing to determine the position  $x_0$  of the transition. The reason for this is that the outer solutions  $y = \pm 1$  are, like that of (3.16), accurate to all orders of  $\varepsilon$ . Consequently, as far as the asymptotic solution is concerned, the location of the boundaries is immaterial, and it is because of this that the interior position of the transition layer is indeterminate. Its position (if a solution exists) presumably depends on the exponentially small correction terms which lie beyond all orders of the expansion in powers of  $\varepsilon$ .

A less degenerate case occurs if we modify the boundary conditions to be

$$y(-1) = -B_-, \quad y(1) = B_+, \quad (3.41)$$

where  $B_{\pm} > 0$ . Suppose that  $B_{\pm} < 1$ , and define

$$B_{\pm} = \tanh A_{\pm}, \quad (3.42)$$

where  $A_{\pm} > 0$ . The previous boundary condition (3.36) corresponds to the double limit  $A_{\pm} \rightarrow \infty$ . The outer solutions are  $y = 0$  or

$$y_{\pm} = \pm \tanh \left( \frac{1 \mp x + 2A_{\pm}}{2} \right). \quad (3.43)$$

Now at an internal transition layer, (3.39) allows solutions

$$y = y^* \tanh y^* X; \quad (3.44)$$

the jump is thus from  $-y^*$  to  $y^*$ , and this allows us to determine  $x_0$  from (3.43):

$$x_0 = A_+ - A_-. \quad (3.45)$$

Thus we obtain a self-consistent asymptotic solution, but only for a restricted range of boundary conditions. We can also see the degeneracy as  $A_{\pm} \rightarrow \infty$ .

This example, apart from illustrating the occurrence of an interior layer, also shows how asymptotic solutions are not always straightforward, and often require guile to discover their intricacies. This is particularly true of nonlinear problems. To some extent, the problem arises here because the boundary conditions (3.36) are exact solutions of the equation. But it is also emblematic of the artificial nature of the example. It is more generally the case that good mathematical models of real physical systems do not have such pathological behaviour.

### 3.3 Multiple scales and averaging

A different kind of singular limit can occur if the domain of integration is infinite. This is typically the case in time dependent problems, and the usual sorts of illustrative example are models of oscillators. As an example, suppose that  $u(t)$  satisfies the initial value problem

$$\ddot{u} + u = \varepsilon u^2, \quad u(0) = 1, \quad \dot{u}(0) = 0, \quad (3.46)$$

where  $\varepsilon \ll 1$ , and  $\dot{u} = du/dt$ . This is the equation of a weakly nonlinear oscillator, a spring with a weak amplitude-dependent spring constant  $\propto 1 - \varepsilon u$ . We seek a perturbation expansion of the form

$$u \sim u_0 + \varepsilon u_1 + \dots, \quad (3.47)$$

and substituting this into (3.46) we have the sequence of equations

$$\begin{aligned} \ddot{u}_0 + u_0 &= 0, \\ \ddot{u}_1 + u_1 &= u_0^2, \\ \ddot{u}_2 + u_2 &= 2u_0 u_1, \end{aligned} \quad (3.48)$$

and so on. The initial conditions at  $t = 0$  for these are

$$\begin{aligned} u_0 &= 1, & \dot{u}_0 &= 0, \\ u_1 &= 0, & \dot{u}_1 &= 0, \\ u_2 &= 0, & \dot{u}_2 &= 0. \end{aligned} \quad (3.49)$$

It is straightforward but tedious to solve these, and we find the resultant expansion for  $u$  to be

$$\begin{aligned} u \sim \cos t + \varepsilon \left[ \frac{1}{2} - \frac{1}{6} \cos 2t - \frac{1}{3} \cos t \right] \\ + \varepsilon^2 \left[ -\frac{1}{3} + \frac{29}{144} \cos t + \frac{1}{9} \cos 2t + \frac{1}{48} \cos 3t + \frac{5}{12} t \sin t \right] \dots \end{aligned} \quad (3.50)$$

This is a perfectly good asymptotic solution for bounded times, but the series becomes disordered at large times; specifically the third term becomes as large as the first when  $t = O(1/\varepsilon^2)$ . This contradicts the assumption that the terms are decreasing asymptotically.

The breakdown in the expansion is associated with the occurrence of the *secular* term  $t \sin t$  at  $O(\varepsilon^2)$ , which arises through a resonance between the basic linear oscillator and the forcing terms in the expansion. Now we know that the solutions are indeed oscillatory, and so we know that the amplitude cannot grow in this way. It is also clear why the method must break down. The period of the nonlinear oscillator is not constant, and depends weakly on the amplitude. But the unperturbed oscillator has fixed period, and our expansion method forces this period on the perturbation solution. It is this incorrect implicit assumption of an unperturbed period which generates the secular terms.

This idea is simply illustrated. The function  $\cos[(1 - \varepsilon)t]$  has period  $2\pi + O(\varepsilon)$ . If we expand in powers of  $\varepsilon$ , we find

$$\cos[(1 - \varepsilon)t] = \cos t + \varepsilon t \sin t + \dots; \quad (3.51)$$

the secular terms arise through a misguided expansion of the oscillatory term.



### 3.3.1 Poincaré-Lindstedt method

There are various ways to fix this breakdown, of which the simplest is the Poincaré-Lindstedt method. Essentially we need to generalise our naïve perturbation method to allow for a drift in the period of the underlying oscillator.

We do this by straining the time variable; we define a new time  $\tau$  via

$$\tau = (1 + \varepsilon\omega_1 + \varepsilon^2\omega_2 + \dots)t, \quad (3.52)$$

so that (3.46) becomes

$$\left[1 + 2\varepsilon\omega_1 + \varepsilon^2(\omega_1^2 + 2\omega_2) + \dots\right] u'' + u = \varepsilon u^2, \quad (3.53)$$

where now  $u' = du/d\tau$ . The constants  $\omega_i$  are to be chosen precisely to eliminate the secular terms in the expansion.

Expanding  $u$  as a perturbation series as before, we find

$$\begin{aligned} u_0'' + u_0 &= 0, \\ u_1'' + u_1 &= u_0^2 - \omega_1 u_0'', \\ u_2'' + u_2 &= 2u_0 u_1 - (\omega_1^2 + 2\omega_2) u_0'', \end{aligned} \quad (3.54)$$

and we solve these successively as before. There are two little tricks which make it less algebraically nightmarish than our earlier calculation. One is to write the general solution of (3.54)<sub>1</sub> as

$$u_0 = Ae^{it} + (\text{cc}), \quad (3.55)$$

where (cc) denotes the complex conjugate, and  $A$  is in general a complex amplitude. The second is not to specify  $A = \frac{1}{2}$ , but to leave it as an undetermined constant at this stage. The point is that we can in principle subsume all the fundamental solutions  $e^{it}$  which occur at each order into the definition of  $A$ , or equivalently we allow  $A$  to have its own expansion in powers of  $\varepsilon$ .

The  $O(\varepsilon)$  terms imply

$$u_1'' + u_1 = A^2 e^{2it} + |A|^2 + \omega_1 A e^{it} + (\text{cc}); \quad (3.56)$$

terms  $\propto e^{it}$  (or its complex conjugate) are secular producing terms, and so we must choose

$$\omega_1 = 0; \quad (3.57)$$

there is no secular term at  $O(\varepsilon)$ , and the (particular) solution is

$$u_1 = |A|^2 - \frac{1}{3} A^2 e^{2it} + (\text{cc}). \quad (3.58)$$

At the next order, we find (taking care to ensure we have all the nonlinear terms)<sup>2</sup>

$$u_2'' + u_2 = 2 \left[ \frac{5}{3} |A|^2 A e^{it} - \frac{1}{3} A^3 e^{3it} + \omega_2 A e^{it} + (\text{cc}) \right]; \quad (3.59)$$

---

<sup>2</sup>Note that  $(a + \bar{a})(b + \bar{b}) = (a + \bar{a})b + (\text{cc})$ , where the overbar denotes the complex conjugate.

suppression of secular terms requires us to choose

$$\omega_2 = -\frac{5}{3}|A|^2, \quad (3.60)$$

and the (particular) solution is

$$u_2 = \frac{1}{12}A^3e^{3it} + (\text{cc}). \quad (3.61)$$

As we surmised, the period  $P$  (in  $t$ ) is weakly amplitude dependent,

$$P = 2\pi \left(1 + \frac{5}{3}\varepsilon^2|A|^2 + \dots\right), \quad (3.62)$$

and the coefficient  $\frac{5}{12}$  if  $|A| = \frac{1}{2}$  is consistent with the secular term in (3.50). This procedure can be carried on sequentially, but usually it is only the leading order frequency correction which is of interest, which is also why it is not really necessary to determine the amplitude of the fundamental beyond leading order.

### 3.3.2 Van der Pol oscillator

The Van der Pol oscillator is the equation for  $x(t)$ ,

$$\ddot{x} + \varepsilon(x^2 - 1)\dot{x} + x = 0, \quad (3.63)$$

and we suppose  $\varepsilon$  is small and positive. Unlike the conservative nonlinear oscillator of the preceding section, there is a unique stable limit cycle of the Van der Pol equation. The Poincaré-Lindstedt method will find this limit cycle, both its leading order amplitude and the frequency correction (see question 3.7), but to prove stability of the limit cycle, a more subtle approach is necessary.

This retains the idea that the leading order solution will be of the form

$$x \approx Ae^{it} + (\text{cc}), \quad (3.64)$$

but rather than locking the time scale to a fixed modification of the original one, we instead allow the amplitude  $A$  to vary slowly in time. This is a generalisation of the Poincaré-Lindstedt method, insofar as we can replace  $A \exp[i(1 + \varepsilon\omega_1 + \dots)t]$  (with constant  $A$ ) by  $A(\tilde{t})e^{it}$ , where the *slow time*  $\tilde{t}$  here could be taken to be  $\varepsilon t$ . Yet more generally, one can choose both a new strained fast time  $t^* = t(1 + \dots)$  as well as a slow time, and for strongly nonlinear equations (unlike the weakly nonlinear (3.63)), fast and slow times may have complicated functional dependence on  $t$ . At higher order, yet more slow time scales may be necessary.

However, in its simplest form, such complications are not necessary. Consideration of (3.63) suggests that we define fast and slow times

$$t^* = \omega t, \quad \tilde{t} = \varepsilon t, \quad (3.65)$$

where

$$\omega = (1 + \varepsilon^2\omega_2 + \dots), \quad (3.66)$$

and we then formally seek the solution as a function  $x(t^*, \tilde{t})$  of the two times  $t^*$  and  $\tilde{t}$ . These give the method of multiple scales its name.

The chain rule relates the  $t$ -derivatives to the partial  $t^*$ - and  $\tilde{t}$ -derivatives, so that the equation (3.63) is embedded in the hyperbolic partial differential equation

$$\omega^2 \frac{\partial^2 x}{\partial t^{*2}} + 2\varepsilon\omega \frac{\partial^2 x}{\partial t^* \partial \tilde{t}} + \varepsilon^2 \frac{\partial^2 x}{\partial \tilde{t}^2} + \varepsilon (x^2 - 1) \left( \omega \frac{\partial x}{\partial t^*} + \varepsilon \frac{\partial x}{\partial \tilde{t}} \right) + x = 0. \quad (3.67)$$

We actually construct a perturbation solution of this equation in the two-dimensional  $(t^*, \tilde{t})$  space, but are only concerned with its value on the characteristic through the origin. There is some flexibility about how we impose initial conditions. If  $x(0) = x_0$  and  $\dot{x}(0) = v_0$ , then the most natural initial data for (3.67) is to specify

$$\begin{aligned} x &= x_0 \quad \text{on} \quad \tilde{t} = 0, \\ \omega \frac{\partial x}{\partial t^*} &= v_0 \quad \text{on} \quad \tilde{t} = 0. \end{aligned} \quad (3.68)$$

As in our discussion of the Poincaré-Lindstedt method, we shall not be much concerned with initial conditions.

We now expand the solution in the usual way,

$$x \sim x_0 + \varepsilon x_1 + \dots, \quad (3.69)$$

where each  $x_i = x_i(t^*, \tilde{t})$ . We then gain the following sequence of problems:

$$\begin{aligned} \frac{\partial^2 x_0}{\partial t^{*2}} + x_0 &= 0, \\ \frac{\partial^2 x_1}{\partial t^{*2}} + x_1 &= -2 \frac{\partial^2 x_0}{\partial t^* \partial \tilde{t}} - (x_0^2 - 1) \frac{\partial x_0}{\partial t^*}, \end{aligned} \quad (3.70)$$

and luckily for us we need go no farther, since secular terms appear already at  $O(\varepsilon)$ . (If they did not, then there would be no need of the  $O(1/\varepsilon)$  slow time scale, and a slower time scale would be relevant.) The solution at leading order is

$$x_0 = A(\tilde{t})e^{it^*} + (\text{cc}), \quad (3.71)$$

and therefore the equation at the next order is

$$\frac{\partial^2 x_1}{\partial t^{*2}} + x_1 = -2iA'e^{it^*} - \left[ iA^3e^{3it^*} + (2|A|^2 - 1)iAe^{it^*} - i|A|^2Ae^{it^*} \right] + (\text{cc}). \quad (3.72)$$

We do not even need to solve this, we simply need to choose  $A$  in order to remove the secular terms  $\propto e^{it^*}$ . We do this by requiring  $A$  to satisfy the Ginzburg–Landau equation

$$2 \frac{dA}{d\tilde{t}} = A - |A|^2 A. \quad (3.73)$$

Multiplying this by  $\bar{A}$  and its conjugate by  $A$  and adding yields the equation for  $|A|$ ,

$$\frac{d|A|^2}{d\tilde{t}} = |A|^2 - |A|^4, \quad (3.74)$$

which shows that  $|A| \rightarrow 1$  as  $t \rightarrow \infty$ , and there is a stable limit cycle

$$x \sim 2 \cos t. \quad (3.75)$$

The amplitude of the limit cycle is two, and the frequency is shifted by  $O(\varepsilon^2)$ , but to find the magnitude of the frequency shift would require evaluation of the  $O(\varepsilon^2)$  terms.

### 3.3.3 Averaging

The method of averaging is another formal method which is particularly useful for weakly nonlinear oscillations, and also can be used rapidly to give a leading order evolution equation for the amplitude of motion. It relies on the fact that the unperturbed system has a constant of integration, which will usually be the energy. We illustrate its use with the Van der Pol equation,

$$\ddot{x} + \varepsilon(x^2 - 1)\dot{x} + x = 0. \quad (3.76)$$

The unperturbed equation is that of simple harmonic motion, with constant amplitude  $a$  and phase velocity  $\dot{\theta}$ . This suggests the exact change of variables

$$x + i\dot{x} = ae^{i\theta}, \quad (3.77)$$

as a consequence of which

$$\frac{1}{2}x^2 + \frac{1}{2}\dot{x}^2 = \frac{1}{2}a^2. \quad (3.78)$$

The energy  $\frac{1}{2}a^2$  is conserved when  $\varepsilon = 0$ , and therefore we can expect that  $a$  will be slowly varying in the motion. Specifically, we have from (3.76)

$$a\dot{a} = -\varepsilon(x^2 - 1)\dot{x}^2; \quad (3.79)$$

using (3.77), this implies

$$\dot{a} = \varepsilon a(1 - a^2 \cos^2 \theta) \sin^2 \theta, \quad (3.80)$$

and we can also deduce that

$$\dot{\theta} = -1 + O(\varepsilon), \quad (3.81)$$

though this is of less concern. The central idea of the averaging method is that since  $a$  and  $\dot{\theta}$  are approximately over the underlying period of the oscillation, we can effectively integrate the amplitude equation over this period keeping  $a$  constant. The result of doing this to (3.80) is to take the average (in  $\theta$ ) of the right hand side, and this yields the approximate amplitude equation

$$\dot{a} \approx \varepsilon a \left( \frac{1}{2} - \frac{1}{8}a^2 \right). \quad (3.82)$$

Putting  $a = 2A$ , we have the same evolution equation as (3.73), and

$$a \rightarrow 2 \quad \text{as} \quad t \rightarrow \infty, \quad (3.83)$$

as we found before.

## 3.4 Notes and references

There is a wealth of books on perturbation methods. A rapid assessment is that by Hinch (1991); longer classics are those by Kevorkian and Cole (1981) and Bender and Orszag (1978); the last of these annotates its text with many illustrative numerical examples.

## Exercises

3.1 Find two terms of a regular perturbation expansion of the form

$$u \sim \sum_0^{\infty} \varepsilon^n u_n$$

to the boundary value problem

$$u'' + u = \varepsilon u^2, \quad u(0) = 0, \quad u(\pi/2) = 1,$$

when  $\varepsilon$  is small and positive.

Show by induction that  $u_n$  can be written in the form

$$u_n = \sum_{-(n+1)}^{n+1} a_k e^{ikx}.$$

Repeat the exercise for the equation

$$u'' + u = \varepsilon u^3,$$

with the same boundary conditions. Can you modify the inductive step to find a general form of expression for  $u_n$ ?

3.2 Suppose that  $x(t)$  and  $y(t)$  are real and satisfy the pair of first order equations

$$\begin{aligned} \dot{x} &= -y - 2\varepsilon axy, \\ \dot{y} &= x + \varepsilon(x^2 - y^2), \end{aligned}$$

in which  $a$  is constant, and  $x$  and  $y$  satisfy initial conditions

$$x(0) = 1, \quad y(0) = 0.$$

By eliminating  $y$ , find a second order equation for  $x$  of the form

$$\ddot{x} + x = \varepsilon f(x, \dot{x}; a),$$

and give an expression for the function  $f$ . Write down the associated initial conditions for  $x$ .

Use perturbation theory to find two terms of an expansion for  $x$  in powers of  $\varepsilon$ .

Find an exact solution for the case  $a = 1$  by writing  $z = x + iy$ , and show that the expansion of this solution in powers of  $\varepsilon$  agrees with the perturbation expansion. Deduce that in this case the perturbation expansion converges for sufficiently small  $\varepsilon$ .

3.3 The functions  $u(t)$  and  $v(t)$  satisfy the initial value problem

$$\begin{aligned}\varepsilon \dot{u} &= v - u, \\ \dot{v} &= -u,\end{aligned}$$

with the initial conditions

$$u(0) = 0, \quad v(0) = 1.$$

Show that if the outer solution is chosen to satisfy the initial condition for  $v$ , then this outer solution is  $v = e^{-t}$ . Show that there must be a boundary layer near  $t = 0$ , and by rescaling  $t$ , write down a suitable version of the equations in this boundary layer, and find a leading order approximation for  $u$  and  $v$  in the boundary layer.

Hence write down a uniform approximation for  $u$ .

3.4 The Michaelis-Menten model of enzyme kinetics is the pair of equations

$$\begin{aligned}\dot{s} &= -s(1 - c) + (K - L)c, \\ \varepsilon \dot{c} &= s - (s + K)c,\end{aligned}$$

where  $\varepsilon \ll 1$  and  $K, L = O(1)$ . The initial conditions are that

$$s(0) = 1, \quad c(0) = 0,$$

and the constants  $\varepsilon$ ,  $K$  and  $L$  are all positive. Find leading order approximations to the solution.

3.5 Find leading order approximations (including boundary layers, where necessary) to the solutions of the boundary value problems

$$\begin{aligned}\varepsilon y'' + 2y' + e^y &= 0, \quad y(0) = y(1) = 0, \\ \varepsilon y'' - x^2 y' - y &= 0, \quad y(0) = y(1) = 1.\end{aligned}$$

3.6 Suppose that  $y(x)$  satisfies the boundary value problem

$$\varepsilon y'' + a(x)y' + b(x)y = 0,$$

with  $0 < \varepsilon \ll 1$  and

$$y(0) = A, \quad y(1) = B.$$

Suppose also that  $a(x)$  and  $b(x)$  are analytic in  $[0, 1]$  (i.e., have convergent Taylor series about any point in this interval).

(i) If  $a > 0$ , find an approximate solution and show that it has a boundary layer at  $x = 0$ .

(ii) If  $a < 0$ , find an approximate solution and show that it has a boundary layer at  $x = 1$ .

(iii) Finally, if  $a < 0$  for  $x < x_0$  and  $a > 0$  for  $x > x_0$ , where  $0 < x_0 < 1$ , show that no boundary layers at the end points can exist, and therefore that an interior layer must exist at  $x_0$ .

Suppose that

$$\beta = \frac{b(x_0)}{a'(x_0)}.$$

Show that as  $x \rightarrow x_0 \pm$ , the outer solutions in  $x \lesseqgtr x_0$  satisfy

$$y \sim c_{\pm} |x - x_0|^{\beta},$$

where the constants  $c_{\pm}$  should be specified.

Hence show that by rescaling  $x$  and  $y$  as

$$y = \{\varepsilon a'(x_0)\}^{\beta/2} Y, \quad x = x_0 + \{\varepsilon a'(x_0)\}^{1/2} X,$$

the equation can be approximately written in the transition region as

$$Y'' + XY' + \beta Y = 0,$$

and explain why suitable matching conditions for the solution of this equation are

$$Y \sim c_{\pm} |X|^{\beta} \quad \text{as} \quad X \rightarrow \pm\infty.$$

3.7 The Van der Pol oscillator is represented by the equation

$$\ddot{x} + \varepsilon (x^2 - 1) \dot{x} + x = 0,$$

and we assume that  $0 < \varepsilon \ll 1$ .

Use the Poincaré-Lindstedt method to find the amplitude and frequency correction of the periodic solution of the equation.

3.8 Duffing's equation is

$$\ddot{x} + x + \varepsilon x^3 = 0.$$

By first finding a first integral, show that all solutions are periodic if  $\varepsilon > 0$ . If  $\varepsilon \ll 1$ , use the Poincaré-Lindstedt method to construct approximate solutions, and find the frequency correction of a periodic solution in terms of  $\varepsilon$  and its amplitude  $A$ .

# Chapter 4

## Stability and oscillations

If we move from first order systems to second order systems of the form

$$\begin{aligned}\dot{x} &= f(x, y), \\ \dot{y} &= g(x, y),\end{aligned}\tag{4.1}$$

more interesting phenomena can occur. There is indeed a hierarchy of complexity which one ascends as the order of the equation increases. As we saw in the preceding chapter, first order equations have steady state solutions which are alternately stable and unstable, and the instability is direct, in the sense that loss of stability as a parameter changes leads to a transient migration towards another fixed point.

In second order systems, a different kind of instability can occur. As well as the direct instability, or exchange of stability, between different fixed points, oscillatory instability can occur, and the consequence of such instabilities is that permanent oscillatory (periodic) solutions may exist: in two dimensions there is dynamic behaviour.<sup>1</sup>

### 4.1 Linear stability

We illustrate the technique of linear stability analysis for (4.1). The analysis applies (and is easy) in two dimensions, but evidently the method applies in  $n$  dimensions, although the classification of behaviour takes its essence from the two-dimensional example.

A steady state of (4.1) is a constant pair  $(x_0, y_0)$  which satisfies the equations, i. e.,  $f(x_0, y_0) = g(x_0, y_0) = 0$ . For small perturbations about this state, we write

$$x = x_0 + X, \quad y = y_0 + Y,\tag{4.2}$$

and linearise the system (4.1) to obtain the approximate equations

$$\begin{pmatrix} \dot{X} \\ \dot{Y} \end{pmatrix} = \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix},\tag{4.3}$$

---

<sup>1</sup>When one proceeds to third or higher order systems, more exotic behaviour can occur: period-doubling, quasi-periodic oscillations, and chaos, for example.



where the partial derivatives are evaluated at  $(x_0, y_0)$ . The matrix

$$M = \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix} \quad (4.4)$$

is sometimes called the community matrix (particularly in applications in population biology), and its trace and determinant will be denoted

$$T = \text{tr } M, \quad D = \det M. \quad (4.5)$$

Because  $M$  is a constant matrix, solutions of (4.3) are of the form

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \mathbf{u} e^{\lambda t}, \quad (4.6)$$

where  $\mathbf{u}$  is an eigenvector of  $M$  and  $\lambda$  is the corresponding eigenvalue; thus  $\lambda$  is a root of the quadratic equation

$$\lambda^2 - T\lambda + D = 0. \quad (4.7)$$

Classification of the different kinds of behaviour follows from the different combinations of pairs of values of  $\lambda$ . For  $D > T^2/4$ , the roots are complex, and the consequent phase portrait is a spiral; unstable if  $T > 0$ , and stable if  $T < 0$ . Examples are shown in figure 4.1.

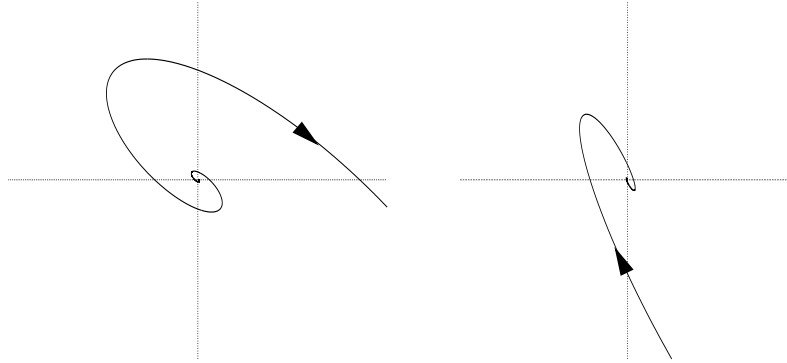


Figure 4.1: Unstable (left) and stable (right) spirals solving (4.3) with  $M = \begin{pmatrix} 1 & 1 \\ -1 & -0.4 \end{pmatrix}$  (left) and  $M = \begin{pmatrix} 1 & 1 \\ -3 & -2 \end{pmatrix}$  (right).

If  $0 < D < T^2/4$ , the fixed point is a node, with two real eigenvalues of  $M$  having the same sign. The node is unstable if  $T > 0$  and stable if  $T < 0$ . An example is shown in figure 4.2.

Finally, a saddle point is shown in figure 4.3. This occurs if  $D < 0$ . The eigenvalues are real and opposite in sign. There is one stable direction and one unstable, and the fixed point is thus unstable. Putting these results together, we see that stability occurs if and only if  $D > 0$  and  $T < 0$ , and the classification of fixed points as spiral, node or saddle can be represented in  $(D, T)$  space as shown in figure 4.4.

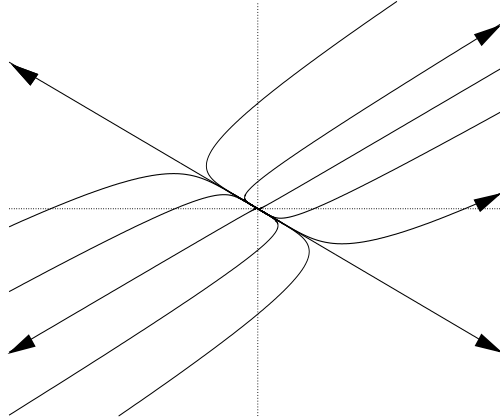


Figure 4.2: Unstable node with  $M = \begin{pmatrix} 1 & 1 \\ 0.5 & 1 \end{pmatrix}$ .

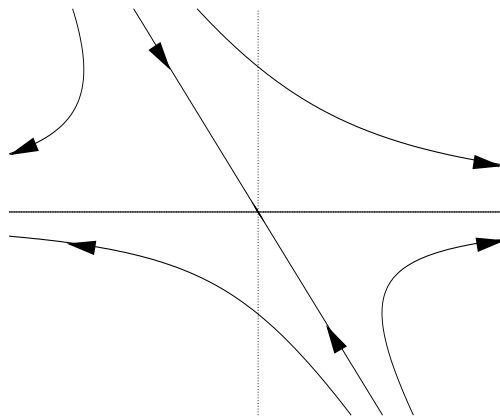


Figure 4.3: Saddle point:  $M = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$ .

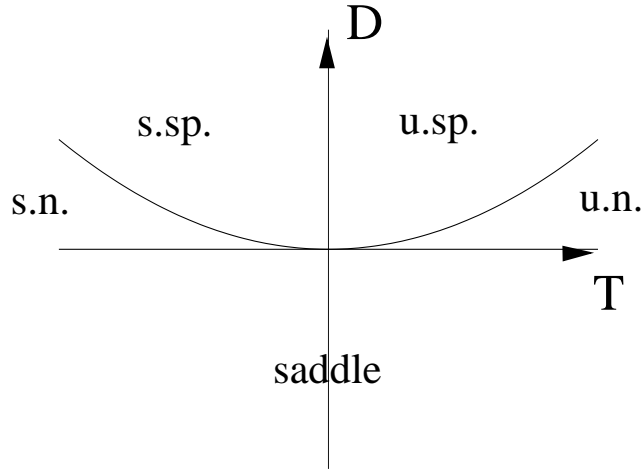


Figure 4.4:  $T$ - $D$  parameter space indicating location of stable and unstable spirals and nodes, and saddles.

## 4.2 Linear systems

It is easy to extend the discussion of two-dimensional systems to systems of  $n$  dimensions. We consider first autonomous systems, i. e., of the form

$$\dot{x} = Ax, \quad (4.8)$$

where  $x \in \mathbf{R}^n$  and  $A$  is an  $n \times n$  constant matrix.<sup>2</sup> Problems of this type arise when studying the stability of steady states of  $n$ -th order ordinary differential equations of the form  $\dot{x} = f(x)$ , when  $A = Df(x_0)$  is the Jacobian of  $f$  at the fixed point  $x_0$ . As before, solutions exist of the form

$$x = e^{\lambda t}u, \quad Au = \lambda u, \quad (4.9)$$

and solutions grow or decay according as  $\text{Re } \lambda \gtrless 0$ .  $x = 0$  is stable if  $\text{Re } \lambda < 0$  for all eigenvalues  $\lambda$ , but is unstable if at least one  $\lambda$  has positive real part.

A succinct representation of the general solution follows from defining the exponential matrix  $e^{tA}$ . This slightly strange looking beast is defined in the obvious way, as

$$e^{tA} = I + tA + \frac{t^2 A^2}{2!} + \dots \quad (4.10)$$

For this definition to make sense, we need to borrow some ideas from linear algebra and functional analysis. We need to define convergence in the space of  $n \times n$  matrices, which requires the definition of an appropriate norm. Matrices being overgrown

---

<sup>2</sup>We are straying into the territory of dynamical systems, in which it is conventional not to use the bold  $\mathbf{x}$  for vectors, and we will follow this convention in this section.

vectors, these are supplied in the usual way. Convergence of the exponential series is then automatic, and indeed the usual exponential laws apply, in particular  $e^M e^N = e^{M+N}$  for matrices  $M$  and  $N$ , and thus also  $(e^{tA})^{-1} = e^{-tA}$ .

The general solution of (4.8) can be written

$$x = e^{tA} x_0, \quad (4.11)$$

where  $x = x_0$  at  $t = 0$ . That this is a solution follows from first principles, since the definition of  $e^{tA}$  implies that  $\frac{d}{dt} e^{tA} = A e^{tA}$ , and Picard's theorem implies it is the only solution.

With the exponential matrix in hand, it is easy to find the solution of the inhomogeneous problem

$$\dot{x} = Ax + g(t). \quad (4.12)$$

Noting that

$$\frac{d}{dt} [e^{-tA} x] = -e^{-tA} Ax + e^{-tA} \dot{x}, \quad (4.13)$$

it quickly follows that

$$x = e^{tA} x_0 + \int_0^t e^{(t-s)A} g(s) ds. \quad (4.14)$$

### 4.2.1 The fundamental solution

We now generalise the preceding discussion to non-autonomous systems of the form (4.8), in which  $A = A(t)$ . Let  $x^{(k)}$  denote the solution of this with

$$x_i^{(k)}(0) = \delta_{ik} = \begin{cases} 1, & i = k, \\ 0, & i \neq k; \end{cases} \quad (4.15)$$

$\delta_{ik}$  is the Kronecker delta. The vectors  $x^{(k)}$  are initially linearly independent, and the matrix whose columns are these vectors is called the fundamental matrix, or fundamental solution,  $\Phi(t)$ , and  $\Phi_{ik} = x_i^{(k)}$ . In particular,

$$\dot{\Phi} = A\Phi, \quad \Phi(0) = I. \quad (4.16)$$

The fundamental matrix plays the rôle of (and, indeed, is) the exponential matrix of an autonomous system. It is easy to show that the solution of the inhomogeneous equation (4.12) is

$$x = \Phi(t)x_0 + \int_0^t \Phi(t)\Phi^{-1}(s)g(s) ds. \quad (4.17)$$

### 4.2.2 Floquet theory

Floquet theory concerns the case where the matrix  $A(t)$  is periodic. Without loss of generality we will suppose that its period is  $2\pi$ . Such systems arise when considering the stability of periodic orbits. Suppose, for example, that the system of ordinary

differential equations  $\dot{x} = f(x)$  has a  $2\pi$ -periodic solution  $x_p(t)$ . To study its stability, we put  $x = x_p + X$ , and linearise the equation for small  $X$ . We obtain

$$\dot{X} = A(t)X, \quad (4.18)$$

where  $A = Df\{x_p(t)\}$  is the Jacobian of  $f$  on the periodic orbit, and is evidently periodic. The question we ask is whether the solutions  $X$  grow or decay in time.

It is not possible to provide such a good recipe as when  $A$  is constant (calculate the eigenvalues of  $A$ ), but thanks to Floquet's theorem we can come close. Again, we let  $\Phi$  be the fundamental solution for  $A$ , thus (4.16) applies. We define the *monodromy operator* to be the matrix

$$M = \Phi(2\pi); \quad (4.19)$$

in the language of dynamical systems it is the Poincaré map on the surface of section  $t \equiv 0 \pmod{2\pi}$ . The difficulty with Floquet theory lies precisely in determining this map. As we shall see, stability is directly determined by the eigenvalues of  $M$ .

We need to show that zero is not an eigenvalue of  $M$ . To do this, we define the Wronskian  $W$  of the system to be the determinant of  $\Phi$ . From first principles, we can deduce Liouville's theorem, that

$$\dot{W} = (\text{tr } A)W, \quad (4.20)$$

and it follows from this that  $W$  is never zero (since  $W = 1$  at  $t = 0$ ). This implies, in particular, that  $\det M \neq 0$ , and since the determinant of a matrix is the product of its eigenvalues, we see that the eigenvalues of  $M$  are non-zero.

If the eigenvalues of  $M$  are  $\mu_s$ , we can therefore define quantities  $\lambda_s$  by

$$\mu_s = e^{2\pi\lambda_s}, \quad (4.21)$$

and we note that  $\text{Re } \lambda > 0$  if and only if  $|\mu| > 1$ . Next we define the diagonal matrix

$$\Lambda = \text{diag}(\lambda_s), \quad (4.22)$$

and we define a matrix  $\Psi(t)$  of fundamental type by

$$\dot{\Psi} = \Lambda\Psi, \quad \Psi(0) = I, \quad (4.23)$$

Now suppose that  $M$  is diagonalisable, as is normally the case, thus there is a matrix  $P$  such that

$$P^{-1}MP = \text{diag}(e^{2\pi\lambda_s}). \quad (4.24)$$

We then have that

$$P\Psi(0) = P = \Phi(0)P, \quad (4.25)$$

and

$$P\Psi(2\pi) = \Phi(2\pi)P, \quad (4.26)$$

and it follows from this that the matrix

$$B(t) = \Phi P \Psi^{-1} P^{-1} \quad (4.27)$$

is  $2\pi$ -periodic. Further, if we define  $X = B(t)Y$ , then

$$\dot{Y} = (P\Lambda P^{-1})Y. \quad (4.28)$$

Therefore the stability of the periodic orbit  $x = x_p(t)$  is determined by the sign of the real part of the eigenvalues  $\lambda_s$  of  $\Lambda$ , or equivalently by whether the characteristic multipliers  $\mu_s$  of  $M$  lie inside or outside the unit circle in the complex plane. Unfortunately, these must in general be computed numerically, since  $M$  must be computed numerically. On the other hand, this is a relatively simple computational task, if the periodic orbit can be computed. Of course, the computation itself then tells us directly whether the periodic orbit is stable or not.

### 4.2.3 Mathieu equation

## 4.3 Nonlinear stability

## 4.4 Phase plane analysis

To go beyond linear stability analysis and complete the phase portrait of solution trajectories is the subject of phase plane analysis, and the most interesting feature of the phase plane is that periodic oscillations can occur. An illuminating example is illustrated in figure 4.5, and is typified by (but is not restricted to) the equations

$$\begin{aligned} \dot{x} &= g(x) - y, \\ \dot{y} &= y - h(x), \end{aligned} \quad (4.29)$$

where the functions  $g$  and  $h$  are as shown in the figure:  $g$  is unimodal (e.g., like  $g = xe^{-x}$ ) and  $h$  is monotonic decreasing (e.g., like  $h = 1/(x - c)$ ). The graphs of  $g(x)$  and  $h(x)$  (and more generally, the curves where  $\dot{x} = 0$  and  $\dot{y} = 0$ ) are called the nullclines of  $x$  and  $y$ , and it is simple to see that where they intersect, there is a steady state solution, and also that in the four regions separated by the nullclines, the trajectories wind round the fixed point in an anti-clockwise manner.

The next issue is whether the fixed point is unstable. If we denote it as  $(x^*, y^*)$ , write  $x = x^* + X$ ,  $y = y^* + Y$ , and linearise for small  $X$  and  $Y$ , then

$$\begin{pmatrix} \dot{X} \\ \dot{Y} \end{pmatrix} \approx \begin{pmatrix} g' & -1 \\ -h' & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}, \quad (4.30)$$

where the derivatives are evaluated at the fixed point. The stability of this two-by-two system with community matrix  $A = \begin{pmatrix} g' & -1 \\ -h' & 1 \end{pmatrix}$  is governed by the trace and determinant of  $A$ , as indicated in figure 4.4. In the present case,  $\text{tr } A = g' + 1$ ,  $\det A = g' - h'$ , so that for the situation shown in figure 4.5, where  $h' < g' < 0$ ,  $\det A > 0$ , and the fixed point is an unstable spiral (or node) if  $g' > -1$ . When  $g' = -1$ , there is a Hopf bifurcation, and if the system has bounded trajectories (as is normal for a model of a physical process) then one expects a stable periodic solution to exist. Figure 4.6 illustrates an example.

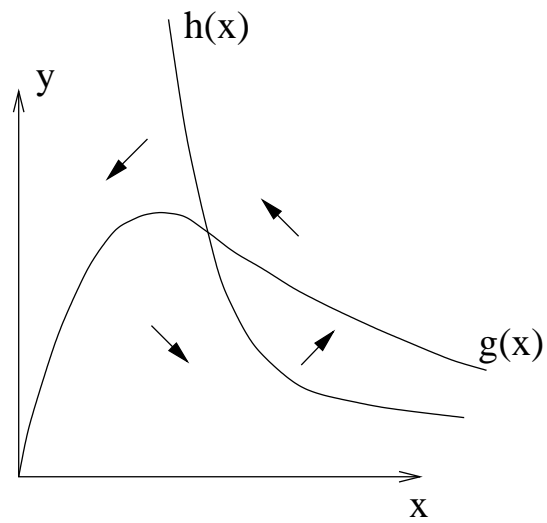


Figure 4.5: Nullclines for (4.29).

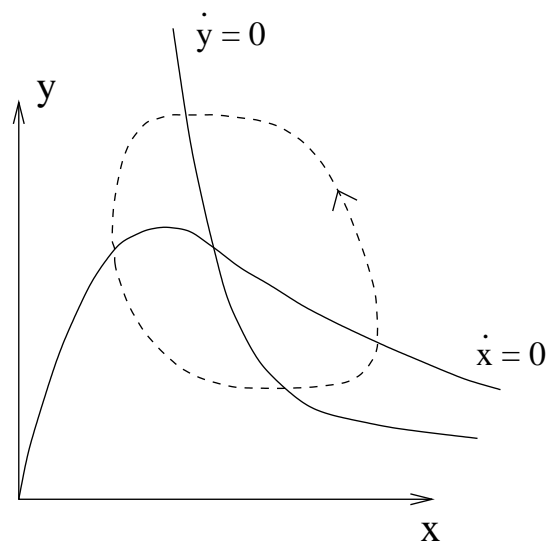


Figure 4.6: Typical form of limit cycle for a system with nullclines as in figure 4.5.

#### 4.4.1 An example of Murray

Murray (2002, p. 116) gives an example of a population interaction model representing aquatic populations of herbivores (e.g., zooplankton) and (phyto-)plankton. Denoting the populations as  $x$  (phytoplankton) and  $y$  (herbivore), the model is

$$\begin{aligned}\dot{x} &= x \left[ k - x - \frac{y}{1+x} \right], \\ \dot{y} &= \alpha y \left[ \frac{x}{1+x} - ay \right],\end{aligned}\tag{4.31}$$

and as normal in population models we are concerned only with the first quadrant  $x \geq 0, y \geq 0$ . We suppose the constants  $k, \alpha$  and  $a$  are positive.

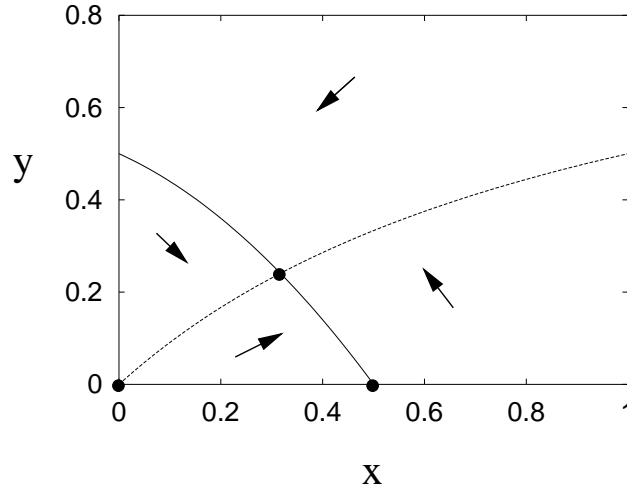


Figure 4.7:  $x$  and  $y$  nullclines for (4.31), with  $a = 1$  and  $k = \frac{1}{2}$ . Arrows indicate the direction of the trajectories in the four sectors bounded by the nullclines.

Define the functions

$$F(x) = (k - x)(1 + x), \quad G(x) = \frac{x}{1 + x}.\tag{4.32}$$

The  $x$  and  $y$  nullclines are  $y = F(x)$  and  $y = G(x)/a$ , respectively, and a typical situation is shown in figure 4.7. There are two fixed points at  $(0, 0)$  and  $(k, 0)$ , and there is always a fixed point  $(x_0, y_0)$  in the positive quadrant. If  $k < 1$  as in the figure, then this fixed point is uniquely defined as there is only one intersection point of the nullclines (in which both  $x$  and  $y$  are positive).

The stability at the origin can be found by expanding for small  $x$  and  $y$ , whence we find

$$\dot{x} \approx kx, \quad \dot{y} \approx \alpha y(x - \alpha y).\tag{4.33}$$

Note that the origin is degenerate in the sense that the linear approximation for  $y$  is just  $\dot{y} \approx 0$ . It is clear from the behaviour of the solutions of (4.33) that the origin is a kind of saddle point. It is similarly easy to show that  $(k, 0)$  is a saddle.



To find the stability of the other fixed point, we write

$$x = x_0 + X, \quad y = y_0 + Y, \quad (4.34)$$

and linearise, so that

$$\begin{pmatrix} \dot{X} \\ \dot{Y} \end{pmatrix} = M \begin{pmatrix} X \\ Y \end{pmatrix}, \quad M = \begin{pmatrix} ay_0 F' & -ay_0 \\ \alpha y_0 G' & -\alpha ay_0 \end{pmatrix}, \quad (4.35)$$

where we use the fact that at  $(x_0, y_0)$ ,  $ay_0 = x_0/(1+x_0)$ ;  $F'$  and  $G'$  are evaluated at  $x_0$ .

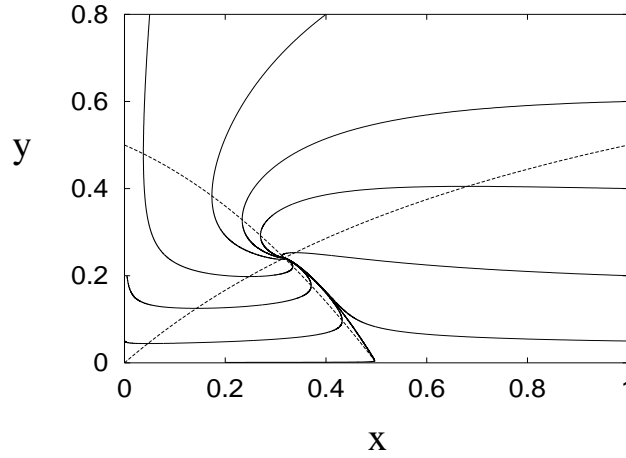


Figure 4.8: Phase plane trajectories for (4.31), with  $a = 1$ ,  $\alpha = 0.4$  and  $k = 0.5$ .

From this we find that

$$\begin{aligned} \text{tr } M &= ay_0(F' - \alpha), \\ \det M &= a\alpha y_0^2[G' - aF']. \end{aligned} \quad (4.36)$$

It is evident from the way the nullclines intersect (since their slopes are just  $G'/a$  and  $F'$ ) that for the situation shown in figure 4.7,  $\det M > 0$ . Therefore stability is associated with the sign of  $\text{tr } M$ , and in particular there is a Hopf bifurcation to a periodic orbit if  $F' > \alpha$ . For the situation of figure 4.7, where  $k < 1$  and thus  $F' < 0$ , we see that this fixed point is a stable node or spiral. Figure 4.8 shows this behaviour.

The possibility of instability and the consequent occurrence of a periodic oscillation occurs if  $F'(x_0) > 0$ . Since  $x_0$  depends only on  $a$  and  $k$ , we see that in this case, there will be instability if  $\alpha$  is small enough. The question then arises, for what  $a$  and  $k$  is  $F'(x_0) > 0$ ? Since  $F'(x_0) = k - 1 - 2x_0$ , we need  $x_0 < \frac{k-1}{2}$  (remember we also must have  $k > 1$ ). Since  $F(x_0) = G(x_0)/a$  and  $G$  is monotonically increasing, we thus also have

$$F(x_0) = \frac{G(x_0)}{a} < \frac{1}{a} \left( \frac{k-1}{k+1} \right), \quad (4.37)$$

and since  $F(x_0) < F(k-1) = \frac{1}{4}(k+1)^2$ , we see that a necessary condition for instability to occur is that

$$0 < a < \frac{4(k-1)}{(k+1)^3}. \quad (4.38)$$

When  $k$  is close to 1, the upper boundary of this region becomes a good estimate of the true upper limit. Figure 4.9 shows an example of periodic behaviour when  $a = 0.05$ ,  $\alpha = 0.9$ , and  $k = 3$ .

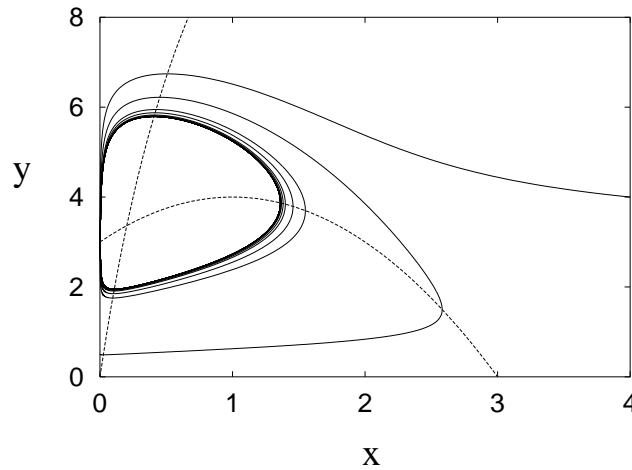


Figure 4.9: A stable periodic solution when  $a = 0.05$ ,  $\alpha = 0.9$ , and  $k = 3$ . The nullclines are shown as the dashed lines.

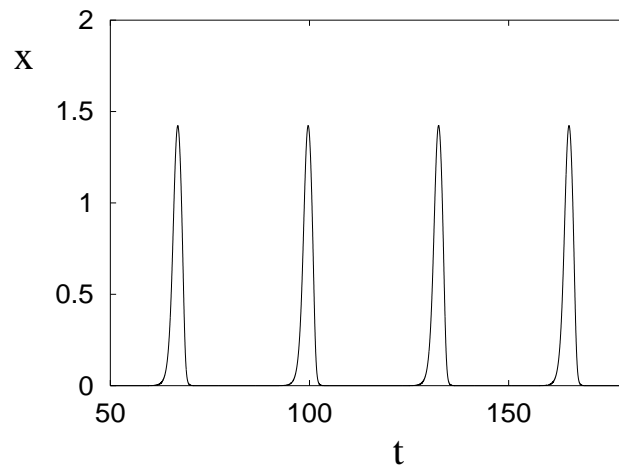


Figure 4.10: Relaxational solution of (4.31) when  $a = 0.02$ ,  $\alpha = 0.9$ , and  $k = 3$ . The graph is that of  $x(t)$ .

It is evident from this phase portrait that the limit cycle has a relaxational character, in the sense that there is a (longish) period when  $x$  is very small. Of the

parameters used in figure 4.9, we see that  $a$  is small, suggesting that the limit  $a \rightarrow 0$  is responsible for this behaviour. This is confirmed in figure 4.10, which shows the time sequence of  $x(t)$  when  $a = 0.02$ ,  $\alpha = 0.9$ , and  $k = 3$ . It is not uncommon to see near extinction oscillations in population models. It is possible, though not entirely straightforward, to analyse such oscillations using singular perturbation methods.

### Multiple steady states

There is always at least one steady state with  $y > 0$ , but inspection of  $F(x)$  and  $G(x)$  shows that there can be more than one. Figure 4.11 shows an example. Preceding discussion indicates that if we label the three equilibria in this case as  $(x_i, y_i)$ , with  $i = 0, 1, 2$  and  $x_0 < x_1 < x_2$ , then the middle steady state  $(x_1, y_1)$  is always an unstable saddle ( $\det M < 0$ ), the upper steady state  $(x_2, y_2)$  is always stable ( $\det M > 0$  and  $\text{tr } M < 0$ ), and the stability of  $(x_0, y_0)$  is determined as before.

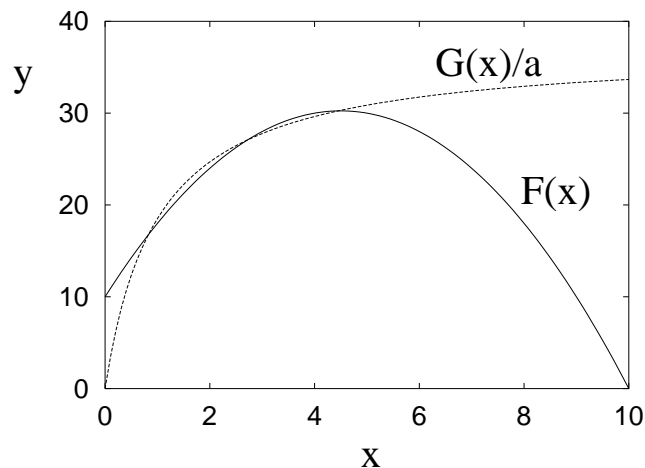


Figure 4.11: Nullclines of  $x$  and  $y$  when  $k = 10$  and  $a = 0.027$ .

It is clear from figure 4.11 that, since  $F$  increases with  $k$ , there will be an interval of  $k$ -values,  $k_- < k < k_+$ , for which multiple steady states occur. The values  $k_{\pm}$  can be evaluated by the condition that there be a double root of  $F(x) = G(x)/a$ , and this leads to the parametric prescription for  $k_{\pm}$  as

$$k = \frac{2x^2}{x-1}, \quad a = \frac{x-1}{(1+x)^3}, \quad (4.39)$$

which describes a narrow sector in the  $(k, a)$  plane, as shown in figure 4.12. The upper and lower boundaries of this sector are given, from (4.39), by

$$a \sim \frac{1}{4k}, \quad a \sim \frac{4}{k^2} \quad \text{as } k \rightarrow \infty. \quad (4.40)$$

Figure 4.12 suggests that oscillations do not occur when there are multiple steady states.

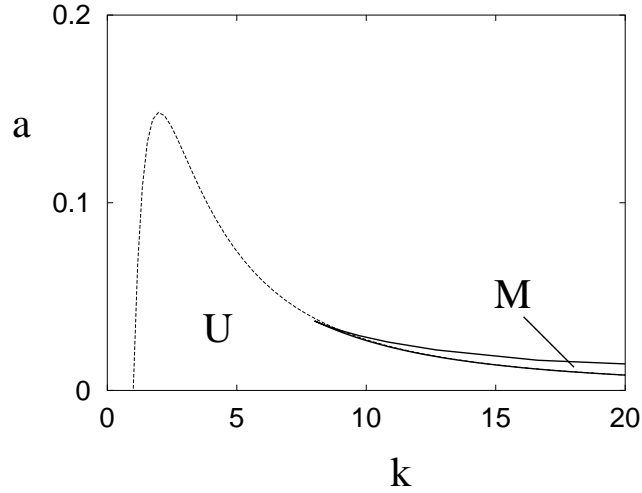


Figure 4.12: The regions in which instability of  $(x_0, y_0)$  may occur (U) and in which there are multiple steady states (M). The figure suggests (but does not prove) that when multiple steady states occur, both the upper and lower branches are stable.

#### 4.4.2 Travelling waves on the Martian north polar ice cap

Phase plane analyses often arise when one seeks travelling wave solutions of partial differential equations, as will be discussed further in chapter 7. An example of this approach occurs in the study of a model describing accretion and sublimation of ice on the north polar ice cap of Mars. As can be seen in figure 4.13, a series of arcuate canyons occur on this ice cap, marked by dust deposits on the steep scarps. The ice cap is some three kilometres deep and 1000 kilometres wide, while the canyons are of typical depths 500 m, and spaced some 50 km apart. Their occurrence is mysterious, and unlike anything seen on the ice sheets of Earth.

One model which has been studied with a view to explaining them monitors the effect of atmospheric dust concentration on sublimation rate. To explain the spiral wave like form of the troughs, travelling wave solutions of the model are sought, and these are described by

$$\begin{aligned} s' &= vs - q, \\ uc' &= \phi q, \end{aligned} \tag{4.41}$$

in which  $q = q(c, s)$ ;  $s$  represents surface slope,  $q$  is the net rate of sublimation, and  $c$  is the atmospheric dust concentration. The prime denotes a spatial derivative in the moving frame of the travelling wave.  $v$  denotes the travelling wave speed, while  $u$  is the atmospheric wind speed, and  $\phi$  is the dust fraction of the ice, and these are all taken as constants;  $u$  and  $\phi$  are positive, but  $v$  is unknown.

The sublimation function  $q$  is taken to be

$$q = f(c) + ms, \tag{4.42}$$

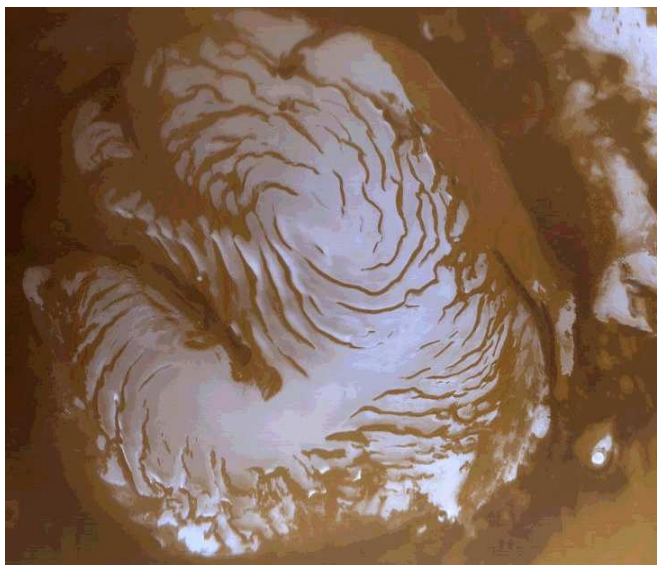


Figure 4.13: The spiral canyons of the Martian north polar ice cap.

with  $m$  being positive and  $f(c)$  having a typical cubic shape exemplified by

$$f(c) = (c - 1)(c - 2) \left(1 - \frac{c}{M}\right). \quad (4.43)$$

The model has three steady states where  $s = 0$ , and  $c = 1$ ,  $c = 2$  or  $c = M$ ; we denote these steady states as  $P$ ,  $Q$ ,  $R$ , respectively, as shown in figure 4.14. The results of a phase plane analysis are then the following. When  $0 < v < m$ ,  $P$  and  $R$  are saddles, while  $Q$  is a node or spiral, which undergoes a Hopf bifurcation at

$$v = v_c = m - \frac{\phi f'}{u}, \quad (4.44)$$

assuming  $v_c$  is positive.  $Q$  is unstable for  $v > v_c$ , and numerical computations suggest that the Hopf bifurcation is supercritical, so that there is a stable periodic solution for  $v > v_c$ . This orbit exists up to a value  $v = v_h$ , at which it becomes heteroclinic to  $P$ , and disappears in a ‘blue skies’ bifurcation. The stable orbit and its disappearance are shown in figures 4.14 and 4.15.

If  $v > m$ ,  $P$  and  $R$  are still saddles, and  $Q$  is an unstable node or spiral. It is possible that the limit cycle may still exist, and this probably depends on the size of  $f$ ; it is also possible for the limit cycle to disappear by coalescence with  $R$ , but this seems unlikely for large  $M$ . In fact, it is found that as  $M \rightarrow \infty$  in (4.43), the limit cycle remains finite and independent of  $M$ .

## 4.5 Relaxation oscillations

It is a general precept of the applied mathematician that there are three kinds of numbers: small, large, and of order one. And the chances of a number being  $O(1)$  are

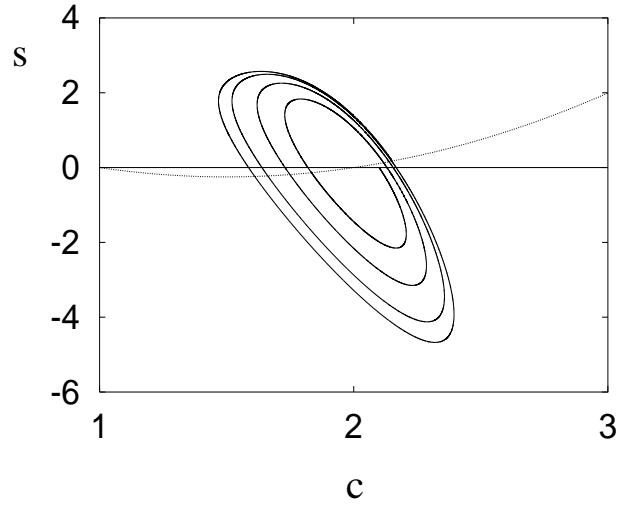


Figure 4.14: Approach to stable limit cycle solution of (4.41) using (4.43) with  $M = 1000$ , and  $v = 0.05$ ,  $\phi = 0.2$ ,  $m = 0.14$ , and  $u = 2$ . The parabolic curve is the nullcline  $q = 0$ .

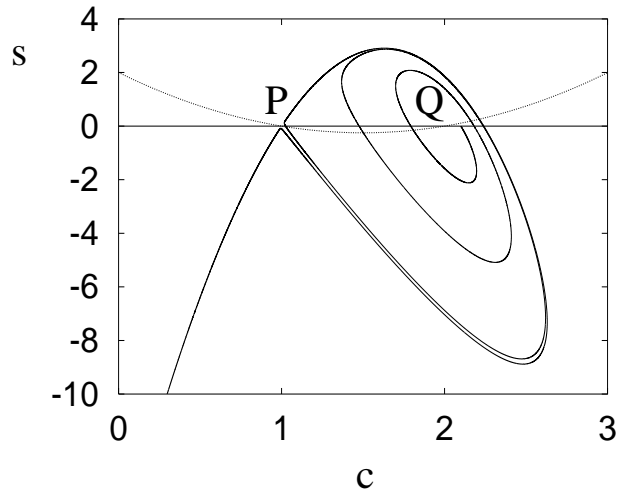


Figure 4.15: Disappearance of limit cycle in heteroclinic connection with  $P$ . Parameters as for figure 4.15, except that  $v = 0.06$ .

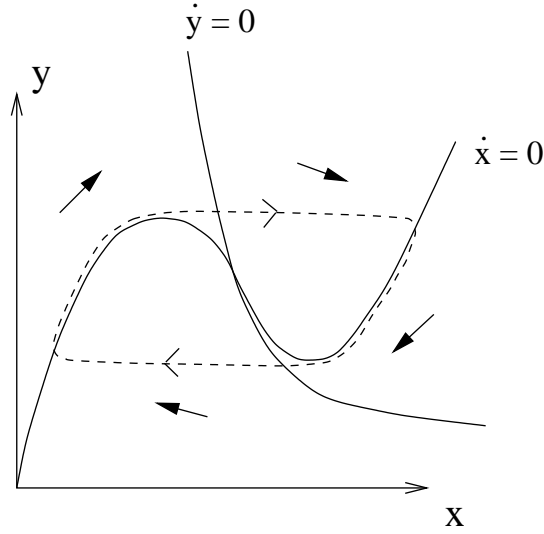


Figure 4.16: Typical form of relaxation oscillation in phase plane for (4.45).

not great. Thus for systems of the form (4.1), it is often the case in practice that the time scales for each equation are different, so that in suitable dimensionless units, a second order system might take the form

$$\begin{aligned}\varepsilon \dot{x} &= y - g(x), \\ \dot{y} &= h(x) - y,\end{aligned}\tag{4.45}$$

where the parameter  $\varepsilon$  is small. The system (4.45) is essentially the same as (4.29) with time reversed, but now suppose that the nullclines are as shown in figure 4.16, i.e.  $g$  has a cubic shape. Trajectories now rotate clockwise, and linearisation about the fixed point yields a community matrix  $A$  with  $\text{tr } A = -(g'/\varepsilon) - 1$ ,  $\det A = (g' - h')/\varepsilon$ , thus with  $g' > h'$ , the fixed point is a spiral or node, and with  $\varepsilon \ll 1$ ,  $\text{tr } A \approx -g'/\varepsilon > 0$ , so it is unstable. Thus we expect a limit cycle, and because  $\varepsilon \ll 1$ , this takes the form of a *relaxation oscillation* in which the trajectory jumps rapidly backwards and forwards between branches of the  $x$  nullcline. For  $\varepsilon \ll 1$ ,  $x$  rapidly jumps to its quasi-equilibrium  $y \approx g(x)$ , and then  $y$  migrates slowly ( $\dot{x} \approx [h(x) - g(x)]/g'(x)$ ) until  $g' = 0$  and  $x$  jumps rapidly to the other branch of  $g$ . Figure 4.17 shows the time series of the resulting oscillation. The motion is called ‘relaxational’ because the fast variable  $x$  ‘relaxes’ rapidly to a quasi-stationary state after each transient excursion.

### 4.5.1 The Van der Pol oscillator

The relaxational form of the van der Pol oscillator is

$$\varepsilon \ddot{x} + (x^2 - 1)\dot{x} + x = 0, \quad \varepsilon \ll 1.\tag{4.46}$$

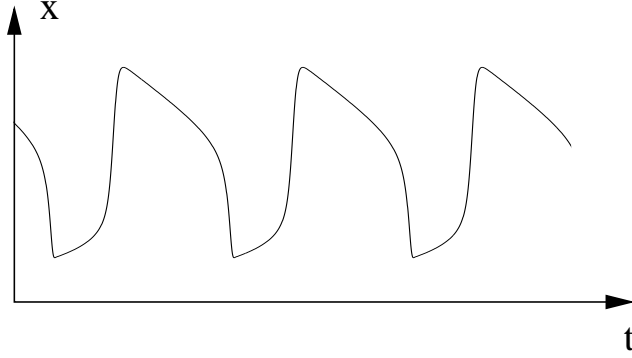


Figure 4.17: Time series for  $x$  corresponding to figure 4.16.

This form of the equation is appropriate when the ‘damping’ (first derivative) term is large. If we define

$$y = \varepsilon \dot{x} + \frac{1}{3}x^3 - x, \quad (4.47)$$

then the appropriate phase plane equations are

$$\begin{aligned} \varepsilon \dot{x} &= y - S(x), \\ \dot{y} &= -x, \end{aligned} \quad (4.48)$$

where

$$S(x) = \frac{1}{3}x^3 - x. \quad (4.49)$$

The  $x$ -nullcline is the cubic  $y = S(x)$ ; above this, trajectories move rapidly to the right, and below it they move rapidly to the left. Thus, from an arbitrary initial location, a trajectory will rapidly approach either upward sloping part of the *slow manifold*  $y = S(x)$ , on which the motion is slow. On the right side, in  $x > 1$ ,  $y$  decreases until it reaches the base of the curve at  $x = 1$ , where it undergoes a rapid transition (an interior layer, in fact) to the left side of the curve in  $x < -1$ . There, it increases until a symmetric transition occurs at  $x = -1$ . In this way, a periodic oscillation occurs. The typical form of the trajectories is shown in figure 4.18, which shows that the relaxational form is already appropriate when  $\varepsilon = 0.1$ . The time series of  $x$  versus  $t$  is very similar to figure 4.17.

It is possible to derive characteristics of the oscillation very simply from this description. The maximum amplitude is  $x = 2$ , following which there is a slow decline to  $x = 1$ . Because of the symmetry of the oscillation, the time to get from  $x = 2$  to  $x = 1$  is approximately  $P/2$ , where  $P$  is the period (since the fast transitions between are very short). Thus the period is

$$P \approx 2 \int_{x=2}^{x=1} dt \approx \int_1^2 \frac{S'(x) dx}{x} \approx 3 - 2 \ln 2. \quad (4.50)$$



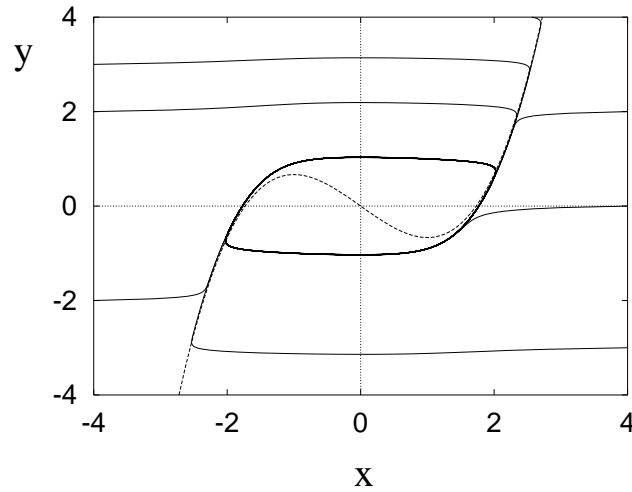


Figure 4.18: Trajectories of (4.46) approaching the limit cycle (thick line) when  $\varepsilon = 0.1$ . Also indicated is the nullcline for  $x$  given by  $y = S(x)$ .

This follows because on the slow manifold,  $dt \approx -dy/x$  and  $y \approx S(x)$ .

## 4.6 Belousov-Zhabotinskii reaction

## 4.7 Notes and references

Phase plane methods are described in many books on ordinary differential equations, for example that of Jordan and Smith (1999).

The blue skies catastrophe and various other bifurcations of exotic type are described in a number of books on dynamical systems, for example that of Thompson and Stewart (1986).

The model of formation of the spiral troughs on the Mars north polar ice cap is the subject of a dissertation by Zammett (2004).

Linear systems of ordinary differential equations are the stuff of many basic textbooks on the subject. A particularly clear and direct discussion is that by Arn'old (1973).

## Exercises

4.1  $u$  and  $v$  satisfy the ordinary differential equations

$$\begin{aligned} \dot{u} &= k_1 - k_2 u + k_3 u^2 v, \\ \dot{v} &= k_4 - k_3 u^2 v, \end{aligned}$$

where  $k_i > 0$ . By suitably scaling the equations, show that these can be written in the dimensionless form

$$\begin{aligned}\dot{u} &= a - u + u^2v, \\ \dot{v} &= b - u^2v,\end{aligned}$$

where  $a$  and  $b$  should be defined. Show that if  $u, v$  are initially positive, they remain so. Draw the nullclines in the positive quadrant, show that there is a unique steady state and examine its stability. Are periodic solutions likely to exist?

4.2 Suppose that

$$\dot{X} = A(t)X, \quad (*)$$

where  $X \in \mathbf{R}^n$  and  $A$  is  $2\pi$ -periodic. Define what is meant by the fundamental matrix and write down the solution of  $(*)$  satisfying  $X = X_0$  at  $t = 0$ .

Define what is meant by the monodromy matrix  $M$ , and show that it has no zero eigenvalues. Describe how it can be used to prove Floquet's theorem, that there is a  $2\pi$ -periodic matrix  $B(t)$  such that  $X = BY$ , and

$$\dot{Y} = CY,$$

where  $C$  is a constant matrix. Explain how the eigenvalues of  $C$  are related to those of  $M$ .

4.3 Use a phase plane analysis (with  $y = \dot{x}$ ) to analyse the solutions of the nonlinear oscillator

$$\ddot{x} + V'(x) = 0,$$

where

$$V(x) = \frac{1}{2}x^2 - \frac{1}{4}x^4.$$

In particular, show that the origin is genuinely a centre.

Construct the phase planes for the systems

$$(i) \quad \dot{x} = y, \quad \dot{y} = -V'(x) - \varepsilon y,$$

$$(ii) \quad \dot{x} = y, \quad \dot{y} = -V'(x) - \varepsilon y^3,$$

where  $\varepsilon$  is small.

Show that if

$$\dot{x} = y, \quad \dot{y} = -V'(x) - \varepsilon D(E)y,$$

where

$$E = \frac{1}{2}y^2 + V(x), \quad D(E) = E - \frac{1}{8},$$

then there is a stable limit cycle in the region  $0 < E < \frac{1}{4}$ , and draw the corresponding phase plane.

- 4.4 An energy balance model for the average atmospheric (absolute) temperature  $T$  of the Earth is given by

$$\rho c_p h \dot{T} = \frac{1}{4} Q (1 - a) - \sigma \gamma T^4,$$

where  $\rho$  is air density,  $c_p$  is the specific heat,  $h$  is the effective height of the atmosphere,  $Q$  is the received short wave solar radiation,  $a$  is the albedo coefficient (reflectivity) of the Earth's surface,  $\sigma$  is the Stefan-Boltzmann constant, and  $\gamma < 1$  is a numerical factor which represents the blanketing (greenhouse) effect of the atmosphere.

Non-dimensionalise the equation using a temperature scale  $T_0$  and a time scale  $t_0$ , so that it takes the dimensionless form

$$\dot{T} = 1 - a - T^4.$$

Show that if the observed albedo is 0.3 and the observed (steady state) temperature is 288 K, then  $T_0 = 315$  K.

Use the values  $Q = 1370 \text{ W m}^{-2}$  and  $\sigma = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$  to find what the greenhouse factor  $\gamma$  must be to obtain this value of  $T_0$ .

Use this value of  $\gamma$ , and the values  $\rho = 1 \text{ kg m}^{-3}$ ,  $c_p = 10^3 \text{ J kg}^{-1} \text{ K}^{-1}$ ,  $h = 10^4 \text{ m}$ , to compute the adjustment time scale  $t_0$ .

The albedo of the Earth varies principally due to the extent of sea and land ice, which in itself depends on temperature. Sea ice varies seasonally, but land-based ice sheets grow over time scales of thousands of years. Suppose that as a consequence the albedo adjusts towards an equilibrium value  $A(T)$  on a time scale  $t_i \sim 1000 \text{ y}$  (years), and is modelled by the equation

$$t_i \dot{a} = A(T) - a.$$

Show that the non-dimensional form of this equation is

$$\dot{a} = \nu [A(T) - a],$$

where  $\nu = t_0/t_i$ .

We can expect  $A(T)$  to be a monotonically decreasing function, with, say,  $A \approx 0.8$  for very low  $T$ , and  $A \approx 0.2$  for very high  $T$ . If  $A$  changes rapidly enough with  $T$  in an appropriate range, show that it is possible to have three steady states, and construct the phase plane for  $(T, a)$  in this case. Hence show that the smallest and largest steady states of  $T$  are stable, and the intermediate state is unstable.

- 4.5 The sizes of two populations, one being a predator and the other its prey, are described by the equations

$$\begin{aligned} \dot{u} &= u(1 - u) - \frac{auv}{u + d}, \\ \dot{v} &= bv \left( 1 - \frac{v}{u} \right), \end{aligned}$$

where  $a$ ,  $b$  and  $d$  are positive constants.

Show that if  $u$  and  $v$  are non-negative, then they remain so.

Show that there is a unique steady state in the first quadrant, and show that there is a region in  $(d, a, b)$  space where this steady state is unstable.

Draw the possible form(s) of the first quadrant phase plane in this case.

4.6 A nonlinear oscillator is given by the equation

$$\varepsilon \ddot{x} + (x^4 - 1)\dot{x} + x = 0, \quad \varepsilon \ll 1.$$

A suitable phase plane is spanned by  $(x, y)$ , where  $y = \varepsilon \dot{x} + \frac{1}{5}x^5 - x$ . Describe the motion in this phase plane, and find, approximately, the period of the relaxation oscillation. What happens if  $\varepsilon < 0$ ?

4.7 The Belousov-Zhabotinskii chemical reaction can be approximately described by the two component pair of ordinary differential equations

$$\varepsilon \dot{X} = X(1 - X) - \left( \frac{X - \delta}{X + \delta} \right) Z,$$

$$\dot{Z} = \gamma X - Z,$$

where  $\varepsilon$  and  $\delta$  are very small, and  $\gamma$  is  $O(1)$ . Show that relaxation oscillations will occur for  $\gamma$  within a certain range  $(\gamma_-, \gamma_+)$ , and give approximations for the values of  $\gamma_{\pm}$ .

# Chapter 5

## Oscillation and comparison methods

The study of the stability of fixed points of systems of ordinary differential equations leads to the computation of the growth rate exponents  $\sigma$  as eigenvalues of matrices. In a similar way, the computation of the frequencies of small oscillations of Hamiltonian systems leads to algebraic eigenvalue problems.

Both of these situations arise because of the underlying linearity of the system, which allows the possibility of finding separable solutions. In *partial* differential equations, the same situation arises, but now the computation of eigenvalues requires the solution of differential, rather than algebraic, equations, and the corresponding solutions of these are called eigenfunctions (as opposed to eigenvectors). In this chapter we discuss properties of these eigenvalue problems.

### 5.1 Derivation of Sturm-Liouville systems

We begin by giving some examples of eigenvalue problems which arise through the technique of separation of variables in partial differential equations. All of these examples take the form of (second order) Sturm-Liouville systems, that is, ordinary differential equations for  $y(x)$  of the form

$$(py')' + (\lambda r - q)y = 0, \quad (5.1)$$

together with suitable homogeneous boundary conditions of the general form

$$y' + \alpha y = 0 \quad \text{at} \quad x = a, b. \quad (5.2)$$

In this chapter we will mostly take  $\alpha = \infty$ , thus

$$y = 0 \quad \text{at} \quad x = a, b. \quad (5.3)$$

Sturm-Liouville systems possess the functional analytic property of being *self-adjoint* (the definition of which will appear in due course). Indeed, our discussion of Sturm-Liouville systems will frequently skirmish with concepts from functional

analysis, but the material will be presented more in the manner of a tour, and will not dwell overmuch on the niceties of detailed proofs. For these, reference must be made to any number of classical textbooks (see section 5.4).

### 5.1.1 Waves in a bucket

The first example we consider is that of free surface fluid flow in a cylindrical container: waves in a bucket, or a petri dish, or a pond. We suppose that the fluid, when undisturbed, occupies the region  $-h < z < 0$ ,  $0 < r < a$ , where  $r$ ,  $z$  (and  $\theta$ ) are cylindrical coordinates. When the fluid is disturbed, the upper (free) surface is perturbed to a surface  $z = \eta(r, \theta, t)$ , whose location forms part of the problem to be solved: it is therefore a free boundary problem.

For irrotational flow, there is a velocity potential  $\phi$ , which satisfies Laplace's equation

$$\nabla^2 \phi = \phi_{rr} + \frac{1}{r} \phi_r + \frac{1}{r^2} \phi_{\theta\theta} + \phi_{zz} = 0 \quad (5.4)$$

together with the boundary conditions of no flow through the base and the sides:

$$\begin{aligned} \phi_r &= 0 & \text{on } r &= a, \\ \phi_z &= 0 & \text{on } z &= -h. \end{aligned} \quad (5.5)$$

On the top surface, the kinematic and Bernoulli conditions imply

$$\phi_t + g\eta = 0, \quad \phi_z = \eta_t \quad \text{on } z = 0, \quad (5.6)$$

where it is assumed that the free surface elevation  $\eta$  is small, so that the boundary conditions can be linearised about  $z = 0$ . In all these equations, subscripts indicate partial differentiation.

The equation and boundary conditions are linear, and are autonomous in  $t$ ,  $\theta$  and  $z$  (though not  $r$ ), so that separable solutions can be found. These take the form

$$\phi = R(r)e^{i\omega t + im\theta} \cosh\{k(z+h)\}, \quad \eta = A(r)e^{i\omega t + im\theta}, \quad (5.7)$$

where, because of the linearity, we allow  $R$  and  $A$  to be complex, and it is understood that we may take the real part of the solutions. Note that we require  $m$  to be an integer, so that the solutions are well-defined.

Substitution of the expressions in (5.7) into (5.4)–(5.6) leads to the ordinary differential equation for  $R$ ,

$$R'' + \frac{1}{r}R' + \left(k^2 - \frac{m^2}{r^2}\right)R = 0, \quad (5.8)$$

where  $k$  is a real constant, which is in essence the wavenumber. Satisfaction of the surface boundary conditions leads to the *dispersion relation*

$$\omega^2 = gk \tanh kh, \quad (5.9)$$

which relates wave speed  $\omega/k$  to wavenumber, but is not of immediate concern here.

The equation (5.8) can be written in the form

$$(rR')' + \left(k^2r - \frac{m^2}{r}\right)R = 0, \quad (5.10)$$

and as such represents a Sturm-Liouville system of the form (5.1) (do not confuse the function  $r(x)$  in (5.1) with the polar radius  $r$  in (5.10)).

The question of appropriate boundary conditions for (5.10) is an interesting one. We require

$$R(a) = 0 \quad (5.11)$$

to satisfy the no flow through condition on the tube wall, but the other condition required is more subtle. Since  $\phi$  satisfies Laplace's equation, we require  $\nabla\phi = \left(\phi_r, \frac{1}{r}\phi_\theta, \phi_z\right)$  to be regular, and in particular, bounded. It follows from this that we require

$$\begin{aligned} R(0) &= 0, & m &\neq 0, \\ R'(0) &= 0, & m &= 0. \end{aligned} \quad (5.12)$$

The solutions of (5.8) are Bessel functions  $R = J_m(kr)$ , to which we will return in due course.

### 5.1.2 Laplace's equation

The second example we consider is that of solving Laplace's equation in spherical polars. This has a variety of applications, the most obvious of which are associated with properties of the Earth. For example, deviations of the Earth's gravity field from a purely radial one are usually expressed in terms of spherical harmonics, which are essentially the separable solutions of Laplace's equation in spherical coordinates.

We suppose

$$\nabla^2\psi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2} = 0, \quad (5.13)$$

where  $(r, \theta, \phi)$  are spherical polar coordinates. Separable solutions exist in the form

$$\psi = r^n e^{im\phi} \Theta(\theta), \quad (5.14)$$

where we require  $m$  to be an integer in order that the solution be single-valued. Substituting (5.14) into (5.13), we find that  $\Theta$  satisfies the ordinary differential equation

$$n(n+1)\Theta + \frac{1}{\sin\theta}[\sin\theta\Theta']' - \frac{m^2}{\sin^2\theta}\Theta = 0. \quad (5.15)$$

It is convenient to write  $\mu = \cos\theta$ , so that  $\frac{d}{d\mu} = -\frac{1}{\sin\theta}\frac{d}{d\theta}$ , and then

$$\left[(1-\mu^2)\Theta'\right]' + \left[n(n+1) - \frac{m^2}{1-\mu^2}\right]\Theta = 0. \quad (5.16)$$

This is Legendre's equation. The solutions are usually denoted as  $P_n^m(\mu)$ ,  $Q_n^m(\mu)$ ; when  $m = 0$ ,  $P_n^0 = P_n$  are the Legendre polynomials, if  $n$  is an integer.

We see that Legendre's equation is of Sturm-Liouville type, with  $\lambda = n(n+1)$ . In spherical polars,  $\nabla\psi = \left(\psi_r, \frac{1}{r}\psi_\theta, \frac{1}{r\sin\theta}\psi_\phi\right)$ , and thus the requirement that  $\nabla\psi$  be bounded implies that  $n = 0$  or  $n \geq 1$ , if the solution domain encloses the origin. If it encloses the poles at  $\mu = 1$  and  $\mu = -1$ , then we require

$$\begin{aligned}\Theta(\pm 1) &= 0 & \text{if } m = 0, \\ \Theta'(\pm 1) &= 0 & \text{if } m \neq 0.\end{aligned}\tag{5.17}$$

### 5.1.3 Instability in thermally activated shear flow

A classical problem of some interest and relevance is that of thermal runaway. In its origins, this problem concerns combustion (why does a match light when you strike it?), but a different application occurs in the flow of fluids with temperature-dependent viscosity. The process of injection moulding (of, for example, telephones) fills a mould with a molten polymer, which then solidifies to form the manufactured object. If the flow rate of the polymer is high enough, charring can occur at the edges, which is evidently not desirable!<sup>1</sup>

The mechanism for this charring is that there is frictional heat released by the viscous flow of the melt, and this depends on the temperature, through the dependence of viscosity on temperature. The heat released increases the temperature, which further increases the heat released. This kind of circular reinforcement is called a *positive feedback*, and provides a mechanism for instability and, in some cases, what is called *blow up*: this is discussed further in chapter 9.

A simple model which represents the injection moulding process is that of an axial flow of an incompressible fluid in a cylindrical tube. We use cylindrical polar coordinates  $(r, \theta, z)$ , and suppose that the tube wall is at  $r = a$ . If there is a purely axial fully developed velocity  $u(r, t)$ , then the only stresses are a shear stress  $\tau$  and the pressure  $p$ , and these satisfy the axial component of the momentum equation, which is

$$\frac{1}{r} \frac{\partial}{\partial r}(r\tau) = \frac{\partial p}{\partial z}.\tag{5.18}$$

Because of the purely radially dependent velocity, mass conservation (incompressibility) is identically satisfied, and the acceleration terms in the momentum equation are identically zero. This latter situation is also approximately true for low Reynolds number (slow) flow.

The viscous flow law is

$$\tau = \mu \frac{\partial u}{\partial r},\tag{5.19}$$

where  $\mu$  is the viscosity. We assume that  $\mu = \mu(T)$ , where  $T$  is the temperature, whose evolution must therefore also be determined. The energy equation takes the

---

<sup>1</sup>This was the subject of an industrial problem brought to Cornell University in 1975 by the Bell telephone company.



form (assuming axial symmetry)

$$\rho c_p \left[ \frac{\partial T}{\partial t} + u \frac{\partial T}{\partial z} \right] = k \left[ \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial T}{\partial r} \right) + \frac{\partial^2 T}{\partial z^2} \right] + \frac{\tau^2}{\mu}, \quad (5.20)$$

where the terms represent successively heat advection, heat conduction and viscous dissipative heating:  $\rho$  is the fluid density,  $c_p$  its specific heat, and  $k$  is the thermal conductivity. The last term in (5.20) is normally completely insignificant in ordinary laboratory scale fluid flows, but is important in the flow of molten polymers and plastics, and also in some larger geophysical flows.

If we suppose that the temperature field is also fully developed, i. e.,  $T = T(r, t)$ , then the  $z$  derivative terms in (5.20) may be ignored. From the momentum equations it follows that  $p' \equiv \frac{\partial p}{\partial z}$  is constant (since also radial momentum implies  $\frac{\partial p}{\partial r} = 0$ ), and therefore

$$\tau = \frac{1}{2} r p', \quad u = -\frac{1}{2} p' \int_r^1 \frac{r \, dr}{\mu}, \quad (5.21)$$

and the temperature satisfies the equation

$$\rho c_p \frac{\partial T}{\partial t} = \frac{k}{r} \frac{\partial}{\partial r} \left( r \frac{\partial T}{\partial r} \right) + \frac{r^2 p'^2}{4\mu}, \quad (5.22)$$

and is evidently uncoupled from the determination of the velocity field. We suppose temperature is prescribed on the tube wall, thus

$$T = T_0 \quad \text{on} \quad r = a. \quad (5.23)$$

### Non-dimensionalisation

We suppose that  $\mu_0$  is a typical value of the viscosity, and  $\Delta T$  is the temperature over which it varies significantly; generally  $\mu$  decreases with  $T$  (think of maple syrup), and to be specific we will suppose

$$\mu = \mu_0 \exp \left[ -\frac{(T - T_0)}{\Delta T} \right]. \quad (5.24)$$

We non-dimensionalise the temperature equation and boundary conditions by writing

$$r = a\xi, \quad T = T_0 + (\Delta T)\theta, \quad t = \frac{a^2 \tau}{\kappa}, \quad (5.25)$$

where  $\kappa = k/\rho c_p$  is the thermal diffusivity. This leads to the dimensionless differential equation

$$\frac{\partial \theta}{\partial t} = \frac{1}{\xi} \frac{\partial}{\partial \xi} \left[ \xi \frac{\partial \theta}{\partial \xi} \right] + \alpha \xi^2 e^\theta, \quad (5.26)$$

with the boundary conditions

$$\begin{aligned} \theta &= 0 \quad \text{on} \quad r = 1, \\ \theta_r &= 0 \quad \text{at} \quad r = 0, \end{aligned} \quad (5.27)$$

this last condition following from the requirement that  $\theta$  be bounded at  $r = 0$ . The single dimensionless parameter  $\alpha$  is given by

$$\alpha = \frac{a^4 p'^2}{4\mu_0 k \Delta T}, \quad (5.28)$$

and we suppose that viscous heating is significant, i. e.,  $\alpha \geq O(1)$ .

### Steady state

The steady state solution  $\theta_0(r)$  of (5.26) satisfies the nonlinear ordinary differential equation

$$(\xi \theta'_0)' + \alpha \xi^3 e^{\theta_0} = 0, \quad (5.29)$$

with boundary conditions

$$\theta_0(1) = 0, \quad \theta'_0(0) = 0. \quad (5.30)$$

We put

$$z = -\ln \xi, \quad \theta_0 = 4z + \psi, \quad (5.31)$$

so that  $\psi$  satisfies

$$\psi'' + \alpha e^\psi = 0, \quad (5.32)$$

with boundary conditions

$$\psi = 0 \quad \text{at} \quad z = 0, \quad \psi' \rightarrow -4 \quad \text{as} \quad z \rightarrow \infty. \quad (5.33)$$

A first integral of this can be found, which is

$$\frac{1}{2} \psi'^2 + \alpha e^\psi = 8, \quad (5.34)$$

and the problem reduces to a quadrature, which determines the solution as

$$\psi = -2 \ln \left[ \sqrt{\frac{\alpha}{8}} \cosh(2z + c) \right], \quad (5.35)$$

where  $c$  is given by

$$c = \pm \cosh^{-1} \sqrt{\frac{8}{\alpha}}. \quad (5.36)$$

There are thus two solutions if  $\alpha < 2\sqrt{2}$ , and none if  $\alpha > 2\sqrt{2}$ . For  $\alpha$  greater than this critical value, the solution to the initial value problem for (5.22) blows up in finite time, and no steady state solution exists. It is this phenomenon which is associated with the charring which is observed if the injection moulding flow is too fast.

### Linear stability

We study the stability of a steady state solution by putting

$$\theta = \theta_0 + \Theta. \quad (5.37)$$

Linear stability follows from assuming that (remember the variables are dimensionless)  $\Theta \ll 1$ , and linearising the resulting evolution equation for  $\Theta$ , to obtain

$$\Theta_\tau = \frac{1}{\xi} \frac{\partial}{\partial \xi} \left[ \xi \frac{\partial \Theta}{\partial \xi} \right] + \alpha \xi^2 e^{\theta_0} \Theta. \quad (5.38)$$

This has separable solutions of the form

$$\Theta = e^{\sigma \tau} P(\xi), \quad (5.39)$$

where  $P$  satisfies

$$(\xi P')' + \left[ -\sigma \xi + \alpha \xi^3 e^{\theta_0(\xi)} \right] P = 0, \quad (5.40)$$

together with boundary conditions

$$P'(0) = 0, \quad P(1) = 0. \quad (5.41)$$

This is of the same form as the Sturm-Liouville problem (5.1). If written in terms of  $z$  given in (5.31), the problem becomes

$$P'' + \left[ -\sigma e^{-2z} + 8 \operatorname{sech}^2(2z + c) \right] P = 0, \quad (5.42)$$

where now  $P' = dP/dz$ ; this is to be solved on  $0 < z < \infty$ , with

$$P(0) = 0, \quad P'(\infty) = 0. \quad (5.43)$$

## 5.2 Sturm-Liouville theory

The general Sturm-Liouville system is written in the form

$$(py')' + (\lambda r - q)y = 0, \quad (5.44)$$

where  $p$ ,  $q$  and  $r$  are functions of the space variable  $x$ , and to be specific we solve the equation in the domain  $[a, b]$  and apply boundary conditions

$$y(a) = y(b) = 0. \quad (5.45)$$

This is an eigenvalue problem: as we shall see, solutions exist only for a discrete set of eigenvalues  $\lambda_1, \lambda_2, \dots$ , for which the corresponding functions  $\phi_1(x), \phi_2(x), \dots$ , are called eigenfunctions. We wish to demonstrate this, and elucidate some of the properties of these eigenpairs.

We will assume that  $p$ ,  $q$  and  $r$  are positive and *smooth*, that is to say continuously differentiable ( $\in C^1$ ), which provides the simplest case. It is possible to relax some of

these assumptions; in particular, it is often found (as in our examples (5.10), (5.16) and (5.40)) that  $p = 0$  at an end point; in this case the Sturm-Liouville system is referred to as a singular system. Depending on the singularity, the basic results which we describe below will still hold, with some modifications.

The system (5.44) and (5.45) is a boundary value problem, which is why it is difficult to prove existence results. We therefore begin with a corresponding initial value problem, that is, (5.44) together with the initial conditions

$$y(a) = 0, \quad y'(a) = 1. \quad (5.46)$$

We define  $z = py'$  and  $\mathbf{u} = \begin{pmatrix} y \\ z \end{pmatrix}$ , so that

$$\mathbf{u}' = M\mathbf{u}, \quad (5.47)$$

where the matrix  $M$  is

$$M = \begin{pmatrix} 0 & \frac{1}{p} \\ q - \lambda r & 0 \end{pmatrix}, \quad (5.48)$$

and

$$\mathbf{u}(a) = \begin{pmatrix} 0 \\ p(a) \end{pmatrix}. \quad (5.49)$$

By our assumptions, the matrix  $M$  is smooth and therefore  $M\mathbf{u}$  is Lipschitz continuous: hence Picard's theorem implies that there exists a solution which is unique so long as  $\mathbf{u}$  is bounded. Being a linear system (with bounded coefficients), the solution can grow no faster than exponentially, and therefore the solution will exist for all  $x$ .

The matrix  $M$  depends on  $\lambda$ , so that the solutions do also, and in particular  $y = y(x, \lambda)$ . The dependence on  $\lambda$  is smooth, and indeed analytic, as can be shown simply by constructing the solution  $\frac{\partial \mathbf{u}}{\partial \lambda}$  of the differentiated system

$$\mathbf{u}'_{\lambda} = M\mathbf{u}_{\lambda} + \begin{pmatrix} 0 \\ -rz \end{pmatrix}, \quad (5.50)$$

using the fundamental solution as in chapter 4. Thus the value of  $y$  at  $x = b$ ,

$$y(b, \lambda) = y_b(\lambda) \quad (5.51)$$

is an analytic function of  $\lambda$ . If  $y = 0$  at  $x = b$  (as well as  $x = a$ ), i.e.,  $y_b(\lambda) = 0$ , then we say  $\lambda$  is an eigenvalue of the Sturm-Liouville system. The solution function  $y(x)$  is called the corresponding eigenfunction.

## 5.2.1 Eigenvalues and eigenfunctions

It will not have escaped the observant reader that Sturm-Liouville systems represent a generalisation of the simple harmonic oscillator

$$y'' + \lambda y = 0, \quad y(0) = y(1) = 0, \quad (5.52)$$

where  $p = r = 1$ ,  $q = 0$ . Evidently the eigensolutions are  $y_n = \sin n\pi x$ , where  $n$  is an integer, and the corresponding eigenvalues are  $\lambda_n = n^2\pi^2$ . We note that the eigenvalues form a denumerable sequence which tends to infinity, and the eigenfunctions are orthogonal, in the sense that  $\int_0^1 y_m y_n dx = 0$  for  $m \neq n$ . It is as a consequence of this that the functions  $\{y_n\}$  form a basis for the expansion of functions in Fourier series, which is in fact *complete*, i. e., any smooth (or in fact piecewise continuous) function can be expanded in a Fourier series. We will find that these results carry across to the general Sturm-Liouville system.

Let us denote  $w = \frac{\partial y}{\partial \lambda}$ . We then have

$$\begin{aligned} (py')' + (\lambda r - q)y &= 0, \\ (pw')' + (\lambda r - q)w &= -ry. \end{aligned} \quad (5.53)$$

Multiplying by  $w$  and  $y$  respectively, subtracting and integrating, we find

$$[p(y'w - yw')]_a^b = \int_a^b ry^2 dx > 0, \quad (5.54)$$

assuming  $r > 0$ . Since  $y = 0$  at  $a, b$ , and  $w = 0$  at  $a$ , it follows that

$$(py'w)|_b = \int_a^b ry^2 dx > 0, \quad (5.55)$$

and therefore  $w = \frac{\partial y}{\partial \lambda} \neq 0$  at  $x = b$ . It follows from this that the zeros of  $y_b(\lambda)$  are isolated, since  $y_b$  is analytic, and if there an infinite number, the only limit point can be at  $\pm\infty$ .

### An eigenvalue exists

It is by no means obvious that any eigenvalues will exist. To show that there is at least one, we consider the initial value problem (5.44) with the initial conditions (5.46). If  $p > 0$ , then we can suppose without loss of generality that  $p = 1$ , and we suppose that  $q$  and  $r$  are positive.

If  $\lambda \leq 0$ , then  $y'' \geq 0$  as long as  $y > 0$ . Since  $y > 0$  for  $x$  close to  $a$ , this shows that  $y$  will continue to increase if  $\lambda \leq 0$ , and therefore  $y_b(\lambda) > 0$  for  $\lambda \leq 0$ . By continuity, this can be extended to state that  $y_b(\lambda) > 0$  for  $\lambda$  positive and sufficiently small.

Next we consider two solutions  $y_1$  and  $y_2$  with different values of  $\lambda = \lambda_1$  and  $\lambda = \lambda_2$ , and we suppose  $\lambda_1 < \lambda_2$ . Multiplying the equations for  $y_1$  and  $y_2$  by  $y_2$  and  $y_1$ , respectively, and subtracting, we find

$$(y_1'y_2 - y_1y_2')' = (\lambda_2 - \lambda_1)ry_1y_2. \quad (5.56)$$

Integrating, and then dividing by  $y_2^2$ , we obtain

$$\left(\frac{y_1}{y_2}\right)' = \frac{(\lambda_2 - \lambda_1) \int_a^x ry_1y_2 dx}{y_2^2} > 0. \quad (5.57)$$

By l'Hôpital's rule, we have  $\frac{y_1}{y_2} = 1$  at  $x = a$ , therefore  $y_1 > y_2$  as long as  $y_2 > 0$ .

This implies that  $y_b(\lambda)$  is monotonically decreasing with  $\lambda$  as long as it is positive, and also that, if  $x_1(\lambda)$  is the first zero of  $y$  (if it exists), then  $x_1$  is monotonically decreasing with  $\lambda$ ; in particular,  $x_1 > b$  for small positive  $\lambda$ .

On the other hand, it is fairly obvious that if  $\lambda \gg 1$ , then by comparison of the solution of (5.44) (with  $p = 1$ ) with that of

$$z'' + \mu z = 0, \quad z(a) = 0, \quad z'(a) = 1, \quad (5.58)$$

where  $\lambda > \frac{q_{\max} + \mu}{r_{\min}}$  and  $\mu$  is positive, we can show that  $x_1(\lambda)$  does indeed exist, and  $x_1 \rightarrow 0$  as  $\lambda \rightarrow \infty$ . Since  $x_1$  varies continuously with  $\lambda$ , is monotonically decreasing with  $\lambda$ , and  $x_1 > b$  for sufficiently small  $\lambda$ , this implies that there is a unique positive value  $\lambda_1$  where  $x_1(\lambda_1) = b$ , and this is an eigenvalue of the Sturm-Liouville system. Extension to the case  $p \neq 1$  is trivial, by defining the new space variable  $X = \int_a^x \frac{d\xi}{p(\xi)}$ , and suitable redefinition of  $q$  and  $r$ .

This proves that there is an eigenvalue. Moreover, we have in passing proved that all the eigenvalues are positive, their only limit point is  $\infty$ , and that the minimum eigenvalue  $\lambda_1$  has a corresponding eigenfunction  $\phi_1(x)$ , say, which is of one sign. In addition, it is simple to repeat the argument for  $\lambda > \lambda_1$  (simply redefine the left endpoint  $a_1$  to be at  $x_1(\lambda)$ ) to show that there must be an infinite sequence of eigenvalues, and the sequence is denumerable since it can be exhaustively constructed in the above way.

## Eigenfunctions are orthogonal

A fundamentally important property of the eigenfunctions of Sturm-Liouville systems is that they are orthogonal. Orthogonality is a property which we associate most easily with vectors. In three dimensions, two vectors **a** and **b** are at right angles if their scalar product  $\sum_i a_i b_i$  is zero, and we extend the definition to less geometrically accessible  $n$ -dimensional vectors simply by allowing the same definition of scalar product.

The language we are using to discuss eigenvalues is very suggestive of the fact that we are here dealing with the same algebraic and geometric structures as occur in finite-dimensional vector space. There is, in fact, a direct analogy: functions are essentially (countably) infinite dimensional vectors; the finite number of eigenvalues of a matrix is replaced by the denumerable sequence we have discussed, and just as the eigenvalues of a real symmetric (or self-adjoint) matrix are real and orthogonal, so also a Sturm-Liouville system is self-adjoint (this will be elaborated in chapter 6), and has real eigenvalues which are orthogonal. The definition of a scalar product for a function space is called an *inner product*, and its definition is the extension of the finite sum  $\sum_i a_i b_i$  to its limit — an integral. To be specific, we define the inner product of two functions  $u$  and  $v$  with (positive) weight function  $r(x)$  to be

$$(u, v) = \int_a^b r u v dx, \quad (5.59)$$

and we say that  $u$  and  $v$  are orthogonal if  $(u, v) = 0$ .

For the eigenfunctions  $y_1$  and  $y_2$  with corresponding eigenvalues  $\lambda_1$  and  $\lambda_2$ , we find from the differential equation that

$$[p(y_1' y_2 - y_1 y_2')] + (\lambda_1 - \lambda_2) r y_1 y_2 = 0. \quad (5.60)$$

Integrating from  $a$  to  $b$  and using the boundary conditions, we find

$$\int_a^b r y_1 y_2 dx = 0, \quad (5.61)$$

that is,  $y_1$  and  $y_2$  are orthogonal.

Just as we construct bases for vector spaces from sequences of orthogonal vectors, so we can attempt to construct ‘bases’ of orthogonal functions for function spaces. This is the realm of Fourier series. We have an infinite sequence of eigenfunctions, and if we take these to be orthonormal, in the sense that  $(\phi_n, \phi_n) = 1$ , then a candidate Fourier expansion of an arbitrary function  $f(x)$  is

$$f = \sum_i (f, \phi_i) \phi_i, \quad (5.62)$$

and the issue of whether this candidate expansion is actually correct resides in whether the eigenfunctions form a ‘complete’ sequence. The answer is that they do, but a complete proof of this lies somewhat beyond the scope of these notes. We will sketch a proof in chapter 6.

## A variational principle

Multiplying the Sturm-Liouville equation for  $y$  by  $y$  and integrating, we find that each eigenvalue  $\lambda$  satisfies

$$\lambda = \frac{\int_a^b (q y'^2 + p y^2) dx}{\int_a^b r y^2 dx}. \quad (5.63)$$

Since this expression is homogeneous, we can take the denominator to be equal to one without loss of generality. In different phrasing, we define the *norm* of a function  $u$  to be

$$||u|| = \left\{ \int_a^b r u^2 dx \right\}^{1/2}, \quad (5.64)$$

that is,  $||u|| = \sqrt{(u, u)}$ , in analogy with the Euclidean vector norm, so that we may take  $||y|| = 1$  in (5.63). In that case

$$\lambda = \int_a^b (q y'^2 + p y^2) dx. \quad (5.65)$$

It is a simple observation of the calculus of variations that the functional  $\lambda[y]$  defined by (5.65) (or (5.63)) takes stationary values when the Sturm-Liouville equation

is satisfied, and in fact is also a local minimum (since  $p$  and  $q$  are positive). The first eigenvalue  $\lambda_1$  is equal to the global minimum of  $\lambda[y]$ , and then the next one  $\lambda_2$  is the global minimum of  $\lambda[y]$  subject to  $(y, y_1) = 0$ , and so on. It is somewhat easier to prove these statements when the Sturm-Liouville system is rewritten as an integral equation, and this is done in chapter 6.

## 5.2.2 Singular Sturm-Liouville systems

## 5.3 Comparison methods

Related to eigenvalue problems are the oscillation and comparison theorems, which apply generally to second order differential equations, and concern the existence and location of zeros of the solutions. We will prove four related results of this type.

Suppose first that

$$\begin{aligned} (py_1')' - q_1 y_1 &= 0, \\ (py_2')' - q_2 y_2 &= 0, \end{aligned} \tag{5.66}$$

and  $q_1 > q_2$ . Then  $y_2$  has at least one zero between two successive zeros of  $y_1$ . For, if this is not the case, let us take  $x_1$  and  $x_2$  to be two successive zeros of  $y_1$ , and suppose without loss of generality that  $y_1 > 0$  for  $x_1 < x < x_2$ . If  $y_2$  is non-zero in the same interval, we may suppose that it is also positive. Using a by now familiar construction, we have that

$$[p(y_1' y_2 - y_1 y_2')] = (q_1 - q_2) y_1 y_2 > 0, \tag{5.67}$$

and integrating this between  $x_1$  and  $x_2$  yields

$$(py_2 y_1')|_{x_2} > (py_2 y_1')|_{x_1}, \tag{5.68}$$

which is impossible since we must have  $y_1'|_{x_2} < 0$  and  $y_1'|_{x_1} > 0$ .

Next we show that the zeros of independent solutions of

$$(py')' - qy = 0 \tag{5.69}$$

interlace. Suppose  $y_1$  and  $y_2$  are independent solutions of (5.69): then their zeros are distinct (since if  $y_1 = y_2 = 0$  at some value of  $x$ , then Picard's theorem implies that one is a multiple of the other). If  $x_1$  and  $x_2$  are successive zeros of  $y_1$ , then as above, we can find that

$$[p(y_1' y_2 - y_1 y_2')]_{x_1}^{x_2} = 0, \tag{5.70}$$

whence

$$(py_2 y_1')|_{x_2} = (py_2 y_1')|_{x_1}. \tag{5.71}$$

Because  $y_1'(x_1)$  and  $y_1'(x_2)$  have opposite signs, it follows that so also do  $y_2(x_1)$  and  $y_2(x_2)$ , and thus that  $y_2$  has a zero between  $x_1$  and  $x_2$ . The same argument works the other way round. This proves that the zeros of  $y_1$  and  $y_2$  interlace.



Now we turn to the Sturm-Liouville eigenvalue problem. If  $y_1$  and  $y_2$  are two eigenfunctions with eigenvalues  $\lambda_1$  and  $\lambda_2$ , so that

$$\begin{aligned}(py_1')' + (\lambda_1 r - q)y_1 &= 0, \\ (py_2')' + (\lambda_2 r - q)y_2 &= 0,\end{aligned}\tag{5.72}$$

let us suppose that  $x_1$  and  $x_2$  are two successive zeros of  $y_1$ , with  $y_1$  positive between them, and suppose also that  $y_2 > 0$  in this interval. If  $\lambda_1 < \lambda_2$ , then

$$[py_1' y_2]_{x_1}^{x_2} = (\lambda_2 - \lambda_1) \int_{x_1}^{x_2} r y_1 y_2 dx > 0,\tag{5.73}$$

which is impossible. It follows that if  $\lambda_1 < \lambda_2$ , then the eigenfunction  $y_2$  has a zero between any two zeros of  $y_1$ .

Finally we prove again the result on the number of eigenvalues. For (5.44), we define

$$\xi = \int_a^x \frac{dt}{p(t)},\tag{5.74}$$

so that with  $y' = dy/d\xi$ , we have

$$y'' + [\lambda pr - pq]y = 0.\tag{5.75}$$

We compare the solutions of this equation with those of

$$z'' + (\lambda R - Q)z = 0,\tag{5.76}$$

where  $R$  and  $Q$  are positive constants such that  $pr > R$  and  $pq < Q$ , and we suppose  $\lambda > \frac{Q}{R} > 0$ . Then  $Q - \lambda R > pq - \lambda pr$ , and the first result in this section (after (5.66)) then implies that  $y$  has a zero between any two of  $z$ . Since by construction,  $z$  has an infinite number of zeros, it follows that so also does  $y$ .

Denote these as  $x_n(\lambda)$ ,  $n = 1, 2, \dots$ . From (5.55) it follows that  $y'y_\lambda > 0$  at  $x_n$ , where also  $y[x_n(\lambda), \lambda] \equiv 0$ . Thus we find

$$x'_n(\lambda) = -\frac{y_\lambda y'}{y'^2} < 0,\tag{5.77}$$

and thus all the zeros decrease as  $\lambda$  increases. From all this we can (again) deduce that there are an infinite number of eigenvalues  $\lambda_n$ , with  $\lambda_n \rightarrow \infty$  as  $n \rightarrow \infty$ .

## 5.4 Notes and references

A good source for the material on separation of variables and the derivation of the main equations of mathematical physics is the classic text by Jeffreys and Jeffreys (1946). This book has been reprinted many times and is still in print.

There are many books on differential equations which cover Sturm-Liouville equations. A classical text is that by Courant and Hilbert (1937). Other books, no less valuable, are those by Mackie (1965), Coddington and Levinson (1972), and Stakgold (2000). Oscillation and comparison theorems are nicely discussed in the little book by Burkill (1956). Simmons (1972) is a very accessible entry level text.

## Exercises

5.1 Show that the second order differential equation

$$w'' + a(x)w' + [b(x) + \lambda c(x)w] = 0$$

can be written in the Sturm-Liouville form

$$(py')' + (\lambda r - q)y = 0,$$

by suitable definitions of the functions  $y$ ,  $p$ ,  $q$  and  $r$ .

If the boundary conditions are

$$w' + \alpha w = 0 \quad \text{at} \quad x = 0, 1,$$

what are the corresponding boundary conditions for  $y$ ?

5.2 Bessel's equation of order  $\nu$  is

$$w'' + \frac{1}{z}w' + \left(1 - \frac{\nu^2}{z^2}\right)w = 0,$$

and the solutions are called Bessel functions of order  $\nu$ . Show that the solutions of the Sturm-Liouville equation

$$(rR')' + \left(k^2r - \frac{m^2}{r}\right)R = 0 \quad (*)$$

can be written as Bessel functions of order  $m$ .

By seeking solutions near  $z = 0$  of the form  $w \sim z^\mu$ , show that if  $\nu \neq 0$ , then independent solutions  $w_\pm$  are possible with  $w_\pm \sim z^{\pm\nu}$ . Deduce that if we require  $R(0) = 0$  when  $m$  is a positive integer in  $(*)$ , then we must have  $R \propto w_+(kr)$ .

( $w_+$  is normally written  $J_\nu$ , and is the Bessel function of the first kind; the Bessel function of the second kind is singular at  $z = 0$  and is denoted  $Y_\nu$ .)

Show that if  $\nu = 0$ , then independent solutions are possible in which

$$J_0 \sim 1 - \frac{1}{4}z^2, \quad Y_0 \sim \ln z.$$

5.3 A function  $u(x, t)$  satisfies the nonlinear reaction-diffusion equation

$$u_t = u_{xx} + f(u) \quad \text{on} \quad [0, 1],$$

with boundary conditions

$$u = 0 \quad \text{on} \quad x = 0, 1.$$

Suppose that  $u_0(x)$  is a steady solution. By writing  $u = u_0 + U$ , derive a linearised equation for small perturbations  $U$ . Show that it has separable solutions of the form  $U = v(x)e^{\sigma t}$ , and deduce that  $v$  satisfies the Sturm-Liouville equation

$$v'' + [s(x) - \sigma]v = 0, \quad v(0) = v(1) = 0,$$

where

$$s(x) = f'[u_0(x)].$$

By consideration of a suitable variational principle, or otherwise, show that if  $s(x) < 0$ , then all eigenvalues  $\sigma$  are negative, and hence the steady state is stable.

If  $s(x) > 0$ , show, by writing  $q = s_{\max} - s$  and  $\lambda = s_{\max} - \sigma$ , that instability can occur ( $\sigma > 0$ ) if

$$\min_v \frac{\int_0^1 (v'^2 - s(x)v^2) dx}{\int_0^1 v^2 dx} < 0.$$

(This shows that the constraint that  $q > 0$  in Sturm-Liouville theory is not necessary.)

- 5.4 A shear thinning fluid (like tomato ketchup or Guinness) has a viscosity which decreases with increasing stress. Suppose such a fluid flows in a cylindrical tube of radius  $a$ , and has viscosity

$$\mu = \frac{\mu_0}{|\tau|^{n-1}} \exp \left[ -\frac{(T - T_0)}{\Delta T} \right],$$

where  $\tau$  is the shear stress, and the power law exponent  $n > 1$  ( $n = 1$  corresponds to Newtonian flow).

Show that a suitable dimensionless formulation of the temperature profile across the tube satisfies the dimensionless equation

$$\frac{\partial \theta}{\partial t} = \frac{1}{\xi} \frac{\partial}{\partial \xi} \left[ \xi \frac{\partial \theta}{\partial \xi} \right] + \beta \xi^{n+1} e^\theta,$$

with the boundary conditions

$$\theta = 0 \quad \text{on} \quad r = 1,$$

and

$$\theta_r = 0 \quad \text{at} \quad r = 0.$$

Give the definition of the parameter  $\beta$ .

Find a steady solution of this problem, and find a Sturm-Liouville eigenvalue problem, whose eigenvalues determine the linear stability of the steady solution.

# Chapter 6

## Integral equations and eigenfunctions

### 6.1 Integral equations

Like differential equations, the formulation and solution of integral equations is of interest in its own right. In this chapter, we will however focus on one particular class of integral equations, those which arise through the reformulation of a Sturm-Liouville differential equation system as a Fredholm integral equation. More generally, we begin by illustrating how this reformulation occurs.

#### 6.1.1 Volterra integral equations

Volterra integral equations are those in which the integral occurs in indefinite form (i. e., the space variable  $x$  occurs as the upper limit of the integral). Typically Volterra integral equations are associated with initial value problems. Consider for example the inhomogeneous differential equation

$$(py')' - qy = -f, \tag{6.1}$$

together with the initial conditions

$$y(a) = y'(a) = 0. \tag{6.2}$$

We can solve this using the method of variation of parameters, as follows. Suppose that  $y_1$  and  $y_2$  are independent solutions of the homogeneous equation (i. e., with  $f(x) = 0$ ) satisfying

$$\begin{aligned} y_1(a) &= 0, & y_1'(a) &= 1, \\ y_2(a) &= 1, & y_2'(a) &= 0. \end{aligned} \tag{6.3}$$

We seek a solution in the form

$$y = A(x)y_1 + B(x)y_2, \tag{6.4}$$

where the coefficient functions  $A$  and  $B$  are to be found. We require two equations for  $A$  and  $B$ , and we take the first of them to be

$$A'y_1 + B'y_2 = 0. \quad (6.5)$$

With this choice, we have that

$$y' = Ay'_1 + By'_2, \quad (6.6)$$

and the initial conditions for  $y$  in (6.2) require  $A$  and  $B$  to satisfy

$$A(a) = B(a) = 0. \quad (6.7)$$

Multiplying (6.6) by  $p$ , differentiating, and substituting into (6.1), we find the second equation for  $A$  and  $B$ :

$$y'_1 A' + y'_2 B' = -\frac{f}{p}. \quad (6.8)$$

Inverting the matrix equation formed from (6.5) and (6.8), we find

$$\begin{pmatrix} A' \\ B' \end{pmatrix} = \frac{1}{pW} \begin{pmatrix} y'_2 & -y_2 \\ -y'_1 & y_1 \end{pmatrix} \begin{pmatrix} 0 \\ -f \end{pmatrix}, \quad (6.9)$$

where  $W = y_1 y'_2 - y'_1 y_2$  is the Wronskian. Directly from the equations for  $y_1$  and  $y_2$ , we know that  $pW = C$  is constant (non-zero, since the solutions are independent), and therefore the solutions for  $A$  and  $B$  satisfying the initial conditions (6.7) are

$$A = \frac{1}{C} \int_a^x y_2(\xi) f(\xi) d\xi, \quad B = -\frac{1}{C} \int_a^x y_1(\xi) f(\xi) d\xi. \quad (6.10)$$

We can thus write the solution compactly as

$$y = \int_a^x K(x, \xi) f(\xi) d\xi, \quad (6.11)$$

where the integral kernel is defined by

$$K(x, \xi) = \frac{y_1(x)y_2(\xi) - y_1(\xi)y_2(x)}{C}. \quad (6.12)$$

This simply defines the solution. An integral equation occurs if we put  $f = \lambda ry$ , thus regaining the Sturm-Liouville equation (with initial conditions) of chapter 5. In this case, (6.11) becomes

$$y = \lambda \int_a^x K(x, \xi) r(\xi) y(\xi) d\xi. \quad (6.13)$$

This is now an integral equation of Volterra type, because of the upper limit  $x$  in the integral. In principle it can be solved using Picard iteration.

More generally, if we seek eigenfunctions satisfying  $y(a) = 0$ ,  $y(b) = 0$ , they can be formally constructed by solving the initial value problem with  $y(a) = 0$ ,  $y'(a) = 1$ , thus

$$y = \lambda \int_a^x K(x, \xi) r(\xi) y(\xi) d\xi + y_1(x), \quad (6.14)$$

and then choosing  $\lambda$  so that  $y(b) = 0$ . Practically, though, this is not a very good idea.

### 6.1.2 Fredholm integral equations

Now we consider the boundary value problem for (6.1), i. e.,

$$Ly \equiv (py')' - qy = -f, \quad y(a) = y(b) = 0. \quad (6.15)$$

The solution for this can be written in terms of a Green's function, which is constructed from independent solutions  $y_1$  and  $y_2$ , which we now take to be defined as solutions of the homogeneous version of (6.15) (i. e.,  $f = 0$ ) with

$$\begin{aligned} y_1(a) &= 0, & y_1'(a) &= 1, \\ y_2(b) &= 0, & y_2'(b) &= 1. \end{aligned} \quad (6.16)$$

We explicitly assume that zero is not an eigenvalue for the corresponding Sturm-Liouville system, so that this choice gives independent solutions.

The Green's function  $G(x, \xi)$  satisfies the homogeneous equation

$$LG \equiv (pG')' - qG = 0 \quad \text{in} \quad x \lessgtr \xi, \quad (6.17)$$

with

$$[G]_{\xi-}^{\xi+} = 0, \quad [pG']_{\xi-}^{\xi+} = -1, \quad (6.18)$$

where the primes denote differentiation with respect to  $x$ , and  $[\ ]_{\xi-}^{\xi+}$  indicates the jump from  $x$  just less than  $\xi$  to  $x$  just greater than  $\xi$ . In terms of the independent homogeneous solutions  $y_1$  and  $y_2$ , the Green's function is explicitly given by

$$G(x, \xi) = \begin{cases} -\frac{y_1(x)y_2(\xi)}{C}, & x < \xi, \\ -\frac{y_1(\xi)y_2(x)}{C}, & x > \xi, \end{cases} \quad (6.19)$$

where  $C = pW$  is constant (and non-zero),  $W$  being the Wronskian for this pair  $y_1, y_2$ .

If now we consider a Sturm-Liouville system in which  $f = \lambda ry$ , then the eigenfunctions  $y$  and eigenvalues  $\lambda$  satisfy the Fredholm integral equation

$$y(x) = \lambda \int_a^b G(x, \xi) r(\xi) y(\xi) d\xi. \quad (6.20)$$

This is called a Fredholm integral equation because the limits on the integral are fixed.

We may also derive an inhomogeneous Fredholm equation by consideration of the inhomogeneous Sturm-Liouville equation

$$(py')' + (\lambda r - q)y = -f, \quad (6.21)$$

subject to  $y(a) = y(b) = 0$ . This leads to the integral equation

$$y(x) = \lambda \int_a^b G(x, \xi) r(\xi) y(\xi) d\xi + g(x), \quad (6.22)$$

where

$$g(x) = \int_a^b G(x, \xi) f(\xi) d\xi. \quad (6.23)$$

### 6.1.3 The delta function

Although not essential to the discussion, it is difficult to go past the derivation of the Green's function without some reference to the delta function  $\delta(x)$ . The delta function is like an administrative tool; it makes things easier to treat, once you have the confidence to use it, but at first sight it is very strange.

For a start it is not a function. It is called a generalised function. It has the property that it is zero everywhere except at  $x = 0$ , where it is 'infinite enough' that

$$\int_{-\infty}^{\infty} \delta(x) dx = 1. \quad (6.24)$$

Obviously this is not really satisfactory. There are two formal ways of thinking about generalised functions. The first is that we think of a generalised function as being defined as a 'limit' (more precisely, an equivalence class) of a family of smooth functions; for example, the delta function is associated with the family of functions  $\sqrt{\frac{n}{\pi}} e^{-nx^2}$  as  $n \rightarrow \infty$ . (Note that there are many such families.)

But the main way of manipulating generalised functions is not to manipulate them directly at all, but to define their action by means of integrals. If  $\{g_n(x)\}$  is a family of functions in the equivalence class of a generalised function  $g(x)$ , then we define the integral

$$\int_{-\infty}^{\infty} f(x)g(x) dx = \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f(x)g_n(x) dx. \quad (6.25)$$

Using this definition, it is possible to prove that one can manipulate generalised functions directly by using this integral property, so that for example, one can define the delta function through the requirement that

$$\int_{-\infty}^{\infty} f(x)\delta(x) dx = f(0) \quad (6.26)$$

for all smooth functions  $f(x)$ .

With this rudimentary description of the delta function in hand, we can more conveniently define the Green's function  $G$  to be the solution of

$$(pG')' - qG = -\delta(x - \xi) \quad (6.27)$$

which satisfies the boundary conditions  $G = 0$  at  $x = a$  and  $x = b$ . Integration of (6.27) across  $x = \xi$  then implies the second equality in (6.18).

### 6.1.4 The Fredholm alternative

The Fredholm alternative concerns the solution of the inhomogenous Fredholm equation (6.22). To put this in more familiar shape, it is convenient to define an integral operator

$$\mathcal{G}u = \int_a^b G(x, \xi)r(\xi)u(\xi) d\xi. \quad (6.28)$$

We should think of this in the same way that we (should) do of functions and linear transformations. Both define mappings: a (real-valued) function of one (real) variable

$f(x)$  defines a mapping from the real numbers to the real numbers; a linear transformation  $M$  on a finite-dimensional vector space maps a vector  $\mathbf{v}$  to another vector  $M\mathbf{v}$ . In a similar way, an integral operator such as  $\mathcal{G}$  in (6.28) or a differential operator such as  $L$  in (6.15) maps a function to another function. These functions reside in a suitable function space, and the issue of what a suitable function space should be is of technical concern, and forms a focal point of functional analysis. It will only be of diversionary interest in these notes, but it is as well to be aware that many of the proofs of the subject really require attention to the detail of what function space the operators are defined on. Unlike, say, linear transformations, it is usual for mappings of functions to be defined on one function space, but to have their range, or image, in another function space. As a simple example, we might define the operator  $L$  in (6.15) as acting on the space  $C^2[a, b]$  of twice continuously differentiable functions on  $[a, b]$ , but its image is then clearly only the continuous functions  $C[a, b]$ . Apart from this, we should think of operators on function spaces in much the same way as we do linear (or nonlinear) transformations on vector spaces.

With the definition of  $\mathcal{G}$  in (6.28), we might take  $u \in C[a, b]$ , and then its image is also in  $C[a, b]$  if, as we suppose,  $\mathcal{G}$  is continuous. The inhomogeneous Fredholm integral equation (6.22) can then be written in the compact form

$$y = \lambda \mathcal{G}y + g. \quad (6.29)$$

The Fredholm alternative is essentially the same statement we would make about this equation if  $y$  and  $g$  were finite-dimensional vectors, and  $\mathcal{G}$  were a matrix. It is this: either the homogeneous equation  $g = 0$  has no non-zero solution, and then (6.29) has a unique solution; or, the homogeneous equation has a solution  $\phi(x)$  (an eigenfunction), and then the inhomogeneous equation has no solution at all unless  $g$  is orthogonal to  $\phi$ , in which case there are an infinite number of solutions (a one parameter family). The particular form of the statement above applies to self-adjoint linear operators  $\mathcal{G}$ ; this was touched on briefly, as was the definition of orthogonality, in chapter 5; it is discussed further in section 6.2.2 below. The proof of the Fredholm alternative will emerge as we discuss eigenfunctions further.

## 6.2 Eigenfunction expansions

### 6.2.1 Neumann series

Just as Picard's iterative method for Volterra integral equations leads to a convergent series form of solution, so also does an iterative method for the inhomogeneous Fredholm integral equation. It is most easily written using the operator notation, thus we solve (6.29) by defining the sequence of functions

$$y_0 = g, \quad y_n = \lambda \mathcal{G}y_{n-1} + g. \quad (6.30)$$



This leads to the sequence  $y_n = \sum_0^n \lambda^n \mathcal{G}^n g$ , and suggests that we may write the solution as the formal series

$$y = \sum_0^\infty \lambda^n \mathcal{G}^n g. \quad (6.31)$$

This is known as a Neumann series, or Neumann expansion. Note that the solution can be written in the formal way

$$y = (I - \lambda \mathcal{G})^{-1} g, \quad (6.32)$$

where  $I$  is the identity operator, and that the Neumann series results from the formal expansion of  $(I - \lambda \mathcal{G})^{-1}$  as a Taylor series. In fact, the issue of convergence of this series (and hence of validity of this formal solution) requires consideration of the norm of the operator  $\mathcal{G}$ . So now we revisit the definitions given in chapter 5 of norms, and we extend their definition from norms of functions to norms of operators.

### 6.2.2 Inner products and norms

We recall the definition of an inner product from chapter 5. If  $r(x) > 0$ , then we can define an inner product on a suitable space of functions as

$$(u, v) = \int_a^b r u v dx. \quad (6.33)$$

In reality, the space is often taken to be  $L^2$ , the space of Lebesgue square-integrable functions, but in practice, we will limit ourselves to considering the space of continuous functions. In general, an inner product  $(u, v)$  is a quantity which is symmetric ( $(u, v) = (v, u)$ ), bilinear (i. e., linear in both  $u$  and  $v$  arguments), and satisfies  $(u, u) \geq 0$ , with equality only if  $u = 0$ .<sup>1</sup>

We have already mentioned the idea that an operator be self-adjoint. The *adjoint*  $L^*$  of an operator  $L$  is defined (uniquely) by the property that

$$(Lu, v) = (u, L^*v), \quad (6.34)$$

and the operator is self-adjoint if  $L = L^*$ . It is easy to see that the Sturm-Liouville operators of chapter 5 and the integral operators of the present chapter are self-adjoint. It is this property which ensures the eigenvalues are real, and the eigenfunctions are orthogonal.

A space which has an inner product is called an inner product space, and an inner product space is also endowed with a norm:

$$\|u\| = (u, u)^{1/2}. \quad (6.35)$$

In general, a norm  $\|u\|$  is a distance-like measure of an element  $u$  of a space which is positive (unless  $u = 0$ ), linear under scalar multiplication, and satisfies the triangle inequality

$$\|u + v\| \leq \|u\| + \|v\| \quad \forall u, v. \quad (6.36)$$

---

<sup>1</sup>As we do throughout this discussion, we ignore the possibility that  $u$  and  $v$  are complex, and restrict ourselves to the case where they are real.

Spaces having norms are called metric spaces, and they allow the notion of convergence. If every convergent sequence of elements converges to an element of the same space, then the space is called complete. A complete, normed linear space  $B$  (i. e., if  $u \in B$  and  $v \in B$ , then  $\alpha u + \beta v \in B$  for all  $\alpha, \beta \in \mathbf{R}$ ) is called a Banach space. A complete inner product space is called a Hilbert space.

In an inner product space, we have that  $\|u + sv\|^2 = \|u\|^2 + 2s(u, v) + s^2\|v\|^2$  is always positive, and from this follows the Cauchy-Schwarz inequality:

$$(u, v) \leq \|u\| \|v\|. \quad (6.37)$$

It is as a consequence of this that the inner product defines an associated norm (6.35), since the triangle inequality can be deduced from the Cauchy-Schwarz inequality.

One of the reasons we bother so much with what space a function is in is that we want convergent sequences to converge in the same space: we want to deal with complete spaces. It is not difficult to see that apparently nice looking spaces, such as the continuous functions on  $[0, 1]$ , denoted  $C[0, 1]$ , are not complete. Endowed with the norm  $\|u\| = \left\{ \int_0^1 u^2 dx \right\}^{1/2}$ , the sequence of functions  $u_n = e^{-nx}$  converges, in the sense that  $\|u_n\| \rightarrow 0$  as  $n \rightarrow \infty$ . But the limit function is clearly  $u(0) = 1$ ,  $u(x) = 0$ ,  $x > 0$ , which is not continuous. So  $C[a, b]$  is not complete with this norm. The reason that the square-integrable function space  $L^2$  is of interest is that it *is* complete with this norm.

Given the property of a norm, we can also define norms of operators. We define the norm of an operator  $\mathcal{G}$  to be

$$\|\mathcal{G}\| = \sup \frac{\|\mathcal{G}u\|}{\|u\|}, \quad (6.38)$$

or equivalently,  $\|\mathcal{G}\| = \sup_{\|u\|=1} \|\mathcal{G}u\|$ . We leave it as an exercise to show that this definition indeed defines a norm. It is then immediate that the Neumann series (6.31) converges if

$$|\lambda| < \frac{1}{\|\mathcal{G}\|}, \quad (6.39)$$

and is an analytic function (or functional) of  $\lambda$  within this circle of convergence. The series convergence is compromised by a singularity on the circle of convergence, and this singularity occurs at an eigenvalue. More generally, analytic continuation of  $(I - \lambda\mathcal{G})^{-1}$  to other complex  $\lambda$  is limited only by singularities at the (real) eigenvalues  $\lambda_1, \lambda_2, \dots$

It remains to show that the integral operator  $\mathcal{G}$  defined by (6.28) is indeed bounded, so that its norm exists. For each  $x$ , we have in fact

$$\mathcal{G}u = (G, u) \leq \|G\| \|u\|, \quad (6.40)$$

and thus

$$\frac{\|\mathcal{G}u\|^2}{\|u\|^2} \leq \int_a^b \int_a^b r(\xi)r(\eta)G^2(\xi, \eta) d\xi d\eta, \quad (6.41)$$

whence  $\mathcal{G}$  is bounded if the right hand side of (6.41) is bounded (for example, if  $G$  is continuous), and then

$$\|\mathcal{G}\| \leq \left\{ \int_a^b \int_a^b r(\xi)r(\eta)G^2(\xi,\eta) d\xi d\eta \right\}^{1/2}. \quad (6.42)$$

This property defines  $\mathcal{G}$  as a *Hilbert-Schmidt* operator.

### 6.2.3 Spectral theory

The discussion of eigenvalues and eigenfunctions sits within the more general functional analytic subject of spectral theory. For a general operator  $\mathcal{G}$ , the values of  $\lambda$  where the equation  $(I - \lambda\mathcal{G})u = 0$  has a non-trivial solution are the eigenvalues, and in the present situation, where we have a Hilbert-Schmidt operator, the eigenvalues constitute the entire *spectrum*. More generally, the eigenvalues constitute the *point spectrum*. If we suppose that  $\mathcal{G}$  maps functions in a space  $U$  to  $U$  (e. g., we might have  $U = C[a, b]$ ), then values of  $\lambda$  for which  $(I - \lambda\mathcal{G})u = 0$  has no non-trivial solution, but where the range of  $(I - \lambda\mathcal{G})$  is not the whole of  $U$ , constitute the continuous and residual spectrum; the continuous spectrum is that part for which  $(I - \lambda\mathcal{G})^{-1}$  is unbounded.

## 6.3 Variational methods

Most of what we have said so far applies to general Fredholm integral equations with real symmetric kernels, and this will continue to be the case. For the particular integral operators which are the inverse of the Sturm-Liouville operators, we already know that the eigenvalues are characterised by a variational principle, specifically

$$\lambda_1 = \min_{\|u\|=1} \left\{ \int_a^b (qu^2 + pu'^2) d\xi \right\}, \quad (6.43)$$

with similar variational principles for the other eigenvalues. It is easiest to prove such assertions when the problem is formulated as an integral equation.

### 6.3.1 A variational principle

Now let us address the Fredholm integral eigenvalue problem

$$\mathcal{G}u = \int_a^b G(x, \xi)r(\xi)u(\xi) d\xi = \frac{u}{\lambda}, \quad (6.44)$$

where we assume  $G$  is real and symmetric, so that  $\mathcal{G}$  is self-adjoint. The Cauchy-Schwarz inequality states that  $(\mathcal{G}u, u) \leq \|\mathcal{G}u\| \|u\| \leq \|\mathcal{G}\|$  if  $\|u\| = 1$ , and it is simple to show that when this upper bound is obtained, then  $\mathcal{G}u = \frac{u}{\lambda_1}$ , and thus we have

$$\sup_{\|u\|=1} (\mathcal{G}u, u) = \frac{1}{\lambda_1} = \|\mathcal{G}\|. \quad (6.45)$$

All that is missing from this characterisation of the first eigenvalue is a proof that the supremum is actually attained. This takes us to the periphery of our exposition. From the definition of a supremum, we have that there is a sequence of functions  $\{u_m\}$  such that  $(\mathcal{G}u_m, u_m) \rightarrow \|\mathcal{G}\| = 1/\lambda_1$  as  $m \rightarrow \infty$ . For  $u_m \in U = C[a, b]$ ,  $\mathcal{G}u_m$  is an equicontinuous, uniformly bounded sequence of functions (an equicontinuous family is one which is uniformly continuous over both  $x$  and  $m$ ). The Arzela-Ascoli theorem states that an infinite equicontinuous, uniformly bounded sequence of functions contains a convergent subsequence. Taking this sequence to be  $\{u_m\}$ , this shows that  $\mathcal{G}u_m \rightarrow \phi_1(x)$ , where  $\phi_1$  is continuous, and it is straightforward to show that  $\mathcal{G}\phi_1 = \frac{\phi_1}{\lambda_1}$ ;  $\phi_1$  is the first eigenfunction.

Now consider the modified integral equation

$$\mathcal{G}_2 u = \int_a^b G_2(x, \xi) r(\xi) u(\xi) d\xi = \frac{u}{\lambda}, \quad (6.46)$$

where

$$G_2(x, \xi) = G(x, \xi) - \frac{\phi_1(x)\phi_1(\xi)}{\lambda_1}. \quad (6.47)$$

As in the process of Gram-Schmidt orthogonalisation, we can write

$$u = v + \alpha\phi_1, \quad (6.48)$$

where

$$v = u - (u, \phi_1)\phi_1, \quad \alpha = (u, \phi_1), \quad (6.49)$$

and thus  $(v, \phi_1) = 0$ . By direct calculation, we find that  $(\mathcal{G}_2 u, u) = (\mathcal{G}v, v)$ , and as before, we can deduce that

$$\sup_{\|u\|=1} (\mathcal{G}_2 u, u) = \sup_{(v, \phi_1)=0} (\mathcal{G}v, v) = \frac{1}{\lambda_2} \quad (6.50)$$

is obtained at a second eigenfunction  $\phi_2$ , which is orthogonal to the first.

We can go on iteratively in this way to construct a sequence of operator equations

$$\mathcal{G}_n u = \int_a^b G_n(x, \xi) r(\xi) u(\xi) d\xi = \frac{u}{\lambda}, \quad (6.51)$$

where

$$G_n(x, \xi) = G(x, \xi) - \sum_{j=1}^{n-1} \frac{\phi_j(x)\phi_j(\xi)}{\lambda_j}, \quad (6.52)$$

and we find that the  $n$ -th eigenvalue  $\lambda_n$  satisfies

$$\frac{1}{\lambda_n} = \sup_{\|u\|=1} (\mathcal{G}_n u, u) = \|\mathcal{G}_n\|, \quad (6.53)$$

and this bound is attained when  $u = \phi_n$ , the  $n$ -th eigenfunction. The eigenfunctions (with  $\|\phi_j\| = 1$ ) form a denumerable orthonormal set.

We already know from chapter 5 that  $\lambda_n \rightarrow \infty$  as  $n \rightarrow \infty$ , at least for the Green's function kernel  $G(x, \xi)$ . We can also show this directly by calculating

$$\int_a^b \int_a^b G_n^2(x, \xi) r(x) r(\xi) dx d\xi = \int_a^b \int_a^b G^2(x, \xi) r(x) r(\xi) dx d\xi - \sum_1^{n-1} \frac{1}{\lambda_j^2}, \quad (6.54)$$

from which it follows that the infinite sum converges, and we have a form of Bessel's inequality:

$$\sum_1^\infty \frac{1}{\lambda_j^2} \leq \int_a^b \int_a^b G^2(x, \xi) r(x) r(\xi) dx d\xi. \quad (6.55)$$

In particular, there is no finite limit point for  $\lambda_j$ , and  $\lambda_j \rightarrow \infty$  as  $j \rightarrow \infty$ .

Since  $\|\mathcal{G}_n\| = 1/\lambda_n \rightarrow 0$  as  $n \rightarrow \infty$ , it follows that for any  $u$  on the unit sphere  $\|u\| = 1$ , the sequence  $\{\|\mathcal{G}_n u\|\} \rightarrow 0$  as  $n \rightarrow \infty$ . As before, Arzela-Ascoli implies that there is a subsequence such that  $\{\mathcal{G}_n u\} \rightarrow 0$ , and since  $\mathcal{G}_n$  is equicontinuous, this implies that  $\mathcal{G}_n \rightarrow 0$ , and thus that  $G_n \rightarrow 0$ ; hence we have the expansion of the kernel in the form

$$G(x, \xi) = \sum_1^\infty \frac{\phi_j(x) \phi_j(\xi)}{\lambda_j}. \quad (6.56)$$

### 6.3.2 Fourier series

Having constructed an orthonormal sequence of eigenfunctions, we would like to be able to solve some of the general initial value problems which give rise to the Sturm-Liouville problems of chapter 5. Whether we can do this depends on whether the orthonormal sequence of eigenfunctions is complete. To see this, let us suppose that the function  $f(x) \in C^2[a, b]$  (this is not actually necessary, but makes the demonstration simpler). With the Sturm-Liouville differential operator defined by (6.15), we have that  $u = -Lf$  is continuous, and its inverse is  $f = \mathcal{G}u$ . We wish to show that  $f$  can be expanded in a Fourier series,

$$f = \sum_1^\infty (f, \phi_j) \phi_j. \quad (6.57)$$

Note that  $(f, \phi_j) = (\mathcal{G}u, \phi_j) = (u, \mathcal{G}\phi_j)$  by the property of self-adjointness, and this equals  $\frac{1}{\lambda_j} (u, \phi_j)$ . Hence we can deduce that

$$f - \sum_1^{m-1} (f, \phi_j) \phi_j = \mathcal{G}[u - \sum_1^{m-1} (u, \phi_j) \phi_j] = \mathcal{G}_m u, \quad (6.58)$$

and thus

$$\|f - \sum_1^{m-1} (f, \phi_j) \phi_j\| = \|\mathcal{G}_m u\| \leq \|\mathcal{G}_m\| \|u\| = \frac{\|u\|}{\lambda_m} \rightarrow 0 \quad \text{as } m \rightarrow \infty. \quad (6.59)$$

Again, equicontinuity of  $\{\mathcal{G}_m\}$  implies that  $f - \sum_1^{m-1} (f, \phi_j) \phi_j$  converges to zero, and hence (since  $f$  is continuous) we have

$$f = \sum_1^{\infty} (f, \phi_j) \phi_j. \quad (6.60)$$

As a corollary (although it is in fact equivalent), we have Parseval's relation:

$$\|f\|^2 = \sum_1^{\infty} |(f, \phi_j)|^2. \quad (6.61)$$

A version of (6.61) with inequality (left hand side  $\geq$  right hand side) is Bessel's inequality.

### 6.3.3 The Fredholm alternative revisited

### 6.3.4 Rayleigh-Ritz method

## 6.4 Notes and references

There are a good number of books on integral equations. Most of the material in this chapter can be found in Courant and Hilbert (1937), a classic exposition, or Stakgold (2000), which has a more up to date flavour. Generalised functions are treated elegantly in the little book by Lighthill (1958).

## Exercises

- 6.1 Show that the Wronskian  $W = y_1 y_2' - y_1' y_2$  of two solutions  $y_1, y_2$  of the differential equation

$$[p(x)y']' - q(x)y = 0, \quad p(x) > 0,$$

satisfies  $pW = C$ , a constant. Show that the solutions are dependent (one a multiple of the other) if  $C = 0$ , and independent if  $C \neq 0$ .

Use the method of variation of parameters to find the solution in integral form of

$$y'' + \omega^2 y = -f(x), \quad y(0) = y'(0) = 0.$$

By constructing the Green's function, find the solution of

$$y'' + \omega^2 y = -f(x), \quad y(0) = y(1) = 0,$$

in the form

$$y = \int_0^1 G(x, \xi) f(\xi) d\xi.$$

6.2 The integral operator  $\mathcal{G}$  is defined on the space of continuous functions on  $[a, b]$ ,  $C[a, b]$ , by

$$\mathcal{G}u = \int_a^b G(x, \xi)r(\xi)u(\xi) d\xi,$$

where  $G$  is real, symmetric and continuous on  $[a, b] \times [a, b]$ , and  $r$  is positive and continuous on  $[a, b]$ .

Suppose that the eigenfunctions  $\phi_n$  with corresponding eigenvalues  $\lambda_n$  of

$$(I - \lambda\mathcal{G})\phi = 0$$

form a complete, orthonormal sequence, and that a continuous function  $f$  has a Fourier expansion

$$f = \sum_i c_i \phi_i.$$

By writing  $(I - \lambda\mathcal{G})^{-1} = \sum_n \lambda^n \mathcal{G}^n$ , show that a formal solution of the inhomogeneous Fredholm integral equation

$$(I - \lambda\mathcal{G})y = f$$

can be found in the form

$$y = \sum_i \frac{c_i \lambda_i \phi_i}{\lambda_i - \lambda}. \quad (*)$$

Show that the derivation is only valid for  $|\lambda| < \frac{1}{\|\mathcal{G}\|}$ , but that the series in  $(*)$  converges for all  $\lambda$  except the eigenvalues (where there are simple poles).

6.3 In deriving the Neumann series

$$y = \sum_0^\infty \lambda^n \mathcal{G}^n g$$

for the solution of

$$(I - \lambda\mathcal{G})y = g,$$

we implicitly assumed that the operator  $\mathcal{G}$  is associative, that is,  $\mathcal{G} \circ \mathcal{G}^n = \mathcal{G}^n \circ \mathcal{G}$  for all positive integral  $n$ . Demonstrate this explicitly, first for  $n = 2$ , and then by induction for all  $n$ , by deriving an expression for the kernel of the integral operator

$$\mathcal{G}^n u = \int_a^b G^{(n)}(x, \xi)u(\xi) d\xi.$$

6.4 An integral operator is defined by

$$\mathcal{G}u = \int_a^b G(x, \xi)r(\xi)u(\xi) d\xi,$$

where  $G$  is real, symmetric and continuous,  $r$  is positive and continuous, and  $u$  is continuous.

Show that

$$\|u\| = \left\{ \int_a^b r u^2 d\xi \right\}^{1/2}$$

defines a norm for  $u$ , and that

$$\|\mathcal{G}\| = \sup \frac{\|\mathcal{G}u\|}{\|u\|}$$

defines a norm for  $\mathcal{G}$ . Hence show that

$$\|\mathcal{G}\| \leq \left\{ \int_a^b \int_a^b r(\xi)r(\eta)G^2(\xi,\eta) d\xi d\eta \right\}^{1/2}.$$

If  $G$  is the Green's function calculated for Bessel's equation of order zero on the interval  $[0, 1]$ , is  $\|\mathcal{G}\|$  bounded? (See also question 5.2.)

By consideration of the kernel

$$G_n(x, \xi) = G(x, \xi) - \sum_1^{n-1} \frac{\phi_j(x)\phi_j(\xi)}{\lambda_j},$$

or otherwise, derive the Bessel inequality

$$\sum_1^\infty \frac{1}{\lambda_j^2} \leq \int_a^b \int_a^b G^2(x, \xi)r(x)r(\xi) dx d\xi,$$

and deduce that  $\lambda_n \rightarrow \infty$  as  $n \rightarrow \infty$ .



# Chapter 7

## Waves and shocks

Any introductory course on partial differential equations will provide the classification of second order partial differential equations into the three categories: elliptic, parabolic, hyperbolic; and one also finds the three simple representatives of these: Laplace's equation  $\nabla^2 u = 0$ , governing steady state temperature distribution (for example); the heat equation  $u_t = \nabla^2 u$ , which describes diffusion of heat (or solute); and the wave equation  $u_{tt} = \nabla^2 u$ , which describes the oscillations of a string or of a drum. These equations are of fundamental importance, as they describe diffusion or wave propagation in many other physical processes, but they are also linear equations; however the way in which they behave carries across to nonlinear equations, but of course nonlinear equations have other behaviours as well.

In the linear wave equation (in one dimension, describing waves on strings)  $u_{tt} = c^2 u_{xx}$ , the general solution is  $u = f(x+ct) + g(x-ct)$ , and represents the superposition of two travelling waves of speed  $c$ . In more than one space dimension, the equivalent model is  $u_{tt} = c^2 \nabla^2 u$ , and the solutions are functions of  $(\mathbf{k} \cdot \mathbf{x} \pm \omega t)$ , where  $\omega$  is frequency and  $\mathbf{k}$  is wave vector; the wave speed is then  $c = \omega/|\mathbf{k}|$ .

Even simpler to discuss is the first order wave equation

$$u_t + cu_x = 0, \quad (7.1)$$

which is trivially solved by characteristics to give

$$u = f(x - ct), \quad (7.2)$$

representing a wave of speed  $c$ . The idea of finding characteristics generalises to systems of the form

$$A\mathbf{u}_t + B\mathbf{u}_x = \mathbf{0}, \quad (7.3)$$

where  $\mathbf{u} \in \mathbf{R}^n$  and  $A$  and  $B$  are constant  $n \times n$  matrices. We can solve this system as follows. The eigenvalue problem

$$\lambda A\mathbf{w} = B\mathbf{w} \quad (7.4)$$

will in general have  $n$  solution pairs  $(\mathbf{w}, \lambda)$ , where each value of  $\lambda$  is one of the roots of the  $n$ -th order polynomial

$$\det(\lambda A - B) = 0. \quad (7.5)$$

Suppose the  $n$  values of  $\lambda$ ,  $\lambda_i$ , are distinct (which is the general case); then the corresponding  $\mathbf{w}_i$  are independent, and the matrix  $P$  formed by the eigenvectors as columns (i.e.,  $P = (\mathbf{w}_1, \dots, \mathbf{w}_n)$ ) satisfies  $BP = APD$ , where  $D$  is the diagonal matrix  $\text{diag}(\lambda_1, \dots, \lambda_n)$ .  $P$  is invertible, and if we write  $\mathbf{v} = P^{-1}\mathbf{u}$ , then  $AP\mathbf{v}_t + BP\mathbf{v}_x = 0$ , whence  $\mathbf{v}_t + D\mathbf{v}_x = 0$ , and the general solution is

$$\mathbf{u} = P \left[ \sum_i f_i(x - \lambda_i t) \right], \quad (7.6)$$

representing the superposition of  $n$  travelling waves with speeds  $\lambda_i$ . This procedure works providing  $A$  is invertible, and also if all the  $\lambda_i$  are real, in which case we say the system is hyperbolic.

More precisely, we can use the above prescription to solve the nonlinear equation

$$A\mathbf{u}_t + B\mathbf{u}_x = c(x, t, \mathbf{u}), \quad (7.7)$$

where we allow  $A$  and  $B$  to depend on  $x$  and  $t$  also. The diagonalisation procedure works exactly as before, leading to

$$A \frac{\partial}{\partial t}(P\mathbf{v}) + B \frac{\partial}{\partial x}(P\mathbf{v}) = c[x, t, P\mathbf{v}]; \quad (7.8)$$

now, however,  $\lambda$ ,  $\mathbf{w}$  and therefore also  $P$  will depend on  $x$  and  $t$ . Thus we find

$$\mathbf{v}_t + D\mathbf{v}_x = P^{-1}A^{-1}\mathbf{c} - [P^{-1}P_t + DP^{-1}P_x]\mathbf{v}, \quad (7.9)$$

and the components of  $\mathbf{v}$  can be solved as a set of coupled ordinary differential equations along the characteristics  $dx/dt = \lambda_i$ .

If  $A$  and  $B$  depend also on  $\mathbf{u}$ , the procedure is less clear for systems. However, the method of characteristics always works in one dimension, so we now return our attention to this case. Consider as an example the nonlinear evolution equation

$$u_t + uu_x = 0, \quad (7.10)$$

to be solved on the whole real axis. The method of characteristics leads to the implicitly defined general solution

$$u = f(x - ut), \quad (7.11)$$

which is analogous to (7.2), and represents a wave whose *speed depends on its amplitude*. Thus higher orders of  $u$  propagate more rapidly, and this leads to the wave steepening depicted in figure 7.1.

In fact, it can be seen that eventually  $u$  becomes multi-valued, and this signifies a break down of the solution. The usual way in which this multi-valuedness is avoided is to allow the formation of a *shock*, which is a point of discontinuity of  $u$ . the characteristic solution applies in front of and behind the shock, and the characteristics intersect at the shock, whose propagation forwards is described by an appropriate *jump condition*: see figure 7.2.

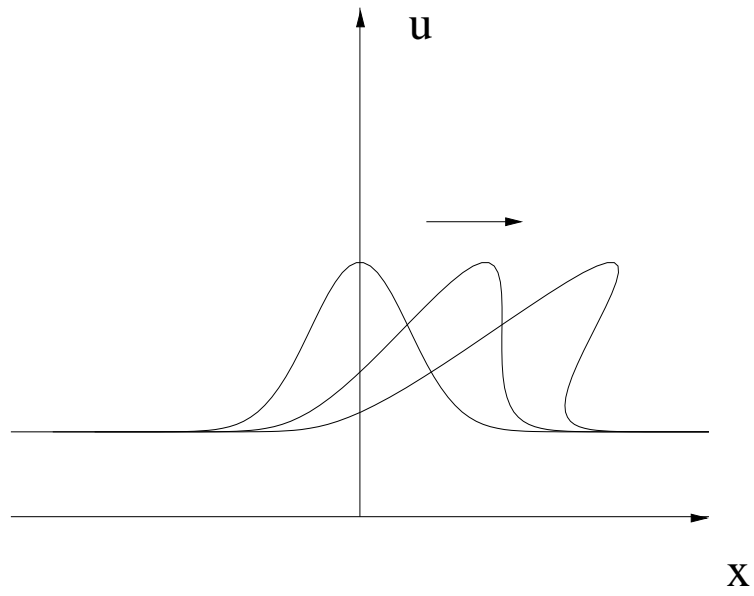


Figure 7.1: Nonlinearity causes wave steepening.

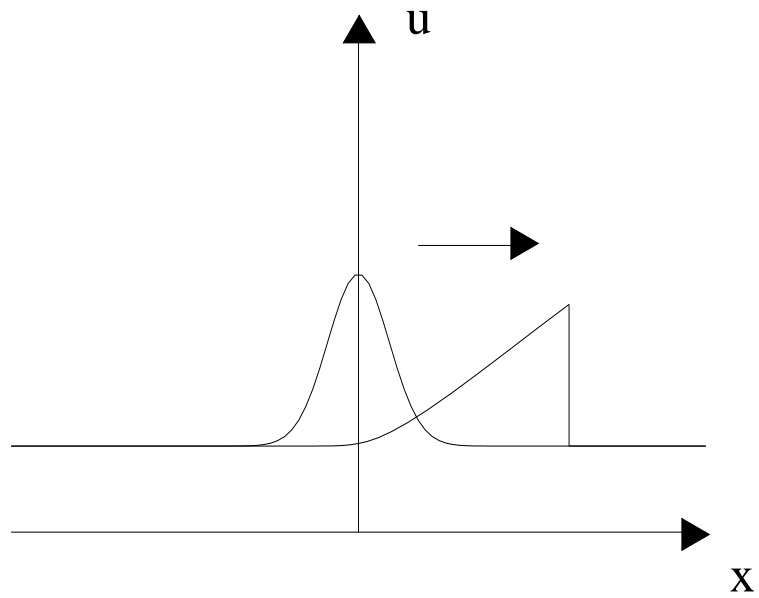


Figure 7.2: Intersection of characteristics leads to shock formation.

This seemingly arbitrary escape route is motivated by the fact that evolution equations such as (7.10) are generally derived from a conservation law of the form

$$\frac{d}{dt} \int_A^B \rho dx = -[q]_A^B, \quad (7.12)$$

where the square-bracketed term represents the jump in  $q$  between  $A$  and  $B$ ; (7.10) might be derived from this equation with  $\rho = u$  and  $q = \frac{1}{2}u^2$ , for example. The deduction of the point form

$$\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x} = 0 \quad (7.13)$$

from (7.12) requires the additional assumption that  $\rho$  and  $q$  are continuously differentiable; however, it is possible to satisfy (7.12) at a point of discontinuity of  $\rho$ . Suppose  $\rho$  is discontinuous at  $x = x_S(t)$ , and denote the jump in any quantity  $r$  across the shock as  $[r]^\pm = r(x_S+, t) - r(x_S-, t)$ . Then let  $A = x_S(t)-$  and  $B = x_S(t + \delta t)+$ .

From first principles we have the change of  $\int_A^B \rho dx$  in the time interval  $\delta t$  as

$$\delta \int_A^B \rho dx \approx [\rho]_+^- \delta x_S \approx [q]_+^- \delta t, \quad (7.14)$$

and from this it follows that

$$\dot{x}_S = \frac{[q]_+^-}{[\rho]_+^-}. \quad (7.15)$$

For the particular case of (7.10), we then find

$$\dot{x}_S = \frac{[\frac{1}{2}u^2]_+^-}{[u]_+^-} = \frac{1}{2}(u_+ + u_-). \quad (7.16)$$

### An example

We illustrate how to solve a problem of this type by considering the initial value

$$u = u_0(x) = \frac{1}{1+x^2} \quad \text{at } t = 0. \quad (7.17)$$

The implicitly defined solution is then

$$u = \frac{1}{1+(x-ut)^2}, \quad (7.18)$$

or, in characteristic form,

$$u = u_0(\xi) = \frac{1}{1+\xi^2}, \quad x = \xi + ut. \quad (7.19)$$

This defines a single-valued function so long as  $u_x$  is finite. Differentiating (7.19) leads to

$$u_x = \frac{u'_0(\xi)}{1+tu'_0(\xi)}, \quad (7.20)$$

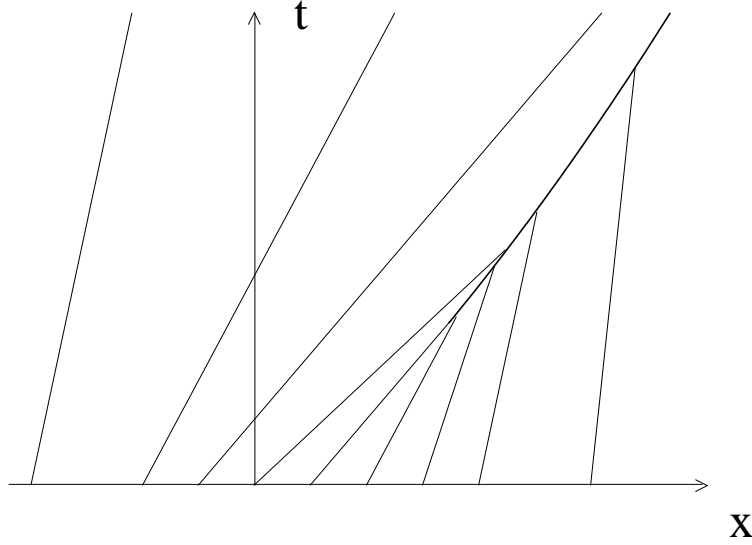


Figure 7.3: Characteristic diagram indicating shock formation.

and this shows that  $u_x \rightarrow -\infty$  as  $t \rightarrow t_c = \min_{\xi: u'_0 < 0} \left[ -\frac{1}{u'_0(\xi)} \right]$ . Since  $-u'_0 = 2\xi/(1 + \xi^2)^2$ , we find the relevant value of  $\xi$  is  $1/\sqrt{3}$ , and thus  $t_c = 8/3\sqrt{3}$  and the corresponding value of  $x$  is  $x_c = \sqrt{3}$ . Thus (7.18) applies while  $t < t_c = 8/3\sqrt{3}$ , and thereafter the solution also applies in  $x < x_S(t)$  and  $x > x_S(t)$ , where

$$\dot{x}_S = \frac{1}{2}[u(x_S+) + u(x_S-)], \quad (7.21)$$

with

$$x_S = \sqrt{3} \quad \text{at} \quad t = \frac{8}{3\sqrt{3}}. \quad (7.22)$$

As indicated in figure 7.3, the characteristics intersect at the shock, and it is geometrically clear from figure 7.1, for example, that  $u_+$  and  $u_-$  are the largest and smallest roots of the cubic (7.18). An explicit solution for  $x_S$  is not readily available, but it is of interest to establish the long term behaviour, and for this we need approximations to the roots of (7.17) when  $t \gg 1$ .

We write the cubic in the form

$$u = \frac{x}{t} \pm \frac{1}{t} \left( \frac{1-u}{u} \right)^{1/2}. \quad (7.23)$$

We know that  $u \leq 1$ , and we expect the largest root, at least, to tend to infinity as  $t \rightarrow \infty$ . In that case  $u \approx x/t$  if  $u = O(1)$ , and the next corrective term gives

$$u \approx \frac{x}{t} \pm \frac{1}{t} \left( \frac{t-x}{x} \right)^{1/2}. \quad (7.24)$$

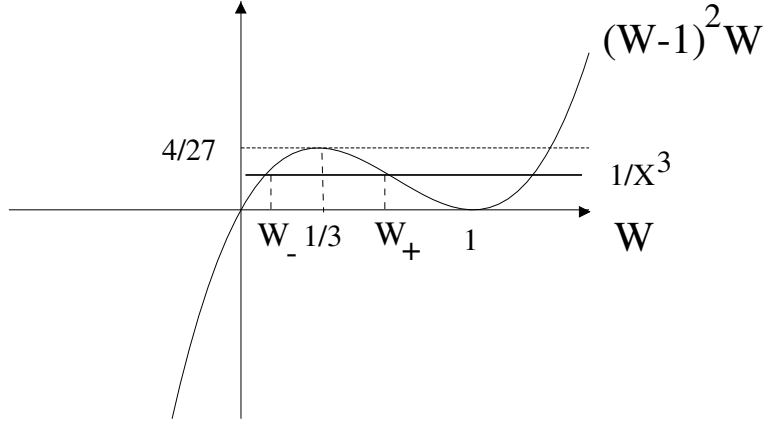


Figure 7.4: Determination of  $W(X)$ .

This evidently gives the upper two roots for  $x < t$  (since they coalesce at  $u = 1$  when  $x = t$ ). For large  $x$ , the other root must have  $u \ll 1$ , and in fact

$$u \approx \frac{1}{x^2}, \quad (7.25)$$

in order that (7.23) not imply (7.24). Alternatively, (7.25) follows from consideration of (7.18) in the form

$$t^2 u^3 - 2xtu^2 + (x^2 + 1)u - 1 = 0, \quad (7.26)$$

providing  $x \gg t^{1/3}$ .

(7.25) gives the right hand nose of the (rightmost) curve in figure 7.1. The left nose can be determined from the observation that the approximation that  $u \approx x/t$  breaks down (from (7.24)) when  $x \sim t^{1/3}$ , which is also where (7.25) becomes invalid. This suggests writing

$$u = \frac{x}{t} W(X), \quad X = \frac{x}{t^{1/3}}, \quad (7.27)$$

and then  $W(X)$  is given approximately, for large  $t$ , by

$$W(W-1)^2 = \frac{1}{X^3}, \quad (7.28)$$

and for  $X = O(1)$  there are three roots providing  $X > 3/2^{2/3}$ ; at  $X = 3/2^{2/3}$ , the two lower roots coalesce at  $W = \frac{1}{3}$ : this is the left nose of the curve.

As  $X$  becomes large, the upper two roots approach  $W = 1$ , i.e.  $u \approx x/t$ , while the lower approaches zero, specifically  $W \approx 1/X^3$ , i.e.  $u \approx 1/x^2$ : see figure 7.4. Thus these roots match to the approximations when  $x \sim t$ . As  $X$  becomes small, the remaining root is given by  $W \approx 1/X$ , i.e.  $u \approx 1/t^{2/3}$ , and (7.26) shows that this is the correct approximation as long as  $|x| \ll t^{1/3}$ . The situation is shown in figure 7.5).

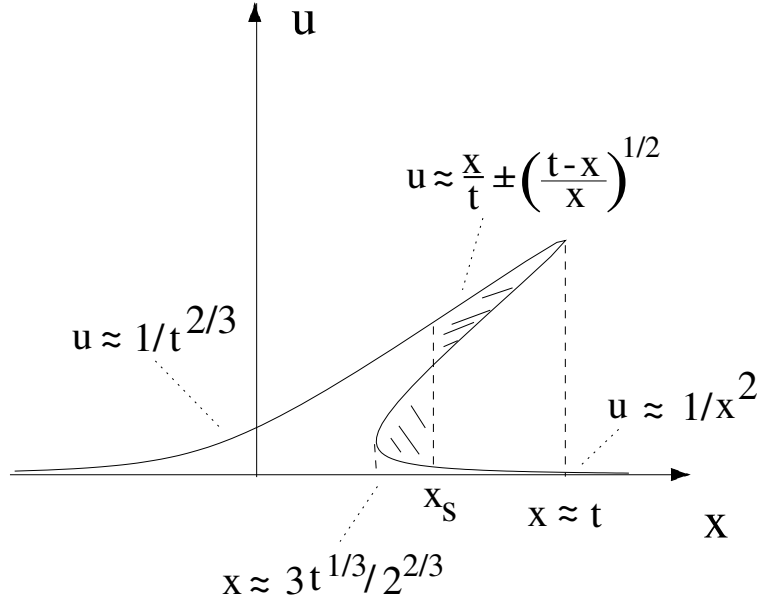


Figure 7.5: Large time solution of the characteristic solution.

In order to determine the shock location  $x_S$ , we make the ansatz that  $t^{1/3} \ll x_S \ll t$ , i. e., that the shock is far from both noses. In that case

$$u_+ \approx \frac{1}{x_S^2}, \quad u_- \approx \frac{x_S}{t}, \quad (7.29)$$

and at leading order we have

$$\dot{x}_S \approx \frac{x_S}{2t}, \quad (7.30)$$

whence

$$x_S \approx at^{1/2}, \quad (7.31)$$

confirming our assumption that  $t^{1/3} \ll x_S \ll t$ .

To determine the coefficient  $a$ , we may use the equal area rule, which follows from conservation of mass, and implies that the two shaded areas in figure 7.5 are equal. We use (7.27) for the left hand area, and (7.24) for the right hand area. then

$$\int_{3t^{1/3}/2^{2/3}}^{at^{1/2}} \frac{x}{t} [w_+(X) - w_-(X)] dx \approx \int_{at^{1/2}}^t \frac{2}{t} \left( \frac{t-x}{x} \right)^{1/2} dx, \quad (7.32)$$

where  $W_+$  and  $W_-$  are the middle and lowest roots of (7.28), as shown in figure 7.4. We write  $x = t^{1/2}\xi$  in the left integral and  $x = t\eta$  in the right, and hence we deduce that

$$a \approx \int_0^1 2 \left( \frac{1-\eta}{\eta} \right)^{1/2} d\eta = \pi. \quad (7.33)$$

## 7.1 Traffic flow

Although the presence of a shock for (7.10) is entirely consistent with the derivation of the equation from an integral conservation law, nature appears generally to avoid discontinuities and singularities, and it is usually the case that in writing (7.10), we have neglected some term which acts to smooth the shock, so that the change of  $u$  is rapid but not abrupt.

A simple example of this arises in the theory of traffic flow on a road. On a one-lane carriageway, the density of cars is  $\rho(x, t)$  (number per unit length, idealised as a continuum, which should be a reasonable assumption for road lengths of many cars);  $x$  is distance along the carriageway, and  $t$  is time. If  $v(x, t)$  is the local traffic vehicular speed, then a conservation law for cars is

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho v) = 0, \quad (7.34)$$

assuming no cars leave or join the carriageway. The car speed must be prescribed, and a simple and sensible prescription is to take  $v = v(\rho)$ : a driver drives at a speed which depends on density. More specifically, we might suppose that there is a maximum speed of  $v_0$  when there is no traffic ( $\rho = 0$ ), but the speed decreases as  $\rho$  increases, reaching zero when the distance between cars is zero, corresponding to a maximum density  $\rho_m$  say. A simple recipe which satisfies these constraints is

$$v = v_0 \left( \frac{\rho}{\rho_m} \right). \quad (7.35)$$

By scaling the variables suitably, we may take  $v_0 = 1$ ,  $\rho_m = 1$ , and then we have

$$\rho_t + c(\rho)\rho_x = 0, \quad (7.36)$$

where

$$c(\rho) = 1 - 2\rho. \quad (7.37)$$

This is equivalent to (7.10), with  $u = 1 - 2\rho$ , and therefore we can in general expect shocks to form. Since  $0 < \rho < 1$ , a difference here is that we can have  $c < 0$  for  $\rho > \frac{1}{2}$ .

Suppose first that  $\rho < \frac{1}{2}$ . A local reduction in  $\rho$  is like a bump for  $u = 1 - 2\rho$ , and will form a forward propagating shock where  $\rho_+ > \rho_-$ ; cars driving through the low density precursor will suddenly hit a traffic jam at high density. Equivalently, a local rise in density leads to a forward-propagating shock with  $\rho_+ > \rho_-$ .

On the other hand, suppose that  $\rho > \frac{1}{2}$ . With  $w = 2\rho - 1$ , then  $w_t - ww_x = 0$ , and a similar discussion arises, with the direction of  $x$  reversed. Thus a local rise in  $\rho$  causes a shock to form in which, again,  $\rho_+ > \rho_-$ , and the shock now propagates backwards. In all cases, the individual driver experiences the shock as a sudden reduction of speed (the car speed is  $1 - \rho$  and is larger than the wave speed  $1 - 2\rho$ : thus cars approach and travel through the shock).

The occurrence of shocks is a common experience to anyone who has driven on a motorway, and can be caused by lane closures, motorway junctions, or speed restrictions. Figure 7.6 shows a beautiful series of shock waves propagating backwards in



morning rush hour traffic on the M25. Beautiful, unless you happen to be in one of the vehicles. The diagram actually suggests that the waves arise through an instability, much like roll waves on rivers, but these cannot be described within the confines of the simple kinematic wave model (7.36).

In general (but not quite always), collisions do not occur, and one may ascribe this to a smoothing mechanism; the question arises, what could this be?

Drivers arguably adjust their speed not only according to the inter-car distance (the direct density effect), but also depending on how much traffic they see ahead. A common exhibition of this is the tailgater: driving at the speed limit on a motorway in the outside lane will attract tailgaters if there is little traffic ahead, but not (as much) if the road is full. That is, at a constant  $v$ ,  $\rho$  may increase above its ideal value if  $\rho$  in front is less, i.e. if  $\partial\rho/\partial x < 0$ . A simple modification for  $v$  which describes this is (in scaled units)

$$v = 1 - \rho - \kappa\rho_x, \quad (7.38)$$

where  $\kappa$  is constant. From this we can deduce the advection-diffusion equation

$$\frac{\partial\rho}{\partial t} + (1 - 2\rho)\frac{\partial\rho}{\partial x} = \kappa\frac{\partial}{\partial x}\left(\rho\frac{\partial\rho}{\partial x}\right), \quad (7.39)$$

and the non-local dependence of  $v$  on  $\rho_x$  is represented as a (nonlinear) diffusion term.

What is the effect of this on the structure of the solutions? If  $\kappa$  is small, we should suppose that it is not much, so that shocks would start to form. However, the neglect of the diffusion term becomes invalid when the derivatives of  $\rho$  become large. In fact, the diffusion term is trying to do the opposite of the advective term. The latter is trying to fold the initial profile together like an accordion, while the former is trying to spread everything apart. We might guess that a balanced position is possible, in which the nonlinear advective term keeps the profile steep, but the diffusion prevents it actually folding over (and hence causing a discontinuity), and this will turn out to be the case.

If we suppose instead of (7.38) that

$$v = 1 - \rho - \frac{\kappa}{\rho}\rho_x, \quad (7.40)$$

which might represent the fact that the strength of tailgating becomes more severe, the emptier the road ahead, then the diffusion term becomes linear, and we have Burgers' equation for  $u = 1 - 2\rho$ :

$$u_t + uu_x = \kappa u_{xx}. \quad (7.41)$$

## 7.2 Burgers' equation

### Shock structure

We suppose  $\kappa \ll 1$ , and that  $u_t + uu_x \approx 0$ , and a shock forms at  $x = x_S(t)$ . Our aim is to show that (7.41) supports a *shock structure*, i.e. a region of radical change near  $x_S$  for  $u_-$  to  $u_+$ .

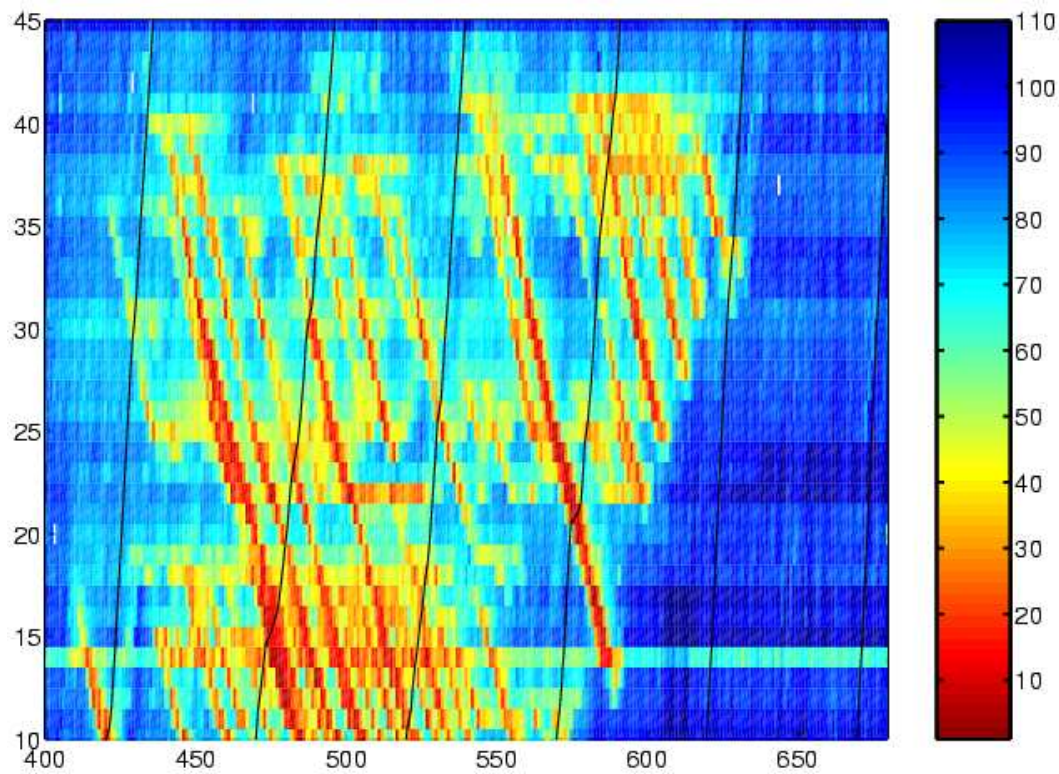


Figure 7.6: Rush hour traffic waves propagating backward on the M25. Horizontal axis is minutes from midnight, vertical axis is “outstation number”. Separation of outstations is 500 m, so the diagram shows 17.5 km of motorway between about 6.40 a.m. and a little after 11 a.m.: this is a typical weekday picture. The picture is built from 1 minute averages of the traffic speed (in lane 2) on the clockwise carriageway of the M25 in the neighbourhood of the M3 and M4 interchanges. (The seed point of the stop and go waves near the top of the picture is the M4 interchange.) Colour is speed in  $\text{km h}^{-1}$ , according to colour bar next to picture. Black lines are “simulated vehicle trajectories”, meaning their gradient at any point is the 1 minute average vehicle speed. The picture was made by Eddie Wilson in Matlab from data belonging to the Highways Agency. The data is captured by their “MIDAS system” which is part of the “Controlled Motorways” project.

To focus on the shock, we need to rescale  $x$  near  $x_S$ , and we do this by writing

$$x = x_S(t) + \kappa X. \quad (7.42)$$

We use the chain rule to rewrite the equations in terms of  $t$  and  $X$  derivatives. To be precise, we consider the change of variables  $t, x \longrightarrow \tau, X$ , where (7.42) defines  $X$  in terms of  $t$  and  $x$ , and  $\tau = t$ . Elementary application of the chain rule for partial derivatives then shows that

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial \tau} - \frac{\dot{x}_S}{\kappa} \frac{\partial}{\partial X}, \quad \frac{\partial}{\partial x} = \frac{1}{\kappa} \frac{\partial}{\partial X}. \quad (7.43)$$

Burgers' equation becomes

$$\kappa u_t - \dot{x}_S u_X + uu_X = u_{XX}, \quad (7.44)$$

where we replace  $\tau$  by  $t$  since these are in fact the same. We expect the characteristic solution (with  $\kappa = 0$ ) to be approximately valid far from  $x_s$ , and so appropriate conditions (technically, these are *matching conditions*) are

$$u \rightarrow u_{\pm} \text{ as } X \rightarrow \pm\infty, \quad (7.45)$$

and we take these values as prescribed from the *outer* solution (i.e., the solution of  $u_t + uu_x = 0$ ).

Since  $\kappa \ll 1$ , (7.44) suggests that  $u$  relaxes rapidly (on a time scale  $t \sim \kappa \ll 1$ ) to a quasi-steady state (quasi-steady, because  $u_+$  and  $u_-$  will vary with  $t$ ) in which

$$-\dot{x}_S u_X + uu_X \approx u_{XX}, \quad (7.46)$$

whence

$$K - \dot{x}_S u + \frac{1}{2} u^2 \approx u_X, \quad (7.47)$$

and prescription of the boundary conditions implies

$$K = \dot{x}_S u_+ - \frac{1}{2} u_+^2 = \dot{x}_S u_- - \frac{1}{2} u_-^2, \quad (7.48)$$

whence

$$\dot{x}_S = \frac{[\frac{1}{2} u^2]_-^+}{[u]_-^+}, \quad (7.49)$$

which is precisely the jump condition we obtained in (7.16). The solution for  $u$  is then

$$u = c - (u_- - c) \tanh \left[ \frac{1}{2} (u_- - c) X \right], \quad (7.50)$$

where  $c = \dot{x}_S$ .

### 7.3 The Fisher equation

In Burgers' equation, a wave arises as a balance between nonlinear advection and diffusion. In Fisher's equation,

$$u_t = u(1 - u) + u_{xx}, \quad (7.51)$$

a wave arises as a mechanism of transferring a variable from an unstable steady state ( $u = 0$ ) to a stable one ( $u = 1$ ). Whereas Burgers' equation balances two transport terms, Fisher's equation balances diffusive transport with an algebraic source term. It originally arose as a model for the dispersal of an advantageous gene within a population, and has taken a plenary rôle as a pedagogical example in mathematical biology of how reaction (source terms) and diffusion can combine to produce travelling waves.

We pose (7.51) with boundary conditions

$$\begin{aligned} u &\rightarrow 1, \quad x \rightarrow -\infty, \\ u &\rightarrow 0, \quad x \rightarrow +\infty. \end{aligned} \quad (7.52)$$

It is found (and can be proved) that any initial condition leads to a solution which evolves into a travelling wave of the form

$$u = f(\xi), \quad \xi = x - ct, \quad (7.53)$$

where

$$f'' + cf' + f(1 - f) = 0, \quad (7.54)$$

and

$$f(\infty) = 0, \quad f(-\infty) = 1. \quad (7.55)$$

In the  $(f, g)$  phase plane, where  $g = -f'$ , we have

$$\begin{aligned} f' &= -g, \\ g' &= f(1 - f) - cg, \end{aligned} \quad (7.56)$$

and a travelling wave corresponds to a trajectory which moves from  $(1, 0)$  to  $(0, 0)$ .

Linearisation of (7.56) near the fixed point  $(f^*, 0)$  via  $f = f^* + F$  leads to

$$\begin{pmatrix} F \\ g \end{pmatrix}' = \begin{pmatrix} 0 & -1 \\ 1 - 2f^* & -c \end{pmatrix} \begin{pmatrix} F \\ g \end{pmatrix}, \quad (7.57)$$

with solutions  $e^{\lambda\xi}$ , where  $\lambda^2 + c\lambda + (1 - 2f^*) = 0$ . We anticipate  $c > 0$ ; then  $(1, 0)$  is a saddle point, while  $(0, 0)$  is a stable node if  $c \geq 2$  (and a spiral if  $c < 2$ ). For  $c \geq 2$ , a connecting trajectory exists as shown in figure 7.7: in practice the minimum wave speed  $c = 2$  is selected. (Connecting trajectories also exist if  $c < 2$ , but because  $(0, 0)$  is a spiral, these have oscillating tails as  $u \rightarrow 0$ , which are unstable and also (for example, if  $u$  is a population) unphysical.)

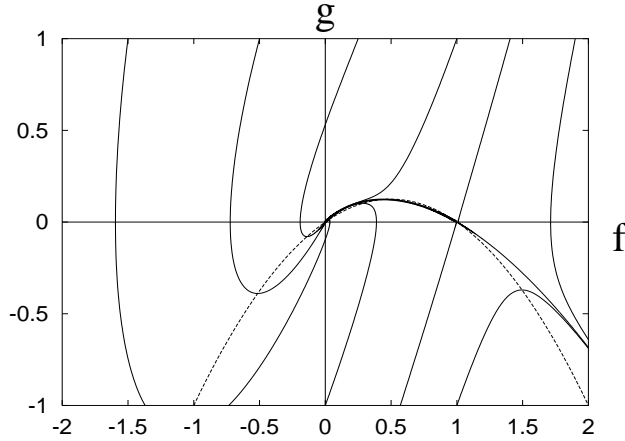


Figure 7.7: Phase portrait of Fisher equation, (7.56), for  $c = 2$ . Note how close the connecting trajectory (thick line) is to the  $g$  nullcline. This is why the large  $c$  approximation is accurate *for this trajectory*.

Explicit solutions for (7.54) are not available, but an excellent approximation is easily available. We put

$$\xi = c\Xi, \quad (7.58)$$

so

$$\nu f'' + f' + f(1 - f) = 0, \quad (7.59)$$

with  $\nu = 1/c^2 = 1/4$  for  $c = 2$ . Taking  $\nu \ll 1$  and writing  $f = f_0 + \nu f_1 + \dots$ , we have

$$\begin{aligned} f_0' + f_0(1 - f_0) &= 0, \\ f_1' + (1 - 2f_0)f_1 &= -f_0'', \end{aligned} \quad (7.60)$$

and thus

$$f_0 = \frac{e^{-\Xi}}{1 + e^{-\Xi}}. \quad (7.61)$$

Also, noting that  $1 - 2f_0 = -f_0''/f_0'$  (differentiate (7.60)<sub>1</sub>),

$$f_1 = f_0(1 - f_0) \ln[f_0(1 - f_0)], \quad (7.62)$$

and so on. Even the first term gives a good approximation, and even for  $c = 2$ .

### 7.3.1 Stability

## 7.4 Solitons

The Fisher wave is an example of a solitary travelling wave. Another type of solitary wave is the soliton, as exemplified by solutions of the Korteweg-de Vries equation

$$u_t + uu_x + u_{xxx} = 0. \quad (7.63)$$

This has travelling wave solutions  $u = f(\xi)$ ,  $\xi = x - ct$ , where

$$f''' + ff' - cf' = 0, \quad (7.64)$$

and solitary waves with  $f \rightarrow 0$  at  $\pm\infty$  satisfy the first integral

$$f'' + \frac{1}{2}f^2 - cf = 0, \quad (7.65)$$

and thus

$$\frac{1}{2}f'^2 + \frac{1}{6}f^3 - \frac{1}{2}cf^2 = 0, \quad (7.66)$$

with solution

$$f = \frac{3}{2}c \operatorname{sech}^2 \left( \frac{\sqrt{c}\xi}{2} \right). \quad (7.67)$$

Thus there is a one parameter family of these solitary waves, and they are called solitons, because they have the remarkable particle-like ability to ‘pass through’ each other without damage, except for a change of relative phase. Despite the nonlinearity, they obey a kind of superposition principle. Soliton equations (of which there are many) have many other remarkable properties, beyond the scope of the present discussion.

Some understanding of the solitary wave arises through an understanding of the balance between nonlinearity ( $uu_x$ ) and dispersion ( $u_{xxx}$ ). The dispersive part of the equation,  $u_t + u_{xxx} = 0$ , is so called because waves  $\exp[ik(x - ct)]$  have wave speed  $c = -k^2$  which depends on wavenumber  $k$ ; waves of different wavelengths ( $2\pi/k$ ) move at different speeds and thus disperse. On the other hand, the nonlinear advection equation  $u_t + uu_x$  has a focussing effect, which (from a spectral point of view) concentrates high wave numbers near shocks (rapid change means large derivatives means high wavenumber). So the nonlinearity tries to move high wavenumber modes in from the left, while the dispersion tries to move them to the left: again a balance is struck, and a travelling wave is the result.

## 7.5 Snow melting

An example of some of the ideas presented so far occurs in the study of melting snow. In particular, it is found that pollutants which may be uniformly distributed in snow (e.g.  $\text{SO}_2$  from sulphur emissions via acid rain) can be concentrated in melt water run-off, with a consequent enhanced detrimental effect on stream pollution. The question then arises, why this should be so.

The model we use is based on the principles of groundwater flow. Suppose we have a snow pack of depth  $h$  in  $0 < z < h$ , where  $z$  is a coordinate pointing downwards from the snow surface. Snow is a porous aggregate of ice crystals, and meltwater formed at the surface can percolate through the snow pack to the base, where run-off occurs. (We ignore effects of re-freezing of meltwater). The flux of water  $u$  downwards is given by Darcy’s law

$$u = \frac{k}{\mu} \left[ -\frac{\partial p}{\partial z} + \rho g \right], \quad (7.68)$$

where  $u$  is measured as a velocity, and represents volume per unit area per unit time;  $p$  is the (pore) water pressure,  $\rho$  is density,  $g$  is gravity,  $\mu$  is viscosity,  $k$  is permeability. The permeability  $k$  is related to the saturation  $S$ , and we will assume

$$k = k_0 S^3, \quad (7.69)$$

where the saturation  $S$  is the volume fraction of the pore space which is occupied by water. For  $S = 1$ , the snow is fully saturated (no air), for  $S = 0$  it is fully dry (no water).

Conservation of the liquid water implies

$$\phi \frac{\partial S}{\partial t} + \frac{\partial u}{\partial z} = 0, \quad (7.70)$$

where  $\phi$  is the porosity (volume of pore space per unit volume of snow). Finally, the water pressure is related to the air pressure ( $p_a$ , taken as constant) by the capillary pressure

$$p_c = p_a - p, \quad (7.71)$$

and this is a function of  $S$ : we take

$$p_c(S) = p_0 \left( \frac{1}{S} - S \right), \quad (7.72)$$

based on typical experimental results.

Suitable boundary conditions in a melting event might be to prescribe the melt flux at the surface:

$$u = u_0 \quad \text{at} \quad z = 0. \quad (7.73)$$

If the base is impermeable, then

$$u = 0 \quad \text{at} \quad z = h. \quad (7.74)$$

This is certainly not realistic if  $S$  reaches 1 at the base, since then ponding must occur and presumably melt drainage via a channelised flow, but we examine the initial stages of the flow using (7.74). Finally, we suppose  $S = 0$  at  $t = 0$ . Again, this is not realistic in the model (it implies  $p_c = \infty$ ) but it is a feasible approximation to make.

Simplification of this model now leads to the Darcy-Richards equation in the form

$$\phi \frac{\partial S}{\partial t} + \frac{3k_0 \rho g}{\mu} S^2 \frac{\partial S}{\partial z} = \frac{k_0 p_0}{\mu} \frac{\partial}{\partial z} \left[ S(1 + S^2) \frac{\partial S}{\partial z} \right], \quad (7.75)$$

which, as we see, is a convective-diffusion equation of Burgers type. The quantity  $K = k_0 \rho g / \mu$  is known as the saturated hydraulic conductivity; it is a velocity, and represents the highest rate at which water can flow through the snow steadily under gravity.

To scale the equation, we note that  $S$  is dimensionless and of  $O(1)$  by definition. We choose scales for  $z$  and  $t$ :

$$z \sim h, \quad t \sim \frac{\phi h}{K}; \quad (7.76)$$

the model then becomes

$$\frac{\partial S}{\partial t} + 3S^2 \frac{\partial S}{\partial z} = \kappa \frac{\partial}{\partial z} \left[ S(1 + S^2) \frac{\partial S}{\partial z} \right], \quad (7.77)$$

where

$$\kappa = \frac{p_0}{\rho g h}. \quad (7.78)$$

if we choose  $p_0 = 1$  kPa,  $\rho = 10^3$  kg m<sup>-3</sup>,  $g = 10$  m s<sup>-2</sup>,  $h = 1$  m, then  $\kappa = 0.1$ . it follows that (7.77) has a propensity to form shocks, these being diffused by the term in  $\kappa$  over a distance  $O(\kappa)$  (by analogy with the shock structure for the Burgers equation).

We want to solve (7.77) with the initial condition

$$S = 0 \quad \text{at} \quad t = 0, \quad (7.79)$$

and the boundary conditions

$$S^3 - \kappa S(1 + S^2) \frac{\partial S}{\partial z} = \frac{u_0}{K} \quad \text{on} \quad z = 0, \quad (7.80)$$

and

$$S^3 - \kappa S(1 + S^2) \frac{\partial S}{\partial z} = 0 \quad \text{at} \quad z = 1. \quad (7.81)$$

Roughly, for  $\kappa \ll 1$ , these are

$$\begin{aligned} S &= S_0 \quad \text{at} \quad z = 0, \\ S &= 0 \quad \text{at} \quad z = 1, \end{aligned} \quad (7.82)$$

where  $S_0 = (u_0/K)^{1/3}$ , which we initially take to be  $O(1)$  (and  $< 1$ , so that surface ponding does not occur).

Neglecting  $\kappa$ , the solution is the step function

$$\begin{aligned} S &= S_0, \quad z < z_f, \\ S &= 0, \quad z > z_f, \end{aligned} \quad (7.83)$$

and the shock front at  $z_f$  advances at a rate  $\dot{z}_f$ , given from the jump condition

$$\dot{z}_f = \frac{[S^3]_-^+}{[S]_-^+} = S_0^2. \quad (7.84)$$

In dimensional terms, the shock front moves at speed  $u_0/\phi S_0$ , which is in fact obvious (given that it has constant  $S$  behind it).

The shock structure is similar to that of Burgers' equation. We put

$$z = z_f + \kappa Z, \quad (7.85)$$

and  $S$  rapidly approaches the quasi-steady solution  $S(Z)$  of

$$-V S' + 3S^2 S' = [S(1 + S^2) S']', \quad (7.86)$$



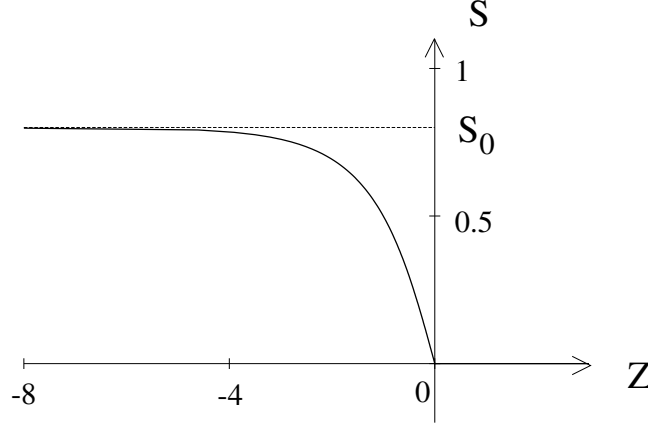


Figure 7.8:  $S(Z)$  given by (7.90); the shock front terminates at the origin.

where  $V = \dot{z}_f$ ; hence

$$S(1 + S^2)S' = -S(S_0^2 - S^2), \quad (7.87)$$

in order that  $S \rightarrow S_0$  as  $Z \rightarrow -\infty$ , and where we have chosen

$$V = S_0^2, \quad (7.88)$$

(as  $S_+ = 0$ ), thus reproducing (7.84). The solution is a quadrature,

$$\int^S \frac{(1 + S^2)dS}{(S_0^2 - S^2)} = -Z, \quad (7.89)$$

with an arbitrary added constant (amounting to an origin shift for  $Z$ ). Hence

$$S - \frac{(1 + S_0^2)}{2S_0} \ln \left[ \frac{S_0 + S}{S_0 - S} \right] = Z. \quad (7.90)$$

The shock structure is shown in figure 7.8, and one particular feature is noteworthy; the profile terminates where  $S = 0$  at  $Z = 0$ . In fact, (7.87) implies that  $S = 0$  or (7.90) applies. Thus when  $S$  given by (7.90) reaches zero, the solution switches to  $S = 0$ . The fact that  $\partial S / \partial Z$  is discontinuous is not a problem because the diffusivity  $S(1 + S^2)$  goes to zero when  $S = 0$ . In fact, this *degeneracy* of the equation is a signpost for fronts with discontinuous derivatives, and we shall encounter this situation again when we study non-linear diffusion. Essentially, the profile can maintain discontinuous gradients at  $S = 0$  because the diffusivity is zero there, and there is no mechanism to smooth the jump away.

Suppose now that  $k_0 = 10^{-9} \text{ m}^2$  (a plausible value) and  $\mu/\rho = 10^{-6} \text{ m}^2 \text{ s}^{-1}$ , then the saturated hydraulic conductivity  $K = k_0 \rho g / \mu = 10^{-2} \text{ m s}^{-1}$ . On the other hand, if a metre thick snow pack melts in ten days, this implies  $u_0 \sim 10^{-6} \text{ m s}^{-1}$ . Thus  $S_0^2 = u_0 / K \sim 10^{-4}$ , and the approximation  $S \approx S_0$  looks less realistic. With

$$S^3 - \kappa S(1 + S^2) \frac{\partial S}{\partial z} = S_0^3, \quad (7.91)$$

and  $S_0 \sim 10^{-2}$  and  $\kappa \sim 10^{-1}$ , it seems that one should assume  $S \ll 1$ . We define

$$S = \left( \frac{S_0^3}{\kappa} \right)^{1/2} s, \quad (7.92)$$

note that  $(S_0^3/\kappa)^{1/2} \sim 0.03$ , so that (7.91) becomes

$$\beta s^3 - s \left[ 1 + \frac{S_0^3}{\kappa} s^2 \right] \frac{\partial s}{\partial z} = 1 \quad \text{on } z = 0, \quad (7.93)$$

and we have  $S_0^3/\kappa \sim 10^{-3}$ ,  $\beta = (S_0/\kappa)^{3/2} \sim 0.3$ .

We neglect the term in  $S_0^3/\kappa$ , so that

$$\beta s^3 - s \frac{\partial s}{\partial z} \approx 1 \quad \text{on } z = 0, \quad (7.94)$$

and substituting (7.92) into (7.77) leads to

$$\frac{\partial s}{\partial \tau} + 3\beta s^2 \frac{\partial s}{\partial z} \approx \frac{\partial}{\partial z} \left[ s \frac{\partial s}{\partial z} \right]. \quad (7.95)$$

A simpler analytic solution is no longer possible, but the development of the solution will be similar. The flux condition (7.94) at  $z = 0$  allows the surface saturation to build up gradually, and a shock will only form if  $\beta \gg 1$  (when the preceding solution becomes valid).

## 7.6 Notes and references

A fundamental reference on the subject of waves in all kinds of media is the book by Whitham (1974). An older book which details waves in fluid is the book by Stoker (1957).

## Exercises

7.1 A simple model for the flow of a mixture of two fluids along a tube (e.g., air and water) is

$$\begin{aligned} \alpha_t + (\alpha v)_z &= 0 \\ -\alpha_t + [(1 - \alpha)u]_z &= 0 \\ \rho_g[(\alpha v)_t + (\alpha v^2)_z] &= -\alpha p_z, \\ \rho_l[\{(1 - \alpha)u\}_t + \{D_l(1 - \alpha)u^2\}_z] &= -(1 - \alpha)p_z, \end{aligned}$$

where  $p$  is pressure,  $u$  and  $v$  are the two fluid velocities,  $\alpha$  is the volume fraction of the fluid with speed  $v$ ,  $\rho_g$  is its density, and  $\rho_l$  is the density of the other fluid. Show that there are two characteristic speeds  $dz/dt = \lambda$ , satisfying

$$(\lambda - u)^2 = (D_l - 1)[u^2 + 2u(\lambda - u)] - s^2(\lambda - v)^2,$$

where

$$s = \left[ \frac{\rho_g(1 - \alpha)}{\rho_l \alpha} \right]^{1/2}.$$

Deduce that the characteristic speeds are real if, when  $D_l - 1 \ll 1$ ,  $s \ll 1$ ,

$$D_l \gtrsim 1 + \left\{ \frac{s(u - v)}{u} \right\}^2.$$

In particular, show that the roots are complex if  $D_l = 1$  and  $u \neq v$ . What does this suggest concerning the well-posedness of the model?

7.2 The function  $u(x, t)$  satisfies

$$u_t + uu_x = \alpha(1 - u^2)$$

for  $-\infty < x < \infty$ , where  $\alpha > 0$ , and with  $u = u_0(x)$  at  $t = 0$ , and  $0 < u_0 < 1$  everywhere. Show that the characteristic solution can be written parametrically in the form

$$u = \frac{u_0(s) + \tanh \alpha t}{1 + u_0(s) \tanh \alpha t}, \quad \exp[\alpha(x - s)] = \cosh \alpha t + u_0(s) \sinh \alpha t.$$

Sketch the form of the characteristics for an initial function such as  $u_0(s) = a/(1 + s^2)$ . Show that, in terms of  $s$  and  $t$ ,  $u_x$  is given by

$$u_x = \frac{[\alpha \operatorname{sech}^2 \alpha t] u'_0(s)}{[1 + u_0(s) \tanh \alpha t][\alpha + \{u'_0(s) + \alpha u_0(s)\} \tanh \alpha t]},$$

and deduce that a shock will form if  $u'_0 + \alpha(1 + u_0)$  becomes negative for some  $s$ . Show that if  $u_0 = a/(1 + s^2)$  and  $a$  is small, this occurs if

$$\alpha \lesssim \frac{3a\sqrt{3}}{8}.$$

7.3 Discuss the formation of shocks and the resulting shock structure for the equation

$$u_t + u^\alpha u_x = \varepsilon [u^\beta u_x]_x,$$

where  $\alpha, \beta > 0$ , and  $\varepsilon \ll 1$ . (Assume  $u > 0$ , and  $u \rightarrow 0$  at  $\pm\infty$ .)

Show that the equation

$$u_t + uu_x = \varepsilon uu_{xx}$$

admits a shock structure joining  $u_-$  to a lower value  $u_+$ , but not one in which the wave speed  $c = [\frac{1}{2}u^2]_-^+ / [u]_-^+$ . Why should this be so?

7.4 Use phase plane methods to study the existence of travelling wave solutions to the equation

$$u_t = u^p(1 - u^q) + [u^r u_x]_x,$$

when (i)  $p = 1, q = 2, r = 0$ ; (ii)  $p = 1, q = 1, r = 1$ .

- 7.5 Two examples of integrable partial differential equations which admit soliton solutions are the nonlinear Schrödinger (NLS) equation

$$iu_t = |u|^2 u + u_{xx},$$

and the sine-Gordon equation

$$u_{tt} - u_{xx} = \sin u.$$

Show that these equations admit solitary wave solutions (which are in fact solitons).

- 7.6 In a model of snow melting, it is assumed that the permeability is  $k = k_0 S^\alpha$ , and the capillary suction is  $p_c(S) = p_0(S^{-\beta} - S)$ , where  $\alpha, \beta > 0$ , and  $S$  is the saturation. How does the choice of different values of  $\alpha$  and  $\beta$  affect the formation and propagation of shock waves, and their internal structure?

# Chapter 8

## Similarity solutions

Another class of particular solutions of partial differential equations is that of similarity solutions. Like travelling wave solutions, similarity solutions are important indicators of solution behaviour. They are appropriate when the statement of the differential equation and its associated boundary and initial conditions make no reference to any intrinsic scale. One way to recognise when this is the case is to see whether the equation and its associated initial/boundary conditions are invariant under a rescaling of the variables. If this is the case, a similarity solution is appropriate, and its form can be inferred from the combinations of variables which are invariant under the transformation.

### 8.1 The heat equation

The classical situation where a similarity solution is appropriate is the heat equation in one space dimension,

$$u_t = u_{xx}, \quad (8.1)$$

which we choose to solve on  $(-\infty, \infty)$  subject to the boundary conditions

$$u \rightarrow 0 \quad \text{as} \quad x \rightarrow \pm\infty, \quad (8.2)$$

and an initial condition to be described below.

The heat or diffusion equation represents the diffusion of a quantity whose density is  $u$ . For example, it may represent heat transfer, where  $u$  is temperature, or salt diffusion, where  $u$  is salt concentration. A standard kind of problem to consider is then the release of a concentrated amount at  $x = 0$  at  $t = 0$ . We can idealise this by supposing that at  $t = 0$  (with  $u$  suitably normalised),

$$u = 0, \quad x \neq 0, \quad \int_{-\infty}^{\infty} u(x) dx = 1. \quad (8.3)$$

This apparently contradictory prescription idealises the concept of a very concentrated local injection of  $u$ . For example, (8.1) with (8.3) could represent the diffusion of sugar in hot (one-dimensional) tea from an initially emplaced sugar grain. (8.3) defines the

delta function  $\delta(x)$ , an example of a generalised function. One can think of generalised functions as being (defined by) the equivalence classes of well-behaved functions  $u_n$  with appropriate limiting behaviour. For example, the delta function is defined by the class of well-behaved functions  $u_n$  for which

$$\int_{-\infty}^{\infty} u_n(x) f(x) dx \rightarrow f(0) \quad (8.4)$$

as  $n \rightarrow \infty$  for all well-behaved  $f(x)$ . As a shorthand, then,

$$\int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0) \quad (8.5)$$

for any  $f$ , but the ulterior definition is really in (8.4). In practice, however, we think of a delta function as a ‘function’ of  $x$ , zero everywhere except for a (very) sharp spike at  $x = 0$ .

A similarity solution is appropriate for this problem because it has no intrinsic space or time scales. The time scale is semi-infinite and the space scale is infinite. The equation fixes a diffusive time scale if there is a length scale, but there is none. It is in this context that one can expect the solution to look the same at different times on different scales, and it is this self-similarity that lends the solution type its name. In general, as  $t$  varies, then the length scale might vary as  $\xi(t)$  and the amplitude of the solution  $u$  might vary as  $U(t)$ . That is, if we look at  $u/U$  as a function of  $x/\xi$ , it will look the same for all  $t$ . This in turn suggests that the solution takes the form

$$u = U(t) f \left[ \frac{x}{\xi(t)} \right], \quad (8.6)$$

and this is one of the forms of a similarity solution.

It is most often the case that  $U$  and  $\xi$  are powers of  $t$ , and the exponents are then to be chosen so that the problem has such a solution. This is best seen by example. If we denote  $\eta = x/\xi(t)$ , and substitute the form (8.6) into (8.1), (8.2) and (8.3), then we find

$$\begin{aligned} U' f - \frac{U \xi'}{\xi} \eta f' &= \frac{U}{\xi^2} f'', \\ f(\pm\infty) &= 0, \\ \xi(0) &= 0, \\ U \xi \int_{-\infty}^{\infty} f d\eta &= 1. \end{aligned} \quad (8.7)$$

Since  $f$  is supposed to be a function of  $\eta$  only, all  $t$  dependence must vanish from this set of equations. The integral constraint requires

$$U \xi = 1 \quad (8.8)$$

(the constant can be arbitrarily chosen), whence the differential equation for  $f$  becomes

$$f'' + \xi \xi' (\eta f)' = 0, \quad (8.9)$$

and we choose  $\xi\xi' = 2$ ; together with  $\xi(0) = 0$  (which follows from the requirement that for  $x \neq 0$ ,  $u = 0$  at  $t = 0$ ), this implies

$$\xi = 2\sqrt{t}, \quad U = \frac{1}{2\sqrt{t}}, \quad (8.10)$$

and the solution for  $f$  satisfying the integral constraint is

$$f = \frac{1}{\sqrt{\pi}} e^{-\eta^2}. \quad (8.11)$$

Finally, we have the fundamental solution of the heat equation,

$$u = \frac{1}{2\sqrt{\pi t}} \exp\left(-\frac{x^2}{4t}\right). \quad (8.12)$$

A feature of similarity solutions is the reduction in order, and an associated reduction in the number of boundary conditions and in the number of independent variables. Thus, apart from the requirement that the equation admit a similarity solution, we need at least two of the initial/boundary conditions to collapse into one. In diffusional problems, this is most commonly the initial condition and the condition at infinity. This is the case above, where we required  $u \rightarrow 0$  both as  $x \rightarrow \pm\infty$  and as  $t \rightarrow 0$ . The partial differential equation is third order in the sense that there is one time derivative and two space derivatives (thus requiring three conditions), but the similarity equation for  $f$  is a second order ordinary differential equation.

There are two other comments to make about the example above. One is that the first integral of the similarity equation is a consequence of the conserved quantity  $\int_{-\infty}^{\infty} u dx$ , whose constancy follows from the equation and boundary conditions. When there is no such conservation law, then generally the equation cannot be solved analytically.

The second comment concerns the usefulness of similarity solutions. They require particular combinations of initial and boundary conditions which, one might suppose, are not generally applicable. An important observation about more general problems is that they will often develop self-similar characteristics in localised regions of space or at large times.

We can illustrate this simply by solving (8.1) and (8.2) with the general initial condition

$$u = u_0(x) \quad \text{at} \quad t = 0. \quad (8.13)$$

Because the problem is linear, we can use the fundamental solution (8.12) to generate a solution.<sup>1</sup> The initial condition can equally be written as

$$u = \int_{-\infty}^{\infty} u_0(\xi) \delta(x - \xi) d\xi, \quad (8.14)$$

and this tells us that the solution satisfying (8.13) is

$$u = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} u_0(\xi) \exp\left(-\frac{(x - \xi)^2}{4t}\right) d\xi. \quad (8.15)$$

---

<sup>1</sup>We are simply using the Green's function for the heat equation here.

Now the idea is this. Suppose  $u_0$  is localised over some region near the origin; for example we might have  $u_0 = \frac{1}{1+x^2}$ ; then at large time, we expect the initial concentration to have diffused a long way, and to have ‘forgotten’ details of the initial condition. It is because of this that a similarity solution becomes relevant. This can be seen by approximating the exact solution (8.15) for  $t \gg 1$ . We rewrite the solution as

$$u = \frac{1}{2\sqrt{\pi t}} \exp\left(-\frac{x^2}{4t}\right) \int_{-\infty}^{\infty} u_0(\xi) \exp\left[-\frac{(\xi^2 - 2\xi x)}{4t}\right] d\xi; \quad (8.16)$$

For large  $t$  and localised  $u_0$ , the exponential in the integrand becomes approximately one for bounded  $\xi$  (and for large  $|\xi|$  we suppose  $u_0$  is negligible), and therefore we regain (8.12) in the form

$$u \approx \frac{1}{2\sqrt{\pi t}} \left( \int_{-\infty}^{\infty} u_0(\xi) d\xi \right) \exp\left(-\frac{x^2}{4t}\right), \quad t \gg 1. \quad (8.17)$$

### 8.1.1 Error function profile

Let us turn to a problem consisting of the same heat equation, but now to be solved on  $0 < x < \infty$ , with

$$\begin{aligned} u &= 0 & \text{at } t &= 0, \\ u &= 1 & \text{at } x &= 0, \\ u &\rightarrow 0 & \text{as } x &\rightarrow \infty. \end{aligned} \quad (8.18)$$

There is again a similarity solution, but the  $x = 0$  boundary gives a scale for  $u$ , and its size is thus fixed. Invariance of the equation again implies that the similarity variable can be taken to be  $x/2\sqrt{t}$ , and therefore we seek a similarity solution of the form

$$u = f(\eta), \quad \eta = \frac{x}{2\sqrt{t}}. \quad (8.19)$$

Substituting this into the heat equation (8.1), together with the initial/boundary conditions (8.18), we find that  $f$  satisfies

$$f'' + 2\eta f' = 0, \quad f(0) = 1, \quad f(\infty) = 0. \quad (8.20)$$

The solution of this is the complementary error function profile

$$f(\eta) = \operatorname{erfc} \eta, \quad (8.21)$$

where the complementary error function is defined by

$$\operatorname{erfc} \eta = \frac{2}{\sqrt{\pi}} \int_{\eta}^{\infty} e^{-s^2} ds; \quad (8.22)$$

it is a monotonically decreasing function as shown in figure 8.1, and is a commonly occurring function in diffusion problems (the error function is defined by  $\operatorname{erf} \eta = 1 - \operatorname{erfc} \eta$ ).



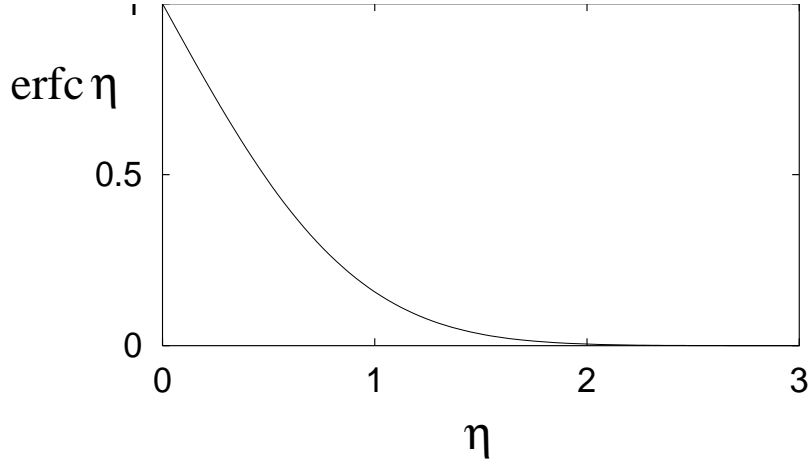


Figure 8.1: The complementary error function.

## 8.2 The porous medium equation

Gas in a porous medium satisfies the mass conservation law

$$\rho_t + \nabla \cdot \mathbf{q} = 0, \quad (8.23)$$

where  $\rho$  is density and  $\mathbf{q}$  is the mass flux. In a porous medium, Darcy's law (which we have already met in section 7.5) relates  $q$  to pressure gradient, specifically

$$\mathbf{q} = -\frac{\rho k}{\mu} \nabla p, \quad (8.24)$$

where  $k$  is the permeability,  $\mu$  is the gas viscosity, and  $p$  is pressure. The density  $\rho$  of a gas is related to its pressure by a constitutive law, which is usually taken to be the perfect gas law

$$p = \frac{\rho R T}{M}, \quad (8.25)$$

where  $M$  is the molecular weight,  $T$  is temperature, and  $R$  is the universal gas constant. In isothermal conditions where the temperature is constant, we thus have  $p \propto \rho$ , and in suitable dimensionless units, we have the porous medium equation

$$\rho_t = \nabla \cdot [\rho \nabla \rho]. \quad (8.26)$$

This is appropriate for slow gas seepage.

If the gas flow is rapid, then the temperature alters as the pressure changes, and in adiabatic conditions where there is no heat loss,  $p \propto \rho^\gamma$ , where  $\gamma > 1$  is the ratio of the specific heats; for air,  $\gamma = 1.4$ . In this case, the appropriate version of (8.26) is

$$\rho_t = \nabla \cdot [\rho^m \nabla \rho], \quad (8.27)$$

where  $m = \gamma > 1$ . Thus both cases are described by (8.27), differing only in the choice of exponent  $m$ . In both cases,  $m \geq 1$ .

Let us consider the seepage of gas in a one-dimensional column, and let us write  $u$  for  $\rho$ , so that

$$u_t = (u^m u_x)_x, \quad (8.28)$$

and let us consider the spread of an initially concentrated dose at  $x = 0$ , so that

$$\begin{aligned} u &\rightarrow 0 \quad \text{as } x \rightarrow \pm\infty, \\ u &= 0 \quad \text{at } t = 0, \quad x \neq 0, \end{aligned} \quad (8.29)$$

and

$$\int_{-\infty}^{\infty} u \, dx = 1 \quad (8.30)$$

for all  $t$ .

We seek a similarity solution of the form

$$u = U(t)f(\eta), \quad \eta = \frac{x}{\xi(t)}. \quad (8.31)$$

The integral constraint requires  $U = 1/\xi$ , and then we find

$$-\xi^{m+1}\xi'(\eta f)' = [f^m f']', \quad (8.32)$$

where  $\xi' = d\xi/dt$  but  $f' = df/d\eta$ . The initial/boundary conditions become

$$f(\pm\infty) = 0, \quad (8.33)$$

and the normalisation condition (8.30) is

$$\int_{-\infty}^{\infty} f \, d\eta = 1. \quad (8.34)$$

The solution requires  $\xi^{m+1}\xi'$  to be constant, and it is algebraically convenient to choose  $\xi^{m+1}\xi' = 2/m$ , thus

$$\eta = x \left[ \frac{m}{2(m+2)t} \right]^{\frac{1}{m+2}}. \quad (8.35)$$

Because  $u$  is conserved with these boundary conditions, a first integral of (8.32) exists, and is

$$f^m f' + \frac{2}{m} \eta f = 0, \quad (8.36)$$

with the constant of integration being zero (because  $f \rightarrow 0$  as  $\eta \rightarrow \pm\infty$ ). Thus either  $f = 0$ , or

$$f = [\eta_0^2 - \eta^2]^{1/m}, \quad (8.37)$$

so that the solution has the form of a cap of finite extent, given by (8.37) (for  $|\eta| < \eta_0$ , and  $f = 0$  for  $|\eta| > \eta_0$ ). The value of  $\eta_0$  is determined from  $\int_{-\infty}^{\infty} f d\eta = 1$ , and is

$$\eta_0 = \frac{1}{\left[2 \int_0^{\pi/2} \cos^{\frac{m+2}{m}} \theta d\theta\right]^{\frac{m}{m+2}}}. \quad (8.38)$$

The finite extent of the profile is due to the degeneracy of the equation when  $m > 0$ . (The limit  $m \rightarrow 0$  regains the Gaussian solution of the heat equation by first putting  $\eta = \sqrt{m}\eta_0\zeta$ ,  $f = F/\sqrt{m}$ , and noting that  $\eta_0 \approx (\pi m)^{-m/2}$  as  $m \rightarrow 0$  (this last following by application of Laplace's method<sup>2</sup> to (8.38)).) The graph of  $f(\eta)$  is shown in figure 8.2.

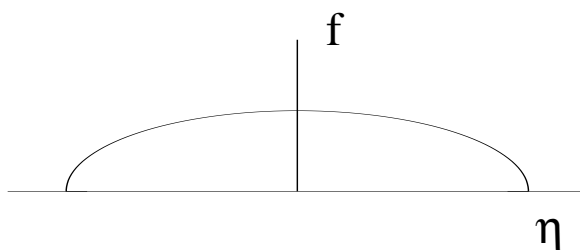


Figure 8.2:  $f(\eta)$  given by (8.37).

### 8.3 Snow melting revisited

Let us return to the snow melting model of chapter 7. Suppose that in the model described by (7.94) and (7.95), the parameter  $\beta \ll 1$ ; then the saturation profile approximately satisfies

$$\begin{aligned} \frac{\partial s}{\partial \tau} &= \frac{\partial}{\partial z} \left[ s \frac{\partial s}{\partial z} \right], \\ -s \frac{\partial s}{\partial z} &= \begin{cases} 1 & \text{on } z = 0, \\ 0 & \text{on } z = 1. \end{cases} \end{aligned} \quad (8.39)$$

At least for small times, the model admits a *similarity solution* of the form

$$s = \tau^\alpha f(\eta), \quad \eta = z/\tau^\beta, \quad (8.40)$$

where satisfaction of the equations and boundary conditions requires  $2\alpha = \beta$  and  $2\beta = 1 = \alpha$ , whence  $\alpha = 1/3$ ,  $\beta = 2/3$ , and  $f$  satisfies

$$(ff')' - \frac{1}{3}(f - 2\eta f') = 0, \quad (8.41)$$

---

<sup>2</sup>Laplace's method for the asymptotic evaluation of integrals is described in Carrier, Krook and Pearson (1966), for example.

and the condition at  $z = 0$  becomes

$$-ff' = 1 \quad \text{at} \quad \eta = 0. \quad (8.42)$$

The condition at  $z = 1$  can be satisfied for small enough  $\tau$ , as we shall see, because the equation (8.41) is degenerate, and  $f$  reaches zero in a finite distance,  $\eta_0$ , say, and  $f = 0$  for  $\eta > \eta_0$ . As  $\eta = 1/\tau^{2/3}$  at  $z = 1$ , then this solution will satisfy the no flux condition at  $z = 1$  as long as  $\tau < \eta_0^{-3/2}$ , when the advancing front will reach  $z = 1$ .

To see why  $f$  behaves in this way, integrate once to find

$$f(f' + \frac{2}{3}\eta) = -1 + \int_0^\eta f d\eta. \quad (8.43)$$

For small  $\eta$ , the right hand side is negative, and  $f$  is positive (to make physical sense), so  $f$  decreases (and in fact  $f' < -\frac{2}{3}\eta$ ). For sufficiently small  $f(0) = f_0$ ,  $f$  will reach zero at a finite distance  $\eta = \eta_0$ , and the solution must terminate. On the other hand, for sufficiently large  $f_0$ ,  $\int_0^\eta f d\eta$  reaches 1 at  $\eta = \eta_1$  while  $f$  is still positive (and  $f' = -\frac{2}{3}\eta_1$  there). For  $\eta > \eta_1$ , then  $f$  remains positive and  $f' > -\frac{2}{3}\eta$  ( $f$  cannot reach zero for  $\eta > \eta_1$  since  $\int_0^\eta f d\eta > 1$  for  $\eta > \eta_1$ ). Eventually  $f$  must have a minimum and thereafter increase with  $\eta$ . This is also unphysical, so we require  $f$  to reach zero at  $\eta = \eta_0$ . This will occur for a range of  $f_0$ , and we have to select  $f_0$  in order that

$$\int_0^{\eta_0} f d\eta = 1, \quad (8.44)$$

which in fact represents global conservation of mass. Figure 8.3 shows the schematic form of solution both for  $\beta \gg 1$  and  $\beta \ll 1$ . Evidently  $\beta \sim 1$  will have a travelling front solution between these two end cases.

## 8.4 Notes and references

An excellent introduction to the subject of partial differential equations from the viewpoint of the applied mathematician is the book by Carrier and Pearson (1976).

## Exercises

8.1 The function  $u(x, t)$  satisfies the heat equation

$$u_t = u_{xx}$$

on  $0 < x < \infty$ , together with the initial condition

$$u = 0 \quad \text{at} \quad t = 0,$$

and the boundary condition

$$u \rightarrow 0 \quad \text{as} \quad x \rightarrow \infty.$$

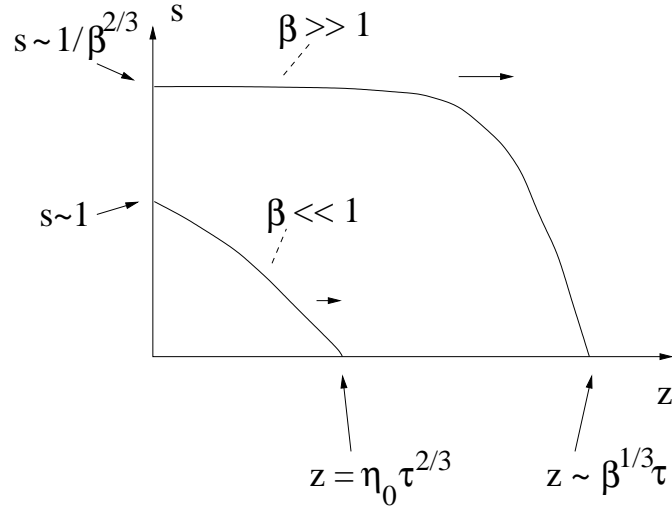


Figure 8.3: Schematic representation of the evolution of  $s$  in (7.95) for both large and small  $\beta$ .

Find similarity solutions for  $u$  when the boundary condition at  $x = 0$ ,  $t > 0$  has each of the following forms:

- (i)  $u = 1$ ;
- (ii)  $u_x = -1$ ;
- (iii)  $uu_x = -1$ ;
- (iiii)  $u_x = -u^\beta$ .

Show in case (iiii) that no similarity solution exists if  $\beta = 1$ . What is it about this case that prevents a similarity solution existing?

8.2 Which of the following equations have similarity solutions? In all cases suppose that the domain of solution is  $0 < x < \infty$ , and that  $u = 0$  at  $t = 0$  and  $u = 1$  at  $x = 0$ . If there is a similarity solution, solve the similarity equation if you can.

- (i)  $u_t = xu_{xx}$ ;
- (ii)  $u_t = (xu_x)_x$ ;
- (iii)  $u_t = u_x^2$ .

8.3 Write down the equation satisfied by a similarity solution of the form  $u = t^\beta f(\eta)$ ,  $\eta = x/t^\alpha$ , for the equation

$$u_t = (u^m u_x)_x \quad \text{in } 0 < x < \infty,$$

where  $m > 0$ , with  $u^m u_x = -1$  at  $x = 0$ ,  $u \rightarrow 0$  as  $x \rightarrow \infty$ ,  $u = 0$  at  $t = 0$ . Show that  $\int_0^\infty f d\eta = 1$ , and show that in fact  $f$  reaches zero at a finite value  $\eta_0$ . Is the requirement that  $m > 0$  necessary?

8.4  $u$  satisfies the equation

$$u_t = [D(u)u_x]_x \quad \text{in } 0 < x < \infty,$$

with  $u = 0$  at  $x \rightarrow \infty$  and  $t = 0$ . For a general function  $D$  (not a power of  $u$ ), for what kind of boundary condition at  $x = 0$  does a similarity solution exist? What if, instead,  $D = D(u_x)$ ? Write down suitable equations and boundary conditions for the similarity function in each case.

# Chapter 9

## Nonlinear diffusion

### 9.1 The viscous droplet

An example of where the nonlinear diffusion equation can arise is in the dynamics of a drop of viscous fluid on a level surface. If the fluid occupies  $0 < z < h(x, y, t)$  and is shallow, then lubrication theory gives the approximation

$$\begin{aligned}\nabla p &= \mu \frac{\partial^2 \mathbf{u}}{\partial z^2}, \\ p_z &= -\rho g,\end{aligned}\tag{9.1}$$

in which  $\mathbf{u} = (u, v, 0)$  is the horizontal component of velocity, and  $\nabla$  is the horizontal gradient  $(\partial/\partial x, \partial/\partial y, 0)$ . With  $p = 0$  at  $z = h$ , we have the hydrostatic pressure  $p = \rho g(h - z)$ , so that  $\nabla p = \rho g \nabla h$ , and three vertical integrations of (9.1)<sub>1</sub> (with zero shear stress  $\partial \mathbf{u}/\partial z = 0$  at  $z = h$  and no slip  $\mathbf{u} = 0$  at  $z = 0$ ) yield the horizontal fluid flux

$$\mathbf{q} = \int_0^h \mathbf{u} dz = -\frac{\rho g}{3\mu} h^2 \nabla h.\tag{9.2}$$

Conservation of fluid volume for an incompressible fluid is  $h_t + \nabla \cdot \mathbf{q} = 0$ , and thus

$$h_t = \frac{\rho g}{3\mu} \nabla \cdot [h^3 \nabla h],\tag{9.3}$$

corresponding to (8.27) (in two space dimensions) with  $m = 3$ . A drop of fluid placed on a table will spread out at a finite rate.

That this does not continue indefinitely is due to surface tension. Rather than having  $p = 0$  at  $z = h$  (where the atmospheric pressure above is taken as zero), the effect of surface tension is to prescribe

$$p = 2\gamma\kappa,\tag{9.4}$$

where  $\gamma$  is the surface tension, and  $\kappa$  is the mean curvature relative to the fluid droplet (i.e.  $\kappa > 0$  if the interface is concave, as illustrated in figure 9.1). The curvature is

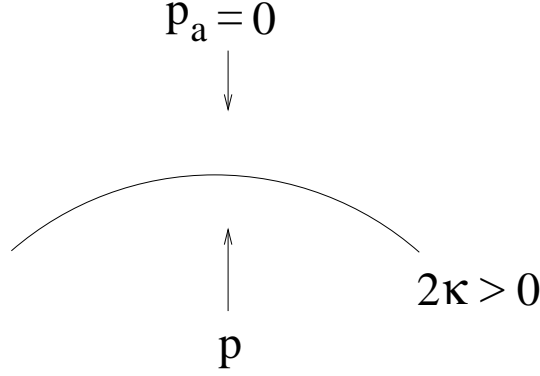


Figure 9.1: The surface shown has positive curvature when the radius of curvature is measured from below the surface; in this case equilibrium requires  $p > p_a$ .

defined as  $2\kappa = \nabla \cdot \mathbf{n}$ , where  $\mathbf{n}$  is the unit normal pointing away from the fluid (i.e., upwards). At least this shorthand definition works if we define

$$\mathbf{n} = \frac{(-h_x, -h_y, 1)}{[1 + |\nabla h|^2]^{1/2}}; \quad (9.5)$$

thus

$$2\kappa = -\nabla \cdot \left[ \frac{\nabla h}{\{1 + |\nabla h|^2\}^{1/2}} \right]. \quad (9.6)$$

It is less obvious that it will work more generally, since there are many ways of defining the interface as  $\phi(x, y, z) = 0$  and thus  $\mathbf{n} = \nabla \phi / |\nabla \phi|$  (that is (9.5) uses  $\phi = z - h$ ); but in fact it does not matter, since we may generally take  $\phi = (z - h)P$ , so that  $\nabla \phi = (-h_x, -h_y, 1)$  on  $z = h$ , and  $\nabla \phi / |\nabla \phi|$  is the same expression as in (9.6).

For shallow flows, we replace  $p = 0$  on  $z = h$  by  $p = -\gamma \nabla^2 h$  there, and thus

$$p \approx \rho g(h - z) - \gamma \nabla^2 h, \quad (9.7)$$

and (via (9.2)), (9.3) is modified to

$$h_t = \nabla \cdot \left[ \frac{h^3}{3\mu} \nabla \{ \rho g h - \gamma \nabla^2 h \} \right]. \quad (9.8)$$

The fourth order term is also ‘diffusive’, insofar as it is a smoothing term: this is most easily seen by considering the fate of modes  $e^{\sigma t + i k x}$  for the linear equation  $h_t = -h_{xxxx}$ :  $\sigma = -k^4$ : high wave number (high gradient) modes are rapidly damped. Surface tension can thus also act to smooth out shocks. The effect of surface tension relative to the diffusional gravity term is given by the Bond number

$$Bo = \frac{\rho g l^2}{\gamma}, \quad (9.9)$$



where  $l$  is the lateral length scale of the drop. This is the (only) dimensionless parameter which occurs when (9.8) is written dimensionlessly.

For a drop released on a table, (9.8) still predicts unending dispersal, but if the full nonlinear curvature term (9.6) is kept, then a steady state will exist, and surface tension keeps the drop of finite extent.

## 9.2 Advance and retreat: waiting times

The similarity solution (8.37) predicts an infinite slope at the margin (where  $f = 0$ ) if  $m > 1$  (and a zero slope if  $m < 1$ ). If one releases a finite quantity at  $t = 0$ , then one expects the long time solution to be this similarity solution. The question then arises as to how this similarity solution is approached, in particular if the initial drop has finite slope at the margin.

This question can be addressed in a more general way by studying the behaviour near the margin  $x = x_S(t)$  of a solution  $h(x, t)$  of the same equation as in (8.28),

$$h_t = (h^m h_x)_x. \quad (9.10)$$

Suppose that  $h \sim c(x_S - x)^\nu$  for  $x$  near  $x_S$ . Then satisfaction of (9.10) requires

$$\dot{x}_S \approx c^m [\nu(m+1) - 1] (x_S - x)^{\nu m - 1}. \quad (9.11)$$

Note that the similarity solution (8.37) has  $\dot{x}_S$  finite when  $\nu = 1/m$ , consistent with (9.11), and more generally we see that the margin will advance at a rate  $\dot{x}_S \approx c^m/m$  if  $h \sim c(x_S - x)^{1/m}$ .

Suppose now that  $m > 1$ , and we emplace a drop with finite slope,  $\nu = 1$ . Then the right hand side of (9.11) is zero at  $x = x_S$ , and thus  $\dot{x}_S = 0$ : the front does not move. What happens in this case is that the drop flattens out: there is transport of  $h$  towards the margin, which steepens the slope at  $x_S$  until it becomes infinite, at which point it will move. This pause while the solution fattens itself prior to margin movement is called a *waiting time*.

Conversely, if  $m < 1$ , then the front moves (forward) if the slope is zero,  $\nu = 1/m$ . If the slope is finite,  $\nu = 1$ , then (9.10) would imply infinite speed. An initial drop of finite margin slope will instantly develop zero front slope as the margin advances.

(9.11) does not allow the possibility of retreat, because it describes a purely diffusive process. The possibility of both advance and retreat is afforded by a model of a viscous drop with accretion, one example of which is the mathematical model of an ice sheet. Essentially, an ice sheet, such as that covering Antarctica or Greenland, can be thought of as a (large) viscous drop which is nourished by an accumulation rate (of ice formed from snow). A general model for such a nourished drop is

$$h_t = (h^m h_x)_x + a, \quad (9.12)$$

where  $a$  represents the accumulation rate. Unlike the pure diffusion process, (9.12) has a steady state

$$h = \left[ \frac{(m+1)a}{2} \right]^{\frac{1}{m+1}} (x_0^2 - x^2)^{\frac{1}{m+1}}, \quad (9.13)$$

where  $x_0$  must be prescribed. (In the case of an ice sheet, we might take  $x_0$  to be at the continental margin.) (9.13) is slightly artificial, as it requires  $a = 0$  for  $x > x_0$ , and allows a finite flux  $-h^m h_x = ax_0$  where  $h = 0$ . More generally, we might allow accumulation and ablation (snowfall and melting), and thus  $a = a(x)$ , with  $a < 0$  for large  $|x|$ . In that case the steady state is

$$h = \left[ (m+1) \int_x^{x_0} s \, dx \right]^{\frac{1}{m+1}}, \quad (9.14)$$

where the balance is

$$s = \int_0^x a \, dx, \quad (9.15)$$

and  $x_0$  is defined to be where accumulation balances ablation,

$$\int_0^{x_0} a \, dx = 0. \quad (9.16)$$

This steady state is actually stable, and both advance and retreat can occur. Suppose the margin is at  $x_S$ , where  $a = a_S = -|a_S|$  ( $a_S < 0$ , representing ablation). If we put  $h \approx c(x_S - x)^\nu$ , then (9.12) implies

$$\nu c \dot{x}_S (x_S - x)^{\nu-1} \approx \nu c^{m+1} [\nu(m+1) - 1] (x_S - x)^{[\nu(m+1)-2]} - |a_S|, \quad (9.17)$$

and there are three possible balances of leading order terms.

The first is as before,

$$\dot{x}_S \approx c^m [\nu(m+1) - 1] (x_S - x)^{\nu m - 1}, \quad (9.18)$$

and applies generally if  $\nu < 1$ . Supposing  $m > 1$ , then we have advance,  $\dot{x}_S \approx c^m/m$  if  $\nu = 1/m$ , but if  $\nu > 1/m$ , this cannot occur, and the margin is stationary if  $1/m < \nu < 1$ . If  $\nu = 1$ , then  $\nu(m+1) - 2 = m - 1 > 0$ , so that

$$\dot{x}_S \approx -|a_S|/c, \quad (9.19)$$

and the margin retreats; if  $\nu > 1$ , then instantaneous adjustment to finite slope and retreat occurs.

The ice sheet exhibits the same sort of waiting time behaviour as the viscous drop without accretion. For  $1/m < \nu < 1$ , the margin is stationary, and if  $x_S < x_0$  then the margin slope will steepen until  $\nu = 1/m$ , and advance occurs. On the other hand, if  $x_S > x_0$ , then the slope will decrease until  $\nu = 1$ , and retreat occurs. In the steady state, a balance is achieved (from (9.14)) when  $\nu = 2/(m+1)$ .

### 9.3 Blow-up

Further intriguing possibilities arise when the source term is nonlinear. An example is afforded by the nonlinear (diffusion) equation

$$u_t = u_{xx} + \lambda e^u, \quad (9.20)$$

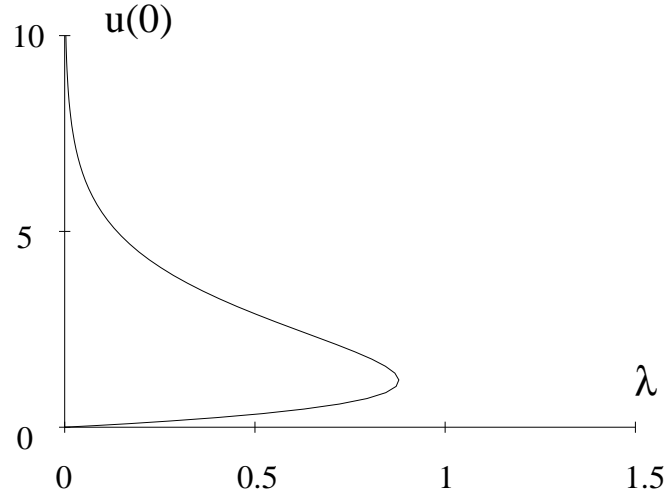


Figure 9.2: Maximum values of  $u$ ,  $u(0)$ , as a function of the parameter  $\lambda$ . Blow-up occurs if  $\lambda \gtrsim 0.878$ .

which arises in the theory of combustion. Indeed, as we saw earlier, combustion occurs through the fact that multiple steady states can exist for a model such as (7.74), and the same is true for (9.20), which can have two steady solutions. In fact, if we solve  $u'' + \lambda e^u = 0$  with boundary conditions  $u = 0$  on  $x = \pm 1$ , then the solutions are

$$u = 2 \ln \left[ A \operatorname{sech} \left\{ \sqrt{\frac{\lambda}{2}} A x \right\} \right], \quad (9.21)$$

where  $A = \exp[u(0)/2]$ , and  $A$  satisfies

$$A = \cosh \left[ \sqrt{\frac{\lambda}{2}} A \right], \quad (9.22)$$

which has two solutions if  $\lambda < 0.878$ , and none if  $\lambda > 0.878$ : the situation is depicted in figure 9.2. If we replace  $e^u$  by  $\exp[u/(1 + \varepsilon u)]$ ,  $\varepsilon > 0$ , we regain the top (hot) branch also.

One wonders what the absence of a steady state for (9.20) if  $\lambda > \lambda_c$  implies. The time-dependent problem certainly has a solution, and an idea of its behaviour can be deduced from the spatially independent problem,  $u_t = \lambda e^u$ , with solution  $u = \ln[1/\{\lambda(t_0 - t)\}]$ :  $u$  reaches infinity in a finite time. This phenomenon is known as *thermal runaway*, and more generally the creation of a singularity of the solution in finite time is called *blow-up*. Numerical solutions of the equation (9.20) including the diffusion term show that blow-up still occurs, but at an isolated point; figure 9.3 shows the approach to blow-up as  $t$  approaches a critical blow-up time  $t_c$ .

In fact, one can prove generally that no steady solutions exist for  $\lambda$  greater than some critical value, and also that in that case, blow-up will occur in finite time. To do this, we use some pretty ideas of higher grade mathematics.

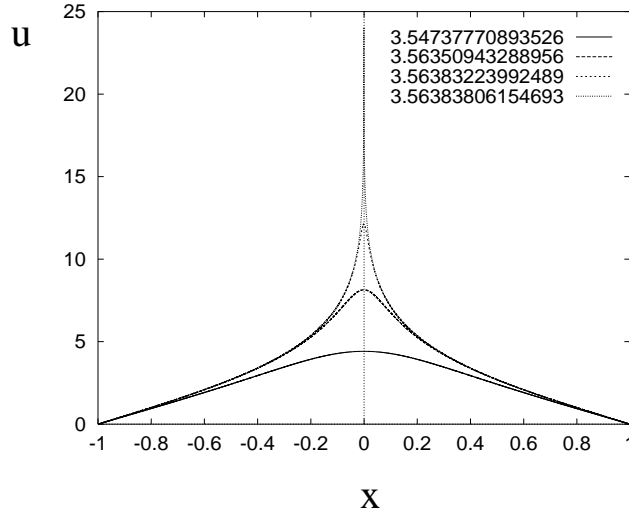


Figure 9.3: Solution of  $u_t = u_{xx} + e^u$  on  $[-1, 1]$ , with  $u = 0$  at  $x = -1, 1$  and  $t = 0$ . The solution is shown for four times close to the blow-up time  $t_c = 3.56384027594971$ . The many decimal places should not be treated too seriously, but they do indicate the logarithmic suddenness of the runaway.

Suppose we want to solve the more general problem

$$u_t = \nabla^2 u + \lambda e^u \quad \text{in } \Omega, \quad (9.23)$$

with  $u = 0$  in the boundary  $\partial\Omega$ , and  $u = 0$  at  $t = 0$  (these conditions are for convenience rather than necessity). We will be able to prove results for (9.23) which are comparable to those for the ordinary differential equation version (cf. (7.81))

$$\dot{w} = -\mu_1 w + \lambda e^w, \quad (9.24)$$

because, in some loose sense, the Laplacian operator  $\nabla^2$  resembles a loss term.

More specifically, we recall some pertinent facts about the (Helmholtz) eigenvalue problem

$$\nabla^2 \phi + \mu \phi = 0 \quad \text{in } \Omega, \quad (9.25)$$

with  $\phi = 0$  on  $\partial\Omega$ . There exists a denumerable sequence of real eigenvalues  $0 < \mu_1 \leq \mu_2 \leq \dots$ , with  $\mu_n \rightarrow \infty$  as  $n \rightarrow \infty$ , and corresponding eigenfunctions  $\phi_1, \phi_2, \dots$  which form an orthonormal set (using the  $L^2$  norm), thus

$$(\phi_i, \phi_j) \equiv \int_{\Omega} \phi_i \phi_j dV = \delta_{ij}, \quad (9.26)$$

where  $\delta_{ij}$  is the Kronecker delta ( $= 1$  if  $i = j$ ,  $0$  if  $i \neq j$ ). These eigenvalues satisfy a variational principle of the form

$$\mu_i = \min \int_{\Omega} |\nabla \phi|^2 dV, \quad (9.27)$$

where  $\phi$  ranges over functions of unit norm,  $\|\phi\|_2 = \{\int \phi^2 dV\}^{1/2} = 1$ , which are orthogonal to  $\phi_j$  for  $j < i$ ; (more generally  $\mu_i = \min \{\int |\nabla \phi|^2 / \int \phi^2\}$  if  $\phi$  is not normalised on to the unit sphere  $\|\phi\|_2 = 1$ ). In particular

$$\mu_1 = \min_{\|\phi\|_2=1} \int_{\Omega} |\nabla \phi|^2 dV, \quad (9.28)$$

and the corresponding  $\phi_1$  is of one sign, let us say positive.

We take the inner product of the equation (9.23) with  $\phi_1$  and divide by  $\int \phi_1 dV$ ; defining

$$v(t) = \frac{\int_{\Omega} u \phi_1 dV}{\int_{\Omega} \phi_1 dV} = \int_{\Omega} u d\omega, \quad (9.29)$$

where  $d\omega = \phi_1 dV / \int_{\Omega} \phi_1 dV$  is a measure on  $\Omega$  (with  $\int_{\Omega} d\omega = 1$ ), and using Green's theorem, we find

$$\dot{v} = \lambda \int_{\Omega} e^u d\omega - \mu_1 v, \quad (9.30)$$

and the equation for  $v$  is close to the ordinary differential equation (9.24).

Now we use Jensen's inequality. This says that if we have an integrable function  $g$  on  $\Omega$  and a convex function  $f(x)$  (i.e. one that bends upwards,  $f'' > 0$ ), then

$$f \left[ \int_{\Omega} g d\omega \right] \leq \int_{\Omega} f[g] d\omega \quad (9.31)$$

for any measure  $\omega$  on  $\Omega$  such that  $\int_{\Omega} d\omega = 1$ . We have chosen  $\omega$  to be so normalised, and  $e^u$  is convex: thus

$$\int_{\Omega} \exp(u) d\omega \geq \exp \left[ \int_{\Omega} u d\omega \right] = e^v, \quad (9.32)$$

so that

$$\dot{v} \geq \lambda e^v - \mu_1 v. \quad (9.33)$$

It is now easy to prove non-existence and blow-up for  $\lambda$  greater than some critical value  $\lambda_c$ . Firstly,  $u$  must be positive, and hence also  $v$ . (For suppose  $u < 0$ : since  $u = 0$  at  $t = 0$  and on  $\partial\Omega$ , then  $u$  attains its minimum in  $\Omega$  at some  $t > 0$ , at which point  $u_t \leq 0$ ,  $u_{xx} \geq 0$ , which is impossible, since then  $u_t - u_{xx} = \lambda e^u \leq 0$ .) For any  $v$ ,  $e^v \geq ev$ , thus  $\dot{v} \geq (\lambda e - \mu_1)v$ . In a steady state,  $\dot{v} = 0$ , and  $v > 0$  (clearly  $u = 0$  is not a solution), so this is impossible if

$$\lambda > \mu_1/e. \quad (9.34)$$

This implies non-existence of a solution for  $\lambda > \lambda_c$ , where  $\lambda_c \leq \mu_1/e$ .

In a similar vein, if  $\lambda > \mu_1/e$ , then

$$\dot{v} > \mu_1[e^{v-1} - v], \quad (9.35)$$

and  $v > w$ , where

$$\dot{w} = \mu_1(e^{w-1} - w), \quad w(0) = 0. \quad (9.36)$$

(This is a standard comparison argument:  $v = w$  at  $t = 0$ , and  $\dot{v} > \dot{w}$  there, so  $v - w$  is initially positive. It remains so unless at some future time  $v - w$  reaches zero again, when necessarily  $\dot{v} - \dot{w} \leq 0$  — which is impossible, since  $\dot{v} > \dot{w}$  whenever  $v = w$ .) But  $w \rightarrow \infty$  in finite time ( $\dot{w} > 0$  so that  $w \rightarrow \infty$  as  $t$  increases, and as  $w \rightarrow \infty$ ,  $e^{-w}\dot{w} \approx \mu e^{-1}$ , so  $e^{-w}$  reaches zero in finite time); therefore also  $v$  reaches infinity in finite time. Finally

$$v = \int_{\Omega} u \, d\omega \leq \sup_{\Omega} u, \quad (9.37)$$

since  $\int_{\Omega} d\omega = 1$ : hence  $u \rightarrow \infty$  in finite time.

In fact  $u \rightarrow \infty$  at isolated points, and usually at one isolated point. As blow-up is approached, one might suppose that the nature of the solution in the vicinity of the blow-up point would become independent of the initial (or boundary) conditions, and thus that some form of similarity solution might be appropriate.

This is indeed the case, although the precise structure is rather complicated. We examine blow-up in one spatial dimension,  $x$ . As a first guess, the logarithmic nature of blow-up in the spatially independent case, together with the usual square-root behaviour of the space variable in similarity solutions for the diffusion equation, suggests that we define

$$\tau = -\ln(t_0 - t), \quad \eta = \frac{x - x_0}{(t_0 - t)^{1/2}}, \quad u = -\ln[\lambda(t_0 - t)] + g(\eta, \tau), \quad (9.38)$$

where blow-up occurs at  $x = x_0$  at  $t = t_0$ ; hence  $g$  satisfies

$$g_{\tau} = g_{\eta\eta} - \frac{1}{2}\eta g_{\eta} + e^g - 1. \quad (9.39)$$

The natural candidate for a similarity solution is then a steady solution  $g(\eta)$  of (9.39), satisfying

$$g'' - \frac{1}{2}\eta g' + (e^g - 1) = 0, \quad (9.40)$$

and matching to a far field solution  $u(x, t_0)$  would suggest

$$g \sim -2 \ln |\eta| \quad \text{as } \eta \rightarrow \pm\infty, \quad (9.41)$$

and solutions of (9.40) with this asymptotic structure do exist — but not at each end. (9.40) admits even solutions, and if we restrict ourselves to these, then we may take

$$g'(0) = 0, \quad g(0) \neq 0. \quad (9.42)$$

(If  $g(0) = 0$ , then  $g \equiv 0$  is the solution.) However, it is found that such solutions have a different asymptotic behaviour as  $\eta \rightarrow \infty$ , namely

$$g \sim -\frac{A}{|\eta|} \exp\left[\frac{1}{4}\eta^2\right], \quad (9.43)$$

and  $A = A[g(0)] > 0$  for  $g(0) \neq 0$  (and  $A(0) = 0$ ), and these cannot match to the outer solution. If one alternately prescribes (9.41) as  $\eta \rightarrow +\infty$ , for example, then the solution is asymmetric, and has the exponential behaviour (9.43) as  $\eta \rightarrow -\infty$ . thus

the appealingly simple similarity structure implied by steady solutions of (9.39) is wrong (and actually, the solution of the initial value problem (9.39) satisfying (9.41) tends to zero as  $\tau \rightarrow \infty$ ).

However, (9.39) itself develops a local similarity structure as  $\tau \rightarrow \infty$ , using a further similarity variable

$$z = \frac{\eta}{\tau^{1/2}} = \frac{x - x_0}{(t_0 - t)^{1/2}[-\ln(t_0 - t)]^{1/2}}. \quad (9.44)$$

Rewriting (9.39) in terms of  $z$  and  $\tau$  yields

$$g_\tau + \frac{1}{2}z g_z + 1 - e^g = \frac{1}{\tau}[g_{zz} + \frac{1}{2}z g_z]. \quad (9.45)$$

At leading order in  $\tau^{-1}$  this has a solution

$$g = -\ln[1 + \frac{1}{4}cz^2], \quad (9.46)$$

where  $c$  is indeterminate, and this forms the basis for a formal expansion. It is algebraically convenient to use (9.46) to define  $c$  as a new variable, and also to write

$$s = \ln \tau; \quad (9.47)$$

Then (9.45) becomes

$$c_z = \frac{2}{\tau z^3} \left[ 2c + 4z c_z + z^2 c_{zz} + z^2 \left\{ -\frac{[c + \frac{1}{2}z c_z]^2}{1 + \frac{1}{4}c z^2} + c + \frac{1}{2}z c_z - c_s \right\} \right]. \quad (9.48)$$

We seek a solution for (9.48) in the form

$$c \sim c_0(z, s) + \frac{1}{\tau} c_1(z, s) + \frac{1}{\tau^2} c_2(z, s) \dots, \quad (9.49)$$

and then, since  $\tau d/d\tau = d/ds$ , we have

$$c_s \sim \dot{c}_0 + \frac{1}{\tau}(\dot{c}_1 - c_1) + \frac{1}{\tau^2}(\dot{c}_2 - 2c_2) + \dots, \quad (9.50)$$

where  $\dot{c}_2 \equiv \partial c_2 / \partial s$ . Substituting this into (9.48) and equating powers of  $\tau$ , we find

$$c_0 = C_0(s), \quad (9.51)$$

where  $C_0$  is arbitrary, and

$$c_{1z} = \frac{2}{z^3} \left[ 2C_0 + z^2 \left\{ -\frac{C_0^2}{1 + \frac{1}{4}C_0 z^2} + C_0 - \dot{C}_0 \right\} \right]. \quad (9.52)$$

The arbitrary function  $C_0$  arises because the order of the approximate equation is reduced. In order to specify it, and other arbitrary functions of  $s$  which arise at each order, we require that the solutions  $c_i$  be smooth, and this requires that there be no term on the right hand side of (9.52) proportional to  $1/z$  as  $z \rightarrow 0$ , in order that

logarithmic singularities not be introduced. Specifically, we require at each stage of the approximation that

$$\frac{\partial c_i}{\partial z} = \frac{2}{z^3} [a_{0i} + a_{1i}z + a_{3i}z^3 + \dots]; \quad (9.53)$$

so that  $z^2 c_i$  is smooth. Applying this to (9.52) requires that

$$\dot{C}_0 = C_0(1 - C_0), \quad (9.54)$$

so that  $C_0 \rightarrow 1$  as  $s \rightarrow \infty$ , and then

$$c_1 = -\frac{2C_0}{z^2} + C_1(s) + C_0^2 \ln[1 + \frac{1}{4}C_0 z^2]. \quad (9.55)$$

At  $O(1/\tau^2)$ , we then have

$$c_{2z} = \frac{2}{z^3} \left[ 2c_1 + 4zc_{1z} + z^2 c_{1zz} + z^2 \left\{ -(\dot{c}_1 - c_1) + (c_1 + \frac{1}{2}zc_{1z}) - \frac{2c_0(c_1 + \frac{1}{2}zc_{1z})}{1 + \frac{1}{4}c_0 z^2} + \frac{1}{4}c_0^2 c_1 z^2 (1 + \frac{1}{4}c_0 z^2)^2 \right\} \right], \quad (9.56)$$

and applying the regularity condition (9.53), we find, after some algebra,

$$\dot{C}_1 = 2(1 - C_0)C_1 + \frac{5}{2}C_0^3, \quad (9.57)$$

so that  $C_1 \rightarrow C_{10} + \frac{5}{2}s$  as  $s \rightarrow \infty$ . Thus finally we obtain the local similarity solution

$$u \approx -\ln \left[ \lambda \left\{ t_0 - t + \frac{c(x - x_0)^2}{4[-\ln(t_0 - t)]} \right\} \right], \quad (9.58)$$

where  $c \approx C_0(s)$ ,  $s = \ln \tau = \ln[-\ln(t_0 - t)]$ .

## 9.4 Notes and references

There is a large literature on the subject of nonlinear diffusion, blow up, and related issues. A thorough examination of some of these problems is contained in the book by Samarskii *et al.* (1995).

## Exercises

9.1 A small droplet satisfies the surface-tension controlled equation

$$h_t = -\frac{\gamma}{3\mu} \nabla \cdot [h^3 \nabla \nabla^2 h].$$

A small quantity  $\int h dV = Q$  is released at time zero at the origin. Find a suitable similarity solution in one and two spatial dimensions.



- 9.2 A gravity-driven drop of fluid spreads out on a flat surface. Its viscosity  $\mu$  is a function of shear rate, so that in (9.1),

$$\rho g \nabla h = \frac{\partial \tau}{\partial z},$$

$$\frac{\partial \mathbf{u}}{\partial z} = A |\tau|^{n-1} \tau.$$

(A constant viscosity fluid has  $n = 1$ .) Show that the horizontal fluid flux is

$$\mathbf{q} = -\frac{A(\rho g)^n}{n+2} |\nabla h|^{n-1} h^{n+2} \nabla h,$$

and deduce that

$$\frac{\partial h}{\partial t} = \frac{A(\rho g)^n}{n+2} \nabla \cdot [h^{n+2} |\nabla h|^{n-1} \nabla h].$$

Find similarity solutions in one and two dimensions for the initial emplacement of a finite volume at the origin. What happens as  $n \rightarrow \infty$  or  $n \rightarrow 0$ ?

- 9.3 Suppose a two-dimensional drop as in question 9.2 is subjected to a spatially varying accumulation  $a(x)$ , with  $xa' < 0$  for  $x \neq 0$ . Find appropriate local behaviour near the right hand margin  $x = x_s > 0$ , where  $h = 0$ , if  $\dot{x}_s > 0$ ,  $\dot{x}_s < 0$ ,  $\dot{x}_s = 0$ .

- 9.4 Let  $u$  satisfy

$$u_t = \lambda u^p + u_{xx},$$

with  $u = 1$  on  $x = \pm 1$  and  $t = 0$ . Prove that if  $\lambda$  is large enough,  $u$  must blow up in finite time if  $p > 1$ . Supposing this happens at time  $t_0$  at  $x = 0$ , show that a possible local similarity structure is of the form

$$u = \frac{1}{(t_0 - t)^\beta} f(\xi), \quad \xi = \frac{x}{(t_0 - t)^{1/2}},$$

and prove that  $\beta = 1/(p-1)$ . Show that in this case,  $f$  would satisfy

$$f'' - \frac{1}{2} \xi f' + \lambda f^p - \beta f = 0,$$

and explain why appropriate boundary conditions would be

$$f \sim |\xi|^{-2\beta} \quad \text{as } \xi \rightarrow \pm\infty,$$

and show that such solutions might be possible. Are any other limiting behaviours possible?

- 9.5 By direct integration, show that the solution of

$$u'' + \lambda e^u = 0$$

satisfying  $u = 0$  on  $x = \pm 1$  is

$$u = 2 \ln \left[ A \operatorname{sech} \left\{ \sqrt{\frac{\lambda}{2}} Ax \right\} \right],$$

and find a transcendental equation for  $A$ . Hence show that no solution exists for  $\lambda > \lambda_c$ , and derive and solve (numerically) an algebraic equation for  $\lambda_c$ .

If the equation is to be solved in  $[0, 1]$ , with  $u' = 0$  on  $x = 0$  and  $u' = -1$  on  $x = 1$ , find the solution, and plot  $u(0)$  as a function of  $\lambda$ . Is there a critical value  $\lambda_c$ ? If so, find it; if not, why not?

9.6 (i) Find an exact solution of the Gel'fand equation

$$\nabla^2 \theta + \lambda e^\theta = 0 \quad \text{in} \quad 0 < r < 1,$$

where  $r$  is the cylindrical polar radius, and  $\theta = 0$  on  $r = 1$ . [*Assume cylindrical symmetry, and a suitable condition of regularity at  $r = 0$ .*] Show that there is a critical parameter  $\lambda_c$  such that no solution exists for  $\lambda > \lambda_c$ , and find its value.

(ii) Write down the ordinary differential equation satisfied by a spherically symmetric solution of the Gel'fand equation in part (i). Suppose that  $\theta = 0$  on  $r = 1$  and  $\theta_r = 0$  on  $r = 0$  (why?). By putting

$$p = \lambda r^2 e^\theta, \quad q = 2 + r\theta_r, \quad r = e^{-t},$$

show that  $p(t)$  and  $q(t)$  satisfy the ordinary differential equations

$$\begin{aligned} \dot{p} &= -pq, \\ \dot{q} &= p + q - 2. \end{aligned}$$

By consideration of trajectories for  $p$  and  $q$  in the phase plane, show that multiple solutions exist for  $\lambda \approx 2$ , and infinitely many at  $\lambda = 2$ . Sketch the corresponding response diagram of  $\theta(0)$  versus  $\lambda$ .

## References

- Arn'old, V. I. 1973 Ordinary differential equations. M. I. T. Press, Cambridge, MA.
- Bender, C. M. and S. A. Orszag 1978 Advanced mathematical methods for scientists and engineers. McGraw-Hill, New York.
- Burkill, J. C. 1956 The theory of ordinary differential equations. Oliver and Boyd, Edinburgh.
- Carrier, G. F., M. Krook and C. E. Pearson 1966 Functions of a complex variable. McGraw-Hill, New York.
- Carrier, G. F. and C. E. Pearson 1976 Partial differential equations. Academic Press, New York.
- Coddington, E. A. and N. Levinson 1972 Theory of ordinary differential equations. Tata McGraw-Hill, New Delhi.
- Courant, R. and D. Hilbert 1937 Methods of mathematical physics, vol. I. Interscience, New York.
- Fowkes, N. D. and J. J. Mahony 1994 An introduction to mathematical modelling. John Wiley, Chichester.
- Fowler, A. C. 1997 Mathematical models in the applied sciences. C. U. P., Cambridge.
- Haberman, R. 1998 Mathematical models. Society for Industrial and Applied Mathematics, Philadelphia.
- Hassell, D. C., D. J. Allwright and A. C. Fowler 1999 A mathematical analysis of Jones's site model for spruce budworm infestations. *J/ Math. Biol.* **38**, 377–421.
- Hinch, E. J. 1991 Perturbation methods. C. U. P., Cambridge.
- Hoppensteadt, F. 1975 Mathematical theories of populations: demographics, genetics and epidemics. SIAM, Philadelphia.
- Howison, S. D. 2005 Practical applied mathematics: modelling, analysis, approximation. C. U. P., Cambridge.
- Jeffreys, H. and B. S. Jeffreys 1946 Methods of mathematical physics. C. U. P., Cambridge.
- Jones, D. D. 1979 The budworm site model. In: Pest management: Proceedings of an International Conference, eds. G. A. Norton and C. S. Holling, pp. 91–156. Pergamon Press, Tarrytown, NY.

- Jordan, D.W. and P. Smith 1999 Nonlinear ordinary differential equations, 3rd ed. O. U. P., Oxford.
- Keener, J.P. 2000 Principles of applied mathematics: transformation and approximation, revised ed. Perseus Books, Cambridge, MA.
- Kermack, W.O. and A.G. McKendrick 1927 Contributions to the mathematical theory of epidemics. *Proc. R. Soc. A*. **115**, 700–721.
- Kevorkian, J. and J.D. Cole 1981 Perturbation methods in applied mathematics. Springer-Verlag, Berlin.
- Lighthill, M.J. 1958 An introduction to Fourier analysis and generalised functions. C. U. P., Cambridge.
- Lin, C.C. and L. A. Segel 1974 Mathematics applied to deterministic problems in the natural sciences. MacMillan, New York.
- Ludwig, D, D.D. Jones and C. S. Holling 1979 Qualitative analysis of insect outbreak systems: the spruce budworm and forest. *J. Anim. Ecol.* **47**, 315–332.
- Mackie, A.G. 1965 Boundary value problems. Oliver and Boyd, Edinburgh.
- Murray, J.D. 2002 Mathematical biology. I: an introduction. 3rd ed. Springer-Verlag, Berlin.
- Ockendon, J.R., S.D. Howison, A. A. Lacey and A. B. Movchan 2003 Applied partial differential equations, revised ed. O. U. P., Oxford.
- Royama, T. 1992 Analytical population dynamics. Chapman and Hall, London.
- Samarskii, A. A., V. A. Galaktionov, S.P. Kurdyumov and A. P. Mikhailov 1995 Blow-up in quasilinear parabolic equations. de Gruyter Expositions in Mathematics, vol. 19. de Gruyter, Berlin.
- Simmons, G. F. 1972 Differential equations with applications and historical notes. McGraw-Hill, New York.
- Stakgold, I. 2000 Boundary value problems of mathematical physics, vol. I. SIAM, Philadelphia.
- Stoker, J. J. 1957 Water waves. Interscience, New York.
- Taylor, A. B. 1986 Mathematical models in applied mechanics. Clarendon Press, Oxford.
- Thompson, J.M. T. and H. B. Stewart 1986 Nonlinear dynamics and chaos. John Wiley, Chichester.
- Whitham, G. B. 1974 Linear and nonlinear waves. John Wiley, New York.

Zammett, R. 2004 The formation and evolution of spiral troughs in the Martian North Polar Ice Cap. M. Sc. dissertation, University of Oxford.