Springer Undergraduate Mathematics Series

S
U
M
S

Thomas Witelski
Mark Bowen

Methods of Mathematical Modelling

Continuous Systems and Differential Equations

Springer

Springer Undergraduate Mathematics Series

Advisory Board

M.A.J. Chaplain, *University of Dundee, Dundee, Scotland, UK*
K. Erdmann, *University of Oxford, Oxford, England, UK*
A. MacIntyre, *Queen Mary, University of London, London, England, UK*
E. Süli, *University of Oxford, Oxford, England, UK*
M.R. Tehranchi, *University of Cambridge, Cambridge, England, UK*
J.F. Toland, *University of Cambridge, Cambridge, England, UK*

More information about this series at http://www.springer.com/series/3423

Thomas Witelski · Mark Bowen

Methods of Mathematical Modelling

Continuous Systems and Differential Equations

Springer

Thomas Witelski
Department of Mathematics
Duke University
Durham, NC
USA

Mark Bowen
International Center for Science and Engineering Programs
Waseda University
Tokyo
Japan

ISSN 1615-2085 ISSN 2197-4144 (electronic)
Springer Undergraduate Mathematics Series
ISBN 978-3-319-23041-2 ISBN 978-3-319-23042-9 (eBook)
DOI 10.1007/978-3-319-23042-9

Library of Congress Control Number: 2015948859

Mathematics Subject Classification: 34-01, 35-01, 34Exx, 34B40, 35Qxx, 49-01, 92-XX

Springer Cham Heidelberg New York Dordrecht London
© Springer International Publishing Switzerland 2015
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

For Yuka and Emma,

For Mom and Hae-Young, and

For students seeking mathematical tools to model new challenges…

Preface

What is Mathematical Modelling?

In order to explain the purpose of modelling, it is helpful to start by asking: *what is a mathematical model*? One answer was given by Rutherford Aris [4]:

> A model is a set of mathematical equations that … provide an adequate description of a physical system.

Dissecting the words in his description, "*a physical system*" can be broadly interpreted as any real-world problem—natural or man-made, discrete or continuous and can be deterministic, chaotic, or random in behaviour. The context of the system could be physical, chemical, biological, social, economic or any other setting that provides observed data or phenomena that we would like to quantify. Being "*adequate*" sometimes suggests having a minimal level of quality, but in the context of modelling it describes equations that are good enough to provide sufficiently accurate predictions of the properties of interest in the system without being too difficult to evaluate.

Trying to include every possible real-world effect could make for a complete description but one whose mathematical form would likely be intractable to solve. Likewise, over-simplified systems may become mathematically trivial and will not provide accurate descriptions of the original problem. In this spirit, Albert Einstein supposedly said, *"Everything should be made as simple as possible, but not simpler"* [107], though ironically this is actually an approximation of his precise statement [34].

Many scientists have expressed views about the importance of modelling and the limitations of models. Some other notable examples are:

- In the opening of his foundational paper on developmental biology, Alan Turing wrote *"This [mathematical] model will be a simplification and an idealisation, and consequently a falsification. It is to be hoped that the features retained for discussion are those of greatest importance ..."* [100]
- George Box wrote *"...all models are wrong, but some are useful."* [17]

- Mark Kac wrote *"Models are, for the most part, caricatures of reality, but if they are good, they portray some features of the real world."* [55]

Useful models strike a balance between such extremes and provide valuable insight into phenomena through mathematical analysis. Every proposed model for a problem should include a description of how results will be obtained—a solution strategy. This suggests an operational definition:

> **model**: a useful, practical description of a real-world problem, capable of providing systematic mathematical predictions of selected properties

Models allow researchers to assess balances and trade-offs in terms of levels of calculational details versus limitations on predictive capabilities.

Concerns about models being "wrong" or "false" or "incomplete" are actually criticisms of the levels of physics, chemistry or other scientific details being included or omitted from the mathematical formulation. Once a well-defined mathematical problem is set up, its mathematical study can be an important step in understanding the original problem. This is particularly true if the model predicts the observed behaviours (a positive result). However, even when the model does not work as expected (a negative result), it can lead to a better understanding of which (included or omitted) effects have significant influence on the system's behaviour and how to further improve the accuracy of the model.

While being mindful of the possible weaknesses, the positive aspects of models should be praised,

> Models are expressions of the hope that aspects of complicated systems can be described by simpler underlying mathematical forms.

Exact solutions can be found for only a very small number of types of problems; seeking to extend systems beyond those special cases often makes the exact solutions unusable. Modelling can provide more viable and robust approaches, even though they may start from counterintuitive ideas, *"... simple, approximate solutions are more useful than complex exact solutions"* [15].

Mathematical models also allow for the exploration of conjectures and hypothetical situations that cannot normally be de-coupled or for parameter ranges that might not be easily accessible experimentally or computationally. Modelling lets us qualitatively and quantitatively dissect problems in order to evaluate the importance of their various parts, which can lead to the original motivating problem becoming a building block for the understanding of more complex systems. Good models provide the flexibility to be systematically developed allowing more accurate answers to be obtained by solving extensions of the model's mathematical equations. In summary, our description of the process is

> **modelling**: a systematic mathematical approach to formulation, simplification and understanding of behaviours and trends in problems.

Levels of Models

Mathematical models can take many different forms spanning a wide range of types and complexity,

$$\boxed{\textit{“Toy”}\text{ problems}} \quad \leq \quad \text{Math Models} \quad \leq \quad \boxed{\text{Complete systems}}$$

At the upper end of complexity are models that are equivalent to the full first-principles scientific description of all of the details involved in the entire problem. Such systems may consist of dozens or even hundreds of equations describing different parts of the problem; computationally intensive numerical simulations are often necessary to investigate the full system.

At the other end of the spectrum are improvised or phenomenological “toy” problems[1] that may have some conceptual resemblance to the original system but have no obvious direct derivation from that problem. These might be only a few equations or just some geometric relations. They are the mathematical modelling equivalents of an “artistic impression” motivated or inspired by the original problem. Their value is that they may provide a simple “proof of concept” prototype for how to describe a key element of the complete system.

Both extremes have drawbacks: intractable calculations in one extreme, and imprecise qualitative results at the other. Mathematical models exist in-between and try to bridge the gap by offering a process for using identifiable assumptions to reduce the full system down to a simpler form, where analysis, calculations and insights are more achievable, but without losing the accuracy of the results and the connection to the original problem.

Classes of Real World Problems

The kinds of questions being considered play an important role in how the model for the problem should be constructed. There are three broad types of questions:

(i) Evaluation questions [also called *Forward problems*]: Given all needed information about the system, can we quantitatively predict its other properties and how the system will function? Examples: What is the maximum attainable speed of this car? How quickly will this disease spread through the population of this city?

(ii) Detection questions [*Inverse problems*] [8]: If some information about a “black box” system is not directly available, can you “reverse engineer” those missing parameters? Examples: How can we use data from CAT scans to

[1]Sometimes also described as *ad hoc* or *heuristic* models.

estimate the location of a tumour? Can we determine the damping of an oscillator from the decay of its time series data?

(iii) Design questions [*Control and optimisation problems*]: Can we create a solution that best meets a proposed goal? Examples: What shape paper airplane flies the furthest? How should a pill be coated to release its drug at a constant rate over an entire day?

There are many routes available to attack such questions that are typically treated in different areas of study. This book will introduce methods for addressing some problems of the forms (i) and (iii) in the context of continuous systems and differential equations.

Stages of the Modelling Process

The modelling process can sometimes start from a creative and inspired toy problem and then seeks to validate the model's connection to the original problem. However, this approach requires having a lot of previous experience with and background knowledge on the scientific area and/or relevant mathematical techniques in order to generate the new model. In this book, we follow the more systematic approach of starting with some version of the complete scientific problem statement and then using mathematical techniques to obtain reduced models that can be simplified to a manageable level of computational difficulty.

The modelling process has two stages, consisting of setting up the problem and then solving it:

- In the *formulation phase*, the problem is described using basic principles or governing laws and assumed relations taken from some branches of knowledge, such as physics, biology, chemistry, economics, geometry, probability or others. Then all side-conditions that are needed to completely define the problem must be identified: geometric constraints, initial conditions, material properties, boundary conditions and design parameter values. Finally, the properties of interest, how they are to be measured, relevant variables, coordinate systems and a system of units must all be decided on.
- Then[2], in the *solution phase*, mathematical modelling provides approaches to reformulating the original problem into a more convenient structure from which it can be reduced into solvable parts that can ultimately be re-assembled to address the main questions of interest for the problem.

In some cases, the reformulated problem may seem to only differ from the original system at a notational level, but these changes can be essential for separating out different effects in the system. At the simplest level, "problem reduction" consists of

[2]Assuming that the problem cannot be easily solved analytically or computed numerically, and hence does not need modelling.

obtaining so-called *asymptotic approximations* of the solution, but for more challenging problems, this will also involve approaches for transforming the problem into different forms that are more tractable for analysis or computation.

The techniques described here are broadly applicable to many branches of engineering and applied science: biology, chemistry, physics, the geosciences and mechanical engineering, to name a few. To keep examples compact and accessible, we present concise reviews of background from different fields when needed (including population biology and chemical reactions in Chap. 1, fluid dynamics in Chap. 2 and classical mechanics in Chap. 3), but we seek to maintain focus on the modelling techniques and the properties of solutions that can be obtained. We direct interested readers to books that present more detailed case studies of problems needing more extensive background in specific application areas [8, 27, 37, 38, 51, 69, 96].

The structure of this book follows the description of the modelling process described above:

- Part I: Formulation of models
  This part consists of four chapters that present fundamental approaches and exact methods for formulating different classes of problems:

  1. Rate equations: simple models for properties evolving in time
  2. Transport equations: models involving structural changes
  3. Variational principles: models based on optimisation of properties
  4. Dimensional analysis: determination of the number of essential system parameters

- Part II: Solution techniques
  This part presents methods for obtaining approximate solutions to some of the classes of problems introduced in Part I.

  5. Similarity solutions: determining important special solutions of PDEs using scaling analysis
  6. Perturbation methods: exploiting limiting parameters to obtain expansions of solutions
  7. Boundary layers: constructing solutions having non-uniform spatial structure
  8. Long-wave asymptotics: reduction of problems on slender domains
  9. Weakly-nonlinear oscillators: predicting cumulative changes over large numbers of oscillations
  10. Fast/slow dynamical systems: separating effects acting over different timescales
  11. Reduced models: obtaining essential properties from simplified versions of partial differential equation problems

- Part III: Case studies: some applications illustrating uses of techniques from Parts I and II.

While this book cannot be an exhaustive introduction to all types of mathematical models, we seek to develop intuition from the ground-up on formulating equations and methods of solving models expressible in terms of differential equations.

The book is written in a concise and self-contained form that should be well-suited for an advanced undergraduate or beginning graduate course or independent study. Students should have a background in calculus and basic differential equations. Each chapter provides references to sources that can provide more detail on topics that readers wish to pursue in greater depth. We note that the examples and exercises are an important part of the book and will introduce readers to many classic models that have become important milestones in applied mathematics for illustrating important or universal solution structures. Some of these highlights include:

- The Burgers equation (Chaps. 2, 4, 5)
- The shallow water equations (Chaps. 2, 4, 6)
- The porous medium equation (Chaps. 5, 8, 11)
- The Korteweg de Vries (KdV) equation (Chaps. 8, 9)
- The Fredholm alternative theorem (Chap. 9)
- The van der Pol equation (Chaps. 9, 10)
- The Michaelis–Menten reaction rate model (Chap. 10)
- The Turing instability mechanism (Chap. 11)
- Taylor dispersion (Chap. 11)

Solutions are provided to many exercises. Readers are encouraged to work through the exercises in order to gain a deeper understanding of the techniques presented.

Durham, NC, USA — Thomas Witelski
Tokyo, Japan — Mark Bowen
June 2015

Acknowledgments

This book follows in the tradition set by the classic text by Lin and Segel [63] that first collected and systematically applied the foundational approaches in modelling. More recent books by Tayler [96], Fowler [37], Haberman [45], Logan [64], Holmes [49] and Howison [51] have made the art of modelling and the tools of applied mathematics more accessible.

This book shares many elements in common with those books but seeks to highlight different connections between topics and to use elementary approaches to make modelling more applicable to students coming from a diverse range of fields seeking to incorporate mathematical modelling in their scientific studies.

The authors thank the many colleagues and students who gave feedback on early versions of the materials in this book. Thomas Witelski also thanks the OCCAM and OCIAM centres in the Mathematical Institute of the University of Oxford for a visiting fellowship.

Durham, NC, USA — Thomas Witelski
Tokyo, Japan — Mark Bowen
June 2015

Contents

Part I Formulation of Models

1 Rate Equations 3
1.1 The Motion of Particles 4
1.2 Chemical Reaction Kinetics 5
1.3 Ecological and Biological Models 8
1.4 One-Dimensional Phase-Line Dynamics 10
1.5 Two-Dimensional Phase Plane Analysis 12
1.5.1 Nullclines 16
1.6 Further Directions 18
1.7 Exercises 19

2 Transport Equations 23
2.1 The Reynolds Transport Theorem 24
2.2 Deriving Conservation Laws 26
2.3 The Linear Advection Equation 28
2.4 Systems of Linear Advection Equations 30
2.5 The Method of Characteristics 32
2.6 Shocks in Quasilinear Equations 34
2.7 Further Directions 39
2.8 Exercises 39

3 Variational Principles 47
3.1 Review and Generalisation from Calculus 47
3.1.1 Functionals 49
3.2 General Approach and Basic Examples 50
3.2.1 The Simple Shortest Curve Problem 51
3.2.2 The Classic Euler–Lagrange Problem 55
3.3 The Variational Formation of Classical Mechanics 56
3.3.1 Motion with Multiple Degrees of Freedom 58

$$m\frac{d^2\mathbf{X}}{dt^2} = \mathbf{F}. \tag{1.4}$$

Note that every nth order autonomous differential equation can always be re-expressed in the form of a system of n (first-order) rate equations. Defining $\mathbf{x} = (X_1, X_2, X_3, P_1, P_2, P_3)$, system (1.3) can be put in form (1.2) as

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}) \qquad \text{with} \qquad \mathbf{f} = \begin{pmatrix} \mathbf{P}/m \\ \mathbf{F}(\mathbf{X}, \mathbf{P}) \end{pmatrix} \in \mathbb{R}^6. \tag{1.5}$$

Newton's laws give a well-defined procedure for constructing the dynamical system for a mechanics problem.[1] For other fields, different principles provide guidance; we will review how to set up rate equations for problems in chemistry and biology.

1.2 Chemical Reaction Kinetics

For chemical systems [6], the fundamental principle for translating chemical reactions into corresponding sets of rate equations is given by the *law of mass action*.

In "simple" (or *elementary*) reactions, generically of the form

$$\text{reactants} \xrightarrow{k} \text{products}, \tag{1.6}$$

the rate of creation of products depends on the concentrations of available reactants and is also characterised by a *rate constant* k. The rate of consumption of reactants also follows from this relation.

We will denote the concentration of chemical 'A' by $A(t) \geq 0$, and the total rate of production of A will have contributions from its creation and/or consumption due to each chemical reaction involving A,

$$\frac{dA}{dt} = +\sum_{n=1}^{N} (\text{creation rate})_n - \sum_{n=1}^{N} (\text{consumption rate})_n. \tag{1.7}$$

The rate equations form a system of first-order equations, one for each chemical in the set of reactions, $\{A, B, C, \ldots\}$, and the rate functions will be polynomials in terms of products of concentrations of the reactant chemicals.

We briefly summarise the basic forms of elementary chemical reactions and the corresponding rate equations that follow from the law of mass action:

[1] In Chap. 3 we will consider a different approach.

(0) **Constant supply**: compound A is pumped into the system at a constant rate k

$$(\text{source}) \xrightarrow{k} A \qquad \Longrightarrow \qquad \frac{dA}{dt} = k. \tag{1.8}$$

This is called a zeroth-order reaction since the rate does not depend on the concentration of any reactants, $k = k \cdot 1 = kA^0$.

(i) **Decay**: substance A transforms into waste at rate k (i.e. A decomposes and is removed from the system)

$$A \xrightarrow{k} (\text{waste}) \qquad \Longrightarrow \qquad \frac{dA}{dt} = -kA. \tag{1.9}$$

This is called a first-order reaction since the reaction rate depends linearly on the concentration of a single reactant.

(ii) **Transformation**: A is consumed with B being produced from A

$$A \xrightarrow{k} B \qquad \Longrightarrow \qquad \frac{dA}{dt} = -kA, \qquad \frac{dB}{dt} = +kA. \tag{1.10}$$

This is the simplest reaction "system", having two distinct concentrations evolving due to a single reaction.

(iii) **Reversible transformation**: A transforms into B and vice versa. Such reactions should be explicitly expanded into separate forward and reverse reactions.

$$A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} B \qquad = \qquad \{\, A \xrightarrow{k_1} B \quad \text{and} \quad A \xleftarrow{k_2} B \,\}$$

$$\Longrightarrow \qquad \frac{dA}{dt} = -k_1 A + k_2 B, \qquad \frac{dB}{dt} = k_1 A - k_2 B. \tag{1.11}$$

In this system, the net rate of production of each substance is obtained by summing the reaction rate from the reaction producing it minus the rate from the reaction consuming it, as in (1.7).

(iv) **Compound formation**: A and B combine to form C

$$A + B \xrightarrow{k} C \qquad \Longrightarrow$$

$$\frac{dA}{dt} = -kAB, \qquad \frac{dB}{dt} = -kAB, \qquad \frac{dC}{dt} = kAB. \tag{1.12}$$

The production rate of C being proportional to the product of the reactant concentrations follows from a probabilistic description of the collision of independent molecules [6]. The probability of forming a molecule of C increases when either of the concentrations of A, B increases (and clearly the reaction will not proceed if either is absent, $A = 0$ or $B = 0$).

However, suppose A and B are the same chemical, then from (1.12)

$$\begin{aligned} A + A \xrightarrow{k} C \quad &= \quad 2A \xrightarrow{k} C \quad \Longrightarrow \\ \frac{dC}{dt} = +kA^2 \quad &\left(\text{but} \quad \frac{dA}{dt} \neq -kA^2 \quad (?)\right) \end{aligned} \tag{1.13}$$

The equation for the rate of consumption of A cannot be correct because it implies that A, C have equal but opposite rates of change, while we know that formation of each molecule of C should consume *two* molecules of A. This is an issue with "double-counting" that points to the need for a more precise definition of the "reaction rate," as will be addressed in the final type of elementary reaction.

(v) **Multiple products**: n molecules of A and m molecules of B react to produce p molecules of C and q molecules of D:

$$n\ A + m\ B \xrightarrow{k} p\ C + q\ D \tag{1.14a}$$

The problem associated with (1.13) is resolved by defining the reaction rate to be

$$\begin{aligned} \text{Reaction Rate} &= -\text{rate of consuming one unit of reactant} \\ &= +\text{rate of creating one unit of product.} \end{aligned}$$

Then for (1.14a) we get

$$\text{Rate} = \frac{1}{p}\frac{dC}{dt} = \frac{1}{q}\frac{dD}{dt} = -\frac{1}{n}\frac{dA}{dt} = -\frac{1}{m}\frac{dB}{dt} = kA^n B^m$$

and can write the rate equations as

$$\text{reactants:} \quad \frac{dA}{dt} = -nkA^n B^m \qquad \frac{dB}{dt} = -mkA^n B^m \tag{1.14b}$$

$$\text{products:} \quad \frac{dC}{dt} = pkA^n B^m \qquad \frac{dD}{dt} = qkA^n B^m \tag{1.14c}$$

Consequently, the rate equation for A for the reaction (1.13) is now correctly given by

$$\frac{dA}{dt} = -2kA^2. \tag{1.15}$$

For systems without losses or sources of chemicals, as in the case of (1.10) or (1.11), physical expectations based on the conservation of mass suggest that $A + B =$ constant, which is validated by evaluating $\frac{d}{dt}(A + B)$ using the rate equations. This is called a *conservation law* linking the products and reactants. Typically, the value for the constant is set by the summed initial concentrations and can be used to express one concentration in terms of the other. For example, if $A + B =$ constant, then

$B(t) = (A_0 + B_0) - A(t)$, allowing us to reduce the number of unknowns in the system. In some cases there may be various combinations of reactants/products that can be used to write a conservation law. For example, for (1.14a), one possible form of mass balance is given by

$$\frac{d}{dt}\left(\frac{1}{n}A + \frac{1}{p}C\right) = 0.$$

1.3 Ecological and Biological Models

Rate equations are also widely used for predicting population growth and the spread of epidemics (namely, the growth of the population of diseased individuals) [18, 74].

The simplest model of population growth, called Malthus' law, describes a growth rate proportional to the current population, and generates exponential growth,

$$\frac{dA}{dt} = kA \qquad \Longrightarrow \qquad A(t) = A_0 e^{kt}. \tag{1.16}$$

In analogy with the description of chemical reactions, using (1.14a) the growth of the population would be due to additions by births, with the parent remaining in the population,

$$A \xrightarrow{\beta} A + A \qquad \Longrightarrow \qquad \frac{dA}{dt} = -\beta A + 2\beta A = \beta A. \tag{1.17}$$

We note that the rate constant β would incorporate the time needed for the birth process (the gestation period) as well as the fraction of the total population who choose to become parents. More generally, the net rate of population growth would be the difference between the birth rate β and the death (decay) rate δ, $k = \beta - \delta$.

The weakness of Malthus' law with respect to predicting unlimited growth of populations lead to an improved model developed by Pierre Verhulst (1804–1849), usually called the *logistic equation*,

$$\frac{dA}{dt} = (\beta - \delta)A - \gamma A^2 = (k - \gamma A)A, \tag{1.18}$$

where the coefficient γ scales the influence of competition effects in decreasing the birth rate (and/or increasing the death rate) in growing populations. Here, the effective growth rate, $\tilde{k}(A) = k - \gamma A$, depends on the population size and changes the dynamics from growth ($\tilde{k} > 0$) for small populations to decay ($\tilde{k} < 0$) for large populations. The borderline between these two cases defines a critical population size, $A_* = k/\gamma$, called the carrying capacity, which allow (1.18) to be written as

$$\frac{dA}{dt} = k\left(1 - \frac{A}{A_*}\right)A. \tag{1.19}$$

The logistic equation is the simplest model that makes progress in capturing the influence of limited external resources (e.g. food) on growth of the population.

A more detailed way to describe the coupling of populations to their use of resources is to introduce additional rate equations to also describe the growth/decay of the resources. Consider a population of rabbits whose number evolves according to the logistic equation (1.18). Let that population of rabbits, $A(t)$, also serve as the food supply (prey) for a population of foxes (predators), $B(t)$. Let the foxes have a constant death rate and reproduce only when food available, analogous to a chemical reaction of the form, $A + B \to B + B$. The influence of the consumption of prey by predators is called *predation* and generates an additional loss term in the logistic equation (1.19) for the rabbit population. Along with the rate equation for the fox population, this yields a Lotka-Volterra predator-prey system [45, 74],

$$\frac{dA}{dt} = (\beta - \delta)A - \gamma A^2 - \rho AB, \qquad \frac{dB}{dt} = -\kappa B + \sigma AB. \tag{1.20}$$

In the context of the spread of diseases [18], simple models of epidemics divide the total population into sub-groups depending on whether individuals are infected ($I(t)$), susceptible to the disease ($S(t)$) or recovering from the disease ($R(t)$). The transitions between these states can be interpreted as reactions

$$I + S \to I + I, \qquad I \to R,$$

yielding a basic version of what are generally called SIR models

$$\frac{dS}{dt} = -kSI, \qquad \frac{dI}{dt} = kSI - \gamma I, \qquad \frac{dR}{dt} = \gamma I. \tag{1.21}$$

This particular system conserves the total population, $N = S + I + R$, but other formulations can allow for growing or declining overall populations. Diseases for which immunity is not achievable can be described by reversible transitions between susceptible and infected states, analogous to (1.11),

$$S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} I \qquad \Longrightarrow \qquad \frac{dS}{dt} = -k_1 S + k_2 I, \qquad \frac{dI}{dt} = k_1 S - k_2 I, \tag{1.22}$$

and are generally called SIS models. Many further extensions are possible, including, for example, subdividing the infected population into individuals that are in an earlier, exposed phase, $E(t)$, versus a later infectious phase, $I(t)$, called *SEIR* models [19].

1.4 One-Dimensional Phase-Line Dynamics

The simplest applications described above were expressed as initial value problems for a single autonomous ODE (e.g. (1.15), (1.19)),

$$\frac{dx}{dt} = f(x), \qquad x(0) = x_0. \tag{1.23}$$

If the rate function $f(x)$ is linear or a polynomial function of x, then various methods, such as separation of variables can be used to determine the exact solution. However, we will see that many problems will yield more complicated rate functions where the solution cannot be explicitly calculated. For this reason, it is very useful to have a general theory that provides an understanding of all solutions of (1.23) without having to find explicit solutions. Further, even when a solution can be written out, it may be complicated to analyse. In contrast, there is a qualitative approach that can be easier to calculate and to interpret.

The dynamics of the solutions of (1.23) can be described by a one-dimensional[2] "phase line" in terms of the graph of the function $f(x)$ (see Fig. 1.1 (left)). Assuming that $f(x)$ is a nice function[3] the behaviour of all solutions, $x(t)$, can be described qualitatively in relation to the properties of $f(x)$ at its zeroes. The values $x = x_*$ where $f(x_*) = 0$ are called *equilibrium points* of (1.23). The values x_* define steady state solutions since where $dx/dt = 0$, solutions starting at x_* remain there forever. The behaviour of solutions starting from within a small neighbourhood around x_* can be analysed by approximating $f(x)$ by its Taylor series about $x = x_*$,

$$f(x) \approx f(x_*) + f'(x_*)(x - x_*) + \tfrac{1}{2} f''(x_*)(x - x_*)^2. \tag{1.24}$$

Writing the separation from the equilibrium point as $u(t) = x(t) - x_*$, (1.24) becomes

$$f(x) \approx 0 + au + bu^2 + cu^3 + \cdots \qquad \text{as } u \to 0, \tag{1.25}$$

where $a = f'(x_*)$. If $a \neq 0$, then as $u \to 0$ (corresponding to $x \to x_*$), (1.23) can be approximated by the linearised equation,

$$\frac{du}{dt} = au \qquad \Rightarrow \qquad u(t) = Ce^{at}, \tag{1.26}$$

and hence yields $x(t) \approx x_* + Ce^{at}$ with $C \neq 0$ for any solution not starting exactly at the equilibrium point. For $a > 0$ the equilibrium is locally 'repelling' since the separation from the equilibrium point increases with time, while conversely for $a < 0$ the separation decreases with time and the equilibrium is locally 'attracting'. More generally, if *any* solution starting near an equilibrium point leaves the neighbourhood

[2] Here one-dimensional indicates that the dynamics of solutions can be understood in terms of a single variable, x.

[3] With f being bounded and sufficiently smooth.

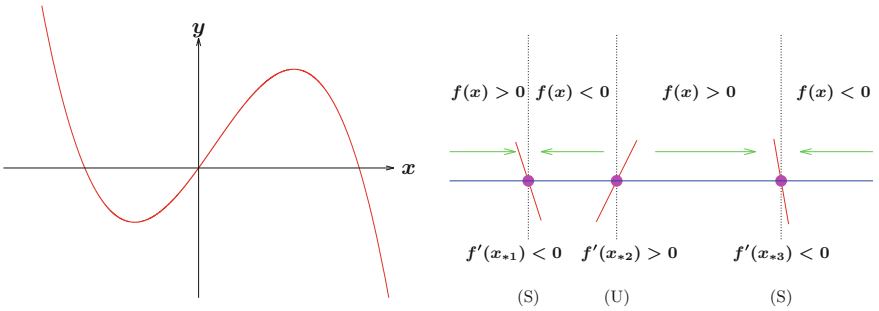


Fig. 1.1 (*Left*) A smooth rate function $f(x)$ with three equilibrium points, (*Right*) corresponding dynamics on the phase line for (1.23) obtained from local properties of f at its equilibrium points

of x_* as $t \to \infty$, then x_* is called *asymptotically unstable*, while if *all* solutions starting within the neighbourhood approach x_* as $t \to \infty$ then the equilibrium is called *asymptotically stable*. The characterisation of the dynamics of the solutions near x_* based on the linearised equation (1.26) is commonly called *linear stability analysis*.

In graphical terms, x_* is stable if the slope $f'(x_*)$ is negative ($a < 0$) and unstable if the slope is positive ($a > 0$). If we consider $x(t)$ to be the position of a point moving along the x-axis in Fig. 1.1 (the phase line) then this follows from f's role as the velocity, or rate of change of position x. If x starts to the right of x_{*3} ($x_0 > x_{*3}$) where $f(x) < 0$ then $x(t)$ will decrease (move to the left, $dx/dt < 0$) back towards x_{*3}. In contrast if $x_0 > x_{*2}$ where $f(x) > 0$ then $x(t)$ would increase with time ($dx/dt > 0$), moving away from x_{*2}.

Linear stability results provide guidance to understanding the global dynamics (not limited to small neighbourhoods of the x_*'s). Since f changes sign only at its zeroes (where we have assumed $f'(x_*)$ to be non-zero), solutions starting at x_0 within intervals between zeroes will either be monotone increasing (moving to the right on the phase line) or decreasing (moving to the left on the phase line) depending on the sign of f. This is consistent with the results of the local stability analysis at each x_* being controlled by the sign of $f'(x_*)$.

The case where $a = f'(x_*) = 0$ is called a degenerate equilibrium point; the local analysis for $u \to 0$ is still addressed using the Taylor series (1.25), but now the first nontrivial term in the expansion is (at least) quadratic and so the rates of growth or decay will be algebraic rather than exponential. If $f''(x_*) \neq 0$, then the ODE

$$\frac{du}{dt} = bu^2 \tag{1.27}$$

can be used to show that a degenerate equilibrium point x_* is unstable for any $b \neq 0$. Similarly, for the case with $f'(x_*) = f''(x_*) = 0$, but $f'''(x_*) \neq 0$ yielding

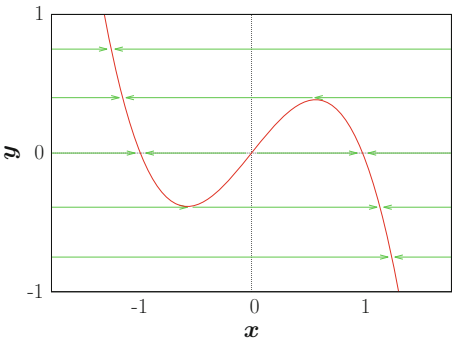


Fig. 1.2 The graph of $y = f(x) = h(x) - k$ for (1.29), giving the dynamics for $x(t)$ on different phase lines, graphically parametrised by $y = k$

$$\frac{du}{dt} = cu^3, \tag{1.28}$$

a third-order degenerate equilibrium point can be shown to be stable if $f'''(x_*) < 0$.

The dependence of solutions on parameters in the system will be an important aspect of many problems. Consider the ODE

$$\frac{dx}{dt} = h(x) - k \qquad \text{where } h(x) = x - x^3, \tag{1.29}$$

where k is a parameter. Matching to (1.23), we identify the rate function as $f(x) = h(x) - k$. Plotting $f(x)$ in Fig. 1.2, we illustrate how the family of problems (1.29) parametrised by k can be understood from this graph. Observe that for different values of the constant k, the horizontal lines $y = k$ form a "stacked" set of phase lines cutting through the curve $y = h(x)$ at the equilibrium points, $f(x_*) = 0$, corresponding to that value of k. The positions of the equilibrium are functions of k, but more significantly, the number and type of the equilibrium points change for different ranges of k, such qualitative changes in the structure of sets of solutions are called *bifurcations*. For (1.29), the dynamics for all times is restricted to a single phase line, with $y = k$ fixed, in the next section we'll review systems with coupling to a second rate equation for $y(t)$ yielding general motion in the xy plane. Despite its restricted form, we will see that problems like (1.29) occur as pseudo-two-dimensional phase plane structures in reductions of more complicated systems in Chap. 10.

1.5 Two-Dimensional Phase Plane Analysis

For systems of two coupled autonomous rate equations (called *phase plane systems*) the approach of the previous section can be extended to similarly give a qualitative understanding of all solutions from just local properties of the rate functions.

For $n = 2$ with $\mathbf{x}(t) = (x(t), y(t))$, Eq. (1.2) can be written as

$$\frac{dx}{dt} = f(x, y), \qquad \frac{dy}{dt} = g(x, y), \tag{1.30a}$$

with initial conditions at $t = 0$,

$$x(0) = x_0, \qquad y(0) = y_0. \tag{1.30b}$$

Interpreting variable t as time, this problem can be interpreted as describing the motion of a point in the xy "phase plane" starting from position $\mathbf{x}_0 = (x_0, y_0)$ subject to a position-dependent velocity $\mathbf{v} = (f, g)$. The solution of (1.30a, 1.30b) is a parametric curve (or *trajectory*) passing through the point $\mathbf{x}_0$.

Using the chain rule to eliminate t from $y(x) = y(x(t))$, (1.30a) leads to the *associated ODE* (also called the *slope field equation*),

$$\frac{dy}{dx} = \frac{g(x, y)}{f(x, y)}. \tag{1.31}$$

Solving this first order ODE for $y(x)$ yields a family of *implicit solutions* called the *integral curves* that trace out the same paths in the plane as the trajectories. The implicit solutions give only the shape of the curves while parametric solutions also give the direction of motion on the curves with respect to increasing t.

Analogous to the analysis of the phase line, equilibrium points (x_*, y_*) are defined by positions where both rate functions vanish,

$$f(x_*, y_*) = 0, \qquad g(x_*, y_*) = 0. \tag{1.32}$$

With the exception of equilibrium points, results on uniqueness of solutions ensure that each point $\mathbf{x}_0$ has a single solution curve passing through it and that those curves cannot cross. Consequently, assuming f and g are smooth, the solutions are smooth curves everywhere in the phase plane except at the equilibrium points. Observe from (1.31) that at an equilibrium point the slope dy/dx is indeterminate (formally $dy/dx =$'0/0') and requires more careful analysis to describe the local structure of solutions.

Near an equilibrium point, define $\mathbf{u}(t) = (x(t)-x_*, y(t)-y_*)$ satisfying $d\mathbf{u}/dt = \mathbf{f}(\mathbf{u} + \mathbf{x}_*)$, and using a multi-variable Taylor series approximation of $\mathbf{f}$ for $|\mathbf{u}| \to 0$ yields the linearised system

$$\frac{d\mathbf{u}}{dt} = \mathbf{A}\mathbf{u}, \tag{1.33}$$

where the matrix $\mathbf{A}$ is the *Jacobian* (or gradient $\nabla\mathbf{f}$), evaluated at (x_*, y_*),

$$\mathbf{A} = \mathbf{J}(x_*, y_*) \equiv \begin{pmatrix} \partial_x f(x_*, y_*) & \partial_y f(x_*, y_*) \\ \partial_x g(x_*, y_*) & \partial_y g(x_*, y_*) \end{pmatrix}. \tag{1.34}$$

Seeking solutions of the form $\mathbf{u}(t) = \mathbf{v}e^{\lambda t}$ then reduces (1.33) to the matrix eigenvalue problem,

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}.$$

At each equilibrium point, linear algebra yields the eigenvalues as the roots of the characteristic polynomial obtained from the setting the determinant to zero,

$$|\mathbf{A} - \lambda\mathbf{I}| = 0.$$

Subsequently, for each eigenvalue, the eigenvectors can be obtained by row-reductions as the nontrivial nullvector of $(\mathbf{A} - \lambda_k\mathbf{I})\mathbf{v}_k = \mathbf{0}$. The general solution of the linearised system is then given by the linear combination of the eigenmodes,

$$\mathbf{x}(t) \approx \mathbf{x}_* + c_1\mathbf{v}_1e^{\lambda_1 t} + c_2\mathbf{v}_2e^{\lambda_2 t} \qquad \text{for } |\mathbf{x} - \mathbf{x}_*| \to 0. \tag{1.35}$$

If $\mathbf{A}$ does not have a complete set of eigenvectors (a possibility with repeated eigenvalues) then this form must be modified.

Extending the discussion of asymptotic stability from Sect. 1.4 in terms of all solutions approaching $\mathbf{x}_*$ or any diverging from it, the stability of solutions starting near $\mathbf{x}_*$ can be understood in terms of the eigenvalues, see Table 1.1. Problems having $\mathrm{Re}(\lambda) = 0$ fall into a degenerate case, called a centre manifold, and must be studied more carefully, somewhat like (1.27, 1.28).

While for the phase line case the behaviour of solutions could be inferred directly from the slope of the rate function, $\lambda = a = f'(x_*)$, the geometry for the phase plane case is more complicated. The stability properties are still set by the derivatives of the rate functions at $\mathbf{x}_*$ but now the different cases are most conveniently expressed in terms of the eigenvalues of $\mathbf{A}$, (1.35). The local geometry of the integral curves near non-degenerate $\mathbf{x}_*$ is then given by the cases outlined in Table 1.2, where the

Table 1.1 Asymptotic stability of equilibrium point $\mathbf{x}_*$ in terms of eigenvalues from the linear stability analysis

λ's	Stability	$t \to \infty$
Both $\mathrm{Re}(\lambda) < 0$	Stable	$\lvert\mathbf{u}(t)\rvert \to 0$
Either $\mathrm{Re}(\lambda) > 0$	Unstable	$\lvert\mathbf{u}(t)\rvert \to \infty$

Table 1.2 Geometry of trajectories near equilibrium point $\mathbf{x}_*$ in terms of the eigenvalues

λ's	Name	Geometry	Stability
λ_1, λ_2 same sign	Node	Rays	$\mathrm{Re}(\lambda)$
λ_1, λ_2 opp. signs	Saddle	Hyperbolas	Unstable
$\lambda = \pm i\beta$	Centre	Circles	Neutral
$\lambda = \alpha \pm i\beta$	Spiral	Spirals	$\mathrm{Re}(\lambda)$

eigenvectors set the orientation of the geometry (most notably the directions of the stable and unstable axes of a saddle point).

If $f(x, y)$ and $g(x, y)$ are linear functions, then there will be a single equilibrium point and the above analysis describes the form of the entire phase plane. Otherwise, this gives a description in a neighbourhood surrounding each $\mathbf{x}_*$. However, using the results on uniqueness of trajectories (i.e. curves may cross only at equilibrium points), trajectories can be smoothly extended and a global structure can be sketched by piecing-together the local linearised geometries around each of the equilibrium points.

In addition to systems of rate equations, phase plane analysis can also be applied to second-order autonomous ODEs for $x(t)$, such as

$$\frac{d^2x}{dt^2} - g\left(x, \frac{dx}{dt}\right) = 0, \tag{1.36}$$

which can be written as an equivalent system by defining an intermediate variable $y(t)$:

$$\frac{dx}{dt} = y, \qquad \frac{dy}{dt} = g(x, y). \tag{1.37}$$

In this system, the relationship $x' = y$ defines the direction of trajectories as follows:

- $x' > 0 : x(t)$ increasing ($\rightarrow$) for $y > 0$ (upper half xy plane)
- $x' < 0 : x(t)$ decreasing ($\leftarrow$) for $y < 0$ (lower half xy plane)

This choice for the intermediate variable is not unique but has a nice physical interpretation in term of $(x, y) =$ (position, velocity) with the x-axis being states at rest (zero velocity). An example of a different choice for $y(t)$ will be given in a later chapter.

As an example, consider the pendulum equation for the angular position of a suspended mass swinging under the influence of gravity [7],

$$\frac{d^2\theta}{dt^2} + \sin\theta = 0. \tag{1.38}$$

Letting $x(t) = \theta(t)$ and following (1.37), we arrive at the phase plane system,

$$\frac{dx}{dt} = y, \qquad \frac{dy}{dt} = -\sin x. \tag{1.39}$$

Applying (1.32), the equilibrium points are given by $y_* = 0$ with $x_* = n\pi$ for $n = 0, \pm 1, \pm 2, \ldots$. From (1.34) the Jacobian matrix at any (x, y) is

$$\mathbf{J}(x, y) = \begin{pmatrix} 0 & 1 \\ -\cos x & 0 \end{pmatrix}.$$

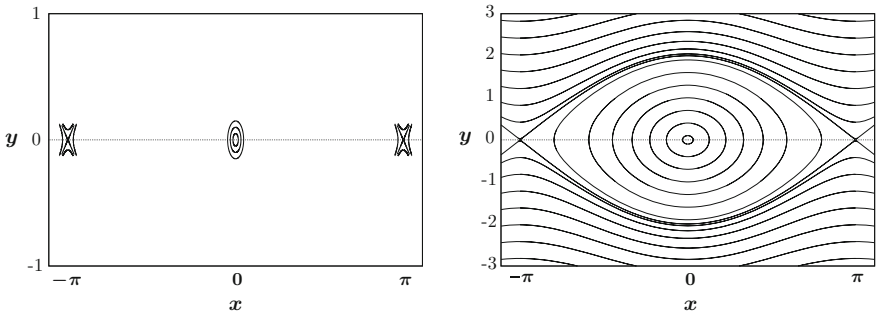


Fig. 1.3 The phase plane for the pendulum (1.39): (*Left*) Sketch of linearised behaviours in neighbourhoods of the equilibrium points and (*Right*) a computed plot of the full phase plane

Evaluating the Jacobian at the equilibrium point $(0, 0)$ yields

$$\mathbf{A} = \mathbf{J}(0,0) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \Longrightarrow \quad \lambda^2 + 1 = 0, \quad \lambda = \pm i \quad \text{(a centre)},$$

while the equilibrium point $(\pi, 0)$ yields

$$\mathbf{A} = \mathbf{J}(\pi,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \Longrightarrow \quad \lambda^2 - 1 = 0, \quad \lambda = \pm 1 \quad \text{(a saddle point)},$$

with eigenvectors $(1, -1)$ and $(1, 1)$. These vectors define the asymptotes of the family of hyperbolas centred at this equilibrium point, called the stable manifold (for $\lambda < 0$) and unstable manifold (for $\lambda > 0$), $y = \pm(x - \pi)$, see Fig. 1.3 (left). Other equilibrium points at $([2k + 1]\pi, 0)$ also have exactly the same form by the periodic nature of $g(x)$. Figure 1.3 shows that the full phase plane smoothly extends the local behaviours at the equilibria to cover all possible solutions. It is notable that trajectories connecting the saddles (called *heteroclinic orbits*) separate small-amplitude periodic oscillations abound $\theta = 0$ (the continuous family of closed curves) from high-speed "spinning" solutions, where the angle increases (or decreases) monotonically for all time.

1.5.1 Nullclines

Nullclines are curves in the xy plane that can provide further understanding of systems by dividing the phase plane into regions with different behaviours. The nullclines are *not* solutions of the system, but they do yield valuable insight on the properties of solution trajectories that pass through the nullcline curves (or lie on one side or the other).

The *x-nullcline* is a curve in the phase plane on which $dx/dt = 0$. On this nullcline $x(t)$ instantaneously has zero rate of change, hence x is fixed while $y(t)$ changes with time, so a trajectory passing through the x-nullcline will have a *vertical* tangent. The x-nullcline is given by the implicitly defined curve $f(x, y) = 0$. The x-nullcline also separates solutions that have $x(t)$ increasing from those that have x decreasing with time,

$$\text{Region where } f(x, y) < 0: \quad \frac{dx}{dt} < 0 \qquad x(t) \text{ decreasing } (\leftarrow)$$
$$\text{Region where } f(x, y) > 0: \quad \frac{dx}{dt} > 0 \qquad x(t) \text{ increasing } (\rightarrow)$$

Analogously, the *y-nullcline* is a curve on which $dy/dt = 0$. Here $y(t)$ instantaneously has zero rate of change and a trajectory passing through the y-nullcline will have a *horizontal* tangent. The y-nullcline is given by the implicitly defined curve $g(x, y) = 0$. The y-nullcline also separates solutions that have $y(t)$ increasing from those that have y decreasing with time,

$$\text{Region where } g(x, y) < 0: \quad \frac{dy}{dt} < 0 \qquad y(t) \text{ decreasing } (\downarrow)$$
$$\text{Region where } g(x, y) > 0: \quad \frac{dy}{dt} > 0 \qquad y(t) \text{ increasing } (\uparrow)$$

Combining the information on horizontal and vertical components of motion from the x, y-nullclines respectively, yields very useful local qualitative information on the directions of trajectories that can be extended to understand the structure of the phase plane. In particular, note that the intersections of the nullclines are equilibrium points and the transitions in values of dx/dt and dy/dt can help identify the type of equilibrium point without doing the linear stability analysis.

As an example consider the system

$$\frac{dx}{dt} = y - \frac{x^2}{\sqrt{2}}, \qquad \frac{dy}{dt} = x^2 + y^2 - 4. \tag{1.40}$$

The first equation gives the x-nullcline as the parabola $y = x^2/\sqrt{2}$. Above the parabola $x(t)$ is increasing with time, while below the parabola $x(t)$ is monotone decreasing. On the parabola, trajectories have vertical tangents. The second equation in (1.40) gives the y-nullcline as the circle $x^2 + y^2 = 4$. Inside the circle $y(t)$ is decreasing, while outside the circle $y(t)$ is increasing. On the circle, trajectories have horizontal tangents.

The intersection of the nullclines makes the locations of the two equilibrium points immediately clear. These points could also be obtained analytically from (1.32) as $(x_*, y_*) = (\pm\sqrt{2}, \sqrt{2})$. Subsequently, the eigenvalues for the linear stability analysis at the equilibrium points can be obtained by solving

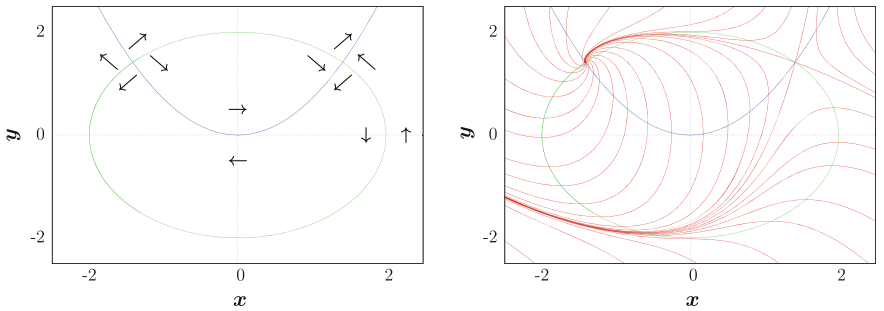


Fig. 1.4 Nullclines for system (1.40): (*Left*) directional information determined from the x- and y-nullclines and its application to local structure near the two equilibrium points, (*Right*) computed trajectories for the system

$$\begin{vmatrix} -\sqrt{2}x_* - \lambda & 1 \\ 2x_* & 2y_* - \lambda \end{vmatrix} = 0.$$

At (x_*^+, y_*) we get $\lambda_1 \approx 3.356$ and $\lambda_2 \approx -2.528$ yielding a saddle point. At (x_*^-, y_*) we get an unstable spiral with $\lambda_{1,2} \approx 2.414 \pm 1.630i$. The nature of these equilibria could be inferred from information provided by the nullclines without these slightly cumbersome algebraic calculations, see Fig. 1.4. The usefulness of nullclines becomes more significant in reducing more complicated systems and constructing proofs about properties of families of solutions.

1.6 Further Directions

The analysis of the linear stability of equilibrium points can be extended to dynamical systems in n-dimensions [70]. However, the phase plane is special because some geometric arguments do not extend in a simple way to curves in space $\mathbb{R}^n$ for $n \geq 3$. Consequently it is very helpful when higher-order systems can be reduced down to phase planes.

The material in this chapter gives only a very brief review of selected fundamental results. For more detailed background on phase planes and further coverage of bifurcations and dynamical systems theory, see, for example [43, 45, 54, 70, 94].

Comprehensive presentations of chemical reaction systems are found in most chemistry textbooks, for example [6]. More concise introductions are given in the applied mathematics books by Holmes [49] and Keener and Sneyd [57].

The models described here by systems of ODE are the simplest type of rate equations, sometimes also called state-space models, with the rates of evolution dependent on the current solution state. In *delay differential equations*, the current rate depends on the solution at an earlier time, $dx/dt = f(x(t-1))$ [35]. *Discrete-time maps*, also called *difference equations* are analogous algebraic equations describing evolv-

ing solutions at discrete times, say year-to-year, as in $x_{n+1} = f(x_n)$ [5]. *Stochastic differential equations* have randomly varying contributions to the rate functions and are sometimes used to represent the variations among the dynamics of individuals in a large population, improving on (1.2) which describes a uniform average behaviour [36]. More detailed models of populations, called *structured population models* divide up the population by age or size; if these are treated as continuous variables, then the rate equations yield partial differential equations. We will touch on some of these topics in later chapters. Haberman's book [45] also provides very accessible introductions to several of these types of models.

1.7 Exercises

1.1 Consider the problem of tracking the vertical position, $z(t)$, of a rocket whose mass changes as it consumes its fuel. If the rocket starts from rest at $z(0) = 0$ with initial mass $m(0) = m_0$ and obeys

$$\frac{d}{dt}\left(m(t)\frac{dz}{dt}\right) = -mg + \tau m, \qquad \frac{dm}{dt} = -m,$$

solve the ODEs to determine $z(t)$ and determine the condition on τ that is necessary for lift-off.

1.2 Use basic solution methods for first order ODEs to solve the elementary reactions (1.8–1.13) for $A(t)$, $B(t)$, $C(t)$ starting from initial conditions A_0, B_0, C_0 respectively.

1.3 Write the four rate equations for chemicals C, E, P, S governed by the reactions[4]

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

1.4 Write the rate equations for $X(t)$, $Y(t)$, $Z(t)$ describing the reaction system

$$(\text{source}) \xrightarrow{A} X \qquad Y \underset{C}{\overset{B}{\rightleftharpoons}} X \qquad Z \underset{E}{\overset{D}{\rightleftharpoons}} Y \qquad 3Y \xrightarrow{G} Z \qquad Z \xrightarrow{H} (\text{waste}).$$

1.5 Consider the dynamics of $x(t)$ satisfying the first order ODE,

$$\frac{dx}{dt} = x - x^3 - k \qquad x(0) = x_0,$$

for different values of the parameter k (see Fig. 1.2).

[4] We delay solving more complicated systems of reactions to Chap. 10. Here we only want to set up the rate equations.

(a) For $k = 0$, obtain the exact solutions starting from $x_0 = 1/2$ and $x_0 = -2$. Show that these solutions match the linearised results from (1.26) at the equilibrium points $x_* = \{-1, 0, 1\}$ for the long-time behaviour ($t \to \infty$), and the "pre-history" of the solution ($t \to -\infty$). Show that for all initial conditions with $x_0 \neq 0$, one of two final states are approached as $t \to \infty$.

(b) Determine the two values of k for which the ODE has only two distinct equilibrium points. Sometimes called *critical values* or *bifurcation points*, such parameter values determine special cases, where different analysis is needed to describe the system; in this case, this will involve second-order degenerate equilibrium points. Determine the coefficient b occurring in the equation for the local behaviour of solutions near the degenerate equilibrium points, $du/dt = bu^2$ with $u(t) = x(t) - x_*$.
Show that for this problem, these bifurcation values separate ranges of k where there are *global attractors* (unique final states approached by all initial conditions for $t \to \infty$) from cases where two stable equilibrium states co-exist (called *bi-stability*).

1.6 Consider the equations for local behaviour at second- and third-order degenerate equilibrium points, (1.27) and (1.28).

(a) Use separation of variables to solve the ODE analytically and subsequently describe the stability of the equilibrium point at $u = 0$ for both choices of the sign of b.

(b) Using only the sign of the velocity on the phase line consider solutions starting from initial conditions with $u_0 \gtrless 0$ to obtain the stability of the equilibrium point without the need for the ODE solutions.

(c) Repeat (a, b) for the dependence of (1.28) on c.

1.7 Consider the second order ODE for $x(t)$,

$$\frac{d^2x}{dt^2} = x - \tfrac{1}{2}x^2.$$

(a) Let $y = dx/dt$ and write the ODE as a phase plane system. Determine the linear stability properties of the two equilibrium points.

(b) Show that the solutions satisfy the equation

$$\tfrac{1}{2}(x')^2 - \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3 = H,$$

where H is a constant of integration, sometimes called the *Hamiltonian*.

(c) For a range of values, $0 < x < M$, this problem has a continuous set of periodic solutions. Show that the maximum and minimum values of $x(t)$ of each periodic solution satisfy a polynomial equation involving H. Define the amplitude of the oscillations as $A = x_{\max} - x_{\min}$. Show that $A = 0$ corresponds to an equilibrium point. Show that the largest amplitude solution has $x_{\min} = 0$; determine its value for $x_{\max}(=M)$. What is the range of values for H?

(d) Use $t = \int dt = \int \frac{dx}{dx/dt}$ to show that the period of oscillation for these solutions is given by

$$P = 2 \int_{x_{\min}}^{x_{\max}} \frac{dx}{\sqrt{x^2 - \frac{1}{3}x^3 + 2H}}.$$

(e) Show that the simpler *piecewise-linear model*

$$\frac{d^2x}{dt^2} = f(x) \qquad \text{with} \quad f(x) = \begin{cases} x & x < 1 \\ 2 - x & x \geq 1 \end{cases}$$

has the same equilibrium points and the same linear stability properties at the equilibria.
Solve the two linear problems

$$\frac{d^2x_A}{dt^2} = x_A \qquad x_A(0) = x_{\min} \qquad x_A'(0) = 0$$

$$\frac{d^2x_B}{dt^2} = 2 - x_B \qquad x_B(P/2) = x_{\max} \qquad x_B'(P/2) = 0$$

and construct a periodic solution of the piecewise-linear model by enforcing smoothness, $x_A' = x_B'$ at $x_A = x_B = 1$.

Chapter 2
Transport Equations

Many systems exhibit evolution over time with properties of interest that vary throughout their spatial domains. Examples of such systems arise in *population dynamics*, which describes the distribution of individuals in some population and how they interact. "Individuals" could refer to molecules, electrons, particles, animals, people, company stocks, or network messages.

One way to study the overall population is to attempt to track each individual, such approaches are sometimes called *individual-based models*. The tracking process is usually very labour intensive and involves collecting a lot of data on the actions of all individuals. If this level of detail is not crucial and a more 'large-scale' view is of interest, then *continuum theory* may be a better option.

Continuum theories yield evolution equations with respect to properties that are averaged over small intervals of time and small regions of space. In such cases, we implicitly assume a *continuum hypothesis* which states that appropriately averaged behaviours of individuals can be generally predicted from trends in the local population. We consequently formulate *continuum models* as partial differential equations (PDE) governing the evolution of density functions ($f(\mathbf{x}, t)$) describing properties of the population at a given position and time. Integrating the density over the entire domain can be used to capture the time-dependence of the property on the whole population, $\mathscr{F}(t) = \int f(\mathbf{x}, t)\, d\mathbf{x}$.

Some specific applications of continuum models include:

- Fluid dynamics: the flow of liquids and gases (density of molecules in space)
- Solid mechanics: the deformation of solids (density of molecules in space)
- Electromagnetics: the flow of electric currents in materials (density of electric charges in space)
- Scattering theory: dynamics due to collisions in high energy particle physics (density of particles having different velocities)

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9_2

- Age-structured population dynamics: birth, death, and aging of a population (distribution of individuals having different ages) [28, 74]
- Size-structured population dynamics: growth and decay of physical sizes of individuals in a population (distribution of individuals having different sizes) [8]

Fluid dynamics and solid mechanics are often grouped together under the heading of *continuum mechanics*. Other population models focus not on change in position, but on properties like stock price in financial models, popularity in social networks or genetic traits in biological systems.

In each of these contexts, PDEs can be used to describe the redistribution of the property of interest over time. Conveying shifts or "motion" in the density, whether with respect to spatial position or with the independent variable $\mathbf{x}$ representing other properties (e.g. velocity, age, size), such PDEs are also broadly called *transport equations*. The common structure shared by these models is having a PDE for the rate of change of the density involving the spatial gradient of a *flux function*, which characterises transport of the property within the population. Transport equations are fundamental for describing problems in many fields extending from theoretical physics, chemical engineering, and mathematical biology to probability theory.

In this chapter we introduce the fundamental approach for formulating transport models (conservation laws and the Reynolds transport theorem). We then go on to describe the method of characteristics, a methodology for constructing exact solutions to basic transport models.

2.1 The Reynolds Transport Theorem

As a conceptual starting point, we consider how we might describe the dynamics of individuals in a large population—describing their motion (change of absolute position as a function of time, $\mathbf{X}(t)$) and deformation (rearrangement or change of relative position within groups of surrounding individuals).

In order to introduce the basic principles, we begin by studying the case of *passive transport* of inert particles carried by an externally imposed flow field, for example, particles of dust cloud in the wind or a pollutant carried in a running river. Here "passive" indicates that the presence of the particles does not influence the flow driving their motion.

Each particle can be uniquely identified by its initial position at time $t = 0$,

$$\mathbf{X}(t = 0) = \mathbf{A} = (A, B, C) \qquad (\, X(t = 0) = A \text{ in 1D} \,) \tag{2.1a}$$

Assume an imposed velocity field, $\mathbf{v}(\mathbf{x}, t)$, is given that specifies the speed and direction that a particle occupying position $\mathbf{x}$ would take at time t. Having an explicit expression for properties in a fixed coordinate system is called an *Eulerian description*. From the definition of velocity as the rate of change of position, the motion of a particle is given by

$$\frac{d\mathbf{X}}{dt} = \mathbf{v}(\mathbf{X}, t). \qquad \left(\frac{dX}{dt} = v(X, t) \text{ in 1D} \right) \tag{2.1b}$$

This initial value problem gives the motion of a particle over time, and is sometimes called the problem for the particle's *pathline*.

An alternative point of view is to recognise that for a particle starting from position (2.1a), the Eulerian description of the entire velocity field is not needed. All that is essential is the velocity along the particle's pathline. The *Lagrangian description* gives a property following the motion of a given particle as a function of time. The Lagrangian velocity for the particle starting from position $\mathbf{A}$ is

$$\mathbf{V}(t; \mathbf{A}) = \mathbf{v}(\mathbf{X}(t; \mathbf{A}), t). \tag{2.2}$$

The problem for the pathline can then be restated in Lagrangian form as

$$\frac{d\mathbf{X}}{dt} = \mathbf{V}(t; \mathbf{A}), \qquad \mathbf{X}(0; \mathbf{A}) = \mathbf{A}. \tag{2.3}$$

We will show that the ability to change between these two equivalent forms will enable us to solve transport equations.

Values of other properties carried by point particles (e.g. density, temperature, radioactivity) can similarly be expressed in both Eulerian $f(\mathbf{x}, t)$ and Lagrangian $F(t)$ forms via

$$F(t; \mathbf{A}) = f(\mathbf{X}(t; \mathbf{A}), t). \tag{2.4}$$

In describing the evolution of any property f following a particular particle, it is necessary to calculate the rate of change of f for the given particle $f_{\text{part}}(t) \equiv F(t; \mathbf{A}) = f(\mathbf{X}(t; \mathbf{A}), t)$. Using the chain rule, we can express this "Lagrangian time derivative" in terms of Eulerian functions,

$$\begin{aligned} \frac{df_{\text{part}}}{dt} &= \frac{\partial f}{\partial t} + \nabla f \cdot \frac{d\mathbf{X}}{dt} \\ &= \boxed{\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f \equiv \frac{Df}{Dt}} \end{aligned} \tag{2.5}$$

where we have used (2.1b). The Eulerian form of this derivative is called the *convective (or material) derivative* and is denoted $\frac{Df}{Dt}$. We will see that the velocity field defining the flow has a special role in transport equations. If $\mathbf{v}$ is given, then the problem for the evolution of the property of interest is called a *kinematics problem* (as in passive transport). If the evolution of $\mathbf{v}$ is coupled to f and must be determined as part of the solution, then it is a more challenging *dynamics problem*.

The next stage in formulating a continuum model is to determine the rate of change of a property evaluated over a *"material blob"*—in other words, we consider a specific set of particles, defined as starting from a set of $\mathbf{A}$ values (2.1a) occupying a region D in space. For example, picture the blob as the fluid in a small droplet. The

cumulative value of property f over the material blob is given by

$$f_{\text{blob}}(t) = \iiint_{D(t)} f(\mathbf{x}, t)\, dV, \tag{2.6}$$

where $D(t)$ is the region occupied by the moving, deforming blob at time t. We would like to know how $f_{\text{blob}}(t)$ varies with time. In one dimension, a "blob" is simply a time-dependent interval, $a(t) \le x \le b(t)$ and (2.6) reduces to

$$f_{\text{blob}}(t) = \int_{a(t)}^{b(t)} f(\mathbf{x}, t)\, dx. \tag{2.7}$$

To find the rate of change, we apply Leibniz's rule to determine the derivative of an integral with time-dependent endpoints,

$$\begin{aligned}\frac{d}{dt}\left(\int_{a(t)}^{b(t)} f(x,t)\,dx\right) &= \int_a^b \frac{\partial f}{\partial t}\,dx + \underbrace{f(b,t)\frac{db}{dt} - f(a,t)\frac{da}{dt}}_{f(x,t)\frac{dx}{dt}\Big|_{x=a}^{x=b}} \\ &= \int_a^b \left[\frac{\partial f}{\partial t} + \frac{\partial}{\partial x}\left(f(x,t)\frac{dx}{dt}\right)\right] dx \\ &= \int_a^b \left[\frac{\partial f}{\partial t} + \frac{\partial}{\partial x}(fv)\right] dx \end{aligned} \tag{2.8}$$

where again, we have made use of (the one-dimensional version) of (2.1b). In two and three dimensions, making use of the divergence theorem, we can generalise this result to give the *Reynolds Transport Theorem* [1],

$$\boxed{\frac{d}{dt}\left(\iiint_{D(t)} f\, dV\right) = \iiint_{D(t)} \left[\frac{\partial f}{\partial t} + \nabla\cdot(f\mathbf{v})\right] dV} \tag{2.9}$$

2.2 Deriving Conservation Laws

In order to apply the Reynolds transport theorem to obtain a transport model, we need the introduction of a *conservation principle*. This is a statement providing information about the rate of change of f_{blob} that applies *to all possible material blobs*.

For example, requiring *conservation of mass*—the principle that total mass can be neither created nor destroyed, the mass of every blob must remain constant in time.

Expressing the mass in terms of the material density, $f \equiv \rho$, in one-dimension this gives

$$m_{\text{blob}}(t) = \int_{a(t)}^{b(t)} \rho(x,t)\,dx, \qquad \frac{dm_{\text{blob}}}{dt} = 0 \quad \forall (a,b). \tag{2.10}$$

In order to change from a statement about properties of blobs to a transport PDE, the final tool involved is an integral result from analysis, sometimes called the *du Bois-Reymond lemma* [49], which we state in its simplest form on the one-dimensional domain $0 \le x \le 1$ as

$$\text{If } \int_a^b g(x)\,dx = 0 \quad 0 \le \boxed{\forall (a,b)} \le 1, \quad \text{then } \boxed{g(x) \equiv 0 \text{ for } 0 \le x \le 1}. \tag{2.11}$$

In other words, if an integral vanishes for all choices of sub-domains, then the integrand must vanish on the whole domain.[1] This result allows us to convert from an integral equation (called a *weak form*, applying on the domain as a whole) to a differential equation (called a *strong form* that applies locally, pointwise at each x in the whole domain), if the integral is valid for all blobs.

Applying the Reynolds Transport theorem to (2.10), we reduce the principle of conservation of mass to a PDE for the density, yielding the (local) *conservation law* for mass density in one dimension,

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho v) = 0, \tag{2.12}$$

also known as the *continuity equation*. In three dimensions, the corresponding result is

$$\frac{d}{dt}\left(\iiint_{D(t)} \rho(\mathbf{x},t)\,dV\right) = 0 \quad \forall D \qquad \Longrightarrow \qquad \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \tag{2.13}$$

The quantity in the parenthesis is called the *flux*, $\mathbf{q}$, and corresponds physically to the rate of ρ passing through a fixed point per unit time (here $\mathbf{q} = \rho \mathbf{v}$).

If we are given information on the rate at which the property of interest is created or destroyed, say due to a chemical reaction as in $d\rho/dt = R$ like (1.8), then the rate of change of the total amount of the chemical in domain D be can expressed by

$$\frac{d}{dt}\left(\iiint_D \rho\,dV\right) = \iiint_D R\,dV. \tag{2.14}$$

This should be true in any subdomain (any material blob). If transport is present then $D = D(t)$ and after applying the Reynolds Transport Theorem to the left integral, we can regroup the resulting integrals together as

[1] Assuming the integrand to be a smooth function.

$$\iiint_{D(t)} \left[\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) - R \right] dV = 0 \qquad \forall D(t). \tag{2.15}$$

Then applying the du Bois-Reymond lemma yields the general conservation law including reaction terms (sometimes called *source* or *sink* terms depending on whether the rate of production is positive or negative),

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = R. \tag{2.16}$$

Equation (2.14) can be applied to Newton's second law, stating that the rate of change of momentum is equal to the sum of the applied forces

$$\frac{d}{dt} \left(\iiint_{D(t)} \rho \mathbf{v} \, dV \right) = \iiint_{D(t)} \mathbf{f} \, dV, \tag{2.17}$$

where $\mathbf{f}$ gives net forces per unit volume. Applying the Reynolds transport theorem, the du Bois Reymond lemma, and expressing the forces in terms of the divergence of a stress tensor, $\mathbf{f} = \nabla \cdot \sigma$, yields the *Cauchy momentum equation*,

$$\frac{\partial (\rho \mathbf{v})}{\partial t} + \nabla \cdot (\rho \mathbf{v}\mathbf{v}) = \nabla \cdot \sigma. \tag{2.18}$$

Encapsulating the principles of conservation of mass and momentum, equations (2.13) and (2.18) are the basis of continuum mechanics [41, 50]. Supplemented by appropriate equations for defining the stress in terms of ρ and $\mathbf{v}$, called *constitutive relations* or *equations of state*, these yield the governing equations for solid mechanics and the Navier–Stokes equations for fluid dynamics.

2.3 The Linear Advection Equation

In this section we will focus on the most fundamental case of (2.16), where $R = 0$ (no reactions) and the velocity is a uniform constant vector, $\mathbf{v} = c\hat{\mathbf{i}}$ (constant speed), yielding the one-dimensional *advection equation*

$$\frac{\partial \rho}{\partial t} + c \frac{\partial \rho}{\partial x} = 0. \tag{2.19}$$

All solutions of the advection equation can be written as constant profile *travelling waves*,

$$\rho(x, t) = P(x - ct), \tag{2.20}$$

where the terminology "travelling wave" refers to a solution moving with a fixed velocity (speed in one dimension) and with a "constant profile" meaning that the

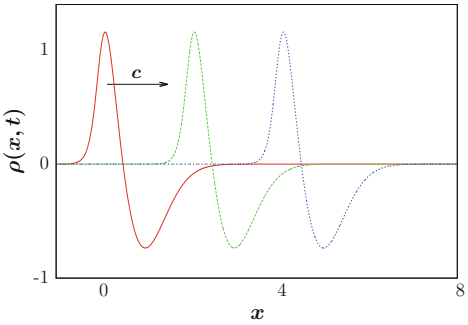


Fig. 2.1 A travelling wave solution (2.20), $p(x,t) = P(x-ct)$, of (2.19) for $c > 0$ at three successive times

wave shape $P(x)$ given at $t = 0$ is maintained for all times. If $c > 0$, the wave profile propagates to the right (and to the left if $c < 0$) (see Fig. 2.1).

While (2.20) is a solution of equation (2.19) for any function P, for the purpose of considering the complexities associated with more general transport models (with other forms of flux), it is helpful to first understand the action of the equation on simple functions. If we can express $P(x)$ as a complex Fourier series (see Appendix A), $P(x) = \sum_k A_k e^{ikx}$, then the travelling wave (2.20) can be written as

$$\rho(x,t) = \sum_{k=-\infty}^{\infty} A_k e^{ik(x-ct)} = \sum_{k=-\infty}^{\infty} A_k e^{i[kx-\omega t]}, \tag{2.21}$$

where k is called the wave number (related to the wavelength between successive crests of $\cos(kx)$ in space, $L = 2\pi/k$) and ω is called the angular frequency (related to the period of oscillation of $\sin(\omega t)$ in time, $T = 2\pi/\omega$). For (2.19), the expansion (2.21) defines $\omega = ck$, but the Fourier decomposition is applicable to other equations as long as a relationship can be found between the frequency and the wavenumber, $\omega = \omega(k)$, called the *dispersion relation*.

Consider the fifth order linear homogeneous constant-coefficient PDE

$$\frac{\partial \rho}{\partial t} + c_1 \frac{\partial \rho}{\partial x} + c_3 \frac{\partial^3 \rho}{\partial x^3} + c_5 \frac{\partial^5 \rho}{\partial x^5} = 0. \tag{2.22}$$

If we substitute in a Fourier mode $\rho_k(x,t) = A_k \exp(i[kx - \omega t])$ (called a uniform *plane wave*) as a trial solution, (2.22) reduces to a relationship between ω and k given by

$$-i\omega + c_1(ik) + c_3(ik)^3 + c_5(ik)^5 = 0,$$

which determines the dispersion relation, $\omega(k) = c_1 k - c_3 k^3 + c_5 k^5$.

Descriptions of wave properties in terms of the dispersion relation are fundamental in many models that arise as generalisations of the advection equation. The *phase speed*, $c_p(k) \equiv \omega(k)/k$, gives the speed of waves with wavenumber k. Equation (2.19) has constant phase speed (being independent of k), and so is called *dispersionless*

because all of its wave modes travel at the same speed, maintaining the steady profile $P(x - ct)$ (2.21). In contrast, for *dispersive* equations with $c_p(k) \neq$ constant, the profile given by the sum $\rho(x,t) = \sum_k A_k e^{ik[x - c_p(k)t]}$ changes with time as the component waves separate ("disperse") according to their different speeds. The growth or decay of waves is described by dispersion relations having imaginary components: wave amplitudes are maintained if $\omega(k)$ is purely real, while if $\omega(k)$ has an imaginary part then *dissipation* occurs. In order to observe this effect, consider what happens when we substitute the plane wave solution into the diffusion equation $\rho_t = \rho_{xx}$; the dispersion relation is found to be $\omega = -ik^2$, so that the (real-valued) phase speed is $c_p = 0$ and the wave will dissipate without propagating: $\rho_k(x,t) = (A_k e^{-k^2 t})e^{ikx}$ (these correspond to so-called 'standing waves').

We now turn our attention to describing methods for obtaining the solutions to various generalisations of the advection equation (2.19).

2.4 Systems of Linear Advection Equations

Consider a system of two coupled linear advection equations for properties $p(x,t)$, $q(x,t)$ on $-\infty < x < \infty$,

$$p_t + ap_x + bq_x = 0 \tag{2.23a}$$

$$q_t + cp_x + dq_x = 0 \tag{2.23b}$$

where a, b, c, d are constants, and the initial conditions are given by

$$p(x,0) = f(x) \qquad q(x,0) = g(x). \tag{2.23c}$$

The typical approach for determining the solutions of (2.23) is to decouple the system into independent advection equations, each with a travelling wave solution of the form

$$\frac{\partial w}{\partial t} + \lambda \frac{\partial w}{\partial x} = 0 \qquad \Rightarrow \qquad w(x,t) = W(x - \lambda t), \tag{2.24}$$

where λ is the wave speed for some appropriate travelling wave profile W.

We start by assuming a solution as a linear combination,

$$w(x,t) = Ap(x,t) + Bq(x,t), \tag{2.25}$$

for some constants A, B to be determined. Forming the linear combination of $A\cdot$ (2.23a) $+ B\cdot$ (2.23b) yields

$$(Ap + Bq)_t + (aA + cB)p_x + (bA + dB)q_x = 0,$$

while substituting w from (2.25) into (2.24) gives

$$(Ap + Bq)_t + (\lambda A)p_x + (\lambda B)q_x = 0.$$

Comparing coefficients in these equations leads to a matrix eigenvalue problem,

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \lambda \begin{pmatrix} A \\ B \end{pmatrix}. \tag{2.26}$$

Solving this system yields the eigenvalues (λ) defining the wave speeds and the components of the eigenvectors defining the relationship of w to the solutions p and q, (2.25). In order for the eigenvalues to be meaningful as speeds, they must be real values. The solutions of the eigenvalue problem, λ_1 and A_1, B_1, and λ_2 and A_2, B_2, determine the travelling wave solutions

$$\begin{aligned} w_1(x,t) &= A_1 p(x,t) + B_1 q(x,t) = W_1(x - \lambda_1 t), \\ w_2(x,t) &= A_2 p(x,t) + B_2 q(x,t) = W_2(x - \lambda_2 t), \end{aligned} \tag{2.27}$$

where the travelling wave profiles $W_1(x)$, $W_2(x)$ are otherwise undetermined so far. Applying initial conditions (2.23c) to evaluate these equations at time $t = 0$ explicitly defines the wave profiles as satisfying

$$\begin{aligned} A_1 f(x) + B_1 g(x) &= W_1(x), \\ A_2 f(x) + B_2 g(x) &= W_2(x). \end{aligned} \tag{2.28}$$

With W_k now known in terms of f, g and the A_k, B_k coefficients given by the eigenvectors, the right-hand side of (2.27) is known, yielding a linear system for $p(x,t)$, $q(x,t)$ in terms of combinations of of $f(x - \lambda_k t)$ and $g(x - \lambda_k t)$ (with $k = 1, 2$).

This approach extends to systems of any number of coupled linear equations,

$$\frac{\partial \mathbf{p}}{\partial t} + \mathbf{M}\frac{\partial \mathbf{p}}{\partial x} = \mathbf{0} \tag{2.29}$$

where $\mathbf{p} = (p, q, r, \ldots)^T \in \mathbb{R}^n$ and $\mathbf{M}$ is an $n \times n$ constant coefficient matrix. Assuming trial solutions of the form $w = \mathbf{u} \cdot \mathbf{p}$ where $\mathbf{u} = (A, B, C, \ldots)^T$ and noting the relationship between (2.26) and the coefficients in (2.23), we see that the general eigenvalue problem can be stated as

$$\mathbf{M}^T \mathbf{u} = \lambda \mathbf{u}. \tag{2.30}$$

If all the eigenvalues of $\mathbf{M}^T$ are real and there is a complete set of eigenvectors, then (2.29) is called a *hyperbolic system* and its solutions can be expressed in terms of sums of travelling waves (see [65, 80, 81, 106] for more details).

2.5 The Method of Characteristics

Semilinear wave equations take the form

$$\frac{\partial p}{\partial t} + c(x,t)\frac{\partial p}{\partial x} = r(x,t,p), \tag{2.31a}$$

where the speed c depends on both position and time, but not on the solution itself, and the reaction rate can depend on all three, but not on derivatives of the solution. The terminology 'semilinear' arises because the left-hand side of the equation is linear with respect to the dependent variable p, while the right-hand side can be nonlinear in p. Problems for this class of equations with initial conditions

$$p(x,0) = f(x), \tag{2.31b}$$

can be solved by converting from the original Eulerian PDE form into a Lagrangian form that can be solved as a set of coupled ODEs.

In order to see how (2.31a) defines a property evolving in a flow, let the Lagrangian form of the solution for each initial material coordinate A be written as

$$P(t;A) = p(X(t;A),t) \qquad X(0;A) = A. \tag{2.32}$$

Using the chain rule, the rate of change of P is

$$\frac{dP}{dt} = \frac{\partial p}{\partial t} + \frac{\partial p}{\partial x}\frac{dX}{dt} \tag{2.33}$$

and this can be matched term by term with the PDE (2.31a) if the Lagrangian variables evolve according to the *characteristic equations*

$$\frac{dP}{dt} = r(X(t),t,P(t)), \qquad \frac{dX}{dt} = c(X(t),t). \tag{2.34a}$$

The initial conditions (2.31b) take the corresponding form

$$P(0;A) = f(A), \qquad X(0;A) = A. \tag{2.34b}$$

The solution of the X-equation subject to its initial condition therefore defines a path in the (x,t) plane known as a *characteristic curve* (or simply a 'characteristic'), and generalises the concept of a pathline. Once $X(t)$ has been determined, it can be substituted into the evolution equation for P $(2.34a)_1$ and can then be solved for $P(t)$.

Thus, the so-called *method of characteristics* replaces the Eulerian PDE with an initial value problem for a pair of ODEs. The Lagrangian solutions $(X(t;A), P(t;A))$ are parameterised by the initial material coordinate A and can sometimes be inverted to give the explicit Eulerian solution $p(x,t)$.

As an example, consider the transport problem for $p(x, t)$,

$$\frac{\partial p}{\partial t} + 2x\frac{\partial p}{\partial x} = xp^2, \qquad p(x, 0) = 2 + \sin(x). \tag{2.35}$$

Following the above analysis, the corresponding characteristic problem is given by

$$\frac{dX}{dt} = 2X, \qquad X(0) = A, \tag{2.36a}$$

$$\frac{dP}{dt} = XP^2, \qquad P(0) = 2 + \sin(A). \tag{2.36b}$$

Solving (2.36a) yields $X(t; A) = Ae^{2t}$, which we can then substitute into the ODE for P (2.36b) to obtain the general solution

$$\frac{dP}{dt} = Ae^{2t}P^2 \qquad \Rightarrow \qquad P(t) = \frac{1}{C - \frac{1}{2}Ae^{2t}}.$$

Imposing the initial condition on P at $t = 0$ (when $X = A$) then yields that

$$P(t; A) = \left(\frac{A}{2}\left(1 - e^{2t}\right) + \frac{1}{2 + \sin(A)}\right)^{-1},$$

(see Fig. 2.2). Inverting $Ae^{2t} = X = x$ gives $A = xe^{-2t}$. Substituting this into $P(t; A)$ and applying $P = p$ results in the final form of the solution

$$p(x, t) = \left(\frac{x}{2}\left(e^{-2t} - 1\right) + \frac{1}{2 + \sin(xe^{-2t})}\right)^{-1}.$$

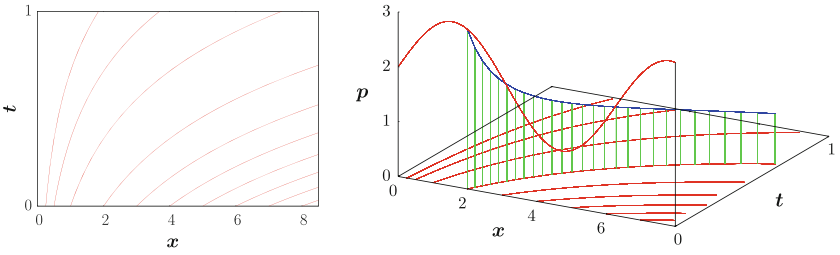


Fig. 2.2 (*Left*) Characteristic curves in the (x, t) plane for problem (2.35). (*Right*) Function values evolving on a characteristic curve

2.6 Shocks in Quasilinear Equations

Quasilinear equations differ from semilinear equations only in having the speed c depend additionally on the solution,

$$\frac{\partial p}{\partial t} + c(x,t,p)\frac{\partial p}{\partial x} = r(x,t,p). \tag{2.37}$$

This makes the characteristic ODEs fully coupled,

$$\frac{dP}{dt} = r(X,t,P), \qquad \frac{dX}{dt} = c(X,t,P). \tag{2.38}$$

For semilinear equations, the equation (2.34a) for the characteristic curves $x = X(t;A)$ decouples from P, and standard existence and uniqueness results guarantee that two curves starting from different initial positions (and having different values of $P(t)$) will never intersect. This is not the case for the coupled equations in (2.38), where characteristics can cross and hence predict several different $P(t;A_j)$ values occurring simultaneously at the same x position.

As an example, consider the *inviscid Burgers equation*,

$$\frac{\partial p}{\partial t} + p\frac{\partial p}{\partial x} = 0, \tag{2.39}$$

with initial conditions

$$p(x,0) = \begin{cases} 1-|x| & |x| \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.40}$$

Equation (2.39) is an important mathematical model that will arise again later in other contexts; one aspect of its importance can be observed by noting that if the property p is the flow velocity, then (2.39) is the convective derivative of the velocity (2.5) (the Lagrangian form of the acceleration), $\frac{Dv}{Dt} = v_t + vv_x = 0$.

The characteristic equations for this example are

$$\frac{dX}{dt} = P, \qquad X(0;A) = A, \tag{2.41a}$$

$$\frac{dP}{dt} = 0, \qquad P(0;A) = \begin{cases} 1-|A| & |A| \le 1, \\ 0 & \text{else.} \end{cases} \tag{2.41b}$$

We note that for this problem P remains constant along each characteristic, so that $X = Pt + A$ and the solution can be expressed as

$$\begin{cases} X = A, & P = 0 & A < -1, \\ X = (1+A)t + A, & P = 1 + A & -1 \le A \le 0, \\ X = (1-A)t + A, & P = 1 - A & 0 \le A \le 1, \\ X = A, & P = 0 & A > 1 \end{cases} \tag{2.42}$$

and restated in an explicit form for $t \ge 0$

$$p(x,t) = \begin{cases} 0 & x < -1, \\ (1+x)/(1+t) & -1 \le x \le t, \\ (1-x)/(1-t) & t \le x \le 1, \\ 0 & x > 1. \end{cases} \tag{2.43}$$

There are several points to note about this representation of the solution—first, the piecewise-defined solutions are not mutually exclusive for $t > 1$. Consequently multiple values are being predicted for p at some locations when $t > 1$. In relation to this, the third subcase becomes undefined at time $t = 1$, changing from negative slopes for $t < 1$ to positive slopes for $t > 1$. Figure 2.3 shows the characteristic curves and $p(x, t)$ profiles given by (2.43). We observe that the portion of the solution starting from $x \in [-1, 0]$ spreads out over an increasingly large region as its characteristics separate from each other (this is sometimes called an *expansion fan* or *rarefaction wave*).

In contrast, the portion of the solution starting from $x \in [0, 1]$ is being compressed into a smaller region (up until $t = 1$) and is referred to as a *compressive wave*. Its slope steepens and overturns for $t > 1$ to yield what could be described as a multi-valued *"breaking wave"*.

Mathematically this part of the solution is predicting three values for a physical quantity (maybe a density, concentration or temperature, for example) that should have a unique value at any point x at a given time t.

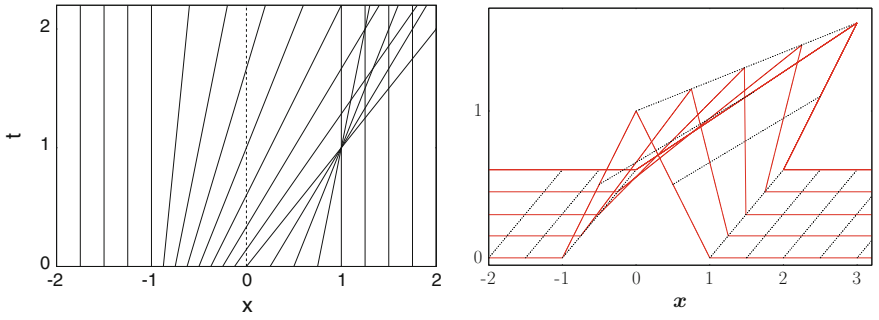


Fig. 2.3 (*Left*) characteristic curves $X(t; A)$ given by (2.42) in the xt plane and a 3D view of the evolving multi-valued solution $p(x, t)$ (2.43) (*Right*)

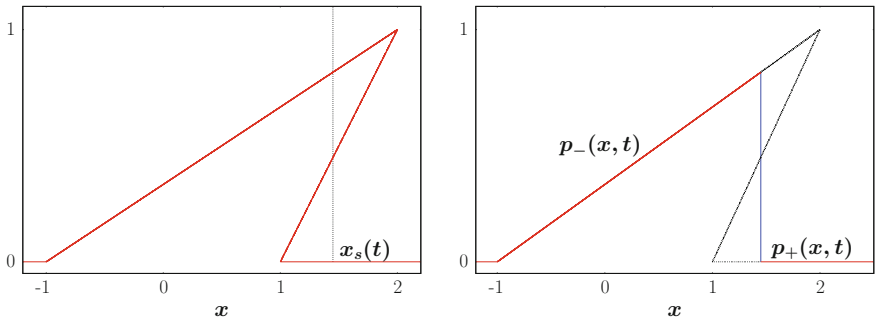


Fig. 2.4 (*Left*) The multivalued solution (2.43) and insertion of a shock at $x = x_s(t)$, (*Right*) reduction to a single-valued solution given by (2.46)

There is a systematic and rigorous approach for correcting this unphysical behaviour and modifying the solution to make it single-valued through the use of *shocks*. Shocks are moving jump discontinuities in the solution that separate one piecewise-defined portion of the solution ahead of the shock, $p_+(x, t)$, from another part behind the shock, $p_-(x, t)$, eliminating the multi-valued behaviour (see Fig. 2.4). Shock waves are common in many physical systems, including acoustics (sonic booms), fluid flow (hydraulic jumps), and traffic flow (moving traffic jams).

The construction of shock-corrected solutions builds on the idea that if a shock is inserted at one position, $x = x_s(t)$, it can appropriately separate the overlapping sets of characteristic curves to produce a well-defined single-valued solution everywhere away from the shock. The constructed solution should satisfy all of the properties expected for the conservation law and we use this to derive the equation for the motion of the shock position.

Consider a general quasilinear equation describing the transport of a property $p(x, t)$ according to a flux function $q = q(p)$,

$$\frac{\partial p}{\partial t} + \frac{\partial q(p)}{\partial x} = 0, \tag{2.44}$$

which has been derived for smooth solutions via the Reynolds transport theorem from a conservation law for p. The integrated form of (2.44) on a fixed domain $a \le x \le b$ includes contributions from fluxes at the ends of the domain,

$$\frac{d}{dt}\left(\int_a^b p\, dx\right) + q(p)\Big|_{x=a}^{x=b} = 0. \tag{2.45}$$

If a shock were inserted at some position $x = x_s(t)$, the piecewise-defined form of the solution becomes

$$p(x,t) = \begin{cases} p_-(x,t) & a \le x < x_s(t), \\ p_+(x,t) & x_s(t) < x \le b. \end{cases} \tag{2.46}$$

Separating (2.45) with respect to dependence on the solution to the left or right of the shock then yields

$$\left[\frac{d}{dt}\left(\int_a^{x_s} p_-\,dx\right) - q(p_-(a,t))\right] + \left[\frac{d}{dt}\left(\int_{x_s}^{b} p_+\,dx\right) + q(p_+(b,t))\right] = 0. \tag{2.47}$$

Applying Leibniz's rule (2.8) then gives

$$\left(\int_a^{x_s} \frac{\partial p_-}{\partial t}\,dx + p_-(x_s,t)\frac{dx_s}{dt}\right) + \left(\int_{x_s}^{b} \frac{\partial p_+}{\partial t}\,dx - p_+(x_s,t)\frac{dx_s}{dt}\right) + [q(p_+(b,t)) - q(p_-(a,t))] = 0.$$

By adding and subtracting $q(p_\pm)(x_s)$ and re-grouping terms, we obtain

$$\left(\int_a^{x_s} \partial_t p_-\,dx + q(p_-(x_s,t)) - q(p_-(a,t))\right) + \left(\int_{x_s}^{b} \partial_t p_+\,dx + q(p_+(b,t)) - q(p_+(x_s,t))\right) - q(p_-(x_s,t)) + q(p_+(x_s,t)) + [p_-(x_s,t) - p_+(x_s,t)]\frac{dx_s}{dt} = 0.$$

The terms in parentheses on the first two lines vanish based on applying (2.45) to the smooth solutions on the sub-intervals $a \le x < x_s$ and $x_s < x \le b$ respectively. The remaining terms give the so-called *Rankine–Hugoniot shock speed relation*

$$\frac{dx_s}{dt} = \frac{q(p_+(x_s,t)) - q(p_-(x_s,t))}{p_+(x_s,t) - p_-(x_s,t)}. \tag{2.48}$$

For the inviscid Burgers equation (2.39), the flux is $q(p) = \frac{1}{2}p^2$, and for our specific example, $p_-(x,t) = (1+x)/(1+t)$ and $p_+(x,t) = 0$ yielding the shock speed equation,

$$\frac{dx_s}{dt} = \frac{p_+(x_s,t) - p_-(x_s,t)}{2} \qquad \Longrightarrow \qquad \frac{dx_s}{dt} = \frac{1+x_s}{2(1+t)}. \tag{2.49}$$

Initial conditions for this equation are determined by the time and position where characteristics first cross, necessitating the insertion of a shock; in this case, $x_s(1) = 1$. Consequently the position of the shock is given by

$$x_s(t) = \sqrt{2(1+t)} - 1 \qquad \text{for } t \ge 1.$$

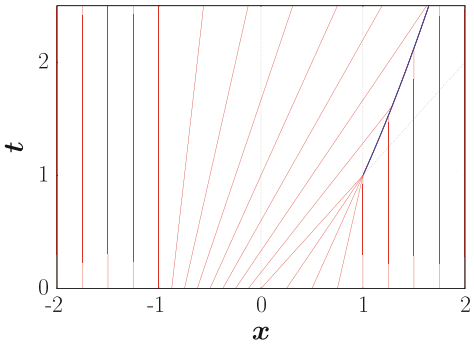


Fig. 2.5 Characteristic curves in the xt truncated by the shock $x = x_s(t)$ corresponding to (2.46)

Figure 2.5 shows the shock in the xt plane in its role in separating families of characteristics that would otherwise intersect. Excluding the shock curve, Fig. 2.5 has a single characteristic curve passing through each (x, t) point and hence describes a single-valued solution of the transport problem. This figure differs from Fig. 2.4 (left) only in the wedge-shaped region bounded by the curves $x = 1$ and $x = t$ that form the boundaries of what is sometimes called the *shock envelope*.

Returning to Fig. 2.3, we note that our solution $p(x, t)$ began as an equilateral triangle on $-1 \le x \le 1$ with height one, and hence area one. As time increases the profile steepens toward the right while maintaining its base and height, and hence area, even as it transitions from being an acute triangle (single-valued solution) to an obtuse triangle (multivalued solution) (also see Fig. 2.4 (left)). The consequence of introducing the shock is to cut out the portion of (2.43) that overturns while modifying the domains on which the other parts of the solution apply for $t \ge 1$,

$$p(x,t) = \begin{cases} 0 & x < -1, \\ (1+x)/(1+t) & -1 \le x \le x_s(t), \\ 0 & x > x_s(t). \end{cases} \tag{2.50}$$

The resulting solution profiles are right triangles on the base $-1 \le x \le x_s(t)$ and p ranging over $0 \le p \le \max p = p_-(x_s(t), t) = \sqrt{2/(1+t)}$. To satisfy the conservation of the integral of p, (2.48) ensures that the shock maintains the area of $A = \frac{1}{2}$ (base)(height) $= \frac{1}{2}(\sqrt{2(1+t)} - 1 - (-1))(1/\sqrt{2(1+t)}) = 1$. This illustrates why the shock selection rule is sometimes called the *equal-area rule*.[2] Later, we will see in Chap. 5 that solution (2.50) can also be obtained as a *similarity solution*.

[2] See Fig. 2.4—the placement of the shock not only conserves the area of the newly-formed right triangle, but also requires that the areas of the two cut-off multi-valued regions from the obtuse triangle to be equal.

2.7 Further Directions

The classic text on linear and nonlinear waves is the book by Whitham [106]. There are additional more recent books on waves [65, 80] and the method of characteristics and its extensions are covered in detail in most books on applied partial differential equations and modelling (see [45, 81], for example).

2.8 Exercises

2.1 Let $x = a(t)$ and $x = b(t)$ be positions of two points that move according to a flow $dx/dt = v(x, t)$, then if

$$f_{\text{avg}}(t) = \frac{1}{b(t) - a(t)} \int_{a(t)}^{b(t)} f(x, t)\, dx,$$

calculate $\lim_{b \to a} \dfrac{df_{\text{avg}}}{dt}$. Hint: Let $b(t) = a(t) + \varepsilon h(t)$ with $\varepsilon \to 0$.

2.2 In one dimension, the *Euler equations* for compressible gas dynamics are

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho v) = 0, \qquad \rho\left(\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x}\right) = -\frac{\partial P}{\partial x},$$

where the gradient of the pressure $P(x, t)$ represents an internal force generated by the gas. Note that the left-hand side of the second equation can be written in terms of the convective derivative, $\rho Dv/Dt$. The first equation is the continuity equation for the conservation of mass. Show that the second equation is consistent with the conservation of momentum with non-constant density,

$$\frac{\partial(\rho v)}{\partial t} + \frac{\partial(\rho v^2)}{\partial x} = -\frac{\partial P}{\partial x}.$$

2.3 Consider the solution of a linear wave equation having the dispersion relation $\omega = \omega(k)$,

$$\rho(x, t) = \cos(kx - \omega(k)t) + \cos([k + \varepsilon]x - \omega(k + \varepsilon)t).$$

(a) Show that for $\varepsilon \to 0$ this can be re-written in terms of the *phase velocity*, c_p, and *group velocity*, c_g, as

$$\rho(x, t) = 2\cos(k[x - c_p(k)t])\cos\left(\tfrac{1}{2}\varepsilon[x - c_g(k)t]\right) + O(\varepsilon)$$

What is the formula for the group velocity for $\varepsilon \to 0$?

(b) Use $\rho(x,t) = \cos(kx - \omega t)$ to determine the dispersion relation for the equation

$$\rho_t + \rho_x - \rho_{xxt} = 0 \tag{2.51}$$

and calculate the group velocity. Also determine the "modified dispersion relation" $\tilde{\omega}(k)$ from $\rho(x,t) = e^{kx - \tilde{\omega}t}$, which will be used in the next exercise.

2.4 *Solitons* (or "solitary waves") are steady profile travelling waves describing a single "pulse" whose size and speed are connected through nonlinear effects. Consider the nonlinear wave equation,

$$\rho_t + \rho_x + 6\rho\rho_x - \rho_{xxt} = 0,$$

called the *Benjamin-Bona-Mahony equation*.

(a) By looking for solutions of the PDE in travelling wave form, $\rho(x,t) = P(x - ct)$, determine the ODE for $P(s)$ with $s = x - ct$.
(b) Show that there is a one-parameter family of solutions of the form

$$P(s) = A\text{sech}^2(Bs)$$

and determine how A and B are related to the speed c.
(c) Show that the nonlinear solitary wave surprisingly satisfies the modified dispersion relation from the linearised equation (2.51).

2.5 Dispersion relations are not limited to being algebraic relations. Consider the following fluid dynamics model of water waves, given in terms of a potential function $\phi(x,y,t)$ (defined on $0 \le y \le 1$) and a wave profile $f(x,t)$ on the surface of the water (at $y = 0$),

$$\begin{aligned} \phi_{xx} + \phi_{yy} &= 0 \quad && \text{on } 0 \le y \le 1, \\ \phi_y &= 0 && \text{at } y = 0, \\ \phi_t + f &= 0 && \text{at } y = 1, \\ f_t &= \phi_y && \text{at } y = 1. \end{aligned}$$

Assume the wave is $f(x,t) = A\cos(kx - \omega t)$. Show that the corresponding potential must be of the form $\phi(x,y,t) = B(y)\sin(kx - \omega t)$ and obtain the dispersion relation $\omega(k)$.

2.6 Obtain explicit solutions $\rho = \rho(x,t)$ for $t \ge 0$ for the following problems,

(a) The initial value problem for $\rho(x,t)$ on $-\infty < x < \infty$:

$$\frac{\partial \rho}{\partial t} + e^{2t}\frac{\partial \rho}{\partial x} = \rho + x + t, \qquad \rho(x, t = 0) = \cos x.$$

(b) The signalling problem[3] for $\rho(x, t)$ on $x \geq 0$:

$$\frac{\partial \rho}{\partial t} + (x+4)\frac{\partial \rho}{\partial x} = -2\rho, \quad \rho(x=0,t) = \cos t, \quad \rho(x, t=0) = e^{-x}.$$

2.7 A one-dimensional compressible fluid blob starts at $t = 0$ with uniform density $\rho \equiv 1$ on $1 \leq x \leq 2$. The blob obeys the conservation of mass equation

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho v)}{\partial x} = 0.$$

with the (Eulerian) velocity field given as $v(x, t) = x^2 e^{-3t}$.

(a) Find the density of the blob for $t \geq 0$ as a function of position and time, $\rho = \rho(x, t)$.
(b) Find the positions of the moving left and right edges of the blob, $x_1(t) \leq x \leq x_2(t)$.
(c) Use your results from parts (a, b) to directly evaluate the integral

$$\int_{x_1(t)}^{x_2(t)} \rho(x, t)\, dx$$

and show that this is consistent with the Reynolds transport theorem.

2.8 (*The method of characteristics in two dimensions*) At time $t = 0$, a radioactive substance is released into a steady two-dimensional flow. The initial concentration of the substance is given by

$$c_0(x, y) = \begin{cases} 1 - (x^2 + y^2) & x^2 + y^2 \leq 1, \\ 0 & \text{else.} \end{cases}$$

In the absence of flow, the concentration would decay according to the rate equation $dc/dt = -c$. The substance is carried by the two-dimensional Eulerian velocity field $\mathbf{v} = (1 + x, y)$.

(a) The conservation law for the decaying substance on any fluid blob $D(t)$ is

$$\frac{d}{dt}\left(\iint_{D(t)} c\, dA\right) = -\iint_{D(t)} c\, dA.$$

Use the Reynolds transport theorem to derive the PDE for the concentration field $c(x, y, t)$.
(b) Instead of writing "$\mathbf{A} = (x_0, y_0)$" as a material coordinate in rectangular coordinates, parametrise the initial data in terms of polar material coordinates (R, Θ).

[3] A wave being specified by a boundary condition from a fixed "signal source" position.

Write the characteristic ODEs (dX/dt, dY/dt, dC/dt) and solve these equations with the given initial conditions.

(c) Obtain the explicit Eulerian solution $c(x, y, t)$.

(d) The boundary of the region over which the substance has spread remains circular for all times; find its radius and the coordinates of its centre as functions of time.

2.9 Consider the system of wave equations

$$p_t + 5p_x - 7q_x = 0,$$
$$q_t + 2p_x - 4q_x = 0.$$

(a) Find the eigenvalues and eigenvectors for travelling waves in this system.

(b) Find the solutions $p(x, t)$ and $q(x, t)$ if the initial conditions at $t = 0$ are

$$p(x, 0) = 5\sin(7x), \qquad q(x, 0) = -9\cos(4x).$$

2.10 *The classic wave equation* is

$$\frac{\partial^2 \phi}{\partial t^2} = c^2 \frac{\partial^2 \phi}{\partial x^2}. \tag{2.52}$$

It can be used to describe small transverse vibrations of a displaced string (like a guitar or violin string) starting from initial conditions for the position and velocity of each point along the string at $t = 0$:

$$\phi(x, 0) = f(x), \qquad \phi_t(x, 0) = g(x).$$

(a) Show that this problem can be written as a system of first order wave equations:

$$p_t - c^2 q_x = 0, \qquad q_t - p_x = 0,$$

where $q = \phi_x$ and $p = \phi_t$.

(b) Solve the system to obtain $p(x, t)$ and $q(x, t)$ in terms of travelling waves.

(c) Show that from (b) we can write the solution of the original problem as

$$\phi(x, t) = A(x - ct) + B(x + ct).$$

(d) Determine the functions $A(x)$ and $B(x)$ in terms of the initial data, functions $f(x)$ and $g(x)$. This form of the solution is called the *D'Alembert solution*.

2.11 *The shallow water equations* describe fluid flow in shallow (long, slender) layers, such as rivers. In simplest form, they are

$$\frac{\partial h}{\partial t} + h\frac{\partial v}{\partial x} + v\frac{\partial h}{\partial x} = 0, \qquad \frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} + g\frac{\partial h}{\partial x} = 0,$$

where $v(x,t)$ is the fluid speed, $h(x,t)$ is the height of the fluid layer, and g is the acceleration due to gravity.
Show that the shallow water equations correspond to the following Lagrangian statements about conservation of mass and change of momentum for all fluid blobs moving with the flow:

$$\frac{d}{dt}\left(\int_{a(t)}^{b(t)} h\,dx\right) = 0, \qquad \frac{d}{dt}\left(\int_{a(t)}^{b(t)} hv\,dx\right) = -G(h)\Big|_{x=a}^{x=b}$$

Determine the function $G(h)$.

2.12 To completely specify the problem for the determination of the position of a shock in the Rankine–Hugoniot equation (2.48), we must provide initial conditions on when/where the shock first forms, $x_s(t_*) = x_*$.

Consider the transport equation $p_t + q(p)_x = 0$ with initial condition $p(x,0) = f(x)$. Assume that $q'(p) > 0$ and $f(x)$ has a local maximum (for example, $f(x) = e^{-x^2}$).

(a) Consider two characteristic curves, $x = X(t;A_0)$ and $x = X(t;A_1)$, and assume that they carry different values of the solution ($P_0 \neq P_1$). Determine the position, $x_{0,1}$, and time, $t_{0,1}$, where they will intersect.
(b) Determine the time when the shock first forms, t_* in terms of q, f, by minimising the (positive) intersection time over all pairs of characteristics. Hint: Consider the limit of two characteristics starting very close together.

2.13 (*Shocks in quasilinear equations*) Consider the inviscid Burgers equation,

$$\frac{\partial p}{\partial t} + p\frac{\partial p}{\partial x} = 0,$$

starting from the initial conditions

$$p(x, t=0) = \begin{cases} 9 - x^2 & |x| \le 3, \\ 0 & |x| > 3. \end{cases}$$

(a) Use the method of characteristics to construct the parametric solution.
(b) Eliminate the parameter from the solution found in part (a) to obtain a multi-valued solution in two parts.
(c) Determine the time t_* and x_*-position where $x(t)$ characteristic curves first intersect (see Exercise 2.12).

(d) Use the results from parts (b, c) along with the shock speed equation (2.48) to write the ODE initial value problem for the shock position $x_s(t)$.
(e) Write the inviscid Burgers equation as a conservation law, state the conserved quantity (with the specific value set by the above initial condition), and integrate using the solution from (b) to produce an algebraic equation involving t and $x_s(t)$.

2.14 Consider the (viscous) Burgers equation

$$\frac{\partial p}{\partial t} + p\frac{\partial p}{\partial x} = \varepsilon^2 \frac{\partial^2 p}{\partial x^2} \qquad \varepsilon \to 0$$

subject to boundary conditions

$$p(x \to -\infty) = 2, \qquad p(x \to \infty) = 1.$$

Determine the first order ODE satisfied by travelling wave solutions, $p(x, t) = P(x - ct)$. Solve the ODE and apply the boundary conditions to determine travelling wave profile $P(s)$ and the wave speed c. Show that this matches the shock speed that would be obtained with $\varepsilon = 0$, for (2.39).

2.15 (*Fully-nonlinear first order PDEs*) The most general first-order PDE involving only first derivatives of a solution $p(x, t)$ can be written as

$$F(p, p_x, p_t, x, t) = 0,$$

where F is a given function. Let s be a parametric variable and parametrise all quantities in terms of s as

$$x = X(s) \qquad t = T(s) \qquad p = P(s) \qquad p_x = R(s) \qquad p_t = Q(s)$$

(a) Starting from $P(s) = p(X(s), T(s))$ take the derivative of P with respect to s and use the chain rule to express dP/ds in terms of R, Q, X', T'.
(b) Take the s-derivative of $F = 0$ using the chain rule and then use the result from part (a) to write the equation in the form

$$\frac{dF}{ds} = \underbrace{\left[A\frac{dR}{ds} + B\frac{dX}{ds}\right]}_{=0} + \underbrace{\left[C\frac{dQ}{ds} + D\frac{dT}{ds}\right]}_{=0} = 0.$$

Note that selecting $X' = A, R' = -B, T' = C, Q' = -D$ satisfies the overall equation.
Obtain a system of five autonomous ODEs for X, T, P, R, Q in terms of those variables and derivatives of F.

(c) Determine the characteristic equations for the general *Hamilton–Jacobi PDE*,

$$p_t + H(p, p_x, x, t) = 0,$$

where H is a given function.

(d) Obtain the parametric solution, $X(A, t), P(A, t)$ for the problem for $p(x, t)$ on $-\infty < x < \infty$ and $t \geq 0$,

$$\frac{\partial p}{\partial t} + \left(\frac{\partial p}{\partial x}\right)^4 = 0, \qquad p(x, 0) = e^{-x^2}.$$

Chapter 3
Variational Principles

In many problems, we need to find an optimal solution, one that maximises a benefit or minimises a cost for example. Various branches of science include formulations based on such principles:

- Fermat's principle of least time in optics
- Hamilton's principle of least action in mechanics
- "paths of least resistance" in electrostatics, hydrology, and other areas.

The criteria selecting an "optimal" solution in a physical system may not always be clear (especially in biological problems). However, if we are given a quantity to be minimised or maximised, the *calculus of variations* provides a natural methodology for reformulating the question in terms of a differential equation problem.

In this chapter we focus on problems that require the determination of a *function*, say $y_*(x)$, as the optimal solution to a given system. A classical example known as the *brachistochrone* problem (meaning "shortest time" in Greek) is motivated by the question: *What shape should a ramp take in order to deliver a mass (moving under the influence of gravity) to a specified final position in the least time?*; Fig. 3.1 shows schematics of some possible trial solutions.

Our approach will extend elementary methods from calculus for finding local optimal points to problems where solution functions, $y_*(x)$, are optimal relative to all small possible variations of that function.

3.1 Review and Generalisation from Calculus

Since our presentation will build on the basic formulations for finding maxima and minima from single- and multi-variable calculus, we briefly review that background as a means of introducing the terminology that we will employ.

Consider the problem of finding local extrema of a smooth function $y = f(x)$, called the *objective function*. If f has a local maximum at $x = x_*$ with value $y_* = f(x_*)$ then y_* must be greater or equal to all values of f achieved in a small neighbourhood

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9_3

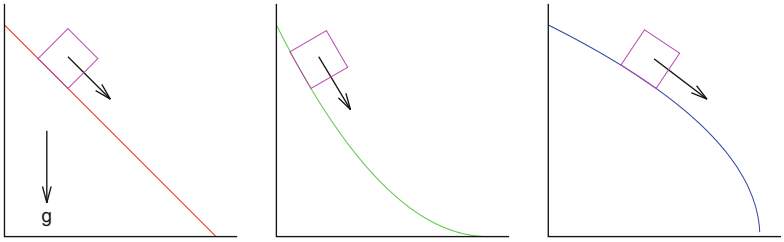


Fig. 3.1 Three trial solutions for the brachistochrone problem for the shape of a ramp to move a mass down from (0, 1) to (1, 0) under the influence of gravity as quickly as possible, $\tilde{y} = (1-x)^{\beta}$ for $\beta = 1, 2, 1/2$

around x_*,

$$f(x_*) \geq f(x_* + \varepsilon) \qquad \text{for all } |\varepsilon| \to 0. \tag{3.1}$$

Taking a Taylor series expansion of f at x_* yields

$$f(x_* + \varepsilon) = f(x_*) + f'(x_*)\varepsilon + \tfrac{1}{2}f''(x_*)\varepsilon^2 + \cdots \qquad \text{as } \varepsilon \to 0. \tag{3.2}$$

If the slope $f'(x_*)$ were nonzero, then for $|\varepsilon| > 0$ either $f(x_* + \varepsilon)$ or $f(x_* - \varepsilon)$ would have a value greater than $f(x_*)$, violating our assumption that $f(x_*)$ is a maximum. Consequently, we must have $f'(x_*) = 0$ at a maximum. An analogous argument applies for local minima of $y = f(x)$, and so local extrema for smooth functions can only arise at *critical points*, where

$$f'(x_*) = 0.$$

Having eliminated the linear term from the Taylor series, the higher order terms in (3.2) must also respect the requirement (3.1). This leads to the standard condition on the second derivative for a local maximum ($f''(x_*) < 0$) (or minimum ($f''(x_*) > 0$)). If $f''(x_*) = 0$ at a critical point, then the next nonzero higher order terms need to be considered in order to determine if the critical point indeed yields a maximum or a minimum of f, or is neither (an inflection point).

Similarly, for a two-variable objective function, having a local maximum of f at (x_*, y_*) with value $z_* = f(x_*, y_*)$ implies that all nearby points satisfy

$$f(x_*, y_*) \geq f(x_* + \varepsilon_1, y_* + \varepsilon_2) \qquad \text{for all } |\varepsilon_1|, |\varepsilon_2| \to 0.$$

The corresponding multi-variable Taylor series for $\varepsilon_1, \varepsilon_2 \to 0$ yields

$$f(x_* + \varepsilon_1, y_* + \varepsilon_2) = f(x_*, y_*) + \underbrace{\partial_x f(x_*, y_*)\varepsilon_1 + \partial_y f(x_*, y_*)\varepsilon_2}_{\text{[first derivative terms]}} + \cdots .$$

As in the single variable case, the linear terms must vanish at a local maximum or minimum, and since this must hold for all choices of ε_1 and ε_2, we arrive at the generalisation of the previous critical point condition,

$$\partial_x f_* = 0, \qquad \partial_y f_* = 0. \tag{3.3}$$

Namely, the first partial derivatives of the objective function with respect to all of its independent variables must vanish, which can be more compactly expressed in terms of the gradient

$$\nabla f_* = \mathbf{0}. \tag{3.4}$$

This definition of critical points in terms of the gradient applies to functions of any number of independent variables. Being a critical point is a necessary condition for a local extrema, but it is not sufficient. Conditions on the second derivatives (represented by the *Hessian matrix* for functions of two or more variables) also extend to arbitrary numbers of variables for determining whether a critical point yields a maximum, minimum or inflection point of the objective function [24, 68]. Whether we are seeking a maximum or a minimum, all local extrema are determined by the critical point conditions and hence we focus on obtaining critical points of the more challenging class of problems that we will consider.

3.1.1 Functionals

In general, a *functional* is a mathematical expression that can be evaluated to give a scalar value corresponding to each function to which it is applied. Two simple examples are provided by the integrals that correspond to the area under a curve and the arclength of a curve,

$$A(y) = \int_0^1 y(x)\,dx, \qquad S(y) = \int_0^1 \sqrt{1 + y'(x)^2}\,dx,$$

for any given function $y(x)$ on $0 \le x \le 1$.

For the most part, we will only consider functionals that are definite integrals of a function and its derivatives,

$$J(y) = \int_a^b L(x, y(x), y'(x), y''(x), \ldots)\,dx. \tag{3.5}$$

Here the integrand function $L(x, y, \ldots)$ is called the *Lagrangian* (named after Joseph Lagrange (1736–1813) who reformulated classical mechanics based on a variational principle—see Sect. 3.3).

For the above geometric examples of area and arclength, the dependence of the functional on the solution is straightforward, while for the brachistochrone problem

from Fig. 3.1, the connection is a little less direct. The time of travel to be minimised can be expressed as

$$T(y) = \int_0^T dt = \int \frac{ds}{v} = \int_0^1 \frac{\sqrt{1 + (y')^2}}{v(y)}\, dx, \tag{3.6}$$

where we have used the relationship between the speed and distance travelled, $v = ds/dt$, and a further relationship is also needed between the speed and position (see Exercise 3.8).

Now consider the minimisation of an *objective functional*; by analogy with the previous analysis, we see that for $J(y)$ to have a local minimum[1] for a particular function $y_*(x)$ the functional must satisfy

$$J(y_*(x)) \leq J(y_*(x) + \varepsilon h(x)),$$

for all $\varepsilon \to 0$ and for all admissible "perturbation functions" $h(x)$ of $y_*(x)$. For the moment, $h(x)$ is unspecified and serves to describe a neighbourhood of the optimal solution $y_*(x)$ within the space of smooth functions. While there are many technical issues underlying these concepts that are deserving of careful analysis, we defer detailed considerations to more advanced texts [71, 99].

We will describe a straightforward procedure (as an extension of the process of finding critical points of an objective function) for determining optimal solutions of a functional. In particular, we will have to explain what it means to take a derivative of a functional with respect to a function, and we will distinguish these from standard derivatives by using the terminology *variational derivatives*.

3.2 General Approach and Basic Examples

Given an objective functional, there is a systematic four-step approach which can be used to obtain smooth solutions that correspond to local extrema of the functional:

(i) Assume the existence of an optimal solution $y_*(x)$ and perform an expansion of the functional in the neighbourhood of $y_*(x)$ using

$$\tilde{y} = y_*(x) + \varepsilon h(x) \qquad \tilde{J} = J(\tilde{y}), \tag{3.7}$$

where the perturbation function $h(x)$ is independent of ε. Formally expanding $\tilde{J}$ as $\varepsilon \to 0$ like a Taylor series yields

$$\tilde{J} = J(y_* + \varepsilon h) = J(y_*) + \text{“}\left(\frac{dJ}{dy}\bigg|_{y_*}\right)(\varepsilon h)\text{”} + O(\varepsilon^2), \tag{3.8}$$

[1] The behaviour at a local maximum of $J(y)$ follows similarly.

where we still have to precisely define the notion of a variational derivative (in this case the evaluation of the derivative of an integral with respect to a function). The $O(\varepsilon)$ term[2] in the Taylor series of J will be called the *first variation*, the $O(\varepsilon^2)$ term the *second variation* and so on.

(ii) Apply the critical point condition: the first variation of the functional must be zero at the solution $y_*(x)$ for all admissible perturbations $h(x)$,

$$\text{“}\left(\frac{dJ}{dy}\bigg|_{y_*}\right)(\varepsilon h)\text{”} = 0 \tag{3.9}$$

We will explain the meaning of 'admissible' solutions and perturbations below.

(iii) Assuming the solution to be smooth, convert the critical point condition on the functional into a differential equation for the optimal[3] solution $y_*(x)$.

(iv) Solve the differential equation problem.

The details of this approach will be discussed and expanded out in following examples. This framework applies to many more complicated problems. However, an important disclaimer is that not every problem can be simplified to a differential equation using this framework, those that can be are called *variational problems*. A number of difficulties can arise to make a problem non-variational, such as the functional not having critical points, or it not being possible to reduce the first variation to a differential equation.

3.2.1 The Simple Shortest Curve Problem

We begin by considering the simple problem of finding the function $y = y(x)$ that gives the shortest path from the origin $(0, 0)$ to the given fixed point $(1, b)$, $b \in \mathbb{R}$, see Fig. 3.2. While we know the answer to be the straight line connecting the two points, it will be instructive to see how the calculus of variations framework constructs this result.

The shortest path is the one that minimises the total arclength and hence our objective functional is

$$J(y) = \int_0^1 \sqrt{1 + (y')^2}\, dx. \tag{3.10a}$$

[2] The $O(\varepsilon^n)$ order symbol will be defined precisely in Chap. 6, but for the current context we use this to refer to the $\varepsilon \to 0$ limit of the remainder of the terms in the series with coefficients ε^N for $N \geq n$.

[3] Showing that a solution given by a critical point is a local maximum or minimum involves evaluating the second variation at $y_*(x)$.

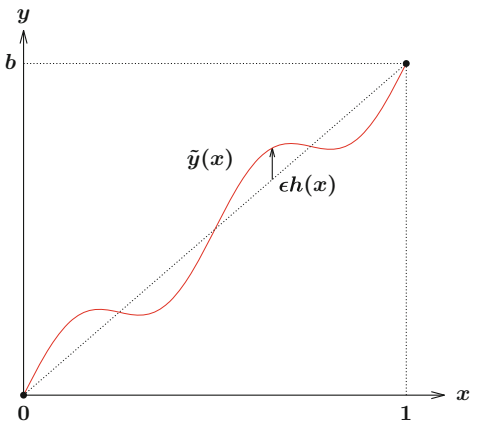


Fig. 3.2 An admissible trial solution $\tilde{y}(x)$ for problem (3.10a, 3.10b) and its perturbation from the optimal solution $y_*(x)$

The statement of the problem identifies boundary conditions that all solutions must satisfy,

$$y(0) = 0, \qquad y(1) = b. \tag{3.10b}$$

Analysis of the problem begins by identifying the properties of the family of perturbations $h(x)$ to the optimal solution $y_*(x)$ that define all admissible trial solutions (3.7), $\tilde{y}(x) = y_*(x) + \varepsilon h(x)$. Applying boundary conditions (3.10b) to the admissible solution $\tilde{y}$ yields that

$$\tilde{y}(0) = y_*(0) + \varepsilon h(0) = 0, \qquad \tilde{y}(1) = y_*(1) + \varepsilon h(1) = b.$$

Since the boundary conditions on the optimal solution are necessarily the same as those on the trial solutions ($y_*(0) = 0$ and $y_*(1) = b$), we obtain that the perturbation functions must satisfy

$$h(0) = 0, \qquad h(1) = 0. \tag{3.11}$$

Next, we substitute (3.7) into (3.10a) giving

$$\tilde{J} = \int_0^1 \sqrt{1 + \left(y_*' + \varepsilon h'\right)^2}\, dx = \int_0^1 \sqrt{1 + (y_*')^2 + 2\varepsilon h' y_*' + \varepsilon^2 (h')^2}\, dx.$$

For investigating the limit $\varepsilon \to 0$, it is convenient to factor the integrand as

$$= \int_0^1 \sqrt{\left(1 + (y_*')^2\right)\left(1 + \frac{2\varepsilon h' y_*'}{1 + (y_*')^2} + \frac{\varepsilon^2 (h')^2}{1 + (y_*')^2}\right)}\, dx$$

and then using the Taylor series expansion $\sqrt{1+z} = 1 + \frac{1}{2}z + O(z^2)$ as $z \to 0$ applied to the second factor allows us to write

$$\tilde{J} = \left(\int_0^1 \sqrt{1 + (y_*')^2}\, dx \right) + \varepsilon \left(\int_0^1 \frac{h' y_*'}{\sqrt{1 + (y_*')^2}}\, dx \right) + O(\varepsilon^2). \tag{3.12}$$

The first term in this expansion gives the expression for the functional at the optimal solution, $J_* = J(y_*)$, but since y_* is unknown, in order to evaluate this integral, we must examine further terms. The $O(\varepsilon)$ term provides the first variation (3.8) that will determine the critical points.

It is important to note that the form of the first variation in (3.8) was chosen to intentionally suggest that the perturbation function $h(x)$ should appear *linearly and undifferentiated*. We will always seek to put the first variation term into this form.

For (3.12), getting to this form can be accomplished by performing integration by parts to shift the derivative off of the h' factor,

$$\int_0^1 \frac{h' y_*'}{\sqrt{1 + (y_*')^2}}\, dx = \frac{y_*' h}{\sqrt{1 + (y_*')^2}} \bigg|_{x=0}^{x=1} - \int_0^1 \frac{d}{dx} \left(\frac{y_*'}{\sqrt{1 + (y_*')^2}} \right) h\, dx;$$

integration by parts will be a fundamental calculational tool in most of our problems. The boundary conditions on h given by (3.11) eliminate the contributions from the boundary terms, and leave the final form of the first variation of the functional as a single integral

$$\frac{\delta J}{\delta y}\bigg|_{y_*} h = - \int_0^1 \frac{d}{dx} \left(\frac{y_*'}{\sqrt{1 + (y_*')^2}} \right) h\, dx, \tag{3.13}$$

where we use the δ-notation to indicate that this is a variational (functional) derivative (also sometimes called a *Frechet derivative*).

The next step is to enforce the critical point condition (3.9) on (3.13) for all admissible perturbations. At this point, we need a basic, but important result from analysis called the *Fundamental Lemma of the Calculus of Variations* which states that

$$\text{If} \quad \int_a^b g(x) h(x)\, dx = 0 \quad \forall h(x) \quad \text{then} \quad \boxed{g(x) \equiv 0 \text{ on } a \le x \le b.} \tag{3.14}$$

In other words, if the integral of a product of functions is zero for all choices of one of the factors, then this is only possible if the other factor is identically zero on the whole interval of integration.[4,5]

Applying the fundamental lemma to the critical point condition for (3.13),

$$-\int_0^1 \frac{d}{dx}\left(\frac{y_*'}{\sqrt{1+(y_*')^2}}\right) h\,dx = 0, \tag{3.15}$$

for all possible h, we obtain the differential equation for $y_*(x)$

$$-\frac{d}{dx}\left(\frac{y_*'}{\sqrt{1+(y_*')^2}}\right) = 0. \tag{3.16}$$

We have consequently replaced an integral condition (the *weak form*, that applies over the whole domain) with an equivalent differential equation (the *strong form*) that applies pointwise on smooth solutions. Together with boundary conditions (3.10b), (3.16) provides us with a complete ODE boundary value problem defining the optimal solution $y_*(x)$.

In this problem it is straightforward to integrate the ODE once yielding a constant of integration C_1, then following some algebra,

$$\frac{y_*'}{\sqrt{1+(y_*')^2}} = C_1 \qquad \Rightarrow \qquad (y_*')^2 = \frac{C_1^2}{1-C_1^2} = C_*^2 \geq 0, \tag{3.17}$$

for some constant C_*. Then using the boundary conditions (3.10b) yields $y_*(x) = bx$, confirming that the shortest path is a straight line.

A second derivative test can be applied to verify that this solution is a local minimum of (3.10) (see Exercise 3.2), but here we can use an understanding of the structure of the problem to show that it is indeed the solution we seek.[6] In many cases, problem-specific intuition can be used to distinguish whether critical point solutions minimise or maximise the functional.

A more general version of this problem, allowing for solutions as parametric curves in the plane is given in Exercise 3.3.

[4] In this simplified statement of this result, we are assuming that $g(x)$ is smooth and hence has no discontinuities.

[5] The du Bois Reymond lemma (2.11) is closely related to this lemma, choosing $h(x) = 1$ on arbitrary sub-intervals, and otherwise $h = 0$.

[6] For the area and arclength examples, the functionals become unbounded for large amplitude solutions and hence there are no local maxima.

3.2.2 The Classic Euler–Lagrange Problem

Consider an objective functional defined in terms of an integral on the interval $a \leq x \leq b$ with an integrand that is a function of the independent variable x, the solution $y(x)$ and its first derivative $y'(x)$

$$J(y) = \int_a^b L(x, y, y')\, dx, \tag{3.18a}$$

subject to prescribed boundary conditions on y given by

$$y(a) = c, \qquad y(b) = d. \tag{3.18b}$$

Parallelling the previous example, we find that admissible perturbation functions must satisfy the homogeneous boundary conditions

$$h(a) = 0, \qquad h(b) = 0. \tag{3.19}$$

The expanded form of $J(y_* + \varepsilon h)$ is given by

$$\tilde{J} = \int_a^b L(x, y_* + \varepsilon h, y_*' + \varepsilon h')\, dx$$

and by applying the multi-variable Taylor series, this can be written for $\varepsilon \to 0$ as

$$\begin{aligned} \tilde{J} &= \int_a^b \left[L(x, y_*, y_*') + \varepsilon \left(\frac{\partial L}{\partial y} h + \frac{\partial L}{\partial y'} h' \right) \Big|_{y=y_*} + O(\varepsilon^2) \right] dx \\ &= J_* + \varepsilon \left(\int_a^b \frac{\partial L}{\partial y} h\, dx + \int_a^b \frac{\partial L}{\partial y'} h'\, dx \right) + O(\varepsilon^2). \end{aligned}$$

In order to address the h' derivative factor in the third term, we apply integration by parts to obtain

$$\int_a^b \frac{\partial L}{\partial y'} h'\, dx = \frac{\partial L}{\partial y'} h \Big|_{x=a}^{x=b} - \int_a^b \frac{d}{dx} \left(\frac{\partial L}{\partial y'} \right) h\, dx.$$

Imposing the boundary conditions (3.19) on the perturbation function, we can then re-group the first variation as

$$\tilde{J} = J_* + \varepsilon \int_a^b \left[\frac{\partial L}{\partial y} - \frac{d}{dx} \left(\frac{\partial L}{\partial y'} \right) \right] h\, dx + O(\varepsilon^2). \tag{3.20}$$

This was an important and necessary step because in order to apply the fundamental lemma (3.14) to the critical point condition, we must have the first variation expressed as a single integral having the perturbation $h(x)$ as a factor,

$$\int_a^b \left[\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'} \right) \right] h \, dx = 0.$$

Since this holds for all admissible choices of $h(x)$, by (3.14) we obtain the differential equation on $a \le x \le b$,

$$\boxed{\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'} \right) = 0} \tag{3.21}$$

Subject to the boundary conditions (3.18b), (3.21) defines a differential equation problem that should have a unique solution for $y = y_*(x)$.

Equation (3.21) is known as the *Euler–Lagrange* equation for the functional (3.18a). In Sect. 3.3, we will see that this particular form is used extensively in problems for mechanical systems.

A very useful consequence of the general form of L in (3.18a) is that the Euler–Lagrange differential equation for any functional with $L = L(x, y, y')$ is given by substituting the specific L into (3.21). For example, the ODE (3.16) for the shortest distance problem (3.10a) is produced by (3.21) with $L = \sqrt{1 + (y')^2}$. Similarly, if our Lagrangian were $L(x, y, y') = f(x)y + g(x)y' + (y')^2$, then

$$\frac{\partial L}{\partial y} = f(x), \qquad \frac{\partial L}{\partial y'} = g(x) + 2y', \qquad \frac{d}{dx}\left(\frac{\partial L}{\partial y'} \right) = g'(x) + 2y''$$

and (3.21) gives the Euler–Lagrange equation for $y_*(x)$ as

$$f(x) - g'(x) - 2\frac{d^2y}{dx^2} = 0.$$

3.3 The Variational Formation of Classical Mechanics

One of the most wide-spread applications of the calculus of variations is in deriving the equations of motion for mechanical systems from the critical points of a functional. The theory is based on a restatement of the previous Euler–Lagrange results—we relabel the variables to describe the motion of a mass:

$$
\begin{aligned}
x &\to t && : \text{independent variable (time)} \\
y(x) &\to y(t) && : \text{solution (position)} \\
y'(x) &\to y'(t) && : \text{derivative (velocity)} \\
L(x, y, y') &\to L(t, y, y') && : \text{still called the Lagrangian} \\
J = \int L\, dx &\to I = \int L\, dt && : \text{functional (action)}
\end{aligned}
$$

The Euler–Lagrange equation (3.21) now takes the form

$$\frac{\partial L}{\partial y} - \frac{d}{dt}\left(\frac{\partial L}{\partial y'}\right) = 0, \tag{3.22}$$

and the key step in the application to mechanics is in identifying the Lagrangian function L as the difference between the kinetic energy T and potential energy V

$$L = T - V. \tag{3.23}$$

William Hamilton (1805–1865) used this form of Lagrangian based on the assumption that mechanical systems should satisfy what come to be called *Hamilton's principle of least action*, namely that the optimal (actual, observable) solution $y_*(t)$ will be the one that minimises the action integral, I.

A very simple example of a mechanical system is the vertical motion of a rigid object with mass m subject to gravity g. The kinetic energy of the point mass is $T = \frac{1}{2}m(y')^2$, while the potential energy due to gravity is given by $V(y) = mgy$. Substituting this form for $L(t, y, y')$ (3.23) into the Euler–Lagrange equation (3.22) yields

$$\frac{dV}{dy} - m\frac{d(y')}{dt} = 0 \quad \Rightarrow \quad m\frac{d^2y}{dt^2} = -\frac{dV}{dy} \quad \Rightarrow \quad m\frac{d^2y}{dt^2} = -mg,$$

exactly as would be obtained from Newton's second law, but without the need to explicitly work out the forces acting on the system. The difference between the Lagrangian and force-based approaches is not significant in this example, but for more complicated systems, the Euler–Lagrange approach can notably simplify the process of deriving the equations of motion.

Thorough treatments of the implications of this formulation are given in books on intermediate and advanced mechanics [40, 62, 67]; we limit ourselves here to the basic concepts that are most immediately useful for the formulation of models for mechanical systems.

3.3.1 Motion with Multiple Degrees of Freedom

One generalisation of the Euler–Lagrange equation of motion is the derivation of the equations of motion for a multi-variable systems. As a specific example, consider the motion of a projectile in two dimensions with position $\mathbf{x} = (x(t), y(t))$ subject to gravity acting in the y-direction. The action can be written in terms of the difference of the kinetic and potential energies,

$$I = \int \left[\tfrac{1}{2} m \left((x')^2 + (y')^2 \right) - mgy \right] dt.$$

The principle of least action then requires that $I(x, y)$ is minimised over all possible choices of the unknown functions $x(t)$ and $y(t)$. We assume that an optimal solution $(x_*(t), y_*(t))$ exists and express all admissible solutions in terms of independent perturbations of the unknowns

$$\tilde{x}(t) = x_*(t) + \varepsilon h_1(t), \qquad \tilde{y}(t) = y_*(t) + \varepsilon h_2(t).$$

Substituting into the action, $I(\tilde{x}, \tilde{y})$ and Taylor expanding in the limit $\varepsilon \to 0$,

$$\tilde{I} = I_* + \varepsilon \int \left[m(x_*' h_1' + y_*' h_2') - mgh_2 \right] dt + O(\varepsilon^2).$$

By application of integration by parts and eliminating the boundary terms, we obtain the $O(\varepsilon)$ term as

$$- \int (mx_*'') h_1 + (my_*'' + mg) h_2 \, dt.$$

Satisfying the critical point condition implies that this integral must be equal to zero for all choices of $h_1(t)$ and $h_2(t)$. One class of possibilities is given by $h_2 \equiv 0$ and $h_1(t)$ arbitrary, this yields the equation $mx_*'' = 0$. Another choice is $h_1 \equiv 0$ with $h_2(t)$ arbitrary, yielding $mg + my_*'' = 0$; the overall result comes from the intersection of these cases using the linear independence of the individual perturbations.

This can be understood as a generalisation of fundamental lemma (3.14) for the dot product of vector functions, $\mathbf{g}(x), \mathbf{h}(x) \in \mathbb{R}^n$,

$$\text{If} \quad \int_a^b \mathbf{g}(x) \cdot \mathbf{h}(x) \, dx = 0 \quad \forall \mathbf{h}(x) \quad \text{then} \quad \mathbf{g}(x) \equiv \mathbf{0} \text{ on } a \le x \le b. \tag{3.24}$$

In other words, if the value of the integral is zero for all choices of component perturbation functions in $\mathbf{h}$ then each component of $\mathbf{g}$ must be identically zero, $g_i(x) \equiv 0$ for $i = 1, 2, \ldots, n$.

Hence we can arrive at the (general) result that the critical point condition will yield separate Euler–Lagrange equations with respect to each variable with an independent perturbation. In this case

$$\forall h_1 : \quad x\text{-Euler–Lagrange:} \qquad \frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial x'}\right) = 0, \tag{3.25a}$$

$$\forall h_2 : \quad y\text{-Euler–Lagrange:} \qquad \frac{\partial L}{\partial y} - \frac{d}{dt}\left(\frac{\partial L}{\partial y'}\right) = 0. \tag{3.25b}$$

Thus systems of differential equations for individual particles moving in two or three dimensions ($\mathbf{x} = (x, y, z)$), or sets of particles ($\mathbf{x}_1, \mathbf{x}_2, \ldots$), can be obtained from the Lagrangian for the complete system, $L(t, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_1', \mathbf{x}_2', \ldots)$, through the application of the Euler–Lagrange equation to each independent unknown function.

3.4 The Influence of Boundary Conditions

In the previous examples, we have seen the important roles played by the boundary conditions in:

(i) specifying boundary conditions for the Euler–Lagrange problem
(ii) determining boundary conditions on the perturbation functions
(iii) eliminating the boundary terms generated by integration by parts in the calculation of the first variation

The last point is perhaps the most important in terms of reducing the problem to a differential equation; if it is not possible to eliminate the boundary terms, then the fundamental lemma cannot be applied. Choices of boundary conditions that are compatible with this requirement are called *natural boundary conditions*.[7]

So far, we have seen how to work with Dirichlet boundary conditions (e.g. $y(a) = c$). We will now consider how other types of boundary conditions affect problems.

3.4.1 Problems with a Free Boundary

Consider a modification of the problem of determining the shortest path we considered in Sect. 3.2.1: find the function $y(x)$ passing through the origin that gives the shortest path to the vertical line $x = 1$.

Being a problem in terms of the minimum distance, the functional to be minimised is the same as in (3.10a),

[7] Given the lack of elementary approaches if the fundamental lemma can not be used, it is tempting to call them *necessary boundary conditions*.

$$J(y) = \int_0^1 \sqrt{1 + (y')^2}\, dx,$$

but the problem statement now provides only one definite boundary condition, $y(0) = 0$. There is no specific requirement on y at $x = 1$ and hence this is called a *free boundary*. In order to see how to deal with such a situation, we revisit (3.7) to consider the form of admissible solutions,

$$\tilde{y}(x) = y_*(x) + \varepsilon h(x).$$

At $x = 0$, we should have zero perturbation, $h(0) = 0$ since $\tilde{y}(0) = y_*(0) = 0$. At $x = 1$, $\tilde{y}(1)$ and $y_*(1)$ are arbitrary, and so we have no information about $h(1)$.

Proceeding as before, we expand $J(y_* + \varepsilon h)$ for $\varepsilon \to 0$ as

$$\tilde{J} = J_* + \varepsilon \left(\frac{y_*' h}{\sqrt{1 + (y_*')^2}} \bigg|_{x=0}^{x=1} - \int_0^1 \frac{d}{dx} \left(\frac{y_*'}{\sqrt{1 + (y_*')^2}} \right) h\, dx \right) + O(\varepsilon^2).$$

We need the total contribution from the boundary terms to vanish, requiring

$$\frac{y_*'(1)h(1)}{\sqrt{1 + y_*'(1)^2}} - 0 = 0,$$

where the $x = 0$ boundary term vanishes due to $h(0) = 0$. Since we do not know anything about $h(1)$, in order to guarantee that the remaining term vanishes, we require that the optimal solution satisfies the natural boundary condition, $y_*'(1) = 0$.

It is important to note that irrespective of the form taken by the natural boundary conditions, they are only involved in eliminating the boundary terms in order to leave the integral in a form compatible with the fundamental lemma. Consequently, the Euler–Lagrange equation is again given by (3.21) and (in this case) leads to the same ODE as found for the prescribed end-point case (3.16). However, the new boundary conditions for the Euler–Lagrange problem, $y_*(0) = 0$ and $y_*'(1) = 0$, now select a different solution, $y_*(x) \equiv 0$ (corresponding to the line from the origin along the x-axis up to the point of intersection with the line $x = 1$).

3.4.2 Problems with a Variable Endpoint

Another version of the "find the shortest path" problem is to find $y(x)$ that gives the shortest path from the origin to a given curve $y = f(x)$ (see Fig. 3.3).

The boundary condition $y_*(0) = 0$ holds as before, but now the other end-point can be any point on the curve $y = f(x)$. Let us denote this unknown point as $x = b$ and then the boundary condition is simply

$$y(b) = f(b). \tag{3.26a}$$

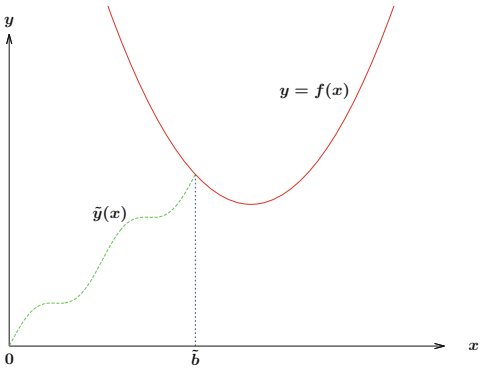


Fig. 3.3 A trial solution, $\tilde{y}(x)$ on $0 \le x \le \tilde{b}$, for the "shortest path to a given *curve*" variable endpoint problem (3.26)

The objective functional is (as before) the total arclength

$$J(y) = \int_0^b \sqrt{1 + (y')^2}\, dx. \tag{3.26b}$$

We begin the analysis as usual, by adding a perturbation to *each unknown*, $y(x)$ and b. As before, the admissible solutions take the form

$$\tilde{y}(x) = y_*(x) + \varepsilon h(x),$$

while the admissible endpoints can be written as

$$\tilde{b} = b_* + \varepsilon c,$$

where c is a perturbation constant. Consequently, we have

$$\tilde{J} = \int_0^{b_*+\varepsilon c} \sqrt{1 + (y_*' + \varepsilon h')^2}\, dx. \tag{3.27}$$

At this point in previous examples, we have Taylor expanded $\tilde{J}$ in the limit $\varepsilon \to 0$. A new complication present in this problem is that a perturbation appears in the limits of integration as well as in the integrand. One effective tool for dealing with this issue is Leibniz's rule for the derivative of an integral

$$\begin{aligned} &\frac{d}{dz}\left(\int_{a(z)}^{b(z)} g(x, y(x), z)\, dx\right) \\ &\quad = \int_a^b \frac{\partial g}{\partial z}\, dx + \left[g(b, y(b), z)\frac{db}{dz} - g(a, y(a), z)\frac{da}{dz}\right]. \end{aligned} \tag{3.28}$$

Applying Leibniz's rule to (3.27) with ε playing the role of z, we obtain the first variation in the form

$$\left.\frac{dJ}{d\varepsilon}\right|_{\varepsilon=0} = \int_0^{b_*} \frac{y_*' h'}{\sqrt{1+(y_*')^2}}\,dx + c\sqrt{1+y_*'(b_*)^2}.$$

After using integration by parts and imposing the critical point condition, this yields

$$\left(c\sqrt{1+y_*'(b_*)^2} + \frac{y_*'(b_*)h(b_*)}{\sqrt{1+y_*'(b_*)^2}}\right) - \int_0^{b_*} \frac{d}{dx}\left(\frac{y_*'}{\sqrt{1+(y_*')^2}}\right) h\,dx = 0.$$

If the boundary terms vanish, then applying the fundamental lemma to the integral will yield the expected Euler–Lagrange equation, but it is not yet clear which boundary conditions on y achieve this since c and $h(b_*)$ are unknown. However, we have not yet fully utilised (3.26a), namely that $\tilde{y}(\tilde{b}) = f(\tilde{b})$. Writing this out more explicitly, we have

$$y_*(b_* + \varepsilon c) + \varepsilon h(b_* + \varepsilon c) = f(b_* + \varepsilon c). \tag{3.29}$$

Expanding (3.29) in the limit $\varepsilon \to 0$, matching first terms in the respective expansions of the left- and right-hand sides, we recover $y_*(b_*) = f(b_*)$ for the optimal solution. Matching $O(\varepsilon)$ terms, we find that

$$y_*'(b_*)c + h(b_*) = f'(b_*)c$$

which can be solved for c

$$c = \frac{h(b_*)}{f'(b_*) - y_*'(b_*)}.$$

Substituting this result into the boundary terms yields

$$\left(\frac{y_*'(b_*)}{\sqrt{1+y_*'(b_*)^2}} + \frac{\sqrt{1+y_*'(b_*)^2}}{f'(b_*) - y_*'(b_*)}\right) h(b_*).$$

Since no constraints are known on the value of $h(b_*)$, in order to ensure that this term vanishes for all admissible perturbations, we require that the terms in parentheses sum to zero, which (after a little algebra) yields the natural boundary condition

$$y'(b_*) = -\frac{1}{f'(b_*)}. \tag{3.30}$$

Recalling that the Euler–Lagrange equation yields straight lines as solutions (3.17), this condition selects the minimum distance to the curve via the line that intersects the curve perpendicularly (recall the negative reciprocal slope relation from analytic geometry). While it may seem unusual to have three boundary conditions

(on $y(0), y(b), y'(b)$) for a second order differential equation, it is not overdetermined since b_* is unknown and has to be determined as part of the solution.

3.5 Optimisation with Constraints

While the solutions of many systems are determined from optimising some property of the system, in some cases, the choice of the optimal solution is constrained further by some implicit structure in the problem—such as a finite amount of building materials limiting the maximum size of a structure, or finite time limiting how thoroughly a task can be performed. These are examples of *constrained optimisation problems*—the specific constraints are crucial in selecting the relevant solution. In simple problems, the constraints can be substituted directly into the objective function in order to obtain a reduced problem describing all achievable solutions (see Exercise 3.5, for example), but for problems where this cannot be done, the method of *Lagrange multipliers* can be employed instead. This approach will be used to investigate several classes of calculus of variations problems in the remaining sections of this chapter,

(i) Isoperimetric problems: optimising a functional subject to a condition on an integral of the solution.
(ii) Holonomic systems: optimising a functional subject to a geometric condition applied pointwise on the solution.
(iii) Optimal control: optimising a functional subject to a differential equation applied pointwise on the solution.

We begin with a brief review of the method of Lagrange multipliers applied to problems from multivariable calculus.

3.5.1 Review of Lagrange Multipliers

In constrained optimisation problems, feasible solutions lie within a subset of the space of all possible solutions. The method of *Lagrange multipliers* identifies the feasible solutions through the same process used in unconstrained optimisation problems—namely obtaining the possible solutions from solving equations determined by the critical point condition (cf. (3.4))

$$\frac{\partial(\text{objective})}{\partial(\text{variables})} = \mathbf{0}.$$

The Lagrange multiplier approach constructs an "augmented objective function" $\mathscr{L}$ incorporating the original objective function, call it $f(x, y)$, along with all constraints so that the critical points are still described by a gradient condition

$$\nabla \mathscr{L} = \mathbf{0} \qquad \Leftrightarrow \qquad \text{all critical point solutions.} \tag{3.31}$$

Consider the fundamental problem of maximising a function of two variables subject to a single constraint:

Find (x_*, y_*) yielding $\max f(x, y)$ from among the points (x, y) on the curve $g(x, y) = 0$,

where $g(x, y) = 0$ is the implicit equation of a given curve.

Suppose that the parametric equations describing the $g = 0$ curve are known, $x = x(t)$, $y = y(t)$, such that

$$g(x(t), y(t)) = 0 \qquad \forall t.$$

Differentiating this equation using the chain rule yields

$$0 = \frac{d}{dt}\big[g(x(t), y(t))\big] = \nabla g \cdot \frac{d\mathbf{x}}{dt} = 0. \tag{3.32}$$

If the parameter t is taken to represent time, the geometric interpretation of this equation is that the "velocity" vector $d\mathbf{x}/dt$ (which is tangent to the curve $\mathbf{x}(t)$) is perpendicular to the level-curve $g = 0$ since the gradient ∇g is orthogonal to contours of constant function-value.

Substituting $x(t)$, $y(t)$ into $f(x, y)$ yields a function of a single variable

$$F(t) = f(x(t), y(t))$$

that is parametrised along the curve $g = 0$. Finding the critical points of $F(t)$ with respect to t gives the critical points of f on g,

$$\frac{dF}{dt} = 0 \qquad \Longrightarrow \qquad \frac{dF}{dt} = \nabla f \cdot \frac{d\mathbf{x}}{dt} = 0. \tag{3.33}$$

Hence, the chain rule shows that at a critical point, ∇f is also orthogonal to the velocity vector. In two dimensions, this forces the two gradients to be a scalar multiple of each other and then we obtain

$$\nabla f = \lambda \nabla g$$

where λ (called the *Lagrange multiplier*) is another unknown that needs to be determined. In component form, the resulting set of equations of three variables in the three unknowns is

$$f_x(x_*, y_*) = \lambda g_x(x_*, y_*), \quad f_y(x_*, y_*) = \lambda g_y(x_*, y_*), \quad g(x_*, y_*) = 0.$$

A typical solution strategy for solving the equations is to first solve the $\nabla f = \lambda \nabla g$ equations to obtain $x(\lambda), y(\lambda)$, and then substitute these into $g(x(\lambda), y(\lambda)) = 0$ to obtain a final equation for the values of λ.

This approach generalises in a straightforward manner to higher dimensions and to multiple constraints ($\mathscr{L} \equiv f - \lambda g - \mu h - \cdots$) [24, 68]. We note that had the parametric equations for $x(t), y(t)$ been available, then solutions could have been obtained directly from (3.33). The Lagrange approach does not actually use these parametric equations, just the assumption of their existence, to reformulate the problem into a convenient form.

The above discussion motivates the introduction of the augmented (constrained) Lagrange objective function satisfying (3.31),

$$\mathscr{L}(x, y, \lambda) \equiv f(x, y) - \lambda g(x, y). \tag{3.34}$$

This is just a convenient form that reproduces the set of equations above as the critical point equations for $\mathscr{L}$ with respect to its three variables,

$$\nabla \mathscr{L} = \mathbf{0} \quad \leftrightarrow \quad \{ \mathscr{L}_x = 0, \quad \mathscr{L}_y = 0, \quad \mathscr{L}_\lambda = 0 \}. \tag{3.35}$$

3.6 Integral Constraints: Isoperimetric Problems

Consider problems where the solution $y_*(x)$ satisfies

$$\max_y \left(J \equiv \int_a^b L(x, y, y')\, dx \right) \quad \text{subject to} \quad G \equiv \int_a^b g(x, y, y')\, dx = 0. \tag{3.36}$$

These are sometimes called *isoperimetric problems* with the name referring back to a classic geometric problem of maximising the area enclosed by a curve constrained to have a fixed perimeter (see Exercise 3.20).

Maximising J subject to the constraint $G = 0$ can be expressed in terms of an augmented Lagrange functional of the form

$$I(y, \lambda) = J - \lambda G, \tag{3.37}$$

where λ is a constant. This can also be restated in terms of an augmented Lagrangian function

$$I = \int_a^b \mathscr{L}\, dx \quad \text{where} \quad \mathscr{L}(x, y, y', \lambda) = L(x, y, y') - \lambda g(x, y, y'). \tag{3.38}$$

Once this functional has been identified, we follow the standard variational process described in Sect. 3.2, starting with the introduction of perturbations to all of the unknowns

$$\tilde{y}(x) = y_*(x) + \varepsilon h(x), \qquad \tilde{\lambda} = \lambda_* + \varepsilon\gamma.$$

Expanding $\tilde{I}$ as a Taylor series for $\varepsilon \to 0$ and enforcing the critical point condition on the $O(\varepsilon)$ term yields

$$\int_a^b \left[\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) - \lambda_* \left\{ \frac{\partial g}{\partial y} - \frac{d}{dx}\left(\frac{\partial g}{\partial y'}\right) \right\} \right] h\,dx + \gamma \int_a^b g(x, y_*, y_*')\,dx = 0.$$

Requiring this to hold with respect to (i) all possible γ perturbations recovers the geometric constraint

$$\int_a^b g(x, y_*, y_*')\,dx = 0, \tag{3.39a}$$

and (ii) all $h(x)$ perturbations using the fundamental lemma yields the Euler–Lagrange equation

$$\frac{\partial \mathscr{L}}{\partial y} - \frac{d}{dx}\left(\frac{\partial \mathscr{L}}{\partial y'}\right) = 0. \tag{3.39b}$$

This again shows that the form of the Euler–Lagrange equation (3.21) generalises to a wide array of problems.

As an example, consider the problem of finding the function $y(x) > 0$ on $0 < x < 1$ satisfying boundary conditions $y(0) = 0$ and $y(1) = 0$, with a given arclength, $S = \int_0^1 \sqrt{1 + (y')^2}\,dx$ that maximises the area under the curve, $A = \int_0^1 y\,dx$.

We can write the augmented Lagrangian $\mathscr{L} = A - \lambda P$ to maximise the enclosed area subject to the constraint of a fixed perimeter, $P = S + 1$ (which includes the contribution from the lower boundary, $y \equiv 0$). Applying (3.39a, 3.39b), we obtain

$$1 + \lambda_* \frac{d}{dx}\left(\frac{y_*'}{\sqrt{1 + (y_*')^2}} \right) = 0, \qquad \int_0^1 \left(\sqrt{1 + (y_*')^2} - S \right) dx = 0.$$

Solving the ODE and imposing the boundary conditions, yields

$$y_*(x) = \sqrt{\lambda_*^2 - (x - \tfrac{1}{2})^2} - \sqrt{\lambda_*^2 - \tfrac{1}{4}}, \tag{3.40}$$

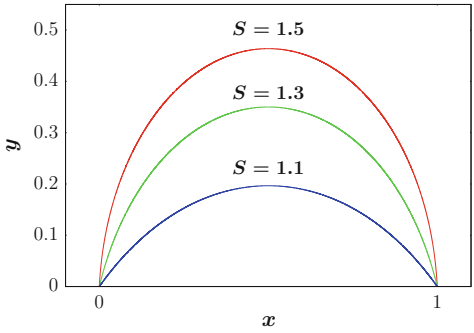


Fig. 3.4 The solution $y_*(x)$ (3.40) maximising the area under the curve on $0 \le x \le 1$ for given arclength S

which can be integrated to yield the resulting area and arclength for a given value of the Lagrange multiplier λ_*,

$$a(\lambda_*) \equiv \int_0^1 y_*(x)\,dx = \lambda_*^2 \arcsin\left(\frac{1}{2\lambda_*}\right) - \sqrt{\lambda_*^2 - \tfrac{1}{4}},$$

$$s(\lambda_*) \equiv \int_0^1 \sqrt{1 + y_*'(x)^2}\,dx = 2\lambda_* \arcsin\left(\frac{1}{2\lambda_*}\right).$$

To complete the solution of the constrained problem, the arclength constraint, $s(\lambda_*) = S$, must be applied to determine a value for λ_* in terms of S.

From (3.40) we see that acceptable (real-valued) solutions must have $\lambda_* \ge \frac{1}{2}$, which limits S in the problem statement to be in the range $1 \le S \le \pi/2$. Noting that (3.40) describes arcs of circles going through the fixed endpoints, with centres given by $(\frac{1}{2}, -\sqrt{\lambda_*^2 - \frac{1}{4}})$, with λ_* being the radius of the circle, see Fig. 3.4. The fact that (3.40) is not valid for $S > \pi/2$ suggests that the assumed form of the solutions (in this case, graphs of functions $y = y(x)$) may be too restrictive and different representations of the problem could be helpful, see Exercise 3.20.

3.7 Geometric Constraints: Holonomic Problems

Consider the problem of finding parametric equations for a curve $(x_*(t), y_*(t)$ on $0 \le t \le T)$ that minimises the integral

$$I = \int_0^T L(t, x, y, x', y')\,dt, \tag{3.41a}$$

subject to the constraint that for each value of t,

$$g(x(t), y(t), t) = 0. \tag{3.41b}$$

Problems of this form are called *holonomic problems* in the context of dynamics of mechanical systems. The constraint (3.41b) imposes a condition that the motion in the system must satisfy at each instant of time, as typically imposed by structural *geometric constraints* (as in the case of a roller coaster car moving on its track).

It can be shown that the appropriate generalisation of (3.34) for holonomic problems is the augmented Lagrangian

$$\mathscr{L}(x, y, \lambda) = L(t, x, y, x', y') - \lambda(t) g(x, y, t) \quad \Longrightarrow \quad \mathscr{I} = \int_0^T \mathscr{L}\, dt. \tag{3.42}$$

Here, a notable difference from (3.38) for isoperimetric problems is that the Lagrange multiplier is a function of t, rather than a constant, and cannot be factored out of the integral as we did in (3.37). Applying independent perturbations to each unknown,

$$\tilde{x}(t) = x_*(t) + \varepsilon h_1(t), \quad \tilde{y}(t) = y_*(t) + \varepsilon h_2(t), \quad \tilde{\lambda}(t) = \lambda_*(t) + \varepsilon \gamma(t),$$

and expanding $\mathscr{I}$ as $\varepsilon \to 0$, we obtain

$$\begin{aligned} \delta \mathscr{I}_* = \int_0^T \Bigg\{ &\left[\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial x'} \right) \right] h_1 + \left[\frac{\partial L}{\partial y} - \frac{d}{dt}\left(\frac{\partial L}{\partial y'} \right) \right] h_2 \\ &- \lambda_* \left(\frac{\partial g}{\partial x} h_1 + \frac{\partial g}{\partial y} h_2 \right) - g\gamma \Bigg\} dt, \end{aligned} \tag{3.43}$$

where the terms in the first line give the contributions from the original objective function and the terms in the second line represent the influence of the constraint. Noting that all perturbations are independently allowable, in enforcing the critical point condition, $\delta \mathscr{I}_* = 0$ (through linear independence, as in (3.24)) we conclude that the coefficient of each perturbation must vanish

$$\begin{aligned} \forall h_1 : &\quad x\text{Euler–Lagrange:} && \frac{\partial \mathscr{L}}{\partial x} - \frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial x'} \right) = 0, \\ \forall h_2 : &\quad y\text{Euler–Lagrange:} && \frac{\partial \mathscr{L}}{\partial y} - \frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial y'} \right) = 0, \\ \forall \gamma : &\quad \text{Geometric constraint:} && g(x_*(t), y_*(t), t) = 0, \end{aligned}$$

where we have combined the L, g terms from (3.43) to write the Euler–Lagrange equations in a more compact form using (3.42). In fact, this system can be condensed

further to be expressed as

$$u \text{ Euler–Lagrange:} \qquad \frac{\partial \mathscr{L}}{\partial u} - \frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial u'}\right) = 0 \qquad u = \{x, y, \lambda\},$$

where u denotes each unknown function of t in the system. In other contexts, systems containing both geometric constraints and differential equations are often also called *differential algebraic equations* (DAE).

3.8 Differential Equation Constraints: Optimal Control

For holonomic problems, we considered optimising a functional subject to constraints on the possible solution applied pointwise over the entire domain. We now extend the analysis to systems where the constraint is given by a differential equation. In such cases, we want to minimise

$$J = \int_0^T L(t, x(t), u(t))\, dt, \tag{3.44a}$$

subject to the constraint that at each t on $0 \leq t \leq T$,

$$\frac{dx}{dt} = f(t, x(t), u(t)), \tag{3.44b}$$

along with given conditions on $x(t)$

$$x(0) = x_0 \qquad x(T) = x_1. \tag{3.44c}$$

Problems of this form arise in *optimal control theory* which seeks to determine a *control function* $u(t)$ that allows the *state function* $x(t)$ to achieve a desired *target value* (or "final state") x_1 while also minimising the "cost" of the overall process in imposing the control on the evolution of the state. The cost will be represented by the integral (3.44a). The ending time T will also be considered as a degree of freedom and used to help reduce the cost and make the target achievable, as some systems might be able to run cheaply, if perhaps slowly.

Equation (3.44b) is called a *state equation*; it describes how the system evolves based on a dynamical model relevant to the problem (e.g. mechanics or chemical kinetics). In the absence of external forcing, described here by the control function $u(t)$, ($u \equiv 0$), then the natural (*uncontrolled*) dynamics starting from the initial condition $(3.44c)_1$ are given by the solution of

$$\frac{d\hat{x}}{dt} = f(\hat{x}, 0, t) \qquad \hat{x}(0) = x_0. \tag{3.45}$$

Under typical assumptions on the rate function f, this initial value problem has a well-defined solution $\hat{x}(t)$. However, this solution might not equal the target value x_1 at any value of $T \geq 0$, or might reach it only after an undesirably long time. Both of these situations would make the uncontrolled solution $\hat{x}$ unacceptable, and hence we need to consider modifying the evolution of the state from (3.45) by forcing the system with some control function $u(t)$ yielding (3.44b).

Appropriate control functions may allow the state to reach the target, but may also involve excessive input of energy, or could still take too long. Hence, the "optimality" in optimal control theory refers to minimising the cost involved in imposing the control. Two basic types of cost functionals that can be used, depending on whether system speed or minimisation of the energy of the control function is the priority, are given respectively by

$$J_{\text{speed}} = \int_0^T 1 \, dt = T, \qquad J_{\text{energy}} = \int_0^T u^2 \, dt. \tag{3.46}$$

Examples of systems that can be described in this form include the input control of chemicals in reaction systems to maintain a steady output, the design of controlled-release timed drug delivery, and car power-steering systems. In addition to uses in "engineered" systems, such models can also be applied to describe how biological systems adapt to their environments.

System (3.44) is a classic optimal control theory problem—its solution involves determining the evolution of the state $x(t)$, the control $u(t)$, and the optimal stopping time T. The imposition of the state equation at each time suggests forming an augmented Lagrangian with a time-dependent Lagrange multiplier, as in (3.42),

$$\mathscr{L}(x, x', u, \lambda) = L(t, x, u) - \lambda(t) \left(\frac{dx}{dt} - f(x, u, t) \right) \tag{3.47}$$

and corresponding augmented functional

$$I = \int_0^T \mathscr{L}(x, x', u, \lambda) \, dt. \tag{3.48}$$

Since the stopping time is also an unknown, we will also draw on the analysis carried out for the variable endpoint problem in Sect. 3.4.2.

Applying perturbations to all of the unknowns

$$\tilde{x} = x_* + \varepsilon h(t) \qquad \tilde{u} = u_* + \varepsilon v(t) \qquad \tilde{\lambda} = \lambda_* + \varepsilon \gamma(t) \qquad \tilde{T} = T_* + \varepsilon S,$$

the cost functional is then

$$\tilde{I} = \int_0^{T_* + \varepsilon S} \mathscr{L}(x_* + \varepsilon h, x_*' + \varepsilon h', u_* + \varepsilon v, \lambda_* + \varepsilon \gamma) \, dt. \tag{3.49}$$

Expanding $\tilde{I}$ as a Taylor series for $\varepsilon \to 0$ gives

$$\tilde{I} = I_* + \varepsilon \frac{dI}{d\varepsilon}\bigg|_{\varepsilon=0} + O(\varepsilon^2),$$

where, using Leibniz' rule (3.28), we obtain the first variation in the form

$$\int_0^{T_*} \left[\frac{\partial \mathscr{L}}{\partial x} h + \frac{\partial \mathscr{L}}{\partial x'} h' + \frac{\partial \mathscr{L}}{\partial u} v + \frac{\partial \mathscr{L}}{\partial \lambda} \gamma \right] dt + \mathscr{L}_*\Big|_{t=T_*} S.$$

Applying integration by parts to the h' term, the first variation becomes

$$= \int_0^{T_*} \left[\left\{ \frac{\partial \mathscr{L}}{\partial x} - \frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial x'} \right) \right\} h + \frac{\partial \mathscr{L}}{\partial u} v + \frac{\partial \mathscr{L}}{\partial \lambda} \gamma \right] dt + \left(\mathscr{L} S + \frac{\partial \mathscr{L}}{\partial x'} h \right)\Big|_{t=T_*}, \tag{3.50}$$

where we have used the initial condition $(3.44c)_1$ to obtain that the perturbation satisfies $h(0) = 0$. Recalling the form of the condition at the variable endpoint (3.26a), we write the target/final value condition, $\tilde{x}(\tilde{T}) = x_1$, as

$$x_*(T_* + \varepsilon S) + \varepsilon h(T_* + \varepsilon S) = x_1. \tag{3.51}$$

Expanding this equation in the limit $\varepsilon \to 0$, we recover $x_*(T_*) = x_1$ by matching leading terms, and at $O(\varepsilon)$ the perturbation function satisfies

$$h(T_*) = -x_*'(T_*) S. \tag{3.52}$$

Substituting this into the first variation (3.50), the boundary term reduces to

$$\left(\mathscr{L} - x' \frac{\partial \mathscr{L}}{\partial x'} \right)\Big|_{t=T_*} S.$$

This combination of terms is called the *Legendre transform* of the Lagrangian with respect to the state variable. It will be convenient to define a new function, the *Hamiltonian*,[8] from this combination,

$$\mathscr{H} \equiv \mathscr{L} - x' \frac{\partial \mathscr{L}}{\partial x'}. \tag{3.53}$$

[8] Note that in a different context, there is another definition of the Hamiltonian having $H = -\mathscr{H}$. Despite the difference in sign conventions, both are called Hamiltonians, see Exercise 3.7.

Since $\mathscr{L} = L - \lambda(x' - f)$, we find that the Hamiltonian is given by

$$\mathscr{H} = L + \lambda f, \tag{3.54}$$

and hence the boundary term becomes $\mathscr{H}|_{t=T_*} S$. Imposing the critical point condition $\delta I_* = 0$ for all independent perturbations ($\forall h$, $\forall v$, $\forall \gamma$, $\forall S$), we obtain the four equations:

$$\frac{\partial \mathscr{L}}{\partial x} - \frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial x'}\right) = 0, \qquad \frac{\partial \mathscr{L}}{\partial u} = 0, \qquad \frac{\partial \mathscr{L}}{\partial \lambda} = 0, \qquad \mathscr{H}(T_*) = 0. \tag{3.55}$$

The last equation provides a natural final condition on the Hamiltonian. Making use of the fact that it can be shown that if $\mathscr{L}$ does not explicitly (non-autonomously) depend on t, then the Hamiltonian is constant (see Exercise 3.7), the condition at T_* determines that $\mathscr{H}(t) \equiv 0$ for $0 \le t \le T_*$.

The third equation in (3.55) returns the state equation (3.44b). Similarly, the first equation yields an ordinary differential equation for the evolution of the Lagrange multiplier,

$$\left(\frac{\partial L}{\partial x} + \lambda_* \frac{\partial f}{\partial x}\right) + \frac{d\lambda_*}{dt} = 0, \tag{3.56}$$

which is called the *co-state equation*, with $\lambda_*(t)$ (responsible for imposing the state-equation constraint) being called the *co-state* in control theory.

Finally, the perturbation with respect to the control, $(3.55)_2$ gives a geometric constraint on the control at each time t on $0 \le t \le T_*$,

$$\frac{\partial L}{\partial u} + \lambda_* \frac{\partial f}{\partial u} = 0. \tag{3.57}$$

We apply the above derivation to a simple example of an optimal control problem. Consider the initial value problem for $x(t)$,

$$\frac{dx}{dt} = -3x + u \qquad x(0) = 2.$$

We immediately see that the uncontrolled system would have the exponentially decaying solution, $\hat{x}(t) = 2e^{-3t}$. As an illustrative problem where the influence of control is crucial, let the final target state for $T > 0$ be

$$x(T) = 5,$$

which is unattainable with the uncontrolled solution. Consider minimising a control-based cost functional, like J_{energy} from (3.46),

$$\min J = \tfrac{1}{2}\int_0^T u^2\,dt.$$

We can directly identify the Lagrangian function from J as $L = \frac{1}{2}u^2$ and the rate function from the state equation as $f = -3x + u$. Hence the augmented Lagrangian is

$$\mathscr{L} = \tfrac{1}{2}u^2 - \lambda(x' + 3x - u),$$

and the Hamiltonian is

$$\mathscr{H} = \tfrac{1}{2}u^2 + \lambda(u - 3x).$$

From the Euler–Lagrange equations (3.55), we then find

$$-3\lambda + \frac{d\lambda}{dt} = 0, \qquad u + \lambda = 0, \qquad \frac{dx}{dt} + 3x - u = 0.$$

Starting with the algebraic relation, we can eliminate the control in terms of the co-state, $u = -\lambda$. The co-state equation for λ can then be re-expressed for u as

$$\frac{du}{dt} = 3u \qquad \Longrightarrow \qquad u(t) = Ae^{3t},$$

where A is a constant of integration. With this expression for $u(t)$, the state equation becomes

$$\frac{dx}{dt} = -3x + Ae^{3t} \qquad \Rightarrow \qquad x(t) = Be^{-3t} + \tfrac{1}{6}Ae^{3t},$$

where B is a second constant of integration. Substituting x, u and λ into the Hamiltonian yields

$$\mathscr{H} = 3AB = 0,$$

so that either A or B is zero. The choice $A = 0$ returns the uncontrolled state, and so we conclude that $B = 0$ and hence $x(t) = 2e^{3t}$. Illustrating the dramatic influence of the control, the optimal controlled solution is exponentially growing, whereas the uncontrolled solution is exponentially decaying. The optimal stopping time can then be obtained directly from the optimal solution for this simple problem as $T_* = \frac{1}{3}\ln(5/2)$.

3.9 Further Directions

There is an extensive literature on both the theory and the applications of the calculus of variations. Some good texts with additional introductory material and other advanced topics include [64, 66, 71, 84, 104]. Further texts also provide more rigorous analysis [99]. Presentations of the applications of the calculus of variations in

mechanics [40, 62, 67] and other areas of applied physics, including electrostatics and quantum mechanics, [104] leading to ODE and PDE problems can also be found in many classic engineering and physics textbooks. There are also numerous books that present constrained optimisation and optimal control problems as extensions of the basic method of the calculus of variations [59, 62, 66, 71, 84, 99].

While the introductory presentation in this chapter as focused on formulating problems leading to ODEs, some of the exercises will show that PDEs can be derived similarly from multiple integrals. More challenging problems arise when assumptions on the smoothness of solutions are removed, then optimal solutions may be composed of multiple piecewise-smooth sections that must satisfy appropriate connection conditions at transition points.

3.10 Exercises

3.1 Consider the functional for $y(x)$ on $0 \le x \le 1$,

$$J = \int_0^1 \tfrac{1}{2}(y')^2 + \frac{k}{x} yy' + x^2 y \, dx,$$

where k is a constant

(a) Determine the ODE for $y_*(x)$ by taking the variation of J.
(b) Show that the ODE can be also obtained directly from the Euler–Lagrange equation (3.21).
(c) For $k = 0$, determine $y_*(x)$ satisfying the boundary conditions $y(0) = 1$, $y(1) = 1$.
(d) For $k = 2$, determine $y_*(x)$ satisfying the boundary conditions $y'(0) = 1$, $y(1) = 1$. Show there is no solution that satisfies $y(0) = 1$.

3.2 (*The second derivative test*) As in single- and multi-variable calculus, whether a critical point is a local maximum or minimum can be determined from the next term in the local expansion of the objective function. Recall the problem of minimising the arclength of the curve $y = y(x)$, (3.10a),

$$J = \int_0^1 \sqrt{1 + (y')^2} \, dx, \qquad y(0) = 0, \qquad y(1) = b.$$

Setting $y(x) = y_*(x) + \varepsilon h(x)$ and expanding J to $O(\varepsilon)$ yielded the Euler–Lagrange equation.

(a) Continue the expansion of J to the $O(\varepsilon^2)$ term to obtain the *second variation*, $\delta^2 J$.
(b) Show that for this problem the solution $y_*(x) = bx$ is indeed a local minimiser by showing that the second variation is positive for all $h(x) \not\equiv 0$.

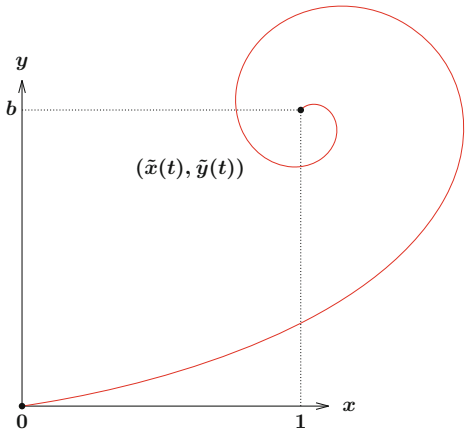


Fig. 3.5 Exercise 3.3: An admissible trial solution in the form of a parametric curve

3.3 Consider finding the shortest path between two points from among all possible smooth parametric curves in the xy plane $\mathbf{x}(t) = (x(t), y(t))$ for $0 \le t \le T$. Let the curve start at $(x, y) = (0, 0)$ at $t = 0$ and end at $(x, y) = (1, b)$ at $t = T$ for T fixed (see Fig. 3.5). The functional for the arclength of the curve is

$$J(x, y) = \int_0^T \sqrt{(x')^2 + (y')^2}\, dt.$$

Show that the Euler–Lagrange equations for x, y give the parametric equations of the straight line found as the solution of (3.10a, 3.10b).

3.4 For problems where the distance between two points is not given by the Euclidean distance, *geodesics* are curves giving the shortest path between two points. Consider the problem of determining a geodesic between two points on a surface $z = F(x, y)$.

(a) For a parametric curve $(x(t), y(t), z(t))$ that minimises the arclength,

$$J(x, y) = \int \sqrt{(x')^2 + (y')^2 + (z')^2}\, dt,$$

observe that the resulting Euler–Lagrange equations for $(x(t), y(t))$ are rather complicated, even if the surface is the simple paraboloid, $z = k(x^2 + y^2)$.

(b) Show that even if we restrict the solutions to be functions, with $y = y(x)$, the Euler–Lagrange equation for $y(x)$ on the paraboloid is still a challenging nonlinear ODE for $k \neq 0$.

(c) Consider the problem of finding the geodesic from $(x, y) = (-1, 0)$ to $(1, 0)$ on the paraboloid. Determine the arclengths of the following trial solutions for $k \ge 0$ (from geometry or by calculating the integral):

(i) the path along the x-axis,
(ii) the semicircular path connecting the points.
What can you infer about the geodesic from the limits $k \to 0$ and $k \to \infty$ of (i, ii)?

3.5 The Euler–Lagrange equations for constrained motion can be obtained by starting with the Lagrangian for general unconstrained motion and then substituting-in the parametric equations describing the geometric constraint (a curve or surface) before taking variations.

(a) Consider the action integral for two-dimensional motion of a particle subject to gravity

$$I = \int \tfrac{1}{2}m\left[x'(t)^2 + y'(t)^2\right] - mgy(t)\,dt.$$

Consider a particle constrained to be on a circle, $x^2+y^2 = \ell^2$. Derive the equation of a pendulum by first plugging the parametric equations $x(t) = \ell \sin\theta(t)$, $y(t) = -\ell\cos\theta(t)$ into I and then applying the principle of least action with respect to $\theta(t)$.

(b) Repeat (a) for the motion of a particle constrained to the curve $y = f(x)$.

(c) Consider the action integral for three-dimensional motion of a particle subject to gravity

$$I = \int \tfrac{1}{2}m\left[x'(t)^2 + y'(t)^2 + z'(t)^2\right] - mgz(t)\,dt$$

Derive the equations of motion for a particle moving on the surface of a cone, $x^2 + y^2 = z^2$, using the parametric equations $x(t) = r(t)\cos\theta(t)$, $y(t) = r(t)\sin\theta(t)$ and $z(t) = r(t)$.

3.6 Consider a pendulum whose suspension point is vertically oscillated. Let the length of the pendulum be ℓ and its mass m. The acceleration due to gravity, g, is acting downward in the $-y$ direction. If the suspension point is oscillating according to $\bar{y}(t) = -\sigma\sin(\omega t)$, then the position of the pendulum's mass is

$$x(t) = \ell\sin\theta(t), \qquad y(t) = -\sigma\sin(\omega t) - \ell\cos\theta(t),$$

where θ is measured from the $-y$ axis.

Write the Lagrangian for this system in terms of θ, θ' and obtain the Euler–Lagrange equation for this problem, called the *parametrically driven pendulum*.

3.7 Define the Hamiltonian, H, in terms of the Lagrangian $L(t, y, y')$, through the Legendre transform as

$$H \equiv -L + y'\frac{\partial L}{\partial y'}. \tag{3.58}$$

(Note the opposite sign convention relative to (3.53).)

(a) If L is independent of time (i.e. $L = L(y, y')$) show that the Hamiltonian is constant. (Hint: show that $dH/dt = 0$)
Further, show that for this case, the second-order Euler–Lagrange ODE (3.21) is equivalent to the first order ODE,

$$-L(y, y') + y' \frac{\partial L}{\partial y'} = C, \tag{3.59}$$

where C is a constant ($=H$); this reduction of the Euler–Lagrange equation to a first order equation is called the *Beltrami identity*, and is a special case of *Noether's theorem* [62].

(b) If the potential energy V does not depend on y', determine the most general form for the kinetic energy $T(t, y, y')$ so that the Hamiltonian is the total energy, $H = T + V$.

3.8 (*The brachistochrone*) Recall the problem of finding the curve of least-time descent under gravity from $(x, y) = (0, 1) \to (1, 0)$, starting from rest (Fig. 3.1 and Eq. (3.6)). We now complete the problem as follows:

(a) Assuming there is no friction, the total mechanical energy, $E = \frac{1}{2}mv^2 + mgy$, (where $v^2 = x'(t)^2 + y'(t)^2$) remains constant, set by its initial value. Use this to determine $v(y)$ and complete the expression of the functional in (3.6).
(b) Apply (3.21) to obtain the Euler–Lagrange second order ODE for $y_*(x)$.
(c) Noting that the integrand in (3.6) does not explicitly depend on x, use the Beltrami identity (3.59) to obtain a simpler first order ODE for $y_*(x)$. What is the value of the constant C in this ODE?
(d) Show that the solution can be expressed in parametric form by the equations of a cycloid with $k > 0$ (see Fig. 3.6),

$$x(\theta) = k(\theta - \sin\theta), \qquad y(\theta) = 1 - k(1 - \cos\theta). \tag{3.60}$$

What is the pair of equations that must be solved numerically to determine the value of k?

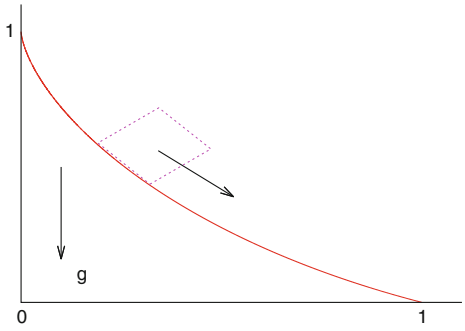


Fig. 3.6 Exercise 3.8: The solution of the brachistochrone problem, given by the cycloid (3.60)

3.9 Consider the functional for $y(x)$ on $1 \le x \le 2$

$$J(y) = \frac{1}{2}\int_1^2 \left[2y^2 + x^2\left(\frac{dy}{dx}\right)^2\right] dx.$$

Find the solution $y_*(x)$ that minimises J and satisfies the boundary condition $y(2) = 17$. What is the natural boundary condition on y_* at $x = 1$?

3.10 Obtain the curve $y = y_*(x)$ on $0 \le x \le b_*$ that starts at the origin, ends on the curve $y = 1 + (x-1)^2$, and minimises the functional

$$J(y) = \frac{1}{2}\int_0^b (y')^2\, dx.$$

3.11 (*Higher-order Euler–Lagrange equations*) Derive the Euler–Lagrange equation for the functional

$$J(y) = \int_a^b L(x, y(x), y''(x))\, dx$$

Describe the kinds of natural boundary conditions that are needed.

3.12 Consider the functional for $y(x)$ on $0 \le x \le 1$,

$$J = \int_0^1 \left[\frac{dy}{dx}\frac{d^3y}{dx^3} - 240xy\right] dx$$

where $y(x)$ satisfies the boundary conditions

$$y'(0) = 0, \qquad y''(1) = 0, \qquad y'''(0) = 0, \qquad y(1) = 5.$$

(a) Write the expression for the first variation, δJ.
(b) Show that the critical point condition can be reduced to an ODE boundary value problem. Justify how each of the six boundary terms in δJ are eliminated. Solve the ODE problem for $y_*(x)$.

3.13 Consider the functional for $y(x)$ on $0 \le x \le 1$,

$$J = \int_0^1 \left[(y')^2 + (1-2x)\left(\int_0^x y^2(t)\, dt\right)\right] dx.$$

(a) Write the expression for the first variation, δJ.
(b) Write the four possible boundary conditions on $y(x)$ (two at each boundary) under which, the critical point condition, $\delta J = 0$, can be reduced to an ODE for $y(x)$.

3.14 Consider the functional for $y(x)$ on $0 \le x \le 1$,

$$J(y) = \int_0^1 \left[y \left(\frac{d^2y}{dx^2} \right)^2 + y^2 \right] dx + y'(0)y'(1)$$

with $y(x)$ satisfying the boundary conditions

$$y(0) = 2, \qquad y(1) = 5.$$

Determine the ODE boundary value problem for solutions that minimise or maximise J.

3.15 *Fermat's principle of least time* states that a beam of light will take a path that minimises its time of travel. The index of refraction, n, of a material gives the ratio of the speed of light in vacuum to the slowed speed of light in the material, $n = c/v \ge 1$.

Consider a layer of glass on $0 \le x \le 1$ whose index of refraction varies with position, $n = n(x)$. Let $y(x)$ describe the path of a beam of light entering the layer at $x = 0$ at a 45° angle, $y'(0) = 1$ (see Fig. 3.7 (left)).

(a) Making use of functional (3.6), derive the second-order ODE problem for $y(x)$ and show it can be reduced to a first-order equation.
(b) If $n(x)$ is piecewise constant, n_1 for $x < \frac{1}{2}$ and n_2 for $x > \frac{1}{2}$, show that part (a) reduces to *Snell's law* [91] (see Fig. 3.7 (right)).
(c) What is the form of the Euler–Lagrange equation if $n = n(x, y)$?

3.16 (*The beam equation*) Hamilton's principle of least action can be applied to derive the time-dependent partial differential equation for the transverse deflections $y = u(x, t)$ of a flexible solid rod or beam. The problem for a one-dimensional beam is characterised by the following properties:

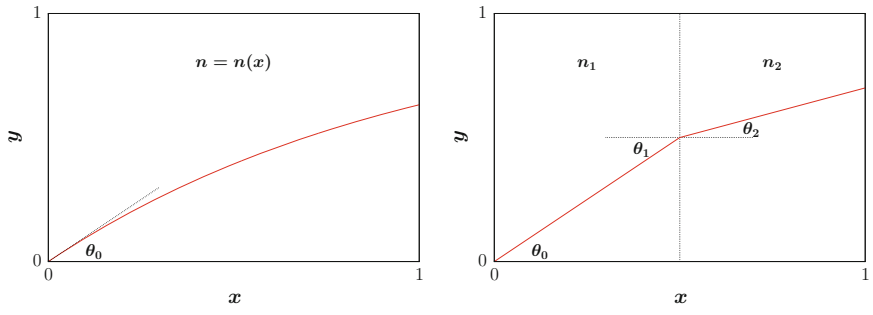


Fig. 3.7 Exercise 3.15: (*Left*) diffraction of the path of a ray of light due the change of the index of refraction with position, $n = n(x)$, (*Right*) The special case when n is piecewise constant

- The length of the beam is ℓ (domain: $0 \le x \le \ell$) with constant mass density ρ (mass per unit length), constant bending stiffness E (like a spring constant) and constant moment of inertia I.
- The overall total kinetic and potential energies of the beam are given by integrals over the length of the beam of the corresponding energy density functions, $e(u)$, $\mathscr{E} = \int_0^\ell e(u)\,dx$.
- The kinetic energy density is $\frac{1}{2}\rho(\partial_t u)^2$.
- The potential energy density due to bending (i.e. transverse deflection or buckling) is $\frac{1}{2}EI(\partial_{xx} u)^2$.

Follow these steps to derive the beam equation:

(a) Write the integrals for the total kinetic T and potential V energies of the beam.
(b) Write the action integral $J(u)$. (Hint: J is a double integral)
(c) Apply the principle of least action to derive the beam equation for $u(x, t)$ and state the choices of natural boundary conditions for u at $x = 0$ and $x = \ell$. Two boundary conditions (out of 2 sets of 2 options for natural boundary conditions) are needed at each boundary.

3.17 For a given positive function $k(x, y) > 0$, consider the integral for the function $u(x, y)$ on a finite two-dimensional domain D,

$$J(u) = \iint_D \tfrac{1}{2}k(x, y)\left(u_x^2 + u_y^2\right)\,dy\,dx.$$

Derive the partial differential equation for the solution $u(x, y)$ and two appropriate forms of natural boundary conditions on $u(x, y)$ that produce minima of $J(u)$.
If k is a constant, show that the PDE reduces to Laplace's equation.
Hint: Recall the vector version of the derivative product (for the divergence of a scalar times a vector), $\nabla \cdot (f\mathbf{g}) = (\nabla f) \cdot \mathbf{g} + f(\nabla \cdot \mathbf{g})$, and then make use of the divergence theorem.

3.18 How does the problem of minimising the arclength of a non-negative function $y(x) \ge 0$ on $0 \le x \le 1$ that encloses a fixed given area $A = \int_0^1 y\,dx$ relate to the isoperimetric example given in Sect. 3.6?

3.19 Find the solution $y(x)$ on $0 \le x \le \pi$ that minimises

$$J(x) = \int_0^\pi 1 + (y')^2\,dx \quad \text{subject to the constraint} \quad \int_0^\pi y^2\,dx = 80$$

and boundary conditions $y(0) = 0$ and $y(\pi) = 1$.

(a) Write the augmented Lagrangian $\mathscr{L}$ for this problem. Determine the Euler–Lagrange problem and solve for $y(x, \lambda)$ subject to the boundary conditions, where λ is the Lagrange multiplier.

(b) Evaluate the integral, $I(\lambda) = \int y^2\,dx$, as a function of λ. Plot this function and estimate the values of λ that yield solutions. If the integral constraint is required to equal one instead of 80, what happens to the solutions? What is the minimum value of I for which solutions exist?

3.20 (*The classic isoperimetric problem*) This problem will lead you through deriving that the circle is the closed curve of fixed perimeter P that encloses the maximum area:
Let $x(t)$, $y(t)$ be the parametric equations for a closed curve with $0 \le t \le 1$ that goes through the origin:

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

(a) Use Green's theorem to show that the area enclosed by the curve is given by

$$\frac{1}{2}\int_0^1 \left[x(t)y'(t) - y(t)x'(t)\right]\,dt.$$

Show that the augmented objective function

$$\mathscr{L}(x, y, \lambda) = \tfrac{1}{2}(xy' - yx') - \lambda\left[\sqrt{(x')^2 + (y')^2} - P\right],$$

with $J = \int_0^1 \mathscr{L}\,dt$, defines the problem of maximising the area for a closed curve with perimeter P.

(b) Obtain the Euler–Lagrange equations for $x(t)$ and $y(t)$.
(c) Integrate each once with respect to t and show that they can be combined to yield the equation for a circle with the radius given in terms of $|\lambda|$.
(d) Determine λ so that the perimeter constraint is satisfied. Determine the possible positions for the centre of the circle.

3.21 (*Sturm-Liouville eigenvalue problems*) For given functions $p(x) > 0$ and $q(x)$, show that the functional for $y(x)$

$$I = \int_0^1 p(x)(y')^2 - q(x)y^2\,dx$$

yields the Euler–Lagrange equation

$$\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + q(x)y = 0.$$

For homogeneous boundary conditions on $y(x)$, the trivial solution $y(x) \equiv 0$, is a critical point for all p, q. Show that seeking nontrivial critical points of I satisfying the normalisation condition,

$$\int_0^1 y^2\sigma(x)\,dx = 1,$$

for a given weight function $\sigma(x) > 0$, yields the Sturm-Liouville equation

$$\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + (q(x) + \lambda\sigma(x))y = 0,$$

where λ are eigenvalues and $y(x)$ are the corresponding eigenfunctions.

3.22 (*The pendulum revisited*) Consider again, the action integral for motion of a point mass in two dimensions subject to gravity (recall Exercise 3.5(a)),

$$I = \int \tfrac{1}{2}m\left[x'(t)^2 + y'(t)^2\right] - mgy(t)\,dt.$$

Suppose that the mass is constrained to move on the circle, $x^2 + y^2 = \ell^2$. Consider this as a problem with a holonomic constraint. Write the Euler–Lagrange equations for $x(t)$, $y(t)$. Note that the Lagrange multiplier appearing in these equations can be interpreted as a force required to keep the mass on the circle. Show that the equations can be reduced to the pendulum equation for $\theta(t)$.

3.23 Consider the problem of finding a locally optimal solution of the functional

$$J = \int_1^2 \left[6y^2 + x^2\left(\frac{dy}{dx}\right)^2 + x^7\right] dx$$

subject to the conditions that

$$y(2) = y(1) + 3, \qquad \int_1^2 24xy\,dx = 5.$$

(a) Write the augmented functional for the constrained optimisation problem.
(b) Determine the natural boundary condition that allows the critical point condition to be reduced to an ODE problem.
(c) Write the general solution of the ODE (homogeneous and particular terms) and the system of equations to determine the three constants in your solution.

3.24 Find the locally optimum solution of the functional

$$J = \int_1^2 \left[\frac{3y^2}{x^5} - \frac{(y')^2}{x^3}\right] dx$$

subject to the conditions that

$$y(1) = 4, \qquad y(2) = -10, \qquad \int_1^2 y\,dx = -3.$$

3.25 The *Pontryagin Maximum Principle* (PMP)[9] is a classic result from optimal control theory. For the classic optimal control problem (3.44) the PMP states that the equations for the optimal solution can be very concisely given in terms of the Hamiltonian, $\mathscr{H} = L + \lambda f$,

$$\frac{dx}{dt} = \frac{\partial \mathscr{H}}{\partial \lambda}, \qquad \frac{d\lambda}{dt} = -\frac{\partial \mathscr{H}}{\partial x}, \qquad \frac{\partial \mathscr{H}}{\partial u} = 0,$$

Subject to $\mathscr{H}(T_*) = 0$. Show that these equations reproduce (3.55) and these equations are consistent with the Hamiltonian being a constant for all times. (Hint: Apply the chain rule to evaluate $d\mathscr{H}/dt$)

3.26 Determine the solution $x(t)$ and the control function $u(t)$ that satisfy the state equation

$$\frac{dx}{dt} = 3x + u \qquad 0 \le t \le T$$

with initial and final conditions

$$x(0) = 2, \qquad x(T) = 1,$$

while minimising the cost functional

$$J = \int_0^T \left(4x^2 + 3xu + u^2\right) dt.$$

(a) Use the Pontryagin principle from Exercise 3.25 for the case where the final time is the optimal stopping time $T = T_*$.
(b) If instead the final time is specified as $T = 1/4$, what is the optimal solution? What is the value of the Hamiltonian?

3.27 (*The brachistochrone revisited*) Recall Exercise 3.5(b), where we determined that the equation of motion for the horizontal position $x(t)$ of a particle sliding down a ramp $y = f(x)$ under the influence of gravity is

$$\frac{d}{dt}\left([1 + f'(x)^2]\frac{dx}{dt}\right) = -gf'(x) + f'(x)f''(x)\left(\frac{dx}{dt}\right)^2.$$

Use this result applied to functions satisfying the boundary conditions $f(0) = 1$ and $f(1) = 0$ to show that the brachistochrone problem can be expressed as an optimal control problem on f subject to minimising travel time T for a particle satisfying $x(0) = 0$ and $x(T) = 1$. Write a Lagrangian analogous to (3.47) and the corresponding functional to determine the Euler–Lagrange problem.

[9] Sometimes called the Pontryagin *minimum* principle, depending on the choice of sign convention used for H versus $\mathscr{H}$, recall page 71.

Chapter 4
Dimensional Scaling Analysis

"Back of the envelope calculations" are simple algebraic calculations that can fit into a small space but provide substantial guidance on understanding problems. Such calculations are often based on *dimensional analysis*, which identifies the relations between system properties that must be present in a problem due to the fundamental types of the properties being considered. It is valuable to understand that insight on how a system can be expected to behave can be obtained from interpretation of appropriate products of the parameters characterising the problem.

One simple type of dimensional analysis result, sometimes called a *Fermi estimate*, provides predictions based on products of dimensional rates, as in the number of widgets produced in a year = (number of widgets produced per day) × (number of work-days per week) × (average number of weeks per month) × (months per year). This result does not take account of the details of the problem, but gives an *order of magnitude estimate* of the expected number. Dimensional analysis can not replace detailed solutions of problems, but it can provide a guide for comparison with expectations and identification of important influences in problems.

In this chapter, our focus will be on the use of dimensional analysis to *nondimensionalise* entire problems in order to provide simplified re-statements of the problems that identify a reduced (minimal) number of essential parameters that solutions should depend on. We will illustrate the use of dimensional analysis through its application to a set of examples before discussing the Buckingham Pi theorem, which gives a mathematical framework to support the results of the elementary scaling calculations.

4.1 Dimensional Quantities

Every physical quantity Q can be expressed as the product of a *dimensional unit*, denoted here by $[\mathsf{Q}]$ (describing the physical nature of a single 'portion' or unit of the quantity), and a *magnitude* q (the number of portions), i.e. $\mathsf{Q} = q[\mathsf{Q}]$. For example, mass can be measured in dimensional units of kilograms, weight in pounds, volume in litres, data in gigabytes, frequency in megahertz, length in meters, time in seconds,

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9_4

temperature in degrees Kelvin, and so on. These choices are not unique: $\mathsf{Q} = 1\,\text{mile} = 1{,}760\,\text{yd} = 5{,}280\,\text{ft} = 63{,}360\,\text{in} = 160{,}934.4\,\text{cm} = 1.7011 \times 10^{-13}$ light years; all of these are equivalent descriptions of the same length with respect to different units of measure and different corresponding magnitudes.

The choice of dimensional units for measurements should be appropriate to the problem under consideration. For example, it would be very inconvenient to measure interplanetary distances in feet (the magnitudes involved being literally astronomically large) or to record the size of a bacterial swarm in kilometers. Having to deal with very large (or very small) numbers is not only cumbersome, but the analysis in some later chapters will crucially rely on separating small from large quantities and a bad choice of measurement scales can obscure this process.

4.1.1 The SI System of Base Units

In many everyday physical problems, it is reasonable to employ dimensional units of meters for displacement ([Length] = meters), kilograms for mass ([Mass] = kilograms), and seconds for time ([Time] = seconds). This is a standard set of units[1] forming part of what is known as the International System (SI) of Base Units used for quantifying distinct physical properties. The seven fundamental (base) dimensional units that make up the SI system are [97]

$$\begin{aligned} &[\text{Length}] = \text{meter}, \quad [\text{Time}] = \text{second}, \quad [\text{Mass}] = \text{kilogram}, \\ &[\text{Temperature}] = \text{Kelvin}, \quad [\text{Electric current}] = \text{ampere}, \\ &[\text{Light intensity}] = \text{candela}, \quad [\text{Material quantity}] = \text{mole}. \end{aligned} \tag{4.1}$$

These dimensional units are *fundamental* in the sense that they describe independent physical properties that cannot be represented in terms of each other. Dimensional units that can be expressed in terms of the fundamental units are known as *derived units*. For example, [speed] = meter/second, [acceleration] = meter/second^2, and [force] = kilogram·meter/second^2. The most commonly occurring derived units are given their own names for convenience, such as [force] = newton, [pressure] = pascal, and [energy] = joule. For some problems, it can be convenient to express quantities with respect to different sets of base units, those of easily-measurable properties, for example, in terms of the pressure of a gas, [force] = [pressure][area].

Derived dimensional units are products of powers the fundamental units

$$[\mathsf{Q}] = [\text{Length}]^{\alpha}[\text{Time}]^{\beta}[\text{Mass}]^{\gamma} \ldots .$$

This result also extends to products of quantities, $[\mathsf{Q}_1^2\mathsf{Q}_2] = [\mathsf{Q}_1]^2[\mathsf{Q}_2]$.

[1] Also called the MKS system from Meters, Kilograms, Seconds.

In what follows, in order to emphasise that it is the nature of a quantity (length, mass, time and so on) that we are interested in rather than its representation in a particular system of base units, we will simply talk about the quantity's *dimensions* instead of explicitly referring to *dimensional units*.

4.2 Dimensional Homogeneity

The *principle of dimensional homogeneity* states the intuitive result that all terms summed to yield an equation modelling a system must have the same dimensions. For example, sums of forces in Newton's second law, or sums of electrical currents in Kirichoff's law, or sums and differences of amounts of money in accounting calculations. In contrast, it not valid to add a velocity to an acceleration, or a time to a mass.

An immediate consequence is that any complicated (non-monomial) function, such as sin, exp, tan, log, and generally any $f(\mathsf{X}) \neq \alpha \mathsf{X}^{\beta}$, must have dependent-arguments (variables) that are *dimensionless*. Namely the dimensional units in the variables must cancel-out identically. To see why this must be true, consider writing the exponential function in terms of its Taylor series expansion

$$e^{\mathsf{X}} = 1 + \mathsf{X} + \tfrac{1}{2}\mathsf{X}^2 + \cdots + \frac{1}{n!}\mathsf{X}^n + \cdots .$$

If X was not dimensionless, then each term in the Taylor expansion would have different dimensions thereby violating the principle of dimensional homogeneity; here all terms must have the same units as the first term in the expansion, [1] which is dimensionless. We note that variables representing angles are dimensionless, as in the formula for the arclength of circle, $\mathsf{S} = \mathsf{R}\theta$, [length] = [length][dimensionless].

As an illustrative example of dimensional homogeneity, consider the motion of a vertically launched projectile. If we denote Y as the projectile's height (with initial height Y_0), T as the time passed, V_0 as the initial velocity, and A_0 as a constant acceleration (typically gravity), the equation of motion for the projectile is simply a quadratic form in T

$$\mathsf{Y} = \mathsf{Y}_0 + \mathsf{V}_0\mathsf{T} + \tfrac{1}{2}\mathsf{A}_0\mathsf{T}^2.$$

Expressed in terms of its dimensions, we see that each term has the same dimensions, namely

$$[\text{Length}] = [\text{Length}] + \left[\frac{\text{Length}}{\text{Time}}\right][\text{Time}] + \left[\frac{\text{Length}}{\text{Time}^2}\right][\text{Time}^2].$$

4.3 The Process of Nondimensionalisation

The importance of dimensional homogeneity lies in the realisation that as each term in an equation must have the same dimensions, we can scale all dependent and independent variables by dimensional constants to yield *dimensionless equations*. Effectively, this means that the dimensional units have been factored out of the original problem, leaving a "clean" mathematical system in terms of only dimensionless magnitudes and *dimensionless parameters*.

The dimensions of the scaling constants are predetermined by dimensional homogeneity, but we have the freedom to select their magnitudes in order to obtain 'convenient forms' of the nondimensional model that will be better-suited for further mathematical analysis.

In the following examples, we will introduce standard conventions on how to select dimensionless parameters as well as highlighting further consequences of dimensional analysis.

4.3.1 Projectile Motion

Consider a projectile of mass M kilograms that is launched vertically with initial speed V_0 m/s, from a position Y_0 meters above the surface of the Earth. Newton's universal law of gravitation coupled with the second law of motion then gives that the height of the projectile $\mathsf{Y}(\mathsf{T})$ varies with time T according to the ordinary differential equation

$$\mathsf{M}\frac{d^2\mathsf{Y}}{d\mathsf{T}^2} = -\frac{\mathsf{G}\mathsf{M}_\mathsf{E}\mathsf{M}}{(\mathsf{R}_\mathsf{E}+\mathsf{Y})^2}, \qquad \mathsf{Y}(0) = \mathsf{Y}_0, \qquad \mathsf{Y}'(0) = \mathsf{V}_0, \tag{4.2}$$

where the Earth's properties (mass and radius) are $\mathsf{M}_\mathsf{E} = 6 \times 10^{24}\,\mathrm{kg}$ and $\mathsf{R}_\mathsf{E} = 6.4 \times 10^6\,\mathrm{m}$, and the universal gravitational constant $\mathsf{G} = 6.7 \times 10^{-11}\,\mathrm{m}^3/(\mathrm{s}^2\,\mathrm{kg})$. We begin by noting that the parameters G and M_E only occur as a product and so can only appear in the solution in the same form. Consequently, we can regard this combination as a single parameter and replace it using $\mathsf{g} = \mathsf{G}\mathsf{M}_\mathsf{E}/\mathsf{R}_\mathsf{E}^2 \approx 9.81\,\mathrm{m/s}^2$, this being the familiar value for acceleration due to gravity at the Earth's surface. The problem then takes the form

$$\mathsf{M}\frac{d^2\mathsf{Y}}{d\mathsf{T}^2} = -\frac{\mathsf{g}\mathsf{R}_\mathsf{E}^2\mathsf{M}}{(\mathsf{R}_\mathsf{E}+\mathsf{Y})^2}, \qquad \mathsf{Y}(0) = \mathsf{Y}_0, \qquad \mathsf{Y}'(0) = \mathsf{V}_0. \tag{4.3}$$

We now introduce the (as yet arbitrary) length and time scales, L and T,

$$\mathsf{Y}(\mathsf{T}) = \mathrm{L}y(t), \qquad \mathsf{T} = \mathrm{T}t, \tag{4.4}$$

so that y and t are dimensionless variables. The scaling constants (L and T) for the different independent dimensional quantities in a problem (here, length and time) are called the *characteristic scales*.

Substituting into (4.3) yields the rescaled problem

$$\frac{\mathrm{L}}{\mathrm{T}^2}\frac{d^2y}{dt^2} = -\mathsf{g}\frac{1}{\left(1+\left(\frac{\mathrm{L}}{\mathsf{R_E}}\right)y\right)^2}, \quad \mathrm{L}y(0) = \mathsf{Y}_0, \quad \frac{\mathrm{L}}{\mathrm{T}}y'(0) = \mathsf{V}_0.$$

Dividing through each expression by the dimensional coefficient of the first term on the left-hand-side yields the *nondimensionalised problem*

$$\frac{d^2y}{dt^2} = -\underbrace{\left(\frac{\mathsf{g}\mathrm{T}^2}{\mathrm{L}}\right)}_{\Pi_1}\frac{1}{\left(1+\underbrace{\left(\frac{\mathrm{L}}{\mathsf{R_E}}\right)}_{\Pi_2}y\right)^2}, \quad y(0) = \underbrace{\left(\frac{\mathsf{Y}_0}{\mathrm{L}}\right)}_{\Pi_3}, \quad y'(0) = \underbrace{\left(\frac{\mathsf{V}_0\mathrm{T}}{\mathrm{L}}\right)}_{\Pi_4}, \tag{4.5}$$

where the Π_i are nondimensional constants. It is important to realise that such dimensionless parameters always represent the ratio of two competing effects. For example, Π_1 represents the relative importance of gravitational acceleration versus the acceleration of the projectile, $\Pi_1 = \mathsf{g}/(\mathrm{L}/\mathrm{T}^2)$. Likewise, Π_2 compares the length scale of interest to the Earth's radius, $\Pi_2 = \mathrm{L}/\mathsf{R_E}$.

We have the freedom to choose the length and time scales L and T to scale the problem into a desired form; here we describe one such form. Selecting the length scale $\mathrm{L} = \mathsf{Y}_0$ sets $\Pi_3 = 1$, while choosing a timescale based on acceleration due to gravity, $\mathrm{T} = (\mathrm{L}/\mathsf{g})^{1/2}$, fixes $\Pi_1 = 1$. Through (4.4), the selection of L and T gives the respective units of measure that all lengths and times in the problem will be calibrated against. Characteristics scales given directly by a single given quantity from the original problem (here $\mathrm{L} = \mathsf{Y}_0$) are sometimes called *imposed scales*. Scales obtained from combinations of given quantities (here T in terms of Y_0 and g) are called *derived scales*.

We note that there are, in fact, many possible scaling choices as we could, for example, multiply either (or both) of the length or time scales by some function $f(\mathsf{Y}_0/\mathsf{R_E})$ and the resulting equation would still be nondimensional. However, it turns out that the normalised forms are often the most convenient choices.

We have employed what we shall refer to as

The 1st general nondimensional scaling principle: Select the characteristic scales so that as many as possible of the Π's are normalised.	(4.6)

For the current example, the nondimensionalised model takes the simplified form

$$\frac{d^2y}{dt^2} = -\frac{1}{(1+\Pi_2 y)^2}, \qquad y(0) = 1, \qquad y'(0) = \Pi_4. \tag{4.7a}$$

with

$$\Pi_2 = \frac{\mathsf{L}}{\mathsf{R_E}} = \frac{\mathsf{Y}_0}{\mathsf{R_E}}, \qquad \Pi_4 = \frac{\mathsf{V}_0\mathsf{T}}{\mathsf{L}} = \frac{\mathsf{V}_0}{(\mathsf{g}\mathsf{Y}_0)^{1/2}}. \tag{4.7b}$$

The solution of (4.7a) will be a function of the independent variable t and the remaining nondimensional constants Π_2 and Π_4, namely

$$y = y(t, \Pi_2, \Pi_4),$$

or equivalently,

$$\mathsf{Y} = \mathsf{Y}_0 y(t, \Pi_2, \Pi_4), \quad \mathsf{T} = (\mathsf{Y}_0/\mathsf{g})^{1/2} t,$$

in contrast to the solution of the original problem (4.3) which takes the general form

$$\mathsf{Y} = \mathsf{Y}(\mathsf{T}, \mathsf{g}, \mathsf{R_E}, \mathsf{Y}_0, \mathsf{V}_0).$$

We have effectively reduced the complexity of the original problem (involving a solution function of five variables) to a solution depending on only three variables.

As will be discussed further, nondimensionalization allows us to examine different limiting forms of problems in a mathematically precise framework. Consider the limit $\Pi_2 = \mathsf{Y}_0/\mathsf{R_E} \to 0$ in the model (4.7). At first sight, this limit may appear to describe increasing the radius of the Earth ($\mathsf{R_E} \to \infty$), or decreasing the initial height ($\mathsf{Y}_0 \to 0$), or both. For the limit $\mathsf{R_E} \to \infty$ (with Y_0 fixed), we can formally set $\Pi_2 = 0$ to reduce (4.7a) to a simpler ODE that can be easily solved,

$$\frac{d^2y}{dt^2} = -1 \qquad \Longrightarrow \qquad y(t) = 1 + \Pi_4 t - \tfrac{1}{2}t^2.$$

This solution can be interpreted as a *first approximation* to the motion of a projectile that is launched relatively close to the surface of a large-radius planet with the same gravity as the Earth. Undoing the nondimensional scaling (4.4) converts this to the dimensional form of the classic quadratic polynomial for projectile motion,

$$\mathsf{Y} = \mathsf{Y}_0 + \mathsf{V}_0\mathsf{T} - \tfrac{1}{2}\mathsf{g}\mathsf{T}^2.$$

However if we consider the limit $\mathsf{Y}_0 \to 0$ (with $\mathsf{R_E}$ fixed) then $\Pi_4 \to \infty$; this is not acceptable as a value of for the initial condition on $y'(0)$ and forces us to question whether we have selected an appropriate choice of characteristic scales.

In fact, if scalings yield any undefined terms in a non-dimensionalized model, then that model in that form will not be solvable. This points to the need to change the scalings to put the model into a well-defined form, which we express as

The 2nd general nondimensional scaling principle:
Select characteristic scales so that NO terms in
the model diverge in the physical limit of interest. (4.8)

The choice of scales is determined once the sub-set of Π's to be normalised is made; hence other choices correspond to picking different Π's. For problems with k characteristic scales, the scales will be determined from normalising k Π's; this will be discussed further in connection with the Buckingham Pi theorem, see Sect. 4.5.

To examine the $\mathsf{Y}_0 \to 0$ limit, consider setting $\Pi_1 = 1$ and $\Pi_4 = 1$ in (4.5). The length and time scales are then $\mathrm{L} = \mathsf{V}_0^2/\mathsf{g}$ and $\mathrm{T} = \mathsf{V}_0/\mathsf{g}$. The revised form of the nondimensionalised problem is

$$\frac{d^2y}{dt^2} = -\frac{1}{(1+\tilde{\Pi}_2 y)^2}, \qquad y(0) = \tilde{\Pi}_3, \qquad y'(0) = 1. \tag{4.9a}$$

$$\tilde{\Pi}_2 = \frac{\mathrm{L}}{\mathsf{R_E}} = \frac{\mathsf{V}_0^2}{\mathsf{g}\mathsf{R_E}}, \qquad \tilde{\Pi}_3 = \frac{\mathsf{Y}_0}{\mathrm{L}} = \frac{\mathsf{Y}_0\mathsf{g}}{\mathsf{V}_0^2}. \tag{4.9b}$$

This choice of scalings satisfies (4.8) since setting $\mathsf{Y}_0 = 0$ sets $\tilde{\Pi}_3 = 0$ but does not cause any other parameters in the problem to diverge,

$$\frac{d^2y}{dt^2} = -\frac{1}{(1+\tilde{\Pi}_2 y)^2}, \qquad y(0) = 0, \qquad y'(0) = 1.$$

If we further specify that $\mathsf{R_E} \to \infty$ then $\tilde{\Pi}_2 = 0$ and the ODE reduces to $y'' = -1$ yielding the solution

$$y(t) = -\tfrac{1}{2}t^2 + t \qquad \Longrightarrow \qquad \mathsf{Y} = \mathsf{V}_0\mathsf{T} - \tfrac{1}{2}\mathsf{g}\mathsf{T}^2.$$

For $\mathsf{V}_0 > 0$ the *time of flight* of the projectile is clearly seen to be $2\mathsf{V}_0/\mathsf{g}$ while the maximum height reached is $\mathsf{V}_0^2/(2\mathsf{g})$, both of which are directly proportional to the scales that we chose.

4.3.2 Terminal Velocity of a Falling Sphere in a Fluid

We now consider further aspects of scaling in nondimensionalizing another example from physics—describing the motion of a solid ball with a given initial velocity moving vertically downward through a viscous fluid.

The ball has radius R m and uniform density ρ kg/m^3 (so that the mass of the ball is given by $\mathsf{M} = \frac{4}{3}\pi\mathsf{R}^3\rho$). The initial speed of the ball is taken to be V_0 m/s. The

fluid has a different density, ρ_f and its viscosity is μ kg/m s, which gives a measure of the frictional resistance to the ball moving through the fluid.

Noting that the forces acting on the ball are due to its weight, buoyancy and the drag of the fluid, we can use Newton's second law to write down a force balance

$$\mathsf{M}\frac{d\mathsf{V}}{d\mathsf{T}} = \mathsf{Mg} - \mathsf{F}_{\mathsf{buoy}} - \mathsf{F}_{\mathsf{drag}}, \qquad \mathsf{V}(0) = \mathsf{V}_0,$$

where we have expressed the acceleration as the derivative of the ball's velocity. The magnitude of the buoyancy force is given by the weight of the fluid displaced by the ball, $\mathsf{F}_{\mathsf{buoy}} = \frac{4}{3}\pi\mathsf{R}^3\rho_f\mathsf{g}$, while the frictional force acting on a sphere is given by $\mathsf{F}_{\mathsf{drag}} = 6\pi\mu\mathsf{RV}$, a classic result derived by Stokes.

Consequently, the problem can be restated as

$$\frac{4}{3}\pi\mathsf{R}^3\rho\frac{d\mathsf{V}}{d\mathsf{T}} = \frac{4}{3}\pi\mathsf{R}^3\rho\mathsf{g} - \frac{4}{3}\pi\mathsf{R}^3\rho_f\mathsf{g} - 6\pi\mu\mathsf{RV}. \tag{4.10}$$

Introducing the nondimensionalization, $\mathsf{V}(\mathsf{T}) = \mathrm{V}v(t)$ with $\mathsf{T} = \mathrm{T}t$, where V and T are characteristic scales for speed and time yields the nondimensional model

$$\frac{dv}{dt} = \underbrace{\left(\frac{(\rho-\rho_f)\mathsf{gT}}{\rho\mathsf{V}}\right)}_{\Pi_1} - \underbrace{\left(\frac{9\mu\mathrm{T}}{2\rho\mathsf{R}^2}\right)}_{\Pi_2} v, \qquad v(0) = \underbrace{\left(\frac{\mathsf{V}_0}{\mathrm{V}}\right)}_{\Pi_3}. \tag{4.11}$$

Setting $\Pi_1 = 1$ and $\Pi_3 = 1$ yields the characteristic scales

$$\mathrm{V} = \mathsf{V}_0, \qquad \mathrm{T} = \frac{\rho\mathsf{V}_0}{(\rho-\rho_f)\mathsf{g}}, \tag{4.12}$$

and the nondimensionalized problem reads:

$$\frac{dv}{dt} = 1 - \mathrm{St}\, v, \qquad v(0) = 1, \tag{4.13}$$

where the Π_2 group has been renamed the *Stokes parameter*

$$\mathrm{St} = \frac{9\mu\mathsf{V}_0}{2(\rho-\rho_f)\mathsf{gR}^2}. \tag{4.14}$$

The "Stokes parameter" name attached to Π_2 is a historical label used for this dimensionless parameter, but this classification has important value since it makes it easier to search for other results on related problems involving this parameter in books, journals, and other sources. Hence dimensionless parameters make it possible to universally compare results from different studies (experiments, simulations, and theory) and communicate in a common terminology across many branches of

Table 4.1 A few commonly used dimensionless parameters

Name	Formula	Competing effects	Area of study
Arrhenius	$E/(\mathsf{RT})$	Activation/potential energy	Thermodynamics
Damkohler	$\mathsf{kL}/(\mathsf{UC}_0)$	Reaction/transport rates	Thermodynamics
Lewis	$\mathsf{D}_T/\mathsf{D}_M$	Thermal/mass diffusivity	Thermodynamics
Mach	V/c	Char. speed/sound speed	Aerodynamics
Peclet	UL/D	Convection/diffusion	Thermodynamics
Reynolds	$\rho\mathsf{UL}/\mu$	Inertia/viscosity	Fluid dynamics
Stokes	(4.14)	Drag/gravity	Fluid dynamics

science and engineering. See Table 4.1 for a short list of some of the many other named parameters, also see [69].

The nondimensionalized problem (4.13) has an explicit solution which consequently depends only on t and the single St parameter:

$$v(t, \mathrm{St}) = \frac{1}{\mathrm{St}}\left(1 - e^{-t\,\mathrm{St}}\right) + e^{-t\,\mathrm{St}}. \tag{4.15}$$

The solution of the original dimensional problem can then be written in terms of (4.15) and (4.12) as

$$\begin{aligned}\mathsf{V}(\mathsf{T}) &= \mathsf{V}_0 v(t/\mathrm{T}, \mathrm{St}) \\ &= \frac{\mathsf{V}_0}{\mathrm{St}} + \mathsf{V}_0\left(1 - \frac{1}{\mathrm{St}}\right)\exp\left(-\mathrm{St}\,\frac{(\rho - \rho_f)\mathsf{g}}{\rho\mathsf{V}_0}\,t\right) \\ &= \frac{2(\rho - \rho_f)\mathsf{g}\mathsf{R}^2}{9\mu}\left(1 - e^{-\frac{9\mu}{2\rho\mathsf{R}^2}\,t}\right) + \mathsf{V}_0 e^{-\frac{9\mu}{2\rho\mathsf{R}^2}\,t}.\end{aligned} \tag{4.16}$$

The last expression is what would have been obtained from solving (4.10) directly. As should be expected, its functional form is equivalent to (4.15), but (4.16) makes it difficult to see that the solution is actually only dependent on a single parameter. Having (4.13) in a less cumbersome form than (4.10) facilitates using it to identify the steady state as $v = 1/\mathrm{St}$; this represents the "terminal" free-fall velocity in this problem ($\mathsf{V}_0/\mathrm{St} = 2(\rho - \rho_f)\mathsf{g}\mathsf{R}^2/(9\mu)$ in dimensional form).

Here, as in most problems, the dimensionless parameters more efficiently and compactly capture the dependence of key properties of the system on its design parameters. In addition, if this model had been compared against experimental data for $\mathsf{V} = \mathsf{V}(\mathsf{T}, \mathsf{R}, \mathsf{V}_0, \rho_f, \rho, \mu, \mathsf{g})$ then while the data could involve many experiments varying the six parameters, Eq. (4.16) shows that all of that data could be captured by the two characteristic scales and the Stokes parameter. Dramatic reduction in re-organising data to show its underlying fundamental structure is often called *collapsing of data*, namely when many data points from different runs are all shown to fall on *universal curves*.

In (4.11), the ODE was normalised by the coefficient of the inertial term, making Π_1 the ratio of buoyancy to inertial effects and Π_2 the ratio of drag to inertial effects for given V, T. Setting the V, T characteristic scales to normalise Π_1, $\Pi_3 = 1$ gives a formula for St in terms of given quantities (4.14) which embodies this ratio,

$$\mathrm{St} = \frac{9\mu \mathsf{V}_0}{2(\rho - \rho_f)\mathsf{g}\mathsf{R}^2} = \frac{6\pi\mu \mathsf{R}\mathsf{V}_0}{\frac{4}{3}\pi \mathsf{R}^3(\rho - \rho_f)\mathsf{g}} = \frac{\text{drag force}}{\text{net gravity force}}, \tag{4.17}$$

We want to re-iterate the very important point that **all** dimensionless parameters are ratios of competing effects in models. This is a consequence of all Π's being formed by dividing equations in the model by scaling coefficients for the influence of one term. In general,

$$\Pi = \frac{\text{strength of effect 1}}{\text{strength of effect 2}}. \tag{4.18}$$

The limiting cases of such ratios have clear interpretations:

- $\Pi \to 0$: effect 1 is very weak (relative to effect 2).
- $\Pi \to \infty$: effect 1 is very strong (relative to effect 2).

Between these extremes other special *critical values* Π_c can exist that separate different *qualitative regimes* for the solution's behaviour. The model (4.13) has a critical Stokes number of $\mathrm{St}_c = 1$:

- For $\mathrm{St} < 1$ solutions describe balls starting from lower speeds, that will accelerate up to the terminal velocity.
- For $\mathrm{St} > 1$ solutions represent balls starting from speed highers than the terminal velocity decelerating for all times.
- For $\mathrm{St} = 1$, the solution starts at and maintains the terminal velocity.

This is a simple example of a critical value of a parameter marking a *bifurcation*. In other systems, critical values can signify a change in the stability of a solution, or act as boundaries of parameter regimes in which different numbers of solutions coexist.

Careful examination of related issues will be given in upcoming chapters (see Chaps. 6–10) using asymptotic analysis and perturbation methods to solve models in small parameter limits. For the moment, we note that apart from any physical interpretations, the limiting cases, $\Pi = 0$ and $\Pi = \infty$, have a clear practical difference—an infinite coefficient precludes any calculations using the "usual methods". In (4.13), having $\mathrm{St} = 0$ yields a well-defined solution,[2] $v(t) = 1 + t$, while (4.13) with $\mathrm{St} = \infty$ is a *singular limit* that can not be sensibly evaluated directly. As already described in connection with (4.8), the remedy for this is to rescale the problem.

[2]This limiting solution could also be obtained using L'Hopital's rule for the $\mathrm{St} \to 0$ limit of (4.15).

We note that the limit $\mathrm{St} \to 0$ in (4.14) corresponds to several different physical limiting situations:

- Low fluid viscosity, $\mu \to 0$,
- Low initial ball velocity, $\mathsf{V}_0 \to 0$,
- High ball density, $\rho \to \infty$,
- Large ball size, $\mathsf{R} \to \infty$.

For any of these physical regimes, (4.13) gives a well-defined model capable of representing the limiting behaviour. Conversely, for the opposite extremes (i.e. high viscosity or high speed or low ball mass), (4.13) would have a divergent coefficient and should not be used in current form.

For example, consider the limit of high fluid viscosity that would be problematic for (4.13). Setting $\Pi_2 = 1$ and $\Pi_3 = 1$ in (4.11) selects the characteristic scales

$$\mathrm{V} = \mathsf{V}_0, \qquad \mathrm{T} = \frac{2\rho \mathsf{R}^2}{9\mu}, \tag{4.19}$$

and the new form of the scaled model:

$$\frac{dv}{dt} = \tilde{\Pi}_1 - v, \qquad v(0) = 1, \tag{4.20}$$

with $\tilde{\Pi}_1 = 1/\mathrm{St}$ and yields the limiting solution $v(t) = e^{-t}$ for $\mathrm{St} = \infty$.

4.3.3 The Burgers Equation

To illustrate nondimensionalization for a partial differential equation, we consider a problem for the *Burgers equation*,

$$\frac{\partial \mathsf{U}}{\partial \mathsf{T}} + \mathsf{U}\frac{\partial \mathsf{U}}{\partial \mathsf{X}} = \mathsf{D}\frac{\partial^2 \mathsf{U}}{\partial \mathsf{X}^2}, \qquad 0 \le \mathsf{X} \le \mathsf{E}, \tag{4.21a}$$

with the boundary and initial conditions,

$$\mathsf{U}(0, \mathsf{T}) = \mathsf{A}, \qquad \partial_{\mathsf{X}}\mathsf{U}(\mathsf{E}, \mathsf{T}) = \mathsf{B}, \qquad \mathsf{U}(\mathsf{X}, 0) = \mathsf{C}, \tag{4.21b}$$

where A–E are given constants with appropriate dimensional units. In Chap. 2, the inviscid Burgers equation (2.39) was introduced, which is a special case of (4.21a) with $\mathsf{D} = 0$. Equation (4.21a) occurs in models for problems in fluid dynamics (where U represents a velocity) and other chemical and physical systems, where U represents other physical quantities. Our analysis of the non-dimensionalization applies irrespective of the units of U.

We must choose dimensional scaling constants U, L, and T in

$$\mathsf{U}(\mathsf{X},\mathsf{T}) = \mathrm{U}u(x,t), \qquad \mathsf{X} = \mathrm{L}\,x, \qquad \mathsf{T} = \mathrm{T}t, \tag{4.22}$$

to yield variables u, x, t that are dimensionless. Using (4.22) in the form $\mathsf{U}(\mathsf{X},\mathsf{T}) = \mathrm{U}u(\mathsf{X}/\mathrm{L}, \mathsf{T}/\mathrm{T})$ and employing the chain rule leads to the transformed problem

$$\left(\frac{\mathrm{U}}{\mathrm{T}}\right)\frac{\partial u}{\partial t} + \left(\frac{\mathrm{U}^2}{\mathrm{L}}\right)u\frac{\partial u}{\partial x} = \left(\frac{\mathsf{D}\mathrm{U}}{\mathrm{L}^2}\right)\frac{\partial^2 u}{\partial x^2}, \qquad 0 \le \mathrm{L}\,x \le \mathsf{E},$$

$$\mathrm{U}u(0,t) = \mathsf{A}, \qquad (\mathrm{U}/\mathrm{L})\partial_x u(\mathsf{E}/\mathrm{L},t) = \mathsf{B}, \qquad \mathrm{U}u(x,0) = \mathsf{C}.$$

Dividing across by one coefficient in each expression, we arrive at the nondimensional form

$$\frac{\partial u}{\partial t} + \underbrace{\left(\frac{\mathrm{UT}}{\mathrm{L}}\right)}_{\Pi_1} u\frac{\partial u}{\partial x} = \underbrace{\left(\frac{\mathsf{D}\mathrm{T}}{\mathrm{L}^2}\right)}_{\Pi_2}\frac{\partial^2 u}{\partial x^2}, \qquad 0 \le x \le \underbrace{\left(\frac{\mathsf{E}}{\mathrm{L}}\right)}_{\Pi_3}, \tag{4.23}$$

$$u(0,t) = \underbrace{\left(\frac{\mathsf{A}}{\mathrm{U}}\right)}_{\Pi_4}, \qquad \partial_x u(\mathsf{E}/\mathrm{L},t) = \underbrace{\left(\frac{\mathsf{B}\mathrm{L}}{\mathrm{U}}\right)}_{\Pi_5}, \qquad u(x,0) = \underbrace{\left(\frac{\mathsf{C}}{\mathrm{U}}\right)}_{\Pi_6}.$$

In (4.22), we have three scaling constants whose values we can choose to set (an arbitrary choice of) three out of the six the Π's in (4.23) to unity, say

$$\begin{aligned} \Pi_3 = 1 \quad &\Longrightarrow \quad \mathrm{L} = \mathsf{E}, \\ \Pi_4 = 1 \quad &\Longrightarrow \quad \mathrm{U} = \mathsf{A}, \\ \Pi_1 = 1 \quad &\Longrightarrow \quad \mathrm{T} = \mathrm{L}/\mathrm{U} = \mathsf{E}/\mathsf{A}. \end{aligned} \tag{4.24}$$

The characteristic length-, solution- and time-scales determined in (4.24) can be separated into two types: L, U are imposed scales, while the timescale T is a derived scale.

The final scaled version of the problem is then

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \Pi_2\frac{\partial^2 u}{\partial x^2}, \qquad 0 \le x \le 1, \tag{4.25a}$$

$$u(0,t) = 1, \qquad u_x(1,t) = \Pi_5, \qquad u(x,0) = \Pi_6, \tag{4.25b}$$

with

$$\Pi_2 = \frac{\mathsf{D}}{\mathsf{A}\mathsf{E}}, \qquad \Pi_5 = \frac{\mathsf{B}\mathsf{E}}{\mathsf{A}}, \qquad \Pi_6 = \frac{\mathsf{C}}{\mathsf{A}}.$$

The solution of this problem will be a function of the independent variables and the three remaining Π parameters and gives the solution of the original problem (4.21) via (4.22) as

$$\mathsf{U} = \mathsf{A}u(x, t, \Pi_2, \Pi_5, \Pi_6), \qquad \mathsf{X} = \mathsf{E}x, \qquad \mathsf{T} = (\mathsf{E}/\mathsf{A})t. \tag{4.26}$$

This gives the most general explicit understanding of how the solution depends on the constants defining the original problem. In contrast, without the use of scaling analysis, the best that can be said in general about the solution of the original problem (4.21) is that it will have some dependence on each parameter and variable,

$$\mathsf{U} = \mathsf{U}(\mathsf{X}, \mathsf{T}, \mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}, \mathsf{E}).$$

The nondimensionalization above proceeded by setting Π_1, Π_3 and Π_4 equal to unity. Setting $\Pi_1 = 1$ yielded $\mathrm{T} = \mathrm{L}/\mathrm{U}$, which is known as the *convective timescale* (the time taken to travel a distance L at velocity U). An alternative choice for a derived time-scale results if we set $\Pi_2 = 1$ instead of $\Pi_1 = 1$; the characteristic time is then $\mathrm{T} = \mathrm{L}^2/\mathsf{D}$ and is called the *diffusive timescale* (the time taken for diffusive effects to propagate throughout the domain). The new form of the scaled problem becomes

$$\frac{\partial u}{\partial t} + \tilde{\Pi}_1 u \frac{\partial u}{\partial x} = \frac{\partial^2 u}{\partial x^2}, \qquad 0 \le x \le 1, \tag{4.27a}$$

$$u(0, t) = 1, \qquad u_x(1, t) = \Pi_5, \qquad u(x, 0) = \Pi_6, \tag{4.27b}$$

with

$$\tilde{\Pi}_1 = \frac{\mathrm{UT}}{\mathrm{L}} = \frac{\mathsf{AE}}{\mathsf{D}}, \qquad \Pi_5 = \frac{\mathsf{BE}}{\mathsf{A}}, \qquad \Pi_6 = \frac{\mathsf{C}}{\mathsf{A}}.$$

It is notable that $\tilde{\Pi}_1$ is the inverse of Π_2 from the first choice of scalings. As explained earlier, a dimensionless parameter always represents the ratio of two competing effects; in this case convection and diffusion. The relative importance of these effects is therefore measured by the ratio

$$\mathrm{Pe} = \frac{\text{convective effects}}{\text{diffusive effects}} = \frac{\mathrm{UL}}{\mathsf{D}}, \tag{4.28}$$

called the Peclet number. We can write $\tilde{\Pi}_1 = \mathrm{Pe}$ or $\Pi_2 = 1/\mathrm{Pe}$ in their respective models to put the problem in forms comparable to other heat and mass transfer problems studied in engineering.

Whether (4.25), (4.27) or some other choice of scaled problem is appropriate for the questions of interest is resolved by the second scaling principle (4.8). We should choose the system (4.27) if we consider the limit $\mathrm{Pe} \to 0$ ($\Pi_1 \to 0$) (convective effects being less important than diffusive effects) and the system (4.25) for the limit $\mathrm{Pe} \to \infty$ ($\Pi_2 \to 0$) (convective effects being more important than diffusive effects).

4.4 Further Applications of Dimensional Analysis

4.4.1 Projectile Motion (Revisited)

Dimensional analysis can sometimes be used to identify the solution structure of problems without even writing down model equations. Although this may sound simpler than the approach we have outlined above, the absence of equations derived from physical laws requires us torely heavily on experience and intuition for guidance. With that caveat in mind, we now construct the form of the solution to a projectile problem without first deriving a model equation.

Consider a projectile of mass M kg launched at an angle α up from the ground with a given speed V m/s. At some later time, the projectile will return to the ground. Suppose that we wish to calculate the time of flight of the projectile T (s). Besides the given parameters describing the launch of the object (M, V, α), we expect the acceleration due to gravity, g m/s^2, to also play a key role.

Consequently, we may expect that

$$\mathsf{T} = \mathsf{F}(\mathsf{M}, \mathsf{V}, \mathsf{g}, \alpha) \tag{4.29}$$

for some dimensional function F. With respect to units of time, the only way to construct a dimensionless parameter in terms of the given quantities is $\Pi_1 = \mathsf{gT}/\mathsf{V}$. Then, corresponding to (4.29), the dimensionless time must given by a dimensionless function f

$$\frac{\mathsf{gT}}{\mathsf{V}} = f(\mathsf{M}, \mathsf{V}, \mathsf{g}, \alpha). \tag{4.30}$$

Of the parameters in f, M, V and g are independent in that we cannot write the dimensions of one of these quantities as a combination of the dimensions of the others (also note that α is an angle and so already a dimensionless quantity). We cannot therefore construct any further nondimensional quantities by combining them. This is not consistent with the requirement that f be a nondimensional function and so implies that f must be independent of *all* of these quantities, yielding simply that

$$\mathsf{T} = \frac{\mathsf{V}}{\mathsf{g}} f(\alpha). \tag{4.31}$$

We have thus simplified a dimensional function of four quantities down to a nondimensional function of just one, and, in addition, have shown that T should be independent of the projectile mass M without reference to any model equation.

Although this analysis does not provide the form of the function $f(\alpha)$, we can immediately see that T is linearly dependent on the initial velocity and inversely dependent on g; which is precisely what we obtained in Sect. 4.3.1.

In the above discussion, we ignored the influence of air resistance on the dynamics of the projectile. If we include the effects of air resistance so that the projectile

experiences a retarding force linearly proportional to the velocity with proportionality constant K (measured in units of kg/s), Eq. (4.29) would now take the form

$$\mathsf{T} = \tilde{\mathsf{F}}(\mathsf{M}, \mathsf{V}, \mathsf{g}, \alpha, \mathsf{K}). \tag{4.32}$$

In contrast with (4.31), it is now possible to combine M, V, g along with K into a dimensionless product. We consequently obtain

$$\frac{\mathsf{gT}}{\mathsf{V}} = \tilde{f}\left(\alpha, \frac{\mathsf{Mg}}{\mathsf{KV}}\right), \tag{4.33}$$

namely a dimensionless function of two dimensionless parameters based on independent combinations of the given quantities describing the problem. Note that we do not combine the dimensionless parameter α with the other dimensionless product; as α could be the exponent of a non-monomial function, say sin or log, and this whole function would remain nondimensional (and undetermined through dimensional analysis).

The introduction of the air resistance constant of proportionality K, also provides another way to generate a quantity with dimensions of time and this allows us to write

$$\frac{\mathsf{KT}}{\mathsf{M}} = \hat{f}\left(\alpha, \frac{\mathsf{Mg}}{\mathsf{KV}}\right). \tag{4.34}$$

At first glance, it appears that we have two functional forms for the same expression, but in fact (4.33) and (4.34) are equivalent, where we have (multiplying through by the nondimensional product)

$$\frac{\mathsf{Mg}}{\mathsf{KV}}\frac{\mathsf{KT}}{\mathsf{M}} = \frac{\mathsf{gT}}{\mathsf{V}} = \frac{\mathsf{Mg}}{\mathsf{KV}}\hat{f}\left(\alpha, \frac{\mathsf{Mg}}{\mathsf{KV}}\right) = \tilde{f}\left(\alpha, \frac{\mathsf{Mg}}{\mathsf{KV}}\right). \tag{4.35}$$

The time of flight T can, in fact, be shown to be given implicitly by

$$\frac{\mathsf{M}}{\mathsf{Kg}}\left(\mathsf{V}\sin\alpha + \frac{\mathsf{Mg}}{\mathsf{K}}\right)(1 - e^{-\mathsf{KT}/\mathsf{M}}) - \frac{\mathsf{MT}}{\mathsf{K}} = 0. \tag{4.36}$$

This equation cannot be inverted to give an explicit form for T. But (4.36) can be expressed in the form $F(\Pi_1, \Pi_2, \Pi_3) = 0$, with $\Pi_1 = \mathsf{KT}/\mathsf{M}$, which subsequently leads to (4.34). Thus (4.34) provides us with an understanding of the qualitative dependence of T on all quantities in the system and a more detailed quantitative solution would involve the numerical solution of (4.36).

4.4.2 Closed Curves in the Plane

For our final example, we touch on how dimensional analysis connects with the more general concept of *similarity* that we will describe in detail later in Sect. 4.5.1.

Simple (non-intersecting) closed curves can be associated with two quantities: an enclosed area A, and a perimeter P. In terms of dimensions, $[\mathsf{A}] = \mathsf{L}^2$ and the perimeter $[\mathsf{P}] = \mathsf{L}$, from which we can form one dimensionless parameter $\Pi_1 = \mathsf{P}^2/\mathsf{A}$ (the dimensionless perimeter to area ratio). For a circle of radius R, this ratio reduces to

$$\Pi_1 = \frac{(2\pi\mathsf{R})^2}{\pi\mathsf{R}^2} = 4\pi. \tag{4.37}$$

Note that this is the only dimensionless parameter for circles and it is independent of the radius of the circle.

For a square with sides of length L, the perimeter to area ratio is

$$\Pi_1 = \frac{(4\mathsf{L})^2}{\mathsf{L}^2} = 16. \tag{4.38}$$

Consequently, all squares are also similar to each other. Noting that $16 > 4\pi$, we observe that for a given perimeter length, circles enclose a greater area than that achieved by squares.[3]

Triangles with sides of length $\mathsf{L}, \mathsf{M}, \mathsf{N}$ have perimeter $\mathsf{P} = \mathsf{L} + \mathsf{M} + \mathsf{N}$ and Heron's formula gives the area as

$$\mathsf{A} = \sqrt{\frac{\mathsf{P}}{2}\left(\frac{\mathsf{P}}{2} - \mathsf{L}\right)\left(\frac{\mathsf{P}}{2} - \mathsf{M}\right)\left(\frac{\mathsf{P}}{2} - \mathsf{N}\right)}. \tag{4.39}$$

Consequently, the perimeter to area ratio can be written as

$$\Pi_1 = \frac{1}{\sqrt{\frac{1}{2}\left(\frac{1}{2} - \frac{\mathsf{L}}{\mathsf{P}}\right)\left(\frac{1}{2} - \frac{\mathsf{M}}{\mathsf{P}}\right)\left(\frac{1}{2} - \frac{\mathsf{N}}{\mathsf{P}}\right)}}.$$

In this expression, we have several dimensionless ratios of lengths: $\Pi_2 = \mathsf{L}/\mathsf{P}$, $\Pi_3 = \mathsf{M}/\mathsf{P}$, and $\Pi_4 = \mathsf{N}/\mathsf{P}$. As Π_1, Π_2 are independent parameters and noting that Π_4 can be written as $\Pi_4 = 1 - \Pi_2 - \Pi_3$ (using the equation for P), we can write

$$\Pi_1 = \frac{1}{\sqrt{\frac{1}{2}\left(\frac{1}{2} - \Pi_2\right)\left(\frac{1}{2} - \Pi_3\right)\left(\Pi_2 + \Pi_3 - \frac{1}{2}\right)}}. \tag{4.40}$$

[3] In Chap. 3, we showed that the calculus of variations can be used to prove that, among all smooth simple closed curves, the circle uniquely minimises the ratio Π_1.

4.5 The Buckingham Pi Theorem

We conclude this chapter by reviewing the mathematical theory behind dimensional analysis, known as the *Buckingham Pi theorem*. While dimensional analysis can be applied directly without this theorem, its results can be used to predict the number of dimensionless parameters in a problem and further reduce calculations involved.

For any system described by n dimensional quantities $\mathsf{Q}_1, \mathsf{Q}_2, \ldots, \mathsf{Q}_n$ involving r independent base dimensional units, $[\mathsf{U}_1], [\mathsf{U}_2], \ldots, [\mathsf{U}_r]$,[4] two results are universal:

(i) A set of dimensionless parameters Π_j, $j = 1, 2, \ldots$ (sometimes also called "dimensionless groups") can be written as monomial products in terms of powers of the Q_i's:

$$\Pi = \mathsf{Q}_1^{\alpha}\mathsf{Q}_2^{\beta}\mathsf{Q}_3^{\gamma}\ldots \tag{4.41}$$

Since each Π must be dimensionless, they must satisfy:

$$[\Pi] = [\mathsf{Q}_1]^{\alpha}[\mathsf{Q}_2]^{\beta}[\mathsf{Q}_3]^{\gamma}\ldots = [\mathsf{U}_1]^0[\mathsf{U}_2]^0\ldots[\mathsf{U}_r]^0 \tag{4.42}$$

This yields a set of r homogeneous linear equations for the exponents $\alpha, \beta, \gamma, \ldots$, one for each of the $[\mathsf{U}_k]$ units, $k = 1, 2, \ldots, r$.
Let $\tilde{r}$ be the number of linearly independent equations in this set, with $\tilde{r} \leq r$. Of the n exponents $(\alpha, \beta, \gamma, \ldots)$, $\tilde{r}$ of them will be determined from the dimensional relations (4.42) in terms of the remaining $n - \tilde{r}$ exponents, which are un-determined from this analysis.

(ii) All relations between the original quantities, $\mathsf{Q}_1, \mathsf{Q}_2, \ldots, \mathsf{Q}_n$ will be represented by corresponding relations among the $n - \tilde{r}$ dimensionless parameters, $\Pi_1, \Pi_2, \ldots, \Pi_{n-\tilde{r}}$:

$$F(\Pi_1, \Pi_2, \ldots, \Pi_{n-\tilde{r}}) = 0. \tag{4.43}$$

The structure of the relations is dependent upon the specific form of each problem. □

At its heart, the Π theorem is a consequence of linear algebra applied to the dimensional units of the Q's, with the exponents in the Π's being analogous to coefficients of a set of linearly independent vectors (whose roles are played by the base units U_k). Different presentations and more detail on the Π-theorem can be found in [9, 10, 49, 64].

[4] The base units, U_k in the Π theorem need not be fundamental units, they can be derived units, just as long they form a linearly independent set.

4.5.1 Mathematical Consequences

While the Buckingham Π theorem and the other observations made earlier in this chapter are straightforward, they lead to important consequences:

(i) For problems, where the solution is a function, the Pi theorem predicts that the non-dimensional form of the solution will depend on any independent variables and the Π groups,

$$u = u(x, t, \Pi_1, \Pi_1, \ldots, \Pi_{n-\tilde{r}}).$$

To write the solution of the original problem in terms of this dimensionless solution, characteristic scales are needed for all of the dependent and independent variables. Analogous to (4.42), choices for these scales in terms of the given quantities of the problem are obtained by solving for the exponents in

$$[\mathsf{Q}_1]^{\alpha}[\mathsf{Q}_2]^{\beta}[\mathsf{Q}_3]^{\gamma} \ldots = [\mathsf{U}_j]^1, \tag{4.44}$$

namely, the options for characteristic scales are determined by solving the linear system with the entry for one dimension set to unity while the other righthand side entries are still zero, as in (4.42).

(ii) If Π_1 is a dimensionless parameter for a system, then so is $\tilde{\Pi}_1 = \alpha \Pi_1^{\beta}$ for any numbers α, β, and even more generally, so is $\hat{\Pi}_1 = h(\Pi_1)$ and $\check{\Pi}_1 = g(\Pi_1, \Pi_2)$ for any functions g and h. This illustrates the very large degree of non-uniqueness in specifying dimensionless parameters. However, this being said, once choices are made for the Π's, they allow for systematic analysis of comparable systems.

(iii) Two systems are called *similar* if they reduce to the same dimensionless problem with the *same values for all of the Π's* in (4.43). Consequently, two similar systems will have the same dimensionless properties and their dimensional properties will be proportionally scaled. For example, in the context of triangles (from Sect. 4.4.2), two triangles are similar if they share the same values of Π_1, Π_2, Π_3; this is consistent with the description for similar triangles in plane geometry.
This principle, sometimes called the *similitude condition* (or "dynamic similitude") is the basis for constructing smaller-scale inexpensive prototypes for testing and predicting the properties of full-sized systems.

(iv) Under broad conditions, when we can assume the implicit function theorem applies to (4.43), then the values of some dimensionless parameters (or system properties) can be expressed as functions of the values of the others. For example, say $F(\Pi_1, \Pi_2, \Pi_3) = 0$, then we may able to obtain

$$\Pi_3 = f(\Pi_1, \Pi_2).$$

In particular, note that (4.40) is an example of a specific relation of the form $\Pi_1 = f(\Pi_2, \Pi_3)$.
The Π-theorem will not determine the solution function f, but it does help greatly by telling us how many independent parameters it depends on.

As described earlier, there are many different possible sets of Π's that can be employed to describe each problem. Useful choices depend on which system properties are known versus which ones are being sought.

4.5.2 Application to the Quadratic Equation

We now apply the Π theorem in detail to an elementary problem to illustrate a broader idea—while the theorem follows from dimensional analysis, it can be applied to any mathematical problem to determine the essential dependence of solutions on the given parameters of the system.

One context in which the quadratic equation,

$$\mathsf{A}\mathsf{T}^2 + \mathsf{B}\mathsf{T} + \mathsf{C} = 0, \tag{4.45}$$

arises is (as mentioned previously) in the kinematic study of an object moving at a constant acceleration, where T would be the time of flight (an unknown, to be solved for) and A, B, C are given dimensional constants related to the acceleration, initial velocity and initial position respectively. In this context, the dimensions of the quantities in (4.45) are

$$[\mathsf{T}] = \mathrm{T}, \qquad [\mathsf{C}] = \mathrm{L}, \qquad [\mathsf{B}] = \mathrm{L/T}, \qquad [\mathsf{A}] = \mathrm{L/T}^2.$$

Following (4.41), we can denote dimensionless parameters in terms of these quantities as

$$\Pi = \mathsf{A}^{\alpha}\mathsf{B}^{\beta}\mathsf{C}^{\gamma}\mathsf{T}^{\delta}, \tag{4.46}$$

for appropriate values of the exponents. Noting that the only dimensions that appear in this problem are L and T, for Π to be dimensionless requires that $[\Pi] = \mathrm{L}^0\mathrm{T}^0$, and so

$$(\mathrm{L/T}^2)^{\alpha}(\mathrm{L/T})^{\beta}\mathrm{L}^{\gamma}\mathrm{T}^{\delta} = \mathrm{L}^0\mathrm{T}^0 \qquad \rightarrow \qquad \mathrm{L}^{\alpha+\beta+\gamma}\mathrm{T}^{-2\alpha-\beta+\delta} = \mathrm{L}^0\mathrm{T}^0.$$

The four unknown exponents α, β, γ and δ must satisfy the two dimension-independence equations

$$\begin{aligned} \mathrm{L}: \quad & \alpha + \beta + \gamma = 0, \\ \mathrm{T}: \quad & -2\alpha - \beta + \delta = 0. \end{aligned}$$

Standard theory from linear algebra then gives that each solution will have two free parameters that can be set arbitrarily, with two variables being determined in terms of the other two variables. This exactly corresponds to the Π theorem's count: 4 quantities -2 independent dimensions $= 2$ parameters.

There are many possible choices for selecting the Π's. However, as explained in Sect. 4.3.1, there are conventions that lead to specific choices that make the mathematical formulation more easily interpretable than the original problem.

For the current example, we seek a dimensionless variable corresponding to T: in (4.46), if we pick $\delta = 1$ and for convenience take $\gamma = 0$, then we have $\alpha = 1$, $\beta = -1$, giving, $\Pi_1 = \mathsf{A}^1\mathsf{B}^{-1}\mathsf{T}^1$. Since $[\Pi_1] = 1$ we see that $[\mathsf{T}] = [\mathsf{B}/\mathsf{A}]$, namely that the ratio of the given quantities B/A determines a characteristic scale.

Excluding the variable T, the example has 3 given quantities (A, B, C)—2 dimensions (L, T) $= 1$ independent system parameter. Setting $\delta = 0$ in (4.46) and for convenience picking $\gamma = 1$ yields $\alpha = 1$, $\beta = -2$ and $\Pi_2 = \mathsf{AC}/\mathsf{B}^2$.

Employing the second result from the Buckingham theorem, the dimensionless groups should all be related, $F(\Pi_1, \Pi_2) = 0$. Applying the implicit function theorem to F, we can write

$$\Pi_1 = f(\Pi_2) \qquad \Longrightarrow \qquad \frac{\mathsf{AT}}{\mathsf{B}} = f(\Pi_2),$$

and hence

$$\mathsf{T} = \left(\frac{\mathsf{B}}{\mathsf{A}}\right) f\left(\frac{\mathsf{AC}}{\mathsf{B}^2}\right). \tag{4.47}$$

We now compare this against the standard result for the solution of the quadratic equation (4.45):

$$\begin{aligned} \mathsf{T} &= \frac{-\mathsf{B} \pm \sqrt{\mathsf{B}^2 - 4\mathsf{AC}}}{2\mathsf{A}} = \frac{-\mathsf{B} \pm \sqrt{\mathsf{B}^2(1 - 4\mathsf{AC}/\mathsf{B}^2)}}{2\mathsf{A}} \\ &= \frac{\mathsf{B}}{\mathsf{A}}\left(-\frac{1}{2} \pm \sqrt{1 - 4\frac{\mathsf{AC}}{\mathsf{B}^2}}\right) = \left(\frac{\mathsf{B}}{\mathsf{A}}\right) f(\Pi_2). \end{aligned} \tag{4.48}$$

Even though the Buckingham theorem cannot predict the details of the function $f(\Pi_2)$, (4.47) has indicated the way in which the solution of (4.45) depends on the given quantities.

4.6 Further Directions

Further use of dimensional analysis for mathematical modelling is well illustrated in Chap. 1 of Holmes [49] and at a more advanced level in the books of Barenblatt [9, 10] and Sedov [88]. In addition, the book by Szirtes [95] contains an extensive

number of worked examples in dimensional analysis taken from a wide range of areas.

Models for projectile motion have been used as a classic example of dimensional analysis in many books on modelling [49, 64] since Lin and Segel [63, 89]. This stems from its broad appeal as a fundamental problem in physics as well as the large number of basic modelling concepts it connects to. Likewise, the Burgers equation is a partial differential equation that plays an important role in many applications in applied mathematics beyond its original setting in fluid dynamics [106], and will be re-visited in several later chapters.

Finally, in our introductory discussion, we have barely touched on the relation of dimensional analysis to similitude and scale models—how to construct a model ship, for example, with properties that when scaled up will be consistent with those of the actual ship. The classic text by Pankhurst [83] has a clear and concise introduction to this topic written from an engineering perspective.

4.7 Exercises

4.1 *(Choosing scalings for different limits)* Consider the dimensional problem for the motion of a projectile launched from close to the surface of the Earth:

$$\frac{d^2\mathsf{Y}}{d\mathsf{T}^2} = -\frac{\mathsf{G}\mathsf{M}_\mathsf{E}}{(\mathsf{R}_\mathsf{E}+\mathsf{Y})^2}, \qquad \mathsf{Y}(0) = 2\,\mathrm{m}, \qquad \mathsf{Y}'(0) = -\mathsf{V}_0\,\mathrm{m/s}.$$

Assume the Earth to be spherical with a uniform density, $\mathsf{M}_\mathsf{E} = \frac{4}{3}\pi\mathsf{R}_\mathsf{E}^3\rho_\mathsf{E}$.
Let $\mathsf{Y}(\mathsf{T}) = \mathrm{L}y(t)$ and $\mathsf{T} = \mathrm{T}t$. Consider the following cases:

(i) The fast projectile limit: R_E = fixed, $\mathsf{V}_0 \to \infty$, ρ_E = fixed.
(ii) The dense Earth limit: R_E = fixed, V_0 = fixed, $\rho_\mathsf{E} \to \infty$.
(iii) The light Earth limit: R_E = fixed, V_0 = fixed, $\rho_\mathsf{E} \to 0$.
(iv) The small Earth limit: $\mathsf{R}_\mathsf{E} \to 0$, V_0 = fixed, M_E = fixed.

For each case:

(a) Choose characteristic scalings L, T to normalise as many terms as possible.
(b) Choose the scalings so that the time it takes for the projectile to fall to should be finite for the given limit. Namely, the speed, acceleration, and initial height should not diverge, nor should falling (say down to 1 m) take infinitely long in time (as in for example if the velocity and acceleration both approach zero) or happen nearly instantaneously (if the initial height approaches zero).
(c) Write the scaled problem and identify all remaining dimensionless parameters.
(d) Identify the limiting small parameter and for each case, write the problem (called *the leading order problem*) when the parameter is set to zero.

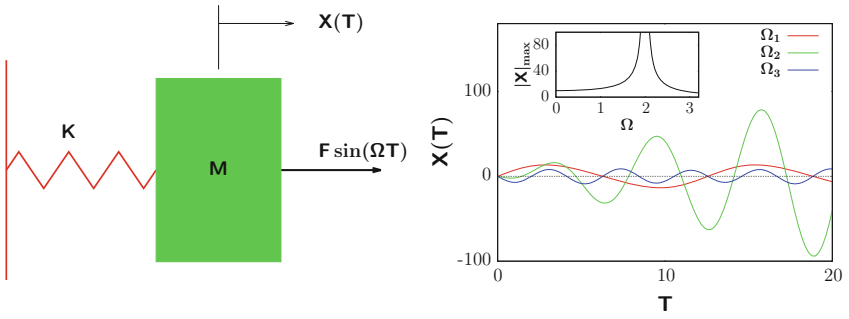


Fig. 4.1 The elementary problem of a mass on spring with an applied force: (*Left*) dimensional parameters, (*Right*) solutions $\mathsf{X}(\mathsf{T})$ for various parameters and (*Inset*) the amplitude in relation to the forcing frequency

4.2 The governing equation for the linear mass-spring system shown in Fig. 4.1 is

$$\mathsf{M}\frac{d^2\mathsf{X}}{d\mathsf{T}^2} + \mathsf{K}\mathsf{X} = \mathsf{F}\sin(\Omega\mathsf{T}), \quad \mathsf{X}(0) = \mathsf{X}_0, \quad \left.\frac{d\mathsf{X}}{d\mathsf{T}}\right|_{\mathsf{T}=0} = \mathsf{V}_0,$$

where $\mathsf{X}(\mathsf{T})$ [m] is the position of the mass as a function of time with mass M [kg], and spring coefficient K [N/m], magnitude of the applied force, F [N], forcing frequency, Ω [s^{-1}], as well as general initial conditions,

Nondimensionalize and select length- and time-scales L, T to normalise the coefficients of the terms on the left side of the ODE and the initial condition for position. Write and solve the scaled problem, and identify the dimensionless parameter for the ratio of the forcing frequency to resonant frequency, where the solution's amplitude grows without bound, as shown for Ω_2 in Fig. 4.1.

4.3 Consider the initial value problem for a damped driven nonlinear oscillator:

$$\mathsf{M}\frac{d^2\mathsf{X}}{d\mathsf{T}^2} + \mathsf{B}\frac{d\mathsf{X}}{d\mathsf{T}} + \mathsf{K}\mathsf{X}^3 = \mathsf{F}\sin(\Omega\mathsf{T}), \quad \mathsf{X}(0) = \mathsf{A}, \quad \left.\frac{d\mathsf{X}}{d\mathsf{T}}\right|_{\mathsf{T}=0} = \mathsf{C}.$$

(a) If the units are $[\mathsf{M}] = \mathrm{kg}$, $[\mathsf{T}] = \mathrm{s}$ and $[\mathsf{X}] = \mathrm{m}$, state the units for $\mathsf{A}, \mathsf{B}, \mathsf{K}, \mathsf{F}, \Omega$.
(b) Consider the limit $\mathsf{M} \to 0$. Let $\mathsf{X}(\mathsf{T}) = \mathsf{L}x(t)$ and $\mathsf{T} = \mathrm{T}t$ where L,T are characteristic scale constants. Determine choices for L,T that yield an ODE for x with all three terms on the left side of the equation normalised.
(c) Identify the dimensionless parameters in this scaled problem.

4.4 Consider the dimensional equation for the damped pendulum

$$\frac{d^2\Theta}{d\mathsf{T}^2} + \mathsf{B}\frac{d\Theta}{d\mathsf{T}} + \frac{\mathsf{g}}{\mathsf{L}}\sin\Theta = 0, \qquad \Theta(0) = \Theta_0, \qquad \Theta'(0) = \Omega_0,$$

where B is the damping coefficient, g gravity, L length, and Θ_0, Ω_0 initial angle and speed. Nondimensionalize in terms of general scalings for the position and time: $\Theta(T) = \bar{\Theta}\theta(t), T = \mathrm{T}t$. Determine the form of the scaled model and the appropriate characteristic scales for:

(i) The weakly damped limit: $\mathsf{B} \to 0$.
(ii) The slow speed limit: $\Omega_0 \to 0$.
(iii) The small amplitude oscillation limit: $\Theta_0 \to 0$.

4.5 Consider dimensional equations describing a system of chemical reactions for the concentrations of three chemicals:

$$\begin{aligned} \frac{d\mathsf{X}}{d\mathsf{T}} &= \mathsf{A} - \mathsf{BY} & \mathsf{X}(0) &= \mathsf{X}_0 \\ \frac{d\mathsf{Y}}{d\mathsf{T}} &= \mathsf{CX} - \mathsf{DZ} & \mathsf{Y}(0) &= \mathsf{Y}_0 \\ \frac{d\mathsf{Z}}{d\mathsf{T}} &= \mathsf{EY} - \mathsf{FY}^2 + \mathsf{GY}^3 - \mathsf{HZ} & \mathsf{Z}(0) &= \mathsf{Z}_0 \end{aligned} \tag{4.49}$$

where $\mathsf{A}, \mathsf{B}, \mathsf{C}, \ldots \mathsf{H}$ are given dimensional constants. By scaling $\mathsf{X}, \mathsf{Y}, \mathsf{Z}, \mathsf{T}$ (i.e. $\mathsf{X}(\mathsf{T}) = \mathrm{X}x(t)$ and similarly for the others) show that these equations can be nondimensionalized in the form:

$$\begin{aligned} \frac{dx}{dt} &= \alpha - y & x(0) &= \mu \\ \beta\frac{dy}{dt} &= x - z & y(0) &= \sigma \\ \gamma\frac{dz}{dt} &= y - y^2 + \frac{1}{3}\delta y^3 - z & z(0) &= \omega \end{aligned} \tag{4.50}$$

(a) Determine the characteristic scalings X, Y, Z, T and the dimensionless parameters in terms of A–H and $\mathsf{X}_0, \mathsf{Y}_0, \mathsf{Z}_0$.
(b) If $\gamma = 0$ show that the problem reduces to a system of two ODEs in terms of $x(t)$ and $y(t)$ only. Find the relation that the initial conditions must satisfy in order to avoid a contradiction.
(c) If $\beta = 0$ show that the problem reduces to single first-order ODE for $y(t)$ alone. Find the relation that the initial conditions must satisfy in order to avoid a contradiction.

4.6 Consider the dimensional problem for $\mathsf{U}(\mathsf{X}, \mathsf{T})$:

$$\frac{\partial \mathsf{U}}{\partial \mathsf{T}} + \mathsf{AU}^2\frac{\partial \mathsf{U}}{\partial \mathsf{X}} + \mathsf{B}\frac{\partial^4 \mathsf{U}}{\partial \mathsf{X}^4} = \mathsf{CX}^2\mathsf{T}^3 \qquad \mathsf{X} \geq 0 \qquad \mathsf{T} \geq 0$$

$$\mathsf{U}(0, \mathsf{T}) = 0, \qquad \mathsf{U}^3(0, \mathsf{T}) + \mathsf{DU}_{\mathsf{XXX}}(0, \mathsf{T}) = \mathsf{E},$$

where $\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}, \mathsf{E}$ are given constants.

(a) If X is measured in meters, T is measured in seconds, and U is a density function, measured in kg/m^3, determine the units for A–E.
(b) Determine L, T, U in terms of A–E in the dimensional scaling $\mathsf{U} = \mathrm{U}u(x,t)$ with $\mathsf{X} = \mathrm{L}x$, $\mathsf{T} = \mathrm{T}t$ so that all of coefficients in the PDE are normalised. Write the complete nondimensionalized problem and identify the dimensionless parameters.

4.7 *The shallow water equations* are a system of two PDEs describing fluid flow in shallow (long, slender) layers, like rivers. In dimensional form, they are

$$\partial_{\mathsf{T}}\mathsf{H} + \mathsf{H}\partial_{\mathsf{X}}\mathsf{U} + \mathsf{U}\partial_{\mathsf{X}}\mathsf{H} = 0,$$
$$\partial_{\mathsf{T}}\mathsf{U} + \mathsf{U}\partial_{\mathsf{X}}\mathsf{U} + \mathsf{g}\partial_{\mathsf{X}}\mathsf{H} = 0,$$

where $\mathsf{U}(\mathsf{X},\mathsf{T})$ is the fluid speed, $\mathsf{H}(\mathsf{X},\mathsf{T})$ is the height of the fluid layer, and g is the acceleration due to gravity.
Nondimensionalize using the average depth of a river H, and average speed U and lengthscale L

$$\mathsf{H} = \mathrm{H}h, \qquad \mathsf{U} = \mathrm{U}u, \qquad \mathsf{X} = \mathrm{L}x, \qquad \mathsf{T} = \mathrm{T}t.$$

Select the timescale T to normalise all terms in the h equation and determine the only remaining dimensionless parameter, called the Froude number (Fr).

4.8 One model for the motion of a projectile of mass M launched from the origin at an angle α to the horizontal with initial velocity V is given by the solution to the vector equation,

$$\mathsf{M}\frac{d^2\mathbf{X}}{d\mathsf{T}^2} = -\mathsf{M}\mathbf{g} - \mathsf{K}\frac{d\mathbf{X}}{d\mathsf{T}},$$

where $\mathbf{X} = (\mathsf{X},\mathsf{Y})$, $\mathbf{g} = (0,\mathsf{g})$ and K is a coefficient of air resistance.

(a) Use dimensional homogeneity to determine the units of K.
(b) In the absence of air resistance ($\mathsf{K} = 0$), show that the solution to this equation is consistent with (4.31).
(c) Show that Eq. (4.36) for the time of flight T can be obtained from this equation when the air resistance is directional proportional to the projectile's velocity ($\mathsf{K} \neq 0$).
(d) Repeat the dimensional analysis from the example in the text, assuming that the air-resistance is directly proportional to the square of the velocity, (note: you will have to redefine the dimensions of K).
(e) Perform a similar analysis in both cases to derive a nondimensional expression for the distance travelled.

4.9 Returning to the relation between area and perimeter from Sect. 4.4.2, consider these questions for different families of shapes:

(a) Show that for all rectangles (with length L and width W), the perimeter-to-area ratio can be expressed in terms of a length-to-width aspect ratio parameter.

(b) For ellipses having semi-major and semi-minor axes of L, W, $(\mathsf{X}/\mathsf{L})^2 + (\mathsf{Y}/\mathsf{W})^2 = 1$, the area is given by $\mathsf{A} = \pi\mathsf{L}\mathsf{W}$ and a formula by Ramanujan [85] gives the approximate perimeter as $\mathsf{P} \approx \pi(3(\mathsf{L}+\mathsf{W}) - \sqrt{(3\mathsf{L}+\mathsf{W})(\mathsf{L}+3\mathsf{W})}\,)$. Relate the dimensionless perimeter to area ratio to the aspect ratio. What can be said about the relations between a rectangle and its inscribed ellipse?
(c) Use (4.40) to show that all triangles have $\Pi_1 \geq 12\sqrt{3}$.

4.10 To show that if different choices are made for Π_1, Π_2 in (4.46), the results still yield the form of the quadratic formula (4.48). Determine $\Pi_1, \Pi_2, f(\Pi_2)$ if $\delta = \gamma = 1$ in Π_1 and $\delta = 0, \gamma = 1$ in Π_2.

4.11 Consider the cubic equation,

$$\mathsf{A}\mathsf{T}^3 + \mathsf{B}\mathsf{T}^2 + \mathsf{C}\mathsf{T} + \mathsf{D} = 0.$$

(a) Assuming the dimensions are $[\mathsf{T}] = \mathrm{T}$ and $[\mathsf{D}] = \mathrm{L}$, use the Buckingham Π theorem to determine the dimensionless groups for this equation.
(b) Using (a) write a formula for how the roots T depend on combinations of the coefficients.
(c) If X_k are the three roots of the equation $2\mathsf{X}^3 + 3\mathsf{X}^2 + 4\mathsf{X} + 5 = 0$, then use (b) to determine new coefficients in the following equation

$$\mathsf{Y}^3 + \tilde{B}\,\mathsf{Y}^2 + \tilde{C}\,\mathsf{Y} + \tilde{D} = 0$$

in terms of the previous A–D such that the roots are double the roots of the first equation, $\mathsf{Y}_k = 2\mathsf{X}_k$. (Do not determine the values of the roots directly.)

4.12 The fluid dynamics problem of how water drips from a faucet depends on several quantities: the acceleration due to gravity $\mathsf{g}[\mathrm{m/s^2}]$, the density of water $\rho[\mathrm{kg/m^3}]$, the viscosity of water $\mu[\mathrm{kg/(ms)}]$, the surface tension of water $\sigma[\mathrm{kg/s^2}]$, the radius of the faucet nozzle $\mathsf{R}[\mathrm{m}]$, and the speed of the water coming out of the faucet $\mathsf{U}[\mathrm{m/s}]$.

(a) Use the Buckingham Π theorem to determine the number of dimensionless parameters and write

$$\Pi = \mathsf{g}^A \rho^B \mu^C \sigma^D \mathsf{R}^E \mathsf{U}^F$$

to determine the equations relating the exponents A–F.
(b) The choice of dimensionless parameters to describe problems is not unique. This problem can be described in terms of the set of Π parameters:

$$\mathrm{Re} = \frac{\rho \mathsf{U}\mathsf{R}}{\mu} \qquad \mathrm{Bo} = \frac{\rho \mathsf{g}\mathsf{R}^2}{\sigma} \qquad \mathrm{Ca} = \frac{\mu \mathsf{U}}{\sigma}$$

(called the Reynolds, Bond, and Capillary numbers) or in terms of the set

$$\mathrm{We} = \frac{\rho \mathsf{U}^2 \mathsf{R}}{\sigma} \qquad \mathrm{Oh} = \frac{\mu}{\sqrt{\rho \mathsf{R} \sigma}} \qquad \mathrm{Ga} = \frac{\rho^2 \mathsf{g} \mathsf{R}^3}{\mu^2}$$

(the Weber, Ohnesorge, and Galileo numbers).

(i) State the choices for A–F for each parameter and show that each of these parameters satisfies the equations from (a).
(ii) Show that the 3 parameters in each set are independent by showing that their vectors of dimensional exponents $\mathbf{v} = (A, B, C, D, E, F)$ are linearly independent.

Part II
Solution Techniques

Chapter 5
Self-Similar Scaling Solutions of Differential Equations

Constructing solutions of ordinary and partial differential equations can be a challenging process (even for the linear case). For linear problems, techniques such as separation of variables, and Fourier and Laplace transforms can often be used to write the general solution of a differential equation that can then be specialised to a particular problem through the imposition of additional (initial and/or boundary) conditions to generate a unique solution. Such solutions can be cumbersome to obtain and may still require substantial further analysis in order to gain a clear insight into the nature of the system's resulting behaviour. For nonlinear problems, techniques to obtain general solutions exist for only a very limited number of types of equations.

In this chapter, we will show that the scaling analysis introduced in the context of dimensional analysis in Chap. 4 can be applied to obtain special solutions of both linear and nonlinear partial differential equations. These solutions are called *similarity solutions* (sometimes also called *self-similar solutions*).

It will be shown that if the PDE can be scaled in such a way as to exactly reproduce its original form (as suggested by the term 'self-similar'), then it will possess a class of solutions that share this property. A consequence of the property of self-similarity is that these solutions can be obtained from a reduced version of the original problem. For a PDE with two independent variables, this leads to an ODE problem for the similarity solution. Despite their nature as exact solutions to a given problem under special conditions, they often provide important understanding of the broader behaviour of all solutions of the system.

The process of constructing a similarity solution to a given problem consists of three stages:

(i) Looking for a scaling "symmetry" of the problem to see if similarity solutions are possible.
(ii) Determining the forms of the similarity variables and functions from the scale-invariant Π groups for the problem.

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9_5

(iii) Solving a reduced problem for the similarity function and then transforming back to give the solution of original problem.

We will illustrate the stages in this process in detail and then go on to apply the method to further examples.

5.1 Finding Scaling-Invariant Symmetries

Scaling analysis can identify whether a problem, say for $u = u(x, t)$, will admit a similarity solution by making use of the change of variables

$$u(x, t) = \mathsf{U}\tilde{u}(\tilde{x}, \tilde{t}), \qquad x = \mathsf{L}\tilde{x}, \qquad t = \mathsf{T}\tilde{t}, \tag{5.1}$$

where $\mathsf{U}, \mathsf{L}, \mathsf{T}$ are un-determined real positive parameters. We call the problem *scale invariant* if relationships exist between the scaling parameters $\mathsf{U}, \mathsf{L}, \mathsf{T}$ in (5.1) that make the scaled problem take exactly the same form as the original problem with at least one scaling parameter remaining undetermined,

$$\text{Problem for } u(x, t) \quad \equiv \quad \text{Problem for } \tilde{u}(\tilde{x}, \tilde{t}\,). \tag{5.2}$$

For evolution equations, it expected that the timescale T is the free parameter and the other scales can be expressed as $\mathsf{U} = \mathsf{U}(\mathsf{T})$ and $\mathsf{L} = \mathsf{L}(\mathsf{T})$.

As an example, consider the inviscid Burgers equation for $u(x, t)$ that was introduced in Chap. 2,

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0, \qquad 0 \le x < \infty, \tag{5.3a}$$

subject to two side conditions,

$$u(0, t) = 0, \qquad \int_0^\infty u(x, t)\, dx = 4. \tag{5.3b}$$

Substituting the change of variables (5.1) into (5.3a, 5.3b) yields,

$$\frac{\partial \tilde{u}}{\partial \tilde{t}} + \left(\frac{\mathsf{UT}}{\mathsf{L}}\right)\tilde{u}\frac{\partial \tilde{u}}{\partial \tilde{x}} = 0, \qquad 0 \le \tilde{x} < \infty, \tag{5.4a}$$

$$\tilde{u}(0, \tilde{t}\,) = 0, \qquad (\mathsf{UL})\int_0^\infty \tilde{u}(\tilde{x}, \tilde{t}\,)\, d\tilde{x} = 4. \tag{5.4b}$$

Setting $\mathsf{UT}/\mathsf{L} = 1$ and $\mathsf{UL} = 1$ eliminates the scaling constants from (5.4a, 5.4b) and makes the scaled system identical with (5.3a, 5.3b) and hence the problem is scale invariant. From these two relationships, we can express U and L in terms of the free parameter T,

$$\mathsf{L} = \mathsf{T}^{1/2}, \qquad \mathsf{U} = \mathsf{T}^{-1/2}. \tag{5.5}$$

Consequently, if we have one solution of (5.3a, 5.3b), say $u = \tilde{u}(\tilde{x}, \tilde{t})$, then other solutions are given by the one-parameter continuous family,

$$u(x,t) = \mathsf{T}^{-1/2}\tilde{u}(\mathsf{T}^{-1/2}x, \mathsf{T}^{-1}t), \tag{5.6}$$

for arbitrary values of T. This transformation is called a *scaling symmetry* of (5.3) [22, 52].

There is a strong resemblance between the scaling forms (5.1) and (4.22), but there is also an important underlying difference. The problems in this chapter will be taken to already be in dimensionless form (having been previously nondimensionalized using the approach described in Chap. 4). The scaling parameters in (5.1) are therefore dimensionless, corresponding to the magnitudes of solution-, length-, time-, or other relevant scales.

5.2 Determining the Form of the Similarity Solution

Setting $\mathsf{T} = 1$ in (5.6) returns the original solution and provides no new insight. But, noting that T is an arbitrary parameter and not part of the original specification of the problem suggests that the solution should be *independent of* T, or more precisely, the solution should be invariant with respect to changes in T. We can use this principle to yield additional information on the form of the solution.

In a manner analogous to that used for dimensionless parameters in Chap. 4, we define a *scale-invariant parameter*, Π, as a monomial product of powers of the variables of the system

$$\Pi = u^a x^b t^c, \tag{5.7}$$

which, upon substitution of (5.1), is independent of the undetermined scaling parameter (T). Analogous to the notation we have previously employed for dimensions of a quantity, let $[\Pi]$ represent the dependence of Π on the scalings $\mathsf{U}, \mathsf{L}, \mathsf{T}$ so that

$$[\Pi] = \mathsf{U}^a\mathsf{L}^b\mathsf{T}^c = \mathsf{T}^0, \tag{5.8}$$

and in particular for (5.5), this yields

$$[\Pi] = \mathsf{T}^{-a/2}\mathsf{T}^{b/2}\mathsf{T}^c = \mathsf{T}^0.$$

We then define a *scale invariant independent variable* (or "*similarity variable*") and the corresponding *scale invariant solution* (or "*similarity function*") as follows.

The *similarity variable* is a scale-invariant product Π that does not depend on the solution u, namely $a = 0$ in (5.7). Generally it is most convenient to choose a linear dependence on the independent spatial variable, $b = 1$, which leads to[1]

$$\Pi_1 = xt^c \qquad \Longrightarrow \qquad [\Pi_1] = \mathsf{L}\mathsf{T}^c = \mathsf{T}^0. \tag{5.9}$$

For our example, from (5.5), $[\Pi_1] = \mathsf{T}^{1/2+c}$ and so we obtain $c = -1/2$. A common notation for the similarity variable is the Greek letter η, hence we relabel Π_1 as η yielding

$$\eta = xt^{-1/2}.$$

The *similarity function* is a scale-invariant Π that is (most simply) linearly related to the solution of the model, namely $a = 1$ in (5.7). We must make choices for the exponents b, c in order to represent the solution in the most convenient form; such choices are typically guided by the side-conditions present in the problem.

For convenience in (5.3), we select $b = 0$ yielding

$$\Pi_2 = ut^c \qquad \Longrightarrow \qquad [\Pi_2] = \mathsf{U}\mathsf{T}^c = \mathsf{T}^0. \tag{5.10}$$

This requires that $[\Pi_2] = \mathsf{T}^{-1/2+c}$ and hence $c = 1/2$; we relabel Π_2 as f (the *similarity function*) which satisfies

$$f = t^{1/2}u.$$

The main result then follows from the analogous result in the second part of the Buckingham Pi theorem (see Sect. 4.5.1)—that all dimensionless parameters must be related to each other. Similarly, this implies that all scale invariant parameters must be related, which we state in the form

$$F(\Pi_1, \Pi_2) = 0 \qquad \Longrightarrow \qquad F(\eta, f) = 0.$$

Applying the implicit function theorem then gives

$$f = f(\eta) \qquad \Longrightarrow \qquad t^{1/2}u = f(xt^{-1/2}).$$

and consequently the final form of the similarity solution of the PDE,

$$u(x,t) = t^{-1/2} f(xt^{-1/2}). \tag{5.11}$$

[1] The choice $b = 1$ leads to a differential equation for the similarity function whose dependence on η-derivatives will mirror the x-derivatives in the original problem.

We note that using this representation of the solution reduces the boundary condition at $x = 0$ for $t > 0$ to the condition $f(0) = 0$. If we had instead selected $c = 0$ in (5.7), that choice would yield $\Pi_2 = ux$ and $u(x,t) = x^{-1}g(xt^{-1/2})$. The boundary condition at $x = 0$ would then be in-determinant (and hence more difficult to work with).

5.3 Solving for the Similarity Function

Once the form of the similarity solution of the PDE has been determined, substituting that form into the original problem will yield a reduced problem for the similarity function $f(\eta)$. Substituting (5.11) into (5.3a, 5.3b) gives

$$-\tfrac{1}{2}t^{-3/2}f(\eta) - \tfrac{1}{2}xt^{-2}\frac{df}{d\eta} + t^{-3/2}f(\eta)\frac{df}{d\eta} = 0,$$

$$t^{-1/2}f(0) = 0, \qquad \int_0^\infty f(\eta)\,d\eta = 4.$$

Noting that the product xt^{-2} can be rewritten as $xt^{-2} = \eta t^{-3/2}$, we are able to factor out all occurrences of the variable t to leave a reduced problem for the similarity function,

$$f + \eta\frac{df}{d\eta} - 2f\frac{df}{d\eta} = 0, \qquad f(0) = 0, \qquad \int_0^\infty f\,d\eta = 4. \tag{5.12}$$

In this case, we have reduced a PDE to an ODE. The separation between the new linearly independent variables t and η is a universal feature of similarity solutions stemming from the scale-invariance of the problem. If (5.12) could not be written in a form independent of t, then an error must have been introduced in one of the previous steps; this is a useful consistency check on the process.

The ODE for $f(\eta)$ in (5.12) can then be put in form $(\eta f)' = (f^2)'$ and directly integrated to give

$$\eta f - f^2 = C.$$

Evaluating this equation at $\eta = 0$ and applying the condition $f(0) = 0$ then fixes the value of the constant of integration $C = 0$ and we are left with two possible solutions,

$$f(\eta) = 0, \qquad \text{or} \qquad f(\eta) = \eta. \tag{5.13}$$

Inserting the nontrivial solution, $f(\eta) = \eta$, into (5.11) we obtain a self-similar solution of the PDE in the form

$$u(x,t) = x/t, \tag{5.14}$$

which is known as a rarefaction fan (see Sect. 2.6, Eq. (2.43) for a shifted version of this solution obtained via the method of characteristics). Finally, we note that satisfying the integral condition in (5.12) necessitates the use of both solutions from (5.13) in constructing the solution of (5.3) (compare with Sect. 2.6).

5.4 Further Comments on Self-Similar Solutions

It is instructive to contrast the above analysis of problem (5.3a, 5.3b) with the scaling and non-dimensionalization performed on problem (4.21a, 4.21b) in Chap. 4. The PDE (4.21a) is very closely related to (5.3a), and we will show that this equation does indeed have similarity solutions of the form (5.11) in Exercise 5.1.

The scaling of (4.21a, b) yielded fixed constants for all of the scaling parameters and hence that problem is not scale invariant. While such fully specified problems set specific scales for all variables, appropriately modified versions of such problems may have similarity solutions. Such modifications may include omitting initial or boundary conditions, changing the domain (say from $0 \le x \le L$ to $0 \le x < \infty$), or even omitting some terms from the governing equation. In Chaps. 6 and 7, we will see how such changes can be motivated by asymptotic analysis. Solutions of such problems may approach the similarity solution in some limit (of a variable or system parameter) and are sometimes called *asymptotically self-similar* solutions.

5.5 Similarity Solutions of the Heat Equation

The heat equation (or diffusion equation)

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \tag{5.15}$$

is a classic example in the investigation of symmetries in PDEs because it admits several classes of symmetries, each of which corresponds to a different similarity solution.

Applying the scaling (5.1) yields

$$\frac{\partial \tilde{u}}{\partial \tilde{t}} = \left(\frac{\mathsf{T}}{\mathsf{L}^2}\right) \frac{\partial^2 \tilde{u}}{\partial \tilde{x}^2}.$$

Note that U scales out due to the linearity of the equation, and must be determined by side conditions that further specify the problem. Making the PDE scale-invariant fixes the relationship between the length- and time-scales for diffusive solutions, $\mathsf{L} = \sqrt{\mathsf{T}}$ (recall the discussion of the diffusive timescale in Sect. 4.3.3). Consequently, we can write $\Pi_1 = \eta = xt^{-1/2}$ as the similarity variable for all typical problems for (5.15).

In the context of the linear diffusion equation, we will now consider how the choice of side conditions can affect the overall structure of the similarity solution.

5.5.1 Source-Type Similarity Solutions

Consider (5.15) on $-\infty < x < \infty$ starting from a non-negative initial condition at $t = 0$ that decays rapidly, $u(|x| \to \infty) \to 0$. Assuming no flux from infinity, the solution will maintain its initial mass (assumed finite),

$$\int_{-\infty}^{\infty} u(x,t)\,dx = \int_{-\infty}^{\infty} u(x,0)\,dx = M.$$

This integral constraint will be scale invariant if $\mathsf{UL} = 1$. Consequently, $\mathsf{U} = \mathsf{T}^{-1/2}$, and $\Pi_2 = ut^{1/2}$, yielding the form of the similarity solution,

$$u(x,t) = t^{-1/2} f(xt^{-1/2}).$$

Substituting this form into (5.15) yields the ODE for $f(\eta)$,

$$-\tfrac{1}{2}(f + \eta f') = f'', \tag{5.16}$$

which can be re-written as $-\frac{1}{2}(\eta f)' = f''$ and integrated by separation of variables. The non-negative solution of this problem is

$$f(\eta) = \frac{M}{\sqrt{4\pi}} e^{-\eta^2/4}$$

and corresponds to the solution of the heat equation,

$$u(x,t) = \frac{M}{\sqrt{4\pi t}} e^{-x^2/(4t)} \tag{5.17}$$

which is known as the *Cauchy similarity solution* or the *source-type similarity solution*, based on its role as the solution of the heat equation starting from a point source (with strength M) at the origin at $t = 0$ (Fig. 5.1). The 'Cauchy solution' label for (5.17) stems from the fact that it solves the *Cauchy problem* for the PDE, namely, an initial value problem on the whole line $-\infty < x < \infty$.

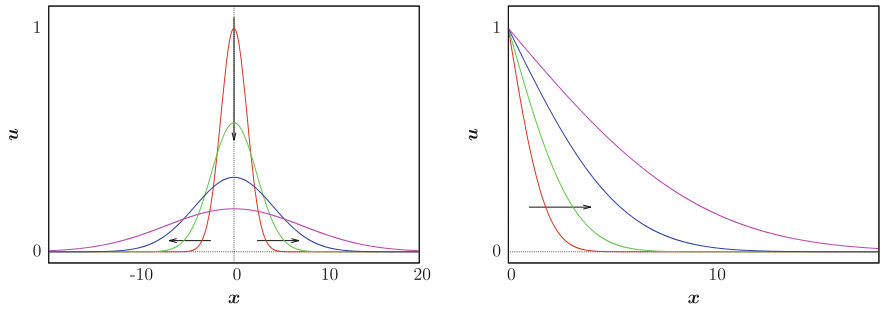


Fig. 5.1 (*Left*) The Cauchy similarity solution (5.17), (*Right*) The Boltzmann similarity solution (5.20)

Note that (5.15) is also invariant under *translations* in space, $x = \tilde{x} + a$, and time, $t = \tilde{t} + b$; namely $\tilde{u} = u(\tilde{x}, \tilde{t})$ is also a solution of the heat equation for any choice of the shifts a and b. A consequence is that (5.17) can be used to describe the long-time behaviour for more general initial conditions, with parameters that depend only on the mass, centre of mass and variance of the initial data (see Chap. 11).

5.5.2 The Boltzmann Similarity Solution

Consider (5.15) on the semi-infinite domain $0 \leq x < \infty$ with the Dirichlet boundary conditions

$$u(x = 0, t) = 1, \qquad u(x \to \infty) \to 0.$$

Having the boundary condition at $x = 0$ be scale invariant requires $\mathsf{U} = 1$. We therefore find that $\Pi_2 = u$ and the similarity solution takes the form

$$u(x, t) = f(xt^{-1/2}),$$

with $f(\eta)$ satisfying the ODE

$$-\tfrac{1}{2}\eta f' = f'', \tag{5.18}$$

which can be integrated once to yield a Gaussian, and then formally again as the integral of the Gaussian,

$$f(\eta) = 1 - \mathrm{erf}(\eta/2),$$

where the error function is defined as

$$\mathrm{erf}(z) \equiv \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2}\, dt, \tag{5.19}$$

yielding the similarity solution

$$u(x,t) = 1 - \operatorname{erf}\big(x/\sqrt{4t}\,\big). \tag{5.20}$$

5.6 A Nonlinear Diffusion Equation

Finally, we consider the similarity solution for the Cauchy problem for a nonlinear diffusion equation,

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(u^3 \frac{\partial u}{\partial x}\right), \tag{5.21a}$$

with

$$u_x(0,t) = 0, \qquad \int_{-\infty}^{\infty} u\,dx = M. \tag{5.21b}$$

This PDE, often called the *porous medium equation*, models the spreading of a puddle of very viscous fluid (such as honey) on a flat solid surface moving under the influence of gravity [1]. Here, $u(x,t)$ represents the height of the fluid layer at location x and time t; physically meaningful solutions of this problem will necessarily have $u(x,t) \geq 0$. The finite-mass similarity solution of this problem will have compact support. In other words, it will be positive on a finite domain,[2] $-x_*(t) < x < x_*(t)$, with $u \equiv 0$ for $|x| \geq x_*(t)$ where $x_*(t)$ is called the *interface* or *contact line*.

Proceeding as before, making the PDE scale invariant, yields the relation $\mathsf{U}^3 = \mathsf{L}^2/\mathsf{T}$ while the integral sets $\mathsf{UL} = 1$. This yields the scalings

$$\mathsf{L} = \mathsf{T}^{1/5} \qquad \mathsf{U} = \mathsf{T}^{-1/5},$$

and determines the form of the similarity solution as

$$u(x,t) = t^{-1/5} f(xt^{-1/5}).$$

Substituting this form back into the original system reduces it to an ODE problem for $f(\eta)$

$$-\tfrac{1}{5}(\eta f)' = (f^3 f')', \qquad f'(0) = 0, \qquad \int_{-\eta_*}^{\eta_*} f\,d\eta = M, \tag{5.22}$$

[2] Without loss of generality, we assume the support to be symmetric relative to the origin.

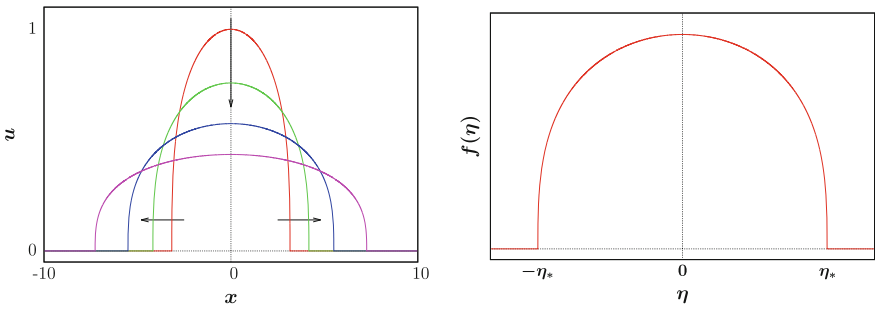


Fig. 5.2 (*Left*) The $u(x,t)$ lowering/spreading self-similar solution profiles (5.23) at several times, (*Right*) The profiles scaled in terms of $t^{1/5}u = f$ and $t^{-1/5}x = \eta$ showing the similarity function

where η_* is a constant to be determined, related to the interface position by $x_*(t) = \eta_* t^{1/5}$. The general solution of the differential equation can be obtained in closed form and, upon application of the boundary condition at $\eta = 0$, reduces to

$$f^3 = C_2 - \frac{1}{10}\eta^2.$$

Finally, for fixed initial mass M, the integral condition combined with continuity of the solution (i.e. $f(\pm\eta_*) = 0$) determines C_2 and η_*, thereby specifying f completely as

$$u(x,t) = \frac{1}{t^{1/5}}\left(C_2 - \frac{x^2}{10t^{2/5}}\right)^{1/3}, \tag{5.23}$$

see Fig. 5.2.

5.7 Further Directions

In the examples, we have seen how a PDE problem can sometimes be reduced to an ODE problem through the determination of a self-similar solution. In other problems, scaling invariance may be used to reduce PDEs in terms of several variables to PDE depending on fewer variables, or to generate autonomous ODEs from non-autonomous ones.

The books by Barenblatt [9, 10] are an excellent introduction to problems (taken from various different fields) that admit similarity solutions. Also see [2, 3, 88] for example. Other books provide extensive discussion on constructing similarity solutions [22, 52], and the theory of symmetries of PDEs [14, 32].

5.8 Exercises

5.1 Show that the Burgers equation,

$$u_t + uu_x = \kappa u_{xx},$$

with $\kappa > 0$ admits the same form of similarity solution (5.11) as problem (5.3) and determine the ODE for $f(\eta)$.

5.2 Find the similarity solution of the inviscid Burgers equation (5.3a) with the integral condition in (5.3b) replaced by

$$\int_0^\infty u^2(x,t)\,dx = 1.$$

Show that the same similarity scalings also apply to the generalised Burgers equation

$$\left(\tfrac{1}{2}u^2\right)_t + \left(\tfrac{1}{3}u^3\right)_x = \kappa u_{xx}.$$

Show that if $\kappa = 0$, this PDE reduces to (5.3a).

5.3 Consider the following problem for $u(x,t)$ on $0 \le x \le \infty$:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = t^\sigma u^4, \qquad u(0,t) = t^3, \qquad u(x,0) = 0.$$

(a) Determine the value of the constant σ so that this problem has a similarity solution.
(b) Determine the scales U, L in terms of the timescale T for the self-similar solution.
(c) Write $u(x,t)$ for the similarity solution as a product of some power of t and a similarity function $f(\eta)$. State the ODE for $f(\eta)$.

5.4 Consider the heat equation for $u(x,t)$ on the half-line, $0 \le x < \infty$

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

For each of the following sets of side-conditions, determine: (i) the invariant scalings U, L in terms of T, (ii) the form of the self-similar solution $u(x,t)$ in terms of $f(\eta)$, and (iii) the ODE and boundary conditions that $f(\eta)$ should satisfy.

(a) $u(0,t) = t^2$ and $u(x \to \infty, t) = 0$.
(b) $u_x(0,t) = 2$ and $u(x \to \infty, t) = 0$.
(c) $u_x(0,t) = -u^2(0,t)$ and $u(x \to \infty, t) = 0$.

(d) $\int_0^\infty x^2 u\,dx = 1$ and $u(x \to \infty, t) = 0$.
(e) $u_x(0,t) = -u(0,t)$ and $u(x \to \infty, t) = 0$.
Hint: Show that this problem does not have a solution following (i, ii), but can still be obtained from solving an ODE problem.

5.5 Determine the scale factors and similarity function ODE for the similarity solution $u(x,t)$ on $-\infty < x < \infty$ for

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u^4, \qquad u(|x| \to \infty) = 0.$$

5.6 Consider the problem for $u(x,t)$ on $-\infty < x < \infty$, see Fig. 5.3,

$$\frac{\partial u}{\partial t} = -\frac{\partial}{\partial x}\left(u^3 \frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial x}\left(u \frac{\partial^3 u}{\partial x^3}\right), \qquad \int_{-\infty}^{\infty} u\,dx = 1.$$

(a) Apply the scaling (5.1) in order to determine the self-similar solution in the form $u(x,t) = t^\alpha f(\eta)$ with $\eta = x/t^\beta$ whose amplitude decays, and width broadens with increasing time as shown in Fig. 5.3 (left). Determine the ODE for the similarity function $f(\eta)$. The corresponding PDE solution is called an *infinite-time defocusing solution* since it spreads indefinitely.
(b) Under other conditions, solutions may have a *critical time* t_c such that, when the critical time is approached the solution will diverge, $u(t \nearrow t_c) \to \infty$. Using (5.1) with $t = t_c + \mathsf{T}\tilde{t}$ (with $\tilde{t} \le 0$), determine the ODE for the similarity function $\hat{f}(\hat{\eta})$. This is called a *finite-time focusing* or *finite-time blow-up* solution; it diverges in amplitude and narrows in width as the critical time is approached.

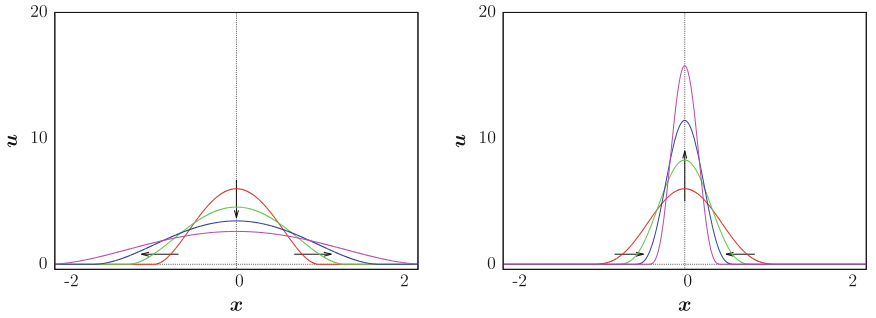


Fig. 5.3 Exercise 5.6: (*Left*) infinite-time spreading solution, (*Right*) finite-time blow-up solution

5.7 The solution $u(r,t)$ of the following problem describes the height of a circular drop of fluid spreadingon a dry surface:

$$\frac{\partial u}{\partial t} = \frac{1}{4r}\frac{\partial}{\partial r}\left(ru\frac{\partial u}{\partial r}\right)$$

on $0 \le r < \infty$ with

$$\left.\frac{\partial u}{\partial r}\right|_{r=0} = 0, \qquad \int_0^\infty u(r,t) r\, dr = 4.$$

The solution is positive over a finite range $0 \le r \le r_*(t)$ with $u(r_*(t),t) = 0$ defining a moving "edge" with no fluid beyond this position. In other words, take the solution beyond the edge $r = r_*(t)$ to be zero.

(a) Show that this problem is scale invariant.
(b) Determine the ODE for the similarity function $f(\eta)$.
(c) Determine the similarity solution $u(r,t) = t^\alpha f(\eta)$ and $r_*(t)$. (Hint: Explain why $\int_0^\infty ur\, dr = \int_0^{r_*} ur\, dr$)

5.8 Consider the system of two coupled PDEs for $h(x,t)$ and $u(x,t)$,

$$\frac{\partial h}{\partial t} + \frac{\partial (hu)}{\partial x} = 0, \qquad \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \frac{\partial^3 h}{\partial x^3} + h^2 u.$$

(a) Let $h = \mathsf{H}\tilde{h}(\tilde{x},\tilde{t})$, $u = \mathsf{U}\tilde{u}(\tilde{x},\tilde{t})$, $x = \mathsf{L}\tilde{x}$, and $t = \mathsf{T}\tilde{t}$. Determine scaling relations for $\mathsf{H}, \mathsf{U}, \mathsf{L}$ in terms of T to show that these PDEs are scale invariant.
(b) Let $h = t^\alpha f(\eta)$, $u = t^\beta g(\eta)$ and $\eta = xt^\gamma$. Find α, β, γ for the similarity solution for this problem and write the ODEs satisfied by $f(\eta)$ and $g(\eta)$.

5.9 Consider the problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + 1, \qquad 0 \le x \le \infty,$$

with boundary conditions

$$u(x=0,t) = 0, \qquad \partial_x u(x \to \infty, t) = 0.$$

Find the similarity solution of this problem and then use this to determine a, b in the far-field growth of the solution, $u(x \to \infty, t) = at^b$ and c, d in the derivative of the solution at the boundary, $u_x(0,t) = ct^d$.
Hint: Express your solution in terms of the error function (5.19), which satisfies the asymptotic approximations

$$\operatorname{erf}(x) \sim 1 - \frac{e^{-x^2}}{x\sqrt{\pi}} \quad \text{as } x \to \infty, \qquad \operatorname{erf}(x) \sim \frac{2x}{\sqrt{\pi}} \quad \text{as } x \to 0.$$

Chapter 6
Perturbation Methods

As we described in Chap. 4, physical problems can always be scaled and restated as dimensionless models. The scaling process identifies the relative importance of different physical effects in terms of the magnitudes of the dimensionless parameters that appear. In the absence of actual parameter values, problem-specific analytical and/or numerical methods are typically necessary to make progress towards a general solution. However, if any dimensionless parameters are known to be relatively large or small, then so-called *perturbation methods* can often be employed in order to generate accurate approximations to the solution.

Perturbation methods provide a systematic approach to constructing approximate solutions to equations such as

$$F(x, \varepsilon) = 0, \qquad \mathbf{G}(x, y, \varepsilon) = \mathbf{0}, \qquad \frac{dx}{dt} = H(x, \varepsilon), \quad \ldots, \tag{6.1}$$

in the limit of a vanishing small *perturbation parameter*, $\varepsilon \to 0$. This is accomplished through the introduction of *asymptotic expansions*, whereby the original problem is decomposed into an ordered sequence of simpler sub-problems. The solutions of the sub-problems are then recombined to form an approximate solution to the full original problem. This may sound very similar to superposition principles that are often used to construct solutions of linear ODE and PDE, but a notable difference here is that the original equation (algebraic, ODE or PDE) may be nonlinear.

6.1 Asymptotic Analysis: Concepts and Notation

Asymptotic analysis provides the mathematical framework that justifies perturbation methods. The term *asymptotic* implies a *limit* process and hence every asymptotic result must be given in terms of a stated parameter approaching a limiting value; for example: $e^{-x} \to 1$ as $x \to 0$ or $e^{-x} \to 0$ as $x \to +\infty$.

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9_6

Two functions $f(\varepsilon)$ and $g(\varepsilon)$ are said to be *asymptotically equivalent* in the limit $\varepsilon \to 0$, if

$$\lim_{\varepsilon \to 0} \frac{f(\varepsilon)}{g(\varepsilon)} = 1. \tag{6.2}$$

This relationship is often written more compactly as "$f \sim g$ as $\varepsilon \to 0$". It is important to note that this definition does not simply just mean that f and g have the same limiting values; for functions that approach infinity or vanish in the limit, equivalence states that they have the same limiting rate of growth or decay. For example, in the limit $\varepsilon \to 0$

$$\frac{1}{\varepsilon} + 1 \sim \frac{1}{\varepsilon}, \qquad \tan\left(\tfrac{\pi}{2} - \varepsilon\right) \sim \frac{1}{\varepsilon}, \qquad \sin(\varepsilon) \sim \varepsilon, \qquad \frac{1}{1 - \varepsilon x} \sim 1 + \varepsilon x,$$

but $\sin(\varepsilon) \nsim \varepsilon^2$ despite each vanishing in the limit. We therefore see that asymptotic equivalence does not imply equality of the functions, but it is a necessary condition for them to have the same limiting behaviour. As a consequence, asymptotic equivalence is not a unique relationship (it effectively defines *equivalence classes* of functions). For example, $\cos(\sqrt{\varepsilon}) \sim (1 - \varepsilon/2) \sim e^{\varepsilon} \sim (1 + \varepsilon)$ as $\varepsilon \to 0$.

A related description of asymptotic behaviour is given by the *order relation*, "$f = O(g)$ as $\varepsilon \to 0$" (read as "f is big-Oh of g") defined by

$$\lim_{\varepsilon \to 0} \frac{f(\varepsilon)}{g(\varepsilon)} = A, \tag{6.3}$$

where A is finite. The statement $f = O(g)$, identifies the function $f(\varepsilon)$ as having comparable growth or decay as $g(\varepsilon)$ as $\varepsilon \to 0$. For example, $\sin(\varepsilon) = O(4\varepsilon)$ as $\varepsilon \to 0$, and $N! = O(N^{N+1/2}e^{-N})$ as $N \to \infty$. The latter result is known as Stirling's formula [13], a central result for estimating large combinatorial values in computing and probability applications. The notation $O(1)$ is often used to describe expressions having finite limiting values and separates quantities that are singular ($f \to \infty$ as in $f = O(\varepsilon^{-1})$) from those that vanish ($f \to 0$ as in $f = O(\varepsilon^2)$) as $\varepsilon \to 0$.

There is also a "little oh" relation, "$f = o(g)$" as $\varepsilon \to 0$ (also written as "$f \ll g$") describing f being *asymptotically smaller than* g if

$$\lim_{\varepsilon \to 0} \frac{f(\varepsilon)}{g(\varepsilon)} = 0. \tag{6.4}$$

The $o(\cdot)$ notation is used to indicate weak effects due to smaller ("*higher order*") terms that can be neglected in comparison with other larger terms. Since, for example, $\varepsilon^2 = o(\varepsilon)$ we may then write $2\varepsilon + 3\varepsilon^2 \sim 2\varepsilon$. The $o(\cdot)$ notation is also often used

to state that the difference between two results is smaller than a specified reference quantity, as in

$$e^{\varepsilon} - \left(1 + \varepsilon + \tfrac{1}{2}\varepsilon^2\right) = o(\varepsilon^2) \quad \Longrightarrow \quad e^{\varepsilon} = 1 + \varepsilon + \tfrac{1}{2}\varepsilon^2 + o(\varepsilon^2). \tag{6.5}$$

We note that our descriptions of these asymptotic relations in terms of limits of ratios are actually convenient simplifications of more rigorous definitions. In particular, $f = O(g)$ if $|f| \leq A|g|$ for all sufficiently small ε [13, 72]. The simplifications are sufficient for most cases, but can fail if the limit is not defined. For example, it is true that $\sin(x) = O(1)$ as $x \to \infty$ since $|\sin(x)|$ is bounded by one, even though is does not have a limit. Asymptotic equivalence is more generally defined along the lines of (6.5), in terms of the difference between the functions being small, $f \sim g$ if $f - g = o(g)$.

6.2 Asymptotic Expansions

Having established the order notation, we can now describe the structure of basic asymptotic expansions. Consider a function $x(t, \varepsilon)$ that can be expanded in terms of a "separation of variables"-type series

$$x(t, \varepsilon) = \delta_0(\varepsilon)x_0(t) + \delta_1(\varepsilon)x_1(t) + \delta_2(\varepsilon)x_2(t) + \cdots, \tag{6.6}$$

where it is assumed that all $x_n = O(1)$ and the *gauge functions* $\{\delta_n(\varepsilon)\}$ are asymptotically ordered in size as $\varepsilon \to 0$ so that

$$\delta_0(\varepsilon) \gg \delta_1(\varepsilon) \gg \delta_2(\varepsilon) \gg \cdots. \tag{6.7}$$

Equations (6.6) and (6.7) define the basic form of an *asymptotic expansion (AE)* and, in particular, (6.6) separates the dependence on the asymptotic parameter ε in each term from the coefficients x_n (which may, in general, also be functions of other independent variables and parameters in the model). The first term in the expansion (6.6) is usually referred to as the *leading order term*, $x \sim \delta_0 x_0$.

There are subtle differences in the way one might write the basic content of the first few terms of an asymptotic expansion based on our previous discussion of asymptotic ordering. Consider

$$\begin{aligned} x &\sim \delta_0 x_0 + \delta_1 x_1 + \delta_2 x_2, \\ x &= \delta_0 x_0 + \delta_1 x_1 + O(\delta_2), \\ x &= \delta_0 x_0 + \delta_1 x_1 + o(\delta_1). \end{aligned} \tag{6.8}$$

The first expression gives the first three terms explicitly, the second gives two terms and an estimate on the asymptotic order of the remainder of the expansion, and the third expression does not predict δ_2 but just states that the omitted contributions are all smaller than the last given $O(\delta_1)$ term which thus effectively provides only the information contained in the two term estimate $x \sim \delta_0 x_0 + \delta_1 x_1$.

A simple example of an asymptotic expansion is provided by the Taylor series of a smooth function $x(t)$ expanded in a neighbourhood of a point t_*

$$x(t) = x(t_*) + \varepsilon x'(t_*) + \tfrac{1}{2!}\varepsilon^2 x''(t_*) + \tfrac{1}{3!}\varepsilon^3 x'''(t_*) + \cdots, \tag{6.9}$$

where $\varepsilon = t - t_*$ (cf. (6.6)); in this case, the small parameter is ε, the separation between the fixed point t_* and the variable t. We know from calculus that in the limit $\varepsilon \to 0$, an accurate estimate of the value of $x(t)$ in a neighbourhood of t_* can be obtained from a limited number of terms in the expansion.

For some functions, the asymptotic expansion can be handled by symbolic algebra programs (such as `Maple` or `Mathematica`). In `Maple`, for example, to calculate the first six terms in the expansion of $x(t) = e^{\tan t}$ for $t \to 0$ (here t is the small parameter), the command is simply

```
> series(exp(tan(t)),t=0);
```

$$1 + t + \frac{1}{2}t^2 + \frac{1}{2}t^3 + \frac{3}{8}t^4 + \frac{37}{120}t^5 + O(t^6).$$

`Maple` can also generate asymptotic expansions in other limits, such as for $x(t) = 1/(1+4t)$ as $t \to \infty$

```
> asympt(1/(1+4*t),t);
```

$$\frac{1}{4}t^{-1} - \frac{1}{16}t^{-2} + \frac{1}{64}t^{-3} - \frac{1}{256}t^{-4} + \frac{1}{1024}t^{-5} + O\left(t^{-6}\right),$$

where the gauge functions are now inverse powers of t; this result can also be derived as a Taylor series expansion with respect to the variable $t = 1/\varepsilon$ in the limit $\varepsilon \to 0$.

We note that asymptotic expansions can take more complex forms than (6.6). For example, the function $x(t, \varepsilon) = \exp(\sin(\varepsilon\sqrt{t})/\varepsilon^2)$, has the asymptotic expansion

$$x \sim e^{\sqrt{t}/\varepsilon}\left(1 - \frac{1}{6}\varepsilon t^{3/2} + \frac{1}{72}\varepsilon^2 t^3 + \cdots\right) \quad \text{as } \varepsilon \to 0, \tag{6.10}$$

which is clearly not of the form (6.6). Our investigations of solutions to equations like (6.1) will be further complicated by the fact that the form of the solution is not known beforehand and so we will need to find the gauge functions as well as the coefficients.

6.2.1 Divergence of Asymptotic Expansions

Example (6.10) indicates that asymptotic expansions exist even in cases where convergent Taylor series do not. In fact, summed over all terms, asymptotic expansions

can diverge, but, as we will shall show, this does not prevent partial sums like (6.8) from providing accurate results in the vanishing limit of the perturbation parameter.

In order to illustrate the above points, we consider an example of a function defined in terms of an integral, namely

$$I(\varepsilon) = \int_0^\infty \frac{e^{-t}}{1+\varepsilon t}\, dt; \tag{6.11}$$

such a definition is rather common and includes Fourier and Laplace integrals, as well as Mellin, Hankel and many other integral transformations, and a number of special functions, such as the Gamma, Beta, Airy and error functions [11, 13, 29, 56].

The Taylor series (6.9) of the integrand in the limit $\varepsilon \to 0$ leads to the asymptotic expansion

$$I(\varepsilon) \sim 1 - \varepsilon + 2\varepsilon^2 - 6\varepsilon^3 + \cdots = \sum_{n=0}^{\infty} (-1)^n n! \varepsilon^n. \tag{6.12}$$

The standard ratio test for power series requires $(n+1)\varepsilon < 1$, so we find that the series has a zero radius of convergence (i.e. it diverges for all $|\varepsilon| > 0$ at sufficiently large values of n). Furthermore, convergence for ε in an interval around zero, $-\delta < \varepsilon < \delta$, should not have been expected since for $\varepsilon < 0$ the singularity at $t = -1/\varepsilon$ is not integrable. Despite these issues, Fig. 6.1 shows that (6.12) truncated at a finite number of terms provides a good estimate of the value of the integral as $\varepsilon \to 0^+$.

To understand this behaviour, consider (6.11) with $\varepsilon = \frac{1}{10}$; the integral can be numerically evaluated to give $I(0.1) \approx 0.915633$. For (6.12), the terms in the expansion remain asymptotically ordered while $(n+1)\varepsilon < 1$, namely for $n < 10$. Recalling that the first term neglected in the asymptotic expansion gives an estimate of the error (cf. (6.8)), we see that truncating the expansion up to and including ten terms will give a decreasing magnitude for the error, but going beyond ten will lead

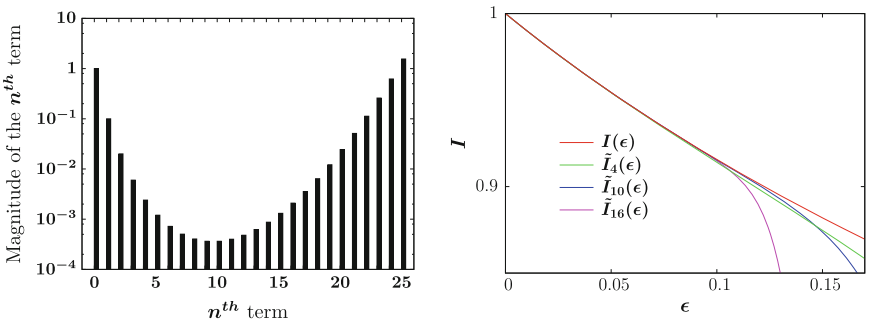


Fig. 6.1 (*Left*) The magnitude of term a_n in (6.12) for $\varepsilon = 0.1$. For $n < 10$, the terms decay in size, but thereafter start to grow. (*Right*) $I(\varepsilon)$ from (6.11) compared with the partial sums, $\tilde{I}_N(\varepsilon)$, of series (6.12) showing the error growing for $\varepsilon \geq 0.1$ for above the optimal truncation

to the error increasing in size (see Fig. 6.1). This is in stark contrast to convergent series, where retaining more terms always reduces the error.

To summarise, asymptotic expansions are always accurate in the limit $\varepsilon \to 0$, but for finite ε, there will be an *optimal truncation* $n = 0, 1, \ldots, N(\varepsilon)$, that minimises the error. In most cases, however, even just the first few terms from the asymptotic expansion can yield an excellent approximation to the true solution, as illustrated by $\tilde{I}_4(\varepsilon)$ in Fig. 6.1. Complications stemming from attempting to interchange limits ($\varepsilon \to 0$ and $n \to \infty$) will appear in several situations and often signal subtle changes of behaviour in the asymptotic expansion.

6.3 The Calculation of Asymptotic Expansions

Rather than calculating asymptotic expansions of given functions, we are usually more interested in constructing an asymptotic expansion of solutions to problems of forms like (6.1).

We begin by calculating asymptotic expansions of solutions to algebraic equations, for which the coefficients $\{x_n\}_{n=0,1,2,\ldots}$ in (6.6) are necessarily constants. It will be useful to classify solutions (and the corresponding asymptotic expansions) in the limit $\varepsilon \to 0$ into the following types:

- *Regular solutions*, which have finite limits: $\lim_{\varepsilon \to 0} x = x_0$ with the leading order gauge function in (6.6) being $\delta_0 \equiv 1$.
- *Vanishing solutions*, which are regular solutions with a zero limit, i.e. $\lim_{\varepsilon \to 0} x = 0$ with the leading order gauge function satisfying $\delta_0(\varepsilon) \ll 1$.
- *Singular solutions*, which have divergent limits: $\lim_{\varepsilon \to 0} |x| = \infty$, corresponding to a finite x_0 with a singular $\delta_0(\varepsilon) \gg 1$.

We now introduce two approaches for constructing solutions in the form of asymptotic expansions known as the *expansion method* and the *iteration method*. We will illustrate both methods applied to a simple example for which the solutions can be found explicitly.

Consider the quadratic equation

$$x^2 - x + \tfrac{1}{4}\varepsilon = 0 \qquad \text{as } \varepsilon \to 0, \tag{6.13}$$

its exact solutions being

$$x_{A,B} = \frac{1 \pm \sqrt{1 - \varepsilon}}{2}. \tag{6.14}$$

Employing the *generalised binomial expansion* for $f(z) = (1+z)^r$, which can be derived by application of Taylor series,

$$(1+z)^r \sim 1 + rz + \tfrac{r(r-1)}{2}z^2 + \tfrac{r(r-1)(r-2)}{3!}z^3 + \cdots \qquad \text{as } z \to 0, \tag{6.15}$$

with $r = \frac{1}{2}$, $z = -\varepsilon$, allows us to write the two solutions as

$$\begin{aligned} x_A &= 1 - \tfrac{1}{4}\varepsilon - \tfrac{1}{16}\varepsilon^2 + \cdots = O(1), \\ x_B &= 0 + \tfrac{1}{4}\varepsilon + \tfrac{1}{16}\varepsilon^2 + \cdots = O(\varepsilon). \end{aligned} \tag{6.16}$$

We note that taking $\varepsilon \ll 1$ directly in (6.13) gives $x^2 - x \approx 0$ yielding approximate solutions $x \approx 0, 1$. In other words, the balance of the first two terms in (6.13) is important in locating the roots, with the third term only slightly adjusting their values (cf. (6.16)). This is the hallmark of a *regular perturbation problem*, for which the limit $\varepsilon \to 0$ only yields regular (i.e. non-singular) solutions with well-defined limits for $\varepsilon \to 0$.

We will begin our analysis with the more concise 'expansion method', which benefits from an a priori assumption on the form for the asymptotic expansion being sought.

6.3.1 The Expansion Method

In order to solve (6.13) without explicit reference to the exact solutions, we assume the solutions are regular and have $\delta_n = \varepsilon^n$ for $n = 0, 1, 2, \ldots$ (i.e. the sequence of gauge functions $\{\delta_n(\varepsilon)\}$ is already known), so that

$$x = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \cdots . \tag{6.17}$$

This is a commonly occurring asymptotic expansion for regular solutions, which can be thought of as being a Taylor series expansion of the solution with respect to the parameter ε around $\varepsilon = 0$.

Substituting (6.17) into Eq. (6.13) yields

$$(x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \cdots)^2 - (x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \cdots) + \frac{1}{4}\varepsilon = 0,$$

and ordering terms in powers of ε yields

$$\left(x_0^2 - x_0\right) + \varepsilon\left(2x_1x_0 - x_1 + \frac{1}{4}\right) + \varepsilon^2(x_1^2 + 2x_0x_2 - x_2) + \cdots = 0 + \varepsilon 0 + \varepsilon^2 0 + \cdots .$$

Assuming that the coefficients are $O(1)$, requiring both sides of the equation to balance leads to the system of equations

$$\begin{array}{lcll} O(\varepsilon^0): x_0^2 - x_0 = 0 & \Rightarrow & x_0 = 1 & \text{or } x_0 = 0, \\ O(\varepsilon^1): 2x_1x_0 - x_1 + \frac{1}{4} = 0 & \Rightarrow & x_1 = -1/4 \;\; | & x_1 = 1/4, \\ O(\varepsilon^2): x_1^2 + 2x_0x_2 - x_2 = 0 & \Rightarrow & x_2 = -1/16 \; | & x_2 = 1/16, \end{array} \tag{6.18}$$

and so on for higher orders. Note that only the leading order equation is nonlinear; subsequent corrections depend on which of the leading order solutions is considered. Comparing the coefficients resulting from (6.18) with (6.16), we observe that we have obtained the asymptotic expansions for both solutions of (6.13).

6.3.2 The Iteration Method

In contrast to the expansion method, we no longer assume a form for the entire asymptotic expansion, but instead only look (initially) at the leading order term $x \sim x_0\delta_0(\varepsilon)$, with both x_0 and $\delta_0(\varepsilon)$ to be determined. There are two fundamental assumptions made in relation to the leading order term:

(i) For every nontrivial solution (i.e. not all $x_n = 0$), the leading order term must be nontrivial: $x_0 \neq 0$ and $\delta_0 \neq 0$.[1]
(ii) The leading order coefficient is finite, $x_0 = O(1)$ as $\varepsilon \to 0$ (i.e. we are looking for regular solutions; we tackle singular solutions later).

Substituting the leading order term into (6.13) yields

$$\underbrace{x_0^2\delta_0^2}_{(1)} - \underbrace{x_0\delta_0}_{(2)} + \underbrace{\tfrac{1}{4}\varepsilon}_{(3)} = 0. \tag{6.19}$$

In order for this equation to hold as $\varepsilon \to 0$, at least two of the terms must balance in asymptotic scales, with the remaining terms being *sub-dominant* (i.e. asymptotically smaller than the retained terms as $\varepsilon \to 0$). The smaller terms can then be neglected in determining the leading order solution. The set of dominant terms that balance to yield the leading order solutions are called the *dominant terms* and this argument is referred to as the *principle of dominant balance*.

Ignoring all $O(1)$ coefficients for the time-being, consider the three possibilities for balancing the asymptotic gauges of the potential dominant terms in (6.19):

$$\begin{array}{lcl} \text{(a) Terms (1, 2): } \delta_0^2 = \delta_0 & \Rightarrow & \delta_0 = 1 \\ \text{(b) Terms (1, 3): } \delta_0^2 = \varepsilon & \Rightarrow & \delta_0 = \sqrt{\varepsilon} \\ \text{(c) Terms (2, 3): } \delta_0 = \varepsilon & \Rightarrow & \delta_0 = \varepsilon \end{array} \tag{6.20}$$

[1] Trivial solutions ($x \equiv 0$) are exact solutions only if they satisfy the full problem for all ε.

In case (a), $\delta_0 = 0$ is excluded by the first fundamental assumption (that δ_0 should be non-zero). It is also clear that for this problem, all three terms cannot be of the same asymptotic order.

For each of the proposed balances, we must now attempt to verify the principle of dominant balance. Namely, it is *essential* to confirm that all sub-dominant terms which have been omitted from the dominant balance are indeed smaller than the dominant terms. When this occurs, the balance is called a *distinguished limit*.

For the three possible balances in (6.20) we see that

$$\begin{array}{lll} \text{(a) Terms (1, 2)} : \delta_0^2 = \delta_0 = O(1) & \gg & \text{Term (3)} : epsilon = O(\varepsilon) \\ \text{(b) Terms (1, 3)} : \delta_0^2 = \varepsilon = O(\varepsilon) & \not\gg & \text{Term (2)} : \delta_0 = O(\sqrt{\varepsilon}) \\ \text{(c) Terms (2, 3)} : \delta_0 = \varepsilon = O(\varepsilon) & \gg & \text{Term (1)} : \delta_0^2 = O(\varepsilon^2) \end{array}$$

In case (b), we see that the second term is not sub-dominant to the first and third terms, and so is not a valid balance. This leaves cases (a) and (c) as valid dominant balances leading to distinguished limits.

The coefficients in each case can now be obtained:

(a) $\underline{\delta_0 = 1}$: Eq. (6.19) becomes $x_0^2 - x_0 + \frac{1}{4}\varepsilon = 0$. Thus, as $\varepsilon \to 0$, the leading order dominant balance determines the leading order coefficient from

$$x_0^2 - x_0 = 0 \qquad \Rightarrow \qquad x_0 = 1,$$

where we reject the root $x_0 = 0$ in line with assumption (i) above. We see that $x_0 = 1$ corresponds to the first term in x_A from (6.16).

(c) $\underline{\delta_0 = \varepsilon}$: Eq. (6.19) takes the form $\varepsilon x_0^2 - x_0 + \frac{1}{4} = 0$ (where we have divided through by the common factor ε). The rescaled leading order equation yields $x_0 = 1/4$, which together with $\delta_0 = \varepsilon$ determines the first term in $x_B \sim \frac{1}{4}\varepsilon$ from (6.16).

Further terms in the expansion of each solution can be obtained by repeating the process: determining the distinguished limits for the gauge functions $\delta_i(\varepsilon)$ and the values for the coefficients x_i using

$$\begin{array}{ll} x_A \sim 1 + \delta_{1A} x_{1A} & \text{and } x_B \sim \frac{1}{4}\varepsilon + \delta_{1B} x_{1B}, \\ x_A \sim 1 + \delta_{1A} x_{1A} + \delta_{2A} x_{2A} & \text{and } x_B \sim \frac{1}{4}\varepsilon + \delta_{1B} x_{1B} + \delta_{2A} x_{2A}, \end{array}$$

and so on, substituting into (6.13) at each step and requiring that the gauge functions be asymptotically ordered, i.e. $1 \gg \delta_{1A} \gg \delta_{2A}$ and $\varepsilon \gg \delta_{1B} \gg \delta_{2B}$. In this manner, we can obtain the expansions of x_A and x_B from (6.16).

Obviously, as more terms are retained in the expansion, the determination of the dominant balance requires an increasing amount of work and this can be a substantial drawback of the iterative approach. Consequently, the expansion method is often favoured in obtaining a rapid result. However, while both roots for this example could be obtained using the expansion (6.17), we will encounter many situations for which the principle of dominant balance will be particularly useful.

6.3.3 Further Examples

Suppose we wish to find the roots of the transcendental equation

$$x^2 - 2x + \varepsilon \sin x = 0 \qquad \text{as} \qquad \varepsilon \to 0. \tag{6.21}$$

This equation does not admit explicit expressions for all roots, but it can be seen that $x = 0$ is a solution for any ε. Setting $\varepsilon = 0$ in (6.21) gives the leading order equation

$$x_0^2 - 2x_0 = 0 \qquad \Rightarrow \qquad x_0 = 0 \text{ or } x_0 = 2. \tag{6.22}$$

We will focus on the expansion for the nontrivial root, see Fig. 6.2.

While we may again proceed by hand, we take this opportunity to demonstrate how symbolic algebra software can be employed to perform some of the computationally intensive calculations. Using `Maple`, we write

```
> eps:=epsilon;
```

$$eps := \varepsilon$$

```
> eq:=x^2-2*x+eps*sin(x);
```

$$eq := x^2 - 2\,x + \varepsilon\,\sin(x)$$

```
> x:=x0+eps*x1+eps^2*x2+eps^3*x3;
```

$$x := x0 + \varepsilon\,x1 + \varepsilon^2 x2 + \varepsilon^3 x3$$

```
> eqser:=series(eq,eps=0,3);
```

$$eqser := x0^2 - 2\,x0 + (2\,x0\,x1 - 2\,x1 + \sin(x0))\,\varepsilon + \left(2\,x0\,x2 + x1^2 - 2\,x2 + \cos(x0)\,x1\right)\varepsilon^2 + O\left(\varepsilon^3\right)$$

```
> order0:=coeff(eqser,eps,0);
```

$$order0 := x0^2 - 2\,x0$$

```
> order1:=coeff(eqser,eps,1);
```

$$order1 := 2\,x0\,x1 - 2\,x1 + \sin(x0)$$

```
> order2:=coeff(eqser,eps,2);
```

$$order2 := 2\,x0\,x2 + x1^2 - 2\,x2 + \cos(x0)\,x1$$

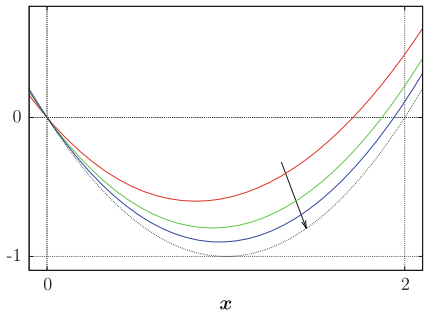


Fig. 6.2 Convergence of the function $f(x, \varepsilon) = x^2 - 2x + \varepsilon \sin x$ for $\varepsilon = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}$ to the $\varepsilon = 0$ limit

The `series()` command performs a Taylor series expansion on the equation. The `order0` equation is simply our leading order problem (6.22). Separating out the coefficient of $O(\varepsilon^1)$ and substituting the value $x_0 = 2$ into `order1` yields the value for $x_1 = -\frac{1}{2}\sin 2$. Continuing on, the `order2` equation gives $x_2 = \frac{1}{4}\sin 2 - \frac{1}{8}\sin^2 2$. The expansion up to $O(\varepsilon^3)$ for the nontrivial root is therefore

$$x = 2 - \frac{1}{2}(\sin 2)\varepsilon + \left(\frac{1}{4}\sin 2 - \frac{1}{8}\sin^2 2\right)\varepsilon^2 + O(\varepsilon^3).$$

To recap, when the expansion method works, it is a very straightforward approach to carry out, either by hand or with the aid of computer software. The limitation of the method is that it will fail to construct solutions that are not of the form assumed for the asymptotic expansion. The trial solution need not be (6.17) (experience can provide good guesses), but the choice of the gauge functions δ_n must be known before this method can be used.

Our next example introduces a problem for which the expansion method using the asymptotic expansion (6.17) fails, and consequently illustrates the strengths of the iteration method. Consider the $\varepsilon \to 0$ limit of the equation

$$(x-1)^2 - 9\varepsilon = 0. \tag{6.23}$$

Substituting the expansion (6.17) into Eq. (6.23) leads to the system of equations

$$\begin{array}{lll} O(\varepsilon^0): & (x_0-1)^2 = 0 & \Rightarrow \quad x_0 = 1 \text{ or } x_0 = 1, \\ O(\varepsilon^1): & 2x_1(x_0-1) - 9 = 0 & \Rightarrow \quad -9 = 0 \quad (\times ?) \\ O(\varepsilon^2): & x_1^2 + 2(x_0-1)x_2 = 0 & \\ \vdots & & \end{array}$$

Substituting the leading order double root into the $O(\varepsilon)$ equation has yielded $-9 = 0$, a clear contradiction. This indicates that our choice for the expansion of x was incorrect, namely that the solutions of (6.23) are not of the assumed form (6.17). In fact, as is trivial to verify, the exact solutions are $x = 1 \pm 3\sqrt{\varepsilon}$.

Applying the iteration method to (6.23) with $x \sim x_0\delta_0(\varepsilon) + x_1\delta_1(\varepsilon)$, we would find directly that the first two gauge functions are given by $\delta_0 = 1$ and $\delta_1 = \sqrt{\varepsilon}$ and the coefficients are $x_0 = 1, 1$ and $x_1 = \pm 3$; looking for higher-order terms would yield coefficients $x_i = 0$ for $i \geq 2$.

Problems having degenerate (repeated) leading order roots should *always* be treated with caution regarding their expansions. There are also many problems requiring non-algebraic gauge functions ($\delta(\varepsilon) \neq \varepsilon^\alpha$); such problems frequently arise from transcendental equations (see [47, 49]).

6.4 A Regular Expansion for a Solution of an ODE Problem

The expansion and iteration methods also extend to solving ODE and PDE problems, the only major difference being that at each order we will have to determine a function rather than just a constant coefficient.

We now illustrate the application of the expansion method to find the solution of an ordinary differential equation using the projectile problem (4.7a) introduced in Chap. 4,

$$\frac{d^2x}{dt^2} = -\frac{1}{(1+\varepsilon x)^2}, \qquad x(0) = 1, \qquad x'(0) = \alpha, \tag{6.24}$$

in the limit $\varepsilon \to 0$. We begin by substituting the asymptotic expansion $x(t) \sim x_0(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + \cdots$ into the ODE and initial conditions. The expansion for the left-hand side of the ODE is simply

$$\frac{d^2x}{dt^2} = x_0''(t) + \varepsilon x_1''(t) + \varepsilon^2 x_2''(t) + \cdots$$

and using the binomial expansion (6.15) with $r = -2$, $z = \varepsilon x$, the right-hand side becomes

$$-\frac{1}{(1+\varepsilon x)^2} = -1 + 2\varepsilon x - 3\varepsilon^2 x^2 + 4\varepsilon^3 x^3 + \cdots \tag{6.25}$$

In fact, this needs to be further expanded using the asymptotic expansion for $x(t)$ to yield

$$\begin{aligned} &= -1 + 2\varepsilon(x_0 + \varepsilon x_1 + \cdots) - 3\varepsilon^2(x_0 + \varepsilon x_1 + \cdots)^2 + \cdots \\ &= -1 + 2\varepsilon x_0 + \varepsilon^2(2x_1 - 3x_0^2) + \varepsilon^3(2x_2 - 6x_0x_1 + \cdots) + O(\varepsilon^4). \end{aligned}$$

The initial conditions provide initial conditions on the coefficient functions $x_n(t)$ of the asymptotic expansion through a comparison at each order in ε

$$x(0) = 1 \quad \Longrightarrow \quad x_0(0) + \varepsilon x_1(0) + \varepsilon^2 x_2(0) + \cdots = 1 + \varepsilon 0 + \varepsilon^2 0 + \cdots,$$

$$x'(0) = \alpha \quad \Longrightarrow \quad x_0'(0) + \varepsilon x_1'(0) + \varepsilon^2 x_2'(0) + \cdots = \alpha + \varepsilon 0 + \varepsilon^2 0 + \cdots.$$

We can now separate the original ODE problem (6.24) into sub-problems for each $x_n(t)$ at $O(\varepsilon^n)$

$$\begin{array}{llllll} O(\varepsilon^0): & x_0'' = -1 & & x_0(0) = 1 & & x_0'(0) = \alpha, \\ O(\varepsilon^1): & x_1'' = 2x_0 & & x_1(0) = 0 & & x_1'(0) = 0, \\ O(\varepsilon^2): & x_2'' = 2x_1 - 3x_0^2 & & x_2(0) = 0 & & x_2'(0) = 0, \end{array}$$

and so on. The solution to the higher order problems depends on the solutions from the lower order ones and so we must solve the sub-problems in sequence. Solving the

$O(\varepsilon^0)$ problem yields the leading order solution, $x_0(t) = -\frac{1}{2}t^2+\alpha t+1$. Substituting x_0 into the $O(\varepsilon^1)$ equation we obtain,

$$x_1'' = -t^2 + 2\alpha t + 2 \qquad \Longrightarrow \qquad x_1(t) = -\tfrac{1}{12}t^4 + \tfrac{\alpha}{3}t^3 + t^2$$

and results at higher orders follow analogously. Reassembling the asymptotic expansion for the solution gives

$$x(t) = \left(-\tfrac{1}{2}t^2 + \alpha t + 1\right) + \varepsilon\left(-\tfrac{1}{12}t^4 + \tfrac{\alpha}{3}t^3 + t^2\right) + O(\varepsilon^2). \tag{6.26}$$

Physically, the flight of the projectile ends when the ground is reached ($0 \le t \le t_*$ for $x(t_*) = 0$), but mathematically, there is nothing stopping us from considering the behaviour predicted by (6.26) for longer times. Note that the asymptotic ordering of the expansion breaks down with $O(x_0) = O(\varepsilon x_1) = O(1/\varepsilon)$ when $t = O(1/\sqrt{\varepsilon})$ and the construction of the solution would also break down at this point, since the assumption that $|\varepsilon x_1| \ll |x_0|$ implicit in (6.25) would be violated. This is a common occurrence for asymptotic expansions and we will also see examples of systems in Chaps. 7 and 8 where such difficulties arise. Such a breakdown indicates a transition in scaling regimes and the resolution of the problem involves identification of the appropriate new scaling. In the next section, we will consider this issue further in the context of determining solutions to singular perturbation problems.

6.5 Singular Perturbation Problems

Problems having one or more solutions that exhibit singular (divergent) behaviour in the limit $\varepsilon \to 0$ are called *singular perturbation problems*. The singular solutions are not obtainable from asymptotic expansions appropriate to regular solutions, such as (6.17); attempts to use such asymptotic expansions will return only a subset of the solutions of the full problem, or at worst, no solutions at all (see Exercise 6.3).

This behaviour is inherent to all classes of singular perturbation problems:

- For algebraic equations, a singularly perturbed Nth degree polynomial will only have M regular solutions, with $M < N$ if the leading order equation reduces to an Mth degree polynomial for x_0 due to terms of the form $\varepsilon^p x^K$ ($p > 0$) vanishing for $M < K \le N$.
- For differential equations, if the limit $\varepsilon \to 0$ causes the highest order derivative to vanish (e.g. $\varepsilon^p d^N x/dt^N$), the leading order solution of the remaining lower order problem will typically not have enough degrees of freedom to satisfy all of the initial or boundary conditions imposed on the original problem.

- Likewise, in other classes of singular problems, the $\varepsilon = 0$ leading order problem may be dramatically reduced from the full problem for $\varepsilon > 0$, for example: a singular PDE reducing to an ODE, a singular ODE reducing to an algebraic equation, or a system of equations reducing to a single equation.

The non-regular solutions are not truly lost by such reductions, and can be recovered through appropriate rescalings of the original problem. Which solutions are obtained depends on the scaling of the problem and the form of the asymptotic expansion assumed for the solution.

As an example, consider the $\varepsilon \to 0$ limit of the (immediately solvable) algebraic equation

$$\varepsilon x^2 - 2x + 1 = 0. \tag{6.27}$$

Substituting $\varepsilon = 0$ into (6.27) gives

$$-2x + 1 = 0 \qquad \Rightarrow \qquad x = \tfrac{1}{2}. \tag{6.28}$$

This leading order equation therefore only yields one of the two roots expected from the second-degree equation (6.27). The exact solutions of (6.27) can be written as

$$x = \frac{1 \pm \sqrt{1-\varepsilon}}{\varepsilon} \sim \frac{1 \pm \left(1 - \frac{1}{2}\varepsilon - \frac{1}{8}\varepsilon^2 + O(\varepsilon^3)\right)}{\varepsilon}, \tag{6.29}$$

from which we see that

$$x_A \sim \frac{2}{\varepsilon} - \frac{1}{2} - \frac{\varepsilon}{8}, \qquad x_B \sim \frac{1}{2} + \frac{\varepsilon}{8} + \frac{\varepsilon^2}{16}. \tag{6.30}$$

Equation (6.28) produced only the leading order term of the regular solution x_B. The x_A solution diverges as $\varepsilon \to 0$ and cannot be expressed in terms of the regular expansion (6.17).

Substituting the leading term from each of these solutions back into (6.27) helps identify the cause of the discrepancy

$$x_A: \quad \varepsilon\left(\tfrac{2}{\varepsilon}\right)^2 - 2\left(\tfrac{2}{\varepsilon}\right) + 1 = 0, \qquad x_B: \quad \varepsilon\left(\tfrac{1}{2}\right)^2 - 2\left(\tfrac{1}{2}\right) + 1 = 0.$$

The solutions are determined by different dominant balances in Eq. (6.27). For x_A, the first two terms balance at $O(1/\varepsilon)$ while the third term is sub-dominant at $O(1)$. For x_B, the second and third terms balance at $O(1)$ with the first term being sub-dominant at $O(\varepsilon)$ (this being the regular solution that was obtained from (6.28)).

6.5.1 Rescaling to Obtain Singular Solutions

We now outline a systematic procedure for how the previously introduced methods for regular perturbation problems can be extended to handle singular solutions:

(1) Substitute $x = \delta_0(\varepsilon)X$ into the given problem.
(2) Choose $\delta_0(\varepsilon)$ so as to produce a consistent dominant balance; verify that all neglected terms are indeed sub-dominant. Different $\delta_0(\varepsilon)$'s lead to different dominant balances, and considering all of the possible choices will yield all of the regular and singular solutions.
(3) Factor out any common ε-scalings to yield a regular perturbation problem in X. Then solve this problem using either the iteration or expansion methods to generate regular solutions corresponding to this distinguished limit.
(4) Rescale X by δ_0 to obtain x in final form.

We will illustrate this methodology on the example from the previous section. Substituting $x = \delta_0 x_0$ into (6.27) yields

$$\underbrace{\varepsilon\delta_0^2 X^2}_{(1)} - \underbrace{2\delta_0 X}_{(2)} + \underbrace{1}_{(3)} = 0. \tag{6.31}$$

We compare the orders of magnitude of the terms (cf. Sect. 6.3.2). There are three possible balances:

$$\begin{array}{lll} \text{(a) Terms (2, 3): } \delta_0 = 1 & \Rightarrow & \delta_0 = 1 \\ \text{(b) Terms (1, 3): } \varepsilon\delta_0^2 = 1 & \Rightarrow & \delta_0 = 1/\sqrt{\varepsilon} \\ \text{(c) Terms (1, 2): } \varepsilon\delta_0^2 = \delta_0 & \Rightarrow & \delta_0 = 1/\varepsilon \end{array}$$

In case (a), the omitted first term is sub-dominant, $O(\varepsilon)$, and so this is a consistent balance; this balance yields the regular solution, x_B. In case (b), the dominant balance is at $O(1)$, while the omitted second term is $O(1/\sqrt{\varepsilon}) \gg 1$ in the limit $\varepsilon \to 0$ and is not sub-dominant. Hence this balance is inconsistent and must be rejected. Finally, in case (c) the balancing terms are $O(1/\varepsilon)$, while the third term is $O(1) \ll 1/\varepsilon$, therefore yielding a second consistent dominant balance; this case yields the singular solution.

In order to investigate the singular solution, we write $x = X/\varepsilon$ and substitute into (6.27) to arrive at the (rescaled) regular problem

$$X^2 - 2X + \varepsilon = 0, \tag{6.32}$$

where we have multiplied through by ε. Expanding $X(\varepsilon)$ as a regular expansion (here $X \sim X_0 + \varepsilon X_1 + \varepsilon^2 X_2$ will work) and seeking the leading order term yields

$$X_0^2 - 2X_0 = 0, \tag{6.33}$$

which has solutions $X_0 = 0$ and $X_0 = 2$. We ignore the root $X_0 = 0$ because it is not a nontrivial leading order term.[2] The other root yields the singular solution $x_A \sim 2/\varepsilon$.

A particular difference between the iteration method described in Sect. 6.3.2 and the current methodology is worth noting. Both seek to determine the leading order gauge function δ_0, and the set of coefficients x_n versus X_n will be identical. The iteration method seeks to identify one successive term in the asymptotic expansion at each iteration. In contrast, the rescaling approach re-casts the entire problem into a new (regular perturbation) form for $X(\varepsilon) = O(1)$ (as in (6.32)); whether to then use the expansion method or iteration method to solve for X is left as a separate decision. In terms of the asymptotic expansion of the solution, we have

$$\begin{aligned} x &\sim \delta_0 X(\varepsilon) \\ &\sim \delta_0 \left(X_0 + \tilde{\delta}_1 X_1 + \tilde{\delta}_2 X_2 + \tilde{\delta}_3 X_3 + \cdots \right) \\ &\sim \delta_0 x_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_3 + \cdots , \end{aligned} \tag{6.34}$$

so that the leading order gauge function δ_0 scales through the $\tilde{\delta}$ gauge functions in the expansion of the rescaled solution.

Singular problems for ODE and PDE introduce additional complexities and are of particular interest as they frequently arise in applications. The following two chapters consider these scenarios in detail.

6.6 Further Directions

There are many variations of the methods used in this chapter. Reference [47] employs an iterative procedure, where the original equation must be written in a form that is compatible with the contraction mapping theorem and allows for greater analysis of the convergence of the asymptotic expansion. The method described in [49] is somewhat more similar to that presented here; there, the gauge functions are assumed to be of the form $\delta_n = \varepsilon^{\alpha_n}$ where α_n need not be an integer. Finally, we note that many further approaches for constructing asymptotic expansions for integrals and differential equations build directly on the perturbation methods for algebraic equations described in this chapter [11, 13, 29, 47, 72, 92].

[2] It is a "ghost" of the x_B regular solution, which is $O(\varepsilon)$ in terms of X (and violates requirement (i) in Sect. 6.3.2).

6.7 Exercises

6.1 Use Taylor series expansions for $\sin y$ and e^y for $y \to 0$ and the basic property $\exp(\sum_k a_k) = \prod_k e^{a_k}$ to derive (6.10) for $\varepsilon \to 0$ from $x(t, \varepsilon) = \exp(\sin(\varepsilon\sqrt{t})/\varepsilon^2)$ when $t = O(1)$, but $x \sim 1 + \sqrt{\varepsilon T} + \frac{1}{2}\varepsilon T$ when $t = \varepsilon^3 T$ with $T = O(1)$.

6.2 Consider the limit of $\varepsilon \to 0$ for the equation

$$(x-3)^3 = 24\varepsilon x^2.$$

Solve by iteration[3] to determine $x \sim \delta_0(\varepsilon)x_0 + \delta_1(\varepsilon)x_1 + \delta_2(\varepsilon)x_2$.

6.3 Consider the algebraic equation for $\varepsilon \to 0$

$$\varepsilon^6 x^3 - 5\varepsilon^3 x^2 - 20\varepsilon x + 60 = 0.$$

Show that there are no regular solutions and determine the leading order nontrivial term in the expansion of each of the three solutions.

6.4 Consider the projectile problem for $\varepsilon \to 0$

$$\frac{d^2x}{dt^2} = -\frac{1}{(1+\varepsilon x)^2}, \qquad x(0) = 1, \qquad x'(0) = 3\varepsilon.$$

(a) Let $t^{\max}$ be the time when the projectile reaches its maximum height. How many terms in the expansion of $x(t)$ will you need to determine $t^{\max} = t_0 + \varepsilon t_1 + O(\varepsilon^2)$?

(b) Show that while (6.26) with $\alpha = 3\varepsilon$ could have been used to determine the solution in part (a), this could not be used with the initial condition $x'(0) = 4/\varepsilon$. Assume a singular form $x(t) = X(t)/\varepsilon \sim X_0(t)/\varepsilon + X_1(t)$ to find this solution.

6.5 Consider the problem

$$\frac{dv}{dt} + \varepsilon v^2 + t = 0, \qquad v(0) = 0, \qquad \varepsilon \to 0.$$

(a) Find the first three terms in the expansion of the solution, $v(t) \sim v_0(t) + \varepsilon v_1(t) + \varepsilon^2 v_2(t)$.

(b) Determine the range of times, $0 \le t < O(\varepsilon^\alpha)$, for which the terms in the expansion retain asymptotic ordering, i.e. $v_0 \gg \varepsilon v_1 \gg \varepsilon^2 v_2 \gg \cdots$.

[3] Use of a computer algebra program (`Maple` or `Mathematica`) is recommended for solving many of the more algebraically intensive problems.

6.6 Consider the system of equations

$$\varepsilon x - y = 1$$
$$\varepsilon^2 x + y = 4$$

in the limit $\varepsilon \to 0$.

Explain why setting $\varepsilon = 0$ does not lead to an acceptable leading order solution. Determine the first two terms in the expansions of $x(\varepsilon)$, $y(\varepsilon)$.

6.7 Use the iteration method with $x \sim \delta_0(\varepsilon)x_0 + \delta_1(\varepsilon)x_1$ to find the singular real solution of the equation

$$2x^2 e^{-5x} = 8\varepsilon, \qquad \varepsilon \to 0.$$

Note: This equation has three real roots; the two vanishing roots are not of interest here (they are $x \sim \pm 2\varepsilon^{1/2} + 10\varepsilon$). The one you should locate has $\delta_0(\varepsilon) \to \infty$ (i.e. it is large and positive).
Hint: Take the logarithm of both sides of the equation before beginning iteration.

6.8 Re-consider the solutions of Eq. (6.21), but now in the limit of $\varepsilon \to \infty$: let $\tilde{\varepsilon} = 1/\varepsilon$ and find the real-valued solutions of

$$\sin x + \tilde{\varepsilon}(x^2 - 2x) = 0, \qquad \tilde{\varepsilon} \to 0.$$

Assume $x \sim x_0 + \tilde{\varepsilon}x_1 + \tilde{\varepsilon}^2 x_2$. Note that for $\tilde{\varepsilon} = 0$ the leading order problem has infinitely many solutions, $x_0 = n\pi$ for any integer n.

(a) Determine x_1, x_2.
(b) Plot $F(x, \tilde{\varepsilon}) = \sin x + \tilde{\varepsilon}(x^2 - 2x)$ for $\tilde{\varepsilon} = \frac{1}{100}$. How many zeroes does it have?
(c) Parts (a, b) show that there is a conflict, as $\varepsilon \to 0$ there is a finite number of solutions, namely there is a maximum value for n, call it $N(\tilde{\varepsilon})$, at each value of $\tilde{\varepsilon}$. Requiring that the expansion for x remain asymptotically ordered ($x_0 \gg \varepsilon x_1 \gg \ldots$) suggests one estimate for N, but show that maintaining ordering in the expansion of the equation, given x_0 and x_1 from (a), yields the correct $N(\tilde{\varepsilon})$.

6.9 In the 1930s, Carleman showed that in an appropriate asymptotic limit a diffusion model could be derived from a system of reactive wave equations. Consider the system of PDEs for $p(x, t), q(x, t)$ with $\varepsilon > 0$:

$$\varepsilon^2 \frac{\partial p}{\partial t} + \varepsilon \frac{\partial p}{\partial x} = q - p, \qquad \varepsilon^2 \frac{\partial q}{\partial t} - \varepsilon \frac{\partial q}{\partial x} = p - q.$$

(a) If the solutions are uniform in space (i.e. $p = p(t), q = q(t)$), then p, q satisfy a reversible transformation reaction $P \rightleftharpoons Q$ (see Sect. 1.2). What is the conserved quantity for this reaction system? What is the reaction rate?
(b) If the reactions (the right-hand side terms) are eliminated, find the travelling waves for the reduced equations for $p(x, t)$ and $q(x, t)$. What are the speeds of these waves?

(c) Define $u = p + q$ and $v = (p - q)/\varepsilon$. Combine the full equations for p and q to obtain two equations for u and v. In the limit $\varepsilon \to 0$, use regular perturbation expansions $u(x,t) \sim u_0 + \varepsilon u_1 + \cdots$ and $v(x,t) \sim v_0 + \varepsilon v_1 + \cdots$ to find the leading order equations for u and v, and then derive a PDE for $u_0(x,t)$ alone.

6.10 The nondimensional form of the shallow water equations is

$$\frac{\partial h}{\partial t} + h\frac{\partial u}{\partial x} + u\frac{\partial h}{\partial x} = 0, \qquad \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{1}{\mathrm{Fr}^2}\frac{\partial h}{\partial x} = 0,$$

where $u(x,t)$ is the fluid speed, $h(x,t)$ is the height of the fluid, and Fr is a dimensionless parameter. A uniform, steady solution is given by $h \equiv 1$ and $u \equiv 1$, which can represent a river having uniform speed and depth.
Consider waves generated due to some small disturbance on this steady state. Let

$$h = 1 + \varepsilon\eta(x,t), \qquad u = 1 + \varepsilon v(x,t).$$

(a) For $\varepsilon \to 0$, determine the $O(\varepsilon)$ linearised wave equations for η, v.
(b) Apply the approach of Sect. 2.4 to the linearised system to determine the critical value of the Froude number at which waves from the disturbance change from spreading in both up- and down-stream directions to having all ripples being swept downstream (called *sub- and super-critical behaviours*).

Chapter 7
Boundary Layer Theory

We have seen in Chap. 6 that singularly perturbed problems can have co-existing regular and singular solutions that scale differently as $\varepsilon \to 0$. In the context of physical systems described by differential equations, such structures yield *multi-scale phenomena*. Everyday life yields countless examples of multi-scale phenomena: violent winds in tornadoes surrounded by relatively calm air over large areas, bands of wake behind ships moving in otherwise still waters, cracks forming in uniform solid materials, spots, stripes and other intricate patterns developing in biological systems. In these, and many other contexts, we can separate the behaviour of the system into regions of rapid variation of quantities of interest compared to other larger scale regions of slow variation. Models that can capture such diverse behaviours will allow for multiple distinguished limits, describing balances between different sets of dominant effects in different regions. The relatively narrow regions of rapid variations are generally called *boundary layers*.

Although boundary layers were originally formulated to describe problems in fluid mechanics and aerodynamics [76] (where they generally occur on the boundary of a solid object passing through a surrounding uniform fluid flow), they also describe solutions in broader sets of contexts.

While attempting to directly find a solution of the full problem on an entire domain directly may be very difficult, constructing "partial solutions" on different regions using perturbation expansions can be straightforward. Following the approach introduced in Sect. 6.5, different dominant balances will re-scale the full model into different forms, leading to different regular or singular solutions. Further analysis is then needed to assemble the partial solutions into a complete solution of the full problem. This is accomplished through *asymptotic matching* and leads to this solution methodology being called *the method of matched asymptotic expansions*.

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9_7

7.1 Observing Boundary Layer Structure in Solutions

In order to illustrate the main principles behind boundary layers and matched asymptotics, we first consider a problem for which we can express the solution exactly and examine how it can be separated into pieces stemming from behaviours at different scales.

Consider the linear, constant coefficient ordinary differential equation

$$\varepsilon \frac{d^2y}{dx^2} + 2\frac{dy}{dx} + y = 0 \qquad \text{for } \varepsilon \to 0, \tag{7.1a}$$

on the domain $0 \leq x \leq 1$, subject to the boundary conditions

$$y(0) = 0, \qquad y(1) = 1. \tag{7.1b}$$

Singular behaviour in the solution should be expected since setting $\varepsilon = 0$ in (7.1a) reduces the equation to a first order ODE, whose solution can satisfy only one of the boundary conditions.

The exact solution of (7.1) for any $\varepsilon > 0$ is given by

$$y(x) = \frac{\exp\left(\frac{-1+\sqrt{1-\varepsilon}}{\varepsilon}\, x\right) - \exp\left(\frac{-1-\sqrt{1-\varepsilon}}{\varepsilon}\, x\right)}{\exp\left(\frac{-1+\sqrt{1-\varepsilon}}{\varepsilon}\right) - \exp\left(\frac{-1-\sqrt{1-\varepsilon}}{\varepsilon}\right)} \tag{7.2}$$

(see Fig. 7.1). If ε is small ($0 < \varepsilon \ll 1$), we can expand the arguments of the exponentials to yield

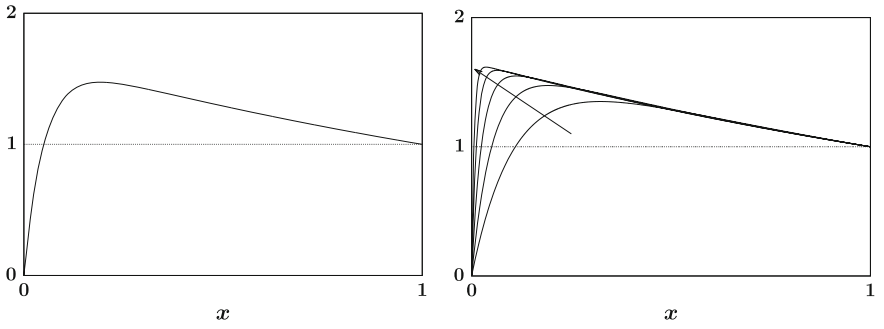


Fig. 7.1 (*Left*) solution (7.2) to problem (7.1) for $\varepsilon = 0.1$. (*Right*) the behaviour of the solution in the limit $\varepsilon \to 0$

$$y(x) = \frac{\exp\left(-\left[\frac{1}{2} + \frac{\varepsilon}{8} + \cdots\right]x\right) - \exp\left(-\left[\frac{2}{\varepsilon} - \frac{1}{2} + \cdots\right]x\right)}{\exp\left(-\left[\frac{1}{2} + \frac{\varepsilon}{8} + \cdots\right]\right) - \exp\left(-\left[\frac{2}{\varepsilon} - \frac{1}{2} + \cdots\right]\right)}$$
$$\sim \frac{\exp\left(-\frac{1}{2}x\right) - \exp\left(-\frac{2}{\varepsilon}x\right)}{\exp\left(-\frac{1}{2}\right) - \exp\left(-\frac{2}{\varepsilon}\right)}, \tag{7.3}$$

The size of the domain is independent of ε, and hence one of the spatial scales should be $x = O(1)$ as $\varepsilon \to 0$, as represented by the $e^{-x/2}$ term in (7.3). The other term there depends on a more rapidly varying spatial scale, x/ε, in which a small ($O(\varepsilon)$) change in x yields a $O(1)$ change in the solution. The distinctness of these two scales makes it possible to construct the solution to (7.1) by seeking its dependence on each scale separately.

Each spatial scale in the problem is associated with a limiting process for the solution of (7.1) as $\varepsilon \to 0$. Fixing $x = O(1)$ in the range $0 < x \leq 1$ gives

$$\lim_{\varepsilon \to 0} y(x) \sim \frac{e^{-x/2} - \text{e.s.t.}}{e^{-1/2} - \text{e.s.t.}} \sim e^{(1-x)/2}, \tag{7.4}$$

where "e.s.t." refers to *exponentially small terms*, of the form $e^{-\alpha/\varepsilon}$ for $\alpha > 0$, that are smaller than all algebraic powers of ε ($\varepsilon^n \gg e^{-\alpha/\varepsilon}$ for $\varepsilon \to 0$) and are treated as negligible in this context. This limiting form of the solution satisfies the boundary condition at $x = 1$, but not the one at $x = 0$ (7.1b).

Note that we have excluded the case $x = 0$ from consideration in (7.4) so that $e^{-2x/\varepsilon}$ is indeed exponentially small. If instead, we consider a small neighbourhood of the origin, $0 \leq x = O(\varepsilon)$, then we must take the dual limit $\varepsilon \to 0$ and $x \to 0$ with the ratio $X = x/\varepsilon$ held fixed. In terms of the new spatial variable X, the limit of (7.3) now becomes

$$\lim_{\varepsilon \to 0} y \sim \frac{e^{-\varepsilon X/2} - e^{-2X}}{e^{-1/2} - e^{-2/\varepsilon}} \sim e^{1/2}(1 - e^{-2X}). \tag{7.5}$$

This limiting form satisfies the boundary condition at $x = 0$ ($X = 0$). The boundary condition at $x = 1$ is not satisfied, but since the right hand boundary position, corresponding to $X = 1/\varepsilon$, violates the assumption $X = O(1)$ made in taking the limit, agreement should not have been expected.

Similar examples are often used in analysis [21] to illustrate functions that have *non-uniform convergence*. In the present context, we see that different limiting properties of the solution are captured by different limiting processes. For the majority of the domain (here, where $x = O(1)$), the solution is given by (7.4) and is called the *outer solution*. In contrast, in the *boundary layer*, or *inner domain*, the solution exhibits singular behaviour that cannot be captured by the outer solution. In this example, the *inner solution* in the boundary layer close to $x = 0$ has a singular derivative, $dy/dx = O(\varepsilon^{-1}) \to \infty$ as $\varepsilon \to 0$.

It is also worth mentioning that problems featuring a separation of scales can be extremely difficult to compute numerically (typically referred to as *stiff problems*). In contrast, solutions of these problems can often be very accurately obtained in terms of inner and outer solutions using perturbation methods.

7.2 Asymptotics of the Outer and Inner Solutions

We now discuss the construction of solutions to singular perturbation problems by calculating perturbation expansions for the outer and inner solutions.

We shall continue to use problem (7.1a, 7.1b) to illustrate the methodology. We begin by attempting to find the outer solution on the *outer domain*, $0 < x \leq 1$, in which we assume that $y(x)$ and all of its derivatives are bounded, smooth and $O(1)$. We assume $y(x)$ can be expanded as a regular perturbation expansion of the form

$$y(x) \sim y_0(x) + \varepsilon y_1(x) + \varepsilon^2 y_2(x) + \cdots \quad \text{as} \quad \varepsilon \to 0. \tag{7.6}$$

Similar to the approach used for (6.24), we substitute this expansion into (7.1a, 7.1b) and separate terms in powers of $\varepsilon \to 0$ yielding the system of sub-problems

$$\begin{array}{lll} O(\varepsilon^0): & 2y_0' + y_0 = 0 & y_0(1) = 1, \\ O(\varepsilon^1): & 2y_1' + y_1 = -y_0'' & y_1(1) = 0, \\ O(\varepsilon^2): & 2y_2' + y_2 = -y_1'' & y_2(1) = 0, \end{array} \tag{7.7}$$

and so on for higher powers of ε. Note that the boundary condition at $x = 0$ from (7.1b) is not included above since $x = 0$ is not within the outer domain. Solving the $O(1)$ problem gives the leading order outer solution

$$y_0(x) = e^{(1-x)/2}, \tag{7.8}$$

which reproduces (7.4). Additional terms in the expansion can be constructed by working through the higher order sub-problems in (7.7) in sequence.

We now turn our attention to constructing the solution in the *inner region*, corresponding to having $x = O(\varepsilon)$. Introducing the rescaling $x = \varepsilon X$ and writing $y(x) = Y(X)$ transforms (7.1a) to

$$\frac{d^2Y}{dX^2} + 2\frac{dY}{dX} + \varepsilon Y = 0, \tag{7.9}$$

with the boundary condition (7.1b) becoming $Y(0) = 0$. Seeking the solution in the form of a regular expansion, $Y(X) = Y_0(X) + \varepsilon Y_1 + \varepsilon^2 Y_2(X) + \cdots$, yields the leading order $O(1)$ problem

$$Y_0'' + 2Y_0' = 0, \qquad Y_0(0) = 0, \tag{7.10}$$

with solution

$$Y_0(X) = A(1 - e^{-2X}). \tag{7.11}$$

Note that one constant, A, remains undetermined in the solution since the ODE in (7.10) is second order, but we are only imposing one side condition. The solutions of the higher order terms, $Y_n(X)$ will similarly add one new undetermined constant at each order in the expansion. Since the inner problem did not uniquely define the inner solution, we examine its relationship with the outer solution in an attempt to fix the unknown constant A.

It is important to note that the inner and outer domains are not mutually exclusive and are valid on regions broader than the strict prescriptions of their spatial scales (here $x = O(\varepsilon)$ and $x = O(1)$ respectively). In the outer solution x is bounded away from zero, but can become small, say $x = O(\sqrt{\varepsilon}) \to 0$. Likewise, in the inner solution x should be small, but $X = x/\varepsilon$ can become large, say $X = O(1/\sqrt{\varepsilon}) \to \infty$.

In fact, it can be shown that there is an *overlap domain*,

$$\text{overlap domain:} \quad \varepsilon \ll x \ll 1, \tag{7.12}$$

between the scales set by the inner and outer domains, where both inner and outer solution are valid (see Fig. 7.2), Loosely speaking, there is a range where x is small (small enough for the inner solution to apply), but not too small (where the outer solution would not apply). Since the original full problem (7.1a, 7.1b) has a unique solution; if the inner and outer solutions are both valid in the overlap domain, they cannot be two distinct solutions and must in fact be two different asymptotic representations of the same solution. This relation is expressed in terms of limits derived from (7.12): (i) $x \ll 1$, the outer variable must approach the inner domain, $x \to 0$ and (ii) $\varepsilon \ll x$ (or after dividing across by ε: $1 \ll X$), the inner variable must approach the outer domain, $X \to \infty$. The resulting limit requirement

$$\lim_{X\to\infty} Y_0(X) = \lim_{x\to 0} y_0(x), \tag{7.13}$$

is called the leading order *asymptotic matching condition* [60, 101]. This principle can be paraphrased as

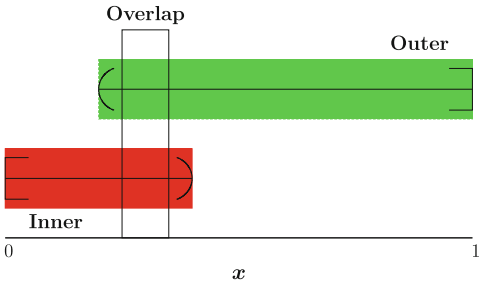


Fig. 7.2 A schematic representation of the inner, outer and overlap domains for problem (7.1a, 7.1b)

"The outer limit of the inner solution equals the inner limit of the outer solution." (7.14)

Applying (7.13) to Y_0 given by (7.11) and y_0 by (7.8) yields

$$\lim_{X\to\infty} A(1-e^{-2X}) = A = \lim_{x\to 0} e^{(1-x)/2} = e^{1/2}, \tag{7.15}$$

thereby determining the constant $A = e^{1/2}$. With this value for A the inner solution (7.11) reproduces the limit (7.5) found from the exact solution.

While we have now determined the leading order inner and outer solutions completely, only a little more work is needed combine the outer and inner solutions to form a *composite* representation of the leading order solution, denoted here by $y_{\text{comp}}(x)$, valid over the entire domain $0 \le x \le 1$. An appropriate form for $y_{\text{comp}}(x)$ is given by the expression

$$y_{\text{comp}}(x) = y_0 + Y_0 - (\text{overlap from matching}), \tag{7.16}$$

where the overlap is simply the contribution found through matching in (7.15). At a formal level, on most of the domain, we have $y \sim y_0$, with the inner solution only becoming significant in the inner domain. Where Y_0 becomes important, we gain equal contributions from both the inner and outer solutions, and so to effectively prevent "double-counting", we must subtract off the overlap.

Another way of expressing (7.16) is to write

$$y_{\text{comp}} = y_0 + Y_{\text{BLC}}, \tag{7.17}$$

where the *boundary layer correction*, Y_{BLC}, is the adjustment to the outer solution made by the boundary layer to satisfy the boundary condition, with

$$Y_{\text{BLC}} \equiv Y_0 - (\text{overlap from matching}), \tag{7.18}$$

where we expect $Y_{\text{BLC}} \to 0$ as $X \to \infty$.

Writing the inner solution (7.11) as $Y_0 = e^{1/2} - e^{(1-4x/\varepsilon)/2}$ and using the overlap from (7.15), the leading order boundary layer correction is

$$Y_{\text{BLC}} = e^{1/2} - e^{(1-4x/\varepsilon)/2} - e^{1/2} = -e^{(1-4x/\varepsilon)/2}, \tag{7.19}$$

which does indeed vanish as $X \to \infty$. Hence, (7.17) gives the leading order solution

$$y_{\text{comp}} = e^{(1-x)/2} - e^{(1-4x/\varepsilon)/2} = \frac{e^{-x/2} - e^{-2x/\varepsilon}}{e^{-1/2}}, \tag{7.20}$$

on the entire domain $0 \leq x \leq 1$. This is comparable to the exact solution (7.2), except for the absence of the exponentially small $e^{-2/\varepsilon}$ term (which cannot be captured in regular expansions such as (7.6)).

7.3 Constructing Boundary Layer Solutions

The presentation in the previous section made use of some prior knowledge of the form of the solution to reduce the overall problem to that of determining the expansions of the inner and outer solutions. In this section, we describe the further steps required to investigate boundary layer problems without any additional given information about the form of the solution.

The full process of constructing a solution involving matched asymptotic expansions requires examining the two additional questions:

- What are the scalings for the inner (and outer) solution(s)?
- Where are the boundary layer(s) located?

The answers to these questions ultimately determine the form of the overall solution, and control which boundary conditions apply to the inner and outer solutions.

For many singularly perturbed differential equations, solutions can be constructed by a step-by-step process:

(1) *The outer solution*: If the problem is in standard form, try a regular expansion, $y(x) = y_0(x) + \varepsilon y_1(x) + \cdots$, for the outer solution. If all of the boundary conditions can be satisfied by this solution, then the problem is complete; otherwise, inner regions will be necessary.

(2) *Find the dominant balances*: The appropriate forms for all of the regular (outer) and singular (inner) solutions of the ODE will be determined by the distinguished limits of the problem. In general, both the independent and dependent variables may need to be scaled to obtain all of the dominant balances,

$$y = \varepsilon^{\beta} Y(X), \qquad X = \frac{x - x_*}{\varepsilon^{\alpha}} \quad \Leftrightarrow \quad x = x_* + \varepsilon^{\alpha} X, \tag{7.21}$$

where the powers α, β and the assumed position of the boundary layer, x_* must all be determined (there may be multiple valid locations for x_*). The outer solution in step (1) assumes $\alpha = 0$, $\beta = 0$.

(3) *The inner solution*: For the singular distinguished limit, write the problem as a rescaled regular problem and seek the solution in the form of an appropriate regular expansion, $Y(X) = Y_0(X) + \varepsilon Y_1(x) + \cdots$.

(4) *Asymptotic matching*: Apply asymptotic matching between the inner/outer solutions (typically via (7.13)) to confirm the consistency of the asymptotic expansion and determine any remaining unknown parameters in the solution.

(5) *The composite solution*: Writing the outer and inner solutions in terms of the original variables and subtracting the overlaps from the matching process to

prevent "double-counting" will produce the leading order solution on the entire domain (7.16). So, in an example with boundary layers at both the left and right boundaries, we write

$$y_{\text{comp}}(x) \sim y_0(x) + Y^L_{\text{BLC}} + Y^R_{\text{BLC}}. \tag{7.22}$$

This description covers broad classes of problems, but as will be seen, in some cases, steps (2, 3, 4) may become a bit intertwined. We also note:

- When boundary layers are necessary, which boundary conditions apply to the outer solution may not be immediately apparent. Hence a general form for the outer solution will be needed initially.
- The boundary conditions as well as the ODE play a role in determining the dominant balances.
- The location of the boundary layer and which boundary conditions apply to the inner solution might not be determined until matching is applied.
- If the inner/outer solutions are not matchable (either limit does not exist, or equation (7.13) cannot be satisfied) then the assumed choice of boundary layer position x_* or dominant balance may be not be right.
- While the term "boundary layer" stems from the fact that the inner domain often occurs at a boundary, in some cases, they can also occur within the domain of a problem, in which case they are sometimes called *interior layers*.

Consequently, the reader should consider steps (1)–(5) as "guidelines" that may need to be adjusted depending on the given problem; this is one of the challenging (and interesting) points of matched asymptotic expansions.

We use an example to illustrate the aspects of the above procedure. Consider the boundary value problem

$$\varepsilon \frac{d^2y}{dx^2} + \frac{dy}{dx} = \cos x \tag{7.23a}$$

in the limit $\varepsilon \to 0$ on the domain $0 \le x \le \pi$, subject to the boundary conditions

$$y(0) = 2, \qquad y(\pi) = -1. \tag{7.23b}$$

7.3.1 The Outer Solution

Assuming that $y(x)$ and its derivatives are bounded as $\varepsilon \to 0$, we write the outer solution as $y \sim y_0(x) + \varepsilon y_1(x) + \varepsilon^2 y_2(x) \cdots$. Substituting into (7.23a) gives the sequence of equations

$$\begin{aligned} O(\varepsilon^0): &\quad y_0' = \cos x, \\ O(\varepsilon^1): &\quad y_1' + y_0'' = 0, \\ O(\varepsilon^2): &\quad y_2' + y_1'' = 0, \end{aligned}$$

and so on for higher order equations. The $O(1)$ problem yields $y_0 = \sin x + A$ and substituting this into the $O(\varepsilon)$ equation gives $y_1(x) = -\cos x + B$. We can proceed in this way to determine as many terms as desired in the expansion of the general outer solution

$$y_{\text{out}} = (\sin x + A) + \varepsilon(-\cos x + B) + O(\varepsilon^2). \tag{7.24}$$

At each order, there is only a single constant of integration, $A, B, \ldots$. Imposing the condition at $x = 0$ from (7.23b) selects $A = 2$, while the condition at $x = \pi$ picks $A = -1$; the outer solution cannot satisfy both at once, and hence a boundary layer will be required. In summary, at this point, we do not know which boundary conditions will apply to the outer and which to the inner solutions.

7.3.2 The Distinguished Limits

To determine the relevant scaling of the singular solution, we write $y(x) = \varepsilon^\beta Y(X)$ and $X = (x - x_*)/\varepsilon^\alpha$ and assume that $Y(X) = O(1)$ for $X = O(1)$. Substituting into (7.23a) yields

$$\underbrace{\varepsilon^{1-2\alpha+\beta} Y''}_{(1)} + \underbrace{\varepsilon^{-\alpha+\beta} Y'}_{(2)} = \underbrace{\cos(x_* + \varepsilon^\alpha X)}_{(3)}, \tag{7.25}$$

where $0 \le x_* \le \pi$. We also note that both boundary conditions (7.23b) take the form $\varepsilon^\beta Y = O(1)$, and hence any solution local to a boundary must have $\beta = 0$. It remains to determine α from the possible dominant balances:

$$\begin{array}{lll} \text{(a) Terms (2, 3): } \varepsilon^{-\alpha} = \varepsilon^0 & \Rightarrow & \alpha = 0, \\ \text{(b) Terms (1, 3): } \varepsilon^{1-2\alpha} = \varepsilon^0 & \Rightarrow & \alpha = 1/2, \\ \text{(c) Terms (1, 2): } \varepsilon^{1-2\alpha} = \varepsilon^{-\alpha} & \Rightarrow & \alpha = 1. \end{array}$$

Option (a) is the regular distinguished limit that corresponds to the outer solution. Option (b) is not a valid balance since the neglected term (2) is not sub-dominant, $O(\varepsilon^{-1/2}) \gg O(1)$. Consequently, the boundary layer must take the form given by (c) where the neglected term (3) is sub-dominant to the leading balance with $O(1) \ll O(\varepsilon^{-1})$. We note that in some problems, the dominant balances can change for different assumed positions of the boundary layer, x_* (most notably for non-autonomous equations), but here term (3) uniformly satisfies $|\cos(x)| \le 1 = O(\varepsilon^0)$. Hence our scaled equation for the inner solution is given by

$$\frac{d^2Y}{dX^2} + \frac{dY}{dX} = \varepsilon \cos(x_* + \varepsilon X), \tag{7.26}$$

where x_* has not yet been determined.

7.3.3 The Inner Solution

Having the inner problem in regular perturbation form, we expand $Y(X)$ as $Y \sim Y_0 + \varepsilon Y_1 + \varepsilon^2 Y_2 + \cdots$ and substitute into (7.26) to give the system of equations

$$\begin{aligned} O(\varepsilon^0): &\quad Y_0'' + Y_0' = 0, \\ O(\varepsilon^1): &\quad Y_1'' + Y_1' = \cos(x_*), \\ O(\varepsilon^2): &\quad Y_2'' + Y_2' = -\sin(x_*)X, \ldots . \end{aligned} \tag{7.27}$$

The $O(1)$ equation yields the leading order inner solution,

$$Y_0(X) = D + Ce^{-X}, \tag{7.28}$$

with the higher order problems producing smaller corrections to this result. The constants of integration C, D must be determined by boundary conditions or by matching with the outer solution, but this, in turn, depends on the location of x_*.

Consider the forms of the inner domain in terms of $X = (x - x_*)/\varepsilon$ for different possible values of x_*

$$\begin{array}{lll} \text{(i)} \quad \text{Left boundary } (x_* = 0) : x \geq 0 & \Rightarrow & 0 \leq X < o(1/\varepsilon) \\ \text{(ii)} \quad \text{Interior} : 0 < x_* < \pi & \Rightarrow & -o(1/\varepsilon) < X < o(1/\varepsilon) \\ \text{(iii)} \quad \text{Right boundary } (x_* = \pi) : x \leq \pi & \Rightarrow & -o(1/\varepsilon) < X \leq 0. \end{array}$$

These options correspond to three possible forms of the composite solution (see Fig. 7.3): (a) a left boundary layer satisfying $y(0) = 2$ matching with an outer solution which has $A = 1$ in order to satisfy the right boundary condition, (b) a narrow interior transition region connecting two outer solutions, with $A^L = 2$ and $A^R = -1$, and (c) a right boundary layer satisfying $y(\pi) = -1$, with an outer solution satisfying $y(0) = 2$.

The location of the boundary layer will be determined by the structure of (7.28) and its limiting behaviour. The exponential term e^{-X} in (7.28) diverges if X is allowed to become large and negative. Such *exponentially diverging* terms *cannot* satisfy the

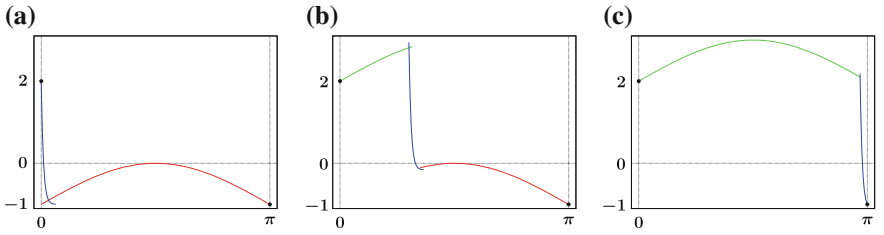


Fig. 7.3 Three hypothetical sketches of the conjectured inner/outer solutions for (7.23a, 7.23b) with a boundary layer **a** at the (*Left*), **b** in the (*Interior*), **c** at the (*Right*) *edge* of the domain

asymptotic matching condition (7.13) (with $X \to -\infty$ being the appropriate form of the 'outer limit' process) and are *un-matchable*.

Consequently options (ii) and (iii) are not feasible, and we conclude that the boundary layer must be at $x_* = 0$ with the left boundary condition from (7.23b) being relevant, namely $Y(0) = 2$. Applying this condition reduces (7.28) to

$$Y_0(X) = 2 + C(e^{-X} - 1), \tag{7.29}$$

where C remains to be determined.

7.3.4 Asymptotic Matching

Having identified the position of the boundary layer as $x_* = 0$, we have the leading order inner solution (7.29), valid on $0 \le x < O(\varepsilon)$, and the outer solution (7.24), valid on $0 < x \le \pi$. Since the right boundary lies in the outer domain, that boundary condition determines $A = -1$ in (7.24), leaving C from (7.29) as the last remaining unknown.

Applying the matching condition (7.13) for $Y_0(X \to \infty)$ and $y_0(x \to 0)$ yields

$$\lim_{X\to\infty} 2 + C(e^{-X} - 1) = 2 - C = \lim_{x\to 0} \sin(x) - 1 = -1 \tag{7.30}$$

and hence $C = 3$.

7.3.5 The Composite Solution

The overlap shared in common by the leading order inner and outer solutions above is -1. Therefore we can form the boundary layer correction as

$$Y_{\text{BLC}} = Y_0 - (-1) = 3e^{-X}.$$

Finally, adding this correction to the outer solution yields the leading order composite solution on $0 \le x \le \pi$ (see Fig. 7.4),

$$y_{\text{comp}} = -1 + \sin(x) + 3e^{-x/\varepsilon}. \tag{7.31}$$

This is in agreement with the exact solution (valid for all $\varepsilon > 0$),

$$y = -1 + \frac{\sin(x) - \varepsilon[1 + \cos(x)]}{1 + \varepsilon^2} + \left(3 + \frac{2\varepsilon}{1 + \varepsilon^2}\right) e^{-x/\varepsilon} + \text{e.s.t.}$$

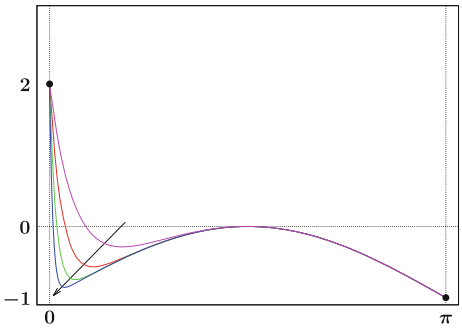


Fig. 7.4 A plot of $y_{\text{comp}}(x)$ (7.31) for a sequence of $\varepsilon \to 0$. The boundary layer becomes narrower as ε decreases

7.4 Further Examples

We present further examples to illustrate other aspects of the method of matched asymptotic expansions.

Consider the problem of obtaining the leading order solution to the ODE problem on $0 \le x \le 1$,

$$\varepsilon \frac{d^2y}{dx^2} - (2 - x^2)y = -1 \qquad \text{for } \varepsilon \to 0, \tag{7.32a}$$

with boundary conditions

$$y'(0) = 0, \qquad y(1) = 0. \tag{7.32b}$$

We begin by seeking the outer solution as a regular perturbation expansion, $y(x) \sim y_0 + \varepsilon y_1(x) + \varepsilon^2 y_2 + \cdots$. The equation for the leading order term is

$$-(2 - x^2)y_0 = -1 \quad \Longrightarrow \quad y_0(x) = \frac{1}{2 - x^2}. \tag{7.33}$$

The leading order outer solution satisfies the $y'(0) = 0$ boundary condition, i.e. $y_0'(0) = 0$, hence no boundary layer is needed there.

The outer solution has no free parameters, and it does not satisfy the boundary condition at $x = 1$. Therefore there must be a boundary layer at $x_* = 1$.

So we seek a singular solution in the form $y(x) = \varepsilon^\beta Y(X)$ with $X = (x-1)/\varepsilon^\alpha$ for $X \le 0$. Unlike the inhomogeneous conditions in (7.23b), the homogeneous boundary condition $y(1) = 0$ does not provide us information on β since $\varepsilon^\beta Y(0) = \varepsilon^\beta 0 = 0$ for any β. However, if we can determine β using some alternative means, it will simplify the process of finding the distinguished limit for the inner solution. Turning to the asymptotic matching condition provides help; here this condition will take the form

$$\lim_{x \to x_*} y_0(x) = \lim_{X \to -\infty} \varepsilon^\beta Y_0(X). \tag{7.34}$$

While we have not determined $Y_0(X)$, it is assumed to be $O(1)$, and for $x \to 1$, we have $y_0(1) = 1$. Since the limit of the outer solution is $O(1)$, so must be the limit of the inner solution, hence $\beta = 0$.

Substituting $y(x) = Y(X)$ and $x = 1 + \varepsilon^\alpha X$ into (7.32a) yields

$$\varepsilon^{1-2\alpha} Y'' - (2 - (1 + \varepsilon^\alpha X)^2) Y = -1$$

and in final form:

$$\underbrace{\varepsilon^{1-2\alpha} Y''}_{(1)} - \underbrace{(1 - 2\varepsilon^\alpha X - \varepsilon^{2\alpha} X^2) Y}_{(2)} = \underbrace{-1}_{(3)}. \tag{7.35}$$

On a finite domain, $\alpha \geq 0$ with $\alpha > 0$ describing boundary layers that narrow like $O(\varepsilon^\alpha)$ as $\varepsilon \to 0$. We note that for $\alpha > 0$ the sum of terms in the parentheses in term (2) have leading order term $(2 - x^2) \sim 1$, so the higher order terms there cannot contribute to a consistent dominant balance.

Balancing (2, 3) yields the distinguished limit for the outer solution, $\alpha = 0$, with term (1) being sub-dominant, $O(\varepsilon) \ll O(1)$. The other distinguished limit is $\alpha = 1/2$, which balances all three terms at $O(1)$.

We now have the form of the inner problem as

$$Y'' - (1 - 2\varepsilon^{1/2} X - \varepsilon X^2) Y = -1, \qquad Y(0) = 0, \tag{7.36}$$

for $X \leq 0$. The presence of the $\varepsilon^{1/2}$ suggests the expansion of the solution should take the form $Y(X) \sim Y_0 + \varepsilon^{1/2} Y_1 + \varepsilon Y_2 + \cdots$. The leading order ODE is

$$Y_0'' - Y_0 = -1, \tag{7.37}$$

with solution

$$Y_0(X) = Ae^{-X} + Be^X + 1. \tag{7.38}$$

For this solution to be matchable to the outer solution as $X \to -\infty$, we must take $A = 0$. Applying the boundary condition, $Y_0(0) = B + 1 = 0$ then gives us the solution,

$$Y_0(X) = 1 - e^X. \tag{7.39}$$

Since neither the leading order inner or outer solutions have any undetermined constants they should match automatically. This is indeed the case, and (7.34) applied to (7.33) and (7.39) shows they match with an overlap limit of 1. Hence we can construct the leading order composite solution (see Fig. 7.5)

$$y_{\text{comp}} \sim \frac{1}{2 - x^2} - e^{(x-1)/\sqrt{\varepsilon}}. \tag{7.40}$$

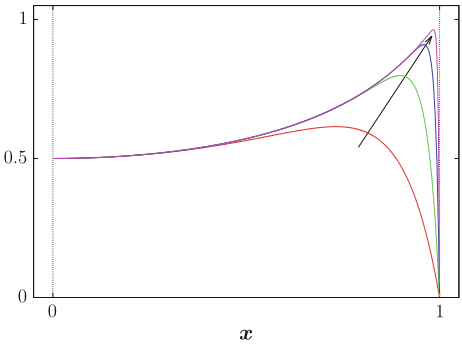


Fig. 7.5 Plot of (7.40) for $\varepsilon = 10^{-n}$ with $n = 2, 3, 4, 5$

Getting a solution accurate to higher orders would involve obtaining further terms in the expansions of the inner and outer solutions. Despite the fact that the expansions have different gauge function (ε^n vs. $\varepsilon^{m/2}$) the solutions must match together. One approach for performing matching to higher order is given in Exercise 7.6.

We now make one change to (7.32a) and illustrate how dramatically the structure of the solution is affected; consider the ODE problem on $0 \le x \le 1$,

$$\varepsilon \frac{d^2y}{dx^2} - (1 - x^2)y = -1 \qquad \text{for } \varepsilon \to 0, \tag{7.41a}$$

with boundary conditions

$$y'(0) = 0, \qquad y(1) = 0. \tag{7.41b}$$

The only change from the previous example is that the coefficient $(2 - x^2)$ in (7.32a) has been replaced by $(1 - x^2)$.

As before, the outer solution can be expressed as a regular expansion and the leading order solution is given by an algebraic equation,

$$y_0 = \frac{1}{1 - x^2}, \tag{7.42}$$

which blows up as $x \to 1$ and does not satisfy the boundary condition $y(1) = 0$. We again conclude that there must be a boundary layer at $x_* = 1$, but now face the problem of determining a boundary layer solution that can match to a diverging outer solution.

We seek an inner solution in the scaled form $y(x) = \varepsilon^\beta Y(X)$ with $X = (x-1)/\varepsilon^\alpha$, where we expect $\beta < 0$ to capture the singular nature of the magnitude of the solution and $\alpha > 0$ for a narrow boundary layer. Substituting into (7.41a) yields

$$\underbrace{\varepsilon^{1-2\alpha+\beta} Y''}_{(1)} + \underbrace{\varepsilon^{\alpha+\beta} X(2 + \varepsilon^\alpha X)Y}_{(2)} = \underbrace{-1}_{(3)}. \tag{7.43}$$

We now consider the options for two-term dominant balances in this equation:

(a) Terms (1, 2) balance if $1 - 2\alpha + \beta = \alpha + \beta$, namely $\alpha = 1/3$. To ensure that the balance is consistent, and these terms are larger than term (3), we need $\alpha + \beta < 0$, i.e. $\beta < -1/3$.
(b) Terms (1, 3) balance if $1 - 2\alpha + \beta = 0$, yielding $\beta = 2\alpha - 1$. Term (2) is sub-dominant if $\alpha + \beta > 0$. Consequently, this and the condition $\beta < 0$ determine the range $1/3 < \alpha < 1/2$.
(c) Terms (2, 3) balance if $\alpha + \beta = 0$, hence $\beta = -\alpha$. Term (1) is sub-dominant when $1 - 2\alpha + \beta > 0$, yielding $0 < \alpha < 1/3$.

It can be useful to visualise these relations in the (α, β) parameter plane in what is called a Newton–Kruskal diagram [105], see Fig. 7.6.

The leading order equations proposed by each of the above respective cases are:

$$Y_0'' + 2XY_0 = 0, \tag{7.44a}$$
$$Y_0'' = -1, \tag{7.44b}$$
$$2XY_0 = -1. \tag{7.44c}$$

Each of these equations can be shown to have some deficiency in trying to describe the inner solution. The solution of (7.44c), $Y_0 = -1/(2X)$, cannot satisfy the boundary condition at $X = 0$. The solution of (7.44b) is a parabola that cannot satisfy the asymptotic matching condition for $X \to -\infty$. Equation (7.44a) is less straightforward; it is a version of Airy's differential equation [11, 105] but it can likewise be shown that its solutions also cannot satisfy the matching condition (7.34).

The above balances are self-consistent, but because they are not the most general dominant balance, they actually have a limited range of validity with the inner domain (it can be shown that (a) holds for $X = O(1)$, (b) holds for $X \to 0$, (c) holds for $X \to -\infty$). The distinguished limit for inner problem is given by the intersection of the three cases, $\alpha = 1/3$, $\beta = -1/3$, with all three terms in (7.43) balancing,

$$Y_0'' + 2XY_0 = -1, \qquad Y_0(0) = 0. \tag{7.45}$$

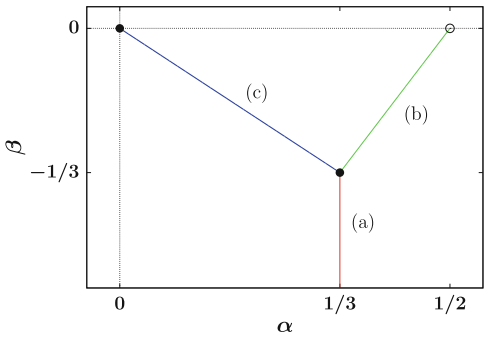


Fig. 7.6 Newton–Kruskal diagram for (7.43) showing possible two-term dominant balances for cases (a, b, c) as line segments

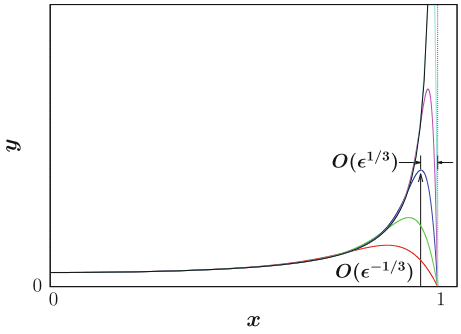


Fig. 7.7 Numerical solutions of (7.41a, 7.41b) for a sequence of $\varepsilon_n \to 0$ (*colour curves*) and the outer solution (7.42) (*black curve*)

Another way to come to this choice of scalings is to use the matching condition (7.34) with the outer solution (7.42) written in terms of X as,

$$y_0 = \frac{1}{1-(1+\varepsilon^\alpha X)^2} \sim -\frac{1}{2\varepsilon^\alpha X} \sim \varepsilon^\beta Y_0, \tag{7.46}$$

which we can recognise as case (c) above (for $X \to -\infty$) with $\beta = -\alpha$. This would reduce (7.43) to an equation for α, having two distinguished limits: $\alpha = 0$ (the outer solution), and the boundary layer given by $\alpha = 1/3$.

Constructing the composite solution would require solving (7.45) (it is an inhomogeneous version of Airy's equation) and carrying out the matching from (7.46) using the approach of Exercise 7.6. Instead, in Fig. 7.7 we show that the numerical solution of the full problem is well characterised by the outer solution on most of the domain with the maximum value of the solution and the width of the boundary layer being well-predicted by the scaling of the inner solution.

7.5 Further Directions

The problems we have considered above provide some insight into how boundary layers and matched asymptotic expansions can separate out some of the delicate behaviours of solutions of singularly perturbed problems. The examples have shown that while the steps outlined in Sect. 7.4 are a good guide, they may be coupled to each other in different ways in each problem—for example: are the scalings of the inner solution determined by the ODE, the boundary conditions, or by matching? is matching needed to set undetermined coefficients in the inner or outer solution (or neither or both)? Many problems require creative application of these steps. Analysis of some more challenging problems remain open research problems.

The form of the solutions obtained to such problem generally provides greater insight into the nature of the system. The dominant balance that determines the leading outer solution provides the simplest essential approximation of the behaviour in the problem. This approximation will be valid everywhere apart from the boundary

layers. Matched asymptotics provides understanding of whether the boundary layers are necessary to pin-down properties of the outer solution or merely correct the solution in narrow regions. Often, knowledge of the physical system being modelled can guide expectations on boundary layer positions and dominant balances; this can sometimes simplify the mathematical steps. Sometimes, the matched asymptotics will uncover unexpected dominant balances that can highlight novel and important behaviours in the problem.

This chapter only hints at the broad array of models that can be studied using matched asymptotics and the types of behaviours that can result. Some of these include boundary layers within boundary layers (nested layers or "triple decks"), boundary layers that begin at higher orders ("corner layers"), and problems with unusual gauge functions. There is a wide array of books that give further studies of boundary layer problems [47, 48, 58, 78, 92] and some primary sources on the theory of asymptotic matching are [60, 101].

7.6 Exercises

7.1 Evaluate $\lim_{\varepsilon\to 0} e^{-1/\varepsilon}/\varepsilon^n$ to demonstrate that exponentially small terms are smaller than all algebraic terms.

7.2 Determine the three possible dominant balances for (7.25) that could occur for $x_* = \pi/2$, noting that $\cos(x) \to 0$ as $x \to x_*$.

7.3 Consider the problem for $y(x)$ on $0 \le x \le 1$ with $\varepsilon \to 0$,

$$\varepsilon \frac{d^2y}{dx^2} - (4 - x^2)y = \cos\left(\tfrac{\pi}{2}x\right) \qquad y(0) = -1 \qquad y(1) = 2$$

(a) Determine the two distinguished limits for this problem.
(b) Write the leading order outer solution $y_0(x)$.
(c) Find the leading order inner solutions $Y_0(X)$.
(d) Write the leading order uniformly-valid solution.

7.4 Consider the initial value problem for $y(x)$ on $0 \le x$ for $\varepsilon \to 0$,

$$\varepsilon \frac{d^2y}{dx^2} + 2\frac{dy}{dx} - 6y = 5x, \qquad y(0) = 0, \qquad y'(0) = \frac{4}{\varepsilon^2}.$$

You are given that the solution has a boundary layer at $x_* = 0$.

(a) Determine the leading order inner solution.
(b) Write the outer limit of the inner solution to determine a necessary matching condition on the outer solution.
(c) Determine the leading order outer solution for $x > 0$.

7.5 Consider the problem for $y(x)$ on $0 \le x \le 1$ with $\varepsilon \to 0$,

$$\varepsilon \frac{d^2y}{dx^2} + 2\frac{dy}{dx} + e^y = 0, \qquad y(0) = 0, \qquad y(1) = 0.$$

(a) Find the general leading order outer solution $y_0(x)$.
(b) Find the leading order inner solution $Y_0(X)$ and determine where the boundary layer occurs.
(c) Write the leading order composite solution.
(d) Determine the next term in the outer solution, $y_1(x)$.

7.6 Consider the problem for $y(x)$ on $0 \le x \le 1$ with $\varepsilon \to 0$,

$$\varepsilon \frac{d^2y}{dx^2} - \frac{dy}{dx} + y = 2x, \qquad y(0) = -2, \qquad y(1) = 1.$$

(a) Determine the distinguished limit for the inner problem for this equation. By solving the leading order inner problem, determine where the boundary layer occurs.
(b) Obtain the *first two terms* in the expansion of the inner solution (with the appropriate boundary conditions imposed),

$$Y(X) \sim Y_0(X) + \varepsilon Y_1(X).$$

(c) Determine the *first two terms* in the expansion of the outer solution (with the appropriate boundary conditions imposed),

$$y(x) \sim y_0(x) + \varepsilon y_1(x).$$

(d) The inner solution will have undetermined constants. These constants can be determined using *higher-order matching using intermediate variables* [60, 101] via the following steps:

(i) Analogous to (7.12), define small parameter η with $\varepsilon^\alpha \ll \eta \ll 1$.
(ii) Define the intermediate variable $\hat{x}$ as $\hat{x} = (x - x_*)/\eta$.
(iii) Use the relations

$$x = x_* + \eta\hat{x}, \qquad X = \frac{\eta}{\varepsilon^\alpha}\hat{x}$$

to write the outer and inner solutions both in terms of the intermediate variable $\hat{x}$ and ε, η.
(iv) Use the relations $\varepsilon^\alpha \ll \eta \ll 1$ to expand out exponential functions or eliminate small terms in the solutions on the overlap region, with $\hat{x} = O(1)$.
(v) Arrange the remaining terms as an ordered asymptotic expansion involving η, ε and determine the remaining constants through matching of terms.

(e) Write the uniform solution valid up through $O(\varepsilon)$ terms.

7.7 Consider the problem for $y(x)$ on $0 \le x \le 2$ with $\varepsilon \to 0$,

$$x^2 + y^2 = 4 - \varepsilon \frac{dy}{dx}, \qquad y(0) = \frac{3}{\varepsilon}.$$

(a) Determine the first two terms in the outer solution, $y \sim y_0 + \varepsilon y_1$.
(b) Boundary layers can occur at $x_* = 0$ and $x_* = 2$. For each case, determine the α, β for each distinguished limit and write its corresponding leading order equation for $Y_0(X)$. Note that there are two singular distinguished limits at $x_* = 0$.

7.8 Consider the problem for $y(x)$ on $0 \le x \le 1$ for $\varepsilon \to 0$,

$$\varepsilon \frac{d^2y}{dx^2} - y = -4 + \frac{\varepsilon^2 y}{(x-1)^3}, \qquad y(0) = 0, \qquad y(1) = 0.$$

The leading order outer solution is $y_0(x) \equiv 4$.

(a) Determine the leading order inner solution for the boundary layer at $x_* = 0$.
(b) At $x_* = 1$, there are <u>two</u> different distinguished limits. Determine α for each and obtain the respective leading order equations for each $Y_0(X)$.
Since these solutions have different α's, one layer is nested inside the other. The "inner-inner" layer solution should satisfy boundary condition at $x = 1$. The (wider) "intermediate inner" layer should asymptotically match to the outer and inner-inner solutions for the limits $X \to -\infty$ and $X \to 0$ respectively in this *triple deck* problem.

Chapter 8
Long-Wave Asymptotics for PDE Problems

In this chapter we will study matched asymptotics and boundary layer theory applied to some classes of multi-dimensional problems for partial differential equations (PDE). Perturbation methods offer an interesting alternative way to construct solutions that will provide different insight into the structure of some PDE problems. We will also see that matched asymptotics will allow us to solve problems that cannot be tackled directly by classical methods.

We will illustrate the analysis in the context of problems for Laplace's equation,

$$\nabla^2 \mathsf{U} \equiv \frac{\partial^2 \mathsf{U}}{\partial \mathsf{X}^2} + \frac{\partial^2 \mathsf{U}}{\partial \mathsf{Y}^2} = 0. \tag{8.1}$$

The Laplacian operator ∇^2 is a fundamental element of equations describing various phenomena such as diffusion, wave propagation, and equilibrium potentials

$$\frac{\partial \mathsf{U}}{\partial \mathsf{T}} = \nabla^2 \mathsf{U}, \qquad \frac{\partial^2 \mathsf{U}}{\partial \mathsf{T}^2} = \nabla^2 \mathsf{U}, \qquad \mathsf{F}(\mathsf{X}, \mathsf{Y}) = \nabla^2 \mathsf{U}.$$

These equations arise in electromagnetism, heat conduction, mass transfer, fluid flow, solid mechanics, and many other areas.

We will focus on electrostatics as an example application—in this case $\mathsf{U}(\mathsf{X}, \mathsf{Y})$ represents the electric potential (voltage) that drives the flow of charges and currents in a conductor. In particular, we will describe the current flow and electric charge density in a piece of wire by solving Eq. (8.1) on a long slender domain, subject to various boundary conditions.

8.1 The Classic Separation of Variables Solution

We begin by briefly reviewing the traditional separation of variables approach to solving boundary value problems for Laplace's equation [23, 44].

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9_8

For elliptic PDE such as Laplace's equation, to uniquely determine a solution, a problem must provide boundary conditions around the entire boundary of the domain. In the context of electrostatics, Dirichlet boundary conditions, where the value of the solution is specified, correspond to imposing a known voltage on the edge of the domain, e.g. $\mathsf{U}(\mathsf{X}=0,\mathsf{Y})=\bar{\mathsf{U}}$. In contrast, Neumann boundary conditions provide the value of the directional derivative of the solution normal to the boundary

$$\frac{\partial \mathsf{U}}{\partial n} \equiv \hat{\mathbf{n}} \cdot \nabla \mathsf{U}, \tag{8.2}$$

where $\hat{\mathbf{n}}$ is the unit normal vector, perpendicular to the boundary, in the outward direction, e.g. $\hat{\mathbf{n}} = -\hat{\mathbf{i}}$ is the normal to the left boundary of a domain given by $\mathsf{X} \geq 0$,

$$\left.\frac{\partial \mathsf{U}}{\partial n}\right|_{\mathsf{X}=0} = -\hat{\mathbf{i}} \cdot (\partial_{\mathsf{X}}\mathsf{U}, \partial_{\mathsf{Y}}\mathsf{U}) = -\left.\frac{\partial \mathsf{U}}{\partial \mathsf{X}}\right|_{\mathsf{X}=0} = \mathsf{J}(\mathsf{Y}).$$

Physically, Neumann conditions specify the current, or *flux* of the solution, out of the domain.

Consider Laplace's equation on a rectangular domain,

$$\frac{\partial^2 \mathsf{U}}{\partial \mathsf{X}^2} + \frac{\partial^2 \mathsf{U}}{\partial \mathsf{Y}^2} = 0 \qquad 0 \leq \mathsf{X} \leq \mathsf{L} \qquad 0 \leq \mathsf{Y} \leq \mathsf{H} \tag{8.3a}$$

subject to Dirichlet boundary conditions,

$$\mathsf{U}(\mathsf{X},0) = 0, \qquad \mathsf{U}(\mathsf{L},\mathsf{Y}) = 0, \qquad \mathsf{U}(0,\mathsf{Y}) = 0, \qquad \mathsf{U}(\mathsf{X},\mathsf{H}) = \mathsf{F}(\mathsf{X}). \tag{8.3b}$$

This is an elementary 'building-block' problem where the form of the solution will be entirely due to the one inhomogeneous boundary condition. Solutions to problems on the same domain, but with inhomogeneous boundary conditions on other edges, can be built-up from linear superposition of combinations of such building block solutions.

Since this is a linear problem, the overall solution can be obtained as a superposition of linearly independent trial solutions. The trial solutions can be sought in *separation of variables* form, as a product of functions of the independent variables:

$$\mathsf{U}(\mathsf{X},\mathsf{Y}) = \sum_{n=1}^{\infty} \mathsf{c}_n \mathsf{U}_n(\mathsf{X},\mathsf{Y}) \quad \text{with} \quad \mathsf{U}_n(\mathsf{X},\mathsf{Y}) = \mathsf{A}_n(\mathsf{X})\mathsf{B}_n(\mathsf{Y}). \tag{8.4}$$

Substituting (8.4) into the homogeneous boundary condition $\mathsf{U}(0,\mathsf{Y}) = 0$ for $0 \leq \mathsf{Y} \leq \mathsf{H}$ with the assumptions that the solution U is nontrivial (not all coefficients $\mathsf{c}_n = 0$) and that the U_n trial solutions are linearly independent yields boundary conditions on the $\mathsf{A}_n(\mathsf{X})$ functions for $n = 1, 2, 3, \ldots,$

$$\sum \mathsf{c}_n \mathsf{A}_n(0)\mathsf{B}_n(\mathsf{Y}) = 0 \qquad \text{for} \quad 0 < \mathsf{Y} < \mathsf{H} \qquad \Longrightarrow \qquad \mathsf{A}_n(0) = 0.$$

Similarly for the other homogeneous boundary conditions, $\mathsf{U}(\mathsf{L},\mathsf{Y}) = 0$ yields $\mathsf{A}_n(\mathsf{L}) = 0$ while $\mathsf{U}(\mathsf{X},0) = 0$ gives $\mathsf{B}_n(0) = 0$. The inhomogeneous boundary condition in (8.3b) will be approached differently.

Substituting $\mathsf{U}_n(\mathsf{X},\mathsf{Y})$ from (8.4) into (8.3a) and requiring each trial solution to satisfy Laplace's equation yields

$$\frac{d^2\mathsf{A}_n}{d\mathsf{X}^2}\mathsf{B}_n(\mathsf{Y}) + \mathsf{A}_n(\mathsf{X})\frac{d^2\mathsf{B}_n}{d\mathsf{Y}^2} = 0$$

which can be re-arranged to give

$$\frac{\mathsf{A}_n''(\mathsf{X})}{\mathsf{A}_n(\mathsf{X})} = -\frac{\mathsf{B}_n''(\mathsf{Y})}{\mathsf{B}_n(\mathsf{Y})} = \mathsf{s}_n = \pm\lambda_n.$$

Since the equality must hold for all independent values of X and Y, both sides must take the same constant value s_n, called a *separation constant*, which here is written as a sign times the undetermined positive constant $\lambda_n \geq 0$, where the subscript indicates that the separation constants are generally distinct between different trial solutions. The main result of this separation of variables approach is to split the PDE into two separate ODEs for $\mathsf{A}_n(\mathsf{X})$ and $\mathsf{B}_n(\mathsf{Y})$ that are linked only through λ_n.

In order to make further progress it is convenient to begin by analysing the ODE problem satisfying homogeneous boundary conditions. In this case this selects the problem in X-direction for $\mathsf{A}_n(\mathsf{X})$. It can be shown that subject to the boundary conditions, obtaining nontrivial solutions of $\mathsf{A}_n'' - \mathsf{s}_n\mathsf{A}_n = 0$ forces s_n to be negative, $\mathsf{s}_n = -\lambda_n$ [23, 44]. The problem for $\{\mathsf{A}_n, \lambda_n\}$ is an *eigenvalue problem*,

$$\mathsf{A}_n'' + \lambda_n\mathsf{A}_n = 0, \qquad \mathsf{A}_n(0) = 0, \qquad \mathsf{A}_n(\mathsf{L}) = 0, \tag{8.5}$$

yielding an infinite sequence of oscillatory eigenfunction solutions

$$\mathsf{A}_n(\mathsf{X}) = \sin\left(\frac{n\pi}{\mathsf{L}}\mathsf{X}\right), \qquad \lambda_n = \frac{n^2\pi^2}{\mathsf{L}^2}, \qquad n = 1, 2, 3, \ldots. \tag{8.6}$$

Having obtained the separation constants, the ODE for $\mathsf{B}_n(\mathsf{Y})$ is now completely specified,

$$\mathsf{B}_n'' - \tfrac{n^2\pi^2}{\mathsf{L}^2}\mathsf{B}_n = 0 \qquad \Longrightarrow \qquad \mathsf{B}_n(\mathsf{Y}) = \mathsf{C}_1\sinh(\tfrac{n\pi}{\mathsf{L}}\mathsf{Y}) + \mathsf{C}_2\cosh(\tfrac{n\pi}{\mathsf{L}}\mathsf{Y}). \tag{8.7}$$

Imposing the $B_n(0) = 0$ boundary condition reduces the general solution to $\mathsf{B}_n(\mathsf{Y}) = \sinh(\frac{n\pi}{\mathsf{L}}\mathsf{Y})$. The full solution (8.4) then takes the form

$$\mathsf{U}(\mathsf{X},\mathsf{Y}) = \sum_{n=1}^{\infty} \mathsf{c}_n \sinh(\tfrac{n\pi}{\mathsf{L}}\mathsf{Y})\sin(\tfrac{n\pi}{\mathsf{L}}\mathsf{X}); \tag{8.8}$$

all that remains is to determine the coefficients, c_n. Applying the inhomogeneous boundary condition to the series at $Y = H$ yields,

$$U(X, H) = \sum_{n=1}^{\infty} \underbrace{c_n \sinh(\tfrac{n\pi}{L} H)}_{f_n} \sin(\tfrac{n\pi}{L} X) = F(X).$$

This equation can be interpreted as the Fourier sine series for the function $F(X)$ (see Appendix A) and determines the coefficients in the expansion,

$$\sum_{n=1}^{\infty} f_n \sin(\tfrac{n\pi}{L} X) = F(X) \qquad \Longrightarrow \qquad f_n = \frac{2}{L} \int_0^L F(X) \sin(\tfrac{n\pi}{L} X)\, dX.$$

Finally, expressing c_n in terms of f_n, we obtain the solution in the form of an infinite series

$$U(X, Y) = \sum_{n=1}^{\infty} \left(\frac{2}{L \sinh(\tfrac{n\pi}{L} H)} \int_0^L F(\tilde{X}) \sin(\tfrac{n\pi}{L} \tilde{X})\, d\tilde{X} \right) \sinh(\tfrac{n\pi}{L} Y) \sin(\tfrac{n\pi}{L} X). \tag{8.9}$$

8.2 The Dirichlet Problem on a Slender Rectangle

Solution (8.9) holds for all choices of H, L but working with an infinite series can be somewhat cumbersome. We will now see how a much more compact, but equivalent form of the solution can be obtained for slender rectangles, where the *aspect ratio* $H/L \ll 1$, with, for the most-part, a dramatic simplification of the PDE problem to a sequence of simple ODE problems.

Consider a problem for Laplace's equation on the rectangular domain, $0 \le X \le L$, $0 \le Y \le H$ (see Fig. 8.1):

$$\frac{\partial^2 U}{\partial X^2} + \frac{\partial^2 U}{\partial Y^2} = 0, \tag{8.10a}$$

subject to the boundary conditions

$$U(X, 0) = 0, \qquad U(X, H) = \bar{U} f(X/L), \tag{8.10b}$$

$$U(0, Y) = \bar{U} g_0(Y/H), \quad U(L, Y) = \bar{U} g_1(Y/H), \tag{8.10c}$$

where $f(x)$, $g_0(y)$, $g_1(y)$ are given functions.

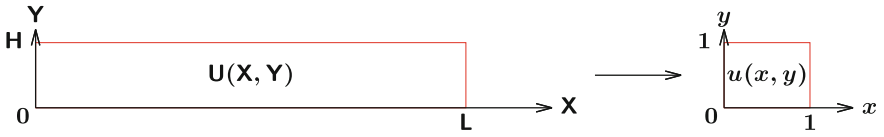


Fig. 8.1 The domain for problem (8.10), a slender rectangle in dimensional coordinates (*Left*) and the rescaled dimensionless domain (*Right*)

We nondimensionalize using the scalings

$$\mathsf{X} = \mathsf{L}x, \qquad \mathsf{Y} = \mathsf{H}y, \qquad \mathsf{U} = \bar{\mathsf{U}}u(x, y), \tag{8.11}$$

yielding the scaled problem on $0 \le x \le 1$ and $0 \le y \le 1$,

$$\varepsilon^2 u_{xx} + u_{yy} = 0 \tag{8.12a}$$

$$u(x, 0) = 0, \quad u(x, 1) = f(x) \qquad \text{for } 0 < x < 1 \tag{8.12b}$$

$$u(0, y) = g_0(y), \qquad u(1, y) = g_1(y) \qquad \text{for } 0 < y < 1 \tag{8.12c}$$

where the aspect ratio, or *slenderness parameter* $\varepsilon = \mathsf{H}/\mathsf{L}$ is small, $\varepsilon \to 0$, corresponding to a long, thin domain.

Using ε as a perturbation parameter, we begin by seeking an outer solution of the form

$$u(x, y) \sim u_0(x, y) + \varepsilon^2 u_1(x, y) + \varepsilon^4 u_2(x, y) + \cdots . \tag{8.13}$$

Note that for this problem it is sufficient to use an expansion having only even powers of ε because the perturbation parameter appears in the problem only as ε^2 in (8.12a). Substituting into (8.12a, 8.12b), we obtain a sequence of problems which are each essentially ODE boundary value problems in the y-direction with x entering only as a secondary parameter. At leading order we get

$$u_{0yy} = 0, \qquad u_0(y = 0) = 0, \qquad u_0(y = 1) = f(x), \tag{8.14a}$$

where the differential equation determines u_0 to be linear with respect to y, $u_0 = C_1 y + C_2$. The C coefficients need not be constants; they must be independent of y but can depend on any other variable(s) present in the problem, namely $u_0(x, y) = C_1(x)y + C_2(x)$. Applying the boundary conditions from (8.14a), we arrive at the leading order solution

$$u_0(x, y) = f(x)y. \tag{8.14b}$$

Similarly, at higher orders,

$$O(\varepsilon^2): \quad u_{1yy} = -u_{0xx}, \qquad u_1(y=0) = 0, \qquad u_1(y=1) = 0$$

$$\Longrightarrow \qquad u_1(x,y) = \tfrac{1}{6} f''(x)(y - y^3).$$

$$O(\varepsilon^4): \quad u_{2yy} = -u_{1xx}, \qquad u_2(y=0) = 0, \qquad u_2(y=1) = 0$$

$$\Longrightarrow \qquad u_2(x,y) = \tfrac{1}{360} f''''(x)(7y - 10y^3 + 3y^5).$$

If the function $f(x)$ in the boundary condition at the top of the domain is smooth, this expansion can be continued to all orders so as to obtain the outer solution in terms of polynomials in y times derivatives of $f(x)$. However, this solution will not in general satisfy the boundary conditions (8.12c) at $x = 0$ and $x = 1$, and hence boundary layer corrections will be needed.

Noting that (8.12a) has a small parameter multiplying the highest order derivative in the x-direction, we recognise it as a singular perturbation problem of the type treated in Chap. 7 and we seek boundary layers at $x = 0$ and $x = 1$. In order to analyse the structure of the inner solution with respect to x, we assume a scaled solution of the form (7.21),

$$X = \frac{x - x_*}{\varepsilon^\alpha}, \qquad u = U(X, y), \tag{8.15}$$

where we have already made use of the Dirichlet boundary conditions, $u = g = O(1)$ for the scaling of U. Substituting this form into (8.12a) yields

$$\varepsilon^{2-2\alpha} U_{XX} + U_{yy} = 0,$$

which selects the distinguished limit $\alpha = 1$ for the inner solution at both $x_*^L = 0$ for the left boundary and $x_*^R = 1$ for the right boundary layer. We will describe the solution for $x_*^L = 0$ with the construction for the right boundary layer following analogously.

While the inner problem with $\alpha = 1$ has brought us back to solving the full Laplace's equation,

$$U_{XX} + U_{yy} = 0, \tag{8.16a}$$

there are subtle changes in the domain and boundary conditions that in this context make the inner problem more tractable than the original full problem. The domain for this inner problem is a semi-infinite strip, $0 \le y \le 1$ with $X \ge 0$. Since $x = \varepsilon X$ with $X = O(1)$ in the left boundary layer, we can re-write the $y = 1$ boundary condition as

$$U(X,1) = f(\varepsilon X) = f(0) + \varepsilon f'(0) X + \tfrac{1}{2}\varepsilon^2 f''(0) X^2 + \cdots$$

yielding boundary conditions on the leading order inner solution as constant values

$$U_0(X,0) = 0, \qquad U_0(X,1) = f(0). \tag{8.16b}$$

The boundary condition at $X = 0$ is unchanged from (8.12c), but now the remaining boundary condition is supplied by the analogue of the asymptotic matching condition with the outer solution (see (7.13))

$$\lim_{X\to\infty} U_0(X,y) = \lim_{x\to 0} u(x,y). \tag{8.16c}$$

Consequently (8.12c), (8.16c) and (8.14b) specify the left and right boundary conditions as

$$U_0(0,y) = g_0(y), \qquad U_0(X\to\infty, y) = f(0)y. \tag{8.16d}$$

Noting the term $f(0)y$ from the outer solution is present in both (8.16b, 8.16d), we observe that the problem can be expressed in terms of the boundary layer correction function $V(X,y)$ defined by

$$U_0(X,y) = V(X,y) + f(0)y, \tag{8.17}$$

where $V(X,y)$ satisfies the problem

$$V_{XX} + V_{yy} = 0, \tag{8.18a}$$

$$V(X,0) = 0, \qquad V(X,1) = 0, \tag{8.18b}$$

$$V(0,y) = g_0(y) - f(0)y, \qquad V(X\to\infty, y) = 0. \tag{8.18c}$$

This matches the form of the Dirichlet "building-block" problem with a single inhomogeneous boundary condition described in Sect. 8.1. Here the homogeneous boundary conditions determine the oscillatory eigenfunctions to be in the y-direction with $\lambda_n = n\pi$ yielding

$$V(X,y) = \sum_{n=1}^{\infty} c_n e^{-n\pi X} \sin(n\pi y), \tag{8.19a}$$

where

$$c_n = 2\int_0^1 [g_0(y) - f(0)y] \sin(n\pi y)\, dy. \tag{8.19b}$$

We note that the $e^{-n\pi X}$ factors resulted from solving the ODE $A_n'' - n^2\pi^2 A_n = 0$ yielding general solutions $A_n(X) = d_0 e^{n\pi X} + d_1 e^{-n\pi X}$ and enforcing the $X \to \infty$ boundary condition to then eliminate the un-matchable exponentially growing modes.

In conclusion, the leading order uniform solution can then be composed from the outer solution and the boundary layer corrections,

$$u_{0\text{unif}}(x, y) = u_0(x, y) + V^L(X^L, y) + V^R(X^R, y), \tag{8.20}$$

where $V^L(X^L, y)$ is given by (8.19a, 8.19b) with $X^L = x/\varepsilon$ and $V^R(X^R, y)$ analogously by the right boundary layer correction with $X^R = (x - 1)/\varepsilon$.

8.3 The Insulated Wire

Current flow through an insulated wire can be represented by replacing the Dirichlet conditions on the upper and lower boundaries with no-flux Neumann boundary conditions (8.2). Using the same scalings as in the previous example (8.11), the nondimensionalized problem on $0 \le x \le 1$ and $0 \le y \le 1$ takes the form,

$$\varepsilon^2 u_{xx} + u_{yy} = 0, \tag{8.21a}$$

$$\left.\frac{\partial u}{\partial y}\right|_{y=0} = 0, \qquad \left.\frac{\partial u}{\partial y}\right|_{y=1} = 0 \qquad \text{for} \quad 0 < x < 1, \tag{8.21b}$$

$$u(x = 0) = g_0(y), \qquad u(x = 1) = g_1(y) \qquad \text{for} \quad 0 < y < 1. \tag{8.21c}$$

This problem describes current flow in a straight insulated wire with prescribed end-voltages.

We begin by constructing the outer solution in the form $u \sim u_0 + \varepsilon^2 u_1 + \varepsilon^4 u_2$. The leading order problem for $u_0(x, y)$ is then given by

$$u_{0yy} = 0, \qquad u_{0y}(y = 0) = 0, \qquad u_{0y}(y = 1) = 0, \tag{8.22}$$

where the general solution of the differential equation is again a linear function of y with x-dependent coefficients, $u_0(x, y) = C_1(x)y + C_2(x)$. Applying the boundary conditions yields

$$u_0(x, y) = C_2(x), \tag{8.23}$$

and so the solution has been shown to be independent of y, but is given by some as-yet undetermined function of x. Often in perturbation methods, undetermined parts of solutions will get pinned down by conditions needed for consistency appearing at higher orders, or from matching to solutions on other parts of the domain. At order $O(\varepsilon^2)$ we have the problem for $u_1(x, y)$,

$$u_{1yy} = -u_{0xx}, \qquad u_{1y}(x, 0) = 0, \qquad u_{1y}(x, 1) = 0. \tag{8.24}$$

Substituting (8.23) for u_0 into this problem, we obtain the general solution

$$u_1 = -\tfrac{1}{2}C_2''(x)y^2 + C_3(x)y + C_4(x).$$

Applying the boundary condition at $y = 0$ determines $C_3(x) \equiv 0$; applying the boundary condition at $y = 1$ yields

$$\frac{d^2C_2}{dx^2} = 0 \qquad \text{for } 0 < x < 1; \tag{8.25}$$

this is an ordinary differential equation for the leading order solution. Such equations determining consistency conditions on lower-order solutions, coming out of parts of higher order problems, are often called *solvability conditions*.

Equation (8.25) has a simple linear general solution,

$$C_2(x) = D_1x + D_2,$$

but in order to determine the constants $D_{1,2}$, we need to impose boundary conditions. Our original problem indeed had boundary conditions at $x = 0$ and $x = 1$, (8.21c), but for general imposed functions $g_0(y)$ and $g_1(y)$, $C_2(x)$ cannot possibly satisfy those y-dependent conditions. Hence we must construct boundary layers to satisfy (8.21c) while determining the effective boundary conditions on $C_2(x)$.

Following the same scaling (8.15) as in the previous problem, we determine the inner distinguished limit $\alpha = 1$ for the boundary layers both at $x_*^L = 0$ and $x_*^R = 1$. We again recover Laplace's equation as the inner problem,

$$U_{XX} + U_{yy} = 0. \tag{8.26a}$$

Analysing the boundary layer at $x_*^L = 0$ (the analysis at $x_*^R = 1$ is analogous), the three boundary conditions we can draw from (8.21b, 8.21c) are straightforward,

$$U(0, y) = g_0(y), \qquad U_y(X, 0) = 0, \qquad U_y(X, 1) = 0, \tag{8.26b}$$

and the final boundary condition is obtained from asymptotic matching to the outer solution, using (8.16c),

$$U(X \to \infty, y) = D_2. \tag{8.26c}$$

Again, it is convenient to recast this problem in terms of the boundary layer correction in order to separate out the overlap from matching, so we write

$$U(X, y) = V(X, y) + D_2, \tag{8.27}$$

with $V(X, y)$ satisfying

$$V_{XX} + V_{yy} = 0, \tag{8.28a}$$

$$V_y(X,0) = 0, \qquad V_y(X,1) = 0, \tag{8.28b}$$

$$V(0,y) = g_0(y) - D_2, \qquad V(X \to \infty, y) = 0. \tag{8.28c}$$

Applying separation of variables to this Neumann boundary value problem yields the solution as a cosine series,

$$V(X,y) = \sum_{n=0}^{\infty} c_n e^{-n\pi X} \cos(n\pi y), \tag{8.29}$$

where the matching condition eliminated the exponentially growing modes, $e^{n\pi X}$. The c_n's are then given by the Fourier coefficients of the $X = 0$ boundary condition,

$$c_0 = \int_0^1 [g_0(y) - D_2]\, dy, \quad \text{and} \quad c_n = 2\int_0^1 [g_0(y) - D_2]\cos(n\pi y)\, dy \tag{8.30}$$

for $n = 1, 2, \ldots$. For $n > 0$ all of the $c_n e^{-n\pi X}$ factors in (8.29) decay to zero as $X \to \infty$; any finite values for c_n would be compatible with the remaining boundary condition, that $V(X \to \infty) = 0$. The only term that contributes to the value of V for $X \to \infty$ is the $n = 0$ term, the uniform constant, $V(X \to \infty) \sim c_0$. In order to satisfy the matching condition (8.28c), we must have $c_0 = 0$ in (8.30). This determines the boundary condition on the outer solution (8.25) in terms of the average of the g_0 boundary function,

$$C_2(0) = D_2 = \int_0^1 g_0(y)\, dy.$$

Analogous analysis of the boundary layer at $x_* = 1$ yields the boundary condition for $C_2(1)$ in terms of the average value of g_1 and determines the outer solution as

$$u_0(x,y) = x\left(\int_0^1 g_1(y) - g_0(y)\, dy\right) + \left(\int_0^1 g_0(y)\, dy\right), \tag{8.31}$$

with the leading order uniform solution taking the same form as (8.20).

It is interesting to compare the 'information flow' or 'structural dependence' of the perturbation solution for the past two problems. For the Dirichlet problem, the outer solution can be determined independently of the boundary layers, and sets matching conditions for the boundary layers,

Problem (8.12): | BL ← Outer → BL |

while in the Neumann problem, the outer solution cannot be completely specified until the boundary layer solutions have been calculated,

Problem (8.21): | BL → Outer ← BL |

8.4 The Nonuniform Insulated Wire

We now extend our previous analysis to consider the problem of an insulated wire whose cross-section is not of constant width,[1] see Fig. 8.2. The main consequence of this change is that the domain is no longer separable[2] so the method of separation of variables from Sect. 8.1 can not be used to construct an exact solution of the whole problem. However, boundary layer theory and the method of matched asymptotics again works, with just minor extensions.

We begin with the dimensional problem for Laplace's equation on a domain with $0 \le \mathsf{X} \le \mathsf{L}$ and $0 \le \mathsf{Y} \le \mathsf{H}f(\mathsf{X}/\mathsf{L})$ where $f(x)$ is a given function describing the width of the wire,

$$\mathsf{U}_{\mathsf{XX}} + \mathsf{U}_{\mathsf{YY}} = 0 \tag{8.32a}$$

along with the same imposed-voltage Dirichlet boundary conditions at $\mathsf{X} = 0$ and $\mathsf{X} = \mathsf{L}$, but it is now notable that the lengths of the ends may differ

$$\mathsf{U}(0, \mathsf{Y}) = \bar{U} g_0(\mathsf{Y}/\mathsf{H}) \qquad \text{for } 0 \le \mathsf{Y} \le Hf(0), \tag{8.32b}$$

$$\mathsf{U}(\mathsf{L}, \mathsf{Y}) = \bar{U} g_1(\mathsf{Y}/\mathsf{H}) \qquad \text{for } 0 \le \mathsf{Y} \le Hf(1). \tag{8.32c}$$

The lower edge of the domain remains straight and the associated no-flux boundary condition corresponds to the first condition in (8.21b)

$$\mathsf{U}_{\mathsf{Y}}(\mathsf{X}, 0) = 0 \qquad \text{for} \quad 0 \le \mathsf{X} \le \mathsf{L}. \tag{8.32d}$$

We must be a little more careful in order to properly determine the form of the no-flux condition, $\hat{\mathbf{n}} \cdot \nabla \mathsf{U} = 0$ along the upper boundary. Recalling from multivariable calculus that the normal to a constant-value contour[3] of a function $\mathsf{B}(\mathsf{X}, \mathsf{Y})$ is given by the gradient of B, we construct $\mathsf{B}(\mathsf{X}, \mathsf{Y}) = \mathsf{Y} - \mathsf{H}f(\mathsf{X}/\mathsf{L})$. This defines our boundary as the zero-level contour of $\mathsf{B}(\mathsf{X}, \mathsf{Y})$. To obtain a unit normal vector, we calculate the normalised gradient:

[1] This is sometimes called a "domain perturbation" [61], also see Hinch [47].

[2] A separable domain being expressible as Cartesian products of intervals in independent variables, for example $(\mathsf{X} \in [0, \mathsf{L}]) \times (\mathsf{Y} \in [0, \mathsf{H}])$.

[3] Sometimes called a level set.

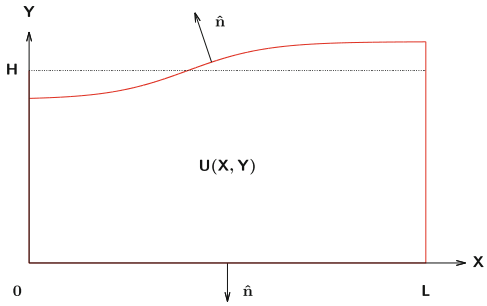


Fig. 8.2 The domain for the non-uniform insulated wire, problem (8.32)

$$\hat{\mathbf{n}} = \pm\frac{\nabla \mathsf{B}}{|\nabla \mathsf{B}|} = \pm\frac{(-(\mathsf{H}/\mathsf{L})f'(\mathsf{X}/\mathsf{L}), 1)}{\sqrt{1+(\mathsf{H}/\mathsf{L})^2 f'(\mathsf{X}/\mathsf{L})^2}},$$

where the $\pm$ sign is chosen so as to correspond to the unit outward normal, pointing out of the domain at the boundary (in this case $+$ for the upper boundary). Finally, taking the dot product with the gradient $\nabla \mathsf{U}$ yields the condition

$$-\frac{\mathsf{H}}{\mathsf{L}} f'(\mathsf{X}/\mathsf{L}) \frac{\partial \mathsf{U}}{\partial \mathsf{X}}\bigg|_{(\mathsf{X},\mathsf{H}f(\mathsf{X}/\mathsf{L}))} + \frac{\partial \mathsf{U}}{\partial \mathsf{Y}}\bigg|_{(\mathsf{X},\mathsf{H}f(\mathsf{X}/\mathsf{L}))} = 0. \tag{8.32e}$$

Employing the previously used scalings (8.11) we arrive at the nondimensional problem on $0 \le x \le 1$ and $0 \le y \le f(x)$,

$$\varepsilon^2 u_{xx} + u_{yy} = 0, \tag{8.33a}$$

subject to the boundary conditions

$$u(0, y) = g_0(y) \qquad 0 \le y \le f(0), \tag{8.33b}$$
$$u(1, y) = g_1(y) \qquad 0 \le y \le f(1), \tag{8.33c}$$

and for $0 \le x \le 1$

$$u_y(x, 0) = 0, \qquad u_y(x, f(x)) - \varepsilon^2 f'(x) u_x(x, f(x)) = 0. \tag{8.33d}$$

Following as in Sect. 8.3, we begin by constructing the outer solution, and at leading order obtain the problem,

$$u_{0yy} = 0, \qquad u_{0y}(x, 0) = 0, \qquad u_{0y}(x, f(x)) = 0, \tag{8.34}$$

yielding $u_0(x, y) = C_2(x)$ as in (8.23). To determine $C_2(x)$ we look to the next order problem,

$$u_{0xx} + u_{1yy} = 0, \qquad u_{1y}(x, 0) = 0, \qquad u_{1y}(x, f) = f'(x)u_{0x}(x, f). \tag{8.35}$$

Substituting in for u_0, we find the general solution,

$$u_1(x, y) = -\tfrac{1}{2}C_2''(x)y^2 + C_3(x)y + C_4(x). \tag{8.36}$$

Imposing the $y = 0$ boundary condition determines $C_3 \equiv 0$. The top boundary condition at $y = f(x)$ then takes the form

$$u_{1y}(x, f(x)) = f'(x)C_2'(x);$$

equating this expression with the y-derivative of the general $u_1(x, y)$ solution given by (8.36) evaluated at $y = f$ yields the compatibility condition

$$-C_2''(x)f(x) = f'(x)C_2'(x).$$

Using the product rule, this result can be expressed more compactly as

$$\frac{d}{dx}\left(f(x)\frac{dC_2}{dx}\right) = 0; \tag{8.37}$$

this is a generalisation of (8.25) for non-constant $f(x)$. Boundary conditions for (8.37) are then determined by solving for the boundary layers at $x_*^L = 0$ and $x_*^R = 1$ using separation of variables as was done in Sect. 8.3 to construct the uniform solution on the whole domain (8.20).

8.5 Further Directions

The main asymptotic reduction considered in this chapter is based on having the aspect ratio as the limiting small parameter, $\varepsilon = H/L \to 0$. Stemming from the historical uses of this approach in the context of many problems in fluid dynamics [1, 61] (involving the flow of oil between machine parts, and other thin layers like tears films on the eye), these types of problems are sometimes called *lubrication models* [51, 79, 96]. They offer valuable simplifications to problems by reducing systems to lower-dimensional domains (in our case from a PDE on a two-dimensional region to an ODE on a one-dimensional interval) by separating out the "thin" directions to leave an underlying governing model on the "wide" spatial directions with the appropriate boundary conditions incorporated. These problems are also sometimes called *slender body asymptotics*, and *reduced-dimension models*.

8.6 Exercises

8.1 Consider the problem for Laplace's equation on $0 \le \mathsf{X} \le \mathsf{L}$ and $0 \le \mathsf{Y} \le \mathsf{H}$,

$$\mathsf{U}_{\mathsf{XX}} + \mathsf{U}_{\mathsf{YY}} = 0,$$

with boundary conditions

$$\begin{aligned} &\mathsf{U}(0,\mathsf{Y}) = 0, \qquad \mathsf{U}(\mathsf{L},\mathsf{Y}) = 0 \qquad \text{for} \quad 0 \le \mathsf{Y} \le \mathsf{H}, \\ &\mathsf{U}(\mathsf{X},0) = 0, \qquad \left.\frac{\partial \mathsf{U}}{\partial \mathsf{Y}}\right|_{\mathsf{Y}=\mathsf{H}} = e^{-3\mathsf{X}/\mathsf{L}} \quad \text{for} \quad 0 \le \mathsf{X} \le \mathsf{L}. \end{aligned}$$

(a) Use separation of variables to construct the exact solution $\mathsf{U}(\mathsf{X},\mathsf{Y})$ (valid for any values of $\mathsf{H}, \mathsf{L} > 0$).
(b) Nondimensionalize the problem and consider the slender limit of $\varepsilon = \mathsf{H}/\mathsf{L} \to 0$ to find the leading order outer solution $u_0(x, y)$.
(c) Rescale the solution from (a) by the natural lengthscales and show that it is equivalent to the solution from (b) in the limit $\varepsilon \to 0$.

8.2 Consider the problem for Laplace's equation on $0 \le \mathsf{X} \le \mathsf{L}$ and $0 \le \mathsf{Y} \le \mathsf{H}$ in the limit that $\mathsf{H} \to 0$ with $\mathsf{L} = O(1)$,

$$\mathsf{U}_{\mathsf{XX}} + \mathsf{U}_{\mathsf{YY}} = 0,$$

with boundary conditions

$$\begin{aligned} &\mathsf{U}(0,\mathsf{Y}) = \sin\left(\frac{\pi \mathsf{Y}}{2\mathsf{H}}\right), \qquad \mathsf{U}(\mathsf{L},\mathsf{Y}) = -(\mathsf{Y}/\mathsf{H})^2 \qquad \text{for} \quad 0 \le \mathsf{Y} \le \mathsf{H}, \\ &\mathsf{U}(\mathsf{X},0) = \sin\left(\frac{5\pi \mathsf{X}}{\mathsf{L}}\right), \qquad \mathsf{U}(\mathsf{X},\mathsf{H}) = \cos\left(\frac{3\pi \mathsf{X}}{\mathsf{L}}\right) \qquad \text{for} \quad 0 \le \mathsf{X} \le \mathsf{L}. \end{aligned}$$

(a) Determine the leading-order outer solution.
(b) Determine the boundary layer corrections for the left and right leading-order inner solutions.
(c) Determine the leading order uniformly valid solution.
(d) Use (c) to show that the boundary layer is crucial for calculating the leading order value of the average flux at the right end of the domain:

$$\mathsf{J}^R = \frac{1}{\mathsf{H}} \int_0^{\mathsf{H}} \left(\frac{\partial \mathsf{U}}{\partial \mathsf{X}}\Big|_{\mathsf{X}=\mathsf{L}}\right) d\mathsf{Y}.$$

Hint: Your solution will involve a sum given by the Riemann zeta function, $\sum_{k=0}^{\infty} 1/(2k+1)^3 = \frac{7}{8}\zeta(3) \approx 1.0518$.

8.3 Consider the problem for Laplace's equation on the non-uniform slender domain, $0 \le \mathsf{X} \le \mathsf{L}$ and $0 \le \mathsf{Y} \le F(\mathsf{X})$ where $F(\mathsf{X}) = 15 + 5\cos(3\pi\mathsf{X}/\mathsf{L})$ in the limit $\mathsf{L} \to \infty$ with $\mathsf{H} = 1$,

$$\mathsf{U}_{\mathsf{XX}} + \mathsf{U}_{\mathsf{YY}} = 0,$$

with boundary conditions

$$\mathsf{U}(0,\mathsf{Y}) = -\tfrac{1}{100}\mathsf{Y}^3, \qquad \mathsf{U}(\mathsf{L},\mathsf{Y}) = 3\mathsf{Y}^2 \qquad \text{for} \quad 0 \le \mathsf{Y} \le F(\mathsf{X}),$$

$$\left.\frac{\partial \mathsf{U}}{\partial \mathsf{Y}}\right|_{\mathsf{Y}=0} = 0, \qquad \left.\hat{\mathbf{n}} \cdot \nabla \mathsf{U}\right|_{\mathsf{Y}=F(\mathsf{X})} = 0 \qquad \text{for} \quad 0 \le \mathsf{X} \le \mathsf{L}.$$

(a) Nondimensionalize the problem.
(b) Write the equations for the outer solution up to $O(\varepsilon^2)$. Determine the first two terms in the expansion of the outer solution.
(c) Determine the boundary layer corrections for the left and right leading-order inner solutions. Hint: What is y^{BL} in $0 \le y \le y^{\mathrm{BL}}$ in each boundary layer?
(d) Use matching to determine the x-boundary conditions for the outer solution and then obtain the leading-order outer solution.

8.4 Consider the problem for Laplace's equation for a curved semicircular arc of wire of varying width on $0 \le \theta \le \pi$ with $\bar{\mathsf{R}} - \mathsf{W}f(\theta) \le \mathsf{R} \le \bar{\mathsf{R}} + \mathsf{W}f(\theta)$ in the limit of a large radius of curvature, $\bar{\mathsf{R}} \to \infty$,

$$\frac{\partial^2 \mathsf{U}}{\partial \mathsf{R}^2} + \frac{1}{\mathsf{R}}\frac{\partial \mathsf{U}}{\partial \mathsf{R}} + \frac{1}{\mathsf{R}^2}\frac{\partial^2 \mathsf{U}}{\partial \theta^2} = 0,$$

with boundary conditions

$$\left.\hat{\mathbf{n}} \cdot \nabla \mathsf{U}\right|_{\mathsf{R}=\bar{\mathsf{R}}-\mathsf{W}f(\theta)} = 0, \qquad \left.\hat{\mathbf{n}} \cdot \nabla \mathsf{U}\right|_{\mathsf{R}=\bar{\mathsf{R}}+\mathsf{W}f(\theta)} = 0 \qquad \text{for} \quad 0 \le \theta \le \pi,$$

$$\mathsf{U}(\mathsf{R},0) = \bar{\mathsf{U}}g_0(\mathsf{R}/\bar{\mathsf{R}}), \qquad \mathsf{U}(\mathsf{R},\pi) = \bar{\mathsf{U}}g_1(\mathsf{R}/\bar{\mathsf{R}}) \qquad \text{for} \quad 0 \le \mathsf{X} \le \mathsf{L}.$$

(a) Nondimensionalize the problem using $\mathsf{R} = \bar{\mathsf{R}}r$, but show that this does not yield a scaled problem with a small parameter.
(b) Let $r = 1 + \varepsilon y$. What should be used for the aspect ratio ε? Write the complete problem satisfied by $u(y,\theta)$.
(c) Write the problems for the outer solution, $u(y,\theta) \sim u_0(y,\theta) + \varepsilon u_1(y,\theta) + \varepsilon^2 u_2(y,\theta)$.

8.5 The derivation of the *Korteweg de Vries* (KdV) equation for waves on the surface of shallow layers of water follows from similar asymptotic reductions of Laplace's equation with appropriate boundary conditions [31, 58, 80]. Consider the following nondimensionalized problem for a time-dependent potential $\phi(x,y,t)$ and the shape of waves on the top-surface of a fluid layer, $F(x,t)$, on $-\infty < x < \infty$,

$$\varepsilon^2\phi_{xx} + \phi_{yy} = 0 \quad \text{on } 0 \le y \le 1 + \varepsilon^2 F \tag{8.38a}$$

$$\phi_y = 0 \quad \text{at } y = 0 \tag{8.38b}$$

$$\phi_t + \tfrac{1}{2}(\varepsilon^2\phi_x^2 + \phi_y^2) + F = 0 \quad \text{at } y = 1 + \varepsilon^2 F \tag{8.38c}$$

$$\varepsilon^2 F_t + \varepsilon^4 \phi_x F_x = \phi_y \quad \text{at } y = 1 + \varepsilon^2 F \tag{8.38d}$$

The first steps in the derivation are:

(a) Expand $\phi = \phi_0 + \varepsilon^2\phi_1 + \varepsilon^4\phi_2 + O(\varepsilon^6)$ and $F = F_0 + \varepsilon^2 F_1 + O(\varepsilon^4)$ for $\varepsilon \to 0$ and use (8.38a, 8.38b) to determine ϕ_0, ϕ_1, ϕ_2 in terms of polynomials in y and three un-determined functions $C_0(x,t), C_1, (x,t), C_2(x,t)$.
(b) Determine the leading order equations for C_0, f_0 obtained from substituting the expansions for ϕ, F into (8.38c, 8.38d).
(c) Determine the equations for the $O(\varepsilon^2)$ next-order equations obtained from (8.38c, 8.38d). (To be concluded in Exercise 9.12.)

8.6 The derivation of the *porous medium equation*, introduced earlier as Eq. (5.21),

$$\frac{\partial h}{\partial t} = \frac{1}{3}\frac{\partial}{\partial x}\left(h^3 \frac{\partial h}{\partial x}\right),$$

follows from a long-wave analysis for the evolution of the height of a layer of a viscous fluid $y = h(x,t)$ [1, 79]. Consider the $\varepsilon \to 0$ limit of the problem

$$\varepsilon^2 u_{xx} + u_{yy} = -h_x \quad \text{on } 0 \le y \le h \tag{8.39a}$$

$$u_x + v_y = 0 \quad \text{on } 0 \le y \le h \tag{8.39b}$$

$$u = v = 0 \quad \text{at } y = 0 \tag{8.39c}$$

$$u_y = 0 \quad \text{at } y = h \tag{8.39d}$$

$$h_t + uh_x = v \quad \text{at } y = h \tag{8.39e}$$

where $u(x,y,t), v(x,y,t)$ are the horizontal and vertical components of the fluid velocity respectively and the last three equations are boundary conditions on the first two.

(a) Show that by using (8.39b, 8.39c) to express v in terms of u, Eq. (8.39e) can be written as

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}\left(\int_0^h u\, dy\right) = 0. \tag{8.40}$$

(b) Determine the leading order horizontal velocity, $u_0(x,y,t)$ in terms of h from (8.39a, 8.39c, 8.39d) to then obtain the porous medium equation from (8.40).

8.7 Note that in all of the problems considered here we assumed a perturbation expansion for the outer solution given in powers of ε^2, (8.13). Sometimes this must be modified. Consider problem (8.12) with the Dirichlet condition $u(x, 1) = x$ on the top boundary. Examine the expansion of the boundary layers to show that if we wish to get a solution that is uniformly valid to $O(\varepsilon^2)$ then the expansion $u \sim u_0 + \varepsilon u_1 + \varepsilon^2 u_2$ is needed.

Chapter 9
Weakly-Nonlinear Oscillators

Many systems function through repeated cycles of operation—from the spinning of gears in a machine, the physiology of heartbeats, biological behaviour on the 24 h circadian cycle, to seasonal climate changes during the motion of the earth around the sun each year. The most basic models for such systems take the form of oscillator equations, having regular predictable periodic solutions; the simplest such model is the *linear oscillator* equation,

$$\frac{d^2x}{dt^2} + \omega_0^2 x = 0, \tag{9.1}$$

with *natural frequency* ω_0. It has the general solution $x(t) = A\sin(\omega_0 t) + B\cos(\omega_0 t)$, with the constants A and B being determined by the initial conditions imposed on the system.

More detailed models of oscillatory systems include additional terms that describe other influences that modify the solutions. If these effects are weak, as indicated by a dimensionless system parameter being small, then the model can be written as

$$\frac{d^2x}{dt^2} + \omega_0^2 x = \varepsilon f\left(x, \tfrac{dx}{dt}, t\right), \quad \varepsilon \to 0. \tag{9.2}$$

Such models are called *weakly-nonlinear oscillators* since they reduce to the linear oscillator (9.1) when the perturbation terms (potentially including nonlinearities) are suppressed. In the context of mechanical systems, Eq. (9.2) can describe the small-amplitude motion of a pendulum, or a mass attached to a nonlinear spring.

In the framework of the earlier chapters on perturbation methods, (9.2) may appear to be a straightforward regular perturbation problem, but we will see that regular perturbation expansions will not be able to address the question of greatest interest for oscillating systems, which is,

> *If we understand the behaviour of the system for a single cycle of the oscillation, can we determine how the perturbation forcing terms cumulatively affect the problem over long times and many oscillation periods?*

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9_9

Two perturbation methods will be described to illustrate how weak influences can be incorporated into the leading order solution to obtain more accurate long-time predictions of oscillatory behaviour.

9.1 Review of Solutions of the Linear Problem

We begin by briefly reviewing the essential results for the linear oscillator equation that form the basis for the perturbation methods for (9.2). As discussed above, the unforced linear oscillator equation is characterised by a natural frequency $\omega_0 > 0$ and has the general homogeneous periodic solution $x_h(t)$,

$$\frac{d^2x}{dt^2} + \omega_0^2 x = 0 \qquad \Longrightarrow \qquad x_h(t) = A\sin(\omega_0 t) + B\cos(\omega_0 t). \tag{9.3}$$

Now consider the inhomogeneous version of this equation, driven by a *periodic* forcing function $f(t)$,

$$x'' + \omega_0^2 x = f(t).$$

The function f can be directly replaced by its Fourier series, $f(t) = \sum_k C_k \sin(\gamma_k t) + D_k \cos(\gamma_k t)$, and then by the linearity of the equation, the overall solution will be the sum of the contributions from each of the terms in the series for f, yielding $x(t) = x_h(t) + \sum_k x_k(t)$. Each Fourier term yields a problem of the form

$$x_k'' + \omega_0^2 x_k = C_k \sin(\gamma_k t) + D_k \cos(\gamma_k t). \tag{9.4}$$

If $\gamma_k \neq \omega_0$ then the general solution is given as a combination of homogeneous and particular solutions as

$$x_k(t) = \underbrace{A\sin(\omega_0 t) + B\cos(\omega_0 t)}_{\text{Homogeneous solution}} + \underbrace{\frac{C_k}{\omega_0^2 - \gamma_k^2}\sin(\gamma_k t) + \frac{D_k}{\omega_0^2 - \gamma_k^2}\cos(\gamma_k t)}_{\text{Particular solution}}.$$

The particular solution can be obtained from the method of undetermined coefficients or other elementary approaches. This solution is not valid for $\gamma_k = \omega_0$, which is called the case of *resonant forcing* (forcing at a natural frequency of the system). For resonant forcing, the method of undetermined coefficients suggests a different form for the particular solution, $x_p(t) = c_1 t\sin(\omega_0 t) + c_2 t\cos(\omega_0 t)$. Substituting this into (9.4) and matching coefficients of corresponding terms yields the general solution as

$$x_{\omega_0}(t) = \underbrace{A\sin(\omega_0 t) + B\cos(\omega_0 t)}_{\text{Homogeneous solution}} + \underbrace{\frac{C}{2\omega_0} t\cos(\omega_0 t) + \frac{D}{2\omega_0} t\sin(\omega_0 t)}_{\text{Resonant forced response}}.$$

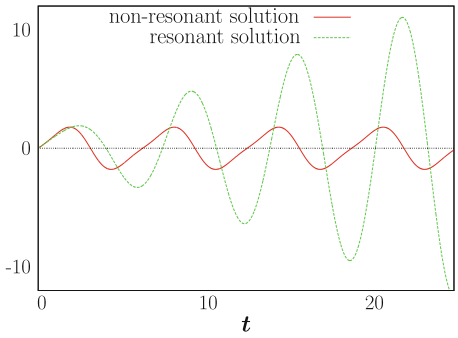


Fig. 9.1 Solutions of the forced linear oscillator equation (9.4) with $\omega_0 = 1$: the non-resonant solution for $\gamma_k = 2$ and resonant solution for $\gamma_k = 1$

The resonant response, while being oscillatory with period $2\pi/\omega_0$, is notable for having an amplitude that exhibits unbounded growth with increasing time. The terms $t\cos\omega_0 t$ and $t\sin\omega_0 t$ are commonly called *secular growth terms* (see Fig. 9.1). We will see that the occurrence of such terms is the central issue that must be addressed in constructing accurate long-time asymptotic solutions for perturbed oscillators.

9.2 The Failure of Direct Regular Expansions

In order to illustrate the shortcomings of the standard regular perturbation expansion approach, we consider two simple perturbed linear oscillator problems in the limit $\varepsilon \to 0$, namely

$$x'' + x = -\varepsilon x, \qquad x(0) = 1, \qquad x'(0) = 0, \tag{9.5}$$

and

$$x'' + x = -\varepsilon x', \qquad x(0) = 1, \qquad x'(0) = -\tfrac{1}{2}\varepsilon. \tag{9.6}$$

Assuming the solutions to be perturbation expansions of the form

$$x(t) = x_0(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + \cdots, \tag{9.7}$$

yields an ordered sequence of initial value problems for each $x_n(t)$ with the leading order for both (9.5) and (9.6) reducing to (9.3) with $\omega_0 = 1$. Solving these sequences of sub-problems up to $O(\varepsilon^2)$, yields the solution of (9.5) as

$$x(t) \sim \cos t - \tfrac{1}{2}\varepsilon t \sin t + \tfrac{1}{8}\varepsilon^2 \left(t\sin t - t^2\cos t\right), \tag{9.8}$$

while the solution of (9.6) is given by

$$x(t) \sim \cos t - \tfrac{1}{2}\varepsilon t \cos t + \tfrac{1}{8}\varepsilon^2 \left(t\sin t + t^2\cos t\right). \tag{9.9}$$

While these two expansions appear quite similar in form, this masks the fact that they come from problems predicting very different behaviours. Problem (9.5) can be re-written in the form (9.3) as $x'' + (1+\varepsilon)x = 0$ and it is straight forward to show that its exact solution is

$$x_{\text{exact}}(t) = \cos(\sqrt{1+\varepsilon}\, t). \tag{9.10}$$

In other words, the exact solution has a slightly perturbed natural frequency, $\omega_0 = \sqrt{1+\varepsilon}$, compared to $\omega_0 = 1$ for the leading order problem, $x_0'' + x_0 = 0$. In contrast, (9.6) is a weakly damped linear oscillator of the form $x'' + \varepsilon x' + x = 0$, and its exact solution is given by

$$x_{\text{exact}}(t) = e^{-\varepsilon t/2} \cos\left(\sqrt{1 - \tfrac{1}{4}\varepsilon^2}\, t\right). \tag{9.11}$$

The respective periodic and decaying oscillatory behaviours of these two solutions are shown in Fig. 9.2 along with plots of the expansions (9.8) and (9.9). While the expansions match their corresponding solutions for early times, both dramatically diverge when their later time behaviours become dominated by the secular growth terms.

While these results might suggest that the perturbation expansions have produced incorrect descriptions, we should not dismiss them too soon. In fact, taking the Taylor series expansions of (9.10) and (9.11) for $\varepsilon \to 0$ at fixed finite times ($t = O(1)$) directly reproduces (9.8) and (9.9). Equations (9.8) and (9.9) are indeed correct, but they must be used with caution.

Applying the fundamental assumption of asymptotic ordering to the terms in (9.7), the expansions are valid only when

$$|x_0(t)| \gg |\varepsilon x_1(t)| \gg |\varepsilon^2 x_2(t)| \gg \cdots, \qquad \varepsilon \to 0. \tag{9.12}$$

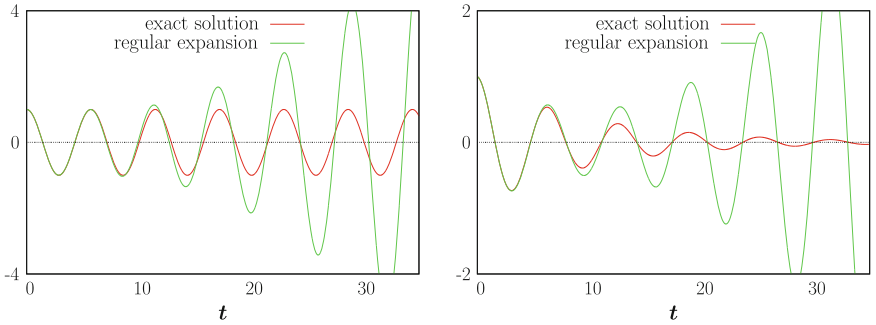


Fig. 9.2 Exact solutions and regular expansions for problem (9.5): (9.10) and (9.8) (*Left*) and for problem (9.6): (9.11) and (9.9) (*Right*), both with $\varepsilon = 1/5$

Considering the first two terms from (9.8), $|\cos t| \gg |\frac{1}{2}\varepsilon t \sin t|$ holds only if $1 \gg \varepsilon t$. In other words, the expansions can only be expected to hold while this product is small ("for a limited time only" is an appropriate phrase, here applying for $0 \leq t \ll 1/\varepsilon$).

These examples are be a bit troubling—we have worked out higher order terms in the asymptotic expansions since the leading order solution by itself, $x_0(t) = \cos t$, does not capture the differences that distinguish these two problems from each other. Yet the additional terms yield contributions that restrict the validity of the expansion to relatively short times, when effects like slow growth or decay or changes in the oscillation frequency have not yet fully developed. This suggests the need for other forms of perturbation expansions that can overcome these limitations.

9.3 Poincare–Lindstedt Expansions

The failure of the regular expansion to describe periodic motion on long timescales in the context of astronomy (for the motion of the planets) motivated the development of an improved approach known as the Poincare–Lindstedt method. A key idea behind the method is that the regular perturbation expansion is too restrictive in its form and does not allow for the possibility that the solution may have a frequency that is shifted from the leading order natural frequency ω_0 (as in (9.10)).

The Poincare–Lindstedt approach begins with a change of variables in terms of a frequency that can depend on the perturbation parameter,

$$x(t) = \tilde{x}(\theta) \qquad \text{with} \qquad \theta = \Omega(\varepsilon)t, \tag{9.13}$$

such that $\Omega(0) = \omega_0$. We then seek regular perturbation expansions in terms of the new unknowns,

$$\tilde{x}(\theta) = \tilde{x}_0(\theta) + \varepsilon\tilde{x}_1(\theta) + \cdots, \qquad \Omega(\theta) = \omega_0 + \varepsilon\omega_1 + \cdots. \tag{9.14}$$

We will now give an example to illustrate how this seemingly minor change allows the Poincare–Lindstedt method to eliminate secular growth terms and obtain perturbation expansions that are valid over longer times for some problems.

Consider a problem similar to (9.5),

$$x'' + 4x = -\varepsilon x, \qquad x(0) = 3, \qquad x'(0) = -8. \tag{9.15}$$

After the change of variables (9.13), we arrive at the modified problem for $\tilde{x}(\theta)$,

$$\Omega^2 \frac{d^2\tilde{x}}{d\theta^2} + 4\tilde{x} = -\varepsilon\tilde{x}, \qquad \tilde{x}(0) = 3, \qquad \Omega \left.\frac{d\tilde{x}}{d\theta}\right|_{\theta=0} = -8. \tag{9.16}$$

Substituting expansions (9.14) into (9.16) and separating by orders of ε yields, at $O(\varepsilon^0)$,

$$\omega_0^2 \frac{d^2\tilde{x}_0}{d\theta^2} + 4\tilde{x}_0 = 0, \qquad \tilde{x}_0(0) = 3, \qquad \omega_0 \frac{d\tilde{x}_0}{d\theta}\bigg|_{\theta=0} = -8, \tag{9.17a}$$

and at $O(\varepsilon^1)$,

$$\omega_0^2 \frac{d^2\tilde{x}_1}{d\theta^2} + 4\tilde{x}_1 = -\tilde{x}_0 - 2\omega_0\omega_1 \frac{d^2\tilde{x}_0}{d\theta^2}, \tag{9.17b}$$

$$\tilde{x}_1(0) = 0, \qquad \omega_0\tilde{x}_1'(0) = -\omega_1\tilde{x}_0'(0),$$

and so on at higher orders. Note that lower-order terms from the expansion of the solution ($\tilde{x}_k$ for $k = 0, 1, \ldots, n-1$) should be shifted to the right-hand side of the equation for $\tilde{x}_n$ and be treated as known parts of the inhomogeneous forcing.

Identifying the natural frequency as $\omega_0 = 2$ from (9.15) with $\varepsilon = 0$, the leading order solution can then be obtained from (9.17a) as

$$\tilde{x}_0(\theta) = -4\sin\theta + 3\cos\theta. \tag{9.18}$$

Substituting these results into the $O(\varepsilon)$ problem (9.17b) yields

$$4\frac{d^2\tilde{x}_1}{d\theta^2} + 4\tilde{x}_1 = [16\omega_1 - 4]\sin\theta + [-12\omega_1 + 3]\cos\theta, \tag{9.19}$$

$$\tilde{x}_1(0) = 0, \qquad 2\tilde{x}_1'(0) = 4\omega_1.$$

At this point ω_1 is an undetermined constant. Problem (9.19) can be solved for any value ω_1 to obtain $\tilde{x}_1(\theta)$. However, noting the presence of the resonant forcing terms, $\sin\theta$ and $\cos\theta$, on the right-hand side of the equation, the solution would include secular growth terms. But, if those resonant terms in (9.19) were eliminated by an appropriate choice of ω_1, then $\tilde{x}_1(\theta)$ would be bounded and $\tilde{x}_0 + \varepsilon\tilde{x}_1$ would remain asymptotically well-ordered for all times. For this problem, this criterion selects $\omega_1 = 1/4$. This choice yields $\tilde{x}_1(\theta) = \frac{1}{2}\sin\theta$, and then using (9.13) and (9.14) the solution of the original problem can be written as

$$x(t) \sim (-2 + \tfrac{1}{2}\varepsilon)\sin([2 + \tfrac{1}{4}\varepsilon]t) + 3\cos([2 + \tfrac{1}{4}\varepsilon]t). \tag{9.20}$$

This solution holds over $0 \le t \ll 1/\varepsilon^2$; compare this with (9.8), which was valid only over $0 \le t \ll 1/\varepsilon$, see Fig. 9.3. In fact, even without going through the work of solving for $\tilde{x}_1(\theta)$, the Poincare–Lindstedt approach has yielded an improved solution by determining ω_1 through suppressing the resonant terms in the $O(\varepsilon)$ equation; the condition on ω_1 is another example of a *solvability condition*.

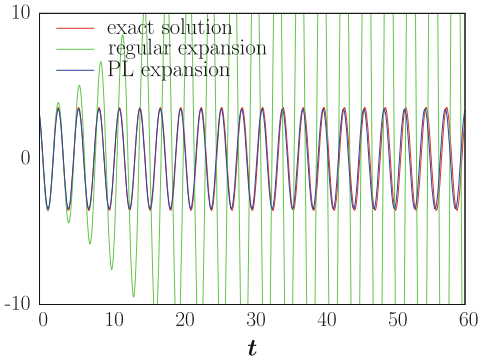


Fig. 9.3 Comparison of the regular expansion and the Poincare–Lindstedt leading order solution against the exact solution of (9.15). Slight shifts between the Poincare–Lindstedt and exact solution are visible for larger times

While the example above was a linear problem, the Poincare–Lindstedt method extends directly to nonlinear equations with perturbation terms of the form $\varepsilon f(x)$. However, readers are right to suspect that the approach has some limitations. Noting that the general solution of the leading order problem has two independent solution terms, $\tilde{x}_0 = A\sin\theta + B\cos\theta$, and that each leads to resonant forcing in (9.19), it should be a little surprising that the coefficients of both resonant terms can be zeroed using only one degree of freedom, ω_1. In fact, this is only possible when $x(t)$ is a *periodic solution*, leading to a degenerate linear system for the coefficients of the resonant forcing terms. Hence, the Poincare–Lindstedt method can only be used to obtain periodic solutions, and cannot, for example, generate the slowly decaying solution (9.11). To overcome this limitation, we consider another related perturbation method.

9.4 The Method of Multiple Time-Scales

Like the Poincare–Lindstedt method, the *method of multiple time-scales* (MMTS) determines solutions to perturbed oscillators by suppressing resonant forcing terms that would yield spurious secular terms in the asymptotic expansions. The method of multiple time-scales makes a less restrictive assumption on the form of the solution than employed by the Poincare–Lindstedt method; it assumes that the solution can be expressed as a function of multiple (for our purposes, just two) time variables,

$$x(t) = X(t, \tau), \tag{9.21}$$

where t is the regular (or "fast") time variable and τ is a new variable describing the "slow-time" dependence of the solution. In a physical context, t could represent a circadian rhythm of a daily cycle, while τ might describe changes to this cycle that are only noticeable over the timespan of years.

The simplest approach to determining the choice of the slow time variable, is to identify the combinations of ε and t present in secular terms in the regular expansion. In the examples from Sect. 9.2, $\tau = \varepsilon t$, so we focus on this case.[1]

The first step is to perform the change of variables (9.21), where by using the chain rule, we arrive at

$$\frac{dx}{dt} = \frac{\partial X}{\partial t} + \frac{\partial X}{\partial \tau}\frac{d\tau}{dt} = \frac{\partial X}{\partial t} + \varepsilon\frac{\partial X}{\partial \tau}, \tag{9.22}$$

and similarly,

$$\frac{d^2x}{dt^2} = \frac{\partial^2 X}{\partial t^2} + 2\varepsilon\frac{\partial^2 X}{\partial t\partial \tau} + \varepsilon^2\frac{\partial^2 X}{\partial \tau^2}. \tag{9.23}$$

The autonomous weakly-nonlinear oscillator equation

$$\frac{d^2x}{dt^2} + x = \varepsilon f\left(x, \tfrac{dx}{dt}\right) \tag{9.24}$$

then becomes

$$\frac{\partial^2 X}{\partial t^2} + 2\varepsilon\frac{\partial^2 X}{\partial t\partial \tau} + \varepsilon^2\frac{\partial^2 X}{\partial \tau^2} + X = \varepsilon f(X, X_t + \varepsilon X_\tau)$$

which can be rearranged into a more useful form as

$$\frac{\partial^2 X}{\partial t^2} + X = \varepsilon\big(f(X, X_t + \varepsilon X_\tau) - 2X_{t\tau}\big) - \varepsilon^2 X_{\tau\tau}. \tag{9.25}$$

We still have a perturbed linear oscillator, but now, we must specify that the oscillation is with respect to the fast time variable t, and we note that the right-hand side perturbation involves additional terms and derivatives with respect to t and τ. It may seem peculiar to replace the simpler ODE (9.24) by a more complicated partial differential equation, but (9.25) can still be solved through application of ODE methods and it provides the degrees of freedom necessary to properly describe the system over longer times.

The next step is to expand the MMTS solution as a regular perturbation expansion,

$$X(t, \tau) = X_0(t, \tau) + \varepsilon X_1(t, \tau) + O(\varepsilon^2), \tag{9.26}$$

and substitute into (9.25). Analogous to the Poincare–Lindstedt method, we want to suppress any resonant terms that occur in order to determine the currently unspecified parts of the solution. The final step is then to reconstruct the relationship between fast and slow times, say $\tau = \varepsilon t$, to obtain the final solution as $x(t) \sim X_0(t, \varepsilon t)$.

[1] Other problems, and additional timescales needed for higher order expansions could involve higher-order (slower) timescales such as $\tau_k = \varepsilon^k t$ for $k = 2, 3, \ldots$.

As an example, we consider a damped nonlinear oscillator,

$$\frac{d^2x}{dt^2} + x = -\varepsilon\kappa\frac{dx}{dt} + \varepsilon x^3, \qquad x(0) = 1, \qquad x'(0) = -2, \tag{9.27}$$

for $\varepsilon \to 0$. Setting $\tau = \varepsilon t$, and following the method of multiple time-scales decomposes the problem into the leading order problem

$$X_{0tt} + X_0 = 0, \qquad X_0(0,0) = 1, \qquad X_{0t}(0,0) = -2, \tag{9.28a}$$

and at $O(\varepsilon)$,

$$X_{0tt} + X_0 = -\kappa X_{0t} + X_0^3 - 2X_{0t\tau}, \tag{9.28b}$$

$$X_1(0,0) = 0, \qquad X_{1t}(0,0) = -X_{0\tau}(0,0).$$

With respect to the fast-time variable t, (9.28a) has a solution that is a linear combination of $\sin t$ and $\cos t$. However, unlike (9.3), since other (slow time) variables appear, the coefficients in the linear combination are not constants, but are functions dependent on the slow-time variable τ,

$$X_0(t,\tau) = A(\tau)\sin t + B(\tau)\cos t. \tag{9.29}$$

From the initial conditions on X_0 at $t = \tau = 0$, we find that

$$A(0) = -2, \qquad B(0) = 1, \tag{9.30}$$

but otherwise, $A(\tau)$ and $B(\tau)$ are as-yet undetermined functions.

Moving on to the $O(\varepsilon)$ problem (9.28b), after substituting in (9.29), we obtain

$$\begin{aligned} X_{1tt} + X_1 = &\left(-\kappa A + \tfrac{3}{4}A^2B + \tfrac{3}{4}B^3 - 2\frac{dA}{d\tau}\right)\cos t \\ &+ \left(\kappa B + \tfrac{3}{4}A^3 + \tfrac{3}{4}AB^2 + 2\frac{dB}{d\tau}\right)\sin t \\ &+ \left(\tfrac{1}{4}B^3 - \tfrac{3}{4}A^2B\right)\cos(3t) - \left(\tfrac{1}{4}A^3 - \tfrac{3}{4}AB^2\right)\sin(3t). \end{aligned} \tag{9.31}$$

Note that many of the terms on the right-hand side of (9.31) are a consequence of the nonlinear term X_0^3. The forcing terms **must** be expanded out as the sum of sines and cosines of the fast time scale with coefficients that can only depend on the slow time scale. For simple nonlinear products, like X_0^3, this can done be through the use of trigonometric identities (see Appendix A). For more complicated types of nonlinear forcing terms, the right side of (9.25) should be replaced by its Fourier series expansion (also see Appendix A). These Fourier series decompositions are essential in separating out the resonant and non-resonant forcing terms. For (9.31), $\cos t$ and $\sin t$ are resonant terms, while $\cos(3t)$, $\sin(3t)$ are non-resonant.

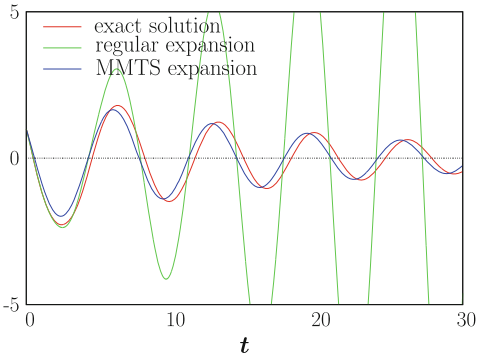


Fig. 9.4 Comparison of the regular expansion and the method of multiple time-scales leading order solution against the exact solution of (9.27)

To eliminate the resonant terms that would cause expansion (9.26) to break-down, we set the respective coefficients to zero,

$$-\kappa A + \tfrac{3}{4}A^2 B + \tfrac{3}{4}B^3 - 2\frac{dA}{d\tau} = 0, \qquad \kappa B + \tfrac{3}{4}A^3 + \tfrac{3}{4}AB^2 + 2\frac{dB}{d\tau} = 0. \tag{9.32}$$

These two equations are the solvability conditions for this problem. They are coupled ordinary differential equations in terms of the slow-time variable that describe the evolution of the amplitude coefficients $A(\tau)$, $B(\tau)$ in solution (9.29); consequently, they are also often called *amplitude equations*.

Solving the amplitude equations to determine $A(\tau)$, $B(\tau)$ subject to their initial conditions (9.30) yields the MMTS leading order approximation of the solution (9.29), as illustrated in Fig. 9.4.

9.5 Further Directions

Perturbation methods for weakly nonlinear oscillators have been developed extensively in connection with many applications ranging from mechanical oscillations and electrical systems to population dynamics. Some more detailed introductory presentations are given in [48, 54, 56, 78], and some advanced treatments are given in [11, 47, 58, 73, 77]. Besides Poincare–Lindstedt and the method of multiple time-scales, other related approaches also exist, including the method of averaging [77, 92, 102] and near-identity transformations [58, 73]. The engineering approach of *harmonic balance* is related to these methods.

The mathematical theory underpinning the solvability conditions coming from the Poincare–Lindstedt and method of multiple time-scales approaches is the *Fredholm alternative theorem* [39, 93]. The Fredholm alternative supplies a criterion for the existence or uniqueness of solutions of inhomogeneous linear problems that applies to these oscillator equations and other classes of perturbation problems (see Exercise 9.13).

9.6 Exercises

9.1 Consider the problem for $x(t)$ with $\varepsilon \to 0$,

$$\frac{d^2x}{dt^2} + x = 32\varepsilon x^3, \qquad x(0) = e^{-\varepsilon}, \qquad x'(0) = 2 + \varepsilon.$$

Obtain the first two terms of the regular expansion, $x(t) = x_0(t) + \varepsilon x_1(t) + O(\varepsilon^2)$. Identify the homogeneous solution, resonant response and non-resonant response in $x_1(t)$.

9.2 For each problem use the Poincare–Lindstedt expansion with the two-term approximation of the solution, $x(t) \sim \tilde{x}_0(\theta) + \varepsilon\tilde{x}_1(\theta)$ with $\theta = (\omega_0 + \varepsilon\omega_1)t$, for $\varepsilon \to 0$ to find $\tilde{x}_0(\theta)$ and ω_0, ω_1.

(a) Are there periodic solutions for every value of $a > 0$ for

$$\frac{d^2x}{dt^2} + 4x = -\varepsilon x^3, \qquad x(0) = a, \qquad x'(0) = 0?$$

(b) Find the value for $a(>0)$ that yields a periodic solution of

$$\frac{d^2x}{dt^2} + 9x = -\varepsilon(x^2 - 1)\frac{dx}{dt}, \qquad x(0) = a, \qquad x'(0) = 0.$$

9.3 Apply the method of multiple scales with $\tau = \varepsilon t$ to the *van der Pol oscillator*

$$\frac{d^2x}{dt^2} + 9x = -\varepsilon(x^2 - 1)\frac{dx}{dt} \qquad \varepsilon \to 0 \tag{9.33}$$

to obtain a solution in the form $x(t) \sim X_0(t, \tau) = A(\tau)\sin(3t) + B(\tau)\cos(3t)$.

(a) Determine the amplitude equations for $A(\tau)$, $B(\tau)$. Determine initial conditions for A, B in terms of $x(0)$, $x'(0)$.
(b) Let $R(\tau) = \sqrt{A^2 + B^2}$. Determine the equation for $dR/d\tau = f(R)$ using the amplitude equations from part (a). Determine the equilibrium values for R.

9.4 Show that the leading order MMTS solution for weakly nonlinear oscillators can be written in the polar form

$$X_0(t, \tau) = R(\tau)\sin(\omega_0 t + \Phi(\tau)). \tag{9.34}$$

(a) Relate the amplitude R and phase Φ to the coefficients A, B introduced in (9.29).

(b) Show that (9.32) can be rewritten as

$$\frac{dR}{d\tau} = -\tfrac{1}{2}\kappa R, \qquad \frac{d\Phi}{d\tau} = -\tfrac{3}{8}R^2$$

and solve this simpler polar system with appropriate initial conditions.

9.5 Use the method of multiple scales to investigate the near-resonant behaviour of the damped, driven oscillator for $x(t)$ for $\varepsilon \to 0$,

$$\frac{d^2x}{dt^2} + \varepsilon\beta\frac{dx}{dt} + x + \varepsilon\alpha x^3 = \varepsilon\cos(t + \gamma\varepsilon t),$$

with given parameters α, β, γ. Use the slow-timescale $\tau = \varepsilon t$. Note the presence of τ in the forcing term on the right-hand side. (Hint: write the forcing as $\cos(t + \gamma\tau)$)

(a) Show that the leading order solution can be written in the complex form

$$X_0(t, \tau) = C(\tau)e^{it} + \overline{C(\tau)}e^{-it},$$

where $\overline{z} = x - iy$ denotes the complex conjugate of $z = x + iy$. Relate the complex-valued function C to the real-valued functions A, B used in (9.29).

(b) Using the result of part (a) in the equation for $X_1(t, \tau)$, find the two solvability conditions. Show that these reduce to a single complex equation for $dC/d\tau$.

(c) *Entrainment* describes a solution locking onto the behaviour entirely set by a forcing term, leaving no direct sign of the natural frequency from the unforced problem (i.e. no homogeneous solution). Setting $C(\tau) = Me^{i\theta}e^{i\gamma\tau}$ in your equation from (b), obtain an equation for M, the real-valued constant amplitude of the entrained solution, with θ being a (real-valued) phase constant. Determine the *detuning relation*, $\gamma = \gamma(M)$, relating the amplitude to the frequency-offset from resonance.

9.6 Apply the method of multiple scales with $\tau = \varepsilon t$ and $\varepsilon \to 0$ to the problem

$$\frac{d^2x}{dt^2} + \varepsilon\left|\frac{dx}{dt}\right|\frac{dx}{dt} + x = 0, \qquad x(0) = 0, \qquad x'(0) = 1.$$

Using the polar form (9.34), derive and solve the amplitude equations for $R(\tau)$ and $\Phi(\tau)$ to obtain the leading order solution $x(t) \sim X_0(t, \tau)$.
Hint: You will need to calculate some terms of a Fourier series. Write the series in terms of the variable $s = t + \Phi$ on $-\pi < s < \pi$, namely $f(s) = \sum_k a_k \sin(ks) + b_k \cos(ks)$.

9.7 In Exercise 3.6, the equation for the parametrically-driven pendulum was derived; for a specific choice of parameters, it can be written as

$$\frac{d^2\theta}{dt^2} + 4\sin\theta = -\varepsilon\sin(4t)\sin\theta.$$

Using the scaling for small-amplitude oscillations, $\theta = \delta x$ for $\delta \to 0$, at leading order this equation yields a form of the *Mathieu equation*,

$$\frac{d^2x}{dt^2} + \left[4 + \varepsilon \sin(4t)\right] x = 0. \tag{9.35}$$

Consider solutions of the Mathieu equation for $\varepsilon \to 0$:

(a) Use the method of multiple scales with $\tau = \varepsilon t$ to determine the amplitude equations for the slowly varying coefficients in the leading order solution.
(b) Solve for the leading order solution that satisfies the initial conditions

$$x(0) = 5, \qquad x'(0) = 6.$$

9.8 Consider the equation for the complex-valued solution $x(t)$ with $\varepsilon \to 0$,

$$\frac{dx}{dt} + i4x = \varepsilon \cos(4t)x^2$$

Apply the method of multiple time scales with $\tau = \varepsilon t$ and $x(t) \sim X_0(t, \tau) + \varepsilon X_1(t, \tau)$.

(a) Write the equations for X_0 and X_1.
(b) Write the general solution of the $O(1)$ equation.
(c) Determine the amplitude equation and explain the condition that selects this result.
(d) Determine the leading order solution for $x(t)$ that satisfies the initial condition

$$x(0) = 1 + i.$$

9.9 For the problem with $\varepsilon \to 0$,

$$\frac{d^2x}{dt^2} + x = \varepsilon x^2, \qquad x(0) = 1, \qquad x'(0) = 0,$$

the slow timescale for MMTS is not the usual one, $\tau \neq \varepsilon t$. Attempt a regular expansion $x \sim x_0 + \varepsilon x_1 + \varepsilon^2 x_2$ to determine the slow time variable. Determine the leading order solution using the Poincare–Lindstedt method.

9.10 For $\varepsilon \to 0$, show that there is a large-amplitude periodic solution of

$$\frac{d^2x}{dt^2} + x = \varepsilon x^2 + \varepsilon \cos t \qquad x'(0) = 0.$$

To do this, let $x(t) \sim \varepsilon^{-\beta} X(t, \tau)$ with $\tau = \varepsilon^\alpha t$ and $X \sim X_0 + \varepsilon^\gamma X_1 + \varepsilon^{2\gamma} X_2$ and select $\alpha, \beta, \gamma > 0$ to ensure that X_0, X_1, X_2 have no secular growth terms.

9.11 *Delay differential equations* (DDE) are equations that involve current properties of the solution coupled to the solution at earlier times [35]. Consider the DDE

$$\frac{d^2x}{dt^2} + x + \varepsilon x(t-2) = 0.$$

(a) This is a linear constant coefficient equation and can be solved in terms of trial solutions of the form $x(t) = e^{\lambda t}$. Write the characteristic equation determining λ and obtain the solutions to $O(\varepsilon)$ for $\varepsilon \to 0$. Are the solutions of this equation stable or unstable? (see Sect. 1.5)
(b) Apply the method of multiple scales to obtain the amplitude equations.
(c) Similarly analyse the *weakly-delayed* linear oscillator equation,

$$\frac{d^2x}{dt^2} + x(t-\varepsilon) = 0.$$

9.12 Now we complete the derivation of the KdV equation that was started in Exercise 8.5. There, we obtained the equations for $C(x,t) = C_0 + \varepsilon^2 C_1 + \cdots$ and $F(x,t) = F_0 + \varepsilon^2 F_1 + \cdots$, as
$O(1)$ equations:

$$\frac{\partial C_0}{\partial t} + F_0 = 0, \qquad \frac{\partial F_0}{\partial t} + \frac{\partial^2 C_0}{\partial x^2} = 0 \tag{9.36}$$

$O(\varepsilon^2)$ equations:

$$\frac{\partial C_1}{\partial t} + F_1 = \tfrac{1}{2}\frac{\partial^3 C_0}{\partial t \partial x^2} - \tfrac{1}{2}\left(\frac{\partial C_0}{\partial x}\right)^2, \tag{9.37a}$$

$$\frac{\partial F_1}{\partial t} + \frac{\partial^2 C_1}{\partial x^2} = \tfrac{1}{6}\frac{\partial^4 C_0}{\partial x^4} - F_0\frac{\partial^2 C_0}{\partial x^2} - \frac{\partial F_0}{\partial x}\frac{\partial C_0}{\partial x}. \tag{9.37b}$$

(a) Show from the $O(1)$ equations that C_0 and F_0 each satisfy the classic wave equation (2.52) with unit speed and having independent left/right moving waves.
(b) We will now restrict attention just to right-moving waves. It can be shown that the $O(\varepsilon^2)$ equations would lead to solutions with secular growth, hence consider a multiple-time scale expansion of the form $C = c_0(z,\tau) + \varepsilon^2 c_1(z,\tau)$ and $F = f_0(z,\tau) + \varepsilon^2 f_1(z,\tau)$ where $z = x - t$ and $\tau = \varepsilon^2 t$. Show that the $O(1)$ equations reduce to a single relation between c_0, f_0 and that the $O(\varepsilon^2)$ equations reduce to a single compatibility condition given by a partial differential equation for $f_0(z,\tau)$, the KdV equation,

$$\frac{\partial f_0}{\partial \tau} + \tfrac{3}{2} f_0 \frac{\partial f_0}{\partial z} + \tfrac{1}{6}\frac{\partial^3 f_0}{\partial z^3} = 0. \tag{9.38}$$

9.13 The *Fredholm alternative theorem* is the key principle that determines the solvability conditions in the Poincare–Lindstedt and MMTS methods, but here we illustrate how it also applied to determining the *eigenvalues of perturbed matrices* [47].

Consider the matrix

$$\mathbf{A}(\varepsilon) = \begin{pmatrix} -e^{\varepsilon} & 3 - 3\varepsilon \\ 3 + \varepsilon & -1 + 2\sin\varepsilon \end{pmatrix}.$$

For the limit $\varepsilon \to 0$, use the following steps to solve the eigenvalue problem,

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}.$$

(a) Write the asymptotic expansion for the matrix in powers of ε, $\mathbf{A} \sim \mathbf{A}_0 + \varepsilon\mathbf{A}_1 + \varepsilon^2\mathbf{A}_2$ and similar expansions for each eigenvector and corresponding eigenvalue $\mathbf{x} \sim \mathbf{x}_0 + \varepsilon\mathbf{x}_1 + \varepsilon^2\mathbf{x}_2$, and $\lambda \sim \lambda_0 + \varepsilon\lambda_1 + \varepsilon^2\lambda_2$.

(b) The leading order problem, $\mathbf{A}_0\mathbf{x}_0 = \lambda_0\mathbf{x}_0$, can be written in the form of a homogeneous problem,

$$(\mathbf{A}_0 - \lambda_0\mathbf{I})\mathbf{x}_0 = \mathbf{0}$$

Obtain the eigensolutions for this matrix, $\{\lambda_0^{(k)}, \mathbf{x}_0^{(k)}\}$.

(c) For this problem, every vector $\mathbf{v}$ can be written as a linear combination of the leading order eigenvectors, $\mathbf{v} = c_{(1)}\mathbf{x}_0^{(1)} + c_{(2)}\mathbf{x}_0^{(2)}$. Write the formulas for $c_{(k)}$ for $k = 1, 2$ in terms of $\mathbf{v}$ and $\mathbf{x}_0^{(k)}$.

(d) Write the $O(\varepsilon)$ equation in the form of an inhomogeneous problem with the same matrix operator as the $O(1)$ equation,

$$(\mathbf{A}_0 - \lambda_0\mathbf{I})\mathbf{x}_1 = \mathbf{f}_1(\mathbf{x}_0, \mathbf{A}_1, \lambda_1).$$

Note the similarities between these equations for $\mathbf{x}_0$, $\mathbf{x}_1$ and the oscillator problems, (9.17a) and (9.17b) or (9.28a) and (9.28b). The Fredholm alternative essentially states that

> *The solution of the $O(\varepsilon^n)$ inhomogeneous problem will exist and be unique if the forcing term has no contribution from the solution of the $O(\varepsilon^0)$ homogeneous problem.*[2]

For oscillator problems, this addresses the existence of *periodic* solutions in the absence of resonant forcing terms.

[2]This is the simplified version for symmetric matrices and self-adjoint differential equations. The general version of the Fredholm alternative is similar:

> *The solution of the non-homogeneous problem $\mathbf{A}_0\mathbf{x} = \mathbf{b}$ will be unique if and only if the adjoint problem $\mathbf{A}_0^{\dagger}\mathbf{y} = \mathbf{0}$ has only the trivial solution* [39, 93].

For matrix equations, non-existence or non-uniqueness will result if the forcing term includes contributions from the $\mathbf{x}_0^{(k)}$ nullvector corresponding to the $\lambda^{(k)}$ in the matrix operator. Select the values of λ_1 to eliminate those contributions and obtain the first two terms in the expansions of the eigenvalues, $\lambda^{(k)} \sim \lambda_0^{(k)} + \varepsilon\lambda_1^{(k)}$ for $k = 1, 2$.

Chapter 10
Fast/slow Dynamical Systems

In previous chapters we employed perturbation methods to study problems in differential equations where the solution of the leading order problem could be obtained in a simple explicit form. For problems where the leading order solution cannot be obtained in a simple form, a different approach is needed. Here we show that ideas from phase plane analysis (Chap. 1) can be used to circumvent these difficulties and allow us to apply boundary layer methods (Chap. 7) to separate problems into dynamics occurring at different scales.

The fundamental principles for formulating rate equations given in Chap. 1 are universally applicable for constructing models. But when the number of equations in the resulting system is large or the leading order problem for $\varepsilon = 0$ is nonlinear, then some of the direct approaches that we have seen in previous chapters become cumbersome and may not be useful.

Similarly, Chap. 9 introduced approaches based on perturbation methods to estimate the behaviour of some models yielding oscillatory behaviours. But those methods were limited to *weakly* nonlinear problems, where the leading order solution satisfied a linear oscillator equation and could be calculated explicitly in terms of sines and cosines. For some more general nonlinear problems, we will show that properties of the system can be determined without the need for a closed-form solution.

The common structure for problems that we will investigate in this chapter is that of singularly-perturbed dynamical systems. In Chap. 7, we've learned that singularly perturbed differential equations exhibit separation of spatial scales, having rapid variations in boundary layers in contrast to slowly-varying outer solutions. Likewise in Chap. 9, a separation of time-scales occurred between the oscillations on the fast time scale and evolution of the amplitude functions on slow scales. Extending this approach to more general dynamical systems, we seek to decompose each problem into two sub-systems:

- *The slow system*: the problem in terms of the original ("slow-time") variable, whose solution will be treated analogously to the outer solution in a boundary layer problem.

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9_10

- *The fast system*: a rescaling of the system describing rapid evolution occurring over shorter times, analogous to the inner solution of a boundary layer problem.

The solution of each sub-system will be sought in the form of a regular perturbation expansion. For singularly-perturbed problems, the sub-systems will have simpler structures than the full problem and allow for the slow- and fast-dynamics to potentially be characterised in terms of reduced phase line or phase plane dynamics.

We will study two problems that could be considered directly in terms of a standard phase plane analysis, but through the use of perturbation theory (that is also applicable to higher order systems), we will show that the underlying dynamics can be separated into more convenient forms.

10.1 Strongly-Nonlinear Oscillators: The van der Pol Equation

For oscillatory systems where the leading problem for $\varepsilon \to 0$ does not yield the harmonic oscillator equation, $x_0''(t) + x_0(t) = 0$, the approaches from Chap. 9 cannot be applied in a straightforward manner. As an example of a "*strongly nonlinear*" oscillator, we will show how to construct solutions of the *van der Pol oscillator*,

$$\varepsilon \frac{d^2x}{dt^2} + (x^2 - 1)\frac{dx}{dt} + x = 0 \qquad \text{for } \varepsilon \to 0. \tag{10.1}$$

This equation is similar to the weakly-nonlinear damped oscillator (9.33), but here the first-order damping term is part of the leading order equation while the second derivative term is a singular perturbation.[1] Since the perturbation parameter here multiplies the highest order derivative, this type of problem could be called a "singularly perturbed oscillator" but the more common name is a *relaxation oscillator*.

We begin by re-writing (10.1) in terms of a convenient formulation for a phase plane analysis but will use a different choice of intermediate variable $y(t)$ than the standard velocity (1.37). Notice that the first two terms in (10.1) can be written as a total derivative,

$$\frac{d}{dt}\left(\varepsilon \frac{dx}{dt} + \tfrac{1}{3}x^3 - x\right) + x = 0.$$

Defining the expression in parentheses as $y = \varepsilon x' + \frac{1}{3}x^3 - x$, we obtain an autonomous system for $(x(t), y(t))$,

$$\varepsilon \frac{dx}{dt} = y + x - \tfrac{1}{3}x^3, \qquad \frac{dy}{dt} = -x, \tag{10.2}$$

[1] Attempting to put this equation in the form (9.2) fails because dividing by ε suggests a very fast oscillation, $\omega_0 = \varepsilon^{-1/2} \to \infty$, and a large ($O(\varepsilon^{-1}) \to \infty$) rather than small perturbation term.

which is called the *Liénard transformation* of the van der Pol equation. The singularly-perturbed nature of this system is clear: for $\varepsilon = 0$, the ODE for $x(t)$ reduces to an algebraic relation.

We begin by determining the outer ("slow") solution using regular perturbation expansions for both x, y:

$$x(t) = x_0(t) + \varepsilon x_1(t) + O(\varepsilon^2), \qquad y(t) = y_0(t) + \varepsilon y_1(t) + O(\varepsilon^2).$$

Collecting the $O(1)$ terms in the expansions of (10.2), we obtain the leading order slow system,

$$0 = y_0 + x_0 - \tfrac{1}{3}x_0^3, \qquad \frac{dy_0}{dt} = -x_0. \tag{10.3}$$

The differential equation for $y_0(t)$ allows us to qualitatively describe the dynamics of the solution in terms of vertical motion in the phase plane as

$$y_0(t) \quad \text{is} \quad \begin{cases} \text{decreasing for } x_0 > 0, \\ \text{increasing for } x_0 < 0; \end{cases} \tag{10.4a}$$

hence the y_0-axis is the y-nullcline curve (see Sect. 1.5.1) for both the full problem and the slow system. The remaining equation in (10.3) restricts the slow-time solution to evolve only on the curve

$$y_0(x_0) = \tfrac{1}{3}x_0^3 - x_0 \equiv S(x_0), \tag{10.4b}$$

which is called the *slow manifold* (or "slow solution curve") and is shown in Fig. 10.1.

Restricting the solution to stay on the slow manifold while following the evolution described by (10.4a) suggests that starting from any $x_0 < 0$ will drive the solution upward to approach the local maximum of the curve at $y_0(-1) = 2/3$, while the dynamics in the right half plane, with $x_0 > 0$ will force solutions to evolve toward the local minimum at $y_0(1) = -2/3$. At those critical points (extrema), the vertical motion still follows (10.4a) even though there is no obvious place to go on the slow

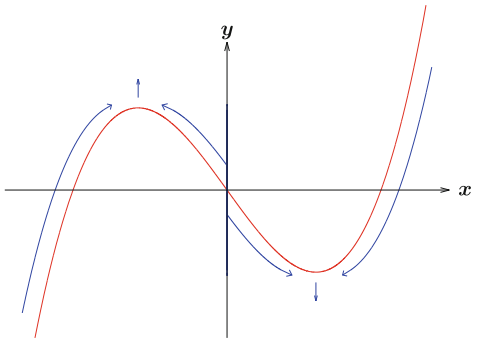


Fig. 10.1 *Arrows* indicating dynamics of the leading order solution on the slow manifold (10.4b) and at the two critical points of the *curve*

manifold. The solution therefore gets *pushed off* the slow manifold[2]; everything off the slow manifold in the phase plane is part of the *fast problem*, to which we now turn our attention.

Following the analysis in Chaps. 6 and 7, in order to find singular solutions, we rescale the original system in terms of new variables. To determine the fast variables, we substitute the general scalings

$$T = \frac{t - t_*}{\varepsilon^\alpha}, \qquad x = \varepsilon^\beta X(T), \qquad y = \varepsilon^\gamma Y(T), \tag{10.5}$$

into (10.2) where the scaling exponents α, β, γ and the time t_* when the fast dynamics occur are to be determined. For this problem, it can be shown that x, y are always $O(1)$, so we take $\beta = \gamma = 0$ and determine the value for $\alpha(\geq 0)$ from dominant balance in

$$\varepsilon^{1-\alpha}\frac{dX}{dT} = Y + X - \tfrac{1}{3}X^3, \qquad \varepsilon^{-\alpha}\frac{dY}{dT} = -X.$$

In applying the method of dominant balance to systems of equations, distinguished limits will typically yield a balance in one equation at a time, while the other equations yield sub-dominant contributions, where the rate of change will be $o(1)$. The distinguished limit $\alpha = 0$ balances the terms in the second equation and reproduces the original slow system we have already considered, (10.2). The choice $\alpha = 1$ balances all of the terms in the first equation,

$$\frac{dX}{dT} = Y + X - \tfrac{1}{3}X^3, \qquad \frac{dY}{dT} = -\varepsilon X, \tag{10.6}$$

where we have multiplied the second equation by ε^α.

Seeking the fast solutions as regular expansions,

$$X(T) = X_0(T) + \varepsilon X_1(T) + O(\varepsilon^2), \qquad Y(T) = Y_0(T) + \varepsilon Y_1(T) + O(\varepsilon^2),$$

yields the leading order fast system,

$$\frac{dX_0}{dT} = Y_0 + X_0 - \tfrac{1}{3}X_0^3, \qquad \frac{dY_0}{dT} = 0. \tag{10.7}$$

The second equation shows that the dynamics in the fast system (everywhere in the phase plane away from the slow manifold) must have $Y_0(T) \equiv$ constant, i.e. only horizontal motion is admissible. With this in mind, the first equation can be interpreted as describing motion on a phase-line for $X_0(T)$,

$$\frac{dX_0}{dT} = Y_0 - S(X_0), \tag{10.8}$$

[2] At which point, the solution is no longer described by the dynamics of the slow system.

where the value of Y_0 selects a "slice" through the slow-manifold analogous to (1.29) from Chap. 1. The equilibrium points of (10.8) are given by the solutions of $S(X_0) = Y_0$ with their stability being determined by the sign of $-S'(X_0)$, while all other values for X_0 give monotonic increasing or decreasing solutions.

In summary, for the slow system, we have identified the slow manifold and described how the dynamics on the slow manifold approach the critical points. Taking those as departure points into the fast system, the subsequent motion is given by the rapid evolution of x on horizontal lines that terminate at points of intersection with the slow curve. This qualitative description allows us to sketch the form of trajectories starting from any point in the (x, y) phase plane and identify the four-stage structure (and global stability) of the stable finite-amplitude *limit cycle* solution, see Figs. 10.2 and 10.3. The general features of the limit cycle (maximum, minimum values and period) can be obtained from these results (see Exercise 10.1). The fast-time dynamics can be asymptotically matched to outer solutions (slow solutions), and so the fast-dynamics act as "interior" boundary layers (in the terminology of Chap. 7).

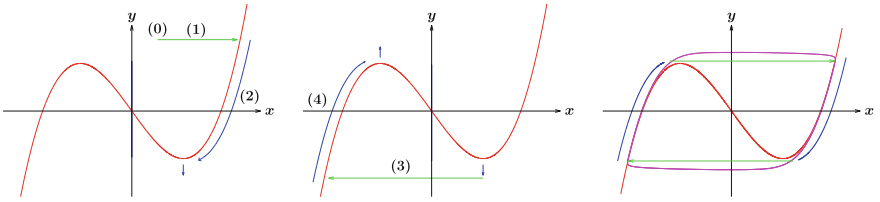


Fig. 10.2 Starting from general initial conditions not on the slow manifold, there will first be a rapid horizontal transition or "jump" *1* to a stable branch of the slow manifold followed by slow evolution down that branch *2* to the equilibrium point. Dropping off the minimum yields another fast jump *3* across to the other stable branch of the slow manifold, where slow dynamics drive the solution upwards *4* to the maximum and on to another fast jump back across to the previous stable branch of the slow manifold. The resulting (periodic) dynamics correspond to an attracting limit cycle

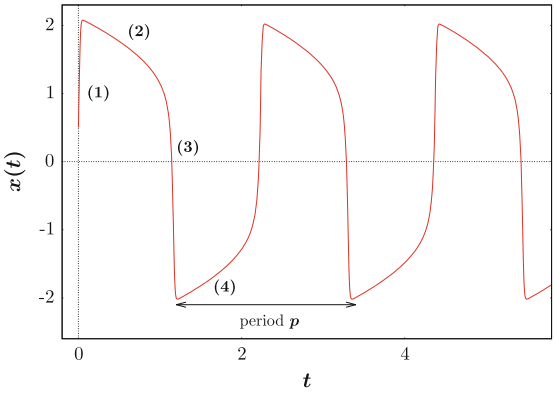


Fig. 10.3 Sketch of the time-profile illustrating the four stages in the van der Pol limit cycle

10.2 Complex Chemical Reactions: The Michaelis-Menten Model

The reactions described in Sect. 1.2 can be said to be "what you see is what you get" elementary reactions if the rate equations obtained from direct applications of the the law of mass action predicts the rates as would be observed in experimental studies of the systems. However, for general reactions, collisions of three or more molecules almost never happen exactly simultaneously (only single-molecule decay or binary collisions typically occur as elementary steps). The term *non-elementary reactions* describes the overall relation between reactants and products, but with the molecular collisions and mass action kinetics involved actually occurring in many "hidden" intermediate sub-stages. For example, a non-elementary version of (1.14a) could be

$$n\,A + m\,B \to (???) \to (???) \to (???) \to p\,C + q\,D.$$

The temporary products from intermediate reactions are called *complexes* and are often unstable compounds that exist only briefly.

To obtain the overall rate of creation of products, we need to analyse the full system with all intermediate stages expanded out as elementary reactions. Part of the very difficult work of chemists and bio-chemists is to determine all of the intermediate reactions. If all of the intermediate reactions are known, then we may have a very complicated system of rate ODEs. Our goal is to try to condense the system down to just the overall reaction

$$\text{reactants} \rightsquigarrow \text{products}$$

and to determine the effective rate equation for the products. The theory of dynamical systems can be applied to accomplish this for many chemical reaction systems. Given a complete system of reactions and initial conditions, we illustrate through an example how to obtain a simple model for the rate equations of the products.

A simple version of an enzyme-mediated chemical reaction process is given by the system,

$$S + E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C \xrightarrow{k_3} P + E, \tag{10.9}$$

where S is the "substrate" reactant and P is the concentration of the desired product. An enzyme (or catalyst) is a compound whose special property is that it allows for intermediate reaction steps that lead to the overall reaction, in this case

$$S \to P, \tag{10.10}$$

or allow it proceed more rapidly than without the enzyme; in some cases this direct transformation might not be possible at all without the enzyme.

Let E be the concentration of the enzyme, and C is the intermediate "complex" of SE bound together. The enzyme is typically a complicated and expensive compound,

but notably it is not actually consumed by (10.9). Consequently, the enzyme may be given in a small initial concentration E_0. The complex C is an unstable temporary state and none would be present before the reaction starts. We will assume that the production process begins with no product, and a large initial supply of reactant, S_0, all to be converted into P. Assuming the reaction rate constants k_1, k_2, k_3 are known finite values, our goal is to determine the rate of production for P and how it can be represented in simplest form.

Beginning by expanding (10.9) into its full set of elementary reactions,

$$S + E \xrightarrow{k_1} C \qquad S + E \xleftarrow{k_2} C \qquad C \xrightarrow{k_3} P + E, \tag{10.11}$$

the law of mass action yields the rate equations for consumption/production of each chemical species in the dimensional form

$$\begin{aligned} \frac{d\mathsf{P}}{d\mathsf{T}} &= \mathsf{k}_3\mathsf{C}, & \frac{d\mathsf{C}}{d\mathsf{T}} &= \mathsf{k}_1\mathsf{S}\mathsf{E} - \mathsf{k}_2\mathsf{C} - \mathsf{k}_3\mathsf{C}, \\ \frac{d\mathsf{S}}{d\mathsf{T}} &= -\mathsf{k}_1\mathsf{S}\mathsf{E} + \mathsf{k}_2\mathsf{C}, & \frac{d\mathsf{E}}{d\mathsf{T}} &= -\mathsf{k}_1\mathsf{S}\mathsf{E} + \mathsf{k}_2\mathsf{C} + \mathsf{k}_3\mathsf{C}, \end{aligned} \tag{10.12a}$$

subject to initial conditions at $t = 0$,

$$\mathsf{S}(0) = \mathsf{S}_0, \qquad \mathsf{E}(0) = \mathsf{E}_0, \qquad \mathsf{C}(0) = 0, \qquad \mathsf{P}(0) = 0. \tag{10.12b}$$

We now nondimensionalize system (10.12a, 10.12b). Given the one-to-one correspondence between S and P and between E and C as reactants/products in (10.11), we scale these pairs by the respective initial concentrations

$$\begin{aligned} \mathsf{S}(\mathsf{T}) = \mathsf{S}_0 s(t), \qquad \mathsf{P}(\mathsf{T}) = \mathsf{S}_0 p(t), \\ \mathsf{E}(\mathsf{T}) = \mathsf{E}_0 e(t), \qquad \mathsf{C}(\mathsf{T}) = \mathsf{E}_0 c(t), \qquad \mathsf{T} = \mathrm{Tt}. \end{aligned} \tag{10.13}$$

We will go on to select the timescale T later below. Through these scalings, the nondimensional concentrations are all normalised such that

$$0 \le \{s(t), p(t), e(t), c(t)\} \le 1.$$

Our previous intuition related to the transformation of enzyme to complex molecules, $E \rightleftharpoons C$, can be made concrete by noting that the total amount of the two substances is conserved for all times, following directly from (10.12a),

$$\frac{d}{d\mathsf{T}}(\mathsf{C} + \mathsf{E}) = 0 \implies \mathsf{C}(\mathsf{T}) + \mathsf{E}(\mathsf{T}) = \mathsf{E}_0,$$

so that $c + e = 1$. This allows us to use $e = 1 - c$ to eliminate the enzyme from the system, leaving the problem in terms of s, c, p.

We take the initial amount of substrate to be finite, $\mathsf{S}_0 = O(1)$, and assume that we begin with a relatively small amount of enzyme, $\mathsf{E}_0 \ll \mathsf{S}_0$. We now select the timescale T in order to make the consumption rate $ds/dt = O(1)$. This choice yields $\mathrm{T} = 1/(\mathsf{k}_1\mathsf{E}_0)$. The resulting scaled system is therefore given by

$$\begin{aligned} \frac{ds}{dt} &= -s(1-c) + \lambda c, & s(0) &= 1, \\ \varepsilon\frac{dc}{dt} &= s(1-c) - \mu c, & c(0) &= 0, \\ \frac{dp}{dt} &= (\mu - \lambda)c, & p(0) &= 0, \end{aligned} \tag{10.14}$$

where we have relabelled dimensionless combinations of parameters as

$$\varepsilon = \frac{\mathsf{E}_0}{\mathsf{S}_0} \ll 1 \qquad \lambda = \frac{\mathsf{k}_2}{\mathsf{k}_1\mathsf{S}_0} = O(1) \qquad \mu = \frac{\mathsf{k}_2 + \mathsf{k}_3}{\mathsf{k}_1\mathsf{S}_0} = O(1). \tag{10.15}$$

We will consider the limit of a very small amount of enzyme in the system ($\varepsilon \to 0$) for our perturbation analysis. Noting that the equations for s, c are independent of p, we first solve for the s, c system and then use the results to determine p from the final equation in (10.14).

Since (10.14) is a singularly perturbed system for $\varepsilon \to 0$, we expect a separation between fast and slow time scales. Beginning with the slow system, we consider regular expansions for the solutions, $s = s_0(t) + O(\varepsilon)$, $c = c_0(t) + O(\varepsilon)$. Substituting these into (10.14) yields the leading order slow system,

$$\frac{ds_0}{dt} = -s_0(1-c_0) + \lambda c_0, \qquad 0 = s_0(1-c_0) - \mu c_0, \tag{10.16}$$

with the second equation giving an algebraic relationship between the complex and substrate concentrations,

$$s_0(c_0) = \frac{\mu c_0}{1-c_0}; \tag{10.17}$$

this is the slow manifold for this problem.

We note that our initial conditions, $c(t=0) = 0$, $s(t=0) = 1$ do not lie on the slow manifold and hence there must be a brief *initial layer* at $t_* = 0$, governed by the dynamics of the fast system, that describes the transition from the initial conditions to the slow manifold. To determine the form of the fast system, we rescale the variables as

$$s = S(T), \qquad c = C(T), \qquad T = \frac{t}{\varepsilon^\alpha},$$

yielding

$$\varepsilon^{-\alpha}\frac{dS}{dT} = -S(1-C) + \lambda C, \qquad \varepsilon^{1-\alpha}\frac{dC}{dT} = S(1-C) - \mu C.$$

The distinguished limit for the fast system is found to occur when $\alpha = 1$ and the leading order fast system is then given by

$$\frac{dS_0}{dT} = 0, \qquad \frac{dC_0}{dT} = S_0(1 - C_0) - \mu C_0, \tag{10.18}$$

with initial conditions $S_0(0) = 1$ and $C_0(0) = 0$. Consequently there is no change to the substrate concentration in the initial layer at leading order, and it remains at its initial value, $S_0(T) \equiv 1$. For completeness, we could determine $C_0(T)$ from this system, but bearing in mind that our final goal is to determine $p(t)$, we only actually need to find $s(t)$ as we can replace $c(t)$ in terms of $s(t)$ (at leading order) in (10.14) using the slow manifold (10.17). Asymptotic matching of the fast and slow solutions for s yields the initial conditions for the slow solution,

$$\lim_{T\to\infty} S_0(T) = 1 = \lim_{t\to 0} s_0(t).$$

Hence, while the initial layer has a dramatic effect on the concentration of the complex, it essentially leaves the initial condition on s from (10.14) unchanged. The initial layer is effectively a boundary layer with respect to time for the initial value problem for c in (10.14). Figure 10.4 shows a comparison of the numerical solution of (10.14) against the leading order fast/slow dynamics (the slow manifold being described by (10.17)).

Substituting (10.17) into (10.14), we get the slow system,

$$\frac{ds_0}{dt} = -\frac{\mu - \lambda}{\mu + s_0} s_0, \qquad \frac{dp_0}{dt} = \frac{\mu - \lambda}{\mu + s_0} s_0, \tag{10.19}$$

and we can finally confirm our expectation that the total of the substrate and product is conserved, $d(s_0 + p_0)/dt = 0$. Using the initial conditions, we have effectively reduced the original problem to solving a single first order ODE for $s_0(t)$, with $p_0(t)$ then being given by $p_0(t) = 1 - s_0(t)$. This effective nonlinear rate law is called the *Michaelis-Menten* law and is used extensively in the modelling of biochemical systems.

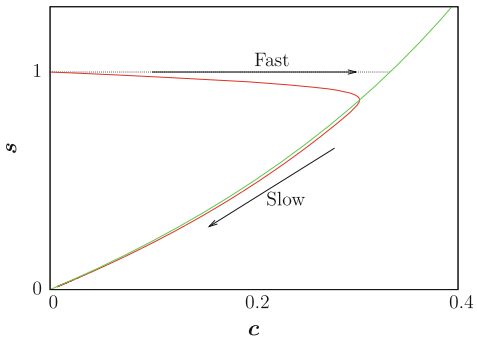


Fig. 10.4 The cs phase plane showing the slow manifold *curve* (10.17) and a numerical solution of the full system (10.14) for $\varepsilon = 1/5$ starting from initial condition $(c, s) = (0, 1)$

10.3 Further Directions

More extensive presentations of dynamical systems are given in many texts, both from more theoretical [43, 70], and more applied standpoints [54, 94]. Many books illustrate in detail the use of perturbation methods for studying dynamical systems [47, 54, 56, 73, 102], with the van der Pol oscillator being a classic example. Modelling and applications of chemical kinetic systems are presented more thoroughly in [37, 49, 57, 74]. Some alternative limiting cases of the enzyme kinetics system are explored in [87, 90].

10.4 Exercises

10.1 Consider the van der Pol equation for $x(t)$ with $0 < \varepsilon \ll 1$,

$$\varepsilon \frac{d^2x}{dt^2} + (3x^2 - 6x - 9)\frac{dx}{dt} + 4x = 0.$$

(a) Determine $f(x)$ so that this equation can be written as a Liénard phase plane system in the form

$$\varepsilon \frac{dx}{dt} = f(x) + 4y, \qquad \frac{dy}{dt} = -x.$$

(b) For fixed $\varepsilon > 0$, find the equilibrium point(s) in the phase plane, find their eigenvalues, and classify their linear stability.

(c) Use the expansions $x(t) = x_0(t) + \varepsilon x_1(t) + O(\varepsilon^2)$, $y(t) = y_0(t) + \varepsilon y_1(t) + O(\varepsilon^2)$, to determine the equations for the leading order slow solution. Sketch the slow manifold, indicate the direction of motion on each part, and identify the two attracting points on the curve.

(d) Use the expansions $x(t) = X_0(T) + \varepsilon X_1(T) + O(\varepsilon^2)$, $y(t) = Y_0(T) + \varepsilon Y_1(T) + O(\varepsilon^2)$ with $T = t/\varepsilon$ to obtain the equations for the leading order fast solution.

(e) Use the phase plane to determine the maximum and minimum values of $x(t)$ during an oscillation, see Fig. 10.2. Sketch $x(t)$ as a function of time, see Fig. 10.3.

(f) Using the time required for the slow motions in (c) (neglecting the short times for the fast solutions (d)), determine the leading order estimate for the period P of oscillation of the limit cycle.
Hint: Find the time spent moving along each of the slow curves by obtaining an equation $dx_0/dt = g(x_0)$ from (c) and then separate variables to write

$$\frac{dx_0}{dt} = g(x_0) \implies P = \int_{t_{\text{start}}}^{t_{\text{end}}} dt = \int_{x_{\text{start}}}^{x_{\text{end}}} \frac{dx}{g(x)}$$

and finally integrate over the ranges in x that are appropriate to each of the two slow segments and adding together those two times.

10.2 We now consider different limits for a dynamical system describing a chemical reaction problem in terms of three variables, see (4.50). Consider the limit $\varepsilon \to 0$ for each of the following cases,

(a) For the system

$$\begin{aligned} \frac{dx}{dt} &= 2 - y, & x(0) &= 1, \\ \frac{dy}{dt} &= x - z, & y(0) &= 3, \\ \varepsilon \frac{dz}{dt} &= y - y^2 + \tfrac{1}{3}y^3 - z, & z(0) &= 0. \end{aligned} \tag{10.20}$$

Identify the surface $z = S(x, y)$ that defines the slow manifold. Find the equilibrium point of the leading order slow phase plane system and show that it is asymptotically stable for $t \to \infty$. Also determine the form of the initial layer that describes the transition from the initial conditions to the slow manifold.

(b) For the system

$$\begin{aligned} \frac{dx}{dt} &= 2 - y, & x(0) &= 0, \\ \varepsilon \frac{dy}{dt} &= x - z, & y(0) &= 3, \\ \frac{dz}{dt} &= y - y^2 + \tfrac{1}{3}y^3 - z, & z(0) &= 1. \end{aligned} \tag{10.21}$$

Show that the slow manifold reduces to a curve that could be written in parametric form as $x = x(z), y = y(z), z = z$. Determine the asymptotic solution for $t \to \infty$. Also determine the form of the initial layer that describes the transition from the initial conditions to the slow manifold.

10.3 Consider the problem of forming a "tri-mer" (a three segment polymer molecule) from three mono-mer molecules,

$$3A \to A_3.$$

This is an example of polymerisation. It is a non-elementary reaction and takes place via intermediate stages. Call the tri-mer "C" and the di-mer (A_2) "B". Suppose that the full reaction mechanism is given by

$$A + A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} B \qquad A + B \xrightarrow{k_3} C$$

(a) Write the dimensional rate equations.
(b) Nondimensionalize using the scalings

$$\mathsf{A}(\mathsf{T}) = \mathsf{A}_0 a(t) \qquad \mathsf{B}(\mathsf{T}) = \mathsf{B}_0 b(t) \qquad \mathsf{C}(\mathsf{T}) = \mathsf{A}_0 c(t) \qquad \mathsf{T} = \mathrm{T}t$$

where A_0 is the initial concentration.
(c) Let $\mathrm{T} = 1/(\mathsf{k}_1\mathsf{A}_0)$ and assume that $\varepsilon = \mathsf{B}_0/\mathsf{A}_0 \to 0$.
Identify the other dimensionless parameters (call them Π_1, Π_2) and state what asymptotic relations on the rate constants k_n must hold if we assume that only the $\varepsilon db/dt$ terms vanishes from the leading order slow system.
Hint: Normalise the coefficient of the $+a^2$ term.
(d) Find the leading order slow manifold and write the dimensional equations for the long-term rate of production, and hence find $\mathsf{G}(\mathsf{A})$ and $\mathsf{F}(\mathsf{A})$ in

$$\frac{d\mathsf{A}}{d\mathsf{T}} = -\mathsf{G}(\mathsf{A}), \qquad \frac{d\mathsf{C}}{d\mathsf{T}} = \mathsf{F}(\mathsf{A}).$$

10.4 Consider the system of chemical reactions

$$A + X \xrightarrow{k_1} Y \qquad A + Y \xrightarrow{k_2} 2X \qquad A \xrightarrow{k_3} Y \qquad 2Y \xrightarrow{k_4} P$$

where the concentration of A is kept constant and k_1, k_2, k_3, k_4 are given.

(a) Write the rate equations for $x(t)$ and $y(t)$.
(b) Nondimensionalize using $\mathsf{T} = \mathrm{T}t$, $\mathsf{X} = \mathrm{X}x$, $\mathsf{Y} = \mathrm{Y}y$. Let $\mathrm{T} = 1/(\mathsf{k}_1\mathsf{A})$. Determine X and Y so that: (1) all of the terms in the dx/dt equation and (2) the y^2 term in the dy/dt equation are normalised.
(c) Determine the remaining independent dimensionless parameters (call them α, β) and write the nondimensionalized equations.
(d) In terms of α, β, determine the concentrations x, y for the positive equilibrium solution.

10.5 In chemistry, a widely-used short-cut avoiding the full scaling and slow/fast perturbation analysis is to jump to the leading order slow equations using the assumption that the rate of production of all intermediates equilibrate (i.e. reduce to steady algebraic relations, like $\varepsilon \frac{dc}{dt} = 0$); this is called the *Quasi-Steady-State Assumption* (QSSA).
Use the QSSA approach to consider the overall reaction for the formation of hydrogen-bromide: $\mathrm{H}_2 + \mathrm{Br}_2 \to 2\mathrm{HBr}$. The reaction takes place through several steps:

$$\mathrm{Br}_2 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 2\mathrm{Br} \qquad \mathrm{Br} + \mathrm{H}_2 \underset{k_4}{\overset{k_3}{\rightleftharpoons}} \mathrm{H} + \mathrm{HBr} \qquad \mathrm{H} + \mathrm{Br}_2 \xrightarrow{k_5} \mathrm{HBr} + \mathrm{Br}$$

Consider the atomic forms H and Br to be unstable intermediates (similar to complexes being unstable intermediates). Apply the QSSA to obtain the dimensional

rate law for the production of HBr (written here using the chemistry-notation of [X] being the concentration of chemical X) [6],

$$\frac{d[\mathrm{HBr}]}{d\mathsf{T}} = \frac{\alpha[\mathrm{H_2}][\mathrm{Br_2}]^{3/2}}{[\mathrm{Br_2}] + \beta[\mathrm{HBr}]}$$

Find α, β. Hint: Write $A = [\mathrm{H_2}], B = [\mathrm{Br_2}], C = [\mathrm{Br}], D = [\mathrm{H}], P = [\mathrm{HBr}]$ for doing your algebra.

Chapter 11
Reduced Models for PDE Problems

In previous chapters, we have considered methods for solving problems in partial differential equations: the method of characteristics in Chap. 2, similarity solutions in Chap. 5, and Fourier series in Chap. 8. Now we pursue another approach that can be applied to problems where we want to obtain certain properties of the solution, but do not need an explicit representation of the entire solution. We will see that in some cases, this leads to substantially shorter calculations.

The approach of *reduced models* is to reformulate a problem into a simpler form that preserves the properties of interest, but decouples them from the calculation of other details of the solution's behaviour. We will describe the *method of moments*, which produces reduced models that can describe the evolution of some quantities for the solution averaged over the domain.[1]

11.1 The Method of Moments

For a solution $\rho(x,t)$ of a PDE problem on a domain D, we can define the *moment integrals*,

$$M_n(t) \equiv \int_D x^n \rho(x,t)\,dx \qquad n = 0, 1, 2, \ldots \tag{11.1}$$

If ρ represents a density, then the $n = 0$ integral corresponds to the total mass. For some problems, we can determine a simple set of equations governing the evolution of the moment integrals (without explicitly finding the solution $\rho(x,t)$) and then use those results to infer information about the behaviour of the solution to the problem.

[1] We have already encountered *reduced dimension models* in Chap. 8, where we developed a methodology for solving PDEs on slender two-dimensional domains by replacing them with ODE problems for an asymptotic solution applying on most of the domain.

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9_11

We illustrate this approach for the heat equation on $-\infty < x < \infty$

$$\frac{\partial \rho}{\partial t} = \frac{\partial^2 \rho}{\partial x^2}, \tag{11.2a}$$

with far-field boundary conditions $u \to 0$ as $|x| \to \infty$ and initial condition

$$\rho(x, t = 0) = f(x), \tag{11.2b}$$

where $f \to 0$ rapidly as $|x| \to \infty$.

Consider the zeroth moment (mass) integral,

$$M_0(t) = \int_{-\infty}^{\infty} \rho(x, t)\, dx.$$

In order to obtain an equation to describe the evolution of $M_0(t)$, we calculate its rate of change with respect to time

$$\frac{dM_0}{dt} = \frac{d}{dt} \int_{-\infty}^{\infty} \rho\, dx = \int_{-\infty}^{\infty} \frac{\partial \rho}{\partial t}\, dx,$$

where we have interchanged the order of differentiation (with respect to time) and integration (over space). Using the PDE (11.2a) to replace the time derivative with the second order spatial derivative and then integrating yields

$$\frac{dM_0}{dt} = \int_{-\infty}^{\infty} \frac{\partial^2 \rho}{\partial x^2}\, dx = \frac{\partial \rho}{\partial x}\bigg|_{-\infty}^{\infty} = 0 - 0 = 0,$$

where we have used that $\rho_x \to 0$ as $|x| \to \infty$ for smooth $\rho \to$ constant (specifically $\rho \to 0$) in evaluating the boundary terms. Consequently, we have

$$\frac{dM_0}{dt} = 0,$$

so that the total mass is constant in the problem and its value for problem (11.2a, 11.2b) is determined by its initial value at $t = 0$,

$$M_0(0) = \int_{-\infty}^{\infty} f(x)\, dx. \tag{11.3}$$

We can proceed similarly for the first moment, $M_1(t) = \int_{-\infty}^{\infty} x\rho\, dx$, for which we find

$$\frac{dM_1}{dt} = \frac{d}{dt}\int_{-\infty}^{\infty} x\rho\, dx = \int_{-\infty}^{\infty} x\frac{\partial \rho}{\partial t}\, dx = \int_{-\infty}^{\infty} x\frac{\partial^2 \rho}{\partial x^2}\, dx.$$

Making use of integration by parts and employing further assumptions about the decay of the solution as $|x| \to \infty$ being sufficiently rapid to eliminate boundary terms from integration by parts, we obtain

$$\frac{dM_1}{dt} = x\frac{\partial \rho}{\partial x}\bigg|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{\partial \rho}{\partial x}\, dx = 0.$$

Therefore, M_1 is also a constant, also set by the initial conditions as

$$M_1(0) = \int_{-\infty}^{\infty} xf(x)\, dx. \tag{11.4}$$

Having these moments allow us to evaluate the centre of mass (the average value of position weighted by the density), defined by the ratio of the first moment to the mass,

$$\bar{x}(t) = \frac{M_1}{M_0} = \frac{\int xf\, dx}{\int f\, dx}. \tag{11.5}$$

Investigating the evolution of $\bar{x}$ can provide an understanding of whether the solution is generally moving in some direction or remaining in a fixed location; in this case, $\bar{x}$ remains fixed at its initial position as both M_0 and M_1 are constants.

Proceeding to the second moment, we find

$$\begin{aligned}\frac{dM_2}{dt} &= \int_{-\infty}^{\infty} x^2\frac{\partial \rho}{\partial t}\, dx = \int_{-\infty}^{\infty} x^2\frac{\partial^2 \rho}{\partial x^2}\, dx \\ &= x^2\frac{\partial \rho}{\partial x}\bigg|_{-\infty}^{\infty} - \int 2x\frac{\partial \rho}{\partial x}\, dx \\ &= -2\left(x\rho\bigg|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \rho\, dx\right),\end{aligned}$$

where the boundary conditions have been used to successively eliminate the boundary terms from the two applications of integration by parts. Consequently, the rate of change of the second moment is given in terms of the mass M_0,

$$\frac{dM_2}{dt} = 2M_0. \tag{11.6}$$

For the heat equation, we have already shown that the mass is constant and so the second moment increases at a constant rate. We can therefore express M_2 in terms of the initial conditions as

$$M_2(t) = 2t \int_{-\infty}^{\infty} f\,dx + \int_{-\infty}^{\infty} x^2 f\,dx. \tag{11.7}$$

The *variance* is defined as the second moment of ρ with respect to the centre of mass and yields a fundamental measure of the spreading of the solution

$$\begin{aligned} V(t) &= \int_{-\infty}^{\infty} (x-\bar{x})^2 \rho\,dx \\ &= \int_{-\infty}^{\infty} x^2\rho\,dx - 2\bar{x}\int_{-\infty}^{\infty} x\rho\,dx + \bar{x}^2\int_{-\infty}^{\infty}\rho\,dx \\ &= M_2 - 2\bar{x}M_1 + \bar{x}^2 M_0 \\ &= M_2 - 2\frac{M_1^2}{M_0} + \frac{M_1^2}{M_0^2}M_0 \\ &= M_2 - \frac{M_1^2}{M_0}. \end{aligned}$$

The variance is also used to define the standard deviation in terms of its square-root, $\sigma = \sqrt{|V|}$, yielding

$$\sigma^2(t) = \left| M_2 - \frac{M_1^2}{M_0} \right|. \tag{11.8}$$

While the terminology used above comes from probability and statistics, the definitions are direct analogues of the formulas for the moment of inertia ($V \to I$) and radius of gyration ($\sigma \to R$) used in mechanics for describing the motion of objects [40, 67].

A somewhat more direct approach for constructing moment equations is to directly the integrate the product of the weight function with the PDE over the domain,

$$\int_D x^n (\text{PDE})\,dx.$$

Consider for example the following problem for the wave equation on the finite domain, $0 \leq x \leq 1$,

$$\frac{\partial^2 \rho}{\partial t^2} = \frac{\partial^2 \rho}{\partial x^2} \tag{11.9a}$$

subject to boundary conditions

$$\left.\frac{\partial \rho}{\partial x}\right|_{x=0} = e^t, \qquad \left.\frac{\partial \rho}{\partial x}\right|_{x=1} = 12t^2, \tag{11.9b}$$

and initial conditions

$$\rho(x,0) = 1 \qquad \left.\frac{\partial \rho}{\partial t}\right|_{t=0} = \cos(3\pi x). \tag{11.9c}$$

Integrating (11.9a) over the domain and the applying boundary conditions yields an ODE for the zeroth moment,

$$\int_0^1 (\rho_{tt} = \rho_{xx})\, dx \quad \Longrightarrow \quad \frac{d^2 M_0}{dt^2} = \left.\frac{\partial \rho}{\partial x}\right|_{x=0}^{x=1} \quad \Longrightarrow \quad \frac{d^2 M_0}{dt^2} = 12t^2 - e^t.$$

Integrating the initial conditions for ρ generates initial conditions for M_0, $M_0(0) = 1$ and $M_0'(0) = 0$ and consequently yields $M_0(t) = t^4 + t + 2 - e^t$.

In general, the method of moments is considered to be successful when it produces a finite, closed set of ODEs that can used to describe selected solution properties. If the moment equations depend on an indefinite number of further moments or other properties of ρ, then the system is not closed. In such cases, the system may be approximated through the use of problem-specific *closure relations*, as used in the modelling of turbulent fluid flows. While this method has been described here as an analytical approach, similar analysis is used to reduce PDE problems to simpler systems of equations that can be evaluated computationally in numerical methods such as *Galerkin projection methods*.

We now go on to consider two classic applied mathematics problems that have different relations to the method of moments.

11.2 Turing Instability and Pattern Formation

In this section, we present an example of how to analyse the development of patterns in PDE models, as occur in important systems in mathematical biology and other applications. The problem considered will also show that predictions from reduced models can sometimes be misleading and must be considered with caution.

We have observed that solutions of the heat equation will generically exhibit diffusive spreading (as implied by $\sigma(t)$ increasing with time from (11.8) and similarity solutions having $x = O(\sqrt{t})$ from Sect. 5.5). Consider the problem for the heat equation on a finite domain with no-flux boundary conditions,

$$\frac{\partial \rho}{\partial t} = \frac{\partial^2 \rho}{\partial x^2}, \qquad 0 \le x \le \pi,$$

$$\left.\frac{\partial \rho}{\partial x}\right|_{x=0} = 0, \qquad \left.\frac{\partial \rho}{\partial x}\right|_{x=\pi} = 0,$$

with a given initial condition

$$\rho(x, 0) = f(x).$$

The mass of the solution will be conserved for all times. Further, it can be shown that for $t \to \infty$, the solution will approach the average value set by the initial condition

$$\rho(x, t \to \infty) \to \bar{f} \qquad \text{where} \quad \bar{f} = \frac{1}{\pi} \int_0^\pi f(x)\, dx.$$

Similarly, adding diffusive effects to a transport equation can generally be expected to cause the solution to both spread and smooth out. Likewise, if diffusion is incorporated into a model for chemical reaction kinetics to yield a partial differential equation, as in

$$\frac{dc}{dt} = -c \qquad \Longrightarrow \qquad \frac{\partial c}{\partial t} = \frac{\partial^2 c}{\partial x^2} - c,$$

then it can be shown that the spatial variation in the initial condition $c(x, 0)$ will eventually level-out as the solution approaches the solution of the original ODE, $c(x, t) \to c(t)$ (see Exercise 11.3).

However, the work of Alan Turing (1912–1954) showed that this intuition is not always a good guide for more complicated systems. In the context of developmental biology [100], he showed that the addition of diffusive effects to a system of reactions could enhance spatial structure rather than suppressing it.

To illustrate this effect, let us consider a simple system involving two chemicals, P and Q, that interact and are also supplied into or drained out of the system according to

$$P + Q \to 2P \qquad P \to (\text{drain}) \qquad Q \to 2Q \qquad (\text{source}) \to Q.$$

These two substances may diffuse at different rates, so we will describe their spatial evolution using diffusion terms with different coefficients. Hence, consider the nondimensionalized governing equations,

$$\frac{\partial p}{\partial t} = 2pq - 4 + \frac{\partial^2 p}{\partial x^2}, \qquad \frac{\partial q}{\partial t} = -2pq + q + 3 + D\frac{\partial^2 q}{\partial x^2} \tag{11.10a}$$

where $D \geq 1$ is the ratio of diffusion coefficients giving the relative spreading rate of Q compared to P. Let the domain be $0 \leq x \leq \pi$, with no-flux boundary conditions,

$$\frac{\partial p}{\partial x} = 0 \quad \text{and} \quad \frac{\partial q}{\partial x} = 0 \qquad \text{at } x = 0 \text{ and } x = \pi \tag{11.10b}$$

and initial conditions at $t = 0$,

$$p(x, 0) = p_0(x), \qquad q(x, 0) = q_0(x). \tag{11.10c}$$

This is a simple example of a *reaction-diffusion system*. Research on similar systems has shown that they can produce surprisingly complicated results, so the general approach to studying reaction-diffusion problems is to start with the simplest possible solutions and build up an understanding from there.

Seeking time-independent steady states reduces the PDEs for $p(x, t), q(x, t)$ to two coupled nonlinear ODEs for $\bar{p}(x), \bar{q}(x)$, which can still be difficult to solve. Hence, we further constrain our search and obtain spatially uniform (constant) solutions, p_*, q_* from the algebraic relations

$$0 = 2p_*q_* - 4, \qquad 0 = -2p_*q_* + q_* + 3, \tag{11.11}$$

which yield the equilibrium solution $p_* = 2, q_* = 1$. To explore the stability of this state to small perturbations, let

$$p = p_* + \varepsilon\tilde{p}(x, t), \qquad q = q_* + \varepsilon\tilde{q}(x, t) \tag{11.12}$$

with $\varepsilon \to 0$. Treating these expression as perturbation expansions, we substitute into (11.10a). At leading order we recover the steady-state equations (11.11), while at $O(\varepsilon)$ we get the linearised version of (11.10a),

$$\frac{\partial \tilde{p}}{\partial t} = 2\tilde{p} + 4\tilde{q} + \frac{\partial^2 \tilde{p}}{\partial x^2}, \qquad \frac{\partial \tilde{q}}{\partial t} = -2\tilde{p} - 3\tilde{q} + D\frac{\partial^2 \tilde{q}}{\partial x^2}. \tag{11.13}$$

Applying the method of moments to this system, we can determine the evolution of the mean values of $\tilde{p}$ and $\tilde{q}$ from the zeroth moment integrals,

$$P_0(t) = \frac{1}{\pi}\int_0^\pi \tilde{p}(x, t)\, dx, \qquad Q_0(t) = \frac{1}{\pi}\int_0^\pi \tilde{q}(x, t)\, dx.$$

Integrating (11.13) over the domain and applying the boundary conditions yields the coupled ODEs,

$$\frac{dP_0}{dt} = 2P_0 + 4Q_0, \qquad \frac{dQ_0}{dt} = -2P_0 - 3Q_0. \tag{11.14}$$

Note that (11.14) corresponds to (11.13) restricted to spatially uniform solutions. A phase plane analysis of this system shows that $(P_0, Q_0) = (0, 0)$ is the only equilibrium point and has eigenvalues $\lambda = (-1 \pm i\sqrt{7})/2$, so it is a stable spiral point. Hence all solutions of (11.14) will converge to $(P_0, Q_0) = (0, 0)$ as $t \to \infty$. While this correctly predicts the evolution of the means of the perturbations $\tilde{p}$ and $\tilde{q}$, further analysis is needed to describe their spatial structure.

The process of reducing a nonlinear problem, like (11.10a), to a linear problem, (11.13), by focusing on the evolution of solutions starting near an equilibrium state is a classic example of linear stability analysis. We applied this approach to ODEs in Chap. 1; this is also one of the most general approaches to exploring the behaviour of nonlinear PDEs [26].

To complete the analysis, we make use of the linearity of system (11.13) to obtain solutions as linear superpositions of eigenmodes, $\tilde{p} = \sum_k \tilde{p}_k(x, t)$, and apply separation of variables to express the eigenmodes as products $\tilde{p}_k(x, t) = f_k(x)g_k(t)$ [44]. Requiring each eigenmode to satisfy the boundary conditions (11.10b) yields

$$\begin{pmatrix} p_k(x, t) \\ q_k(x, t) \end{pmatrix} = \begin{pmatrix} a_k \\ b_k \end{pmatrix} \cos(kx)e^{\lambda_k t}, \tag{11.15}$$

where a_k, b_k, λ_k are constants depending on the wavenumber $k = 0, 1, 2, \ldots$. Substituting each independent mode into (11.13) and eliminating common factors, we arrive at the matrix eigenvalue problem

$$\lambda_k \begin{pmatrix} a_k \\ b_k \end{pmatrix} = \begin{pmatrix} 2 - k^2 & 4 \\ -2 & -3 - k^2 D \end{pmatrix} \begin{pmatrix} a_k \\ b_k \end{pmatrix}. \tag{11.16}$$

For $k = 0$, this is the eigenvalue problem that would be obtained from the stability analysis of the ODE system (11.14). More generally, for $k \geq 0$, we can determine the stability of the eigenmodes from solving the characteristic polynomial for λ_k

$$(\lambda_k + k^2 - 2)(\lambda_k + k^2 D + 3) + 8 = 0. \tag{11.17}$$

The relation between the spatial wavenumber and the real part of the dominant eigenvalue, $\sigma(k) = \mathrm{Re}(\lambda_k^+)$ is often called the *dispersion relation* and concisely conveys the PDE stability results with respect to all admissible types of perturbations, as represented by the eigenmodes with different values of k.

For the case where Q diffuses much more rapidly than P ($D \gg 1$) it can be shown that $\mathrm{Re}(\lambda_k^+)$ can become positive and hence some eigenmodes will grow in amplitude; this is known as a *Turing instability*. The growth rates for system (11.13) when $D = 100$ are plotted in Fig. 11.1, where it can be seen that $k = 1$ is selected as the dominant growing mode (the only integer-valued k with a positive growth rate). Consequently, this mode will eventually dominate the dynamics of the system and produce a pattern

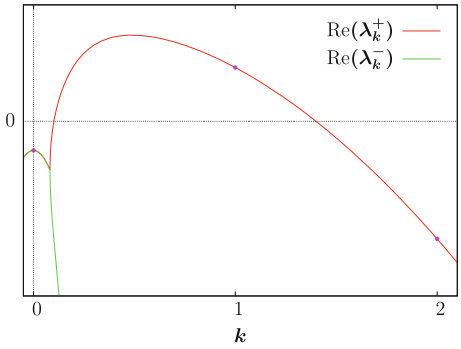


Fig. 11.1 The dispersion relation showing growth rates of eigenmodes (11.15) with $k = 0, 1, 2, \ldots$ of the linearised system (11.13) with $D = 100$

resembling $\cos(x)$ unless the initial conditions (11.10c) have no contribution from the $k = 1$ unstable mode, i.e. $\int_0^\pi (p_0, q_0)\cos(x)\,dx = (0, 0)$. We note that the influence of the parameter D (giving the relative rates of diffusion) dramatically changes the stability of the equilibrium solution for different wavenumbers: a stable spiral for $k = 0$, a saddle point for $k = 1$, and stable node for $k = 2, 3, \ldots$.

For different choices of system parameters, there may be a band of unstable wavenumbers, $k \in \{1, 2, \ldots, k_c\}$ with $\sigma(k) > 0$, that will all grow until nonlinear coupling effects take over to determine the further dynamics of the solution. Systems having $\sigma(k) \geq 0$ below some critical wavenumber, $0 \leq k \leq k_c$, are often called *long-wave unstable* since they involve spatial variations involving only the longest wavelength eigenmodes. The Turing instability has been used in many studies as a model for biological pattern formation, such as the spots and stripes on animal coats [26, 42, 74].

We note that in attempting to use the method of moments for this problem, we considered the zeroth moments of the perturbations in the linearised system (11.13) rather than in the original nonlinear system. There would have been a difficulty in integrating (11.10a) directly, as it might not be immediately clear how to analyse the integral of the nonlinear term, $\int_0^\pi pq\,dx$. In the next section, we will see how this type of issue can be dealt with in the context of a different system.

11.3 Taylor Dispersion and Enhanced Diffusion

We now describe a classic model from fluid dynamics that illustrates the effectiveness of reduction via the moment-type approach.

If a substance (called the *solute*) is released into a channel containing a flowing stream of fluid, the substance will spread out as it is carried downstream by the flow. Part of this behaviour is due to standard "molecular diffusion" effects, but spatial variations in the flow velocity also contribute since the different transport speeds will broaden the area of distribution of the solute.

This problem was studied by G.I. Taylor (1886–1975), who explained how the full transport problem for the solute could be reduced to a one-dimensional model. Taylor's model showed that convective effects can increase the effect of diffusion; this has come to be known as *Taylor dispersion* and has been applied to describe the spread of pollutants in rivers, drugs in blood flow as well as in many other settings. Taylor's original paper [98] was presented in his unique and very physically intuitive style; it is deceptively short and challenging to follow. Here we present a derivation of Taylor's result that is very similar to the approach given by Leal [61, Sect. 3H] (also see [20]).

Consider a dimensional transport problem for the concentration of a solute $\mathsf{C}(\mathsf{X}, \mathsf{Y}, \mathsf{T})$, given by the advection-diffusion equation

$$\frac{\partial \mathsf{C}}{\partial \mathsf{T}} + \nabla \cdot (\mathsf{C}\mathbf{U}) = \mathsf{D}\nabla^2 \mathsf{C}, \tag{11.18a}$$

where $\mathbf{U}(\mathsf{X}, \mathsf{Y}, \mathsf{T})$ is the fluid velocity field and D is the constant of molecular diffusion. We choose the domain to be a two-dimensional long slender channel, $-\infty < X < \infty$, $-\mathrm{H} \le \mathsf{Y} \le \mathrm{H}$, and let the velocity field take the classic *Poiseuille flow* parabolic profile

$$\mathbf{U} = \mathrm{U}_0 \left(1 - \frac{\mathsf{Y}^2}{\mathrm{H}^2}\right) \hat{\mathbf{I}} \qquad -\mathrm{H} \le \mathsf{Y} \le \mathrm{H}, \tag{11.18b}$$

where U_0 is the maximum speed of the flow (see Fig. 11.2 for a schematic of the problem). We assume that there is no flux of the solute out of the sides of the domain, $\hat{\mathbf{n}} \cdot \nabla C = 0$, which for the uniform width channel, simplifies to

$$\left.\frac{\partial \mathsf{C}}{\partial \mathsf{Y}}\right|_{\mathsf{Y}=\pm\mathrm{H}} = 0. \tag{11.18c}$$

While in most common situations, diffusion occurs at the same rate in every direction (so-called *isotropic* behaviour), to help distinguish different effects, we label the diffusion coefficients in the direction of and transverse to the imposed flow as D_x

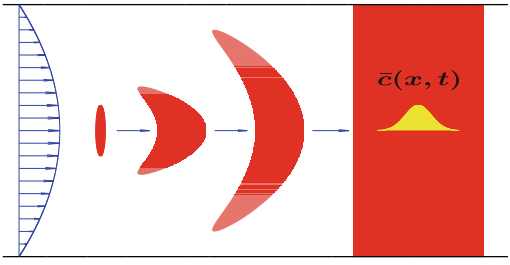


Fig. 11.2 Schematic behaviour of solute dispersion (*green* contours) in a long channel starting from a small release area due to a velocity field (*red*)

and D_y respectively. In this *anisotropic* context, the transport equation (11.18a) takes the form

$$\partial_T \mathsf{C} + \nabla \cdot (\mathsf{C}\mathbf{V}) = 0, \qquad \mathbf{V} = \mathbf{U} - \mathbf{D}\nabla \mathsf{C},$$

where $\mathbf{D}$ is a diffusion coefficient matrix, $\mathbf{D} = \begin{pmatrix} D_x & 0 \\ 0 & D_y \end{pmatrix}$. For the isotropic case, where $D_x = D_y = D$, this equation reduces to (11.18a).

We nondimensionalize by taking

$$\mathsf{C} = C_0 c, \qquad \mathsf{X} = L x, \qquad \mathsf{Y} = H y, \qquad \mathsf{T} = (L/U_0) t,$$

where C_0 is a concentration scale that could be set by the initial conditions and L is a typical length scale along the channel. Consequently, the scaled problem on $-\infty < x < \infty$, $-1 \le y \le 1$ becomes

$$\frac{\partial c}{\partial t} + (1 - y^2)\frac{\partial c}{\partial x} = \frac{1}{\mathrm{Pe}_x}\frac{\partial^2 c}{\partial x^2} + \frac{1}{\mathrm{Pe}_y}\frac{\partial^2 c}{\partial y^2}, \tag{11.19}$$

with boundary conditions $\partial_y c|_{y=\pm 1} = 0$. The Peclet numbers,

$$\mathrm{Pe}_x = \frac{U_0 L}{D_x}, \qquad \mathrm{Pe}_y = \frac{U_0 H^2}{D_y L},$$

give the relative importance of the convective flow along the channel versus diffusive effects in the x and y directions respectively. As the flow carries the solute along the channel, the concentration will disperse across the full width of the channel. Taylor showed that the average concentration across the channel at each x position could be estimated from a reduced model.

We note that the concentration can be separated into the average across the channel and deviation from the average

$$c(x, y, t) = \bar{c}(x, t) + \tilde{c}(x, y, t), \tag{11.20a}$$

where the average $\bar{c}$ is defined by

$$\bar{c}(x, t) \equiv \frac{1}{2}\int_{-1}^{1} c(x, y, t)\, dy. \tag{11.20b}$$

A consequence of this definition of $\bar{c}$ is that the deviation $\tilde{c}$ has zero-mean, i.e.

$$\int_{-1}^{1} \tilde{c}(x, y, t)\, dy = 0. \tag{11.20c}$$

Similarly, we can separate the flow, $u(y) = 1 - y^2$, into mean flow and deviation from the mean,

$$u(y) = \bar{u} + \tilde{u}(y), \tag{11.21a}$$

$$\bar{u} = \frac{1}{2}\int_{-1}^{1} 1 - y^2\, dy = \frac{2}{3}, \qquad \tilde{u}(y) = \frac{1}{3} - y^2. \tag{11.21b}$$

The PDE (11.19) can now be expanded as

$$\begin{aligned} &\frac{\partial \bar{c}}{\partial t} + \frac{\partial \tilde{c}}{\partial t} + \bar{u}\frac{\partial \bar{c}}{\partial x} + \tilde{u}\frac{\partial \bar{c}}{\partial x} + \bar{u}\frac{\partial \tilde{c}}{\partial x} + \tilde{u}\frac{\partial \tilde{c}}{\partial x} \\ &\qquad = \frac{1}{\mathrm{Pe}_x}\frac{\partial^2 \bar{c}}{\partial x^2} + \frac{1}{\mathrm{Pe}_x}\frac{\partial^2 \tilde{c}}{\partial x^2} + \frac{1}{\mathrm{Pe}_y}\frac{\partial^2 \tilde{c}}{\partial y^2}. \end{aligned} \tag{11.22}$$

This equation can be integrated term by term across the width of the channel,

$$\frac{1}{2}\int_{-1}^{1} (11.2)\, dy,$$

to give an evolution equation for $\bar{c}(x, t)$. The integrals of the deviations vanish, thus

$$\int_{-1}^{1} \tilde{c}_t\, dy = \bar{u}\int_{-1}^{1} \tilde{c}_x\, dy = \bar{c}\int_{-1}^{1} \tilde{u}\, dy = 0.$$

The boundary conditions on c yield corresponding no-flux boundary conditions on $\tilde{c}$ and hence the integral of the last term in the PDE also vanishes,

$$\left.\frac{\partial \tilde{c}}{\partial y}\right|_{y=\pm 1} = 0 \qquad \Longrightarrow \qquad \int_{-1}^{1} \frac{\partial^2 \tilde{c}}{\partial y^2}\, dy = \left.\frac{\partial \tilde{c}}{\partial y}\right|_{-1}^{1} = 0. \tag{11.23}$$

The remaining terms determine the evolution of the mean concentration,

$$\frac{\partial \bar{c}}{\partial t} + \bar{u}\frac{\partial \bar{c}}{\partial x} + \frac{1}{2}\int_{-1}^{1} \tilde{u}\frac{\partial \tilde{c}}{\partial x}\, dy = \frac{1}{\mathrm{Pe}_x}\frac{\partial^2 \bar{c}}{\partial x^2}. \tag{11.24}$$

In the absence of the integral term, this would be a linear advection-diffusion equation for $\bar{c}(x, t)$. While the integral involves only perturbation terms, it is not valid to neglect it since while the mean of each deviation factor is zero, the average of a product is generally not equal to the product of the averages. Consequently, additional analysis on the properties of $\tilde{c}$ is needed.

Subtracting the averaged Eq. (11.24) from the full problem (11.22) we obtain an equation for $\tilde{c}(x, y, t)$,

$$\frac{\partial \tilde{c}}{\partial t} + \tilde{u}\frac{\partial \bar{c}}{\partial x} + \bar{u}\frac{\partial \tilde{c}}{\partial x} + \tilde{u}\frac{\partial \tilde{c}}{\partial x} - \frac{1}{2}\int_{-1}^{1} \tilde{u}\frac{\partial \tilde{c}}{\partial x}\, dy = \frac{1}{\mathrm{Pe}_x}\frac{\partial^2 \tilde{c}}{\partial x^2} + \frac{1}{\mathrm{Pe}_y}\frac{\partial^2 \tilde{c}}{\partial y^2}. \tag{11.25}$$

Taylor's approach can then be expressed in terms of two assumptions: (i) the deviation is generally smaller than the mean, $\tilde{c} \ll \bar{c}$, and (ii) the channel can be assumed to be slender (similarly to problems in Chap. 8) and the y-derivative terms will dominate the corresponding x-derivatives, $\partial_y \gg \partial_x$. The leading order dominant balance of the largest terms from the left and right sides of (11.25) then yields

$$\left(\tfrac{1}{3} - y^2\right)\frac{\partial \bar{c}}{\partial x} = \frac{1}{\mathrm{Pe}_y}\frac{\partial^2 \tilde{c}}{\partial y^2}. \tag{11.26}$$

Since $\bar{c}$ does not depend on y, we can integrate this equation once with respect to y to get

$$\frac{\partial \tilde{c}}{\partial y} = \mathrm{Pe}_y \frac{\partial \bar{c}}{\partial x}\left(\tfrac{1}{3}y - \tfrac{1}{3}y^3 + A_1(x,t)\right).$$

Applying the no-flux boundary conditions (11.23) at $y = \pm 1$ determines the function of integration to be $A_1(x,t) \equiv 0$, so we can proceed by integrating again to yield

$$\tilde{c} = \mathrm{Pe}_y \frac{\partial \bar{c}}{\partial x}\left(\frac{1}{6}y^2 - \frac{1}{12}y^4 + A_2(x,t)\right),$$

where A_2 is a second function of integration. By definition, the perturbation must have zero mean, namely, $\int_{-1}^{1} \tilde{c}\, dy = 0$. Applying this condition determines $A_2 \equiv -7/180$ and hence we conclude that

$$\tilde{c} = \mathrm{Pe}_y \frac{\partial \bar{c}}{\partial x}\left(\frac{1}{6}y^2 - \frac{1}{12}y^4 - \frac{7}{180}\right). \tag{11.27}$$

We can now use this to calculate the integral term from (11.24),

$$\frac{1}{2}\int_{-1}^{1} \tilde{u}\frac{\partial \tilde{c}}{\partial x}\, dy = -\frac{8\mathrm{Pe}_y}{945}\frac{\partial^2 \bar{c}}{\partial x^2}.$$

Consequently, we can complete our model equation for the evolution of the average density across the channel (11.24) as

$$\frac{\partial \bar{c}}{\partial t} + \bar{u}\frac{\partial \bar{c}}{\partial x} = \alpha \frac{\partial^2 \bar{c}}{\partial x^2} \qquad \alpha = \frac{1}{\mathrm{Pe}_x}\left(1 + \frac{8}{945}\frac{\mathrm{U}_0^2 \mathrm{H}^2}{\mathrm{D}_x \mathrm{D}_y}\right), \tag{11.28}$$

where α is called the *Taylor enhanced diffusion coefficient*.

In comparison with the original two-dimensional problem (11.19), we have obtained a reduced model. Equation (11.28) is a one-dimensional linear convection-diffusion equation. Considering the left side of the model, we can identify a convective derivative with constant speed $\bar{u}$ and in fact, making a change of variables with this wave speed, $\bar{c}(x,t) = F(x - \bar{u}t, t)$, yields the classical diffusion equation for F,

$$\frac{\partial F}{\partial t} = \alpha \frac{\partial^2 F}{\partial z^2}. \tag{11.29}$$

As we have seen in Chap. 5, this can be solved using similarity solutions, the one appropriate for a point source (consistent with initial conditions describing release of solute at a single location) is $F(z,t) = M_0 \exp(-z^2/(4\alpha t))/\sqrt{4\pi\alpha t}$ and hence we can predict the large-time behaviour as

$$\bar{c}(x,t) \sim \frac{M_0}{\sqrt{4\pi\alpha t}} \exp\left(-\frac{(x - \bar{u}t)^2}{4\alpha t}\right), \tag{11.30}$$

where M_0 is the mass of solute (determined from the initial conditions).

11.4 Further Directions

In this chapter, we have sought to present approaches for obtaining information about the behaviour of solutions of PDEs that do not rely as heavily on perturbation methods as some of the previous chapters. The derivation of the Taylor dispersion results can be made more rigorous in terms of an asymptotic analysis similar to that used in Chap. 8 (see Fowler [37, pp. 222–223] and [58]).

11.5 Exercises

11.1 In Chap. 5 we derived the Gaussian self-similar solution of the heat equation $\rho_t = \rho_{xx}$. The most general form of this solution is

$$\rho(x,t) = \frac{C_1}{\sqrt{4\pi(t + C_2)}} \exp\left(-\frac{(x - C_3)^2}{4(t + C_2)}\right),$$

which contains three arbitrary constants. Evaluate the M_0, M_1, M_2 moment integrals for this solution and determine C_1, C_2, C_3 in terms of the results for (11.2a, 11.2b). This can be shown to select the $t \to \infty$ asymptotic solution of the initial value problem for the heat equation [108].

11.2 Consider the convection-diffusion-reaction equation

$$\frac{\partial \rho}{\partial t} + 2\frac{\partial \rho}{\partial x} = 3\frac{\partial^2 \rho}{\partial x^2} + 4\rho$$

subject to the following conditions:

(a) On the domain $-\infty < x < \infty$ with initial condition $\rho(x, 0) = e^{-x^2}$ and boundary conditions $\rho \to 0$ as $|x| \to \infty$. Define the moment integrals as

$$M_n(t) = \int_{-\infty}^{\infty} x^n \rho(x, t)\, dx.$$

Find $M_0(t)$, $M_1(t)$, $M_2(t)$.

(b) For the same equation on a semi-infinite domain, $0 \le x < \infty$, with initial condition $\rho(x, 0) = e^{-x}$ and boundary conditions

$$\rho(0, t) = 1, \qquad \rho(x \to \infty, t) = 0.$$

Define the moment integrals as

$$M_n(t) = \int_{0}^{\infty} x^n \rho(x, t)\, dx.$$

Write the differential equations and initial conditions for dM_0/dt and dM_1/dt. Write what additional information you would need in order to solve for $M_0(t)$ and $M_1(t)$.

11.3 Consider the reaction-diffusion equation

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

on $0 \le x \le \pi$ with no-flux boundary conditions,

$$\left.\frac{\partial \rho}{\partial x}\right|_{x=0} = 0, \qquad \left.\frac{\partial \rho}{\partial x}\right|_{x=\pi} = 0,$$

and the initial condition, $\rho(x, 0) = f(x)$.

(a) Derive the problem for the evolution of the mass $M_0(t)$, but show that the equations for the higher moments are not closed.
(b) Show that separation of variables can be used to write the solution in the form

$$\rho(x,t) = \sum_{k=0}^{\infty} a_k e^{-\lambda_k t} \cos(kx).$$

Determine the coefficients a_k in terms of the initial condition. Show that

$$\rho(x,t) \sim \frac{M_0(t)}{\pi} \quad \text{as} \quad t \to \infty.$$

(c) If the domain is changed to be $-\infty < x < \infty$, and the first three moment integrals of $f(x)$ converge, the exact solution can be written as

$$\rho(x,t) = \frac{e^{-t}}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} f(x-y) e^{-y^2/(4t)}\, dy.$$

Using results on integrals of the Gaussian and basic properties of double integrals, show that moments of this formula for the exact solution reproduce the results for M_0, M_1, M_2 that can be obtained from using the PDE alone.

11.4 The *Von Foerster/McKendrick model* describes the evolution of populations where age-distribution of individuals is of interest (called an *age-structured population model*). Consider a population described by a density function $\rho(a,t)$ of individuals with ages $a \geq 0$ [74]. Let the density for $0 < a < \infty$ evolve according to the transport equation

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial a} = -2\rho \tag{11.31}$$

with $\rho(a \to \infty) \to 0$ (describing that no-one lives forever).

(a) Suppose that the new births at any time are proportional to the total population size,

$$\rho(0,t) = \beta \int_0^{\infty} \rho(a,t)\, da.$$

Show that the model can be reduced to a single ODE for the total population, $M_0(t) = \int_0^{\infty} \rho(a,t)\, da$.
(b) One possible refinement of the birth condition could be to account for the fact that reproductive activity of individuals tends to decrease with increasing age. Consider the revised birth condition

$$\rho(0,t) = \int_0^\infty e^{-3a}\rho(a,t)\,da.$$

Cushing [28] describes an approach that can reduce age-structured models to ODE systems in terms of weighted moment integrals. Show that the model with this birth condition can be reduced to a system of two ODEs for

$$M_0(t) = \int_0^\infty \rho\,da, \qquad M_1(t) = \int_0^\infty e^{-3a}\rho\,da.$$

11.5 In Chap. 5 a similarity solution of the porous medium equation was derived using the same approach that was used for the heat equation. To see how the method of moments applies to nonlinear equations, consider

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left(\rho^3 \frac{\partial \rho}{\partial x}\right),$$

with $\rho \to 0$ for $|x| \to \infty$ and initial condition $\rho(x,0) = f(x) \geq 0$ on $-\infty < x < \infty$.

(a) Show that M_0, M_1 are constant for this problem.
(b) Show that while the moment model is not closed for M_2, $M_2(t)$ is a strictly increasing function.

11.6 While describing the spreading of the solute in the transverse direction depends on the presence of diffusion, justify the curved shape of the solute contours sketched in Fig. 11.2 using the method of characteristics for the PDE

$$\frac{\partial c}{\partial t} + (1-y^2)\frac{\partial c}{\partial x} = 0$$

starting from the initial conditions

$$c(x,y,t=0) = \begin{cases} \frac{1}{4} - (x^2+y^2) & x^2+y^2 \leq \frac{1}{4} \\ 0 & \text{elsewhere} \end{cases}$$

Hint: Recall Exercise 2.8.

11.7 Examine how the Taylor dispersion derivation changes for different velocity fields:

(a) for the uniform "plug flow" $\mathbf{U} = U_0\hat{\mathbf{I}}$,
(b) for linear shear flow $\mathbf{U} = U_0(1 - Y/H)\hat{\mathbf{I}}$,

11.8 Derive the Taylor diffusion model for laminar axisymmetric flow in a circular pipe of radius $\mathsf{R} = \mathrm{R}_0$:

$$\frac{\partial \mathsf{C}}{\partial \mathsf{T}} + \mathsf{U}_0 \left(1 - \frac{\mathsf{R}^2}{\mathrm{R}_0^2}\right) \frac{\partial \mathsf{C}}{\partial \mathsf{X}} = \mathrm{D}_x \frac{\partial^2 \mathsf{C}}{\partial \mathsf{X}^2} + \frac{\mathrm{D}_r}{\mathsf{R}} \frac{\partial}{\partial \mathsf{R}} \left(\mathsf{R} \frac{\partial \mathsf{C}}{\partial \mathsf{R}}\right)$$

with no flux boundary conditions at the pipe wall.

11.9 The key detail obtained in the derivation of the Taylor diffusion model (11.28) was the coefficient 8/945. The corresponding correction factor in Exercise 11.8 or for other geometries will be different constants. For the problem of two-dimensional laminar flows (11.19), some books and articles may give the "magic number" 2/105 instead. Show that this result does not clash with 8/945 given in the derivation above. Hint: Consider a different choice for the scaling of the flow velocity.

Part III
Case Studies

Chapter 12
Modelling in Applied Fluid Dynamics

This chapter differs substantially from the previous chapters in that we do not aim to explain any new methodologies here. Instead, we will make use the approaches developed earlier to illustrate two case studies of how mathematical models of physical problems can be formulated and analysed.

Our two examples applications come from problems in fluid dynamics, on air bearing sliders and wedge-rivulet flow. We begin by deriving a common core-model, called *lubrication theory*, that forms the basis for both applications. While the derivation of lubrication theory draws extensively on the methods for long-wave asymptotics that were introduced in Chap. 8, we will not attempt to individually identify the many other connections between the material in this chapter and the methods given in the other earlier chapters. Having this background, readers should be comfortable with the presentations here, which are at levels that are typical of basic modelling in current applied research and case studies in more advanced books on modelling [27, 37, 96]. For further examples of case studies in mathematical modelling see [51, 53, 69].

12.1 Lubrication Theory

The Navier-Stokes equations comprise the fundamental continuum mechanics mathematical model for the dynamics of fluids having viscosity (namely, realistic dissipation or internal friction). These partial differential equations are conservation laws for mass and momentum of fluids and universally applicable. However, for a number of analytical and computational reasons, they can be very challenging to solve for most problems. Consequently, to make progress, most applied studies make use of specific forms of their problems to further reduce the Navier-Stokes equations to more tractable models. For problems that involve slender layers of viscous fluids, *lubrication models* can be derived from the Navier-Stokes equations using asymptotics in terms of the dominance of viscous effects and the slenderness of the layer. Dat-

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9_12

ing back to the work of Osborne Reynolds (1842–1912), such models have become essential for advancing the analysis of problems from a diverse range of applications.

We begin with the Navier-Stokes equations for a compressible viscous Newtonian fluid. Written in vector form, the continuity equation and momentum balance are

$$\frac{\partial \tilde{\rho}}{\partial \mathsf{T}} + \nabla \cdot (\tilde{\rho}\mathbf{V}) = 0, \qquad \tilde{\rho}\left(\frac{\partial \mathbf{V}}{\partial \mathsf{T}} + \mathbf{V} \cdot \nabla \mathbf{V}\right) = -\nabla \mathsf{P} + \mu \nabla^2 \mathbf{V}, \tag{12.1}$$

where μ is the viscosity coefficient, $\tilde{\rho}(\mathsf{X}, \mathsf{Y}, \mathsf{Z}, \mathsf{T})$ is the fluid density, $\mathsf{P}(\mathsf{X}, \mathsf{Y}, \mathsf{Z}, \mathsf{T})$ is the pressure, and the fluid velocity field is $\mathbf{V}(\mathsf{X}, \mathsf{Y}, \mathsf{Z}, \mathsf{T}) = (\mathsf{U}, \mathsf{V}, \mathsf{W})$. Expanded out componentwise (12.1) yields a system of four coupled nonlinear PDEs,

$$\tilde{\rho}_{\mathsf{T}} + (\tilde{\rho}\mathsf{U})_{\mathsf{X}} + (\tilde{\rho}\mathsf{V})_{\mathsf{Y}} + (\tilde{\rho}\mathsf{W})_{\mathsf{Z}} = 0, \tag{12.2a}$$

$$\tilde{\rho}(\mathsf{U}_{\mathsf{T}} + \mathsf{U}\mathsf{U}_{\mathsf{X}} + \mathsf{V}\mathsf{U}_{\mathsf{Y}} + \mathsf{W}\mathsf{U}_{\mathsf{Z}}) = -\mathsf{P}_{\mathsf{X}} + \mu(\mathsf{U}_{XX} + \mathsf{U}_{YY} + \mathsf{U}_{ZZ}), \tag{12.2b}$$

$$\tilde{\rho}(\mathsf{V}_{\mathsf{T}} + \mathsf{U}\mathsf{V}_{\mathsf{X}} + \mathsf{V}\mathsf{V}_{\mathsf{Y}} + \mathsf{W}\mathsf{V}_{\mathsf{Z}}) = -\mathsf{P}_{\mathsf{Y}} + \mu(\mathsf{V}_{XX} + \mathsf{V}_{YY} + \mathsf{V}_{ZZ}), \tag{12.2c}$$

$$\tilde{\rho}(\mathsf{W}_{\mathsf{T}} + \mathsf{U}\mathsf{W}_{\mathsf{X}} + \mathsf{V}\mathsf{W}_{\mathsf{Y}} + \mathsf{W}\mathsf{W}_{\mathsf{Z}}) = -\mathsf{P}_{\mathsf{Z}} + \mu(\mathsf{W}_{XX} + \mathsf{W}_{YY} + \mathsf{W}_{ZZ}). \tag{12.2d}$$

For convenience, we simplify our presentation to the two-dimensional problem[1] where $\mathsf{W} \equiv 0$ and all properties are independent of Z (i.e. $\partial_{\mathsf{Z}} \equiv 0$) and hence consider the system,

$$\tilde{\rho}_{\mathsf{T}} + (\tilde{\rho}\mathsf{U})_{\mathsf{X}} + (\tilde{\rho}\mathsf{V})_{\mathsf{Y}} = 0, \tag{12.3a}$$

$$\tilde{\rho}(\mathsf{U}_{\mathsf{T}} + \mathsf{U}\mathsf{U}_{\mathsf{X}} + \mathsf{V}\mathsf{U}_{\mathsf{Y}}) = -\mathsf{P}_{\mathsf{X}} + \mu(\mathsf{U}_{XX} + \mathsf{U}_{YY}), \tag{12.3b}$$

$$\tilde{\rho}(\mathsf{V}_{\mathsf{T}} + \mathsf{U}\mathsf{V}_{\mathsf{X}} + \mathsf{V}\mathsf{V}_{\mathsf{Y}}) = -\mathsf{P}_{\mathsf{Y}} + \mu(\mathsf{V}_{XX} + \mathsf{V}_{YY}). \tag{12.3c}$$

We consider the dynamics of a thin layer of viscous fluid spreading on a flat solid surface, as shown in Fig. 12.1, with the average height given by $\bar{\mathrm{H}}$ and the lateral lengthscale given by $\bar{\mathrm{L}}$. We formally nondimensionalize using the scalings

$$\mathsf{X} = \bar{\mathrm{L}}x, \qquad \mathsf{Y} = \bar{\mathrm{H}}y, \qquad \mathsf{T} = \bar{\mathrm{T}}t, \tag{12.4a}$$

$$\mathsf{U} = \bar{\mathrm{U}}u, \qquad \mathsf{V} = \bar{\mathrm{V}}v, \qquad \mathsf{P} = \bar{\mathrm{P}}p, \qquad \tilde{\rho} = \bar{\rho}\rho. \tag{12.4b}$$

We go on to use the aspect ratio $\varepsilon = \mathrm{H}/\mathrm{L}$ as an asymptotic parameter, $\varepsilon \to 0$.

Applying the scalings to (12.3a) yields the nondimensional form

$$\frac{\partial \rho}{\partial t} + \left(\frac{\bar{\mathrm{U}}\bar{\mathrm{T}}}{\bar{\mathrm{L}}}\right)\frac{\partial}{\partial x}(\rho u) + \left(\frac{\bar{\mathrm{V}}\bar{\mathrm{T}}}{\bar{\mathrm{H}}}\right)\frac{\partial}{\partial y}(\rho v) = 0.$$

Conservation of mass is a fundamental property and it should hold exactly, independent of the geometry or any considerations of flow speed. Hence, independent of the final choice of characteristic scales, we should retain the general form given

[1] The derivation for the full three-dimensional system follows analogously.

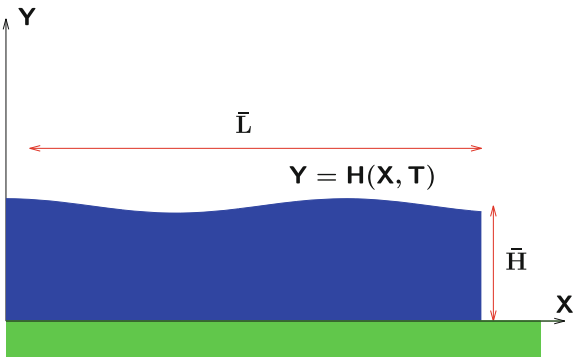


Fig. 12.1 A thin film of a viscous fluid coating a flat solid surface

by (12.3a). This motivates selecting scales to make the continuity equation scale-invariant; namely we pick a convective timescale $T = L/U$ and a vertical velocity scale based on the horizontal velocity and the aspect ratio, $V = \varepsilon U$, to yield the final dimensionless equation,

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho u) + \frac{\partial}{\partial y}(\rho v) = 0. \tag{12.5}$$

Turning to the momentum equations, using the scalings, the non-dimensional forms of (12.3b, 12.3c) are

$$\left(\varepsilon^2 \frac{\bar{\rho}\bar{U}\bar{L}}{\mu}\right)(u_t + uu_x + vu_y) = -\left(\varepsilon^2 \frac{\bar{P}\bar{L}}{\mu\bar{U}}\right) p_x + \varepsilon^2 u_{xx} + u_{yy}, \tag{12.6a}$$

$$\left(\varepsilon^2 \frac{\bar{\rho}\bar{U}\bar{L}}{\mu}\right)(v_t + uv_x + vv_y) = -\left(\frac{\bar{P}\bar{L}}{\mu\bar{U}}\right) p_y + \varepsilon^2 v_{xx} + v_{yy}. \tag{12.6b}$$

Apart from the aspect ratio, these equations contain two dimensionless parameters. The reduced Reynolds number gives the ratio of effects of inertial acceleration relative to viscous forces,

$$\widetilde{\mathrm{Re}} = \varepsilon^2 \frac{\bar{\rho}\bar{U}\bar{L}}{\mu} \ll 1, \tag{12.7}$$

To assert the dominance of viscous effects, we assume that this parameter is negligibly small.
Consequently, (12.6a, 12.6b) reduces to

$$0 = -\left(\varepsilon^2 \frac{\bar{P}\bar{L}}{\mu\bar{U}}\right) p_x + \varepsilon^2 u_{xx} + u_{yy}, \tag{12.8a}$$

$$0 = -\left(\frac{\bar{P}\bar{L}}{\mu\bar{U}}\right) p_y + \varepsilon^2 v_{xx} + v_{yy}. \tag{12.8b}$$

We note that this reduction of the Navier Stokes equations (having removed the nonlinear inertial acceleration terms) is called the *Stokes equations*, and for the incompressible (constant density) case, their dimensional form is

$$\mathbf{0} = -\nabla \mathsf{P} + \mu \nabla^2 \mathbf{V}, \qquad \nabla \cdot \mathbf{V} = 0. \tag{12.9}$$

The choice of the pressure scaling in (12.8a, 12.8b) remains. One option is to apply dominant balance to the y-momentum equation, yielding the derived scale $\bar{\mathsf{P}} = \mu \bar{\mathrm{U}}/\bar{\mathrm{L}} = O(1)$, which is called a viscous pressure scale since it balances viscous and pressure-gradient effects. However it can be shown from the resulting leading order x-equation, $u_{0yy} = 0$, that for many problems, this choice can be too restrictive and may to lead to solutions that will not capture the full structure of the flow. The other option is dominant balance in the x-equation, yielding another version of the viscous pressure scale, $\bar{\mathsf{P}} = \mu \bar{\mathrm{U}}/(\varepsilon^2 \bar{\mathrm{L}}) = O(\varepsilon^{-2})$. This leads to the leading order system,

$$\frac{\partial p_0}{\partial x} = \frac{\partial^2 u_0}{\partial y^2}, \qquad \frac{\partial p_0}{\partial y} = 0. \tag{12.10}$$

The second equation determines that p_0 is independent of y, namely $p_0 = p_0(x, t)$. Consequently the first equation can be integrated with respect to y to yield a parabolic form, sometimes called the *Nusselt velocity profile*,

$$u_0 = \tfrac{1}{2} \frac{\partial p_0}{\partial x} y^2 + C_1 y + C_2, \tag{12.11}$$

where C_1, C_2 are constants of integration with respect to y. These constants are determined by boundary conditions on the lateral velocity u at the top and bottom of the fluid layer for the particular problem at hand.

For viscous fluids in contact with solids, the *no-slip boundary condition* states that the velocity of the fluid tangential to the solid must match that of the solid. In this problem the solid surface is $y = 0$, and the no-slip condition specifies that $u_0(0) = 0$, which determines $C_2 = 0$ in (12.11). At the surface of the layer, $y = h(x, t)$, if the fluid were not subjected to any forces (surface stresses) then the lateral speed at the surface should be the same as the bulk of the fluid; the simplest form of the *stress-free boundary condition* is then

$$\left. \frac{du_0}{dy} \right|_{y=h} = 0 \qquad \Longrightarrow \qquad C_1 = -p_{0x}(x, t) h(x, t). \tag{12.12}$$

Consequently, we have determined that for films spreading on flat surfaces,

$$u_0(x, y, t) = \tfrac{1}{2} p_{0x} (y^2 - 2hy). \tag{12.13}$$

To complete the derivation of the lubrication model, we return to the equation for the conservation of mass, (12.5), and integrate it over the thickness of the layer,

$$\int_0^h \left[\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho u) + \frac{\partial}{\partial y}(\rho v) \right] dy = 0.$$

The final term can be integrated directly to give a relation between the fluxes through the bottom and top of the layer,

$$\rho v \Big|_{y=0}^{y=h} = -\int_0^h \left[\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho u) \right] dy. \tag{12.14}$$

At $y = 0$, the fluid rests on an impermeable solid surface, so there should be no flux through it, hence the no-flux boundary condition gives $v(y = 0) = 0$. At the top of the layer, the fluid has a free boundary, and the vertical component of the velocity there should follow the motion of the surface $y = h(x, t)$. Consider a particle on the surface that is carried by the flow, having position $(X(t), Y(t))$, and remains on the surface for all times, so $Y(t) = h(X(t), t)$. Then, using the chain rule and recalling the Lagrangian description of the velocity, the rate of change of the vertical position of any point on a free boundary is given by

$$\frac{dY}{dt} = \frac{\partial h}{\partial t} + \frac{\partial h}{\partial x}\frac{dX}{dt} \qquad \Longrightarrow \qquad v\Big|_{y=h} = \frac{\partial h}{\partial t} + u\frac{\partial h}{\partial x}\Big|_{y=h}, \tag{12.15}$$

which is called the *kinematic boundary condition*. Applying these two boundary conditions to (12.14) yields

$$\left(\rho \frac{\partial h}{\partial t} + \rho u \frac{\partial h}{\partial x} \right) \Big|_{y=h} + \int_0^h \left[\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho u) \right] dy = 0.$$

Applying Leibniz's rule (in reverse, with respect to x and t separately) then gives a transport equation describing conservation of mass

$$\frac{\partial}{\partial t}\left(\int_0^h \rho \, dy \right) + \frac{\partial}{\partial x}\left(\int_0^h \rho u \, dy \right) = 0. \tag{12.16}$$

If we take $\rho = \rho(x, t)$ then this reduces to

$$\frac{\partial (\rho h)}{\partial t} + \frac{\partial}{\partial x}\left(\rho \int_0^h u \, dy \right) = 0. \tag{12.17}$$

Finally, using the velocity profile (12.13) to evaluate the integral yields

$$\frac{\partial(\rho h)}{\partial t} = \frac{1}{3}\frac{\partial}{\partial x}\left(\rho h^3 \frac{\partial p}{\partial x}\right). \tag{12.18}$$

This is a form of the *Reynolds equation* that forms the basis of lubrication models in a wide array of applications involving free-surface thin film flows [25, 75, 79, 82] and fluid-structure interactions involving lubricating fluid layers [12, 46].

12.2 Dynamics of an Air Bearing Slider

One of the historical motivations for the development of lubrication theory has been the use of fluids as cushioning layers between moving surfaces—an everyday application is oil lubricating parts of an engine to allow relative motion without metal surfaces scratching against each other.

Lubrication theory is also used in modern technology [12]—in computer hard disk drives, air acts as a lubricating gas layer separating the electronic read/write head from the rapidly rotating rigid data disks. To maximize the data density on the disk, the write head must be kept very close to the surface, and to maximize data access speed, the disk speed should be high. However both of these effects might suggest the system could be sensitive to any variations in disk speed, external forces and motion of the system, or variations of pressure that might lead the head to collide with the disk. Hence it is very important to have a model of the system that can guide the design process to configurations that are stable to perturbations.

The one-dimensional[2] version of the geometry we are considering is shown in Fig. 12.2. The read/write head, sometimes also called a *slider bearing* [79], has a particular length $\bar{\mathrm{L}}$ and has its lateral position fixed, but it is allowed to move vertically. The disk is a flat surface ($\mathsf{Y} = 0$) moving horizontally at local speed $\bar{\mathrm{U}}$. The motion of the disk will generate a flow of air under the slider due to the no-slip boundary condition on the gas. This flow will generate a lift force on the slider that will balance against the weight of the slider and any other applied downward forces. The desired average gap height $\bar{\mathrm{H}}$ sets a vertical lengthscale. The gap height between the disk and slider surfaces is given by

$$\mathsf{H}(\mathsf{X}, \mathsf{T}) = \mathsf{A}(\mathsf{T}) + \mathsf{S}(\mathsf{X}), \tag{12.19}$$

where A is the vertical position of the leading edge of the slider and $\mathsf{S}(\mathsf{X})$ gives the shape of the slider's lower surface, with $\mathsf{S}(0) = 0$. Outside of the region under the slider, the air pressure will be assumed to be the usual atmospheric pressure, $\mathrm{P}_{\mathrm{atm}}$.

[2] Referring to the number of lateral dimensions.

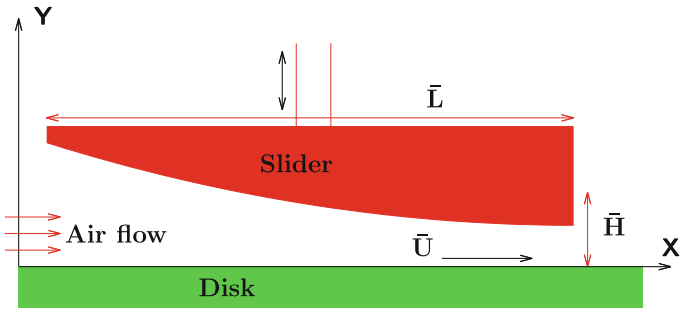


Fig. 12.2 The geometry of the problem for an air bearing slider. Lubrication theory can be applied to model the air flow in the thin gap between the moving disk and slider surfaces

Under the assumption that the reduced Reynolds number is small, (12.7), and that the aspect ratio of the gap is small, $\varepsilon \equiv \bar{H}/\bar{L} \ll 1$, lubrication theory can be applied. Written in dimensional form, the parabolic velocity profile (12.11) in the gap is

$$U = \frac{1}{2\mu}\frac{\partial P}{\partial X}Y^2 + C_1 Y + C_2. \tag{12.20}$$

The scalings and boundary conditions appropriate to this problem must be applied to this general form. At the disk and slider surfaces, no-slip boundary conditions impose the lateral speeds on the gas flow,

$$U(Y = 0) = \bar{U}, \qquad U(Y = H) = 0.$$

These conditions select the constants of integration in (12.20) to yield

$$U = \frac{1}{2\mu}\frac{\partial P}{\partial X}(Y^2 - HY) + \bar{U}\left(1 - \frac{Y}{H}\right), \tag{12.21}$$

where we have separated the contributions due to the pressure-gradient driven Poiseuille flow term and the linear shear-flow Couette flow term. The next step is to use this velocity profile in the equation of conservation of mass to derive the appropriate form of the Reynolds equation for this problem.

While Eq. (12.16) was derived under a different type of boundary condition on the upper surface of the fluid layer, it can be shown that the same equation is obtained when the kinematic boundary condition is replaced by a no-flux condition appropriate to the interface with the slider surface at $Y = H$ [46]. This could be expected from physical considerations of (12.17) as a transport equation; since the mass of air in any column, $\tilde{\rho}H$, is conserved, its rate of change can only be due to transport to other positions in the gap by the flux. Consequently, we obtain the dimensional equation

$$\frac{\partial(\tilde{\rho}H)}{\partial T} + \frac{\bar{U}}{2}\frac{\partial(\tilde{\rho}H)}{\partial X} = \frac{1}{12\mu}\frac{\partial}{\partial X}\left(\tilde{\rho}H^3\frac{\partial P}{\partial X}\right). \tag{12.22}$$

To nondimensionalize this equation, we need to select characteristic scales. For the velocity scale, it is natural to use the imposed disk speed $\bar{\mathrm{U}}$. For the lateral length scale, the length of the slider, $\bar{\mathrm{L}}$, is a convenient choice; the derived convective timescale is then $\bar{\mathrm{T}} = \bar{\mathrm{L}}/\bar{\mathrm{U}}$. In the vertical direction, the reference height $\bar{\mathrm{H}}$ can be used. Equation (12.22) is linear with respect to $\tilde{\rho}$, so whatever the choice of the characteristic density, it will scale out. The boundary conditions on the pressure at the edges of the slider give an imposed pressure scale, $\bar{\mathrm{P}} = \mathrm{P}_{\mathrm{atm}}$. Using these scalings in (12.4) and writing $\mathsf{H} = \bar{\mathrm{H}}h$ we obtain the dimensionless equation [30, 46]

$$\frac{\partial(\rho h)}{\partial t} + \tfrac{1}{2}\frac{\partial(\rho h)}{\partial x} = \frac{1}{\Lambda}\frac{\partial}{\partial x}\left(\rho h^3 \frac{\partial p}{\partial x}\right),$$

containing one dimensionless parameter, called the *bearing number*,

$$\Lambda = \frac{12\mu\bar{\mathrm{L}}\bar{\mathrm{U}}}{\mathrm{P}_{\mathrm{atm}}\bar{\mathrm{H}}^2}, \tag{12.23}$$

which gives the ratio of the relative effects of the convective (Couette) to diffusive (Poiseuille) terms, analogous to a Peclet number. For applications of interest, small gap heights, small aspect ratios and high speeds all point to large bearing numbers; hence we define $\delta = 1/\Lambda$ as a small parameter.

To complete the formulation of the model, we need to specify a relation between the pressure and the gas density. The simplest choice for the equation of state is the ideal gas law, which states that the pressure is proportional to the product of the density with the temperature. Assuming the temperature to be held fixed, this gives $\rho = kp$ and reduces the Reynolds equation to a PDE for the evolution of the product $p(x,t)h(x,t)$,

$$\frac{\partial(ph)}{\partial t} + \tfrac{1}{2}\frac{\partial(ph)}{\partial x} = \delta\frac{\partial}{\partial x}\left(ph^3\frac{\partial p}{\partial x}\right), \tag{12.24a}$$

on the domain $0 \le x \le 1$ and is subject to the boundary conditions on the pressure,

$$p(0,t) = 1, \qquad p(1,t) = 1. \tag{12.24b}$$

For $\delta \to 0$, this is a singularly perturbed boundary value problem and can be shown to generate a boundary layer in the pressure at $x_* = 1$, the trailing edge of the slider. On the rest of the domain, $0 \le x < 1$, we can approximate the solution by the solution of the leading order outer problem,

$$\frac{\partial(p_0 h)}{\partial t} + \tfrac{1}{2}\frac{\partial(p_0 h)}{\partial x} = 0. \tag{12.25}$$

This is a linear advection equation and yields the solution as a traveling wave with speed one half, $F(x - \frac{1}{2}t)$, which can also be expressed as

$$p_0(x,t)h(x,t) = f(t-2x), \tag{12.26}$$

for some choice of function $f(s)$. Applying the boundary condition on the pressure at $x = 0$ and evaluating the gap height there, from (12.19), $h(0,t) = a(t)$, and we determine $f(t) \equiv a(t)$. Consequently the solution of this signaling-type wave problem for the pressure can be expressed as

$$p_0(x,t) = \frac{a(t-2x)}{a(t)+s(x)}, \tag{12.27}$$

from $p_0 = f/h$. We can now use this representation of the pressure to return to the primary question of the stability and dynamics of the slider.

The motion of the slider is characterized by the height of the leading edge, $y = a(t)$, hence we seek an evolution equation for $a(t)$. This will be provided by a force balance in the vertical direction. As described above, the air flow generates a lift force on the slider, this is given by the integral of the excess pressure over the domain of the slider,

$$F_L(t) = \int_0^1 (p(x,t)-1)\,dx.$$

This balances against the downward applied load imposed on the slider by structural constraints, F_S to yield the force balance equation

$$m\frac{d^2a}{dt^2} = \int_0^1 \frac{a(t-2x)}{a(t)+s(x)}\,dx - (F_S+1), \tag{12.28}$$

where m is the scaled mass of the slider. We note that by neglecting the trailing edge boundary layer in the pressure, we only make a small error, $O(\Lambda^{-1})$ in calculating the lift integral.

We can obtain a time-independent steady-state, $a = \bar{a}$, by solving an algebraic equation,

$$\bar{a}\int_0^1 \frac{dx}{\bar{a}+s(x)} = F_S + 1.$$

Then, we can examine the linear stability of this steady solution to small deviations by assuming a perturbed solution, $a(t) = \bar{a} + \sigma e^{\lambda t}$, for $\sigma \to 0$. Substituting this form into (12.28) yields an equation for the exponential growth rate λ at $O(\sigma)$,

$$m\lambda^2 = \int_0^1 \frac{e^{-2\lambda x}}{\bar{a}+s(x)}\,dx - \int_0^1 \frac{\bar{a}}{(\bar{a}+s(x))^2}\,dx. \tag{12.29}$$

If the governing equation for $a(t)$ were an autonomous ODE, then λ would be given by the roots of a characteristic polynomial. However due to the unusual shifted depen-

dence on the solution in the integral, (12.28) can be related to a delay-differential equation and yields a more challenging transcendental equation for the linear stability analysis. Careful consideration of this problem was explored in [109].

12.3 Rivulets in a Wedge Geometry

In the previous section, we considered a scenario where the substrate was uniformly flat. Here, we investigate a lubrication-type problem for a rivulet (slender thread) of viscous fluid constrained within a wedge geometry[3] (see Fig. 12.3). In particular, we will study the so-called *large-time* dynamics of the problem using asymptotic approximations and self-similar solutions. As well as being a well-observed physical phenomena, interesting in its own right, the results that we obtain also have application to the study of foam drainage, where the flow takes place within the "triangular" shaped region of the Plateau borders[4] (see [103] for a general discussion of the dynamics of foams).

We shall take X to represent the distance along the wedge, with Y as the vertical upward distance from the base of the wedge, and Z as the transverse distance from the centreline of the wedge (along with the respective velocities U, V and W).

We begin with the dimensional Navier-Stokes equations in three-dimensions (12.2); the dimensional density $\tilde{\rho}$ is taken to be constant. We apply the nondimensionalisation

$$\mathsf{X} = \bar{\mathrm{L}}x, \qquad \mathsf{Y} = \bar{\mathrm{H}}y, \qquad \mathsf{Z} = \bar{\mathrm{H}}z, \qquad \mathsf{T} = \bar{\mathrm{T}}t, \tag{12.30a}$$

$$\mathsf{U} = \bar{\mathrm{V}}u/\varepsilon, \qquad \mathsf{V} = \bar{\mathrm{V}}v, \qquad \mathsf{W} = \bar{\mathrm{V}}w, \qquad \mathsf{P} = \bar{\mathrm{P}}p, \tag{12.30b}$$

where $\bar{\mathrm{H}}$ is a typical fluid depth and $\bar{\mathrm{V}} = \varepsilon^2\gamma/\mu$ is a representative velocity scale (here, γ is the coefficient of surface tension for the free surface of the liquid). The expressions (12.30) imply that we consider the magnitude of the flows in the vertical and transverse directions to be of the same order, and (as in the previous section) the aspect ratio $\varepsilon = \mathrm{H}/\mathrm{L} \ll 1$. After substitution and collecting terms in orders of ε, the leading order equations are found to be

$$\frac{\partial p}{\partial x} = \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}, \qquad \frac{\partial p}{\partial y} = 0, \qquad \frac{\partial p}{\partial z} = 0, \tag{12.31}$$

being a reduced version of the Stokes equations (12.9). The last two equations of (12.31) show that to leading order the pressure field is independent of both y and z, and hence $p = P(x,t)$. Furthermore, geometric considerations can be used to show

[3] This analysis is based on the paper [16] to which we refer the interested reader for more details.

[4] The sides of a Plateau border are actually circular arcs and there is no free-surface, but the model equations that we derive here are still applicable. This is an example of how modelling and investigating one problem can also provide useful information for other related problems.

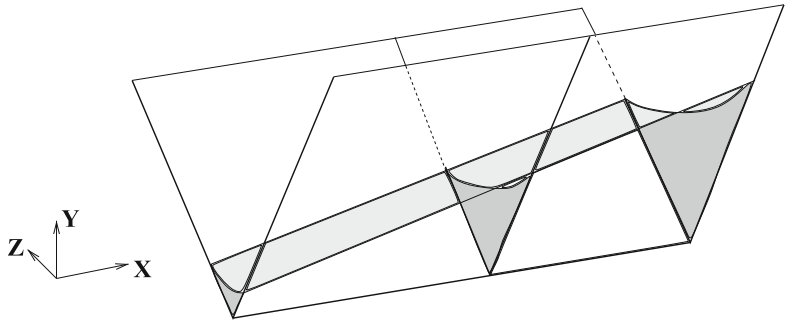


Fig. 12.3 Schematic of the problem geometry

that $P \propto A^{-1/2}$, where $A(x,t)$ is the cross-sectional area of the rivulet at location x and time t. The first equation of (12.31) can be solved subject to no-slip boundary conditions on the substrate and pressure jump conditions across the free surface of the liquid to yield $u \propto A^2 \partial P/\partial x$.

The conservation of mass result for the cross-sectional area $A(x,t)$ takes the form

$$\frac{\partial A}{\partial t} + \frac{\partial}{\partial z}\left(\int_D u\, dy\, dz\right) = 0, \tag{12.32}$$

where D is the area of the (y,z) plane occupied by liquid at fixed x and t. Substituting in the relevant expressions for u and P from above consequently results in a PDE for the evolution of $A(x,t)$ which takes the (porous-medium-equation-type) form

$$\frac{\partial A}{\partial t} = \frac{\partial}{\partial x}\left(A^{1/2}\frac{\partial A}{\partial x}\right), \quad 0 \le x < \infty \tag{12.33}$$

and is the basic equation that we will further analyse. For initial data, we impose

$$A(x,0) = \begin{cases} M & 0 \le x < 1, \\ 0 & x \ge 1, \end{cases} \tag{12.34}$$

corresponding to a constant cross-section of fluid over $0 \le x < 1$ with a dry wedge for $x > 1$; the initial volume of liquid is consequently also given by M.

Once the liquid is in motion ($t > 0$), there will be a moving free-boundary $x = s(t)$ (with $A \equiv 0$ for $x > s(t)$), on which we impose the physically sensible conditions $A = 0$ (zero height) and $A^{1/2}A_x = 0$ (zero flux across the interface). At $x = 0$, we also impose zero flux (corresponding to a solid wall), so that the total mass of liquid in the wedge is conserved for all time, i.e.

$$\frac{d}{dt}\int_0^{s(t)} A(x,t)\,dx = 0 \quad \Longrightarrow \quad \int_0^{s(t)} A(x,t)\,dx = \int_0^{s(t)} A(x,0)\,dx = M. \tag{12.35}$$

The nonlinear diffusion equation (12.33) coupled with the conservation law (12.35) admits a similarity solution

$$A = t^{-2/5} f(\eta), \quad \eta = \frac{x}{t^{2/5}}, \tag{12.36}$$

where $f(\eta)$ satisfies

$$-\frac{2}{5}\left(f + \eta \frac{df}{d\eta}\right) = \frac{d}{d\eta}\left(f^{1/2}\frac{df}{d\eta}\right),$$

with solution

$$f(\eta) = \frac{1}{100}(\eta_0^2 - \eta^2)_+^2; \tag{12.37}$$

the subscripted plus sign here denotes that we are only interested in the non-negative part of the solution, with the moving interface $x = s(t)$ corresponding to $\eta = \eta_0$ in terms of the similarity variables. Of course, (12.36) is singular in the limit $t \to 0^+$ and so cannot satisfy the initial condition (12.35). However, it is well-known that similarity solutions often act as large-time attractors for the dynamics of nonlinear diffusion equations and so we can expect $A(x,t)$ to approach (12.36) and (12.37) at large times.

12.3.1 Imbibition in a Vertical Wedge

Suppose now that we orient the wedge so that it is vertical, with the base located at $x = 0$. In such a geometry, we can no longer neglect gravity and it will be the competing effects of gravity (acting downwards) and capillarity (acting upwards) that drives the motion of the fluid.

Taking gravity to act equally everywhere on the fluid leads to a convective flow in the downward direction that requires the evolution Eq. (12.33) to be adjusted to read

$$\frac{\partial A}{\partial t} - 2A\frac{\partial A}{\partial x} = \frac{\partial}{\partial x}\left(A^{1/2}\frac{\partial A}{\partial x}\right), \quad 0 \le x < \infty; \tag{12.38}$$

we note that the left-hand-side terms of (12.38) correspond to a quasilinear convection equation that can be tackled by the method of characteristics (starting from the initial data (12.34)) and shown to have smooth solutions for all $t > 0$.

Physically, we expect the amount of liquid that can be pulled up the wedge by capillarity to be small compared to the bulk mass. We therefore assume that the solution in the latter region will effectively be a steady-state solution of (12.38) and consequently find

$$A = \frac{4}{(x + x_0)^2} \quad \text{for} \quad x = O(1). \tag{12.39}$$

The value of x_0 is set (to leading order) by the mass constraint (12.35), by which we calculate $x_0 = 4/M$. The steady-state *inner solution* (12.39) will match to the *outer solution* governed by capillary action.

In the capillary flow region, we cannot neglect the time-dependence in (12.38). We do not attempt to derive an exact solution of (12.38), but note that it admits a similarity solution of the form

$$A = t^{-2/3} g(\mu), \quad \mu = \frac{x}{t^{1/3}}, \tag{12.40}$$

with $g(\mu)$ satisfying

$$-\frac{2}{3} g - \frac{1}{3} \mu g_\mu = \left(g^{1/2} g_\mu - g^2 \right)_\mu .$$

Close to $x = s(t)$, with $s(t) = \mu_0 t^{1/3}$ in terms of the similarity variables,

$$g \sim \frac{\mu_0^2}{36} (\mu_0 - \mu)^2 \quad \text{as} \quad \mu \to \mu_0^-,$$

so that the solution has zero contact angle at $\mu = \mu_0$. As $\mu \to 0^+$ we must match with (12.39) as $x \to \infty$ and this requires that

$$g \sim \frac{4}{\mu^2} \quad \text{as} \quad \mu \to 0^+.$$

A schematic of the overall (large-time) solution structure is shown in Fig. 12.4.

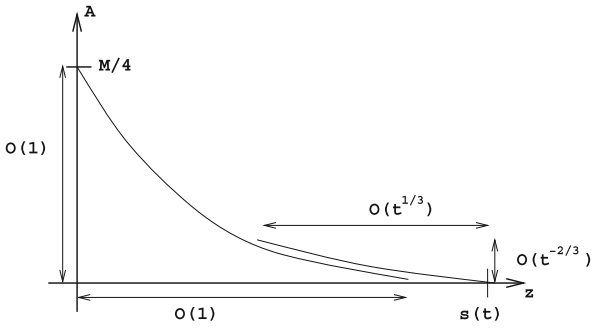


Fig. 12.4 Schematic of the large-time solution for gravity opposing capillarity based on the solution of Eq. (12.38)

12.3.2 Draining in a Vertical Wedge

In this scenario, $x = 0$ is taken to be at the top of the wedge, with the liquid moving downwards (again starting from (12.34)) under the now combined effects of gravity and capillarity. The evolution equation takes the same form as (12.38), except for a different sign on the convective term

$$\frac{\partial A}{\partial t} + 2A\frac{\partial A}{\partial x} = \frac{\partial}{\partial x}\left(A^{1/2}\frac{\partial A}{\partial x}\right), \quad 0 \le x < \infty; \tag{12.41}$$

notably, in contrast to the case of imbibition, the quasilinear equation allows shocks to form and we will observe this phenomena in the large time solution to (12.41).

Away from $x = 0$, the dynamics of the (*outer region*) solution are primarily controlled by the effects of gravity and the reduced PDE is consequently given by

$$\frac{\partial A}{\partial t} \sim -2A\frac{\partial A}{\partial x}, \tag{12.42}$$

with a solution of self-similar form

$$A = t^{-1/2}h(v), \quad v = \frac{x}{t^{1/2}}. \tag{12.43}$$

Here, $h(v)$ satisfies

$$\frac{1}{2}\frac{d}{dv}(hv) = \frac{d}{dv}(h^2),$$

so that

$$h = \begin{cases} \frac{1}{2}v & 0 < v < v_0, \\ 0 & v > v_0. \end{cases} \tag{12.44}$$

Standard matching arguments imply that $A \sim z/2t$ as $v \to 0^+$, while conservation of mass (12.35) determines the shock location $v_0 = 2\sqrt{M}$. Across the shock, the effects of capillarity are important and smooth the solution over a narrow *interior layer* with scalings

$$x = s(t) + t^{1/4}\xi, \quad s \sim v_0 t^{1/2}, \quad A \sim t^{-1/2}\Phi(\xi).$$

As a result, we find

$$\Phi(\xi) = \frac{v_0}{2}\tanh^2\left(\frac{1}{2}\sqrt{\frac{v_0}{2}}\xi\right), \tag{12.45}$$

which corresponds to a travelling wave solution of (12.41).

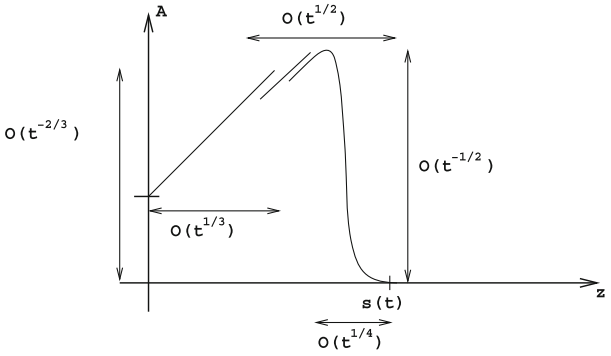


Fig. 12.5 Schematic of the large-time solution for gravity assisting capillarity based on the solution of Eq. (12.41)

The (*inner solution*) behaviour close to $x = 0$ is again given by (12.40), with $g \sim \mu/2$ as $\mu \to \infty$ in order to match with the outer solution. Finally, as $u \to 0^+$, we have $g \sim g_0 + g_0^{3/2} v$ as a consequence of the zero-flux boundary condition imposed at $x = 0$; numerical calculations yield $g_0 \approx 0.5885$. The entire large-time solution structure is illustrated schematically in Fig. 12.5.

A number of extensions of this problem are possible, such as the introduction of an influx of liquid at $x = 0$ and the investigation of different geometries for the wedge shape (see [16] for a detailed description of the previous analysis and further extensions).

Epilogue

The approaches and methods described in this book have long histories and have been employed to make major advances in research on many challenging fundamental problems. These approaches to modelling and problem reduction still form the basis for a great deal of current research on more advanced problems. As one example, the quasi-steady state assumption from Chap. 10 is used as the basis for handling large systems of complex bio-chemical reactions.

Current directions in modelling of physical systems have also put a new focus on complex systems [86] and *multiscale modelling* [33]. These are models that seek to constructively incorporate layers of understanding that come from different physical scales, often referred to as micro-, meso-, and macro-scales. Examples include (i) improved models for materials properties in continuum model settings that are derived from microscale models at the atomic scale and (ii) systems biology models of organs or entire physiological systems building from models of cells and biochemical reactions.

One of many active forums for current work on modelling is the journal *Multiscale Modeling and Simulation* published by the Society of Industrial and Applied Mathematics (SIAM). Many multiscale models make use of asymptotic approaches[1] to implement matching of descriptions at different scales in combination with numerical computations to make progress on fundamental questions in a broad array of application areas. We hope that this book has given readers a strong starting point for moving on to such advanced and challenging problems.

TW and MB

[1] Including boundary layers and other extensions of multiple scales, like averaging [73, 102] and homogenisation theory [49].

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9

Appendix A
Trigonometric Identities and Fourier Series

A brief summary of useful trigonometric identities:

$$\sin^2 x + \cos^2 x = 1$$

$$\sin x \cos y + \cos x \sin y = \sin(x+y) \qquad \cos x \cos y - \sin x \sin y = \cos(x+y)$$

$$\sin x \cos y = \tfrac{1}{2}\left(\sin(x-y) + \sin(x+y)\right) \qquad \sin x \sin y = \tfrac{1}{2}\left(\cos(x-y) - \cos(x+y)\right)$$

$$\cos x \cos y = \tfrac{1}{2}\left(\cos(x-y) + \cos(x+y)\right)$$

$$\sin(2t) = 2\sin t \cos t \qquad \cos(2t) = 2\cos^2 t - 1$$

$$\sin^2 t = \frac{1-\cos 2t}{2} \qquad \cos^2 t = \frac{1+\cos 2t}{2}$$

$$\sin^3 t = \frac{3\sin t - \sin 3t}{4} \qquad \cos^3 t = \frac{3\cos t + \cos 3t}{4}$$

$$\sin^2 t \cos t = \frac{\cos t - \cos 3t}{4} \qquad \sin t \cos^2 t = \frac{\sin t + \sin 3t}{4}$$

All trigonometric identities can be derived from successive applications of the formulas for the sum or difference of angles $(x \pm y)$ and further algebra.

The need for these identities can be eliminated by using Euler's formula, $e^{i\theta} = \cos\theta + i\sin\theta$, to replace cosine and sine by their complex representations,

$$\cos t = \mathrm{Re}(e^{it}) = \frac{e^{it} + e^{-it}}{2} \qquad \sin t = \mathrm{Im}(e^{it}) = \frac{e^{it} - e^{-it}}{2i},$$

then all results follow from algebra and re-grouping $e^{\pm int}$ to determine coefficients of $\cos(nt)$ and $\sin(nt)$ terms.

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9

A.1 Trigonometric Fourier Series

Analogous to the way that every nth order polynomial can be expressed as a finite Taylor polynomial, the trigonometric identities allow every nth order product of sines and cosines to be written as a sum of sines and cosines. Fourier series generalise this to represent all integrable periodic functions in terms of an infinite series of sines and cosines.

Expansions of given periodic functions, $f(x)$ on $-L < x < L$ can be written as

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(\tfrac{n\pi}{L}x) + \sum_{n=1}^{\infty} b_n \sin(\tfrac{n\pi}{L}x) \tag{A.1a}$$

where for $n = 0, 1, 2, \ldots$

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos(\tfrac{n\pi}{L}x)\,dx \qquad b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin(\tfrac{n\pi}{L}x)\,dx. \tag{A.1b}$$

The Fourier series for a function f will converge to the value of the function at all points where $f(x)$ is continuous.

The Fourier expansion can also be written in *complex form* as

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{in\pi x/L} \qquad c_n = \frac{1}{2L}\int_{-L}^{L} f(x)e^{-in\pi x/L}\,dx. \tag{A.2}$$

The *Fourier cosine series* for a function $F(x)$ given on $0 \le x < L$ can be derived from the general Fourier series by defining $f(x)$ to be the even extension of $F(x)$:

$$f(x) = \begin{cases} F(x) & 0 \le x < L, \\ F(-x) & -L < x \le 0 \end{cases}$$

$$F(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(\tfrac{n\pi}{L}x) \qquad a_n = \frac{2}{L}\int_{0}^{L} F(x)\cos(\tfrac{n\pi}{L}x)\,dx. \tag{A.3}$$

Similarly, the *Fourier sine series* for a function $F(x)$ given on $0 \le x < L$ can be derived from the general Fourier series by defining $f(x)$ to be the odd extension of $F(x)$:

$$f(x) = \begin{cases} F(x) & 0 \le x < L, \\ -F(-x) & -L < x \le 0 \end{cases}$$

$$F(x) = \sum_{n=1}^{\infty} b_n \sin(\tfrac{n\pi}{L}x) \qquad b_n = \frac{2}{L}\int_{0}^{L} F(x)\sin(\tfrac{n\pi}{L}x)\,dx. \tag{A.4}$$

Solutions to Selected Problems

Chapter 1

1.1 Solving the ODE for m first yields $m(t) = m_0 e^{-t}$. After substituting $m(t)$ into the equation for $z(t)$, this equation can be integrated twice and with $z'(0) = 0$ yields $z(t) = (\tau - g)(e^t - t - 1)$. From this solution we can see that lift-off occurs ($z > 0$ for $t > 0$) if $\tau > g$.

1.2 For (1.8), direct integration yields $A(t) = A_0 + kt$. For (1.9), separation of variables yields $A(t) = A_0 e^{-kt}$ and for (1.10) substitution into the equation for $B(t)$ yields $B(t) = B_0 + A_0 - A_0 e^{-kt}$. For (1.12), note that $dA/dt = dB/dt$ so $B(t) = A(t) + c$ where from the IC's $c = B_0 - A_0$, then substituting into the ODE for A, separation of variables yields $A(t) = A_0(B_0 - A_0)/(B_0 \exp(-k(A_0 - B_0)t) - A_0)$.

1.3 See (10.12a).

1.4 See Exercise 4.5 and set $\mathsf{F} = 0$ in (4.49).

1.5 (a) All solutions starting from $x_0 \neq 0$ approach either $x = \pm 1$. The *basin of attraction* of $x = -1$ is $x_0 < 0$ and all solutions having $x(t \to \infty) \to 1$ have $x_0 > 0$. (b) Second-order equilibria occur where $f = 0$ has a double root, $f'(x_*) = 1 - 3x_*^2 = 0$, namely $x_* = \pm 1/\sqrt{3}$. Substituting these x_* into $f(x_*) = 0$ yields $k_\pm = \pm 2\sqrt{3}/9$. Plotting directions on the phase lines shows these bifurcation values correspond to the changes in qualitative behaviours.

Chapter 2

2.1 Using the hint, we can express the problem as

$$\frac{df_{\text{avg}}}{dt} = \lim_{\varepsilon \to 0} \frac{d}{dt}\left(\frac{1}{\varepsilon h(t)} \int_{a(t)}^{a(t)+\varepsilon h(t)} f(x,t)\, dx \right).$$

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9

We are taking a derivative of a product of functions with the second factor being an integral, hence we apply Leibniz's rule (with respect to t) to it,

$$= \frac{1}{\varepsilon}\left(-\frac{h'}{h^2}\int_a^{a+\varepsilon h} f\,dx + \frac{1}{h}\left[\int_a^{a+\varepsilon h}\frac{\partial f}{\partial t}\,dx + f(a+\varepsilon h,t)\frac{d(a+\varepsilon h)}{dt} - f(a,t)\frac{da}{dt}\right]\right)$$

Now we are ready to consider the limit for $\varepsilon \to 0$. We will expand out the first two terms in the Taylor series for $\varepsilon \to 0$, using Leibniz's rule again, to take the derivative of integrals (with respect to ε),

$$\begin{aligned}\approx \frac{1}{\varepsilon}\Bigg(&-\frac{h'}{h^2}\left[\int_a^a f\,dx + \varepsilon f(a,t) - 0\right]\\ &+\frac{1}{h}\left[\int_a^a \frac{\partial f}{\partial t}\,dx + \varepsilon\frac{\partial f}{\partial t}h\bigg|_{x=a} + \left(f+\varepsilon h\frac{\partial f}{\partial x}\right)\bigg|_{x=a}\left(\frac{da}{dt}+\varepsilon\frac{dh}{dt}\right) - f(a,t)\frac{da}{dt}\right]\Bigg)\end{aligned}$$

Eliminating null integrals, $\int_a^a g\,dx = 0$, and cancelling terms reduces this to

$$\begin{aligned}&= \frac{1}{\varepsilon}\left(-\frac{\varepsilon f h'}{h} + \frac{1}{h}\left[\varepsilon\frac{\partial f}{\partial t}h + \varepsilon f\frac{dh}{dt} + \varepsilon h\frac{\partial f}{\partial x}\frac{da}{dt} + \varepsilon^2 h\frac{\partial f}{\partial x}\frac{dh}{dt}\right]\bigg|_{x=a}\right)\\ &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\frac{da}{dt} + \varepsilon\frac{\partial f}{\partial x}\frac{dh}{dt}\bigg|_{x=a}\end{aligned}$$

Hence we obtain the limit as the convective derivative at $x = a(t)$,

$$\lim_{\varepsilon\to 0}\frac{df_{\text{avg}}}{dt} = \frac{\partial f}{\partial t} + v\frac{\partial f}{\partial x}\bigg|_{x=a}$$

In three dimensions it can be shown that

$$f_{\text{avg}}(t) = \frac{\iiint_{D(t)} f\,dV}{\iiint_{D(t)} dV} \qquad \Longrightarrow \qquad \lim_{D(t)\to 0}\frac{df_{\text{avg}}}{dt} = \frac{\partial f}{\partial t} + \mathbf{v}\cdot\nabla f\bigg|_{x=a}.$$

2.2 Expand the left-hand side of the conservation of momentum equation using the product rule

$$\begin{aligned}\frac{\partial(\rho v)}{\partial t} + \frac{\partial(\rho v^2)}{\partial x} &= \rho\frac{\partial v}{\partial t} + v\frac{\partial \rho}{\partial t} + \rho v\frac{\partial v}{\partial x} + v\frac{\partial(\rho v)}{\partial x}\\ &= \rho\frac{\partial v}{\partial t} + \rho v\frac{\partial v}{\partial x} + v\underbrace{\left[\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v)}{\partial x}\right]}_{=0}\\ &= \rho\left(\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x}\right)\end{aligned}$$

where the terms on the second line vanish by virtue of the continuity equation (the conservation of mass).

2.3 (a) Use the trigonometric identity $\cos a \cos b = 2\cos(\frac{1}{2}[a+b])\cos(\frac{1}{2}[a-b])$ to re-write ρ as

$$\rho = 2\cos\left(\tfrac{1}{2}[(k+k+\varepsilon)x - (\omega(k)+\omega(k+\varepsilon))t]x\right)\cos\left(\tfrac{1}{2}[(k+\varepsilon-k)x - (\omega(k+\varepsilon)-\omega(k))t]\right)$$

For $\varepsilon \to 0$ the Taylor series gives $\omega(k+\varepsilon) = \omega(k) + \omega'(k)\varepsilon + \cdots$; using this, for $\varepsilon \to 0$ we get

$$\omega(k+\varepsilon) - \omega(k) \approx \omega'(k)\varepsilon, \qquad \omega(k+\varepsilon) + \omega(k) \approx 2\omega(k)$$

and then the first term in the expansion of ρ is

$$\rho \approx 2\cos(kx - \omega(k)t)\cos\left(\tfrac{1}{2}\varepsilon[x - \omega'(k)t]\right) = 2\cos\left(k\left[x - \frac{\omega(k)}{k}t\right]\right)\cos\left(\tfrac{1}{2}\varepsilon[x-\omega'(k)t]\right),$$

Matching terms to the prescribed form identifies the phase and group velocities as

$$c_p(k) = \frac{\omega(k)}{k}, \qquad c_g(k) = \frac{d\omega}{dk}.$$

(b) Substituting $\rho = \cos(kx - \omega t)$ in the PDE yields

$$\rho_t + \rho_x - \rho_{xxt} = (\omega - k + \omega k^2)\sin(kx - \omega t) = 0$$

and forcing the coefficient to vanish yields the dispersion relation,

$$\omega(k) = \frac{k}{1+k^2}.$$

Similarly, substituting $\rho = \exp(kx - \tilde{\omega}t)$ yields

$$\rho_t + \rho_x - \rho_{xxt} = (-\tilde{\omega} + k + \tilde{\omega}k^2)\exp(kx - \tilde{\omega}t) = 0$$

yielding the modified dispersion relation

$$\tilde{\omega}(k) = \frac{k}{1-k^2}.$$

2.4 (a) Substituting $\rho = P(x - ct)$ in the BBM PDE yields

$$-c\frac{dP}{ds} + \frac{dP}{ds} + 6P\frac{dP}{ds} + c\frac{d^3P}{ds^3} = 0$$

(b) Note that we can integrate the ODE to yield

$$d - cP + P + 3P^2 + c\frac{d^2P}{ds^2} = 0$$

Now considering $P(s) = A\text{sech}^2(Bs)$; for $|s| \to \infty$, $P \to 0$ and similarly all derivatives $P', P'', \ldots \to 0$. Hence, evaluating the ODE for $|s| \to \infty$ we can conclude that the constant of integration is $d = 0$,

$$(1 - c)P + 3P^2 + c\frac{d^2P}{ds^2} = 0$$

Substituting in the sech^2 solution form, after some algebra the ODE gives,

$$(4A + 16cAB^2 - 4cA)\cosh(2Bs) + (24A^2 + 4A - 32cAB^2 - 4cA) = 0$$

The equation is satisfied for all s if the coefficients (in parentheses) are each zero. The first coefficient yields

$$(1 + c) + 4cB^2 = 0 \qquad \Longrightarrow \qquad B = \sqrt{\frac{c-1}{4c}},$$

then the second coefficient is obtained from

$$6A + 1 - 8cB^2 - c = 0 \qquad \Longrightarrow \qquad A = \frac{c-1}{2}.$$

Hence the soliton is

$$\begin{aligned}\rho(x,t) = P(x - ct) &= \frac{c-1}{2}\text{sech}^2\left(\sqrt{\frac{c-1}{4c}}\,(x - ct)\right) \\ &= \frac{c-1}{2}\text{sech}^2\left(\frac{1}{2}\left[\sqrt{\frac{c-1}{c}}\,x - \sqrt{c^2 - c}\,t\right]\right),\end{aligned}$$

where the final line is meant to be of the wavenumber-frequency form, $\rho = G(kx - \tilde{\omega}t)$. To check if this satisfies the modified dispersion relationship found in the previous exercise, consider if

$$k(c) = \sqrt{\frac{c-1}{c}} \qquad \tilde{\omega}(c) = \sqrt{c^2 - c},$$

are parametric equations for

$$\tilde{\omega}(k) = \frac{k}{1-k^2} = \frac{\sqrt{(c-1)/c}}{1-(c-1)/c} = c\sqrt{\frac{c-1}{c}} = \sqrt{c^2-c} = \tilde{\omega}(c),$$

so yes, this is verified. Also, the definition of the phase speed, $c = \tilde{\omega}/k = c$ holds. One of many special properties of solitons is that they satisfy dispersion relations obtained from the linearised version of the PDE while having their form determined by the full, nonlinear equation.

2.5 Using the given forms, the third equation, $\phi_t(x, 1, t) = -f = -A\cos(kx - \omega t)$ reduces to $-\omega B(1) = -A$. Similarly, the fourth equation, $f_t = \phi_y = B'(1)\cos(kx - \omega t)$ yields $\omega A = B'(1)$. The second equation $\phi_y = B'(0)\sin(kx - \omega t)$ reduces to $B'(0) = 0$. Finally, substituting ϕ into $\phi_{xx} + \phi_{yy} = 0$ yields

$$-k^2 B + \frac{d^2 B}{dy^2} = 0 \qquad \Longrightarrow \qquad B(y) = c_1 e^{ky} + c_2 e^{-ky}.$$

Applying the boundary condition $B'(0) = 0$ and $B'(1) = \omega A$ to the general solution of the ODE yields

$$c_1 k - c_2 k = 0, \qquad c_1 k e^k - c_2 k e^{-k} = \omega A \qquad\qquad c_1 = c_2 = \frac{\omega A}{2k\sinh(k)},$$

then $B(1) = A/\omega$ yields the dispersion relation,

$$\frac{\omega A}{k\sinh(k)}\cosh(k) = \frac{\omega A}{k}\tanh(k) = \frac{A}{\omega} \qquad \Longrightarrow \qquad \omega(k) = \sqrt{k\tanh(k)}.$$

2.6 (a) The initial value problems for the characteristic ODEs are

$$\frac{dX}{dt} = e^{2t} \qquad X(0) = A, \qquad \frac{dP}{dt} = P + X + t \qquad P(0) = \cos(A)$$

First solving for X yields $X(t) = A + \frac{1}{2}(e^{2t} - 1)$ which can then be plugged into the equation for P to yield

$$P(t) = (A + \cos(A))e^t - A - t - \tfrac{1}{2}(1 - e^{2t})$$

On each characteristic curve, we can invert $X(t, A)$ to get $A = x + \frac{1}{2}[1 - e^{2t}]$, then substituting into P yields

$$\rho(x,t) = \left(x + \tfrac{1}{2}[1-e^{2t}] + \cos(x + \tfrac{1}{2}[1-e^{2t}])\right)e^t - \left(x + \tfrac{1}{2}[1-e^{2t}]\right) - t - \tfrac{1}{2}(1-e^{2t}).$$

(b) The characteristic ODEs are

$$\frac{dX}{dt} = X + 4, \qquad \frac{dP}{dt} = -2P.$$

The general solutions of these equations are

$$X(t) = c_1 e^t - 4, \qquad P(t) = c_2 e^{-2t}.$$

The side conditions give data in two parts.

For $A > 0$, we have $X(0) = A$, $P(0) = e^{-A}$ at $t = 0$. Applying this to the general solution yields

$$X(t) = (A + 4)e^t - 4, \qquad P(t) = e^{-A-2t}.$$

Inverting $x = X(t, A)$ yields $A = (x + 4)e^{-t} - 4 > 0$ and substituting this into $P(t)$ yields

$$\rho(x, t) = \exp\left(4 - (x + 4)e^{-t} - 2t\right) \qquad \text{for } x > 4(e^t - 1).$$

The second part of the data is given at $x = 0 \implies X(T) = 0$ with $P(T) = \cos(T)$ for $t > 0 \implies T > 0$. Applying these conditions to the general solution yields

$$X(t) = 4e^{t-T} - 4, \qquad P(t) = \cos(T)e^{-2t+2T}.$$

Inverting $x = X(t, T)$ yields $T = t - \ln(1 + x/4) > 0$ and substituting this into $P(t)$ yields

$$\rho(x, t) = \cos\left(t - \ln\left(1 + \frac{x}{4}\right)\right)\left(1 + \frac{x}{4}\right)^{-2} \qquad \text{for } x < 4(e^t - 1).$$

2.7 (a) Expanding out the product rule and using the given velocity yields the PDE

$$\frac{\partial \rho}{\partial t} + x^2 e^{-3t}\frac{\partial \rho}{\partial x} = -2xe^{-3t}\rho$$

and consequently the characteristic ODEs

$$\frac{dX}{dt} = X^2 e^{-3t}, \qquad X(0) = A, \qquad \text{for } 1 \le A \le 2.$$

and

$$\frac{dP}{dt} = -2XPe^{-3t}, \qquad P(0) = 1.$$

The problem for $X(t)$ yields

$$X(t, A) = \left(\frac{1}{A} + \tfrac{1}{3}e^{-3t} - \tfrac{1}{3}\right)^{-1},$$

and the problem for $P(t)$ yields

$$P(t, A) = \frac{A^2}{9}\left(e^{-3t} - 1 + \frac{3}{A}\right)^2.$$

Inverting $x = X(t, A)$ for A yields

$$A = \left(\frac{1}{x} + \tfrac{1}{3} - \tfrac{1}{3}e^{-3t}\right)^{-1},$$

and substituting this into $P(t, A)$ yields

$$\rho(x, t) = \left(\frac{x}{3}[e^{-3t} - 1] - 1\right)^{-2}.$$

2.10 (a) Substituting in the definitions of p, q in the first equation yields $\phi_{tt} - c^2 \phi_{xx} = 0$ (the wave equation) and $\phi_{xt} - \phi_{tx} = 0$, always true by the identity of mixed partial derivatives of smooth functions.
(b) The system can be written as

$$\frac{\partial}{\partial t}\begin{pmatrix} p \\ q \end{pmatrix} + \underbrace{\begin{pmatrix} 0 & -c^2 \\ -1 & 0 \end{pmatrix}}_{\mathbf{M}} \frac{\partial}{\partial x}\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Following the approach given in Sect. 2.4, we obtain the wavespeeds from

$$|\mathbf{M}^T - \lambda \mathbf{I}| = \begin{vmatrix} -\lambda & -1 \\ -c^2 & -\lambda \end{vmatrix} = \lambda^2 - c^2 = 0 \qquad \Longrightarrow \qquad \lambda = \pm c.$$

The corresponding eigenvectors are then

$$\begin{aligned} \lambda_1 &= c \qquad W_1(x - ct) = p(x, t) - cq(x, t) \\ \lambda_2 &= -c \qquad W_2(x + ct) = p(x, t) + cq(x, t) \end{aligned}$$

(b) Then the solutions can be expressed in terms of W_1, W_2 are

$$p(x, t) = \frac{1}{2}\left(W_1(x - ct) + W_2(x + ct)\right) \qquad q(x, t) = \frac{1}{2c}\left(W_2(x + ct) - W_1(x - ct)\right).$$

(c) We can relate A, B to W_1, W_2 from

$$p = \frac{\partial \phi}{\partial t} = \frac{1}{2}\left(W_1(x-ct) + W_2(x+ct)\right) = -cA'(x-ct) + cB'(x+ct)$$
$$q = \frac{\partial \phi}{\partial x} = \frac{1}{2c}\left(W_2(x+ct) - W_1(x-ct)\right) = A'(x-ct) + B'(x+ct)$$

Both of which yield

$$A'(x-ct) = -\frac{1}{2c}W_1(x-ct) \qquad B'(x+ct) = \frac{1}{2c}W_2(x+ct).$$

(d) At $t = 0$ the initial condition $\phi(x,0) = f(x)$ implies $\phi_x(x,0) = q(x,0) = f'(x)$ and $\phi_t(x,0) = p(x,0) = g(x)$, and consequently

$$W_1(x) = g(x) - cf'(x) \qquad W_2(x) = g(x) + cf'(x)$$

Hence, from part (c)

$$A'(x) = -\frac{1}{2c}\left(g(x) - cf'(x)\right)$$
$$A(x) = \int_0^x A'(\tilde{x})\,d\tilde{x} + a = \frac{f(x)-f(0)}{2} - \frac{1}{2c}\int_0^x g(\tilde{x})\,d\tilde{x} + a$$
$$B'(x) = \frac{1}{2c}\left(g(x) + cf'(x)\right)$$
$$B(x) = \int_0^x B'(\tilde{x})\,d\tilde{x} + b = \frac{f(x)-f(0)}{2} + \frac{1}{2c}\int_0^x g(\tilde{x})\,d\tilde{x} + b,$$

where a, b are constants of integration.

So using $\phi(x,t) = A(x-ct) + B(x+ct)$,

$$\begin{aligned}\phi &= \tfrac{1}{2}\left(f(x-ct) - f(0)\right) - \frac{1}{2c}\int_0^{x-ct} g(\tilde{x})\,d\tilde{x} + a \\ &\quad + \tfrac{1}{2}\left(f(x+ct) - f(0)\right) + \frac{1}{2c}\int_0^{x+ct} g(\tilde{x})\,d\tilde{x} + b \\ &= \tfrac{1}{2}\left(f(x+ct) + f(x-ct)\right) + \frac{1}{2c}\int_{x-ct}^{x+ct} g(\tilde{x})\,d\tilde{x} + (a+b-f(0)).\end{aligned}$$

Checking the initial condition, $\phi(x,0) = f(x)$,

$$\phi(x,0) = \tfrac{1}{2}(f(x) + f(x)) + \frac{1}{2c}\underbrace{\int_x^x g(\tilde{x})\,d\tilde{x}}_{=0} + (a+b-f(0)),$$

so the constants of integration should be chosen so that $a + b - f(0) = 0$, yielding the D'Alembert solution.

2.12 The conservation law $p_t + q(p)_x = 0$ can be re-written as $p_t + q'(p)p_x = 0$, from which we can write the characteristic equations as

$$\frac{dP}{dt} = 0 \qquad \Longrightarrow \qquad P = f(A) = \text{constant}$$
$$\frac{dX}{dt} = q'(P) \qquad \Longrightarrow \qquad X = q'(f(A))t + A$$

(a) The two characteristic curves will intersect when $X(t_{0,1}, A_0) = X(t_{0,1}, A_1)$, namely

$$q'(f(A_0))t_{0,1} + A_0 = q'(f(A_1))t_{0,1} + A_1 \qquad \Longrightarrow \qquad t_{0,1} = -\frac{A_1 - A_0}{q'(f(A_1)) - q'(f(A_0))}$$

and subsequently using this value of $t_{0,1}$, $x_{0,1} = q'(f(A_0))t_{0,1} + A_0$ (or equivalently in terms of A_1).
(b) Minimising $t_{0,1}$ over all possible A_0, A_1 to get t_*, consider $A_1 = A_0 + \varepsilon$,

$$t_* = \min_{A_0, A_1} -\frac{A_1 - A_0}{q'(f(A_1)) - q'(f(A_0))} = \min_{A_0, \varepsilon} -\frac{\varepsilon}{q'(f(A_0 + \varepsilon)) - q'(f(A_0))}$$

Taking the limit $\varepsilon \to 0$ and recalling the limit definition of the derivative,

$$t_* = \min_x \left[-\frac{1}{q''(f(x))f'(x)} \right] \geq 0.$$

2.13 (a) The characteristic ODEs are $dX/dt = P$, $dP/dt = 0$, as in (2.41). Applying the given initial conditions yields the parametric solutions

$$\begin{cases} X(t, A) = (9 - A^2)t + A, & P(t, A) = 9 - A^2 & |A| \leq 3 \\ X(t, A) = A, & P(t, A) = 0 & |A| > 3 \end{cases}$$

(b) To invert $X(t, A) = x$ for the nontrivial part of the solution, we recognise it as a quadratic equation for A,

$$A^2 t - A + (x - 9t) = 0 \qquad \Longrightarrow \qquad A = \frac{1 \pm \sqrt{1 - 4(xt - 9t^2)}}{2t}$$

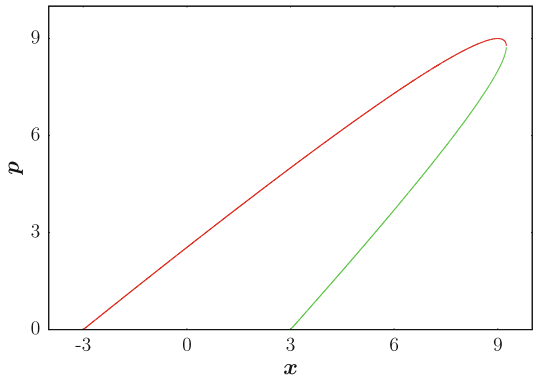
then we can substitute into P to obtain

$$p(x,t) = 9 - \left(\frac{1 \pm \sqrt{1 - 4(xt - 9t^2)}}{2t} \right)^2$$

For more general initial conditions, $p(x,0) = f(x)$, for the inviscid Burgers' equation we have $X = Pt + A$ with P on a characteristic curve, so we can still write $A = x - Pt$ to obtain an implicit equation for the solution, $p = f(x - pt)$, here $p = 9 - (x - pt)^2$, to yield the equivalent form

$$p(x,t) = \frac{2xt - 1 \pm \sqrt{36t^2 - 4xt + 1}}{2t^2}$$

See below for a graph of this multi-valued solution at time $t = 1$.



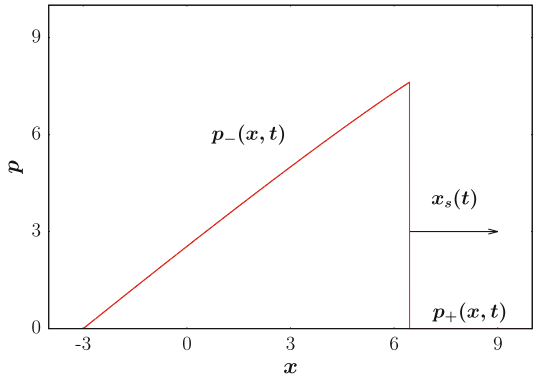
(c) The result of Exercise 2.12 gives that the time at which the shock forms is determined by the initial conditions,

$$t_* = \min_x \left(-\frac{1}{f'(x)} \right) \qquad \Longrightarrow \qquad t_* = \min_{x \in [-3,3]} \left(-\frac{1}{2x} \right) = \frac{1}{6},$$

and the shock will first form at $x_* = 3$.
(d) The Rankine-Hugoniot equation for the inviscid Burgers' equation (2.49) and the result from part (c) give the ODE problem for the shock position as

$$\frac{dx_s}{dt} = -\frac{2x_s t - 1 + \sqrt{36t^2 - 4x_s t + 1}}{4t^2}, \qquad x_s(1/6) = 3,$$

where the solution ahead of the shock is $p_+(x,t) \equiv 0$ and the solution behind the shock, $p_-(x,t)$ is given by the solution above with the positive sign on the square root.



(e) The inviscid Burgers equation is a conservation law for the integral, $\int p\,dx$,

$$p_t + \left(\tfrac{1}{2}p^2\right)_x = 0 \qquad \Longrightarrow \qquad \frac{d}{dt}\left(\int p\,dx\right) = 0$$

Hence the conserved value is set by the area under the initial condition $\int (9-x^2)\,dx = 36$. Using the form of the solution with the shock, the equation that $\int p(x,t)\,dx = 36$ becomes (for $t > 1/6$, after the shock forms),

$$\int_{-3}^{x_s(t)} \frac{2xt - 1 + \sqrt{36t^2 - 4xt + 1}}{2t^2}\,dx = 36.$$

Evaluating the integral yields an implicit algebraic equation for $x_s(t)$,

$$\frac{1}{12t^3}\left(6x_s^2t^2 - (1 - 4x_st + 36t^2)^{3/2} - 6x_st + 216t^3 + 54t^2 + 1\right) = 36.$$

Chapter 3

3.1 (a, b) Both approaches give

$$-y'' + \frac{k}{x^2}y + x^2 = 0.$$

Multiplying across by x^2, this equation can be recognised as an inhomogeneous Cauchy-Euler equation, $x^2y'' - ky = -x^4$, solvable in terms of a sum of homogeneous and particular solutions.
(c) For $k = 0$ the general solution is $y(x) = x^4/12 + c_1x + c_2$. Applying the boundary conditions yields $c_1 = -1/12$, $c_2 = 1$.
(d) For $k = 2$ the general solution is $y(x) = x^4/10 + c_1x^2 + c_2/x$. Applying the first boundary conditions yields $c_1 = 9/10$, $c_2 = 0$. The condition $y(0) = 1$ cannot be satisfied because unless $c_2 = 0$ that term would diverge and the other terms in the general solution vanish for $x = 0$.

3.2

$$\delta^2 J = \tfrac{1}{2}\int_0^1 \frac{(h')^2}{(1+(y')^2)^{3/2}}\,dx \qquad \Longrightarrow \qquad \delta^2 J_* = \tfrac{1}{2}\int_0^1 \frac{(h')^2}{(1+b^2)^{3/2}}\,dx \geq 0,$$

where $\delta^2 J_*$ will be strictly positive unless $h'(x) \equiv 0$, but from the boundary conditions on y, the boundary conditions on h are $h(0) = h(1) = 0$. Hence $h' \equiv 0$ only if $h \equiv 0$.

3.4 (a) The $x(t)$ Euler-Lagrange equation can be simplified to

$$(y''x'y' - x''(y')^2)(1 + 4k^2(x^2+y^2)) + 4k^2((x')^2 + (y')^2)(x'y - xy')y' = 0.$$

The y Euler-Lagrange equation takes the same form after interchanging $x \leftrightarrow y$.
(b) The Euler-Lagrange equation for $y(x)$ can be reduced to

$$y'' = \frac{4k^2(1+(y')^2)(xy'-y)}{1+4k^2(x^2+y^2)}.$$

(c) For (ii), since z is constant on the semicircle, the distance is just the arclength of the semi-circle, $J_{ii} = \pi$ for all k.

For (i), we have $y = 0$ and $z = kx^2$, hence the arclength can be calculated from

$$J_i(k) = \int_{-1}^{1} \sqrt{1+4k^2x^2}\,dx = \sqrt{1+4k^2} + \frac{1}{2k}\operatorname{arcsinh}(2k).$$

For $k = 0$, $J_i(0) = 2$ giving the length of the straight line in the xy plane, while for $k > 0$ J_i is an unbounded monotone increasing function of k, $J_i(k \to \infty) \to \infty$. Let k_* be the value of k for which $J_i(k_*) = \pi$. Then for $0 \leq k < k_*$, the straight-line path will be shorter than the semi-circle detours. But for $k > k_*$, the semi-circle is shorter.

3.5 (a) After the substitution, the action in terms of $\theta(t)$ is

$$I = \int \tfrac{1}{2}m\ell^2(\theta')^2 + mg\ell\cos\theta\,dt$$

and from (3.21) applied to $\theta(t)$ the resulting Euler-Lagrange is the equation of the pendulum,

$$\theta'' + \frac{g}{\ell}\sin\theta = 0.$$

(b) See Exercise 3.27 for the Euler-Lagrange equation for $x(t)$.

3.6 Substituting the parametric equation for x, y into the general Lagrangian, $L = \frac{1}{2}m((x')^2 + (y')^2) - mgy$ yields

$$L = \tfrac{1}{2}m\ell^2(\theta')^2 + mg\ell\cos\theta - \sigma m\left(\ell\omega\theta'\cos(\omega t)\sin\theta - g\sin(\omega t)\right) + \tfrac{1}{2}m\sigma^2\omega^2\cos^2(\omega t)$$

Applying (3.21) then yields the (vertically-oscillated) parametrically-driven pendulum equation,

$$\theta'' + \frac{g}{\ell}\sin\theta = -\frac{\omega^2\sigma}{\ell}\sin(\omega t)\sin\theta.$$

3.7 (a) Starting from $H = -L + y'\partial_{y'}L$,

$$\begin{aligned}\frac{dH}{dt} &= \frac{d}{dt}\left(-L + y'\frac{\partial L}{\partial y'}\right)\\ &= -\left(\frac{\partial L}{\partial t} + \frac{\partial L}{\partial y}\frac{dy}{dt} + \frac{\partial L}{\partial y'}\frac{d^2y}{dt^2}\right) + \frac{d^2y}{dt^2}\frac{\partial L}{\partial y'} + y'\frac{d}{dt}\left(\frac{\partial L}{\partial y'}\right)\\ &= -\frac{\partial L}{\partial t} - y'\left[\frac{\partial L}{\partial y} - \frac{d}{dt}\left(\frac{\partial L}{\partial y'}\right)\right] = 0,\end{aligned}$$

where the final step makes use of the assumption that $\partial_t L = 0$ to leave $-y'$ times the Euler-Lagrange equation. Hence $H' = 0$ and H is a constant, justifying (3.59).
(b) Writing the Lagrangian as $L = T(t, y, y') - V(t, y)$ the condition that the Hamiltonian is the total energy is

$$H = -T + V + y'\frac{\partial T}{\partial y'} = T + V \qquad \Longrightarrow \qquad y'\frac{\partial T}{\partial y'} = 2T$$

This is a first order separable ODE for T as a function of y' and yields $T(t, y, y') = A(t, y)(y')^2$ where $A(t, y)$ is any function of t and y.

3.8 (a) At $t = 0$, the mass starts at $x = 0$, $y = 1$ from rest, $v = 0$, hence the initial energy is $E_0 = mg$. Equating the energy at later times with E_0 yields

$$\tfrac{1}{2}mv^2 + mgy = mg \qquad \Longrightarrow \qquad v(y) = \sqrt{2g(1-y)}$$

Consequently the functional for the time of travel is

$$T = \int_0^1 \sqrt{\frac{1 + (y')^2}{2g(1-y)}}\,dx \qquad L(y, y') = \sqrt{\frac{1 + (y')^2}{2g(1-y)}}$$

(b) The Euler-Lagrange equation for $y(x)$ can be simplified down to

$$\tfrac{1}{2}\left(\frac{1+(y')^2}{(1-y)^3}\right)^{1/2} - \frac{d}{dx}\left(\frac{y'}{\sqrt{(1+(y')^2)(1-y)}}\right) = 0,$$

but this is a difficult looking equation, so we turn to a different approach.
(c) The Beltrami identity (3.59) is applicable since L does not explicitly depend on x, so

$$\begin{aligned} H &= L - y'\frac{\partial L}{\partial y'} \\ &= \sqrt{\frac{1+(y')^2}{2g(1-y)}} - \frac{(y')^2}{\sqrt{2g(1+(y')^2)(1-y)}} \\ &= \frac{1}{\sqrt{2g(1+(y')^2)(1-y)}} = C \end{aligned}$$

At $x = 0$, $y = 1$ and to make C finite, the slope would have to diverge, $y' \to -\infty$, so this form is indeterminate. At $x = 1$, we have $y = 0$, so $C = 1/\sqrt{2g(1+y'(1)^2)}$, but we don't know $y'(1)$. Still, we can write a singular ODE problem for $y(x)$,

$$\frac{dy}{dx} = -\sqrt{\frac{1}{2gC^2(1-y)} - 1}, \qquad y(0) = 1, \qquad y(1) = 0.$$

(d) Substituting the parametric equations of the cycloid (3.60) into the ODE gives a necessary relation between C and k, $1 = 4gC^2k$. The initial conditions at $\theta = 0$, $x(0) = 0$ and $y(0) = 1$ are satisfied automatically for any k. The final state, $x = 1$ and $y = 0$ at some $\theta = \theta_*$ yield two coupled equations for finding (k, θ_*):

$$k(\theta_* - \sin\theta_*) = 0 \qquad 1 - k(1 - \cos\theta_*) = 0.$$

3.10 Following the approach given in Sect. 3.4.2 with $f(x) = 1+(x-1)^2$, $\tilde{y} = y_* + \varepsilon h$ and $\tilde{b} = b_* + \varepsilon c$, we obtain the first variation as

$$\delta J_* = y_*' h\Big|_0^{b_*} - \int_0^{b_*} y_*'' h\, dx + \tfrac{1}{2} y_*'(b)^2 c.$$

The boundary condition at the origin determines that $h(0) = 0$. Expanding out the boundary condition $\tilde{y}(\tilde{b}) = f(\tilde{b})$ yields

$$y_*(b_*) = f(b_*), \qquad y_*'(b_*)c + h(b_*) = f'(b_*)c$$

consequently, $c = h(b_*)/(f'(b_*) - y'_*(b_*))$. Using this, the first variation can be re-written as

$$\delta J_* = y'_*(b_*)\left(1 + \frac{y'_*(b_*)}{2(f'(b_*) - y'_*(b))}\right)h(b_*) - \int_0^{b_*} y''_* h\,dx.$$

There are two possible choices for natural boundary conditions that would make the boundary term vanish for all possible choices of $h(b)$: either $y'_*(b_*) = 0$ or (after some algebra) $y'_*(b_*) = 2f'(b_*)$ (we will explore both possibilities). Subject to natural boundary conditions, applying the fundamental lemma to the critical point condition $\delta J_* = 0$, we obtain the Euler-Lagrange equation $y''_* = 0$. After applying the boundary condition $y(0) = 0$, viable solutions are of the form $y_*(x) = Ax$. The boundary condition $y'_*(b_*) = 0$ would yield $A = 0$, but this option must be rejected since the solution would not reach the curve $y = f(x)$. Hence, using $y'_*(b_*) = 2f'(b_*)$, we get $A = 4(b_* - 1)$ and then $y_*(b_*) = f(b_*)$ yields the quadratic equation, $3b_*^2 - 2b_* - 2 = 0$ and the solution

$$y_*(x) = \frac{4}{3}(\sqrt{7} - 2)x \qquad 0 \le x \le \frac{1 + \sqrt{7}}{3}.$$

3.12 Beginning by substituting $\tilde{y} = y_* + \varepsilon h$ and expanding for $\varepsilon \to 0$ yields

$$\tilde{J} = J_* + \varepsilon \int_0^1 h' y'''_* + y'_* h''' - 240xh\,dx + \cdots$$

Using integration by parts to get the derivatives off of the h's, we arrive at a form where the fundamental lemma can be applied to critical point condition, yielding the ODE,

$$-2y'''_* - 240x = 0.$$

Note that all of the boundary terms vanish thanks to

$$y'(0) \text{ given} \implies h'(0) = 0 \qquad y''(1) \text{ given} \implies h''(1) = 0$$

$$y'''(0) \text{ given} \implies h'''(0) = 0 \qquad y(1) \text{ given} \implies h(1) = 0$$

The ODE can be integrated directly to give the solution as a polynomial, after applying the four boundary conditions, the final solution is

$$y_*(x) = -x^5 + 10x^2 - 4.$$

3.13 Beginning by substituting $\tilde{y} = y_* + \varepsilon h$ and expanding for $\varepsilon \to 0$ yields

$$\tilde{J} = J_* + \varepsilon \int_0 \left[2y_*' h' + (1-2x) \int_0^x 2y_*(t)h(t)dt \right] dx + \cdots$$

Applying integration by parts yields the first variation as

$$\left(2y_*' h + (x - x^2) \int_0^x 2y_* h \, dt \right) \Big|_0^1 - \int_0^1 \left(2y_*'' + 2(x-x^2)y_* \right) h dx$$

The second boundary term vanishes automatically at $x = 0$ and $x = 1$ due to the $x(1-x)$ factor. The first boundary term vanishes under the natural boundary conditions

$$2y_*'(0)h(0) = 0 \qquad \Longrightarrow \qquad \{y_*'(0) = 0 \quad \text{or} \quad h(0) = 0 \implies y_*(0) = A\}$$

and

$$2y_*'(1)h(1) = 1 \qquad \Longrightarrow \qquad \{y_*'(1) = 0 \quad \text{or} \quad h(1) = 0 \implies y_*(1) = B\}$$

3.14 Beginning by substituting $\tilde{y} = y_* + \varepsilon h$ and expanding for $\varepsilon \to 0$ yields

$$\tilde{J} = J_* + \varepsilon \left(\int_0^1 2y_* y_*'' h'' + (y_*'')^2 h + 2y_* h \, dx + y_*'(0)h'(1) + y_*'(1)h'(0) \right) + \cdots$$

Integrating by parts yields the first variation as

$$\left(2y_*(1)y_*''(1) + y_*'(0) \right) h'(1) + \left(-2y_*(0)y_*''(0) + y_*'(1) \right) h'(0) + \int_0^1 \left(2(y_* y_*'')'' + (y_*'')^2 + 2y \right) h \, dx$$

Hence the Euler-Lagrange equation is

$$\left(\frac{d^2 y_*}{dx^2} \right)^2 + 2 \frac{d^2}{dx^2} \left(y_* \frac{d^2 y_*}{dx^2} \right) + 2y_* = 0$$

and the further natural boundary conditions needed are

$$\{2y_*(1)y_*''(1) + y_*'(0) = 0 \quad \text{or} \quad h'(1) = 0 \implies y_*'(1) = A\}$$

and

$$\{-2y_*(0)y_*''(0) + y_*'(1) = 0 \quad \text{or} \quad h'(0) = 0 \implies y_*'(0) = B\}$$

3.15 (a) Using $v = c/n(x)$, functional (3.6) becomes

$$T(y) = \frac{1}{c}\int_0^1 n(x)\sqrt{1+(y')^2}\,dx \qquad \Longrightarrow \qquad L(x,y') = \frac{n(x)}{c}\sqrt{1+(y')^2}$$

Applying (3.21) to this Lagrangian yields

$$0 - \frac{d}{dx}\left(\frac{n(x)y'}{\sqrt{1+(y')^2}}\right) = 0,$$

which can be integrated once, and applying the initial condition yields

$$\frac{n(x)y'}{\sqrt{1+(y')^2}} = C = \frac{n(0)}{\sqrt{2}} \qquad \Longrightarrow \qquad \frac{dy}{dx} = \frac{n(0)}{\sqrt{2n(x)^2 - n(0)^2}}$$

(b) Yes, $n_1 \sin\theta_1 = n_2 \sin\theta_2$.
(c) $L(x, y, y') = n(x, y)\sqrt{1+(y')^2}/c$ yields the Euler-Lagrange equation

$$\sqrt{1+(y')^2}\,\frac{\partial n}{\partial y} - \frac{d}{dx}\left(\frac{n(x,y)y'}{\sqrt{1+(y')^2}}\right) = 0.$$

3.16 (a) The total kinetic and potential energies are given by

$$T = \int_0^\ell \tfrac{1}{2}\rho u_t^2\,dx \qquad V = \int_0^\ell \tfrac{1}{2}EIu_x^2\,dx$$

(b) $L = T - V$ then the action is

$$J = \int_{t_0}^{t_1} L\,dt = \int_{t_0}^{t_1}\int_0^\ell \left(\tfrac{1}{2}\rho u_t^2 - \tfrac{1}{2}EIu_x^2\right)\,dx\,dt$$

(c) Substituting $\tilde{u}(x,t) = u_*(x,t) + \varepsilon h(x,t)$ and expanding for $\varepsilon \to 0$ yields

$$\tilde{J} = J_* + \varepsilon\int_{t_0}^{t_1}\int_0^\ell (\rho u_{*t}h_t - EIu_{*xx}h_{xx})\,dx\,dt + \cdots$$

Splitting the $O(\varepsilon)$ term into separate integrals, we interchange the order of integration on the first,

$$\int_0^\ell\int_{t_0}^{t_1} \rho u_{*t}h_t\,dt\,dx - \int_{t_0}^{t_1}\int_0^\ell EIu_{*xx}h_{xx}\,dx\,dt$$

Applying integration by parts with respect to the inner integrals respectively yields

$$\rho\left(\int_0^\ell \left[u_{*t}h\Big|_{t_0}^{t_1}\right] dx - \int_{t_0}^{t_1}\int_0^\ell u_{*tt}h\,dx\,dt\right)$$
$$+ EI\left(\int_{t_0}^{t_1}\left[u_{*xx}h_x - u_{*xxx}h\Big|_0^\ell\right] dt - \int_{t_0}^{t_1}\int_0^\ell u_{*xxxx}h\,dx\,dt\right)$$

The perturbation in the solution at initial and final times $t = t_0, t_1$ can be assumed to be zero $h(x, t_0) \equiv h(x, t_1) \equiv 0$ to eliminate the first boundary terms. Eliminating the second boundary term defines natural boundary conditions to remove the $u_{*xx}h_x$ and $u_{*xxx}h$ terms at each time,

$$\{u_{*xx}(0,t) = 0 \quad \text{or} \quad h_x(0) = 0 \quad \Longrightarrow \quad u_{*x}(0,t) = A(t)\}$$

and

$$\{u_{*xxx}(0,t) = 0 \quad \text{or} \quad h(0) = 0 \quad \Longrightarrow \quad u_*(0,t) = B(t)\}$$

and similarly for the boundary conditions at $x = 1$.

Applying the fundamental lemma to $\iint(\rho u_{*tt} + EIu_{*xxxx})h\,dx\,dt = 0$ yields the beam equation PDE as the Euler-Lagrange equation,

$$\rho\frac{\partial^2 u}{\partial t^2} + EI\frac{\partial^4 u}{\partial x^4} = 0.$$

3.17 Substituting $\tilde{u}(x, y) = u_*(x, y) + \varepsilon h(x, y)$ and expanding for $\varepsilon \to 0$ yields

$$\tilde{J} = J_* + \varepsilon\iint_D k\left(u_{*x}h_x + u_{*y}h_y\right) dA + \cdots$$

The first variation can be written in vector form as $\iint k\nabla u_* \cdot \nabla h\,dA$ then matching to the product rule with $\mathbf{g} = k\nabla u_*$ and $f = h$, it can be expressed as

$$\delta J_* = \iint_D \nabla\cdot(hk\nabla u_*)\,dA - \iint_D \nabla\cdot(k\nabla u_*)h\,dA$$

Applying the divergence theorem to the first integral changes it to an integral on the boundary

$$\delta J_* = \oint_{\partial D} h(k\mathbf{n}\cdot\nabla u_*)\,ds - \iint_D \nabla\cdot(k\nabla u_*)h\,dA$$

Eliminating the boundary integral leads to two choices for natural boundary conditions: (i) homogeneous Neumann conditions, $\mathbf{n} \cdot \nabla u = \frac{\partial u}{\partial n} = 0$, or (ii) specified Dirichlet conditions, $h(\partial D) = 0 \implies u(\partial D) = f(\partial D)$ given. Subsequently, the fundamental lemma can be applied to the remaining double integral to yield the elliptic PDE

$$\nabla \cdot (k \nabla u) = 0.$$

3.19 (a) The augmented Lagrangian for this problem is

$$\mathscr{L} = 1 + (y')^2 - \lambda \left(y^2 - \tfrac{80}{\pi} \right),$$

then the Euler-Lagrange problem is

$$\frac{d^2y}{dx^2} + \lambda y = 0, \qquad y(0) = 0, \qquad y(\pi) = 1, \qquad \int_0^\pi y^2 \, dx = 80.$$

The ODE is a linear-constant coefficient equation; it breaks down into two cases, depending on the value of $\lambda \gtrless 0$.

If $\lambda = -\alpha^2 < 0$ then the solution of the boundary value problem for the ODE is

$$y_-(x) = \frac{\sinh(\alpha x)}{\sinh(\alpha \pi)}.$$

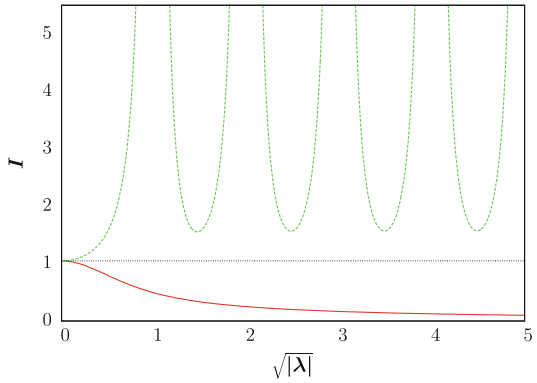
If $\lambda = \alpha^2 > 0$ then the solution is

$$y_+(x) = \frac{\sin(\alpha x)}{\sin(\alpha \pi)}.$$

In the two cases, the integral becomes

$$I_\pm(\alpha) = \int_0^\pi y_\pm^2 \, dx = \left\{ \frac{2\alpha\pi - \sin(2\alpha\pi)}{4\alpha \sin^2(\alpha\pi)}, \ \frac{\sinh(2\alpha\pi) - 2\alpha\pi}{4\alpha \sinh^2(\alpha\pi)} \right\}$$

$I_+(\alpha) \geq \pi/3$ and has multiple solutions for any value of the constraint greater than $\pi/3 \approx 1.047$. Meanwhile $0 \leq I_-(\alpha) \leq \pi/3$ and is monotone decreasing, so there is a single solution for each value of the constraint.



3.20 (a) Recall Green's theorem,

$$\oint_{\partial D} P(x,y)\,dx + Q(x,y)\,dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA$$

We have

$$\frac{1}{2}\int_0^1 \left[x(t)y'(t) - y(t)x'(t)\right] dt = \oint \tfrac{1}{2}x\,dy - \tfrac{1}{2}y\,dx.$$

We can identify $P = -\frac{1}{2}y$, $Q = \frac{1}{2}x$, therefore,

$$\frac{1}{2}\int_0^1 \left[x(t)y'(t) - y(t)x'(t)\right] dt = \iint \left(\tfrac{1}{2} + \tfrac{1}{2}\right) dA = \text{Area}.$$

(b) The Euler-Lagrange equations are

$$\frac{d}{dt}\left(y + \frac{\lambda x'}{\sqrt{(x')^2 + (y')^2}}\right) = 0 \qquad \frac{d}{dt}\left(-x + \frac{\lambda y'}{\sqrt{(x')^2 + (y')^2}}\right) = 0$$

(c) Integrating once yields

$$y + \frac{\lambda x'}{\sqrt{(x')^2 + (y')^2}} = c_1 \qquad -x + \frac{\lambda y'}{\sqrt{(x')^2 + (y')^2}} = c_2$$

and can be re-arranged to yield

$$(c_1 - y)^2 + (c_2 + x)^2 = \lambda^2 \left(\frac{(x')^2 + (y')^2}{(x')^2 + (y')^2}\right)^2 = \lambda^2$$

Hence we have the equation of a circle with radius $\lambda = P/(2\pi)$ and centre $(-c_2, c_1)$. Having the origin on the circle means that $c_1^2 + c_2^2 = \lambda^2$, hence we can take $c_1 = \lambda \sin\theta$, $c_2 = \lambda\cos\theta$ for any θ.

3.23 (a) The augmented functional is

$$I = \int_1^2 \left[6y^2 + x^2 \left(\frac{dy}{dx}\right)^2 + x^7 - \lambda(24xy - 5) \right] dx.$$

(b) Substituting $\tilde{y}(x) = y_*(x) + \varepsilon h(x)$ and expanding for $\varepsilon \to 0$ yields the first variation as

$$\delta I = \int_1^2 12y_* h + 2x^2 y_*' h' - 24xh\, dx$$

Examining the boundary condition $\tilde{y}(2) = \tilde{y}(1) + 3$ gives that $h(2) = h(1)$, and applied to the boundary terms produced by integration by parts, $2x^2 y_*' h|_1^2 = 2(4y_*'(2) - y_*'(1))h(1)$, hence we determine the natural boundary condition

$$4y_*'(2) - y_*'(1) = 0$$

and then the Euler-Lagrange equation is

$$12y - (2x^2y')' - 24\lambda x = 0 \qquad \Longrightarrow \qquad x^2y'' + 2xy' - 6y = -12\lambda x.$$

This is an inhomogeneous Cauchy-Euler equation with solution

$$y = c_1 x^2 + c_2 x^{-3} + 3\lambda x$$

(c) Substituting into the boundary condition, the natural boundary condition and the integral constraint yields three equations for c_1, c_2, λ,

$$3c_1 - \frac{7}{8}c_2 + 3\lambda = 3, \qquad 14c_1 + \frac{9}{4}c_2 + 9\lambda = 0, \qquad 90c_1 + 12c_2 + 168\lambda = 5.$$

3.24 The augmented functional for the constrained problem is

$$I = \int_1^2 \frac{3y^2}{x^5} - \frac{(y')^2}{x^3} - \lambda(y+3)\, dx$$

The resulting Euler-Lagrange equation for $y = y_*(x)$ is

$$x^2y'' - 3xy' + 3y = \tfrac{1}{2}\lambda x^5.$$

The solution of this inhomogeneous Cauchy-Euler equation is

$$y = c_1 x + c_2 x^3 + \frac{\lambda}{16} x^5.$$

Applying the boundary conditions and the integral constraint determine $c_1 = 11$, $c_2 = -8, \lambda = 16$.

3.26 (a) The Hamiltonian is

$$\mathscr{H} = L + \lambda f = 4x^2 + 3xu + u^2 + \lambda(3x + u).$$

The PMP yields

$$\frac{dx}{dt} = 3x + u, \qquad \frac{d\lambda}{dt} = -8x - 3u - 3\lambda, \qquad 3x + 2u + \lambda = 0.$$

Eliminating λ reduces the problem to a system of two linear ODEs in x, u,

$$\frac{dx}{dt} = 3x + u \qquad \frac{du}{dt} = -5x - 3u$$

having the general solution

$$x(t) = c_1 e^{2t} + c_2 e^{-2t} \qquad u(t) = -c_1 e^{2t} - 5c_2 e^{-2t}$$

Substituting these into the Hamiltonian yields $\mathscr{H} = 16 c_1 c_2$. The Hamiltonian is indeed a constant, and to enforce $\mathscr{H} = 0$ we need either $c_1 = 0$ or $c_2 = 0$. The possibility of $c_1 = c_2 = 0$ can be excluded because it would give the trivial solution and could not satisfy the initial condition $x(0) = 2$. Consider $c_2 = 0$, then $c_1 = 2$; this yields a monotone increasing function for $x(t) \geq 2$ and could never satisfy the target condition $x(T) = 1$. Hence $c_1 = 0$ yielding $x(t) = 2e^{-2t}$ and the target condition determines $T_* = \frac{1}{2} \ln 2$.

(b) Everything remains the same, but since $T = \frac{1}{4}$ is imposed, we lose the natural boundary condition, that $\mathscr{H} = 0$, so imposing $x(0) = 2$ and $x(\frac{1}{4}) = 1$ determine the constants, and

$$x(t) = \left(\frac{e^{1/2} - 2}{e - 1} \right) e^{2t} + \left(\frac{2e - e^{1/2}}{e - 1} \right) e^{-2t} \qquad u(t) = -10 e^{-2t}.$$

For this solution, the value of the Hamiltonian is

$$\mathscr{H} = -\frac{16(e-2)(2e^{1/2} - 1)}{(e-1)^2} \approx -7.21 < 0.$$

This illustrates the *maximum* in PMP, at the optimal solution (having the optimal stopping time), the value of the Hamiltonian will be maximised; in general $\mathscr{H} \leq 0$.

Chapter 4

4.1 For cases (i, ii, iii) the scaled problem is

$$\frac{d^2y}{dt^2} = -\frac{\Pi_1}{(1+\Pi_2 y)^2} \qquad y(0) = \Pi_3 \qquad y'(0) = -\Pi_4$$
$$\Pi_1 = \frac{4\pi G\rho_{\mathsf{E}}\mathsf{R}_{\mathsf{E}}\mathrm{T}^2}{3\mathrm{L}} \qquad \Pi_2 = \frac{\mathrm{L}}{\mathsf{R}_{\mathsf{E}}} \qquad \Pi_3 = \frac{2}{\mathrm{L}} \qquad \Pi_4 = \frac{\mathsf{V}_0\mathrm{T}}{\mathrm{L}}$$

(i) Starting with $\Pi_4 = 1$ we get $\mathrm{L} = \mathsf{V}_0\mathrm{T}$ then $\Pi_3 = 1$ sets $\mathrm{L} = 2, \mathrm{T} = 2/\mathsf{V}_0$, $\Pi_1 = 8\pi G\rho_{\mathsf{E}}\mathsf{R}_{\mathsf{E}}/(3\mathsf{V}_0^2) = \varepsilon \to 0$

$$\frac{d^2y}{dt^2} = -\frac{\varepsilon}{(1+\Pi_2 y)^2} \qquad y(0) = 1 \qquad y'(0) = -1$$

(ii) Starting with $\Pi_1 = 1$ we get $\mathrm{L} = 4\pi G\rho_{\mathsf{E}}\mathsf{R}_{\mathsf{E}}\mathrm{T}^2/3$ then $\Pi_3 = 1$ sets $\mathrm{T} = \sqrt{3/(2\pi G\rho_{\mathsf{E}}\mathsf{R}_{\mathsf{E}})}$ and $\mathrm{L} = 2$ and $\Pi_4 = \mathsf{V}_0\sqrt{3/(8\pi G\rho_{\mathsf{E}}\mathsf{R}_{\mathsf{E}})} = \varepsilon \to 0$

$$\frac{d^2y}{dt^2} = -\frac{1}{(1+\Pi_2 y)^2} \qquad y(0) = 1 \qquad y'(0) = -\varepsilon$$

(iii) Starting with $\Pi_3 = 1$, we get $\mathrm{L} = 2$, then $\Pi_4 = 1$ sets $\mathrm{T} = 2/\mathsf{V}_0$ and $\Pi_1 = 8\pi G\rho_{\mathsf{E}}\mathsf{R}_{\mathsf{E}}/(3\mathsf{V}_0^2) = \varepsilon \to 0$

$$\frac{d^2y}{dt^2} = -\frac{\varepsilon}{(1+\Pi_2 y)^2} \qquad y(0) = 1 \qquad y'(0) = -1$$

For (iv) the scaled problem is

$$\frac{d^2y}{dt^2} = -\frac{\Pi_1}{(y+\Pi_2)^2} \qquad y(0) = \Pi_3 \qquad y'(0) = -\Pi_4$$
$$\Pi_1 = \frac{G\mathsf{M}_{\mathsf{E}}\mathrm{T}^2}{\mathrm{L}^3} \qquad \Pi_2 = \frac{\mathsf{R}_{\mathsf{E}}}{\mathrm{L}} \qquad \Pi_3 = \frac{2}{\mathrm{L}} \qquad \Pi_4 = \frac{\mathsf{V}_0\mathrm{T}}{\mathrm{L}}$$

Setting $\Pi_3 = \Pi_4 = 1$ yields $\mathrm{L} = 2$ and $\mathrm{T} = 2/\mathsf{V}_0$ and $\Pi_2 = \mathsf{R}_{\mathsf{E}}/2 = \varepsilon \to 0$

$$\frac{d^2y}{dt^2} = -\frac{\Pi_1}{(y+\varepsilon)^2} \qquad y(0) = 1 \qquad y'(0) = -1$$

4.2 Defining the nondimensionalized solution as $\mathsf{X} = \mathrm{L}x(t)$ with $\mathsf{T} = \mathrm{T}t$, we can write the scaled problem as

$$x'' + x = \Pi_1 \sin(\Pi_2 t), \qquad x(0) = 1, \qquad x'(0) = \Pi_3$$

with $\mathrm{L} = \mathsf{X}_0$ being an imposed scale set by the IC and derived timescale $\mathrm{T} = \sqrt{\mathsf{M}/\mathsf{K}}$ being the inverse of the natural frequency $\omega_0 = \sqrt{\mathsf{K}/\mathsf{M}}$ and

$$\Pi_1 = \frac{\mathsf{F}}{\mathsf{K}\mathsf{X}_0}, \qquad \Pi_2 = \frac{\Omega}{\omega_0}, \qquad \Pi_3 = \frac{\omega_0 \mathsf{V}_0}{\mathsf{X}_0}.$$

4.3 The choice of scalings $\mathrm{L} = \mathsf{B}/\sqrt{\mathsf{K}\mathsf{M}}, \mathrm{T} = \mathsf{M}/\mathsf{B}$ yields the nondimensional problem

$$x'' + x' + x^3 = \Pi_1 \sin(\Pi_2 t), \qquad x(0) = \Pi_3, \qquad x'(0) = \Pi_4,$$

with parameters

$$\Pi_1 = \frac{\mathsf{F}\mathsf{K}^{1/2}\mathsf{M}^{3/2}}{\mathsf{B}^3}, \qquad \Pi_2 = \omega\frac{\mathsf{M}}{\mathsf{B}}, \qquad \Pi_3 = \frac{\mathsf{A}\mathsf{K}^{1/2}\mathsf{M}^{1/2}}{\mathsf{B}^1}, \qquad \Pi_4 = \frac{\mathsf{C}\mathsf{K}^{1/2}\mathsf{M}^{3/2}}{\mathsf{B}^2}.$$

4.5 (a) After some algebra, we get the scalings

$$\mathrm{X} = \frac{\mathsf{D}\mathsf{E}^2}{\mathsf{C}\mathsf{F}\mathsf{H}}, \qquad \mathrm{Y} = \frac{\mathsf{E}}{\mathsf{F}}, \qquad \mathrm{Z} = \frac{\mathsf{E}^2}{\mathsf{F}\mathsf{H}}, \qquad \mathrm{T} = \frac{\mathsf{D}\mathsf{E}}{\mathsf{B}\mathsf{C}\mathsf{H}},$$

then the dimensionless parameters follow as

$$\alpha = \frac{\mathsf{A}\mathsf{F}}{\mathsf{B}\mathsf{E}}, \qquad \beta = \frac{\mathsf{B}\mathsf{C}\mathsf{H}^2}{\mathsf{D}^2\mathsf{E}^2}, \qquad \gamma = \frac{\mathsf{B}\mathsf{C}}{\mathsf{D}\mathsf{E}}, \qquad \delta = \frac{3\mathsf{G}\mathsf{E}}{\mathsf{F}^2}$$

and $\mu = \mathsf{X}_0/\mathrm{X}, \sigma = \mathsf{Y}_0/\mathrm{Y}, \omega = \mathsf{Z}_0/\mathrm{Z}$
(b) $x' = \alpha - y$, $\beta y' = x - z$ and $0 = y - y^2 + \frac{1}{3}\delta y^3 - z$. Solving the last equation for $z = f(y)$ yields a phase plane system for (x, y)

$$x' = \alpha - y \qquad \beta y' = x - y + y^2 - \tfrac{1}{3}\delta y^3$$

A mismatch will occur unless the initial conditions satisfy $\omega = f(\sigma)$.
(c) $x' = \alpha - y$, $0 = x - z$, $\gamma z' = y - y^2 + \frac{1}{3}\delta y^3 - z$. Using the second equation, the other two reduce to

$$x' = \alpha - y \qquad \beta x' = y - y^2 + \tfrac{1}{3}\delta y^3 - x.$$

Equating the two expressions for x' yields $\gamma(\alpha - y) = y - y^2 + \frac{1}{3}\delta y^3 - x$. Finally, implicitly differentiating and using the equation for x' yields

$$y' = \frac{\alpha - y}{1 + \gamma - 2y + \delta y^2}$$

A mismatch will occur unless the initial conditions satisfy $\omega = \mu$.

4.7 The final nondimensionalized system is

$$h_t + hu_x + uh_x = 0 \qquad u_t + uu_x + \frac{1}{\mathrm{Fr}^2} h_x = 0$$

where the Froude number is defined as $\mathrm{Fr} = \mathrm{U}/\sqrt{g\mathrm{H}}$.

4.9 (a) Let $\delta = \mathsf{W}/\mathsf{L}$ then $\Pi_1 = 4(1+\delta)(1+1/\delta)$.
(b) $\Pi_1 = \pi((3(1+\delta) - \sqrt{(3+\delta)(1+3\delta)})(3(1+1/\delta) - \sqrt{(3/\delta+1)(1/\delta+3)})$. $\Pi_{1,\mathrm{ellipse}} > \Pi_{1,\mathrm{rect}}$ if the aspect ratio is sufficiently large or small.
(c) Let $x = \Pi_2, y = \Pi_3$ and $z = 1/\Pi_1^2$ for a more convenient calculations: $z = g(x,y) = \frac{1}{2}(\frac{1}{2} - x)(\frac{1}{2} - y)(x + y - \frac{1}{2})$. Using multivariable calculus, determine the critical points of g: $g_x(x,y) = g_y(x,y) = 0$ yielding $(x,y) = (0, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, 0)$ or $(\frac{1}{3}, \frac{1}{3})$. Of these possibilities, $(\frac{1}{3}, \frac{1}{3})$ maximises $g = 1/432$ and hence minimises $\Pi_{1,\mathrm{tri}} \geq 12\sqrt{3}$. Note that $(12\sqrt{3} \approx 20.78) > (12.56 \approx 4\pi)$ as is expected from the result that the circle minimises the perimeter-to-area ratio, $\Pi_{1,\mathrm{circ}} = 4\pi$.

4.11 $\Pi_1 = \mathsf{AT}/\mathsf{B}, \Pi_2 = \mathsf{AC}/\mathsf{B}^2, \Pi_3 = \mathsf{A}^2\mathsf{D}/\mathsf{B}^3, \mathsf{T} = (\mathsf{B}/\mathsf{A})f(\mathsf{AC}/\mathsf{B}^2, \mathsf{A}^2\mathsf{D}/\mathsf{B}^3)$

4.12 (a) The equations for the dimensional exponents for cancelling out units in the Π's are:

$$\begin{aligned} A - 3B - C + E + F &= 0 \quad [\mathrm{m}] \\ B + C + D &= 0 \quad [\mathrm{kg}] \\ -2A - C - 2D - F &= 0 \quad [\mathrm{s}] \end{aligned}$$

These three equations are linearly independent (as can be seen by reducing them to echelon form), so $\tilde{r} = 3$. There are 6 given quantities and 3 base units ($r = 3$), so we get $n - \tilde{r} = 6 - 3 = 3$ free parameters.

Chapter 5

5.1 Using the rescaled solution (5.1), we can rewrite Burgers' equation as

$$u_t + \left(\frac{\mathsf{UT}}{\mathsf{L}}\right) uu_x = \left(\frac{\mathsf{T}}{\mathsf{L}}\right) \kappa u_{xx}.$$

To make this scale invariant, the coefficient factors need to be set to one. The normalising the first yields $\mathsf{U} = \mathsf{L}/\mathsf{T}$. Normalising the second yields $\mathsf{L} = \mathsf{T}^{1/2}$ and

subsequently $\mathsf{U} = \mathsf{T}^{-1/2}$. A scale-invariant similarity variable can be obtained from $\Pi_1 = \mathsf{L}\mathsf{T}^c = \mathsf{T}^{1/2}\mathsf{T}^{1/2} = \mathsf{T}^0$, hence $c = 1/2$ and hence determines $\eta = xt^{-1/2}$. Likewise a scale-invariant similarity function is determined by $\Pi_2 = \mathsf{U}\mathsf{T}^d = \mathsf{T}^{-1/2}\mathsf{T}^d = \mathsf{T}^0$, hence $d = 1/2$ and $f(\eta) = t^{1/2}u$ yielding the similarity solution form $u(x, t) = t^{-1/2}f(\eta)$. Substituting this into Burgers' equation yields

$$\tfrac{1}{2}\left(f + \eta\frac{df}{d\eta}\right) + f\frac{df}{d\eta} = \kappa\frac{d^2f}{d\eta^2}.$$

5.2 Applying (5.1)–(5.3) and the integral yields

$$u_t + \left(\frac{\mathsf{U}\mathsf{T}}{\mathsf{L}}\right)uu_x = 0, \qquad (\mathsf{U}^2\mathsf{L})\int_0^\infty u^2\,dx = 1.$$

Making the PDE scale invariant selects the scaling relation $\mathsf{U} = \mathsf{L}/\mathsf{T}$. Then making the integral scale-invariant determines $\mathsf{L} = \mathsf{T}^{2/3}$ and consequently $\mathsf{U} = \mathsf{T}^{-1/3}$. Further, this determines the similarity variable $\eta = x/t^{2/3}$ and the similarity solution as $u(x, t) = t^{-1/3}f(\eta)$. Substituting this into the inviscid Burgers equation yields the ODE for $f(\eta)$,

$$-\tfrac{1}{3}\left(f + \eta\frac{df}{d\eta}\right) + f\frac{df}{d\eta} = 0 \qquad \int_0^\infty f^2\,d\eta = 1.$$

The similarity solution likewise reduces the generalised Burgers equation, $u(u_t + uu_x) = \kappa u_{xx}$, to the ODE

$$-\tfrac{1}{3}f\left(f + 2\eta\frac{df}{d\eta}\right) + f^2\frac{df}{d\eta} = \kappa\frac{d^2f}{d\eta^2}.$$

5.3 (a, b) Applying (5.1) to the PDE yields

$$\frac{\partial u}{\partial t} + \left(\frac{\mathsf{U}\mathsf{T}}{\mathsf{L}}\right)u\frac{\partial u}{\partial x} = (\mathsf{T}^{\sigma+1}\mathsf{U}^3)t^\sigma u^4.$$

To make this equation scale invariant, we need $\mathsf{U} = \mathsf{L}/\mathsf{T}$ (from the first coefficient) and $\mathsf{T}^{\sigma+1}\mathsf{U}^3 = 1$ in the second one. The boundary condition must be scale invariant as well,

$$u(0, t) = \left(\frac{\mathsf{T}^3}{\mathsf{U}}\right)t^3 \qquad \Longrightarrow \qquad \mathsf{U} = \mathsf{T}^3,$$

consequently $\mathsf{L} = \mathsf{T}^4$ (from $\mathsf{L} = \mathsf{U}\mathsf{T}$). Satisfying the second condition for the scale invariance of the PDE, $\mathsf{T}^{\sigma+1}\mathsf{U}^3 = \mathsf{T}^{\sigma+1}\mathsf{T}^9 = \mathsf{T}^0$ determines $\sigma = -10$.

(c) The scale-invariant similarity variable is $\eta = x/t^4$ and the form of the similarity solution is $u(x,t) = t^3 f(\eta)$. Substituting this form into the PDE reduces it to the ODE,

$$3f - 4\eta \frac{df}{d\eta} + f\frac{df}{d\eta} = f^4.$$

5.4 Satisfying scale-invariance in the heat equation determines $\mathsf{L} = \mathsf{T}^{1/2}$ and the similarity variable $\eta = x/t^{1/2}$.
(a) The boundary condition will be scale invariant for $\mathsf{U} = \mathsf{T}^2$. This yields the form of the similarity solution as $u(x,t) = t^2 f(\eta)$. Substituting in the PDE yields the ODE problem

$$2f - \tfrac{1}{2}\eta f' = f'', \qquad f(0) = 1, \qquad f(\eta \to \infty) \to 0.$$

(b) Here the boundary condition will be scale-invariant if $\mathsf{U} = \mathsf{L} = \mathsf{T}^{1/2}$ yielding the similarity solution $u(x,t) = t^{1/2} f(\eta)$, satisfying the ODE problem,

$$\tfrac{1}{2}(f - \eta f') = f'', \qquad f'(0) = 2, \qquad f(\eta \to \infty) \to 0.$$

(c) Here the boundary condition will be scale-invariant if $\mathsf{U} = 1/\mathsf{L} = \mathsf{T}^{-1/2}$ yielding the similarity solution $u(x,t) = t^{-1/2} f(\eta)$, satisfying the ODE problem,

$$-\tfrac{1}{2}(f + \eta f') = f'', \qquad f'(0) = f(0)^2, \qquad f(\eta \to \infty) \to 0.$$

(d) Here the integral condition will be scale-invariant if $\mathsf{L}^3\mathsf{U} = 1$, or $\mathsf{U} = \mathsf{T}^{-3/2}$ yielding the similarity solution $u(x,t) = t^{-3/2} f(\eta)$, satisfying the ODE problem,

$$-\tfrac{1}{2}(3f + \eta f') = f'', \qquad \int_0^\infty \eta^2 f \, d\eta = 1, \qquad f(\eta \to \infty) \to 0.$$

(e) Here $\mathsf{L} = \mathsf{T} = 1$ but U is a free parameter. The solution can be written as $u(x,t) = Ce^t e^{-x}$, which is actually a travelling wave, $u = Ce^{-(x-t)}$.

5.5 Applying (5.1) to the PDE yields

$$\frac{\partial u}{\partial t} = \left(\frac{\mathsf{T}}{\mathsf{L}^2}\right)\frac{\partial^2 u}{\partial x^2} + \left(\mathsf{U}^3\mathsf{T}\right)u^4.$$

The two coefficients must be normalised in order to make the PDE scale-invariant. The first coefficient determines $\mathsf{L} = \mathsf{T}^{1/2}$ and the second sets $\mathsf{U} = \mathsf{T}^{-1/3}$. The scale-invariant similarity variable is $\eta = x/t^{1/2}$ and the form of the similarity solution is $u(x,t) = t^{-1/3} f(\eta)$, satisfying the ODE

$$-\tfrac{1}{3}f - \tfrac{1}{2}\eta f' = f'' + f^4.$$

5.6 (a) Applying (5.1) to the PDE yields

$$\frac{\partial u}{\partial t} = -\left(\frac{\mathsf{U}^3\mathsf{T}}{\mathsf{L}^2}\right)\frac{\partial}{\partial x}\left(u^3\frac{\partial u}{\partial x}\right) - \left(\frac{\mathsf{U}\mathsf{T}}{\mathsf{L}^4}\right)\frac{\partial}{\partial x}\left(u\frac{\partial^3 u}{\partial x^3}\right)$$

The two coefficients must be normalised in order to make the PDE scale-invariant. The first coefficient determines $\mathsf{L}^2 = \mathsf{U}^3\mathsf{T}$; substituting this into $\mathsf{L}^4 = \mathsf{U}\mathsf{T}$ determines $\mathsf{U} = \mathsf{T}^{-1/5}$ and subsequently $\mathsf{L} = \mathsf{T}^{1/5}$. The scale-invariant similarity variable is then $\eta = x/t^{1/5}$ and the form of the similarity solution is $u(x, t) = t^{-1/5}f(\eta)$, satisfying the ODE

$$-\tfrac{1}{5}(f + \eta f') = -(f^3 f')' - (ff''')'.$$

Note that the form and scaling of the similarity solution has already been determined, the additional integral condition happens to be consistent, $\int f\, d\eta = 1$.
(b) Applying

$$u(x, t) = \mathsf{U}\tilde{u}(\tilde{x}, \tilde{t}\,), \qquad x = \mathsf{L}\tilde{x}, \qquad t = t_c + \mathsf{T}\tilde{t}$$

to the PDE yields the same scalings, $\mathsf{U} = \mathsf{T}^{-1/5}$ and $\mathsf{L} = \mathsf{T}^{1/5}$. However the similarity variable and solution now take a modified form,

$$\eta = \frac{x}{(t_c - t)^1/5} \qquad u(x, t) = (t_c - t)^{-1/5}f(\eta),$$

and the similarity function satisfies the modified ODE

$$\tfrac{1}{5}(f + \eta f') = -(f^3 f')' - (ff''')'.$$

In part (a), the similarity solution evolves to a limiting behaviour as $\mathsf{T} \to \infty\,(t \to \infty)$, while for this finite-time blow-up case, divergence occurs as $\mathsf{T} \to 0\,(t \to t_c)$.

5.8 The similarity solutions are

$$h(x, t) = t^{-1/2}f(xt^{-3/8}), \qquad u(x, t) = t^{-5/8}g(xt^{-3/8}).$$

Chapter 6

6.2 At leading order it is straightforward to identify $\delta_0 = 1$ and $x_0 = 3$ as a triple root of $(x_0-3)^3 = 0$. The next iteration, $x \sim 3+\delta_1 x_1$ yields $\delta_1^3 x_1^3 = 216\varepsilon$ hence $\delta_1 = \varepsilon^{1/3}$ and x_1 is given by the one of the three cube roots of 216, namely $x_1 = 6e^{i2\pi k/3}$ for $k = 0, 1, 2$. Going on, we will get $\delta_2 = \varepsilon^{2/3}$ and

$$x \sim 3 + 6\varepsilon^{1/3} + 8\varepsilon^{2/3}, \qquad x \sim 3 - (3 \pm i3\sqrt{3})\varepsilon^{1/3} - (4 \mp i4\sqrt{3})\varepsilon^{2/3}$$

6.3 Observe that setting $\varepsilon = 0$ in the equation yields a contradiction ('$60 = 0$'), hence the solutions must be singular in order to yields valid balances. Let $x = \delta X$ to yield

$$\underbrace{\varepsilon^6\delta^3X^3}_{(1)} - \underbrace{5\varepsilon^3\delta^2X^2}_{(2)} - \underbrace{20\varepsilon\delta X}_{(3)} + \underbrace{60}_{(4)} = 0.$$

Consider the different possibilities for dominant balances to determine the distinguished limits yielding solutions. This is a third order polynomial, so there must be exactly three roots that must represented within the set of distinguished limits. These distinguished limits are given by

- Balancing terms (3, 4): $\delta = \varepsilon^{-1}$, $x \sim 3/\varepsilon$
- Balancing terms (2, 3): $\delta = \varepsilon^{-2}$, $x \sim -4/\varepsilon^2$
- Balancing terms (1, 2): $\delta = \varepsilon^{-3}$, $x \sim 5/\varepsilon^3$

6.5 (a) The solution is given by

$$v(t) \sim -\tfrac{1}{2}t^2 - \tfrac{1}{20}\varepsilon t^5 - \tfrac{1}{160}\varepsilon^2 t^8.$$

(b) Note that after factoring out t^2, the magnitude of the terms follows

$$O(1) \gg O(\varepsilon t^3) \gg O(\varepsilon^2 t^6)$$

with the ordering being preserved if $\varepsilon t^3 \ll 1$, namely $0 \le t \ll O(\varepsilon^{-1/3})$.

6.6 Observe that setting $\varepsilon = 0$ in the equations yields $-y = 1$ and $y = 4$, hence we get a contradiction (y is overdetermined). Rescale each variable independently, $x = \delta(\varepsilon)X$ and $y = \sigma(\varepsilon)Y$ and carry out dominant balance of the system to obtain $\delta = \varepsilon^{-1}$ (from a dominant balance in the first equation) and $\sigma = \varepsilon^0$ (from a dominant balance in the second equation). Subsequently,

$$x \sim \frac{1}{\varepsilon}(5 - 5\varepsilon) \qquad y \sim 4 - 5\varepsilon.$$

6.7 Taking the logarithm of the equation, we get

$$\ln(2) + 2\ln(x) - 5x = 3\ln(2) + \ln(\varepsilon)$$

It is easy to identify $\ln(\varepsilon) \to -\infty$ as the dominant term on the righthand side. Now, substitute $x = \delta_0 x_0$ with the assumption that $\delta_0 \to \infty$,

$$\ln(2) + 2\ln(\delta_0) + 2\ln(x_0) - 5\delta_0 x_0 = 3\ln(2) + \ln(\varepsilon)$$

Since $x_0 = O(1)$, so is its log. Since $z \gg \ln(z)$ as $z \to \infty$ the $5\delta_0 x_0$ term is the largest term on the left-hand side of the equation. Hence we get that $x_0 = 1/5$ and $\delta_0 = -\ln(\varepsilon)$. Note that we have put the negative sign in δ_0 rather than in the coefficient so that the gauge function is positive ($\ln(z) \to -\infty$ for $z \to 0$). In general for dealing with logarithms, it may be better to write them as $\delta_0 = \ln(\frac{1}{\varepsilon})$. The next iteration, $x \sim \delta_0 x_0 + \delta_1 x_1$ (with $\delta_0 \gg \delta_1$) yields a dominant balance between terms

coming from the original $2\ln(x)$ and $5x$ terms to determine $\delta_1 = \ln(\ln(\frac{1}{\varepsilon}))$ and the solution as

$$x \sim \tfrac{1}{5}\ln(\tfrac{1}{\varepsilon}) + \tfrac{2}{5}\ln(\ln(\tfrac{1}{\varepsilon})).$$

6.10 (a) Substituting $h = 1 + \varepsilon\eta$ and $u = 1 + \varepsilon v$ into the PDEs and then collecting $O(\varepsilon)$ terms yields the linearised system

$$\eta_t + v_x + \eta_x = 0 \qquad v_t + v_x + \mathrm{Fr}^{-2}\eta_x = 0$$

(b) We can then write the system in the form

$$\frac{\partial}{\partial t}\begin{pmatrix}\eta\\ v\end{pmatrix} + \underbrace{\begin{pmatrix}1 & 1\\ \mathrm{Fr}^{-2} & 1\end{pmatrix}}_{\mathbf{M}}\frac{\partial}{\partial x}\begin{pmatrix}\eta\\ v\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix}$$

Consequently the determinant condition for the wavespeeds is

$$|\mathbf{M}^T - \lambda\mathbf{I}| = \begin{vmatrix}1-\lambda & \mathrm{Fr}^{-2}\\ 1 & 1-\lambda\end{vmatrix} = 0 \qquad \Longrightarrow \qquad (1-\lambda)^2 = \mathrm{Fr}^{-2}$$

and hence the waves generated will have speeds

$$\lambda = 1 \pm \mathrm{Fr}^{-1}$$

and if $\mathrm{Fr} > 1$ then $\lambda_\pm > 0$ and all waves will move to the right (downstream, subcritical), while if $\mathrm{Fr} < 1$ then $\lambda_\pm \gtrless 0$ and the two classes of waves move in opposite directions.

Chapter 7

7.3 Noting that the inhomogeneous boundary conditions are $O(1)$, the dominant is finite and the coefficient $3 \le (4 - x^2) \le 4$ is bounded on the domain, we expect the solution to be finite and bounded, hence we will assume $\beta = 0$ and seek solutions of the form $y = Y(X)$. Letting $X = (x - x_*)/\varepsilon^\alpha$ for $\alpha \ge 0$, equivalently we have $x = x_* + \varepsilon^\alpha X$. The inhomogeneous term then becomes

$$\cos(\tfrac{\pi}{2}x) = \cos(\tfrac{\pi}{2}x_* + \varepsilon^\alpha\tfrac{\pi}{2}X)$$

There are two cases that need to be considered for the scaling of this term with respect to the position of the boundary layer:

$$\begin{aligned} x_* = 0 \qquad &\Longrightarrow \qquad \cos(\tfrac{\pi}{2}x) = \cos(\varepsilon^\alpha\tfrac{\pi}{2}X) \sim 1\\ x_* = 1 \qquad &\Longrightarrow \qquad \cos(\tfrac{\pi}{2}x) = -\sin(\varepsilon^\alpha\tfrac{\pi}{2}X) \sim -\varepsilon^\alpha\tfrac{\pi}{2}X \end{aligned}$$

To determine the dominant balances, examine

$$\underbrace{\varepsilon^{1-2\alpha}Y''}_{(1)} + \underbrace{\varepsilon^0(4-x_*^2)Y}_{(2)} = \underbrace{\cos(\tfrac{\pi}{2}x)}_{(3)},$$

where term (3) is either $O(\varepsilon^0)$ or $O(\varepsilon^\alpha)$ as described above.

For $\alpha = 0$, we recover the outer distinguished limit which yields the leading order outer solution

$$y_0(x) = -\frac{\cos(\frac{\pi}{2}x)}{4-x^2}.$$

Note that this solution does not satisfy either boundary condition (with $y_0(0) = -1/4$ and $y_0(1) = 0$, so we will need boundary layers at both $x_* = 0$ and $x_* = 1$. First consider the case that the boundary layer is on the left edge of the domain, $x_* = 0$, making term (3) be $O(1)$. Then the dominant balance for the singular distinguished limit is between terms (1, 2) yielding $\alpha = \frac{1}{2}$. Term (3) is of the same order, $O(1)$, and hence also takes part in the dominant balance. The resulting leading order inner problem for $y = Y(X)$ with $X = x/\varepsilon^{1/2}$ is

$$Y_0'' - 4Y_0 = 1, \qquad Y_0(0) = -1 \qquad \Longrightarrow \qquad Y_0^L = -\frac{3}{4}e^{-2X} - \frac{1}{4},$$

where we note that we have eliminated an un-matchable exponentially growing term for $X \to \infty$. Applying the matching condition between this inner solution and the outer solution,

$$\lim_{X\to\infty} Y_0^L = \lim_{x\to 0} y_0 = -\frac{1}{4},$$

which fortunately yields consistency and gives the overlap as $-1/4$.

For the boundary layer at $x_* = 1$, term (3) is $O(\varepsilon^\alpha)$. The singular distinguished limit is still $\alpha = \frac{1}{2}$ but now only terms (1, 2) are involved in the dominant balance, yielding the leading order inner problem for $y = Y(X)$ with $X = (x-1)/\varepsilon^{1/2}$,

$$Y_0'' - 3Y_0 = 0, \qquad Y_0(0) = 2 \qquad \Longrightarrow \qquad Y_0^R = 2e^{\sqrt{3}X},$$

and again, an unmatchable exponentially growing term (for $X \to -\infty$) has be excluded. Here, applying the matching condition between this inner solution and the outer solution,

$$\lim_{X\to-\infty} Y_0^R = \lim_{x\to 1} y_0 = 0,$$

again yielding consistency of the construction process, here with zero overlap.

Forming the left and right boundary layer corrections by subtracting the respective matching overlaps from their inner solutions, we can write the composite solution (7.22) as

$$y \sim -\frac{\cos(\frac{\pi}{2}x)}{4-x^2} - \frac{3}{4}e^{-2x/\sqrt{\varepsilon}} + 2e^{\sqrt{3/\varepsilon}(x-1)}.$$

7.4 (a) Given that the boundary occurs at $x_* = 0$, the singular solution will have the form $y = \varepsilon^\beta Y(X)$ with $X = x/\varepsilon^\alpha$ and $\alpha > 0$. The homogeneous initial condition provides no information. The second initial condition gives

$$y'(0) = \frac{4}{\varepsilon^2} \qquad \Longrightarrow \qquad \varepsilon^{\beta-\alpha}Y'(0) = 4\varepsilon^{-2},$$

hence $Y'(0) = 0$ and $\beta - \alpha = -2$. Using this in the ODE, we get

$$\underbrace{\varepsilon^{-1-\alpha}Y''}_{(1)} + \underbrace{2\varepsilon^{-2}Y'}_{(2)} - \underbrace{6\varepsilon^{-2+\alpha}Y}_{(3)} = \underbrace{5\varepsilon^\alpha X}_{(4)}.$$

It can be shown that the only consistent singular dominant balance is between terms (1, 2), yielding $\alpha = 1$, hence $\beta = -1$. Consequently, the inner problem is

$$Y'' + 2Y' - 6\varepsilon Y = 5\varepsilon^3 X, \qquad Y(0) = 0, \qquad Y'(0) = 4,$$

which reduces to the leading order problem

$$Y_0'' + 2Y_0' = 0 \qquad \Longrightarrow \qquad Y_0(X) = 2(1 - e^{-2X}),$$

satisfying both initial conditions.
(b) Applying the matching condition,

$$\lim_{X\to\infty} \varepsilon^{\beta_{\text{inner}}} Y_{\text{inner}}(X) = \lim_{x\to 0} \varepsilon^{\beta_{\text{outer}}} y_{\text{outer}}(x),$$

we get from the inner solution that the limit is $2/\varepsilon$ hence the outer solution must be scaled by $\beta_{\text{outer}} = -1$ and satisfy initial condition $y_{\text{outer}}(0) = 2$.
(c) Using $\beta = -1$, $\alpha = 0$ for the scaling of the outer solution yields the equation for $y = y_{\text{outer}}$

$$\varepsilon^0 y'' + 2\varepsilon^{-1}y' - 6\varepsilon^{-1}y = 5x, \qquad y(0) = 2$$

which at leading order reduces to

$$2y_0' - 6y_0 = 0 \qquad y_0(x) = 2e^{3x}.$$

Consequently, we can construct the composite solution from the outer solution plus the inner solution minus the overlap as

$$y \sim \frac{2}{\varepsilon}e^{3x} + \frac{2}{\varepsilon}\left(1 - e^{-2x/\varepsilon}\right) - \frac{2}{\varepsilon} = \frac{2}{\varepsilon}\left(e^{3x} - e^{-2x/\varepsilon}\right).$$

7.5 (a) The leading order outer solution is given by

$$2y' = e^{y_0} = 0 \qquad \Longrightarrow \qquad y_0(x) = -\ln\left(\tfrac{1}{2}(x-c)\right),$$

where the choice of the constant of integration depends on which boundary condition applies to the outer solution.
(b) Assuming the solution to be finite, $y = O(1)$, hence $\beta = 0$, dominant balances are determined by

$$\underbrace{\varepsilon^{1-2\alpha}Y''}_{(1)} + \underbrace{2\varepsilon^{-\alpha}Y'}_{(2)} - \underbrace{\varepsilon^0 e^Y}_{(3)} = 0.$$

The outer distinguished limit is obtained by balancing terms (2,3) at $\alpha = 0$. The singular distinguished limited follows from balancing terms (1,2) at $\alpha = 1$. The leading order inner problem and its general solution is

$$Y_0'' + 2Y_0' = 0 \qquad \Longrightarrow \qquad Y_0(X) = C_1 + C_2 e^{-2X}.$$

As yet we have not determined the position of the boundary layer (x_*), but from the form of $Y_0(X)$ we can see that only the choice $x_* = 0$ can yield a nontrivial inner solution. Consequently, applying the boundary condition

$$Y_0(0) = 0 \qquad \Longrightarrow \qquad Y_0(X) = C_1(1 - e^{-2X}).$$

(c) Meanwhile we know that the other boundary condition, $y(1) = 0$ must apply to the outer solution (since there is no boundary possible at $x_* = 1$), hence the final form of the outer solution becomes

$$y_0(x) = -\ln\left(\tfrac{1}{2}(x+1)\right).$$

Asymptotic matching to the inner solution then determines that $C_1 = \ln(2)$. Consequently the leading order composite solution is

$$y \sim -\ln\left(\tfrac{1}{2}(x+1)\right) - \ln(2)e^{-2x/\varepsilon}.$$

(d) Expanding the outer solution as $y \sim y_0 + \varepsilon y_1$, the exponential term can be written as

$$e^y \sim e^{y_0+\varepsilon y_1} = e^{y_0}e^{\varepsilon y_1} \sim e^{y_0}\left(1 + \varepsilon y_0 + \tfrac{1}{2}\varepsilon^2 y_1^2 + \cdots\right).$$

Collecting $O(\varepsilon)$ terms in the ODE, the equation for y_1 is then

$$y_0'' + 2y_1' + y_1 e^{y_0} = 0 \qquad \Longrightarrow \qquad y_1(x) = -\frac{\ln\left(\frac{1}{2}(x+1)\right)}{2(x+1)}.$$

7.7 (a) The outer solution is given by

$$y \sim \pm\sqrt{4-x^2} + \frac{\varepsilon}{2(4-x^2)}.$$

(b) At $x_* = 2$, the limiting behaviour of the outer solution for $x \to 2$ determines that $\beta = \frac{1}{2}\alpha$ and the dominant balance of terms in the ODE determines the singular distinguished limit as $\alpha = \frac{2}{3}$, $\beta = \frac{1}{3}$ with the inner problem being

$$4X + \varepsilon^{2/3}X^2 + Y^2 = -Y' \qquad X = \frac{x-2}{\varepsilon^{2/3}}.$$

At $x_* = 0$, there is one inner solution that must be matchable to the outer solution, $y(x \to 0) \sim \pm 2 = O(1)$, hence it must have $\beta = 0$. A dominant balance in the ODE yields $\alpha = 1$ with the inner problem

$$\varepsilon^2 X^2 + Y^2 = 4 - Y', \qquad X = \frac{x}{\varepsilon}.$$

At $x_* = 0$ there is a different inner solution that satisfies the initial condition $y(0) = 3/\varepsilon$, hence having $\beta = -1$. Determination of the dominant balance with this value of β yields $\alpha = 2$ and the ODE

$$\varepsilon^6 X^2 + Y^2 = 4\varepsilon^2 - Y' \qquad X = \frac{x}{\varepsilon^2},$$

since this solution exists on a narrower domain ($\varepsilon^2 \ll \varepsilon$), this is sometimes called an *inner-inner* solution.

Chapter 8

8.1 (a) Using the method of separation of variables, we get (8.8),

$$\mathsf{U}(\mathsf{X},\mathsf{Y}) = \sum_{n=1}^{\infty} \mathsf{c}_n \sinh\left(\tfrac{n\pi}{\mathsf{L}}\mathsf{Y}\right) \sin\left(\tfrac{n\pi}{\mathsf{L}}\mathsf{X}\right)$$

Evaluating the boundary condition along $\mathsf{Y} = \mathsf{H}$ yields

$$\left.\frac{\partial \mathsf{U}}{\partial \mathsf{Y}}\right|_{\mathsf{Y}=\mathsf{H}} = \sum_{n=1}^{\infty} \left(\frac{n\pi \mathsf{c}_n}{\mathsf{L}} \cosh(\tfrac{n\pi}{\mathsf{L}}\mathsf{H}) \right) \sin(\tfrac{n\pi}{\mathsf{L}}\mathsf{X}) = e^{-3\mathsf{X}/\mathsf{L}}$$

and the Fourier sine series determines the final form of the solution as

$$\mathsf{U}(\mathsf{X},\mathsf{Y}) = \sum_{n=1}^{\infty} \left(\frac{2}{n\pi \cosh\left(\frac{n\pi \mathsf{H}}{\mathsf{L}}\right)} \int_0^L e^{-3\tilde{\mathsf{X}}/\mathsf{L}} \sin(\tfrac{n\pi}{\mathsf{L}}\tilde{\mathsf{X}})\, d\tilde{\mathsf{X}} \right) \sinh(\tfrac{n\pi}{\mathsf{L}}\mathsf{Y}) \sin(\tfrac{n\pi}{\mathsf{L}}\mathsf{X})$$

(b) The leading order outer problem, scaled with $\bar{\mathsf{U}} = 1\,\mathrm{m}$ is

$$u_{0yy} = 0, \qquad u_0(x,0) = 0, \qquad u_{0y} = \mathsf{H} e^{-3x}$$

yielding $u_0(x,y) = \mathsf{H} e^{-3x} y$.
(c) Substituting $\mathsf{X} = \mathsf{L}x$, $\mathsf{Y} = \mathsf{H}y$ and taking the $\varepsilon \to 0$ limit of the separation of variables solution yields $\mathsf{U} =$ (Fourier sine series of e^{-3x})$\mathsf{H}y$.

8.2 (a) The leading order outer problem is

$$u_{0yy} = 0, \qquad u_0(x,0) = \sin(5\pi x), \qquad u_0(x,1) = \cos(3\pi x)$$

yielding $u_0(x,y) = (\cos(3\pi x) - \sin(5\pi x))y + \sin(5\pi x)$.
(b) Since $\lim_{x\to 0} u_0(x,y) = y$, the boundary layer correction at $x_*^L = 0$ can be expressed as $V^L(X,y) = U(x,y) - y$ with $V^L(0,y) = \sin(\frac{1}{2}\pi y) - y$ and homogeneous Dirichlet boundary conditions on rest of the boundary of the semi-infinite strip, $0 \le y \le 1$, $x \ge 0$. Then

$$V^L(X^L,y) = -2\sum_{n=1}^{\infty} \frac{(-1)^n}{(4n^2-1)n\pi} e^{-n\pi X^L} \sin(n\pi y) \qquad X^L = \frac{x}{\varepsilon} \ge 0.$$

Similarly, the right boundary layer correction can be determined to be

$$V^R(X^R,y) = \sum_{k=0}^{\infty} \frac{8}{(2k+1)^3\pi^3} e^{(2k+1)\pi X^R} \sin((2k+1)\pi y) \qquad X^R = \frac{x-1}{\varepsilon} \le 0.$$

(d) The nondimensional form of the flux is

$$J^R = \frac{1}{\mathsf{L}} \int_0^1 u_x(1,y)\, dy \sim \frac{1}{\mathsf{L}} \int_0^1 5\pi(y-1) + \frac{1}{\varepsilon} \sum_{k=0}^{\infty} \frac{8\sin((2k+1)\pi y)}{(2k+1)^2\pi^2}\, dy.$$

The value of the flux is dominated by the $O(1/\varepsilon)$ contribution from the boundary layer correction

$$J^R \sim \frac{1}{\varepsilon \mathsf{L}} \sum_{k=0}^{\infty} \frac{16}{(2k+1)^3 \pi^3} \approx \frac{16}{\mathsf{H}\pi^3}(1.0518) \approx \frac{0.543}{\mathsf{H}}.$$

8.3 Let $f(x) = F(\mathsf{X})$. (a) The scaled form of the no-flux condition on the top boundary is

$$u_y + \varepsilon^2 15\pi \sin(3\pi x) u_x = 0 \qquad \text{at } y = f(x) = 15 + 5\cos(3\pi x).$$

(b) The $O(1)$ outer problem is

$$u_{0yy} = 0 \qquad u_{0y}(x,0) = 0 \qquad u_{0y}(x, f(x)) = 0$$

The $O(\varepsilon)$ outer problem is

$$u_{0xx} + u_{1yy} = 0 \qquad u_{1y}(x,0) = 0 \qquad u_{1y}(x,f) + 15\pi \sin(3\pi x) u_{0x}(x,f) = 0$$

yielding the solutions

$$u_0(x,y) = C_2(x), \qquad u_1(x,y) = -\tfrac{1}{2} C_2''(x) y^2 + C_4(x)$$

where the top boundary condition gives the compatibility condition

$$-\frac{d}{dx}\left((15 + 5\cos(3\pi x)) \frac{dC_2}{dx} \right) = 0$$

(d) Applying matching of the outer solution to the boundary layers implies zeroing the $n = 0$ coefficients in the cosine series expansions of the boundary layer corrections yields the boundary conditions on $C_2(x)$,

$$c_0 = \frac{1}{20} \int_0^{20} \left[-\frac{y^3}{100} - C_2(0) \right] dy = 0 \qquad \Longrightarrow \qquad C_2(0) = -20,$$
$$d_0 = \frac{1}{10} \int_0^{10} \left[3y^2 - C_2(1) \right] dy = 0 \qquad \Longrightarrow \qquad C_2(1) = 100.$$

8.5 (a) The $O(1)$ problem is

$$\phi_{0yy} = 0, \qquad \phi_{0y}(x,0) = 0 \qquad \Longrightarrow \qquad \phi_0(x,y) = C_0(x,t).$$

The $O(\varepsilon^2)$ problem is

$$\phi_{1yy} + \phi_{0xx} = 0, \qquad \phi_{1y}(x,0) = 0 \qquad \Longrightarrow \qquad \phi_1(x,y) = -\tfrac{1}{2} C_{0xx}(x,t) + C_1(x,t).$$

The $O(\varepsilon^4)$ problem is

$$\phi_{2yy} + \phi_{1xx} = 0, \qquad \phi_{2y}(x,0) = 0 \quad \Longrightarrow \quad \phi_1(x,y) = \tfrac{1}{24} C_{0xxxx}(x,t) - \tfrac{1}{2} C_{1xx}(x,t) y^2 + C_2(x,t).$$

(b, c) See Exercise 9.12.

Chapter 9

9.1 Using the initial conditions, $x_0(t) = \cos(t) + 2\sin(t)$. Using the $O(\varepsilon)$ equation with $x_0(t)$ specified and the initial conditions $x_1(0) = -1, x_1'(0) = 1$, $x_1(t) = [115\sin(t) - 12\cos(t)] + [60t\sin(t) - 120t\cos(t)] + [11\cos(3t) + 2\sin(3t)]$ with the homogeneous, secular and non-resonant respond terms grouped respectively.

9.2 (a) $\tilde{x}_0 = a\cos\theta$ with $\theta \sim (2 + \frac{3}{16}\varepsilon a^2)t$ works for any $a > 0$. (b) $\tilde{x}_0 = a\cos\theta$ with $\omega_0 = 3, \omega_1 = 0$ satisfies the problem only for $a = 2$.

9.3 (a) The amplitude equations are

$$\frac{dA}{d\tau} = \tfrac{1}{2}A - \frac{A}{8}(A^2 + B^2), \qquad \frac{dB}{d\tau} = \tfrac{1}{2}B - \frac{B}{8}(B^2 + A^2).$$

(b) Note that $\frac{d(R^2)}{d\tau} = 2A\frac{dA}{d\tau} + 2B\frac{dB}{d\tau} = 2R\frac{dR}{d\tau}$. Expanding out $2R\frac{dR}{d\tau}$ yields

$$\frac{dR}{d\tau} = \tfrac{1}{2}R(1 - \tfrac{1}{4}R^2)$$

which has $R_* = 2$ as an equilibrium. Note that this matches the periodic solution determined in Exercise 9.2(b), but now, using MMTS, we have a prediction for rate of convergence to the limit cycle.

9.5 (b) Suppressing the $e^{\pm it}$ resonant forcing terms in the $O(\varepsilon)$ equation for X_1 yields

$$\frac{dC}{d\tau} = -\tfrac{1}{2}\beta C + \frac{3}{2}i\alpha|C|^2 C + \tfrac{1}{2}e^{i\gamma\tau},$$

with the second solvability condition being exactly the complex conjugate of the above equation (and hence yielding no independent information).
(c) Substituting into the amplitude ODE for C reduces to the algebraic equation

$$\gamma = \tfrac{1}{2}i\beta + \frac{3}{2}\alpha M^2 - \frac{1}{4M}e^{-i\theta}.$$

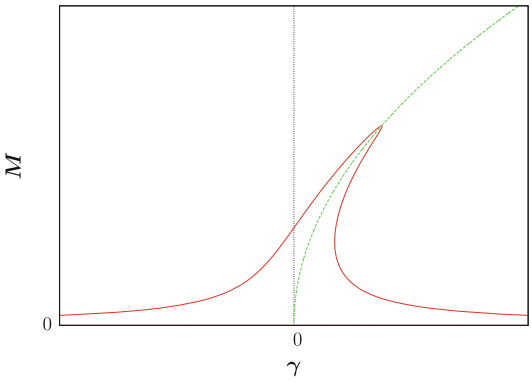
Separating real and imaginary parts of this equation yields

$$\gamma = \tfrac{3}{2}\alpha M^2 - \frac{\cos\theta}{4M}, \qquad 0 = \tfrac{1}{2}\beta + \frac{\sin\theta}{4M}.$$

For fixed α, β, expressing $\sin\theta = \pm\sqrt{1-\cos^2\theta}$, we can solve the second equation to get $\cos\theta = \pm\sqrt{1-4\beta^2M^2}$. Substituting this into the first equation gives the final (real-valued) form of the detuning relation:

$$\gamma(M) = \tfrac{3}{2}\alpha M^2 \pm \frac{\sqrt{1-4\beta^2M^2}}{4M}.$$

Entrained solutions exist up to a critical amplitude set by the damping, $M \le M_c = 1/(2\beta)$. Noting that the terms $f = x+\varepsilon\alpha x^3$ can be interpreted as a restoring force due to a spring, the case of $\alpha > 0$ is called a *hardening spring* since the restoring force is stronger than for the corresponding linear Hooke's law spring [54]. Conversely, the case $\alpha < 0$ is called a *softening spring*. In contrast to the case of a linear oscillator near resonance (see Fig. 4.1), for $\alpha \neq 0$, the amplitude of the forced solution is shifted away from the oscillator's natural frequency by the amplitude-dependent detuning, $\gamma(M)$. For hardening (softening) springs, the resonant peak gets shifted above (below) the (zero-amplitude) linear natural frequency. See below for a plot of the detuning relation plotted for a hardening spring case, $\alpha > 0$ along with the (undamped) resonant "backbone" curve $\gamma = \frac{3}{2}\alpha M^2$.



9.6 The amplitude equations are $\frac{d\Phi}{d\tau} = 0$ and $\frac{dR}{d\tau} = -\frac{4R^2}{3\pi}$ yielding the solution

$$x(t) \sim \frac{\sin(t)}{1+\frac{4\varepsilon t}{3\pi}}.$$

9.7 The amplitude equations are $\frac{dA}{d\tau} + \frac{1}{8}A = 0$, $\frac{dB}{d\tau} - \frac{1}{8}B = 0$.

9.8 (c, d) The amplitude equation is determined by forcing the coefficient of the resonant forcing term (e^{-i4t}) in the equation for X_1 to vanish, $\frac{dA}{d\tau} - \frac{1}{2}A^2 = 0$, yielding the solution $x \sim \frac{2}{1-i-\varepsilon t} e^{-i4t}$.

9.10 Substituting the expansion into the ODE, at leading order we get

$$X_{0tt} + X_0 = \varepsilon^{1-\beta} X_0^2 + \varepsilon^{1+\beta} \cos t.$$

The inhomogeneous forcing terms will be higher order if $-1 < \beta < 1$ to yield $X_0(t, \tau) = A(\tau)e^{it} + B(\tau)e^{-it}$. In the equation for X_1, deferring the resonant terms to higher order requires $\alpha - \gamma > 0$ and $1 + \beta - \gamma > 0$ with $1 - \beta - \gamma = 0$ needed to make the remaining forcing term be $O(1)$. Finally, in the equation for X_2 balancing resonant terms leads to $\alpha = 4/3$, $\beta = 1/3$, $\gamma = 2/3$ with amplitude equations

$$-2i\frac{dA}{d\tau} + \frac{1}{2} + \frac{10}{3}A^2 B = 0 \qquad 2i\frac{dB}{d\tau} + \frac{1}{2} + \frac{10}{3}AB^2 = 0,$$

yielding a stable limit cycle with $X_0(t) = -(6/5)^{1/3} \cos t$.

9.11 (a) The trivial solution $x \equiv 0$ is an equilibrium state. Linear stability analysis yields $\lambda^2 + 1 + \varepsilon e^{-2\lambda} = 0$ with $O(1)$ solutions $\lambda \sim \pm i \pm \varepsilon \frac{1}{2} i e^{\mp 2i}$, both of these have $\mathrm{Re}(\lambda) > 0$ so the trivial solution is unstable. (b) The amplitude equations are

$$\frac{dA}{d\tau} - \tfrac{1}{2}(A \sin 2 + B \cos 2) = 0, \qquad \frac{dB}{d\tau} + \tfrac{1}{2}(A \cos 2 - B \sin 2) = 0.$$

9.12 (a) Taking ∂_t of the first equation in (9.36) yields $F_{0t} = -C_{0tt}$ and hence $C_{0tt} = C_{0xx}$ from the second equation. Taking ∂_{xx} of the first equation yields $F_{0xx} = -C_{0txx}$ then using this in ∂_t of the second equation yields $F_{0tt} = F_{0xx}$.
(b) The $O(1)$ equations determine that $f_0 = c_{0z}$ at leading order. There is one $\varepsilon^2 \partial_\tau$ term from the $O(1)$ equations that gets added to the $O(\varepsilon^2)$ (9.37) equations. The sum of ∂_z of the first equation plus the second cancels the c_1, f_1 terms to yield the KdV equation (9.38) depending only on $f_0(z, \tau)$.

9.13 $\lambda^{(1)} \sim 2 - \frac{1}{2}\varepsilon$, $\lambda^{(2)} \sim -4 + \frac{3}{2}\varepsilon$

Chapter 10

10.1 (a) $f(x) = -x^3 + 9x + 3x^2$,
(b) $(x_*, y_*) = (0, 0)$ is the only equilibrium point; the eigenvalues from its linear stability are $\lambda = (9 \pm \sqrt{81 - 16\varepsilon})/(2\varepsilon) > 0$, so it is an unstable node.
(c) The slow manifold is $y_0(x_0) = (x_0^3 - 3x_0^2 - 9x_0)/4$.
(e) The slow manifold has local extrema at $y(-1) = 5/4$ and $y(3) = -27/4$. In the fast evolution stages, y is constant and x evolves from the extrema value to the other value on the slow manifold, satisfying $y = S(x)$. Solving the cubic $S(x) = 5/4$ yields $x_{\max} = 5$ and solving $S(x) = -27/4$ yields $x_{\min} = -3$.

(f) Implicitly differentiating the equation for the slow manifold, $\frac{d}{dt}(y = S(x))$, and applying $\frac{dy}{dt} = -x$ yields the equation for evolution on the slow manifold,

$$\frac{dx}{dt} = -\frac{4x}{3x^2 - 6x - 9} \quad \rightarrow \quad \int \frac{dx}{g(x)} = -\int \frac{3x^2 - 6x - 9}{4x}\, dx$$

The limit cycle is composed of two fast and two slow stages, $P = T_{\text{fast},1} + T_{\text{slow},2} + T_{\text{fast},3} + T_{\text{slow},4} \sim T_{\text{slow},2} + T_{\text{slow},4}$. Stage 2 starts at $x = 5$ and ends at $x = 3$,

$$T_{\text{slow},2} = \frac{1}{4}\int_5^3 \frac{9}{x} + 6 - 3x\, dx = 3 + \frac{9}{4}\ln(3/5).$$

Similarly, stage 4 starts at $x = -3$ and ends at $x = -1$,

$$T_{\text{slow},4} = \frac{1}{4}\int_{-3}^{-1} \frac{9}{x} + 6 - 3x\, dx = 6 - \frac{9}{4}\ln(3).$$

Therefore the period is $P \sim 9 - (9/4)\ln(5) \approx 5.38$.

10.2 (a) Setting $\varepsilon = 0$ in the z' equation yields the slow manifold as $z = S(y) = y - y^2 + \frac{1}{3}y^3$ and then the leading order slow system is

$$\frac{dx_0}{dt} = 2 - y_0 \qquad \frac{dy_0}{dt} = x_0 - y_0 + y_0^2 - \tfrac{1}{3}y_0^3$$

The only equilibrium point of the slow system is $(x_*, y_*) = (\frac{2}{3}, 2)$, which is a stable spiral point with eigenvalues $\lambda = \frac{1}{2}(-1 \pm 3i)$. Consequently the $t \to \infty$ solution will be $(x, y, z) \sim (\frac{2}{3}, 2, \frac{2}{3} = S(2))$.

Note that the initial condition is not on the slow manifold since $0 \neq S(3)$, therefore there will be an initial layer governed by the fast system (for the distinguished limit with $\alpha = 1$ set by the z' equation),

$$\frac{dX}{dT} = \varepsilon(2 - Y), \qquad \frac{dY}{dT} = \varepsilon(X - Z), \qquad \frac{dZ}{dT} = Y - Y^2 + \tfrac{1}{3}Y^3 - Z$$

At leading order, the initial layer will have $X_0(T) \equiv 1$, $Y_0(T) \equiv 3$, constants set by the initial conditions. Then the last relation is $Z_0' = 3 - Z_0$ yielding $Z_0(T) = 3(1 - e^{-T})$ which connections the initial condition to the point $(x, y, z) = (1, 3, 3)$ on the slow manifold.

(b) Setting $\varepsilon = 0$ in the y' equation yields the relation $x = z$. Using that result in the x' gives a second expression for z' that could be matched to the z' equation to yield $2 - y = y - y^2 + \frac{1}{3}y^3 - z$. The expression for $z = z(y)$ is monotone increasing and hence could be inverted to yield $y = y(z)$ to describe the slow manifold as a parametric curve in 3D. Implicitly differentiation gives motion on this curve as

$$\frac{dy}{dt} = \frac{2-y}{2-2y+y^2},$$

which has a stable equilibrium point at $y_* = 2$. Consequently the solution for $t \to \infty$ will approach $(x, y, z) \sim (\frac{2}{3} = x(\frac{2}{3}), 2, \frac{2}{3} = z(2))$.

Since the initial condition does not lie on the slow manifold, there will be an initial layer governed by the fast system,

$$\frac{dX}{dT} = \varepsilon(2-Y), \qquad \frac{dY}{dT} = X - Z, \qquad \frac{dZ}{dT} = \varepsilon\left(Y - Y^2 + \tfrac{1}{3}Y^3 - Z\right).$$

At leading order, the initial layer will have $X_0(T) \equiv 0, Z_0(T) \equiv 1$, constants set by the initial conditions. Then the last relation is $Y_0' = -1$, describing linear decrease of Y_0 from the initial value $Y_0(0) = 3$ until it reaches the slow manifold, at $z(y) = 1$ (at $y \approx 2.1517$).

10.3 (a) The dimensional rate equations are

$$\frac{d\mathsf{A}}{d\mathsf{T}} = -2\mathsf{k}_1\mathsf{A}^2 + 2\mathsf{k}_2\mathsf{B} - \mathsf{k}_3\mathsf{AB}, \qquad \frac{d\mathsf{B}}{d\mathsf{T}} = \mathsf{k}_1\mathsf{A}^2 - \mathsf{k}_2\mathsf{B} - \mathsf{k}_3\mathsf{AB}, \qquad \frac{d\mathsf{C}}{d\mathsf{T}} = \mathsf{k}_3\mathsf{AB}.$$

(c) The scaled problem is

$$\frac{da}{dt} = -2a^2 + 2\Pi_1 b - \Pi_2 ab, \qquad \varepsilon\frac{db}{dt} = a^2 - \Pi_1 b - \Pi_2 ab, \qquad \frac{dc}{dt} = \Pi_2 ab$$

where

$$\varepsilon = \frac{\mathsf{B}_0}{\mathsf{A}_0}, \qquad \Pi_1 = \frac{\mathsf{k}_2\mathsf{B}_0}{\mathsf{k}_1\mathsf{A}_0^2}, \qquad \Pi_2 = \frac{\mathsf{k}_3\mathsf{B}_0}{\mathsf{k}_1\mathsf{A}_0}.$$

(d) By setting $\varepsilon = 0$ in the b' equation, we get the slow manifold, $b = S(a) = a^2/(\Pi_1 + \Pi_2 a)$. Substituting the slow manifold for b in the rate equations for a', c' yields

$$\frac{da}{dt} = -\frac{3\Pi_2 a^3}{\Pi_1 + \Pi_2 a} \qquad \frac{dc}{dt} = \frac{\Pi_2 a^3}{\Pi_1 + \Pi_2 a},$$

which is consistent with the overall expectation of the sum $c + \frac{1}{3}a$ being conserved. Undoing the dimensional scalings, we get

$$\frac{d\mathsf{A}}{d\mathsf{T}} = -\frac{3\mathsf{k}_1\mathsf{k}_3\mathsf{A}^3}{\mathsf{k}_2 + \mathsf{k}_3\mathsf{A}} = -\mathsf{G}(\mathsf{A}) \qquad \text{and } \mathsf{F}(\mathsf{A}) = \tfrac{1}{3}\mathsf{G}(\mathsf{A}).$$

10.5 The reactions take the form

$$A + B \to 2P \qquad B \rightleftharpoons 2C \qquad C + A \rightleftharpoons D + P \qquad D + B \to P + C$$

The rate equations are

$$\frac{d\mathsf{A}}{d\mathsf{T}} = -\mathsf{k}_3\mathsf{AC} + \mathsf{k}_4\mathsf{DP} \qquad \frac{d\mathsf{B}}{d\mathsf{T}} = -\mathsf{k}_1\mathsf{B} + \mathsf{k}_2\mathsf{C}^2 - \mathsf{k}_5\mathsf{BD}$$
$$\frac{d\mathsf{C}}{d\mathsf{T}} = 2\mathsf{k}_1\mathsf{B} - 2\mathsf{k}_2\mathsf{C}^2 - \mathsf{k}_3\mathsf{AC} + \mathsf{k}_4\mathsf{DP} + \mathsf{k}_5\mathsf{BD}$$
$$\frac{d\mathsf{D}}{d\mathsf{T}} = \mathsf{k}_3\mathsf{AC} - \mathsf{k}_4\mathsf{DP} - \mathsf{k}_5\mathsf{BD} \qquad \frac{d\mathsf{P}}{d\mathsf{T}} = \mathsf{k}_3\mathsf{AC} - \mathsf{k}_4\mathsf{DP} + \mathsf{k}_5\mathsf{BD}$$

Using the QSSA applied to the intermediates sets $C' = D' = 0$ yielding the slow manifold-type relations

$$\mathsf{C} = \sqrt{\frac{\mathsf{k}_1\mathsf{B}}{\mathsf{k}_2}} \qquad \mathsf{D} = \frac{\mathsf{k}_3\mathsf{A}}{\mathsf{k}_4\mathsf{P} + \mathsf{k}_5\mathsf{B}}\sqrt{\frac{\mathsf{k}_1\mathsf{B}}{\mathsf{k}_2}}$$

Substituting these into the rate equation for P' yields

$$\frac{d\mathsf{P}}{d\mathsf{T}} = \frac{\alpha\mathsf{AB}^{3/2}}{\mathsf{B} + \beta\mathsf{A}} \qquad \alpha = 2\mathsf{k}_3\sqrt{\frac{\mathsf{k}_1}{\mathsf{k}_2}} \qquad \beta = \frac{\mathsf{k}_4}{\mathsf{k}_5}.$$

Chapter 11

11.1 Integrals of the Gaussian yield

$$\int_{-\infty}^{\infty} \rho(x,t)\,dx = C_1, \qquad \int_{-\infty}^{\infty} x\rho(x,t)\,dx = C_1C_3, \qquad \int_{-\infty}^{\infty} x^2\rho(x,t)\,dx = C_1\left(C_3^2 + 2(t + C_2)\right).$$

Matching these results with (11.3), (11.4) and (11.7) determines

$$C_1 = M_0 = \int f\,dx, \qquad C_3 = \frac{M_1}{M_0} = \frac{1}{M_0}\int xf\,dx, \qquad C_2 = \frac{1}{2M_0}\int x^2 f\,dx - \frac{M_1^2}{2M_0^2}.$$

11.2 (a) Integrating the PDE against x^n and applying the boundary conditions yields (noting that $\rho(|x| \to \infty) \to 0$ implies that $\partial_x\rho(|x| \to \infty) \to 0$)

$$\frac{dM_0}{dt} = 4M_0, \qquad \frac{dM_1}{dt} - 2M_0 = 4M_1, \qquad \frac{dM_2}{dt} - 4M_1 = 6M_0 + 4M_2.$$

Solving this system of linear ODEs and applying the initial conditions, $M_0(0) = \sqrt{\pi}, M_1(0) = 0, M_2(0) = \frac{1}{2}\sqrt{\pi}$ yields

$$M_0(t) = \sqrt{\pi} e^{4t}, \qquad M_1(t) = 2\sqrt{\pi} t e^{4t}, \qquad M_2(t) = \sqrt{\pi}\left(4t^2 + 6t + \tfrac{1}{2}\right) e^{4t}.$$

(b) Integrating the PDE against x^0, x^1 and applying the boundary conditions yields

$$\frac{dM_0}{dt} - 2 = -3\frac{\partial \rho}{\partial x}\bigg|_{x=0} + 4M_0, \qquad \frac{dM_1}{dt} - 2M_0 = 3 + 4M_1.$$

11.3 (a) Integrating the PDE against x^0, x^1 and applying the boundary conditions yields

$$\frac{dM_0}{dt} = -M_0, \qquad \frac{dM_1}{dt} = -M_1 - (\rho(\pi, t) - \rho(0, t)),$$

(b) Substituting the separation of variables form into the PDE yields that $\lambda_k = 1 + k^2$ for $k = 0, 1, 2, \ldots$. Namely, all modes decay exponentially since $\lambda_k > 0$, but $\lambda_0 = 1$ decays the least rapidly.

Examining the initial condition at $t = 0$, $\rho(x, 0) = f(x)$ and obtaining the Fourier cosine series coefficients (see Appendix A) yields

$$a_0 = \frac{1}{\pi}\int_0^\pi f\, dx, \qquad a_k = \frac{2}{\pi}\int_0^\pi f \cos(kx)\, dx.$$

Retaining the slowest-decaying mode from the series for $\rho(x, t)$ then yields

$$\rho(x, t) \sim \frac{e^{-t}}{\pi}\int_0^\pi f\, dx \qquad \text{as } t \to \infty.$$

(b) Integrating the PDE against x^0, x^1 and applying the boundary conditions ($\rho \to 0$ as $|x| \to \infty$) yields

$$\frac{dM_0}{dt} = -M_0, \qquad \frac{dM_1}{dt} = -M_1, \qquad \frac{dM_2}{dt} = 2M_0 - M_2.$$

11.4 (a) Integrating the PDE directly yields $M_0'(t) + \rho(\infty, t) - \rho(0, t) = -2M_0(t)$, then applying the boundary conditions yields

$$\frac{dM_0}{dt} = (\beta - 2)M_0.$$

(b) Integrating the PDE directly yields the same initial form as in part (a), but the new birth condition yields the ODE

$$\frac{dM_0}{dt} = -2M_0 + M_1.$$

Integrating the PDE against e^{-3a} similarly yields

$$\frac{dM_1}{dt} = -4M_1.$$

11.5 (a) $M_0 = \int f\,dx$, $M_1 = \int xf\,dx$.
(b) Note that the porous medium equation can be re-written as $\rho_t = \frac{1}{4}(\rho^4)_{xx}$ to make integration by parts more convenient,

$$\frac{dM_2}{dt} = \tfrac{1}{4}\int x^2(\rho^4)_{xx}\,dx = \tfrac{1}{4}\left(x^2(\rho^4)_x - 2x\rho^4\Big| + 2\int \rho^4\,dx\right) \geq 0,$$

where the boundary terms vanish due to the boundary conditions and the integral is positive-definite.

References

1. D.J. Acheson, *Elementary Fluid Dynamics* (Oxford University Press, New York, 1990)
2. W.F. Ames, *Nonlinear Partial Differential Equations in Engineering* (Academic Press, New York, 1965)
3. W.F. Ames, *Nonlinear Partial Differential Equations in Engineering*, vol. II (Academic Press, New York, 1972)
4. R. Aris, *Mathematical Modelling Techniques* (Dover Publications Inc., New York, 1994)
5. D.K. Arrowsmith, C.M. Place, *An Introduction to Dynamical Systems* (Cambridge University Press, Cambridge, 1990)
6. P.W. Atkins, *Physical Chemistry* (W.H. Freeman, New York, 1986)
7. G.L. Baker, J.A. Blackburn, *The Pendulum* (Oxford University Press, Oxford, 2005)
8. H.T. Banks, H.T. Tran, *Mathematical and Experimental Modeling of Physical and Biological Processes*. Textbooks in Mathematics (CRC Press, Boca Raton, 2009)
9. G.I. Barenblatt, *Scaling, Self-similarity, and Intermediate Asymptotics* (Cambridge University Press, Cambridge, 1996)
10. G.I. Barenblatt, *Scaling* (Cambridge University Press, Cambridge, 2003)
11. C.M. Bender, S.A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (Springer, New York, 1999)
12. B. Bhushan, *Principles and Applications of Tribology* (Wiley, New York, 2013)
13. N. Bleistein, R.A. Handelsman, *Asymptotic Expansions of Integrals*, 2nd edn. (Dover Publications Inc., New York, 1986)
14. G.W. Bluman, J.D. Cole, *Similarity Methods for Differential Equations* (Springer, New York, 1974)
15. J.M. Borwein, R.E. Crandall, Closed forms: what they are and why we care. Not. AMS **60**(1), 50–65 (2013)
16. M. Bowen, J.R. King, Dynamics of a viscous thread on a non-planar substrate. J. Eng. Math. **80**, 39–62 (2013)
17. G.E.P. Box, N.R. Draper, *Empirical Model-Building and Response Surfaces* (Wiley, New York, 1987)
18. F. Brauer, C. Castillo-Chavez, *Mathematical Models in Population Biology and Epidemiology*. Texts in Applied Mathematics, 2nd edn. vol. 40 (Springer, New York, 2012)
19. F. Brauer, C. Castillo-Chavez, *Mathematical Models for Communicable Diseases*. CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 84 (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2013)
20. M.P. Brenner, H.A. Stone, Modern classical physics through the work of G.I. Taylor. Phys. Today **53**(5), 30–35 (2000)

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9

21. R.C. Buck, *Advanced Calculus*, 3rd edn. (McGraw-Hill Book Co., New York, 1978)
22. B.J. Cantwell, *Introduction to Symmetry Analysis* (Cambridge University Press, Cambridge, 2002)
23. R.V. Churchill, J.W. Brown, *Fourier Series and Boundary Value Problems* (McGraw-Hill Book Co., New York, 1987)
24. S.J. Colley, *Vector Calculus*, 3rd edn. (Prentice Hall, New Jersey, 2006)
25. R.V. Craster, O.K. Matar, Dynamics and stability of thin liquid films. Rev. Mod. Phys. **81**(3), 1131 (2009)
26. M.C. Cross, P.C. Hohenberg, Pattern-formation outside of equilibrium. Rev. Mod. Phys. **65**(3, 2), 851–1112 (1993)
27. E. Cumberbatch, A. Fitt, *Mathematical Modeling: Case Studies from Industry* (Cambridge University Press, New York, 2001)
28. J.M. Cushing, *An Introduction to Structured Population Dynamics* (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1998)
29. N.G. de Bruijn, *Asymptotic Methods in Analysis*, 3rd edn. (Dover Publications Inc., New York, 1981)
30. R.C. DiPrima, Asymptotic methods for an infinitely long slider squeeze-film bearing. J. Tribol. **90**(1), 173–183 (1968)
31. P.G. Drazin, R.S. Johnson, *Solitons: An Introduction* (Cambridge University Press, Cambridge, 1989)
32. L. Dresner, *Applications of Lie's Theory of Ordinary and Partial Differential Equations* (Institute of Physics Publishing, Bristol, 1999)
33. E. Weinan, *Principles of Multiscale Modeling* (Cambridge University Press, Cambridge, 2011)
34. A. Einstein, On the method of theoretical physics. Philos. Sci. **1**(2), 163–169 (1934)
35. T. Erneux, *Applied Delay Differential Equations* (Springer, New York, 2009)
36. L.C. Evans, *An Introduction to Stochastic Differential Equations* (American Mathematical Society, Providence, 2013)
37. A.C. Fowler, *Mathematical Models in the Applied Sciences* (Cambridge University Press, Cambridge, 1997)
38. A. Friedman, W. Littman, *Industrial Mathematics* (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1994)
39. B. Friedman, *Principles and Techniques of Applied Mathematics* (Dover Publications Inc., New York, 1990)
40. H. Goldstein, *Classical Mechanics* (Addison-Wesley, London, 1980)
41. O. Gonzalez, A.M. Stuart, *A First Course in Continuum Mechanics*. Cambridge Texts in Applied Mathematics (Cambridge University Press, Cambridge, 2008)
42. P. Grindrod, *The Theory and Applications of Reaction-Diffusion Equations* (Oxford University Press, New York, 1996)
43. J. Guckenheimer, P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Springer, New York, 1990)
44. R. Haberman, *Elementary Applied Partial Differential Equations* (Prentice Hall Inc., Englewood Cliffs, 1987)
45. R. Haberman, *Mathematical Models* (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1998)
46. B.J. Hamrock, S.R. Schmid, B.O. Jacobson, *Fundamentals of Fluid Film Lubrication* (CRC press, Boca Raton, 2004)
47. E.J. Hinch, *Perturbation Methods* (Cambridge University Press, Cambridge, 1991)
48. M.H. Holmes, *Introduction to Perturbation Methods* (Springer, New York, 1995)
49. M.H. Holmes, *Introduction to the Foundations of Applied Mathematics* (Springer, New York, 2009)
50. P. Howell, G. Kozyreff, J. Ockendon, *Applied Solid Mechanics*. Cambridge Texts in Applied Mathematics (Cambridge University Press, Cambridge, 2009)
51. S. Howison, *Practical Applied Mathematics* (Cambridge University Press, Cambridge, 2005)

52. P.E. Hydon, *Symmetry Methods for Differential Equations* (Cambridge University Press, Cambridge, 2000)
53. R. Illner, C.S. Bohun, S. McCollum, T. van Roode, *Mathematical Modelling: A Case Studies Approach*, vol. 27 (American Mathematical Society, Providence, 2005)
54. D.W. Jordan, P. Smith, *Nonlinear Ordinary Differential Equations*, 3rd edn. (Oxford University Press, Oxford, 1999)
55. M. Kac, Some mathematical models in science. Science **166**, 695–699 (1969)
56. P.B. Kahn, *Mathematical Methods for Scientists and Engineers: Linear and Nonlinear Systems* (Dover Publications, New York, 2004)
57. J. Keener, J. Sneyd, *Mathematical Physiology* (Springer, New York, 1998)
58. J. Kevorkian, J.D. Cole, *Multiple Scale and Singular Perturbation Methods* (Springer, New York, 1996)
59. D.E. Kirk, *Optimal Control Theory, An Introduction* (Dover, New York, 1998)
60. P.A. Lagerstrom, *Matched Asymptotic Expansions* (Springer, New York, 1988)
61. L.G. Leal, *Advanced Transport Phenomena* (Cambridge University Press, Cambridge, 2007)
62. M. Levi, *Classical Mechanics with Calculus of Variations and Optimal Control*. Student Mathematical Library, vol. 69 (American Mathematical Society, Providence, 2014)
63. C.C. Lin, L.A. Segel, *Mathematics Applied to Deterministic Problems in the Natural Sciences* (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1988)
64. J.D. Logan, *Applied Mathematics*, 3rd edn. (Wiley-Interscience, Hoboken, 2006)
65. J.D. Logan, *An Introduction to Nonlinear Partial Differential Equations*, 2nd edn. (Wiley-Interscience, Hoboken, 2008)
66. C.R. MacCluer, *Calculus of Variations* (Pearson Prentice Hall, Upper Saddle River, 2005)
67. J.B. Marion, *Classical Dynamics of Particles and Systems* (Academic Press, New York, 1970)
68. J.E. Marsden, A.J. Tromba, *Vector Calculus*, 5th edn. (W.H. Freeman, New York, 2003)
69. R.M.M. Mattheij, S.W. Rienstra, J.H.M. ten Thije Boonkkamp, *Partial Differential Equations: Modeling, Analysis, Computation* (SIAM, Philadelphia, 2005)
70. J.D. Meiss, *Differential Dynamical Systems* (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2007)
71. M. Mesterton-Gibbons, *A Primer on the Calculus of Variations and Optimal Control Theory* (American Mathematical Society, Providence, 2009)
72. P.D. Miller, *Applied Asymptotic Analysis*. Graduate Studies in Mathematics, vol. 75 (American Mathematical Society, Providence, 2006)
73. J.A. Murdock, *Perturbations* (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1999)
74. J.D. Murray, *Mathematical Biology*, 2nd edn. (Springer, Berlin, 1993)
75. T.G. Myers, Thin films with high surface tension. SIAM Rev. **40**(3), 441–462 (1998)
76. R. Narasimha, Divide, conquer and unify. Nature **432**, 807 (2004)
77. A.H. Nayfeh, *Perturbation Methods* (Wiley-Interscience, New York, 1973)
78. A.H. Nayfeh, *Introduction to Perturbation Techniques* (Wiley-Interscience, New York, 1981)
79. H. Ockendon, J.R. Ockendon, *Viscous Flow* (Cambridge University Press, Cambridge, 1995)
80. H. Ockendon, J.R. Ockendon, *Waves and Compressible Flow* (Springer, New York, 2004)
81. J. Ockendon, S. Howison, A. Lacey, A. Movchan, *Applied Partial Differential Equations* (Oxford University Press, Oxford, 2003)
82. A. Oron, S.H. Davis, S.G. Bankoff, Long-scale evolution of thin liquid films. Rev. Mod. Phys. **69**(3), 931 (1997)
83. R.C. Pankhurst, *Dimensional Analysis and Scale Factors* (Chapman and Hall, London, 1964)
84. E.R. Pinch, *Optimal Control and the Calculus of Variations* (Oxford University Press, New York, 1993)
85. S. Ramanujan, *Modular Equations and Approximations to* π. Q. J. Math. **45** (1914), 350–372. Collected Papers of Srinivasa Ramanujan (AMS Chelsea Publication, Providence, 2000), pp. 23–39
86. A .J. Roberts, *Model Emergent Dynamics in Complex Systems* (SIAM, Philadelphia, 2014)

87. S. Schnell, P.K. Maini, Enzyme kinetics at high enzyme concentration. Bull. Math. Biol. **62**, 482–499 (2000)
88. L.I. Sedov, *Similarity and Dimensional Methods in Mechanics* (Academic Press, New York, 1959)
89. L.A. Segel, Simplification and scaling. SIAM Rev. **14**, 547–571 (1972)
90. L.A. Segel, M. Slemrod, The quasi-steady-state assumption: a case study in perturbation. SIAM Rev. **31**(3), 446–477 (1989)
91. R.A. Serway, *Physics for Scientists and Engineers* (Brooks/Cole, Boston, 2014)
92. D.R. Smith, *Singular-Perturbation Theory: An Introduction with Applications* (Cambridge University Press, Cambridge, 1985)
93. I. Stakgold, *Green's Functions and Boundary Value Problems* (Wiley, New York, 1998)
94. S.H. Strogatz, *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry, and Engineering*. Advanced Book Program (Westview Press, Cambridge, 1994)
95. T. Szirtes, *Applied Dimensional Analysis and Modelling* (McGraw-Hill, New York, 1997)
96. A.B. Tayler, *Mathematical Models in Applied Mechanics* (Oxford University Press, Oxford, 2001)
97. B.N. Taylor, A. Thompson (eds.) The International System of Units. National Institute of Standards and Technology, U.S. Department of Commerce 2008, Publication 330
98. G.I. Taylor, Dispersion of soluble matter in solvent flowing slowly through a tube. Proc. R. Soc. Lond. Ser. A **219**, 186–203 (1953)
99. J.L. Troutman, *Variational Calculus and Optimal Control* (Springer, New York, 1996)
100. A.M. Turing, The chemical basis of morphogenesis. Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci. **237**(641), 37–72 (1952)
101. M. Van Dyke, *Perturbation Methods in Fluid Mechanics* (The Parabolic Press, Stanford, 1975)
102. F. Verhulst, *Methods and Applications of Singular Perturbations* (Springer, New York, 2005)
103. D. Weaire, S. Hutzler, *The Physics of Foams* (Clarendon Press, Oxford, 1999)
104. R. Weinstock, *Calculus of Variations* (Dover Publications Inc., New York, 1974)
105. R.B. White, *Asymptotic Analysis of Differential Equations*, Revised edn. (Imperial College Press, London, 2010)
106. G.B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1999)
107. Wikiquote, Albert Einstein—Wikiquote (2015). http://en.wikiquote.org/w/index.php?title=Albert_Einstein&oldid=1859225. Accessed 15 March 2015
108. T.P. Witelski, A.J. Bernoff, Self-similar asymptotics for linear and nonlinear diffusion equations. Stud. Appl. Math. **100**(2), 153–193 (1998)
109. T.P. Witelski, Dynamics of air bearing sliders. Phys. Fluids **10**(3), 698–708 (1998)

Index

A
Advection equation, 28
Amplitude equation, 194
Asymptotic equivalence, 127
Asymptotic expansion, 129
Asymptotic matching, 247
Asymptotic stability, 11

B
Basin of attraction, 20
Beam equation, 79
Beltrami identity, 77
Benjamin-Bona-Mahony equation, 40
Bifurcation, 12
Bifurcation points, 20
Binomial expansion, 133
Bi-stability, 20
Blow-up solution, 124
Boundary layer, 149
Brachistochrone, 47, 77
Breaking wave, 35
Buckingham Pi theorem, 101
Burgers equation, 44, 95, 105
  inviscid, 34, 43, 114

C
Carleman's model, 144
Carrying capacity, 8
Cauchy momentum equation, 28
Characteristic equations, 32
Characteristic scale, 104
Conservation law, 27
  chemical, 7
Constitutive relation, 28
Continuity equation, 27
Continuum hypothesis, 3, 23
Control function, 69
Convective derivative, 25
Couette flow, 241

D
D'Alembert solution, 42
Defocusing solution, 124
Delay differential equation, 198, 244
Derived scale, 89
Differential algebraic equations, 69
Dimensional homogeneity, 87
Dimensionless parameter, 89, 92
Dimensionless variable, 89
Dispersion relation, 29, 222
Divergence theorem, 26
Dominant balance, 134
Du Bois-Reymond lemma, 27
Dynamical systems, 4

E
Entrainment, 196
Error function, 121, 125
Euler equations, 39
Eulerian description, 24
Euler–Lagrange equation, 56
Expansion fan, 35, 118
Exponentially small terms, 149

F
Fermat's principle, 47, 79
Fermi estimate, 85
Focusing solution, 124
Fourier series, 193

© Springer International Publishing Switzerland 2015

T. Witelski and M. Bowen, *Methods of Mathematical Modelling*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-23042-9

Fredholm alternative theorem, 194, 199
Free boundary, 60
Froude number, 108
Functional, 49
Fundamental Lemma of the Calculus of Variations, 53
  generalised, 58

G
Gauge function, 129
Geodesic, 75
Group velocity, 39

H
Hamiltonian, 72, 76
Hamilton-Jacobi equation, 45
Hamilton's principle, 47, 57
Holonomic problems, 68
Hyperbolic system, 31

I
Imposed scale, 89
Initial layer, 208
Inner region, 149
Inner solution, 149
Isoperimetric problems, 65

K
Korteweg de Vries equation, 181, 198

L
Lagrange multiplier, 63
Lagrangian, 49, 55
Lagrangian derivative, 25
Lagrangian description, 25
Law of mass action, 5
Legendre transform, 71, 76
Leibniz's rule, 26, 37, 61
Liénard transformation, 203
Limit cycle, 205
Linear stability analysis, 11, 222
Logistic equation, 8
Long-wave unstable, 223
Lotka-Volterra model, 9

M
Maple, 130
Material blob, 25
Material derivative, 25
Mathieu equation, 197
McKendrick model, 230
Method of multiple time scales, 191
Michaelis-Menten rate law, 209
Moment integral, 215
Monomial form, 86

N
Natural boundary condition, 59
Navier-Stokes equations, 28
Newton's second law, 4
Noether's theorem, 77
Nullcline, 203

O
Objective function, 47
Optimal control theory, 69
Order symbol
  $O()$, 51
Outer solution, 149

P
Passive transport, 24
Peclet number, 97, 225, 242
Pendulum, 106
Phase plane analysis, 4
Phase velocity, 29, 39
Plane wave, 29
Poincare–Lindstedt, 189
Poiseuille flow, 224, 241
Pontryagin maximum principle, 83
Population dynamics, 8, 23, 230
Population, age-structured, 230
Porous medium equation, 121, 182, 231, 245
Projectile motion, 87

Q
Quasi-steady state assumption (QSSA), 212, 297
Quasilinear wave equation, 34

R
Rankine-Hugoniot equation, 37
Rarefaction wave, 35, 118
Rate constant, 5
Relaxation oscillator, 202
Resonant forcing, 186, 187
Reynolds equation, 240
Reynolds transport theorem, 26

S
Scaling constants, 96
Scaling principles, 89, 90
Scaling symmetry, 115
Secular growth, 187
Semilinear wave equation, 32
Separation of variables, 222
Shock, 36
Signalling problem, 40, 243
Similarity function, 116
Similarity variable, 116
SIR model, 9
Slender body limit, 179
Slow manifold, 203
Solvability condition, 175, 190, 194
Source-type similarity solution, 119, 229, 246
State equation, 69
Strong form, 27, 54
Sub-dominant terms, 134
Surface waves, 40, 181

T
Taylor series, 10, 130
Translation invariance, 120
Transport equation, 24
Travelling wave, 28, 248
Triple deck, 163, 165

V
Van der Pol equation, 195, 202, 210
Variation
  first, 51
  second, 51, 74
Verhulst equation, 8
von Foerster model, 230

W
Weak form, 27, 54
Weakly nonlinear oscillator, 185