Hans-Görg Roos
Martin Stynes
Lutz Tobiska

SPRINGER SERIES
IN COMPUTATIONAL MATHEMATICS 24

Robust Numerical Methods for Singularly Perturbed Differential Equations

Second Edition

Springer

Springer Series in Computational Mathematics

24

Hans-Görg Roos
Martin Stynes
Lutz Tobiska

Robust Numerical Methods for Singularly Perturbed Differential Equations

Convection-Diffusion-Reaction and Flow Problems

Second Edition
With 41 Figures

Springer

Hans-Görg Roos
Technische Universität Dresden
Fakultät Mathematik und
Naturwissenschaften
Institut für Numerische Mathematik
01062 Dresden
Germany
hans-goerg.roos@tu-dresden.de

Martin Stynes
Department of Mathematics
University College Cork
Western Road
Cork
Ireland
m.stynes@ucc.ie

Lutz Tobiska
Otto-von-Guericke-Universität
Magdeburg
Fakultät für Mathematik
Institut für Analysis und Numerik
PF 4120
39016 Magdeburg
Germany
tobiska@ovgu.de

ISBN 978-3-540-34466-7 e-ISBN 978-3-540-34467-4

DOI 10.1007/978-3-540-34467-4

Springer Series in Computational Mathematics ISSN 0179-3632

Library of Congress Control Number: 2008930796

Mathematics Subject Classification (2000): Primary: 65N12, 65M12, 76D05; Secondary: 35B25, 65M06, 65M15, 65M60, 65N06, 65N15, 65N30, 76M10

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

The analysis of singular perturbed differential equations began early in the twentieth century, when approximate solutions were constructed from asymptotic expansions. (Preliminary attempts appear in the nineteenth century – see [vD94].) This technique has flourished since the mid-1960s and its principal ideas and methods are described in several textbooks; nevertheless, asymptotic expansions may be impossible to construct or may fail to simplify the given problem and then *numerical approximations* are often the only option.

The systematic study of numerical methods for singular perturbation problems started somewhat later – in the 1970s. From this time onwards the research frontier has steadily expanded, but the exposition of new developments in the analysis of these numerical methods has not received its due attention. The first textbook that concentrated on this analysis was [DMS80], which collected various results for ordinary differential equations. But after 1980 no further textbook appeared until 1996, when *three* books were published: Miller et al. [MOS96], which specializes in upwind finite difference methods on Shishkin meshes, Morton's book [Mor96], which is a general introduction to numerical methods for convection-diffusion problems with an emphasis on the cell-vertex finite volume method, and [RST96], the first edition of the present book. Nevertheless many methods and techniques that are important today, especially for partial differential equations, were developed after 1996. To give some examples, layer-adapted special meshes are frequently used, new stabilization techniques (such as discontinuous Galerkin methods and local subspace projections) are prominent, and there is a growing interest in the use of adaptive methods. Consequently contemporary researchers must comb the literature to gain an overview of current developments in this active area. In this second edition we retain the exposition of basic material that underpinned the first edition while extending its coverage to significant new numerical methods for singularly perturbed differential equations.

Our purposes in writing this introductory book are twofold. First, we present a structured and comprehensive account of current ideas in the numerical analysis of singularly perturbed differential equations. Second, this important area has many open problems and we hope that our book will stimulate their investigation. Our choice of topics is inevitably personal and reflects our own main interests.

We have learned a great deal about singularly perturbed problems from other researchers. We thank those colleagues who helped and influenced us; these include V.B. Andreev, A.E. Berger, P.A. Farrell, A. Felgenhauer, E.C. Gartland, Ch. Großmann, A.F. Hegarty, V. John, R.B. Kellogg, N. Kopteva, G. Lube, N. Madden, G. Matthies, J.J.H. Miller, K.W. Morton, F. Schieweck, G.I. Shishkin, E. Süli, and R. Vulanović; in particular Herbert Goering and Eugene O'Riordan guided our initial steps in the area. Our research colleague T. Linß deserves additional thanks for providing many of the figures in this book.

Our work was supported by the Deutsche Forschungsgemeinschaft and by the Boole Centre for Research in Informatics at the National University of Ireland, Cork. We are grateful to them, to the Mathematisches Forschungsinstitut in Oberwolfach for its hospitality, and to Springer-Verlag for its cooperation.

Contents

Notation

I	identity
L	differential operator
L^*	adjoint operator
$a(\cdot,\cdot)$	bilinear form
g, G	Green's function
V, V^*	Banach space and the corresponding dual space
V_h	finite-dimensional subspace of V
$\|\|\cdot\|\|_V$	norm on the space V
$\|\|\cdot\|\|_{*,d}$	discrete version of the norm $\|\|\cdot\|\|_*$
$r \cdot s$	scalar product of vectors in $\mathbb{R}^d$
$(\cdot,\cdot)$	scalar product in Hilbert space
$f(v)$ or $\langle f, v\rangle$	functional f applied to v
$\|\|f\|\|_*$	norm of the functional f
$U \hookrightarrow V$	continuous embedding of U in V
Ω	given space variable(s) domain
$\partial\Omega = \Gamma$	boundary of Ω
$meas\,(\Omega)$	measure of Ω
n	outward-pointing unit vector normal to $\partial\Omega$
t, T	time with $t \in (0, T)$
$Q = \Omega \times (0, T)$	given domain for nonstationary problems
$C^l(\Omega)$, $C^{l,\alpha}(\Omega)$	function spaces
$L_p(\Omega)$	function space, $1 \le p \le \infty$
$\|\|\cdot\|\|_{0,p}$	norm in $L_p(\Omega)$
$\|\|\cdot\|\|_{L_p,d}$	discrete norm in $L_p(\Omega)$
$W^{m,p}(\Omega), \|\cdot\|_{m,p,\Omega}$	Sobolev spaces and their norms
$H^l(\Omega), H_0^l(\Omega)$	Sobolev spaces $W^{1,2}(\Omega)$
$\|\|\cdot\|\|_l, \|\cdot\|_l$	norm and seminorm in $H^l(\Omega)$
$\|\|\cdot\|\|_{l,E}$	H^l-norm restricted to $E \subset \Omega$
ε	singular perturbation parameter
C	generic constant, independent of ε

$\|\| \cdot \|\|_\varepsilon$	ε-weighted $H^1(\Omega)$ norm
$\|\| \cdot \|\|_{gr}$	graph norm
∇ or $grad$	gradient
div, div $c = \nabla \cdot c$	divergence operator
$\mathcal{O}(\cdot),\ o(\cdot)$	Landau symbols
P_r	polynomials of degree at most r
P_r^{disc}	piecewise polynomials of degree at most r, discontinuous across element boundaries
Q_r	products of polynomials of degree at most r
Q_r^{disc}	products of polynomials of degree at most r, discontinuous across element boundaries
h, h_i	mesh parameter in space
τ, τ_j	mesh parameter in time
L_h	difference operator
D^+, D^-, D^0	difference quotients
$\triangle, \triangle_h$	Laplacian and its discretization
$\omega_h,\ \Omega_h$	set of meshpoints
$u, u_h, u_i, u_i^j, u_{ij}$	unknown(s)
u_0	reduced solution
I_h	interpolation operator
$u^I = I_h\, u$	nodal interpolant of u
$\pi_h u, \Pi_h u, i_h u$	quasi-interpolant of u, defined for non-smooth functions u

mesh-dependent norms are written with three vertical lines: $\|\|\| \cdot \|\|\|$

$\|\|\| \cdot \|\|\|_{SD}$	norm used in streamline diffusion finite element method
$\|\|\| \cdot \|\|\|_{CIP}$	norm used in continuous interior penalty finite element method
$\|\|\| \cdot \|\|\|_{LPS}$	norm used in local projection stabilization finite element method
$\|\|\| \cdot \|\|\|_{dG}$	norm used in discontinuous Galerkin finite element method
$\|\|\| \cdot \|\|\|_{GLS}$	norm used in the Galerkin least-squares finite element method

Introduction

Imagine a river – a river flowing strongly and smoothly. Liquid pollution pours into the water at a certain point. What shape does the pollution stain form on the surface of the river?

Two physical processes operate here: the pollution *diffuses* slowly through the water, but the dominant mechanism is the swift movement of the river, which rapidly *convects* the pollution downstream. Convection alone would carry the pollution along a one-dimensional curve on the surface; diffusion gradually spreads that curve, resulting in a long thin curved wedge shape.

When convection and diffusion are both present in a linear differential equation and convection dominates, we have a *convection-diffusion problem.*

The simplest mathematical model of a convection-diffusion problem is a two-point boundary value problem of the form

$$-\varepsilon u''(x) + a(x)u'(x) + b(x)u(x) = f(x) \quad \text{for } 0 < x < 1,$$

with $u(0) = u(1) = 0$, where ε is a small positive parameter and a, b and f are some given functions. Here the term u'' corresponds to diffusion and its coefficient $-\varepsilon$ is small. The term u' represents convection, while u and f play the rôles of a source and driving term respectively. (Spriet and Vansteenkiste [SV82] explain why diffusion and convection should be modelled by second-order and first-order derivatives respectively.)

Example 0.1. Consider the problem

$$-\varepsilon u''(x) + u'(x) = 1 \qquad \text{for} \quad 0 < x < 1, \tag{0.1}$$

with $u(0) = u(1) = 0$ and $0 < \varepsilon \ll 1$.

Suppose that we set formally $\varepsilon = 0$ here. This yields

$$u'(x) = 1 \qquad \text{for} \quad 0 < x < 1, \tag{0.2}$$

with $u(0) = u(1) = 0$. Unlike (0.1) this problem has no solution in $C^1[0, 1]$. We infer that when ε is near zero, the solution of (0.1) is badly behaved in some way. ♣

Problems like (0.1) form the subject matter of this book. They are differential equations (ordinary or partial) that depend on a small positive parameter ε and whose solutions (or their derivatives) approach a discontinuous limit as ε approaches zero. Such problems are said to be *singularly perturbed*, where we regard ε as a perturbation parameter. In more technical terms, one cannot represent the solution of a singularly perturbed differential equation as an asymptotic expansion in powers of ε.

The solutions of singular perturbation problems typically contain *layers*. Ludwig Prandtl introduced the terminology *boundary layer* at the Third International Congress of Mathematicians in Heidelberg in 1904. (Prandtl's paper, "Über Flüssigkeitsbewegung bei sehr kleiner Reibung", is one of the most influential applied mathematics papers of the 20th century.) To see how such layers arise, consider the following time-dependent Navier-Stokes problem in two space variables x and y:

$$\frac{\partial u}{\partial t} - \frac{1}{\text{Re}}\triangle u + (u \cdot \nabla)u = -\nabla p \quad \text{in the upper half-plane } y > 0, \tag{0.3a}$$

$$\nabla \cdot u = 0 \quad \text{in the same domain,} \tag{0.3b}$$

$$u = 0 \quad \text{on the boundary } y = 0, \tag{0.3c}$$

at large Reynolds number Re. One can regard the boundary $y = 0$ as a fixed plate, and we assume that the velocity u at $y = \infty$ is parallel to the x-axis with magnitude U. We seek a flow, at constant pressure p, whose velocity is parallel to the plate and independent of x. Then equation (0.3a) reduces to

$$\frac{\partial u}{\partial t} = \varepsilon \frac{\partial^2 u}{\partial y^2}, \quad \text{where} \quad \varepsilon = \frac{1}{\text{Re}}.$$

Set $\eta = y/(2\sqrt{\varepsilon t})$ and let $u(y,t) = U\, f(\eta)$. A computation leads to

$$u = 2U\,\text{erf}(\eta), \quad \text{where} \quad \text{erf}(\eta) = \frac{1}{\sqrt{\pi}}\int_0^\eta e^{-s^2} ds. \tag{0.4}$$

Equation (0.4) shows that there is a narrow region near $y = 0$ where u departs significantly from the constant flow U. We say that u has a *boundary layer* at $y = 0$. See [CM93] for a detailed discussion. Linearization of (0.3) yields an equation of the form

$$\frac{\partial u}{\partial t} - \varepsilon\triangle u + b \cdot \nabla u + cu = f,$$

where b is independent of u. Such convection-diffusion equations model many fluid flows [Hir88, KL04]; they appear in the well-known Oseen equations and in related subjects like water pollution problems [REI$^+$07], simulation of oil extraction from underground reservoirs [Ewi83], flows in chemical reactors [Alh07] and convective heat transport problems with large Péclet numbers [Jak59].

Of course, convection-diffusion equations do not arise only in fluid flows; the next illustration comes from semiconductor device simulation.

Example 0.2. The "continuity equation" for electrons [PHSM87] in a steady-state scaled model of a one-dimensional semiconductor – with several simplifying assumptions – is

$$\frac{d^2n}{dx^2} - \frac{d}{dx}\left[n\frac{d}{dx}(\psi + \log n)\right] = 0, \tag{0.5}$$

where the unknown function n is the electron concentration, and ψ (which is computed from another part of the model) is the electrostatic potential. Now $d\psi/dx$ is typically very large (perhaps 10^5) on part of its domain (see [PHSM87, Figure 2]), so the unit coefficient of the diffusion term d^2n/dx^2 will be dominated there by the convection term coefficient. That is, equation (0.5) is a convection-diffusion problem. ♣

Singularly perturbed differential equations appear in several branches of applied mathematics. (We have seen only two examples, albeit significant ones.) The analysis and numerical solution of convection-diffusion problems deservedly attracts substantial attention.

In this book, we discuss the nature of solutions of various singularly perturbed differential equations before presenting methods for their numerical solution. Thus Part I begins with an exposition of the technique of matched asymptotic expansions, which is then used to examine various classes of two-point boundary value problems. In Part II we move on to time-dependent problems with one space dimension. Elliptic and parabolic problems in several space dimensions come in Part III. Finally, Part IV discusses finite element methods for a significant applied model: the Navier-Stokes equations.

If any discretization technique is applied to a parameter-dependent problem, then the behaviour of the discretization depends on the parameter. For singularly perturbed problems, conventional techniques often lead to discretizations that are worthless if the singular perturbation parameter is close to some critical value. We are interested in *robust* methods that work for all values of the singular perturbation parameter. We therefore track carefully the dependence on this parameter of those constants that arise in consistency, stability and error estimates. Thus the philosophy of this book emphasizes *realistic error estimates.* This contrasts sharply with much published research whose analysis ignores the effect of parameter dependence. There is a growing awareness of the dangers of this neglect; in the particular case of the incompressible Navier-Stokes equations, Johnson, Rannacher and Boman [JRB95a] observe that existing analyses often contain constants that depend on exp(Re), where Re is the Reynolds number, and conclude that "in the majority of cases of interest, the existing error analysis has no meaning". We hope that the careful approach that is followed here will provide a serviceable foundation for future work.

Discretization leads to a linear or nonlinear system of equations with a large number of unknowns. Iterative methods are commonly used to solve these systems. It is important to realize that these solvers, like the underlying discretization, should be robust with respect to the singular perturbation parameter. The discretization of a convection-diffusion problem usually produces a nonsymmetric system of equations and this asymmetry complicates the linear algebra analysis. No attempt is made in this book to discuss these issues; instead the recent textbook of Elman, Silvester and Wathen [ESW05] is recommended.

In general standard notation is used for function spaces, norms, etc. (see the notation list on page XIII), but two special conventions should be noted. First, the unknown u in a singular perturbation problem depends, of course, on the perturbation parameter ε. While one must always bear this dependence in mind, it is not included in our notation; that is, we write $u(x)$ instead of, for instance, $u(x,\varepsilon)$ or $u_\varepsilon(x)$. This simplifies the notation, especially when the discretization requires the use of some indices that depend on the mesh. On the other hand, an expression like $\lim_{\varepsilon\to 0} u(x)$ then looks odd, but one should remember that the unknown u does depend on ε. Every notation has its advantages and disadvantages! Second, in our analysis it is important to declare whether or not each constant depends on ε. Thus we denote by C (sometimes subscripted or superscripted) a *generic constant* that is always *independent of the perturbation parameter and of any mesh used.* Other letters are used to denote other "constants" when such a dependence is present.

The following example illustrates our system of numbering and internal cross-referencing. In Part I, Theorem 1.4 lies in Chapter 1 (hence the numbering "1.∗"). In Part I it is referred to as "Theorem 1.4", but we call it "Theorem I.1.4" when it's referred to from outside Part I. A similar convention is used for equations, Lemmas, etc.

We assume that the reader is familiar with the basic theory of ordinary and partial differential equations, and with the jargon and usage of finite difference and finite element methods.

Finally, despite our best efforts, mistakes are undoubtedly present in this book. We invite each reader to email us `[rst-book@ovgu.de]` any corrections that s/he notices, and this information will be made publicly available at the website `[www.rst-book.ovgu.de]`.

Part I

Ordinary Differential Equations

Part I of this book deals with singularly perturbed two-point boundary value problems. This field of research is of interest in its own right and also serves as an introduction to the more complicated problems posed in higher dimensions that we shall meet later in Parts II, III and IV. An initial discussion of analytical techniques such as maximum principles, asymptotic expansions and stability estimates for the solution of the boundary value problem provides the background needed for the numerical analysis of these ordinary differential equations. Then finite difference, finite element and finite volume methods are presented and analysed, error estimates are derived in various norms, and the relevance of mesh selection is examined. The material here is explained in detail in order to lead the reader gently into this fascinating world.

1

The Analytical Behaviour of Solutions

We begin with a general form of the problem that will occupy our attention throughout most of Part I. Consider the linear two-point boundary value problem

$$Lu := -\varepsilon u'' + b(x)u' + c(x)u = f(x) \quad \text{for } x \in (d, e),$$

with the boundary conditions

$$\alpha_d u(d) - \beta_d u'(d) = \gamma_d,$$
$$\alpha_e u(e) - \beta_e u'(e) = \gamma_e.$$

Assume that the functions b, c and f are continuous. The constants α_d, α_e, β_d, β_e, γ_d and γ_e are given, and the parameter ε satisfies $0 < \varepsilon \ll 1$.

In general, one can assume homogeneous boundary conditions $\gamma_d = \gamma_e = 0$ by subtracting from u a smooth function ψ that satisfies the original boundary conditions. For example, given Dirichlet boundary conditions $u(d) = \gamma_d$ and $u(e) = \gamma_e$, take

$$\psi(x) = \gamma_d \frac{x-e}{d-e} + \gamma_e \frac{x-d}{e-d}$$

and set $u^*(x) = u(x) - \psi(x)$. Then u^* is the solution of a differential equation of the same type but with homogeneous boundary conditions.

One can also assume without loss of generality that $x \in [0, 1]$ by means of the linear transformation

$$x \mapsto \frac{x-d}{e-d}.$$

The analytical behaviour of the solution of a singularly perturbed boundary value problem depends on the nature of the boundary conditions. From the numerical analyst's point of view, the most difficult case is when these conditions are Dirichlet. We consequently pay scant attention to other boundary conditions. Thus Sections 1.1 and 1.2 investigate the singularly perturbed problem

$$Lu := -\varepsilon u'' + b(x)u' + c(x)u = f(x) \quad \text{for } x \in (0,1), \tag{1.1a}$$
$$u(0) = u(1) = 0, \quad \text{with } c(x) \geq 0 \quad \text{for } x \in [0,1]\,, \tag{1.1b}$$

under the conditions on ε, b, c and f stated earlier. This is a typical *convection-diffusion problem* (see the Introduction) because in general we assume that b is not identically zero.

We begin our study by stating three closely-related properties of differential operators $M : C^2(0,1) \to C(0,1)$. Let $w \in C^2(0,1) \cap C[0,1]$. Then M is said to be *inverse-monotone* if the inequalities

$$Mw(x) \geq 0 \quad \text{for all } x \in (0,1), \quad w(0) \geq 0, \quad w(1) \geq 0$$

together imply that $w(x) \geq 0$ for all $x \in [0,1]$. To see that the operator L of (1.1) is inverse-monotone, one argues by contradiction [GT83].

We say that M satisfies a *maximum principle* if $Mu(x) = 0$ for all $x \in (0,1)$ implies that

$$\min\{u(0), u(1), 0\} \leq u(x) \leq \max\{u(0), u(1), 0\} \quad \text{for all } x \in [0,1].$$

Inverse-monotonicity implies that L satisfies a maximum principle. It also implies that L satisfies the following *comparison principle* which for our purposes is the most useful of the three properties.

Lemma 1.1 (Comparison principle). *Let $v, w \in C^2(0,1) \cap C[0,1]$ satisfy*

$$Lw(x) \geq Lv(x) \quad \text{for all } x \in (0,1)$$

and $w(0) \geq v(0),\ w(1) \geq v(1)$. Then

$$w(x) \geq v(x) \quad \text{for all } x \in [0,1].$$

We then say that w is a *barrier function* for v. A fairly complete discussion of maximum and comparison principles for second-order elliptic problems can be found in [GT83]. Unfortunately the terminology in the literature is inconsistent, in the sense that each of the three properties above is sometimes called a maximum principle.

Lemma 1.1 implies immediately the uniqueness of classical solutions of the boundary value problem (1.1). In this one-dimensional case, the existence of a classical solution follows. The condition $c \geq 0$ cannot in general be discarded, as is evident from the problem

$$-\varepsilon u'' + \lambda u = 0 \quad \text{on } (0,1), \quad u(0) = u(1) = 0,$$

which has multiple solutions when $\lambda = -\varepsilon k^2 \pi^2,\ k = 1, 2, \ldots.$

1.1 Linear Second-Order Problems Without Turning Points

Existence and uniqueness of the classical solution u of (1.1) are now guaranteed, but the behaviour of u when ε is small is still obscure. To gain an initial insight into the structure of u when ε is near zero, we study a simple example.

Example 1.2. The boundary value problem

$$-\varepsilon u'' + u' = 1 \quad \text{on } (0,1), \quad u(0) = u(1) = 0,$$

has the solution

$$u(x) = x - \frac{\exp(-\frac{1-x}{\varepsilon}) - \exp(-\frac{1}{\varepsilon})}{1 - \exp(-\frac{1}{\varepsilon})} \,.$$

Hence, for $a \in [0,1)$,

$$\lim_{x\to a}\lim_{\varepsilon\to 0} u(x) = a = \lim_{\varepsilon\to 0}\lim_{x\to a} u(x),$$

but

$$1 = \lim_{x\to 1}\lim_{\varepsilon\to 0} u(x) \neq \lim_{\varepsilon\to 0}\lim_{x\to 1} u(x) = 0.$$

The presence of a point ($x = 1$ in this example) where such an inequality appears means that the problem is *singularly perturbed*. The inequality implies that the solution $u(x)$ changes abruptly as x approaches 1 – we say that there is a *boundary layer* at $x = 1$. See Figure 1.1. ♣

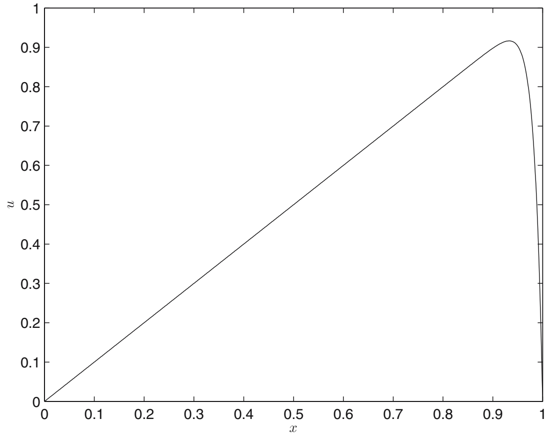

Fig. 1.1. Solution of Example 1.2 with a boundary layer at $x = 1$

1.1.1 Asymptotic Expansions

Can we approximate the solution u of (1.1) by a simple known function? Yes, by means of a standard technique in singular perturbation theory called the *method of matched asymptotic expansions*; see, for instance, [Eck73, O'M91]. The function u_{as} constructed by this technique is an *asymptotic expansion* of u; it illuminates the nature of u and thus is valuable information.

The function u_{as} is an *asymptotic expansion* of order m of u (in the maximum norm) if there is a constant C such that

$$|u(x) - u_{as}(x)| \leq C\varepsilon^{m+1} \quad \text{for all } x \in [0,1] \text{ and all } \varepsilon \text{ sufficiently small.}$$

Here we remind the reader that throughout the book C denotes a generic constant that is independent of ε. In the construction of u_{as} for (1.1), we assume that b, c and f are sufficiently smooth on [0,1].

The first step is to try to find a *global expansion* (or *regular expansion* or *outer expansion*) u_g. This function will be a good approximation of u away from any layer(s), i.e., on nearly all of the domain $[0, 1]$. We set

$$u_g(x) = \sum_{\nu=0}^{m} \varepsilon^\nu u_\nu(x), \tag{1.2}$$

where the $u_\nu(x)$ are yet to be determined. (Here, as for regular perturbations, we try to expand the solution in a Taylor-type series.) Define the operator L_0 by formally setting $\varepsilon = 0$ in L, viz.,

$$L_0 v := bv' + cv.$$

Substituting u_g into (1.1) and equating coefficients of like powers of ε yields

$$\begin{aligned} L_0 u_0 &= f, \\ L_0 u_\nu &= u''_{\nu-1} \quad \text{for} \quad \nu = 1, ..., m. \end{aligned}$$

If $b(x)$ has any zero in the interval $[0, 1]$, this causes difficulty in defining the coefficients u_ν of the global expansion because the operator L_0 then becomes singular. Zeros of b are called *turning points*. We exclude such phenomena here and defer their examination to Section 1.2.

Suppose that $b(x) \neq 0$ for all $x \in [0, 1]$. Then in principle one can calculate $u_0, u_1, \dots, u_m$ explicitly, provided that there is some additional condition on each of these functions that ensures its uniqueness. One of the boundary conditions in (1.1b) should be used to define u_0, and the crucial question is: which boundary condition should we discard? Guided by Example 1.2, we state the following *cancellation law*, which specifies the boundary condition to discard (see Section 1.4.1 for a more general formulation):

- If $b > 0$ then the boundary layer is located at $x = 1$ and to define u_0 one omits the boundary condition at $x = 1$. If $b < 0$ then the boundary layer is located at $x = 0$ and the boundary condition at $x = 0$ is dropped.

The transformation $x \mapsto 1 - x$ reduces the case $b < 0$ to $b > 0$; thus it suffices to study the case $b > 0$ in detail. The coefficients in the global expansion u_g are defined by

$$L_0 u_0 = f, \qquad u_0(0) = 0, \tag{1.3a}$$
$$L_0 u_\nu = u''_{\nu-1}, \quad u_\nu(0) = 0 \quad \text{for } \nu = 1, ..., m. \tag{1.3b}$$

We call equation (1.3a) the *reduced problem* and u_0 is the *reduced solution*. The condition $u_0(0) = 0$ comes from (1.1b), while the conditions $u_\nu(0) = 0$ for $\nu \geq 1$ ensure that $u_g(0) = u(0)$.

The aim of the method of matched asymptotic expansions is to construct an approximation of u that is valid for all $x \in [0, 1]$. But u_g cannot be such an approximation since it fails to satisfy the boundary condition at $x = 1$. Therefore one adds a local correction to u_g near $x = 1$. First, observe that the difference $w := u - u_g$ satisfies

$$Lw = \varepsilon^{m+1} u''_m,$$
$$w(0) = 0, \quad w(1) = -\sum_{\nu=0}^{m} \varepsilon^\nu u_\nu(1).$$

Write $L = \varepsilon L_1 + L_0$. Recalling that a local correction is needed near $x = 1$, where the solution u has a boundary layer, we stretch the scale there in the x direction by introducing the local variable

$$\xi = \frac{1 - x}{\delta}, \quad \text{where } \delta > 0 \text{ is small and yet to be specified.}$$

One chooses δ such that L_0 and εL_1 have formally the same order with respect to ε after the independent variable is transformed from x to ξ. That is, since $b \neq 0$, one sets

$$\varepsilon \delta^{-2} \approx \delta^{-1}.$$

This leads to the choice $\delta = \varepsilon$.

In terms of the new variable ξ, use Taylor expansions to write

$$b(1 - \varepsilon\xi) = \sum_{\nu=0}^{\infty} b_\nu \varepsilon^\nu \xi^\nu \quad \text{with } b_0 = b(1),$$
$$c(1 - \varepsilon\xi) = \sum_{\nu=0}^{\infty} c_\nu \varepsilon^\nu \xi^\nu \quad \text{with } c_0 = c(1).$$

Consequently, for any sufficiently differentiable function g, we can express L in terms of ξ as

$$\varepsilon L_1 g + L_0 g = \frac{1}{\varepsilon} \sum_{\nu=0}^{\infty} \varepsilon^\nu L_\nu^* g,$$

with

$$L_0^* := -\frac{d^2}{d\xi^2} - b_0\frac{d}{d\xi},$$
$$L_1^* := \qquad -b_1\xi\frac{d}{d\xi} + c_0,$$

etc. Now introduce the local expansion

$$v_{loc}(\xi) = \sum_{\mu=0}^{m+1} \varepsilon^\mu v_\mu(\xi). \tag{1.4}$$

In order that v_{loc} approximates $w = u - u_g$, the local corrections v_μ should satisfy the *boundary layer equations*

$$L_0^* v_0 = 0, \tag{1.5a}$$
$$L_0^* v_\mu = -\sum_{\kappa=1}^{\mu} L_\kappa^* v_{\mu-\kappa}, \quad \text{for } \mu = 1, ..., m+1. \tag{1.5b}$$

To obtain the correct boundary condition at $x = 1$, one takes $v_\kappa(0) = -u_\kappa(1)$ for $\kappa = 0, 1, \ldots, m$. As the differential equations (1.5) are of second order, a further boundary condition is also needed. To ensure the local character of the local correction, one requires that $\lim_{\xi\to\infty} v_\mu(\xi) = 0$. With these two boundary conditions the problem (1.5) has a unique solution, because the characteristic equation corresponding to L_0^* (which is a differential operator with *constant* coefficients) is

$$-\lambda^2 - b(1)\lambda = 0,$$

which has exactly one negative root. For example, the first-order correction is

$$v_0(\xi) = -u_0(1)e^{-b(1)\xi}.$$

Remark 1.3. A critical question in this method is whether or not the equations (1.5) for the local correction possess a number of decaying solutions that is equal to the number of boundary conditions that are not satisfied by the global approximation. If one cancels the wrong boundary condition when defining the reduced problem, this can lead to boundary layer equations without decaying solutions and the method then fails. ♣

Boundary layers are classified according to the nature of the boundary layer equations. The simplest layers are *exponential boundary layers* (which are sometimes called *ordinary boundary layers*), where the solutions of the boundary layer equations are decaying exponential functions. The solution of (1.1) usually has a layer of this type at $x = 1$ when $b > 0$ on $[0, 1]$.

Theorem 1.4. *If the coefficients and the right-hand side of the boundary value problem (1.1) are sufficiently smooth and* $b(x) > \beta > 0$ *on* $[0, 1]$*, then its solution* u *has a matched asymptotic expansion of the form*

$$u_{as}(x) = \sum_{\nu=0}^{m} \varepsilon^\nu u_\nu(x) + \sum_{\mu=0}^{m} \varepsilon^\mu v_\mu \left(\frac{1-x}{\varepsilon}\right), \tag{1.6}$$

such that for any sufficiently small fixed constant ε_0 one has

$$|u(x) - u_{as}(x)| \le C\varepsilon^{m+1} \quad \text{for } x \in [0,1] \text{ and } \varepsilon \le \varepsilon_0.$$

Here C is independent of x and ε.

Proof. Consider

$$u_{as}^*(x) := \sum_{\nu=0}^{m} \varepsilon^\nu u_\nu(x) + \sum_{\mu=0}^{m+1} \varepsilon^\mu v_\mu \left(\frac{1-x}{\varepsilon}\right),$$

which has an additional term for $\mu = m+1$ compared with (1.6). (This is a standard trick: if the transformed problem in the local variables has a leading term that is $O(\varepsilon^{-l})$, one considers $\sum_{\mu=0}^{m+l}$.) Our construction of the u_ν and v_μ yields

$$L(u - u_{as}^*) = O(\varepsilon^{m+1}),$$
$$(u - u_{as}^*)(0) = O(\varepsilon^\kappa), \quad (u - u_{as}^*)(1) = O(\varepsilon^{m+1}),$$

where $\kappa > 0$ is arbitrary. Now apply the comparison principle of Lemma 1.1, with the barrier function $w(x) = C\varepsilon^{m+1}(1+x)$ – this choice of w exploits the property $b \ge b_0 > 0$. We get

$$|(u - u_{as}^*)(x)| \le |w(x)| \le C\varepsilon^{m+1} \quad \text{for all } x \in [0,1].$$

But $|u_{as}(x) - u_{as}^*(x)| = |\varepsilon^{m+1} v_{m+1}((1-x)/\varepsilon)| \le C\varepsilon^{m+1}$, so a triangle inequality completes the argument. □

A formal differentiation of (1.6) leads to the following conjecture:
If b, c and f are sufficiently smooth and $b > 0$ (so turning points are excluded), the solution u of the boundary value problem (1.1) satisfies

$$|u^{(i)}(x)| \le C\left[1 + \varepsilon^{-i} \exp\left(-b(1)\frac{1-x}{\varepsilon}\right)\right].$$

A rigorous proof of the validity of this differentiation is possible [O'M91], but it is not simple. In Section 1.1.3 we shall prove a similar bound on $u^{(i)}(x)$ without using an asymptotic expansion.

Remark 1.5. (Effect of boundary conditions on the layer) In the case $b > 0$, suppose that the boundary conditions in (1.1b) are replaced by

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

Then the method of matched asymptotic expansions yields a local correction of the type

$$v_{loc}(\xi) = \varepsilon \sum_{\mu=0}^{m} \varepsilon^{\mu} v_{\mu}(\xi)$$

because, for example,

$$-\frac{\varepsilon}{b(1)} u_0'(1) e^{-b(1)\xi}$$

corrects the boundary condition at $x = 1$. One can show that:
A Dirichlet boundary condition at $x = 1$ causes a boundary layer there with

$$u'(1) = O(\varepsilon^{-1}) \quad \text{as } \varepsilon \to 0,$$

but a Neumann boundary condition at $x = 1$ causes a less severe boundary layer, since then

$$u'(1) = O(1) \text{ and } u''(1) = O(\varepsilon^{-1}) \quad \text{as } \varepsilon \to 0.$$

For example, the exact solution of

$$-\varepsilon u'' + u' = 1, \quad u(0) = 0 \text{ and } u'(1) = 0$$

is $u(x) = x - \varepsilon[e^{-(1-x)/\varepsilon} - e^{-1/\varepsilon}]$.

Under special circumstances, a different weakening of the boundary layer can occur. If, for example, the boundary condition at $x = 1$ were

$$b(1)u'(1) + c(1)u(1) = f(1)$$

– which is satisfied by the reduced solution u_0 of (1.3a) – then the asymptotic expansion of u starts with $u_0 + \varepsilon u_1 + \varepsilon^2 v_2$ because one can choose $v_0 \equiv v_1 \equiv 0$. In this particular case one has

$$u''(1) = 0 \text{ and } u'''(1) = O(\varepsilon^{-1}) \quad \text{as } \varepsilon \to 0,$$

while $u(x), u'(x)$ and $u''(x)$ are all bounded uniformly on $[0, 1]$ as $\varepsilon \to 0$. ♣

1.1.2 The Green's Function and Stability Estimates

Assume that $b(x) \geq \beta > 0$ on $[0, 1]$. The comparison principle of Lemma 1.1 provides a simple proof of the *stability estimate*

$$\|v\|_\infty \leq C\|Lv\|_\infty \quad \text{for all } v \in C^2[0, 1] \text{ with } v(0) = v(1) = 0, \tag{1.7}$$

where

$$\|z\|_\infty := \max_{x \in [0,1]} |z(x)|.$$

To prove (1.7), use $w(x) = \|Lv\|_\infty (1 + x)/\beta$ as a barrier function for v.

Note that the stability constant C in (1.7) is independent of ε. When applied to the solution u of (1.1), inequality (1.7) yields

$$\|u\|_\infty \le C\|f\|_\infty.$$

This is typical: a stability inequality implies an *a priori estimate* for the exact solution. This a priori estimate tells us that u is bounded, uniformly with respect to ε, in the maximum norm.

For the analysis of numerical methods, especially on non-equidistant meshes and in the context of a posteriori error estimates, it is very useful to have stronger stability results that use other norms. Let $(A, \|\cdot\|_A)$ and $(B, \|\cdot\|_B)$ be normed linear spaces with $M : A \to B$. Then M is said to be *uniformly* (A, B)*-stable* if

$$\|v\|_A \le C\,\|Mv\|_B \quad \text{for all } v \in A \tag{1.8}$$

with a stability constant C that is independent of ε. If $A = B$, we say simply that M is A-stable.

In this section we shall derive stability results for the convection-diffusion problem (1.1) under the hypotheses that b is continuous and does not vanish in $[0, 1]$. The (L_∞, L_1) stability result (1.19) comes from [Gar89], while the negative norm stability estimate (1.20) is in [And01, Kop01b]. We follow the presentation of [Lin02a].

Consider the boundary value problem (1.1):

$$\begin{aligned} Lu := -\varepsilon u'' + bu' + cu &= f, \\ u(0) = u(1) &= 0, \end{aligned}$$

where $b \ge \beta > 0$. Additionally, to simplify certain arguments, assume that

$$c \ge 0 \qquad \text{and} \quad c - b' \ge 0. \tag{1.9}$$

Remark 1.6. Because $b > 0$ the conditions (1.9) can always be guaranteed for ε smaller than some threshold value ε_0 by making a change of variable $u(x) = \hat{u}(x)\exp(kx)$ with the constant k chosen appropriately. ♣

The standard *Green's function* $G(x, \xi)$ associated with L and homogeneous Dirichlet boundary conditions is for each fixed $\xi \in [0, 1]$ the solution of

$$(L\,G(\cdot, \xi))(x) = \delta(x - \xi) \text{ for } x \in (0, 1), \quad G(0, \xi) = G(1, \xi) = 0, \tag{1.10}$$

where δ is the Dirac-δ distribution. Equivalently, to avoid introducing distributions, for fixed ξ one seeks a classical solution in $C^2((0, 1) \setminus \{\xi\}) \cap C[0, 1]$ that satisfies

$$(L\,G(\cdot, \xi))(x) = 0 \text{ for } x \in (0, 1) \setminus \{\xi\}, \quad G(0, \xi) = G(1, \xi) = 0, \tag{1.11}$$

and the jump condition

$$-\varepsilon[G(\cdot,\xi)'](\xi) = 1,$$

where the notation $[v](d) := v(d+0) - v(d-0)$ denotes the jump of a discontinuous function $v(x)$ at $x = d$.

In terms of the adjoint operator $L^* v := -\varepsilon v'' - (bv)' + cv$, for fixed x the Green's function $G(x,\xi)$ satisfies

$$(L^* G(x,\cdot))(\xi) = \delta(\xi - x) \text{ for } \xi \in (0,1), \quad G(x,0) = G(x,1) = 0. \tag{1.12}$$

To derive stability estimates we shall use the solution representation

$$v(x) = \int_0^1 G(x,\xi)(L\,v)(\xi)\,d\xi \tag{1.13}$$

which is valid for all v satisfying $v(0) = v(1) = 0$. Thus some bounds on G are needed.

Similarly to the classical comparison principle of Lemma 1.1, one has: if the functions v and w in $C^2((0,1) \setminus \{\xi\}) \cap C[0,1]$ satisfy

$$\begin{aligned} v(0) &\le w(0), \\ v(1) &\le w(1), \\ \mathcal{L}v(x) &\le \mathcal{L}w(x) \quad \text{in } (0,1) \setminus \{\xi\}, \\ -\varepsilon[v'](\xi) &\le -\varepsilon[w'](\xi), \end{aligned}$$

then $v(x) \le w(x)$ for all $x \in [0,1]$. This piecewise comparison principle can be found in [Mey98]; it is well known in the field of enclosing discretization methods but is rarely stated explicitly in the literature. Using the comparison principle with the barrier functions $\hat{G}_1 \equiv 0$ and

$$\hat{G}_2 = \begin{cases} (1/\beta)\exp(-\beta(\xi - x)/\varepsilon) & \text{for } 0 \le x \le \xi, \\ 1/\beta & \text{for } \xi \le x \le 1, \end{cases}$$

we get the following bounds for the Green's function:

$$0 \le G(x,\xi) \le \frac{1}{\beta} \quad \text{for } (x,\xi) \in [0,1] \times [0,1]. \tag{1.14}$$

The representation (1.13) then implies that for any function $v \in W^{2,1}(0,1)$ with $v(0) = v(1) = 0$, the stability estimate (1.7) has been sharpened to the (L_∞, L_1) estimate

$$\|v\|_\infty \le \frac{1}{\beta}\|L\,v\|_{L_1} \text{ for } v \in W_0^{1,1}(0,1) \cap W^{2,1}(0,1).$$

Here we used the notation $W^{m,p}(0,1)$ for the Sobolev space of functions defined on $[0,1]$ whose derivatives of order m are in L_p. Functions in $W_0^{1,p}(0,1)$ vanish at $x = 0$ and $x = 1$. See [Ada78] for a thorough discussion of Sobolev spaces.

We want to go one step further. For each $v \in W_0^{1,\infty}(0,1)$ let the auxiliary function $V \in L_\infty(0,1)$ satisfy $V' = L\,v$. Then an integration by parts gives

$$v(x) = -\int_0^1 G_\xi(x,\xi)V(\xi)d\xi \tag{1.15}$$

and

$$v'(x) = -\int_0^1 G_{x\xi}(x,\xi)V(\xi)d\xi. \tag{1.16}$$

These formulas are well defined: piecewise existence of $G_{x\xi}$ follows from explicit representations of G in [And01] or, alternatively, from the piecewise existence of G_{xx} and $G_{\xi\xi}$.

To extract the desired stability estimates from these representations, we need more information about the Green's function.

Since $G \geq 0$ and G satisfies the boundary conditions of (1.12), one has $G_\xi(x,0) \geq 0$ and $G_\xi(x,1) \leq 0$. Rearranging (1.12) shows that $v(\cdot) := G_\xi(x,\cdot)$ satisfies

$$\varepsilon v_\xi + bv = (c - b_\xi)G \geq 0 \quad \text{for } \xi \in (0,x) \tag{1.17}$$

where we used (1.9). As $v(0) \geq 0$, an integration of (1.17) yields $v \geq 0$ on $[0,x]$, so $G(x,\cdot)$ increases monotonically on $[0,x]$. Integrating (1.12) over $[\xi,1]$ with $\xi > x$ gives

$$\varepsilon G_\xi(x,\xi) - \varepsilon G_\xi(x,1) + b(\xi)G(x,\xi) - b(1)G(x,1) = -\int_\xi^1 c(s)G(x,s)ds.$$

Hence, using $c \geq 0$ from (1.9),

$$\varepsilon G_\xi(x,\xi) \leq \varepsilon G_\xi(x,1) - b(\xi)G(x,\xi) \leq 0\,.$$

Thus $G(x,\cdot)$ decreases monotonically on $[x,1]$.

One can prove similarly that $G_x(x,\xi) \geq 0$ for $0 \leq x < \xi \leq 1$ and $G_x(x,\xi) \leq 0$ for $0 \leq \xi < x \leq 1$. Consequently

$$G_{x\xi}(x,0) \leq 0 \quad \text{and} \quad G_{x\xi}(x,1) \leq 0 \quad \text{for } x \in (0,1).$$

For $\xi < x$ we see that $w = G_{x\xi}(x,\cdot)$ satisfies

$$\varepsilon w_\xi + bw = (c - b_\xi)G_x \leq 0.$$

It now follows from $w(0) \leq 0$ that $G_{x\xi} \leq 0$ for $\xi < x$. For $\xi > x$, differentiate the above identity:

$$\varepsilon G_{x\xi}(x,\xi) - \varepsilon G_{x\xi}(x,1) + b(\xi)G_x(x,\xi) - b(1)G_x(x,1) = -\int_\xi^1 c(s)G_x(x,s)ds.$$

This gives $G_{x\xi} \leq 0$ for $\xi > x$.

The next step is to bound the L_1 norms of G_ξ and $G_{x\xi}$ using the above monotonicity properties and the L_∞ bound (1.14). First, we get

$$\|G_\xi(x,\cdot)\|_{L_1} = \int_0^x G_\xi(x,\xi)d\xi - \int_x^1 G_\xi(x,\xi)d\xi = 2G(x,x) \le \frac{2}{\beta}. \tag{1.18}$$

A related argument shows that

$$\|G_{x\xi}(x,\cdot)\|_{L_1} = \frac{2}{\varepsilon},$$

on taking account of the singularity caused by $G_x(x,x+0)-G_x(x,x-0) = 1/\varepsilon$. These bounds can be combined with (1.15) and (1.16) to produce new stability estimates. In summary, introducing the norm

$$\|v\|_* := \inf_{V:V'=v} \|V\|_\infty,$$

the stability results we have proved in this section are the following:

Theorem 1.7. *The operator L satisfies the stability estimates*

$$\|v\|_\infty \le \frac{1}{\beta}\|L\,v\|_{L_1} \text{ for } v \in W_0^{1,1}(0,1) \cap W^{2,1}(0,1) \tag{1.19}$$

and

$$\frac{\beta}{2}\|v\|_\infty + \frac{\varepsilon}{2}\|v'\|_\infty \le \|L\,v\|_* \quad \text{for } v \in \ W_0^{1,\infty}(0,1). \tag{1.20}$$

The space $W^{-1,\infty} = (W_0^{1,1})'$ is isometrically isomorphic to the space of distributions generated by integrals of L_∞ functions and equipped with the norm $\|\cdot\|_*$; see [Ada78, Theorem 3.10]. In this sense, the norm $\|\cdot\|_*$ is the $W^{-1,\infty}$-norm and we say that (1.20) is a negative-norm stability estimate.

Now $L_1[0,1] \subset W^{-1,\infty} = (W_0^{1,1})'$. Andreev [And01, Lemma 2.6] observed that

$$\|f\|_{-1,\infty} = \sup_{1=\|v\|_{W_0^{1,1}}} \left|\int_0^1 fv\,dx\right| = \inf_C \left\|\int^1 f(s)\,ds + C\right\|_\infty = \|f\|_*.$$

Note that since

$$\|v\|_* \le \|v\|_{L_1} \le \|v\|_\infty,$$

the negative-norm bound is the strongest of our stability results.

In [And01] an assumption of the type (1.9) was not used, which makes the analysis more difficult; this paper begins with a differential equation in conservation form (assuming a different sign for the convective term)

$$\mathcal{L}v := -\varepsilon v'' - (bv)' + cv,$$

then goes on to the more complicated case where the equation is not in conservation form.

1.1.3 A Priori Estimates for Derivatives and Solution Decomposition

The numerical analysis of discretization methods requires information about higher-order derivatives of u, the solution of (1.1). Theorem 1.7 tells us that

$$|u^{(k)}(x)| \leq C\varepsilon^{-k} \quad \text{for}\, x \in [0,1],\ k = 0,1.$$

Hence, by repeated differentiation of the differential equation (1.1a), we obtain

$$|u^{(k)}(x)| \leq C\varepsilon^{-k} \quad \text{for}\, x \in [0,1],\ k = 0,1,\ldots,q,$$

where q depends on the smoothness of the data.

In general, crude bounds like these are inadequate for the job of analysing discretization methods. We now use the argument of [KT78, Lemma 2.3] to deduce a sharper estimate directly from (1.1); no asymptotic expansion is used.

Lemma 1.8. *Assume that $b(x) > \beta > 0$ and b, c, f are sufficiently smooth. Then for $i = 1, 2, \ldots, q$, the solution u of (1.1) satisfies*

$$|u^{(i)}(x)| \leq C\left[1 + \varepsilon^{-i}\exp\left(-\beta\frac{1-x}{\varepsilon}\right)\right] \quad \text{for } 0 \leq x \leq 1,$$

where the maximal order q depends on the smoothness of the data.

Proof. Set $h = f - cu$. Using an integrating factor we integrate $-\varepsilon u'' + bu' = h$ twice, obtaining

$$u(x) = u_p(x) + K_1 + K_2\int_x^1 \exp[-\varepsilon^{-1}(B(1) - B(t))]\,dt,$$

where

$$u_p(x) := -\int_x^1 z(t)\,dt, \quad z(x) := \int_x^1 \varepsilon^{-1}h(t)\exp[-\varepsilon^{-1}(B(t) - B(x))]\,dt,$$
$$B(x) := \int_0^x b(t)\,dt;$$

here the constants of integration (K_1 and K_2) may depend on ε.

The boundary condition $u(1) = 0$ implies that $K_1 = 0$. One can also see that $u'(1) = -K_2$. Now $u(0) = 0$ gives

$$K_2\int_0^1 \exp[-\varepsilon^{-1}(B(1) - B(t))]\,dt = -u_p(0). \tag{1.21}$$

The bound $\|u\|_\infty \leq C$ implied by (1.7) leads to

$$|z(x)| \le C\varepsilon^{-1} \int_x^1 \exp[-\varepsilon^{-1}(B(t) - B(x))]\, dt.$$

Applying the inequality

$$\exp[-\varepsilon^{-1}(B(t) - B(x)] \le \exp[-\beta\varepsilon^{-1}(t - x)] \quad \text{for} \quad x \le t,$$

we obtain

$$|z(x)| \le C\varepsilon^{-1} \int_x^1 \exp[-\beta\varepsilon^{-1}(t - x)]\, dt \le C.$$

Hence $|u_p(0)| \le C$. Set $\|b\|_\infty = \max_{x\in[0,1]} b(x)$. Then

$$\int_0^1 \exp[-\varepsilon^{-1}(B(1) - B(t)]\, dt) \ge \int_0^1 \exp[-\|b\|_\infty \varepsilon^{-1}(1 - t)]\, dt \ge C\varepsilon.$$

It then follows from (1.21) that $|K_2| \le C\varepsilon^{-1}$.

Now

$$u'(x) = z(x) - K_2 \exp[-\varepsilon^{-1}(B(1) - B(x))]$$

implies that

$$|u'(x)| \le C\left[1 + \varepsilon^{-1} \exp\left(-\frac{\beta(1-x)}{\varepsilon}\right)\right].$$

The bound on $u^{(i)}(x)$ for $i > 1$ follows by induction on i and repeated differentiation of (1.1a). □

A classical asymptotic expansion like that of Theorem 1.4 decomposes the solution u into a smooth part (i.e., a function for which certain low-order derivatives are bounded uniformly in ε), a layer part and a remainder. We now construct a decomposition of u into a sum of a smooth part and a layer part, with no remainder. This type of decomposition is helpful in the analysis of certain numerical methods.

The standard asymptotic expansion of Theorem 1.4 gives

$$u = u_0 + \varepsilon u_1 + \ldots + \varepsilon^k u_k + v_0 + \varepsilon v_1 + \ldots + \varepsilon^k v_k + \varepsilon^{k+1} R,$$

where R satisfies a boundary value problem similar to (1.1). Set

$$S^* := u_0 + \varepsilon u_1 + \ldots + \varepsilon^k u_k + \varepsilon^{k+1} R,$$
$$E^* := v_0 + \varepsilon v_1 + \ldots + \varepsilon^k v_k,$$

The crude estimate $\|R^{(m)}\|_\infty \le C\varepsilon^{-m}$ yields

$$|S^{*(l)}(x)| \le C \quad \text{for } l \le k + 1. \tag{1.22}$$

For the boundary layer functions, the construction of Section 1.1 leads to

$$|E^{*(l)}(x)| \le C\varepsilon^{-l} \exp\left(-\frac{\beta(1-x)}{\varepsilon}\right). \tag{1.23}$$

We call a decomposition $u = S^* + E^*$ with the properties (1.22) and (1.23) an *S-type decomposition.*

A minor modification of this construction yields an *S-decomposition*; this splitting of u enjoys the extra property that the layer part lies in the null space of L. Decompositions of this type were introduced by Shishkin in the analysis of difference schemes on piecewise equidistant meshes; see Section 2.4.2. Write

$$u = u_0 + \varepsilon u_1 + \ldots + \varepsilon^k u_k + \varepsilon^{k+1} u_{k+1}^* + v_0 + \varepsilon v_1 + \ldots + \varepsilon^k v_k + \varepsilon^{k+1} v_{k+1}^*,$$

where $u_0, \ldots, u_k, v_0, \ldots, v_k$ are the standard terms of the asymptotic expansion whereas u_{k+1}^* and v_{k+1}^* are defined by

$$Lu_{k+1}^* = u_k'', \quad u_{k+1}^*(0) = u_{k+1}^*(1) = 0$$

and

$$\begin{aligned} Lv_{k+1}^* &= -\varepsilon^{-(k+1)} L(v_0 + \varepsilon v_1 + \ldots + \varepsilon^k v_k), \\ v_{k+1}^*(0) = 0, \quad v_{k+1}^*(1) &= -(v_0 + \varepsilon v_1 + \ldots + \varepsilon^k v_k)(1). \end{aligned}$$

Now set

$$\begin{aligned} S &:= u_0 + \varepsilon u_1 + \ldots + \varepsilon^k u_k + \varepsilon^{k+1} u_{k+1}^*, \\ E &:= v_0 + \varepsilon v_1 + \ldots + \varepsilon^k v_k + \varepsilon^{k+1} v_{k+1}^*, \end{aligned}$$

and putting $q = k + 1$ we obtain

Lemma 1.9. ***(S-decomposition)*** *Let q be some positive integer. Consider the boundary value problem (1.1) with $b(x) > \beta > 0$ and sufficiently smooth data. Its solution u can be decomposed as $u = S + E$, where the smooth part S satisfies $LS = f$ and*

$$|S^{(l)}(x)| \le C \quad \text{for } 0 \le l \le q,$$

while the layer part E satisfies $LE = 0$ and

$$|E^{(l)}(x)| \le C\varepsilon^{-l} \exp\left(-\frac{\beta(1-x)}{\varepsilon}\right) \text{ for } 0 \le l \le q.$$

Clearly Lemma 1.9 implies the bounds of Lemma 1.8. Conversely, the S-decomposition of Lemma 1.9 can in fact be deduced from Lemma 1.8, as we now show. Assume the bounds of Lemma 1.8. Let $x^* = 1 - (q\varepsilon/\beta) \ln 1/\varepsilon$. Set $S(x) = u(x)$ in $[0, x^*]$. Then Lemma 1.8 implies that

$$|S^{(l)}(x)| \le C \quad \text{on } [0, x^*] \text{ for } 0 \le l \le q$$

since $e^{-\beta(1-x^*)/\varepsilon} = \varepsilon^q$. Thus one can extend the definition of S to all of $[0, 1]$ with $|S^{(l)}(x)| \le 2C$ on $[0, 1]$ for $0 \le l \le q$.

Now consider $E := u - S$. Then $E \equiv 0$ in $[0, x^*]$, while in $(x^*, 1]$ one has

$$|E^{(q)}(x)| \le |u^{(q)}(x)| + |S^{(q)}(x)| \le C\left(1 + \varepsilon^{-q}e^{-\beta(1-x)/\varepsilon}\right) \le C\varepsilon^{-q}e^{-\beta(1-x)/\varepsilon}.$$

Integrating $E^{(k)}$ for $k = q, q-1, \dots, 1$, we get inductively

$$\begin{aligned} |E^{(k-1)}(x)| &= \left| \int_{x^*}^{x} E^{(k)}(s)ds \right| \\ &\le C \int_{x^*}^{x} \varepsilon^{-k} e^{-\beta(1-s)/\varepsilon} ds \le C\varepsilon^{-(k-1)} e^{-\beta(1-x)/\varepsilon}. \end{aligned}$$

Thus $S + E$ is an S-decomposition of u.

In [Lin02b] Linß shows how to construct an S-decomposition under minimal regularity hypotheses.

Remark 1.10. (Reaction-Diffusion Problems) Consider the reaction-diffusion problem

$$-\varepsilon u'' + c(x)u = f(x) \quad \text{on } (0,1)$$

with Dirichlet boundary conditions. Assume that $c > \gamma > 0$ on $[0,1]$. Then in general the solution u contains exponential boundary layers of the form $\exp(-\sqrt{\gamma}x/\sqrt{\varepsilon})$ and $\exp(-\sqrt{\gamma}(1-x)/\sqrt{\varepsilon})$; note that these layers depend on $\sqrt{\varepsilon}$ and are present at both $x = 0$ and $x = 1$. An S-decomposition of u can be found in [MOS96, Chapter 6].

The stability properties of the reaction-diffusion operator are very different from those of the convection-diffusion operator. For instance, the Green's function of the reaction-diffusion problem with homogeneous Dirichlet conditions satisfies

$$\|G\|_\infty \le \frac{C}{\sqrt{\varepsilon}}$$

and is not bounded as $\varepsilon \to 0$. ♣

Remark 1.11. (Two-parameter convection-diffusion-reaction problems) Consider the two-parameter problem

$$-\varepsilon_1 u'' + \varepsilon_2 b(x)u' + c(x)u = f(x)$$

where ε_1 and ε_2 are small positive parameters, $b > 0$ and $c > 0$. It is shown in [LR04] that the nature of the solution decomposition depends on the relative sizes of ε_1 and ε_2. The associated Green's function satisfies

$$\|G\|_\infty \le \frac{C}{\sqrt{\varepsilon_1 + \varepsilon_2^2}};$$

see [RU03]. ♣

1.2 Linear Second-Order Turning-Point Problems

In second-order singularly perturbed differential equations, isolated points where the coefficient of u' vanishes are called *turning points*. We first look at the case of a single turning point in the interior of the domain. For convenience, the differential equation is posed on $(-1, 1)$ with its turning point placed at $x = 0$. That is, we consider

$$Lu := -\varepsilon u'' + xb(x)u' + c(x)u = f(x) \quad \text{in} \quad (-1, 1), \tag{1.24a}$$

$$u(-1) = u(1) = 0, \tag{1.24b}$$

under the following hypotheses:

$$(i) \quad b(x) \neq 0 \quad \text{on} \quad [-1, 1], \tag{1.25a}$$

$$(ii) \quad c(x) \geq 0, \quad c(0) > 0. \tag{1.25b}$$

The assumption $c(0) > 0$ simplifies the problem, as will be seen later. As in the cancellation law of page 12, the location of any boundary layer(s) depends on the sign of the convection term. From our previous experience, we expect a boundary layer at $x = -1$ if the coefficient $xb(x)$ of the convection term is negative at $x = -1$, and a boundary layer at $x = 1$ if the same coefficient is positive at $x = 1$.

If $b(x)$ is positive on $[-1, 1]$, we have $xb(x)|_{x=-1} < 0$ and $xb(x)|_{x=1} > 0$. Consequently, *if b is positive on $[-1, 1]$, then the solution u has two boundary layers.* In this case, the reduced solution is the smooth solution of

$$L_0 u_0 := xb(x)u_0' + c(x)u_0 = f(x) \quad \text{for } -1 < x < 1,$$

with no additional boundary condition! The function u_0 is well defined: use $c(0) > 0$ and a Taylor expansion about the singular point $x = 0$. Combining u_0 with two boundary layer corrections, we obtain a first-order asymptotic expansion of u, and it is straightforward to prove a result analogous to Theorem 1.4.

If the condition $c(0) > 0$ is removed, this changes the nature of the problem. In the example

$$-\varepsilon u'' + xu' = x, \quad u(-1) = u(1) = 0,$$

one finds that

$$u_0(x) = x + A,$$

with a constant A that is not determined by the method of matched asymptotic expansions. This is called a *resonance case*. The difficulty arises because $\mu_1 \to 0$ as $\varepsilon \to 0$, where μ_1 is an eigenvalue of

$$-\varepsilon w'' + xw' + \mu w = 0, \quad w(-1) = w(1) = 0.$$

See [dG76] for details of the asymptotic behaviour in this situation.

We return to the case $c(0) > 0$. Our experience in Section 1.1 leads us to expect that *if b is negative on $[-1, 1]$, then boundary layers will not occur.* In this case the reduced solution u_0 satisfies

$$L_0 u_0 = f \quad \text{in} \quad (-1, 0), \quad u_0(-1) = 0,$$

and

$$L_0 u_0 = f \quad \text{in} \quad (0, 1), \quad u_0(1) = 0.$$

The behaviour of u_0 near the turning point $x = 0$ depends strongly on the parameter $\lambda := -c(0)/b(0) > 0$. This is clearly demonstrated by the example

$$xbu_0' + cu_0 = bx^k \quad (\text{constants } b < 0 < c, \text{ integer } k > 0),$$

whose solution is

$$u_0(x) = \begin{cases} (|x|^k - |x|^\lambda)/(k - \lambda), & \text{if } \lambda \neq k, \\ x^k \ln |x|, & \text{if } \lambda = k. \end{cases}$$

At $x = 0$ the solution has an *interior layer*.

Once more, we digress to the case where $c(0) > 0$ does not hold. If $\lambda = 0$, then an interior *shock layer* in u exists, i.e., u_0 is *discontinuous*. For example, the solution of

$$-xu_0' = x$$

that satisfies $u_0(-1) = u_0(1) = 0$ is

$$u_0(x) = \begin{cases} 1 - x & \text{for } 0 < x \leq 1, \\ -1 - x & \text{for } -1 \leq x < 0. \end{cases}$$

Returning to the case $\lambda > 0$, we state without proof a result of Berger et al. [BHK84] on the behaviour of the derivatives of u (see [CL93] for a simpler argument in the case $0 < \lambda < 1$).

Lemma 1.12. *In the turning-point problem (1.24), assume that $b(x)$ is negative and λ is not an integer. Assume also that b, c and f are sufficiently smooth. Write $\lambda = m + \beta$, where m is a non-negative integer and $0 < \beta < 1$. Then the solution u of (1.24) satisfies*

$$|u^{(l)}(x)| \leq C \quad \text{on } (-1, 1) \quad \text{for } l \leq m, \tag{1.26}$$

and for $-1 < x < 1$ and $l = m + 1, m + 2, \ldots, q$,

$$|u^{(l)}(x)| \leq C \left(1 + |x| + \varepsilon^{1/2}\right)^{\lambda - l} \quad \text{on } (-1, 1). \tag{1.27}$$

Here the value of q depends on the smoothness of b, c and f.

The interior layer in u is called a *cusp layer* because it can be modelled approximately by the cusp-like function $(x^2+\varepsilon)^{\lambda/2}$. If one defines the local variable ξ in the layer by $\xi := x/\varepsilon^{1/2}$, one obtains the interior layer equation

$$-\frac{d^2v}{d\xi^2} + b(0)\xi\frac{dv}{d\xi} + c(0)v = 0.$$

The solution of this equation can be expressed in terms of parabolic cylinder functions; see [BHK84].

The problem analysed in Lemma 1.12, where the coefficient of u' has a simple zero, has a *simple turning point* at $x = 0$. If the problem has a finite number of simple turning points in $(-1, 1)$, then the result of this lemma is valid in a neighbourhood of each of these turning points. There are few stability estimates for turning-point problems in the literature; see [Doe98] for some (L_∞, L_∞) and (L_1, L_1) estimates in certain situations for simple turning points. For *multiple turning-point* problems, where the coefficient of u' has a multiple zero, less is known; see [VF93], where such a problem is discussed.

We close this section with a general L_1-norm bound on the derivative of the solution u of (1.1). No assumption is made on the sign of b so this result applies also to solutions of (1.24a).

Theorem 1.13. *For the boundary value problem (1.1), assume that b, c and f are smooth and $c(x) \ge c_0 > 0$ for $0 \le x \le 1$. Then there exists a constant C such that*

$$\int_0^1 |u'(x)|\,dx \le C. \tag{1.28}$$

Proof. The argument uses Lorenz's technique [Lor82, Nii84]. First, write (1.1) in the form

$$-\varepsilon u'' + (bu)' + (c - b')u = f$$

and differentiate, to get

$$(c - b')u' = \varepsilon u''' - (bu)'' + f' - (c' - b'')u. \tag{1.29}$$

An integration by parts then yields

$$\begin{aligned}\int_0^1 (c-b')u'\,dx &= [\varepsilon u'' - (bu)']_0^1 + \int_0^1 [f' - (c'-b'')u]\,dx \\ &= [(c-b')u - f]_0^1 + \int_0^1 [f' - (c'-b'')u]\,dx.\end{aligned}$$

Since $c(x) \ge c_0 > 0$, a comparison principle and barrier function argument gives $\|u\|_\infty \le \|f\|_\infty / c_0 = C$. Hence

$$\left|\int_0^1 (c-b')u'\,dx\right| \le C.$$

Unfortunately, this is not exactly the desired estimate and we have to modify the simple argument presented above. Thus, before integrating (1.29), multiply by $\mathrm{sgn}(u')$, where

$$\mathrm{sgn}(z) := \begin{cases} -1 & \text{if } z < 0, \\ 0 & \text{if } z = 0, \\ 1 & \text{if } z > 0. \end{cases}$$

This gives

$$\int_0^1 (c - b')|u'|\, dx = \varepsilon \int_0^1 u''' \mathrm{sgn}(u')\, dx - \int_0^1 (bu)'' \mathrm{sgn}(u')\, dx + \int_0^1 [f' - (c' - b'')u] \mathrm{sgn}(u')\, dx.$$

We would like to integrate by parts as before, but this is impossible because the function sgn is not differentiable. Thus replace sgn by a differentiable approximation s_μ, where μ is a positive parameter and $s_\mu \to \mathrm{sgn}$ as $\mu \to 0^+$. This is done by defining

$$s_\mu(z) = \begin{cases} -1 & \text{for} \quad z \le -\mu, \\ -1 + (z/\mu + 1)^2 & \text{for} \quad -\mu < z \le 0, \\ 1 - (z/\mu - 1)^2 & \text{for} \quad 0 < z < \mu, \\ 1 & \text{for} \quad z \ge \mu, \end{cases}$$

for each $\mu > 0$. For later use, observe that

$$\left| \frac{ds_\mu(z)}{dz} \right| \le \frac{C^*}{\mu} \quad \text{for all } z \in (-1, 1).$$

Replacing s above by s_μ, one obtains

$$\int_0^1 (c - b')u' s_\mu(u')\, dx = \varepsilon \int_0^1 u''' s_\mu(u')\, dx - \int_0^1 (bu)'' s_\mu(u')\, dx + \int_0^1 [f' - (c' - b'')u] s_\mu(u')\, dx.$$

Since

$$\int_0^1 u''' s_\mu(u')\, dx = u'' s_\mu(u')|_0^1 - \int_0^1 (u'')^2 \frac{ds_\mu(z)}{dz}|_{z=u'}\, dx$$

and $ds_\mu(z)/dz \ge 0$, it follows that

$$\int_0^1 u''' s_\mu(u')\, dx \le u'' s_\mu(u')|_0^1.$$

Now letting $\mu \to 0^+$ gives

$$\int_0^1 (c - b')|u'|\,dx \le \varepsilon u'' s(u')|_0^1 + \lim_{\mu\to 0^+} E + C \tag{1.30}$$

with

$$E = -\int_0^1 (bu)'' s_\mu(u')\,dx.$$

Integrating by parts, write

$$E = -(bu)' s_\mu(u')|_0^1 + E_1 + E_2, \quad \text{with} \quad E_2 = \int_0^1 b' u (s_\mu(u'))'\,dx.$$

By Lebesgue's dominated convergence theorem one has

$$\lim_{\mu\to 0+} E_2 = b' u s(u')|_0^1 - \int_0^1 b'|u'|\,dx - \int_0^1 b'' u s(u')\,dx.$$

We will show below that $\lim_{\mu\to 0^+} E_1 = 0$. Assuming this for the moment, it follows from (1.30) that

$$\int_0^1 (c - b')|u'|\,dx \le (\varepsilon u'' - bu')s(u')|_0^1 - \int_0^1 b'|u'|\,dx + C,$$

whence

$$\int_0^1 |u'|\,dx \le C,$$

since $c(x) \ge c_0 > 0$ and $\varepsilon u'' - bu' = cu - f$.

To complete the proof, consider $\lim_{\mu\to 0^+} E_1$. Now $|bu''| \le K$, where K may depend on ε, and $|(d/dz)(s_\mu(z))| \le C^*/\mu$. Hence

$$\begin{aligned} |E_1| &= \left| \int_{|u'|<\mu} bu'u'' \frac{d}{dz} s_\mu(z)|_{z=u'}\,dx \right| \\ &\le C^* K(\varepsilon)\ \text{meas}\{x \in [0,1]\ :\ 0 < |u'(x)| < \mu\}, \end{aligned}$$

which implies that $\lim_{\mu\to 0^+} E_1 = 0$. □

Theorem 1.13 is quite powerful because it makes no assumption regarding the location or multiplicity of turning points.

1.3 Quasilinear Problems

We now move on to the more general quasilinear boundary value problem

$$-\varepsilon u''(x) + b(x, u(x))u'(x) + c(x, u(x)) = 0, \quad \text{for } x \in (0,1), \tag{1.31a}$$

$$u(0) = A, \quad u(1) = B. \tag{1.31b}$$

Unlike the previous sections, inhomogeneous boundary conditions are assumed here since a transformation to homogeneous boundary conditions would alter slightly the nonlinear differential operator. In the *semilinear* case, i.e., when $b(x,u) = b(x)$, results similar to those of Sections 1.1 and 1.2 are valid.

Assume that

$$\frac{\partial c}{\partial s}(x,s) \geq \mu > 0 \quad \text{for all } x \in (0,1) \text{ and all } s \in R. \tag{1.32}$$

Then Nagumo's theory of upper and lower solutions [CH84] yields existence of a solution u of (1.31) with

$$|u(x)| \leq \max\left\{\frac{1}{\mu}\max_{x\in[0,1]}|c(x,0)|,\ |A|,\ |B|\right\} \quad \text{for all } x \in [0,1].$$

This solution is unique [O'M91].

If $b(\cdot,\cdot)$ has constant sign – say $b < 0$ – then, as in Section 1.1, we expect a boundary layer at $x = 0$. The theory is more complicated than in the linear case: one must include a pertinent *boundary layer stability assumption*, as we describe below. For the moment assume that u has a boundary layer at $x = 0$. Then the reduced solution u_R is defined by

$$b(x,u_R)u_R' + c(x,u_R) = 0 \quad \text{on } (0,1) \quad \text{with } u_R(1) = B,$$

where we assume that

$$b(x,u_R(x)) \leq -\kappa < 0 \quad \text{for all } x \in [0,1] \text{ and some } \kappa > 0.$$

With the aim of finding a boundary layer correction v_0 at $x = 0$, set $\xi = x/\varepsilon$. Then v_0 should satisfy

$$-\frac{d^2v_0}{d\xi^2} + b(0,u_R(0)+v_0)\frac{dv_0}{d\xi} = 0, \quad v_0(0) = A - u_R(0).$$

In the linear case, one can compute v_0 explicitly and see that it is exponentially decaying. But in the nonlinear case, the existence of exponentially boundary layers v_0 depends on $|A - u_R(0)|$. One needs the following additional boundary layer stability assumption [CH84, VBK95], which guarantees that the boundary layer jump $|A - u_R(0)|$ belongs to the domain of influence of the asymptotically stable solution $v_0 \equiv 0$:

$$\int_\eta^{u_R(0)} b(0,s)\,ds < 0 \quad \text{if } A < \eta < u_R(0) \tag{1.33a}$$

and

$$\int_{u_R(0)}^{\eta} b(0,s)\,ds < 0 \quad \text{if} \quad u_R(0) < \eta < A. \tag{1.33b}$$

The necessity of the inequalities (1.33) can be deduced from the implicit representation

$$\xi = \int_{v_0}^{A-u_R(0)} \frac{ds}{q(s)} \quad \text{where } q(s) = -\int_0^s b(0, u_R(0)+t)dt.$$

These conditions say essentially that the jump $|A - u_R(0)|$ should not be too large; if they are violated, then we cannot construct a boundary layer correction at $x = 0$.

A rigorous analysis leads to the following classical result [O'M91], which is due to Coddington and Levinson.

Theorem 1.14. *Assume that b and c are sufficiently smooth. Define the reduced solution u_R by*

$$b(x, u_R)u_R' + c(x, u_R) = 0 \text{ on } (0,1) \text{ with } u_R(1) = B.$$

Assume that $b(x, u_R(x)) \leq -\kappa < 0$ and that the boundary layer stability conditions (1.33) are satisfied. Then for $0 < x < 1$ one has

$$\begin{aligned} u(x) &= u_R(x) + O\left(|A - u_R(0)| \exp(-\kappa x/\varepsilon)\right) + O(\varepsilon), \\ u'(x) &= u_R'(x) + O\left(\varepsilon^{-1} \exp(-\kappa x/\varepsilon)\right) + O(\varepsilon). \end{aligned}$$

The hypotheses of Theorem 1.14 can be weakened. In particular, one can replace the condition $b(x, u_R(x)) \leq -\kappa < 0$ by the hypothesis that u_R is *globally stable*, viz., that $b(x, u_R(x)) < 0$ for $0 < x \leq 1$; see [How78, Theorem 5.5]. Analogously, if u_L is defined by

$$b(x, u_L)u_L' + c(x, u_L) = 0 \quad \text{with } u_L(0) = A,$$

we say that u_L is globally stable if $b(x, u_L(x)) > 0$ for $0 \leq x < 1$.

One can verify that the conditions of Theorem 1.14 are satisfied in the example

$$-\varepsilon u'' - e^u u' + \frac{\pi}{2} \sin \frac{\pi x}{2} e^{2u} = 0, \quad u(0) = A, \ u(1) = 0,$$

without any restriction on the boundary layer jump.

In the example

$$-\varepsilon u'' - uu' + u = 0, \quad u(0) = -2, \ u(1) = 1.5, \tag{1.34}$$

both $u_R(x) = x + 0.5$ and $u_L(x) = -2 + x$ are globally stable, but neither boundary layer stability condition (the condition for u_L is analogous to (1.33)) is satisfied:

$$\int_{A=-2}^{u_R(0)=0.5} (-s)\,ds \not< 0 \quad \text{and} \quad \int_{u_L(1)=-1}^{B=1.5} (-s)\,ds \not> 0.$$

Thus a boundary layer cannot exist at $x = 0$ nor at $x = 1$ because the boundary layer jump is too large! That is, the solution u has an interior layer but no boundary layer.

As with the linear turning-point problems of Section 1.2, we expect interior layers if no boundary layer is present. In the nonlinear case the analysis can be much more complicated than before. It is not easy to find the location(s) of possible interior layers, and the reduced equation may have more than one solution – then it is not clear which of these is the correct limit (as $\varepsilon \to 0$) of the exact solution u in a given subinterval and where a transition from one reduced solution to another takes place. A discontinuous transition will cause a *shock* layer in the solution u, and a continuous transition a *corner* layer.

We sketch the situation for the problem

$$-\varepsilon u'' + b(u)u' + c(x,u) = 0 \quad \text{for } x \in (0,1), \tag{1.35a}$$

$$u(0) = A, \quad u(1) = B, \tag{1.35b}$$

under the hypothesis (1.32). It is easier to handle (1.35) than (1.31) because the convection term can be written in the conservation form

$$b(u)u' = (e(u))', \quad \text{with } e(u) := \int^u b(s)ds.$$

The principal approach used to find the reduced solution $u_0(x) := \lim_{\varepsilon\to 0} u(x)$ is a standard technique in the theory of conservation laws (see [LeV90]); these are equations of the form $u_t + (e(u))' = 0$, where t is a time variable.

Introduce the *entropy flux* $E(\cdot)$ and the convex *entropy function* $U(\cdot)$, which depend on $e(\cdot)$ above. These functions are related by

$$\frac{dE}{dz} = \frac{dU}{dz}\frac{de}{dz}\,.$$

A simple example is $U(z) = z^2/2,\ E(z) = \int^z se'(s)ds$. Another important choice is due to Kruzkov [LeV90]: set

$$U(z) = |z-k| \quad \text{and} \quad E(z) = [e(z) - e(k)]\,\mathrm{sgn}(u-k),$$

where k is an arbitrary constant. Multiplying the differential equation (1.35a) by $U'(u)$, one writes it in the form

$$\frac{d}{dx}E(u) + U'(u)c(x,u) = \varepsilon\frac{d^2}{dx^2}U(u) - \varepsilon U''(u)\left(\frac{du}{dx}\right)^2.$$

Now multiply by a smooth function φ, integrate by parts, and take the limit as $\varepsilon \to 0$. This steers us to the inequality

$$\int_0^1 \left[-E(u_0)\varphi' + U'(u_0)c(x,u_0)\varphi\right]dx \le -E(u_0)\varphi|_0^1.$$

That is, Kruzkov's choice yields

$$\int_0^1 \operatorname{sgn}(u_0 - k)\left[(e(u_0) - e(k))\varphi' - c(x, u_0)\varphi\right] dx$$
$$\geq \sum_{i=0,1} (-1)^i \operatorname{sgn}\left(u_0(i) - k\right)\left(e(u_0(i)) - e(k)\right)\varphi(i).$$

If one chooses special test functions φ, this yields [Lor84] the following convenient characterization of the reduced solution u_0 :

Theorem 1.15. *For* $0 \leq x \leq 1$, *set* $u_0(x) = \lim_{\varepsilon \to 0} u(x)$, *where* u *is the solution of (1.35). Then*
(i) If u_0 *is smooth in a subinterval, it satisfies the reduced equation*

$$b(u_0)u_0' + c(x, u_0) = 0.$$

(ii) At the boundaries $x = 0$ *and* $x = 1$, u_0 *satisfies*

$$\operatorname{sgn}\big(u_0(0) - A\big) \int_k^{u_0(0)} b(s)\, ds \leq 0 \text{ for all } k \text{ between } A \text{ and } u_0(0),$$

$$\operatorname{sgn}\big(u_0(1) - B\big) \int_k^{u_0(1)} b(s)\, ds \geq 0 \text{ for all } k \text{ between } B \text{ and } u_0(1).$$

(iii) At a discontinuity $x_* \in (0, 1)$ *of* u_0, *the following jump condition is satisfied:*

$$\operatorname{sgn}\big(u_0(x_*^+) - u_0(x_*^-)\big) \int_k^{u_0(x_*)} b(s)\, ds \geq 0$$
$$\text{for all } k \text{ between } u_0(x_*^+) \text{ and } u_0(x_*^-).$$

Part (ii) of Theorem 1.15 is closely related to the boundary layer stability conditions (1.33), and the characterization (iii) allows us to find the position of interior layers.

For example, consider the case where u_L and u_R are globally stable but no boundary layer exists. For convenience we assume that $u_L < 0 < u_R$. We expect that

$$u_0(x) = \begin{cases} u_L(x) & \text{for} \quad 0 \leq x < x_*, \\ u_R(x) & \text{for} \quad x_* < x \leq 1, \end{cases}$$

but x_* is unknown. Theorem 1.15 (iii) tells us that

$$J(x_*) = 0, \quad \text{where } J(x) := \int_{u_L(x)}^{u_R(x)} b(s)\, ds. \tag{1.36}$$

Because no boundary layer is present,

$$J(0) = \int_A^{u_R(0)} b(s) ds > 0 \quad \text{and} \quad J(1) = \int_{u_L(1)}^B b(s) ds < 0.$$

Furthermore, for some $\zeta \in (u_L, u_R)$,

$$J'(x) = b(u_R)u'_R - b(u_L)u'_L = c(x, u_L) - c(x, u_R) = c_u(x, \zeta)(u_L - u_R) < 0.$$

Hence x_* is uniquely determined by (1.36). In example (1.34),

$$J(x) = \int_{-2+x}^{x+0.5} (-s)ds = -\frac{1}{2}(5x - 3.75),$$

which delivers the value $x_* = 0.75$.

Suppose now that we know only that $b(x, u_L(x)) > 0$ on $[0, x_L)$ for some $x_L \in (0, 1)$ (i.e., u_L is stable only on $[0, x_L)$), and $b(x, u_R(x)) < 0$ on $(x_R, 1]$ for some $x_R \in (x_L, 1)$. Then one expects that

$$u_0(x) = \begin{cases} u_L(x) \text{ for } 0 \le x \le x_L, \\ u_s(x) \text{ for } x_L \le x \le x_R, \\ u_R(x) \text{ for } x_R \le x \le 1, \end{cases}$$

with u_s a smooth solution of the reduced equation and corner layers at x_L and x_R. If example (1.34) is modified to

$$-\varepsilon u'' - uu' + u = 0, \quad u(0) = -\frac{1}{2}, \; u(1) = \frac{1}{3},$$

then one gets $u_L(x) = -1/2 + x$ with $x_L = 1/2$, and $u_R(x) = x - 2/3$ with $x_R = 2/3$. In this example, $u_s \equiv 0$ and

$$u_0(x) = \begin{cases} x - \frac{1}{2} & \text{for} \quad 0 \le x \le \frac{1}{2}, \\ 0 & \text{for} \quad \frac{1}{2} \le x \le \frac{2}{3}, \\ x - \frac{2}{3} & \text{for} \quad \frac{2}{3} \le x \le 1. \end{cases}$$

We end with a stability result from [Lor82] and an *a priori* bound on the first-order derivative of the exact solution of the quasilinear problem (1.35). Define the operator T by

$$Tv := -\varepsilon v'' + b(v)v' + c(x, v).$$

Theorem 1.16. *In the boundary value problem (1.31) assume that*

$$\frac{\partial c}{\partial s}(x, s) \ge \mu > 0 \quad \textit{for all } x \in (0, 1) \textit{ and all } s \in R.$$

Then for all v and w in $C^2(0, 1)$ that satisfy $v(0) = w(0)$ and $v(1) = w(1)$, one has

$$\|v - w\|_{L_1} \le \frac{1}{\mu} ||Tv - Tw||_{L_1}.$$

Furthermore,

$$\int_0^1 |u'(x)|dx \le C.$$

The proof of the stability result uses the Green's function of the linearized problem, while the proof of the *a priori* bound for u' resembles the proof of Theorem 1.13.

1.4 Linear Higher-Order Problems and Systems

1.4.1 Asymptotic Expansions for Higher-Order Problems

Consider the linear differential equation

$$Lu := \varepsilon^{m-n}u^{(m)} + \sum_{\nu=0}^{n} a_\nu(x)u^{(\nu)} = f(x), \quad \text{for } 0 < x < 1, \tag{1.37}$$

subject to the boundary conditions

$$u^{(\mu_i)}(0) = 0, \quad \text{for } i = 1, ..., r, \tag{1.38a}$$

$$u^{(\mu_i)}(1) = 0, \quad \text{for } i = r+1, ..., m. \tag{1.38b}$$

Here m and n are positive integers with $m > n$, so the order of the differential equation decreases if one sets $\varepsilon = 0$. The boundary conditions are ordered so that $m > \mu_1 > \mu_2 > ... > \mu_r \geq 0$ and $m > \mu_{r+1} > \mu_{r+2} > ... > \mu_m \geq 0$. Furthermore, we exclude turning points by assuming that

$$a_n(x) \neq 0 \quad \text{for all } x \in [0, 1]. \tag{1.39}$$

Applying the method of matched asymptotic expansions, the leading part u_0 of the global expansion satisfies the n^{th}-order equation

$$L_0 u_0 := \sum_{\nu=0}^{n} a_\nu(x)u_0^{(\nu)} = f.$$

It is natural to attach n boundary conditions to this differential equation. That is, $m - n$ of the original m boundary conditions will be discarded and we must decide which conditions to retain.

Introduce the local variable $\xi = x/\varepsilon$ to investigate possible boundary layers at $x = 0$ (one could similarly explore the behaviour of u near $x = 1$). The leading term in the local correction is a differential equation with constant coefficients. Its characteristic equation is

$$\lambda^n \left(\lambda^{m-n} + a_n(0)\right) = 0.$$

Suppose that σ roots of this equation have negative real part and τ roots have positive real part. Two possible situations can occur [O'M91]: in the *nonexceptional* case, $\sigma + \tau = m - n$, while in the *exceptional* case there are two pure imaginary roots so $\sigma + \tau = m - n - 2$. The corresponding *cancellation law* is:

- Cancel σ boundary conditions at $x = 0$ and τ boundary conditions at $x = 1$, choosing those with the highest-order derivatives.
- In the exceptional case, also cancel from the remaining boundary conditions those two with the highest-order derivatives, provided that they belong to the same endpoint and that the selection is without ambiguity.

After the application of the cancellation law, the reduced solution is required to satisfy the remaining n boundary conditions; this defines the *reduced problem*. If the cancellation law and reduced problem are well defined then the method of matched asymptotic expansions works, but the cancellation law is not well defined in all cases.

For example, consider the boundary value problem

$$\varepsilon^2 u^{(4)} - u'' = f(x) \quad \text{for } x \in (0,1),$$

subject to the boundary conditions

$$u'''(0) = u(0) = u'(1) = u(1) = 0.$$

Here we have $\sigma = \tau = 1$ and the cancellation law is well defined. The reduced problem is

$$-u_0'' = f \quad \text{with} \quad u_0(0) = u_0(1) = 0.$$

This has a unique solution. We find that $u(x)$ has an asymptotic expansion of the form

$$u_{as}(x) = \sum_{\nu=0}^{m} u_\nu(x)\varepsilon^\nu + \varepsilon^3 \left(\sum_{\mu=0}^{m} v_\mu(\xi)\varepsilon^\mu \right) e^{-x/\varepsilon} + \varepsilon \left(\sum_{\mu=0}^{m} w_\mu(\zeta)\varepsilon^\mu \right) e^{-(1-x)/\varepsilon},$$

for arbitrary m, with $\xi = x/\varepsilon$ and $\zeta = (1-x)/\varepsilon$. This expansion can be formally differentiated to get information about derivatives of u; see [O'M91].

Little is known about higher-order problems with turning points.

1.4.2 A Stability Result

Stability is an essential property of every discretization method and to get some insight into this property one must study the stability properties of the given continuous problem. Furthermore, asymptotic expansions require high smoothness of the coefficients of the problem; consequently, they may fail to provide sufficient information about derivatives of the exact solution for the analysis of discretization methods.

We consider the boundary value problem (1.37)–(1.38), under the assumption (1.39), for the case $n = m-1$. That is, the order of the differential equation decreases by one if $\varepsilon = 0$. We introduce the abbreviation

$$Bu = (B_1 u, B_2 u, ..., B_m u) = 0$$

for the m boundary conditions (1.38) and define the norm

$$\|v\|_{\varepsilon,m-1,\infty} := \max\left\{ \|v\|_\infty, \|v'\|_\infty, ..., \|v^{(m-2)}\|_\infty, \varepsilon\|v^{(m-1)}\|_\infty \right\}.$$

Remark 1.17. It is possible to replace $\varepsilon\|v^{(m-1)}\|_\infty$ by $\|v^{(m-1)}\|_{L_1}$. ♣

Niederdrenk and Yserentant [NY83] prove the following stability estimate for continuous coefficients, and Gartland [Gar91] extends it to the case

$$a_{m-1} \in L_\infty, \quad a_0, a_1, ..., a_{m-2} \in L_1. \tag{1.40}$$

Theorem 1.18. *Assume that the boundary conditions are bounded with respect to the norm* $\|\cdot\|_{\varepsilon,m-1,\infty}$*, in the sense that*

$$|B_\nu(v)| \leq C\|v\|_{\varepsilon,m-1,\infty} \text{ for } \nu = 1, ..., m.$$

Suppose that (1.40) is satisfied. If there exists a fundamental system $\{\phi_\nu\}$ *for* $L\phi = 0$ *that satisfies*

$$\|\phi_\nu\|_{\varepsilon,m-1,\infty} \leq C,$$

and the $m \times m$ *matrix* $[B_\mu(\phi_\nu)]$ *has an inverse whose norm (induced by the discrete* L_1 *norm) can be bounded independently of* ε*, then we have the stability inequality*

$$\|v\|_{\varepsilon,m-1,\infty} \leq C\left(\|Lv\|_{L_1} + |Bv|\right).$$

The theorem is also valid for more general boundary condition functionals. Note that for (1.38), the boundedness of the boundary conditions with respect to the norm $\|\cdot\|_{\varepsilon,m-1,\infty}$ requires that

$$\mu_1 \leq m-2 \quad \text{and} \quad \mu_{r+1} \leq m-2;$$

thus the boundary conditions cannot contain the $(m-1)^{\text{th}}$ derivative.

The conditions on the fundamental system and on the inverse of the matrix $[B_\mu(\phi_\nu)]$ are opposing constraints, as can be seen from a careful study of the following example.

Example 1.19. Consider the differential operator and boundary conditions

$$Lu := \varepsilon u^{(4)} + u''', \quad Bu := (u(0), u''(0), u(1), u''(1)).$$

Then the fundamental system $\{1, x, x^2, \varepsilon^2 e^{-x/\varepsilon}\}$ satisfies the conditions of Theorem 1.18. With homogeneous boundary data, the theorem gives not only stability but also the *a priori* estimate

$$\|u\|_\infty + \|u'\|_\infty + \|u''\|_\infty + \varepsilon\|u'''\|_\infty + \|u'''\|_{L_1} \leq C\|f\|_{L_1}.$$

If, however, the boundary conditions are

$$Bu := (u(0), u'(0), u(1), u'(1)),$$

then Theorem 1.18 does *not* apply and stability holds only in some weaker norm. ♣

Little attention has been paid in the literature to the case $n \leq m-2$ for $m > 2$. See [SS95a] for some results when $n = m-2$.

1.4.3 Systems of Ordinary Differential Equations

Systems of ordinary differential equations are often discussed in books on asymptotic expansions for singularly perturbed problems: see, e.g., [O'M91, Chapter 3], [VB90, Chapter 2] or [Was65, Chapter 7]. Nevertheless in the past relatively little attention was paid to their numerical solution, although the papers [Bak69] (reaction-diffusion systems) and [AKK74] (convection-diffusion systems) are worth noting. In recent years interest in this area has grown, as we now describe.

Consider a general system of M equations:

$$L\mathbf{u} := -\varepsilon\mathbf{u}'' + B\mathbf{u}' + A\mathbf{u} = \mathbf{f} \quad \text{on } \Omega := (0,1), \tag{1.41a}$$

$$\mathbf{u}(0) = \mathbf{g}_0, \; \mathbf{u}(1) = \mathbf{g}_1, \tag{1.41b}$$

where $\mathbf{u} = (u_1, u_2, \ldots, u_M)^T$ is the unknown solution while $\mathbf{f} = (f_1, \ldots, f_M)^T$, $\mathbf{g}_0$ and $\mathbf{g}_1$ are constant column vectors, and $A = (a_{ij})$ and $B = (b_{ij})$ are $M \times M$ matrices.

The system (1.41) is said to be *weakly coupled* if the convection coupling matrix B is diagonal, i.e., the i^{th} equation of the system is

$$-\varepsilon u_i'' + b_{ii}u_i' + \sum_{j=1}^{M} a_{ij}u_j = f_i, \tag{1.42}$$

so the system is coupled only through the lower-order reaction terms.

Linß [Lin07b] allows different diffusion coefficients in different equations: $\varepsilon = \varepsilon_i$ in the i^{th} equation for $i = 1, \ldots, M$. Assume that $b_{ii}(x) \geq \beta_i > 0$ and $a_{ii}(x) \geq \alpha > 0$ on $[0,1]$ for each i. (In [Lin07b] the weaker hypothesis $|b_{ii}(x)| \geq \beta_i > 0$ is used, which permits layers in $\mathbf{u}$ at both ends of [0,1], but for brevity we won't consider this here.) Rewrite (1.42) as

$$-\varepsilon_i u_i'' + b_{ii}u_i' + a_{ii}u_i = -\sum_{j \neq i} a_{ij}u_j + f_i, \tag{1.43}$$

Then $\|u_i\|_\infty \leq \|(-\sum_{j\neq i} a_{ij}u_j + f_i)/a_{ii}\|_\infty$ by a standard maximum principle argument. Rearranging, one gets

$$\|u_i\|_\infty - \sum_{j \neq i} \left\| \frac{a_{ij}}{a_{ii}} \right\|_\infty \|u_j\|_\infty \leq \left\| \frac{f_i}{a_{ii}} \right\|_\infty \quad \text{for } i = 1, \ldots, M.$$

Define the $M \times M$ matrix $\Gamma = (\gamma_{ij})$ by $\gamma_{ii} = 1$, $\gamma_{ij} = -\|a_{ij}/a_{ii}\|_\infty$ for $i \neq j$. Assume that Γ is inverse-monotone, i.e., that $\Gamma^{-1} \geq 0$. It follows that $\|\mathbf{u}\|_\infty \leq C\|\mathbf{f}\|_\infty$ for some constant C, where $\|\mathbf{v}\|_\infty = \max_i \|v_i\|_\infty$ for $\mathbf{v} = (v_1, \ldots, v_M)^T$. One can now apply the scalar-equation analysis of Lemma 1.8 to (1.43) for each i and get

$$|u_i^{(k)}(x)| \leq C\left[1 + \varepsilon_i^{-k} e^{-\beta_i(1-x)/\varepsilon_i}\right] \quad \text{for } x \in [0,1] \text{ and } k = 0, 1.$$

Thus there is no strong interaction between the layers in the first-order derivatives of different components u_i; nevertheless the domains of these layers can overlap and this influences the construction of numerical methods for (1.41).

The system (1.41) is said to be *strongly coupled* if for some $i \in \{1, \dots, M\}$ one has $b_{ij} \neq 0$ for some $j \neq i$. Such systems do not satisfy a maximum principle of the usual type. One now gets stronger interactions between layers; see [AKK74, Lin07a, OS, OSS]. For each i assume $b_{ii}(x) \geq \beta_i > 0$ and $a_{ii}(x) \geq 0$ on $[0, 1]$. Rewrite the i^{th} equation as

$$L_i u := -\varepsilon u_i'' + b_{ii}u_i' + a_{ii}u_i = f_i + \sum_{\substack{j=1 \\ j\neq i}}^{m} \left[(b_{ij}u_j)' - (b_{ij}' + a_{ij})u_j\right], \quad (1.44a)$$

$$u_i(0) = u_i(1) = 0. \quad (1.44b)$$

For the scalar problem $L_i v = \phi$ and $v(0) = v(1) = 0$, one has by (1.20) – see [AK98, And02] for the case where (1.9) is not satisfied – the stability result $\|v\|_\infty \leq C_i \|\phi\|_{W^{-1,\infty}}$ for a certain constant C_i that depends only on b_{ii} and a_{ii}. Apply this result to (1.44) then, similarly to the analysis of (1.43), gather the $\|u_j\|_\infty$ terms to the left-hand side. Define the $M \times M$ matrix $\Upsilon = (\gamma_{ij})$ by $\gamma_{ii} = 1$, $\gamma_{ij} = -C_i[\|b_{ij}' + a_{ij}\|_{L_1} + \|b_{ij}\|_\infty]$ for $i \neq j$. Assuming that Υ is inverse monotone, we get an a priori bound on $\|\mathbf{u}\|_\infty$. Using this bound, it is shown in [OSS] that one can decompose each component of $\mathbf{u}$ similarly to (1.22) and (1.23).

For the analysis of systems of reaction-diffusion equations (i.e., $B \equiv 0$ in (1.41)), see [Bak69, LM, MS03].

2

Numerical Methods for Second-Order Boundary Value Problems

2.1 Finite Difference Methods on Equidistant Meshes

2.1.1 Classical Convergence Theory for Central Differencing

This section examines linear two-point boundary value problems that are not singularly perturbed, in order to introduce the classical terminology of finite difference methods. Thus consider the problem

$$Lu := -u'' + b(x)u' + c(x)u = f(x), \quad u(0) = u(1) = 0, \tag{2.1}$$

under the assumptions that b, c, f are smooth and $c(x) \geq 0$.

Finite difference methods will be studied on an *equidistant* grid with *mesh size* $h = 1/N$; that is, set

$$x_i = ih \quad \text{for } i = 0, 1, ..., N, \quad \text{with } x_0 = 0 \text{ and } x_N = 1.$$

(We could work equally well with almost-equidistant meshes, but for simplicity restrict ourselves to the equidistant case. See Section 2.4 for a classification of meshes and for extensions of the theory to meshes that are not almost equidistant.)

A finite difference method is a discretization of the differential equation using the *grid points* x_i, where the unknowns u_i (for $i = 0, \ldots, N$) are approximations of the values $u(x_i)$. It is natural to approximate $u'(x)$ by the *central* difference

$$(D^0u)(x) := [u(x+h) - u(x-h)]/(2h).$$

Composing the *forward* and *backward* differences

$$(D^+u)(x) := [u(x+h) - u(x)]/h \text{ and } (D^-u)(x) := [u(x) - u(x-h)]/h,$$

yields the following central approximation for $u''(x)$:

$$(D^+D^-u)(x) := [u(x+h) - 2u(x) + u(x-h)]/h^2.$$

The *order of accuracy* of every finite difference approximation depends on the smoothness of u. For instance, Taylor's formula yields

$$u(x \pm h) = u(x) \pm hu'(x) + h^2 \frac{u''(x)}{2} \pm h^3 \frac{u'''(x)}{6} + R_4,$$

with

$$R_4 = \int_x^{x \pm h} [u'''(\xi) - u'''(x)] \frac{(x \pm h - \xi)^2}{2} \, d\xi.$$

Hence

$$\left|(D^+ D^- u)(x) - u''(x)\right| \leq Kh^2 \quad \text{if } u \in C^4, \tag{2.2}$$

– this condition can be weakened to the Lipschitz continuity of u''' – and we say that D^+D^- is second-order accurate, which is sometimes written as $O(h^2)$ accurate. The order decreases if u is less smooth; for example, if one only has $u \in C^3$, then D^+D^- is first-order accurate. Using the notation

$$g_i = g(x_i), \quad \text{where } g \text{ can be } b, \, c \text{ or } f,$$

the classical *central difference scheme* for the boundary value problem (2.1) is

$$-D^+ D^- u_i + b_i D^0 u_i + c_i u_i = f_i \quad \text{for } i = 1, ..., N-1, \tag{2.3a}$$

$$u_0 = u_N = 0. \tag{2.3b}$$

This is a tridiagonal system of linear equations:

$$r_i u_{i-1} + s_i u_i + t_i u_{i+1} = f_i \quad \text{for } i = 1, ..., N-1, \quad \text{with } u_0 = u_N = 0, \tag{2.4}$$

where

$$r_i = -\frac{1}{h^2} - \frac{1}{2h} b_i, \quad s_i = c_i + \frac{2}{h^2}, \quad t_i = -\frac{1}{h^2} + \frac{1}{2h} b_i. \tag{2.5}$$

Two questions must now be tackled: what properties does the discrete problem (2.3) enjoy? What can we say about the errors $|u(x_i) - u_i|$?

Classical convergence theory for finite difference methods is based on the complementary concepts of *consistency* and *stability*. First, formally write (2.3) (or any difference scheme) as

$$L_h u_h = f_h, \tag{2.6}$$

where L_h is a matrix,

$$u_h := (u_h(x_0), u_h(x_1), ..., u_h(x_N))^T := (u_0, u_1, \ldots, u_N)^T,$$

and $f_h := (f(x_0), f(x_1), \ldots, f(x_N))^T$. Functions defined on the grid, such as u_h and f_h, are called *grid functions*. The restriction of a function $v \in C[0,1]$ to a grid function is denoted by $R_h v$, viz., $R_h v = (v(x_0), v(x_1), ..., v(x_N))$. We sometimes omit R_h when the meaning is clear. The discrete maximum norm on the space of grid functions is

$$\|v_h\|_{\infty,d} := \max_i |v_h(x_i)|.$$

Definition 2.1. *Consider a difference scheme of the form* $L_h u_h = R_h(Lu)$*, where we incorporate the boundary conditions into the scheme by taking the first and last rows of* L_h *to be identical to the first and last rows respectively of the identity matrix, with* $(R_h Lu)_0 = u_0$ *and* $(R_h Lu)_N = u_N$*. This scheme is consistent of order* k *in the discrete maximum norm if*

$$\|L_h R_h u - R_h Lu\|_{\infty,d} \leq K h^k,$$

where the positive constants K *and* k *are independent of* h*.*

One could define consistency analogously with respect to an arbitrary norm. As in (2.2), one can apply Taylor's formula to prove

Lemma 2.2. *Under the assumption* $u \in C^4[0,1]$*, the central difference scheme (2.3) is consistent of order two.*

Applying the discrete operator L_h to the error at the interior grid points yields

$$L_h(R_h u - u_h) = L_h R_h u - f_h = L_h R_h u - R_h Lu. \tag{2.7}$$

In order to estimate $R_h u - u_h$ from (2.7) and the consistency order, it is natural to introduce the concept of *stability.*

Definition 2.3. *A discrete problem* $L_h u_h = f_h$ *is stable in the discrete maximum norm, if there exists a constant* K *(the stability constant) that is independent of* h*, such that*

$$\|u_h\|_{\infty,d} \leq K \|L_h u_h\|_{\infty,d} \tag{2.8}$$

for all mesh functions u_h*.*

Note that, analogously to the continuous case, one could generalize this to (A, B) stability which is particularly important for non-equidistant meshes. Thus, to be precise, Definition 2.3 deals with (L_∞, L_∞) stability.

Our final ingredient is

Definition 2.4. *A difference method for* (2.1) *is convergent (of order* k*) in the discrete maximum norm if there exist positive constants* K *and* k *that are independent of* h *for which*

$$\|u_h - R_h u\|_{\infty,d} \leq K h^k.$$

The main result of classical convergence theory for finite difference methods now follows immediately:

$$Consistency + Stability \Longrightarrow Convergence.$$

The investigation of the order of consistency is usually based on Taylor's formula and is straightforward. But to prove stability one needs some new tools.

In classical finite difference analyses, it is standard to use the theory of *M-matrices*, which is now described; see [Boh81, OR70] for further information.

The material that follows uses the natural ordering of vectors, viz., $x \le y$ if and only if $x_i \le y_i$ for all i. Sometimes we simply write $z \ge 1$ when we mean that $z_i \ge 1$ for all i. For each matrix $A = (a_{ij})$, the inequality $A \ge 0$ means that $a_{ij} \ge 0$ for all i and j.

A matrix A for which A^{-1} exists with $A^{-1} \ge 0$ is called *inverse-monotone.*

Lemma 2.5 (Discrete comparison principle). *Let A be inverse-monotone. Then $Av \le Aw$ implies that $v \le w$.*

Proof. The argument is simple: multiply $A(v - w) = b \le 0$ by A^{-1} and use $A^{-1} \ge 0$. □

The class of M-matrices is an important subset of the class of inverse-monotone matrices.

Definition 2.6. *A matrix A is an M-matrix if its entries a_{ij} satisfy $a_{ij} \le 0$ for $i \ne j$ and its inverse A^{-1} exists with $A^{-1} \ge 0$.*

The diagonal entries of an M-matrix satisfy $a_{ii} > 0$.

While the condition $a_{ij} \le 0$ is easy to check, it may be difficult to verify directly the inequality $A^{-1} \ge 0$. Fortunately, several equivalent but more tractable characterizations of M-matrices are known. The following result is frequently used in the context of discretization methods (see [Boh81] or [AK90] for a proof).

Theorem 2.7 (M-criterion). *Let the matrix A satisfy $a_{ij} \le 0$ for $i \ne j$. Then A is an M-matrix if and only if there exists a vector $e > 0$ such that $Ae > 0$. Furthermore, we have*

$$\|A^{-1}\|_{\infty,d} \le \frac{\|e\|_{\infty,d}}{\min_k (Ae)_k} . \tag{2.9}$$

Here the matrix norm is the norm induced by the corresponding vector norm.

In Theorem 2.7 the vector e is called a *majorizing element* for the matrix A. This theorem allows us to verify that the coefficient matrix of a given discretization is an M-matrix while simultaneously estimating the stability constant from (2.9) — provided that we are able to find a majorizing element. The following recipe for construction of this element is often successful:

- Find a function $e > 0$ such that $Le(x) > 0$ for $x \in (0, 1)$ – this is a majorizing element for the differential operator L.
- Restrict e to a grid function e_h.

In general, if the first step in this method is feasible then the method will work (at least for sufficiently small h) provided the discretization is consistent to some positive order.

For homogeneous boundary conditions one usually eliminates the variables u_0 and u_N before applying Theorem 2.7.

Example 2.8. Consider the special case where $b(x) \equiv 0$ in the differential operator L of (2.1). Choose $e(x) := x(1-x)/2$. Then

$$Le(x) = 1 + c(x)e(x) \geq 1.$$

On setting $e_h := R_h e$ one obtains

$$L_h e_h \geq (1, ..., 1)^T.$$

since D^+D^- discretizes quadratic functions exactly at the interior grid points. Now inequality (2.9) provides a stability constant of 1/8. ♣

In the general case of (2.1), the construction of a majorizing element is slightly more complicated. Define $e(x)$ to be the solution of the boundary value problem

$$-w'' + b(x)w' = 1, \quad w(0) = w(1) = 0.$$

Then $e(x) > 0$ for $x \in (0,1)$ and $e(x)$ is bounded. The inequality $c(x) \geq 0$ and the consistency of the discretization imply that at the interior grid points one has

$$L_h e_h = R_h Le + (L_h e_h - R_h Le) \geq 1/2$$

for all sufficiently small h, because $R_h Le = 1$. This proves

Lemma 2.9. *For all sufficiently small h, the central difference scheme for the boundary value problem (2.1) is stable in the discrete maximum norm; moreover, the corresponding coefficient matrix is then an M-matrix.*

One can clearly combine Lemmas 2.2 (consistency) and 2.9 (stability) to obtain a second-order convergence result.

Remark 2.10. In general, the proof of stability via M-matrices is inapplicable to higher-order difference schemes that are based on stencils with more than three points. It may nevertheless be possible to use the property of strong diagonal dominance (see, e.g., [Her90]) or to factor a matrix as a product of M-matrices [Lor75] or to use special splittings [AK90]. For a general stability theory of difference schemes see [Gri85a]. ♣

2.1.2 Upwind Schemes

This subsection and its two successors study difference schemes for the singularly perturbed boundary value problem

$$Lu := -\varepsilon u'' + b(x)u' + c(x)u = f(x) \quad \text{on } (0,1), \quad u(0) = u(1) = 0, \quad (2.10)$$

when turning points are excluded, i.e., when $b(x) \neq 0$ for all $x \in [0,1]$. We also assume that $c \geq 0$ on [0,1] and that the functions b, c and f are smooth. Recall that for $b > 0$ there is an exponential boundary layer at $x = 1$, and for

$b < 0$ the boundary layer is at $x = 0$. The conditions "$b < 0$" and "$b > 0$" are equivalent: the change of variable $x \mapsto 1-x$ transforms the problem from one formulation to the other.

Suppose that $\varepsilon > 0$ is small. If u exhibits a boundary layer, this adversely affects both consistency and stability. If instead the boundary conditions are such that u has no layer, then the consistency error improves but stability may still be a problem.

To begin, the central difference scheme is applied to the example

$$-\varepsilon u'' + u' = 0 \text{ on } (0,1), \quad u(0) = 0,\ u(1) = 1.$$

A transformation $u(x) = x + v(x)$ would give homogeneous boundary conditions, but one can use the scheme directly with inhomogeneous conditions. The discrete problem is

$$-\varepsilon D^+ D^- u_i + D^0 u_i = 0, \quad u_0 = 0,\ u_N = 1.$$

It is easy to solve this exactly:

$$u_i = \frac{r^i - 1}{r^N - 1}, \quad \text{with} \quad r = \frac{2\varepsilon + h}{2\varepsilon - h}.$$

If $h \gg 2\varepsilon$, then $r \approx -1$ so this computed solution oscillates badly and is not close to the true solution

$$u(x) = \frac{e^{-(1-x)/\varepsilon} - e^{-1/\varepsilon}}{1 - e^{-1/\varepsilon}}.$$

Figure 2.1 shows the oscillations of the central scheme on an uniform mesh if ε is small compared with h. On the other hand if $h < 2\varepsilon$, then the central difference scheme works — but from the practical point of view this assumption is unsatisfactory when, for instance, $\varepsilon = 10^{-5}$. *A fortiori*, in two or three dimensions such a mesh restriction would lead to unacceptably large numbers of mesh points, as for small ε the dimension of the algebraic system generated would be too large for computer solution.

Returning to the general problem (2.10), write the central difference scheme in the form of (2.5), viz.,

$$r_i = -\frac{\varepsilon}{h^2} - \frac{1}{2h} b_i, \quad s_i = c_i + \frac{2\varepsilon}{h^2}, \quad t_i = -\frac{\varepsilon}{h^2} + \frac{1}{2h} b_i.$$

This gives an M-matrix and hence stability if we assume that

$$h \le h_0(\varepsilon) = \frac{2\varepsilon}{\|b\|_\infty},$$

which generalizes the observation of the example above. Note that $h_0(\varepsilon) \to 0$ if $\varepsilon \to 0$. This conclusion is not confined to the central difference scheme: *Classical numerical methods on equidistant grids yield satisfactory numerical*

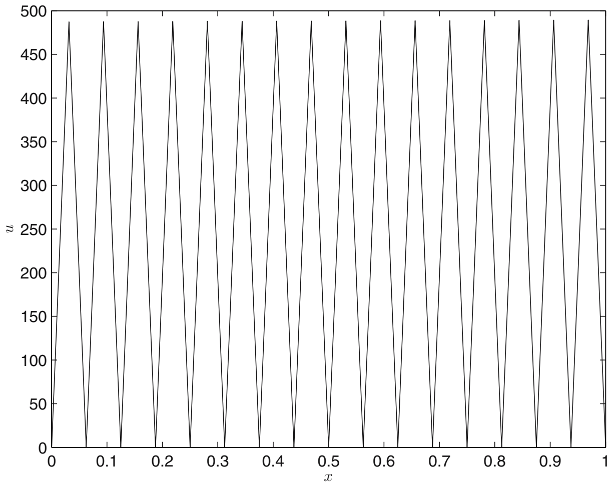

Fig. 2.1. Oscillations of the central difference scheme

solutions for singularly perturbed boundary value problems only if one uses an unacceptably large number of grid points. In this sense, *classical methods fail.*

An alternative heuristic explanation for the failure of central differencing in the above example is that when $\varepsilon \ll h$ the scheme is essentially $D^0 u_i = 0$, which implies in particular that $u_{N-2} \approx u_N = 1$, so u_{N-2} is a poor approximation to $u(x_{N-2}) \approx 0$.

This argument also shows that we would do well to avoid any difference approximation of $u'(x_{N-1})$ that uses u_N. The simplest candidate meeting this requirement is the approximation

$$u'(x_i) \approx \frac{u_i - u_{i-1}}{h}\,. \tag{2.11}$$

An inspection of the signs of the matrix entries of the earlier discrete problem, with the aim of modifying the difference scheme in order to generate an M-matrix, also motivates (2.11).

Thus for the general case where the sign of b may be positive or negative, consider the scheme

$$-\varepsilon D^+ D^- u_i + b_i D^{\aleph} u_i + c_i u_i = f_i \quad \text{for } i = 1, ..., N-1, \tag{2.12a}$$

$$u_0 = u_N = 0, \tag{2.12b}$$

with

$$D^{\aleph} = \begin{cases} D^+ & \text{if} \quad b < 0, \\ D^- & \text{if} \quad b > 0. \end{cases} \tag{2.12c}$$

This is the *simple upwind scheme.* (We saw in the Introduction that convection dominates the problem and assigns a direction to the flow; *upwind* means that the finite difference approximation of the convection term is taken on the upstream side of each mesh point.) The numerical behaviour of the upwind scheme is much better than the central scheme: see Figure 2.2.

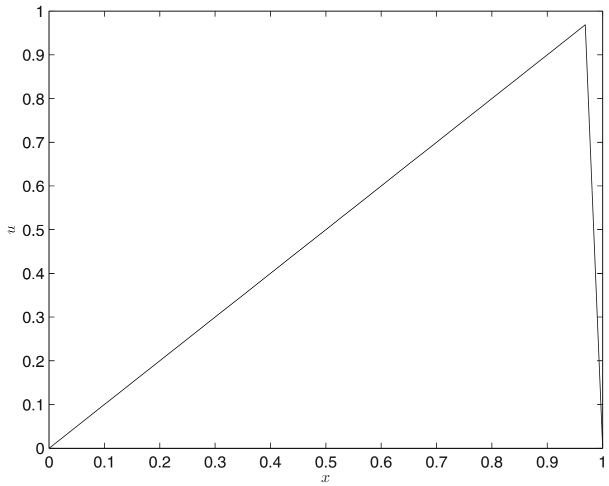

Fig. 2.2. Solution of the upwind scheme on an equidistant mesh

We now begin our analysis of the upwind scheme. Write L_h for the matrix of the scheme after eliminating u_0 and u_N. In the form (2.4), the coefficients of the discrete problem are

$$r_i = -\frac{\varepsilon}{h^2} - \frac{1}{h}\max\{0, b_i\}, \quad s_i = c_i + \frac{2\varepsilon}{h^2} + \frac{1}{h}|b_i|,$$
$$t_i = -\frac{\varepsilon}{h^2} + \frac{1}{h}\min\{0, b_i\}.$$

Now the off-diagonal matrix entries are non-positive, irrespective of the relative sizes of h and ε.

Lemma 2.11. *Assume that $b(x) \neq 0$ for all $x \in [0,1]$. Then the coefficient matrix L_h for the upwind scheme (2.12) is an M-matrix and the upwind scheme is uniformly stable with respect to the perturbation parameter:*

$$\|u_h\|_{\infty,d} \leq C\|L_h u_h\|_{\infty,d},$$

with a stability constant C that is independent of ε and h.

Proof. For definiteness assume that $b(x) \geq \beta > 0$. We construct a suitable majorizing vector. Choose $e(x) := x$, so $Le(x) \geq \beta$. A direct computation yields $L_h e_h \geq \beta$. By Theorem 2.7 the matrix is an M-matrix and one gets the desired stability bound with stability constant $C = 1/\beta$. □

This stability result for the upwind scheme remains valid on arbitrary meshes. Moreover, introducing mesh analogues of the norms previously seen, one can also prove $(L_{\infty,d}, L_{1,d})$ and $(L_{\infty,d}, W_{-1,\infty,d})$ stability results which are useful when analysing the scheme on layer-adapted meshes, as will be seen later.

In ensuring the stability of the upwind scheme, we have paid a certain price in accuracy: D^+ and D^- are only $O(h)$ approximations of the first-order derivative whereas the central difference D^0 is an $O(h^2)$ approximation. The precise analysis of the consistency error and convergence behaviour of the upwind scheme that now follows is based on [KT78] and draws on the bounds of Lemma 1.8 on derivatives of the exact solution u.

Theorem 2.12. *Assume that $b > \beta > 0$ and $c \geq 0$. Then there exists a positive constant β^*, which depends only on β, such that the error of the simple upwind scheme (2.12) at the inner grid points $\{x_i : \ i = 1, \ldots, N-1\}$ satisfies*

$$|u(x_i) - u_i| \leq \begin{cases} Ch\left[1 + \varepsilon^{-1}\exp(-\beta^*(1-x_i)/\varepsilon)\right] & \text{if } h \leq \varepsilon, \\ Ch + C\exp(-\beta^*(1-x_{i+1})/\varepsilon) & \text{if } h \geq \varepsilon. \end{cases}$$

Proof. As for the central scheme in Section 2.1.1, the consistency error is estimated using Taylor's formula. At each grid point x_i one obtains

$$|\tau_i| := |L_h u(x_i) - f(x_i)| \leq C \int_{x_{i-1}}^{x_{i+1}} \left(\varepsilon |u^{(3)}(t)| + |u^{(2)}(t)|\right) dt. \tag{2.13}$$

The crude bound $|u^{(k)}| \leq C\varepsilon^{-k}$ combined with the stability result of Lemma 2.11 yields only $|u(x_i) - u_i| \leq Ch/\varepsilon^2$, so a more precise bound on $|u^{(k)}|$ is needed. Invoking Lemma 1.8 yields the inequality

$$\begin{aligned} |\tau_i| &\leq Ch + C\varepsilon^{-2} \int_{x_{i-1}}^{x_{i+1}} \exp\left(-\beta(1-t)/\varepsilon\right) dt \\ &\leq Ch + C\varepsilon^{-1} \sinh\left(\frac{\beta h}{\varepsilon}\right) \exp\left(- \frac{\beta(1-x_i)}{\varepsilon}\right). \end{aligned}$$

Consider first the case when $h \leq \varepsilon$. Then $\beta h/\varepsilon$ is bounded. Now $\sinh t \leq Ct$ when t is bounded, so

$$|\tau_i| \leq Ch\left[1 + \varepsilon^{-2} \exp\left(- \frac{\beta(1-x_i)}{\varepsilon}\right)\right].$$

At first sight, this inequality seems unable to deliver the desired power of ε (viz., ε^{-1} instead of ε^{-2}) when Lemma 2.11 is applied. But if one considers the boundary value problem

$$-\varepsilon w'' + bw' + cw = C\varepsilon^{-1} \exp\left(- \frac{\beta(1-x)}{\varepsilon} \right), \quad w(0) = w(1) = 0,$$

then using the barrier function

$$w^*(x) = C \exp\left(- \frac{\beta^*(1-x)}{\varepsilon} \right)$$

where $\beta^* > \beta$, the comparison principle of Lemma 1.1 yields the estimate

$$|w(x)| \leq C \exp\left(- \frac{\beta^*(1-x)}{\varepsilon} \right)$$

– where w has gained a power of ε compared with Lw! The same calculation at the discrete level, using the discrete comparison principle of Lemma 2.5, completes the proof of the theorem when $h \leq \varepsilon$.

In the more difficult case $h \geq \varepsilon$, we decompose the solution as

$$u(x) = -u_0(1) \exp\left(- \frac{b(1)(1-x)}{\varepsilon} \right) + z(x).$$

By imitating the proof of Lemma 1.8 one finds that

$$|z^{(i)}(x)| \leq C \left[1 + \varepsilon^{1-i} \exp\left(- \frac{b(1)(1-x)}{\varepsilon} \right) \right].$$

Set

$$v(x) = -u_0(1) \exp\left(- \frac{b(1)(1-x)}{\varepsilon} \right)$$

and define v_h and z_h by

$$L_h v_h = Lv \quad \text{and} \quad L_h z_h = Lz,$$

where v_h and z_h agree with v and z, respectively, at x_0 and x_N. Then

$$|u(x_i) - u_i| = |v(x_i) + z(x_i) - (v_i + z_i)| \leq |v(x_i) - v_i| + |z(x_i) - z_i|.$$

For the consistency error associated with z, similarly to before one gets

$$|\tau_i(z)| \leq Ch + C \sinh\left(\frac{\beta h}{\varepsilon} \right) \exp\left(- \frac{\beta(1-x_i)}{\varepsilon} \right).$$

As now $h \geq \varepsilon$, we use the inequality $\sinh t \leq Ce^t$. Hence

$$|\tau_i(z)| \leq Ch + C \exp\left(- \frac{\beta(1-x_{i+1})}{\varepsilon} \right).$$