Chapter 4
Finite Difference Schemes for Convection-Diffusion Problems

This chapter is concerned with finite-difference discretisations of the stationary linear convection-diffusion problem

$$\mathcal{L}u := -\varepsilon u'' - bu' + cu = f \quad \text{in } (0,1), \;\; u(0) = \gamma_0, \;\; u(1) = \gamma_1, \tag{4.1}$$

with $b \geq \beta > 0$ on $[0,1]$. For the sake of simplicity we shall assume that

$$c \geq 0 \quad \text{and} \quad b' \geq 0 \quad \text{on } [0,1]. \tag{4.2}$$

Using (4.1) as a model problem, a general convergence theory for certain first- and second-order upwinded difference schemes on arbitrary and on layer-adapted meshes is derived. The close relationship between the differential operator and its upwinded discretisations is highlighted.

4.1 Notation

Meshes and mesh functions

Throughout this chapter let $\bar{\omega} : 0 = x_0 < x_1 < \cdots < x_N = 1$ be an arbitrary partition of $[0,1]$ with mesh intervals $I_i := [x_{i-1}, x_i]$. The set of inner mesh points is denoted by ω. The midpoint of I_i is $x_{i-1/2} := (x_i - x_{i+1})/2$ and its length $h_i := x_i - x_{i-1}$. Let $h := \max_{i=1,\dots,N} h_i$ be the maximal mesh size.

We shall identify mesh functions $v : \bar{\omega} \to \mathbb{R} : x_i \mapsto v_i$ with vectors $v \in \mathbb{R}^{N+1}$ and with spline functions

$$v \in V^\omega := \mathcal{S}_1^0(\bar{\omega}) := \left\{ w \in C^0[0,1] : w|_{I_i} \in \Pi_1 \text{ for } i = 1, \dots, N \right\}.$$

Let $\mathbb{R}_0^{N+1} := \left\{ v \in \mathbb{R}^{N+1} : v_0 = v_N = 0 \right\}$ be the space of mesh functions that vanish at the boundary. Furthermore, $V_0^\omega := \mathcal{S}_1^0(\bar{\omega}) \cap H_0^1(0,1)$.

T. Linß, *Layer-Adapted Meshes for Reaction-Convection-Diffusion Problems*, Lecture Notes in Mathematics 1985, DOI 10.1007/978-3-642-05134-0_4,
© Springer-Verlag Berlin Heidelberg 2010

Difference operators

In our notation of difference operators we follow Samarski's text book [146]. For any mesh function $v \in I\!R^{N+1}$ set

$$v_{x;i} := \frac{v_{i+1} - v_i}{h_{i+1}}, \quad v_{\bar{x};i} := v_{x;i-1} = \frac{v_i - v_{i-1}}{h_i}, \quad v_{\breve{x};i} := \frac{v_i - v_{i-1}}{h_{i+1}}$$

$$v_{\hat{x};i} := \frac{v_{i+1} - v_i}{\hbar_i}, \quad v_{\check{x};i} := \frac{v_i - v_{i-1}}{\hbar_i}, \quad v_{\mathring{x};i} := \frac{v_{i+1} - v_{i-1}}{2\hbar_i}$$

with the weighted mesh increment $\hbar$ defined by

$$\hbar_0 := \frac{h_1}{2}, \quad \hbar_i := \frac{h_i + h_{i+1}}{2}, \;\; i = 1, \dots, N-1, \quad \text{and} \quad \hbar_N := \frac{h_N}{2}.$$

To simplify the notation we set $g_i := g(x_i)$ for any $g \in C[0,1]$.

Further, less frequently used, difference operators will be introduced when needed.

Discrete norms and inner products

For any mesh function $v \in I\!R^{N+1}$ define the ℓ_∞ (semi-)norms

$$\|v\|_{\infty,\omega} := \max_{i=1,\dots,N-1} |v_i|, \quad \|v\|_{\infty,\bar{\omega}} := \max_{i=0,\dots,N} |v_i|,$$

$$\|[v\|_{\infty,\omega} := \max_{i=0,\dots,N-1} |v_i|, \quad |||v|||_{\varepsilon,\infty,\omega} := \max\Big\{\varepsilon \,\|[v_x\|_{\infty,\omega}, \beta\|v\|_{\infty,\bar{\omega}}\Big\},$$

the ℓ_1 norm

$$\|v\|_{1,\omega} := \sum_{j=0}^{N-1} h_{j+1}|v_j|$$

and the $w^{-1,\infty}$ norm

$$\|v\|_{-1,\infty,\omega} := \min_{V:V_x=v} \|V\|_{\infty,\bar{\omega}} = \min_{c\in I\!R} \Big\| \sum_{j=\cdot}^{N-1} h_{j+1}v_j + c \Big\|_{\infty,\bar{\omega}}.$$

We shall also use the following discrete inner products:

$$[w,v)_\omega := \sum_{i=0}^{N-1} h_{i+1}w_iv_i \quad \text{and} \quad (w,v)_\omega := \sum_{i=1}^{N-1} h_{i+1}w_iv_i.$$

4.2 A Simple Upwind Difference Scheme

In this section we study a first-order difference scheme for the discretisation of (4.1) on arbitrary meshes. Find $u^N \in I\!R^{N+1}$ such that

$$\left[Lu^N\right]_i = f_i \quad \text{for } i = 1, \ldots, N-1, \; u_0^N = \gamma_0, \; u_N^N = \gamma_1 \tag{4.3}$$

with

$$[Lv]_i := -\varepsilon v_{\bar{x}x;i} - b_i v_{x;i} + c_i v_i \quad \text{for } v \in I\!R^{N+1}.$$

At first glance the discretisation of the second-order derivative is a bit non-standard, because on non-uniform meshes it is not consistent in the maximum norm, but it has advantages that become clearer in the course of our analysis. More frequently used is the central difference approximation $u_i'' \approx u_{\bar{x}\hat{x};i}$. An upwind scheme based on this discretisation of the second-order derivative will be studied in Sect. 4.2.6, because the technique used there becomes more important in 2D, see Sect. 9.1.

The difference scheme (4.3) can be generated by a finite-element approach. To this end consider (4.1) with homogeneous boundary conditions. Its weak formulation is: Find $u \in H_0^1(0,1)$ such that

$$a(u,v) = f(v) \quad \text{for all } v \in H_0^1(0,1)$$

with

$$a(w,v) := \varepsilon\,(w',v') - (bw' - cw, v)\,, \quad f(v) := (f,v)$$

and the $L_2(0,1)$-scalar product $(w,v) := \int_0^1 \big(wv\big)(s)\mathrm{d}s$.

A standard FEM approximation is: Find $u^N \in V_0^\omega$ such that

$$a(u^N,v) = f(v) \quad \text{for all } v \in V_0^\omega.$$

The integrals in the bilinear form and the linear functional have to approximated. Use the left-sided rectangle rule $\int_{I_i} g(s)\mathrm{d}s \approx h_i g_{i-1}$, to arrive at: Find $u^N \in V_0^\omega$ such that

$$a_u\big(u^N,v\big) = f_u\big(v\big) \quad \text{for all } v \in V_0^\omega, \tag{4.4}$$

where

$$a_u\,(w,v) := \varepsilon\,[w_x, v_x)_\omega - (bw_x - cw, v)_\omega \quad \text{and} \quad f_u\,(v) := (f,v)_\omega\,.$$

Taking as test functions v the standard hat-function basis in V_0^ω, we see that (4.3) and (4.4) are equivalent. In particular,

$$a_u(w, v) = (Lw, v)_\omega = (w, L^* v)_\omega \quad \text{for all } w, v \in V_0^\omega,$$

with the adjoint operator L^* given by

$$[L^* v]_j = -\varepsilon v_{\bar{\xi}\xi;j} + (bv)_{\breve{\xi};j} + c_j v_j.$$

This is verified using summation by parts; cf. [146].

4.2.1 Stability of the Discrete Operator

The matrix associated with the difference operator L is a L_0-matrix because all off-diagonal entries are non-positive. Application of the M-criterion (Lemma 3.14) with a test vector with components $e_i = 2 - x_i$, $i = 0, \dots, N$ establishes the inverse monotonicity of L. Thus, L satisfies a comparison principle: For any mesh functions $v, w \in I\!R^{N+1}$

$$\left.\begin{array}{l} Lv \le Lw \quad \text{on } \omega, \\ v_0 \le w_0, \\ v_N \le w_N \end{array}\right\} \implies v \le w \quad \text{on } \bar{\omega}. \tag{4.5}$$

This comparison principle and Lemma 3.17 give the stability inequality

$$\left|u_i^N\right| \le \max\{|\gamma_0|, |\gamma_1|\} + (1 - x_i) \|f/b\|_{\infty,\omega} \quad \text{for } i = 0, \dots, N.$$

Thus, the operator L is $(\ell_\infty, \ell_\infty)$-stable with

$$\|v\|_{\infty,\bar{\omega}} \le \|Lv/b\|_{\infty,\omega} \quad \text{for all } v \in I\!R_0^{N+1}.$$

Alternatively, if $c > 0$ on $[0, 1]$, then Lemma 3.17 with $\boldsymbol{e} \equiv 1$ yields

$$\|v\|_{\infty,\bar{\omega}} \le \|Lv/c\|_{\infty,\omega} \quad \text{for all } v \in I\!R_0^{N+1}.$$

Note the analogy with (3.28).

Green's function estimates

Using the discrete Green's function $G : \bar{\omega}^2 \to I\!R : (x_i, \xi_j) \mapsto G_{i,j} = G(x_i, \xi_j)$ associated with L and Dirichlet boundary conditions, any mesh function $v \in I\!R_0^{N+1}$ can be represented as

$$v_i = a_u(v, G_{i,\cdot}) = (Lv, G_{i,\cdot})_\omega = (v, L^* G_{i,\cdot})_\omega \quad \text{for } i = 1, \dots, N-1.$$

Taking for v the standard basis in V_0^ω, we see that for fixed $i = 1, \dots, N-1$

$$[L^* G_{i,\cdot}]_j = \delta_{i,j} \quad \text{for } j = 1, \dots, N-1, \;\; G_{i,0} = G_{i,N} = 0, \tag{4.6}$$

where

$$\delta_{i,j} := \begin{cases} h_{i+1}^{-1} & \text{if } i = j, \\ 0 & \text{otherwise,} \end{cases}$$

is a discrete equivalent of the Dirac-δ distribution. As function of the first argument G solves, for fixed $j = 1, \dots, N-1$,

$$[L G_{\cdot,j}]_i = \delta_{i,j} \quad \text{for } i = 1, \dots, N-1, \;\; G_{0,j} = G_{N,j} = 0.$$

Theorem 4.1. *Suppose* (4.2) *holds true. Then the Green's function G associated with the discrete operator L and Dirichlet boundary conditions satisfies*

$$0 \le G_{i,j} \le \frac{1}{\beta} \begin{cases} 1 & \textit{for } 0 \le i \le j \le N, \\ \displaystyle\prod_{k=j+1}^{i} \left(1 + \frac{\beta h_{k+1}}{\varepsilon}\right)^{-1} & \textit{for } 0 \le j < i \le N, \end{cases}$$

$$G_{x;i,j} \le 0, \quad G_{\xi;i,j} \ge 0 \quad \textit{for } 0 \le j < i < N,$$
$$G_{x;i,j} \ge 0, \quad G_{\xi;i,j} \le 0 \quad \textit{for } 0 \le i \le j < N,$$
$$G_{x\xi;ij} \le 0 \quad \textit{for } 0 \le i, j < N, \; i \ne j$$

and

$$0 \le G_{x\xi;ii} \le \frac{1}{\varepsilon h_{i+1}} \; \textit{for } i = 0, \dots, N-1.$$

Proof. The upper and lower bounds on G are verified using (4.5).

Since $G \ge 0$ on $\bar{\omega}^2$ and $G_{i,0} = 0$ for $i = 0, \dots, N$, we have $G_{\xi;i,0} \ge 0$ for $i = 0, \dots, N$. By multiplying (4.6) by h_{j+1} and summing over j, we get

$$-\varepsilon G_{\xi;i,j} + \varepsilon G_{\xi;i,0} + b_j G_{i,j} = -\sum_{k=1}^{j} h_{k+1} c_k G_{i,k} \quad \text{for } j = 1, \dots, i-1.$$

Hence,

$$\varepsilon G_{\xi;i,j} \ge \varepsilon G_{\xi;i,0} + b_j G_{i,j} \ge 0 \quad \text{for } j = 1, \dots, i-1,$$

since $G_{i,j} \geq 0$ and $G_{\xi;i,0} \geq 0$. On the other hand, $G_{\xi;i,N-1} \leq 0$ for $i = 0, \ldots, N$ because $G \geq 0$ on $\bar{\omega}^2$ and $G_{i,N} = 0$ for $i = 0, \ldots, N$. By inspecting the difference equation (4.6), we see that $v_j := G_{\xi;i,j}$ satisfies, for $i < j < N$,

$$-\frac{\varepsilon}{h_{j+1}}\left(v_j - v_{j-1}\right) + \frac{h_j b_{j-1}}{h_{j+1}} v_{j-1} = -\big(b_{\breve{\xi};j} + c_j\big) G_{i,j} \leq 0, \tag{4.7}$$

by (4.2). Since $v_{N-1} \leq 0$, induction for decreasing j yields $G_{\xi;i,j} = v_j \leq 0$ for $i \leq j < N$.

Similarly, one can prove that $G_{x;i,j} \geq 0$ for $0 \leq i < j-1$ and $G_{x;i,j} \leq 0$ for $j \leq j < N$. Thus,

$$G_{x\xi;i,0} \leq 0 \;\text{ and }\; G_{x\xi;i,N-1} \leq 0 \;\text{ for }\; i = 0, \ldots, N-1.$$

because $G_{x;i,0} = G_{x;i,N} = 0$ for $0 \leq i < N$. Taking differences of (4.6) with respect to i and summing over j, we get

$$-\varepsilon G_{x\xi;i,j} + \varepsilon G_{x\xi;i,0} + b_j G_{x;i,j} + \sum_{k=1}^{j} h_{k+1} c_k G_{x;i,k} = -\delta_{i,j} \quad \text{for } 0 < j \leq i.$$

Therefore,

$$G_{x\xi;i,j} \leq 0 \quad \text{for} \quad 0 \leq j < i < N \quad \text{and} \quad G_{x\xi;i,i} \leq \frac{1}{\varepsilon h_{i+1}} \quad \text{for} \quad 0 \leq i < N$$

because $G_{x;i,j} \geq 0$, $G_{x\xi;i,0} \leq 0$ and $G_{x;i,0} = 0$.

For $i < j$, take differences of (4.7) to see that $v_j = G_{x\xi;i,j}$ satisfies

$$-\frac{\varepsilon}{h_{j+1}}\left(v_j - v_{j-1}\right) + \frac{h_j b_{j-1}}{h_{j+1}} v_{j-1} = -\big(b_{\breve{\xi};j} + c_j\big) G_{x;i,j} \leq 0 \qquad \text{for } j = i+2, \ldots, N-1.$$

because $b', c \geq 0$ and $G_{x;i,j} \geq 0$ for $i < j$. We get $G_{x\xi;i,j} \leq 0$ for $0 \leq i < j < N$. Finally, for $i = j$, use

$$\sum_{j=0}^{N-1} h_{j+1} G_{x\xi;i,j} = G_{x;i,N} - G_{x;i,0} = 0$$

in order to obtain $h_{i+1} G_{x\xi;i,i} \geq 0$. □

Mimicking the arguments of Theorem 3.23 we obtain its discrete counterpart.

Theorem 4.2. *Suppose* (4.2) *holds true. Then the Green's function G associated with the discrete operator L satisfies*

$$\|G_{i,\cdot}\|_{1,\omega} \leq \frac{1}{\beta}, \quad \|G_{\xi;i,\cdot}\|_{1,\omega} \leq \frac{2}{\beta}, \quad \|G_{x;\cdot,j}\|_{1,\omega} \leq \frac{2}{\beta}$$

and

$$\|G_{x\xi;i,\cdot}\|_{1,\omega} \le \frac{2}{\varepsilon}.$$

for all $i, j = 1, \dots, N-1$.

The ℓ_1-norms bounds are used to establish stability properties for L that resemble those of Theorem 3.45 for the differential operator $\mathcal{L}$.

Theorem 4.3. *Suppose* (4.2) *holds true. Then the operator* L *satisfies*

$$\|v\|_{\infty,\omega} \le \min\left\{\|Lv/b\|_{\infty,\omega}, \|Lv/c\|_{\infty,\omega}\right\}, \tag{4.8a}$$

$$\|v\|_{\infty,\omega} \le \beta^{-1}\|Lv\|_{1,\omega}, \quad \|v_x\|_{1,\omega} \le 2\beta^{-1}\|Lv\|_{1,\omega} \tag{4.8b}$$

and

$$|||v|||_{\varepsilon,\infty,\omega} \le 2\|Lv\|_{-1,\infty,\omega} \tag{4.8c}$$

for all $v \in I\!R_0^{N+1}$.

Remark 4.4. Similar to Remark 3.22 we have

$$2\,\|v\|_{-1,\infty,\omega} \le \|v\|_{1,\omega} \le \|v\|_{\infty,\omega}\,.$$

Therefore, the $(\ell_\infty, w^{-1,\infty})$-stability (4.8c) is the strongest of the three stability inequalities of Theorem 4.3. It was first derived by Andreev and Kopteva [11], though their derivation is different. A systematic approach can be found in [5], where stability of both the continuous operator $\mathcal{L}$ and of its discrete counterpart L is investigated. So far the $(\ell_\infty, w^{-1,\infty})$-stability inequality gives the sharpest error bounds for one-dimensional problems. But unlike the (ℓ_∞, ℓ_1) stability, it is unclear whether it can be generalised to higher dimensions. ♣

Remark 4.5. The same stability results hold true if the convection-diffusion problem in conservative form

$$\mathcal{L}^c u := -\varepsilon u'' - (bu)' + cu = f \quad \text{in } (0,1), \;\; u(0) = \gamma_0, \;\; u(1) = \gamma_1, \tag{4.9}$$

is discretised by

$$\begin{gathered} \left[L^c u^N\right]_i := -\varepsilon u^N_{\bar{x}x;i} - \left(bu^N\right)_{x;i} + c_i u^N_i = f_i \quad \text{for } i = 1, \dots, N-1, \\ u^N_0 = \gamma_0, \;\; u^N_N = \gamma_1. \end{gathered} \tag{4.10}$$

♣

Remark 4.6. The (ℓ_∞, ℓ_1) stability (4.8b) was first given by Andreev and Savin [12] for a modification of Samarskii's scheme [144]. It has been used in a number of publications to establish uniform convergence on S-type and B-type meshes; see, e. g., [10, 12, 106, 154]. Details of a convergence analysis can be found in Sect. 4.2.5. This stability result can be generalised to study two-dimensional problems; see Sect. 9.3.2. ♣

Corollary 4.7. *By Theorem* (4.3) *there exists a unique solution of* (4.3) *and it satisfies*

$$\left\|u^N\right\|_{\infty,\bar\omega} \le \min\left\{\|f/b\|_{\infty,\omega}, \|f/c\|_{\infty,\omega}\right\}.$$

4.2.2 A Priori Error Bounds

Let us consider the approximation error of the simple upwind scheme (4.3) applied to the boundary value problem (4.1). We give a convergence analysis based on the negative-norm stability of Theorem 4.3.

Introduce the continuous and discrete operators and functions

$$\big(\mathcal{A}v\big)(x) := \varepsilon v'(x) + \big(bv\big)(x) + \int_x^1 \big((b'+c)v\big)(s)\mathrm{d}s, \quad \mathcal{F}(x) := \int_x^1 f(s)\mathrm{d}s$$

and

$$[Av]_i := \varepsilon v_{\bar x;i} + b_i v_i + \sum_{k=i}^{N-1} h_{k+1}\left(b_{x;k}v_{k+1} + c_k v_k\right), \quad F_i := \sum_{k=i}^{N-1} h_{k+1} f_k.$$

Note that $\mathcal{L}v = -(\mathcal{A}v)'$ and $f = -\mathcal{F}'$ on $(0,1)$, and $Lv = -(Av)_x$ and $f = -F_x$ on ω. Thus,

$$\mathcal{A}u - \mathcal{F} \equiv \alpha \;\text{ on }\; (0,1) \;\text{ and }\; Au^N - F \equiv a \;\text{ on }\; \omega \tag{4.11}$$

with constants α and a.

In view of the stability inequality (4.8c) we have

$$\left|\left|\left|u-u^N\right|\right|\right|_{\varepsilon,\infty,\omega} \le 2\left\|L\left(u-u^N\right)\right\|_{-1,\infty,\omega} = 2\min_{c\in I\!R}\left\|A\left(u-u^N\right)+c\right\|_{\infty,\omega}.$$

Taking $c = a - \alpha$, where a and α are the constants from (4.11), we get

$$\left|\left|\left|u-u^N\right|\right|\right|_{\varepsilon,\infty,\omega} \le 2\|Au - \mathcal{A}u - F + \mathcal{F}\|_{\infty,\omega}. \tag{4.12}$$

Furthermore,

$$\begin{aligned}(Au - \mathcal{A}u - F + \mathcal{F})_i &= \varepsilon\left(u_{\bar{x}} - u'\right)_i + \sum_{k=i}^{N-1} h_{k+1} b_{x;k} u_{k+1} - \int_{x_i}^{x_N} \left(b'u\right)(x)\mathrm{d}x \\ &\quad + \sum_{k=i}^{N-1} h_{k+1}\left(c_k u_k - f_k\right) - \int_{x_i}^{x_N} (cu - f)(x)\mathrm{d}x.\end{aligned} \tag{4.13}$$

Taylor expansions with the integral form of the remainder give

$$h_{k+1}\left(c_k u_k - f_k\right) - \int_{I_{k+1}} (cu - f)(x)\mathrm{d}x = \int_{I_{k+1}} \int_x^{x_k} \big(cu - f\big)'(s)\mathrm{d}s\mathrm{d}x,$$
$$h_{k+1} b_{x;k} u_{k+1} - \int_{I_{k+1}} \left(b'u\right)(x)\mathrm{d}x = \int_{I_{k+1}} b'(x) \int_x^{x_{k+1}} u'(s)\mathrm{d}s\mathrm{d}x$$

and

$$\varepsilon\left(u_{\bar{x}} - u'\right)_k = \frac{\varepsilon}{h_k} \int_{I_k} \int_x^{x_k} u''(s)\mathrm{d}s\mathrm{d}x = \frac{1}{h_k} \int_{I_k} \int_{x_k}^{x} \big(bu' - cu + f\big)(s)\mathrm{d}s\mathrm{d}x,$$

by (4.1). Combining these representations with (4.12) and (4.13) we get the following general convergence result.

Theorem 4.8. *Let u be the solution of* (4.1) *and u^N that of* (4.3). *Then*

$$\left|\left|\left|u - u^N\right|\right|\right|_{\varepsilon,\infty,\omega} \le 2 \max_{k=1,\dots,N} \int_{I_k} \left(C_1 \left|u'(x)\right| + C_2 |u(x)| + C_3\right)\mathrm{d}x$$

with the constants

$$C_1 := \|c\|_\infty + \|b'\|_\infty + \|b\|_\infty, \quad C_2 := \|c\|_\infty + \|c'\|_\infty$$

and

$$C_3 := \|f\|_\infty + \|f'\|_\infty.$$

Remark 4.9. A similar result is given in [85] for the discretisation of the conservative form (4.9). When using the conservative form, the last two terms in (4.13) which involve b_x and b' disappear. ♣

Corollary 4.10. *Theorem 4.8 and the a priori bounds* (3.30) *yield*

$$\left|\left|\left|u - u^N\right|\right|\right|_{\varepsilon,\infty,\omega} \le C\vartheta_{cd}^{[1]}(\bar{\omega}),$$

where the characteristic quantity $\vartheta_{cd}^{[p]}(\bar{\omega})$ has been defined on p. 6:

$$\vartheta_{cd}^{[p]}(\bar{\omega}) := \max_{i=1,\dots,N} \int_{I_i} \left(1 + \varepsilon^{-1} \mathrm{e}^{-\beta s/p\varepsilon}\right) \mathrm{d}s.$$

Remark 4.11. The mesh function u^N can be extended to a piecewise linear function on the mesh $\bar{\omega}$. For convenience we denote this extended function by u^N also. Then

$$|||u - u^N|||_{\varepsilon,\infty} \le C\vartheta_{cd}^{[1]}(\bar{\omega})$$

follows from a triangle inequality and our bounds for the interpolation error. ♣

Remark 4.12. Corollary 4.10 allows to immediate deduce (almost) first-order uniform convergence for particular meshes. Suppose the mesh parameter σ in the definition of the meshes (see Sect. 2.1) satisfies $\sigma \ge 1$. Then

$$|||u - u^N|||_{\varepsilon,\infty} \le \begin{cases} CN^{-1} & \text{for Bakhvalov meshes,} \\ C\left(h + N^{-1} \max\limits_{\xi\in[0,q]} |\psi'(\xi)|\right) & \text{for S-type meshes and} \\ CN^{-1}\ln N & \text{for Shishkin meshes,} \end{cases}$$

by (2.6) and (2.9). ♣

A numerical example

Table 4.1 displays numerical results for the upwind scheme (4.3) on a Bakhvalov mesh applied to the test problem

$$-\varepsilon u'' - u' + 2u = \mathrm{e}^{x-1}, \quad u(0) = u(1) = 0. \tag{4.14}$$

Table 4.1 Simple upwinding on a Bakhvalov mesh ($q = 1/2$)

	$\sigma = 0.2$		$\sigma = 0.4$		$\sigma = 0.6$		$\sigma = 0.8$		$\sigma = 1.0$	
N	error	rate	error	rate	error	rate	error	rate	error	rate
2^7	2.246e-2	0.23	1.173e-2	0.47	6.856e-3	0.69	4.658e-3	0.87	3.995e-3	0.97
2^8	1.913e-2	0.22	8.482e-3	0.45	4.258e-3	0.68	2.547e-3	0.88	2.036e-3	0.98
2^9	1.641e-2	0.21	6.201e-3	0.44	2.662e-3	0.67	1.388e-3	0.87	1.030e-3	0.99
2^{10}	1.416e-2	0.21	4.576e-3	0.43	1.675e-3	0.66	7.586e-4	0.87	5.193e-4	0.99
2^{11}	1.224e-2	0.20	3.403e-3	0.42	1.062e-3	0.65	4.155e-4	0.87	2.611e-4	0.99
2^{12}	1.063e-2	0.20	2.545e-3	0.41	6.784e-4	0.64	2.281e-4	0.86	1.310e-4	1.00
2^{13}	9.226e-3	0.20	1.911e-3	0.41	4.361e-4	0.63	1.256e-4	0.86	6.568e-5	1.00
2^{14}	8.030e-3	0.20	1.439e-3	0.41	2.819e-4	0.62	6.937e-5	0.85	3.290e-5	1.00
2^{15}	6.969e-3	0.21	1.086e-3	0.40	1.830e-4	0.62	3.846e-5	0.85	1.647e-5	1.00
2^{16}	6.026e-3	—	8.207e-4	—	1.193e-4	—	2.139e-5	—	8.245e-6	—

In our computations we have fixed the parameter q and varied σ to illustrate the sharpness of our theoretical results. The errors are measured in the discrete maximum norm $\|\cdot\|_{\infty,\bar{\omega}}$. Apparently, choosing $\sigma < 1$ adversely affects the order of convergence. Similar observations can be made for the Shishkin mesh and other meshes.

4.2.3 Error Expansion

In the previous section we have seen that the error of the simple upwind scheme (4.3) satisfies

$$|||u - u^N|||_{\varepsilon,\infty,\omega} \le C\vartheta_{cd}^{[1]}(\bar{\omega}).$$

Now an expansion of the error of this scheme is constructed. We shall show there exists a function ψ, the leading term of the error, such that

$$u - u^N = \psi + \text{second order terms}.$$

This result can be applied to analyse, e.g., derivative approximations, defect correction and Richardson extrapolation, see Sect. 4.2.9 and 4.3.3.

For the sake of simplicity, we will study the conservative form (4.9), i. e.,

$$\mathcal{L}^c u := -\varepsilon u'' - (bu)' + cu = f \quad \text{in } (0,1), \;\; u(0) = \gamma_0, \;\; u(1) = \gamma_1,$$

and its discretisation by (4.10):

$$\left[L^c u^N\right]_i := -\varepsilon u^N_{\bar{x}x;i} - \left(bu^N\right)_{x;i} + c_i u_i^N = f_i \quad \text{for } i = 1,\dots,N-1,$$
$$u_0^N = \gamma_0, \;\; u_N^N = \gamma_1.$$

Corresponding to Sect. 4.2.2 we introduce

$$\left(\mathcal{A}^c v\right)(x) := \varepsilon v'(x) + \left(bv\right)(x) + \int_x^1 \left(cv\right)(s)\mathrm{d}s$$

and

$$[A^c v]_i := \varepsilon v_{\bar{x};i} + b_i v_i + \sum_{k=i}^{N-1} h_{k+1} c_k v_k.$$

Note that $\mathcal{L}^c v = -(\mathcal{A}^c v)'$ on $(0,1)$ and that $L^c v = -(A^c v)_x$ on ω.

4.2.3.1 Construction of the Error Expansion

We define the leading term of the error expansion as the solution of

$$\mathcal{L}^c\psi = \Psi' \quad \text{in } (0,1), \quad \psi(0) = \psi(1) = 0, \tag{4.15}$$

where

$$\Psi(x) = \varepsilon\frac{h(x)}{2}u''(x) - \int_x^1 \left(hg'\right)(s)\mathrm{d}s,$$

with

$$h(x) = x - x_{k-1} \quad \text{for } x \in (x_{k-1}, x_k) \quad \text{and} \quad g = f - cu.$$

Note that Ψ is discontinuous at the mesh nodes. Hence, $\mathcal{L}^c\psi$ is a generalised function. Therefore, (4.15) has to be interpreted in the context of distributions. Alternatively, one may seek a solution $\psi \in C^2((0,1)\setminus\omega) \cap C[0,1]$ such that

$$\mathcal{L}^c\psi = \Psi' \quad \text{in } (0,1)\setminus\omega, \;\; \psi(0) = \psi(1) = 0,$$

and

$$-\varepsilon[\psi'](x_i) = [\Psi](x_i) = -\varepsilon\frac{h_i}{2}u''(x_i) \quad \text{for } x_i \in \omega.$$

Since $\mathcal{A}^c\psi = -\Psi$ on $(0,1)\setminus\omega$, we have

$$\left[A^c\psi\right]_i = \varepsilon\left(\psi_{\bar{x};i} - \psi'_{i-0}\right) + \sum_{k=i}^{N-1} h_{k+1}c_k\psi_k - \int_{x_i}^1 (c\psi)(s)\mathrm{d}s + \Psi_{i-0}.$$

Thus,

$$\begin{aligned}\left[A^c\left(u-\psi-u^N\right)\right]_i = \varepsilon&\left(u_{\bar{x};i} - u'_i + \frac{h_i}{2}u''_i\right) - \varepsilon\left(\psi_{\bar{x};i} - \psi'_{i-0}\right)\\ &+ \int_{x_i}^1 \left(g - hg'\right)(x)\mathrm{d}x - \sum_{k=i}^{N-1} h_{k+1}g_k\\ &- \sum_{k=i}^{N-1} h_{k+1}c_k\psi_k + \int_{x_i}^1 (c\psi)(s)\mathrm{d}s.\end{aligned} \tag{4.16}$$

The function ψ has been designed such that the terms on the right-hand side that involve u are of second order. Those involving ψ are formally only first-order terms, but second order is gained since ψ itself is first order.

In order to bound the terms on the right-hand side, bounds for the derivatives of u up to order three are needed. These are provided by (3.30). The following theorem gives bounds for the leading term ψ of the error expansion and its derivatives up to order two, which are also required. Because of the number of technical details, its proof is deferred to the end of this section.

Lemma 4.13. *Let ψ be the solution of the boundary-value problem* (4.15). *Assume that $b, c \in C^2[0,1]$ and $f \in C^1[0,1]$. Then*

$$\left|\psi^{(k)}(x)\right| \le C\vartheta_{cd}^{[2]}(\bar{\omega})\left(1+\varepsilon^{-k}\mathrm{e}^{-\beta x/2\varepsilon}\right) \quad \textit{for } x \in (0,1)\setminus\omega, \;\; k = 0, 1, \tag{4.17a}$$

and

$$\varepsilon\left|\psi''(x)\right| \le C\vartheta_{cd}^{[2]}(\bar{\omega})\left(1+\varepsilon^{-1}\mathrm{e}^{-\beta x/2\varepsilon}\right) \quad \textit{for } x \in (0,1)\setminus\omega. \tag{4.17b}$$

Later we shall also show that (3.30), (4.16) and Lemma 4.13 yield

$$\left\|A^c(u-\psi-u^N)\right\|_{\infty,\omega} \le C\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2. \tag{4.18}$$

Then Theorem 4.3 yields our main result of this section.

Theorem 4.14. *Let u, u^N and ψ be the solutions of* (4.9), (4.10) *and* (4.15), *respectively. Assume that $b, c, f \in C^2[0,1]$. Then*

$$|||u-\psi-u^N|||_{\varepsilon,\infty,\omega} \le C\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2.$$

4.2.3.2 Detailed Proofs

Proof of Lemma 4.13

Now we derive bounds for the derivatives of the leading term ψ in the error expansion. The following auxiliary result will be used several times in the subsequent analysis.

Proposition 4.15. *Let $x \in (x_{k-1}, x_k)$ and $\sigma > 0$ be arbitrary. Then*

$$h(x)\left(1+\varepsilon^{-1}\mathrm{e}^{-\beta x/\sigma\varepsilon}\right) \le \int_{x_{k-1}}^{x}\left(1+\varepsilon^{-1}\mathrm{e}^{-\beta s/\sigma\varepsilon}\right)\mathrm{d}s$$

Proof. Let

$$F(x) := h(x)\left(1+\varepsilon^{-1}\mathrm{e}^{-\beta x/\sigma\varepsilon}\right) \quad \text{and} \quad G(x) := \int_{x_{k-1}}^{x}\left(1+\varepsilon^{-1}\mathrm{e}^{-\beta s/\sigma\varepsilon}\right)\mathrm{d}s.$$

Clearly $F(x_{k-1}) = G(x_{k-1}) = 0$ and

$$F'(x) = 1 + \varepsilon^{-1}\mathrm{e}^{-\beta x/\varepsilon} - \frac{h(x)\beta}{\sigma\varepsilon^2}\mathrm{e}^{-\beta x/\sigma\varepsilon} \le 1 + \varepsilon^{-1}\mathrm{e}^{-\beta x/\sigma\varepsilon} = G'(x)$$

for $x \in (x_{k-1}, x_k)$. The result follows. □

First (3.30) implies

$$|\Psi(x)| \le C\varepsilon h(x)\left(1 + \varepsilon^{-2}\mathrm{e}^{-\beta x/\varepsilon}\right) + C\int_x^1 h(s)\left(1 + \varepsilon^{-1}\mathrm{e}^{-\beta s/\varepsilon}\right)\mathrm{d}s.$$

This inequality, (3.29c) and Prop. 4.15 yield (4.17a) for $k = 0$.

Next, we derive bounds on ψ'. Set

$$B(x) := \frac{1}{\varepsilon}\int_0^x b(s)\mathrm{d}s, \quad a(x) := \Psi'(x) + (c - b')(x)\psi(x)$$

and

$$\chi(x) := \frac{1}{\varepsilon}\int_0^x a(s)\mathrm{e}^{B(s)-B(x)}\mathrm{d}s.$$

Then ψ can be written as

$$\psi(x) = \int_x^1 \chi(s)\mathrm{d}s + \kappa\int_x^1 \mathrm{e}^{-B(s)}\mathrm{d}s \quad \text{with} \quad \kappa = -\frac{\int_0^1 \chi(s)\mathrm{d}s}{\int_0^1 \mathrm{e}^{-B(s)}\mathrm{d}s}.$$

For ψ' we get

$$\psi'(x) = -\chi(x) - \kappa\mathrm{e}^{-B(x)}. \tag{4.19}$$

Apparently the critical point is to derive bounds on χ. Integration by parts and the definition of Ψ yield

$$2\chi(x) = (hu'')(x) - \zeta(x) \tag{4.20}$$

with

$$\zeta(x) := \frac{1}{\varepsilon}\int_0^x \left(hbu'' - 2h(f - cu)' - 2(c - b')\psi\right)(s)\mathrm{e}^{B(s)-B(x)}\mathrm{d}s.$$

For the first term on the right-hand side of (4.20) we have by (3.30) and Prop. 4.15

$$\left|(hu'')(x)\right| \le Ch(x)\left(1 + \varepsilon^{-1}\mathrm{e}^{-\beta x/2\varepsilon}\right)^2 \le C\vartheta_{cd}^{[2]}(\bar{\omega})\left(1 + \varepsilon^{-1}\mathrm{e}^{-\beta x/2\varepsilon}\right). \tag{4.21}$$

To bound $\zeta(x)$, the second term in (4.20), we use (3.30), (4.17a) for $k = 0$ and (4.21):

$$\begin{aligned}|\zeta(x)| &\le \frac{C}{\varepsilon}\int_0^x \left[h(s)\left(1+\varepsilon^{-2}\mathrm{e}^{-\beta s/\varepsilon}\right)+\vartheta_{cd}^{[2]}(\bar\omega)\right]\mathrm{e}^{\beta(s-x)/\varepsilon}\mathrm{d}s\\ &\le C\vartheta_{cd}^{[2]}(\bar\omega)\int_0^x\left(1+\varepsilon^{-1}\mathrm{e}^{\beta s/2\varepsilon}\right)\mathrm{e}^{\beta(s-x)/\varepsilon}\mathrm{d}s\\ &\le C\vartheta_{cd}^{[2]}(\bar\omega)\left(1+\varepsilon^{-1}\mathrm{e}^{-\beta x/2\varepsilon}\right).\end{aligned}$$

This, eq. (4.20) and inequality (4.21) give

$$|\chi(x)| \le C\vartheta_{cd}^{[2]}(\bar\omega)\left(1+\varepsilon^{-1}\mathrm{e}^{-\beta x/2\varepsilon}\right). \tag{4.22}$$

Integrating (4.22), we obtain

$$|\kappa| \le C\varepsilon^{-1}\vartheta_{cd}^{[2]}(\bar\omega), \tag{4.23}$$

since $\int_0^1 \mathrm{e}^{-B(s)}\mathrm{d}s \ge \varepsilon/\|b\|_\infty$. Combining (4.19)-(4.23), we get (4.17a) for $k = 1$.

Finally, the bound (4.17b) for the second-order derivative of ψ follows from (4.15), (3.30), (4.17a) and Prop. 4.15.

Proof of (4.18)

We now bound the terms on the right-hand side of (4.16). For the first two terms a Taylor expansion with the integral form of the remainder yields

$$\varepsilon\left|u_{\bar x;i}-u_i'+\frac{h_i}{2}u_i''\right| \le C\int_{I_i}(x-x_{i-1})\left(1+\varepsilon^{-2}\mathrm{e}^{-\beta x/\varepsilon}\right)\mathrm{d}x$$

by (3.30). To estimate the right-hand side we use the following result from [24].

Lemma 4.16. *Let g be a positive monotonically decreasing function on $[a,b]$. Let $p \in I\!N^+$. Then*

$$\int_a^b g(\xi)\,(\xi-a)^{p-1}\,\mathrm{d}\xi \le \frac{1}{p}\left\{\int_a^b g(\xi)^{1/p}\mathrm{d}\xi\right\}^p.$$

Proof. Consider the two integrals as functions of the upper integration limit. □

We get

$$\varepsilon\left|u_{\bar x;i}-u_i'+\frac{h_i}{2}u_i''\right| \le C\left\{\int_{I_i}\left(1+\varepsilon^{-1}\mathrm{e}^{-\beta x/2\varepsilon}\right)\mathrm{d}x\right\}^2 \le C\left(\vartheta_{cd}^{[2]}(\bar\omega)\right)^2. \tag{4.24}$$

Next we bound the third term in (4.16). Assuming $c, f \in C^2[0,1]$, we have

$$\int_{I_{k+1}} \big(g(x) - (x - x_k)g'(x)\big)\mathrm{d}x - h_{k+1}g_k = \int_{I_{k+1}} \int_{x_k}^{x} (s - x_k)g''(s)\mathrm{d}s.$$

Thus,

$$\left|\int_{I_{k+1}} (g(x) - (x - x_k)g')\,\mathrm{d}x - h_{k+1}g_k\right| \\ \le Ch_{k+1} \int_{I_{k+1}} (s - x_k)\left(1 + \varepsilon^{-2}\mathrm{e}^{-\beta s/\varepsilon}\right)\mathrm{d}s \le Ch_{k+1}\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2,$$

by (3.30) and Lemma 4.16. Hence

$$\left|\int_{x_i}^{1} \big(g - hg'\big)(x)\mathrm{d}x - \sum_{k=i}^{N-1} h_{k+1}g_k\right| \le C\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2. \tag{4.25}$$

To bound the remaining terms we use the bounds on ψ and its derivatives from Lemma 4.13. A Taylor expansion and (4.17b) yield

$$\varepsilon\left|\psi_{\bar{x};k} - \psi'_{k-0}\right| \le \varepsilon \int_{x_{k-1}}^{x_k} |\psi''(x)|\,\mathrm{d}x \le C\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2. \tag{4.26}$$

Finally,

$$\left|\int_{x_k}^{x_{k+1}} (c\psi)(s)\mathrm{d}s - h_{k+1}(c\psi)_k\right| \\ \le h_{k+1} \int_{I_{k+1}} |(c\psi)'(s)|\,\mathrm{d}\xi\mathrm{d}s \le Ch_{k+1}\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2,$$

by (4.17a). Therefore,

$$\left|\sum_{k=i}^{N-1} h_{k+1}c_k\psi_k - \int_{x_i}^{1} (c\psi)(s)\mathrm{d}s\right| \le C\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2. \tag{4.27}$$

Applying (4.24)–(4.27) to (4.16) and taking the maximum over $i = 0, \dots, N-1$, we get (4.18).

4.2.4 A Posteriori Error Estimation and Adaptivity

In Sect. 4.2.2 the stability of the *discrete operator* L was used to bound the error in the *discrete* maximum norm in terms of the derivative of the *exact solution*.

Now, in the first part of this section, roles are interchanged and the stability of the *continuous operator* $\mathcal{L}$ is used to bound the error in the *continuous* maximum norm in terms of finite differences of the *numerical solution*. We follow [64].

4.2.4.1 A Posteriori Error Bounds

Let u^N be the piecewise-linear function that solves (4.3). Then (3.29c) yields

$$|||u - u^N|||_{\varepsilon,\infty} \le 2\left\|\mathcal{L}\left(u - u^N\right)\right\|_{-1,\infty} = 2\min_{c\in I\!R}\left\|\mathcal{A}\left(u - u^N\right) + c\right\|_\infty .$$

Clearly

$$\min_{c\in I\!R}\left\|\mathcal{A}\left(u - u^N\right) + c\right\|_\infty \le \left\|\mathcal{A}\left(u - u^N\right) + a - \alpha\right\|_\infty , \tag{4.28}$$

where a and α are the constants from (4.11). Furthermore, for any $x \in (x_{i-1}, x_i)$,

$$\mathcal{A}\left(u - u^N\right) + a - \alpha = [Au^N]_i - \left(\mathcal{A}u^N\right)(x) - F_i + \mathcal{F}(x).$$

We bound the two terms on the right-hand side.

Since $\left(u^N\right)' = u^N_{\bar{x},i}$ for all $x \in (0,1)\setminus\omega$, we have

$$\begin{aligned}
&[Au^N]_i - \left(\mathcal{A}u^N\right)(x)\\
&\quad = \sum_{k=i}^{N-1} h_{k+1} b_{x;k} u^N_{k+1} - \int_{x_k}^1 \left(b' u^N\right)(s)\mathrm{d}s + \int_x^{x_i} b(s)\left(u^N\right)'(s)\mathrm{d}s\\
&\qquad - \int_x^{x_i}\left(cu^N\right)(s)\mathrm{d}s - \int_{x_i}^1 \left(cu^N\right)(s)\mathrm{d}s + \sum_{k=i}^{N-1} h_{k+1} c_k u^N_k ,
\end{aligned}$$

by the definitions of A and $\mathcal{A}$ and by integration by parts. For the terms on the right-hand side, Taylor expansions give

$$\left|h_{k+1} b_{x;k} u^N_{k+1} - \int_{I_{k+1}} \left(b' u^N\right)(s)\mathrm{d}s\right| \le h_{k+1}\|b'\|_\infty \left|u^N_{k+1} - u^N_k\right| ,$$

$$\left|\int_x^{x_k} b(s)\left(u^N\right)'(s)\mathrm{d}s\right| \le \|b\|_\infty \left|u^N_k - u^N_{k-1}\right| ,$$

$$\begin{aligned}
&\left|h_{k+1} c_k u^N_k - \int_{I_{k+1}} \left(cu^N\right)(s)\mathrm{d}s\right|\\
&\qquad \le h^2_{k+1}\|c'\|_\infty \max\left\{|u^N_{k+1}|, |u^N_k|\right\} + h_{k+1}\|c\|_\infty \left|u^N_{k+1} - u^N_k\right|
\end{aligned}$$

and

$$\left|\int_x^{x_k} \left(cu^N\right)(s)\mathrm{d}s\right| \le h_k\|c\|_\infty \max\left\{\left|u_k^N\right|, \left|u_{k-1}^N\right|\right\}.$$

Thus,

$$\left|[Au^N]_i - \left(\mathcal{A}u^N\right)(x)\right| \le C_1 \max_{k=0,\dots,N-1} |u_{k+1}^N - u_k^N| + C_2 h\|u^N\|_{\infty,\omega} \quad (4.29)$$

with the constants C_1 and C_2 form Theorem 4.8.

Next bound $F - \mathcal{F}$.

$$\left|\int_x^{x_k} f(s)\mathrm{d}s\right| \le h_k\|f\|_\infty$$

and

$$\left|h_{k+1}f_k - \int_{I_{k+1}} f(s)\mathrm{d}s\right| \le h_{k+1}^2\|f'\|_\infty$$

yield

$$|F_k - \mathcal{F}(x)| \le C_3 h.$$

Combining this with (4.28) and (4.29), then taking the supremum over all $x \in (0,1) \setminus \omega$, we get

$$\left\|\mathcal{L}\left(u - u^N\right)\right\|_{-1,\infty} \le C_1 \max_{i=0,\dots,N-1} |u_{i+1}^N - u_i^N| + h\left(C_2\|u^N\|_{\infty,\omega} + C_3\right)$$

with the constants C_1, C_2 and C_3 from Theorem 4.8.

Finally, use (3.29c) in order to obtain the main result of this section.

Theorem 4.17. *Let u be the solution of* (4.1) *and u^N that of* (4.3)*. Then*

$$\left|\left|\left|u - u^N\right|\right|\right|_{\varepsilon,\infty} \le 2C_1 \max_{i=0,\dots,N-1} |u_{i+1}^N - u_i^N| + 2h\left(C_2\|u^N\|_{\infty,\omega} + C_3\right).$$

Corollary 4.18. *Theorem 4.17 and Corollary 4.7 yield*

$$\left|\left|\left|u - u^N\right|\right|\right|_{\varepsilon,\infty} \le C \max_{i=0,\dots,N-1} h_{i+1}\left(1 + |u_{x;i}^N|\right).$$

Note the analogy of these results to Theorem 4.8 and to Corollary 4.10.

4.2.4.2 An Adaptive Method

From Theorem 4.8 it is easily concluded that the error of our upwind scheme satisfies

$$|||u - u^N|||_{\varepsilon,\infty,\omega} \le C \max_{i=1,\dots,N} \int_{I_i} \sqrt{1 + u'(x)^2}\mathrm{d}x.$$

On the other hand,

$$\int_0^1 \sqrt{1 + u'(x)^2}\mathrm{d}x \le C,$$

by (3.30). Thus, if the mesh is designed so that

$$\int_{I_{i-1}} \sqrt{1 + u'(x)^2}\mathrm{d}x = \int_{I_i} \sqrt{1 + u'(x)^2}\mathrm{d}x \tag{4.30}$$

for $i = 1, \dots, N-1$, i. e., if the mesh equidistributes the arc length of the exact solution, then

$$|||u - u^N|||_{\varepsilon,\infty,\omega} \le CN^{-1}. \tag{4.31}$$

However, u' is not available. An idea that leads to an adaptive method is to approximate the integrals in (4.30) by the mid-point quadrature rule, and $u'(x_{k-1/2})$ by a central difference quotient and finally to replace u by the numerical solution u^N. We get

$$\int_{I_i} \sqrt{1 + u'(x)^2}\mathrm{d}x \approx h_i \sqrt{1 + \left(u^N_{\bar{x};i}\right)^2}.$$

Thus setting

$$Q_i = Q_i\left(u^N, \omega\right) := \sqrt{1 + \left(u^N_{\bar{x};i}\right)^2},$$

we can replace (4.30) by

$$h_i Q_i = \frac{1}{N} \sum_{j=1}^{N} h_j Q_j \quad \text{for } i = 1, \dots, N-1. \tag{4.32}$$

Now solving the difference equation (4.3) and the discretised equidistribution principle (4.32) simultaneously, we get an adaptive method.

Kopteva and Stynes [69] proved that the nonlinear system of equations (4.3) and (4.32) possesses a solution and the error of the solution u^N obtained satisfies (4.31). An essential ingredient in the analysis is the a posteriori error bound of

Theorem 4.17. They proceed by considering a mesh movement algorithm, originally due to de Boor [25], which starts with a uniform mesh and aims to construct a mesh that solves the equidistribution problem (4.32).

In [69] it is shown that (4.32) does not need to be enforced strictly. The de Boor algorithm, which we are going to describe now, can be stopped when the relaxed equidistribution principle

$$Q_i h_i \leq \frac{\gamma}{N} \sum_{j=1}^{N} h_j Q_j \quad \text{for } i = 1, \ldots, N,$$

with a user-chosen constant $\gamma > 1$ is satisfied.

Algorithm:

1. Initialisation: Fix N and choose the constant $C_0 > 1$. The initial mesh $\omega^{[0]}$ is uniform with mesh size $1/N$.
2. For $k = 0, 1, \ldots$, given the mesh $\omega^{[k]}$, compute the discrete solution $u^{N,[k]}$ on this mesh. Set $h_i^{[k]} = x_i^{[k]} - x_{i-1}^{[k]}$ for each i. Let the piecewise-constant monitor function $\tilde{M}^{[k]}$ be defined by

$$\tilde{M}^{[k]}(x) := Q_i^{[k]} := Q_i \left(u^{N,[k]}, \omega^{[k]} \right) \quad \text{for} \quad x \in \left(x_{i-1}^k, x_i^k \right).$$

   Then the total integral of the monitor function $\tilde{M}^{[k]}$ is

$$I^{[k]} := \int_0^1 \tilde{M}^{[k]}(t)\mathrm{d}t = \sum_{j=1}^{N} h_j^{[k]} Q_j^{[k]}.$$

3. Test mesh: If

$$\max_{j=1,\ldots,N} h_j^{[k]} Q_j^{[k]} \leq \gamma \frac{I^{[k]}}{N},$$

   then go to Step 5. Otherwise, continue to Step 4.
4. Generate a new mesh by equidistributing the monitor function $\tilde{M}^{[k]}$ of the current computed solution: Choose the new mesh $\omega^{[k+1]}$ such that

$$\int_{x_{i-1}^{[k+1]}}^{x_i^{[k+1]}} \tilde{M}^{[k]}(t)\mathrm{d}t = \frac{I^{[k]}}{N}, \quad i = 0, \ldots, N.$$

   (Since $\int_0^x M^{[k]}(t)\mathrm{d}t$ is increasing in x, the above relation clearly determines the mesh $\omega^{[k+1]}$ uniquely.) Return to Step 2.
5. Set $\omega^* = \omega_N^{[k]}$ and $u^{N,*} = u^{N,[k]}$ then stop.

In [69] it is shown that the stopping criterion is met after $\mathcal{O}(|\ln \varepsilon|)$ iterations and the error of the numerical solution obtained satisfies (4.31) with a constant $C = C(\gamma)$.

Remark 4.19. Beckett [19] notes that when γ is chosen close to 1 the algorithm becomes numerically unstable. The mesh starts to oscillate: Mesh points moved into the layer region in one iteration are moved back out of it in the next iteration. Thus, the parameter γ must not be chosen too small. Values used in various publications are 2 and 1.2, but may be problem dependent. ♣

To avoid these oscillations Linß [86] rewrites (4.31) as

$$\begin{aligned}(x_i - x_{i-1})^2 &+ (u_i^N - u_{i-1}^N)^2 \\ &= (x_{i+1} - x_i)^2 + (u_{i+1}^N - u_i^N)^2 \quad \text{for } i = 1, \dots, N-1.\end{aligned}$$

Then he treats the system of (4.3) and (4.31) as a map

$$(0,1] \to I\!R^{2(N+1)} : \varepsilon \mapsto \left(\bar{\omega}_\varepsilon, u_\varepsilon^N\right)$$

and applies a continuation method combining an explicit Euler method (predictor) with a Newton method (corrector). The iteration matrices in each Newton step are seven diagonal and in an example the numerical costs are approximately of order $N\,|\ln(N\varepsilon)|$. However, convergence of this method is not proved in [86].

4.2.5 An Alternative Convergence Proof

In this section we shall demonstrate how the (ℓ_∞, ℓ_1) stability (4.8b) can be exploited to study convergence of the scheme (4.3) on S-type meshes. The results are less general than those of Sect. 4.2.2, but can be generalised to two dimensions; cf. Sect. 9.3.2. In our presentation we follow [106].

By (4.8b), we have

$$\left\|u - u^N\right\|_{\infty,\bar{\omega}} \le \beta^{-1} \left\|Lu - f\right\|_{1,\omega}. \tag{4.33}$$

Thus, the maximal nodal error is bounded by a discrete ℓ_1 norm of the truncation error $\zeta := Lu - f$:

$$\|\zeta\|_{1,\omega} = \sum_{j=0}^{N-1} h_{j+1}|\zeta_j|.$$

Using the solution decomposition $u = v + w$ of Theorem 3.48 and a triangle inequality, we can bound the truncation error pointwise:

$$|\zeta_i| \le \left|\left[Lv\right]_i - f_i\right| + \left|\left[Lw\right]_i\right|.$$

Separate Taylor expansions for the two solution components and the derivative bounds of Theorem 3.48 yield

$$h_{i+1}\left|\zeta_i\right| \le C\left(h_{i+1} + h_i + \mathrm{e}^{-\beta x_{i-1}/\varepsilon}\right) \tag{4.34a}$$

and

$$\begin{aligned} h_{i+1}\left|\zeta_i\right| \le C\Big\{&\left|h_{i+1} - h_i\right|\left(1 + \varepsilon^{-1}\mathrm{e}^{-\beta x_{i-1}/\varepsilon}\right) \\ &+ \left(h_i^2 + h_{i+1}^2\right)\left(1 + \varepsilon^{-2}\mathrm{e}^{-\beta x_{i-1}/\varepsilon}\right)\Big\}. \end{aligned} \tag{4.34b}$$

In the analysis, assume the mesh generating function $\tilde{\varphi}$ of the S-type mesh satisfies (2.8) and that $\sigma \ge 2$. For the sake of simplicity suppose $\tilde{\varphi}'$ is nondecreasing. This leads to a mesh that does not condense on $[0,\tau]$ as we move away from the layer, i.e., $h_i \le h_{i+1}$ for $i = 1,\ldots,qN-1$, which is reasonable for the given problem.

Now let us bound the ℓ_1 norm of the truncation error. Apply (4.34a) to bound $h_{i+1}\left|\zeta_i\right|$ for $i = qN, qN+1$ and (4.34b) otherwise. We get

$$\begin{aligned} \|\zeta\|_{1,\omega} \le C\sum_{i=1}^{qN-1}\Big\{&\left(h_{i+1} - h_i\right)\left(1 + \varepsilon^{-1}\mathrm{e}^{-\beta x_i/\varepsilon}\right) \\ &+ \left(h_i^2 + h_{i+1}^2\right)\left(1 + \varepsilon^{-2}\mathrm{e}^{-\beta x_{i-1}/\varepsilon}\right)\Big\} \\ &+ C\left(h + \mathrm{e}^{-\beta x_{qN-1}/\varepsilon} + \mathrm{e}^{-\beta x_{qN}/\varepsilon}\right) \\ &+ C\sum_{i=qN+2}^{N-1} N^{-2}\left(1 + \varepsilon^{-2}\mathrm{e}^{-\beta x_{i-1}/\varepsilon}\right). \end{aligned} \tag{4.35}$$

We bound the terms on the right-hand side separately in reverse order.

Let H denote the (constant) mesh size on $[\tau, 1]$. Then for $i = qN+2,\ldots,N-1$

$$\varepsilon^{-2}\mathrm{e}^{-\beta x_{i-1}/\varepsilon} \le \varepsilon^{-2}\mathrm{e}^{-\beta H/\varepsilon}\mathrm{e}^{-\beta\tau/\varepsilon} \le C\big(H/\varepsilon\big)^2\mathrm{e}^{-\beta H/\varepsilon} \le C,$$

since $x_{i-1} \ge x_{N/2} + H = \tau + H$ and $\sigma \ge 2$. Thus,

$$\sum_{i=qN+2}^{N-1} N^{-2}\left(1 + \varepsilon^{-2}\mathrm{e}^{-\beta x_{i-1}/\varepsilon}\right) \le CN^{-1}. \tag{4.36}$$

Furthermore,

$$\begin{aligned} &h + \mathrm{e}^{-\beta x_{qN-1}/\varepsilon} + \mathrm{e}^{-\beta x_{qN}/\varepsilon} \\ &\quad \le h + \left(1 + \mathrm{e}^{\beta h_{qN}/\varepsilon}\right) \mathrm{e}^{-\beta x_{qN}/\varepsilon} \le h + CN^{-\sigma}, \end{aligned} \tag{4.37}$$

by (2.11).

Next we bound the first sum in (4.35). We have

$$\sum_{i=1}^{qN-1} \left(h_{i+1} - h_i + h_i^2 + h_{i+1}^2\right) \le 3h \tag{4.38}$$

and

$$\begin{aligned} &\sum_{i=1}^{qN-1} (h_{i+1} - h_i)\, \mathrm{e}^{-\beta x_i/\varepsilon} \\ &\quad = -h_1 \mathrm{e}^{-\beta x_1/\varepsilon} + \sum_{i=2}^{qN-1} h_i \left(\mathrm{e}^{-\beta x_{i-1}/\varepsilon} - \mathrm{e}^{-\beta x_i/\varepsilon}\right) + h_{qN} \mathrm{e}^{-\beta x_{qN-1}/\varepsilon}. \end{aligned}$$

The mean value theorem, (2.11) and (2.12) imply

$$\left|\mathrm{e}^{-\beta x_{i-1}/\varepsilon} - \mathrm{e}^{-\beta x_i/\varepsilon}\right| \le h_i \frac{\beta}{\varepsilon} \mathrm{e}^{-\beta x_{i-1}/\varepsilon} \le C\varepsilon N^{-1} \max|\psi'| \mathrm{e}^{-\beta x_{i-1}/(2\varepsilon)}.$$

Therefore, it follows that

$$\varepsilon^{-1} \sum_{i=1}^{qN-1} (h_{i+1} - h_i)\, \mathrm{e}^{-\beta x_i/\varepsilon} \le CN^{-1} \max|\psi'| \sum_{i=1}^{qN} \frac{h_i}{\varepsilon} \mathrm{e}^{-\beta x_{i-1}/(2\varepsilon)}.$$

Ineq. (2.11) also gives

$$\sum_{i=1}^{qN} \frac{h_i}{\varepsilon} \mathrm{e}^{-\beta x_{i-1}/(2\varepsilon)} \le C \int_0^\tau \varepsilon^{-1} \mathrm{e}^{-\beta x/(2\varepsilon)} \mathrm{d}x \le C.$$

Hence,

$$\varepsilon^{-1} \sum_{i=1}^{qN-1} (h_{i+1} - h_i)\, \mathrm{e}^{-\beta x_i/\varepsilon} \le CN^{-1} \max|\psi'|. \tag{4.39}$$

Similar calculations yield

$$\varepsilon^{-2} \left| \sum_{i=1}^{qN-1} \left(h_i^2 + h_{i+1}^2\right) \mathrm{e}^{-\beta x_{i-1}/\varepsilon} \right| \le CN^{-1} \max|\psi'|. \tag{4.40}$$

Substituting (4.36)–(4.40) into (4.35) and applying (4.33), we get the uniform error bound

$$\left\|u - u^N\right\|_{\infty,\bar\omega} \le C\left(h + N^{-1}\max|\psi'|\right).$$

In [106] the authors proceed—using more detailed bounds on the discrete Green's function—to prove the sharper bound

$$\left\|u - u^N\right\|_{\infty,\bar\omega\cap[\tau,1]} \le CN^{-1}$$

for the error outside of the layer region provided that (2.14) is satisfied by the mesh generating function.

4.2.6 The Truncation Error and Barrier Function Technique

We now consider the convection-diffusion problem

$$-\varepsilon u'' - bu' + cu = f \;\text{ in } (0,1), \;\; u(0) = u(1) = 0 \tag{4.1}$$

discretised by

$$\begin{gathered}\left[\hat L u^N\right]_i := -\varepsilon u^N_{\bar x\hat x;i} - b_i u^N_{x;i} + c_i u^N_i = f_i \quad \text{for } \; i = 1,\dots,N-1,\\ u^N_0 = \gamma_0, \;\; u^N_N = \gamma_1.\end{gathered} \tag{4.41}$$

In contrast to the scheme (4.3) this method is first-order consistent in the mesh nodes on arbitrary meshes.

The analysis of this section uses the truncation error and barrier function technique developed by Kellogg and Tsan [61]. This was adapted to the analysis of Shishkin meshes by Stynes and Roos [153] and later used for other meshes also [137]. This technique can be used for problems in two dimensions too; see Sect. 9.1 or [81, 107]. We demonstrate this technique by sketching the convergence analysis for S-type meshes. For more details the reader is referred to [137].

The matrix associated with $\hat L$ is an M-matrix. Similar to (4.5), we have the following comparison principle for two mesh functions $v, w \in I\!R^{N+1}$:

$$\left.\begin{aligned}\hat L v &\le \hat L w \;\text{ on } \omega,\\ v_0 &\le w_0,\\ v_N &\le w_N\end{aligned}\right\} \quad\Longrightarrow\quad v \le w \;\text{ on } \bar\omega. \tag{4.42}$$

Theorem 4.20. *Let ω be a S-type mesh with $\sigma \geq 2$; see Sect. 2.1.3. Assume that the function $\tilde{\varphi}$ is piecewise differentiable and satisfies (2.8) and (2.14). Then the error of the simple upwind scheme satisfies*

$$\left|u_i - u_i^N\right| \leq \begin{cases} C\left(h + N^{-1}\max|\psi'|\right) & \textit{for } i = 0, \dots, qN-1, \\ C\left(h + N^{-1}\right) & \textit{for } i = qN, \dots, N. \end{cases}$$

Proof. The numerical solution u^N is split analogously to the splitting of $u = v + w$ of Theorem 3.48: $u^N = v^N + w^N$ with

$$[\hat{L}v^N]_i = f_i \quad \text{for } i = 1, \dots, N-1, \;\; v_0^N = v(0), \;\; v_N^N = v(1) = \gamma_1$$

and

$$[\hat{L}w^N]_i = 0 \quad \text{for } i = 1, \dots, N-1, \;\; w_0^N = w(0), \;\; w_N^N = w(1) = 0.$$

Then the error is $u - u^N = (v - v^N) + (w - w^N)$ and we can estimate the error in v and w separately. For the regular solution component v Taylor expansions and (3.34a) give

$$\left|\hat{L}(v - v^N)_i\right| = \left|[\hat{L}v]_i - (\mathcal{L}v)_i\right| \leq Ch \quad \text{for } i = 1, \dots, N-1.$$

Furthermore, $(v - v^N)_0 = (v - v^N)_N = 0$. Then the comparison principle (4.42) yields

$$\left|v_i - v_i^N\right| \leq C(1 - x_i)h$$

with some constant C, which is independent of ε. Thus,

$$\left\|v - v^N\right\|_{\infty,\omega} \leq Ch. \tag{4.43}$$

Using (4.42), one can show that

$$\left|w_i^N\right| \leq \bar{w}_i^N := C\prod_{k=1}^{i}\left(1 + \frac{\beta h_k}{2\varepsilon}\right)^{-1} \quad \text{for } i = 0, \dots, N. \tag{4.44}$$

For $\xi \geq 0$ we have $\ln(1+\xi) \geq \xi - \xi^2/2$ which implies

$$\bar{w}_i^N \leq \bar{w}_{qN}^N \leq N^{-\sigma/2}\exp\left(\frac{1}{2}\sum_{k=1}^{qN}\left(\frac{\beta h_k}{2\varepsilon}\right)^2\right) \leq CN^{-1} \quad \text{for } i = qN, \dots, N;$$

see Remark 2.4. Hence,

$$\left|w_i - w_i^N\right| \le |w_i| + \left|w_i^N\right| \le CN^{-1} \quad \text{for } i = qN, \dots, N, \tag{4.45}$$

where we have used (3.34b).

For the truncation error with respect to the layer part w, Taylor expansions and (3.34b) give

$$\begin{aligned}\left|[\hat{L}(w - w^N)]_i\right| = \left|[\hat{L}w]_i\right| &\le C\varepsilon^{-2}\left(h_i + h_{i+1}\right)\mathrm{e}^{-\beta x_{i-1}/\varepsilon} \\ &\le C\varepsilon^{-1}\mathrm{e}^{-\beta x_i/(2\varepsilon)} N^{-1} \max|\psi'| \\ &\le C\varepsilon^{-1}\bar{w}_i^N N^{-1} \max|\psi'| \quad \text{for } i = 1, \dots, qN-1,\end{aligned}$$

by (2.11) and (2.12). Note that $w_0 - w_0^N = 0$ and $\left|w_{qN} - w_{qN}^N\right| \le CN^{-1}$. Therefore, (4.42) yields

$$\left|(w - w^N)_i\right| \le C\left\{N^{-1} + \bar{w}_i^N N^{-1} \max|\psi'|\right\}, \quad \text{for } i = 0, \dots, qN-1,$$

for C chosen sufficiently large. Thus,

$$\left|w_i - w_i^N\right| \le CN^{-1} \max|\psi'| \quad \text{for } i = 0, \dots, qN-1.$$

Combine (4.43) and (4.45) with the last inequality to complete the proof. □

Remark 4.21. We are not aware of any results for B-type meshes that make use of this truncation error and barrier function technique. Also note that this technique needs $\sigma \ge 2$, while in Sect. 4.2.2 only $\sigma \ge 1$ was assumed. ♣

Remark 4.22. The technique of Sect. 4.2.2 also provides error estimates for the approximation of the first-order differences:

$$\varepsilon\left\|\left[(u - u^N)_x\right\|_{\infty,\omega} \le C\vartheta_{cd}^{[1]}(\bar{\omega}).$$

In [33] the authors use the barrier function technique to establish that the upwind scheme (4.41) on standard Shishkin meshes satisfies

$$\varepsilon\left|\left(u^N - u\right)_{x;i}\right| \le \begin{cases} CN^{-1}\ln N & \text{for } i = 0, \dots, qN-1, \\ CN^{-1} & \text{for } i = qN, \dots, N-1. \end{cases}$$

However, the technique in [33] makes strong use of the piecewise uniformity of the mesh. ♣

4.2.7 Discontinuous Coefficients and Point Sources

Consider the convection-diffusion problem in conservative form with a point source or a discontinuity of the convection coefficient at $d \in (0,1)$:

$$\begin{aligned} \mathcal{L}^c u := -\varepsilon u'' - (bu)' + cu = f + \alpha\delta_d, \quad \text{in } (0,1), \\ u(0) = \gamma_0, \;\; u(1) = \gamma_1, \end{aligned} \tag{4.46}$$

where δ_d is the shifted Dirac-delta function $\delta_d(x) = \delta(x-d)$ with $d \in (0,1)$. Assume that $b \geq \beta_1 > 0$ on $(0,d)$ and $b \geq \beta_2 > 0$ on $(d,1)$ and set $\beta = \min\{\beta_1, \beta_2\}$. For the sake of simplicity, also assume that $c \geq 0$ and $c - b' \geq 0$ on $[0,1]$. Properties of the exact solution are studied in Sect. 3.4.1.3.

Following [88], we consider the upwinded finite difference method: Find $u^N \in I\!R^{N+1}$ with

$$\begin{aligned} \left[L^c u^N\right]_i := -\varepsilon u^N_{\bar{x}x;i} - (b^- u^N)_{x;i} + c_i u^N_i = f_i + \varDelta_{d,i} \quad \text{for } i = 1,\dots,N-1, \\ u^N_0 = \gamma_0, \;\; u^N_N = \gamma_1, \end{aligned} \tag{4.47}$$

where $v^-(x) := \lim\limits_{s \to x-0} v(s)$, $v_i^- := v^-(x_i)$ and

$$\varDelta_{d;i} := \begin{cases} h_{i+1}^{-1} & \text{if } d \in [x_i, x_{i+1}), \\ 0 & \text{otherwise} \end{cases}$$

is an approximation of the shifted *Dirac*-delta function.

The discrete operator L^c enjoys the stability property (4.8c). Therefore, it is sufficient to derive bounds for the truncation error $\left\|L^c(u - u^N)\right\|_{-1,\infty,\omega}$. Adapting the notation from Sect. 4.2.2, we set

$$\begin{aligned} \big(\mathcal{A}^c v\big)(x) &:= \varepsilon v'(x) + \big(b^- v\big)(x) + \int_x^1 \big(cv\big)(s)\mathrm{d}s, \\ \mathcal{F}(x) &:= \int_x^1 f(s)\mathrm{d}s + \begin{cases} \alpha & \text{if } x_i \leq d, \\ 0 & \text{otherwise}, \end{cases} \\ [A^c v]_i &:= \varepsilon v_{\bar{x};i} + \big(b^- v\big)_i + \sum_{k=i}^{N-1} h_{k+1}\,(cv)_k \end{aligned}$$

and

$$F_i := \sum_{k=i}^{N-1} h_{k+1} f_k + \begin{cases} \alpha & \text{if } x_i \leq d, \\ 0 & \text{otherwise.} \end{cases}$$

Inspecting (4.46) and (4.47), we see

$$\mathcal{A}^c u - \mathcal{F} \equiv \text{const} \quad \text{on } (0,1) \quad \text{and} \quad A^c u^N - F \equiv \text{const} \quad \text{on } \omega.$$

Then, analogously to (4.13), we obtain

$$\begin{aligned}
&(A^c u - \mathcal{A}^c u - F + \mathcal{F})_i \\
&\quad = \varepsilon \left(u_{\bar{x}} - u'\right)_i + \sum_{k=i}^{N-1} h_{k+1} \left(c_k u_k - f_k\right) - \int_{x_i}^{x_N} (cu - f)(x)\mathrm{d}x
\end{aligned}$$

since the contributions from the δ functions and its discretisation cancel.

Proceeding along the lines of Sect. 4.2.2, we get.

Theorem 4.23. *Let u be the solution of* (4.46). *Then the error of the simple upwind scheme* (4.47) *satisfies*

$$|||u - u^N|||_{\varepsilon,\infty,\omega} \le C \max_{i=1,\dots,N} \int_{I_i} \bigl(1 + |u'(x)|\bigr)\mathrm{d}x.$$

Corollary 4.24. *Theorem 4.23 and the derivative bounds* (3.36) *yield*

$$|||u - u^N|||_{\varepsilon,\infty,\omega} \le C\vartheta^{[1]}_{cd^i}(\bar{\omega}),$$

where

$$\vartheta^{[p]}_{cd^i}(\bar{\omega}) := \max_{k=1,\dots,N} \int_{I_k} \left(1 + \varepsilon^{-1}\mathrm{e}^{-\beta_1 s/p\varepsilon} + H_d(s)\varepsilon^{-1}\mathrm{e}^{-\beta_2(s-d)/p\varepsilon}\right) \mathrm{d}s,$$

and H_d is the shifted Heaviside function.

Remark 4.25. Layer-adapted meshes for (4.46) have been introduced in Sect. 2.1.5. We have the uniform error bounds

$$|||u - u^N|||_{\varepsilon,\infty,\omega} \le \begin{cases} CN^{-1}\ln N & \text{for the Shishkin mesh and} \\ CN^{-1} & \text{for the Bakhvalov mesh} \end{cases}$$

if $\sigma_1 \ge 1$ and $\sigma_2 \ge 1$; see Sect. 2.1.5 for the bounds on $\vartheta^{[p]}_{cd^i}(\bar{\omega})$. ♣

Numerical results

Let us verify experimentally the theoretical result of Theorem 4.23. Our test problem is

$$-\varepsilon u'' - u' = x + \delta_{1/2} \ \text{ in } (0,1), \ \ u(0) = u(1) = 0. \tag{4.48}$$

Table 4.2 The upwind difference scheme for (4.48); errors in the discrete maximum norm

	Bakhvalov mesh		Shishkin mesh	
N	error	rate	error	rate
2^7	2.822e-2	0.95	3.898e-2	0.78
2^8	1.458e-2	0.97	2.277e-2	0.81
2^9	7.447e-3	0.98	1.299e-2	0.84
2^{10}	3.779e-3	0.99	7.280e-3	0.85
2^{11}	1.909e-3	0.99	4.027e-3	0.87
2^{12}	9.610e-4	0.99	2.204e-3	0.88
2^{13}	4.828e-4	1.00	1.197e-3	0.89
2^{14}	2.422e-4	1.00	6.454e-4	0.90
2^{15}	1.214e-4	1.00	3.462e-4	0.91
2^{16}	6.080e-5	—	1.848e-4	—

The results presented in Table 4.2 are in fair agreement with Theorem 4.23. Again the Bakhvalov mesh gives more accurate results than the Shishkin mesh.

Further remarks

The traditional truncation error and barrier function technique of Sect. 4.2.6 can also be applied to problems with interior layers. Farrell et al. [35] consider the problem of finding $u \in C^2((0,d) \cap (d,1)) \cup C^1[0,1]$ such that

$$-\varepsilon u'' - bu' = f \text{ in } (0,d) \cup (d,1), \quad u(0) = u(1) = 0,$$

where at the point $d \in (0,1)$ the convection coefficient changes sign:

$$b(x) > 0 \text{ for } x \in (0,d), \quad b(x) < 0 \text{ for } x \in (d,1) \text{ and } |b(x)| \geq \beta > 0.$$

The solution u and its derivatives satisfy

$$\left|u^{(k)}(x)\right| \leq C \left\{1 + \varepsilon^{-k} \mathrm{e}^{-\beta|x-d|/\varepsilon}\right\} \quad \text{for } k = 0, 1, \dots, q \text{ and } x \in [0,1],$$

where the maximal order q depends on the smoothness of the data. Using the barrier function technique of Sect. 4.2.6, in [35] the authors establish the error bound

$$\left\|u - u^N\right\|_{\infty,\omega} \leq CN^{-1} \ln N$$

for the simple upwind scheme (4.41) on a Shishkin mesh.

4.2.8 Quasilinear Problems

We now extend the results of Sect. 3.4.1 and 4.2.1 to the class of quasilinear problems described by

$$\begin{gathered} \mathcal{T}^c u := -\varepsilon u'' - b(\cdot,u)' + c(\cdot,u) = 0 \quad \text{in } (0,1), \\ u(0) = \gamma_0, \;\; u(1) = \gamma_1 \end{gathered} \tag{4.49}$$

with $0 < \varepsilon \ll 1$, $\partial_u b \geq \beta > 0$ and $\partial_u c \geq 0$ and its simple upwind discretisation

$$\begin{gathered} \left[T^c u^N\right]_i := -\varepsilon u^N_{\bar{x}x;i} - b(\cdot,u^N)_{x;i} + c(\cdot,u^N)_i = 0 \quad \text{for } i = 1,\dots,N-1, \\ u^N_0 = \gamma_0, \;\; u^N_N = \gamma_1. \end{gathered}$$

First, for the solution u of (4.49) and its derivatives, the bounds (3.30) hold true too; see [166]:

$$\left|u^{(k)}(x)\right| \leq C\left\{1 + \varepsilon^{-k}\mathrm{e}^{-\beta x/\varepsilon}\right\} \quad \text{for } k = 0,1,\dots,q \text{ and } x \in [0,1],$$

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

Next we use a standard linearisation technique to study stability properties of $\mathcal{T}^c$. For any two functions $v, w \in W^{1,\infty}(0,1)$ define the linear operator

$$\mathcal{L}^c y = \mathcal{L}^c[v,w]y := -\varepsilon y'' - (py)' + qy,$$

with

$$p(x) = \int_0^1 \partial_u b\big(x, w(x) + s(v-w)(x)\big)\mathrm{d}s \geq \beta$$

and

$$q(x) = \int_0^1 \partial_u c\big(x, w(x) + s(v-w)(x)\big)\mathrm{d}s \geq 0.$$

The linearised operator $\mathcal{L}^c$ is constructed such that $\mathcal{L}^c(v-w) = \mathcal{T}^c v - \mathcal{T}^c w$ on $(0,1)$. The analysis of Sect. 3.4.1 can be applied to $\mathcal{L}^c$. We get

$$|||v-w|||_{\varepsilon,\infty} \leq \left\|\mathcal{T}^c v - \mathcal{T}^c w\right\|_{-1,\infty} \qquad \text{for all } v,w \in W^{1,\infty} \text{ with } v-w \in W^{0,\infty}_1.$$

Similarly, we linearise T^c. For arbitrary mesh functions $v, w \in I\!R^{N+1}$ set

$$[L^c y]_i = \left[L^c[v,w]y\right]_i := -\varepsilon y_{\bar{x}x;i} - (py)_{x;i} + q_i y_i$$

with p and q defined above. Again the linearised operator L^c is constructed such that $L^c(v-w) = T^c v - T^c w$ on ω. The technique from Sect. 4.2.1 can be used to obtain

$$|||v-w|||_{\varepsilon,\infty,\omega} \le \|Tv - Tw\|_{-1,\infty,\omega} \quad \text{for all } v,w \in I\!R^{N+1} \text{ with } v-w \in I\!R_0^{N+1}.$$

In order to conduct an error analysis, take $v = u$ and $w = u^N$ and proceed as in Sect. 4.2.2 and 4.2.4 to get a priori and a posteriori error bounds.

Remark 4.26. Discretisations of quasilinear problems can also be analysed using the truncation error and barrier function technique of Sect. 4.2.6 or using the (ℓ_∞, ℓ_1) stability (4.8b); see [36, 149] and [106], respectively. ♣

4.2.9 Derivative Approximation

In a number of applications the user is more interested in the approximation of the gradient or of the flow than in the solution itself. In Sect. 4.2.2 the following error bound for the weighted derivative was established:

$$\varepsilon \left\| \left(u - u^N\right)' \right\|_\infty \le C\vartheta_{cd}^{[1]}(\bar\omega).$$

Note that $u'(0) \approx \varepsilon^{-1}$ by (3.30). Therefore, multiplying by ε in this estimate is the correct weighting. However, looking at the bounds (3.30) for the derivative of u, we see that the derivative is bounded uniformly away from the layer, where we therefore expect that a similar bound holds without the weighting by ε.

Theorem 4.27. *Let u be the solution of* (4.1) *and u^N that of* (4.3). *Then*

$$\left|u_i' - u_{\bar x;i}^N\right| \le Ch_i^{-1}\left(\vartheta_{cd}^{[2]}(\bar\omega)\right)^2 \quad \textit{for } i = 1,\dots,N.$$

Proof. We work from the error expansion of Sect. 4.2.3:

$$\left(u-u^N\right)_{\bar x;i} = \frac{u_i - \psi_i - u_i^N - \left(u_{i-1} - \psi_{i-1} - u_{i-1}^N\right)}{h_i} + \frac{\psi_i - \psi_{i-1}}{h_i}$$

Then

$$\left|\left(u-u^N\right)_{\bar x;i}\right| \le Ch_i^{-1}\left(\vartheta_{cd}^{[2]}(\bar\omega)\right)^2, \tag{4.50}$$

by Lemma 4.13 and Theorem 4.14. Furthermore,

$$\begin{aligned}|u_i' - u_{\bar{x};i}| &= \frac{1}{h_i}\left|\int_{I_i}(s-x_{i-1})u''(s)\mathrm{d}s\right| \\ &\le \frac{C}{h_i}\int_{I_i}(s-x_{i-1})\left(1+\varepsilon^{-2}\mathrm{e}^{-\beta s/\varepsilon}\right)\mathrm{d}s \le Ch_i^{-1}\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2,\end{aligned}$$

by (3.30) and Lemma 4.16. Finally, a triangle inequality yields the assertion. □

Layer-adapted meshes.

Let us illustrate Theorem 4.27 by applying it to two standard layer-adapted meshes.

Bakhvalov meshes (Sect. 2.1.1) can be generated by equidistributing

$$M_{Ba}(\xi) = \max\left\{1, \frac{K\beta}{\varepsilon}\exp\left(-\frac{\beta\xi}{\sigma\varepsilon}\right)\right\} \quad \text{for } \xi \in [0,1].$$

Clearly M_{Ba} is continuous and monotonically decreasing. Therefore,

$$\frac{1}{N}\int_0^1 M_{Ba}(s)\mathrm{d}s = \int_{I_i} M_{Ba}(s)\mathrm{d}s \le h_i M_{Ba}(x_{i-1})$$

and

$$\frac{1}{h_i} \le CNM_{Ba}(x_{i-1}) = CN\max\left\{1, \frac{K\beta}{\varepsilon}\exp\left(-\frac{\beta x_{i-1}}{\sigma\varepsilon}\right)\right\}.$$

Now, (2.6) and Theorem 4.27 yield

$$\left|u_i' - u_{\bar{x};i}^N\right| \le CN^{-1}\max\left\{1, \frac{K\beta}{\varepsilon}\exp\left(-\frac{\beta x_{i-1}}{\sigma\varepsilon}\right)\right\} \quad \text{if } \sigma \ge 2.$$

A very similar result was established by Kopteva and Stynes [72] through a different technique.

Shishkin meshes (Sect. 2.1.3). For these meshes the local step sizes satisfy

$$h_i = \frac{\sigma\varepsilon}{q\beta}\frac{\ln N}{N} \quad \text{for } i = 1,\dots,qN \quad \text{and} \quad h_i \ge N^{-1} \quad \text{for } i = qN+1,\dots,N.$$

Hence, Theorem 4.27 gives

$$\left|u_i' - u_{\bar{x};i}^N\right| \le \begin{cases} C\varepsilon^{-1}N^{-1}\ln N & \text{for } i = 1,\dots,qN-1, \\ CN^{-1}\ln^2 N & \text{for } i = qN,\dots,N. \end{cases}$$

Outside the layer region this result is slightly suboptimal. Both in [43] and in [72] it was shown by means of barrier function techniques that the approximation is a factor of $\ln N$ better, i.e.,

$$\left|u'_{i+1} - u^N_{x;i}\right| \le CN^{-1}\ln N \quad \text{for } i = qN+1,\dots,N.$$

4.3 Second-Order Difference Schemes

As simple upwinding yields only low accuracy, it is natural to look for higher-order alternatives. For one-dimensional problems inverse-monotone schemes exist that are second-order accurate. One will be studied in Sect. 4.3.1. However, the construction of inverse-monotone difference schemes in two or more dimensions is an open problem.

Sect. 4.3.2 summarises stability and convergence results for an unstabilised central difference scheme.

Possible other approaches to higher-order schemes include:

- the combination of two (or more) approximations by a first-order upwind scheme on nested meshes by means of the Richardson extrapolation technique.
- their combination of simple upwinding with higher-order unstabilised schemes using defect correction.

Both approaches have the advantage that linear problems involving only stabilised operators have to be solved. Sect. 4.3.3 is devoted to these techniques.

Finally, we like to mention the HODIE technique which was used by Clavero et al. [28] to construct and analyse second- and third-order compact schemes on Shishkin meshes.

4.3.1 Second-Order Upwind Schemes

Because of their stability properties, they can be analysed with the techniques similar to those of Sect. 4.2. Consider the convection-diffusion problem in conservative form:

$$\mathcal{L}^c u := -\varepsilon u'' - (bu)' + cu = f \quad \text{in } (0,1), \;\; u(0)=\gamma_0, \;\; u(1)=\gamma_1. \tag{4.9}$$

Let ρ_i, $i = 1,\dots,N$ be arbitrary with $\rho_i \in [1/2,1]$. Define the weighted step sizes

$$\chi_i = \rho_{i+1}h_{i+1} + (1-\rho_i)h_i \quad \text{for } i = 1,\dots,N-1, \quad \chi_0 = \chi_N = 0.$$

Then following Andreev and Kopteva [11], our discretisation is: Find $u^N \in I\!R^{N+1}$ such that

$$\left[L^\rho u^N\right]_i = f_{\rho;i} \quad \text{for } \; i = 1, \ldots, N-1, \quad u_0^N = \gamma_0, \;\; u_N^N = \gamma_1, \tag{4.51}$$

where

$$[L^\rho v]_i := -\varepsilon v_{\bar{x}\acute{x};i} - (\rho b v + (1-\rho)(bv)_-)_{\acute{x};i} + (cv)_{\rho;i},$$
$$v_{\acute{x};i} = \frac{v_{i+1} - v_i}{\chi_i}, \;\; v_{-;i} = v_{i-1}$$

and

$$v_{\rho;i} = \frac{\rho_{i+1} v_{i+1} + (1 - \rho_{i+1} + \rho_i) v_i + (1 - \rho_i) v_{i-1}}{2}.$$

The approximation of the first-order derivative is a weighted combination of upwinded and downwinded operators. At first glance the approximation of the lowest-order term and of the right-hand side seems to be very non-standard. It is chosen such that

$$\chi_i g_{\rho;i} \quad \text{is a second-order approximation of} \quad \int_{x_{\rho;i-1/2}}^{x_{\rho;i+1/2}} g(x)\mathrm{d}x$$

with $x_{\rho;i-1/2} = x_{i-1} + \rho_i h_i$. For $\rho \equiv 1/2$ we obtain a central difference scheme, while for $\rho \equiv 1$ the mid-point upwind scheme is recovered.

This second-order upwind scheme is very similar to the streamline-diffusion FEM, which is studied in Sect. 5.3.2.

4.3.1.1 Stability of the Discrete Operator

The stability analysis of the operator L^ρ is complicated by the positive contribution of the discretisation $(cu^N)_{\rho;i}$ of the lowest order term to the offdiagonal entries of the system matrix. It is difficult to ensure the correct sign pattern for the application of the M-matrix criterion (Lemma 3.14). Instead we follow [85] which adapts the technique from [11].

Set

$$[A^\rho v]_i := \varepsilon v_{\bar{x};i} + \rho(bv)_i + (1-\rho_i)(bv)_{i-1} - \sum_{j=1}^{i-1} \chi_j (cv)_{\rho;j}, \quad i = 1, \ldots, N.$$

This operator is related to L^ρ by $(A^\rho v)_{\acute{x}} = -L^\rho v$. Then any function $v \in I\!R_0^{n+1}$ can be represented as

$$v_i = \frac{W_N}{V_N} V_i - W_i,$$

where V and W are the solution of the difference equations

$$[A^\rho V]_i = 1, \quad i = 1, 2, \ldots, N,\ V_0 = 0$$

and

$$[A^\rho W]_i = [A^\rho v]_i + c, \quad i = 1, 2, \ldots, N,\ W_0 = 0$$

for any constant $c \in \mathbb{R}$.

Proposition 4.28. *Assume that*

$$1 \geq \rho_i \geq \max\left\{\frac{1}{2}, 1 - \frac{\varepsilon}{b_{i-1}h_i}\right\} \ \textit{for}\ \ i = 1, \ldots, N, \tag{4.52a}$$

and

$$\|c\|_\infty h \leq \beta/4. \tag{4.52b}$$

Then the matrix associated with A^ρ is an M-matrix.

Proof. First (4.52a) ensures that the offdiagonal entries of A^ρ are nonpositive, while (4.52b) implies that the diagonal entries are positive.

For any monotonically increasing mesh function $z_i \geq 0$ we have

$$[A^\rho z]_i > \rho_i b_i z_i - \frac{\|c\|_\infty}{2} \sum_{j=1}^{i-1} \chi_j \left(z_{j+1} + z_j\right) \geq \frac{\beta}{4} z_i - \|c\|_\infty \sum_{j=1}^{i-2} \chi_j z_{j+1},$$

by (4.52).

Now let

$$z_0 = z_1 = z_2 = 1, \ \text{and}\ \ z_i = \prod_{k=3}^{i} \left(1 + \frac{4\|c\|_\infty}{\beta} \chi_{k-2}\right) \ \ \text{for}\ \ i = 3, \ldots, N. \tag{4.53}$$

Clearly $z_i \leq \mathrm{e}^{4\|c\|_\infty/\beta}$ and

$$\frac{\beta}{4} z_i - \|c\|_\infty \chi_{i-2} z_{i-1}, \geq \frac{\beta}{4} z_{i-1}, \ \ \text{by (4.52b)}.$$

Then induction for i yields

$$[A^\rho z]_i > \frac{\beta}{4} \ \ \text{for}\ \ i = 1, \ldots, N.$$

Finally, application of Lemma 3.14 with the test function $e_i = z_i$ completes the proof.

The M-matrix property of A^ρ and the function z from (4.53) can now be used to establish bounds on V and W:

$$0 < V_i \le \frac{4}{\beta} z_i \le \frac{4}{\beta} \mathrm{e}^{4\|c\|/\beta} \quad \text{and} \quad |W_i| \le V_i \, \|A^\rho v + c\|_{\infty,\omega}\,, \;\; i = 1, \dots, N.$$

We get our main stability result.

Theorem 4.29. *Let ρ and h satisfy* (4.52). *Then the operator L^ρ is $(\ell_\infty, w^{-1,\infty})$-stable with*

$$\|v\|_{\infty,\omega} \le \frac{8}{\beta} \mathrm{e}^{4\|c\|_\infty/\beta} \min_{c \in \mathbb{R}} \|A^\rho v + c\|_{\infty,\omega} \quad \text{for all} \;\; v \in \mathbb{R}_0^{N+1}.$$

Remark 4.30. The (ℓ_∞, ℓ_1) stability

$$\|v\|_{\infty,\omega} \le C \sum_{k=1}^{N-1} \chi_k \left| [L^\rho v]_k \right|$$

is an immediate consequence of the negative-norm stability.

Analyses of second-order upwind schemes based on this type of stability inequality were given by Andreev and Savin [12] for a modification of Samarskii's scheme on a Shishkin mesh [12], and on Bakhvalov meshes [4] and by Linß [87] for quasilinear problems discretised on S-type meshes. ♣

4.3.1.2 Error Analysis

We now study the approximation error of the scheme (4.51). Following [11, 85], we base our analysis on the $(l_\infty, w^{-1,\infty})$ stability of Theorem 4.29.

Choose

$$\rho_i = \begin{cases} 1/2 & \text{if } \; h_i \le 2\varepsilon / b_{i-1}, \\ 1 & \text{otherwise.} \end{cases} \tag{4.54}$$

This choice satisfies the assumptions of Theorem 4.29. Therefore,

$$\left\|u - u^N\right\|_{\infty,\omega} \le C \min_{c \in \mathbb{R}} \left\|A^\rho (u - u^N) + c\right\|_{\infty,\omega}. \tag{4.55}$$

Set

$$\left(\mathcal{A}^c v\right)(x) := \varepsilon v'(x) + \big(bv\big) + \int_x^{x_{\rho;1/2}} \big(cv\big)(s)\mathrm{d}s, \quad \mathcal{F} := \int_x^{x_{\rho;1/2}} f(s)\mathrm{d}s$$

and

$$F_i^\rho := -\sum_{k=1}^{i-1} \chi_k f_{\rho;k}$$

Inspecting (4.9) and (4.51), we see that

$$\mathcal{A}^c u - \mathcal{F} \equiv \alpha \ \text{ on } \ (0,1) \ \text{ and } \ A^\rho u^N - F^\rho \equiv a \ \text{ on } \ \omega$$

with constants α and a because $\mathcal{L}^c v = -(\mathcal{A}^c v)'$ and $f = -\mathcal{F}'$ on $(0,1)$, and $L^\rho v = (A^\rho v)_{\acute{x}}$ and $f = F^\rho_{\acute{x}}$ on ω. Take $c = a - \alpha$ in (4.55) in order to get

$$\left\|u - u^N\right\|_{\infty,\omega} \le C \max_{i=1,\dots,N} \left| [A^\rho u]_i - (\mathcal{A}^c u)(x_{\rho;i}) + \mathcal{F}(x_{\rho;i}) - F_i^\rho \right|. \tag{4.56}$$

Set $g := cu - f$,

$$[B^\rho u]_i := \varepsilon u_{\bar{x};i} + \rho_i b_i u_i + (1-\rho_i) b_{i-1} u_{i-1}$$

and

$$\mathcal{B}(x) := \varepsilon u'(x) + \big(bu\big)(x).$$

Then

$$\begin{aligned} &[A^\rho u]_i - (\mathcal{A}^c u)(x_{\rho;i}) + \mathcal{F}(x_{\rho;i}) - F_i^\rho \\ &\qquad = [B^\rho u]_i - (\mathcal{B}^c u)(x_{\rho;i}) + \int_{x_{\rho;1/2}}^{x_{\rho;i-1/2}} g(s)\mathrm{d}s - \sum_{j=1}^{i-1} \chi_j g_{\rho;j}. \end{aligned} \tag{4.57}$$

When bounding the first term on the right-hand side of (4.57), we have to distinguish two cases: $\sigma_i = 1$ and $\sigma_i = 1/2$.

For $\sigma_i = 1$ we have

$$\big[B^\rho u\big]_i - \big(\mathcal{B}u\big)(x_{\rho;i}) = \varepsilon \left\{ \frac{u_i - u_{i-1}}{h_i} - u_i' \right\} = \frac{\varepsilon}{h_i} \int_{I_i} u''(t)(t - x_{i-1})\mathrm{d}t.$$

Thus,

$$\left| \big[B^h u\big]_i - \big(Bu\big)(x_{\rho,i}) \right| \le C \int_{I_i} \left(1 + \varepsilon^{-2} \mathrm{e}^{-\beta t/\varepsilon}\right)(t - x_{i-1})\mathrm{d}t, \tag{4.58}$$

by (3.30) and because $\varepsilon/h_i < \|b\|_\infty/2$ for $\rho_i = 1$.

Next, consider $\sigma_i = 1/2$. Then

$$\begin{aligned} &\big[B^\rho u\big]_i - \big(\mathcal{B}u\big)(x_{\rho,i}) \\ &\qquad = \varepsilon \left\{ \frac{u_i - u_{i-1}}{h_i} - u_{i-1/2}' \right\} + \frac{b_i u_i + b_{i-1} u_{i-1}}{2} - b_{i-1/2} u_{i-1/2}, \end{aligned}$$

where $u_{i-1/2} = u(x_{i-1/2})$. Taylor expansions for u and u' about x_i give

$$\varepsilon \left| \frac{u_i - u_{i-1}}{h_i} - u_{i-1/2}' \right| \le \frac{3\varepsilon}{2} \int_{I_i} \big|u'''(t)\big|(t - x_{i-1})\mathrm{d}t$$

and

$$\left|\frac{b_i u_i + b_{i-1} u_{i-1}}{2} - b_{i-1/2} u_{i-1/2}\right| \leq \frac{3}{2} \int_{I_i} \left|(bu)''(t)\right| (t - x_{i-1}) \mathrm{d}t.$$

From this and (3.30) we see that (4.58) holds for $\sigma_i = 1/2$ too.

Finally, we bound the second term of the right-hand side of (4.57):

$$\int_{x_{\rho,j-1/2}}^{x_{\rho,j+1/2}} g(s) \mathrm{d}s - \chi_{\rho,j} g_{\rho,j} = \int_{x_{\rho,j-1/2}}^{x_{\rho,j+1/2}} \left(g(s) - g_{\rho,j}\right) \mathrm{d}s.$$

The representation

$$g(s) = g_{j+1} - g'_{j+1}(x_{j+1} - s) + \int_s^{x_{j+1}} g''(t)(t - s) \mathrm{d}t$$

yields

$$\left|g_{\rho,j} - g(s) - (x_{\rho,j} - s) g'_{j+1}\right| \leq 2 \int_{x_{j-1}}^{x_{j+1}} \left|g''(t)\right| (t - x_{j-1}) \mathrm{d}t.$$

Next,

$$\begin{aligned}\left|\int_{x_{\rho,j-1/2}}^{x_{\rho,j+1/2}} g(s) \mathrm{d}s - \chi_{\rho,j} g_{\rho,j}\right| &\leq 2(h_j + h_{j+1}) \int_{x_{j-1}}^{x_{j+1}} \left|g''(t)\right| (t - x_{j-1}) \mathrm{d}t \\ &\leq C(h_j + h_{j+1}) \int_{x_{j-1}}^{x_{j+1}} \left(1 + \varepsilon^{-2} \mathrm{e}^{-\beta t/\varepsilon}\right) (t - x_{j-1}) \mathrm{d}t,\end{aligned}$$

by (3.30) and because $g = cu - f$.

Combining this estimate with (4.56), (4.55) and (4.58), we get

$$\left\|u - u^N\right\|_{\infty,\omega} \leq C \max_{i=1,\ldots,N-1} \int_{x_{i-1}}^{x_{i+1}} \left(1 + \varepsilon^{-2} \mathrm{e}^{-\beta t/\varepsilon}\right) (t - x_{i-1}) \mathrm{d}t.$$

Finally, Lemma 4.16 gives the following convergence result.

Theorem 4.31. *Let u^N be the approximate solution to* (4.9) *obtained by the difference scheme* (4.51) *with ρ chosen according to* (4.54). *Assume $\|c\|_\infty h \leq \beta/4$. Then*

$$\left\|u - u^N\right\|_{\infty,\omega} \leq C \left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2.$$

Quasilinear problems

The conclusion of the Theorem also holds when (4.51) is adapted to discretise the quasilinear problem

$$-\varepsilon u'' - b(\cdot, u)' + c(\cdot, u) = 0 \quad \text{in } (0,1), \; u(0) = \gamma_0, \; u(1) = \gamma_1$$

with $0 < \varepsilon \ll 1$, $\partial_u b \geq \beta > 0$ and $\partial_u c \geq 0$. The scheme reads: Find $u^N \in I\!R^{n+1}$ such that

$$-\varepsilon u^N_{\bar{x}\acute{x};i} - \left(\rho b(\cdot, u^N) + (1-\rho) b(\cdot, u^N)_-\right)_{\acute{x};i} + c(x_{\rho;i}, u^N_{\rho;i}) = 0 \quad \text{on } \omega,$$
$$u^N_0 = \gamma_0, \; u^N_N = \gamma_1$$

with the stabilisation parameter chosen to satisfy, e. g,

$$\rho_i = \begin{cases} 1/2 & \text{if } h_i \leq 2\varepsilon/\|b\|_\infty, \\ 1 & \text{otherwise.} \end{cases}$$

Discontinuous coefficients and point sources

Consider the convection-diffusion problem (4.46) with a point source:

$$\mathcal{L}^c u := -\varepsilon u'' - (bu)' + cu = f + \alpha\delta_d, \; \text{in } (0,1), \; u(0) = \gamma_0, \; u(1) = \gamma_1,$$

with the shifted Dirac-delta function $\delta_d(x) = \delta(x-d)$. The coefficient b may also have a discontinuity at $x = d$. Assume that $b \geq \beta_1 > 0$ on $(0,d)$ and $b \geq \beta_2 > 0$ on $(d,1)$.

Using (4.51) we seek an approximation $u^N \in I\!R^{n+1}$ with

$$\left[L^\rho u^N\right]_i = f_{\rho;i} + \Delta_{d,\rho;i} \quad \text{for } i = 1, \ldots, N-1, \; u^N_0 = \gamma_0, \; u^N_N = \gamma_1$$

with

$$\Delta_{d,\rho;i} := \begin{cases} \chi_i^{-1} & \text{if } d \in [x_{\rho;i-1/2}, x_{\rho;i+1/2}), \\ 0 & \text{otherwise.} \end{cases}$$

Then the above technique and the a priori bounds (3.36) for the derivatives of u yield the error estimate [88]

$$\left\|u - u^N\right\|_{\infty,\omega} \leq C \left(\vartheta^{[2]}_{cd^i}(\bar{\omega})\right)^2,$$

where $\vartheta^{[2]}_{cd^i}(\bar{\omega})$ has been defined in Sect. 2.1.5.

Remark 4.32. Roos and Zarin [143] study the difference scheme generated by the streamline diffusion FEM on Shishkin and on Bakhvalov-Shishkin meshes for the discretisation of a problem with a point source. They prove (almost) second-order convergence in the discrete maximum norm too. ♣

A posteriori error estimates

in the maximum norm for (4.9) discretised by (4.51) can be derived using the $(L_\infty, W^{-1,\infty})$-stability (3.29c) of the continuous operator $\mathcal{L}^c$. However, compared to Sect. 4.2.4 the analysis becomes more technical. Therefore, we refer the reader to the article by Kopteva [64]. A flavour of the technique is given in Sect. 4.3.3.1 where a defect-correction method is analysed.

4.3.1.3 The Barrier Function Technique

Stynes and Roos [153] study a hybrid difference scheme on a Shishkin mesh (with $q = 1/2$ and $\sigma > 4$). Their scheme uses central differencing on the fine part of the mesh and the mid-point upwind scheme on the coarse part.

Let us consider the convection-diffusion problem

$$\mathcal{L}u := -\varepsilon u'' - bu' + cu = f \quad \text{in } (0,1), \; u(0) = \gamma_0, \; u(1) = \gamma_1, \tag{4.1}$$

with $b \geq \beta > 0$ and $c \geq 0$ on $[0,1]$. This is discretised on a Shishkin mesh—see Sect. 2.1.3—using the difference scheme

$$\left[Lu^N\right]_i = \tilde{f}_i \quad \text{for } i = 1, \ldots, N-1, \; u_0^N = \gamma_0, \; u_N^N = \gamma_1 \tag{4.59}$$

with

$$[Lv]_i := \begin{cases} -\varepsilon v_{\bar{x}\hat{x};i} - b_i v_{\mathring{x};i} + c_i v_i & \text{if } b_i h_i \leq 2\varepsilon, \\ -\varepsilon v_{\bar{x}\hat{x};i} - b_{i+1/2} v_{x;i} + (c_i v_i + c_{i+1} v_{i+1})/2 & \text{otherwise,} \end{cases}$$

and

$$\tilde{f}_i := \begin{cases} f_i & \text{if } b_i h_i \leq 2\varepsilon, \\ f_{i+1/2} & \text{otherwise.} \end{cases}$$

For N larger than a certain threshold value N_0, the matrix associated with L is an M-matrix and central differencing is used exclusively on the fine part of the mesh.

Theorem 4.33. *Let ω be a Shishkin mesh with $\sigma \geq 2$; see Sect. 2.1.3. Then the error of the upwinded scheme* (4.59) *applied to* (4.1) *satisfies*

$$\left|u_i - u_i^N\right| \leq \begin{cases} CN^{-2}\ln^2 N & \textit{for } i = 0, \dots, qN-1, \\ CN^{-2} & \textit{for } i = qN, \dots, N, \end{cases}$$

if N is larger than a certain threshold value.

Remark 4.34. A similar scheme generated by streamline-diffusion stabilisation was analysed by Stynes and Tobiska [154] with special emphasis on the choice of the mesh parameter σ. There it was first established that the mesh parameter σ should be chosen equal (or greater than) to the formal order of the scheme. ♣

Proof (of Theorem 4.33). Start with the truncation error. When $2\varepsilon < b_i h_i$ we have the bound

$$\left|[Lg]_i - (\mathcal{L}g)_{i+1/2}\right| \leq C\left\{\varepsilon \int_{x_{i-1}}^{x_{i+1}} |g'''(s)|\,\mathrm{d}s + h_{i+1}\int_{x_i}^{x_{i+1}} \left[|g'''(s)| + |g''(s)|\right]\mathrm{d}s\right\}, \tag{4.60a}$$

otherwise we use

$$\left|[Lg]_i - (\mathcal{L}g)_i\right| \leq C\int_{x_{i-1}}^{x_{i+1}} \left[\varepsilon|g'''(s)| + |g''(s)|\right]\mathrm{d}s \tag{4.60b}$$

and, if $h_i = h_{i+1}$,

$$\left|[Lg]_i - (\mathcal{L}g)_i\right| \leq Ch_i\int_{x_{i-1}}^{x_{i+1}} \left[\varepsilon|g^{(4)}(s)| + |g'''(s)|\right]\mathrm{d}s. \tag{4.60c}$$

For the analysis we split the numerical solution u^N analogously to the splitting $u = v + w$ of Theorem 3.48 and Remark 3.50: $u^N = v^N + w^N$ with

$$\left[Lv^N\right]_i = \tilde{f}_i \quad \text{for } i = 1, \dots, N-1, \;\; v_0^N = v(0), \; v_N^N = v(1)$$

and

$$\left[Lw^N\right]_i = 0 \quad \text{for } i = 1, \dots, N-1, \;\; w_0^N = w(0), \; w_N^N = w(1).$$

Then the error is $u - u^N = (v - v^N) + (w - w^N)$ and we estimate the error in v and w separately.

For the regular solution component Theorem 3.48, Remark 3.50 and (4.60) give

$$\left|\left[L(v-v^N)\right]_i\right| = \left|[Lv]_i - \tilde{f}_i\right| \le \begin{cases} CN^{-1} & \text{for } i = qN, \\ CN^{-2} & \text{otherwise.} \end{cases}$$

Note that $(v-v^N)_0 = (v-v^N)_N = 0$. Now set

$$\varphi_i = \begin{cases} 1 & \text{for } i = 0,\dots,qN, \\ \prod_{k=qN+1}^{i} \left(1+\frac{\beta h_k}{\varepsilon}\right)^{-1} & \text{for } i = qN+1,\dots,N. \end{cases}$$

Clearly $\varphi_0 \ge 0$ and $\varphi_N \ge 0$. Furthermore,

$$[L\varphi]_i \ge \begin{cases} 0 & \text{for } i \ne qN, \\ \dfrac{\beta}{2h_{qN+1}} \ge \dfrac{\beta(1-q)}{2}N & \text{for } i = qN. \end{cases}$$

Application of a comparison principle with the barrier function $CN^{-2}(1-x_i+\varphi_i)$ yields

$$\left\|v-v^N\right\|_{\infty,\omega} \le CN^{-2}, \tag{4.61}$$

since the matrix associated with L is inverse monotone as mentioned before.

Next, consider the layer component w. Let

$$\psi_i := \begin{cases} \prod_{k=1}^{i}\left(1+\frac{\beta h_k}{\varepsilon}\right)^{-1} + \prod_{k=1}^{qN}\left(1+\frac{\beta h_k}{\varepsilon}\right)^{-1} & \text{for } i = 1,\dots,qN, \\ 2\prod_{k=1}^{i}\left(1+\frac{\beta h_k}{\varepsilon}\right)^{-1} & \text{for } i = qN,\dots,N. \end{cases}$$

The inverse monotonicity of the discrete operator L yields

$$\left|w_i^N\right| \le |v(0)-\gamma_0|\,\psi_i \quad \text{for } i = 0,\dots,N,$$

because $L\psi \ge 0$. Furthermore, $|w_i| \le C\mathrm{e}^{-\beta x_i/\varepsilon} \le C\psi_i$. Thus,

$$\left|w_i - w_i^N\right| \le C\psi_i \quad \text{for } i = 0,\dots,N.$$

Now the argument that lead to (4.45) for the first-order scheme is imitated to establish

$$\left|w_i - w_i^N\right| \le CN^{-2} \ \text{ for } i = qN,\dots,N, \tag{4.62}$$

if $\sigma \ge 2$ in the construction of the Shishkin mesh (Sect. 2.1.3).

For $i = 1, \ldots, qN - 1$ the truncation error with respect to w satisfies

$$\left|\left[L(w - w^N)\right]_i\right| \le CN^{-2} \ln^2 N \varepsilon^{-1} \mathrm{e}^{-\beta x_{i-1}/\varepsilon} \le CN^{-2} \ln^2 N \varepsilon^{-1} \tilde{\psi}_i,$$

by (4.60c), Theorem 3.48 and Remark 3.50, where

$$\tilde{\psi}_i := \prod_{k=1}^{i} \left(1 + \frac{\beta h_k}{2\varepsilon}\right)^{-1}.$$

The inverse monotonicity of L gives

$$\left|(w - w^N)_i\right| \le CN^{-2} \ln^2 N \tilde{\psi}_i \quad \text{for } i = 1, \ldots, qN - 1, \tag{4.63}$$

because

$$\left[L\tilde{\psi}\right]_i \ge C\varepsilon^{-1}\tilde{\psi}_i \quad \text{for } i = 1, \ldots, qN - 1$$

and because both $\left|w_0 - w_0^N\right| \le CN^{-2}$ and $\left|w_{qN} - w_{qN}^N\right| \le CN^{-2}$.

Combining (4.61), (4.62) and (4.63), we are finished. □

4.3.2 Central Differencing

In numerical experiments [34, 50, 120] it was observed that central differencing on Shishkin meshes yields almost second-order accuracy.

A drawback of central difference approximations is their lack of stability. The discretisations are not maximum-norm stable. It will be seen in Sect. 4.3.2.1 that the use of layer-adapted meshes induce some additional stability. However, the discrete systems remain difficult to solve efficiently by means of iterative solvers. The system matrices have eigenvalues with large imaginary parts. This becomes a particularly important issue when solving higher-dimensional problems.

We shall consider the discretisation

$$\left[Lu^N\right]_i = f_i \quad \text{for } i = 1, \ldots, N - 1, \;\; u_0^N = \gamma_0, \;\; u_N^N = \gamma_1 \tag{4.64}$$

of (4.9), where

$$[Lv]_i := -\varepsilon v_{\bar{x}\hat{x};i} - (bv)_{\mathring{x};i} + c_i v_i.$$

Similar to (4.4) this scheme is equivalent to a FEM with piecewise linear trial and test functions, but with the trapezium rule

$$\int_{I_i} g(s)\mathrm{d}s \approx h_i \frac{g_{i-1} + g_i}{2}$$

used to approximate the integrals on each subinterval I_i of the partition.

4.3.2.1 Stability

A first analysis was conducted by Andreev and Kopteva [10] who prove that central differencing on a Shishkin mesh is (l_∞, l_1) stable. This result was later generalised by Kopteva [70].

Theorem 4.35. *Assume that*

$$\left| \prod_{i=1}^{N} \left(\frac{\varepsilon}{h_i b_{i-1}} - \frac{1}{2} \right) \Big/ \left(\frac{\varepsilon}{h_i b_i} + \frac{1}{2} \right) \right| \leq \frac{1}{4} \tag{4.65}$$

and that $h_i \leq \mu h_j$ *for* $i \leq j$ *with some constant* μ. *Then the central difference operator* L *is* (l_∞, l_1) *stable with*

$$\|v\|_{\infty,\omega} \leq \frac{81}{4\beta} \sum_{i=1}^{N-1} \hbar_i \left| [Lv]_i \right|. \tag{4.66}$$

Furthermore, let m *be such that* $h_i \leq 2\varepsilon/b_{i-1}$ *for* $i = 1, \ldots, m$ *and* $h_{m+1} > 2\varepsilon/b_m$. *Then the operator* L *is* $(l_\infty, w^{-1,\infty})$*-stable with*

$$\|v\|_{\infty,\omega} \leq \frac{11}{2\beta} \max_{j=1,\ldots,N-1} \left| \sum_{k=j}^{N-1} \hbar_k \, [Lv]_k \right|.$$

for any mesh function v *with* $[Lv]_i = 0$ *for* $i > m$.

Proof. The argument is very technical and therefore not presented here. Instead the reader is referred to the original work by Kopteva [70]. □

4.3.2.2 A Priori Error Bounds

Based on Theorem 4.35 Kopteva [70] established convergence results for central differencing on two types of layer-adapted meshes:

$$\left\| u - u^N \right\|_{\infty,\omega} \leq \begin{cases} CN^{-2} & \text{for Bakhvalov meshes with } \sigma > 2, \\ CN^{-2} \ln^2 N & \text{for Shishkin meshes with } \sigma > 2; \end{cases} \tag{4.67}$$

The (l_∞, l_1) stability (4.66) was used by Roos and Linß [138] to prove

$$\left\| u - u^N \right\|_{\infty,\omega} \leq C \left(h + \max |\psi'| \, N^{-1} \right)^2 \tag{4.68}$$

on S-type meshes with $\sigma \geq 3$. A similar result was given by Kopteva and Linß [71] for certain quasilinear problems of type (4.49).

Another approach to study central differencing on Shishkin meshes is that of Lenferink [76, 77]. He eliminates every other unknown to get a scheme whose system matrix is an M-matrix.

4.3.2.3 Derivative Approximation

For the central-difference scheme (4.64) on S-type meshes with $\sigma \geq 3$ we have the second order bound

$$\varepsilon \left| u^N_{\bar{x};i} - u'_{i-1/2} \right| \leq C\big(h + N^{-1} \max |\psi'|\big)^2.$$

The proof in [138] uses the bound (4.68) for the discretisation error, then interprets the scheme as a finite element method with inexact integration and finally applies a finite element technique [173] to get the bound for the derivative approximation.

4.3.3 Convergence Acceleration Techniques

In the early 1980s Hemker [51] proposed the use of defect-correction methods when solving singularly perturbed problems. However, the first rigorous proof of uniform convergence of a defect-correction scheme was not published before 2001 (Fröhner et al. [43]). Various analyses by Nikolova and Axelsson [16, 126] are at least not rigorous with regard to the robustness, i. e. the ε-independence of the error constants, while the analysis by Fröhner and Roos [44] turned out to be technically unsound [42].

4.3.3.1 Defect Correction

Let us consider the defect correction method from [43] for our model convection-diffusion problem in conservative form:

$$\mathcal{L}^c u := -\varepsilon u'' - (bu)' + cu = f \ \text{ in } \ (0,1), \ \ u(0) = \gamma_0, \ \ u(1) = \gamma_1. \tag{4.9}$$

It is based on the upwind scheme

$$\left[L^c u^N\right]_i := -\varepsilon u^N_{\bar{x}x;i} - \left(bu^N\right)_{x;i} + c_i u^N_i = f_i \tag{4.10}$$

combined with the unstabilised second-order central difference scheme

$$\left[\widehat{L}^c u^N\right]_i := -\varepsilon u^N_{\bar{x}\hat{x};i} - \left(bu^N\right)_{\mathring{x};i} + c_i u^N_i = f_i.$$

With this notation we can formulate the defect correction method. This two-stage method is the following:

1. Compute an initial first-order approximation $\tilde{u}^N$ using simple upwinding:

$$\left[L^c\tilde{u}^N\right]_i = f_i \quad \text{for } i = 1,\dots,N-1,\ \ \tilde{u}^N_0 = \gamma_0,\ \ \tilde{u}^N_N = \gamma_1. \qquad (4.69a)$$

2. Estimate the defect ζ in the differential equation by means of the central difference scheme:

$$\zeta_i = [\widehat{L}^c\tilde{u}^N]_i - f_i \quad \text{for } i = 1,\dots,N-1. \qquad (4.69b)$$

3. Compute the defect correction Δ by solving

$$[L^c\Delta]_i = \kappa_i\zeta_i, \qquad \text{for } i = 1,\dots,N-1,\ \ \Delta_0 = \Delta_N = 0. \qquad (4.69c)$$

   with $\kappa_i = \hbar_i / h_{i+1}$.
4. Then the final computed solution is

$$u^N = \tilde{u}^N - \Delta. \qquad (4.69d)$$

Remark 4.36. At first glance both the upwind discretisation and the particular weighting of the residual in (4.69c) appear a bit non-standard. No justification for these choices is provided by [43, 93]. An argument that suggests this particular choice is presented in [101].

Furthermore, the weighting becomes the standard $\kappa_i = 1$ on uniform meshes; however, when used on non-uniform meshes, $\kappa_i = 1$ might reduce the order of convergence which is illustrated by numerical experiments in [101]. ♣

In the analysis of the method we use the following notation:

$$(\mathcal{A}^c v)(x) := \varepsilon v'(x) + \big(bv\big)(x) + \int_x^1 \big(cv\big)(s)\mathrm{d}s, \quad \mathcal{F}(x) := \int_x^1 f(s)\mathrm{d}s,$$

$$\left[A^c v\right]_i := \varepsilon v_{\bar{x};i} + (bv)_i + \sum_{k=i}^{N-1} h_{k+1}\,(cv)_k\,, \quad F_i := \sum_{k=i}^{N-1} h_{k+1} f_k$$

and

$$\left[\widehat{A}^c v\right]_i := \varepsilon v_{\bar{x};i} + \frac{(bv)_i + (bv)_{i-1}}{2} + \sum_{k=i}^{N} \hbar_k (cv)_k,\ \ \widehat{F}_i := \sum_{k=i}^{N} \hbar_k f_k.$$

The differential equation (4.9) yields

$$\mathcal{A}^c u - \mathcal{F} \equiv \alpha = \text{const}, \qquad (4.70)$$

while (4.69b) and (4.69c) imply

$$A^c\Delta - (\widehat{A}^c\tilde{u}^N - \widehat{F}) \equiv a = \text{const}\,. \tag{4.71}$$

A priori analysis

The negative norm stability (4.8c) of the operator L^c yields for the error of the defect-correction method

$$\begin{aligned}\left|\left|\left|u-u^N\right|\right|\right|_{\varepsilon,\infty,\omega} \le 2\left\|(A^c-\widehat{A}^c)(u-\tilde{u}^N)\right\|_{\infty,\omega} \\ + 2\left\|\widehat{A}^c u - \widehat{F} + \alpha\right\|_{\infty,\omega},\end{aligned} \tag{4.72}$$

by (4.70) and (4.71), where α is the constant from (4.70).

The second term in (4.72) is the truncation error of the central difference scheme. It is formally of second order. The first term is the so called **relative consistency error**. While the error $u-\tilde{u}^N$ of the simple upwind scheme is only of first order, the hope is that A^c and $\widehat{A}^c$ are sufficiently close to gain second order in this term too.

Consider the relative consistency error first. Let $\eta := u - \tilde{u}^N$ denote the error of the simple upwind scheme. A straight-forward calculation and summation by parts give

$$\left[(A^c-\widehat{A}^c)\eta\right]_i = \frac{(b\eta)_i-(b\eta)_{i-1}}{2} + \sum_{k=i+1}^{N-1} h_{k+1}\frac{(c\eta)_{k-1}-(c\eta)_k}{2} - \frac{h_i}{2}(c\eta)_i,$$

which can be bounded by

$$\begin{aligned}\left|\left[(A^c-\widehat{A}^c)\eta\right]_i\right| \le \left(\|b\|_\infty + \frac{\|c\|_\infty}{2}\right)\max_{i=1,\dots,N}|\eta_i-\eta_{i-1}| \\ + h\left(\|b'\|_\infty + \frac{\|c'\|_\infty+\|c\|_\infty}{2}\right)\|\eta\|_{\infty,\omega}.\end{aligned}$$

Thus,

$$\left\|(A^c-\widehat{A}^c)\eta\right\|_{\infty,\omega} \le C\left(\max_{i=1,\dots,N}|\eta_i-\eta_{i-1}| + h\|\eta\|_{\infty,\omega}\right) \le C\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2, \tag{4.73}$$

by (4.50) and because $h \le \vartheta_{cd}^{[1]}(\bar{\omega}) \le \vartheta_{cd}^{[2]}(\bar{\omega})$.

Remark 4.37. The first term, the maximum difference of the error of the upwind scheme in two adjacent mesh points, constituted the main difficulty in [43]. With the error expansion of Sect. 4.2.3 this has become a simple task. ♣

Next, let us bound the truncation error of the central difference scheme. By (4.70) we have $(\widehat{A}^c u - \widehat{F})_i - \alpha = (\widehat{A}^c u - \widehat{F})_i - (\mathcal{A}u - \mathcal{F})_{i-1/2}$. Hence,

$$\begin{aligned} \left|(\widehat{A}^c u - \widehat{F})_i - \alpha\right| \le \varepsilon \left|u_{\bar{x};i} - u'_{i-1/2}\right| + \left|\frac{(bu)_i + (bu)_{i-1}}{2} - (bu)_{i-1/2}\right| \\ + \left|\sum_{k=i}^{N} \hbar_k g_k - \int_{x_{i-1/2}}^{1} g(s)\mathrm{d}s\right| \end{aligned} \tag{4.74}$$

with $g = f - cu$. Using Taylor expansions for u, u' and $(bu)'$ about $x = x_i$, we obtain

$$\varepsilon \left|u_{\bar{x};i} - u'_{i-1/2}\right| \le \frac{3\varepsilon}{2} \int_{I_i} (s - x_{i-1})\left|u'''(s)\right|\mathrm{d}s \le C\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2 \tag{4.75}$$

and

$$\left|\frac{(bu)_i + (bu)_{i-1}}{2} - (bu)_{i-1/2}\right| \le \frac{3}{2} \int_{I_i} (s - x_{i-1})\left|(bu)''(s)\right|\mathrm{d}s \le C\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2,$$

by (4.9), (3.30) and Lemma 4.16.

For the last term in (4.74) a Taylor expansion gives

$$\begin{aligned} \left|\frac{h_k}{2} g_k - \frac{h_k^2}{8} g'_{k-1/2} - \int_{x_{k-1/2}}^{x_k} g(s)\mathrm{d}s\right| &\le \frac{h_k^3}{8} \|g''\|_{\infty,(x_{k-1/2},x_k)} \\ &\le C h_k^3 \left(1 + \varepsilon^{-2} \mathrm{e}^{-\beta x_{k-1/2}/\varepsilon}\right) \le C h_k \left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2, \end{aligned} \tag{4.76a}$$

where we have used (3.30) and Proposition 4.15 with $x = x_{k-1/2}$ and $\sigma = 2$. Furthermore, we have

$$\begin{aligned} \left|\frac{h_{k+1}}{2} g_k + \frac{h_{k+1}^2}{8} g'_{k+1/2} - \int_{x_k}^{x_{k+1/2}} g(s)\mathrm{d}s\right| \\ \le h_{k+1} \int_{x_k}^{x_{k+1/2}} (\sigma - x_k)\left|g''(\sigma)\right|\mathrm{d}\sigma \le C h_{k+1} \left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2, \end{aligned} \tag{4.76b}$$

by (3.30) and Lemma 4.16. Combine these two estimates:

$$\begin{aligned} \left|\sum_{k=i}^{N} \hbar_k g_k - \int_{x_{i-1/2}}^{1} g(s)\mathrm{d}s\right| \\ \le C\left\{\vartheta_{cd}^{[2]}(\bar{\omega})^2 + h_i^2\left(1 + \varepsilon^{-2}\mathrm{e}^{-\beta x_{i-1/2}/\varepsilon}\right)\right\} \le C\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2, \end{aligned}$$

by Proposition 4.15.

Therefore,

$$\left\|\widehat{A}^c u - \widehat{F} - \alpha\right\|_{\infty,\omega} \le C \left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2 ,$$

by (4.9), (3.30) and Lemma 4.16.

Collect (4.72), (4.73) and the last inequality to get the main result of this section.

Theorem 4.38. *Let u be the solution of (4.9) and u^N that of the defect correction method (4.69). Then*

$$|||u - u^N|||_{\varepsilon,\infty,\omega} \le C \left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2 .$$

Derivative approximation

For $x \in I_i$ a triangle inequality gives

$$\begin{aligned} \varepsilon \left|u'(x) - u^N_{\bar{x};i}\right| &\le \varepsilon \left|u'(x) - u_{\bar{x};i}\right| + \varepsilon \left|(u^I - u^N)_{\bar{x};i}\right| \\ &\le \varepsilon \left|u'(x) - u_{\bar{x};i}\right| + C \left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2 , \end{aligned}$$

by Theorem 4.38. The representation

$$u'(x) - u_{\bar{x};i} = \frac{1}{h_i} \int_{I_i} \int_s^x u''(t)\mathrm{d}t\mathrm{d}s. \tag{4.77}$$

yields

$$\varepsilon \left|u'(x) - u_{\bar{x};i}\right| \le C\vartheta_{cd}^{[1]}(\bar{\omega}), \quad \text{by (3.30)}.$$

Thus, in general we only have a first-order approximation for the ε-weighted derivative:

$$\varepsilon \left|u'(x) - u_{\bar{x};i}\right| \le C\vartheta_{cd}^{[2]}(\bar{\omega}) \quad \text{for } x \in I_i.$$

This result is sharp.

For the midpoint $x_{i-1/2}$ of the mesh interval I_i we expand (4.77) to give

$$u'(x_{i-1/2}) - u_{\bar{x};i} = \frac{1}{h_i} \int_{I_i} \int_s^x \int_{x_{i-1/2}}^t u'''(\xi)\mathrm{d}\xi\mathrm{d}t\mathrm{d}s.$$

The right-hand side can be bounded using (3.30) and Lemma 4.16. We get

$$\varepsilon \left|u'(x_{i-1/2}) - u_{\bar{x};i}\right| \le C \left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2 ,$$

and finally

$$\varepsilon \left|u'(x_{i-1/2}) - u^N_{\bar{x};i}\right| \le C \left(\vartheta^{[2]}_{cd}(\bar{\omega})\right)^2 .$$

This means that the midpoints of the mesh intervals are superconvergence points for the derivative and we can define a recovery operator R for the derivative. For a given mesh function $v \in I\!R^{N+1}$, let Rv be that function that is piecewise linear on the mesh $\hat{\omega} = \left\{0, x_{1+1/2}, x_{2+1/2}, \ldots, x_{N-1-1/2}, 1\right\}$ and satisfies

$$(Rv)_{i-1/2} = v_{\bar{x};i} \quad \text{for } \ i = 1, \ldots, N.$$

Then one can prove

$$\varepsilon \left\|u' - Ru^N\right\|_\infty \le C \left(\vartheta^{[2]}_{cd}(\bar{\omega})\right)^2 .$$

A posteriori analysis

The next result is an extension of Lemma 2.2 in [64] which gave bounds in the mesh points only. It is an essential ingredient for the analysis of second-order approximations.

Theorem 4.39. *Let the hypothesis of Theorem 3.45 be satisfied. Let ψ be the solution of the boundary value problem*

$$\mathcal{L}^c \psi = -\Psi' \quad \textit{in } \ (0,1), \ \ u(0) = u(1) = 0$$

with

$$\Psi(x) = A_{i-1/2}(x - x_{i-1/2}) \quad \textit{for } \ x \in (x_{i-1}, x_i).$$

Then

$$\|\psi\|_\infty \le C^* \max_{i=1,\ldots,N} \left\{ \left|A_{i-1/2}\right| \min\left[\frac{h_i}{\|b\|_\infty}, \frac{h_i^2}{4\varepsilon}\right] \right\},$$

where

$$C^* = \frac{2\|b\|_\infty + \|c\|_\infty + \beta}{2\beta}.$$

Proof. Let $x \in (0,1)$ be arbitrary, but fixed. The Green's function representation gives

$$u(x) = \int_0^1 \partial_\xi \mathcal{G}(x,\xi) F(\xi) \mathrm{d}\xi = \sum_{i=1}^N A_{i-1/2} J_i$$

with

$$J_i := \int_{I_i} \partial_\xi \mathcal{G}(x,\xi)(\xi - x_{i-1/2})\mathrm{d}\xi.$$

A first bound for the J_i's is

$$|J_i| \le \left| \int_{I_i} \partial_\xi \mathcal{G}(x,\xi)(\xi - x_{i-1/2})\mathrm{d}\xi \right| \le \frac{h_i}{2} \int_{I_i} |\partial_\xi \mathcal{G}(x,\xi)| \,\mathrm{d}\xi, \tag{4.78}$$

while integration by parts yields

$$J_i = \int_{I_i} \partial_\xi^2 \mathcal{G}(x,\xi) \left[\frac{h_i^2}{8} - \frac{(\xi - x_{i-1/2})^2}{2} \right] \mathrm{d}\xi.$$

Hence

$$\begin{aligned} |J_i| &\le \frac{h_i^2}{8} \int_{I_i} \left| \partial_\xi^2 \mathcal{G}(x,\xi) \right| \mathrm{d}\xi \\ &\le \frac{h_i^2}{8\varepsilon} \int_{I_i} \left[\delta_x(\xi) + \|b\|_\infty \, |\partial_\xi \mathcal{G}(x,\xi)| + c(\xi)\mathcal{G}(x,\xi) \right] \mathrm{d}\xi. \end{aligned}$$

This estimate is combined with (4.78) to give

$$|J_i| \le \min \left[\frac{h_i^2}{8\varepsilon}, \frac{h_i}{2\|b\|_\infty} \right] \int_{I_i} \left[\delta_x(\xi) + \|b\|_\infty \, |\partial_\xi \mathcal{G}(x,\xi)| + \|c\|_\infty \mathcal{G}(x,\xi) \right] \mathrm{d}\xi.$$

Multiply by $\left|A_{i-1/2}\right|$, take sums for $i = 1, \dots, N$, use a discrete Hölder inequality and note that

$$\int_0^1 \left[\delta_x(\xi) + \|b\|_\infty \, |\partial_\xi \mathcal{G}(x,\xi)| + \|c\|_\infty \mathcal{G}(x,\xi) \right] \mathrm{d}\xi \le 1 + \frac{2\|b\|_\infty + \|c\|_\infty}{\beta},$$

by Theorem 3.20. This completes the proof. □

With these stability results at hand we can now derive our a posteriori error bounds. We shall identify any mesh function v with its piecewise linear nodal interpolant.

Theorem 4.40. *Let the hypothesis of Theorem 3.45 be satisfied. Set $g := f - cu^N$. Then the error of the defect-correction method satisfies*

$$\left\| u - u^N \right\|_\infty \le \eta := \eta_1 + \eta_2 + \eta_3 + \eta_4 + \eta_5$$

with

$$\eta_1 := C^* \max_{i=1,\dots,N} \min\left\{\frac{h_i}{\|b\|_\infty}, \frac{h_i^2}{\varepsilon}\right\} \left|g_{i-1/2} + \left(bu^N\right)_{\bar{x},i}\right|,$$

$$\eta_2 := \frac{1}{\beta} \max_{i=1,\dots,N} h_i \left|(b\Delta)_{\bar{x},i}\right|, \quad \eta_3 := \frac{1}{\beta} \max_{i=1,\dots,N-1} \left|\sum_{k=i}^{N-1} \frac{h_{k+1}-h_k}{2} c_k \Delta_k\right|,$$

$$\eta_4 := \frac{1}{6\beta} \sum_{i=1}^{N} h_i^3 \left\|g''\right\|_{\infty,I_i}$$

and

$$\eta_5 := \frac{3}{4\beta} \max_{i=1,\dots,N} h_i^2 \left\{2\left\|g'\right\|_{\infty,I_i} + \left\|(bu^N)''\right\|_{\infty,I_i}\right\}.$$

Proof. By (4.70) and (4.71) we have, for $x \in (x_{i-1}, x_i)$,

$$\begin{aligned}&\mathcal{A}^c\left(u-u^N\right)(x)\\ &\quad= \mathcal{F}(x) - \widehat{F}_i + \left[\widehat{A}^c u^N\right]_i - \left(\mathcal{A}^c u^N\right)(x) - \left[\left(A^c - \widehat{A}^c\right)\Delta\right]_i + \alpha - a.\end{aligned}$$

Recalling the definitions of $\mathcal{F}$, $\widehat{F}$, $\mathcal{A}^c$, A^c and $\widehat{A}^c$, we obtain the representation

$$\begin{aligned}&\mathcal{A}^c\left(u-u^N\right)(x)\\ &\quad= \int_x^1 g(s)\mathrm{d}s - \sum_{k=i}^{N} \hbar_k g_k + \frac{\left(bu^N\right)_i + \left(bu^N\right)_{i-1}}{2} - \left(bu^N\right)(x)\\ &\qquad - \frac{h_i}{2}\left(b\Delta\right)_{\bar{x},i} - \sum_{k=i}^{N-1} \frac{h_{k+1}-h_k}{2} c_k \Delta_k + \alpha - a.\end{aligned} \tag{4.79}$$

Taylor expansions yield

$$\int_x^1 g(s)\mathrm{d}s - \sum_{k=i}^{N} \hbar_k g_k = \int_{x_i}^1 (g-g^I)(s)\mathrm{d}s + (x_{i-1/2}-x)g_{i-1/2} + \mu_i(x),$$

where g^I is the piecewise linear interpolant of g, and

$$\frac{\left(bu^N\right)_i + \left(bu^N\right)_{i-1}}{2} - \left(bu^N\right)(x) = (x_{i-1/2}-x)\left(bu^N\right)_{\bar{x},i} + \tilde{\mu}_i(x)$$

with

$$\|\mu_i\|_{\infty,I_i} \le \frac{3h_i^2}{4}\|g'\|_{\infty,I_i} \quad \text{and} \quad \|\tilde{\mu}_i\|_{\infty,I_i} \le \frac{3h_i^2}{8}\left\|(bu^N)''\right\|_{\infty,I_i}$$

Substitute the above two equations into (4.79). We get

$$\begin{aligned}
&\mathcal{A}\left(u-u^N\right)(x)\\
&\quad = \int_{x_i}^1 (g-g^I)(s)\mathrm{d}s + \left(x_{i-1/2}-x\right)\left(g_{i-1/2}+\left(bu^N\right)_{\bar{x},i}\right)\\
&\qquad - \frac{h_i}{2}\left(b\Delta\right)_{\bar{x},i} - \sum_{k=i}^{N-1}\frac{h_{k+1}-h_k}{2}c_k\delta_k + \left(\mu_i+\tilde{\mu}_i\right)(x) + \alpha - a.
\end{aligned}$$

Furthermore,

$$\left|\int_{x_i}^1 (g-g^I)(s)\mathrm{d}s\right| \le \frac{1}{12}\sum_{i=1}^N h_i^3\,\|g''\|_{\infty,I_i}\,.$$

Finally, note that $(\mathcal{A}v)' = -\mathcal{L}v$. Use Theorems 3.45 and 4.39 to complete the proof. □

Remark 4.41. The error estimate of Theorem 4.40 contains terms, namely η_4 and η_5, that in general have to be approximated, for example

$$g' \approx \frac{g_i - g_{i-1}}{h_i}, \quad g'' \approx 4\frac{g_i - 2g_{i-1/2} + g_{i-1}}{h_i^2}$$

and

$$(bu^N)'' \approx 4\frac{(bu^N)_i - 2(bu^N)_{i-1/2} + (bu^N)_{i-1}}{h_i^2}.$$

The additional errors introduced this way are of third order and therefore decay rapidly when the mesh is refined. ♣

An adaptive mesh algorithm

Based on Theorem 4.40, the de Boor algorithm described in Sect. 4.2.4.2 can be adapted for the defect-correction method by choosing

$$Q_i = Q_i(u^N, \Delta, \omega) := \left\{\rho_0 + \sum_{k=1}^5 \rho_k\eta_{k;i}\right\}^{1/2} \tag{4.80}$$

with

$$\eta_{1;i} := \min\left\{\frac{h_i}{\|b\|_\infty}, \frac{h_i^2}{4\varepsilon}\right\}\left|g_{i-1/2} + \left(bu^N\right)_{\bar{x},i}\right|, \quad \eta_{2;i} := h_i\left|(b\Delta)_{\bar{x},i}\right|,$$
$$\eta_{3;i} := \left|\sum_{k=i}^{N-1}\frac{h_{k+1}-h_k}{2}c_k\Delta_k\right|, \quad \eta_{4;i} := \frac{2}{3\beta}\left|g_i - 2g_{i-1/2} + g_{i-1}\right|,$$
$$\eta_{5;i} := \frac{1}{2\beta}h_i\left|g_i - g_{i-1}\right| + \left|(bu^N)_i - 2(bu^N)_{i-1/2} + (bu^N)_{i-1}\right|$$

and non-negative weights ρ_ℓ.

Remark 4.42. The square root in (4.80) is necessary because the underlying method is formally of second order. ♣

Remark 4.43. The numerical experiments in [101] indicate that η_1 contains sufficient information to steer the mesh adaptation. Therefore η_k, $k = 2, \dots, 5$, can be set to zero, however ρ_0 must not in order to avoid mesh starvation in regions where the solution does not vary much. This reduces the computational costs in the remeshing phase of the de Boor algorithm. ♣

4.3.3.2 Richardson Extrapolation

Richardson extrapolation on layer-adapted meshes was first analysed by Natividad and Stynes [124]. They study a simple first-order upwind scheme on a Shishkin mesh and prove that Richardson extrapolation improves the accuracy to almost second order, although the underlying scheme is only of first order. The analysis in [124] is based on comparison principles and barrier function techniques.

Here we shall pursue an alternative approach similar to the one in [93] that is based on the $(l_\infty, w^{-1,\infty})$ stability and on the error expansion of Sect. 4.2.3. Again consider the conservative form of our model problem:

$$\mathcal{L}^c u := -\varepsilon u'' - (bu)' + cu = f \quad \text{in } (0,1), \;\; u(0) = \gamma_0, \;\; u(1) = \gamma_1. \tag{4.9}$$

Given an arbitrary mesh $\bar{\omega}$, let $\bar{\omega}' : 0 = x_{1/2} < x_1 < x_{1+1/2} < \cdots < x_N = 1$ be the mesh obtained by uniformly bisecting $\bar{\omega}$. Let $\tilde{u}^N$ be the solution of the upwind scheme (4.10) on $\bar{\omega}$ and

$$u^{2N} = \left(u_0^{2N}, u_{1/2}^{2N}, u_1^{2N}, \dots, u_{N-1/2}^{2N}, u_N^{2N}\right)$$

that of the difference scheme on $\tilde{\omega}$. Since (4.10) is a first-order scheme we combine $\tilde{u}^N$ and $\tilde{u}^{2N}$ by

$$u_i^N := 2\tilde{u}_i^{2N} - \tilde{u}_i^N \quad \text{for } i = 0, \dots, N,$$

in order to get a second-order approximation defined on the coarser mesh $\bar{\omega}$.

In addition to the notation introduced on p. 122 set

$$\left[\tilde{A}^c v\right]_i := 2\varepsilon \frac{v_i - v_{i-1}}{h_i} + b_i v_i + \sum_{k=i}^{N-1} h_{k+1} \frac{c_k v_k + c_{k+1/2} v_{k+1/2}}{2}$$

and

$$\tilde{F}_i := \sum_{k=i}^{N-1} h_{k+1} \frac{f_k + f_{k+1/2}}{2}.$$

The differential equation (4.9) and the difference equation (4.10) yield

$$\mathcal{A}^c u - \mathcal{F} \equiv \alpha, \quad A^c \tilde{u}^N - F \equiv a \quad \text{and} \quad \tilde{A}^c \tilde{u}^{2N} - \tilde{F} \equiv \tilde{a}.$$

A direct calculation gives

$$\begin{aligned} & A^c(2\tilde{u}^{2N} - u^N - u)_i + (\mathcal{A}^c u - \mathcal{F})_{i-1/2} \\ &\quad = -\varepsilon \left(\frac{u_i - u_{i-1}}{h_i} - u'_{i-1/2} \right) \\ &\qquad + \left\{ b_i(\tilde{u}_i^{2N} - u_i) - b_{i-1/2}(\tilde{u}_{i-1/2}^{2N} - u_{i-1/2}) \right\} \\ &\qquad - \left\{ \frac{h_i}{2} (c\tilde{u}^{2N} - cu)_{i-1/2} \right. \\ &\qquad\qquad \left. + \sum_{k=i}^{N-1} h_{k+1} \left[(c\tilde{u}^{2N} - cu)_{k+1/2} - (c\tilde{u}^{2N} - cu)_k \right] \right\} \\ &\qquad + \left\{ \int_{i-1/2}^{1} g(s) \mathrm{d}s - \frac{h_i}{2} g_{i-1/2} - \sum_{k=i}^{N-1} h_{k+1} g_{k+1/2} \right\} \end{aligned}$$

with $g = cu - f$. The first term on the right-hand side is bounded by $C\vartheta_{cd}^{[2]}(\bar{\omega})^2$, see (4.75) The second and third term can be bounded by $C\vartheta_{cd}^{[2]}(\tilde{\omega})^2$ using the technique that gave (4.73). The last term is also bounded by $C\vartheta_{cd}^{[2]}(\bar{\omega})^2$, since similar to (4.76) we have

$$\left| \frac{h_k}{2} g_{k-1/2} - \frac{h_k^2}{8} g'_{k-1/2} - \int_{x_{k-1/2}}^{x_k} g(s) \mathrm{d}s \right| \le C h_k \left(\vartheta_{cd}^{[2]}(\bar{\omega}) \right)^2,$$

and

$$\left| \frac{h_{k+1}}{2} g_{k+1/2} + \frac{h_{k+1}^2}{8} g'_{k+1/2} - \int_{x_k}^{x_{k+1/2}} g(s) \mathrm{d}s \right| \le C h_{k+1} \left(\vartheta_{cd}^{[2]}(\bar{\omega}) \right)^2.$$

Finally, using the stability inequality (4.8c) we obtain the following convergence result.

Theorem 4.44. *Let u^N be the approximate solution to* (4.9) *obtained by the Richardson extrapolation technique applied to the simple upwind scheme* (4.10)*. Then*

$$\left|\left|\left|u - u^N\right|\right|\right|_{\varepsilon,\infty,\omega} \le C\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2.$$

Corollary 4.45. *Theorem 4.44 and interpolation error bounds (see Sect. 5.1) give*

$$\left\|u - u^N\right\|_\infty + \varepsilon\left\|u' - Ru^N\right\|_\infty \le C\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2,$$

where R is the recovery operator from Sect. 4.3.3.1.

4.3.4 A Numerical Example

The following tables contain the results of test computations for test problem (4.14):

$$-\varepsilon u'' - u' + 2u = \mathrm{e}^{x-1}, \quad u(0) = u(1) = 0.$$

We test the performance of:

- second-order upwinding
- central differencing
- defect correction and
- Richardson extrapolation.

In the experiments we have chosen $\varepsilon = 10^{-8}$. The meshes have been constructed with parameters $\sigma = 2$, $\beta = 1$ and $q = K = 1/2$.

The numerical results in Tables 4.3 and 4.4 are clear illustrations of the convergence estimates of Theorems 4.31, 4.38 and 4.44 and of (4.67). Furthermore, as the theory predicts, all four methods give higher accuracy on Bakhvalov meshes than on Shishkin meshes.

Finally, we consider a modified Shishkin mesh which is constructed as follows. Pick the transition point $\tau = 2\varepsilon\beta^{-1}\ln N$ as usual. Set $h = 2\tau/N$ and $H = 2(1-\tau)/N$. Then the mesh is defined by

$$h_i = \begin{cases} h & \text{if } i \le N/2, \\ 4H/3 & \text{if } i \text{ is odd and } i > N/2, \\ 2H/3 & \text{if } i \text{ is even and } i > N/2. \end{cases}$$

Thus, instead of a uniform mesh on each of the two subdomains $[0,\tau]$ and $[\tau,1]$ we use non-uniform, though very regular sub meshes. A similar mesh was considered in [26].

Table 4.3 Second-order schemes on Shishkin meshes

N	second order upwinding		central differencing		defect correction		Richardson extrapolation	
2^7	3.29e-04	1.61	1.37e-04	1.55	2.32e-04	1.64	2.33e-04	1.57
2^8	1.07e-04	1.66	4.66e-05	1.62	7.47e-05	1.68	7.86e-05	1.64
2^9	3.40e-05	1.70	1.52e-05	1.67	2.33e-05	1.72	2.53e-05	1.69
2^{10}	1.05e-05	1.73	4.79e-06	1.70	7.06e-06	1.75	7.86e-06	1.72
2^{11}	3.17e-06	1.75	1.47e-06	1.73	2.10e-06	1.78	2.38e-06	1.75
2^{12}	9.43e-07	1.77	4.43e-07	1.76	6.12e-07	1.79	7.07e-07	1.77
2^{13}	2.77e-07	1.79	1.31e-07	1.78	1.76e-07	1.81	2.07e-07	1.79
2^{14}	8.02e-08	1.80	3.83e-08	1.79	5.03e-08	1.82	5.97e-08	1.81
2^{15}	2.30e-08	1.81	1.11e-08	1.81	1.42e-08	1.83	1.71e-08	1.82
2^{16}	6.54e-09	1.83	3.16e-09	1.82	4.00e-09	1.84	4.84e-09	1.83
2^{17}	1.85e-09	1.84	8.95e-10	1.83	1.12e-09	1.85	1.36e-09	1.84
2^{18}	5.17e-10	—	2.52e-10	—	3.10e-10	—	3.81e-10	—

Table 4.4 Second-order schemes on Bakhvalov meshes

N	second order upwinding		central differencing		defect correction		Richardson extrapolation	
2^7	5.81e-04	2.36	9.67e-05	1.99	2.74e-04	1.94	1.16e-04	1.94
2^8	1.13e-04	2.11	2.43e-05	2.02	7.14e-05	1.97	3.01e-05	1.97
2^9	2.62e-05	2.05	5.98e-06	1.99	1.83e-05	1.98	7.67e-06	1.99
2^{10}	6.34e-06	2.27	1.51e-06	2.00	4.63e-06	1.99	1.93e-06	1.99
2^{11}	1.31e-06	2.36	3.77e-07	2.00	1.16e-06	2.00	4.86e-07	2.00
2^{12}	2.55e-07	2.11	9.45e-08	2.01	2.92e-07	2.00	1.22e-07	2.00
2^{13}	5.90e-08	2.05	2.34e-08	1.99	7.30e-08	2.00	3.05e-08	2.00
2^{14}	1.43e-08	2.05	5.89e-09	2.00	1.83e-08	2.00	7.63e-09	2.00
2^{15}	3.46e-09	2.00	1.47e-09	2.00	4.57e-09	2.00	1.91e-09	2.00
2^{16}	8.64e-10	2.00	3.68e-10	2.01	1.14e-09	2.00	4.77e-10	2.00
2^{17}	2.16e-10	2.00	9.16e-11	2.00	2.86e-10	2.00	1.19e-10	2.00
2^{18}	5.40e-11	—	2.30e-11	—	7.14e-11	—	2.98e-11	—

For this mesh $\vartheta_{cd}^{[2]}(\bar{\omega}) \leq CN^{-1}\ln N$ and almost second order convergence is guaranteed for the upwind scheme, defect correction and Richardson extrapolation. This order of convergence is observed in our computational experiments; see Table 4.5. However, for central differencing the observed rate is only one. Thus, on this mesh the assumption (4.65) must be violated. In [26] it is shown that for $b = \text{const}$ and $c \equiv 0$ the stability constant in (4.66) blows up for $N \to \infty$.

Remark 4.46. This means that for central differencing a general result like

$$\left\|u - u^N\right\|_{\infty,\omega} \leq C\left(\vartheta_{cd}^{[2]}(\bar{\omega})\right)^2$$

cannot hold. ♣

Table 4.5 Second-order schemes on modified Shishkin meshes

N	second order upwinding		central differencing		defect correction		Richardson extrapolation	
2^7	1.30e-03	1.61	6.73e-03	0.98	5.94e-04	1.46	7.66e-04	1.46
2^8	4.27e-04	1.66	3.41e-03	0.99	2.16e-04	1.57	2.77e-04	1.58
2^9	1.35e-04	1.70	1.72e-03	0.99	7.25e-05	1.65	9.25e-05	1.65
2^{10}	4.17e-05	1.72	8.63e-04	1.00	2.31e-05	1.70	2.94e-05	1.70
2^{11}	1.26e-05	1.75	4.32e-04	1.00	7.11e-06	1.74	9.04e-06	1.74
2^{12}	3.76e-06	1.77	2.16e-04	1.00	2.13e-06	1.77	2.71e-06	1.76
2^{13}	1.10e-06	1.79	1.08e-04	1.00	6.24e-07	1.79	7.97e-07	1.79
2^{14}	3.20e-07	1.80	5.41e-05	1.00	1.81e-07	1.81	2.31e-07	1.80
2^{15}	9.18e-08	1.81	2.70e-05	1.00	5.16e-08	1.82	6.63e-08	1.82
2^{16}	2.61e-08	1.82	1.35e-05	1.00	1.46e-08	1.83	1.88e-08	1.83
2^{17}	7.37e-09	1.84	6.74e-06	1.02	4.11e-09	1.84	5.31e-09	1.84
2^{18}	2.07e-09	—	3.33e-06	—	1.15e-09	—	1.49e-09	—

This observation has far reaching consequences, in particular in higher dimensions, where it is difficult, if not impossible, to construct uniform or nearly uniform meshes. Therefore, stabilisation in regions where the mesh is coarse becomes essential.

4.4 Systems

We now leave the scalar convection-diffusion equation and move on to systems of equations of this type.

4.4.1 Weakly Coupled Systems in One Dimension

Consider the weakly coupled problem from Sect. 3.4.2:

$$\begin{aligned} \boldsymbol{\mathcal{L}u} := -\operatorname{diag}(\boldsymbol{\varepsilon})\boldsymbol{u}'' - \operatorname{diag}(\boldsymbol{b})\boldsymbol{u}' + \boldsymbol{Au} = \boldsymbol{f} \quad \text{on } (0,1), \\ \boldsymbol{u}(0) = \boldsymbol{u}(1) = \boldsymbol{0}, \end{aligned} \tag{4.81}$$

where $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_\ell)^T$ and the small parameter ε_k is in $(0,1]$ for $k = 1, \dots, \ell$. Assume that for each k one has $a_{kk} \geq 0$ and either $b_k \geq \beta_k$ or $b_k \leq -\beta_k$ on $[0,1]$ with positive constants b_k.

We follow [96] and discretise (4.81) by means of the simple upwind scheme that was studied in detail in Sect. 4.2: Find $\boldsymbol{u} \in \left(I\!R_0^{N+1}\right)^\ell$ such that

$$\left[\boldsymbol{L}\boldsymbol{u}^N\right]_i = \boldsymbol{f}_i \quad \text{for } k = 1, \dots, N-1, \tag{4.82}$$

where $\boldsymbol{f}_i = \boldsymbol{f}(x_i) = (f_{1;i}, f_{2;i}, \ldots, f_{\ell,i})^T$, $\boldsymbol{L}\boldsymbol{v} := \big((\boldsymbol{L}\boldsymbol{v})_1, (\boldsymbol{L}\boldsymbol{v})_2, \ldots, (\boldsymbol{L}\boldsymbol{v})_\ell\big)^T$,

$$(L\boldsymbol{v})_k := \Lambda_k v_k + \sum_{\substack{m=1\\ m\neq k}}^{\ell} a_{km} v_m$$

and

$$[\Lambda_k v]_i := \begin{cases} -\varepsilon_k v_{\bar{x}x;i} - b_{k;i} v_{x;i} + a_{kk;i} v_i & \text{if } b_k > 0, \\ -\varepsilon_k v_{x\bar{x};i} - b_{k;i} v_{\bar{x};i} + a_{kk;i} v_i & \text{if } b_k < 0. \end{cases}$$

4.4.1.1 Stability

The stability analysis for the discrete operator is conducted along the lines of the continuous analysis. By the definition of $\boldsymbol{L}$ and the Λ_k's we have, for any vector-valued mesh function $\boldsymbol{v} \in \left(\mathbb{R}_0^{N+1}\right)^\ell$,

$$\Lambda_k v_k = -\sum_{\substack{m=1\\ m\neq k}}^{\ell} a_{km} v_m + (\boldsymbol{L}\boldsymbol{v})_k \quad \text{on } \omega, \ \ k = 1, \ldots, \ell. \tag{4.83}$$

Then Theorem 4.3 yields

$$\|v_k\|_{\infty,\omega} + \sum_{\substack{m=1\\ m\neq k}}^{\ell} \tilde{\gamma}_{km} \|u_m\|_{\infty,\omega} \le \min\left\{ \left\| \frac{(\boldsymbol{L}\boldsymbol{v})_k}{a_{kk}} \right\|_{\infty,\omega}, \left\| \frac{(\boldsymbol{L}\boldsymbol{v})_k}{b_k} \right\|_{\infty,\omega} \right\}$$

for, $k = 1, \ldots, \ell$, where the $\ell \times \ell$ constant matrix $\tilde{\boldsymbol{\Gamma}} = \tilde{\boldsymbol{\Gamma}}(\boldsymbol{A}, \boldsymbol{b}) = (\tilde{\gamma}_{km})$ is—as in Sect. 3.4.2—defined by

$$\tilde{\gamma}_{kk} = 1 \quad \text{and} \quad \tilde{\gamma}_{km} = -\min\left\{ \left\| \frac{a_{km}}{a_{kk}} \right\|_{\infty,\omega}, \left\| \frac{a_{km}}{b_k} \right\|_{\infty,\omega} \right\} \quad \text{for } k \neq m.$$

We reach the following stability result.

Theorem 4.47. *Assume that the matrix $\boldsymbol{A}$ has non-negative diagonal entries. Assume that $\tilde{\boldsymbol{\Gamma}}(\boldsymbol{A})$ is inverse-monotone. Then for $i = 1, \ldots, \ell$ one has*

$$\|v_i\|_{\infty,\bar{\omega}} \le \sum_{k=1}^{\ell} \big(\tilde{\boldsymbol{\Gamma}}^{-1}\big)_{ik} \min\left\{ \left\| \frac{(\boldsymbol{L}\boldsymbol{v})_k}{a_{kk}} \right\|_{\infty,\omega}, \left\| \frac{(\boldsymbol{L}\boldsymbol{v})_k}{b_k} \right\|_{\infty,\omega} \right\}$$

for any mesh function $\boldsymbol{v} \in \left(\mathbb{R}_0^{N+1}\right)^\ell$.

Remark 4.48. Theorem 4.47 implies

$$\|\boldsymbol{v}\|_{\infty,\bar{\omega}} := \max_{k=1,\dots,\ell} \|v_k\|_{\infty,\bar{\omega}} \le C \|\boldsymbol{L}\boldsymbol{v}\|_{\infty,\omega} \quad \text{for all } \boldsymbol{v} \in \left(I\!R_0^{N+1}\right)^\ell,$$

i.e., the operator $\boldsymbol{L}$ is $(\ell_\infty, \ell_\infty)$-stable although it does not obey a comparison principle. ♣

Corollary 4.49. *Under the hypotheses of Theorem 4.47 the discrete problem* (4.81) *has a unique solution* $\boldsymbol{u}^N$*, and*

$$\left\|\boldsymbol{u}^N\right\|_{\infty,\bar{\omega}} \le C \|\boldsymbol{f}\|_{\infty,\omega}$$

for some constant C.

Remark 4.50. One can also use (4.8c) when bounding $(\boldsymbol{L}\boldsymbol{v})_k$ in (4.83) to establish that

$$\|\boldsymbol{v}\|_{\infty,\bar{\omega}} \le C \max_{k=1,\dots,\ell} \|(\boldsymbol{L}\boldsymbol{v})_k\|_{-1,\infty} \quad \text{for all } \boldsymbol{v} \in \left(I\!R_0^{N+1}\right)^\ell.$$

This allows to analyse the difference scheme (4.82) when applied to problems whose source terms consist of generalised functions like the δ-distribution. ♣

4.4.1.2 A Priori Error Analysis

Following [96], we split the error $\boldsymbol{\eta} := \boldsymbol{u} - \boldsymbol{u}^N$ into two parts $\boldsymbol{\psi}, \boldsymbol{\varphi} \in (I\!R_0^{N+1})^\ell$ as $\boldsymbol{\eta} = \boldsymbol{\psi} + \boldsymbol{\varphi}$ with

$$\Lambda_k \psi_k = (\boldsymbol{L}\boldsymbol{\eta})_k \quad \text{and} \quad \Lambda_k \varphi_k = -\sum_{\substack{m=1\\ m\neq k}}^{\ell} a_{km}\eta_m \ \text{ on } \omega, \quad k = 1, \dots, \ell.$$

A triangle inequality and Theorem 4.3 yield

$$\|\eta_k\|_{\infty,\omega} \le \|\psi_k\|_{\infty,\omega} - \sum_{\substack{m=1\\ m\neq k}}^{\ell} \tilde{\gamma}_{km} \|\eta_m\|_{\infty,\omega}, \quad k = 1, \dots, \ell.$$

Assuming that $\tilde{\boldsymbol{\Gamma}}$ is an M-matrix, we obtain

$$\left\|\boldsymbol{u} - \boldsymbol{u}^N\right\|_{\infty,\omega} \le C \|\boldsymbol{\psi}\|_{\infty,\omega},$$

and we are left with bounding $\boldsymbol{\psi}$.

The components of $\boldsymbol{\psi}$ are the solutions of scalar problems to which the technique of Sect. 4.2.2 can be applied. The following general error bound is obtained.

Theorem 4.51. *Assume that the matrix $\boldsymbol{A}$ has non-negative diagonal entries. Assume that $\tilde{\boldsymbol{\Gamma}}(\boldsymbol{A})$ is inverse-monotone. Let $\boldsymbol{u}$ and $\boldsymbol{u}^N$ be the solutions of* (4.81) *and* (4.82)*. Then*

$$\left\|\boldsymbol{u}-\boldsymbol{u}^N\right\|_{\infty,\bar{\omega}} \leq C \max_{i=1,\ldots,N} \int_{I_i} \left[1+\sum_{m=1}^{\ell} \left|u_m'(s)\right|\right] \mathrm{d}s.$$

Corollary 4.52. *The a priori bounds on the u_m' in Sect. 3.4.2.2 can be used to derive more explicit error bounds:*

$$\left\|\boldsymbol{u}-\boldsymbol{u}^N\right\|_{\infty,\bar{\omega}} \leq C\vartheta_{cd,\ell}^{[1]}(\bar{\omega}).$$

Note, the quantity $\vartheta_{cd,\ell}^{[p]}(\bar{\omega})$, $p>0$, has been defined in Sect. 2.1.6.

Remark 4.53. We immediately obtain, for example,

$$\left\|u-u^N\right\|_{\infty,\omega} \leq \begin{cases} CN^{-1} & \text{for Bakhvalov meshes,} \\ CN^{-1}\ln N & \text{for Shishkin meshes,} \end{cases}$$

when the mesh parameters satisfy $\sigma_m \geq 1$. ♣

4.4.1.3 A Posteriori Error Analysis

Alternatively, one can appeal to the strong stability (3.29c) of the scalar continuous operators and combine the arguments of Sect. 4.2.4 and 4.4.1.1, in order to get the a posteriori bound

$$\left\|\boldsymbol{u}-\boldsymbol{u}^N\right\|_{\infty} \leq C \max_{k=0,\ldots,N-1} h_{k+1} \left[1+\sum_{m=1}^{\ell} \left|u_{m,x;k}^N\right|\right].$$

The constant(s) involved in this error bound can be specified explicitly; cf. Sect. 4.2.4.

4.4.2 Strongly Coupled Systems

We now consider strongly coupled systems of convection-diffusion type, i.e., for each k one has $b_{km} \not\equiv 0$ for some $k \neq m$. Strong coupling causes interactions between boundary layers that are not fully understood at present. The main papers on this problem are [1, 100, 127, 128].

The general strongly coupled two-point boundary-value problem in conservative form is: Find $\boldsymbol{u} \in \left(C^2(0,1) \cap C[0,1]\right)^\ell$ such that

$$\begin{aligned} \boldsymbol{\mathcal{L}}^{\boldsymbol{c}}\boldsymbol{u} := -\operatorname{diag}(\boldsymbol{\varepsilon})\boldsymbol{u}'' - (\boldsymbol{B}\boldsymbol{u})' + \boldsymbol{A}\boldsymbol{u} = \boldsymbol{f} \quad \text{in } \Omega := (0,1), \\ \boldsymbol{u}(0) = \boldsymbol{u}(1) = \boldsymbol{0}, \end{aligned} \tag{4.84}$$

where as before $\boldsymbol{f} = (f_1, \ldots, f_\ell)^T$, while $\boldsymbol{A} = (a_{km})$ and $\boldsymbol{B} = (b_{km})$ are $\ell \times \ell$ matrices, and the $\ell \times \ell$ matrix $\operatorname{diag}(\boldsymbol{\varepsilon})$ is diagonal with k^{th} entry ε_k for all k. Furthermore, let $b_{kk} \geq \beta_k$ or $b_{kk} \leq -\beta_k$ on $[0,1]$ with positive constants β_k.

We follow [100] and discretise (4.84) using the upwind scheme of Sect. (4.2) for each equation of the system: Find $\boldsymbol{u}^N \in (\mathbb{R}_0^{N+1})^\ell$ such that

$$\left[\boldsymbol{L}^{\boldsymbol{c}}\boldsymbol{u}^N\right]_i = \boldsymbol{f}_i \quad \text{for } i = 1, \ldots, N-1, \tag{4.85}$$

where $\boldsymbol{L}^{\boldsymbol{c}}\boldsymbol{v} := \left((\boldsymbol{L}^{\boldsymbol{c}}\boldsymbol{v})_1, (\boldsymbol{L}^{\boldsymbol{c}}\boldsymbol{v})_2, \ldots, (\boldsymbol{L}^{\boldsymbol{c}}\boldsymbol{v})_\ell\right)^T$,

$$(\boldsymbol{L}^{\boldsymbol{c}}\boldsymbol{v})_k := \Lambda_k^c v_k - \sum_{\substack{m=1 \\ m \neq k}}^{\ell} (b_{km}v_m)_x + \sum_{\substack{m=1 \\ m \neq k}}^{\ell} a_{km}v_m \quad \text{if } b_{kk} > 0,$$

$$(\boldsymbol{L}^{\boldsymbol{c}}\boldsymbol{v})_k := \Lambda_k^c v_k - \sum_{\substack{m=1 \\ m \neq k}}^{\ell} (b_{km}v_m)_{\bar{x}} + \sum_{\substack{m=1 \\ m \neq k}}^{\ell} a_{km}v_m \quad \text{if } b_{kk} < 0,$$

and

$$[\Lambda_k^c v]_i := \begin{cases} -\varepsilon_k v_{\bar{x}x;i} - (b_{kk}v)_{x;i} + a_{kk;i}v_i & \text{if } b_{kk} > 0, \\ -\varepsilon_k v_{x\bar{x};i} - (b_{kk}v)_{\bar{x};i} + a_{kk;i}v_i & \text{if } b_{kk} < 0. \end{cases}$$

4.4.2.1 Stability

The stability analysis for the difference operator $\boldsymbol{L}$ is analogous to that for the continuous operator $\boldsymbol{\mathcal{L}}$ in Sect. 3.4.3.

Define the $\ell \times \ell$ matrix $\boldsymbol{\Upsilon}_\omega = \boldsymbol{\Upsilon}_\omega(\boldsymbol{A}, \boldsymbol{B}) = (\gamma_{km})$ by

$$\gamma_{kk} = 1 \quad \text{and} \quad \gamma_{km} = -\frac{2\,\|b_{km}\|_{\infty,\omega} + \|a_{km}\|_{1,\omega}}{\beta_k} \quad \text{for } k \neq m.$$

Introduce the discrete maximum norms

$$|||v|||_{\varepsilon_k,\infty,\omega} := \varepsilon_k \,\|v_x\|_{\infty,\omega} + \beta_k \,\|v\|_{\infty,\bar{\omega}} \quad \text{for } v \in \mathbb{R}^{N+1}$$

and

$$|||\boldsymbol{v}|||_{\boldsymbol{\varepsilon},\infty,\omega} := \max_{k=1,\dots,\ell} |||v_k|||_{\varepsilon_k,\infty,\omega} \quad \text{for } \boldsymbol{v} \in \left(\mathbb{R}^{N+1}\right)^{\ell}.$$

Theorem 4.54. *Assume that for each $k = 1, \dots, \ell$*

$$b_{kk} \le -\beta_k \quad \textit{or} \quad b_{kk} \ge \beta_k \quad \textit{on } [0,1]$$

with positive constants β_k and that

$$a_{kk} \ge 0 \quad \textit{and} \quad b'_{kk} \ge 0 \quad \textit{on } [0,1].$$

Suppose $\boldsymbol{\Upsilon}_\omega(\boldsymbol{A}, \boldsymbol{B})$ is inverse-monotone. Then the operator $\boldsymbol{L^c}$ is $(\ell_\infty, w^{-1,\infty})$ stable with

$$|||v_k|||_{\varepsilon_k,\infty,\omega} \le \sum_{m=1}^{\ell} \left(\boldsymbol{\Upsilon}_\omega^{-1}\right)_{km} \|(\boldsymbol{L^c v})_m\|_{-1,\infty,\omega} \quad \textit{for } k = 1, \dots, \ell,$$

and for all $\boldsymbol{v} \in (\mathbb{R}_0^{N+1})^\ell$.

Proof. For the sake of simplicity in the presentation, we restrict ourselves to the case when $b_{kk} > 0$ for all k.

Let $\boldsymbol{v} \in (\mathbb{R}_0^{N+1})^\ell$ be arbitrary. Then the definition of $\boldsymbol{L^c}$ and the Λ_k^c yields

$$\Lambda_k^c v_k = (\boldsymbol{L^c v})_k + \sum_{\substack{m=1\\m\ne k}}^{\ell} (b_{km} v_m)_x - \sum_{\substack{m=1\\m\ne k}}^{\ell} a_{km} v_m$$

Apply Theorem 4.3 and Remark 4.4 to get

$$|||v_k|||_{\varepsilon_k,\infty,\omega} \le \sum_{\substack{m=1\\m\ne k}}^{\ell} \Big(2\,\|b_{km}\|_{\infty,\omega} + \|a_{km}\|_{1,\omega}\Big)\, \|v_m\|_{\infty,\omega} + \|(\boldsymbol{L^c v})_k\|_{-1,\infty,\omega}\,.$$

Recall the definition of $\boldsymbol{\Upsilon}_\omega$ and rearrange the last inequality

$$\sum_{m=1}^{\ell} \gamma_{km}\, |||v_m|||_{\varepsilon_m,\infty,\omega} \le \|(\boldsymbol{L^c v})_k\|_{-1,\infty,\omega} \quad \text{for } k = 1, \dots, \ell.$$

Using the inverse monotonicity of $\boldsymbol{\Upsilon}_\omega$, we are finished. □

Corollary 4.55. *Suppose the hypotheses of Theorem 4.54 are satisfied. Then the difference equation (4.85) possesses a unique solution $\boldsymbol{u}^N$, with*

$$|||\boldsymbol{u}^N|||_{\boldsymbol{\varepsilon},\infty,\omega} \le C \max_{m=1,\dots,\ell} \|f_m\|_{-1,\infty,\omega}$$

for some constant C that is independent of both ε and the mesh.

Remark 4.56. In general, the operator $\boldsymbol{L}^c$ does not obey a comparison principle. Nonetheless it is $(\ell_\infty, \ell_\infty)$-stable, i.e.,

$$\|\boldsymbol{v}\|_{\infty,\bar{\omega}} \le C\,\|\boldsymbol{L}^c\boldsymbol{v}\|_{\infty,\omega} \quad \text{for all } \boldsymbol{v} \in \left(I\!R_0^{N+1}\right)^\ell ,$$

by Theorem 4.54. ♣

4.4.2.2 A Priori Error Analysis

Adapt the argument of Sect. 4.4.1.2 as in [100] and split the error $\boldsymbol{\eta} := \boldsymbol{u}^N - \boldsymbol{u}$ into two parts $\boldsymbol{\psi}, \boldsymbol{\varphi} \in (I\!R_0^{N+1})^\ell$ as $\boldsymbol{\eta} = \boldsymbol{\psi} + \boldsymbol{\varphi}$ with

$$\Lambda_k^c \psi_k = (\boldsymbol{L}\boldsymbol{\eta})_k \quad \text{and} \quad \Lambda_k^c \varphi_k = \sum_{\substack{m=1\\ m\ne k}}^{\ell} \left((b_{km}\eta_m)_x - a_{km}\eta_m\right) \quad \text{on } \omega.$$

Recalling the definition of $\boldsymbol{\Upsilon}_\omega = (\gamma_{km})$, we use a triangle inequality and Theorem 4.3 to obtain

$$\sum_{m=1}^{\ell} \gamma_{km}\, |||\eta_m|||_{\varepsilon_m\infty,\omega} \le |||\psi_k|||_{\varepsilon_k,\infty,\omega} \quad \text{for } k = 1,\dots,\ell.$$

Next, if $\boldsymbol{\Upsilon}$ is inverse monotone then

$$|||\boldsymbol{u} - \boldsymbol{u}^N|||_{\boldsymbol{\varepsilon},\infty,\omega} \le C\,|||\boldsymbol{\psi}|||_{\boldsymbol{\varepsilon},\infty,\omega} ,$$

and we are left with bounding $|||\psi_k|||_{\varepsilon_k,\infty,\omega}$, $k = 1,\dots,\ell$.

The components of $\boldsymbol{\psi}$ are the solutions of scalar problems. In [100] the technique of Sect. 4.2.2 is used to obtain the following general error bound.

Theorem 4.57. *Let the hypothesis of Theorem 4.54 be satisfied. Then the error of the upwind scheme (4.85) applied to (4.84) satisfies*

$$|||\boldsymbol{u} - \boldsymbol{u}^N|||_{\boldsymbol{\varepsilon},\infty,\omega} \le C \max_{i=1,\dots,N} \int_{I_i} \left[1 + \sum_{m=1}^{\ell} |u_m'(s)|\right] \mathrm{d}s.$$

Corollary 4.58. *By (3.56) we have $\|u_k'\|_1 \le C$ for $k = 1, \dots, \ell$. Therefore, there exists a mesh ω^* such that*

$$\int_{I_i} \left(1 + \sum_{m=1}^{\ell} |u_m'(x)|\right) \mathrm{d}x = \frac{1}{N} \int_0^1 \left(1 + \sum_{m=1}^{\ell} |u_m'(x)|\right) \mathrm{d}x \le CN^{-1}$$

and on this mesh one consequently has

$$|||\boldsymbol{u} - \boldsymbol{u}^N|||_{\boldsymbol{\varepsilon},\infty,\omega^*} \le CN^{-1}.$$

Remark 4.59. As satisfactory pointwise bounds on $|u_k'|$ are unavailable, this result does not give an immediate explicit convergence result on, e.g., a Bakhvalov or Shishkin mesh. ♣

Remark 4.60. When $\varepsilon_k = \varepsilon$ for $k = 1, \dots, \ell$, Theorem 3.57 yields

$$|||\boldsymbol{u} - \boldsymbol{u}^N|||_{\boldsymbol{\varepsilon},\infty,\omega} \le C \max_{i=1,\dots,N} \int_{I_i} \left(1 + \mathrm{e}^{-\beta x/\varepsilon}\right) \mathrm{d}x = C\vartheta_{cd}^{[1]}(\bar{\omega}).$$

The system behaves like the scalar equation of Section 4.2 and appropriately adapted meshes can be constructed as for scalar problems.

In [127] one also finds an error analysis for a system with a single parameter, but the analysis is limited to Shishkin meshes and uses a more traditional truncation error and barrier function argument. Furthermore, higher regularity of the solution is required. On the other hand, in certain situations the analysis of [127] is valid under less restrictive hypotheses on the entries of the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ than the requirement that $\boldsymbol{\Upsilon}_\omega$ be inverse-monotone. ♣

4.4.2.3 A Posteriori Error Bounds

Using the strong stability results of Theorem 3.54, we can follow [100] to obtain the a posteriori error bound

$$|||\boldsymbol{u} - \boldsymbol{u}^N|||_{\boldsymbol{\varepsilon},\infty} \le C \max_{i=1,\dots,N} h_i \left(1 + \sum_{m=1}^{\ell} |u_{m;\bar{x};i}^N|\right).$$

Remark 4.61. The de Boor algorithm (see Sect. 4.2.4.2) can be used to adaptively generate meshes for (4.84) by choosing

$$Q_i = h_i \left(1 + \sum_{m=1}^{\ell} \left(u_{m;\bar{x};i}^N\right)^2\right)^{1/2}.$$

Numerical examples are presented in [100], but a complete analysis of the adaptive algorithm is not given. ♣

4.4.2.4 Numerical Results

We now present the results of two numerical experiments in order to illustrate the conclusions of Theorem 4.57.

First, we consider a test problem with two equations.

$$-\operatorname{diag}(\boldsymbol{\varepsilon})\,\boldsymbol{u}'' - \begin{pmatrix} 4+x & -1 \\ 1-2x & -3 \end{pmatrix} \boldsymbol{u}' + \begin{pmatrix} 0 & -x^2 \\ 1 & 1 \end{pmatrix} \boldsymbol{u} = \begin{pmatrix} \mathrm{e}^{1-2x} \\ \cos 2x \end{pmatrix} \text{ in } (0,1), \quad (4.86)$$

subject to homogeneous Dirichlet boundary conditions $\boldsymbol{u}(0) = \boldsymbol{u}(1) = \mathbf{0}$.

For this problem our convergence theory applies since

$$\boldsymbol{\Gamma} = \frac{1}{12}\begin{pmatrix} 12 & -7 \\ -12 & 12 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Gamma}^{-1} = \frac{1}{5}\begin{pmatrix} 12 & 7 \\ 12 & 12 \end{pmatrix} \geq 0.$$

We have $b_1 \geq 4$ and $-b_2 \geq 3$. Therefore, one expects that the solution exhibits two layers: one at $x = 0$ behaving like $\mathrm{e}^{-4x/\varepsilon_1}$ and the other at $x = 1$ that behaves like $\mathrm{e}^{-3(1-x)/\varepsilon_2}$. Also the first-order derivative of $\boldsymbol{u}$ is expected to satisfy

$$|u_i'(x)| \leq C\left\{1 + \varepsilon_1^{-1} e^{-4x/\varepsilon_1} + \varepsilon_2^{-1} e^{-3(1-x)/\varepsilon_2}\right\}, \quad i = 1, 2.$$

Note, we do not have proper proof for these derivative bounds. The difficulties in proving them have been explained in Section 3.4.3.1.

The exact solution of (4.86) is not available. Therefore, we compare the numerical solution with that obtained by Richardson extrapolation as before. We consider Bakhvalov and Shishkin meshes and the adaptive de Boor algorithm; see Remark 4.61. The construction of layer-adapted meshes for overlapping layers is explained in Section 2.1.6.

The results of our test computations are contained in Table 4.6. For both *a priori* adapted meshes the expected (almost) first order of uniform convergence is confirmed. For the adaptive algorithm first order is also observed although the numerical rates are “less stable”.

Table 4.6 Simple upwinding for a system of two convection-diffusion equations

N	Shishkin mesh		Bakhvalov mesh		adaptive algorithm	
	η_N	ρ^N	η_N	r^N	η_N	r^N
$3 \cdot 2^{7}$	3.131e-02	0.83	8.194e-03	0.99	7.600e-03	0.96
$3 \cdot 2^{8}$	1.843e-02	0.88	4.119e-03	0.99	3.903e-03	1.05
$3 \cdot 2^{9}$	1.051e-02	0.91	2.073e-03	0.99	1.880e-03	0.99
$3 \cdot 2^{10}$	5.870e-03	0.94	1.044e-03	0.99	9.477e-04	0.98
$3 \cdot 2^{11}$	3.229e-03	0.95	5.240e-04	1.00	4.821e-04	0.94
$3 \cdot 2^{12}$	1.758e-03	0.97	2.620e-04	1.00	2.505e-04	1.00
$3 \cdot 2^{13}$	9.491e-04	0.98	1.308e-04	1.00	1.254e-04	1.15
$3 \cdot 2^{14}$	5.092e-04	0.98	6.540e-05	1.00	5.643e-05	0.91
$3 \cdot 2^{15}$	2.717e-04	0.99	3.269e-05	1.00	3.006e-05	0.99
$3 \cdot 2^{16}$	1.444e-04	—	1.634e-05	—	1.516e-05	—

Table 4.7 Simple upwinding for a convection-diffusion problem with three equations

N	Shishkin mesh		Bakhvalov mesh		adaptive algorithm	
	η_N	ρ^N	η_N	r^N	η_N	r^N
$4 \cdot 2^3$	2.453e-01	0.49	1.780e-01	0.86	1.850e-01	0.64
$4 \cdot 2^4$	1.809e-01	0.63	9.777e-02	0.93	1.187e-01	0.75
$4 \cdot 2^5$	1.218e-01	0.75	5.124e-02	0.97	7.062e-02	0.87
$4 \cdot 2^6$	7.612e-02	0.84	2.623e-02	0.98	3.863e-02	0.96
$4 \cdot 2^7$	4.505e-02	0.90	1.327e-02	0.99	1.991e-02	0.92
$4 \cdot 2^8$	2.569e-02	0.94	6.673e-03	1.00	1.051e-02	1.02
$4 \cdot 2^9$	1.430e-02	0.97	3.346e-03	1.00	5.167e-03	0.97
$4 \cdot 2^{10}$	7.824e-03	0.98	1.676e-03	1.00	2.642e-03	1.06
$4 \cdot 2^{11}$	4.233e-03	1.00	8.384e-04	1.00	1.270e-03	1.05
$4 \cdot 2^{12}$	2.272e-03	—	4.193e-04	—	6.115e-04	—

The second test problem consists of three convection-diffusion equations.

$$-\operatorname{diag}(\boldsymbol{\varepsilon})\,\boldsymbol{u}'' - \begin{pmatrix} 3 & 1 & 0 \\ -x^2 & 5+x & -1 \\ 1-x & 0 & -5 \end{pmatrix} \boldsymbol{u}' = \begin{pmatrix} \mathrm{e}^x \\ \cos x \\ 1+x^2 \end{pmatrix} \quad \text{in } (0,1),$$

with boundary conditions $\boldsymbol{u}(0) = \boldsymbol{u}(1) = \mathbf{0}$. This time:

$$\boldsymbol{\Gamma} = \frac{1}{15}\begin{pmatrix} 15 & 10 & 0 \\ 7 & 15 & 6 \\ 9 & 0 & 15 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Gamma}^{-1} = \frac{1}{119}\begin{pmatrix} 225 & 150 & 60 \\ 159 & 225 & 90 \\ 135 & 90 & 155 \end{pmatrix} \geq 0.$$

We expect layers $\mathrm{e}^{-3x/\varepsilon_1}$, $\mathrm{e}^{-5x/\varepsilon_2}$ and $\mathrm{e}^{-5(1-x)/\varepsilon_3}$ to form and adapt the mesh accordingly.

Table 4.7 gives the numerical results for our second example. A comparison with Table 4.6 reveals that the behaviour of the method is similar to that for Example (4.86).

4.5 Problems with Turning Point Layers

This section considers linear convection-diffusion problems with a boundary turning point: Find $u \in C^2(0,1) \cap C[0,1]$ such that

$$\mathcal{L}u := -\varepsilon u'' - pbu' + pcu = f \quad \text{in } (0,1), \;\; u(0) = \gamma_0, \; u(1) = \gamma_1, \qquad (4.87)$$

where $p(x) = x^\kappa$, $\kappa > 0$, $b \geq \beta > 0$ and $c \geq 0$ in $(0,1)$.

We are aware of four publications analysing numerical methods for this problem with $\kappa = 1$. Liseikin [113] constructs a special transformation and solves the transformed problem on a uniform mesh. The method obtained is proved to be first-order uniformly convergent in the discrete maximum norm. Vulanović [167] studies an upwind-difference scheme on a layer-adapted Bakhvalov-type mesh and proves convergence in a discrete ℓ_1 norm. This result is generalised in [170] for quasilinear problems. However, this norm fails to capture the layers present in the solution. Therefore, the problem is not singularly perturbed in the sense of Def. 1.1. In [112] the authors establish almost first-order convergence for an upwind-difference scheme on a Shishkin mesh. Here we follow [92] and study (4.87) with arbitrary $\kappa > 0$.

4.5.1 A First-Order Upwind Scheme

The boundary-value problem (4.87) is discretised using simple upwinding: Find $u \in I\!R^{N+1}$ such that

$$\left[Lu^N\right]_i = p_i f_i \quad \text{for } i = 1, \dots, N-1, \;\; u_0^N = \gamma_0, \; u_N^N = \gamma_1 \tag{4.88}$$

with

$$[Lv]_i := -\varepsilon v_{\bar{x}x;i} - p_i b_i v_{x;i} + p_i c_i v_i .$$

4.5.1.1 Stability of the Discretisation

The matrix associated with it is an L_0-matrix. Lemma 3.14 with the test function $e_i = 1 - x_i$ verifies that it is an M-matrix. Therefore, the operator L satisfies a discrete comparison principle. That is, for any mesh functions $v, w \in I\!R^{N+1}$

$$\left.\begin{aligned} Lv &\le Lw \quad \text{on } \omega, \\ v_0 &\le w_0, \\ v_N &\le w_N \end{aligned}\right\} \implies v \le w \quad \text{on } \bar{\omega}.$$

Lemma 3.17 with $e_i = 1 - x_i$ yields

$$\|v\|_{\infty,\bar{\omega}} \le \|Lv/pb\|_{\infty,\omega} \quad \text{for all } v \in I\!R_0^{N+1}.$$

Alternatively, if $c > 0$ on $[0,1]$, then Lemma 3.17 with $\boldsymbol{e} \equiv 1$ gives

$$\|v\|_{\infty,\bar{\omega}} \le \|Lv/pc\|_{\infty,\omega} \quad \text{for all } v \in I\!R_0^{N+1}.$$

Thus, the operator L is $(\ell_\infty, \ell_{\infty,1/p})$ stable in the $1/p$-weighted maximum norm.

For the solution u^N of (4.88) this implies

$$\left\|u_i^N\right\|_{\infty,\bar\omega} \le \max\{|\gamma_0|,|\gamma_1|\} + \min\left\{\left\|\frac{f}{b}\right\|_{\infty,\omega}, \left\|\frac{f}{c}\right\|_{\infty,\omega}\right\}.$$

Green's function.

Lemma 4.62. *Assume that*

$$p > 0, \quad p' \ge 0, \quad b \ge \beta > 0 \quad \textit{and} \quad c \ge 0 \ \textit{on} \ (0,1). \tag{4.89}$$

Then

$$0 \le G_{i,j} \le \hat G_{i,j} := \frac{1}{p_j\beta}\begin{cases} 1 & \textit{for } i = 0,\dots,j, \\ \prod_{\nu=j+1}^{i}\left(1+\frac{\beta p_\nu h_{\nu+1}}{\varepsilon}\right)^{-1} & \textit{for } i = j+1,\dots,N. \end{cases}$$

Proof. Let j be arbitrary, but fixed. $G_{\cdot,j}$ solves

$$[LG_{\cdot,j}]_i = \delta_{i,j}, \quad i = 1,\dots,N-1, G_{0,j} = G_{N,j} = 0$$

with

$$\delta_{i,j} = \begin{cases} h_{i+1}^{-1} & \text{for } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

We shall show that $\hat G_{\cdot,j}$ is a barrier function for $G_{\cdot,j}$. Clearly $\hat G_{0,j} \ge 0$ and $\hat G_{N,j} \ge 0$.

Next, verify that

$$\hat G_{x;i,j} = \hat G_{i,j}\begin{cases} 0 & \text{for } i = 0,\dots,j-1, \\ -\dfrac{\beta p_i}{\varepsilon+\beta p_i h_{i+1}} & \text{for } i = j,\dots,N-1, \end{cases}$$

and

$$\hat G_{\bar x;i,j} = \hat G_{i,j}\begin{cases} 0 & \text{for } i = 1,\dots,j, \\ -\dfrac{\beta p_{i-1}}{\varepsilon} & \text{for } i = j+1,\dots,N. \end{cases}$$

Hence,

$$\left[L\hat G_{\cdot,j}\right]_i \ge p_i c_i \hat G_{i,j} \ge 0 \quad \text{for } i = 1,\dots,j-1,$$

$$\left[L\hat{G}_{\cdot,j}\right]_j \geq \left(\frac{\beta pj}{h_{j+1}} + p_j c_j\right)\hat{G}_{j,j} \geq \frac{1}{h_{i+1}},$$

and

$$\left[L\hat{G}_{\cdot,j}\right]_i \geq \left(\beta\frac{p_i - p_{i-1}}{h_{i+1}} + p_i c_i\right)\hat{G}_{i,j} \geq 0 \quad \text{for } i = j+1,\dots,N-1,$$

because $p' \geq 0$ on $[0,1]$. Thus, $\hat{G}$ is a barrier function for G. □

Remark 4.63. The proof of Lemma 4.62 simplifies the argument in [112] where a barrier function was constructed for the adjoint problem. ♣

Remark 4.64. Numerical results indicate that when $p(x) = x^\kappa$, $\kappa \geq 0$, one has the sharper bound

$$G_{i,j} \leq C\left(\varepsilon^{1/(\kappa+1)} + \xi_j\right)^{-\kappa} \quad \text{for all } i,j = 1,\dots,N-1,$$

but, we do not have a rigorous proof for this.

Theorem 4.65. *Assume the data satisfies* (4.89)*. Then the discrete operator L is $(\ell_\infty, \ell_{1,1/p})$ stable*

$$\|v\|_{\infty,\omega} \leq \beta^{-1}\|Lv/p\|_{1,\omega}$$

for all $v \in I\!R_0^{N+1}$ with a $1/p$-weighted ℓ_1 norm.

Proof. For any function $v \in I\!R_0^{N+1}$ we have

$$v_i = \sum_{j=1}^{N} h_{j+1} G_{i,j}\left[Lv\right]_j, \quad i = 1,\dots,N-1.$$

Then Lemma 4.62 yields the assertion of the theorem. □

Remark 4.66. An immediate consequence of Theorem 4.65 for the simple upwind scheme is

$$\left\|u - u^N\right\|_{\infty,\omega} \leq \beta^{-1}\left\|\left(Lu - pf\right)/p\right\|_{1,\omega}.$$

Thus, the error of the numerical solution in the maximum norm is bounded by an ℓ_1-type norm of the truncation error weighted with the inverse of the coefficient of the convection term. This was used in [112] to establish uniform almost first-order convergence on Shishkin meshes for $\kappa = 1$. ♣

4.5.2 Convergence on Shishkin Meshes

In [112] convergence of the upwind scheme (4.88) applied to (4.87) with $\kappa = 1$ was studied. Starting from the observation that for any fixed $m > 0$ there exists a constant $C = C(m)$ such that

$$\exp\left(-\frac{\tilde{\beta}x^2}{2\varepsilon}\right) \le C \exp\left(-m\frac{x}{\sqrt{\varepsilon}}\right)$$

a piecewise uniform mesh is constructed as follows: Fix the mesh transition point

$$\tau = \min\left\{q, \frac{2\sqrt{\varepsilon}}{m}\ln N\right\}.$$

Then divide $[0,\tau]$ uniformly into qN subintervals and $[\tau,1]$ into $(1-q)N$ subintervals.

Using the stability inequality of Theorem 4.65, it is then shown that

$$\left\|u - u^N\right\|_{\infty,\omega} \le CN^{-1}(\ln N)^2.$$

The details of the analysis are similar to the argument in Sect. 4.2.5.

The general case of an arbitrary $\kappa > 0$ has been considered in [92]. We will give a brief summary of that paper now. This time the transition point τ is chosen as follows:

$$\tau = \min\left\{q, \left(\sigma\frac{\varepsilon(\kappa+1)}{\tilde{\beta}}\ln N\right)^{1/(\kappa+1)}\right\}.$$

Here we shall consider $\tau < q$ which is the interesting case.

First Lemma 4.62 is sharpened to

$$G_{i,j} \le \begin{cases} C\tau\varepsilon^{-1} & \text{for } j = 1,\dots,qN-1, \\ \beta^{-1}\xi_j^{-\kappa} & \text{for } j = qN,\dots,N-1. \end{cases}$$

Note that $\tau\varepsilon^{-1} = C\mu^{-\kappa}(\ln N)^{1/(\kappa+1)}$.

Remark 4.67. It is argued in [92] based on numerical evidence that the logarithmic factor is superfluous and one has

$$G_{i,j} \le C(\mu + \xi_j)^{-\kappa} \quad \text{for } i,j = 1,\dots,N-1,$$

but no rigorous analysis is provided. ♣

The analysis in [92] proceeds along the lines of Sect. 4.2.5 using the solution decomposition in Theorem 3.63 to establish

$$\left\|u - u^N\right\|_{\infty,\omega} \leq CN^{-1} (\ln N)^{2/(\kappa+1)} \tag{4.90}$$

if $\sigma \geq 2$.

Remark 4.68. If one had $G_{i,j} \leq C (\mu + \xi_j)^{-\kappa}$, then (4.90) could be sharpened to

$$\left\|u - u^N\right\|_{\infty,\omega} \leq CN^{-1} (\ln N)^{1/(\kappa+1)} .$$

Also we do not have any theory for arbitrary meshes. This is due to a lack of stronger negative-norm stability inequalities for both the continuous and the discrete operators. More work in this direction is required. ♣

4.5.3 A Numerical Example

We verify experimentally the convergence result of (4.90). Our test problem is the semilinear differential equation

$$-\varepsilon u''(x) - x^\kappa (2-x)u'(x) + x^\kappa \mathrm{e}^{u(x)} = 0 \quad \text{for } x \in (0,1),$$
$$u(0) = u(1) = 0.$$

The exact solution of this problem is not available. We therefore estimate the accuracy of the numerical solution by comparing it with the numerical solution on a higher order method: Richardson extrapolation. For our tests we take $\tilde{\beta} = 1$, $q = 1/2$ and $\varepsilon = 10^{-12}$.

Table 4.8 displays the results of the numerical test. For comparison reasons Table 4.9 contains the rates which can be expected if the error bound

Table 4.8 Simple upwinding on Shishkin meshes for turning point problems

	$\kappa = 1/2$		$\kappa = 1$		$\kappa = 2$		$\kappa = 3$	
N	error	rate	error	rate	error	rate	error	rate
2^7	6.171e-3	0.88	5.335e-3	0.92	4.675e-3	0.95	4.411e-3	0.96
2^8	3.358e-3	0.90	2.829e-3	0.93	2.426e-3	0.96	2.270e-3	0.97
2^9	1.803e-3	0.91	1.484e-3	0.94	1.249e-3	0.96	1.160e-3	0.97
2^{10}	9.592e-4	0.92	7.737e-4	0.95	6.401e-4	0.97	5.899e-4	0.98
2^{11}	5.069e-4	0.93	4.014e-4	0.95	3.269e-4	0.97	2.993e-4	0.98
2^{12}	2.666e-4	0.93	2.075e-4	0.96	1.666e-4	0.98	1.516e-4	0.98
2^{13}	1.396e-4	0.94	1.070e-4	0.96	8.473e-5	0.98	7.669e-5	0.98
2^{14}	7.292e-5	0.94	5.506e-5	0.96	4.305e-5	0.98	3.876e-5	0.99
2^{15}	3.798e-5	0.94	2.828e-5	0.96	2.185e-5	0.98	1.958e-5	0.99
2^{16}	1.973e-5	—	1.451e-5	—	1.108e-5	—	9.881e-6	—

Table 4.9 Expected "convergence rates" for $(\ln N)^{p/(\kappa+1)}\, N^{-1}$

	$\kappa = 1/2$		$\kappa = 1$		$\kappa = 2$		$\kappa = 3$	
N	$p = 2$	$p = 1$	$p = 2$	$p = 1$	$p = 2$	$p = 1$	$p = 2$	$p = 1$
2^7	0.74	0.87	0.81	0.90	0.87	0.94	0.90	0.95
2^8	0.77	0.89	0.83	0.92	0.89	0.94	0.92	0.96
2^9	0.80	0.90	0.85	0.92	0.90	0.95	0.92	0.96
2^{10}	0.82	0.91	0.86	0.93	0.91	0.95	0.93	0.97
2^{11}	0.83	0.92	0.87	0.94	0.92	0.96	0.94	0.97
2^{12}	0.85	0.92	0.88	0.94	0.92	0.96	0.94	0.97
2^{13}	0.86	0.93	0.89	0.95	0.93	0.96	0.95	0.97
2^{14}	0.87	0.93	0.90	0.95	0.93	0.97	0.95	0.98
2^{15}	0.88	0.94	0.91	0.95	0.94	0.97	0.95	0.98

is $(\ln N)^{p/(\kappa+1)}\, N^{-1}$ for $p = 2$ and $p = 1$. The rates observed are closer to those expected for $p = 1$. This supports the hypothesis of Remark 4.68.