Chapter 6
Discretisations of Reaction-Convection-Diffusion Problems

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

$$-\varepsilon_d u'' - \varepsilon_c b u + cu = f \quad \text{in } (0,1), \ u(0) = \gamma_0, \ u(1) = \gamma_1,$$

with $b \geq 1$ and $c \geq 1$ on $[0,1]$.

In particular, we shall study the special case of scalar reaction-diffusion problems

$$-\varepsilon^2 u'' + cu = f \quad \text{in } (0,1), \ u(0) = \gamma_0, \ u(1) = \gamma_1,$$

and its vector-valued counterpart

$$-\boldsymbol{E}^2 \boldsymbol{u''} + \boldsymbol{A}\boldsymbol{u} = \boldsymbol{f} \quad \text{in } (0,1), \quad \boldsymbol{u}(0) = \boldsymbol{\gamma}_0, \ \boldsymbol{u}(1) = \boldsymbol{\gamma}_1.$$

6.1 Reaction-Diffusion

This section is concerned with scalar reaction-diffusion problems

$$\mathcal{L}u := -\varepsilon^2 u'' + cu = f \quad \text{in } (0,1), \ u(0) = \gamma_0, \ u(1) = \gamma_1, \tag{6.1}$$

where $c \geq \rho^2$ on $[0,1]$ and $\rho > 0$ is a constant.

Analytical properties of (6.1) were studied in Sect. 3.3, while layer-adapted meshes for it have been introduced in Sect. 2.2. The crucial quantity for reaction-diffusion problems is

$$\vartheta_{rd}^{[p]}(\bar{\omega}) := \max_{k=1,\dots,N} \int_{I_k} \left\{ 1 + \varepsilon^{-1} \left(\mathrm{e}^{-\rho x/p\varepsilon} + \mathrm{e}^{-\rho(1-x)/p\varepsilon} \right) \right\} \mathrm{d}x.$$

Using (6.1) as a model problem, a convergence analysis is conducted for variants of the linear FEM with convergence established in the energy norm.

Next a general convergence theory in the maximum norm is derived for central differencing on arbitrary meshes. The close relationship between the differential

T. Linß, *Layer-Adapted Meshes for Reaction-Convection-Diffusion Problems*, Lecture Notes in Mathematics 1985, DOI 10.1007/978-3-642-05134-0_6,

operator and its discretisation is highlighted. We then move on to a maximum-norm error analysis for the linear FEM and a special 4th-order scheme. Finally, central differencing for systems of reaction-diffusion equations is studied.

6.1.1 Linear Finite Elements

Consider (6.1) with homogeneous boundary conditions. Its weak formulation is: Find $u \in H_0^1(0,1)$ such that

$$a(u,v) := \varepsilon^2 \left(u', v'\right) + r\left(u, v\right) = f(v) \quad \text{for all } v \in H_0^1(0,1) \tag{6.2}$$

with $r(u,v) := (cu,v)$ and $f(v) := (f,v) := \int_0^1 \big(fv\big)(s)\mathrm{d}s$.

Given a mesh $\bar{\omega}$, let V_0^ω be the space of continuous functions that are piecewise linear on the mesh $\bar{\omega}$. Clearly $V_0^\omega \subset H_0^1(0,1)$. The standard Galerkin-FEM approximation is: Find $u^N \in V_0^\omega$ such that

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

Typically, the integrals in the bilinear form $r(\cdot,\cdot)$ and in the linear functional $f(\cdot)$ cannot be evaluated exactly. Therefore, they have to be approximated:

$$r(w,v) \approx \hat{r}(w,v) \quad \text{and} \quad f(v) \approx \hat{f}(v).$$

Different approximations yield different variants of the FEM. Later we shall consider four possible choices for $r(\cdot,\cdot)$ and $f(\cdot)$.

With this notation our FEM is: Find $u^N \in V_0^\omega$ such that

$$\hat{a}(u^N, v) := \varepsilon^2 \left((u^N)', v'\right) + \hat{r}\left(u^N, v\right) = \hat{f}(v) \quad \text{for all } v \in V_0^\omega. \tag{6.3}$$

The norm naturally associated with the weak formulation is the energy norm

$$|||v|||_{\varepsilon^2} := \left\{\varepsilon^2 \left\|v'\right\|_0^2 + \rho^2 \left\|v\right\|_0^2\right\}^{1/2}.$$

It is typically used in the convergence analysis of FEMs. Clearly $a(\cdot,\cdot)$ is coercive in the energy norm:

$$|||v|||_{\varepsilon^2}^2 \le a(v,v) \quad \text{for all } v \in H_0^1(0,1).$$

However, the coercivity of $\hat{a}(\cdot,\cdot)$ depends on the approximation used for the reaction term and has to be investigated separately. It is one ingredient in the error analysis for (6.3). The other ingredient is bounds for the interpolation error.

6.1.1.1 The Interpolation Error

Let again w^I denote the piecewise linear interpolant of w. Throughout this section let us assume the function $\psi \in C^2[0,1]$ admits the derivative bounds

$$|\psi''(x)| \le C\left\{1+\varepsilon^{-2}\left(\mathrm{e}^{-\rho x/\varepsilon}+\mathrm{e}^{-\rho(1-x)/\varepsilon}\right)\right\}. \tag{6.4}$$

For example, the solution u of the boundary-value problem (6.1) belongs to this class of functions, see Sect. 3.3.1.2. And so does $cu - f = \varepsilon^2 u$ whose interpolation error will appear in the later analysis too.

Proposition 6.1. *Suppose ψ satisfies (6.4). Then*

$$\left\|\psi-\psi^I\right\|_{\infty,I_i} \le C\left[\int_{I_i}\left\{1+\varepsilon^{-1}\left(\mathrm{e}^{-\rho x/2\varepsilon}+\mathrm{e}^{-\rho(1-x)/2\varepsilon}\right)\right\}\mathrm{d}x\right]^2$$

for all mesh intervals $I_i = [x_{i-1}, x_i]$.

Proof. For $x \in I_i$ and an arbitrary integrable functions χ set

$$\left(\mathcal{J}_i\chi\right)(x) := \frac{1}{h_i}\int_{I_i}\int_{x_{i-1}}^{x}\int_{\xi}^{s}\chi(t)\mathrm{d}t\mathrm{d}\xi\mathrm{d}s.$$

For triple integrals of this structure we have the two bounds

$$\left|\left(\mathcal{J}_i\chi\right)(x)\right| \le \int_{I_i}\left(\xi-x_{i-1}\right)|\chi(\xi)|\,\mathrm{d}\xi \tag{6.5a}$$

and

$$\left|\left(\mathcal{J}_i\chi\right)(x)\right| \le \int_{I_i}\left(x_i-s\right)|\chi(s)|\,\mathrm{d}s. \tag{6.5b}$$

These integrals can be further bounded using Lemma 4.16. Let $\chi : I_j \to I\!\!R$ be any function with $\chi \ge 0$ on I_i. Then

$$\left(\mathcal{J}_i\chi\right)(x) \le \frac{1}{2}\left\{\int_{I_i}\chi(t)^{1/2}\mathrm{d}t\right\}^2 \quad \text{if } \chi \text{ is decreasing,} \tag{6.6a}$$

and

$$\left(\mathcal{J}_i\chi\right)(x) \le \frac{1}{2}\left\{\int_{I_i}\chi(t)^{1/2}\mathrm{d}t\right\}^2 \quad \text{if } \chi \text{ is increasing.} \tag{6.6b}$$

For the interpolation error on I_i we have the representation

$$\left(\psi - \psi^I\right)(x) = \left(\mathcal{J}_i(\psi'')\right)(x) \quad \text{for } x \in I_i. \tag{6.7}$$

Next we would like to apply (6.5) and (6.6). Therefore, we split ψ'' into two parts that can be bounded by monotone functions—one decreasing and the other increasing. Set

$$\bar{\psi}_I := \psi'' - \bar{\psi}_D \quad \text{and} \quad \bar{\psi}_D(x) := \begin{cases} \psi''(x) & \text{for } x \le 1/2, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly,

$$\left|\bar{\psi}_D(x)\right| \le C\left\{1 + \varepsilon^{-2}\mathrm{e}^{-\rho x/\varepsilon}\right\} \quad \text{and} \quad \left|\bar{\psi}_I(x)\right| \le C\left\{1 + \varepsilon^{-2}\mathrm{e}^{-\rho(1-x)/\varepsilon}\right\},$$

by (6.4). Hence, using (6.5) and (6.6), we obtain

$$\|\mathcal{J}_i(\psi'')\|_{\infty,I_i} \le C\int_{I_i}\left(1 + \varepsilon^{-1}\mathrm{e}^{-\rho x/2\varepsilon} + \varepsilon^{-1}\mathrm{e}^{-\rho(1-x)/2\varepsilon}\right)\mathrm{d}x.$$

Recalling (6.7), we are finished. □

Theorem 6.2. *Suppose ψ satisfies (6.4). Then*

$$\left\|\psi - \psi^I\right\|_0 \le \left\|\psi - \psi^I\right\|_\infty \le C\left(\vartheta_{rd}^{[2]}(\bar{\omega})\right)^2$$

and

$$\left|\left|\left|\psi - \psi^I\right|\right|\right|_{\varepsilon^2} \le C\left(\varepsilon^{1/2} + \vartheta_{rd}^{[2]}(\bar{\omega})\right)\vartheta_{rd}^{[2]}(\bar{\omega}).$$

Proof. The bound on the L_∞ error is an immediate consequence of Prop. 6.1 and the definition of $\vartheta_{rd}^{[2]}$.

For the error in the H^1 norm we proceed as follows using integration by parts

$$\left\|\left(\psi - \psi^I\right)'\right\|_0^2 = \int_0^1\left(\left(\psi - \psi^I\right)'(x)\right)^2\mathrm{d}x = -\int_0^1\psi''(x)\left(\psi - \psi^I\right)(x)\mathrm{d}x.$$

Next, a Hölder inequality gives

$$\left\|\left(\psi - \psi^I\right)'\right\|_0^2 \le \left\|\psi - \psi^I\right\|_\infty\int_0^1\left|\psi''(x)\right|\mathrm{d}x \le C\varepsilon^{-1}\left(\vartheta_{rd}^{[2]}(\bar{\omega})\right)^2,$$

by (6.4). Finally, combine this with the bound for the L_2 norm of the interpolation error to obtain the energy-norm estimate. □

Remark 6.3. The interpolation error in the energy norm is an order of magnitude better than for convection-diffusion problems. This is because this norm fails to capture the wider boundary layers in reaction-diffusion problems:

$$\left|\left|\left|\mathrm{e}^{-\rho x/\varepsilon}\right|\right|\right|_{\varepsilon^2} = \mathcal{O}\left(\varepsilon^{1/2}\right) \quad \text{for } \varepsilon \to 0.$$

Therefore, the reaction-diffusion problem is not singularly perturbed in the energy norm in the sense of Def. 1.1. ♣

Remark 6.4. For Lagrange interpolation with polynomials of arbitrary degree $p \geq 0$, cf. Remark 5.5, we have

$$\|I_p u - u\|_\infty \leq C\left(\vartheta_{rd}^{[p+1]}(\bar{\omega})\right)^{p+1}$$

for the solution u of (6.1). ♣

6.1.1.2 Convergence in the Energy Norm

With interpolation error bounds at hand, we can now return to the convergence analysis for the FEM (6.3).

Assume the bilinear form $\hat{a}(\cdot,\cdot)$ is V_0^ω-coercive with respect to the energy norm. That is, there exists a positive constant γ such that

$$\gamma\, |||v|||_{\varepsilon^2}^2 \leq \hat{a}(v,v) \quad \text{for all } v \in V_0^\omega. \tag{6.8}$$

Set $\eta := u^I - u$ and $\chi := u^I - u^N$. Then by a triangle inequality

$$\left|\left|\left|u - u^N\right|\right|\right|_{\varepsilon^2} \leq |||\eta|||_{\varepsilon^2} + |||\chi|||_{\varepsilon^2}.$$

Theorem 6.2 yields

$$|||\eta|||_{\varepsilon^2} \leq C\left(\varepsilon^{1/2} + \vartheta_{rd}^{[2]}(\bar{\omega})\right)\vartheta_{rd}^{[2]}(\bar{\omega}),$$

and we are left with bounding $|||\chi|||_{\varepsilon^2}$.

Starting from (6.8), we get

$$\gamma\, |||\chi|||_{\varepsilon^2}^2 \leq \hat{a}(\chi,\chi) = \hat{r}\left(u^I,\chi\right) - r(u,\chi) + f(\chi) - \hat{f}(\chi), \tag{6.9}$$

where we have used (6.2) and (6.3).

We shall consider four variants of the FEM characterised by different approximations of the reaction term and of the source term:

FEM-0: $\hat{r}(w,v) = r(w,v)$ and $\hat{f}(v) = f(v)$, i.e., no quadrature is used.
FEM-1: $\hat{r}(w,v) = (c^I w, v)$ and $\hat{f}(v) = (f^I, v)$,

FEM-2: $\hat{r}(w,v) = ((cw)^I, v)$ and $\hat{f}(v) = (f^I, v)$,
FEM-3: $\hat{r}(w,v) = (cw,v)_\omega$ and $\hat{f}(v) = (f,v)_\omega$ with the discrete ℓ_2-product

$$(w,v)_\omega := \sum_{i=1}^{N-1} \hbar_i w_i v_i.$$

This method is generated by applying the trapezium rule

$$\int_{I_i} g(s)\mathrm{d}s \approx \frac{h_{i+1}}{2}\left(g_i + g_{i+1}\right) = \int_{I_i} g^I(s)\mathrm{d}s$$

to (cw,v) and (f,v). It is equivalent to the standard central difference scheme which will be subject of Sect. 6.1.2.

Remark 6.5. A direct calculation shows that $(w,v)_\omega = (w,v)$ for all $w,v \in V_0^\omega$. ♣

Theorem 6.6. *Let u be the solution of (6.1) and u^N its approximation by FEM-0, FEM-1 or FEM-2. Then*

$$\left|\left|\left|u^I - u^N\right|\right|\right|_{\varepsilon^2} \le C\left(\vartheta_{rd}^{[2]}(\bar{\omega})\right)^2, \tag{6.10}$$

and

$$\left|\left|\left|u - u^N\right|\right|\right|_{\varepsilon^2} \le C\left(\varepsilon^{1/2} + \vartheta_{rd}^{[2]}(\bar{\omega})\right)\vartheta_{rd}^{[2]}(\bar{\omega}).$$

Corollary 6.7. *For FEM-0, FEM-1 and FEM-2 we have the uniform first-order convergence result*

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

However, the worst case is not when ε is small, but when $\varepsilon = 1$. This is observed in numerical experiments [74].

We give short proofs of Theorem 6.6 for the various FEMs now.

FEM-0: $\hat{r}(w,v) = r(w,v),\ \ \hat{f}(v) = f(v)$.

Clearly $\hat{a}(\cdot,\cdot) = a(\cdot,\cdot)$. Therefore, (6.8) holds with $\gamma = 1$. Inequality (6.9) yields

$$|||\chi|||_{\varepsilon^2}^2 \le r\left(u^I - u, \chi\right) = \left(c(u^I - u), \chi\right)$$

and

$$|||\chi|||_{\varepsilon^2} \le \|c\|_\infty \left\|u^I - u\right\|_0 \le C\left(\vartheta_{rd}^{[2]}(\bar{\omega})\right)^2,$$

by Theorem 6.2.

FEM-1: $\hat{r}(w,v) = (c^I w, v)$, $\hat{f}(v) = (f^I, v)$.

The coercivity of $\hat{a}(\cdot,\cdot)$ is verified upon noting that

$$\left(c^I w, w\right) \geq \left(\rho^2 w, w\right) = \rho^2 \left\|w\right\|_0^2.$$

Thus, (6.8) is satisfied with $\gamma = 1$. Starting from (6.9) again, we get

$$|||\chi|||_{\varepsilon^2}^2 \leq \left((c^I - c)u^I, \chi\right) + \left(c(u^I - u), \chi\right) + \left(f - f^I, \chi\right).$$

Appealing to Theorem 6.2 again, we get (6.10) for FEM-1.

FEM-2: $\hat{r}(w,v) = ((cw)^I, v)$, $\hat{f}(v) = (f^I, v)$.

This time establishing coercivity is slightly more involved. Let $w \in V_0^\omega$ be arbitrary. A direct calculation gives

$$\hat{r}(w,w) = \sum_{i=0}^{N-1} \frac{h_{i+1}}{3}\left(c_i w_i^2 + c_{i+1} w_{i+1}^2\right) + \sum_{i=0}^{N-1} \frac{h_{i+1}}{6}\left(c_i + c_{i+1}\right) w_i w_{i+1}.$$

We bound the second term from below:

$$\begin{aligned}
\left(c_i + c_{i+1}\right) w_i w_{i+1} &\geq -\frac{c_i + c_{i+1}}{2}\left(w_i^2 + w_{i+1}^2\right) \\
&\geq -\left(c_i + \frac{h_{i+1}}{2}\|c'\|_\infty\right) w_i^2 - \left(c_{i+1} + \frac{h_{i+1}}{2}\|c'\|_\infty\right) w_{i+1}^2.
\end{aligned}$$

Thus, if the maximum step size h is sufficiently small, dependent on κ, but independent of ε, then

$$\begin{aligned}
\hat{r}(w,w) &\geq \sum_{i=0}^{N-1} \frac{h_{i+1}}{8}\left(c_i w_i^2 + c_{i+1} w_{i+1}^2\right) \\
&\geq \rho^2 \sum_{i=0}^{N-1} \frac{h_{i+1}}{8}\left(w_i^2 + w_{i+1}^2\right) = \frac{\rho^2}{4}\|w\|_0^2.
\end{aligned}$$

Hence, $\hat{a}(\cdot,\cdot)$ is coercive and (6.8) holds true for $\gamma = 1/4$.

Next, (6.9) and the Cauchy-Schwarz inequality yield

$$\frac{1}{4}\,|||\chi|||_{\varepsilon^2} \leq \left\|q^I - q\right\|_0, \quad \text{with } q := cu - f.$$

Theorem 6.2 applies to q and we obtain (6.10) for FEM-2.

FEM-3: $\hat{r}(w, v) = (cw, v)_\omega$, $\hat{f}(v) = (f, v)_\omega$.

In view of Remark 6.5, ineq. (6.8) holds true with $\gamma = 1$. Then by (6.9)

$$\begin{aligned}\|\!|\chi|\!\|_{\varepsilon^2}^2 \le (q, \chi)_\omega - (q, \chi) &= \int_0^1 \left((q\chi)^I - q\chi\right)(x)\mathrm{d}x \\ &= \left(q^I - q, \chi\right) + \sum_{i=0}^{N-1} \frac{h_{i+1}^2}{6}\left(q_{i+1} - q_i\right)\chi_{x;i},\end{aligned}$$

where again $q = cu - f = \varepsilon^2 u''$. The first term on the right-hand has just been bounded when analysing FEM-2. Unfortunately, in view of the last term—in particular the presence of the discrete derivative χ_x—it seems impossible to obtain a convergence result as general as Theorem 6.6. On a S-type mesh, one might reason as in Sect. 5.2.1 by using an inverse inequality on the coarse-mesh region, but rely on the small mesh sizes inside the layer to gain the necessary powers of ε.

6.1.2 *Central Differencing*

Given an arbitrary mesh $\bar{\omega}$ the most frequently considered finite-difference approximation of (6.1) is: Find $u^N \in I\!R^{N+1}$ such that

$$\left[Lu^N\right]_i = f_i \quad \text{for } i = 1, \ldots, N-1, \;\; u_0^N = \gamma_0, \;\; u_N^N = \gamma_1 \tag{6.11}$$

with

$$[Lv]_i := -\varepsilon^2 v_{\bar{x}\hat{x};i} + c_i v_i \quad \text{for } v \in I\!R^{N+1}.$$

The difference operators were introduced in Sect. 4.1. As mentioned before this difference scheme is equivalent to FEM-3 in the preceding section: Find $u^N \in V_0^\omega$ such that

$$a_c\big(u^N, v\big) = f_c\big(v\big) := (f, v)_\omega \quad \text{for all } v \in V_0^\omega, \tag{6.12}$$

where

$$a_c(w, v) := \varepsilon^2 (w', v') + (cw, v)_\omega = \varepsilon^2 [w_x, v_x)_\omega + (cw, v)_\omega,$$

and

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

Taking as test functions v the standard hat-function basis in V_0^ω, we see that (6.11) and (6.12) are equivalent. In particular, using summation by parts (cf. [146]) it is verified that

$$a_c(w,v) = (Lw,v)_\omega = (w,Lv)_\omega \quad \text{for all } w,v \in V_0^\omega.$$

Apparently the operator L is self-adjoint as is its continuous counterpart $\mathcal{L}$.

6.1.2.1 Stability

The matrix associated with the difference operator L is an 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 $e_i = 1$ establishes the inverse monotonicity of L. Thus, L satisfies a comparison principle: For any mesh functions $v, w\mathbb{R}^{N+1}$

$$\left.\begin{aligned} Lv &\le Lw \quad \text{on } \omega, \\ v_0 &\le w_0, \\ v_N &\le w_N \end{aligned}\right\} \implies v \le w \quad \text{on } \bar{\omega}. \tag{6.13}$$

This comparison principle and Lemma 3.17 provide a priori bounds for the solution of (6.11):

$$\left\|u^N\right\|_{\bar{\omega}} \le \max\{|\gamma_0|, |\gamma_1|\} + \|f/c\|_{\infty,\omega} \quad \text{for } i = 0,\dots,N.$$

It also gives the stability inequality

$$\|v\|_{\infty,\bar{\omega}} \le \|Lv/c\|_{\infty,\omega} \quad \text{for all } v \in \mathbb{R}_0^{N+1}. \tag{6.14}$$

Green's function estimates

Using the discrete Green's function $G : \bar{\omega}^2 \to \mathbb{R} : (x_i, \xi_j) \mapsto G_{i,j} = G(x_i, \xi_j)$ associated with L and Dirichlet boundary conditions, any mesh function $v \in \mathbb{R}_0^{N+1}$ can be represented as

$$v_i = a_c(v, G_{i,\cdot}) = (Lv, G_{i,\cdot})_\omega = (v, LG_{i,\cdot})_\omega \quad \text{for } i = 1,\dots,N-1, \tag{6.15}$$

because the operator L is self-adjoint, i.e., $L = L^*$. This also implies $G_{i,j} = G_{j,i}$ for all $i,j = 0,\dots,N$.

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

$$[LG_{i,\cdot}]_j = \delta_{i,j} \quad \text{for } j = 1,\dots,N-1, \;\; G_{i,0} = G_{i,N} = 0, \tag{6.16}$$

where

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

is the discrete equivalent of the Dirac-δ distribution.

Lemma 6.8. *For any fixed* $\mu > 0$

$$\|v\|_\infty^2 \le \mu \left[v_x, v_x\right)_\omega + \left(\mu^{-1} + 1\right) (v, v)_\omega \quad \textit{for all } v \in V_0^\omega .$$

Proof. This is Theorem 2 in chap. V, §4 of [145]. □

Theorem 6.9. *The Green's function G associated with the discrete operator L and Dirichlet boundary conditions satisfies*

$$0 \le G_{i,j} \le \frac{1 + \varepsilon\rho^{-1}}{\varepsilon\rho} \quad \textit{for } 0 \le i, j \le N,$$

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

$$G_{\xi;i,j} \le 0 \quad \textit{for } 0 \le i \le j < N.$$

Proof. The positivity of G follows from the inverse monotonicity of L.

The upper bound on G is proved using an argument from [147]. Let i be arbitrary, but fixed. Set $\Gamma := G_{i,\cdot}$. Then Lemma 6.8 with $\mu = \varepsilon/\rho$ yields

$$\begin{aligned} \|\Gamma\|_{\infty,\omega}^2 &\le \frac{1 + \varepsilon\rho^{-1}}{\rho} \Big(\varepsilon \left[\Gamma_\xi, \Gamma_\xi\right)_\omega + \varepsilon^{-1}\gamma^2 \left(\Gamma, \Gamma\right)_\omega \Big) \\ &\le \frac{1 + \varepsilon\rho^{-1}}{\rho} \Big(\varepsilon \left[\Gamma_x, \Gamma_x\right)_\omega + \varepsilon^{-1} \left(c\Gamma, \Gamma\right)_\omega \Big) \\ &= \frac{1 + \varepsilon\rho^{-1}}{\rho\varepsilon} \Gamma_i , \end{aligned}$$

by (6.15). The upper bound on G follows.

When proving the monotonicity, note that

$$\varepsilon^2 G_{\xi;i,j} = \varepsilon^2 G_{\xi;i,j-1} + \hbar_j c_j G_{i,j} \quad \text{for } 0 < j < i$$

and

$$\varepsilon^2 G_{\xi;i,j} = \varepsilon^2 G_{\xi;i,j+1} - \hbar_{j+1} c_{j+1} G_{i,j+1} \quad \text{for } i < j < N-1.$$

The non-negativity of G implies $G_{\xi;i,0} \ge 0$ and $G_{\xi;i,N-1} \le 0$. Then the piecewise monotonicity of G follows by induction. □

Theorem 6.10. *The Green's function G associated with the discrete operator L satisfies*

$$\|cG_{i,\cdot}\|_{1,\omega} \le 1, \quad \|G_{\xi;i,\cdot}\|_{1,^*\omega} \le \frac{2\left(1+\varepsilon\rho^{-1}\right)^2}{\varepsilon\rho}, \quad \|G_{\xi\xi;i,\cdot}\|_{1,\omega} \le \frac{2}{\varepsilon^2}$$

for all $i = 1, \ldots, N-1$.

Proof. First multiply (6.16) by $\hbar_j$ and sum for $j = 1, \ldots, N-1$.

$$\sum_{j=1}^{N-1} \hbar_j c_j G_{i,j} = 1 + \varepsilon^2 \left(G_{\bar{\xi};i,N} - G_{\xi;i,0}\right) \le 1,$$

because $G_{\xi;i,0} \ge 0$ and $G_{\xi;i,N-1} \le 0$. This is the bound on the c-weighted ℓ_1 norm.

Next

$$\|G_{\xi;i,\cdot}\|_{1,^*\omega} = \sum_{j=0}^{i-1} h_{j+1} G_{\xi;i,j} - \sum_{j=i}^{N-1} h_{j+1} G_{\xi;i,j} = 2G_{i,i} \le \frac{2\left(1+\varepsilon\rho^{-1}\right)^2}{\varepsilon\rho}, \tag{6.17}$$

by Theorem 6.9.

Finally, a triangle inequality, (6.16) and (6.17) give the bound on $G_{\xi\xi}$. □

6.1.2.2 A Priori Error Bounds

The analysis follows [95, 103]. By (6.15) we have for the error $u - u^N$ in the mesh node x_i

$$\left(u - u^N\right)_i = a_c\left(u - u^N, G_{i,\cdot}\right) = a_c\left(u, G_{i,\cdot}\right) - f_c\left(G_{i,\cdot}\right).$$

For simplicity we set $\varGamma := G_i$. We identify the mesh function $\varGamma$ with that function from V_0^ω that coincides with $\varGamma$ at the mesh nodes. Using the weak form of the differential equation, we get

$$\begin{aligned}\left(u - u^N\right)_i &= a_c\left(u, \varGamma\right) - a\left(u, \varGamma\right) + f\left(\varGamma\right) - f_c\left(\varGamma\right)\\ &= \left(cu - f, \varGamma\right)_\omega - \left(cu - f, \varGamma\right).\end{aligned}$$

Note, if $v_0 = v_N = 0$ then

$$(w, v)_\omega = \int_0^1 (wv)^I(x)\,\mathrm{d}x, \tag{6.18}$$

where w^I denotes again the piecewise linear interpolant of w.

Setting $q := cu - f = \varepsilon^2 u''$, we obtain the error representation

$$\left(u - u^N\right)_i = \int_0^1 \left(\left(q\Gamma\right)^I - q\Gamma\right)(x)\mathrm{d}x. \tag{6.19}$$

We are left with bounding the interpolation error for $q\Gamma$. To this end we shall avail of the derivative bounds derived in Sect. 3.3.1.2 and repeat some of the arguments from Sect. 6.1.1.1.

By (6.7), for $x \in [x_j, x_{j+1}]$ we have

$$\left(q\Gamma\right)(x) - \left(q\Gamma\right)^I(x) = 2\Gamma_{x;j}\left(\mathcal{J}_j(q')\right)(x) + \left(\mathcal{J}_j(q''\Gamma)\right)(x). \tag{6.20}$$

Next, we wish to apply (6.5) and (6.6) to the right-hand side of (6.20). Therefore, we split q' into two parts that can be bounded by monotone functions—one decreasing and the other increasing. Set

$$q_I := q' - q_D \quad \text{and} \quad q_D(x) := \begin{cases} q'(x) & \text{for } x \le 1/2, \\ 0 & \text{otherwise.} \end{cases}$$

From Sect. 3.3.1.2 we have

$$\varepsilon^{-1}\left|q_D(x)\right| \le C\left\{1 + \varepsilon^{-2}\mathrm{e}^{-\rho x/\varepsilon}\right\}$$

and

$$\varepsilon^{-1}\left|q_I(x)\right| \le C\left\{1 + \varepsilon^{-2}\mathrm{e}^{-\rho(1-x)/\varepsilon}\right\}.$$

Hence, using (6.5) and (6.6), we obtain

$$\|\mathcal{J}_j(q')\|_{\infty,I_j} \le C\varepsilon\left(\vartheta_{rd}^{[2]}(\bar{\omega})\right)^2. \tag{6.21}$$

The second integral in (6.20) is bounded in a similar manner. Set

$$\bar{q}_I := q'' - \bar{q}_D \quad \text{and} \quad \bar{q}_D(x) := \begin{cases} q''(x) & \text{for } x \le 1/2, \\ 0 & \text{otherwise.} \end{cases}$$

From Sect. 3.3.1.2:

$$\left|\left(\Gamma\bar{q}_D\right)(x)\right| \le C\left(\Gamma_j + \Gamma_{j+1}\right)\left\{1 + \varepsilon^{-2}\mathrm{e}^{-\rho x/\varepsilon}\right\}$$

and

$$\left|\left(\Gamma\bar{q}_I(x)\right)\right| \le C\left(\Gamma_j + \Gamma_{j+1}\right)\left\{1 + \varepsilon^{-2}\mathrm{e}^{-\rho(1-x)/\varepsilon}\right\},$$

since $\varGamma$ is piecewise linear and positive. Thus,

$$\|\mathcal{J}_j(q''\varGamma)\|_{\infty,I_j} \le C\,(\varGamma_j + \varGamma_{j+1})\left(\vartheta_{rd}^{[2]}(\bar\omega)\right)^2. \tag{6.22}$$

Applying (6.21) and (6.22) to (6.20), we obtain

$$\left|\left(\varGamma q - (\varGamma q)^I\right)(x)\right| \le C\left(\vartheta_{rd}^{[2]}(\bar\omega)\right)^2 (\varepsilon\varGamma_{\bar x;j} + \varGamma_{j+1} + \varGamma_j) \quad \text{for } x \in (x_j, x_{j+1}).$$

Integrate over $[0,1]$ to get

$$\left|\int_0^1 \left((\varGamma q) - (\varGamma q)^I\right)(x)\mathrm{d}x\right| \le C\left(\vartheta_{rd}^{[2]}(\bar\omega)\right)^2 \left(\varepsilon\,\|\varGamma_x\|_{1,*\omega} + \|\varGamma\|_{1,\omega}\right).$$

Finally, recall (6.19) and Theorem 6.10. We arrive at the main convergence result of this section.

Theorem 6.11. *Let u be the solution of the reaction-diffusion problem* (6.1) *and u^N its central difference approximation* (6.11)*. Suppose $c, f \in C^2[0,1]$. Then*

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

Corollary 6.12. *Theorems 6.2 and 6.11 yield*

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

by a triangle inequality and because $\left\|u^I - u^N\right\|_\infty = \left\|u^I - u^N\right\|_{\infty,\omega}$.

6.1.2.3 A Posteriori Error Analysis

The first a posteriori analysis for central differencing was conducted by Kopteva; see [66]. We slightly modify her argument. It is based on the bounds for the Green's function $\mathcal{G}$ associated with the continuous operator $\mathcal{L}$; see Theorem 3.31.

Let $x \in (0,1)$ be arbitrary, but fixed. Set $\varGamma = \mathcal{G}(x,\cdot)$. Then

$$\begin{aligned}
\left(u - u^N\right)(x) &= a\left(u - u^N, \varGamma\right) = f(\varGamma) - a\left(u^N, \varGamma\right) \\
&= f(\varGamma) - f_c(\varGamma) + a_c(u^N, \varGamma) - a\left(u^N, \varGamma\right) \\
&= \left(cu^N - f, \varGamma\right)_\omega - \left(cu^N - f, \varGamma\right),
\end{aligned}$$

by (6.12). Setting $\hat q := cu^N - f$, we have

$$\left(u - u^N\right)(x) = \int_0^1 \left((\hat q\varGamma)^I - \hat q\varGamma\right)(s)\mathrm{d}s,$$

where (6.18) was used. Expand the integrand

$$\left(\hat{q}\Gamma\right)^I - \hat{q}\Gamma = \left(\hat{q}^I - \hat{q}\right)\Gamma + \hat{q}^I\left(\Gamma^I - \Gamma\right) + \left(\hat{q}\Gamma\right)^I - \hat{q}^I\Gamma^I$$

in order to obtain the error representation

$$\begin{aligned}\left(u - u^N\right)(x) &= \left(\hat{q}^I - \hat{q}, \Gamma\right) + \left(\hat{q}^I, \Gamma^I - \Gamma\right) \\ &\quad + \int_0^1 \left(\left(\hat{q}\Gamma\right)^I - \hat{q}^I\Gamma^I\right)(s)\mathrm{d}s.\end{aligned} \tag{6.23}$$

The terms on the right-hand side are estimated separately.

A Hölder inequality and Theorem 3.31 give

$$\left|\left(\hat{q} - \hat{q}^I, \Gamma\right)\right| \le \rho^{-2}\left\|\hat{q} - \hat{q}^I\right\|_\infty .$$

For the second term in (6.23), integration by parts yields

$$\int_{I_k} \hat{q}^I(s)\left(\Gamma^I - \Gamma\right)(s)\mathrm{d}s = \int_{I_k} \left(s - x_{k+1}\right)\left(s - x_k\right)Q_k(s)\Gamma''(s)\mathrm{d}s,$$

where

$$Q_k(s) = \frac{\hat{q}_{k+1} + \hat{q}_k}{4} + \frac{\hat{q}_{k+1} - \hat{q}_k}{6h_{k+1}}\left(s - x_{k+1/2}\right), \quad x_{k+1/2} = \frac{x_k + x_{k+1}}{2}.$$

Thus

$$\left|\int_{I_k} \hat{q}^I(x)\left(\Gamma^I - \Gamma\right)(s)\mathrm{d}s\right| \le \frac{h_{k+1}^2}{8}\max\left\{\left|\hat{q}_k\right|, \left|\hat{q}_{k+1}\right|\right\}\int_{I_k}\left|\Gamma''(s)\right|\mathrm{d}s.$$

By Theorem 3.31, we have $\|\Gamma''\|_1 \le 2\varepsilon^{-2}$. Hence,

$$\left|\left(\hat{q}^I, \Gamma^I - \Gamma\right)\right| \le \max_{k=0,\ldots,N-1}\frac{h_{k+1}^2}{4\varepsilon^2}\max\left\{\left|\hat{q}_k\right|, \left|\hat{q}_{k+1}\right|\right\}.$$

For the last term in (6.23) a direct calculation yields

$$\begin{aligned}\int_{I_k}\left(\left(\hat{q}\Gamma\right)^I - \hat{q}^I\Gamma^I\right)(s)\mathrm{d}s &= \frac{h_{k+1}}{6}\left(\hat{q}_{k+1} - \hat{q}_k\right)\left(\Gamma_{k+1} - \Gamma_k\right) \\ &= \frac{h_{k+1}}{6}\left(\hat{q}_{k+1} - \hat{q}_k\right)\int_{I_k}\Gamma'(s)\mathrm{d}s.\end{aligned}$$

Thus

$$\left|\int_0^1 \left((\hat{q}\Gamma)^I - \hat{q}^I \Gamma^I \right)(s)\mathrm{d}s\right| \le \max_{k=0,\dots,N-1} \frac{h_{k+1}}{6\varepsilon\rho} \left|\hat{q}_{k+1} - \hat{q}_k\right|,$$

because $\|\Gamma'\|_1 \le 1/(\varepsilon\rho)$, by Theorem 3.31.

All terms in (6.23) have been bounded. We get

Theorem 6.13. *Let u be the solution of the reaction-diffusion problem* (6.1) *and u^N its central difference approximation* (6.11). *Let $\hat{q} := cu^N - f$. Then*

$$\left\|u - u^N\right\|_\infty \le \eta_1 + \eta_2 + \eta_3,$$

where

$$\eta_1 := \frac{1}{\rho^2}\left\|\hat{q} - \hat{q}^I\right\|_\infty, \quad \eta_2 := \max_{k=0,\dots,N-1} \frac{h_{k+1}^2}{4\varepsilon^2} \max\left\{\left|\hat{q}_k\right|, \left|\hat{q}_{k+1}\right|\right\}$$

and

$$\eta_3 := \max_{k=0,\dots,N-1} \frac{h_{k+1}}{6\varepsilon\rho} \left|\hat{q}_{k+1} - \hat{q}_k\right|.$$

Remark 6.14. The term $\left\|\hat{q} - \hat{q}^I\right\|_\infty$ in the a posteriori bound can be further expanded as follows

$$\hat{q} - \hat{q}^I = f - f^I - \left(c - c^I\right) u^N - c^I u^N + \left(cu^N\right)^I.$$

Thus

$$\begin{aligned}\left\|\hat{q} - \hat{q}^I\right\|_\infty \le \left\|f - f^I\right\|_\infty + \left\|c - c^I\right\|_\infty \left\|u^N\right\|_\infty \\ + \frac{1}{4} \max_{k=0,\dots,N-1} \left|u_{k+1}^N - u_k^N\right| \cdot \left|c_{k+1} - c_k\right|.\end{aligned}$$

The first two terms involve (continuous) norms of the data. These have to be approximated numerically with sufficient accuracy. At least $\mathcal{O}(h^2)$ is required. However, higher order is desirable to ensure all non explicitly computable terms in the error estimator are of higher order and decay rapidly as the mesh is refined. ♣

Remark 6.15. Invoking the difference equation (6.11), we see that η_3 implicitly contains third-order discrete derivatives of u^N.

$$\frac{h_{k+1}}{6\varepsilon\rho} \left|\hat{q}_{k+1} - \hat{q}_k\right| = \frac{\varepsilon h_{k+1}^2}{6\rho} \left|u_{\bar{x}\hat{x}x;k}^N\right|.$$

The a posteriori error bound in [66, Theorem 3.3] is given using these higher-order operators.

The argument in [66] proceeds to show that

$$\left\|u-u^N\right\|_\infty \le C \max_{i=0,\dots,N-1} \left\{ h_{k+1}\left[1+\left|u_{\bar{x}\hat{x};k}\right|^{1/2}+\left|u_{\bar{x}\hat{x};k+1}\right|^{1/2}\right]\right\}^2 .$$

The constant C is independent of ε, but not specified in [66]. Therefore, this latest inequality cannot be used for reliable a posteriori error estimation. Nonetheless, it is useful for steering adaptive mesh generation. ♣

Adaptive mesh generation

Based on Theorem 6.13, the de Boor algorithm in Sect. 4.2.4.2 can be adapted for the problem and method under consideration. One only needs to redefine the Q_i:

$$Q_i = Q_i(u^N, \varDelta, \omega) := \left\{ \kappa_0 + \kappa_1 \eta_{1;i} \kappa_2 \eta_{2;i} + \kappa_3 \eta_{3;i} \right\}^{1/2}$$

with

$$\eta_{1;i} := \left|\hat{q}_{i-1} - 2\hat{q}_{i-1/2} + \hat{q}_i\right|, \quad \eta_{2;i} := \frac{h_i^2}{\varepsilon^2} \max\left\{\left|\hat{q}_{i-1}\right|, \left|\hat{q}_i\right|\right\},$$
$$\eta_{3;i} := \frac{h_i}{\varepsilon}\left|\hat{q}_i - \hat{q}_{i-1}\right|$$

and non-negative weights κ_k. Note that $\max_i \eta_{1;i}$ is a second order approximation of η_1 in Theorem 6.13.

In view of Remark 6.15 one can also use the de Boor algorithm with

$$Q_i = 1 + \left|u_{\bar{x}\hat{x};i-1}\right|^{1/2} + \left|u_{\bar{x}\hat{x};i}\right|^{1/2} .$$

Numerical experiment for this variant of the algorithm are documented in [68].

6.1.2.4 An Alternative Convergence Proof

Traditional finite difference analysis aims at directly exploiting the maximum-norm stability or using barrier function techniques. In higher dimensions they are often the only tool available, because of a lack of stronger stability results.

We now present an error analysis for central differencing on a Bakhvalov mesh that solely uses (6.14). In the layer regions this mesh is not approximately equidistant. Consequently, the truncation error of the difference scheme is apparently only first order at points in the layers, but a more delicate analysis given in [18] shows that the truncation error at every mesh point is in fact $\mathcal{O}(N^{-2})$ uniformly in ε. We use the description of the mesh by a mesh generating functions, see Sect. 2.1.1.

For any $\psi \in C^4[0,1]$, Taylor expansions show that

$$\left|[\mathcal{L}\psi - L\psi]_i\right| \le C\varepsilon^2 \left\|\psi''\right\|_{[x_{i-1},x_{i+1}]} \tag{6.24a}$$

and

$$\left|[\mathcal{L}\psi - L\psi]_i\right| \le C\varepsilon^2 \left|h_i - h_{i+1}\right| \left|\psi_i'''\right| + (h_i + h_{i+1})^2 \left\|\psi^{(4)}\right\|_{[x_{i-1},x_{i+1}]}. \quad (6.24b)$$

When $\sigma\varepsilon \ge \rho q$ the mesh is uniform with mesh size N^{-1}. Furthermore, $\varepsilon^{-1} \le C$. Thus,

$$\|\mathcal{L}u - Lu\|_\omega \le CN^{-2},$$

by Theorem 3.35 and (6.24).

Now consider the case $\sigma\varepsilon < \rho q$. For simplicity we will consider only the layer at $x = 0$ and assume that $x_i = \varphi(t_i) \le 1/2$.

From the construction of φ one must have $\tau < q$. It follows that $1 < \chi'(\tau) < \bar{q}$, where we set $\bar{q} = 1/(1-2q)$. Define the auxiliary points τ_1 and τ_2 in $(0, q)$ by $\chi'(\tau_1) = \bar{q}$ and $\chi'(\tau_2) = 1$. Then $\tau_2 = q - \sigma\varepsilon/\rho < \tau < \tau_1 = q - \sigma\varepsilon(1-2q)/\rho$ because $\chi'' > 0$ on $[0, q)$.

(i) $\varphi'(t) \le \chi'(\tau) \le \tilde{q}$ for $t \in [0, 1]$. Thus,

$$h_i = \int_{t_{i-1}}^{t_i} \varphi'(t)\mathrm{d}t \le \tilde{q}N^{-1} \quad \text{for } i = 1, \dots, N. \quad (6.25)$$

(ii) For $t \le t_i < q$ we have $\varphi'(t) \le \chi'(t) = \sigma\varepsilon/\rho(q-t) \le \sigma\varepsilon/\rho(q-t_i)$. Hence, for $t_i \le q - N^{-1}$,

$$h_i = \int_{t_{i-1}}^{t_i} \varphi'(t)\mathrm{d}t \le N^{-1}\varphi'(t_i) \le \frac{\sigma\varepsilon}{\rho N(q-t_i)} \le \frac{2\sigma\varepsilon}{\rho N(q-t_{i-1})}\,. \quad (6.26)$$

(iii) $h_{i+1} - h_i = x_{i+1} - 2x_i + x_{i-1} = \varphi''(t_i^*)N^{-2}$ for some $t_i^* \in [t_{i-1}, t_{i+1}]$. Now

$$\varphi''(t) \le \chi''(\tau) = \frac{\sigma\varepsilon}{\rho(q-\tau)^2} \quad \text{and} \quad \frac{1}{q-\tau} \le \frac{1}{q-\tau_1} = \frac{\rho\bar{q}}{\sigma\varepsilon},$$

which gives

$$|h_{i+1} - h_i| \le \frac{\rho\bar{q}^2}{\sigma\varepsilon} N^{-2}. \quad (6.27)$$

Furthermore, we have the bound

$$\varphi''(t_i^*) \le \frac{\sigma\varepsilon}{\rho(q-t_{i+1})^2} \le \frac{4\sigma\varepsilon}{\rho(q-t_i)^2} \quad \text{for } t_i \le q - \frac{2}{N}\,,$$

which yields

$$|h_{i+1} - h_i| \le \frac{4\sigma\varepsilon}{N^2\rho(q-t_i)^2} \quad \text{for } t_i \le q - \frac{2}{N}\,. \quad (6.28)$$

(iv)

$$\mathrm{e}^{-\rho x_i/\varepsilon} = \left(\frac{q-t_i}{q}\right)^{\sigma} \quad \text{for } t_i \le \tau \tag{6.29}$$

and

$$\mathrm{e}^{-\rho x_i/\varepsilon} \le \left(\frac{\sigma\varepsilon}{\rho q}\right)^{\sigma} \quad \text{for } t_i \ge \tau_2. \tag{6.30}$$

Henceforth let $\sigma \ge 2$. Using (6.24), Theorem 3.35, (6.25) and (6.30), we get

$$\left|[\mathcal{L}u - Lu]_i\right| \le CN^{-2} \quad \text{for } \tau_2 \le t_{i-1},$$

which is the region outside the layer. For $t_i \le q - 2N^{-1}$ (the layer region), from (6.24) and Theorem 3.35 one arrives at

$$\begin{aligned}\left|[\mathcal{L}u - Lu]_i\right| \le C\left|h_i - h_{i+1}\right|\varepsilon^2 + C\left|h_i - h_{i+1}\right|\varepsilon^{-1}\mathrm{e}^{-\rho x_i/\varepsilon} \\ + C\left(h_i + h_{i+1}\right)^2\varepsilon^2 + C\left(h_i + h_{i+1}\right)^2\varepsilon^{-2}\mathrm{e}^{-\rho x_{i-1}/\varepsilon}.\end{aligned}$$

To bound the first term on the right-hand side use (6.27); for the second term, use (6.28) and (6.29); for the third term, use (6.25); and for the fourth, use (6.26), (6.29) and $q - t_{i-1} \le 3(q-t_i)/2$. This yields

$$\left|[\mathcal{L}u - Lu]_i\right| \le CN^{-2} \quad \text{for } t_i \le q - 2N^{-1}.$$

We are left with the transition region where $\tau_2 > t_{i-1}$ and $t_i > q - 2N^{-1}$. Thus,

$$q - \frac{2}{N} < t_i < \tau_2 + \frac{1}{N} = q - \frac{\sigma\varepsilon}{\rho} + \frac{1}{N} < q + \frac{1}{N}.$$

Notice that the first two inequalities here imply that $\varepsilon < 3\rho/(\sigma N)$. Use (6.24):

$$\left|[\mathcal{L}u - Lu]_i\right| \le C\left(\varepsilon^2 + \mathrm{e}^{-\rho x_{i-1}/\varepsilon}\right) \le CN^{-2},$$

by (6.30) and $\varepsilon \le CN^{-1}$.

Thus, on a Bakhvalov mesh the truncation error in the maximum norm is bounded uniformly by $\mathcal{O}\left(N^{-2}\right)$. Application of the stability inequality (6.14) gives the uniform error bound

$$\left\|u - u^N\right\|_{\infty,\omega} \le CN^{-2}, \quad \text{if } \sigma \ge 2.$$

We have recovered Theorem 6.11 for a Bakhvalov mesh by means of a different kind of analysis.

6.1.2.5 Discontinuous Data

Assume the reaction coefficient c or the source term f possess a discontinuity in a point $d \in (0,1)$. Then (6.1) reads: Find $u \in C^1[0,1] \cap C^2\big((0,1) \setminus \{d\}\big)$ such that

$$-\varepsilon^2 u'' + cu = f \quad \text{in } (0,1) \setminus \{d\}, \;\; u(0) = \gamma_0, \;\; u(1) = \gamma_1.$$

How should the central differencing scheme (6.11) *be generalised to deal with the discontinuous data?*

Assuming $d = x_\kappa \in \omega$, i.e., the discontinuity is in a mesh point, a naive finite difference approach would seek a mesh function u^N with

$$\left[Lu^N\right]_i = f_i \quad \text{for } x_i \in \omega \setminus \{x_\kappa\}, \quad u^N_{x;\kappa} = u^N_{\bar{x};\kappa}, \;\; u^N_0 = \gamma_0, \;\; u^N_N = \gamma_1.$$

This method on a Shishkin mesh was analysed in [37]. Only first order (up to a logarithmic factor) was established. The numerical experiments in [37] show that when ε is moderate at best first order can be achieved.

The continuity of the derivative in $x_\kappa = d$ is discretised by imposing

$$u^N_{x;\kappa} = u^N_{\bar{x};\kappa}.$$

However, the one-sided difference operators are first-order approximation only. This might explain the drop in accuracy.

Instead, we start from the variational formulation (6.12): Find $u^N \in V_0^\omega$ such that

$$\varepsilon^2 \left(\left(u^N\right)', v'\right) - \left(cu^N, v\right)_\omega = (f, v)_\omega \quad \text{for all } v \in V_0^\omega,$$

where

$$(w, v)_\omega := \sum_{i=1}^{N-1} \hbar_i w_i v_i = \int_0^1 (wv)^I(s)\mathrm{d}s.$$

For discontinuous functions the nodal interpolant does not exist. However, since the point d of discontinuity is a mesh point, we can define w^I element wise:

$$w^I(x) = \frac{x_{k+1} - x}{h_{k+1}} w(x_k + 0) + \frac{x - x_k}{h_{k+1}} w(x_{k+1} - 0) \quad \text{for } x \in I_k.$$

Clearly, when $w \in C[0,1]$, we recover the standard linear nodal interpolant.

Evaluating $\int_0^1 (wv)^I(s)\mathrm{d}s$, we get

$$(w,v)_\omega = \sum_{i=1}^{N-1} \hbar_i \tilde{w}_i v_i, \quad \text{with} \quad \tilde{w}_i = \frac{h_i w(x_i - 0) + h_{i+1} w(x_i + 0)}{2\hbar_i}.$$

Note that $\tilde{w}_i = w(x_i)$ for $i \neq \kappa$.

The resulting difference scheme is: Find $u^N \in I\!R^{N+1}$ such that

$$Lu^N := -\varepsilon^2 u^N_{\bar{x}\hat{x}} + \tilde{c}u^N = \tilde{f} \quad \text{on } \omega, \ u_0^N = \gamma_0, \ u_N^N = \gamma_1.$$

This scheme was analysed by Boglaev and Pack [22]. They establish uniform convergence of first order. Roos and Zarin [142] consider the scheme with $\tilde{w}_i = (w(x_i - 0) + w(x_i + 0))/2$, but in the critical point x_κ they have $h_\kappa = h_{\kappa+1}$. In that article uniform second order convergence is proved for Shishkin meshes and for Bakhvalov-Shishkin meshes.

Using the derivative bounds derived in Sect. 3.3.1.3, the analysis of Sect. 6.1.2.2 needs only minor modifications to get the pointwise error bound

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

where $\vartheta^{[p]}_{rd^i}(\bar{\omega})$ has been defined in Sect. 2.2.1.

6.1.3 A Non-Monotone Scheme

In this section we shall present maximum-norm error bounds for a FEM applied to (6.1). We consider FEM-2. It generates the difference scheme

$$\begin{aligned} \left[Lu^N\right]_i &:= -\varepsilon^2 u^N_{\bar{x}\hat{x};i} + \widehat{cu^N}_i = \widehat{f}_i \qquad \text{for } i = 1, \dots, N-1, \\ u_0^N &= \gamma_0, \ u_N^N = \gamma_1, \end{aligned} \tag{6.31}$$

where

$$\widehat{w}_i := \frac{h_i}{\hbar_i}\frac{w_{i-1}}{6} + \frac{2w_i}{3} + \frac{h_{i+1}}{\hbar_i}\frac{w_{i+1}}{6}.$$

We see that the discretisation of the reaction term cu generates positive off-diagonal entries in the system matrix. This results in a scheme that is—unlike the central difference scheme studied before—not inverse monotone. Nonetheless, a maximum-norm error analysis can be conducted. We follow [97].

6.1.3.1 Stability

Although L is not inverse-monotone, it possesses a core that is:

$$[\Lambda v]_i := -\varepsilon^2 v_{\bar{x}\hat{x};i} + \frac{2c_i}{3} v_i.$$

This is the standard central finite difference approximation of $-\varepsilon^2 u'' + \frac{2}{3}cu$ and can be generated by means of the bilinear form

$$\varepsilon^2 (w', v') + \frac{2}{3} \int_0^1 (cwv)^I (s) \mathrm{d}s.$$

By (6.14) we have

$$\|v\|_{\infty,\bar{\omega}} \le \frac{3}{2} \left\| \frac{\Lambda v}{c} \right\|_{\infty,\omega} \quad \text{for all } v \in I\!R_0^{N+1}. \tag{6.32}$$

Theorem 6.16. *Suppose $c \in C^{0,\alpha}[0,1]$. Let $\kappa \in (0,1)$ be arbitrary, but fixed. Then*

$$\|v\|_{\infty,\bar{\omega}} \le \frac{3}{1-\kappa} \left\| \frac{Lv}{c} \right\|_{\infty,\omega} \quad \textit{for all } v \in I\!R_0^{N+1},$$

provided that the maximum step size h is smaller than some threshold value that is independent of ε.

Remark 6.17. Theorem 6.16 means the non-monotone scheme (6.31) is $(\ell_\infty, \ell_\infty)$-stable although the underlying operator L is not inverse monotone and does not satisfy a maximum principle. ♣

Proof. Let $v \in I\!R_0^{N+1}$ be an arbitrary mesh function. Then

$$[\Lambda v]_i = -\frac{h_i}{\hbar_i} \frac{c_{i-1}}{6} v_{i-1} - \frac{h_{i+1}}{\hbar_i} \frac{c_{i+1}}{6} v_{i+1} + [Lv]_i \, .$$

Thus,

$$|[\Lambda v]_i| \le \frac{h_i c_{i-1} + h_{i+1} c_{i+1}}{6\hbar_i} \|v\|_{\infty,\bar{\omega}} + |[Lv]_i| \quad \text{for } i = 1, \dots, N-1,$$

and the stability inequality (6.32) yields

$$\|v\|_{\infty,\bar{\omega}} \le \frac{3}{2} \max_{i=1,\dots,N-1} \frac{h_i c_{i-1} + h_{i+1} c_{i+1}}{6 c_i \hbar_i} \|v\|_{\infty,\bar{\omega}} + \frac{3}{2} \left\| \frac{Lv}{c} \right\|_{\infty,\omega} . \tag{6.33}$$

Since $c \in C^{0,\alpha}[0,1]$, there exists a constant M with

$$|c(x) - c(\xi)| \le M |x - \xi|^\alpha \quad \text{for all } x, \xi \in [0,1]. \tag{6.34}$$

Therefore,

$$\frac{h_i c_{i-1} + h_{i+1} c_{i+1}}{6 c_i \hbar_i} \le \frac{1}{3} + \frac{M h^\alpha}{3\rho^2} \le \frac{1+\kappa}{3},$$

provided h is smaller than some threshold value that is independent of ε. Now, the proposition follows from (6.33). □

For our convergence analysis, we shall also require bounds on the discrete Green's function G_i associated with L and the mesh node x_i. With Theorem 6.10 we have

$$\int_0^1 (cG_i)^I (s)\mathrm{d}s = \sum_{j=1}^{N-1} \hbar_j c_j G_{i,j} \le \frac{3}{2} \quad \text{and} \quad \int_0^1 G_i(s)\mathrm{d}s \le \frac{3}{2\rho^2}\,. \tag{6.35}$$

6.1.3.2 A Priori Error Analysis

Theorem 6.18. *Let u be the solution of (6.1) and u^N that of (6.31). Then*

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

provided that h is smaller than some threshold value that is independent of ε.

Proof. By a triangle inequality

$$\left\|u - u^N\right\|_\infty \le \left\|u - u^I\right\|_\infty + \left\|u - u^N\right\|_{\infty,\bar\omega}\,.$$

The interpolation error was studied in Sect. 6.1.1.1.

Let $\eta = u^N - u$ and $q := f - cu$. Then the Green's-function representation and Eqs. (6.1) and (6.31) yield—after some calculations—

$$\eta_i = \left(q^I - q, G_i\right) - \left(\left(c\eta\right)^I, G_i\right) + \frac{2}{3}\int_0^1 (ceG_i)^I \tag{6.36}$$

where we used $\left((u^I)', G_i'\right) = (u', G_i')$, because G_i' is piecewise constant. The first term on the right-hand side of (6.36) is bounded using a Hölder inequality, Theorem 6.2 and (6.35):

$$\left|\left(q^I - q, G_i\right)\right| \le C \left(\vartheta_{rd}^{[2]}(\bar\omega)\right)^2.$$

For the remaining terms in (6.36) a straight-forward calculation gives

$$\frac{2}{3}\int_0^1 (c\eta G_i)^I - \left(\left(c\eta\right)^I, G_i\right) = -\frac{1}{3}\sum \frac{h_k}{2}\left(c_k \eta_k G_{i,k-1} + c_{k-1}\eta_{k-1} G_{i,k}\right).$$

Let $\kappa \in (0,1)$ be arbitrary, but fixed. Then, using (6.34) and (6.35), we can estimate as follows:

$$\left| \frac{2}{3} \int_0^1 (c\eta G_i)^I - \left((c\eta)^I , G_i \right) \right| \le \left(\frac{1}{2} + \frac{Mh^\alpha}{2\rho^2} \right) \|\eta\|_{\infty,\bar{\omega}} \le \frac{1+\kappa}{2} \|\eta\|_{\infty,\bar{\omega}} .$$

Thus,

$$|\eta_i| \le C \left(\vartheta_{rd}^{[2]}(\bar{\omega}) \right)^2 + \frac{1+\kappa}{2} \|\eta\|_{\infty,\bar{\omega}} .$$

Taking the maximum over $i = 1, \dots, N-1$, we get the general error bound of the theorem. □

Remark 6.19. In contrast to the analysis for central differencing, only bounds for the L_1 norm of the Green's function have been used, but no bounds on its derivative. Also no third-order derivative of u is required. Only the second-order derivative is used when Theorem 6.2 is invoked. ♣

Remark 6.20. The proof is easily adapted to deal with discontinuities in the right-hand side or in the reaction coefficient. ♣

6.1.3.3 A Posteriori Error Analysis

The analysis in [97] is along the lines of the analysis for central differencing in Sect. 6.1.2.3.

Set $\Gamma = \mathcal{G}(x,\cdot)$ and $\hat{q} := cu^N - f$. Then

$$\left(u - u^N \right)(x) = \left(\hat{q}^I - \hat{q}, \Gamma \right) + \left(\hat{q}^I, \Gamma^I - \Gamma \right) . \tag{6.37}$$

Note that compared with (6.23) the term

$$\int_0^1 \left((\hat{q}\Gamma)^I - \hat{q}^I \Gamma^I \right)(s) \mathrm{d}s$$

does not appear in (6.37).

Both terms on the right-hand side of (6.37) have been bounded in Sect. 6.1.2.3.

Theorem 6.21. *Let u be the solution of the reaction-diffusion problem* (6.1) *and u^N its approximation by* (6.31). *Let $\hat{q} := cu^N - f$. Then*

$$\left\| u - u^N \right\|_\infty \le \frac{1}{\rho^2} \left\| \hat{q} - \hat{q}^I \right\|_\infty + \max_{k=0,\dots,N-1} \frac{h_{k+1}^2}{4\varepsilon^2} \max\{ |\hat{q}_k| , |\hat{q}_{k+1}| \} .$$

Remark 6.22. Compared with central differencing, we see the term

$$\max_{k=0,\dots,N-1} \frac{h_{k+1}}{6\varepsilon\rho} \left|\hat{q}_{k+1} - \hat{q}_k\right|$$

does not feature in the a posteriori estimate for the non-monotone scheme. By Remark 6.15 this term corresponds to a discrete third-order derivative of u^N.

Also note that in the analysis, no bounds on the derivative of the Green's function are required. ♣

6.1.4 A Compact Fourth-Order Scheme

In this section we consider a compact finite difference scheme of order four. Given an arbitrary mesh $\bar{\omega}$ we seek a mesh function $u^N \in I\!R^{N+1}$ satisfying

$$\begin{aligned}\left[Lu^N\right]_i &:= c_i^l u_{i-1}^N + c_i^c u_i^N + c_i^r u_{i+1}^N \\ &= q_i^l f_{i-1} + q_i^c f_i + q_i^r f_{i+1} =: [Qf]_i\,, \quad i = 1,\dots,N-1.\end{aligned}$$

The coefficients c and q are determined so that the scheme is exact for polynomials up to degree four. That is $[Lp]_i = [Q(\mathcal{L}p)]_i$ for all $p \in \varPi_4$. This construction yields a difference scheme whose system matrix is not inverse monotone.

A similar approach was used in [27] where in order to ensure inverse monotonicity, the high-order approximation is used only when the local mesh size is small enough to give non-positive off-diagonal entries. In all other mesh points central differencing is used. This hybrid method is shown to be third-order convergent uniformly in ε on a Shishkin mesh; see [27, § 2]. An alternative approach to obtain a higher-order difference approximation while maintaining the M-matrix property can be found in [46]. However, the construction of that scheme requires explicit knowledge of the derivatives of the data (c and f) and subtle modifications in those points where the mesh is non-uniform.

We shall follow the analysis in [98] and see that inverse monotonicity is not a prerequisite for the maximum-norm error analysis of higher-order schemes.

6.1.4.1 Discretisation

The exactness of the scheme for polynomials up to degree four and the normalisation condition $q_{l;i} + q_{c;i} + q_{r;i} = 1$, $i = 1,\dots,N-1$, yield the difference scheme: Find $u^N \in I\!R^{N+1}$ such that

$$\begin{aligned}\left[Lu^N\right]_i := -\varepsilon^2 u_{\bar{x}\hat{x},i}^N + \left[Q(cu^N)\right]_i = [Qf]_i \quad \text{for } i = 1,\dots,N-1, \\ u_0^N = \gamma_0,\ \ u_N^N = \gamma_1,\end{aligned} \tag{6.38}$$

where

$$[Qv]_i := \frac{1-\mu_i^-}{6} v_{i-1} + \frac{4+\mu_i^- + \mu_i^+}{6} v_i + \frac{1-\mu_i^+}{6} v_{i+1}$$

with

$$\mu_i^- := \frac{h_{i+1}^2}{2h_i \hbar_i} \quad \text{and} \quad \mu_i^+ := \frac{h_i^2}{2h_{i+1}\hbar_i} \; .$$

6.1.4.2 Stability

For the stability analysis we consider an arbitrary mesh $\bar{\omega}$ with maximal step size h. Although L is not inverse-monotone, it possesses a core that is:

$$[\Lambda v]_i := -\varepsilon^2 \left[\delta_x^2 v\right]_i + \frac{\alpha_i^-}{6} r_i v_{i-1} + \frac{4+\mu_i^- + \mu_i^+}{6} r_i v_i + \frac{\alpha_i^+}{6} r_i v_{i+1}$$

where

$$\alpha_i^- = \min\left\{0, 1-\mu_i^-\right\}, \quad \alpha_i^+ = \min\left\{0, 1-\mu_i^+\right\}.$$

The matrix associated with Λ is an L_0 matrix with row sums $\beta_i/6$, where

$$\beta_i := 4 + \alpha_i^- + \alpha_i^+ + \mu_i^- + \mu_i^+ .$$

Therefore, it is an M-matrix and we can conclude that

$$\|v\|_{\bar{\omega}} \leq \max_{i=1,\dots,N-1} \left| \frac{6\,[\Lambda v]_i}{c_i \beta_i} \right| \quad \text{for all } v \in I\!R_0^{N+1}, \tag{6.39}$$

by Lemma 3.17.

Theorem 6.23. *Suppose $c \in C^1[0,1]$. Let $\kappa \in (0,1)$ be arbitrary, but fixed. Then*

$$\|v\|_{\bar{\omega}} \leq \frac{3}{2-\kappa} \left\| \frac{Lv}{c} \right\|_{\omega} \quad \textit{for all } \; v \in I\!R_0^{N+1} \; ,$$

provided that h is smaller than some threshold value that depends on κ and c only.

Proof. First, note that

$$\mu_i^- + \mu_i^+ = \frac{h_{i+1}^3 + h_i^3}{h_i h_{i+1}(h_i + h_{i+1})} \geq 1, \tag{6.40}$$

by the relation between cubic, arithmetic and geometric means.

Furthermore,

$$h_{i+1} \geq h_i \implies \alpha_i^+ = 0, \qquad \text{and} \qquad h_{i+1} \leq h_i \implies \alpha_i^- = 0.$$

Therefore, at least one of α_i^- and α_i^+ is zero. Without loss of generality we assume $\alpha_i^- = 0$. This implies $0 \leq 1 - \mu_i^- \leq 1$.

We distinguish two cases $\alpha_i^+ = 0$ and $\alpha_i^+ = 1 - \mu_i^+$.

(i) Suppose $0 = \alpha_i^+$. Then $0 \leq 1 - \mu_i^+ \leq 1$ and

$$\begin{aligned}[\Lambda v]_i = [Lv]_i &- \frac{1-\mu_i^-}{6} c_i v_{i-1} - \frac{1-\mu_i^+}{6} c_i v_{i+1} \\ &- \frac{1-\mu_i^-}{6} \left(c_{i-1} - c_i\right) v_{i-1} - \frac{1-\mu_i^+}{6} \left(c_{i+1} - c_i\right) v_{i+1}.\end{aligned}$$

We estimate as follows:

$$\left|\left(1-\mu_i^-\right)\left(c_{i-1} - c_i\right)\right| \leq h_i \|c'\|_\infty \leq \kappa c_i \ .$$

Similarly,

$$\left|\left(1-\mu_i^+\right)\left(c_{i+1} - c_i\right)\right| \leq h_{i+1} \|c'\|_\infty \leq \kappa c_i \ ,$$

provided h is sufficiently small—depending on κ and c only. We also have

$$\left|1-\mu_i^-\right| + \left|1-\mu_i^+\right| = 2 - \mu_i^- - \mu_i^+ \leq 1, \quad \text{by (6.40)}.$$

Therefore,

$$\left|[\Lambda v]_i\right| \leq \left|[Lv]_i\right| + \frac{1+2\kappa}{6} \|v\|_{\bar\omega} \,. \tag{6.41}$$

Note that $\beta_i = 4 + \mu_i^- + \mu_i^+ \geq 5$, by (6.40). Hence,

$$\left|\frac{6\,[\Lambda v]_i}{c_i \beta_i}\right| \leq \frac{6}{5} \left|\frac{[L_\varepsilon v]_i}{c_i}\right| + \frac{1+2\kappa}{5} \|v\|_{\bar\omega} \,. \tag{6.42}$$

(ii) If $\alpha_i^+ = 1 - \mu_i^+ \leq 0$ then

$$\begin{aligned}[\Lambda v]_i = [Lv]_i - \frac{1-\mu_i^-}{6} c_i v_{i-1} - \frac{1-\mu_i^-}{6} \left(c_{i-1} - c_i\right) v_{i-1} \\ - \frac{1-\mu_i^+}{6} \left(c_{i+1} - c_i\right) v_{i+1} \ .\end{aligned}$$

The second and third term on the right-hand side are bounded as in case (i). For the last term note that $1-\mu_i^+ \le 0$ yields $\left|1-\mu_i^+\right| \le \mu_i^+$ and therefore,

$$\left|\left(1-\mu_i^+\right)(c_{i+1}-c_i)\right| \le \mu_i^+ h_{i+1}\|c'\|_\infty \le h_i\|c'\|_\infty \le \kappa c_i,$$

for sufficiently small h. Thus, (6.41) holds for this case too. Furthermore, for β_i we have $\beta_i = 5+\mu_i^- \ge 5$. Consequently, (6.42) holds for all $i=1,\dots,N-1$.

Combining (6.39) with (6.42), we are finished. □

Remark 6.24. The discretisation (6.38) is $(\ell_\infty,\ell_\infty)$-stable although the underlying operator is in general not inverse monotone and therefore does not satisfy a maximum principle. ♣

Remark 6.25. The argument presented here sharpens Theorem 1 in [98] by giving a smaller stability constant. ♣

6.1.4.3 Nodal Error Analysis

We shall apply the difference scheme (6.38) on a Shishkin mesh with $\sigma \ge 4$. For this we have $h \le N^{-1}/(1-2q)$.

Let $\eta = u - u^N$ denote the error of the scheme. We start our analysis by decomposing the error of the scheme as $\eta = \psi + \varphi$, where the two parts $\psi, \varphi \in I\!R_0^{N+1}$ solve

$$[\varLambda\psi]_i = [L\eta]_i = \varepsilon^2\left[Q(u'') - u_{\bar{x}\hat{x}}\right]_i \quad \text{on } \omega,$$

and

$$\begin{aligned}[\varLambda\varphi]_i = &-\frac{c_{i-1}}{6}\eta_{i-1} - \frac{c_{i+1}}{6}\eta_{i+1} \\ &+ \frac{\mu_i^-}{6}(c_{i-1}-c_i)\,\eta_{i-1} + \frac{\mu_i^+}{6}(c_{i+1}-c_i)\,\eta_{i+1} \quad \text{on } \omega.\end{aligned}$$

Let $\kappa \in (0,1)$ be arbitrary, but fixed. Then using arguments from our stability analysis, we get

$$\left|\frac{6\,[\varLambda\varphi]_i}{\beta_i c_i}\right| \le \frac{1+2\kappa}{5}\,\|\eta\|_{\bar\omega} \quad \text{for } i=1,\dots,N-1,$$

if N is greater than some threshold value that is independent of ε. The stability inequality (6.39) yields

$$\|\varphi\|_\omega \le \frac{1+2\kappa}{5}\,\|\eta\|_{\bar\omega}\,.$$

Application of a triangle inequality gives

$$\|\eta\|_{\bar\omega} \le \|\psi\|_{\bar\omega} + \|\varphi\|_{\bar\omega} \le \|\psi\|_{\bar\omega} + \frac{1+2\kappa}{5}\|\eta\|_{\bar\omega}.$$

Hence, the error can be bounded in terms of ψ:

$$\left\|u - u^N\right\|_{\bar\omega} \le \frac{5}{4-2\kappa}\|\psi\|_{\bar\omega}. \tag{6.43}$$

We are left with estimating ψ which will be done using a truncation error and barrier function technique.

Let $g \in C^6[x_{i-1}, x_{i+1}]$. Then Taylor expansions give

$$\left|[Q(g'') - g_{\bar x\hat x}]_i\right| \le \begin{cases} C\,\|g''\|_{[x_{i-1},x_{i+1}]} & \text{if } h_i = h_{i+1}, \\ C\,\|g''\|_{[x_i,x_{i+1}]} + C\mu_i^- h_i\,\|g'''\|_{[x_{i-1},x_i]} & \text{if } h_i \le h_{i+1}, \\ C\,(h_i + h_{i+1})^3\left\|g^{(5)}\right\|_{[x_{i-1},x_{i+1}]} & \text{and} \\ Ch_i^4\left\|g^{(6)}\right\|_{[x_{i-1},x_{i+1}]} & \text{if } h_i = h_{i+1}. \end{cases} \tag{6.44}$$

We consider the two distinct cases for the mesh transition point: $\tau = q$ and $\tau < q$. In the first case, the mesh is uniform with $h_i = 1/N$ for $i = 1, \dots, N$. Moreover, $\varepsilon^{-1} \le C \ln N$. Thus, $\left\|u^{(6)}\right\| \le C\varepsilon^{-2}\ln^4 N$, by Theorem 3.35. Now the fourth bound of (6.44) yields

$$\left|[\varLambda\psi]_i\right| = \varepsilon^2\left|[Q(u'') - u_{\bar x\hat x}]_i\right| \le CN^{-4}\ln^4 N, \quad i = 1, \dots, N-1.$$

Hence,

$$\|\psi\|_{\bar\omega} \le CN^{-4}\ln^4 N \quad \text{if } \tau = q, \tag{6.45}$$

by (6.39).

The case when $\tau < q$ requires a more detailed argument employing the decomposition of u.

(i) For $x_i \in (0,\tau) \cup (1-\tau, 1)$, use the fourth bound of (6.44), $\left\|u^{(6)}\right\| \le C\varepsilon^{-6}$ and $h_i = h_{i+1} \le C\varepsilon N^{-1}\ln N$ to get

$$\left|[\varLambda\psi]_i\right| = \varepsilon^2\left|[Q(u'') - u_{\bar x\hat x}]_i\right| \le CN^{-4}\ln^4 N \quad \text{for } x_i \in (0,\tau)\cup(1-\tau,1).$$

(ii) For $x_i \in (\tau, 1-\tau)$, split the truncation error according to the decomposition of u:

$$\varepsilon^2\left|[Q(u'') - u_{\bar x\hat x}]_i\right| = \varepsilon^2\left|[Q(v'') - v_{\bar x\hat x}]_i\right| + \varepsilon^2\left|[Q(w'') - w_{\bar x\hat x}]_i\right|.$$

The first term is bounded using the fourth estimate of (6.44), $\varepsilon^2 \left\|v^{(6)}\right\| \le C$ and $h_i = h_{i+1} \le CN^{-1}$. We obtain $\varepsilon^2 \left|[Q(v'') - v_{\bar{x}\hat{x}}]_i\right| \le CN^{-4}$. In order to bound the truncation error with respect to the layer part w, use the first estimate of (6.44) to get $\varepsilon^2 \left|[Q(w'') - w_{\bar{x}\hat{x}}]_i\right| \le CN^{-\sigma}$. Hence,

$$\left|[\Lambda\psi]_i\right| \le CN^{-4} \quad \text{for } x_i \in (\tau, 1-\tau).$$

(iii) For $x_i \in \{\tau, 1-\tau\}$, split the truncation error again. For the regular solution component v the third estimate of (6.44) gives

$$\varepsilon^2 \left|[Q(v'') - v_{\bar{x}\hat{x}}]_i\right| \le C\varepsilon N^{-3},$$

while for w we have by the second bound of (6.44)

$$\begin{aligned}\varepsilon^2 \left|[Q(w'') - w_{\bar{x}\hat{x}}]_{qN}\right| &\le C\left(\mathrm{e}^{-\rho x_{qN}/\varepsilon} + \mu_{qN}^- N^{-1}\ln N \mathrm{e}^{-\rho(1-x_{(1-q)N+1})/\varepsilon}\right)\\ &\le CN^{-\sigma} + CN^{\sigma/qN} N^{-\sigma-1}\mu_{qN}^- \ln N\\ &\le CN^{-\sigma} + CN^{-\sigma-1}\mu_{qN}^- \ln N,\end{aligned}$$

with an analogous estimate for $i = (1-q)N$. Collecting the various bounds for the truncation error, we get (for $\tau < q$)

$$\begin{aligned}&\varepsilon^2 \left|[Q(u'') - u_{\bar{x}\hat{x}}]_i\right|\\ &\quad \le CN^{-4}\ln^4 N + \begin{cases} C\varepsilon N^{-3} + CN^{-5}\mu_i^- \ln N & \text{if } i = qN,\\ C\varepsilon N^{-3} + CN^{-5}\mu_i^+ \ln N & \text{if } i = (1-q)N,\\ 0 & \text{otherwise.}\end{cases}\end{aligned} \tag{6.46}$$

Finally, we use an idea from [122]. Define the mesh function $z \in I\!R_0^{N+1}$ by

$$z_i := \begin{cases} x_i/\tau & \text{for } i = 0, \dots, qN,\\ 1 & \text{for } i = qN, \dots, (1-q)N, \text{ and}\\ (1-x_i)/\tau & \text{for } i = (1-q)N, \dots, N.\end{cases}$$

A direct calculation verifies

$$[\Lambda z]_i = \frac{2c_i}{3} z_i + \frac{1}{qN}\begin{cases} \frac{\varepsilon^2}{h_i \hbar_i} + \frac{\mu_i^- r_i}{6} & \text{if } i = qN,\\ \frac{\varepsilon^2}{h_{i+1}\hbar_i} + \frac{\mu_i^+ r_i}{6} & \text{if } i = (1-q)N \text{ and}\\ 0 & \text{otherwise.}\end{cases}$$

Note that

$$\frac{\varepsilon^2}{h_{qN}\hbar_{qN}}, \frac{\varepsilon^2}{h_{(1-q)N+1}\hbar_{(1-q)N}} \geq \frac{\varepsilon\rho(1-2q)N}{2\sigma \ln N}.$$

Therefore, recalling the truncation error bound (6.46), we can use the discrete comparison principle (6.13) with the barrier function $CN^{-4}\left(\ln^4 N + z_i \ln N\right)$ to see that (6.45) holds for $\tau < q$ too.

Finally, (6.43) yields the main convergence result of this section.

Theorem 6.26. *Let u be the solution of the boundary value problem (6.1) and u^N that of the finite difference scheme (6.38) on a Shishkin mesh with $\sigma \geq 4$. Then*

$$\left\|u - u^N\right\|_{\infty,\bar{\omega}} \leq CN^{-4}\ln^4 N,$$

provided that N is larger than some threshold value that depends on c only.

Corollary 6.27. *Clustering three adjacent and equidistant mesh intervals and fitting a cubic function through the numerical approximation on the four associated mesh points, we obtain a cubic C^0-spline $\mathcal{S}u^N$ that approximates u on the whole domain:*

$$\left\|u - \mathcal{S}u^N\right\|_{\infty} \leq CN^{-4}\ln^4 N.$$

Remark 6.28. Approximations of the derivatives can be obtained by differentiating $\mathcal{S}u^N$:

$$\varepsilon^k\left\|(u - \mathcal{S}u^N)^{(k)}\right\| \leq C\left(N^{-1}\ln N\right)^{(4-k)} \quad \text{for } k = 1, 2, 3.$$

Note that in those mesh points where two different cubic functions are concatenated to give $\mathcal{S}u^N$, we have different left- and right-sided derivatives, however for both, the above bound holds.

Better approximations of the second-order derivative are obtained by appealing to (6.1):

$$u''(x) \approx m(x) := \varepsilon^{-2}\left(c\mathcal{S}u^N - f\right)(x).$$

We have the error bound

$$\varepsilon^2\left\|u'' - m\right\| \leq CN^{-4}\ln^4 N.$$

This readily follows from Theorem 6.26. ♣

6.1.4.4 Numerical Results

To illustrate the theoretical results, we consider the problem

$$-\varepsilon^2 u'' + (1 + x^2 + \cos x)u = x^{9/2} + \sin x, \quad \text{in } (0,1), \tag{6.47}$$

with homogeneous Dirichlet boundary conditions.

The exact solution is not available. Therefore, we estimate the error for u^N and $\mathcal{S}u^N$ by comparing them to the numerical solution $\tilde{u}^N$ obtained on a mesh $\tilde{\omega}$ that is seven times as fine. More precisely $\tilde{\omega}$ is constructed from ω by uniformly dividing each of its mesh intervals into seven equidistant subintervals. We take the estimates

$$\left\|u - u^N\right\|_{\infty,\omega} \approx \eta^N := \left\|\tilde{u}^N - u^N\right\|_{\infty,\omega}$$

and

$$\left\|u - \mathcal{S}u^N\right\|_{\infty} \approx \tilde{\eta}^N := \left\|\tilde{u}^N - \mathcal{S}u^N\right\|_{\infty,\tilde{\omega}}.$$

Since we have an error bound of the form $C\left(N^{-1}\ln N\right)^p$, we also compute the "Shishkin" rates of convergence:

$$p^N = \frac{\ln \eta^N - \ln \eta^{2N}}{\ln(2\ln N) - \ln(\ln(2N))}.$$

We choose N divisible by three and define $\mathcal{S}u^N$ on macro intervals $[x_{3k}, x_{3(k+1)}]$, $k = 0, \ldots, N/3 - 1$.

The left half of Table 6.1 displays the results of our test computations for the Shishkin mesh. They are in good agreement with Theorem 6.26 and Corollary 6.27. The right half of the table contains results for a modified Bakhvalov mesh: First

Table 6.1 Compact fourth order scheme for (6.47), $\varepsilon = 10^{-4}$ (identical numbers for $\varepsilon = 10^{-4k}$, $k = 2, \ldots, 6$)

	Shishkin mesh				Bakhvalov mesh			
N	η^N	p^N	$\tilde{\eta}^N$	$\tilde{p}^N$	η^N	π^N	$\tilde{\eta}^N$	$\tilde{\pi}^N$
$3 \cdot 2^7$	1.151e-05	3.99	3.979e-04	3.66	2.644e-08	4.01	4.915e-07	3.99
$3 \cdot 2^8$	1.123e-06	4.00	4.701e-05	3.81	1.641e-09	4.00	3.090e-08	4.00
$3 \cdot 2^9$	1.045e-07	4.00	4.890e-06	3.89	1.024e-10	4.00	1.936e-09	4.00
$3 \cdot 2^{10}$	9.375e-09	4.00	4.672e-07	3.94	6.394e-12	4.00	1.212e-10	4.00
$3 \cdot 2^{11}$	8.160e-10	4.00	4.212e-08	3.97	3.996e-13	4.00	7.580e-12	4.00
$3 \cdot 2^{12}$	6.925e-11	4.00	3.644e-09	3.98	2.497e-14	4.00	4.739e-13	4.00
$3 \cdot 2^{13}$	5.750e-12	4.00	3.058e-10	3.99	1.561e-15	4.00	2.963e-14	4.00
$3 \cdot 2^{14}$	4.686e-13	4.00	2.506e-11	4.00	9.756e-17	4.00	1.852e-15	4.00
$3 \cdot 2^{15}$	3.756e-14	4.00	2.014e-12	4.00	6.097e-18	4.00	1.157e-16	4.00
$3 \cdot 2^{16}$	2.967e-15	—	1.594e-13	—	3.811e-19	—	7.234e-18	—

we construct a standard Bakhvalov mesh with $N/3$ mesh intervals, which are then subdivided into three subintervals of equal length. This gives our computational mesh ω. This modification is necessary because the stability constant of $\mathcal{S}$ depends on the local ratio of the mesh sizes, which on a Bakhvalov mesh depends on ε. For this mesh, we expect uniform convergence of order N^{-4}. This is clearly observed in the numerical experiments.

6.2 Systems of Reaction-Diffusion Type

We now leave the scalar problems and move on to systems of reaction-diffusion equations: Find $\boldsymbol{u} \in \left(C^2(0,1) \cap C[0,1]\right)^\ell$ such that

$$\boldsymbol{\mathcal{L}u} := -\boldsymbol{E}^2\boldsymbol{u}'' + \boldsymbol{Au} = \boldsymbol{f} \quad \text{in } (0,1), \quad \boldsymbol{u}(0) = \boldsymbol{u}(1) = \boldsymbol{0}, \tag{6.48}$$

where $\boldsymbol{E} = \operatorname{diag}(\varepsilon_1, \ldots, \varepsilon_\ell)$ and the small parameters ε_k are in $(0,1]$.

The analytical behaviour of the solution to (6.48) has been studied in Sect. 3.3.2. Again we assume the coupling matrix $\boldsymbol{A}$ has positive diagonal entries and is diagonally dominant with

$$\beta := \max_{k=1,\ldots,\ell} \sum_{\substack{m=1\\ m\neq k}}^{\ell} \left\| \frac{a_{km}}{a_{kk}} \right\|_\infty < 1. \tag{6.49}$$

Define $\kappa > 0$ by

$$\kappa^2 := (1-\beta) \min_{k=1,\ldots,\ell} \min_{x\in[0,1]} a_{kk}(x).$$

For simplicity in our presentation we assume that

$$\varepsilon_1 \geq \varepsilon_2 \geq \cdots \geq \varepsilon_\ell \quad \text{and} \quad \varepsilon_1 \leq \frac{\kappa}{4}\,.$$

The first chain of inequalities can always be achieved by renumbering the equations, while the last inequality provides a threshold value for the validity of our analysis.

6.2.1 The Interpolation Error

We consider piecewise linear interpolation. Using the derivative bounds of Theorem 3.43 for the solution of (6.48), we can apply the technique in Section 6.1.1.1 to establish interpolation error bounds for the components of $\boldsymbol{u}$.

Theorem 6.29. *Suppose the assumptions of Theorem 3.43 hold true. Then*

$$\left\|u_k - u_k^I\right\|_0 \le \left\|u_k - u_k^I\right\|_\infty \le C\left(\vartheta_{rd,k}^{[2]}(\bar{\omega})\right)^2, \quad k = 1, \dots, \ell,$$

and

$$\varepsilon_k \left|u_k - u_k^I\right|_1 \le C\varepsilon_k^{1/2}\vartheta_{rd,k}^{[2]}(\bar{\omega}), \quad k = 1, \dots, \ell,$$

where

$$\vartheta_{rd,k}^{[p]}(\bar{\omega}) := \max_{i=0,\dots,N-1} \int_{I_i} \left(1 + \sum_{m=1}^{k} \varepsilon_m^{-1}\left(\mathrm{e}^{-\kappa s/p\varepsilon_m} + \mathrm{e}^{-\kappa(1-s)/p\varepsilon_m}\right)\right) \mathrm{d}s$$

for $k = 1, \dots, \ell$.

6.2.2 Linear Finite Elements

In view of Lemma 3.40 we may assume without loss of generality that $\boldsymbol{A}$ is coercive, i.e., there exists a positive constant α such that

$$\boldsymbol{v}^T \boldsymbol{A}(x)\boldsymbol{v} \ge \alpha \boldsymbol{v}^T \boldsymbol{v} \quad \text{for all } x \in [0,1], \text{ and } \boldsymbol{v} \in I\!R^\ell. \tag{6.50}$$

As usual with finite element discretisations, we consider the weak formulation of (6.48): Find $\boldsymbol{u} \in H_0^1(0,1)^\ell$ such that

$$B(\boldsymbol{u}, \boldsymbol{v}) = F(\boldsymbol{v}) \quad \text{for all } \boldsymbol{v} \in H_0^1(0,1)^\ell,$$

with

$$B(\boldsymbol{u}, \boldsymbol{v}) := \sum_{m=1}^{\ell} \varepsilon_m^2 (u_m', v_m') + \sum_{m=1}^{\ell}\sum_{k=1}^{\ell} (a_{mk}u_k, v_m)$$

and

$$F(\boldsymbol{v}) := \sum_{m=1}^{\ell} (f_m, v_m), \quad (w, v) = \int_0^1 \big(wv\big)(s)\mathrm{d}s.$$

The natural norm on $H_0^1(0,1)^\ell$ associated with the bilinear form $B(\cdot,\cdot)$ is the energy norm $|||\cdot|||_{\boldsymbol{\varepsilon}^2}$ defined by

$$|||\boldsymbol{v}|||_{\boldsymbol{\varepsilon}^2}^2 := \sum_{m=1}^{\ell} \varepsilon_m^2 \left|v_m\right|_1^2 + \alpha \left\|\boldsymbol{v}\right\|_0^2, \quad \left\|\boldsymbol{v}\right\|_0^2 := \sum_{m=1}^{\ell} \left\|v_m\right\|_0^2.$$

Because of (6.50) the bilinear form $B(\cdot,\cdot)$ is coercive with respect to the energy norm, i.e.,

$$|||\boldsymbol{v}|||_{\boldsymbol{\varepsilon}^2}^2 \le B(\boldsymbol{v},\boldsymbol{v}) \quad \text{for all } \ \boldsymbol{v} \in H_0^1(0,1)^\ell .$$

If $\boldsymbol{f} \in L^2(0,1)^\ell$ then F is a linear continuous functional on $H_0^1(0,1)^\ell$. Therefore, the Lax-Milgram Lemma ensures the existence of a unique solution $\boldsymbol{u} \in H_0^1(0,1)^\ell$ of the variational formulation of (6.48).

Given a mesh $\bar{\omega}$, we seek an approximation $\boldsymbol{u}^N \in (V_0^\omega)^\ell$ such that

$$\hat{B}(\boldsymbol{u}^N,\boldsymbol{v}) = \hat{F}(\boldsymbol{v}) \quad \text{for all } \ \boldsymbol{v} \in (V_0^\omega)^\ell , \tag{6.51}$$

where

$$\hat{B}(\boldsymbol{u},\boldsymbol{v}) := \sum_{m=1}^{\ell} \varepsilon_m^2 (u_m', v_m') + \sum_{m=1}^{\ell}\sum_{k=1}^{\ell} (a_{mk}^I u_k, v_m)$$

and

$$\hat{F}(\boldsymbol{v}) := \sum_{m=1}^{\ell} (f_m^I, v_m).$$

Thus, our FEM incorporates quadrature to approximate the integrals.

The coercivity of $\boldsymbol{A}$, i.e. (6.50), implies the coercivity of its piecewise linear interpolant $\boldsymbol{A}^I$:

$$\boldsymbol{v}^T \boldsymbol{A}^I(x)\boldsymbol{v} \ge \alpha \boldsymbol{v}^T\boldsymbol{v} \quad \text{for all } \ x \in [0,1], \ \text{ and } \ \boldsymbol{v} \in I\!R^\ell .$$

Consequently, the bilinear form $\hat{B}(\cdot,\cdot)$ is also coercive with

$$|||\boldsymbol{v}|||_{\boldsymbol{\varepsilon}^2}^2 \le \hat{B}(\boldsymbol{v},\boldsymbol{v}) \quad \text{for all } \ \boldsymbol{v} \in H_0^1(0,1)^\ell . \tag{6.52}$$

Set $\boldsymbol{\eta} := \boldsymbol{u}^I - \boldsymbol{u}$ and $\boldsymbol{\chi} := \boldsymbol{u}^I - \boldsymbol{u}^N$. Then by a triangle inequality

$$\left|\left|\left|\boldsymbol{u} - \boldsymbol{u}^N\right|\right|\right|_{\boldsymbol{\varepsilon}^2} \le |||\boldsymbol{\eta}|||_{\boldsymbol{\varepsilon}^2} + |||\boldsymbol{\chi}|||_{\boldsymbol{\varepsilon}^2} .$$

Theorem 6.29 yields

$$|||\boldsymbol{\eta}|||_{\boldsymbol{\varepsilon}^2} \le C\left(\varepsilon_1^{1/2} + \vartheta_{rd,\ell}^{[2]}(\omega)\right)\vartheta_{rd,\ell}^{[2]}(\omega),$$

and we are left with bounding $|||\boldsymbol{\chi}|||_{\boldsymbol{\varepsilon}^2}$.

Starting from (6.52), we get

$$|||\boldsymbol{\chi}|||_{\boldsymbol{\varepsilon}^2}^2 \le \hat{B}(\chi,\chi) = \sum_{m=1}^{\ell}\left(\sum_{k=1}^{\ell}\left(a_{mk}^I u_k^I - a_{mk}u_k\right) + f_m - f_m^I, \chi_m\right).$$

Use

$$a_{mk}^I u_k^I - a_{mk} u_k = \left(a_{mk}^I - a_{mk}\right) u_k^I + a_{mk} \left(u_k^I - u_k\right),$$

the Cauchy-Schwarz inequality and Theorem 6.29 to get

$$|||\boldsymbol{\chi}|||_{\boldsymbol{\varepsilon}^2}^2 \le C \left(\varepsilon_1^{1/2} + \vartheta_{rd,\ell}^{[2]}(\omega)\right) \vartheta_{rd,\ell}^{[2]}(\omega) \left\|\boldsymbol{\chi}\right\|_0.$$

We obtain the following convergence results in the energy norm.

Theorem 6.30. *Let* $\boldsymbol{u}$ *be the solution of* (6.48) *and* $\boldsymbol{u}^N$ *its approximation by* (6.51)*. Then*

$$\left|\left|\left|\boldsymbol{u}^I - \boldsymbol{u}^N\right|\right|\right|_{\boldsymbol{\varepsilon}^2} \le C \left(\vartheta_{rd,\ell}^{[2]}(\bar{\omega})\right)^2,$$

and

$$\left|\left|\left|\boldsymbol{u} - \boldsymbol{u}^N\right|\right|\right|_{\boldsymbol{\varepsilon}^2} \le C \left(\varepsilon_1^{1/2} + \vartheta_{rd,\ell}^{[2]}(\bar{\omega})\right) \vartheta_{rd,\ell}^{[2]}(\bar{\omega}).$$

Remark 6.31. A similar result is given in [99], but there the effect of numerical integration is not taken into account. ♣

Remark 6.32. Like in the scalar case, the energy norm $|||\cdot|||_{\boldsymbol{\varepsilon}^2}$ fails to capture the layers present in the solution. ♣

6.2.3 Central Differencing

We consider the discretisation of (6.48) by standard central differencing on meshes $\bar{\omega}$ that for the moment are arbitrary. That is, we seek $\boldsymbol{u}^N \in \left(I\!R_0^{N+1}\right)^\ell$ such that

$$\begin{gathered}\left[\boldsymbol{L}\boldsymbol{u}^N\right]_i := -\operatorname{diag}(\boldsymbol{E})^2 \boldsymbol{u}_{\bar{x}\hat{x};i}^N + \boldsymbol{A}_i \boldsymbol{u}_i^N = \boldsymbol{f}_i \quad \text{for } i = 1, \dots, N-1, \\ \boldsymbol{u}_0^N = \boldsymbol{u}_N^N = \boldsymbol{0}.\end{gathered} \tag{6.53}$$

6.2.3.1 Stability

Our analysis follows that of [104] and is based on the stability properties of Section 6.1.2.1 for scalar operators.

Let $\boldsymbol{v} \in \left(I\!R_0^{N+1}\right)^\ell$ be arbitrary. Then

$$-\varepsilon_k^2 v_{k;\bar{x}\hat{x}} + a_{kk} v_k = (\boldsymbol{L}\boldsymbol{v})_k - \sum_{\substack{m=1 \\ m \ne k}}^{\ell} a_{km} v_m \quad \text{for } k = 1, \dots, \ell.$$

The stability inequality (6.14) and a triangle inequality then yield

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

Define the $\ell \times \ell$ constant matrix $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}(\boldsymbol{A}) = (\gamma_{km})$ by

$$\gamma_{kk} = 1 \quad \text{and} \quad \gamma_{km} = -\left\| \frac{a_{km}}{a_{kk}} \right\|_{\infty} \quad \text{for } k \neq m.$$

Suppose that $\boldsymbol{\Gamma}$ is inverse-monotone, i.e., that $\boldsymbol{\Gamma}$ is invertible and $\boldsymbol{\Gamma}^{-1} \geq 0$. Then we obtain the following stability result for the difference operator $\boldsymbol{L}$.

Theorem 6.33. *Assume the matrix $\boldsymbol{A}$ has positive diagonal entries. Assume that $\boldsymbol{\Gamma}(\boldsymbol{A})$ is inverse-monotone. Then for $k = 1, \dots, \ell$ one has*

$$\|v_k\|_{\infty,\bar\omega} \le \sum_{m=1}^{\ell} \left(\boldsymbol{\Gamma}^{-1}\right)_{km} \left\| \frac{(\boldsymbol{L}\boldsymbol{v})_m}{a_{mm}} \right\|_{\infty,\omega}$$

for any function mesh function $\boldsymbol{v} \in \left(I\!R_0^{N+1}\right)^{\ell}$.

Corollary 6.34. *Under the hypotheses of Theorem 6.33 the discrete problem (6.53) 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.*

Thus, the operator $\boldsymbol{L}$ is $(\ell_\infty, \ell_\infty)$ stable, or maximum-norm stable although in general it is not inverse-monotone.

6.2.3.2 A Priori Error Bounds

Let $\boldsymbol{\eta} := \boldsymbol{u} - \boldsymbol{u}^N$ denote the error of the discrete solution. We decompose the solution error as $\boldsymbol{\eta} = \boldsymbol{\varphi} + \boldsymbol{\psi}$, where the components φ_k and ψ_k of $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ respectively are the solutions of

$$-\varepsilon_k^2 \varphi_{k;\bar x\hat x} + a_{kk}\varphi_k = -\varepsilon_k^2 \left(u_{k,\bar x\hat x} - u_k''\right) \quad \text{on } \omega, \quad \varphi_{k;0} = \varphi_{k;N} = 0$$

and

$$-\varepsilon_i^2 \psi_{k;\bar x\hat x} + a_{kk}\psi_k = -\sum_{\substack{m=1\\ m\neq k}}^{\ell} a_{km}\eta_m \quad \text{on } \omega, \quad \psi_{k;0} = \psi_{k;N} = 0.$$

Assume that the matrix $\boldsymbol{\Gamma}(\boldsymbol{A})$ is inverse-monotone. Then for each k one has

$$\|\eta_i\|_{\infty,\bar\omega} \le \|\varphi_i\|_{\infty,\bar\omega} + \|\psi_i\|_{\infty,\bar\omega} \le \|\varphi_i\|_{\infty,\bar\omega} + \sum_{\substack{m=1\\ m\neq k}}^{\ell} \left\| \frac{a_{km}}{a_{kk}} \right\|_{\infty} \|\eta_m\|_{\infty,\bar\omega} ,$$

by (6.14). Gathering together the η terms and invoking the inverse-monotonicity of $\boldsymbol{\Gamma}(\boldsymbol{A})$, we get

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

Each component φ_i of $\boldsymbol{\varphi}$ is the solution of a scalar problem and can be analysed using the technique in Sect. 6.1.2.2 combined with the derivative bounds of Theorem 3.43. This was done in [104, § 3.2] to deduce the following result:

Theorem 6.35. *Let the matrix $\boldsymbol{A}$ and the source term $\boldsymbol{f}$ be twice continuously differentiable. Assume $\boldsymbol{A}$ possesses positive diagonal entries and satisfies* (6.49). *Then the error in the solution of* (6.53) *satisfies*

$$\left\|\boldsymbol{u}-\boldsymbol{u}^N\right\|_{\infty,\bar{\omega}} \le C\left(\vartheta^{[2]}_{rd,\ell}(\bar{\omega})\right)^2.$$

Corollary 6.36. *Identifying $\mathbf{u}^N$ with a piecewise linear function on $\bar{\omega}$, we have*

$$\left\|\boldsymbol{u}-\boldsymbol{u}^N\right\|_{\infty} \le C\left(\vartheta^{[2]}_{rd,\ell}(\bar{\omega})\right)^2,$$

by Theorems 6.29 and 6.35.

Remark 6.37. An alternative analysis based on comparison principles with special barrier functions was used in [102, 118]. This technique has the more restrictive condition that $\boldsymbol{A}$ be an M-matrix and up to now has been applied successfully only to Shishkin meshes. ♣

6.2.3.3 Numerical Results

We now present the results of some numerical experiments in order to illustrate the conclusions of the preceding section, and to check if the error estimates of Theorem 6.35 are sharp.

The test problem is

$$\begin{aligned}
-\varepsilon_1^2 u_1'' + 3u_1 + (1-x)(u_2-u_3) &= \mathrm{e}^x, & u_1(0) = u_1(1) = 0,\\
-\varepsilon_2^2 u_2'' + 2u_1 + (4+x)u_2 - u_3 &= \cos x, & u_2(0) = u_2(1) = 0,\\
-\varepsilon_3^2 u_3'' + 2u_1 + 3u_3 &= 1 + x^2, & u_3(0) = u_3(1) = 0.
\end{aligned}$$

In the construction of the Bakhvalov and the Shishkin mesh (see Sect. 2.2.2) we take $\kappa/p = 0.8$.

The exact solutions to the test problems is not available, so we estimate the accuracy of the numerical solution by comparing it to the numerical solution of the Richardson extrapolation method, which is of higher order: Let $\boldsymbol{u}_{\boldsymbol{\varepsilon}}^N$ be the solution of the difference scheme on the original mesh and $\tilde{\boldsymbol{u}}_{\boldsymbol{\varepsilon}}^{2N}$ that on the mesh obtained by uniformly bisecting the original mesh. This yields the estimated error

$$\eta_{\boldsymbol{\varepsilon}}^{N} := \frac{4}{3}\left\|\boldsymbol{u}_{\boldsymbol{\varepsilon}}^{N} - \tilde{\boldsymbol{u}}_{\boldsymbol{\varepsilon}}^{2N}\right\|_{\infty,\omega}.$$

The uniform errors for a fixed N are estimated by taking the maximum error over a wide range of ε, namely

$$\eta^{N} := \max_{\varepsilon_1,\dots,\varepsilon_\ell=10^{-0},\dots,10^{-24}} \eta_{\boldsymbol{\varepsilon}}^{N}.$$

For the Shishkin mesh we have an error bound of the form $C\left(N^{-1}\ln N\right)^{p}$. Therefore, we compute the “Shishkin” rates of convergence using the formula

$$p^{N} = \frac{\ln \eta^{N} - \ln \eta^{2N}}{\ln(2\ln N) - \ln(\ln(2N))},$$

while for the Bakhvalov mesh, the standard formula

$$\rho^{N} = \left(\ln \eta^{N} - \ln \eta^{2N}\right) / \ln 2$$

is used.

The results of our test computations displayed in Table 6.2 are in agreement with Theorem 6.35. The Bakhvalov mesh gives second order accuracy, while the results for the Shishkin mesh are spoiled by the typical logarithmic factor.

Table 6.2 Central differencing on layer-adapted meshes for systems of reaction-diffusion type

	Shishkin mesh		Bakhvalov mesh	
N	η^N	p^N	η^N	ρ^N
$8 \cdot 2^3$	4.895e-02	1.22	4.910e-03	2.08
$8 \cdot 2^4$	2.276e-02	1.50	1.157e-03	2.05
$8 \cdot 2^5$	8.854e-03	1.67	2.790e-04	2.04
$8 \cdot 2^6$	3.091e-03	1.77	6.771e-05	2.01
$8 \cdot 2^7$	1.014e-03	1.83	1.676e-05	2.00
$8 \cdot 2^8$	3.201e-04	1.88	4.179e-06	2.00
$8 \cdot 2^9$	9.831e-05	1.91	1.044e-06	2.00
$8 \cdot 2^{10}$	2.955e-05	1.94	2.610e-07	2.00
$8 \cdot 2^{11}$	8.725e-06	1.96	6.527e-08	2.00
$8 \cdot 2^{12}$	2.537e-06	—	1.632e-08	—

6.2.3.4 A Posteriori Error Bounds

By interpreting the components of $\boldsymbol{u}^N$ as piecewise linear functions, one can conduct an a posteriori error analysis that combines the technique in Sect. 6.1.2.3 with Theorem 3.36 to give

$$\left\|u_k - u_k^N\right\|_\infty \le \sum_{m=1}^{\ell} \left(\boldsymbol{\Gamma}^{-1}\right)_{km} \left(\eta_{1;m} + \eta_{2;m} + \eta_{3;m}\right), \quad k = 1, \dots, \ell,$$

where

$$\eta_{1;m} := \frac{\left\|\hat{q}_m - \hat{q}_m^I\right\|_\infty}{\rho_m^2}, \quad \eta_{2;m} := \max_{i=0,\dots,N-1} \frac{h_{i+1}^2}{4\varepsilon_k^2} \max\left\{\left|q_{k;i}\right|, \left|q_{k;i+1}\right|\right\}$$

$$\eta_{3;m} := \max_{i=0,\dots,N-1} \frac{h_{i+1}}{6\varepsilon_m \rho_m} \left|\hat{q}_{m;i+1} - \hat{q}_{m;i}\right|$$

$$\hat{q}_m = f_m - \sum_{\nu=1}^{\ell} a_{m\nu} u_\nu^N \quad \text{and} \quad \rho_m = \min_{x\in[0,1]} a_{mm}(x)^{1/2}.$$

Remark 6.38. Since the a posteriori analysis uses the stability of the differential operator, a posteriori bounds can also be derived for non-monotone discretisations, like schemes generated by FEMs. If for example, FEM-2 (see p. 187) is used, then the η_3 terms in the above estimate will disappear; cf. Sect. 6.1.3.3. ♣

6.3 Reaction-Convection-Diffusion

Consider the scalar reaction-convection-diffusion problem

$$\mathcal{L}u := -\varepsilon_d u'' - \varepsilon_c b u' + cu = f \quad \text{in } (0,1), \;\; u(0) = \gamma_0, \;\; u(1) = \gamma_1, \tag{6.54}$$

where $b \ge 1$ and $c \ge 1$ on $[0,1]$, $\varepsilon_d \in (0,1]$ and $\varepsilon_c \in [0,1]$.

Analytical properties of (6.54) were studied in Sect. 3.2, while layer-adapted meshes for it have been introduced in Sect. 2.2.

We briefly recall some of the notation. Let $\lambda_0 < 0$ and $\lambda_1 > 0$ be the solution of

$$-\varepsilon_d \lambda(x)^2 - \varepsilon_c b(x)\lambda(x) + c(x) = 0. \tag{6.55}$$

Set

$$\mu_0 := \max_{x\in[0,1]} \lambda_0(x) < 0 \quad \text{and} \quad \mu_1 := \min_{x\in[0,1]} \lambda_1(x) > 0.$$

The quantity used for presenting a priori error bounds is

$$\vartheta_{rd}^{[p]}(\bar{\omega}) := \max_{i=1,\ldots,N} \int_{I_i} \left\{ 1 + |\mu_0| \, \mathrm{e}^{\mu_0 s/p\varepsilon} + \mu_1 \mathrm{e}^{-\mu_1(1-s)/p\varepsilon} \right\} \mathrm{d}s.$$

6.3.1 The Interpolation Error

Let w^I denote the piecewise linear interpolant of w again. Adapting the techniques from Sects. 5.1 and 6.1.1.1, and using the derivative bounds of Theorem 3.29 we obtain the following result.

Proposition 6.39. *Let u be the solution of (6.54). Let $p > 2$ be arbitrary, but fixed. Then*

$$\left\| u - u^I \right\|_{\infty, I_i} \le C \left\{ \int_{I_i} \left(1 + |\mu_0| \, \mathrm{e}^{\mu_0 s/p\varepsilon} + \mu_1 \mathrm{e}^{-\mu_1(1-s)/p\varepsilon} \right) \mathrm{d}s \right\}^2$$

for all mesh intervals $I_i = [x_{i-1}, x_i]$.

Define the ε_d-weighted energy norm by

$$|||v|||_{\varepsilon_d}^2 := \varepsilon_d \, |v|_1^2 + \|v\|_0^2 \, .$$

Theorem 6.40. *Let u be the solution of (6.54). Let $p > 2$ be arbitrary, but fixed. Then*

$$\left\| u - u^I \right\|_\infty \le C \left(\vartheta_{rcd}^{[p]}(\bar{\omega}) \right)^2$$

and

$$\left|\left|\left| u - u^I \right|\right|\right|_{\varepsilon_d} \le C \left(\mu_1^{-1/2} + \vartheta_{rcd}^{[p]}(\bar{\omega}) \right) \vartheta_{rcd}^{[p]}(\bar{\omega}).$$

for any $p > 2$.

Proof. The bound on the L_∞ error is an immediate consequence of Prop. 6.39 and the definition of $\vartheta_{rcd}^{[p]}$.

For the error in the H^1 semi-norm use integration by parts to get

$$\left| u - u^I \right|_1^2 = - \int_0^1 u''(x) \left(u - u^I \right)(x) \mathrm{d}x.$$

Note, that

$$\varepsilon_d \int_0^1 |u''(x)| \, \mathrm{d}x \le C \varepsilon_d \, |\mu_0| \le C \mu_1^{-1}.$$

The assertion of the theorem follows. □

Remark 6.41. The results of Theorem 6.40 hold for arbitrary $p > 2$, but not for $p = 2$. This is because of the derivative bounds provided by Theorem 3.29.

6.3.2 Simple Upwinding

This section studies a simple first-order upwind difference scheme for (6.54). The analysis essentially follows [95].

Our scheme is: Find $u^N \in I\!R^{N+1}$ such that

$$\begin{aligned} \left[Lu^N\right]_i := -\varepsilon_d u^N_{\bar{x}x;i} - \varepsilon_c b_i u^N_{x;i} + c_i u^N_i = f_i \quad \text{for } i = 1, \dots, N-1, \\ u^N_0 = \gamma_0, \ \ u^N_N = \gamma_1. \end{aligned} \tag{6.56}$$

The difference operators were introduced in Sect. 4.1.

The variational formulation of (6.56) is: Find $u^N \in V^\omega$ with $u^N_0 = \gamma_0$ and $u^N_N = \gamma_1$ such that

$$a_u\left(u^N, v\right) = f_u\left(v\right) := (f, v)_\omega \quad \text{for all } v \in V_0^\omega , \tag{6.57}$$

where

$$a_u\left(w, v\right) := \varepsilon_d \left[w_x, v_x\right)_\omega - \left(\varepsilon_c b u_x - c w, v\right)_\omega$$

and

$$\left[w, v\right)_\omega := \sum_{i=0}^{N-1} h_{i+1} w_i v_i, \quad \left(w, v\right)_\omega := \sum_{i=1}^{N-1} h_{i+1} w_i v_i.$$

Taking as test functions v the standard hat-function basis in V_0^ω, we see that (6.56) and (6.57) are equivalent. In particular, using summation by parts it is verified that

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

where

$$\left[L^* v\right]_i := -\varepsilon_d v_{\bar{x}x;i} + \varepsilon_c \left(bv\right)_{\breve{x};i} + c_i v_i \ .$$

is the adjoint operator to L with respect to the scalar product $(\cdot, \cdot)_\omega$.

6.3.2.1 Stability

The matrix associated with the difference operator L is an L_0-matrix. Application of Lemma 3.14 with a constant positive test function establishes the inverse monotonicity of L. Thus, L satisfies a comparison principle. This comparison principle gives the $(\ell_\infty, \ell_\infty)$-stability inequality

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

and provides a priori bounds for the solution of (6.56):

$$\left\|u^N\right\|_{\bar{\omega}} \le \max\left\{|\gamma_0|, |\gamma_1|\right\} + \|f/c\|_{\infty,\omega} \quad \text{for } i = 0, \dots, N.$$

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

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

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

$$\left[L^* G_{i,\cdot}\right]_j = \delta_{i,j} \quad \text{for } j = 1, \dots, N-1, \;\; G_{i,0} = G_{i,N} = 0, \tag{6.58}$$

where

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

is the discrete equivalent of the Dirac-δ distribution.

Theorem 6.42. *Assume that $b \ge 1$ and $c \ge 1$ on $[0,1]$, then there exists a positive constant C such that*

$$0 \le G_{i,j} \le K\mu_1 \quad \textit{for } i, j = 0, \dots, N,$$

where

$$K := \begin{cases} \|b\|_\infty \|c\|_\infty \max\left\{\|b\|_\infty, \|c\|_\infty\right\} & \textit{if } \varepsilon_c > 0, \\ \|c\|_\infty^{3/2} & \textit{if } \varepsilon_c = 0. \end{cases}$$

Furthermore, if $b' \ge 0$ on $[0,1]$ then

$$\begin{aligned} G_{\xi;i,j} &\ge 0 \quad \textit{for } 0 \le j < i < N, \\ G_{\xi;i,j} &\le 0 \quad \textit{for } 0 \le i \le j < N. \end{aligned}$$

Proof. Let j be fixed. Set

$$\Gamma_i := \begin{cases} \prod_{k=i+1}^{j} \left(1+\mu_1 h_k\right)^{-1} & \text{for } 0 \le i < j \le N, \\ 1 & \text{for } 0 \le i = j \le N, \\ \prod_{k=j+1}^{i} \left(1-\mu_0 h_k\right)^{-1} & \text{for } 0 \le j < i \le N. \end{cases}$$

We shall use the discrete comparison principle to show that $K\mu_1\Gamma$ is an upper bound for $G_{\cdot,j}$.

Clearly $\Gamma_0 \ge 0$ and $\Gamma_N \ge 0$.

Next, a direct calculation gives

$$\Gamma_{x;i} = \Gamma_i \begin{cases} \mu_1 & \text{for } 0 \le i < j \le N, \\ \dfrac{\mu_0}{1-\mu_0 h_{i+1}} & \text{for } 0 \le j \le i < N \end{cases}$$

and

$$\Gamma_{\bar{x};i} = \Gamma_i \begin{cases} \dfrac{\mu_1}{1+\mu_1 h_{i+1}} & \text{for } 0 < i \le j \le N, \\ \mu_0 & \text{for } 0 \le j < i \le N. \end{cases}$$

For $i < j$ we have

$$\left[L\Gamma\right]_i = \Gamma_i \left(\frac{-\varepsilon_d \mu_1^2}{1+\mu_1 h_{i+1}} - \varepsilon_c b_i \mu_1 + c_i \right) \ge \Gamma_i \left(-\varepsilon_d \mu_1^2 - \varepsilon_c b_i \mu_1 + c_i\right) \ge 0,$$

by Proposition 3.19.

For $i > j$

$$\left[L\Gamma\right]_i = \Gamma_i \left(\frac{-\varepsilon_d \mu_0^2 - \varepsilon_c b_i \mu_0}{1-\mu_0 h_{i+1}} + c_i \right) \ge \Gamma_i \frac{-\varepsilon_d \mu_0^2 - \varepsilon_c b_i \mu_0 + c_i}{1-\mu_0 h_{i+1}} \ge 0,$$

where Proposition 3.19 was used again.

For $i = j$

$$\begin{aligned} \left[L\Gamma\right]_j &= -\frac{\varepsilon_d}{h_{j+1}} \left(\frac{\mu_0}{1-\mu_0 h_{j+1}} - \frac{\mu_1}{1+\mu_1 h_j} \right) - \varepsilon_c b_j \frac{\mu_0}{1-\mu_0 h_{j+1}} + c_j \\ &\ge \frac{-\varepsilon_d \mu_0 - \left(\varepsilon_c b_j \mu_0 - c_j\right) h_{j+1}}{h_{j+1}\left(1-\mu_0 h_{j+1}\right)}. \end{aligned}$$

The function $x \mapsto \lambda_0(x)$ is continuous. Therefore there exists a $\xi \in [0,1]$ such that $\lambda_0(\xi) = \mu_0$. Furthermore, recall $b \geq 1$, $c \geq 1$ and $\mu_0 < 0$. Hence,

$$\begin{aligned}
\left[L\Gamma\right]_j &\geq \frac{1}{\max\{b_j, c_j\}} \frac{-\varepsilon_d\mu_0 - \big(\varepsilon_c b(\xi)\mu_0 - c(\xi)\big)h_{j+1}}{h_{j+1}\left(1 - \mu_0 h_{j+1}\right)} \\
&= \frac{-\varepsilon_d\mu_0}{h_{j+1}\max\{b_j, c_j\}} = \frac{c(\xi)}{h_{j+1}\max\{b_j, c_j\}\,\lambda_1(\xi)}, \quad \text{by (6.55)} \\
&\geq \frac{1}{h_{j+1}K\mu_1}, \quad \text{by (3.6).}
\end{aligned}$$

Thus, $K\mu_1\Gamma$ is a barrier function for $G_{\cdot,j}$, for all j.

In order to verify the monotonicity of $G_{i,\cdot}$ use the argument from the proof of Theorem 4.1. □

Theorem 6.43. *Assume that $b \geq 1$ and $c \geq 1$ on $[0,1]$, then there exists a positive constant C such that*

$$\|cG_{i,\cdot}\|_{1,\omega} \leq 1 \quad \textit{for } i = 0, \dots, N,$$

Furthermore, if $b' \geq 0$ on $[0,1]$ then

$$\|G_{\xi;i,\cdot}\|_{1,\omega} \leq C\mu_1 \quad \textit{for } i = 0, \dots, N.$$

Proof. To verify the bound on the ℓ_1-norm of G multiply (6.58) by h_{j+1} and sum for $j = 1, \dots, N-1$. Then use the positivity of G.

For the bound on G_ξ use the piecewise monotonicity of G_ξ to get

$$\|G_{\xi;i,\cdot}\|_{1,\omega} \leq 2G_{i,i} \leq C\mu_1,$$

by Theorem 6.42. □

6.3.2.2 A Priori Error Analysis

Theorem 6.44. *Let the assumptions of Theorems 3.29 and 6.43 be satisfied. Let u be the solution of (6.54) and u^N its approximation by (6.56). Then*

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

for any $p > 1$.

Proof. Let $\varGamma := G_{i,\cdot} \in V_0^\omega$ be the Green's function associated with L and the mesh node x_i. Then

$$\begin{aligned} u_i^N - u_i &= a_u\left(u^N - u, \varGamma\right) = f_u(\varGamma) - a_u(u, \varGamma) \\ &= f_u(\varGamma) - f(\varGamma) + a(u, \varGamma) - a_u(u, \varGamma), \quad \text{by (6.54)} \\ &= (\varepsilon_c b u_x - cu + f, \varGamma)_\omega - (\varepsilon_c b u' - cu + f, \varGamma)\,. \end{aligned}$$

We obtain the representation for the nodal error:

$$\begin{aligned} u_i^N - u_i &= \varepsilon_c \sum_{i=0}^{N-1} \int_{I_{i+1}} u'(s)\big[(b\varGamma)(s) - b_i \varGamma_i\big]\mathrm{d}s \\ &\quad + \sum_{i=0}^{N-1} \int_{I_{i+1}} \Big[(cu - f)_i \varGamma_i - (cu - f)(s)\varGamma(s)\Big]\mathrm{d}s \\ &=: T_1 + T_0\ . \end{aligned} \tag{6.59}$$

To bound the first term we proceed as follows:

$$|(b\varGamma)(s) - b_i\varGamma_i| \le \|b\|_\infty\, |\varGamma_{i+1} - \varGamma_i| + \|b'\|_\infty h_{i+1} \varGamma_i\,,$$

because $\varGamma$ is piecewise linear. Thus,

$$|T_1| \le C\varepsilon_c \vartheta_{rcd}^{[p]}(\bar{\omega})\Big\{\|b\|_\infty\, \|\varGamma_x\|_{1,\omega} + \|b'\|_\infty\, \|\varGamma\|_{1,\omega}\Big\} \tag{6.60}$$

for any $p > 1$, by Theorem 3.29.

Next we estimate T_0.

$$\begin{aligned} T_0 = \sum_{i=0}^{N-1} \int_{I_{i+1}} &\big[(cu - f)_i - (cu - f)(s)\big]\varGamma_i \mathrm{d}s \\ &+ \sum_{i=0}^{N-1} \int_{I_{i+1}} (cu - f)(s)\big[\varGamma_i - \varGamma(s)\big]\mathrm{d}s. \end{aligned}$$

Hence,

$$\begin{aligned} |T_0| \le \|\varGamma\|_{1,\omega} &\max_{i=1,\dots,N} \int_{I_i} \big|(f - cu)'(s)\big|\mathrm{d}s \\ &+ \|\varGamma_x\|_{1,\omega} \max_{i=1,\dots,N} \int_{I_i} \big|\varepsilon_d u''(s) + \varepsilon_c b(s) u'(s)\big|\mathrm{d}s\,, \end{aligned}$$

where (6.54) was used. To bound u' and u'' apply Theorem 3.29. Then, for any $p > 1$,

$$|T_0| \le C\vartheta_{rcd}^{[p]}(\bar{\omega}) \left(\|\mathit{\Gamma}\|_{1,\omega} + \mu_1^{-1} \|\mathit{\Gamma}_x\|_{1,\omega} \right),$$

where we have used that $\varepsilon_d|\mu_0| + \varepsilon_d|\mu_1| \le C\mu_1^{-1}$ and $\varepsilon_c \le C\mu_1^{-1}$.

Finally, combining the last inequality with (6.59) and (6.60), we get

$$\left|u_i^N - u_i\right| \le C\vartheta_{rcd}^{[p]}(\bar{\omega}) \left(\|\mathit{\Gamma}\|_{1,\omega} + \mu_1^{-1} \|\mathit{\Gamma}_x\|_{1,\omega} \right).$$

Application of Theorem 6.43 completes the analysis. □

Remark 6.45. The truncation error and barrier function technique (see Sect. 4.2.6) was used in [80, 105] when studying the difference scheme

$$-\varepsilon_d u_{\bar{x}\hat{x};i}^N - \varepsilon_c b_i u_{x;i}^N + c_i u_i^N = f_i \tag{6.61}$$

on a Shishkin mesh. The authors establish the typical rate of $N^{-1} \ln N$ if the critical mesh parameter satisfies $\sigma > 2$. In [80, 105] hybrid difference schemes were also studied that raise the order of convergence to almost second order. ♣

Remark 6.46. Gracia et al. [47] combine (6.61) with the mid-point upwind scheme

$$-\varepsilon_d u_{\bar{x}\hat{x};i}^N - \varepsilon_c b_{i+1/2} u_{x;i}^N + c_{i+1/2}\left(u_i^N + u_{i+1}^N\right)/2 = f_{i+1/2}$$

and central differencing

$$-\varepsilon_d u_{\bar{x}\hat{x};i}^N - \varepsilon_c b_i u_{\mathring{x};i}^N + c_i u_i^N = f_i$$

in order to obtain an inverse monotone method that is second-order consistent in the maximum norm for all values of ε_d, ε_c and N.

In [47] that scheme is shown to be uniformly convergent on a Shishkin mesh with

$$\left\|u^N - u\right\| \omega \le CN^{-2} \ln^3 N.$$

The analysis in [47], using solution decompositions and distinguishing different parameter regimes, is very tedious. Results for general meshes are not available.

Surla et al. [158] design an inverse monotone spline difference scheme. They prove uniform convergence on a Shishkin mesh with

$$\left\|u^N - u\right\| \omega \le CN^{-2} \ln^2 N,$$

if $\varepsilon_d, \varepsilon_c \le CN^{-1}$. ♣

6.3.2.3 A Numerical Example

In this section a brief illustration for the a priori error bounds of Theorem 6.44 is shown. The test problem is:

$$-\varepsilon_d u'' + \varepsilon_c u' + u = \cos \pi x \ \text{ in } \ (0,1), \quad u(0) = u(1) = 0.$$

Its solution is easily computed using, e. g., MAPLE.

Indicating by $u^N_{\varepsilon_d,\varepsilon_c}$ that the numerical approximation depends on N, ε_d and ε_c, and by $u_{\varepsilon_d,\varepsilon_c}$ that the exact solution depends on ε_d and ε_c, we estimate the uniform error by

$$\eta^N := \max_{\substack{\varepsilon_d=1,10^{-1},\dots,10^{-12}\\ \varepsilon_c=1,10^{-1},\dots,10^{-12},0}} \left\| u^N_{\varepsilon_d,\varepsilon_c} - u^I_{\varepsilon_d,\varepsilon_c} \right\|_\infty .$$

The rates of convergence are computed using the standard formula $r^N = \log_2\left(\eta^N/\eta^{2N}\right)$. We also compute the constant in the error estimate, i. e., for a given theoretical error bound $\eta^N \le C\vartheta(N)$, we approximate the constant in the error estimate by $C^N = \eta^N/\vartheta(N)$.

Table 6.3 displays the results of our test computations. The numbers confirm the results of Theorem 6.44 for both the Shishkin and the Bakhvalov mesh.

6.3.2.4 A Posteriori Error Analysis

The analysis for all schemes starts from the Green's-function representation (3.1) and utilises Theorem 3.23, when bounds on the norms of the Green's function $\mathcal{G}$ are required.

Table 6.3 Uniform nodal errors of the upwind difference scheme applied to a model reaction-convection-diffusion problem

	Shishkin mesh			Bakhvalov mesh		
N	η^N	r^N	C^N	η^N	r^N	C^N
2^9	1.227e-2	0.83	1.007	3.404e-3	0.99	1.743
2^{10}	6.905e-3	0.85	1.020	1.713e-3	1.00	1.754
2^{11}	3.824e-3	0.87	1.027	8.592e-4	1.00	1.760
2^{12}	2.093e-3	0.88	1.031	4.302e-4	1.00	1.762
2^{13}	1.136e-3	0.89	1.032	2.153e-4	1.00	1.763
2^{14}	6.119e-4	0.90	1.033	1.077e-4	1.00	1.764
2^{15}	3.279e-4	0.91	1.033	5.384e-5	1.00	1.764
2^{16}	1.749e-4	0.91	1.033	2.692e-5	1.00	1.764
2^{17}	9.290e-5	0.92	1.033	1.346e-5	1.00	1.764
2^{18}	4.922e-5	—	1.034	6.731e-6	—	1.764

Let

$$q := f - cu^N + \varepsilon_c b \left(u^N\right)'.$$

Clearly q may have discontinuities at the mesh points because $u^N \in V^\omega$. Therefore, set

$$q_i^+ = \lim_{x \to x_i + 0} q(x) \quad \text{and} \quad q_i^- = \lim_{x \to x_i - 0} q(x)$$

for all mesh nodes x_i.

Fix $x \in (0,1)$ and set $\Gamma := \mathcal{G}(x,\cdot)$. Then (3.1) and (6.54) yield

$$\begin{aligned}\left(u - u^N\right)(x) &= a\left(u - u^N, \Gamma\right) = f(\Gamma) - a\left(u^N, \Gamma\right) \\ &= \left(f - f_u\right)(\Gamma) + \left(a_u - a\right)\left(u^N, \Gamma\right) \\ &= \sum_{i=1}^N \left[q_{i-1}^+ \Gamma_{i-1} h_i - \int_{I_i} \left(q\Gamma\right)(s)\mathrm{d}s\right].\end{aligned}$$

We see the approximation error is bounded by the error of the one-sided rectangle rule used for integrating $q\Gamma$. A Taylor expansion gives

$$\begin{aligned}&\left| h_i q_{i-1}^+ \Gamma_{i-1} - \int_{I_i} \left(b\Gamma\right)(\xi)\mathrm{d}\xi \right| \\ &\quad \le h_i \left\{ \int_{I_i} |q'(s)| \Gamma(s)\mathrm{d}s + \int_{I_i} |q(s)|\, |\Gamma'(s)|\mathrm{d}s \right\} \\ &\quad \le h_i \left\{ \|q'/c\|_{\infty,I_i} \int_{I_i} c(s)\Gamma(s)\mathrm{d}s + \|q\|_{\infty,I_i} \int_{I_i} |\Gamma'(s)|\mathrm{d}s \right\}.\end{aligned}$$

Using the bounds on the Green's function from Theorem 3.23, we arrive at the main result of this section.

Theorem 6.47. *Suppose* (3.8) *is satisfied, then the error of the first-order upwind scheme* (6.56) *satisfies*

$$\left\|u - u^N\right\|_\infty \le \eta_1^u + \eta_2^u,$$

where $\eta_k^u := \max_{i=1,\dots,N} \eta_{k,i}^u$ *and*

$$\eta_{1,i}^u := h_i \left\|q'/c\right\|_{\infty,I_i}, \qquad \eta_{2,i}^u := \frac{2h_i}{\varepsilon_d\left(\mu_1 - \mu_0\right)} \|q\|_{\infty,I_i}.$$

Remark 6.48. The error has been bounded in terms of the numerical solution u^N and of the data of the problem. The two parts of the error bound can be further expanded. For example:

$$h_i \left\| \frac{q'}{c} \right\|_{\infty,I_i} \le \left\{ h_i \left\| \frac{f' - c'u^N}{c} \right\|_{\infty,I_i} + \left\| \frac{\varepsilon_c b'}{c} - 1 \right\|_{\infty,I_i} \left| u_{i+1}^N - u_i^N \right| \right\}. \quad (6.62)$$

Apparently, sampling of the data is inevitable. However, instead of sampling (6.62) it seems advisable to sample the η_i^u:

$$\eta_{1,i}^u \approx \tilde{\eta}_{1,i}^u := \left| \frac{q_i^- - q_{i-1}^+}{c_{i-1/2}} \right| \quad \text{and} \quad \eta_{2,i}^u \approx \tilde{\eta}_{2,i}^u := \frac{2h_i}{\varepsilon_d(\mu_1 - \mu_0)} \left| q_{i-1/2} \right|$$

This avoids the use of a triangle inequality and therefore gives in general sharper upper bounds for the error. ♣

Remark 6.49. The leading term in the estimate of Theorem 6.47 is a discrete first-order derivative. Therefore, the de Boor algorithm of Sect. 4.2.4.2 equidistributing the arc-length of the numerical solution can be used for adaptive mesh generation, when the simple upwind scheme (6.56) is applied to the boundary-value problem (6.54). ♣