Chapter 5
Finite Element and Finite Volume Methods

In this chapter we consider finite element and finite volume discretisations of

$$\mathcal{L}u := -\varepsilon u'' - bu' + cu = f \quad \text{in } (0,1), \;\; u(0) = u(1) = 0, \tag{5.1}$$

with $b \geq \beta > 0$. Its associated variational formulation is: Find $u \in H_0^1(0,1)$ such that

$$a(u,v) = f(v) \quad \text{for all } v \in H_0^1(0,1), \tag{5.2}$$

where

$$a(u,v) := \varepsilon\,(u',v') - (bu',v) + (cu,v)$$

and

$$f(v) := (f,v) := \int_0^1 \big(fv\big)(x)\mathrm{d}x. \tag{5.3}$$

Throughout assume that

$$c + b'/2 \geq \gamma > 0. \tag{5.4}$$

This condition guaranties the coercivity of the bilinear form in (5.2):

$$|||v|||_\varepsilon^2 := \varepsilon\,\|v'\|_0^2 + \gamma\,\|v\|_0^2 \leq a(v,v) \quad \text{for all } v \in H_0^1(0,1).$$

This is verified using standard arguments, see e.g. [141]. If $b \geq \beta > 0$ then (5.4) can always be ensured by a transformation $\bar{u}(x) = u(x)\mathrm{e}^{\delta x}$ with δ chosen appropriately. We assume this transformation has been carried out.

We start our investigations with interpolation-error estimates and a Galerkin discretisations of (5.1)—including aspects of convergence, superconvergence, and postprocessing of the derivatives. Then stabilised finite element methods are considered. We finish with an upwinded finite volume method.

T. Linß, *Layer-Adapted Meshes for Reaction-Convection-Diffusion Problems*, Lecture Notes in Mathematics 1985, DOI 10.1007/978-3-642-05134-0_5,

5.1 The Interpolation Error

In this section we study the error in linear interpolation. The argument follows [84]. Let $\bar{\omega}$ be an arbitrary mesh. Let w^I denote the piecewise-linear function that interpolates to w at the nodes of $\bar{\omega}$.

In this section let us assume the function $\psi \in C^2[0,1]$ admits the derivative bounds

$$|\psi''(x)| \leq C\left\{1+\varepsilon^{-2}\mathrm{e}^{-\beta x/\varepsilon}\right\}. \tag{5.5}$$

For example, the solution u of the boundary-value problem (5.1) belongs to this class of functions, see Sect. 3.4.1.2.

Proposition 5.1. *Suppose ψ satisfies (5.5). Then*

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

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

Proof. For the interpolation error on I_i we have the representation

$$\left(\psi^I-\psi\right)(x) = \frac{1}{h_i}\int_{I_i}\int_{x_{i-1}}^{x}\int_{\xi}^{s}\psi''(t)\mathrm{d}t\mathrm{d}\xi\mathrm{d}s.$$

The right-hand side can be estimated to give

$$\left|\left(\psi^I-\psi\right)(x)\right| \leq \int_{I_i}\left(\xi-x_{i-1}\right)\left|\psi''(\xi)\right|\mathrm{d}\xi.$$

Using Lemma 4.16 and (5.5) to bound the right-hand side, we are finished. □

Theorem 5.2. *Suppose ψ satisfies (5.5). Then*

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

and

$$\left|\left|\left|\psi^I-\psi\right|\right|\right|_\varepsilon \leq C\vartheta_{cd}^{[2]}(\bar{\omega}).$$

Remark 5.3. The quantity

$$\vartheta_{cd}^{[p]}(\bar{\omega}) := \max_{k=0,\dots,N-1}\int_{I_k}\left(1+\varepsilon^{-1}\mathrm{e}^{-\beta s/p\varepsilon}\right)\mathrm{d}s,$$

was introduced in Sect. 2.1, where bounds on $\vartheta_{cd}^{[p]}(\bar{\omega})$ for various layer-adapted meshes are given too. ♣

Proof (of Theorem 5.2). The bound on the L_∞ error is an immediate consequence of Prop. 5.1 and the definition of $\vartheta_{cd}^{[2]}$.

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

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

Thus,

$$\left\| \left(\psi^I - \psi\right)' \right\|_0^2 \leq \left\| \psi^I - \psi \right\|_\infty \int_0^1 \left|\psi''(x)\right| \mathrm{d}x \leq C\varepsilon^{-1} \left\| \psi^I - \psi \right\|_\infty$$

by a Hölder inequality and (5.5). Finally, combine this with the bound for the L_2 norm of the interpolation error to obtain the energy-norm estimate. □

Remark 5.4. Proposition 5.1 can be used to give local estimates for the interpolation error too. For example on S-type meshes (see Sect. 2.1.3) one has

$$\left\| \psi^I - \psi \right\|_{0,[\tau,1]} \leq \left\| \psi^I - \psi \right\|_{\infty,[\tau,1]} \leq CN^{-2}, \quad \text{if } \sigma \geq 2,$$

for the interpolation error outside the layer region. This is in general a sharper bound than that implicitly given by Theorem 5.2. ♣

Remark 5.5. The maximum-norm interpolation error bound can be generalised to Lagrange interpolation with polynomial of arbitrary degree $p \geq 0$.

Fix $0 \leq \xi_0 < \xi_1 < \cdots < \xi_p \leq 1$. Define an interpolant $I_p\psi$ of ψ by

$$I_p\psi\big|_{I_i} \in \Pi_p$$

and

$$\left(I_p\psi\right)\left(x_{i-1} + \xi_k h_i\right) = \psi\left(x_{i-1} + \xi_k h_i\right) \quad \text{for } i = 1, \dots, N, \;\; k = 0, \dots, p.$$

If

$$\left|\psi^{(p+1)}(x)\right| \leq C\left\{1 + \varepsilon^{-(p+1)} \mathrm{e}^{-\beta x/\varepsilon}\right\}$$

then

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

This result applies, for example, to the solution u of (5.1). ♣

5.2 Linear Galerkin FEM

We start from the weak formulation (5.2). Let $\bar{\omega}$ be an arbitrary mesh and let V^ω denote the space of continuous, piecewise linear functions on ω that vanish for $x = 0$ and $x = 1$. Then our discretisation is: Find $u^N \in V^\omega$ such that

$$a(u^N, v) = f(v) \quad \text{for all } \ v \in V^\omega. \tag{5.6}$$

The coercivity of $a(\cdot,\cdot)$ guarantees the existence of unique solutions of both (5.2) and of (5.6).

Notation. *Throughout this section we use $\|\cdot\|_1$ to denote the L_1 norm. This cannot be confused with the H_1 norm because we only use the weighted H_1 norm $|||\cdot|||_\varepsilon$.*

5.2.1 Convergence

Based on the interpolation error bounds of Sect. 5.1 we conduct an error analysis for the Galerkin FEM on S-type meshes (see Sect. 2.1.3). The technique we shall use was developed by Stynes and O'Riordan [152] for standard Shishkin meshes and later generalised for S-type meshes by Linß and Roos [82, 137]. The technique can be extended to discretisations of two-dimensional problems using triangular or rectangular elements on tensor-product S-type meshes; see Sect. 9.2.2.1.

Theorem 5.6. *Let $\bar{\omega}$ be an S-type mesh with $\sigma \geq 2$ whose mesh generating function $\tilde{\varphi}$ satisfies (2.8) and*

$$\max|\psi'| \ln^{1/2} N \leq CN. \tag{5.7}$$

Then

$$|||u - u^N|||_\varepsilon \leq C\left(h + N^{-1}\max|\psi'|\right)$$

for the error of the Galerkin FEM.

Remark 5.7. The additional assumption (5.7) does not constitute a major restriction. For example both the standard Shishkin mesh and the Bakhvalov-Shishkin mesh satisfy this condition. ♣

Proof (Proof of Theorem 5.6). Let $\eta = u^I - u$ and $\chi = u^I - u^N$. For the interpolation error η, we get from Sect. 5.1, the derivative bounds (3.30) and from (2.9) that

$$|||\eta|||_\varepsilon \leq C\left(h + N^{-1}\max|\psi'|\right). \tag{5.8}$$

To bound χ we start from the coercivity of $a(\cdot,\cdot)$ and the orthogonality of the Galerkin method:

$$|||\chi|||_\varepsilon^2 \le a(\chi,\chi) = a(\eta,\chi) = \varepsilon(\eta',\chi') + (b\eta,\chi') + ((c+b')\eta,\chi)\,.$$

Apply the Cauchy-Schwarz inequality to the diffusion and reaction terms and the Hölder inequality to the convection term to estimate

$$|||\chi|||_\varepsilon^2 \le C\,|||\eta|||_\varepsilon\,|||\chi|||_\varepsilon + C\left(\|\eta\|_{\infty,[0,\tau]}\,\|\chi'\|_{1,[0,\tau]} + \|\eta\|_{\infty,[\tau,1]}\,\|\chi'\|_{1,[\tau,1]}\right),$$

where τ is the mesh transition point in the S-type mesh. On $[0,\tau]$ we use

$$\|\chi'\|_{1,[0,\tau]} \le C\tau^{1/2}\,\|\chi'\|_{0,[0,\tau]} \le C\ln^{1/2}N\,|||\chi|||_\varepsilon\,,$$

while on $[\tau,1]$ we have by an inverse inequality

$$\|\chi'\|_{1,[\tau,1]} \le CN\,\|\chi\|_{1,[\tau,1]} \le CN\,\|\chi\|_{0,[\tau,1]} \le CN\,|||\chi|||_\varepsilon\,.$$

These two bounds and the interpolation results of Sect. 5.1 yield

$$|||\chi|||_\varepsilon \le C\Big\{h + N^{-1}\max|\psi'| \\ \qquad + \left(h + N^{-1}\max|\psi'|\right)^2\ln^{1/2}N + N^{-1}\Big\}.$$

Thus,

$$|||\chi|||_\varepsilon \le C\left(h + N^{-1}\max|\psi'|\right),$$

where we have used (5.7). Applying a triangle inequality and the bounds for $|||\eta|||_\varepsilon$ and $|||\chi|||_\varepsilon$, we complete the proof. □

Remark 5.8. Sun and Stynes [157] use a similar technique to study the Galerkin-FEM on standard Shishkin meshes for higher-order problems. ♣

Remark 5.9. We are not aware of a general convergence theory for the Galerkin FEM on arbitrary layer-adapted meshes.

Roos [136] proves the optimal uniform error estimate

$$\left|\left|\left|u - u^N\right|\right|\right|_\varepsilon \le CN^{-1}$$

for the Galerkin FEM on a special B-type mesh under the assumption that $\varepsilon \le N^{-1}$. The key ingredient in his analysis is the use of a special quasi-interpolant with an improved stability property. However, he points out that this technique cannot be extended to higher dimensions. ♣

5.2.2 Supercloseness

In the preceding section we have seen that the Galerkin FEM is (almost) first-order convergent in the ε-weighted energy norm. Now we prove that $|||u^I - u^N|||_\varepsilon$ converges faster than $|||u - u^N|||_\varepsilon$. This means the numerical approximation is *closer to the interpolant* of the exact solution than to the solution itself. This phenomenon is called **supercloseness**. Our analysis follows [83, 176] where two-dimensional problems are studied.

Theorem 5.10. *Let $\bar{\omega}$ be an S-type mesh with $\sigma \geq 5/2$ whose mesh generating function $\tilde{\varphi}$ satisfies* (2.8). *Then*

$$|||u^I - u^N|||_\varepsilon \leq C \left(h^2 \ln^{1/2} N + N^{-2} \max |\psi'|^2 \right) \tag{5.9}$$

for the solution of the Galerkin FEM.

Proof. For the sake of simplicity we assume that b is constant. Let again $\eta = u^I - u$ and $\chi = u^I - u^N$. Then

$$a(\eta, \chi) = \varepsilon(\eta', \chi') - (b\eta', \chi) + (c\eta, \chi)$$

For the diffusion term, integration by parts gives

$$\int_{I_i} \eta'\chi' = \eta\chi'|_{x_{i-1}}^{x_i} - \int_{I_i} \eta\chi'' = 0,$$

because $\eta(x_{i-1}) = \eta(x_i) = 0$ and because χ is linear. Thus, $(\eta', \chi') = 0$. The reaction term is easily bounded using the Cauchy-Schwarz inequality:

$$|(c\eta, \chi)| \leq C \|\eta\|_0 \|\chi\|_0 \leq C \left(h + N^{-1} \max |\psi'|\right)^2 \|\chi\|_0 ,$$

by Theorem 5.2.

We are left with the convection term. Recalling the decomposition (3.34), we split as follows:

$$\begin{aligned} (\eta', \chi) = -\int_0^\tau (w^I - w)\chi' - \int_0^\tau (v^I - v)\chi' \\ - \int_\tau^1 (w^I - w)\chi' + \int_\tau^1 (v^I - v)'\chi. \end{aligned} \tag{5.10}$$

The four terms on the right-hand side are bounded separately.

(i) By a standard interpolation error result

$$\|w^I - w\|_{0,I_i} \leq Ch_i^2 \|w''\|_{0,I_i} .$$

This and the bounds for the derivatives of w give

$$\left\|w^I - w\right\|_{0,(0,\tau)}^2 \le C \sum_{i=1}^{qN} h_i^4 \int_{I_i} \varepsilon^{-4} \mathrm{e}^{-2\beta x/\varepsilon} \mathrm{d}x \le C \sum_{i=1}^{qN} \left(\frac{h_i}{\varepsilon}\right)^4 h_i \mathrm{e}^{-2\beta x_{i-1}/\varepsilon}$$
$$\le C \left(N^{-1} \max|\psi'|\right)^4 \sum_{i=1}^{qN} h_i \mathrm{e}^{(4/\sigma-2)\beta x_i/\varepsilon},$$

by (2.12). Next, from (2.11)

$$h_i \mathrm{e}^{(4/\sigma-2)\beta x_i/\varepsilon} \le C h_i \mathrm{e}^{(4/\sigma-2)\beta x_{i-1}/\varepsilon} \le C \int_{I_i} \mathrm{e}^{(4/\sigma-2)\beta x/\varepsilon} \mathrm{d}x.$$

Therefore,

$$\left\|w^I - w\right\|_{0,(0,\tau)}^2 \le C \left(N^{-1} \max|\psi'|\right)^4 \int_0^\tau \mathrm{e}^{(4/\sigma-2)\beta x/\varepsilon} \mathrm{d}x$$
$$\le C\varepsilon \left(N^{-1} \max|\psi'|\right)^4,$$

where we have used $\sigma > 2$. This result and the Cauchy-Schwarz inequality yield

$$\left|\int_0^\tau (w^I - w)\chi'\right| \le C \left(N^{-1} \max|\psi'|\right)^2 |||\chi|||_\varepsilon . \tag{5.11}$$

(ii) To bound the second term we proceed as follows:

$$\left\|v^I - v\right\|_{0,(0,\tau)}^2 \le C \sum_{i=1}^{qN} h_i^4 \left\|v''\right\|_{0,I_i}^2 \le Ch^4 \left\|v''\right\|_{0,(0,\tau)}^2 \le Ch^4 \varepsilon \ln N,$$

since $|v''| \le C$ on $[0,1]$. Hence,

$$\left|\int_0^\tau (v^I - v)\eta\chi'\right| \le Ch^2 \ln^{1/2} N \, |||\chi|||_\varepsilon . \tag{5.12}$$

(iii) Now we consider $\int_\tau^1 (w - w^I)\chi'$. The argument splits the integral once more, but first let us recall that the mesh on $(\tau, 1)$ is uniform with mesh diameter H: $N^{-1} \le H \le N^{-1}/(1-q)$. We have

$$\left\|w^I - w\right\|_{0,I_{qN}}^2 \le CN^{-1} \mathrm{e}^{-2\beta\tau/\varepsilon} \le CN^{-6}$$

since $\sigma \ge 5/2$. Thus,

$$\left|\int_\tau^{x_{qN+1}} (w - w^I)\chi'\right| \le CN^{-2} \|\chi\|_0,$$

by an inverse inequality. Next we have

$$\begin{aligned}\left\|w^I - w\right\|^2_{0,(x_{qN+1},1)} &\le 2 \sum_{i=qN+1}^{N-1} H \left\|w\right\|_{\infty,I_i} \\ &\le C \sum_{i=qN+1}^{N-1} H \mathrm{e}^{-2\beta x_i/\varepsilon} \le C \int_\tau^{x_{N-1}} \mathrm{e}^{-2\beta x/\varepsilon} \mathrm{d}x \le C\varepsilon N^{-5}.\end{aligned}$$

Thus,

$$\left| \int_\tau^1 (w - w^I)\chi' \right| \le CN^{-2} \left|\!\left|\!\left| \chi \right|\!\right|\!\right|_\varepsilon . \tag{5.13}$$

(iv) To bound the last term in (5.10) we use the integral identity

$$\begin{aligned}\int_{I_i} \left(v - v^I\right)' \chi = \frac{1}{6} \int_{I_i} v''' \left(E_i^2\right)' \chi' \\ - \frac{1}{3}\left(\frac{h_i}{2}\right)^2 \int_{I_i} v'''\chi + \frac{1}{3}\left(\frac{h_i}{2}\right)^2 v''\chi\Big|_{x_{i-1}}^{x_i}\end{aligned} \tag{5.14}$$

with

$$E_i(x) = \frac{1}{2}\left((x - x_{i-1/2})^2 - \left(\frac{h_i}{2}\right)^2 \right) = \frac{1}{2}\left(x - x_{i-1}\right)\left(x - x_i\right).$$

This expansion formula holds true for arbitrary functions $v \in W^{3,\infty}(I_i)$ and linear functions χ; cf. [78]. We get

$$\int_\tau^1 \left(v - v^I\right)' \chi = \frac{1}{6} \int_\tau^1 v''' \left(E^2\right)' \chi' - \frac{H^2}{12}\left(v''\chi\right)(\tau) - \frac{H^2}{12} \int_\tau^1 v''\chi.$$

Assuming more regularity of the data, the decomposition (3.34) can be sharpened to give $|v'''| \le C$. This yields

$$\begin{aligned}\left| \int_\tau^1 \left(v - v^I\right)' \chi \right| &\le CH^3 \left\|\chi'\right\|_{1,(\tau,1)} + CH^2 \left|\chi(\tau)\right| + CH^2 \left\|\chi\right\|_{1,(\tau,1)} \\ &\le CH^2 \left(\left|\!\left|\!\left|\chi\right|\!\right|\!\right|_\varepsilon + \left|\chi(\tau)\right|\right),\end{aligned}$$

by an inverse inequality. Finally, we estimate

$$\left|\chi(\tau)\right| = \left| \int_0^\tau \chi'(s)\mathrm{d}s \right| \le \tau^{1/2} \left\|\chi'\right\|_{0,(0,\tau)} \le C \ln^{1/2} N \left|\!\left|\!\left|\chi\right|\!\right|\!\right|_\varepsilon .$$

Thus,

$$\left| \int_{\tau}^{1} \left(v - v^I\right)' \chi \right| \leq C H^2 \ln^{1/2} N \, |||\chi|||_{\varepsilon} \, . \tag{5.15}$$

Combine (5.10)-(5.15) to get for the convection term

$$|(\eta', \chi)| \leq C \left(h^2 \ln^{1/2} N + N^{-2} \max |\psi'|^2 \right) |||\chi|||_{\varepsilon} \, .$$

This inequality, the bounds for the diffusion and reaction terms and the coercivity of $a(\cdot,\cdot)$ yield the proposition of the theorem. □

Remark 5.11. Surprisingly, the major difficulty in the proof does not arise from the layer term, but from the regular solution component. To cope with this, the special integral expansion formula (5.14) by Lin had to be used. ♣

Remark 5.12. Another attempt at a supercloseness result for convection-diffusion problems is [175]. In that paper, finite elements with piecewise polynomials of degree $p \geq 1$ are used on a piecewise uniform mesh with transition point

$$\tau = \min \left\{ \frac{1}{2}, \frac{\varepsilon(p + 3/2)}{\beta} \ln(N+1) \right\} .$$

It was established that when the regular solution component v lies in the finite element space then

$$\left|\left|\left| Q_p u - u^N \right|\right|\right|_{\varepsilon} \leq C \left(\frac{\ln(N+1)}{N} \right)^{p+1} ,$$

where $Q_p u$ is the $(p+1)$-point Gauss-Lobatto interpolant of u. This is a supercloseness result, because in general, one has for the interpolation error

$$|||Q_p u - u|||_{\varepsilon} \leq C \left(\frac{\ln(N+1)}{N} \right)^{p} .$$

However, the assumption that the regular solution component lies in the finite element space is not very reasonable. If it were to hold for two different values of the mesh parameter N then $v \in \Pi_p$, because Shishkin meshes for different N are not nested.

This too illustrates the technical difficulties with the regular solution component just mentioned. Unfortunately—unlike (5.14) for linear elements—no expansion formulae for the convection term are available for quadratic or higher-order elements. ♣

5.2.3 Gradient Recovery and a Posteriori Error Estimation

Supercloseness results like Theorem 5.10 are basic ingredients for the superconvergent recovery of gradients, see for instance [2]. Furthermore, if a superconvergent recovery operator is available, then it is possible to define an a posteriori error estimator that is asymptotically exact. The presentation follows [138], where further details can be found.

First, we define for a given $v \in V^\omega$ a recovery operator for the derivative. Set

$$(Rv)(x) := \alpha_{i-1}\frac{x_i - x}{h_i} + \alpha_i \frac{x - x_{i-1}}{h_i} \quad \text{for } x \in I_i,\ i = 2, \ldots, N-1,$$

where α_i denotes the weighted average of the constant values of v' on the subintervals adjacent to x_i:

$$\alpha_i := \frac{h_{i+1} v'\big|_{I_i} + h_i v'\big|_{I_{i+1}}}{h_i + h_{i+1}}.$$

For the boundary intervals we simply extrapolate the well-defined linear function of the adjacent interval.

Our aim is to prove a superconvergence estimate for $\varepsilon^{1/2}\left\|u' - Ru^N\right\|_0$ that is superior to that of Theorem 5.6 for $\varepsilon^{1/2}\left\|u' - (u^N)'\right\|_0$. The key ingredients are the supercloseness property of the Galerkin solution, i.e. Theorem 5.10, and the consistency and stability of the recovery operator R.

Consistency: Let v be a quadratic function on $\tilde{I}_i$, the union of I_i and its adjacent mesh intervals. Then

$$R\left(v^I\right) = v' \quad \text{on } I_i. \tag{5.16a}$$

Stability:

$$\|Rv\|_{0,I_i} \le C\,\|v'\|_{0,\tilde{I}_i} \quad \text{for all } v \in V^\omega. \tag{5.16b}$$

We start our analysis from a triangle inequality:

$$\left\|u' - Ru^N\right\|_0 \le \left\|u' - R\left(u^I\right)\right\|_0 + \left\|R\left(u^I - u^N\right)\right\|_0.$$

The second term in this inequality can be bounded using Theorem 5.10 and (5.16b). Thus, we are left with the problem of estimating $\left\|u' - R\left(u^I\right)\right\|_0$. To take advantage of the consistency property (5.16a) we introduce a quadratic approximation of u on $\tilde{I}_i$: $Q_i u$. Using a triangle inequality, we obtain

$$\begin{aligned}\left\|u' - R\left(u^I\right)\right\|_{0,I_i} &\le \left\|u' - (Q_i u)'\right\|_{0,I_i} \\ &\quad + \left\|(Q_i u)' - R\left((Q_i u)^I\right)\right\|_{0,I_i} + \left\|R\left((Q_i u - u)^I\right)\right\|_{0,I_i}.\end{aligned}$$

The second term vanishes because of (5.16a). The last term can be bounded using (5.16b) and the stability of the linear interpolation in H^1, i. e., $|v^I|_1 \leq C|v|_1$. We get

$$\left\|u' - R\left(u^I\right)\right\|_0^2 \leq C \sum_{i=1}^{N} \left\|u' - (Q_i u)'\right\|_{0,\tilde{I}_i}^2 .$$

Note this H^1 stability of the interpolation operator holds true only in the one-dimensional case. In two dimensions the L_∞ stability of the interpolation operator has to be used instead, see Sect. 9.2.2.5.

Choosing $Q_i u$ to be, e. g., that bilinear function that coincides with u at the midpoint and both endpoints of $\tilde{I}_i$ and estimating the interpolation error carefully, see [138], we obtain

$$\varepsilon \sum_{i=1}^{N} \left\|u' - (Q_i u)'\right\|_{0,\tilde{I}_i}^2 \leq C \left(h + N^{-1} \max|\psi'|\right)^4 \quad \text{if } \sigma \geq 2.$$

Combining these estimates, we get the following result.

Theorem 5.13. *Let $\bar{\omega}$ be a S-type mesh with $\sigma \geq 5/2$ whose mesh generating function $\tilde{\varphi}$ satisfies (2.8). Then the error of the recovered gradient of the Galerkin FEM satisfies*

$$\varepsilon^{1/2}\left\|u' - Ru^N\right\|_0 \leq C\left(h^2 \ln^{1/2} N + N^{-2} \max|\psi'|^2\right).$$

Remark 5.14. Using Ru^N instead of u', we get an asymptotically exact error estimator for the weighted H^1-seminorm of the finite element error $\varepsilon^{1/2}\left\|u' - U'\right\|_0$ on S-type meshes:

$$\varepsilon^{1/2}\left\|u' - (u^N)'\right\|_0 = \varepsilon^{1/2}\left\|Ru^N - (u^N)'\right\|_0 + \mathcal{O}\left(h^2 \ln^{1/2} N + N^{-2} \max|\psi'|^2\right).$$

In the generic case one has

$$\varepsilon^{1/2}\left\|u' - (u^N)'\right\|_0 = \mathcal{O}\left(h + N^{-1} \max|\psi'|\right).$$

Thus, the above error estimator is asymptotically exact for $N \to \infty$. ♣

Remark 5.15. There are various other means of postprocessing to obtain superconvergent approximations for the derivatives. ♣

5.2.4 A Numerical Example

Let us briefly illustrate our theoretical results for the linear Galerkin FEM on S-type meshes when applied to the test problem

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

For our tests we take $\varepsilon = 10^{-8}$ which is a sufficiently small choice to bring out the singularly perturbed nature of the problem.

We consider three different S-type meshes: the original Shishkin mesh, the Bakhvalov-Shishkin mesh and a mesh with a rational mesh characterising function ψ. The results of our test computations are presented in Tables 5.1, 5.2 and 5.3. They are clear illustrations of the a priori error bounds given in Theorems 5.6, 5.10 and 5.13.

Table 5.1 Galerkin FEM on a S-type mesh with a rational ψ ($m = 2$)

	$\vert\vert\vert u - u^N \vert\vert\vert_\varepsilon$		$\vert\vert\vert u^I - u^N \vert\vert\vert_\varepsilon$		$\varepsilon^{1/2}\Vert u' - Ru^N\Vert_0$	
N	error	rate	error	rate	error	rate
2^8	1.158e-2	0.50	5.889e-4	0.99	3.289e-3	0.98
2^9	8.213e-3	0.50	2.959e-4	1.00	1.673e-3	0.99
2^{10}	5.817e-3	0.50	1.483e-4	1.00	8.435e-4	0.99
2^{11}	4.117e-3	0.50	7.424e-5	1.00	4.236e-4	1.00
2^{12}	2.912e-3	0.50	3.714e-5	1.00	2.123e-4	1.00
2^{13}	2.060e-3	0.50	1.858e-5	1.00	1.062e-4	1.00
2^{14}	1.456e-3	—	9.290e-6	—	5.315e-5	—

Table 5.2 Galerkin FEM on a standard Shishkin mesh

	$\vert\vert\vert u - u^N \vert\vert\vert_\varepsilon$		$\vert\vert\vert u^I - u^N \vert\vert\vert_\varepsilon$		$\varepsilon^{1/2}\Vert u' - Ru^N\Vert_0$	
N	error	rate	error	rate	error	rate
2^8	6.166e-3	0.83	1.624e-4	1.66	8.045e-4	1.64
2^9	3.470e-3	0.85	5.151e-5	1.69	2.585e-4	1.69
2^{10}	1.928e-3	0.86	1.592e-5	1.72	8.025e-5	1.72
2^{11}	1.060e-3	0.87	4.818e-6	1.75	2.432e-5	1.75
2^{12}	5.784e-4	0.88	1.434e-6	1.77	7.242e-6	1.77
2^{13}	3.133e-4	0.89	4.211e-7	1.79	2.126e-6	1.79
2^{14}	1.687e-4	—	1.221e-7	—	6.164e-7	—

Table 5.3 Galerkin FEM on a Bakhvalov-Shishkin mesh

	$\vert\vert\vert u - u^N \vert\vert\vert_\varepsilon$		$\vert\vert\vert u^I - u^N \vert\vert\vert_\varepsilon$		$\varepsilon^{1/2}\Vert u' - Ru^N\Vert_0$	
N	error	rate	error	rate	error	rate
2^8	1.357e-3	1.00	5.382e-6	1.99	4.173e-5	2.00
2^9	6.800e-4	1.00	1.353e-6	2.00	1.043e-5	2.00
2^{10}	3.403e-4	1.00	3.393e-7	2.00	2.610e-6	2.00
2^{11}	1.702e-4	1.00	8.497e-8	2.00	6.528e-7	2.00
2^{12}	8.514e-5	1.00	2.126e-8	2.00	1.632e-7	2.00
2^{13}	4.258e-5	1.00	5.317e-9	2.01	4.082e-8	2.00
2^{14}	2.129e-5	—	1.321e-9	—	1.020e-8	—

5.3 Stabilised FEM

We have seen that the Galerkin FEM on S-type meshes has good approximation properties. Unfortunately the linear systems generated are difficult to solve iteratively. Therefore, stabilisation is essential. We shall restrict ourselves to artificial viscosity stabilisation and to the streamline-diffusion FEM. Other stabilisation techniques, including:

- discontinuous Galerkin FEM (dGFEM),
- continuous interior penalties (CIP) and
- local projection stabilisation (LPFEM),

have been considered in the literature. However, with regard to the classification of layer-adapted meshes these contributions are negligible. Nonetheless, some of the results for these methods will be mentioned in Sect. 9.2.

5.3.1 Artificial Viscosity Stabilisation

The simplest way to stabilise discretisation methods for convection-diffusion problems consists of altering the diffusion coefficient a priori, the extra diffusion added being called **artificial viscosity**. Typically artificial viscosity proportional to the stepsize is used.

Let $\kappa > 0$ be an arbitrary constant. Then our stabilised FEM is: Find $u^N \in V^\omega$ such that

$$a_\kappa(u^N, v) = f(v) \quad \text{for all } \; v \in V^\omega,$$

where

$$a_\kappa(u,v) := \big((\varepsilon + \kappa\hbar)u', v'\big) - (bu' - cu, v)$$

and

$$\hbar(x) \equiv h_i \quad \text{for } \; x \in I_i.$$

The bilinear form $a_\kappa(\cdot,\cdot)$ is coercive with respect to the norm

$$|||v|||_\kappa := \Big\{\big((\varepsilon + \kappa\hbar)v', v'\big) + \gamma\|v\|_0^2\Big\}^{1/2},$$

which is stronger than the ε-weighted energy norm. This is the reason for the improved stability of the method.

Because of the added artificial viscosity, the method does not satisfy the orthogonality property which slightly complicates the convergence analysis. Assume a S-type mesh is used. Let $\eta = u^I - u$ and $\chi = u^I - u^N$ again. Then

$$\begin{aligned} |||\chi|||_\kappa^2 &\le a_\kappa(\chi,\chi) = a(\eta,\chi) + \left(\kappa\hbar(u^I)',\chi'\right) \\ &= a(\eta,\chi) + \kappa\left(\hbar\eta',\chi'\right) + \kappa\left(\hbar u',\chi'\right). \end{aligned}$$

Bounds for the first term have been derived in Sect. 5.2. The second term $(\hbar\eta',\chi')$ vanishes, while the last term, which is the inconsistency of the method, satisfies

$$\kappa\left|(\hbar u',\chi')\right| \le C\kappa\left(h\ln^{1/2}N + N^{-1}\max|\psi'|\right)|||\chi|||_\varepsilon .$$

The proof recycles some ideas from Sect. 5.2.1 and 5.2.2 and is therefore omitted. For more details see also [148]. We get

$$\begin{aligned} \left|\left|\left|u^I - u^N\right|\right|\right|_\kappa \le C\Big\{ & h(h+\kappa)\ln^{1/2}N \\ & + \left(\kappa + N^{-1}\max|\psi'|\right)N^{-1}\max|\psi'|\Big\}. \end{aligned} \tag{5.17}$$

If we choose $\kappa = \mathcal{O}(1)$, i. e., we add artificial viscosity proportional to the local mesh size, we get

$$\left|\left|\left|u^I - u^N\right|\right|\right|_\kappa + \left|\left|\left|u - u^N\right|\right|\right|_\varepsilon \le C\left(h\ln^{1/2}N + N^{-1}\max|\psi'|\right),$$

by the interpolation error estimate (5.8).

Comparing (5.9) and (5.17), we see that the order of accuracy of the Galerkin FEM is not affected if we take $\kappa = \mathcal{O}\left(N^{-1}\right)$. This results in improved stability compared to the Galerkin method and the discrete systems—in particular for higher-dimensional problems—are slightly easier to solve by means of standard iterative methods. On the other hand, the method is not as stable as if $\kappa = \mathcal{O}(1)$ were chosen.

5.3.2 *Streamline-Diffusion Stabilisation*

The most popular and most frequently studied stabilised FEM is the streamline-diffusion finite element method (SDFEM) which is also referred to as the streamline-upwind Petrov-Galerkin method (SUPG). This kind of stabilisation was introduced by Hughes and Brooks [54]. Given a mesh ω and a finite element space V^ω, this method can be written as: Find $u^N \in V^\omega$ such that

$$a(u^N,v) + \sum_{i=1}^N \delta_i \int_{I_i}\left(f - \mathcal{L}u^N\right)bv' = (f,v) \quad \text{for all } v \in V^\omega, \tag{5.18}$$

where the stabilisation parameters δ_i are chosen according to the local mesh Peclét number:

$$\delta_i = \begin{cases} \kappa_0 h_i & \text{if } Pe_i > 1, \\ \kappa_1 h_i^2 \varepsilon^{-1} & \text{if } Pe_i \le 1, \end{cases} \qquad \text{with} \quad Pe_i = \frac{\|b\|_{\infty, I_i} h_i}{2\varepsilon}$$

with user chosen positive constants κ_0 and κ_1. In contrast to the artificial-viscosity stabilisation, this method is consistent with (5.1) since u satisfies (5.18) for all $v \in H_0^1(0,1)$. Another advantage—though it becomes relevant only in higher dimensions—is the reduction of crosswind smear because artificial viscosity is added only in the streamline direction.

The second-order upwind schemes of Sect. 4.3.1 may be regarded as versions of the SDFEM with linear test and trial functions and inexact integration. While in the one-dimensional case it is always possible to choose the stabilisation parameters δ_i such that the resulting scheme is inverse monotone, this is in general impossible in higher dimensions. Therefore, alternative techniques have to be developed to study the SDFEM. Here we shall consider convergence in the streamline-diffusion norm $|||\cdot|||_{SD}$ naturally associated with the bilinear form of the method. This technique can be extended to two-dimensional problems; see Sect. 9.2.4 or [155].

5.3.2.1 Energy-Norm Error Estimates

We study the SDFEM on S-type meshes $\bar{\omega}$. For the sake of simplicity we consider (5.1) with constant b. Let $V_0^\omega \subset H_0^1(0,1)$ be the linear space of piecewise-affine functions on $\bar{\omega}$ that vanish at the boundary. We rewrite (5.18) as: Find $u^N \in V_0^\omega$ such that

$$a_{SD}(u^N, v) := a(u^N, v) + a_{stab}(u^N, v) = f(v) + f_{stab}(v) \quad \text{for all } v \in V_0^\omega$$

where $a(\cdot,\cdot)$ is the bilinear form of the Galerkin FEM,

$$a_{stab}(w, v) := -\delta \sum_{i=qN+1}^{N} \int_{I_i} (-\varepsilon w'' - bw' + cw) bv',$$

$$f_{stab}(v) := -\delta \sum_{i=qN+1}^{N} \int_{I_i} fbv'$$

and

$$\delta = \begin{cases} \kappa_0 H & \text{if } bH/2\varepsilon > 1, \\ \kappa_1 H^2/\varepsilon & \text{otherwise.} \end{cases}$$

Here H denotes again the mesh size on the coarse part of the mesh.

The streamline-diffusion norm naturally associated with $a_{SD}(\cdot,\cdot)$ is

$$|||v|||_{SD}^2 := \varepsilon \left\| v' \right\|_0^2 + \gamma \left\| v \right\|_0^2 + \left\| \delta^{1/2} b v' \right\|_{0,(\tau,1)}^2.$$

Provided the maximum step size h is smaller than some threshold value, the bilinear form $a_{SD}(\cdot,\cdot)$ is coercive with respect to the streamline-diffusion norm:

$$a_{SD}(v,v) \geq \tfrac{1}{2} \, |||v|||_{SD}^2 \quad \text{for all } \ v \in V_0^\omega;$$

see [141]. The bilinear form also satisfies the Galerkin-orthogonality property

$$a_{SD}(u - u^N, v) = 0 \quad \text{for all } \ v \in V_0^\omega.$$

This is the starting point of our error analysis. Let $\eta = u^I - u$ and $\chi = u^I - u^N$. Then

$$\tfrac{1}{2} \, |||\chi|||_{SD}^2 \leq a(\eta,\chi) + a_{stab}(\eta,\chi). \tag{5.19}$$

For the first term we have from the proof of Theorem 5.10

$$|a(\eta,\chi)| \leq C \left(h^2 \ln^{1/2} + N^{-2} \max |\psi'|^2 \right) |||\chi|||_\varepsilon \, .$$

It remains to bound $a_{stab}(\eta,\chi)$. We have

$$a_{stab}(\eta,\chi) = \delta \int_\tau^1 \left(\varepsilon u'' + b\eta' + c\eta \right) b\chi'.$$

Element-wise integration by parts yields $\int_\tau^1 \eta'\chi' = 0$. Furthermore,

$$\begin{aligned} \left| \delta \int_\lambda^1 c\eta b\chi' \right| &\leq C\delta^{1/2} \|\eta\|_{0,(\tau,1)} \|\delta^{1/2} b\chi'\|_0 \\ &\leq C\delta^{1/2} N^{-2} \|\delta^{1/2} b\chi'\|_0 \leq C N^{-2} \, |||\chi|||_{SD} \, , \end{aligned}$$

by the bounds for the interpolation error from Sect. 5.1.

To bound the remaining term $\int_\tau^1 u''\chi'$ we use the decomposition of u into a regular and a layer component; see Theorem 3.48. For the regular component v we have

$$\int_\tau^1 v''\chi' = -\int_\tau^1 v'''\chi - \int_0^\tau v''\chi'.$$

Hence,

$$\left| \int_\tau^1 v''\chi' \right| \leq C\|\chi\|_0 + C \left(\varepsilon \ln N \right)^{1/2} \|\chi'\|_0,$$

by the bounds for the derivatives of v. Thus,

$$\left|\varepsilon\delta\int_\tau^1 v''\chi'\right| \le CN^{-2}\ln^{1/2}N\,|||\chi|||_{SD}\,,$$

since the choice of δ implies $\varepsilon\delta \le CH^2 \le CN^{-2}$.

For the layer component w we estimate as follows:

$$\begin{aligned}\varepsilon\delta\left|\int_\tau^1 w''b\chi'\right| &\le \varepsilon\delta^{1/2}\|w''\|_{1,(\tau,1)}\|\delta^{1/2}b\chi'\|_{\infty,(\tau,1)}\\ &\le C\delta^{1/2}N^{-5/2}H^{-1/2}\|\delta^{1/2}b\chi'\|_{0,(\tau,1)},\end{aligned}$$

by an inverse inequality and because $\sigma \ge 5/2$. We get

$$\varepsilon\delta\left|\int_\tau^1 w''b\chi'\right| \le CN^{-2}\,|||\chi|||_{SD}\,.$$

Collecting these results, the second term in (5.19) is bounded by

$$|a_{stab}(\eta,\chi)| \le CN^{-2}\ln^{1/2}N\,|||\chi|||_\varepsilon\,.$$

We summarise our results.

Theorem 5.16. *Let $\bar\omega$ be an S-type mesh with $\sigma \ge 5/2$ whose mesh generating function $\tilde\varphi$ satisfies* (2.8)*. Then the error of the SDFEM satisfies*

$$\left|\left|\left|u^I - u^N\right|\right|\right|_{SD} \le C\left(h^2\ln^{1/2}N + N^{-2}\max|\psi'|^2\right).$$

Remark 5.17. This is a superconvergence result like Theorem 5.10. Similar to Sect. 5.2.3 it is possible to construct recovery operators to obtain a second-order approximations of the gradient of the exact solution. ♣

5.3.2.2 Maximum-Norm Error Estimates

Chen and Xu [26] consider a modification of (5.18). Find $\tilde u^N \in V_0^\omega$

$$a(\tilde u^N, v) + \sum_{i=1}^N \int_{I_i}\tilde\delta_i\left(f - \mathcal{L}\tilde u^N\right)bv' = (f,v) \quad \text{for all } v \in V_0^\omega,$$

where

$$\tilde\delta_i(x) = \min\left\{\frac{h_i^2}{2\varepsilon}, \frac{h_i}{b}\right\}(x_i - x)(x - x_{i-1}) \quad \text{for } x \in I_i,$$

i.e., $\tilde\delta_i$ is a bubble function on I_i instead of a constant.

For this modified SDFEM it is shown in [26] that if $b = \text{const.}$ and $c \equiv 0$ then

$$\left\| u - \tilde{u}^N \right\|_\infty \le C \min_{v \in V_0^\omega} \|u - v\|_\infty .$$

Thus, the method is quasi-optimal in the maximum norm.

Suppose $b = \text{const.}$ Then

$$\int_{I_i} \tilde{\delta}_i b w' b v' = \delta_i \int_{I_i} b w' b v' \quad \text{for all } \; w, v \in V_0^\omega$$

with

$$\delta_i = \frac{1}{6} \min \left\{ \frac{h_i^2}{2\varepsilon}, \frac{h_i}{b} \right\},$$

i. e., δ_i is the mean value of $\tilde{\delta}_i$ on I_i. Therefore, the modified and the original SDFEM generate the same difference stencil, but different discretisations of the right-hand side. It follows that

$$a_{SD}\left(u^N - \tilde{u}^N, v\right) = \sum_{i=1}^N \int_{I_i} \left(\tilde{\delta}_i - \delta_i\right)\left(f - f^I\right) b v'.$$

Next, the stability of the discretisation, which was established in [26], implies

$$\left\| u^N - \tilde{u}^N \right\|_\infty \le C \left\| f - f^I \right\|_\infty .$$

Finally, the triangle inequality yields for the solution of the original SDFEM

$$\left\| u - u^N \right\|_\infty \le C \left(\min_{v \in V_0^\omega} \|u - v\|_\infty + \left\| f - f^I \right\|_\infty \right).$$

5.4 An Upwind Finite Volume Method

Let us finish this chapter by considering finite volume discretisations of (5.1).

Although the construction of finite volume methods differs from finite difference and finite element methods, they are typically analysed as special finite difference methods or—more often—as nonconforming finite element methods. Here we like to highlight both approaches. In particular, this section is intended to prepare our later investigation of the FVM in two dimensions in Sect. 9.3. There a detailed construction of the method can be found too.

We shall assume

$$c \ge \gamma > 0, \quad c + b' \ge \gamma > 0 \qquad \text{when studying the FVM as a FDM} \qquad (5.20)$$

and

$$c + b'/2 \ge \gamma > 0 \qquad \text{in the FEM context.} \tag{5.21}$$

In the latter case the variational formulation (5.2) will be used.

Given an arbitrary mesh $\bar{\omega}$ our FVM reads: Find $u^N \in I\!R_0^{N+1}$ such that

$$\left[L_\rho u^N\right]_i = f_i \quad \text{for } i = 1, \dots, N-1, \tag{5.22}$$

where

$$[L_\rho v]_i := -\varepsilon v_{\bar{x}\hat{x};i} - \rho\left(-\mu_{i+1/2}\right) b_{i+1/2} v_{\hat{x};i} - \rho\left(\mu_{i-1/2}\right) b_{i-1/2} v_{\check{x};i} + c_i u_i^N$$

with $\mu_{i+1/2} = b_{i+1/2} h_{i+1}/\varepsilon$ and $b_{i+1/2} = b\big(x_{i+1/2}\big)$.

The method can also be written in variational form: Find $u^N \in I\!R_0^{N+1}$ such that

$$a_\rho(u^N, v) := \left(L_\rho u^N, v\right)_\omega = (f, v)_\omega =: f_\rho(v) \quad \text{for all } v \in I\!R_0^{N+1},$$

where

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

The crucial point is the choice of the controlling function $\rho : I\!R \to [0, 1]$. It has to provide the correct weighting between the two one-sided difference approximations for the first-order derivative. Possible choices for ρ include:

$$\rho_I(t) = \begin{cases} \dfrac{1}{t}\left(1 - \dfrac{t}{\exp t - 1}\right) & \text{for } t \ne 0, \\ \dfrac{1}{2} & \text{for } t = 0, \end{cases}$$

$$\rho_S(t) = \begin{cases} 1/(2+t) & \text{for } t \ge 0, \\ (1-t)/(2-t) & \text{for } t < 0, \end{cases}$$

and, with $m \ge 0$,

$$\rho_{U,m}(t) = \begin{cases} 0 & \text{for } t > m, \\ \frac{1}{2} & \text{for } t \in [-m, m], \\ 1 & \text{for } t < -m; \end{cases}$$

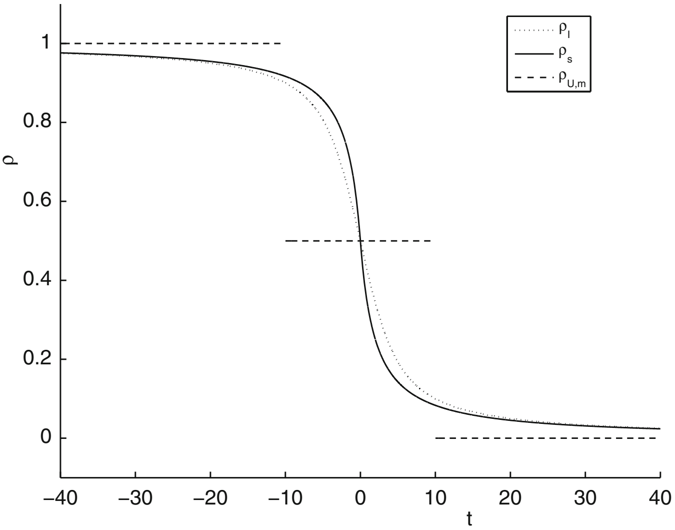

Fig. 5.1 The stabilising functions ρ_I, ρ_S and $\rho_{U,m}$

see Fig. 5.1. The full upwind stabilisation $\rho_{U,0}$ is due to Baba and Tabata [17], while $\rho_{U,m}$ with $m > 0$ was introduced by Angermann [13]. For ρ_I and ρ_S we get slight modifications of the schemes of Il'in [55] and of Samarski [144]. Further choices of ρ are mentioned in [13] and [62] where also a detailed derivation of the method in two dimensions can be found.

The constant choice $\rho \equiv \frac{1}{2}$ generates a central difference scheme, while the choice $\rho_{U,0}$ gives a scheme with upwinded one-sided difference approximation of the first-order derivative which is very similar to the upwind scheme analysed in Sect. 4.2. If a different ρ is used—in particular when ρ is Lipschitz continuous in a neighbourhood of 0—then the first-order derivatives are approximated by weighted combinations of upwinded and downwinded operators. This weighting provides an adaptive transition from an upwinded to a central difference approximation when the local mesh size is small enough. In this case higher accuracy is achieved while retaining the good stability of the scheme.

Important properties of ρ are

$$(\rho_0) \qquad t \mapsto t\rho(t) \text{ is Lipschitz continuous,}$$

$$(\rho_1) \qquad \big[\rho(t) + \rho(-t) - 1\big]t = 0 \quad \text{for all } t \in \mathbb{R},$$

$$(\rho_2) \qquad \big[1/2 - \rho(t)\big]t \geq 0 \quad \text{for all } t \in \mathbb{R},$$

$$(\rho_3) \qquad 1 - t\rho(t) \geq 0 \quad \text{for all } t \in \mathbb{R}.$$

Condition (ρ_1) ensures both the consistency of the scheme and the local conservation of the fluxes, while (ρ_2) guarantees the coercivity of the bilinear form $a_\rho(\cdot,\cdot)$ and (ρ_3) the inverse monotonicity of the scheme.

5.4.1 Stability of the FVM

Coercivity of the bilinear form $a_\rho(\cdot,\cdot)$

The consistency condition (ρ_1) and summation by parts yield

$$\begin{aligned} a_\rho(v,v) = \varepsilon \sum_{i=1}^{N} \frac{(v_i - v_{i-1})^2}{h_i} \\ + \sum_{i=1}^{N} \left[\frac{1}{2} - \rho\left(\mu_{i-1/2}\right)\right] b_{i-1/2}(v_i - v_{i-1})^2 \\ + \sum_{i=1}^{N-1} \left[\hbar_i c_i + \frac{1}{2}\left(b_{i+1/2} - b_{i-1/2}\right)\right] v_i^2. \end{aligned} \tag{5.23}$$

Assume b' is Hölder continuous with coefficient $\alpha \in (0,1]$. Then

$$\left|b_{i+1/2} - b_{i-1/2} - \hbar_i b_i'\right| \le \hbar_i h^\alpha \|b\|_{C^{1,\alpha}[0,1]}\,. \tag{5.24}$$

Thus, if (5.21) is satisfied and if the maximum mesh size h is smaller than some threshold value h^* then

$$\sum_{i=1}^{N-1} \left[\hbar_i c_i + \frac{1}{2}\left(b_{i+1/2} - b_{i-1/2}\right)\right] v_i^2 \ge \frac{\gamma}{2} \sum_{i=1}^{N-1} \hbar_i v_i^2. \tag{5.25}$$

Let

$$|||v|||_\rho^2 := \varepsilon\, |v|_{1,\omega}^2 + |v|_{\rho,\omega}^2 + \frac{\gamma}{2}\, \|v\|_{0,\omega}^2$$

with

$$|v|_{1,\omega}^2 := \sum_{i=1}^{N} \frac{(v_i - v_{i-1})^2}{h_i}, \quad \|v\|_{0,\omega} := \sum_{i=1}^{N-1} \hbar_i v_i^2$$

and

$$|v|_{\rho,\omega}^2 := \sum_{i=1}^{N} \left[\frac{1}{2} - \rho\left(\mu_{i-1/2}\right)\right] b_{i-1/2}(v_i - v_{i-1})^2$$

Note, $|||\cdot|||_\rho$ is a well-defined norm when (ρ_2) is satisfied. The coercivity of the discrete bilinear form $a_\rho(\cdot,\cdot)$ follows from (5.23) and (5.25).

Theorem 5.18. *Assume conditions* (ρ_1), (ρ_2) *and* (5.21) *are satisfied. Let* $b \in C^{1,\alpha}[0,1]$ *with Hölder exponent* $\alpha \in (0,1]$. *Then the bilinear form* $a_\rho(\cdot,\cdot)$ *is coercive with respect to the FV-norm* $|||\cdot|||_\rho$, *i.e.,*

$$a_\rho(v,v) \geq |||v|||_\rho^2 \quad \textit{for all} \;\; v \in I\!R_0^{N+1},$$

provided the maximum mesh size h *is smaller than some threshold value which is independent of the perturbation parameter* ε.

Remark 5.19. Note when $\rho \equiv \frac{1}{2}$ the stabilisation is switched off. Nonetheless Theorem 5.18 states coercivity of the bilinear form with respect to the discrete ε-weighted energy norm $|||v|||_{\varepsilon,\omega}^2 := \varepsilon |v|_{1,\omega}^2 + \frac{\gamma}{2}\|v\|_{0,\omega}^2$. However, in the case $\rho \not\equiv \frac{1}{2}$ the scheme is coercive with respect to a stronger norm which results in enhanced stability of the method. ♣

Inverse monotonicity

Let r^+, r^-, $q \geq \gamma > 0$ and $\chi > 0$ be arbitrary mesh functions with

$$r_i^+ \geq \frac{\beta h_{i+1}}{\alpha\varepsilon} \quad \text{and} \quad 1 \geq r_i^- \geq 0 \quad \text{for } i = 1,\ldots,N-1 \tag{5.26}$$

with a constant $\alpha > 0$. Consider the difference operator

$$[L_\chi v]_i := -\frac{\varepsilon}{\chi_i}\left[1 + r_i^+\right] v_{x;i} + \frac{\varepsilon}{\chi_i}\left[1 - r_i^-\right] v_{\bar{x};i} + q_i v_i. \tag{5.27}$$

We study this more general situation because it will also serve as an auxiliary result in Sect. 9.3.2 when two dimensional problems will be investigated. The FVM (5.22) belongs to this class of schemes provided that (ρ_3) holds.

Clearly, $1+r_i^+ \geq 1$ and $1-r_i^- \geq 0$. Hence, the system matrix associated with L_χ possesses nonnegative offdiagonal entries and is therefore a L_0-matrix. Application of Lemma 3.14 with the test vector $\boldsymbol{e} = \mathbf{1}$ yields the inverse monotonicity of L_χ. In particular, we get the stability inequality

$$\|v\|_{\infty,\bar{\omega}} \leq \|L_\chi v/q\|_{\infty,\omega} \quad \text{for all} \;\; v \in I\!R_0^{N+1}.$$

The inverse monotonicity can be used to study the Green's function associated with L_χ and derive stability inequalities similar to those of Sect. 4.2.1.

The (ℓ_∞,ℓ_1) *stability*

The Green's function $G_{\cdot,j}$ associated with the mesh node x_j satisfies

$$[L_\chi G_{\cdot,j}]_i = \delta_{i,j} \quad \text{for} \;\; i = 1,\ldots,N-1, \;\; G_{0,j} = G_{N,j} = 0,$$

where

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

Given G, we can represent any function $v \in I\!R_0^{N+1}$ as

$$v_i = (G_{i,\cdot}, L_\chi v)_\chi\,, \tag{5.28}$$

with the scalar product

$$(w, v)_\chi := \sum_{j=1}^{N-1} \chi_j w_j v_j\,.$$

Let

$$\hat{G}_{i,j} = \begin{cases} \dfrac{\alpha}{\beta} & \text{for } i = 0, \ldots, j, \\ \dfrac{\alpha}{\beta} \displaystyle\prod_{k=j+1}^{i} \left(1 + \frac{\beta h_k}{\alpha\varepsilon}\right)^{-1} & \text{for } i = j+1, \ldots, N. \end{cases}$$

Lemma 5.20. *Suppose* (5.26) *holds true. Then the Green's function G associated with L_χ satisfies*

$$0 \le G_{i,j} \le \hat{G}_{i,j} \le \alpha/\beta \quad \textit{for} \quad i, j = 0, \ldots, N.$$

Proof. If (5.26) holds then the operator L_χ is inverse monotone and therefore satisfies a discrete comparison principle. The lower bound on G is easily verified using the barrier function $v \equiv 0$.

In order to establish the upper bound, it is sufficient to show that for fixed j we have $G_{0,j} \le \hat{G}_{0,j}$, $G_{N,j} \le \hat{G}_{N,j}$ and

$$\left[L_\chi G_{\cdot,j}\right]_i \le \left[L_\chi \hat{G}_{\cdot,j}\right]_i \quad \text{for } i = 1, \ldots, N-1,$$

(i) First we check the boundary conditions. Clearly $\hat{G}_{i,j} > 0$ for all i, j. Thus,

$$0 = G_{0,j} \le \hat{G}_{0,j} \text{ and } 0 = G_{N,j} \le \hat{G}_{N,j} \text{ for } j = 1, \ldots, N-1.$$

(ii) Next, for $i < j$ we have $\left[L_\chi \hat{G}_{\cdot,j}\right]_i = q_i \hat{G}_{i,j} \ge 0$ since both q and $\hat{G}$ are positive.

(iii) For $i > j$ we have

$$\left[\hat{G}_{\cdot,j}\right]_{x;i} = \frac{\beta}{\alpha\varepsilon + \beta h_{i+1}} \hat{G}_{i,j} \quad \text{and} \quad \left[\hat{G}_{\cdot,j}\right]_{\bar{x};i} = -\frac{\beta}{\alpha\varepsilon} \hat{G}_{i,j}.$$

Thus,

$$\left[L_\chi \hat{G}_{\cdot,j}\right]_i \geq \frac{\varepsilon}{\chi_i}\left\{\frac{\left(1+r_i^+\right)\beta}{\alpha\varepsilon+\beta h_{i+1}} - \frac{\left(1-r_i^-\right)\beta}{\alpha\varepsilon}\right\}\hat{G}_{i,j} \geq 0,$$

by (5.26).

(iv) For $i = j$ a combination of the arguments from *(ii)* and *(iii)* yields

$$\left[L_\chi \hat{G}_{\cdot,j}\right]_j \geq \frac{\varepsilon}{\chi_j}\frac{\left(1+r_i^+\right)\beta}{\alpha\varepsilon+\beta h_{j+1}}\hat{G}_{j,j} \geq \frac{1}{\chi_j} = \left[L_\chi G_{\cdot,j}\right]_j,$$

by (5.26).

This Lemma and (5.28) give the (ℓ_∞, ℓ_1) stability of the method:

Theorem 5.21. *Suppose* (5.26) *holds. Then the operator* L_χ *defined in* (5.27) *satisfies the stability inequality*

$$\|v\|_{\infty,\omega} \leq \frac{\alpha}{\beta}\sum_{i=1}^{N-1}\chi_i\left|\left[L_\chi v\right]_i\right| \quad \textit{for all } v \in I\!R_0^{N+1}.$$

Remark 5.22. An error analysis of the upwind FVM using this (ℓ_∞, ℓ_1) stability can be conducted along the lines of Sect. 4.2.5; see also [90]. ♣

The $(\ell_\infty, w_{-1,\infty})$ *stability*

Now let us restrict our attention to difference operators of the type

$$\left[L_\kappa v\right]_i := -\frac{\varepsilon}{\chi_i}\left[1+\rho_i^+\right]v_{x;i} + \frac{\varepsilon}{\chi_i}\left[1-\rho_i^-\right]v_{\bar{x},i} + c_i v_i$$

with

$$\rho_i^+ := \rho\left(-\frac{b_{i+\kappa}h_{i+1}}{\varepsilon}\right)\frac{b_{i+\kappa}h_{i+1}}{\varepsilon}, \quad \rho_i^- := \rho\left(\frac{b_{i-1+\kappa}h_i}{\varepsilon}\right)\frac{b_{i-1+\kappa}h_i}{\varepsilon}$$

and

$$\kappa \in [0,1], \quad \chi_i = \kappa h_{i+1} + (1-\kappa)h_i \quad \text{and} \quad b_{i+\kappa} = b(x_i + \kappa h_{i+1}).$$

The FVM (5.22) is recovered for $\kappa = 1/2$, while the finite difference scheme of Sect. 4.2 is obtained when $\kappa = 1$ and $\rho = \rho_{U,0}$.

Remark 5.23. Condition (5.26) with $\alpha = \sup_{t<0} 1/\rho(t)$ follows from (ρ_3). ♣

Assuming that (ρ_1) holds, the Green's function G solves for fixed i

$$\left[L_\kappa^* G_{i,\cdot}\right]_j = \delta_{i,j} \qquad \text{for } j = 1,\ldots,N-1, \;\; G_{i,0} = G_{i,N} = 0$$

where L_κ^* is the adjoint operator to L_κ with respect to the scalar product $(\cdot,\cdot)_\chi$. Note (ρ_1) implies $\rho_i^+ + \rho_{i+1}^- = \varepsilon^{-1} b_{i+\kappa} h_{i+1}$. Then it is verified that

$$\left[L_\kappa^* v\right]_j = -\frac{\varepsilon}{\chi_j}\left\{\left[1-\rho_{j+1}^-\right] v_{x;j} - \left[1+\rho_{j-1}^+\right] v_{\hat{x};j}\right\} + \left(c_j + \frac{b_{j+\kappa} - b_{j-1+\kappa}}{\chi_j}\right) v_j.$$

Assume $c + b' \geq \gamma > 0$ and let b' be Hölder continuous with coefficient $\lambda \in (0,1]$. Then

$$c_j + \frac{b_{j+\kappa} - b_{j-1+\kappa}}{\chi_j} \geq 0$$

if the maximum step size h is sufficiently small, independent of ε; cf. (5.24). Proceeding as in Sect. 4.2.1, one can show

$$G_{i,j} \geq G_{i,j-1} \quad \text{for} \quad j = 1,\dots,i \quad \text{and} \quad G_{i,j} \leq G_{i,j-1} \quad \text{for} \quad j = i+1,\dots,N,$$

i.e., $G_{i,\cdot}$ is piecewise monotone.

Theorem 5.24. *Suppose* (ρ_1), (ρ_3) *and* (5.20) *hold true. Assume* $b \in C^{1,\lambda}[0,1]$ *with Hölder exponent* $\lambda \in (0,1]$. *Then the operator* L_κ *satisfies the stability inequality*

$$\|v\|_{\infty,\omega} \leq \frac{2\alpha}{\beta} \min_{C \in I\!R} \left\| \sum_{j=\cdot}^{N-1} \chi_j \left[L_\kappa v\right]_j + C \right\|_{\infty,\omega} \quad \textit{for all } v \in I\!R_0^{N+1},$$

with $\alpha = 1/\inf_{t<0} \rho(t) \leq 2$, *provided the maximum step size* h *is smaller than some threshold value that is independent of the perturbation parameter* ε.

5.4.2 Convergence in the Energy Norm

In this section we study the convergence in the energy norm $|||\cdot|||_\rho$ of the finite volume method on S-type meshes (see Sect. 2.1.3) with $\sigma \geq 2$. The controlling function ρ is assumed to satisfy conditions (ρ_0), (ρ_1) and (ρ_2).

Our analysis starts from coercivity of $a_\rho(\cdot,\cdot)$ (see Theorem 5.18) and follows the standard approach of the Strang Lemma. Let $\eta = u^I - u$ and $\chi = u^I - u^N$, where we use u^N for both the pointwise defined solution of (5.22) and its piecewise-linear extension on the mesh $\bar{\omega}$.

From (5.2) and (5.22) we get

$$a_\rho(\chi,\chi) = a(\eta,\chi) + a_\rho(u^I,\chi) - a(u^I,\chi) + f(\chi) - f_\rho(\chi).$$

Then Theorem 5.18 yields

$$\begin{aligned}|||\chi|||_\rho^2 &\le |a(\eta,\chi)| + |f_\rho(\chi) - f(\chi)| \\ &\quad + \left|r(u^I,\chi) - r_\rho(u^I,\chi)\right| + \left|c(u^I,\chi) - c_\rho(u^I,\chi)\right|\end{aligned} \tag{5.29}$$

with

$$r(u^I,\chi) = \int_0^1 c u^I \chi, \quad r_\rho(u^I,\chi) = \sum_{i=1}^{N-1} \hbar_i c_i u_i \chi_i, \quad c(u^I,\chi) = -\int_0^1 b(u^I)'\chi$$

and

$$\begin{aligned}c_\rho(u^I,\chi) = -\sum_{i=1}^{N-1} \Big\{ &\rho\left(\mu_{i+1/2}\right) b_{i+1/2}\left(u_{i+1} - u_i\right) \\ &+ \rho\left(-\mu_{i-1/2}\right) b_{i-1/2}\left(u_i - u_{i-1}\right)\Big\}\chi_i.\end{aligned}$$

The four terms on the right-hand side of (5.29) will be bounded separately.

(i) The first term has been analysed in Sect. 5.2. We have under the assumptions of Theorem 5.6

$$|a(\eta,\chi)| \le C\left(h + N^{-1}\max|\psi'|\right) |||\chi|||_{\varepsilon,\omega}, \tag{5.30}$$

because the discrete and continuous energy norms are equivalent for functions from V^ω.

(ii) Next we bound the error arising from the discretisation of the right-hand side f. Denoting by φ_i the usual basis functions for linear finite elements, we have

$$\begin{aligned}&\left|\int_{I_i} (f\varphi_i)(x)\mathrm{d}x - \frac{h_i}{2} f_i\right| \\ &\quad = \left|\int_{I_i} \left\{ f_i + \int_{x_i}^x f'(s)\mathrm{d}s\right\} \varphi_i(x)\mathrm{d}x - \frac{h_i}{2} f_i\right| \le \frac{h_i^2}{2}\|f'\|_\infty.\end{aligned}$$

Thus,

$$\begin{aligned}|f(\chi) - f_h(\chi)| &= \left|\sum_{i=1}^{N-1} \chi_i \left\{\int_{x_{i-1}}^{x_{i+1}} \left(f\varphi_i\right)(x)\mathrm{d}x - \hbar_i f_i\right\}\right| \\ &\le \|f'\|_\infty h \sum_{i=1}^{N-1} \hbar_i |\chi_i| \le \|f'\|_\infty h \|\chi\|_{0,\omega}.\end{aligned} \tag{5.31}$$

(iii) The next term in line is $r(u^I,\chi) - r_\rho(u^I,\chi)$. By the definition of $r_\rho(\cdot,\cdot)$ and $r(\cdot,\cdot)$, we have

$$r_\rho(u^I,\chi) - r(u^I,\chi) = \sum_{i=1}^{N-1} s_i\chi_i,$$

where

$$s_i := \int_{x_{i-1}}^{x_{i+1}} \left(cu^I\varphi_i\right)(x)\mathrm{d}x - \hbar_i c_i u_i = \int_{x_{i-1}}^{x_{i+1}} \left[\left(cu^I\right)(x) - c_i u_i\right]\varphi_i(x)\mathrm{d}x.$$

We have

$$\begin{aligned}&\left|\left(cu^I\right)(x) - c_i u_i\right| \\ &\qquad \le \left|\int_{x_i}^{x} \left(cu^I\right)'\mathrm{d}s\right| \le C\int_{x_{i-1}}^{x_{i+1}}\left\{1+\varepsilon^{-1}\mathrm{e}^{-\beta s/\varepsilon}\right\}\mathrm{d}s \le C\vartheta_{cd}^{[1]}(\bar\omega)\end{aligned}$$

for $x\in[x_{i-1},x_{i+1}]$. (The quantity $\vartheta_{cd}^{[p]}(\bar\omega)$ has been introduced in Sect. 2.1.) Hence, $|s_i| \le C\vartheta_{cd}^{[1]}(\bar\omega)\hbar_i$ and we obtain

$$\left|r_\rho(u^I,\chi)_i - r(u^I,\chi)\right| \le C\vartheta_{cd}^{[1]}(\bar\omega)\|\chi\|_{0,\omega}. \tag{5.32}$$

(iv) Finally consider the convection term. We have

$$\begin{aligned}&c_\rho(u^I,\chi) - c(u^I,\chi) \\ &\quad = \sum_{i=1}^{N}\left\{\int_{I_i}\left(b(u^I)'\chi\right)(x)\mathrm{d}x\right. \\ &\qquad\qquad \left. - \left[\rho\left(-\mu_{i-1/2}\right)\chi_{i-1} + \rho\left(\mu_{i-1/2}\right)\chi_i\right] b_{i-1/2}\left(u_i - u_{i-1}\right)\right\}\end{aligned}$$

and

$$\begin{aligned}\int_{I_i}\left(b(u^I)'\chi\right)(x)\mathrm{d}x &= b_{i-1/2}\left(u_i - u_{i-1}\right)\frac{\chi_i+\chi_{i-1}}{2} \\ &\qquad + \int_{I_i}\left\{\int_{x_{i-1/2}}^{x} b'(s)\mathrm{d}s\,\frac{u_i-u_{i-1}}{h_i}\chi(x)\right\}\mathrm{d}x.\end{aligned}$$

Combine these two equations and use (ρ_1) to get

$$\begin{aligned} c_\rho(u^I,\chi) - c(u^I,\chi) &= \sum_{i=1}^{N} \left[\frac{1}{2} - \rho\left(\mu_{i-1/2}\right)\right] \left(\chi_i - \chi_{i-1}\right)\left(u_i - u_{i-1}\right) b_{i-1/2} \\ &\quad + \sum_{i=1}^{N} \int_{I_i} \left\{ \int_{x_{i-1/2}}^{x} b'(s)\mathrm{d}s \frac{u_i - u_{i-1}}{h_i} \chi(x) \right\} \mathrm{d}x. \end{aligned} \tag{5.33}$$

Note that $|u_i - u_{i-1}| \le C\vartheta_{cd}^{[1]}(\bar\omega)$. Thus, for the second term in (5.33) one has

$$\left| \sum_{i=1}^{N} \int_{I_i} \left\{ \int_{x_{i-1/2}}^{x} b'(s)\mathrm{d}s \frac{u_i - u_{i-1}}{h_i} \chi(x) \right\} \mathrm{d}x \right| \le C\vartheta_{cd}^{[1]}(\bar\omega) \left\|\chi\right\|_{0,\omega}. \tag{5.34}$$

Next we bound the first sum in (5.33). For $i \le qN$ use $|u_i - u_{i-1}| \le C\vartheta_{cd}^{[1]}(\bar\omega)$ again to obtain

$$\begin{aligned} &\left| \sum_{i=1}^{qN} \left[\frac{1}{2} - \rho\left(\mu_{i-1/2}\right)\right] \left(\chi_i - \chi_{i-1}\right)\left(u_i - u_{i-1}\right) b_{i-1/2} \right| \\ &\qquad \le C\vartheta_1(\omega) \sum_{i=1}^{qN} |\chi_i - \chi_{i-1}| \le C\vartheta_1(\omega)\varepsilon^{1/2} \ln^{1/2} N \left|\chi\right|_{1,\omega}, \end{aligned} \tag{5.35}$$

by a discrete Cauchy-Schwarz inequality.

For $i > qN$ we use the splitting $u = v + w$ of the exact solution according to Theorem 3.48. Starting with the layer term w, we have $w_i \le CN^{-2}$ for $i \ge qN$. Hence,

$$\begin{aligned} &\left| \sum_{i=qN+1}^{N} \left[\frac{1}{2} - \rho\left(\mu_{i-1/2}\right)\right] \left(\chi_i - \chi_{i-1}\right)\left(w_i - w_{i-1}\right) b_{i-1/2} \right| \\ &\qquad \le CN^{-2} \sum_{i=qN+1}^{N} \left(|\chi_i| + |\chi_{i-1}|\right) \le CN^{-1}\|\chi\|_{0,\omega}, \end{aligned} \tag{5.36}$$

by a discrete Cauchy-Schwarz inequality and an inverse inequality that can be used because $N^{-1} \le h_i$ for $i > qN$. Finally, consider the regular solution component v. To simplify the notation let

$$\gamma_{i-1/2} := b_{i-1/2} \left[\frac{1}{2} - \rho\left(\frac{b_{i-1/2}h_i}{\varepsilon}\right)\right].$$

Summation by parts yields

$$\sum_{i=qN+1}^{N} \gamma_{i-1/2}\big(v_i - v_{i-1}\big)\big(\chi_i - \chi_{i-1}\big)$$
$$= \gamma_{qN+1/2}\left(v_{qN} - v_{qN+1}\right)\chi_{qN}$$
$$- \sum_{i=qN+1}^{N-1} \gamma_{i-1/2}\left(v_{i+1} - 2v_i + v_{i-1}\right)\chi_i$$
$$+ \sum_{i=qN+1}^{N-1} \big(\gamma_{i-1/2} - \gamma_{i+1/2}\big)\left(v_{i+1} - v_i\right)\chi_i.$$

Taylor expansions give $|v_{i+1} - 2v_i + v_{i-1}| \le CN^{-2}$ and $|v_i - v_{i-1}| \le CN^{-1}$, while (ρ_0) implies $\left|\gamma_{i-1/2} - \gamma_{i+1/2}\right| \le CN^{-1}$. Thus,

$$\left|\sum_{i=qN+1}^{N} \gamma_{i-1/2}\left(v_i - v_{i-1}\right)\left(\chi_i - \chi_{i-1}\right)\right|$$
$$\le CN^{-1}\Big(\|\chi\|_{0,\omega} + |\chi_{qN}|\Big) \le CN^{-1}\ln^{1/2} N\, |||\chi|||_{\rho},$$

because

$$|\chi_{qN}| \le \sum_{i=1}^{qN} |\chi_i - \chi_{i-1}| \le C\ln^{1/2} N\varepsilon^{1/2}\, |\chi|_{1,\omega}\,.$$

Collecting (5.33)–(5.37), we get

$$\left|c_\rho(u^I,\chi) - c(u^I,\chi)\right| \le C\vartheta_1(\omega)\ln^{1/2} N\, |||\chi|||_{\rho}\,. \tag{5.37}$$

Now all terms on the right-hand side of (5.29) have been bounded; see (5.30), (5.32), (5.31), (5.34) and (5.37). Divide by $|||\chi|||_{\rho}$. Then recall the interpolation error bounds of Sect. 5.1 and note that $\vartheta_{cd}^{[1]}(\bar\omega) \le C\left(h + \max|\psi'|\right)$ for S-type meshes with $\sigma \ge 1$. We get the main result of this section.

Theorem 5.25. *Let $\bar\omega$ be an S-type mesh with $\sigma \ge 2$ whose mesh generating function $\tilde\varphi$ satisfies* (2.8) *and* $\max|\psi'|\ln^{1/2} N \le CN$*. Assume* (ρ_0)*,* (ρ_1) *and* (ρ_2) *hold. Then*

$$\left|\left|\left|u - u^N\right|\right|\right|_{\varepsilon} + \left|\left|\left|u^I - u^N\right|\right|\right|_{\rho} \le C\left(h + N^{-1}\max|\psi'|\right)\ln^{1/2} N$$

for the error of the upwind FVM (5.22)*.*

5.4.3 Convergence in the Maximum Norm

With the results of Sect. 5.4.1 at hand, the simplest maximum-norm analysis is based on the $(\ell_\infty, w_{-1,\infty})$ stability. Set

$$[A_\rho v]_i := \varepsilon\Big\{1 + \mu_{i-1/2}\big[\rho\left(-\mu_{i-1/2}\right) - \rho\left(\mu_{i-1/2}\right)\big]\Big\}\frac{v_i - v_{i-1}}{h_i} + b_{i-1/2}\frac{v_i + v_{i-1}}{2} + \sum_{j=i}^{N-1}\left(\hbar_j c_j + b_{j+1/2} - b_{j-1/2}\right) v_j .$$

If condition (ρ_1) is satisfied then $L_\rho v = -(A_\rho v)_{\hat{x}}$ and Theorem 5.24 yields

$$\|v\|_{\infty,\omega} \le \frac{4}{\beta}\min_{a\in\mathbb{R}}\|A_\rho v + a\|_{\infty,\omega} \quad \text{for all } v \in \mathbb{R}_0^{N+1}.$$

Integrate (5.1) to see that

$$\varepsilon u'_{i-1/2} + (bu)_{i-1/2} + \int_{x_{i-1/2}}^{x_{N-1/2}} \left((c+b')\,u - f\right)(s)\mathrm{d}s \equiv \alpha \quad \text{for all} \quad i = 1, \dots, N.$$

Thus,

$$\left\|u - u^N\right\|_{\infty,\omega} \le \frac{4}{\beta}\max_{i=1,\dots,N}|M_i|\,, \tag{5.38a}$$

where

$$\begin{aligned} M_i := {} & \varepsilon\left(\frac{u_i - u_{i-1}}{h_i} - u'_{i-1/2}\right) + b_{i-1/2}\left(\frac{u_i + u_{i-1}}{2} - u_{i-1/2}\right) \\ & + \sum_{j=i}^{N-1}\left(\hbar_j c_j + b_{j+1/2} - b_{j-1/2}\right) u_j - \int_{x_{i-1/2}}^{x_{N-1/2}} (c+b')\,(s)u(s)\mathrm{d}s \\ & - \sum_{j=i}^{N-1}\hbar_j f_j + \int_{x_{i-1/2}}^{x_{N-1/2}} f(s)\mathrm{d}s \\ & + b_{i-1/2}\left[\rho\left(-\frac{b_{i-1/2}h_i}{\varepsilon}\right) - \rho\left(\frac{b_{i-1/2}h_i}{\varepsilon}\right)\right](u_i - u_{i-1})\,. \end{aligned} \tag{5.38b}$$

All terms except for the last one can be bounded by $\vartheta_{cd}^{[1]}(\bar{\omega})$ using the technique from Sect. 4.2.2. When bounding the last term note that $\rho(t) \in [0,1]$ for all $t \in \mathbb{R}$ and that

$$|u_i - u_{i-1}| \le \int_{I_i}|u'(s)|\,\mathrm{d}s \le C\vartheta_{cd}^{[1]}(\bar{\omega}).$$

Theorem 5.26. *Suppose* (ρ_1) *and* (ρ_3) *hold. Then the error of the upwind finite volume method* (5.22) *satisfies*

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

It was mentioned earlier that the accuracy of the scheme is improved when the function ρ is Lipschitz continuous in a neighbourhood of $t = 0$, say on an interval $[-m, m]$. We will briefly illustrate this using a standard Shishkin mesh with mesh parameter $\sigma \ge 2$.

We work from (5.38). The arguments from Sect. 4.3.3 are used to bound the first four terms by $\vartheta_{cd}^{[2]}(\bar{\omega})^2 \le CN^{-2}\ln^2 N$.

We are left with the last term in (5.38). On a Shishkin mesh with $\sigma \ge 1$:

$$|u_i - u_{i-1}| \le \begin{cases} CN^{-1}\ln N & \text{for } i = 1, \dots, qN, \\ CN^{-1} & \text{for } i = qN+1, \dots, N. \end{cases}$$

This can be verified using the decomposition of Theorem 3.48.

Furthermore, if $h_i \le m\varepsilon\beta^{-1}$ then

$$\left|\rho\left(-\frac{b_{i-1/2}h_i}{\varepsilon}\right) - \rho\left(\frac{b_{i-1/2}h_i}{\varepsilon}\right)\right| \le C\frac{h_i}{\varepsilon},$$

because ρ is Lipschitz continuous on $[-m, m]$. Hence, in the layer region of the Shishkin mesh we have

$$\left|\rho\left(-\frac{b_{i-1/2}h_i}{\varepsilon}\right) - \rho\left(\frac{b_{i-1/2}h_i}{\varepsilon}\right)\right| \le CN^{-1}\ln N \quad \text{for } \ i = 1, \dots, qN,$$

while on the coarse mesh region $\rho(t) \in [0, 1]$ for all $t \in \mathbb{R}$ is used. We obtain

$$\left|\rho\left(-\frac{b_{i-1/2}h_i}{\varepsilon}\right) - \rho\left(\frac{b_{i-1/2}h_i}{\varepsilon}\right)\right| |u_i - u_{i-1}| \le CN^{-1}.$$

All terms in (5.38) have been bounded by either CN^{-1} or by $CN^{-2}\ln^2 N$.

Theorem 5.27. *Assume* ρ *is Lipschitz continuous on* $[-m, m]$. *Let* (ρ_1) *and* (ρ_3) *hold. Then the error of the upwind FVM* (5.22) *on a standard Shishkin mesh satisfies*

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

if N *is larger than some threshold value which is independent of the perturbation parameter* ε.

Remark 5.28. On a standard Shishkin mesh the use of a Lipschitz continuous function ρ improves the accuracy from $\mathcal{O}\left(N^{-1}\ln N\right)$ to $\mathcal{O}\left(N^{-1}\right)$. ♣

Table 5.4 The upwind FVM on a standard Shishkin mesh

	$\rho_{U,0}$		$\rho_{U,10}$		ρ_S		ρ_I	
N	error	rate	error	rate	error	rate	error	rate
2^7	4.236e-3	0.84	3.855e-3	0.99	3.855e-3	0.99	3.855e-3	0.99
2^8	2.364e-3	0.86	1.942e-3	0.99	1.942e-3	0.99	1.942e-3	0.99
2^9	1.303e-3	0.87	9.745e-4	1.00	9.745e-4	1.00	9.745e-4	1.00
2^{10}	7.111e-4	0.89	4.882e-4	1.00	4.882e-4	1.00	4.882e-4	1.00
2^{11}	3.850e-4	0.89	2.443e-4	1.00	2.443e-4	1.00	2.443e-4	1.00
2^{12}	2.071e-4	0.90	1.222e-4	1.00	1.222e-4	1.00	1.222e-4	1.00
2^{13}	1.108e-4	0.91	6.115e-5	1.00	6.115e-5	1.00	6.115e-5	1.00
2^{14}	5.904e-5	0.91	3.059e-5	1.00	3.059e-5	1.00	3.059e-5	1.00
2^{15}	3.133e-5	0.92	1.531e-5	1.00	1.531e-5	1.00	1.531e-5	1.00
2^{16}	1.658e-5	—	7.672e-6	—	7.672e-6	—	7.672e-6	—

5.4.4 A Numerical Example

Table 5.4 displays the results of test computations using the upwind FVM with various stabilising functions ρ, when applied to the test problem (4.14) and contains the maximum nodal errors. For our tests we have chosen a standard Shishkin mesh with $\sigma = 1$ and $q = 1/2$. The results of the numerical tests are in agreement with Theorem 5.26 and 5.27. Comparing the numbers for $\rho_{U,0}$ with those for other ρs, we clearly see an improvement in the accuracy when ρ is Lipschitz continuous in a neighbourhood of $t = 0$. Also notice there is no (visible) difference in using either of those Lipschitz continuous ρs.