Chapter 9
Convection-Diffusion Problems

This chapter is devoted to numerical methods for the convection-diffusion problem

$$-\varepsilon \Delta u - \boldsymbol{b} \nabla u + cu = f \ \text{ in } \ \Omega = (0,1)^2, \ \ u|_{\partial\Omega} = 0, \tag{9.1}$$

with $b_1 \geq \beta_1 > 0$, $b_2 \geq \beta_2 > 0$ on $[0,1]^2$, i.e., problems with regular boundary layers at the outflow boundary $x = 0$ and $y = 0$. The analytical behaviour of the solution of (9.1) was studied in Sect. 7.3.1.

Results for problems with characteristic layers will only be mentioned briefly.

9.1 Upwind Difference Schemes

We shall consider discretisations of (9.1) on a tensor product mesh $\bar{\omega} = \bar{\omega}_x \times \bar{\omega}_y$ with N mesh intervals in both coordinate directions.

The simple upwind scheme for (9.1) is: Find $u^N \in I\!R_0^{(N+1)^2}$ such that

$$\left[Lu^N\right]_{ij} = f_{ij} \quad \text{for } \ i,j = 1,\ldots,N-1 \tag{9.2}$$

with

$$\left[Lu^N\right]_{ij} := -\varepsilon \left(u^N_{\bar{x}\hat{x};ij} + u^N_{\bar{y}\hat{y};ij}\right) - b_{1;ij} u^N_{x;ij} - b_{2;ij} u^N_{y;ij} + c_{ij} u^N_{ij},$$

and

$$v_{x;ij} = \frac{v_{i+1,j} - v_i}{h_{i+1}}, \quad v_{\bar{x};ij} = \frac{v_{ij} - v_{i-1,j}}{h_i} \quad \text{and} \quad v_{\hat{x};ij} = \frac{v_{i+1,j} - v_{ij}}{\hbar_i},$$

$\hbar_i = (h_i + h_{i+1})/2$ and analogous definitions for $v_{y;ij}$, $v_{\bar{y};ij}$, $v_{\hat{y};ij}$ and $\hbar_j$.

This scheme on layer-adapted meshes was first studied by Shishkin who established the maximum-norm error estimate

$$\|u - u^N\|_{\infty,\omega} \leq CN^{-1} \ln^2 N$$

T. Linß, *Layer-Adapted Meshes for Reaction-Convection-Diffusion Problems*, Lecture Notes in Mathematics 1985, DOI 10.1007/978-3-642-05134-0_9,

on Shishkin meshes; see [121]. He also proved [151, §3, Theorem 2.3]

$$\|u - u^N\|_{\infty,\omega} \le C\big(N^{-1}\ln^2 N\big)^p$$

with $p = 1/4$ and $p = 1/8$ (depending on the precise assumptions on the data) if the solution is less smooth.

Here we shall present the technique from [107] which gives a sharper error estimate. This technique is an extension of the truncation error and barrier function technique from Sect. 4.2.6 to two dimensions.

9.1.1 Stability

The matrix associated with L is an L_0-matrix. Application of the M-criterion (Lemma 3.14) with the test function $v_{ij} = (1 - x_i)/\beta_1$ establishes the inverse monotonicity of L. As a consequence, we have the $(\ell_\infty, \ell_\infty)$-stability inequality

$$\|v\|_{\infty,\bar\omega} \le \min\left\{ \left\|\frac{Lv}{b_1}\right\|_{\infty,\omega}, \left\|\frac{Lv}{b_2}\right\|_{\infty,\omega}, \left\|\frac{Lv}{c}\right\|_{\infty,\omega} \right\} \quad \text{for all } v \in \big(\mathbb{R}_0^{N+1}\big)^2,$$

by Lemma 3.17.

Because of the inverse monotonicity we also have the following discrete comparison principle. For any two mesh functions $v, w \in \big(\mathbb{R}^{N+1}\big)^2$

$$\left.\begin{array}{c} Lv \le Lw \quad \text{in } \omega \text{ and} \\ v \le w \quad \text{on } \partial\omega \end{array}\right\} \quad \Longrightarrow \quad v \le w \quad \text{on } \bar\omega.$$

It will be used repeatedly in the convergence analysis.

9.1.2 Pointwise Error Bounds

Consider the upwind scheme (9.2) on a tensor-product Shishkin-type mesh $\bar\omega = \bar\omega_x \times \bar\omega_y$ where the two one-dimensional meshes in x- and y-direction are constructed as described in Sect. 2.1.3. Give mesh parameters $\sigma > 0$ and $q > 0$ the mesh uses transition points in the S-type mesh are

$$\tau_x := \min\left\{q, \frac{\sigma\varepsilon}{\beta_1}\ln N\right\} \quad \text{and} \quad \tau_y := \min\left\{q, \frac{\sigma\varepsilon}{\beta_2}\ln N\right\}.$$

Fig. 2.19 displays a plot of the resulting mesh.

Theorem 9.1. *Assume the solution u of (9.1) can be decomposed as in Theorem 7.17 with $\alpha = 1$ and $n = 3$. Let the mesh be a tensor-product S-type mesh with $\sigma \geq 2$. Suppose the mesh generating function $\tilde{\varphi}$ satisfies (2.8) and (2.14). Then the error of the simple upwind scheme (9.2) satisfies*

$$\left|u_{ij} - u_{ij}^N\right| \leq \begin{cases} C\left(h + N^{-1}\right) & \text{for } i,j = qN,\ldots,N, \\ C\left(h + N^{-1}\max|\psi'|\right) & \text{otherwise.} \end{cases}$$

Proof. We adapt the truncation error and barrier function technique of Sect. 4.2.6 to two space dimensions.

Recalling the decomposition of Theorem 7.17, we split the numerical solution in a similar manner:

$$u^N = v^N + w_1^N + w_2^N + w_{12}^N,$$

where

$$Lv^N = \mathcal{L}v, \quad Lw_1^N = \mathcal{L}w_1, \quad Lw_2^N = \mathcal{L}w_2, \quad Lw_{12}^N = \mathcal{L}w_{12} \quad \text{on } \omega,$$

and

$$v^N = v, \quad w_1^N = w_1, \quad w_2^N = w_2, \quad w_{12}^N = w_{12} \quad \text{on } \partial\omega.$$

For the regular solution component a Taylor expansion, the derivative bounds of Theorem 7.17 and the inverse monotonicity of L give

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

For the term representing the layer at $x = 0$, we have, similarly to (4.44),

$$0 \leq w_{1;ij}^N \leq \bar{w}_{1;i}^N := C\prod_{k=1}^{i}\left(1 + \frac{\beta_1 h_k}{2\varepsilon}\right)^{-1} \quad \text{for } i,j = 0,\ldots,N.$$

Thus

$$\left|w_{1;ij} - w_{1;ij}^N\right| \leq CN^{-1} \quad \text{for } i = qN,\ldots,N,\ j = 0,\ldots,N;$$

see the argument that led to (4.45). Now let $i < qN$. Taylor expansions and Theorem 7.17 give

$$\begin{aligned} \left|L(w_1 - w_1^N)_{ij}\right| &\leq C\left(h + \varepsilon^{-2}\left(h_i + h_{i+1}\right)\mathrm{e}^{-\beta_1 x_{i-1}/\varepsilon}\right) \\ &\leq C\left(h + \varepsilon^{-1}\bar{w}_{1;i}^N N^{-1}\max|\psi'|\right). \end{aligned}$$

Application of the comparison principle with the barrier function

$$C\left(N^{-1} + h + \bar{w}^N_{1;i} N^{-1} \max|\psi'|\right)$$

with C sufficiently large yields

$$\left|w_{1;ij} - w^N_{1;ij}\right| \le C\left(h + N^{-1}\max|\psi'|\right) \quad \text{for } i = 0,\dots,qN-1,\ j = 0,\dots,N.$$

For the boundary layer at $y = 0$, the same type of argument is used in order to obtain

$$\left|w_{2;ij} - w^N_{2;ij}\right| \le CN^{-1} \quad \text{for } i = 0,\dots,N,\ j = qN,\dots,N$$

and

$$\left|w_{2;ij} - w^N_{2;ij}\right| \le C\left(h + N^{-1}\max|\psi'|\right) \quad \text{for } i = 0,\dots,N,\ j = 0,\dots,qN-1.$$

Finally, for the corner layer term one first shows

$$\left|w_{12;ij} - w^N_{12;ij}\right| \le \bar{w}^N_{12;ij} := C\prod_{k=1}^{i}\left(1 + \frac{\beta_1 h_k}{2\varepsilon}\right)^{-1}\prod_{l=1}^{j}\left(1 + \frac{\beta_2 k_l}{2\varepsilon}\right)^{-1} \quad \text{for } i,j = 0,\dots,N,$$

which implies

$$\left|w_{12;ij} - w^N_{12;ij}\right| \le CN^{-1} \quad \text{if } i \ge qN \text{ or } j \ge qN.$$

In a second step the truncation error is estimated using Taylor expansions:

$$\left|L(w_{12} - w^N_{12})_{ij}\right| \le C\varepsilon^{-1}\bar{w}^N_{12;ij} N^{-1}\max|\psi'|.$$

And the discrete comparison principle yields

$$\left|w_{12;ij} - w^N_{12;ij}\right| \le CN^{-1}\max|\psi'| \quad \text{for } i,j = 0,\dots,qN-1.$$

Collecting the bounds for the various components, we are finished. □

Remark 9.2. We are not aware of any a priori error estimates for arbitrary meshes similar to those of Sect. 4.2.2 for one-dimensional problems. This seems to be due to a lack of strong negative-norms stability inequalities. ♣

Remark 9.3. In [107] a modified, hybrid scheme on a standard Shishkin mesh is considered. It is based on simple upwinding, but employs central differencing whenever the mesh allows one to do this without losing stability. For this scheme the above technique gives the maximum-norm error bound

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

The improved bound is because central differencing improves the error terms of order $N^{-1}\max|\psi'|$ in the above proof to order $N^{-2}\max|\psi'|^2$. ♣

A numerical example

We briefly illustrate our theoretical findings for the simple upwind difference scheme on S-type meshes and for the hybrid scheme when applied to the test problem

$$-\varepsilon\Delta u - (2+x)u_x - (3+y^2)u_y + u = f \ \text{ in } \ \Omega = (0,1)^2, \tag{9.3a}$$

$$u = 0 \ \text{ on } \ \Gamma = \partial\Omega, \tag{9.3b}$$

where the right-hand side is chosen such that

$$u(x,y) = \cos\frac{\pi x}{2}\left(1 - \mathrm{e}^{-2x/\varepsilon}\right)(1-y)^3\left(1 - \mathrm{e}^{-3y/\varepsilon}\right) \tag{9.3c}$$

is the exact solution. This function exhibits typical boundary layer behaviour. For our tests we take $\varepsilon = 10^{-8}$, which is a sufficiently small choice to bring out the singularly perturbed nature of the problem. Table 9.1 displays the results of our text computations. They are in agreement with the theoretical findings.

Table 9.1 Upwind and hybrid difference scheme on S-type meshes

	simple upwinding						hybrid scheme	
	standard Shishkin mesh		Shishkin mesh with 2 transition points		Bakhvalov-Shishkin mesh		standard Shishkin mesh	
N	error	rate	error	rate	error	rate	error	rate
16	9.6379e-2	0.50	9.0430e-2	0.73	9.3261e-2	0.74	1.1072e-1	0.88
32	6.8194e-2	0.59	5.4533e-2	0.76	5.5803e-2	0.90	5.9962e-2	0.94
64	4.5364e-2	0.66	3.2138e-2	0.79	2.9916e-2	0.93	3.1328e-2	0.97
128	2.8636e-2	0.72	1.8606e-2	0.84	1.5665e-2	0.97	1.6031e-2	0.98
256	1.7360e-2	0.77	1.0416e-2	0.87	8.0140e-3	0.98	8.1081e-3	0.99
512	1.0182e-2	0.80	5.6941e-3	0.90	4.0529e-3	0.99	4.0768e-3	1.00
1024	5.8286e-3	0.83	3.0602e-3	0.91	2.0379e-3	1.00	2.0440e-3	1.00
2048	3.2776e-3	—	1.6247e-3	—	1.0219e-3	—	1.0234e-3	—

9.1.3 Error Expansion

Kopteva [65] derives an error expansion for the simple upwind scheme (9.2) on standard Shishkin meshes. Let h_c and h_f denote the coarse and fine mesh sizes in the Shishkin mesh. Provided that $\varepsilon \le CN^{-1}$, she proves that the error can be expanded as

$$u_{ij}^N - u_{ij} = h_c \Phi_{ij} + \frac{h_f}{\varepsilon} \Psi_{ij} + R_{ij}$$

with

$$\begin{aligned} \Phi(x,y) = \varphi(x,y) &+ \varphi(0,0) \exp\left(-\frac{b_1(0,0)x + b_2(0,0)y}{\varepsilon}\right) \\ &- \varphi(0,y) \exp\left(-\frac{b_1(0,y)x}{\varepsilon}\right) - \varphi(x,0) \exp\left(-\frac{b_2(x,0)y}{\varepsilon}\right) \end{aligned}$$

and

$$\Psi(x,y) = \frac{x}{\varepsilon} \frac{b_1^2(0,y)\tilde{w}_1 + b_1^2(0,0)\tilde{w}_{12}}{2} + \frac{y}{\varepsilon} \frac{b_2^2(x,0)\tilde{w}_2 + b_2^2(0,0)\tilde{w}_{12}}{2},$$

where the $\tilde{w}$'s satisfy bounds similar to those of Theorem 7.17 and $\|\varphi\|_{C^{1,1}} \le C$, while for the remainder we have

$$R_{ij} \le \begin{cases} CN^{-2} & \text{for } i,j = qN,\ldots,N, \\ CN^{-2}\ln^2 N & \text{otherwise.} \end{cases}$$

This expansion is used in [65] to derive error bounds for Richardson extrapolation and for the approximation of derivatives.

Richardson extrapolation

Let $\tilde{u}^N$ be the upwind difference solution on the mesh obtained by uniformly bisecting the original mesh $\bar{\omega}$ and let $\Pi\tilde{u}^N$ be the obvious restriction of $\tilde{u}^N$ to $\bar{\omega}$. Then

$$\left| \left(\left[2\Pi\tilde{u}^N - u^N \right] - u \right)_{ij} \right| \le C \begin{cases} N^{-2} & \text{for } i,j = qN,\ldots,N-1, \\ N^{-2}\ln^2 N & \text{otherwise [65].} \end{cases}$$

These results are neatly illustrated by the numbers in Table 9.2 which display the results of Richardson extrapolation applied to our test problem (9.3).

Table 9.2 Richardson extrapolation on a Shishkin mesh

	fine mesh region		coarse mesh region	
N	error	rate	error	rate
16	1.3869e-2	1.08	3.7171e-3	1.44
32	6.5448e-3	1.23	1.3733e-3	1.74
64	2.7918e-3	1.38	4.1086e-4	1.87
128	1.0703e-3	1.49	1.1271e-4	1.93
256	3.8049e-4	1.58	2.9616e-5	1.96
512	1.2701e-4	1.64	7.5975e-6	1.98
1024	4.0623e-5	—	1.9234e-6	—

Derivative approximation

In [65] the bounds

$$\left|(u^N-u)_{x;ij}\right| \leq C \begin{cases} N^{-1} & \text{for } i,j = qN,\dots,N-1, \\ N^{-1}\ln^2 N & \text{for } i = qN,\dots,N-1,\ j = 0,\dots,qN-1, \\ \varepsilon^{-1}N^{-1}\ln N & \text{otherwise} \end{cases}$$

are given with analogous results for $(u^N-u)_y$.

9.2 Finite Element Methods

This section is concerned with finite element discretisations of (9.1).

The variational formulation of (9.1) is as follows. Find $u \in H_0^1(\Omega)$ such that

$$a(u,v) = f(v) \quad \text{for all } v \in H_0^1(\Omega), \tag{9.4}$$

where

$$a(u,v) = \varepsilon(\nabla u, \nabla v) - (\boldsymbol{b}\cdot\nabla u, v) + (cu, v) \quad \text{and} \quad f(v) = (f,v)$$

with

$$(u,v) := \int_\Omega u(x,y)v(x,y)\mathrm{d}x\mathrm{d}y.$$

If

$$c + \tfrac{1}{2}\operatorname{div}\boldsymbol{b} \geq \gamma > 0 \tag{9.5}$$

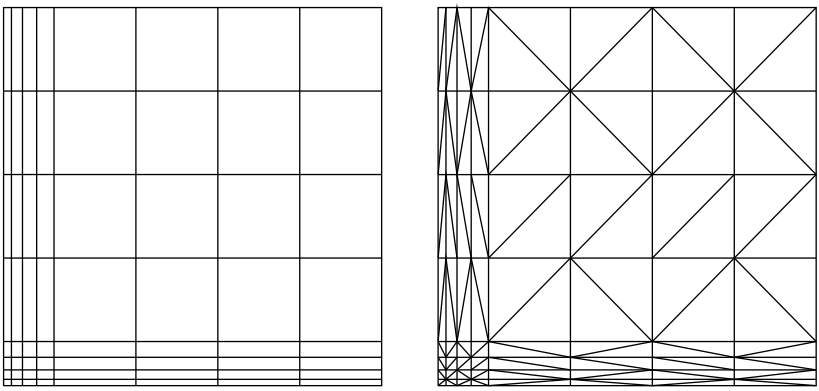


Fig. 9.1 Triangulations into rectangles and triangles on tensor-product layer-adapted meshes

then the bilinear form $a(\cdot,\cdot)$ is coercive, i.e.,

$$a(v,v) \geq |||v|||_\varepsilon^2 := \varepsilon \left(\|\partial_x v\|_0^2 + \|\partial_y v\|_0^2 \right) + \gamma \|v\|_0^2 \quad \text{for all } \ v \in H_0^1(\Omega),$$

and (9.4) possesses a unique solution $u \in H_0^1(\Omega)$.

We shall restrict ourselves to tensor-product meshes $\bar{\omega} := \bar{\omega}_x \times \bar{\omega}_y$ as in Sect. 9.1. Set $I_i := [x_{i-1}, x_i]$, $J_j := [y_{j-1}, y_j]$ and $T_{ij} := I_i \times J_j$. We shall consider both bilinear elements on rectangles and linear elements on triangles with the triangulation obtained by drawing either diagonal in each of the mesh rectangles; see Fig. 9.1.

9.2.1 The Interpolation Error

The first important results are bounds for the interpolation error. We denote by ψ^I the piecewise bilinear function that interpolates to ψ at the nodes of the mesh $\bar{\omega}$. The meshes we consider are characterised by high aspect ratios of the mesh elements. Because of this anisotropy, standard interpolation theory cannot be applied. There have been a number of contributions to extend the theory to anisotropic elements, e.g., [15, 171, 172]. The first uniform interpolation error estimates for layer-adapted meshes, namely Shishkin meshes, were derived by Stynes and O'Riordan [152] and Dobrowolski and Roos [30]. Here we shall give the more general results from [84].

Set

$$\Theta_{cd}^{[p]}(T_{ij}) := \int_{I_i} \left(1 + \varepsilon^{-1} \mathrm{e}^{-\beta_1 x/(p\varepsilon)} \right) \mathrm{d}x + \int_{J_j} \left(1 + \varepsilon^{-1} \mathrm{e}^{-\beta_2 y/(2\varepsilon)} \right) \mathrm{d}y$$

and

$$\vartheta_{cd}^{[p]}(\bar{\omega}) := \max_{i,j=1,\ldots,N} \Theta_{cd}^{[p]}(T_{ij}).$$

Theorem 9.4. *Suppose the assumptions of Theorem 7.17 are satisfied. Then the maximum-norm error of bilinear interpolation on a tensor-product mesh satisfies*

$$\left\|u^I - u\right\|_{\infty,T_{ij}} \leq C\left(\Theta^{[2]}_{cd}(T_{ij})\right)^2, \tag{9.6}$$

$$\varepsilon\left\|\nabla(u^I - u)\right\|_{\infty,T_{ij}} \leq C\Theta^{[1]}_{cd}(T_{ij}) \tag{9.7}$$

and for the ε-weighted energy norm

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

Proof. First Theorem 7.17 implies

$$\left|\partial_x^i\partial_y^j u(x,y)\right| \leq C\Big(1+\varepsilon^{-i}\mathrm{e}^{-\beta_1 x/\varepsilon}\Big)\times\Big(1+\varepsilon^{-j}\mathrm{e}^{-\beta_2 y/\varepsilon}\Big) \tag{9.8}$$

for $i+j\leq 2$.

(i) Let $(x,y)\in T_{ij}$. Then

$$\begin{aligned}\left(u-u^I\right)(x,y) = {}& \frac{1}{h_i}\int_{I_i}\int_{x_{i-1}}^{x}\int_{\sigma}^{\xi}\partial_x^2 u(\tau,y_{j-1})\mathrm{d}\tau\mathrm{d}\xi\mathrm{d}\sigma\\ &+\frac{1}{k_j}\int_{J_j}\int_{y_{j-1}}^{y}\int_{\sigma}^{\xi}\partial_y^2 u(x_{i-1},\tau)\mathrm{d}\tau\mathrm{d}\xi\mathrm{d}\sigma\\ &+\int_{x_{i-1}}^{x}\int_{y_{j-1}}^{y}\partial_x\partial_y u(\xi,\tau)\mathrm{d}\tau\mathrm{d}\xi\\ &-\frac{x-x_{i-1}}{h_i}\frac{y-y_{j-1}}{k_j}\int_{I_i}\int_{J_j}\partial_x\partial_y u(\xi,\tau)\mathrm{d}\tau\mathrm{d}\xi.\end{aligned}$$

Applying the technique from Prop. 5.1, we immediately see that the first two terms are bounded by $\Theta^{[2]}_{cd}(T_{ij})^2$. The third and fourth term are clearly bounded by $\Theta^{[1]}_{cd}(T_{ij})^2$. Ineq. (9.6) follows.

(ii) Next, we have

$$\begin{aligned}\partial_x\left(u-u^I\right)(x,y) = {}& \frac{1}{h_i}\int_{I_i}\int_{\sigma}^{x}\partial_x^2 u(\tau,y_{j-1})\mathrm{d}\tau\mathrm{d}\sigma+\int_{y_{j-1}}^{y}\partial_x\partial_y u(x,\tau)\mathrm{d}\tau\\ &-\frac{1}{h_i}\frac{y-y_{j-1}}{k_j}\int_{I_i}\int_{J_j}\partial_x\partial_y u(\xi,\tau)\mathrm{d}\tau\mathrm{d}\xi.\end{aligned}$$

Thus,

$$\left|\partial_x\left(u-u^I\right)(x,y)\right| \leq \int_{I_i}\left|\partial_x^2 u(\tau,y_{j-1})\right|\mathrm{d}\tau+2\int_{J_j}\left|\partial_x\partial_y u(x,\tau)\right|\mathrm{d}\tau$$

and (9.7) follows from (9.8).

(iii) In order to bound the interpolation error in the H^1 seminorm, use integration by parts:

$$\begin{aligned}\left\|\partial_x(u^I-u)\right\|_0^2 = &\int_\Omega \partial_x^2 u(x,y)\big(u^I-u\big)(x,y)\mathrm{d}x\mathrm{d}y \\ &+\sum_{i=1}^{N-1}\int_0^1 \big(u^I-u\big)(x_i,y)K_i(y)\mathrm{d}y\,,\end{aligned} \tag{9.9}$$

where

$$K_i(y) := \partial_x u^I(x_i-0,y) - \partial_x u^I(x_i+0,y)\,.$$

For $y\in J_j$ we have

$$K_i(y) = \frac{y-y_{j-1}}{k_j}\left(u_{\bar{x};ij}-u_{x;i+1,j}\right) + \frac{y_j-y}{k_j}\left(u_{\bar{x};i,j-1}-u_{x;i+1,j-1}\right).$$

By the mean-value theorem there exists a $\xi_{i,j}\in I_i$, such that $u_{\bar{x};i,j}=\partial_x u(\xi_{i,j},y_j)$. Therefore,

$$|u_{\bar{x};ij}-u_{x;i+1,j}| = |\partial_x u(\xi_{i,j},y_j)-\partial_x u(\xi_{i+1,j},y_j)| \le \int_{x_{i-1}}^{x_{i+1}} \left|\partial_x^2 u(\xi,y_j)\right|\mathrm{d}\xi\,.$$

We get

$$|K_i(y)| \le \max_{y\in[0,1]}\int_{x_{i-1}}^{x_{i+1}} \left|\partial_x^2 u(\xi,y)\right|\mathrm{d}\xi\,.$$

This and a Hölder inequality applied to (9.9) yield

$$\begin{aligned}&\left\|(u^I-u)_x\right\|_0^2 \\ &\quad\le \left\|u^I-u\right\|_\infty\left\{\int_\Omega \left|\partial_x^2 u(x,y)\right|\mathrm{d}x\mathrm{d}y + 2\max_{y\in[0,1]}\int_0^1 \left|\partial_x^2 u(x,y)\right|\mathrm{d}x\right\} \\ &\quad\le C\varepsilon^{-1}\left\|u^I-u\right\|_\infty,\end{aligned}$$

by (9.8). The interpolation error in the L_2 norm is bounded by its L_∞ norm. We get the second bound of the theorem. □

Remark 9.5. The second part of the proof when the H^1 seminorm is considered works for bilinear elements, but not for linear ones. Nonetheless, for S-type meshes and linear elements, the conclusions of the theorem hold too; see [82, 137]. ♣

Remark 9.6. Error bounds for particular layer-adapted meshes can be immediately concluded using the results from Sections 2.1.1 and 2.1.3. ♣

9.2.2 Galerkin FEM

Let $V_0^\omega \subset H_0^1(\Omega)$ be the space of piecewise linear/bilinear functions on the given triangulation that vanish on the boundary of Ω. Then our discretisation is as follows: Find $u^N \in V_0^\omega$ such that

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

The coercivity of $a(\cdot,\cdot)$ guarantees the existence of a unique solution $u^N \in V_0^\omega$.

9.2.2.1 Convergence

Convergence of the (bi)linear Galerkin FEM on standard Shishkin meshes was first studied by Stynes and O'Riordan [152]. Their technique was later adapted by Linß and Roos to the analysis of more general S-type meshes [82, 137].

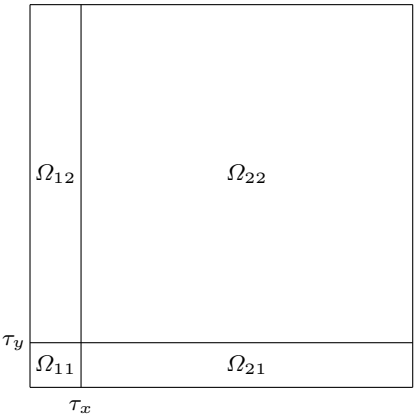
The mesh transition points in the S-type mesh are

$$\tau_x := \min\left\{q, \frac{\sigma\varepsilon}{\beta_1}\ln N\right\} \quad \text{and} \quad \tau_y := \min\left\{q, \frac{\sigma\varepsilon}{\beta_2}\ln N\right\}$$

with mesh parameters $\sigma > 0$ and $q \in (0,1)$ arbitrary, but fixed and with $qN \in \mathbb{N}$. Divide the domain Ω as in Fig. 9.2: $\bar{\Omega} = \Omega_{11} \cup \Omega_{21} \cup \Omega_{12} \cup \Omega_{22}$.

Corollary 9.7. *Let $\bar{\omega} = \bar{\omega}_x \times \bar{\omega}_y$ be a S-type mesh with $\sigma \geq 2$. Then Theorem 9.4 implies*

$$\left\|u - u^I\right\|_{\infty,\Omega\setminus\Omega_{22}} \leq C\left(h + N^{-1}\max|\psi'|\right)^2, \quad \left\|u - u^I\right\|_{\infty,\Omega_{22}} \leq CN^{-2},$$



$$\Omega_{11} = [0,\tau_x] \times [0,\tau_y],$$
$$\Omega_{21} = [\tau_x,1] \times [0,\tau_y],$$
$$\Omega_{12} = [0,\tau_x] \times [\tau_y,1],$$
$$\Omega_{22} = [\tau_x,1] \times [\tau_y,1].$$

Fig. 9.2 Dissection of Ω for tensor-product S-type meshes

and

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

Furthermore, by the Cauchy-Schwarz inequality,

$$\left\|u - u^I\right\|_{0,\Omega\setminus\Omega_{22}} \le C\varepsilon^{1/2}\ln^{1/2} N\left(h + N^{-1}\max|\psi'|\right)^2$$

and

$$\left\|u - u^I\right\|_{0,\Omega_{22}} \le CN^{-2} .$$

Where we have used the results of Sect. 2.1.3 too.

Theorem 9.8. *Let $\bar\omega = \bar\omega_x \times \bar\omega_y$ be a tensor-product S-type mesh with $\sigma \ge 2$, whose mesh generating function $\tilde\varphi$ satisfies (2.8) and*

$$\left(h + N^{-1}\max|\psi'|\right)\ln^{1/2} N \le C. \tag{9.10}$$

Then

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

for both linear elements on triangles and bilinear elements on rectangles.

Proof. The proof is along the lines of Sect. 5.2.1 exploiting the tensor-product structure of the mesh and the solution decomposition of Theorem 7.17; see also [82]. Let $\eta = u^I - u$ and $\chi = u^I - u^N$. Bounds for the interpolation error η are provided by Corollary 9.7.

Bounding χ, we start from the coercivity of $a(\cdot,\cdot)$ and the orthogonality of the Galerkin method, i. e.,

$$\begin{aligned}\left|\!\left|\!\left|\chi\right|\!\right|\!\right|_\varepsilon^2 \le a(\chi,\chi) = a(\eta,\chi) &= \varepsilon(\nabla\eta,\nabla\chi) + (\eta, \boldsymbol{b}^T\nabla\chi) + \big((\operatorname{div} b + c)\eta,\chi\big)\\ &\le C\left|\!\left|\!\left|\eta\right|\!\right|\!\right|_\varepsilon \left|\!\left|\!\left|\chi\right|\!\right|\!\right|_\varepsilon + C\Big(\|\eta\|_{0,\Omega_{22}}\|\nabla\chi\|_{0,\Omega_{22}}\\ &\qquad + \|\eta\|_{L_\infty(\Omega\setminus\Omega_{22})}\|\nabla\chi\|_{L_1(\Omega\setminus\Omega_{22})}\Big).\end{aligned}$$

On $\Omega\setminus\Omega_{22}$ the Cauchy-Schwarz inequality yields

$$\|\nabla\chi\|_{L_1(\Omega\setminus\Omega_{22})} \le C\sqrt{\tau_x+\tau_y}\,\|\nabla\chi\|_0 \le C\ln^{1/2} N\left|\!\left|\!\left|\chi\right|\!\right|\!\right|_\varepsilon ,$$

while on Ω_{22} an inverse inequality yields

$$\|\nabla\chi\|_{0,\Omega_{22}} \le CN\|\chi\|_{0,\Omega_{22}} \le CN\left|\!\left|\!\left|\chi\right|\!\right|\!\right|_\varepsilon .$$

because $h_i \geq N^{-1}$ and $k_j \geq N^{-1}$ for $T_{ij} \in \Omega_{22}$. These two bounds and Corollary 9.7 give

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

Thus,

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

where we have used (9.10). A triangle inequality and the bounds for $\left|\!\left|\!\left|\eta\right|\!\right|\!\right|_\varepsilon$ and $\left|\!\left|\!\left|\chi\right|\!\right|\!\right|_\varepsilon$ complete the proof. □

9.2.2.2 Supercloseness

Similar to the one dimensional case, the Galerkin FEM using bilinear elements on rectangular S-type meshes enjoys a superconvergence property; see [83, 176]. Note that this superconvergence result generally does not hold for linear elements on triangles as numerical experiments confirm [109].

In contrast to the one-dimensional case where we have $((u^I - u)', \chi') = 0$ for arbitrary $\chi \in V^\omega$, we do *not* have $(\nabla(u^I - u^N), \nabla\chi) = 0$ here because $u^I - u$ vanishes in the mesh points only, but not at the inter-element boundaries. This complicates the analysis and requires higher regularity of the solution. In particular, we shall assume that the solution u can be decomposed as $u = v + w_1 + w_2 + w_{12}$, where

$$\left|\partial_x^i \partial_y^j v(x,y)\right| \leq C, \tag{9.11a}$$

$$\left|\partial_x^i \partial_y^j w_1(x,y)\right| \leq C\varepsilon^{-i} \mathrm{e}^{-\beta_1 x/\varepsilon}, \tag{9.11b}$$

$$\left|\partial_x^i \partial_y^j w_2(x,y)\right| \leq C\varepsilon^{-j} \mathrm{e}^{-\beta_2 y/\varepsilon} \tag{9.11c}$$

and

$$\left|\partial_x^i \partial_y^j w_{12}(x,y)\right| \leq C\varepsilon^{-(i+j)} \mathrm{e}^{-(\beta_1 x + \beta_2 y)/\varepsilon} \tag{9.11d}$$

for $i + j \leq 3$ and $x, y \in [0,1]$.

Theorem 9.9. *Let $\bar{\omega} = \bar{\omega}_x \times \bar{\omega}_y$ be a tensor-product S-type mesh with $\sigma \geq 5/2$ that satisfies (2.8). Then the error of the Galerkin FEM u^N satisfies*

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

Proof. The coercivity and Galerkin orthogonality of $a(\cdot,\cdot)$ give

$$\begin{aligned}\left|\!\left|\!\left|u^I-u^N\right|\!\right|\!\right|_\varepsilon^2 &\le \left|a\big(u-u^I,u^I-u^N\big)\right| \\ &\le \varepsilon\left|\big(\nabla(u-u^I),\nabla(u^I-u^N)\big)\right| \\ &\quad + \left|\Big(\boldsymbol{b}^T\nabla(u-u^I)-c(u-u^I),u^I-u^N\Big)\right|.\end{aligned}$$

In the Sect. 9.2.2.3 we shall show that for all $\chi\in V^\omega$

$$\varepsilon\left|\big(\nabla(u-u^I),\nabla\chi\big)\right| \le C\left(h^2+N^{-2}\max|\psi'|^2\right)\left|\!\left|\!\left|\chi\right|\!\right|\!\right|_\varepsilon \tag{9.12}$$

and

$$\begin{aligned}&\left|\Big(\boldsymbol{b}^T\nabla(u-u^I)-c(u-u^I),\chi\Big)\right| \\ &\qquad \le C\left(h^2\ln^{1/2}N+N^{-2}\max|\psi'|^2\right)\left|\!\left|\!\left|\chi\right|\!\right|\!\right|_\varepsilon.\end{aligned} \tag{9.13}$$

Thus,

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

Divide by $\left|\!\left|\!\left|u^I-u^N\right|\!\right|\!\right|_\varepsilon$ to complete the proof.

Theorem 9.9 yields

$$\left|\!\left|\!\left|u^I-u^N\right|\!\right|\!\right|_\varepsilon \le \begin{cases} CN^{-2}\ln^2 N & \text{for the standard Shishkin mesh and} \\ C(\varepsilon^2+N^{-2})\ln^{1/2}N & \text{for the Bakhvalov-Shishkin mesh;}\end{cases}$$

see Sect. 2.1.3 for the bounds on h and $\max|\psi'|$.

Remark 9.10. Another superconvergence result was established by Zhang [176], who studied convergence of the Galerkin FEM on Shishkin meshes in a discrete version of the energy norm, where $\nabla(u-u^N)$ is replaced by a piecewise-constant approximation based on the midpoints of the rectangles of the triangulation. ♣

9.2.2.3 Detailed Analysis, Proofs of (9.12) and (9.13)

In the analysis we require error estimates for interpolation on anisotropic elements, which were derived by Apel and Dobrowolski [15]. Furthermore, a sharp bound for the L_2-norm error of the interpolation error for the layer terms is needed. We shall also use special error expansion formulae derived by Lin [78].

Preliminaries

Let $T_{ij} = I_i \times J_j$ be an element of the triangulation. Set

$$F_i(x) = \frac{(x - x_{i-1/2})^2}{2} - \frac{h_i^2}{8} \qquad \text{and} \qquad G_j(y) = \frac{(y - y_{j-1/2})^2}{2} - \frac{k_j^2}{8},$$

where $(x_{i-1/2}, y_{j-1/2})$ is the midpoint of the mesh rectangle T_{ij}. Denote the east, north, west and south edges of T_{ij} by $l_{k;ij}$ for $k = 1, \ldots, 4$, respectively.

Lemma 9.11 (Lin Identities [78]). *For any function $g \in C^3(\overline{T}_{ij})$ and any $\chi \in V^\omega$ we have the identities*

$$\int_{T_{ij}} \partial_x(g - g^I)\chi_x = \int_{T_{ij}} \left[G_j \partial_x \chi - \frac{1}{3}(G_j^2)' \partial_x \partial_y \chi\right] \partial_x^2 \partial_y g \,, \tag{9.14a}$$

$$\begin{aligned} \int_{T_{ij}} \partial_x(g - g^I)\chi_y &= \left(\int_{l_{4;ij}} - \int_{l_{2;ij}}\right) F_j \partial_x^2 g \chi_x \\ &\quad + \int_{T_{ij}} \left[G_j \partial_x \partial_y^2 g \left(\partial_y \chi - F_i' \partial_x \partial_y \chi\right) + F_j \partial_x^2 \partial_y g \partial_x \chi\right] \end{aligned} \tag{9.14b}$$

and

$$\begin{aligned} &\int_{T_{ij}} \partial_x(g - g^I)\chi \\ &\quad = \int_{T_{ij}} \left[G_j \left(\chi - F_i' \partial_x \chi\right) - \frac{1}{3}\left(G_j^2\right)' \left(\partial_y \chi - F_i' \partial_x \partial_y \chi\right)\right] \partial_x \partial_y^2 g \\ &\qquad + \left(\int_{l_{1;ij}} - \int_{l_{2;ij}}\right) \frac{h_i^2}{12} \chi \partial_x^2 g + \int_{T_{ij}} \left[\frac{1}{6}\left(F_i^2\right)' \partial_x \chi - \frac{h_i^2}{12}\chi\right] \partial_x^3 g \,. \end{aligned} \tag{9.14c}$$

An immediate consequence of (9.14a) is

$$\left|\partial_x \left((g - g^I), \partial_x \chi\right)_{T_{ij}}\right| \le \frac{k_j^2}{8} \int_{T_{ij}} |\partial_x \chi| \, \left|\partial_x \partial_y^2 g\right| + \frac{k_j^3}{24} \int_{T_{ij}} \left|\partial_x \partial_y \chi\right| \, \left|\partial_x \partial_y^2 g\right| .$$

with the Cauchy-Schwarz and an inverse inequality giving

$$\left|\partial_x \left((g - g^I), \partial_x \chi\right)_{T_{ij}}\right| \le C k_j^2 \left\|\partial_x \chi\right\|_{0,T_{ij}} \left\|\partial_x \partial_y^2 g\right\|_{0,T_{ij}} . \tag{9.15}$$

Lemma 9.12 ([15, Theorem 3]). *Let $T_{ij} \in \Omega^N$ and $p \in [1, \infty]$. Assume that g lies in $W_p^2(T_{ij})$. Let g^I denote the bilinear function that interpolates to g at the vertices of T_{ij}. Then*

$$\left\|\partial_x(g - g^I)\right\|_{L_p(T_{ij})} \le C\Big\{h_i \|\partial_x^2 g\|_{L_p(T_{ij})} + k_j \|\partial_x \partial_y g\|_{L_p(T_{ij})}\Big\}.$$

Proposition 9.13. *Let $\bar{\omega} = \bar{\omega}_x \times \bar{\omega}_y$ be a tensor-product S-type mesh that satisfies (2.8). Then for $w = w_1 + w_2 + w_{12}$*

$$\left\|w - w^I\right\|_{0,\Omega_{22}} \leq C\left(\varepsilon^{1/2}N^{-\sigma} + N^{-\sigma-1/2}\right) \tag{9.16a}$$

and

$$\left\|w - w^I\right\|_{0,\Omega\setminus\Omega_{22}} \leq C\varepsilon^{1/2}\left(h + N^{-1}\max|\psi'|\right)^2 \quad \textit{if } \sigma > 2. \tag{9.16b}$$

Proof. (i) When proving (9.16a), we bound $\|w\|_{0,\Omega_{22}}$ and $\left\|w^I\right\|_{0,\Omega_{22}}$ separately and apply a triangle inequality. Clearly,

$$\|w\|_{0,\Omega_{22}} \leq C\varepsilon^{1/2}N^{-\sigma}, \tag{9.17}$$

by (9.11) and a direct calculation.
In order to bound the L_2 norm of w^I we split Ω_{22} into two subdomains

$$S := [x_{qN+1}, 1] \times [y_{qN+1}, 1] \quad \text{and} \quad \Omega_{22} \setminus S.$$

Note that $\Omega_{22} \setminus S$ consists of only one ply of mesh rectangles along the interface between the coarse and the fine mesh regions. We have

$$\left\|w^I\right\|^2_{0,\Omega_{22}\setminus S} \leq \big(2(1-q)N - 1\big)h_{qN+1}k_{qN+1}\left\|w^I\right\|^2_{\infty,\Omega_{22}}.$$

Hence,

$$\left\|w^I\right\|_{0,\Omega_{22}\setminus S} \leq CN^{-\sigma-1/2}. \tag{9.18}$$

For $T_{ij} \subset S$ we estimate as follows:

$$\begin{aligned}\left\|w^I\right\|^2_{0,T_{ij}} &\leq h_i k_j \left\|w^I\right\|^2_{\infty,T_{ij}} \\ &\leq C\int_{I_{i-1}}\int_{J_{j-1}}\left(\mathrm{e}^{-2\beta_1 x/\varepsilon} + \mathrm{e}^{-2\beta_2 y/\varepsilon} + \mathrm{e}^{-2(\beta_1 x+\beta_2 y)/\varepsilon}\right),\end{aligned}$$

by (9.11) and since the mesh on Ω_{22} is uniform. We get

$$\left\|w^I\right\|^2_{0,S} \leq C\int_{\tau_x}^1\int_{\tau_y}^1\left(\mathrm{e}^{-2\beta_1 x/\varepsilon} + \mathrm{e}^{-2\beta_2 y/\varepsilon} + \mathrm{e}^{-2(\beta_1 x+\beta_2 y)/\varepsilon}\right)$$

and

$$\left\|w^I\right\|_{0,S} \leq C\varepsilon^{1/2}N^{-\sigma}. \tag{9.19}$$

Collecting (9.17), (9.18) and (9.19), we obtain (9.16a).

(ii) Before starting the proof of (9.16b) note that by (2.11) and (2.12)

$$\frac{\beta h_i}{\sigma\varepsilon} \le N^{-1} \max|\psi'| \mathrm{e}^{\beta_1 x_i/\sigma\varepsilon} \le CN^{-1} \max|\psi'| \, \mathrm{e}^{\beta_1 x/\sigma\varepsilon} \quad \text{for } x \in I_i.$$

(α) First let us study $w_1 - w_1^I$. For $T_{ij} \subset \Omega_{12} \cup \Omega_{11}$ we have by Lemma 9.12 and (9.11)

$$\begin{aligned}\left\|w_1 - w_1^I\right\|_{0,T_{ij}}^2 &\le C\Bigg\{h_i^4 k_j \int_{I_i} \varepsilon^{-4} \mathrm{e}^{-2\beta_1 x/\varepsilon} \mathrm{d}x \\ &\qquad + h_i^2 k_j^3 \int_{I_i} \varepsilon^{-2} \mathrm{e}^{-2\beta_1 x/\varepsilon} \mathrm{d}x + k_j^5 \int_{I_i} \mathrm{e}^{-2\beta_1 x/\varepsilon} \mathrm{d}x\Bigg\} \\ &\le C\left(N^{-1} \max|\psi'| + h\right)^4 k_j \int_{I_i} \mathrm{e}^{-(2-4/\sigma)\beta_1 x/\varepsilon} \mathrm{d}x.\end{aligned}$$

Summing over all elements in $\Omega_{11} \cup \Omega_{12}$, we get

$$\left\|w_1 - w_1^I\right\|_{0,\Omega_{11}\cup\Omega_{12}}^2 \le C\left(N^{-1} \max|\psi'| + h\right)^4 \int_0^{\tau_x} \mathrm{e}^{-(2-4/\sigma)\beta_1 x/\varepsilon} \mathrm{d}x.$$

Thus,

$$\left\|w_1 - w_1^I\right\|_{0,\Omega_{11}\cup\Omega_{12}} \le C\varepsilon^{1/2}\left(h + N^{-1} \max|\psi'|\right)^2$$

because $\sigma > 2$ is assumed.
On Ω_{21} we estimate as follows

$$\begin{aligned}\left\|w_1 - w_1^I\right\|_{0,\Omega_{21}} &\le \sqrt{\operatorname{meas}\Omega_{21}} \left\|w_1 - w_1^I\right\|_{\infty,\Omega_{21}} \le \sqrt{\operatorname{meas}\Omega_{21}} \left\|w_1\right\|_{\infty,\Omega_{21}} \\ &\le C\varepsilon^{1/2} \ln^{1/2} N \mathrm{e}^{-\beta_1\tau_x/\varepsilon} \le C\varepsilon^{1/2} N^{-2}\end{aligned}$$

because $\sigma > 2$.

Therefore,

$$\left\|w_1 - w_1^I\right\|_{0,\Omega\setminus\Omega_{22}} \le C\varepsilon^{1/2}\left(h + N^{-1} \max|\psi'|\right)^2 \tag{9.20}$$

since $\max|\psi'| \ge 1$.

(β) Clearly a similar argument yields

$$\left\|w_2 - w_2^I\right\|_{0,\Omega\setminus\Omega_{22}} \le C\varepsilon^{1/2}\left(h + N^{-1} \max|\psi'|\right)^2. \tag{9.21}$$

(γ) Finally, let us bound $w_{12} - w_{12}^I$. For $T_{ij} \subset \Omega_{11}$ we have by Lemma 9.12 and (9.11)

$$\begin{aligned}\left\|w_{12} - w_{12}^I\right\|_{0,T_{ij}}^2 &\le C\left(h_i^2 + k_j^2\right)^2 \int_{I_i}\int_{J_j} \varepsilon^{-4} \mathrm{e}^{-2(\beta_1 x + \beta_2 y)/\varepsilon} \mathrm{d}y\mathrm{d}x \\ &\le C\left(h + N^{-1}\max|\psi'|\right)^4 \int_{I_i}\int_{J_j} \mathrm{e}^{-(2-4/\sigma)(\beta_1 x + \beta_2 y)/\varepsilon} \mathrm{d}y\mathrm{d}x\ .\end{aligned}$$

Summing over all elements in Ω_{11}, we get

$$\begin{aligned}&\left\|w_{12} - w_{12}^I\right\|_{0,\Omega_{11}}^2 \\ &\quad\le C\left(N^{-1}\max|\psi'| + h\right)^4 \int_0^{\tau_x}\int_0^{\tau_y} \mathrm{e}^{-(2-4/\sigma)(\beta_1 x + \beta_2 y)/\varepsilon} \mathrm{d}y\mathrm{d}x\ .\end{aligned}$$

Hence,

$$\left\|w_{12} - w_{12}^I\right\|_{0,\Omega_{11}} \le C\varepsilon\left(N^{-1}\max|\psi'| + h\right)^2. \tag{9.22}$$

On $\Omega_{12} \cup \Omega_{21}$ we estimate as follows:

$$\begin{aligned}\left\|w_{12} - w_{12}^I\right\|_{0,\Omega_{12}\cup\Omega_{21}} &\le \sqrt{\operatorname{meas}\Omega_{12}\cup\Omega_{21}}\left\|w_{12} - w_{12}^I\right\|_{\infty,\Omega_{12}\cup\Omega_{21}} \\ &\le \varepsilon^{1/2} N^{-\sigma} \ln^{1/2} N \left\|w_{12}\right\|_{\infty,\Omega_{12}\cup\Omega_{21}} \\ &\le C\varepsilon^{1/2} N^{-\sigma} \ln^{1/2} N \le C\varepsilon^{1/2} N^{-2}\ ,\end{aligned} \tag{9.23}$$

because $\sigma > 2$.

Collect (9.20)–(9.23) to complete the proof. □

Proof of (9.12)

(i) Using (9.15), we obtain for $T_{ij} \subset \Omega_{11} \cup \Omega_{21}$

$$\begin{aligned}&\left|\left(\partial_x(u - u^I), \partial_x\chi\right)_{T_{ij}}\right| \\ &\quad\le Ck_j^2 \left\|\left(1 + \varepsilon^{-1}\mathrm{e}^{-\beta_1 x/\varepsilon}\right)\left(1 + \varepsilon^{-2}\mathrm{e}^{-\beta_2 y/\varepsilon}\right)\right\|_{0,T_{ij}} \|\partial_x\chi\|_{0,T_{ij}} \\ &\quad\le Ck_j^2 \left(1 + \varepsilon^{-2}\mathrm{e}^{-\beta_2 y_{j-1}/\varepsilon}\right)\left\|1 + \varepsilon^{-1}\mathrm{e}^{-\beta_1 x/\varepsilon}\right\|_{0,T_{ij}} \|\partial_x\chi\|_{0,T_{ij}} \\ &\quad\le C\left(h + N^{-1}\max|\psi'|\right)^2 \left\|1 + \varepsilon^{-1}\mathrm{e}^{-\beta_1 x/\varepsilon}\right\|_{0,T_{ij}} \|\partial_x\chi\|_{0,T_{ij}}\ ,\end{aligned}$$

by (2.12) and since $\mathrm{e}^{\beta_2 k_j/\varepsilon} \leq C$ because of (2.8). Application of the Cauchy-Schwarz inequality yields

$$\begin{aligned}\varepsilon &\left|\left(\partial_x(u-u^I),\partial_x\chi\right)_{\Omega_{11}\cup\Omega_{21}}\right| \\ &\leq C\varepsilon\left(h+N^{-1}\max|\psi'|\right)^2\left\|1+\varepsilon^{-1}\mathrm{e}^{-\beta_1 x/\varepsilon}\right\|_{0,\Omega_{11}\cup\Omega_{21}}\|\partial_x\chi\|_{0,\Omega_{11}\cup\Omega_{21}}.\end{aligned}$$

Hence,

$$\varepsilon\left|\left(\partial_x(u-u^I),\partial_x\chi\right)_{\Omega_{11}\cup\Omega_{21}}\right| \leq C\left(h+N^{-1}\max|\psi'|\right)^2|||\chi|||_\varepsilon\ . \tag{9.24}$$

(ii) An argument similar to *(i)* gives

$$\begin{aligned}\varepsilon &\left|\left(\partial_x((v+w_1)-(v+w_1)^I),\partial_x\chi\right)_{\Omega_{12}\cup\Omega_{22}}\right| \\ &\leq C\varepsilon h^2\left\|1+\varepsilon^{-1}\mathrm{e}^{-\beta_1 x/\varepsilon}\right\|_{0,\Omega_{12}\cup\Omega_{22}}\|\partial_x\chi\|_{0,\Omega_{12}\cup\Omega_{22}} \\ &\leq Ch^2|||\chi|||_\varepsilon\ .\end{aligned} \tag{9.25}$$

(iii) Next we consider $w := w_2 + w_{12}$ for $T_{ij} \subset \Omega_{12}$. The stability of the interpolation operator and our bounds on the derivatives of w_2 and w_{12} yield

$$\begin{aligned}&\left\|\partial_x(w-w^I)\right\|_{\infty,T_{ij}} \\ &\quad\leq \|\partial_x w\|_{\infty,T_{ij}} + \left\|\partial_x w^I\right\|_{\infty,T_{ij}} \leq C\|\nabla w\|_{\infty,T_{ij}} \leq C\varepsilon^{-1}N^{-\sigma}.\end{aligned}$$

Thus,

$$\varepsilon\left|\left(\partial_x(w-w^I),\partial_x\chi\right)_{\Omega_{12}}\right| \leq CN^{-\sigma}\|\partial_x\chi\|_{1,\Omega_{12}} \leq CN^{-\sigma}\varepsilon^{1/2}\ln^{1/2}N\,|||\chi|||_\varepsilon\ ,$$

since $\operatorname{meas}\Omega_{12} = \mathcal{O}(\varepsilon\ln N)$. Therefore,

$$\varepsilon\left|\partial_x\left((w_2+w_{12})-(w_2+w_{12})^I\right),\partial_x\chi\right)_{\Omega_{12}}\right| \leq CN^{-2}|||\chi|||_\varepsilon\,, \tag{9.26}$$

because $\sigma > 2$.

(iv) Finally, let us bound the terms involving w_2 and w_{12} on Ω_{22}. Using Lemma 9.12 and (9.11) we get

$$\left\|\partial_x(w_2-w_2^I)\right\|_{0,\Omega_{22}} \leq C\varepsilon^{-1/2}N^{-\sigma}$$

and

$$\left\|\partial_x(w_{12}-w_{12}^I)\right\|_{0,\Omega_{22}} \leq C\varepsilon^{-1}N^{-2\sigma}\ .$$

Thus,

$$\varepsilon \left| \left(\partial_x (w_2 - w_2^I), \partial_x \chi\right)_{\Omega_{22}} \right| \leq C N^{-2} \left\| \chi \right\|_{\varepsilon} \tag{9.27}$$

and

$$\varepsilon \left| \left(\partial_x (w_{12} - w_{12}^I), \partial_x \chi\right)_{\Omega_{22}} \right| \leq C N^{-2\sigma} \left\| \partial_x \chi \right\|_{0,\Omega_{22}} \leq C N^{-2\sigma+1} \left\| \chi \right\|_{0,\Omega_{22}}, \tag{9.28}$$

by an inverse inequality.

Collect (9.24)-(9.28) to obtain

$$\varepsilon \left| \left(\partial_x (u - u^I), \partial_x \chi\right) \right| \leq C \left(h + N^{-1} \max |\psi'|\right)^2 \left\| \chi \right\|_{\varepsilon} \quad \text{for all } \chi \in V^{\omega}.$$

Obviously, we have an identical bound for $\left| \left(\partial_y (u - u^I), \partial_y \chi\right) \right|$ which completes the proof of (9.12).

Proof of (9.13)

Recalling the decomposition (9.11), we set $w = w_1 + w_2 + w_{12}$. Then integration by parts yields

$$\begin{aligned} &\left| - \left(\boldsymbol{b}^T \nabla (u - u^I), \chi\right) + \left(c(u - u^I), \chi\right) \right| \\ &\quad \leq \left| \left(\boldsymbol{b}^T \nabla (v - v^I), \chi\right) \right| + \left| \left(w - w^I, \boldsymbol{b}^T \nabla \chi\right) \right| \\ &\qquad + \left| \left(c(v - v^I), \chi\right) + \left((c + \operatorname{div} \boldsymbol{b})(w - w^I), \chi\right) \right| \end{aligned} \tag{9.29}$$

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

First

$$\begin{aligned} &\left| \left(c(v - v^I), \chi\right) + \left((c + \operatorname{div} \boldsymbol{b})(w - w^I), \chi\right) \right| \\ &\qquad \leq C \big(\|v - v^I\|_0 + \|w - w^I\|_0 \big) \|\chi\|_0 . \end{aligned}$$

Adapting the technique from Sect. 9.2.1 it is shown that

$$\|v - v^I\|_0 + \|w - w^I\|_0 \leq C \left(h + N^{-1} \max |\psi'|\right)^2,$$

since v and w satisfy derivative bounds similar to those of u. Thus,

$$\left| \left(c(v - v^I), \chi\right) + \left((c + \operatorname{div} \boldsymbol{b})(w - w^I), \chi\right) \right| \leq C \left(h + N^{-1} \max |\psi'|\right)^2 \left\| \chi \right\|_{\varepsilon} . \tag{9.30}$$

Next let us bound the second and third term in (9.29). The Cauchy-Schwarz inequality and Proposition 9.13 yield

$$\begin{aligned}
&\left|\left(w - w^I, \boldsymbol{b}^T\nabla\chi\right)\right| \\
&\quad\le C\|w - w^I\|_{0,\Omega_{22}}\|\nabla\chi\|_{0,\Omega_{22}} + C\|w - w^I\|_{0,\Omega\setminus\Omega_{22}}\|\nabla\chi\|_{0,\Omega\setminus\Omega_{22}} \\
&\quad\le C\left(\varepsilon^{1/2}N^{-5/2} + N^{-3}\right)\|\nabla\chi\|_{0,\Omega_{22}} \\
&\qquad\qquad + C\varepsilon^{1/2}\left(h + N^{-1}\max|\psi'|\right)^2\|\nabla\chi\|_{0,\Omega\setminus\Omega_{22}} \\
&\quad\le C\left(h + N^{-1}\max|\psi'|\right)^2|||\chi|||_\varepsilon\,, \qquad (9.31)
\end{aligned}$$

where we have used an inverse inequality and that on Ω_{22} the mesh is uniform with mesh size $\mathcal{O}\left(N^{-1}\right)$.

Finally, we study the term $\left(\boldsymbol{b}^T\nabla(v - v^I), \chi\right)$. Let $b_{1;ij} = b_1(x_i, y_j)$ for all i, j. Using the second identity of Lemma 9.11, we get

$$\begin{aligned}
&\left(b_1\partial_x(v - v^I), \chi\right) \\
&\quad= \sum_{T_{ij}\in\Omega^N}\left\{\left(b_{1;ij}\partial_x(v - v^I), \chi\right)_{T_{ij}} + \left((b_1 - b_{1;ij})\partial_x(v - v^I), \chi\right)_{T_{ij}}\right\} \\
&\quad= \sum_{T_{ij}\in\Omega^N} b_{1;ij}\int_{T_{ij}}\left\{\left[\frac{1}{6}\left(F_i^2\right)'\partial_x\chi - \frac{1}{12}h_i^2\chi\right]\partial_x^3 v\right. \\
&\qquad\qquad\left. + \left[G_j\left(\chi - F_i'\partial_x\chi\right) - \frac{1}{3}\left(G_j^2\right)'\left(\partial_y\chi - F_{ij}'\partial_x\partial_y\chi\right)\right]\partial_x\partial_y^2 v\right\} \\
&\qquad + \frac{1}{12}\sum_{i=1}^{N-1}\sum_{j=1}^{N}\left(b_{1;i+1,j}h_{i+1}^2 - b_{1;ij}h_i^2\right)\int_{J_j}\left(\chi\partial_x^2 v\right)(x_i, y)\mathrm{d}y \\
&\qquad + \sum_{T_{ij}\in\Omega^N}\left((b_1 - b_{1;ij})\partial_x(v - v^I), \chi\right)_{T_{ij}} \\
&\quad=: I_1 + I_2 + I_3\,. \qquad (9.32)
\end{aligned}$$

Use (9.11) to obtain

$$\begin{aligned}
|I_1| \le C\sum_{T_{ij}\in\Omega^N}&\left\{\left(h_i^2 + k_j^2\right)\left(\|\chi\|_{1,T_{ij}} + h_i\|\partial_x\chi\|_{1,T_{ij}}\right)\right. \\
&\left. + k_j^2\left(k_j\|\partial_y\chi\|_{1,T_{ij}} + h_ik_j\|\partial_x\partial_y\chi\|_{1,T_{ij}}\right)\right\}.
\end{aligned}$$

Thus,

$$|I_1| \le C\sum_{T_{ij}\in\Omega^N}\left(h_i^2 + k_j^2\right)\|\chi\|_{1,T_{ij}} \le Ch^2\|\chi\|_{L_1(\Omega)} \le Ch^2\|\chi\|_0\,. \qquad (9.33)$$

For I_2 we proceed as follows: First

$$\int_{J_j} \big(\chi v_{xx}\big)(x_i,y)\mathrm{d}y = \sum_{k=1}^{i} \int_{J_j}\int_{I_k} \big(\partial_x\chi\partial_x^2 v + \chi\partial_x^3 v\big)(x,y)\mathrm{d}x\mathrm{d}y$$

yields

$$\begin{aligned}\sum_{i=1}^{qN} \big(b_{1;i+1,j}h_{i+1}^2 - b_{1;ij}h_i^2\big) \int_{J_j} \big(\chi\partial_x^2 v\big)(x_i,y)\mathrm{d}y \\ = \sum_{i=1}^{qN} \big(b_{1;qN+1,j}h_{qN+1}^2 - b_{1;ij}h_i^2\big) \int_{J_j}\int_{I_i} \big(\partial_x\chi\partial_x^2 v + \chi\partial_x^3 v\big)(x,y)\mathrm{d}x\mathrm{d}y\,.\end{aligned}$$

Thus,

$$\begin{aligned}\Big|\sum_{i=1}^{qN}\sum_{j=1}^{N} \big(b_{1;i+1,j}h_{i+1}^2 - b_{1;ij}h_i^2\big) \int_{J_j} \big(\chi\partial_x^2 v\big)(x_i,y)\mathrm{d}y\Big| \\ \le Ch^2\big\|\partial_x\chi\partial_x^2 v + \partial_x^3\chi v\big\|_{1,\Omega_{11}\cup\Omega_{12}} \\ \le Ch^2\varepsilon^{1/2}\ln^{1/2}N\big(\|\partial_x\chi\|_0 + \|\chi\|_0\big) \\ \le Ch^2\ln^{1/2}N\,|||\chi|||_\varepsilon\,, \end{aligned} \tag{9.34}$$

by (9.11). Furthermore, for $i = qN+1,\dots,N$, we have

$$\Big|\big(b_{1;i+1,j}h_{i+1}^2 - b_{1;ij}h_i^2\big) \int_{J_j} \big(\chi\partial_x^2 v\big)\,(x_i,y)\mathrm{d}y\Big| \le Ch_i^2\|\chi\|_{1,T_{ij}},$$

because $\big|b_{1;i+1,j} - b_{1;ij}\big| \le Ch_{i+1}$, $h_i = h_{i+1} \le h$ and

$$\int_{J_j} \big(\chi\partial_x^2 v\big)(x_i,y)\mathrm{d}y \le Ch_i^{-1}\|\chi\|_{1,T_{ij}},$$

by an inverse inequality. We get

$$\left|\sum_{i=qN+1}^{N-1}\sum_{j=1}^{N} \big(b_{1;i+1,j}h_{i+1}^2 - b_{1;ij}h_i^2\big) \int_{J_j} \big(\chi\partial_x^2 v\big)(x_i,y)\mathrm{d}y\right| \le Ch^2\|\chi\|_0. \tag{9.35}$$

For I_3 we have the following bound:

$$\begin{aligned}|I_3| &\le \sum_{T_{ij}\in\Omega_N} \big\|b_1 - b_{1;ij}\big\|_{\infty,T_{ij}} \Big(h_i\big\|\partial_x^2 v\big\|_{\infty,T_{ij}} + k_j\big\|\partial_x\partial_y v_{xy}\big\|_{\infty,T_{ij}}\Big)\big\|\chi\big\|_{1,T_{ij}} \\ &\le Ch^2\|\chi\|_0, \end{aligned} \tag{9.36}$$

by Lemma 9.12 and (9.11).

Collect (9.32)–(9.36) to obtain

$$\left|\left(b_1\partial_x(v-v^I),\chi\right)\right| \le Ch^2\ln^{1/2}N\,|||\chi|||_\varepsilon \tag{9.37}$$

with the analogously bound

$$\left|\left(b_2\partial_y(v-v^I),\chi\right)\right| \le Ch^2\ln^{1/2}N\,|||\chi|||_\varepsilon\,. \tag{9.38}$$

Substituting (9.30), (9.31), (9.37) and (9.38) into (9.29), we are finished.

9.2.2.4 Maximum-Norm Error Bounds

In this section we use Theorem 9.9 and the interpolation error bounds from Sect. 9.2.1 to obtain bounds for the error of the Galerkin method in the maximum norm.

Start with the region Ω_{22}, where the mesh is quasi-uniform with mesh size $\mathcal{O}\left(N^{-1}\right)$:

$$\begin{aligned}\left\|u^I-u^N\right\|_{\infty,\Omega_{22}} &\le CN\left\|u^I-u^N\right\|_{0,\Omega_{22}}\\ &\le C\left(Nh^2\ln^{1/2}N+N^{-1}\max|\psi'|^2\right).\end{aligned}$$

Thus, on a standard Shishkin mesh, where $h=\mathcal{O}\left(N^{-1}\right)$, one gets

$$\left\|u-u^N\right\|_{\infty,\Omega_{22}} \le CN^{-1}\ln^2 N,$$

while for the Bakhvalov-Shishkin mesh we have

$$\left\|u-u^N\right\|_{\infty,\Omega_{22}} \le CN^{-1}\ln^{1/2} \quad \text{if } \varepsilon \le CN^{-1},$$

because for this mesh $h=\mathcal{O}\left(\max\{N^{-1},\varepsilon\}\right)$.

Now let (x_i,y_j) be any mesh node in Ω_{21}. Then following [152, pp. 11,12] we obtain

$$\begin{aligned}&\left|(u^I-u^N)(x_i,y_j)\right|\\ &\quad=\left|\int_0^{x_i}(u^I-u^N)(x,y_j)\mathrm{d}x\right| \le CN\int_0^{\tau_x}\int_{J_j}\left|\partial_x(u^I-u^N)\right|\\ &\quad\le CN\left(\varepsilon N^{-1}\ln N\right)^{1/2}\left\|\nabla(u^I-u^N)\right\|_{0,[0,\tau_x]\times J_j}\\ &\quad\le CN^{1/2}\ln^{1/2}N\,\left|\left|\left|u^I-u^N\right|\right|\right|_\varepsilon.\end{aligned}$$

Thus,

$$\|u - u^N\|_{\infty,\Omega_{21}} \leq CN^{1/2} \ln^{1/2} N \left(h^2 \ln^{1/2} N + N^{-2} \max |\psi'|^2 \right),$$

by Theorems 9.4 and Theorem 9.9. Clearly identical bounds hold on Ω_{12}.

Apply this result to get bounds for particular S-type meshes:

$$\left\|u - u^N\right\|_{\infty,\Omega_{12}\cup\Omega_{21}} \leq \begin{cases} CN^{-3/2} \ln^{5/2} N & \text{for standard Shishkin meshes,} \\ CN^{-3/2} \ln N & \text{for Bakhvalov-Shishkin meshes} \\ & \text{with } \varepsilon \leq CN^{-1}. \end{cases}$$

9.2.2.5 Gradient Recovery

Similar to Sect. 5.2.3 a gradient recovery operator can be defined for the bilinear Galerkin FEM, that gives approximations of the gradient which are superior to those of Theorem 9.8. We follow [139].

Notation. *In this section let* $\|\cdot\|_{1,D}$ *denote the standard norm in* $H^1(D)$.

Let T be a rectangle of Ω^N and let $\widetilde{T}$ be the patch associated with T, consisting of all rectangles that have a common corner with T (see Fig. 9.3). We define for $v \in V^\omega$ the recovered gradient $\boldsymbol{R}v$ as follows: First we compute the gradient of v at the midpoints of the mesh rectangles $\left(\boldsymbol{\gamma}_{ij} := \nabla v(x_{i-1/2}, y_{j-1/2})\right)$. Then these values are bilinearly interpolated to give the values of Rv at the mesh points of the triangulation, viz.,

$$\left(\boldsymbol{R}v\right)_{ij} = \boldsymbol{\alpha}_{ij} := \sum_{m,n=0}^{1} \frac{h_{i+1-m}}{h_i + h_{i+1}} \frac{k_{j+1-n}}{k_j + k_{j+1}} \boldsymbol{\gamma}_{i+m,j+n} \,. \tag{9.39}$$

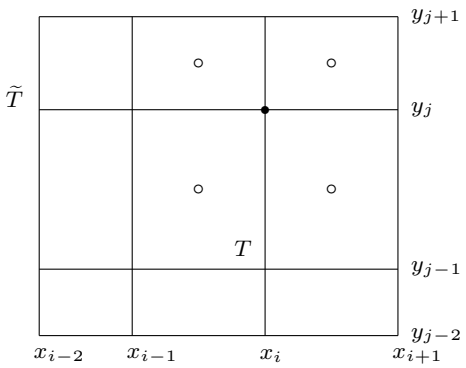


Fig. 9.3 Mesh rectangle T and associated patch $\widetilde{T}$

Bilinear interpolation is again used to extend the recovered gradient from the mesh nodes to the whole of Ω:

$$\begin{aligned}\big(\boldsymbol{R}w^N\big)(x,y) := \boldsymbol{\alpha}_{i-1,j-1}\frac{x_i-x}{h_i}\frac{y_j-y}{k_j} + \boldsymbol{\alpha}_{i,j-1}\frac{x-x_{i-1}}{h_i}\frac{y_j-y}{k_j}\\ + \boldsymbol{\alpha}_{i-1,j}\frac{x_i-x}{h_i}\frac{y-y_{j-1}}{k_j} + \boldsymbol{\alpha}_{i,j}\frac{x-x_{i-1}}{h_i}\frac{y-y_{j-1}}{k_j}\\ \text{for } (x,y)\in T_{ij},\ \ i,j=2,\dots,N-1.\end{aligned}$$

For the boundary rectangles, we simply extrapolate the well-defined bilinear function of the adjacent rectangles.

Lemma 9.14. *$\boldsymbol{R} : V^\omega \to V^\omega \times V^\omega$ is a linear operator with the following properties:*

(locality) $\boldsymbol{R}v$ *on T depends only on values of v on the patch $\widetilde{T}$,*

(stability)
$$\big\|\boldsymbol{R}v\big\|_{\infty,T} \le C\big\|v\big\|_{1,\infty,\widetilde{T}} \quad \textit{for all } v\in V^\omega, \tag{9.40a}$$
$$\big\|\boldsymbol{R}v\big\|_{0,T} \le C\big\|v\big\|_{1,\widetilde{T}} \quad \textit{for all } v\in V^\omega, \tag{9.40b}$$

(consistency)
$$\boldsymbol{R}v^I = \nabla v \quad \textit{on } T \quad \textit{for all } v \textit{ that are quadratic on } \widetilde{T}. \tag{9.40c}$$

Proof. The first three properties are immediate consequences of the definition of $\boldsymbol{R}$, while (9.40c) is verified by a Taylor expansion of v.

Now, given any continuous function v on $\widetilde{T}$, we denote by Qv that quadratic function on $\widetilde{T}$ with

$$\big(Qv\big)(P_k) = v(P_k) \quad \text{for } k=1,\dots,6 \text{ (see Fig. 9.4).}$$

This set of degrees of freedom is unisolvent and thus our Lagrange interpolant Qv is well defined.

The decomposition (9.11) and a careful analysis yield the following bounds for quadratic interpolation.

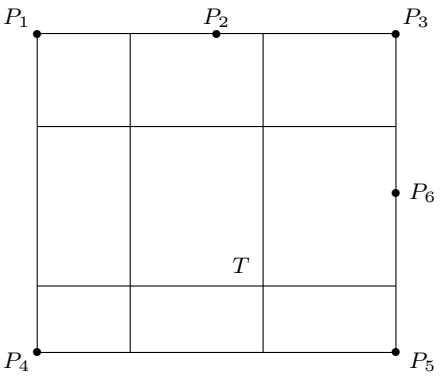


Fig. 9.4 Definition of the quadratic interpolant on the patch $\widetilde{T}$

Lemma 9.15. *Let $\bar{\omega} = \bar{\omega}_x \times \bar{\omega}_y$ be a tensor-product S-type mesh with $\sigma \geq 3$ that satisfies (2.8). Assume that the solution u of (9.1) can be decomposed as in (9.11). Then*

$$\varepsilon \big|u - Qu\big|_{1,\infty,\tilde{T}} \leq \begin{cases} C\big(h + N^{-1}\max|\psi'|\big)^2 & \text{for } T \subset \Omega \setminus \Omega_{22}, \\ CN^{-2} & \text{for } T \subset \Omega_{22}, \end{cases}$$

$$\big\|u - Qu\big\|_{\infty,\tilde{T}} \leq \begin{cases} C\big(h + N^{-1}\max|\psi'|\big)^3 & \text{for } T \subset \Omega \setminus \Omega_{22}, \\ CN^{-3} & \text{for } T \subset \Omega_{22} \end{cases}$$

and

$$\varepsilon^{1/2}\big\|u - Qu\big\|_{1,\tilde{T}} \leq \begin{cases} C(\operatorname{meas}\tilde{T})^{1/2}\big(h + N^{-1}\max|\psi'|\big)^2 & \text{for } T \subset \Omega \setminus \Omega_{22}, \\ CN^{-3} & \text{for } T \subset \Omega_{22}. \end{cases}$$

We would like to estimate the difference between the gradient and the recovered gradient in the ε-weighted H^1 seminorm. We start from

$$\varepsilon^{1/2}\big\|\nabla u - \boldsymbol{R}u^N\big\|_0 \leq \varepsilon^{1/2}\big\|\nabla u - \boldsymbol{R}u^I\big\|_0 + \varepsilon^{1/2}\big\|\boldsymbol{R}(u^I - u^N)\big\|_0, \tag{9.41}$$

by a triangle inequality. For the second term in (9.41), the stability property (9.40b) of the recovery operator and the superconvergence result of Theorem 9.9 yield

$$\varepsilon^{1/2}\big\|\boldsymbol{R}(u^I - u^N)\big\|_0 \leq C\left(h + N^{-1}\max|\psi'|\right)^2 \ln^{1/2} N. \tag{9.42}$$

In the next result we estimate the first term in (9.41).

Proposition 9.16. *Let $\bar{\omega} = \bar{\omega}_x \times \bar{\omega}_y$ be a tensor-product S-type mesh with $\sigma \geq 3$ that satisfies (2.8). Assume that the solution u of (9.1) can be decomposed as in (9.11) and that*

$$\min\{h_{qN}, k_{qN}\} \geq C\varepsilon N^{-1}. \tag{9.43}$$

Then

$$\varepsilon^{1/2}\big\|\nabla u - \boldsymbol{R}u^I\big\|_0 \leq C\big(h + N^{-1}\max|\psi'|\big)^2 \ln^{1/2} N.$$

Proof. For any $T \in \Omega^N$, the consistency property (9.40c) of the recovery operator yields

$$\big\|\nabla u - \boldsymbol{R}u^I\big\|_{0,T} \leq \big\|\nabla(u - Qu)\big\|_{0,T} + \big\|\boldsymbol{R}(u - Qu)^I\big\|_{0,T}, \tag{9.44}$$

since $\boldsymbol{R}(Qu)^I = \nabla Qu$. For the interpolation operator we can use the stability estimates

$$\left\|v^I\right\|_{\infty,T} \le C\|v\|_{\infty,T} \quad \text{and} \quad \left\|v^I\right\|_{1,\infty,T} \le C\left\|v\right\|_{1,\infty,T}.$$

To estimate the second term in (9.44), we bound the L_2 norm by the L_∞ norm and apply the stability property (9.40a) of the recovery operator:

$$\begin{aligned}\left\|\boldsymbol{R}(u-Qu)^I\right\|_{0,T} &\le (\operatorname{meas} T)^{1/2}\left\|\boldsymbol{R}(u-Qu)^I\right\|_{\infty,T}\\ &\le C(\operatorname{meas} T)^{1/2}\left\|(u-Qu)^I\right\|_{1,\infty,\tilde{T}}.\end{aligned} \tag{9.45}$$

Thus, for $T \notin \Omega_{22}$ we have

$$\left\|\boldsymbol{R}(u-Qu)^I\right\|_{0,T} \le C\varepsilon^{-1}(\operatorname{meas} T)^{1/2}\left(h + N^{-1}\max|\psi'|\right)^2, \tag{9.46}$$

by Lemma 9.15.

Next we consider $T \in \Omega_{22}$. We apply an inverse inequality and the L_∞ stability of bilinear interpolation to (9.45) to get

$$\left\|\boldsymbol{R}(u-Qu)^I\right\|_{0,T} \le CN^{-1}\left(\min_{\tilde{T}} h\right)^{-1}\left\|u-u^*\right\|_{\infty,\tilde{T}}.$$

If $\min_{\tilde{T}} h = \mathcal{O}\left(N^{-1}\right)$ then

$$\left\|\boldsymbol{R}(u-Qu)^I\right\|_{0,T} \le C\left\|u-Qu\right\|_{\infty,\tilde{T}} \le CN^{-3}, \tag{9.47}$$

by Lemma 9.15. Otherwise — for the elements T along the transition from the fine to the coarse mesh — we have to estimate more carefully:

$$\left\|\boldsymbol{R}(u-Qu)^I\right\|_{0,T} \le \left\|(u-Qu)^I\right\|_{1,\tilde{T}} \le \sum_{T\in\tilde{T}} \frac{(\operatorname{meas} T)^{1/2}}{\min_T h}\left\|u-Qu\right\|_{\infty,T}.$$

From (9.43), we have

$$\varepsilon^{1/2}\left\|\boldsymbol{R}(u-Qu)^I\right\|_{0,T} \le C\left\|u-Qu\right\|_{\infty,\tilde{T}} \le CN^{-3}, \tag{9.48}$$

by Lemma 9.15. Combining (9.46), (9.47) and (9.48), we have

$$\begin{aligned}&\varepsilon^{1/2}\left\|\boldsymbol{R}(u-Qu)^I\right\|_{0,T}\\ &\quad\le \begin{cases} C\varepsilon^{-1/2}(\operatorname{meas} T)^{1/2}\left(h + N^{-1}\max|\psi'|\right)^2 & \text{for } T \subset \Omega \setminus \Omega_{22},\\ CN^{-3} & \text{for } T \subset \Omega_{22}.\end{cases}\end{aligned}$$

We use the last estimate of Lemma 9.15 and (9.44) to obtain

$$\varepsilon^{1/2}\left\|\nabla u - \boldsymbol{R}u^I\right\|_{0,T} \leq \begin{cases} C\varepsilon^{-1/2}(\operatorname{meas} T)^{1/2}\left(h + N^{-1}\max|\psi'|\right)^2 & \text{for } T \subset \Omega \setminus \Omega_{22}, \\ CN^{-3} & \text{for } T \subset \Omega_{22}\,. \end{cases}$$

Recalling that

$$\left\|\nabla u - \boldsymbol{R}u^I\right\|_0^2 = \sum_{T\in\Omega^N} \left\|\nabla u - \boldsymbol{R}u^I\right\|_{0,T}^2$$

and $\operatorname{meas}\left(\Omega \setminus \Omega_{22}\right) = \mathcal{O}\left(\varepsilon \ln N\right)$, the proof is complete. □

Remark 9.17. The condition (9.43) is satisfied if for example, $\tilde{\varphi}'$ in Sect. 2.1.3 is bounded from below by a positive constant independently of ε and N. Both the original Shishkin mesh and the Bakhvalov-Shishkin mesh satisfy this condition. ♣

As a consequence of (9.41), (9.42) and Proposition 9.16 we have the following result:

Theorem 9.18. *Let $\bar{\omega} = \bar{\omega}_x \times \bar{\omega}_y$ be a tensor-product S-type mesh with $\sigma \geq 5/2$ that satisfies (2.8) and (9.43). Assume that the solution u of (9.1) can be decomposed as in (9.11). Then*

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

Remark 9.19. Similar to the one-dimensional case, using $\boldsymbol{R}u^N$ instead of ∇u^N, we get an asymptotically exact error estimator for the weighted H^1-seminorm of the finite element error $\varepsilon^{1/2}\left\|\nabla(u - u^N)\right\|_0$ on S-type meshes. ♣

9.2.2.6 Numerical Tests

Let us verify our theoretical results for the Galerkin FEM using bilinear trial and test functions on S-type meshes when applied to the test problem (9.3). In our computations we have chosen $\varepsilon = 10^{-8}$ and $\sigma = 3$ for the meshes. In the tables we compare both the error in the ε-weighted energy norm $\left|\!\left|\!\left|u - u^N\right|\!\right|\!\right|_\varepsilon$ with the error in the discrete energy norm $\left|\!\left|\!\left|u^I - u^N\right|\!\right|\!\right|_\varepsilon$, and the accuracy of the gradient approximation ∇u^N with that of the recovered gradient approximation $\boldsymbol{R}u^N$. The errors are estimated using a 4^{th}-order Gauß-Legendre formula on each mesh rectangle. The rates of convergence are computed in the usual way. Tables 9.3 and 9.4 are clear illustrations of Theorems 9.8, 9.9 and 9.18.

Table 9.3 Shishkin mesh

	$\lvert\lvert\lvert u-u^N\rvert\rvert\rvert_\varepsilon$		$\lvert\lvert\lvert u^I-u^N\rvert\rvert\rvert_\varepsilon$		$\varepsilon^{1/2}\lVert\nabla u-\nabla u^N\rVert_0$		$\varepsilon^{1/2}\lVert\nabla u-\boldsymbol{R}u^N\rVert_0$	
N	error	rate	error	rate	error	rate	error	rate
16	2.6900e-1	0.63	5.2110e-2	1.25	2.6898e-1	0.63	9.5425e-1	2.86
32	1.7359e-1	0.72	2.1896e-2	1.43	1.7359e-1	0.72	1.3141e-1	1.81
64	1.0556e-1	0.77	8.1467e-3	1.53	1.0556e-1	0.77	3.7507e-2	1.48
128	6.1881e-2	0.80	2.8137e-3	1.60	6.1881e-2	0.80	1.3479e-2	1.56
256	3.5421e-2	0.83	9.2543e-4	1.65	3.5421e-2	0.83	4.5685e-3	1.64
512	1.9936e-2	0.85	2.9398e-4	1.69	1.9936e-2	0.85	1.4687e-3	1.69
1024	1.1078e-2	—	9.0961e-5	—	1.1078e-2	—	4.5612e-4	—

Table 9.4 Bakhvalov-Shishkin mesh

	$\lvert\lvert\lvert u-u^N\rvert\rvert\rvert_\varepsilon$		$\lvert\lvert\lvert u^I-u^N\rvert\rvert\rvert_\varepsilon$		$\varepsilon^{1/2}\lVert\nabla u-\nabla u^N\rVert_0$		$\varepsilon^{1/2}\lVert\nabla u-\boldsymbol{R}u^N\rVert_0$	
N	error	rate	error	rate	error	rate	error	rate
16	1.2475e-1	1.00	7.9084e-3	2.00	1.2471e-1	1.00	5.0012e-1	3.43
32	6.2574e-2	1.00	1.9800e-3	2.00	6.2569e-2	1.00	4.6315e-2	3.09
64	3.1312e-2	1.00	4.9620e-4	2.00	3.1311e-2	1.00	5.4227e-3	2.43
128	1.5659e-2	1.00	1.2425e-4	2.00	1.5659e-2	1.00	1.0044e-3	2.08
256	7.8298e-3	1.00	3.1096e-5	2.00	7.8298e-3	1.00	2.3690e-4	2.01
512	3.9149e-3	1.00	7.7789e-6	2.00	3.9149e-3	1.00	5.8638e-5	2.00
1024	1.9575e-3	—	1.9460e-6	—	1.9575e-3	—	1.4624e-5	—

9.2.3 Artificial Viscosity Stabilisation

In Sect. 5.3 we studied a FEM with artificial viscosity stabilisation in one dimension. It can be generalised to two dimensions as follows: Set

$$\boldsymbol{\hbar} := \operatorname{diag}(\hbar, \hbar) \quad \text{with} \quad \hbar := h_i \quad \text{in } I_i \times (0,1) \text{ and } \hbar := k_j \quad \text{in } (0,1) \times J_j$$

and let $\kappa \geq 0$ be an arbitrary constant. Then we add artificial viscosity of order $\kappa\hbar$ in x-direction and of order $\kappa\hbar$ in y-direction, i. e., we consider the following discretisation. Find $u^N \in V_0^\omega$ such that

$$a_\kappa(u^N, v) := a(u^N, v) + \kappa \left(\boldsymbol{\hbar}\nabla u^N, \nabla v\right) = (f, v) \quad \text{for all } v \in V_0^\omega.$$

The norm naturally associated with $a_\kappa(\cdot,\cdot)$ is

$$\lvert\lvert\lvert v\rvert\rvert\rvert_\kappa := \left[\lvert\lvert\lvert v\rvert\rvert\rvert_\varepsilon^2 + \kappa\left(\boldsymbol{\hbar}\nabla v, \nabla v\right)\right]^{1/2} \geq \lvert\lvert\lvert v\rvert\rvert\rvert_\varepsilon, \quad \text{for all } v \in H^1(\Omega).$$

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

$$a_\kappa(v, v) \geq \lvert\lvert\lvert v\rvert\rvert\rvert_\kappa^2 \quad \text{for all } v \in H_0^1(\Omega). \tag{9.49}$$

In our analysis we follow Schneider et al. [148], but refine it by explicitly monitoring the dependence on κ. Let again $\eta = u^I - u$ denote the interpolation error and

$\chi = u^I - u^N$ the difference between interpolated and exact solution. Because of the added artificial viscosity the discretisation does not satisfy the Galerkin orthogonality condition, but we have

$$a_\kappa(\chi,\chi) = a(\eta,\chi) + \kappa(\boldsymbol{\hbar}\nabla\eta,\nabla\chi) + \kappa(\boldsymbol{\hbar}\nabla u,\nabla\chi)\ . \tag{9.50}$$

(i) For the first term we have two bounds from Sections 9.2.2.1 and 9.2.2.2:

$$\begin{aligned} &|a(\eta,\chi)| \\ &\quad\le C\,|||\chi|||_\varepsilon \begin{cases} h + N^{-1}\max|\psi'| & \text{for general linear and,} \\ & \text{bilinear elements,} \\ h^2\ln^{1/2} N + N^{-2}\max|\psi'|^2 & \text{for bilinear elements.} \end{cases} \end{aligned} \tag{9.51}$$

(ii) Next we bound $\kappa\big(\boldsymbol{\hbar}\nabla\eta,\nabla\chi\big)$. Let T_{ij} be arbitrary. Then

$$\big(\hbar\eta_x,\chi_x\big)_{T_{ij}} = h_i\int_{J_j}\int_{I_i}\partial_x\eta\partial_x\chi = \int_{J_j}\int_{I_i}\partial_x\eta\int_{I_i}\partial_x\chi.$$

Thus,

$$\Big|\big(\hbar\partial_x\eta,\partial_x\chi\big)_{T_{ij}}\Big| \le 2\|\eta\|_{\infty,T_{ij}}\|\partial_x\chi\|_{1,T_{ij}}.$$

Consequently, we have

$$\begin{aligned} &\big|\big(\hbar\partial_x\eta,\partial_x\chi\big)\big| \\ &\quad\le C\Big\{N^{-2}\|\partial_x\chi\|_{0,\Omega_{22}} + \big(N^{-1}\max|\psi'|\big)^2\big(\varepsilon\ln N\big)^{1/2}\|\partial_x\chi\|_{\Omega\setminus\Omega_{22}}\Big\} \\ &\quad\le CN^{-1}\max|\psi'|\ln^{1/2} N\,|||\chi|||_\varepsilon\,, \end{aligned}$$

by an inverse inequality and (9.5). An analogous estimate holds for $\big|\big(\hbar\partial_y\eta,\partial_y\chi\big)\big|$. Hence,

$$\kappa\,\big|\big(\boldsymbol{\hbar}\nabla\eta,\nabla\chi\big)\big| \le C\kappa N^{-1}\max|\psi'|\ln^{1/2} N\,|||\chi|||_\varepsilon\,. \tag{9.52}$$

(iii) Finally, $\big(\boldsymbol{\hbar}\nabla u,\nabla\chi\big)$ has to be considered. We restrict ourselves to bounding $\big(\hbar\partial_x u,\partial_x\chi\big)$ since the term $\big(\hbar\partial_y u,\partial_y\chi\big)$ can be treated analogously. Using the decomposition of Theorem 7.17, we get

$$\begin{aligned} &\big(\hbar\partial_x u,\partial_x\chi_x\big) \\ &\quad= \big(\hbar\partial_x(v+w_2),\partial_x\chi\big)_{\Omega_{11}\cup\Omega_{12}} + \big(\hbar\partial_x(v+w_2),\partial_x\chi\big)_{\Omega_{21}\cup\Omega_{22}} \\ &\quad+ \big(\hbar\partial_x(w_1+w_{12}),\partial_x\chi\big)_{\Omega_{11}\cup\Omega_{12}} + \big(\hbar\partial_x(w_1+w_{12}),\partial_x\chi\big)_{\Omega_{21}\cup\Omega_{22}}\ . \end{aligned} \tag{9.53}$$

The Cauchy-Schwarz inequality and Theorem 7.17 yield

$$\begin{aligned}&\left|\left(\hbar\partial_x(v+w_2),\partial_x\chi\right)_{\Omega_{11}\cup\Omega_{12}}\right|\\&\quad\le Ch\big(\varepsilon\ln N\big)^{1/2}\left\|\partial_x\chi\right\|_{\Omega_{11}\cup\Omega_{12}}\le Ch\ln^{1/2}N\,|||\chi|||_\varepsilon\,.\end{aligned}\tag{9.54}$$

On $\Omega_{21}\cup\Omega_{22}$ we have

$$\begin{aligned}&\left(\hbar\partial_x(v+w_2),\partial_x\chi\right)_{\Omega_{21}\cup\Omega_{22}}\\&\quad=H\int_0^1\int_{\tau_x}^1\partial_x(v+w_2)\partial_x\chi\mathrm{d}x\mathrm{d}y\\&\quad=-H\int_0^1\Big\{\big(\partial_x(v+w_2)\chi\big)(\tau_x,y)+\int_{\tau_x}^1\partial_x^2(v+w_2)\chi\mathrm{d}x\Big\}\mathrm{d}y.\end{aligned}$$

Thus,

$$\left|\left(\hbar\partial_x(v+w_2),\partial_x\chi\right)_{\Omega_{21}\cup\Omega_{22}}\right|\le CN^{-1}\left\{\|\chi\|_0+\int_0^1|\chi(\tau_x,y)|\,\mathrm{d}y\right\}.\tag{9.55}$$

Note that

$$\int_0^1|\chi(\tau_x,y)|\,\mathrm{d}y=\int_0^1\left|\int_0^{\tau_x}\partial_x\chi\mathrm{d}x\right|\mathrm{d}y\le\|\partial_x\chi\|_{1,\Omega_{11}\cup\Omega_{12}}\le C\ln^{1/2}N\,|||\chi|||_\varepsilon\,.$$

We apply this inequality to (9.55) to obtain

$$\left|\left(\hbar\partial_x(v+w_2),\partial_x\chi\right)_{\Omega_{21}\cup\Omega_{22}}\right|\le CN^{-1}\ln^{1/2}N\,|||\chi|||_\varepsilon\,.\tag{9.56}$$

Now we bound the last two terms in (9.53). Using Theorem 7.17 we get, for any $T_{ij}\in\Omega^N$,

$$\begin{aligned}&\left|\left(\hbar\partial_x(w_1+w_{12}),\partial_x\chi\right)_{T_{ij}}\right|\\&\quad\le C\int_{J_j}\left\{\int_{I_i}\varepsilon^{-1}\mathrm{e}^{-\beta_1x/\varepsilon}\mathrm{d}x\int_{I_i}\left|\partial_x\chi\right|\mathrm{d}x\right\}\mathrm{d}y.\end{aligned}$$

This implies that

$$\begin{aligned}&\left|\left(\hbar\partial_x(w_1+w_{12}),\partial_x\chi\right)_{T_{ij}}\right|\\&\quad\le\begin{cases}CN^{-1}\max|\psi'|\left\|\partial_x\chi\right\|_{1,T_{ij}} & \text{for } T_{ij}\subset\Omega_{11}^N\cup\Omega_{12}^N,\\ CN^{-2}\left\|\partial_x\chi\right\|_{1,T_{ij}} & \text{for } T_{ij}\subset\Omega_{21}^N\cup\Omega_{22}^N.\end{cases}\end{aligned}$$

Therefore,

$$\left|\left(\hbar\partial_x(w_1+w_{12}),\partial_x\chi\right)_{\Omega_{11}\cup\Omega_{12}}\right| \leq CN^{-1}\max|\psi'|\left\|\chi_x\right\|_{1,\Omega_{11}\cup\Omega_{12}} \leq CN^{-1}\max|\psi'|\ln^{1/2}N\,|||\chi|||_\varepsilon$$

and

$$\left|\left(\hbar\partial_x(w_1+w_{12}),\partial_x\chi\right)_{\Omega_{21}\cup\Omega_{22}}\right| \leq CN^{-1}\left\|\chi\right\|_0 ,$$

by an inverse inequality.

Combine the last two bounds with (9.53), (9.54) and (9.56) to get

$$\left|\left(\hbar\partial_x u,\partial_x\chi\right)\right| \leq CN^{-1}\max|\psi'|\ln^{1/2}N\,|||\chi|||_\varepsilon .$$

With an analogous estimate for $\left(\hbar\partial_y u,\partial_y\chi\right)$ we have

$$\kappa\left|(\boldsymbol{\hbar}\nabla u,\nabla\chi)\right| \leq C\kappa N^{-1}\max|\psi'|\ln^{1/2}N\,|||\chi|||_\varepsilon . \tag{9.57}$$

Finally, combine (9.49)–(9.52) and (9.57) in order to obtain the main result of this section.

Theorem 9.20. *Let $\bar\omega := \bar\omega_x\times\bar\omega_y$ be a tensor-product S-type mesh with $\sigma\geq 2$ that satisfies (2.8). Then the upwind-FEM solution u^N satisfies*

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

and, for bilinear elements and $\sigma\geq 5/2$,

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

A consequence of Theorem 9.20 and Sect. 9.2.1 is the following bound of the error in the ε-weighted energy norm:

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

Remark 9.21. The supercloseness property of the Galerkin FEM with bilinear elements is not affected if we take $\kappa=\mathcal{O}\left(N^{-1}\right)$. However, for the efficient treatment of the discrete systems, the choice $\kappa=\mathcal{O}(1)$ is more appropriate which then results in a loss of the supercloseness property. ♣

Remark 9.22. The $|||\cdot|||_\kappa$ bounds imply that the method gives uniform convergent approximations of the gradient on the coarse mesh region Ω_{22}. For example, for a Shishkin mesh, where $\max|\psi'|\leq C\ln N$ and $h\leq 2N^{-1}$, we have

$$\kappa^{1/2}N^{-1/2}\left\|\nabla\left(u^I-u^N\right)\right\|_{0,\Omega_{22}} \leq C\left\{\kappa N^{-1}\ln^{3/2}+N^{-2}\ln^2 N\right\}.$$

Thus,

$$\left\|\nabla\left(u^I - u^N\right)\right\|_{0,\Omega_{22}} \leq \begin{cases} CN^{-1/2}\ln^{3/2} & \text{if } \kappa = \mathcal{O}(1), \\ CN^{-1}\ln^2 & \text{if } \kappa = \mathcal{O}\left(N^{-1}\right). \end{cases}$$

Note that in contrast to the streamline-diffusion FEM, we have full control of the gradient, while for SDFEM one has uniform bounds for the streamline derivative $\|\boldsymbol{b}\cdot\nabla(u^I - u^N)\|_{0,\Omega_{22}}$ only; see Sect. 9.2.4. ♣

Remark 9.23. Suboptimal maximum-norm error bounds on Ω_{22} can be obtained by application of the discrete Sobolev inequality

$$\|\chi\|_{\infty,\Omega_{22}} \leq C\ln^{1/2} N \|\nabla\chi\|_{0,\Omega_{22}}, \tag{9.58}$$

that holds true for piecewise-polynomial functions χ that vanish on a part of the boundary of finite length, see [160, Lemma 5.4] or [63]. We get

$$\left\|u - u^N\right\|_{\infty,\Omega_{22}} \leq \begin{cases} CN^{-1/2}\ln^2 N & \text{if } \kappa = \mathcal{O}(1), \\ CN^{-1}\ln^{5/2} N & \text{if } \kappa = \mathcal{O}\left(N^{-1}\right). \end{cases}$$

Bounds for the maximum-norm error on $\Omega_{21}\cup\Omega_{12}$ can be obtained using the technique from Sect. 9.2.2.4. ♣

9.2.4 Streamline-Diffusion FEM

Introduced by Hughes and Brooks [54], this method is the most commonly used stabilised FEM for the discretisation of convection-diffusion and related problems. Starting from the weak formulation (9.4), we add weighted residuals in order to stabilise the method. Then the SDFEM reads: Find $u^N \in V_0^\omega$ such that

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

with

$$a_{stab}(u^N, v) := \sum_{T\in\Omega^N} \delta_T(\mathcal{L}u^N, -\boldsymbol{b}\cdot\nabla v)_T$$

and

$$f_{SD}(v) := f(v) + \sum_{T\in\Omega^N} \delta_T(f, -\boldsymbol{b}\cdot\nabla v)_T$$

and user chosen stabilisation parameters $\delta_T \geq 0$. We clearly have the Galerkin orthogonality property

$$a_{SD}(u - u^N, v) = 0 \quad \text{for all } v \in V_0^\omega. \tag{9.59}$$

Let V_0^ω be our finite element space consisting of piecewise (bi)linear functions that vanish on $\partial\Omega$. It is shown in, e.g., [141, §III.3.2.1], that if

$$0 \leq \delta_T \leq \gamma \|c\|_{\infty,T}^{-2} \quad \text{for all } T \in \Omega^N, \tag{9.60}$$

then

$$a_{SD}(v, v) \geq \frac{1}{2} |||v|||_{SD}^2 \quad \text{for all } v \in V_0^\omega, \tag{9.61}$$

with the streamline-diffusion norm

$$|||v|||_{SD}^2 := |||v|||_\varepsilon^2 + \sum_{T \in \Omega^N} \delta_T (\boldsymbol{b} \cdot \nabla v, \boldsymbol{b} \cdot \nabla v)_T \ .$$

9.2.4.1 Convergence in the Streamline-Diffusion Norm

Stynes and Tobiska [155] analyse the SDFEM using piecewise bilinear finite elements on standard Shishkin meshes for problems with regular layers. Here we shall extend the technique from [155] to our more general class of S-type meshes, but still consider piecewise bilinear test and trial functions.

Partition the domain $\bar{\Omega}$ as in Fig. 9.2. We follow standard recommendations [141, p. 307] and set

$$\delta_T := \begin{cases} \delta & \text{if } T \subset \Omega_{22}, \\ 0 & \text{otherwise,} \end{cases} \tag{9.62a}$$

and

$$\delta := \begin{cases} \delta_0 N^{-1} & \text{if } \varepsilon \leq N^{-1}, \\ \delta_1 \varepsilon^{-1} N^{-2} & \text{otherwise.} \end{cases} \tag{9.62b}$$

with positive constants δ_0 and δ_1. Clearly $\delta \leq \max\{\delta_0, \delta_1\} N^{-1}$ and therefore (9.60) is satisfied for N sufficiently large, independent of ε.

Note that in the layer regions $\Omega \backslash \Omega_{22}$, the stabilisation is switched off because there the streamline-diffusion stabilisation would be negligible compared to the natural stability induced by the discretisation of the diffusion term.

Our error analysis again starts from the coercivity (9.61) and the Galerkin orthogonality (9.59). Let again $\eta = u^I - u$ and $\chi = u^I - u^N$. Then

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

For the first term we have

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

see Sect. 9.2.2.3, while the stabilisation term

$$a_{stab}(\eta,\chi) = \delta \sum_{T\subset\Omega_{22}} (\varepsilon\Delta u + \boldsymbol{b}\cdot\nabla\eta - c\eta, \boldsymbol{b}\cdot\nabla\chi)_T$$

still has to be analysed. This was done in [155]. Using (9.14b) as a crucial ingredient, Stynes and Tobiska derive the bound

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

Eventually we get the following convergence results.

Theorem 9.24. *Let $\omega_x \times \omega_y$ be a tensor-product S-type mesh with $\sigma \ge 5/2$ that satisfies (2.8). Then the SDFEM solution u^N satisfies*

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

Remark 9.25. Theorems 9.4 and 9.24 give

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

Thus, Theorem 9.24 is a supercloseness result for the SDFEM.

Furthermore,

$$\varepsilon\left\|\nabla u - \boldsymbol{R}u^N\right\|_0 \le C\left(h^2 + N^{-2}\max|\psi'|^2\right)\ln^{1/2} N,$$

where $\boldsymbol{R}$ is the recovery operator from Sect. 9.2.2.5. ♣

9.2.4.2 Maximum-Norm Error Bounds

Clearly the technique for the Galerkin FEM from Sect. 9.2.2.4 can be applied to give pointwise error bounds for the SDFEM with bilinear test and trial functions within the layer regions Ω_{12} and Ω_{21}, while on the coarse mesh region Ω_{22}, we can employ (9.58). We get

$$\|u - u^N\|_{\infty,\Omega\setminus\Omega_{11}} \le \begin{cases} CN^{-3/2}\ln^{5/2} N & \text{for standard Shishkin meshes} \\ CN^{-3/2}\ln N & \text{for Bakhvalov-Shishkin meshes} \\ & \text{with } \varepsilon \le CN^{-1}. \end{cases} \tag{9.63}$$

Adapting Niijima's technique [125], Linß & Stynes [110] study the SDFEM with piecewise linear test and trial functions on Shishkin meshes. For technical reasons a modified version of the SDFEM with artificial crosswind diffusion added on Ω_{22} is studied. Furthermore, it is assumed that the convective field $\boldsymbol{b}$ is constant. The method reads as follows. Find $u^N \in V_0^\omega$ such that

$$a_{SD}(u^N, v) + (\varepsilon^* \boldsymbol{b}^\perp \cdot \nabla u^N, \boldsymbol{b}^\perp \cdot \nabla v) = f_{SD}(v) \quad \text{for all } v \in V_0^\omega$$

with

$$\boldsymbol{b}^\perp := \frac{1}{\|\boldsymbol{b}\|} \begin{pmatrix} -b_2 \\ b_1 \end{pmatrix} \quad \text{and} \quad \varepsilon^* := \begin{cases} \max\left\{0, N^{-3/2} - \varepsilon\right\} & \text{on } \Omega_{22}, \\ 0 & \text{otherwise.} \end{cases}$$

If $\varepsilon \le N^{-3/2}$, then for any point $(x, y) \in \Omega$ the analysis in [110] yields

$$\left|(u - u^N)(x, y)\right| \le \begin{cases} CN^{-1/2} \ln^{3/2} N & \text{if } (x, y) \in \Omega_{22}, \\ CN^{-3/4} \ln^{3/2} N & \text{if } (x, y) \in \Omega \setminus \Omega_{22}, \\ CN^{-11/8} \ln^{1/2} N & \text{if } (x, y) \in (\lambda^*, 1)^2, \end{cases}$$

where $\lambda^* = \mathcal{O}\left(N^{-3/4} \ln N\right)$. The analysis in [110] includes more detailed results and also deals with the case $\varepsilon \ge N^{-3/2}$. Numerical experiments in [109] show convergence of almost second order on the coarse part of the mesh, while inside the boundary layers, the rates are smaller than one. For bilinear elements, almost second-order convergence in the maximum norm is observed globally, but no rigorous analysis is yet available.

9.2.4.3 A Numerical Example

Let us verify the theoretical results when the SDFEM is applied to our test problem (9.3). In the computations we have chosen $\varepsilon = 10^{-8}$ and $\sigma = 3$.

The tables display the error in the ε-weighted energy norm $\left|\!\left|\!\left| u - u^N \right|\!\right|\!\right|_\varepsilon$, in the discrete SD-norm $\left|\!\left|\!\left| u^I - u^N \right|\!\right|\!\right|_{SD}$ and in the maximum-norm. Tables 9.5 and 9.6 clearly illustrate Theorem 9.24, while for the maximum-norm errors (9.63) appears to be suboptimal: instead of convergence of order (almost) $3/2$ we observe (almost) 2^{nd} order.

9.2.4.4 Higher-Order Elements

In [156] Stynes and Tobiska study the SDFEM with Q_p elements, $p > 1$ on tensor-product Shishkin meshes.

Table 9.5 The SDFEM on a Shishkin mesh

	$\|\|\|u-u^N\|\|\|_\varepsilon$		$\|\|\|u^I-u^N\|\|\|_{SD}$		$\|u-u^N\|_\infty$	
N	error	rate	error	rate	error	rate
16	3.3542e-1	0.75	2.0654e-1	1.04	1.7673e-1	1.14
32	1.9932e-1	0.82	1.0021e-1	1.33	8.0261e-2	1.41
64	1.1259e-1	0.83	3.9957e-2	1.50	3.0251e-2	1.51
128	6.3418e-2	0.83	1.4151e-2	1.59	1.0635e-2	1.61
256	3.5718e-2	0.84	4.6849e-3	1.65	3.4956e-3	1.66
512	1.9989e-2	0.85	1.4886e-3	1.69	1.1063e-3	1.70
1024	1.1087e-2	—	4.5993e-4	—	3.4131e-4	—

Table 9.6 The SDFEM on a Bakhvalov-Shishkin mesh

	$\|\|\|u-u^N\|\|\|_\varepsilon$		$\|\|\|u^I-u^N\|\|\|_{SD}$		$\|u-u^N\|_\infty$	
N	error	rate	error	rate	error	rate
16	1.3415e-1	1.07	4.9909e-2	1.92	5.1204e-2	1.89
32	6.3934e-2	1.02	1.3161e-2	1.98	1.3793e-2	1.96
64	3.1488e-2	1.01	3.3354e-3	2.00	3.5346e-3	1.99
128	1.5681e-2	1.00	8.3621e-4	2.00	8.8983e-4	2.00
256	7.8326e-3	1.00	2.0910e-4	2.00	2.2291e-4	2.00
512	3.9153e-3	1.00	5.2263e-5	2.00	5.5756e-5	2.00
1024	1.9575e-3	—	1.3063e-5	—	1.3940e-5	—

The transition points in the Shishkin mesh are

$$\tau_x := \min\left\{q, \frac{(p+1)\varepsilon}{\beta_1}\ln N\right\} \quad \text{and} \quad \tau_y := \min\left\{q, \frac{(p+1)\varepsilon}{\beta_2}\ln N\right\},$$

otherwise the construction of the mesh is unchanged. The stabilisation parameters δ_T are chosen as in (9.62).

We introduce a special vertices-edges-element interpolant [79] as follows. Let v be a given function. On each element $T \in \Omega^N$ the interpolant Iv is defined by

$$\begin{aligned}
&(Iv)(\boldsymbol{x}_i) = v(\boldsymbol{x}_i) \quad \text{for } i = 1,\dots,4,\\
&\int_{\ell_i} (Iv)\varphi = \int_{\ell_i} v\varphi \quad \text{for all } \varphi \in P_{p-2}(\ell_i) \text{ and for } i = 1,\dots,4,
\end{aligned}$$

and

$$\int_T (Iv)\psi = \int_T v\psi \quad \text{for all } \psi \in Q_{p-2}(T),$$

where $\boldsymbol{x}_i$ are the four vertices of T and ℓ_i are the four edges of T. $P_k(\ell_i)$ is the space of polynomials of degree at most k in the single variable whose axis is parallel to the edge ℓ_i.

In [156] the authors proceed by deriving interpolation error bounds for Iu on anisotropic meshes. For example, the pointwise interpolation error is shown to satisfy

$$|(u - Iu)(\boldsymbol{x})| \leq \begin{cases} CN^{-(p+1)} & \text{if } \boldsymbol{x} \in \bar{\Omega}_{22}, \\ C\left(N^{-1}\ln N\right)^{p+1} & \text{otherwise.} \end{cases}$$

However, the main result in [156] is

$$\left|\!\left|\!\left| u^N - Iu \right|\!\right|\!\right|_{SD} \leq CN^{-(p+1/2)}.$$

This is a supercloseness result, because in general for Q_p elements, one can expect at best

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

Postprocessing can be used to obtain an approximation of u of order $p + 1/2$.

Remark 9.26. Matthies [119] considers a different approach to stabilised FEM: local projection stabilisation on Shishkin-type meshes. Using Q_p elements inside the layers and enriched Q_p elements in the region where the mesh is coarse, he proves convergence and supercloseness results that resemble those by Stynes and Tobiska for SDFEM. ♣

9.2.5 Characteristic Layers

We now consider (9.1) with parabolic layers. This is, we assume the convective field is $\boldsymbol{b} = (b, 0)$ and seek a solution u to

$$-\varepsilon\Delta u - bu_x + cu = f \ \text{ in } \Omega = (0,1)^2, \ \ u = 0 \ \text{ on } \partial\Omega \tag{9.64}$$

with $b \geq \beta > 0$ and $c \geq 0$ on $\bar{\Omega}$. The main contributions here are by Franz et al. [38–41].

Analytical properties of this problem have been studied in Sect. 7.3.2. There will be an exponential layer at $x = 0$ and parabolic layers along the boundaries $y = 0$ and $y = 1$.

When discretising (9.64), we use tensor-product Shishkin-type meshes with N mesh intervals in each coordinate direction. The mesh in x-direction is a mesh for one-dimensional convection-diffusion equations (Sect. 2.1.3), while the mesh in y-direction is a mesh for a reaction-diffusion problem (Sect. 2.2). The transition parameters for these meshes are

$$\lambda_x := \min\left\{q, \frac{\sigma\varepsilon}{\beta}\ln N\right\} \quad \text{and} \quad \lambda_y := \min\left\{\frac{q}{2}, \sigma\sqrt{\varepsilon}\ln N\right\}$$

with the mesh parameters $\sigma > 0$ and $q \in (0, 1/2)$. A plot of the resulting mesh is displayed in Fig. 2.21 on p. 29. The mesh characterising function is again denoted by ψ.

Interpolation error

Adapting the technique from Sect. 9.2.1, we obtain for all mesh rectangles T_{ij}

$$\left\|u - u^I\right\|_{\infty,T_{ij}} \leq C\left\{\int_{I_i}\left(1+\varepsilon^{-1}\mathrm{e}^{-\beta x/2\varepsilon}\right)\mathrm{d}x + \int_{J_j}\left(1+\varepsilon^{-1/2}\mathrm{e}^{-y/2\sqrt{\varepsilon}}\right)\mathrm{d}y\right\}^2$$

and

$$\left\|u - u^I\right\|_{\infty} \leq C\left\{\max_{i=1,\ldots,N}\int_{I_i}\left(1+\varepsilon^{-1}\mathrm{e}^{-\beta x/2\varepsilon}\right)\mathrm{d}x + \max_{j=1,\ldots,N}\int_{J_j}\left(1+\varepsilon^{-1/2}\mathrm{e}^{-y/2\sqrt{\varepsilon}}\right)\mathrm{d}y\right\}^2.$$

Bounds for particular meshes (Bakhvalov or Shishkin meshes) can immediately be concluded as has been done before.

Energy-norm bounds. The L_2 part of the norm is easily bounded by the maximum norm. Therefore, we consider the H_1 part only. Reasoning as in Sect. 9.2.1 and using Theorem 7.20, we get

$$\varepsilon^{1/2}\left\|\partial_x\left(u-u^I\right)\right\|_0 \leq C\left\|u-u^I\right\|_{\infty}^{1/2}$$

and

$$\varepsilon^{1/2}\left\|\partial_y\left(u-u^I\right)\right\|_0 \leq C\varepsilon^{1/4}\left\|u-u^I\right\|_{\infty}^{1/2}.$$

The last estimate highlights a problem of the energy norm applied to problems with characteristic layers: It fails to capture these layers. Nonetheless, it is the natural norm associated with the weak formulation of (9.64).

Galerkin FEM

In [38, 39] for finite elements with piecewise bilinear test and trial functions the following bounds in the ε-weighted energy norm are given.

Theorem 9.27. *Let u^N be the piecewise bilinear Galerkin approximation on a S-type mesh with $\sigma \geq 5/2$ then*

$$\left|\!\left|\!\left| u^I - u^N \right|\!\right|\!\right|_\varepsilon \leq C \left(\left(h + N^{-1}\right) \ln^{1/4} N + k + N^{-1} \max |\psi'| \right)^2$$

and

$$\left|\!\left|\!\left| u - u^N \right|\!\right|\!\right|_\varepsilon \leq C \left(\varepsilon^{1/4} h + k + N^{-1} \max |\psi'| \right),$$

where h and k are the maximum step sizes in x- and y-direction, resp.

Remark 9.28. The first bound is a supercloseness result and allows for postprocessing that gives higher-order accurate approximations of the gradient. ♣

Streamline-diffusion FEM

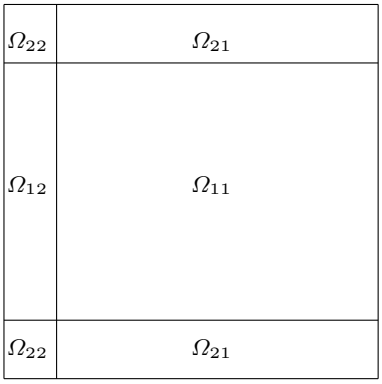
The choice of S-type meshes naturally divides the domain Ω into four (six) subregions, see Fig. 9.5. Ω_{12} covers the exponential layer, Ω_{21} the parabolic layer and Ω_{22} the corner layer and Ω_{11} the remaining region which does not have layers. On each of the four subdomains we allow the streamline-diffusion parameter δ to take different values: δ_{ij} on Ω_{ij}.

Theorem 9.29. *Let u^N be the piecewise bilinear streamline-diffusion approximation on a S-type mesh with $\sigma \geq 5/2$. Suppose the stabilisation parameters satisfy*

$$\delta_{12} \leq \delta_0 \varepsilon \left(h + N^{-1} \max |\psi'|\right)^2, \quad \delta_{21} \leq \delta_0 \varepsilon^{-1/4} N^{-2}, \quad \delta_{22} \leq \delta_0 \varepsilon^{3/4} N^{-2}$$

and

$$\delta_{11} \leq \delta_0 \min \left\{ N^{-1}, \varepsilon^{-1} N^{-2} \right\}$$



$$\Omega_{11} := [\lambda_x, 1] \times [\lambda_y, 1 - \lambda_y]$$
$$\Omega_{12} := [0, \lambda_x] \times [\lambda_y, 1 - \lambda_y]$$
$$\Omega_{21} := [\lambda_x, 1] \times ([0, \lambda_y] \cup [1 - \lambda_y, 1])$$
$$\Omega_{22} := [0, \lambda_x] \times ([0, \lambda_y] \cup [1 - \lambda_y, 1])$$

Fig. 9.5 Dissection of Ω

with a constant δ_0 that is independent of ε. Then

$$\left|\!\left|\!\left| u^I - u^N \right|\!\right|\!\right|_{SD} \le C \left(\left(h + N^{-1}\right) \ln^{1/4} N + k + N^{-1} \max |\psi'| \right)^2$$

and

$$\left|\!\left|\!\left| u - u^N \right|\!\right|\!\right|_{\varepsilon} \le C \left(\varepsilon^{1/4} h + k + N^{-1} \max |\psi'| \right),$$

where h and k are the maximum step sizes in x- and y-direction, resp.

Proof. See [38, 40]. □

Remark 9.30. The amount of stabilisation inside the exponential layer Ω_{12} and in the corner Ω_{22} is negligible and can be switched off. On the coarse-mesh region Ω_{11} the stabilisation is standard. However, inside the characteristic layer, i.e. on Ω_{21} the negative power of ε in the upper bound for δ_{21} is surprising.

This is essentially due to the aforementioned weakness of the SD-norm and the ε-weighted energy norm, which fail to capture the parabolic layer. For the term w_2 in the decomposition of Theorem 7.20 one has $\left|\!\left|\!\left| w_2 \right|\!\right|\!\right|_{\varepsilon} = \left|\!\left|\!\left| w_2 \right|\!\right|\!\right|_{SD} = \mathcal{O}\left(\varepsilon^{1/4}\right)$.

An alternative, though heuristic approach, in [94] using residual free bubbles suggests the choice $\delta_{22} = \mathcal{O}\left(N^{-2}\right)$. ♣

Local projection stabilisation

LPFEM for problems with characteristic layers is studied by Franz and Matthies [41]. The results are similar to those for streamline-diffusion stabilisation.

9.3 Finite Volume Methods

In this section we consider an inverse-monotone finite volume discretisation for problem (7.8). This scheme was introduced by Baba and Tabata [17] and later generalised by Angermann [13, 14]. For a detailed derivation of the method, the reader is referred to [13, 62].

When working on arbitrary partitions we follow Angermann [13]. Further stability results for tensor-product meshes and uniform convergence of the method in both a discrete energy norm and in the maximum norm are due to other authors. References will be given when appropriate.

For the moment let $\Omega \subset I\!R^2$ be an arbitrary domain with polygonal boundary. Consider the problem

$$-\varepsilon \Delta u - \boldsymbol{b} \nabla u + cu = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \Gamma = \partial\Omega \tag{9.65}$$

with $0 < \varepsilon \ll 1$ and

$$c + \frac{1}{2} \operatorname{div} \boldsymbol{b} \geq \gamma > 0 \quad \text{on} \quad \Omega. \tag{9.66}$$

Let $\bar{\omega} = \{\boldsymbol{x}_i\} \subset \bar{\Omega}$ be a set of mesh points. Let Λ and $\partial\Lambda$ be the sets of indices of interior and boundary mesh points, i.e.,

$$\Lambda := \{i : \boldsymbol{x}_i \in \Omega\} \quad \text{and} \quad \partial\Lambda := \{i : \boldsymbol{x}_i \in \partial\Omega\}.$$

Set $\bar{\Lambda} := \Lambda \cup \partial\Lambda$. Partition the domain Ω into subdomains

$$\Omega_i := \left\{\boldsymbol{x} \in \Omega : \|\boldsymbol{x} - \boldsymbol{x}_i\| < \|\boldsymbol{x} - \boldsymbol{x}_j\| \ \text{ for all } j \in \bar{\Lambda} \text{ with } i \neq j\right\} \ \text{ for } i \in \bar{\Lambda},$$

where $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^2$. We define $\Gamma_{ij} = \partial\Omega_i \cap \partial\Omega_j$ and we say that two mesh nodes $\boldsymbol{x}_i \neq \boldsymbol{x}_j$ are adjacent if and only if $m_{ij} := \operatorname{meas}_{1D} \Gamma_{ij} \neq 0$. By Λ_i we denote the set of indices of all mesh nodes that are adjacent to $\boldsymbol{x}_i$. Furthermore, set $d_{ij} := \|\boldsymbol{x}_i - \boldsymbol{x}_j\|$ and $m_i = \operatorname{meas}_{2D} \Omega_i$. We denote by n_{ij} the outward normal on the boundary part Γ_{ij} of Ω_i. Let h, the mesh size, be the maximal distance between two adjacent mesh nodes. Set $N_{ij} := -n_{ij} \cdot \boldsymbol{b}\,((\boldsymbol{x}_i + \boldsymbol{x}_j)/2)$; see Fig. 9.6.

For a reasonable discretisation of the boundary conditions, we shall assume that $\Gamma \subset \bigcup_{i \in \partial\Lambda} \bar{\Omega}_i$.

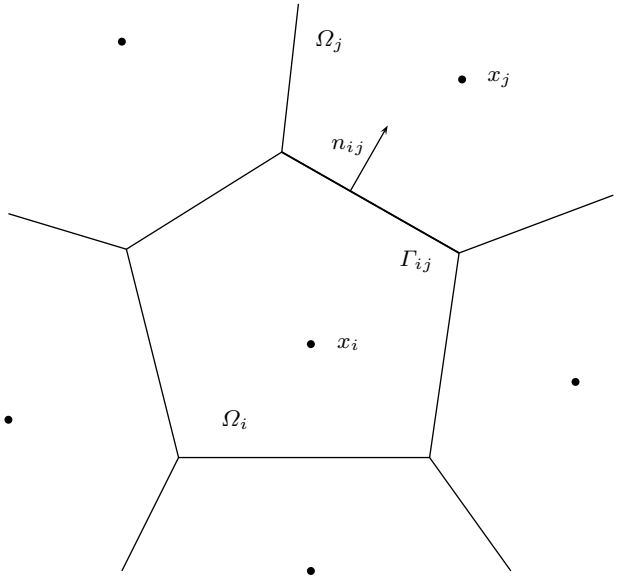


Fig. 9.6 Mesh cell of the FVM

Then our discretisation of (9.65) is as follows. Find $u^N : \bar{\Lambda} \to \mathbb{R}$ such that

$$[L_\rho u^N]_i = f_i m_i \quad \text{for } i \in \Lambda, \;\; u_i^N = 0 \quad \text{for } i \in \partial\Lambda, \tag{9.67a}$$

with u_i^N being the numerical approximation of $u(\boldsymbol{x}_i)$,

$$[L_\rho v]_i := \sum_{j \in \Lambda_i} m_{ij} \left(\frac{\varepsilon}{d_{ij}} - N_{ij}\rho_{ij} \right) (v_i - v_j) + c_i m_i v_i, \tag{9.67b}$$

$\rho_{ij} = \rho(N_{ij} d_{ij}/\varepsilon)$ and a function $\rho : \mathbb{R} \to [0,1]$. Possible choices for ρ are given in Sect. 5.4 which studies the one-dimensional version of the FVM. Again we shall assume that ρ satisfies

$$\begin{aligned}
&(\rho_0) && t \mapsto t\rho(t) \text{ is Lipschitz continuous,} \\
&(\rho_1) && \big[\rho(t) + \rho(-t) - 1\big]t = 0 \quad \text{for all } t \in \mathbb{R}, \\
&(\rho_2) && \big[1/2 - \rho(t)\big]t \geq 0 \quad \text{for all } t \in \mathbb{R}, \\
&(\rho_3) && 1 - t\rho(t) \geq 0 \quad \text{for all } t \in \mathbb{R}.
\end{aligned}$$

Note that the constant choice $\rho \equiv \frac{1}{2}$, which generates a generalised central difference scheme, satisfies conditions (ρ_1) and (ρ_2), but not (ρ_3). Conditions (ρ_1) and (ρ_2) guarantee the coercivity of the weak formulation associated with (9.67), while (ρ_3) ensures the inverse monotonicity of the scheme when the coefficient c is strictly positive.

9.3.1 Coercivity of the Method

The FVM can be written in variational form: Find

$$u^N \in V_0^\omega := \big\{ v \in \mathbb{R}^{\operatorname{card} \bar{\Lambda}} : v_k = 0 \text{ for } k \in \partial\Lambda \big\}$$

such that

$$a_\rho(u^N, v) = f_\rho(v) \quad \text{for all } \; v \in V_0^\omega,$$

with

$$a_\rho(w, v) := \sum_{i \in \Lambda} [L_\rho w]_i v_i \quad \text{and} \quad f_\rho(v) := \sum_{i \in \bar{\Lambda}} f_i m_i v_i.$$

When studying the coercivity of the scheme we split the bilinear form into three parts representing the diffusion, convection and reaction terms:

$$a_\rho(w,v) = \varepsilon d_\rho(w,v) + c_\rho(w,v) + r_\rho(w,v)$$

with

$$d_\rho(w,v) = \sum_{i\in\bar\Lambda}\sum_{j\in\Lambda_i} \frac{m_{ij}}{d_{ij}}(w_i - w_j)v_i,$$
$$c_\rho(w,v) = -\sum_{i\in\bar\Lambda}\sum_{j\in\Lambda_i} m_{ij}N_{ij}\rho_{ij}\,(w_i - w_j)\,v_i$$

and

$$r_\rho(w,v) = \sum_{i\in\bar\Lambda} c_i m_i w_i v_i.$$

These three terms will be studied separately.

Changing the order of summation and renaming the indices yields

$$\sum_{i\in\bar\Lambda}\sum_{j\in\Lambda_i} \frac{m_{ij}}{d_{ij}}(v_i - v_j)v_i = -\sum_{i\in\bar\Lambda}\sum_{j\in\Lambda_i} \frac{m_{ij}}{d_{ij}}(v_i - v_j)v_j$$

Therefore,

$$d_\rho(v,v) = \frac{1}{2}\sum_{i\in\bar\Lambda}\sum_{j\in\Lambda_i} \frac{m_{ij}}{d_{ij}}(v_i - v_j)^2 =: |v|^2_{1,\omega} \tag{9.68}$$

which is a positive definite term.

Remark 9.31. Given a mesh function $v \in V_0^\omega$ define a function $\tilde v \in H_0^1(\Omega)$ that coincides with v in the mesh points, and that is piecewise linear on a Delaunay triangulation associated with the set of mesh points $\bar\omega$. Then

$$|v|^2_{1,\omega} = (\nabla\tilde v, \nabla\tilde v) = |\tilde v|^2_1;$$

see [13]. ♣

Next consider the convection term. By definition we have $m_{ij} = m_{ji}$, $d_{ij} = d_{ji}$ and $N_{ij} = -N_{ji}$. Furthermore, (ρ_1) implies $N_{ji}\rho_{ji} = N_{ij}(\rho_{ij} - 1)$. Hence,

$$\sum_{i\in\bar\Lambda}\sum_{j\in\Lambda_i} m_{ij}N_{ij}\rho_{ij}(v_i - v_j)v_i = -\sum_{i\in\bar\Lambda}\sum_{j\in\Lambda_i} m_{ij}N_{ij}\,(\rho_{ij} - 1)\,(v_i - v_j)v_j$$

and

$$c_\rho(v,v) = \frac{1}{2}\sum_{i\in\bar{\Lambda}}\sum_{j\in\Lambda_i} m_{ij}N_{ij}\left(\frac{1}{2}-\rho_{ij}\right)(v_i-v_j)^2 - \frac{1}{4}\sum_{i\in\bar{\Lambda}}\sum_{j\in\Lambda_i} m_{ij}N_{ij}\left(v_i^2-v_j^2\right).$$

Introducing

$$|v|^2_{\rho,\omega} := \frac{1}{2}\sum_{i\in\bar{\Lambda}}\sum_{j\in\Lambda_i} m_{ij}N_{ij}\left(\frac{1}{2}-\rho_{ij}\right)(v_i-v_j)^2,$$

which is a well-defined semi-norm when (ρ_2) is satisfied, we have

$$c_\rho(v,v) = |v|^2_{\rho,\omega} - \frac{1}{2}\sum_{i\in\bar{\Lambda}} v_i^2 \sum_{j\in\Lambda_i} m_{ij}N_{ij}.$$

This and (9.68) yield

$$a_\rho(v,v) = \varepsilon|v|^2_{1,\omega} + |v|^2_{\rho,\omega} + \sum_{i\in\bar{\Lambda}} m_i v_i^2\left(c_i - \frac{1}{2}\sum_{j\in\Lambda_i} m_{ij}N_{ij}\right).$$

Note that

$$m_i \operatorname{div}\boldsymbol{b}_i + \sum_{j\in\Lambda_i} m_{ij}N_{ij} = \mathcal{O}(h).$$

This implies

$$a_\rho(v,v) \geq \varepsilon|v|^2_{1,\omega} + |v|^2_{\rho,\omega} + \frac{\gamma}{2}\|v\|^2_{0,\omega} =: |||v|||^2_\rho \quad \text{with} \quad \|v\|^2_{0,\omega} := \sum_{i\in\bar{\Lambda}} m_i v_i^2,$$

provided h is sufficiently small, independent of the perturbation parameter ε.

We summarise the result of our stability analysis.

Theorem 9.32. *Assume the discretisation* (9.67) *satisfies conditions* (ρ_1) *and* (ρ_2). *Suppose* (9.66) *holds true. Then the bilinear form* $a_\rho(\cdot,\cdot)$ *is coercive with respect to the norm* $|||\cdot|||_\rho$, *i.e.,*

$$a_\rho(v,v) \geq |||v|||^2_\rho \quad \textit{for all} \;\; v \in V_0^\omega$$

if h *is sufficiently small, independent of the perturbation parameter* ε.

Remark 9.33. When $\rho \equiv \frac{1}{2}$, i.e. when the stabilisation is switched off, the bilinear form is coercive with respect to the discrete ε-weighted energy norm

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

However, when $\rho \not\equiv \frac{1}{2}$ then we have coercivity of the scheme in a stronger norm, which results in enhanced stability properties of the FVM. ♣

9.3.2 Inverse Monotonicity

Let the function ρ, which describes the FVM method, satisfy (ρ_1) and (ρ_3). Furthermore, assume that $c > 0$ on $\bar{\Omega}$. Then recalling the definition (9.67), we have

$$\frac{\varepsilon}{d_{ij}} - N_{ij}\rho_{ij} \geq 0.$$

Hence, the diagonal entries of the matrix associated with L_ρ are positive while the off-diagonal ones are non-positive. Thus, the system matrix is an L_0 matrix. Next note that for $v \equiv 1$ we have $[L_\rho v]_i = c_i m_i > 0$. Therefore, application of the M-criterion (Lemma 3.14) verifies the inverse monotonicity of L_ρ, while Lemma 3.17 gives the $(\ell_\infty, \ell_\infty)$-stability inequality

$$\|v\|_{\infty,\omega} \leq \|f/c\|_{\infty,\omega}\,.$$

Since L_ρ is inverse monotone, it enjoys a comparison principle. That is if two mesh functions v and w satisfy

$$\left.\begin{array}{ll} [L_\rho v]_i \leq [L_\rho w]_i & \text{for all } i \in \Lambda, \\ v_i \leq w_i & \text{for all } i \in \partial\Lambda \end{array}\right\} \quad \Longrightarrow \quad v_i \leq w_i \quad \text{for all } i \in \bar{\Lambda}.$$

Remark 9.34. These results hold true with no restrictions imposed on the convective field $\boldsymbol{b}$ and with (ρ_2) possibly violated. ♣

The Green's function on a tensor-product mesh

Using the inverse monotonicity of L_ρ, we now study the Green's functions associated with L_ρ and derive an anisotropic stability inequality on a general tensor-product mesh $\bar{\omega} := \bar{\omega}_x \times \bar{\omega}_y$, with N mesh intervals in each coordinate direction. A stability result of this kind was first established by Andreev for a simple upwind difference scheme; see [7].

Setting $\hbar_i := (h_{i+1} + h_i)/2$,

$$\rho^+_{1;ij} := \rho\left(-\frac{b_{1;i+1/2,j}h_{i+1}}{\varepsilon}\right) b_{1;i+1/2,j}, \quad \rho^-_{1;ij} := \rho\left(\frac{b_{1;i-1/2,j}h_i}{\varepsilon}\right) b_{1;i-1/2,j},$$

$$v_{\bar{x};ij} := \frac{v_{ij} - v_{i-1,j}}{h_i}, \quad v_{\hat{x};ij} = \frac{v_{i+1,j} - v_{ij}}{\hbar_i} \quad \text{and} \quad v_{\check{x};ij} = \frac{v_{i,j} - v_{i-1,j}}{\hbar_i}$$

with analogous definitions for ρ^+_2, ρ^-_2, $v_{\bar{y}}$, $v_{\hat{y}}$, $v_{\check{y}}$ and $\hbar_k$, we can rewrite (9.67) as: Find $u^N \in \left(I\!R^{N+1}_0\right)^2$ such that

$$\begin{aligned}\left[L_\rho u^N\right]_{ij} := -\varepsilon\left(u^N_{\bar{x}\hat{x};ij} + u^N_{\bar{y}\hat{y};ij}\right) - \rho^+_{1,ij}u^N_{\hat{x},ij} - \rho^-_{1,ij}u^N_{\check{x},ij} \\ - \rho^+_{2,ij}u^N_{\hat{y},ij} - \rho^-_{2,ij}u^N_{\check{y},ij} + c_{ij}u^N_{ij} = f_{ij}\end{aligned}$$

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

Any mesh function v that vanishes on the boundary can be represented using the Green's function:

$$v_{ij} = (v, G_{ij,..})_\rho := \sum_{k,l=1}^{N-1} \hbar_k \hbar_l G_{ij,kl} \left[Lv\right]_{kl}, \tag{9.69}$$

where $G_{ij,kl} = G(x_i, y_j, \xi_k, \eta_l)$ solves for fixed k and l

$$\left[L_\rho G_{..,kl}\right]_{ij} = \delta_{x;ik}\,\delta_{y;jl} \quad \text{on } \omega, \quad G_{ij,kl} = 0 \quad \text{on } \partial\omega$$

with

$$\delta_{x;ik} = \begin{cases}\hbar_i^{-1} & \text{if } i = k, \\ 0 & \text{otherwise,}\end{cases} \qquad \text{and} \qquad \delta_{y;jl} = \begin{cases}\hbar_j^{-1} & \text{if } j = l, \\ 0 & \text{otherwise.}\end{cases}$$

The adjoint operator to L_ρ is

$$\begin{aligned}\left[L^*_\rho v\right]_{kl} = -\varepsilon\left(v_{\bar{\xi}\hat{\xi};kl} + v_{\bar{\eta}\hat{\eta};kl}\right) + \left(\rho^+_1 v\right)_{\check{\xi};kl} + \left(\rho^-_1 v\right)_{\hat{\xi};kl} \\ + \left(\rho^+_2 v\right)_{\check{\eta};kl} + \left(\rho^-_2 v\right)_{\hat{\eta};kl} + c_{kl}v_{kl}\end{aligned}$$

and the Green's function solves, for fixed i and j,

$$\left[L^* G_{ij,..}\right]_{kl} = \delta_{x;ik}\,\delta_{y;jl} \quad \text{on } \omega, \quad G_{ij,kl} = 0 \quad \text{on } \partial\omega. \tag{9.70}$$

In our subsequent analysis the following mean value theorem is used.

Lemma 9.35. *Let $\varphi, g \in \mathbb{R}^{N+1}$ with $g_j \geq 0$ and $m \leq \varphi_j \leq M$ for $j = 1, \ldots, N-1$. Then there exists a constant $\tilde{\varphi} \in [m, M]$ with*

$$\sum_{j=1}^{N-1} \hbar_j \varphi_j g_j = \tilde{\varphi} \sum_{j=1}^{N-1} \hbar_j g_j.$$

Let i and j be fixed. First, the inverse monotonicity of L yields $G_{ij,kl} \geq 0$. Next, multiplying (9.70) by $\hbar_l$ and summing for $l = 1, \ldots, N-1$, we obtain the one-dimensional equation

$$-\varepsilon \left(\sum_{l=1}^{N-1} \hbar_l G_{ij,\cdot l} \right)_{\bar{\xi}\check{\xi},k} + \left(\sum_{l=1}^{N-1} \hbar_l \rho^+_{1,\cdot l} G_{ij,\cdot l} \right)_{\check{\xi},k} + \left(\sum_{l=1}^{N-1} \hbar_l \rho^-_{1,\cdot l} G_{ij,\cdot l} \right)_{\hat{\xi},k} + \sum_{l=1}^{N-1} \hbar_l c_{kl} G_{ij,kl} = \delta_{x;ik} - F_k \ ,$$

where

$$F_k = -\varepsilon \left[1 + \frac{\rho^+_{2;k,N-1} h_N}{\varepsilon} \right] G_{\bar{\eta};ij,kN} + \varepsilon \left[1 - \frac{\rho^+_{2;k,0} h_1}{\varepsilon} \right] G_{\bar{\eta};ij,k1} \geq 0,$$

by (ρ_3) and since $G \geq 0$.

Defining

$$\tilde{G}_k := \sum_{l=1}^{N-1} \hbar_l G_{ij,kl} = \| G_{ij,\cdot l} \|_{1,\omega} \ , \quad \text{for } k = 0, \ldots, N,$$

we see that according to Lemma 9.35 there exist mesh functions $\tilde{\rho}^+, \tilde{\rho}^-, \tilde{c}$ with $\tilde{\rho}^+ \geq \beta_1$, $\tilde{\rho}^- \geq \beta_1$ and $\tilde{c} \geq \gamma$ such that

$$-\varepsilon \tilde{G}_{\bar{\xi}\check{\xi},k} + \left(\tilde{\rho}^+ \tilde{G} \right)_{\check{\xi},k} + \left(\tilde{\rho}^- \tilde{G} \right)_{\hat{\xi},k} + \tilde{c}_k \tilde{G}_k = \delta_{x;ik} - F_k.$$

Let $\Gamma = \Gamma_{m,k}$ be the Green's function of the operator

$$[Lv]_k = -\varepsilon v_{\bar{\xi}\check{\xi},k} - \tilde{\rho}^+_k v_{\check{\xi},k} - \tilde{\rho}^-_k v_{\hat{\xi},k} + \tilde{c}_k v_k.$$

Then $\tilde{G}$ can be written as

$$\tilde{G}_k = \Gamma_{i,k} - \sum_{m=1}^{N-1} \hbar_m \Gamma_{m,k} F_m \ ,$$

The nonnegativity of Γ and F gives

$$\tilde{G}_k \le \Gamma_{i,k} \le \frac{1}{\beta_1 \inf_{t<0} \rho(t)},$$

by Lemma 5.20 and Remark 5.23. We get the first inequality of the following theorem. The other one is proved analogously.

Theorem 9.36. *Suppose the control function ρ enjoys properties (ρ_0) and (ρ_3). Then the Green's function associated with L satisfies*

$$\max_{i,j,k=1,\dots,N-1} \sum_{l=1}^{N-1} \hbar_l G_{ij,kl} \le \frac{\alpha}{\beta_1} \quad \textit{and} \quad \max_{i,j,l=1,\dots,N-1} \sum_{k=1}^{N-1} \hbar_k G_{ij,kl} \le \frac{\alpha}{\beta_2}$$

with $\alpha = 1/\inf_{t<0} \rho(t) \le 2$.

Finally, we use these bounds on the Green's function to derive stability estimates for the operator L_ρ. For any mesh function $v : \bar{\omega} \to I\!R$ that vanish on $\partial\omega$, introduce the norm

$$\|v\|_A := \sum_{k=1}^{N-1} \hbar_k \max_{l=1,\dots,N-1} |v_{kl}| .$$

Its dual norm with respect to the discrete scalar product $(\cdot,\cdot)_\rho$ is

$$\|v\|_{A^*} = \max_{k=1,\dots,N-1} \sum_{l=1}^{N-1} \hbar_k |v_{kl}| ,$$

cf. [21, Theorem 2]. The representation (9.69) gives

$$|v_{ij}| \le \|G_{ij,\cdot\cdot}\|_{A^*} \|v\|_A .$$

Application of Theorem 9.36 yields our final stability result which is an extension of the (ℓ_∞, ℓ_1) stability of Sect. 4.2.5 and a generalisation of [7].

Theorem 9.37. *Suppose the control function ρ enjoys properties (ρ_0) and (ρ_3). Then the operator L_ρ is $(\ell_\infty, \ell_1 \otimes \ell_\infty)$ stable with*

$$\|v\|_{\infty,\omega} \le \frac{\alpha}{\beta_1} \|L_\rho v\|_{\ell_1 \otimes \ell_\infty} \quad \textit{and} \quad \|v\|_{\infty,\omega} \le \frac{\alpha}{\beta_2} \|L_\rho v\|_{\ell_\infty \otimes \ell_1}$$

with $\alpha = 1/\inf_{t<0} \rho(t) \le 2$,

$$\|w\|_{\ell_1\otimes\ell_\infty} := \sum_{k=1}^{N-1} \hbar_k \max_{l=1,\dots,N-1} \left|w_{kl}\right|$$

and

$$\|w\|_{\ell_\infty\otimes\ell_1} := \sum_{l=1}^{N-1} \hbar_l \max_{k=1,\dots,N-1} \left|w_{kl}\right|.$$

9.3.3 Convergence

Energy norm

Starting from the coercivity of the bilinear form $a_\rho(\cdot,\cdot)$, see Theorem 9.32, the analysis proceeds along the lines of Sect. 5.4.2 resembling many of the details also used for the Galerkin FEM in two dimensions, see Sect. 9.2.2. Eventually one gets

$$\left|\!\left|\!\left|u-u^N\right|\!\right|\!\right|_\rho + \left|\!\left|\!\left|u-u^N\right|\!\right|\!\right|_\varepsilon \le CN^{-1}\max|\psi'|\ln^{1/2} N$$

for tensor-product meshes of Shishkin-type with $\sigma \ge 2$; see also [174].

Maximum norm

The pointwise errors can be bounded using the hybrid stability inequalities from Theorem 9.37, see [90]. We give a very brief outline of the argument.

The truncation error is split according to the decomposition of Theorem 7.17. Then either of the two bounds from Theorem 9.37 is applied. Section 4.2.5 gives a flavour of the technical details. For a S-type mesh with $\sigma \ge 2$ we obtain

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

If ρ is Lipschitz continuous in $(-m,m)$ with $m>0$, then there exists an $N_m>0$ independent of the perturbation parameter ε such that on a standard Shishkin mesh with $\sigma \ge 2$

$$\left\|u-u^N\right\|_{\infty,\omega} \le CN^{-1} \quad \text{for } N \ge N_m.$$

In the latter case the stabilisation is reduced when the local mesh size is small enough, thus giving higher accuracy inside the layers.

9.3.3.1 Numerical Tests

We verify our theoretical results for the upwind FEM on Shishkin meshes when applied to the test problem (9.3). 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 test the method for three different choices of the controlling function ρ. The errors are measured in the discrete energy and maximum norm and in the FVM-norm.

For $\rho_{U,0}$ (see Table 9.7) we observe convergence of almost first order, namely $N^{-1} \ln N$, in all three norms, while for both $\rho_{U,2}$ and ρ_I—which are Lipschitz continuous—the errors behave like $\mathcal{O}(N^{-1})$; see Tables 9.8 and 9.9. Note, this is covered by our analysis for the maximum norm only.

Table 9.7 FVM on Shishkin meshes, $\rho = \rho_{U,0}$

	$\lvert\!\lvert\!\lvert u - u^N \rvert\!\rvert\!\rvert_\rho$		$\lVert u - u^N \rVert_{\varepsilon,\omega}$		$\lVert u - u^N \rVert_{\infty,\omega}$	
N	error	rate	error	rate	error	rate
16	2.7575e-1	0.68	2.0623e-1	0.55	1.8112e-1	0.62
32	1.7198e-1	0.75	1.4052e-1	0.66	1.1770e-1	0.71
64	1.0230e-1	0.79	8.9046e-2	0.73	7.1880e-2	0.76
128	5.8999e-2	0.83	5.3575e-2	0.79	4.2537e-2	0.80
256	3.3292e-2	0.85	3.1081e-2	0.82	2.4483e-2	0.83
512	1.8493e-2	0.87	1.7579e-2	0.85	1.3786e-2	0.85
1024	1.0153e-2	0.88	9.7672e-3	0.87	7.6456e-3	0.87
2048	5.5247e-3	—	5.3576e-3	—	4.1908e-3	—

Table 9.8 FVM on Shishkin meshes, $\rho = \rho_I$

	$\lvert\!\lvert\!\lvert u - u^N \rvert\!\rvert\!\rvert_\rho$		$\lVert u - u^N \rVert_{\varepsilon,\omega}$		$\lVert u - u^N \rVert_{\infty,\omega}$	
N	error	rate	error	rate	error	rate
16	1.5894e-1	0.83	8.9598e-2	0.80	7.5370e-2	0.70
32	8.9627e-2	0.92	5.1417e-2	0.90	4.6297e-2	0.84
64	4.7445e-2	0.96	2.7514e-2	0.95	2.5790e-2	0.92
128	2.4388e-2	0.98	1.4222e-2	0.98	1.3610e-2	0.96
256	1.2360e-2	0.99	7.2279e-3	0.99	6.9899e-3	0.98
512	6.2219e-3	1.00	3.6430e-3	0.99	3.5418e-3	0.99
1024	3.1214e-3	1.00	1.8288e-3	1.00	1.7827e-3	1.00
2048	1.5633e-3	—	9.1618e-4	—	8.9431e-4	—

Table 9.9 FVM on Shishkin meshes, $\rho = \rho_{U,2}$

	$\lvert\!\lvert\!\lvert u - u^N \rvert\!\rvert\!\rvert_\rho$		$\lVert u - u^N \rVert_{\varepsilon,\omega}$		$\lVert u - u^N \rVert_{\infty,\omega}$	
N	error	rate	error	rate	error	rate
16	1.5359e-1	0.81	8.2430e-2	0.77	7.6384e-2	0.72
32	8.7574e-2	0.91	4.8263e-2	0.88	4.6337e-2	0.85
64	4.6686e-2	0.95	2.6272e-2	0.93	2.5790e-2	0.92
128	2.4120e-2	0.98	1.3773e-2	0.96	1.3610e-2	0.96
256	1.2270e-2	0.99	7.0752e-3	0.98	6.9899e-3	0.98
512	6.1928e-3	0.99	3.5935e-3	0.99	3.5418e-3	0.99
1024	3.1122e-3	1.00	1.8132e-3	0.99	1.7827e-3	1.00
2048	1.5605e-3	—	9.1144e-4	—	8.9431e-4	—