Robust Numerical Methods for Singularly Perturbed Differential Equations–Supplements

Hans-G. Roos, TU Dresden

November 10, 2018

Abstract

The second edition of the book [106] appeared 10 years ago and was for many years a reliable guide into the world of numerical methods for singularly perturbed problems. Since then many new results came into the game, we present some selected ones and the related sources.

AMS subject classification: 65 L10, 65 L12, 65 L50, 65 N30, 65 N50

1 Introduction

Since the appearance of [106] many new results improved our understanding of the complicated task to solve singularly perturbed and flow problems numerically. The survey [96] discussed new developments in the years 2008-2012, and in a paper 2015 [105] we discussed some open problems in analysis and numerics of singularly perturbed problems. Recently, John, Knobloch and Novo also published a survey on recent numerical methods for convection-dominated equations and incompressible flow problems and sketched some open problems [52].

During the BAIL-Conference in Beijing 2016 the idea was born to think about a third version of the book, and in Glasgow 2017 we fixed a working plan for the realization of that idea.

We planned to include some new material, for instance, on systems of singularly perturbed equations (*see Section 2 and 4*), on flux-corrected transport (see [14] and the references given there for the history of AFC schemes and the analysis presented recently in three papers 2015-2017 by Barrenechea, John and Knobloch), on balanced norms (see the survey paper [98] for second-order problems and [40] for fourth-order problems, [99] for systems), on adaptive methods and a posteriori error estimation (*see Section 6*) and on pressure-robust schemes, a new aspect in the FEM analysis of incompressible flows, discussed especially in papers initiated by Linke [2, 53, 65, 67–69].

Moreover, an update of all other sections was planned, including an overview on all kinds of layer-adapted meshes (*see Section 3.1*), a comment to the tailored finite point

method reviewed in [45], hints to improved error estimates for layer-adapted meshes based on superconvergence results on special triangulations, for instance, in [80–82, 132, 133], a presentation of the technique to prove error estimates for finite elements on Duran-Lombardi meshes (*see Section 3.2*), and some remarks on discontinuous Galerkin methods (*see Section 5*).

Unfortunately, the progress made at the end of 2018 was more or less restricted to the parts presented here. Because it is not clear, whether or not the Glasgow plan will be fully realized in the future, I decided to present the existing material to people who are interested to study it.

2 Systems of ordinary differential equations

Systems of ordinary differential equations are often discussed in books on asymptotic expansions for singularly perturbed problems: see, e.g., [119, Chapter 2] or [129, Chapter 7]. Nevertheless in the past relatively little attention was paid to their numerical solution, although the papers [12] (on reaction-diffusion systems) and [1] (on convection-diffusion systems) are worth noting. In recent years interest in this area has grown, as we now describe.

Consider a general system of M equations:

$$L\vec{u} := -\varepsilon\vec{u}'' + B\vec{u}' + A\vec{u} = \vec{f} \quad \text{on } \Omega := (0,1), \tag{2.1a}$$

$$\vec{u}(0) = \vec{g}_0,\ \vec{u}(1) = \vec{g}_1, \tag{2.1b}$$

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

The system (2.1) is said to be *weakly coupled* if the convection coupling matrix B is diagonal, i.e., the i^{th} equation of the system is

$$-\varepsilon u_i'' + b_{ii}u_i' + \sum_{j=1}^{M} a_{ij}u_j = f_i, \tag{2.2}$$

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

Linß [71] allows different diffusion coefficients in different equations: $\varepsilon = \varepsilon_i$ in the i^{th} equation for $i = 1, \ldots, M$. Assume that $b_{ii}(x) \geq \beta_i > 0$ and $a_{ii}(x) \geq \alpha > 0$ on $[0,1]$ for each i. (In [71] the weaker hypothesis $|b_{ii}(x)| \geq \beta_i > 0$ is used, which permits layers in $\vec{u}$ at both ends of [0,1], but for brevity we won't consider this here.) Rewrite (2.2) as

$$-\varepsilon_i u_i'' + b_{ii}u_i' + a_{ii}u_i = -\sum_{j\neq i} a_{ij}u_j + f_i, \tag{2.3}$$

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

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

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

$$|u_i^{(k)}(x)| \leq C\left[1 + \varepsilon_i^{-k} e^{-\beta_i(1-x)/\varepsilon_i}\right] \quad \text{for } x \in [0,1] \text{ and } k = 0, 1. \tag{2.4}$$

Thus there is no strong interaction between the layers in the first-order derivatives of different components u_i; nevertheless the domains of these layers can overlap and this influences the construction of numerical methods for (2.1). But for higher order derivatives additionally weak layers come into the play. In [103] it is proved

$$|u_i''(x)| \leq C\left(1 + \varepsilon_i^{-2} e^{-\beta(1-x)/\varepsilon_i} + \frac{1}{\delta}\sum_{j=i+1}^{M} \varepsilon_j^{-1} e^{-\beta(1-\delta)(1-x)/\varepsilon_j}\right),$$

here $\delta > 0$ is arbitrary, assuming $b_{kk} \geq \beta > 0$. The estimate shows that the i-th component is overlapped by weak layers generated by other components, but only these for $j > i$ due to the increase in diffusion $\varepsilon_1 < \varepsilon_2 < \cdots$.

The system (2.1) is said to be *strongly coupled* if for some $i \in \{1, \ldots, M\}$ one has $b_{ij} \neq 0$ for some $j \neq i$. Such systems do not satisfy a maximum principle of the usual type. One now gets stronger interactions between layers even in the case $\varepsilon_i = \varepsilon$; see [1,70,92,93,101].

We start with that case and the example

$$\begin{aligned} -\varepsilon u_1'' - 3u_1' - 4u_2' &= 1, \\ -\varepsilon u_2'' - 4u_1' + 3u_2' &= 2 \end{aligned}$$

and homogeneous boundary conditions at $x = 0$ and $x = 1$. Canceling of exponentially small terms yields the asymptotic approximation

$$\begin{aligned} u_1^{as} &= 8/25 - 11/25x - 8/25e^{-5x/\varepsilon} + 3/25e^{-5(1-x)/\varepsilon}, \\ u_2^{as} &= 4/25 + 2/25x - 4/25e^{-5x/\varepsilon} - 6/25e^{-5(1-x)/\varepsilon}. \end{aligned}$$

Both solution components do have strong layers at $x = 0$ and $x = 1$ which correspond to the eigenvalues of the matrix B. Moreover, the reduced solution

$$\begin{aligned} u_1^0 &= 8/25 - 11/25x, \\ u_2^0 &= 4/25 + 2/25x \end{aligned}$$

does not satisfy any of the given boundary conditions, but it holds

$$(u_1 - 2u_2)(0) = 0, \quad (2u_1 + u_2)(1) = 0,$$

which corresponds to the eigenvectors of B.

If we assume that all eigenvalues of B do not change the sign in the given interval, every component u_i has, in general, M layers but their location depend on the sign pattern of the eigenvalues (assuming for simplicity, that B is symmetric).

In the case that all eigenvalues have one sign, all overlapping boundary layers are located at one boundary. Such a case is studied in [92, 93] without assuming symmetry of B. We follow [92] to derive a stability result.

For each i assume $b_{ii}(x) \geq \beta_i > 0$ and $a_{ii}(x) \geq 0$ on $[0,1]$. Rewrite the i^{th} equation as

$$L_i u := -\varepsilon u_i'' + b_{ii} u_i' + a_{ii} u_i = f_i + \sum_{\substack{j=1 \\ j \neq i}}^{m} \left[(b_{ij} u_j)' - (b_{ij}' + a_{ij}) u_j \right], \tag{2.8a}$$

$$u_i(0) = u_i(1) = 0. \tag{2.8b}$$

For the scalar problem $L_i v = \phi$ and $v(0) = v(1) = 0$, one has – see [8,9] – the stability result $\|v\|_\infty \leq C_i \|\phi\|_{W^{-1,\infty}}$ for a certain constant C_i that depends only on b_{ii} and a_{ii}. Apply this result to (2.8) then, similarly to the analysis of (2.3), gather the $\|u_j\|_\infty$ terms to the left-hand side. Define the $M \times M$ matrix $\Upsilon = (\gamma_{ij})$ by $\gamma_{ii} = 1$, $\gamma_{ij} = -C_i[\|b_{ij}' + a_{ij}\|_{L_1} + \|b_{ij}\|_\infty]$ for $i \neq j$. Assuming that Υ is inverse monotone, we get an a priori bound on $\|\vec{u}\|_\infty$. Remark that this assumption on B implies that B is strictly diagonal dominant. In the symmetric case, then for $b_{ii} > 0$ the positive definitness of B follows.

Using this bound, it is shown in [92] that one can decompose each component of $\vec{u}$ into smooth and layer components.

The case of eigenvalues of different sign for a symmetric matrix with a special structure is already studied in [1], see also [101] for a system of two equations. Based on the assumption that $A + 1/2B'$ is positive semidefinite, a stability result in the energy norm allows the proof of the existence of a solution decomposition (see [101], Theorem 2.5).

For strongly coupled systems and different small parameters ε_i the situation is even more complicated. For first results concerning layer structure, layer location and characterization of the reduced problem see [90, 97].

For systems of reaction-diffusion equations

$$\mathcal{L}\vec{u} := -\varepsilon^2 \vec{u}'' + A\vec{u} = \vec{f} \quad \text{on } \Omega := (0,1), \tag{2.9a}$$

$$\vec{u}(0) = \vec{g}_0, \ \vec{u}(1) = \vec{g}_1, \tag{2.9b}$$

often the assumption

$$\xi^T A \xi \geq \alpha^2 \xi^T \xi \tag{2.10}$$

with a constant $\alpha > 0$ is used. Then, $\vec{u}$ can be decomposed into a smooth part $\vec{v}$ and a layer part $\vec{w}$, where the layer part satisfies

$$|w_i^{(j)}| \leq C\varepsilon^{-j} \left(e^{-dx/\varepsilon} + e^{-d(1-x)/\varepsilon} \right) \tag{2.11}$$

with some positive constant d. This fact was already used by Bakhvalov [12] to analyse a finite difference method for (2.9) on a special mesh.

In [76] the authors study L_∞ stability (assumption (2.10) implies stability in some energy norm) of the system

$$\mathcal{L}\vec{u} := -\mathcal{E}\vec{u}'' + A\vec{u} = \vec{f} \quad \text{on } \Omega := (0,1), \tag{2.12a}$$

$$\vec{u}(0) = \vec{g}_0,\ \vec{u}(1) = \vec{g}_1, \tag{2.12b}$$

with $\mathcal{E} := \mathrm{diag}(\varepsilon_1^2, \cdots, \varepsilon_M^2)$ and $\varepsilon_1 \le \varepsilon_2 \le \cdots \le \varepsilon_M$. Assuming $a_{kk} > 0$, from

$$-\varepsilon_k^2 u_k'' + a_{kk} u_k = f_k - \sum_{m \neq k} a_{km} u_m$$

follows as above for strongly coupled systems L_∞ stability if

$$\Gamma^{-1} \ge 0 \quad \text{with} \quad \gamma_{ii} = 1, \quad \gamma_{ij} = -\|\frac{a_{ij}}{a_{ii}}\|_\infty \quad \text{for } i \neq j. \tag{2.13}$$

Assumption (2.13) is satisfied if A is strongly diagonal dominant. In the symmetric case then (2.10) follows.

Now let A be strongly diagonal dominant and

$$\sum_{k \neq i} \|\frac{a_{ik}}{a_{ii}}\|_\infty < \zeta < 1.$$

Define

$$\kappa^2 := (1-\zeta) \min_i \min_{x \in [0,1]} a_{ii}(x) \quad \text{and } \mathcal{B}_\varepsilon(x) := e^{-\kappa x/\varepsilon} + e^{-\kappa(1-x)/\varepsilon}.$$

Then (see [76], Theorem 2.4), $\vec{u}$ can be decomposed in $\vec{v} + \vec{w}$, where the layer part satisfies

$$|w_i^{(k)}| \le C \sum_m \varepsilon_m^{-k} \mathcal{B}_{\varepsilon_m}(x) \quad \text{for } k = 0,1,2 \tag{2.14}$$

and

$$|w_i^{(k)}| \le C \varepsilon_i^{2-k} \sum_m \varepsilon_m^{-2} \mathcal{B}_{\varepsilon_m}(x) \quad \text{for } k = 3,4. \tag{2.15}$$

In the case $M = 2$ and analytic data assuming (2.10), full asymptotic expansions that are explicit in the perturbation parameters and the expansion order are presented in [49].

3 Special meshes and the analysis of FEM on DL-meshes

3.1 Special meshes for problems with layers

We now present the construction of meshes suitable for a problem posed in $[0,1]$ with a layer term $E = \exp(-\gamma\, x/\varepsilon)$ (of course, the layer could also be located at $x = 1$ or at both endpoints of the interval).

Our aim is to achieve *uniform convergence* in the discrete maximum norm; that is, the computed solution $\{u_i^N\}_{i=0}^N$ satisfies

$$\|u - u^N\|_{\infty,d} := \max_{i=0,\ldots,N} |u_i - u_i^N| \leq CN^{-\alpha} \tag{3.1}$$

for some positive constants C and α that are independent of ε and of N. A power of N is a suitable measure of the error $u-u^N$ for the particular families of meshes that are discussed, but a bound of this type is inappropriate for an arbitrary family of meshes; see [114].

The aim to achieve uniform convergence in the maximum norm is demanding and leads to meshes which sometimes do not have desirable properties. Therefore, we will take into account as well meshes where the constant C in (3.1) will weakly depend on ε. Such a desirable property is, for example, the local quasi-uniformity of the mesh.

The simplest meshes which we discuss first are piecewise equidistant, but we also consider in detail graded meshes. Because for $x = \mu\frac{\varepsilon}{\gamma}\ln\frac{1}{\varepsilon}$ the layer term is of order $O(\varepsilon^\mu)$, a simple idea is to use a fine equidistant mesh in the interval $[0, \mu\frac{\varepsilon}{\gamma}\ln\frac{1}{\varepsilon}]$ and a coarse equidistant mesh in $[\mu\frac{\varepsilon}{\gamma}\ln\frac{1}{\varepsilon}, 1]$. But with such an so called A-mesh, in general, it is not possible to achieve uniform convergence.

Shishkin spread the idea to choose as a transition point σ from the fine to the coarse mesh the point defined by

$$\sigma = \min\{1/2, \mu\frac{\varepsilon}{\gamma}\ln N\} \tag{3.2}$$

and to subdivide each of the intervals $[0,\sigma]$ and $[\sigma,1]$ by an equidistant mesh with $N/2$ subintervals. Then, for small ε, one has $E(\sigma) = N^{-\mu}$, and μ is chosen in dependence of the order of the method used.

Remark 1 *It is not vital that one has exactly the same number of subintervals in* $[0,\sigma]$ *and* $[\sigma,1]$. *All that the theory demands is that as* $N \to \infty$ *the number of subintervals in each of these two intervals is bounded below by* CN *for some constant* $C > 0$. ♣

The coarse part of this Shishkin mesh has spacing $H = 2(1-\sigma)/N$, so $N^{-1} \leq H \leq 2N^{-1}$. The fine part has spacing $h = 2\sigma/N = (4/\gamma)\varepsilon N^{-1}\ln N$, so $h \ll \varepsilon$. Thus there is a very abrupt change in mesh size as one passes from the coarse part to the fine part. The mesh is not locally quasi-equidistant, uniformly in ε.

Remark 2 (A key property of the Shishkin mesh) *Nonequidistant meshes for convection-diffusion problems are sometimes described as "layer-resolving" meshes. One might infer from this terminology that wherever the derivatives of* u *are large, the mesh is chosen so fine that the truncation error of the difference scheme is controlled. But the Shishkin mesh does not fully resolve the layer: for*

$$|u'(x)| \approx C\varepsilon^{-1}\exp(-\gamma\, x/\varepsilon),$$

so

$$|u'(\sigma)| \approx C\varepsilon^{-1}\exp(-\mu \ln N) = C\varepsilon^{-1}N^{-\mu},$$

which can be large. That is, $|u'(x)|$ can still be large on part of the first coarse-mesh interval $[x_{N/2}, x_{N/2+1}]$.

At first sight this incomplete resolution of the boundary layer seems like a flaw, but it is in fact the key property of the mesh! Shishkin's insight was that one could achieve satisfactory theoretical and numerical results without resolving all of the layer and as a consequence his mesh permits us to use a fixed number of mesh points that is independent of ε. If one set out to "repair" the Shishkin mesh by constructing a two-stage piecewise-equidistant mesh as we have done, but with the additional requirement that the mesh be fine enough to control the local truncation error wherever the derivatives of u are very large, then the number of mesh points required would have to grow like $|\ln \varepsilon|$ as $\varepsilon \to 0$. See [130] and [38, Section 3.6].

Although the number of mesh points is fixed independently of ε, nevertheless numerical analysis on Shishkin meshes does pay a price for the nature of the construction: typically the trickiest part of the domain to handle is the first coarse mesh interval – because the derivatives of u can be large there. ♣

If a method for a problem with a smooth solution has the order α, due to the fine mesh size $h = O(\varepsilon N^{-1} \ln N)$ in the case $u^{(k)} \approx \varepsilon^{-k}$ we can expect that the error on a Shishkin mesh is of the order $O((N^{-1} \ln N)^{\alpha})$. Especially for higher order methods the logarithmic factor is not nice. An optimal mesh should generate an error of the order $O(N^{-\alpha})$.

One possibility to generate an optimal mesh is to introduce a graded mesh in the fine subinterval $[0, \sigma]$ using a *mesh-generating function*. Assuming the function $\lambda : \ [0, 1/2] \mapsto [0, \ln N]$ to be strictly increasing, set

$$x_i = \frac{\mu\varepsilon}{\gamma}\lambda(i/N), \quad i = 0, 1, \cdots, N/2.$$

We call such meshes *Shishkin-type meshes*. It turns out (see the next Section) that in error estimates for Shishkin-type meshes often the factor $\max|\psi'(\cdot)|$ appears, here ψ is the *mesh-characterizing function* defined by

$$\psi := e^{-\lambda} : \quad [0, 1/2] \mapsto [1, 1/N].$$

For the original Shishkin mesh we have $\max|\psi'(\cdot)| = O(\ln N)$. A popular optimal mesh is the *Bakhvalov-Shishkin mesh* with

$$\psi(t) = 1 - 2t(1 - N^{-1}) \quad \text{and} \quad \max|\psi'(\cdot)| \leq 2.$$

The mesh points of the fine mesh are given by

$$x_i = -\frac{\mu\varepsilon}{\gamma} \ln\left(1 - 2(1 - N^{-1})\frac{i}{N}\right), \quad i = 0, 1, \cdots, N/2. \tag{3.3}$$

For other possibilities to choose λ, see [100].

Remark 3 *In [41] we find a slight generalization of Shishkin-type meshes based on the property $\lambda(1/2) = \ln(\theta N)$ with some additional parameter θ. This allows to characterize the so called eXp-mesh as generalized Shishkin-type mesh.* ♣

If one chooses the transition point from a fine to the coarse mesh by

$$\sigma^* = \min\{1/2, \mu\frac{\varepsilon}{\gamma}\ln\frac{1}{\varepsilon}\}, \tag{3.4}$$

one has to use a graded fine mesh to achieve uniform convergence. Similarly as (3.3), a *Bakhvalov-type mesh* is given by

$$x_i = -\frac{\mu\varepsilon}{\gamma}\ln\left(1 - 2(1-\varepsilon)\frac{i}{N}\right), \quad i = 0, 1, \cdots, N/2. \tag{3.5}$$

In $[\sigma^*, 1]$ the mesh is equidistant.

Bakhvalov's original mesh [12] was a little more complicated. The mesh points near $x = 0$ were defined by

$$q(1 - \exp(-\frac{\gamma x_i}{\mu\varepsilon})) = \frac{i}{N}$$

in some interval $[0, \tau]$, here q is a parameter. That means, Bakhvalov proposed the mesh-generating function

$$\phi(t) = \begin{cases} -\frac{\mu\varepsilon}{\gamma}\ln\frac{q-t}{t} & \text{for} \quad t \in [0, \tau] \\ \phi(\tau) + \phi'(\tau)(t-\tau) & \text{for} \quad t \in [\tau, 1]. \end{cases}$$

Remarkably, τ is defined by the requirement that the mesh-generating function is C^1. Thus, τ has to solve the nonlinear equation

$$\phi'(\tau) = \frac{1 - \phi(\tau)}{1 - \tau}.$$

Bakhvalov-type meshes are simpler than the Bakhvalov meshes, the mesh-generating function is not longer C^1. But both meshes are not locally quasi-equidistant. In some cases for these meshes optimal error estimates are known, but the analysis is often more complicated than for Shishkin-type meshes. In some papers Vulanović [124–127] simplified the mesh-generating function of a Bakhvalov mesh such that the resulting equation for τ is easy to solve.

As Linß pointed out [74], a Bakhvalov-mesh can also be generated by equidistributing the monitor function

$$M(s) = \max\left(1, \tilde{K}\gamma\varepsilon^{-1}e^{-\frac{\gamma s}{\mu\varepsilon}}\right).$$

Remark 4 *In several papers (see [79] and its references) Liseikin examines the convergence of finite difference methods when using mesh generating functions $\lambda(t)$ of the given independent variable that satisfy $|\lambda'(t)| \le C$ for all $t \in [0, 1]$. This approach generates a*

graded grid of Bakhvalov type. His book [78] develops a general theory of grid generation. The analysis in these sources is written in terms of "layer-resolving transformations"; their relationship to mesh generating functions in a singular perturbation context is discussed in [128]. ♣

Remark 5 *In [115] we find the proposal to generate a mesh by the implicitly defined function*

$$\xi(t) - e^{\frac{\gamma\xi(t)}{\mu\varepsilon}} + 1 - 2t = 0. \tag{3.6}$$

The mesh has the advantage that it is not necessary to use different mesh generating functions in different regions. But (3.6) *is not so easy to solve, however, a solution based on the use of Lambert's W-function is possible. Some difference schemes (and finite elements) on that mesh can be analyzed similarly as on Bakhvalov-type meshes [107].* ♣

So far we constructed a graded fine mesh using a mesh-generating function. Alternatively, it is also relatively popular to use a *recursive formula*.

Gartland [43] graded a mesh in the following way:

$$x_0 = 0, \quad x_1 = \varepsilon H, \quad x_{i+1} = x_i + h_i$$

with

$$h_i = \min\left(H, \varepsilon H e^{\frac{\gamma x_i}{2\varepsilon}}, eh_{i-1}.\right) \tag{3.7}$$

The restriction $h_i \le eh_{i-1}$ ensures that the mesh is locally quasi-equidistant.

Remark 6 *If simple upwinding for a convection-diffusion problem is uniformly convergent in the sense of* (3.1) *for some constant $\alpha > 0$, and the mesh is locally quasi-equidistant (uniformly in ε), then the number N of mesh intervals must increase as $\varepsilon \to 0$. To see this, observe that the arguments of [117] are still valid when slightly modified by considering a limit as $N \to \infty$ with $\varepsilon \ge h_1$ and $i = 1$; one then arrives at the conclusion of that paper that $h_1 = o(\varepsilon)$. (There are some minor extra mesh assumptions such as existence of $\lim_{N\to\infty} h_1/h_2$ and $\lim_{N\to\infty} h_2/h_1$.) But the mesh diameter is at least $1/N$, so the locally quasi-equidistant property implies that $\varepsilon K^N \ge 1/N$, where K is the constant in*

$$h_i \le Kh_j \quad \text{for} \quad |i-j| \le 1.$$

Hence $NK^N \ge 1/\varepsilon$, so $N \approx \log_K(1/\varepsilon)$. ♣

Introducing the transition points x^*, x' by

$$x^* \approx K\varepsilon \ln\frac{K}{H}, \quad x' \approx K\varepsilon \ln\frac{K}{\varepsilon}$$

Gartland observed that the number of mesh points in the inner region $[0, x^*]$ as well in the outer region $[x', 1]$ is of order $O(1/H)$, but in the transition region $[x^*, x']$ of order $O(\ln \ln \frac{H}{\varepsilon})$.

We call the modification of the mesh where (3.7) is replaced by

$$h_i = \min \left(H, \varepsilon H e^{\frac{\gamma x_i}{2\varepsilon}}. \right) \tag{3.8}$$

Gartland-type mesh. The number of mesh points is now independent of ε and the mesh allows optimal error estimates. The mesh is not locally quasi-eqidistant. Gartland-type meshes are studied in [104], [83–85].

Much simpler is the *Duran-Lombardi mesh* [35] defined by

$$\begin{array}{lll} x_0 = 0, \ x_i = i\kappa H\varepsilon & \text{for} & 1 \le i \le \frac{1}{\kappa H} + 1 \\ x_{i+1} = x_i + \kappa H x_i \varepsilon & \text{for} & \frac{1}{\kappa H} + 1 \le i \le M-2, \quad x_M = 1. \end{array}$$

Here M is chosen such that $x_{M-1} < 1$ but $x_{M-1} + \kappa H x_{M-1} \ge 1$, assuming that the last interval is not extremely small.

The mesh is locally quasi-equidistant and glitters by its simplicity, there is no need to define a transition point. Uniform error estimates with respect to H are possible, but the number of mesh-points is proportional to $\frac{1}{H} \ln \frac{1}{\varepsilon}$. Numerical experiments show that the mesh is more robust than other meshes with respect to the use of a mesh designed for some given ε for larger values of the parameter.

If the mesh is given by a recursive formula, one can define a mesh-generating function by interpolation of the values given in the mesh points. That gives also a possibility to analyze discretization methods on such a mesh, see, for instance, [108].

Remark 7 *When analyzing hp finite element methods for singularly perturbed problems it is common to use an hp boundary layer mesh, see [89]. For such methods it is possible to prove exponential convergence.* ♣

3.2 The analysis of finite element methods on DL meshes

The analysis of finite element methods for standard convection-diffusion problems with exponential boundary layers for Shishkin-type meshes is well known. One obtains for linear or bilinear elements

$$\|u - u^N\|_\varepsilon \le CN^{-1} \max |\psi'| \tag{3.9}$$

with the mesh characterizing function ψ introduced in the previous subsection. Consequently, a Bakhvalov-Shishkin mesh or a Gartland-Shishkin mesh yield the optimal estimate

$$\|u - u^N\|_\varepsilon \le CN^{-1}. \tag{3.10}$$

But for Bakhvalov meshes the proof of an optimal error estimate in 2D is an open problem, the analysis in Roos/Schopf [102] still yields a weak dependence on ε.

Next we sketch as an example for the analysis on recursively defined meshes the analysis on Duran-Lombardi meshes. For the analysis presented it is sufficient to assume
for $0 \le k \le 2$

$$|\frac{\partial^k u}{\partial x^k}| \le C(1 + \varepsilon^{-k} \exp(-\beta_1 x/\varepsilon)),$$
$$|\frac{\partial^k u}{\partial y^k}| \le C(1 + \varepsilon^{-k} \exp(-\beta_2 y/\varepsilon))$$

and

$$|\frac{\partial^2 u}{\partial x \partial y}| \le C(1 + \varepsilon^{-1} \exp(-\beta_1 x/\varepsilon) + \varepsilon^{-1} \exp(-\beta_2 y/\varepsilon) + \varepsilon^{-2} \exp(-\beta_1 x/\varepsilon) \exp(-\beta_2 y/\varepsilon)).$$

(for that analysis we consider layers located at $x = 0$ and at $y = 0$)

Consider bilinear elements on a mesh defined by

$$x_0 = 0, \ x_1 = \kappa H \varepsilon$$
$$x_{i+1} = x_i + \kappa H x_i \quad \text{for} \quad 1 \le i \le M-2, \quad x_M = 1.$$

Here M is chosen such that $x_{M-1} < 1$ but $x_{M-1} + \kappa H x_{M-1} \ge 1$, assuming that the last interval is not extremely small. Remark that this mesh was the first proposal of Duran and Lombardi, the modification presented is characterized by an increasing mesh size and can be analyzed similarly. As mentioned before the number of mesh points N is proportional to $H^{-1}|\ln \varepsilon|$. Analogously we define the mesh points y_j in the y direction.

First one obtains for the interpolation error on that mesh

$$\|u - u^I\|_0 \le CH^2 \quad \text{and } \varepsilon^{1/2}|u - u^I|_1 \le CH.$$

Let us sketch, for instance, the proof of the estimate in the H^1 semi norm. Consider an element $K_{ij} = (x_{i-1}, x_i) \times (y_{j-1}, y_j)$. Then the anisotropic interpolation error estimate reads

$$\|(u - u^I)_x\|^2_{0,K_{ij}} \le C \left(h_i^2 \|u_{xx}\|^2_{0,K_{ij}} + h_j^2 \|u_{yy}\|^2_{0,K_{ij}} \right).$$

On K_{11} the desired estimate follows immediately because $h_1 = k_1 = H\varepsilon$.
Consider next the case $i, j \ge 2$. Then, the use of the recursive formula $h_i = \kappa H x_{i-1}$ allows the estimate

$$\|(u - u^I)_x\|^2_{0,K_{ij}} \le CH^2 \left(\|x u_{xx}\|^2_{0,K_{ij}} + \|y u_{yy}\|^2_{0,K_{ij}} \right).$$

Now the analysis is based on the simple observation that

$$\int_0^1 e^{-x/\varepsilon} = O(\varepsilon) \quad \text{but} \ \int_0^1 x e^{-x/\varepsilon} = O(\varepsilon^2) \quad \text{and} \ \int_0^1 x^{1/2} e^{-x/\varepsilon} = O(\varepsilon^{3/2}).$$

Using that observation our estimates for the derivatives of u yield, for instance,

$$\|u_{xx}\|_0 \leq C\varepsilon^{-3/2} \quad \text{and} \quad \|xu_{xx}\|_0 \leq C\varepsilon^{-1/2},$$

similarly

$$\|u_{xy}\|_0 \leq C\varepsilon^{-1} \quad \text{and} \quad \|y^{1/2}u_{xy}\|_0 \leq C\varepsilon^{-1/2}.$$

From these estimates the desired estimate for $\|(u-u^I)_x\|_0$ follows.
The remaining elements K_{ij} can be handled similarly.

The error analysis follows the same lines as on Shishkin meshes, the only critical part is to estimate

$$I_{conv} = \int_\Omega b\nabla(u-u^I)(u^I-u_H).$$

Introducing the subdomain Ω_3 by

$$\bar{\Omega}_3 = \cup\left\{\bar{K}_{ij}: \quad x_{i-1} \geq c\varepsilon|\ln\varepsilon|, \quad y_{j-1} \geq c\varepsilon|\ln\varepsilon|\right\}$$

we get on Ω_3

$$|\int_{\Omega_3} b\nabla(u-u^I)(u^I-u_H)| \leq CH\|u^I-u_h\|_0$$

because the behavior of u implies $|u-u^I|_{1,\Omega_3} \leq CH$.

Consider next, for example, the strip

$$\bar{\Omega}_1 = \cup\left\{\bar{K}_{ij}: \quad x_{i-1} < c\varepsilon|\ln\varepsilon|\right\}.$$

The Poincaré-Friedrichs inequality

$$\|u^I-u_H\|_{0,\Omega_1} \leq C\varepsilon\ln\frac{1}{\varepsilon}\|\frac{\partial}{\partial x}(u^I-u_H)\|_{0,\Omega_1}$$

yields

$$|\int_{\Omega_1} b\nabla(u-u^I)(u^I-u_H)| \leq C\varepsilon^{1/2}\|u^I-u\|_1\ln\frac{1}{\varepsilon}\|u^I-u_H\|_\varepsilon,$$

resulting finally in the error estimate

$$\|u-u_H\|_\varepsilon \leq CH\ln\frac{1}{\varepsilon}.$$

Introducing the number of mesh points used, one has

$$\|u-u_H\|_\varepsilon \leq CN^{-1}(\ln\frac{1}{\varepsilon})^2.$$

This estimate on the Duran-Lombardi mesh is not uniform in ε. But the analysis requires only estimates for derivatives, and two further advantages of the mesh are its simplicity (no definition of a transition point is necessary) and its robustness (a mesh defined for some ε can also be used for larger values of the parameter).

4 Numerical Methods for Systems

Consider again the system

$$L\vec{u} := -\varepsilon\vec{u}'' + B\vec{u}' + A\vec{u} = \vec{f} \quad \text{on } \Omega := (0,1), \tag{4.1a}$$

$$\vec{u}(0) = \vec{g}_0,\ \vec{u}(1) = \vec{g}_1. \tag{4.1b}$$

Based on the analytical results presented in Section 2 we now study numerical methods, extending some results from the survey [77].

In [1] the authors consider simple and midpoint upwind on a uniform mesh in a case of strong coupling, assuming B is symmetric and has the special structure

$$B = \begin{pmatrix} B^1 & 0 \\ 0 & B^2 \end{pmatrix}.$$

Here, for instance, B^1 is positive definite, B^2 negative definite. Far from the layers error estimates are proven, but, of course, one cannot obtain uniform convergence.

In [48] the Il'in-Allen-Southwell scheme is generalized for systems. Let us start from

$$-\varepsilon\vec{u}''(x_i) + B(x_i)\vec{u}'(x_i) + A(x_i)\vec{u}(x_i) = \vec{f}(x_i)$$

and assume the matrix $B(x_i)$ to be symmetric with eigenvalues λ_j^i. Then, there exists a matrix P_i which diagonalizes $B(x_i)$. Setting

$$\vec{v} = P_i^{-1}\vec{u}, \quad \vec{g} = P_i^{-1}\vec{f}$$

we obtain for the j-th component of $\vec{v}$ the equation

$$-\varepsilon v_j''(x_i) + \lambda_j^i v_j'(x_i) + \sum_k a_{jk}^* v_k(x_i) = g_j(x_i).$$

Next we introduce the scalar version of the Il'in-Allen-Southwell scheme:

$$-\varepsilon\sigma_j^i D^+D^- v_j^i + \lambda_j^i D^0 v_j^i + \sum_k a_{jk}^* v_k^i = g_j(x_i)$$

with $\sigma_j^i = \frac{\lambda_j^i h}{2}\coth\frac{\lambda_j^i h}{2\varepsilon}$. Collecting the σ_j^i in a vector, the back transformation yields

$$-\varepsilon P_i\sigma^i P_i^{-1}\vec{u^i} + B(x_i)D^0\vec{u^i} + A(x_i)\vec{u^i} = \vec{f}(x_i). \tag{4.2}$$

If B is not diagonalizable one can use the Jordan canonical form, the resulting scheme is a little more complicated. The numerical experiments in [48] promise first order uniform convergence, but there exists no analysis.

Next we sketch error estimates for layer-adapted meshes, mostly for Shishkin meshes. Of course, to create these meshes we need a priori information on the location of layers and

precise estimates on the behavior of derivatives, which we discussed for systems in Section 2.

Under the assumptions formulated in that Section for weakly coupled systems, even with different diffusion coefficients, one can analyze the upwind finite difference scheme based on the information for the first order derivatives given in (2.4). By the transformation $x := 1 - x$ we shift the layer to $x = 0$. Then, the mesh is equidistant and coarse away from $x = 0$, and piecewise equidistant with successively finer meshes, as one approaches $x = 0$. More precisely, we have mesh transition points τ_k defined by

$$\tau_{M+1} = 1, \quad \tau_k = \min\left(\frac{k\tau_{k+1}}{k+1}, \frac{\sigma\varepsilon_k}{\beta}\ln N\right) \quad \text{and} \quad \tau_0 = 0. \tag{4.3}$$

Then, for $k = 0, \cdots, M$, the mesh is obtained by dividing each of the intervals $[\tau_k, \tau_{k+1}]$ into $N/(M+1)$ subintervals of equal length.

Similarly as in the scalar case now from a $(\|\cdot\|_{\infty,d}, \|\cdot\|_{-1,\infty,d})$ stability result one can prove for simple upwind

$$\|\vec{u} - \vec{u^N}\|_{\infty,d} \le C \max_k \int_{x_{k-1}}^{x_k} (1 + |\vec{u'}|)dx,$$

and

$$\|\vec{u} - \vec{u^N}\|_{\infty,d} \le CN^{-1}\ln N \quad \text{for Shishkin meshes}$$

and first order uniform convergence for Bakhvalov meshes follows [72].

Results in the maximum norm for higher order schemes are not known. In [103] linear finite elements are analyzed in the energy norm on a modified Shishkin mesh, but also studied numerically in the maximum norm. It seems that the weak layers for higher order derivatives reduce the convergence order two in the maximum norm observed on the modified mesh if the standard Shishkin mesh is used instead.

For strongly coupled system we assume $\varepsilon_i = \varepsilon$ for all i. In the case of overlapping layers at one boundary with an analogous technique as for weakly coupled systems one can prove for the upwind scheme on a Shishkin mesh

$$\|\vec{u} - \vec{u^N}\|_{\infty,d} \le CN^{-1}\ln N,$$

see [92]. In the case $M = 2$, where both solution components do have layers at $x = 0$ and at $x = 1$, linear finite elements are analyzed in [101] in some energy norm.

Consider finally a system of reaction-diffusion equations:

$$\mathcal{L}\vec{u} := -\mathcal{E}\vec{u}'' + A\vec{u} = \vec{f} \quad \text{on } \Omega := (0,1), \tag{4.4a}$$

$$\vec{u}(0) = \vec{g}_0, \ \vec{u}(1) = \vec{g}_1, \tag{4.4b}$$

On a uniform mesh, in [66] a method is proposed which produces solutions without numerical oscillations. The method starts with a least-squares functional

$$\mathcal{F}(\vec{u}, \vec{p}) = \|D((\vec{u}, \vec{p})) - F\|_0^2$$

where $F = [0, f]^T$ and

$$D((\vec{u}, \vec{p})) = [\varepsilon_1(p_1 - \nabla u_1), \cdots, \varepsilon_M(p_M - \nabla u_M), -\mathcal{E}\nabla\vec{p} + A\vec{u}]^T.$$

A dG method is used to discretize the least-square problem with linear finite elements, see [66].

Under the assumptions formulated in Section 2 we have bounds for derivatives of the solution and can construct a layer adapted mesh. Consider the Shishkin mesh of the type (4.3), now we use that mesh at $x = 0$ and at $x = 1$.

It is well known that L_∞ stability is not sufficient to analyze the standard (central) finite difference scheme on a Shishkin mesh. For a system of two equations in [75] the maximum principle and special barrier functions are used to prove

$$\|\vec{u} - \vec{u^N}\|_{\infty,d} \leq C(N^{-1} \ln N)^2.$$

Later in the general case [76] the authors started again with the discrete L_∞ stability, but used next a representation of the consistency error via the discrete Green's function:

$$\eta_{i,j} = \sum_k \bar{h}_k G_k \varepsilon_i^2 (D^+ D^- u_i - u_i'').$$

A tricky manipulation of this representation which uses the properties of the discrete Green's function and the bounds (2.14), (2.15) leads to

$$\|\vec{u} - \vec{u^N}\|_{\infty,d} \leq C \left(\max_k \int_{x_{k-1}}^{x_k} (1 + \sum_m \varepsilon_m^{-1} B_{2\varepsilon_m}(x)) dx \right)^2.$$

On a Shishkin mesh follows

$$\|\vec{u} - \vec{u^N}\|_{\infty,d} \leq C(N^{-1} \ln N)^2.$$

It is also possible to obtain the corresponding slightly better results on Bakhvalov meshes.

Linear finite elements on layer-adapted meshes in the energy norm are analyzed in [73], robust exponential convergence of hp FEM for a set of two equations in [50]. So far error estimates for systems in a balanced norm are rare.

5 Remarks to the discontinuous Galerkin method

In [106] we presented a short introduction into the discontinuous Galerkin method because during that time the method was still not thus popular as today. Since then five books on dGFEM were published which proves the attractivity of the method [30, 32, 46, 55, 95].

We discussed 2008 mainly the primal formulation of dGFEM and, especially, the symmetric and non-symmetric method with interior penalties, SIP and NIP. In this Section we shortly introduce the very popular local discontinuous Galerkin method LDG, hybrid methods HdG and comment discontinuous Petrov-Galerkin methods dPG with optimal test functions.

5.1 The local discontinuous Galerkin method LDG

Consider the convection-diffusion problem

$$Lu := -\varepsilon\Delta u - b\cdot\nabla u + cu = f \qquad \text{in } \Omega \tag{5.1a}$$

$$u = 0 \qquad \text{on } \Gamma = \partial\Omega. \tag{5.1b}$$

Alternatively to the primal formulation one has the flux formulation of the dGFEM, which starts from the formulation of (5.1) as

$$\theta = \nabla u, \qquad -\varepsilon\nabla\cdot\theta + cu = f.$$

A corresponding weak form is

$$\int_\kappa \theta\cdot\tau = -\int_\kappa u\ \nabla\cdot\tau + \int_{\partial\kappa} u\,\mu_\kappa\cdot\tau,$$
$$\varepsilon\int_\kappa \theta\cdot\nabla v + \int_\kappa cuv = \int_\kappa fv + \int_{\partial\kappa}\theta\cdot\mu_\kappa\, v.$$

This generates the following discretization: find u_h, θ_h such that

$$\int_\kappa \theta_h\cdot\tau_h = -\int_\kappa u_h\ \nabla\cdot\tau_h + \int_{\partial\kappa}\hat{u}_\kappa\,\mu_\kappa\cdot\tau_h,$$
$$\varepsilon\int_\kappa \theta_h\cdot\nabla v_h + \int_\kappa cu_hv_h = \int_\kappa fv_h + \int_{\partial\kappa}\hat{\theta}_\kappa\cdot\mu_\kappa\, v_h.$$

Here the choice of the numerical fluxes $\hat{\theta}_\kappa$ and $\hat{u}_\kappa$ that approximate $\theta = \nabla u$ and u on $\partial\kappa$ is very important. In [11] one finds a thorough discussion of 9 variants of the dGFEM that are characterized by different choices of $\hat{\theta}_\kappa$ and $\hat{u}_\kappa$. For each of these methods, the properties of the associated primal formulation that is obtained by eliminating θ_h are discussed.

Now the very popular local discontinuous Galerkin method is characterized by the following choice of the numerical fluxes on interior edges:

$$\hat{u}_\kappa = \{u_h\} + C_{12}[u_h],$$
$$\hat{\theta}_\kappa = \{\theta_h\} - C_{12}[\theta_h] - C_{11}[u_h].$$

Here $C_{12} \geq 0$, and an adequate choice of C_{11} ensures stability of the method. See [30], Chapter 4.4 or [32], Chapter 3.3 for details and [135, 136] for the analysis of LDG for convection-diffusion problems on layer-adapted meshes.

5.2 Hybrid dG (HdG)

Introducing again $\theta = \nabla u$, we start from the mixed formulation

$$\int_\Omega \theta\cdot\tau + \int_\Omega u\,\mathrm{div}(\tau) = 0 \quad \text{for all } \tau\in H(div,\Omega),$$
$$-\varepsilon\int_\Omega \mathrm{div}(\theta)v + \int_\Omega cuv = (f,v) \quad \text{for all } v\in L^2(\Omega).$$

Now, instead of requiring the discrete fluxes to be in $H(div,\Omega)$, completely discontinuous elements can be used ensuring the continuity of the normal fluxes over element interfaces by adding appropriate constraints. The corresponding discrete problem reads:
Find $(\theta_h, u_h, \lambda_h)$ (λ_h lives on the set of faces), such that

$$\begin{aligned}\int_\Omega \theta_h \cdot \tau_h + \int_\Omega u_h \mathrm{div}(\tau_h) + \langle \lambda_h, \tau_h \cdot n\rangle_{\partial\mathcal{T}_h} &= 0,\\ -\varepsilon \int_\Omega \mathrm{div}(\theta_h) v_h + \int_\Omega c u_h v_h &= (f,v),\\ \langle \theta_h \cdot n, \mu_h\rangle_{\partial\mathcal{T}_h} &= 0.\end{aligned}$$

This method, called MH-dG [36], can alternatively be generated by the HdG method of [42] choosing the stabilization parameter used in that method suitably, see also [22,51]. In [17] the name discontinuous Petrov-Galerkin method was used for a hybrid technique, today this name is mostly related to a hybrid method with optimal test functions (see next subsection).

Let us finally notice that the *weak Galerkin method* (WG) is closely related to the HdG framework [21].

5.3 Discontinuous Petrov-Galerkin methods (dPG) with optimal test functions

Demkowicz and Gopalakrishnan developed in the years 2009-2011 a general finite element frame work which allows to combine discontinuous Galerkin methods with the early methodology proposed by Morton (mentioned in Section 2.2.2 of the book in 2008) of optimal test functions. First pure convection was studied [27], then beginning with [19] and several other papers [18, 91] also convection-diffusion was taken into consideration. For reaction-diffusion in the singularly perturbed case, see [47].

Next we sketch the basic philosophy following the survey [28]. Assume that we want to solve the following problem:
Find $u \in U$ such that

$$A(u,v) = f(v) \quad \text{for all} \quad v \in V. \tag{5.2}$$

We propose that the assumptions of the Babuska-Brezzi theory are satisfied. If $U_h \subset U$ is some finite element space, we define the test space of a Petrov-Galerkin method by $\Theta(U_h)$, where the trial-to-test operator $\Theta : U \mapsto V$ is given by

$$\langle \Theta u, v\rangle = A(u,v) \quad \text{for all} \quad v \in V. \tag{5.3}$$

Here $\langle\cdot\rangle$ denotes a scalar product in V. Then, the Petrov-Galerkin method

$$A(u_h, v_h) = f(v_h) \quad \text{for all} \quad v_h \in \Theta(U_h) \tag{5.4}$$

with $u_h \in U_h$ has the optimal property

$$\|u - u_h\|_E = \inf_{w_h \in U_h} \|u - w_h\|_E$$

with respect to the norm

$$\|w\|_E := \sup \frac{A(w,v)}{\|v\|_V}.$$

The success of this strategy depends on how easily one can compute the test functions. To get a local, element-by-element computation discontinuous methods are used. But still there exist many possibilities of, in general, hybrid formulations and choices of the broken Sobolev spaces spaces U and V and corresponding broken norms [28]. So far we described the *ideal* dPG method, practically one has to approximate the test functions. In the singularly perturbed case it is open to realize that in a robust way.

Remark that there also exist close relations of ideal dPG to a method of Cohen, Dahmen, Huang, Schwab and Welper [23, 25].

6 Adaptive methods

6.1 The stationary case

In the book 2008 we described briefly four types of estimators:

- residual estimators
- estimators based on the solution of local problems
- estimators based on averaging/pre-processing of the flux
- goal-oriented estimators (or the DWR method: dual weighted residuals).

In 2008, estimators based on averaging/pre-processing of the flux were still not thus popular, let us shortly introduce these estimators. The starting point is the introduction of some $q \in H(\mathrm{div}; \Omega)$ in the residual equation:

$$\mathrm{Res}(v) = (f + \mathrm{div}\, q, v) - (\nabla u_h - q, \nabla v) \quad \forall v \in V. \tag{6.1}$$

In general, q is designed in such a way that it satisfies the *equilibration condition*

$$\mathrm{div}\, q + Pf = 0 \tag{6.2}$$

where Pf is some projection of f. Then, for non-singularly perturbed problems the first term in (6.1) is small and the second term generates a good estimator depending on the concrete realization of q. In [16] the authors present five methods to define q and discuss the estimators generated.

For other estimators and more detailed investigations see [4,122]; for the DWR method see in particular [13].

From theoretical considerations it is clear that any good mesh for boundary or interior layers must be anisotropic. Thus an adaptive procedure designed for problems with layers should include an anisotropic refinement strategy. While several anisotropic mesh adaptation strategies do exist (see [33] and its bibliography), all are more or less heuristic. We do not know of any strategy for convection-diffusion problems in two dimensions where it is proved that, starting from some standard mesh, the refinement strategy is guaranteed to lead to a mesh that allows robust error estimates.

Micheletti, Perotto and others [26,39] combine SDFEM, the DWR method and anisotropic interpolation error estimates to get an *a posteriori* error estimate for some target functional. They then use this information to implement a metric-based algorithm for mesh generation [24] that creates an "optimal" mesh. The numerical results obtained are interesting but the second step of the approach has a heuristic flavour.

Of course, strong results require an error estimator suitable for an anisotropic mesh. Several authors use in the theoretical foundation of the estimators the *alignment measure*, introduced by Kunert [63, 64], see, for instance, [44, 94, 134]. But the use of the unknown, not computable alignment measure means that the initial mesh has more or less already to reflect the anisotropy of the solution.

Kopteva designed different estimators (residual [57, 58], flux equilibration [59]) for anisotropic meshes, even in different norms (energy and maximum norm) without using any alignment measure in proving upper error bounds.

Very important is the *robustness* of the estimators. In [109] Sangalli proves the robustness of a certain estimator for the residual-free bubble method applied to convection-diffusion problems. The analysis uses the norm

$$||w||_{San} := \|w\|_\varepsilon + \|b\cdot\nabla w\|_*, \quad \text{where} \quad \|\varphi\|_* = \sup \frac{\langle \varphi, v\rangle}{\|v\|_\varepsilon}\,. \tag{6.3}$$

Although Sangalli's approach is devoted to residual-free bubbles, the same analysis works for the Galerkin method and the SDFEM. For the convection-diffusion problem, the residual error estimator is robust with respect to the dual norm; see [121].

Angermann's graph norm and the norm $||\cdot||_{San}$ above are defined only implicitly by an infinite-dimensional variational problem and cannot be computed exactly in practice. In [111] Sangalli pointed out that the norm (6.3) seems to be not optimal in the convection-dominated regime. He proposes an improved estimator that is robust with respect to his natural norm [110] for the advection-diffusion operator, but studied only the one-dimensional case. The relation to another new improved dual norm is in detail studied in [34].

Today the dual norm or its modification plays an important role in many papers on robust a posteriori error estimation for convection-diffusion problems.

In [118], Tobiska and Verfürth proved in the dual norm that the same robust a posteriori error estimator can be used for a range of stabilized methods such as streamline diffusion, local projection schemes, subgrid-scale techniques and continuous interior penalty methods.

Nonconforming methods are studied in [134]. Variants of discontinuous Galerkin methods are discussed in [37, 44, 86, 112, 137]. Vohralik [123] presents a very general concept of a posteriori error estimation based on potential and flux reconstructions.

If one avoids to use a dual norm, often additional hypothesis are necessary to prove robustness. See, for instance, [54] for the analysis of an estimator for SDFEM in a natural SDFEM norm.

For reaction-diffusion problems the theoretical situation is clearer than for convection-diffusion problems. The residual-based estimator and the related estimator based on the solution of auxiliary local problems are both robust with respect to the associated energy norm [120]. A modification of the equilibrated residual method of Ainsworth and Babuska is also robust for reaction-diffusion problems [3].

For the flux reconstruction technique in the singularly perturbed case equation (6.1) corresponds to

$$Res(v) = (f - cu_h + \operatorname{div} q, v) - (\varepsilon \nabla u_h - q, \nabla v) \tag{6.4}$$

and one has carefully to estimate both terms of (6.4). With some numerical flux q constructed, for instance, in [20], the first term yields the residual part of the estimator

$$\eta_{T,res} := m_T \| f - cu_h + \operatorname{div} q \|_{0,T}$$

with the same weight m_T as in Verfürth's estimator. The second term generates a diffusive flux estimator η_{DF}, see [20] for details. Flux equilibration is also studied in [5, 6], for nonconforming methods see [131].

As in [134] is pointed out, the estimator of [20] is not robust on *anisotropic meshes*. A modification of that estimator is presented but the proof of the upper bound uses the alignment measure. A recent result of Kopteva [59] for flux equilibration on anisotropic meshes avoids the use of that ingredient.

Stevenson proved in [116] the uniform convergence of a special adaptive method for the reaction-diffusion equation in the energy norm.

It is unclear that the energy norm is a suitable norm for these problems because for small ε it is unable to distinguish between the typical layer function of reaction-diffusion problems and zero. It would be desirable to get robust *a posteriori* error estimates in a stronger norm, for instance, some balanced norm or the L_∞ norm.

The first result with respect to the maximum norm is the *a posteriori* error estimate of Kopteva [56] for the standard finite difference method on an arbitrary rectangular mesh. Next we sketch the ideas of [29] for a posteriori error estimation for finite elements of arbitrary order on *isotropic* meshes in the maximum norm.

Using the Green's function of the continuous operator with respect to a point x, the error in that point can be represented by

$$e(x) = \varepsilon^2 (\nabla u_h, \nabla G) + (cu_h, G).$$

For some $G_h \in V_h$ we obtain

$$e(x) = \varepsilon^2 (\nabla u_h, \nabla (G - G_h)) + (cu_h, G - G_h).$$

Integration by parts yields

$$e(x) = \frac{1}{2} \sum_{T \in \mathcal{T}_h} \int_{\partial T} \varepsilon^2 (G - G_h) n_T \cdot [\nabla u_h] + \sum_{T \in \mathcal{T}_h} (cu_h - f - \varepsilon^2 \triangle u_h, G - G_h)_T.$$

Choosing for G_h the Scott-Zhang interpolant of G, one needs sharp estimates for G to control the interpolation error. These are collected in Theorem 1 of [29]. Thus, one obtains finally with $l_h := \ln(2 + \tilde{\varepsilon} \underline{h}^{-1})$ (the constant $\tilde{\varepsilon}$ is of order ε and $\underline{h} = \min h_T$)

$$\begin{aligned} \|u - u_h\|_\infty \le C \max_{T \in \mathcal{T}_h} (&\min(\tilde{\varepsilon}, l_h h_T) \|[\nabla u_h]\|_{\infty, \partial T} \\ &+ \min(1, l_h h_T^2 \varepsilon^{-2}) \|cu_h - f - \varepsilon^2 \triangle u_h\|_{\infty, T}). \end{aligned} \tag{6.5a}$$

On *anisotropic meshes*, Kopteva also derives an a posteriori error estimator in the maximum norm [57], now for *linear* finite elements. Suppose that the triangulation satisfies the maximum angle condition. Then the first result of [57] gives

$$\begin{aligned} \|u - u_h\|_\infty \le C\, l_h \max_{z \in \mathcal{N}} (&\min(\varepsilon, h_z) \|[\nabla u_h]\|_{\infty, \partial \omega_z} \\ &+ \min(1, h_z^2 \varepsilon^{-2}) \|cu_h - f\|_{\infty, \omega_z}). \end{aligned} \tag{6.6a}$$

Here ω_z is the patch of the elements surrounding some knot z of the triangulation, h_z the diameter of ω_z. In a further estimator the second term of (6.6), which has isotropic character, is replaced by a sharper result with more anisotropic nature.

To prove (6.6) two difficulties arise. First, it is necessary to use scaled trace bounds. Moreover, instead of using the Scott-Zhang interpolant of the Green's function (which applicability is restricted on anisotropic meshes) Kopteva uses some standard Lagrange interpolant for some continuous approximation of G. But the construction is based on the following additional assumption on the mesh. Let us introduce $\Omega_1 := \{T : h_T \ge c_1 \varepsilon\}$ and $\Omega_2 := \{T : h_T \le c_2 \varepsilon\}$ with some positive $c_1 < c_2$. Then, the additional assumption requires that the distance of Ω_1 and Ω_2 is at least some $c_3 \varepsilon$ with $c_3 > 0$.

The last condition excludes an too abrupt change of the mesh size, typically for Shishkin meshes. But other layer-adapted meshes satisfy that condition, for instance, Bakhvalov meshes or Bakhvalov-Shishkin meshes.

6.2 Time-dependent problems

We start from the problem

$$Lu := u_t - \varepsilon \Delta u + b \cdot \nabla u + cu = f \quad \text{on } Q, \tag{6.7a}$$

with initial-boundary conditions

$$u(x, y, 0) = s(x, y) \quad \text{on } \Omega, \tag{6.7b}$$

$$u(x, y, t) = 0 \quad \text{on } \partial\Omega \times (0, T], \tag{6.7c}$$

assuming $c - 1/2 \nabla \cdot b \ge \beta > 0$.

The weak form of (6.7) consists in finding $u \in L^2(0,T;H_0^1(\Omega))$ such that $u_t \in L^2(0,T;H^{-1}(\Omega))$, $u(\cdot,0) = s$ in $L^2(\Omega)$, and for almost every $t \in (0,T)$ and $v \in H_0^1(\Omega)$

$$(u_t, v) + \varepsilon(\nabla u, \nabla v) + (b \cdot \nabla u + cu, v) = (f, v). \tag{6.8}$$

In this subsection we mainly discuss two approaches for obtaining a posteriori error estimates: residual estimates and estimates based on elliptic reconstruction.

Let us study the discretisation of (6.7) by linear finite elements in space and backward Euler in time. Set $\tau_n := t_n - t_{n-1}$, with every intermediate time t_n we associate a shape-regular partition $\mathcal{T}_{h,n}$ and a corresponding finite element space $X_{h,n}$. Additionally, some transition condition guarantees that the mesh at $t = t_n$ is not dramatically different from the mesh at $t = t_{n-1}$.
Then, the space-time discretisation reads: Find $u_h^n \in X_{h,n}$ such that

$$u_h^0 = \pi_h s \quad (L_2 \text{ projection}) \tag{6.9}$$

and

$$\begin{aligned}
&(\frac{u_h^n - u_h^{n-1}}{\tau_n}, v_h) + \varepsilon(\nabla u_h^n, \nabla v_h) + (b^n \cdot \nabla u_h^n + c^n u_h^n, v_h) \\
&\qquad + \sum_{K \in \mathcal{T}_{h,n}} \delta_K \big(\frac{u_h^n - u_h^{n-1}}{\tau_n} - \varepsilon \Delta u_h^n + b^n \cdot \nabla u_h^n + c^n u_h^n, b^n \cdot \nabla v_h\big)_K \\
&\qquad = (f^n, v_h) + \sum_{K \in \mathcal{T}_{h,n}} \delta_K (f^n, b^n \cdot \nabla v_h)_K. \qquad (6.10a)
\end{aligned}$$

The choice $\delta_K = 0$ yields standard Galerkin, otherwise we have SDFEM in space. Remark that we follow [122], where more general the θ-scheme in time is studied ($\theta = 1$ gives backward Euler).

The sequence $\{u_h^n\}$ defines a function $u_{h,\tau}$, piecewise affine on $(t_{n-1}, t_n]$ with $u_{h,\tau}(t_n) = u_h^n$. We equip the space

$$X = \{u \in L^2(0,T;H_0^1(\Omega)) \cap L^\infty(0,T;L^2(\Omega)) : \ u_t + b \cdot \nabla u \in L^2(0,T;H^{-1}(\Omega))\}$$

with the norm defined by

$$\|v\|_X^2 := \sup \|v(\cdot,t)\|_0^2 + \int_0^T \|v(\cdot,t)\|_\varepsilon^2 dt + \int_0^T \|(v_t + b \cdot \nabla v)(\cdot,t)\|_*^2 dt.$$

As in the stationary case, residual based a posteriori error estimation uses the equivalence of residual and error. More precisely, the norm of the residual $Res(u_{h,\tau}) \in L^2(0,T;H^{-1}(\Omega))$ defined by

$$< Res(u_{h,\tau}), v > := (f, v) - \big((\partial u_{h,\tau}, v) + \varepsilon(\nabla u_{h,\tau}, \nabla v) + (b \cdot \nabla u_{h,\tau} + c u_{h,\tau}, v)\big) \tag{6.11}$$

is equivalent to $\|u - u_{h,\tau}\|_X$, see Proposition 6.14 in [122]. Here we used

$$\partial u_{h,\tau} := \frac{u_h^n - u_h^{n-1}}{\tau_n} \quad \text{in } (t_{n-1}, t_n].$$

Next we split the residual into a temporal and a spatial part. For simplicity, we restrict ourselves to time-independent data, otherwise additional a data residual is necessary. Set

$$< Res_h(u_{h,\tau}), v >:= (f, v) - \big((\frac{u_h^n - u_h^{n-1}}{\tau_n}, v) + \varepsilon(\nabla u_h^n, \nabla v) + (b \cdot \nabla u_h^n + c u_h^n, v)\big)$$

and

$$< Res_\tau(u_{h,\tau}), v >:= \varepsilon(\nabla(u_h^n - u_{h,\tau}), \nabla v) + (b \cdot \nabla(u_h^n - u_{h,\tau}) + (c\,(u_h^n - u_{h,\tau}), v).$$

Consider first the spatial residual. Analogously as in the stationary case, elementwise residuals R_K and edge residuals R_E are introduced, and with the same weights α_K and α_E as in the stationary case one generates the estimator

$$\eta_h^n := \big\{ \sum_K \alpha_K^2 \|R_K\|^2_{L^2(K)} + \sum_E \varepsilon^{-1/2} \alpha_E^2 \|R_E\|^2_{L^2(E)} \big\}^{1/2}.$$

The direct estimation of the temporal residual yields the estimator

$$\hat{\eta}_h^n := \{\|u_h^n - u_h^{n-1}\|_\varepsilon^2 + \|b \cdot \nabla(u_h^n - u_h^{n-1})\|_*^2\}^{1/2}.$$

Together one gets a robust estimator, with, for instance, the following bound from above [122]:

$$\|u - u_{h,\tau}\|_X \le C\{\|s - \pi_h s\|_0^2 + \sum_n \tau_n((\eta_h^n)^2 + (\hat{\eta}_h^n)^2)\}^{1/2}.$$

There exist also an estimate from below.

Unfortunately, $\|\cdot\|_*$ is not computable, and standard approaches, for instance inverse inequalities, lead to estimates which are not robust with respect to ε. A similar but computable estimator in [10] is also not fully robust. Therefore, Verfürth introduces the auxiliary problem

$$(\varepsilon(\nabla \tilde{u}_h^n, \nabla v_h) + \beta(\tilde{u}_h^n, v_h) = (b \cdot \nabla(u_h^n - u_h^{n-1}), v_h). \tag{6.12}$$

One can show [122], that finally one can replace $\|b \cdot \nabla(u_h^n - u_h^{n-1})\|_*$ by the computable quantity

$$\|\tilde{u}_h^n\|_\varepsilon + \{\sum_K \alpha_K^2 \|b \cdot \nabla(u_h^n - u_h^{n-1}) + \varepsilon \Delta \tilde{u}_h^n - \beta \tilde{u}_h^n\|^2_{0,K} + \sum_E \varepsilon^{-1/2} \alpha_E \|[n \cdot \nabla \tilde{u}_h^n]_E\|^2_{0,E}\}^{1/2}.$$

The price for the final estimator is the need to solve the discrete stationary reaction-diffusion problem (6.12) at each time level.

Remark that in [118] the authors study not only SDFEM but a wide range of stabilization methods, while in [7] for semilinear equations a fully adaptive Newton-Galerkin time stepping algorithm is designed.

Next we sketch the fundamental idea to use *elliptic reconstruction operators* in a posteriori error estimation for parabolic problems. Instead of comparing directly the exact solution with the numerical one, an appropriate auxiliary function $\mathcal{R}u_{h,\tau}$ is defined. Then, we decompose the error into

$$u - u_{h,\tau} = (u - \mathcal{R}u_{h,\tau}) + (\mathcal{R}u_{h,\tau} - u_{h,\tau}).$$

The elliptic reconstruction $\mathcal{R}u_{h,\tau}$ is constructed in such a way that $u_{h,\tau}$ is the finite element solution of an elliptic problem whose exact solution is $\mathcal{R}u_{h,\tau}$. Consequently, $\mathcal{R}u_{h,\tau} - u_{h,\tau}$ can be estimated by any available a posteriori error estimator for elliptic problems. Moreover, $u - \mathcal{R}u_{h,\tau}$ satisfies a variant of the original PDE with a right-hand side that can be controlled a posteriori. Then $u - \mathcal{R}u_{h,\tau}$ can be estimated using well-known a priori estimates for the given time-dependent problem.

For simplicity let us start with the semi-discretisation of

$$u_t + Au = f \quad \text{on } Q, \tag{6.13a}$$

with initial-boundary conditions

$$u(\cdot, 0) = s \quad \text{on } \Omega,$$
$$u = 0 \quad \text{on } \partial\Omega \times (0, T]$$

by

$$(u_h)_t + A_h u_h = \pi_h f. \tag{6.14}$$

Define the elliptic reconstruction of u_h by [87]

$$A(\mathcal{R}u_h) = A_h u_h. \tag{6.15}$$

It follows:

$$a(\mathcal{R}u_h, v_h) = (A_h u_h, v_h) = a(u_h, v_h).$$

That means: u_h is the Ritz projection of $\mathcal{R}u_h$.

Next we derive an equation for $u - \mathcal{R}u_h$. Starting from

$$(u - \mathcal{R}u_h)_t + A(u - \mathcal{R}u_h) = f - (\mathcal{R}u_h)_t - A_h u_h$$

we get the error equation

$$(u - \mathcal{R}u_h)_t + A(u - \mathcal{R}u_h) = f - \pi_h f + (u_h - \mathcal{R}u_h)_t. \tag{6.16}$$

Thus, we have the two properties required: u_h is the Ritz projection of $\mathcal{R}u_h$, and $u - \mathcal{R}u_h$ satisfies a variant of the original equation with a right-hand side which can be controlled.

Based on available a posteriori information on $u_h - \mathcal{R}u_h$ in certain norms, from (6.16) we can bound $u - \mathcal{R}u_h$ using standard a priori estimates in energy norms or in the L^2 norm or, alternatively, in the L_∞ norm using estimates for the Green's function of the given problem.

As an example for discretisation in space and time we follow [88] and consider backward Euler in time:

$$\frac{U^n - U^{n-1}}{\tau_n} + A^n U^n = \pi^n f^n. \tag{6.17}$$

The continuous, piecewise linear interpolant of $\{U^N\}$ in time we denote by U, analogously we define $\mathcal{R}U$ by the interpolate of $\{\mathcal{R}^n U^n\}$. Here

$$A(\mathcal{R}^n v) = A^n v.$$

To derive the main parabolic error equation for $u - \mathcal{R}U$, we start from the discrete problem in the form

$$(\frac{U^n - U^{n-1}}{\tau_n}, \pi^n\phi) + a(U^n, \pi^n\phi) - (\pi^n f^n, \pi^n\phi) = 0$$

or

$$(\frac{U^n - \pi^n U^{n-1}}{\tau_n} + A^n U^n - \pi^n f^n, \pi^n\phi) = 0.$$

Because the quantity $(\frac{U^n - \pi^n U^{n-1}}{\tau_n} + A^n U^n - \pi^n f^n)$ lies in the finite element space, it follows

$$(\frac{U^n - \pi^n U^{n-1}}{\tau_n} + A^n U^n - f^n, \phi) = 0 \quad \forall\phi \in H_0^1(\Omega).$$

Equivalently

$$(\frac{U^n - U^{n-1}}{\tau_n} + A^n U^n - \pi^n f^n, \phi) - (\frac{\pi^n U^{n-1} - U^{n-1}}{\tau_n}, \phi) = 0 \quad \forall\phi \in H_0^1(\Omega).$$

Introducing the elliptic projection we can write

$$(\frac{U^n - U^{n-1}}{\tau_n}, \phi) + a(\mathcal{R}U^n, \phi) - (\pi^n f^n, \phi) - (\frac{\pi^n U^{n-1} - U^{n-1}}{\tau_n}, \phi) = 0. \tag{6.18}$$

From that equation we substract the continuous problem

$$(u_t, \phi) + a(u, \phi) = (f, \phi)$$

and obtain with the notation $\rho = \mathcal{R}U - u$ and $\kappa = \mathcal{R}U - U$ the error equation: for all $\phi \in H_0^1(\Omega)$ it holds on (t_{n-1}, t_n)

$$(\rho_t, \phi) + a(\rho, \phi) = (\kappa_t, \phi) + a(\mathcal{R}U - \mathcal{R}U^n, \phi) + (\pi^n f^n - f, \phi) + (\frac{\pi^n U^{n-1} - U^{n-1}}{\tau_n}, \phi). \tag{6.19}$$

So we can repeat the statement from above:
Based on available a posteriori information on $U - \mathcal{R}U$ in certain norms, from (6.19) we can bound $u - \mathcal{R}U$ using standard a priori estimates in energy norms or in the L^2 norm, alternatively in the L_∞ norm using estimates for the Green's function of the given problem. See especially [60–62] for details concerning the technique described to obtain L_∞ estimates for several discretizations in time (backward Euler, Crank-Nicolson, dG). Backward Euler in combination with discontinuous Galerkin in space is studied in [15].

Remark that a different very general framework for robust a posteriori error estimation in unsteady problems is presented in [31]. The authors use a special error measure (which cannot be computed easily in practice), but obtain upper and lower bounds even in the case where the actual numerical scheme to obtain $u_{h,\tau}$ need not to be specified. The upper bound for the error depends on an convection-diffusion flux reconstruction and its local space-time equilibration, the lower bound requires a local approximation property on this flux.

Goal-oriented a posteriori error control (the DWR method) is discussed in [113]. For SDFEM in space and discontinuous Galerkin in time the authors use the strategy first to dualize and then to stabilize.

References

[1] L. R. Abrahamsson, H. B. Keller, and H. O. Kreiss. Difference approximations for singular perturbations of systems of ordinary differential equations. *Numer. Math.*, 22:367–391, 1974.

[2] N. Ahmed, A. Linke, and C. Merdon. Towards pressure robust mixed methods for the incompressible Navier-Stokes equations. *Appl. Math.*, 2017.

[3] M. Ainsworth and I. Babuska. Reliable and robust a posteriori error estimation for singularly perturbed reaction-diffusion problems. *SIAM J. Numer. Anal.*, 36:331–353, 1999.

[4] M. Ainsworth and J. T. Oden. *A posteriori error estimation in finite element analysis.* John Wiley, New York, 2000.

[5] M. Ainsworth and T. Vejchodsky. Robust error bounds for finite element approximation of reaction-diffusion problems with non-constant reaction coefficient in arbitrary space dimension. *Comput. Math. Appl. Mech. Engrg.*, 119:219–243, 2011.

[6] M. Ainsworth and T. Vejchodsky. Robust error bounds for finite element approximation of reaction-diffusion problems with non-constant reaction coefficient in arbitrary space dimension. *Comput. Math. Appl. Mech. Engrg.*, 281:184–199, 2014.

[7] M. Amrein and T. P. Wihler. An adaptive space-time newton-galerkin approach for semilinear singularly perturbed parabolic evolution equations. *IMA J. Num. Anal.*, 37:2004–2019, 2017.

[8] V. B. Andreev. A priori estimates for solutions of singularly perturbed two-point boundary value problems. *Mat. Model.*, 14(5):5–16, 2002. Second International Conference OFEA'2001 "Optimization of Finite Element Approximation, Splines and Wavelets" (Russian) (St. Petersburg, 2001).

[9] V. B. Andreev and N. V. Kopteva. On the convergence, uniform with respect to a small parameter, of monotone three-point difference schemes. *Differential Equations*, 34:921–929, 1998.

[10] R. Araya and P. Venegas. An a posteriori error estimator for an unsteady advection-diffusion-reaction problem. *Computers and Math. with Appl.*, 66:2456–2476, 2014.

[11] D. Arnold, F. Brezzi, B. Cockburn, and D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39:1749–1779, 2002.

[12] A. S. Bakhvalov. On the optimization of methods for solving boundary value problems with boundary layers (in Russian). *Zh. Vychisl. Mat. i Mat. Fis.*, 9:841–859, 1969.

[13] W. Bangerth and R. Rannacher. *Adaptive finite element methods for differential equations*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003.

[14] G. Barrenechea, V. John, P. Knobloch, and R. Rankin. A unified analysis of algebraic flux correction schemes for convection-diffusion equations. *SeMA*, 75:655–685, 2018.

[15] A. Cangiani, E. H. Georgoulis, and S. Metcalfe. Adaptive discontinuous Galerkin methods for nonstationary convection-diffusion problems. *IMA J. Num. Anal.*, 34:1578–1597, 2014.

[16] F. Carstensen and C. Merdon. Estimator competition for Poisson problems. *J. Comput. Math.*, 28:309–330, 2010.

[17] P. Causin, R. Sacco, and C.L. Bottasso. Flux-upwind stabilization of the discontinuous Petrov-Galerkin formulation with Lagrange multipliers for advection-diffusion problems. *M2AN*, 39:1087–1114, 2005.

[18] J. Chan, J. A. Evans, and W. Qiu. A dual Petrov-Galerkin finite element method for the convection-diffusion equation. *Computers Math. with Appl.*, 68:1513–1529, 2014.

[19] J. Chan, N. Heuer, T. Bui-Than, and L. Demkowicz. A robust DPG method for convection-dominated problems ii. *Computers Math. with Appl.*, 67:771–795, 2014.

[20] I. Cheddadi, R. Fucik, M. I. Prieto, and M. Vohralik. Guaranteed and robust a posteriori error estimates for singularly perturbed reaction-diffusion problems. *Math. Modelling and Num. Anal.*, 43:867–888, 2009.

[21] G. Chen, M. Feng, and X. Xie. A robust WG method for convection-diffusion-reaction equations. *J. Comput. Appl. Math.*, 315:107–125, 2017.

[22] B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of discontinuous Galerkin, mixed and continuous methods for second order elliptic problems. *SIAM J. Num. Anal.*, 47:1319–1365, 2009.

[23] A. Cohen, W. Dahmen, and G. Welper. Adaptivity and variational stabilization for convection-diffusion equations. *ESAIM*, 46:1247–1273, 2012.

[24] F. Courty, D. Leservoisier, P. L. George, and A. Deivieux. Continuous metrics and mesh optimization. *Appl. Numer. Math.*, 56:117–145, 2006.

[25] W. Dahmen, C. Huang, C. Schwab, and G. Welper. Adaptive Petrov-Galerkin methods for first order transport equations. *SIAM J. Num. Anal.*, 50:2420–2445, 2012.

[26] L. Dede, S. Micheletti, and S. Perotto. Anisotropic error control for environmental applications. *Appl. Numer. Math.*, 58:1320–1339, 2008.

[27] L. Demkowicz and J. Gopalakrishnan. A class of discontinuous Petrov-Galerkin methods i. *Comput. Meth. Appl. Mech. Engrg.*, 199:1558–1572, 2010.

[28] L. Demkowicz and J. Gopalakrishnan. Discontinuous Petrov-Galerkin (DGP) method. ICES Report 15-20, University of Texas, 2015.

[29] A. Demlov and N. Kopteva. Maximum-norm a posteriori error estimates for singularly perturbed elliptic reaction-diffusion problems. *Numer. Math.*, 133:707–742, 2016.

[30] D. A. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*. Springer, Heidelberg, 2012.

[31] V. Dolejsi, A. Ern, and M. Vohralik. A framework for robust a posteriori error control in unsteady nonlinear advection-diffusion problems. *SIAM J. Num. Anal.*, 51:773–793, 2013.

[32] V. Dolejší and M. Feistauer. *Discontinuous Galerkin method*. Springer, Heidelberg, 2015.

[33] V. Dolejsi and J. Felcman. Anisotropic mesh adaptation for numerical solution of boundary value problems. *Num. Meth. for Part. Diff. Equ.*, 20:576–608, 2004.

[34] S. Du and Z. Zhang. A robust residual-type a posteriori error estimator for convection-diffusion equations. *J. Sci. Comput.*, 65:138–170, 2015.

[35] R. G. Durán and A. L. Lombardi. Finite element approximation of convection diffusion problems using graded meshes. *Appl. Numer. Math.*, 56:1314–1325, 2006.

[36] H. Egger and J. Schoeberl. Discontinuous Galerkin methods. *IMA J. Num. Anal.*, 30:1206–1234, 2010.

[37] A. Ern, A. F. Stephansen, and M. Vohralik. Guaranteed and robust discontinuous Galerkin a posteriori error estimates for convection-diffusion-reaction problems. *J. Comput. Appl. Math.*, 234:114–130, 2010.

[38] P. A. Farrell, A. F. Hegarty, J. J. Miller, E. O'Riordan, and G. I. Shishkin. *Robust Computational Techniques for Boundary Layers*. Chapman & Hall/CRC, Boca Raton, 2000.

[39] L. Formaggia, S. Micheletti, and S. Perotto. Anisotropic mesh adaption in computational fluid dynamics: application to the advection-diffusion-reaction and Stokes problems. *Appl. Numer. Math.*, 51:511–533, 2004.

[40] S. Franz and H.-G. Roos. Robust error estimation in energy and balanced norms for singularly perturbed fourth-order problems. *Comput. and Math. with Appl.*, 72:233–247, 2016.

[41] S. Franz and C. Xenophontos. A short note on the connection between layer-adapted exponentially graded and S-type meshes. *CMAM*, 18:199–202, 2017.

[42] G. Fu, W. Qiu, and W. Zhang. An analysis of HDG methods for convection-dominated diffusion problems. *ESAIM*, 49:225–256, 2015.

[43] E. C. Gartland. Graded-mesh difference schemes for singularly perturbed two-point boundary value problems. *Math. Comp.*, 51:631–657, 1988.

[44] S. Giani, D. Schötzau, and L. Zhu. An a posteriori error estimate for hg-adaptive DG methods for convection-diffusion problems on anisotropically refined meshes. *Computers and Math. with Appl.*, 67:869–887, 2014.

[45] H. Han and Z. Huang. The tailored finite point method. *Comput. Meth. in Appl. Math.*, 14:321–345, 2014.

[46] J.S. Hesthaven and T. Warburton. *Nodal discontinuous Galerkin methods*, volume 54 of *Texts in Applied Mathematics*. Springer, New York, 2008.

[47] N. Heuer and M. Karkulik. A robust DPG method for singularly perturbed reaction-diffusion problems. *SIAM J. Num. Anal.*, 55:1218–1242, 2017.

[48] P. W. Hsieh, S. Y. Yang, and C.S. You. A robust finite difference scheme for strongly coupled systems of singularly perturbed convection-diffusion equations. *Num. Meth. Part. Diff. Equ.*, 34:121–144, 2018.

[49] C. Xenophontos J. M. Melenk and L. Oberbroeckling. Analytic regularity for a singularly perturbed system of reaction-diffusion equations with multiple scales. arXiv:1108.2002v2, 12 Oct. 2012, 2012.

[50] C. Xenophontos J. M. Melenk and L. Oberbroeckling. Robust exponential convergence of hp FEM for singularly perturbed reaction-diffusion systems with multiple scales. *IMA J. Num. Anal.*, 33:609–628, 2013.

[51] Y. Jin. Local analysis of HDG methods for convection-dominated diffusion problems. *ZAMM*, 304:57–72, 2016.

[52] V. John, P. Knobloch, and J. Novo. Finite elements for scalar convection-dominated equations and incompressible flow problems-a never ending story? *Comput. Vis. Sci.*, pages –, 2018.

[53] V. John, A. Linke, C. Merdon, and C. Neilan. On the divergence constraint in mixed finite element methods for incompressible flows. *SIAM review*, 59:492–544, 2017.

[54] V. John and J. Novo. A robust SUPG norm a posteriori error estimator for stationary convection-diffusion equations. *Comput. Meth. Appl. Mech. Engrg.*, 255:289–305, 2013.

[55] G. Kanschat. *Discontinuous Galerkin Methods for Viscous Incompressible Flow*. Vieweg, 2008.

[56] N. Kopteva. Maximum norm a posteriori error estimate for a 2d singularly perturbed reaction-diffusion problem. *SIAM J. Numer. Anal.*, 46:1602–1618, 2008.

[57] N. Kopteva. Maximum-norm a posteriori error estimates for singularly perturbed reaction-diffusion problems on anisotropic meshes. *SIAM J. Num. Anal.*, 53:2519–2544, 2015.

[58] N. Kopteva. Energy norm a posteriori error estimates for singularly perturbed reaction-diffusion problems on anisotropic meshes. *Numer. Math.*, 137:607–642, 2017.

[59] N. Kopteva. Fully cumputable a posteriori error estimator using anisotropic flux equilibration on anisotropic meshes. pages –, 2018.

[60] N. Kopteva and T. Linß. Maximum norm a posteriori error estimation for a time-dependent reaction-diffusion problem. *Comput. Meth. Appl. Math.*, 12:189–205, 2012.

[61] N. Kopteva and T. Linß. Maximum norm a posteriori error estimation for parabolic problems using elliptic reconstructions. *SIAM J. Num. Anal.*, 51:1494–1524, 2013.

[62] N. Kopteva and T. Linß. Im proved maximum-norm a posteriori error estimates for linear and semilinear parabolic equations. *Adv. Comput. Math.*, pages –, 2017.

[63] G. Kuhnert. A posteriori error estimation for anisotropic tetrahedral and triangular finite element meshes. PhD thesis, Univ. Chemnitz, 1999.

[64] G. Kuhnert. A posteriori H1 error estimation for a singularly perturbed reaction-diffusion problem on anisotropic meshes. *IMA J. Numer. Anal.*, 25:408–428, 2005.

[65] P. Lederer, A. Linke, C. Merdon, and J. Schoebel. Divergence-free reconstruction operators for the pressure-robust stokes discretizations with continuous pressure finite elements. *SIAM J. Num. Anal.*, 55:1291–1314, 2017.

[66] R. Lin. A discontinuous Galerkin least-sqares finite element method for soling coupled singularly perturbed reaction-diffusion equations. *J. Comp. Appl. Math.*, 307:134–142, 2016.

[67] A. Linke, G. Matthies, and L. Tobiska. Robust arbitrary order mixed finite element methods for the incompressible stokes equations with pressure independent velocity errors. *ESAIM*, 50:289–309, 2016.

[68] A. Linke and C. Merdon. On velocity errors due to irrational forces in the Navier Stokes momentum balance. *J. Comput. Phys.*, 313:654–661, 2016.

[69] A. Linke and C. Merdon. Pressure-robustness and discrete helmholtz projectors in mixed finite element methods for the incompresible Navier-Stokes equations. *Comput. methods appl. mech. engineering*, 311:304–326, 2016.

[70] T. Linß. Analysis of a system of singularly perturbed convection-diffusion equations with strong coupling. Technical Report MATH-NM-02-2007, Technical University of Dresden, 2007.

[71] T. Linß. Analysis of an upwind finite-difference scheme for a system of coupled singularly perturbed convection-diffusion equations. *Computing*, 79:23–32, 2007.

[72] T. Linß. Analysis of an upwind finite-difference scheme for a system of coupled singularly perturbed convection-diffusion equations. *Computing*, 79:23–32, 2007.

[73] T. Linß. Analysis of a FEM for a coupled system of singularly perturbed reaction-diffusion equations. *Num. Algor.*, 50:283–291, 2009.

[74] T. Linß. *Layer-adapted meshes for reaction-convection-diffusion problems*. Springer, Berlin, 2010.

[75] T. Linß and N. Madden. A finite element analysis of a coupled system of singularly perturbed reaction-diffusion problems. *Appl. Math. Comp.*, 148:869–880, 2004.

[76] T. Linß and N. Madden. Layer-adapted meshes for a linear system of coupled singularly perturbed reaction-diffusion problems. *IMA J. Num. Anal.*, 29:109–125, 2009.

[77] T. Linß and M. Stynes. Numerical solution of systems of singularly perturbed differential equations. *Comput. Meth. Appl. Math.*, 9:165–191, 2009.

[78] V. D. Liseĭkin. *Grid generation methods*. Scientific Computation. Springer-Verlag, Berlin, 1999.

[79] W. D. Liseĭkin and W. E. Petrenko. An adaptive-invariant method for the numerical solution of problems with boundary and interior layers (in Russian). Computer Center, Academy of Sci., Novosibirsk, 1989.

[80] L.-B. Liu, H. Leng, and G. Long. Analysis of the SDFEM for singularly perturbed differential-difference equations. *Calcolo*, 2018.

[81] L.-B. Liu and J. Zhang. Analysis of the SDFEM in a streamline diffusion norm for singularly perturbed convection-diffusion problems. *Appl. Math. Letters*, 69:61–66, 2017.

[82] L.-B. Liu and J. Zhang. Pointwise estimates of SDFEM on Shishkin triangular meshes for problems with characteristic layers. *Numerical Algorithms*, 78:465–483, 2018.

[83] S.-T. Liu, B. Wu, and Y. Xu. High order Galerkin methods with graded meshes for two-dimensional reaction-diffusion problems. *Int. J. Num. Anal. Mod.*, 13:319–343, 2016.

[84] S.-T. Liu and Y. Xu. Galerkin methods based on Hermite splines for singular perturbation problems. *SIAM J. Num. Anal.*, 43:2607–2623, 2006.

[85] S.-T. Liu and Y. Xu. Graded Galerkin methods for the high-order convection-diffusion problem. *Num. Meth. Part. Diff. Equ.*, 25:1262–1282, 2009.

[86] A. L. Lombardi, P. Pietra, and M. Prieto. A posteriori error estimator for exponentially fitted discontinuous Galerkin approximation of advection dominated problems. *Calcolo*, 53:83–103, 2016.

[87] C. Makridakis and R. H. Nochetto. Elliptic reconstruction and a posteriori error estimates for parabolic problems. *SIAM J. Num. Anal.*, 41:1585–1594, 2003.

[88] G. Matthies and G. Lube. On streamline-diffusion methods of inf-sup stable discretisations of the generalised Oseen problem. Preprint 2007-02, Institut für Numerische und Angewandte Mathematik, Univ. Göttingen, 2007.

[89] J. M. Melenk. *hp-Finite Element Methods for Singular Perturbations.* Springer, Heidelberg, 2002.

[90] J. M. Melenk. Remarks on asymptotic expansions for singularly perturbed second order odes with multiple scales. Workshop in honor of Martin Stynes, Dresden, 2011.

[91] A. Niemi, N. O. Collier, and V. M. Calo. Automatically stable discontinuous Petrov-Galerkin methods for stationary transport problems: quasi-optimal test space norm. *Computers Math. with Appl.*, 66:2096–2113, 2013.

[92] E. O'Riordan, J. Stynes, and M. Stynes. A parameter-uniform finite difference method for a coupled system of convection-diffusion two-point boundary value problems. *Numer. Math. Theor. Meth. Appl.*, 1:176–197, 2008.

[93] E. O'Riordan and M. Stynes. Numerical analysis of a strongly coupled system of two singularly perturbed convection-diffusion problems. *Adv. Comput. Math.*, 30:101–121, 2009.

[94] M. Picasso. An anisotropic error indicator based on the Z-Z error estimator: application to elliptic and parabolic problems. *SIAM J. Sci. Comp.*, 24:1328–1355, 2003.

[95] B. Rivière. *Discontinuous Galerkin methods for solving elliptic and parabolic equations*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[96] H.-G. Roos. Robust numerical methods for singularly perturbed differential equations: a survey covering 2008-2012. *ISRN Applied Mathematics*, 2012, ID 379547, 2012.

[97] H.-G. Roos. Special features of strongly coupled systems of convection-diffusion equations with two small parameters. *Appl. Math. Letters*, 25:1127–1130, 2012.

[98] H.-G. Roos. Error estimates in balanced norms of finite element methods on layer-adapted meshes for second-order reaction-diffusion problems. In Z. Huang, M. Stynes, and Z. Zhang, editors, *Boundary and Interior Layers BAIL 2016*, pages 1–18. Springer, 2017.

[99] H.-G. Roos. Remarks on balanced norm estimates for systems of reaction-diffusion equations. *Appl. of Math.*, 63:273–280, 2018.

[100] H.-G. Roos and T. Linß. Sufficient conditions for uniform convergence on layer-adapted grids. *Computing*, 63:27–45, 1999.

[101] H.-G. Roos and C. Reibiger. Numerical analysis of strongly coupled systems of convection-diffusion equations with full layer interaction. *ZAMM*, 91:537–543, 2011.

[102] H.-G. Roos and M. Schopf. Analysis of finite element methods on bakhvalov-type meshes for linear convection-diffusion problems in 2D. *Appl. of Math.*, 57:97–108, 2012.

[103] H.-G. Roos and M. Schopf. Layer structure and the Galerkin finite element method for a system of weakly coupled singularly perturbed convection-diffusion equations with multiple scales. *ESAIM*, 49:108–117, 2015.

[104] H.-G. Roos and T. Skalický. A comparison of the finite element method on Shishkin and Gartland-type meshes for convection-diffusion problems. *CWI Quarterly*, 10:277–300, 1997.

[105] H.-G. Roos and M. Stynes. Some open problems in the numerical analysis of singularly perturbed differential equations. *CMAM*, 15:531–550, 2015.

[106] H.-G. Roos, M. Stynes, and L. Tobiska. *Robust numerical methods for singularly perturbed problems*. Springer, Berlin, 2008.

[107] H.-G. Roos, L. Teofanov, and Z. Uzelac. A modified Bakhvalov mesh. *Appl. Math. Letters*, 31:7–11, 2014.

[108] H.-G. Roos, L. Teofanov, and Z. Uzelac. Graded meshes for higher order FEM. *J. Comput. Math.*, 33:1–16, 2015.

[109] G. Sangalli. A robust a posteriori estimate for the residual free bubbles method aplied to advection-dominated problems. *Numer. Math.*, 89:379–399, 2001.

[110] G. Sangalli. A uniform analysis of nonsymmetric and coercive linear operators. *SIAM J. Math. Anal.*, 36(6):2033–2048, 2005.

[111] G. Sangalli. Robust a-posteriori estimator for advection-diffusion-reaction problems. *Math. Comp.*, 77(261):41–70, 2008.

[112] D. Schötzau and L. Zhu. A robust a posteriori error estimator for discontinuous Galerkin methods for convection-diffusion equations. *Appl. Num. Math.*, 59:2236–2255, 2009.

[113] K. Schwegler and M. Bause. Goal-oriented a posteriori error control for nonstationary convection-dominated transport problems. 2018.

[114] P. M. Selwood and A. J. Wathen. Convergence rates and classification for one-dimensional finite-element meshes. *IMA J. Numer. Anal.*, 16:65–74, 1996.

[115] G. Soederlind and A. S. Yadaw. The impact of smooth W-grids in the numerical solution of singular perturbation two-point boundary value problems. *Appl. Math. Comp.*, 218:6045–6055, 2012.

[116] R. Stevenson. The uniform saturation property for a singularly perturbed reaction-diffusion equation. *Num. Math.*, 101: 355–379, 2005.

[117] M. Stynes. A jejune heuristic mesh theorem. *Comput. Meth. Appl. Math.*, 3:488–492, 2003.

[118] L. Tobiska and R. Verfürth. Robust a posteriori error estimates for stabilized finite element methods. *IMA J. Num. Anal.*, 35:1652–1671, 2015.

[119] A. B. Vassiljewa and W. F. Butusov. *Asymptotic expansions in the theory of singular perturbations (in Russian)*. Higher School, Moscow, 1990.

[120] R. Verfürth. Robust a posteriori error estimators for a singularly perturbed reaction-diffusion equation. *Numer. Math.*, 78:479–493, 1998.

[121] R. Verfürth. Robust a posteriori error estimates for stationary convection-diffusion equations. *SIAM J. Numer. Anal.*, 43:1766–1782, 2005.

[122] R. Verfürth. *A posteriori error estimation techniques for finite element methods.* Oxford University press, Oxford, 2013.

[123] M. Vohralik. A posteriori error estimates for efficiency and error control in numerical simulations. Lecture Notes, University Pierre et Marie Curie, 2012.

[124] R. Vulanović. On a numerical solution of a type of singularly perturbed boundary value problem by using a special discretization mesh. *Univ. u Novom Sadu Zb. Rad. Prirod.-Mat. Fak. Ser. Mat.*, 13:187–201, 1983.

[125] R. Vulanović. Mesh construction for discretization of singularly perturbed boundary value problems. Doctoral dissertation. University of Novi Sad, 1986.

[126] R. Vulanović. A uniform numerical method for quasilinear singular perturbation problems without turning points. *Computing*, 41:97–106, 1989.

[127] R. Vulanović. Non-equidistant finite difference methods for elliptic singular perturbation methods. In J.J.H. Miller, editor, *Computational methods for boundary and interior layers in several dimensions*, pages 203–223. Boole Press, Dublin, 1991.

[128] R. Vulanović. The layer-resolving transformation and mesh generation for quasilinear singular perturbation problems. *J. Comput. Appl. Math.*, 203:177–189, 2007.

[129] W. Wasow. *Asymptotic expansions for ordinary differential equations.* John Wiley, New York, 1965.

[130] P. Wesseling. Uniform convergence of discretization error for a singular perturbation problem. *Numer. Methods Partial Differential Equations*, 12(6):657–671, 1996.

[131] B. Zhang, S. Chen, and J. Zhao. Guaranteed a posteriori error estimation for nonconforming finite element approximations to singularly perturbed reaction-diffusion problem. *Appl. Num. Math.*, 94:1–15, 2015.

[132] J. Zhang and L.-B. Liu. Supercloseness of the SDFEM on Shishkin triangular meshes for problems with exponential layers. *Adv. Comput. Math.*, 43:759–775, 2017.

[133] J. Zhang and L.-B. Liu. Pointwise estimates of SDFEM on Shishkin triangular meshes for problems with exponential layers. *BIT*, 58:221–246, 2018.

[134] J. Zhao and S. Chen. Guaranteed a posteriori error estimation for nonconforming finite element approximations to singularly perturbed reaction-diffusion problem. *Adv. Comput. Math.*, 40:797–818, 2014.

[135] H. Zhu and Z. Zhang. Convergence analysis of the LDG method applied to singularly perturbed problems. *Numerical Methods for Partial Differential Equations*, pages 396–421, 2012.

[136] H. Zhu and Z. Zhang. Uniform convergence of the LDG method for a singularly perturbed problem with the exponential boundary layer. *Math. Comp.*, pages 635–663, 2014.

[137] L. Zhu and D. Schötzau. A robust a posteriori error estimate for hg-adaptive DG methods. *IMA J. Num. Anal.*, 31:971–1005, 2011.