Mathematical Models and Methods in Applied Sciences
Vol. 12, No. 2 (2002) 205–228

NUMERICAL EVALUATION OF FEM WITH APPLICATION TO THE 1-D ADVECTION–DIFFUSION PROBLEM

GIANCARLO SANGALLI

Dipartimento di Matematica "F. Casorati",
Università di Pavia, Via Ferrata 1, 27100 Pavia, Italy
sangalli@dimat.unipv.it

Received 15 January 2001
Revised 15 May 2001
Communicated by F. Brezzi

In this paper we present a numerical procedure to evaluate the efficiency of finite element numerical methods. We improve some of the ideas proposed in previous works and give a partly theoretical, partly empirical justification in a general framework. The proposed procedure performs an eigenvalue computation, and requires the knowledge of the behavior of the exact operator in order to choose proper norms for the evaluations. In the experiments we focus our attention on the 1-D advection–diffusion problem: we show that our numerical procedure actually gives very sharp indications about the optimality of the tested numerical methods.

Keywords: f.e.m.; stabilization.

AMS Subject Classification: 65N30, 35Q35

1. Introduction

It is well known that there exist differential problems that cannot be solved properly with standard finite element methods. In this work we shall consider as a prototype the very simple one-dimensional advection–diffusion operator:

$$\mathcal{L}w \equiv \mathcal{L}_\varepsilon w := -\varepsilon w'' + w', \tag{1.1}$$

and the related Dirichlet boundary value problem:

$$\begin{cases} \mathcal{L}u = f \qquad \text{in } (0, 2\pi)\,, \\ u(0) = 0\,, \\ u(2\pi) = 0\,. \end{cases} \tag{1.2}$$

Though it is well known how to fix the standard Galerkin numerical approach for this model, for more complex problems which we meet in real applications — for example in computational fluid dynamics — a theoretical construction and analysis of a numerical scheme is difficult. Sometimes there are different methods for the

same problem, and the methods need to be tuned — we have to choose constants or meshes — in order to give a satisfying performance. It is usual in this case to test and compare methods by solving some particular model problems, with particular source terms.

In this work we present an automatic numerical procedure which examines a given method, and computes how far the method is from the *optimal* behavior. Our procedure is inspired by the well-known Babuška[1]–Brezzi[3] theory of finite element methods. The procedure consists of an eigenvalue computation and is related to the ones proposed by Bathe *et al.*[2] and by Sangalli[10]; here we give a justification of it based on partly theoretical and partly empirical arguments, and we apply it to the model problem (1.1)–(1.2), obtaining very sharp results.

Procedures for similar purposes were also proposed for *mixed* finite element formulations, in particular for the Stokes problem[5] and for the plate problem.[7]

In Sec. 2 we present our procedure in a general framework. In order to perform the evaluation of a given numerical scheme, we shall choose a couple of norms for which the exact variational formulation of the problem under consideration is well-posed. For the model problem we propose two different choices in Sec. 3 (fractional order norms) and in Sec. 4 (weighted norms): We develop the analysis for the continuous operator and perform evaluation on the standard Galerkin method and on the well-known SUPG method. Moreover, we shall show that either using fractional order norms or weighted norms, our procedure allows us to tune up the SUPG method: indeed, in this simple example, the best choice for the amount of *streamline diffusion* is known theoretically, and we recover the same indications.

2. The General Framework

Let W and V be Hilbert spaces; let

$$\mathcal{L} : W \to V'$$

be a linear differential operator and let

$$a : W \times V \to \mathbb{R}$$

be the related bilinear form, i.e.

$$a(w, v) := {}_{V'}\langle \mathcal{L}w, v \rangle_V \,; \tag{2.1}$$

the associated variational problem reads

$$\begin{cases} \text{find } u \in W \text{ such that} \\ a(u, v) = {}_{V'}\langle f, v \rangle_V \qquad \forall\, v \in V \,. \end{cases} \tag{2.2}$$

In the following we shall consider the advection–diffusion operator

$$\mathcal{L} \equiv \mathcal{L}_\varepsilon : H_0^1(0, 2\pi) \to H^{-1}(0, 2\pi) \,,$$

defined in (1.1), as the prototype of a parameter-dependent operator. Therefore we get:

$$a(w,v) = \varepsilon \int_0^{2\pi} w'(x)v'(x)\,dx + \int_0^{2\pi} w'(x)v(x)\,dx\,. \tag{2.3}$$

This leads to a singularly perturbed problem: When the parameter ε becomes small the solution u of (1.1) and (1.2) exhibits a boundary layer near 2π: actually in the limit case, for $\varepsilon \to 0$, the differential operator has a different order and the boundary condition in 2π becomes meaningless.

In the sequel we shall make use of norms $\|\cdot\|_S$ and $\|\cdot\|_T$ on W and V respectively, that differ from their natural norms. We assume that $(W, \|\cdot\|_S)$ and $(V, \|\cdot\|_T)$ are Hilbert spaces. We also assume that when a norm appears in a denominator, its argument does not vanish. The following well-known theorem[1] characterizes well-posed problems (2.2) in terms of $a(\cdot,\cdot)$.

Theorem 2.1. *The linear operator $\mathcal{L} : (W, \|\cdot\|_S) \to (V, \|\cdot\|_T)'$ is an algebraic and topological isomorphism if and only if the following three conditions are verified*:

$$\sup_{w\in W}\sup_{v\in V} \frac{a(w,v)}{\|w\|_S\|v\|_T} \le \kappa < +\infty\,; \tag{2.4}$$

$$\inf_{w\in W}\sup_{v\in V} \frac{a(w,v)}{\|w\|_S\|v\|_T} \ge \gamma > 0\,; \tag{2.5}$$

$$\forall\ v \in V,\ \exists\ w \in W \text{ such that } a(w,v) \neq 0\,. \tag{2.6}$$

Actually (2.4) states the continuity of $\mathcal{L}$, the inf–sup condition (2.5) states that $\mathcal{L}$ is an injective map and its rank is closed, and finally (2.6) gives the density of the rank. Then Theorem 2.1 is a variant of the Banach's closed rank theorem, and (2.4)–(2.6) also give a measure of the well-posedness of problem (2.2): for any perturbation δf of f, denoting by δu the related variation of the solution u we have

$$\frac{\|\delta u\|_S}{\|u\|_S} \le \frac{\kappa}{\gamma} \cdot \frac{\|\delta f\|_{T'}}{\|f\|_{T'}}, \tag{2.7}$$

where $\|\cdot\|_{T'}$ is the dual norm of $\|\cdot\|_T$.

The model problem here considered depends on the parameter ε and we are particularly interested in its arbitrarily small values. Actually the left-hand sides of (2.4) and (2.5) may depend on ε. Nevertheless we look for κ and γ which do not depend on it, so that (2.7) states a uniform well-posedness: for this purpose we shall define in the sequel suitable norms $\|\cdot\|_S$ and $\|\cdot\|_T$.

We therefore assume that (2.4)–(2.6) are proven as specified above, and we turn our attention to numerical methods. A generic finite element discretization reads

$$\begin{cases} \text{find } u_h \in W_h \text{ such that} \\ a_h(u_h, v_h) = {}_{V_h'}\langle \mathcal{D}_h f, v_h\rangle_{V_h}, \qquad \forall\ v_h \in V_h\,, \end{cases} \tag{2.8}$$

where $W_h \in W$ and $V_h \in V$ are spaces with the same finite dimension, a_h is a bilinear full-ranked form on $W_h \times V_h$ and the linear operator $\mathcal{D}_h : V' \to V_h'$ gives

the discrete source term. We assume that the method is consistent, i.e. when the solution of (2.2) belongs to W_h, then it also solves (2.8); in other words

$$a_h(w_h, v_h) = {}_{V'}\langle \mathcal{D}_h(\mathcal{L}w_h), v_h\rangle_V, \qquad \forall\, w_h \in W_h, \qquad \forall\, v_h \in V_h\,. \tag{2.9}$$

We present now a simple variation of the usual technique for error analysis in FEM, which has the advantage of allowing a computer implementation. We denote by $\Pi_h : W \to W_h$ the orthogonal projection onto W_h, i.e. $\Pi_h u$ is the best possible approximation of u in W_h with respect to the norm $\|\cdot\|_S$; moreover we denote by $\mathcal{P}_h : W \to W_h$ the Galerkin projection, i.e. following our previous notation, $\mathcal{P}_h(u) := u_h$. As proven by Xu and Zikatanov,[12] when u varies in W the maximum ratio between the error of the method $\|u - \mathcal{P}_h(u)\|_S$ and the error for the best approximation $\|u - \Pi_h(u)\|_S$ is given by $\|\mathcal{P}_h\|$, where

$$\|\mathcal{P}_h\| := \sup_{u\in W} \frac{\|\mathcal{P}_h u\|_S}{\|u\|_S}\,. \tag{2.10}$$

Therefore the larger is (2.10) — the worst scenario (with respect to the chosen norm $\|\cdot\|_S$ and for the chosen finite element space W_h). Our aim would be to compute (2.10), but unfortunately it is a difficult task — as we discuss in Remark 2.2.1 below — and in order to approximate it, we empirically assume that

$$\|\mathcal{P}_h\| \approx \|\mathcal{P}_{h|U_h}\| := \sup_{u\in U_h} \frac{\|\mathcal{P}_h u\|_S}{\|u\|_S}\,, \tag{2.11}$$

where $U_h \in W$ is a set of solutions related to a set $\Phi_h \in V'$ of representative source terms (i.e. $\mathcal{L}U_h = \Phi_h$). Moreover, instead of comparing directly $\|\mathcal{P}_h u\|_S$ with $\|u\|_S$, we compare $\|\mathcal{P}_h u\|_S$ with $\|\mathcal{L}u\|_{T'}$ and $\|\mathcal{L}u\|_{T'}$ with $\|u\|_S$. This last step is given by Theorem 2.1: indeed (2.4) and (2.5) give $\gamma\|u\|_S \le \|\mathcal{L}u\|_{T'} \le \kappa\|u\|_S$. Assume for a moment that we can approximate the norm $\|\cdot\|_{T'}$ with its discrete counterpart:

$$\|\mathcal{L}u\|_{T'} \approx \sup_{v_h\in V_h} \frac{{}_{V'}\langle \mathcal{L}u, v_h\rangle_V}{\|v_h\|_T}, \text{ when } \; u \in U_h, \tag{2.12}$$

and assume that, given a generic $v_h \in V_h$, we are able to find $\tilde{v}_h \in V_h$ such that

$${}_{V'}\langle \phi_h, v_h\rangle_V = {}_{V_h'}\langle \mathcal{D}_h\phi_h, \tilde{v}_h\rangle_{V_h}, \qquad \forall\, \phi_h \in \Phi_h\,; \tag{2.13}$$

then, in order to evaluate (2.11), it remains to compute the (reciprocal of the) saddle point value

$$\begin{aligned}
\inf_{u\in U_h} \frac{1}{\|\mathcal{P}_h u\|_S} \sup_{v_h\in V_h} \frac{{}_{V'}\langle \mathcal{L}u, v_h\rangle_V}{\|v_h\|_T} &= \inf_{u\in U_h} \sup_{v_h\in V_h} \frac{{}_{V_h'}\langle \mathcal{D}_h\mathcal{L}u, \tilde{v}_h\rangle_{V_h}}{\|\mathcal{P}_h u\|_S\|v_h\|_T} \\
&= \inf_{u\in U_h} \sup_{v_h\in V_h} \frac{a_h(\mathcal{P}_h u, \tilde{v}_h)}{\|\mathcal{P}_h u\|_S\|v_h\|_T} \\
&= \inf_{u_h\in W_h} \sup_{\tilde{v}_h\in V_h} \frac{a_h(u_h, \tilde{v}_h)}{\|u_h\|_S\|\tilde{v}_h\|_{T_h}}\,,
\end{aligned}$$

where we have set $\|\tilde{v}_h\|_{T_h} := \|v_h\|_T$. From (2.13) we get $\tilde{v}_h = (\mathcal{D}_h(\mathcal{R}_{|\Phi_h})^{-1})^t v_h$, where $(\cdot)^t$ denotes the *transpose* operator and $\mathcal{R} : V' \to V_h'$ denotes the restriction ${}_{V_h'}\langle \mathcal{R}f, v_h\rangle_{V_h} := {}_{V'}\langle f, v_h\rangle_V$, for all $v_h \in V_h$ — therefore $\mathcal{R}_{|\Phi_h}$ maps Φ_h into V_h'.

We give a precise statement in the next proposition, which justifies the later procedure; for the reader's convenience a detailed proof is provided.

Proposition 2.1. *Following the previous notation (with κ and γ given by* (2.4) *and* (2.5)), *assume that $\mathcal{D}_h$ and $\mathcal{R}$ are both bijective maps from $\Phi_h \in V'$ onto V_h'. Define $\|v_h\|_{T_h} := \|(\mathcal{D}_h(\mathcal{R}_{|\Phi_h})^{-1})^t\, v_h\|_T, \forall\ v_h \in V_h$. Moreover set*

$$\gamma_h := \inf_{w_h \in W_h} \sup_{v_h \in V_h} \frac{a_h(w_h, v_h)}{\|w_h\|_S \|v_h\|_{T_h}}, \tag{2.14}$$

$$\alpha_h := \inf_{\phi_h \in \Phi_h} \sup_{v_h \in V_h} \frac{{}_{V'}\langle \phi_h, v_h\rangle_V}{\|\phi_h\|_{T'} \|v\|_T}\,. \tag{2.15}$$

Then

$$\gamma \alpha_h \gamma_h^{-1} \le \|\mathcal{P}_{h|U_h}\| \le \kappa \gamma_h^{-1}\,. \tag{2.16}$$

Proof. Given a generic $u \in U_h$ we have

$$\begin{aligned}
\gamma_h \|\mathcal{P}_h u\|_S &\le \sup_{v_h \in V_h} \frac{a_h(\mathcal{P}_h u, v_h)}{\|v_h\|_{T_h}} && \text{(using (2.14))}\\
&= \sup_{v_h \in V_h} \frac{{}_{V_h'}\langle \mathcal{D}_h \mathcal{L} u, v_h\rangle_{V_h}}{\|v_h\|_{T_h}} && \text{(by definition of } \mathcal{P}_h)\\
&= \sup_{v_h \in V_h} \frac{{}_{V_h'}\langle (\mathcal{D}_h \mathcal{R}^{-1}_{|\Phi_h}) \mathcal{R}\mathcal{L} u, v_h\rangle_{V_h}}{\|v_h\|_{T_h}} && \text{(since } \mathcal{L}u \in \Phi_h)\\
&= \sup_{v_h \in V_h} \frac{{}_{V_h'}\langle \mathcal{R}\mathcal{L} u, (\mathcal{D}_h \mathcal{R}^{-1}_{|\Phi_h})^t v_h\rangle_{V_h}}{\|(\mathcal{D}_h \mathcal{R}^{-1}_{|\Phi_h})^t v_h\|_T} && \text{(by definition of } \|\cdot\|_{T_h})\\
&= \sup_{v_h \in V_h} \frac{{}_{V_h'}\langle \mathcal{R}\mathcal{L} u, v_h\rangle_{V_h}}{\|v_h\|_T} && \text{(since } (\mathcal{D}_h \mathcal{R}^{-1}_{|\Phi_h})^t \text{ is one-to-one)}\\
&\le \sup_{v \in V} \frac{{}_{V'}\langle \mathcal{L} u, v\rangle_V}{\|v\|_T} && \text{(since } V_h \subset V)\\
&= \sup_{v \in V} \frac{a(u, v)}{\|v\|_T} && \text{(using (2.1))}\\
&\le \kappa \|u\|_S && \text{(using (2.4))}
\end{aligned}$$

and the second inequality in (2.16) follows. Moreover, thanks to our hypotheses, $\mathcal{P}_h$ turns out to be a one-to-one map from U_h onto W_h; therefore, setting $\bar{u} \in U_h$ such that

$$\gamma_h \|P_h \bar{u}\|_S = \sup_{v_h \in V_h} \frac{a_h(\mathcal{P}_h \bar{u}, v_h)}{\|v_h\|_{T_h}}\,,$$

we get

$$\begin{aligned}
\gamma_h \|\mathcal{P}_h \bar{u}\|_S &= \sup_{v_h \in V_h} \frac{{}_{V'}\langle \mathcal{L}\bar{u}, v_h\rangle_V}{\|v_h\|_T} && \text{(similarly as before)} \\
&\geq \alpha_h \sup_{v \in V} \frac{{}_{V'}\langle \mathcal{L}\bar{u}, v\rangle_V}{\|v\|_T} && \text{(using (2.15))} \\
&\geq \alpha_h \gamma \|\bar{u}\|_S && \text{(using (2.5))}
\end{aligned}$$

that completes the proof of (2.16). □

In Proposition 2.1 we assumed to be able to find a proper space Φ_h, included in V', which turns out to have the same dimension of the finite element space V_h and such that the bilinear forms

$$\Phi_h \times V_h \ni (\phi_h, v_h) \mapsto {}_{V_h'}\langle \mathcal{D}_h \phi_h, v_h\rangle_{V_h} \tag{2.17}$$

and

$$\Phi_h \times V_h \ni (\phi_h, v_h) \mapsto {}_{V_h'}\langle \mathcal{R} \phi_h, v_h\rangle_{V_h}\,, \tag{2.18}$$

are both full-ranked.

Note also that the value of α_h depends only on the norm $\|\cdot\|_T$ and on the discrete spaces V_h and Φ_h; it is related to the equivalence

$$\alpha_h \|\phi_h\|_{T'} \leq \sup_{v_h \in V_h} \frac{{}_{V'}\langle \phi_h, v_h\rangle_V}{\|v_h\|_T} \leq \|\phi_h\|_{T'}, \qquad \forall\ \phi_h \in \Phi_h\,; \tag{2.19}$$

that is in fact (2.12); roughly speaking, α_h should have approximately unitary order of magnitude for a *good* choice of the norms and discrete spaces.

Given a differential operator in the context of Theorem 2.1, and given a numerical method, we can therefore evaluate its optimality for a set of different meshsizes h, or a set of different parameters (either inherent to the continuous operator or to its discretizations):

- we choose a proper set of sources Φ_h, in order to verify the hypotheses of Proposition 2.1; we also check that

  $$\alpha_h \text{ stays uniformly away from zero.} \tag{2.20}$$

  We can check (2.20) by hand, or obtain (2.15) by computer (possibly approximating it, if we cannot obtain a computable expression for $\|\cdot\|_{T'}$).
- we get γ_h by solving numerically the related saddle point problem (2.14) and we plot

  $$\rho := \kappa \gamma_h^{-1}\,. \tag{2.21}$$

  The more ρ is near to 1, the better the method behaves.

Remark 2.1. It is worth noting that our restriction $u \in U_h$ (i.e. $\mathcal{L}u \in \Phi_h$) in (2.11) and in the following analysis is due to the use of $(\mathcal{D}_h \mathcal{R}_{|\Phi_h}^{-1})^t : V_h \to V_h$ instead of $\mathcal{D}_h^t : V_h \to V$; this is because $\mathcal{D}_h^t$ could be difficult to compute, while one can easily get $(\mathcal{D}_h \mathcal{R}_{|\Phi_h}^{-1})^t$ numerically. Also note that in the proof of Proposition 2.1 we have used a weaker version of (2.4) and (2.5), with the same restriction $u \in U_h$ (instead of $u \in W$); we shall take advantage of it in Sec. 4.

We now turn our attention to the model problem (1.1)–(1.2): we consider a uniform subdivision of $(0, 2\pi)$ into open elements T_i of size h:

$$T_i \equiv T_{i,h} := \{x : (i-1)h < x < ih\}\,, \qquad \forall\ i = 1, 2, \ldots, 2\pi h^{-1}\,,$$

and the corresponding space of continuous piecewise linear elements:

$$W_h \equiv V_h := v \in H_0^1(0, 2\pi) : v|_{T_i} \text{ is affine}, \qquad \forall\ i = 1, \ldots, 2\pi h^{-1}\,;$$

in this case we shall indicate both spaces with V_h. We deal with the *standard Galerkin* (SG) method and the *streamline upwind Petrov–Galerkin* (SUPG) method: the SG formulation does not introduce any modification in the operator a in (2.3); this leads to our general framework (2.8) setting:

$$a_h(w, v) \equiv a_h^{\mathrm{SG}}(w, v) := a(w, v),$$
$$\mathcal{D}_h = \mathcal{D}_h^{\mathrm{SG}} = \mathcal{R}\,.$$

On the other hand, the SUPG method, proposed by Hughes and coworkers,[6] adds a *weighted residual stabilization*:

$$a_h(w_h, v_h) \equiv a_h^{\mathrm{SUPG}}(w_h, v_h) := a(w_h, v_h) + \sum_{i=1}^{2\pi h^{-1}} \tau \int_{T_i} (\mathcal{L}w_h)(x)\, v_h'(x)\, dx\,,$$

$${}_{V_h'}\langle \mathcal{D}_h f, v_h \rangle_{V_h} \equiv {}_{V_h'}\langle \mathcal{D}_h^{\mathrm{SUPG}} f, v_h \rangle_{V_h} := {}_{V'}\langle f, v_h \rangle_V + \sum_{i=1}^{2\pi h^{-1}} \tau \int_{T_i} f(x)\, v_h'(x)\, dx;$$

this last definition requires f to be regular in the interior of the elements: this is not restrictive for real applications, as it is not restrictive for our procedure (after all, we only require the evaluation of $\mathcal{D}_h$ on Φ_h). The amount of *streamline diffusion* τ is a parameter of the method; for 1-D problems there is an ideal choice, referred to as *exponential fitting*:

$$\tau^{\mathrm{ef}} := \frac{h}{2} \coth\left(\frac{h}{2\varepsilon}\right) - \varepsilon\,; \tag{2.22}$$

in this case the numerical solution u_h is the interpolant of the exact solution u at the mesh points x_i (see, for instance, the paper of Brezzi and Russo[4] for more details).

Proposition 2.1, although complicated at first sight, leads to a quite simple numerical procedure. We postpone the particular choice of the norms $\|\cdot\|_S$ and $\|\cdot\|_T$ to later sections: indeed the usual choice $\|\cdot\|_S \equiv \|\cdot\|_T \equiv \|\cdot\|_{H_0^1(0,2\pi)}$ is inadequate, since

we get $\gamma \approx \varepsilon^{1/2}$ in (2.5). We shall try two possibilities: first we shall consider norms of fractional order, which allow boundary layers, and then we shall use weighted norms, that cut off the layer region. Given these norms, we then set $\Phi_h = V_h$ (by the L^2 scalar product representation): obviously $(\phi_h, v_h) \mapsto \int \phi_h(x)\, v_h(x)\, dx$ is a full rank positive form and, for what concerns the SUPG method, $\mathcal{D}_h \equiv \mathcal{D}_h^{\mathrm{SUPG}}$ leads to the bilinear form $(\phi_h, v_h) \mapsto \int \phi_h(x)\, (v_h(x) + \tau v_h'(x))\, dx$, which is a skew symmetric modification of the previous one and therefore it is also positive definite. With this choice for Φ_h and V_h, (2.20) is a reasonable property of good discrete norms, that we shall discuss later. Finally, introducing a basis for V_h, we can write the matrix representations: $\mathbf{A}$ for a_h, $\mathbf{S}$ and $\mathbf{T}$ for the scalar products related to $\|\cdot\|_S$ and $\|\cdot\|_T$, $\mathbf{M}$ and $\mathbf{D}$ for the L^2 scalar product and for the representation of $\mathcal{D}_h^{\mathrm{SUPG}}$, respectively . Then, in the case of SUPG we get $\tilde{\mathbf{T}} := (\mathbf{M}^{-1}\mathbf{D})^{\mathrm{T}}\mathbf{T}(\mathbf{M}^{-1}\mathbf{D})$ as scalar product related to $\|\cdot\|_{T_h}$, by computer. Following Bathe *et al.*,[2] for the standard Galerkin method γ_h equals the square root of the smallest generalized eigenvalue $\lambda_{\min}$ in

$$\mathbf{A}^{\mathrm{T}}\mathbf{T}^{-1}\mathbf{A}\mathbf{x} = \lambda \mathbf{S}\mathbf{x}\,, \tag{2.23}$$

while for the SUPG case it is substituted by

$$\mathbf{A}^{\mathrm{T}}\tilde{\mathbf{T}}^{-1}\mathbf{A}\mathbf{x} = \lambda \mathbf{S}\mathbf{x}\,. \tag{2.24}$$

3. Fractional Order Norms

3.1. *Analysis of the continuous operator*

Since we consider a simple one-dimensional problem, we introduce fractional order norms by Fourier expansions, avoiding any theoretical and computational complication.

Given a function $v \in L^2(0, 2\pi)$ we have

$$v(x) = \frac{a_0^v}{2} + \sum_{k=1}^{\infty} a_k^v \cos(kx) + b_k^v \sin(kx)\,, \tag{3.1}$$

where

$$a_k^v = \pi^{-1} \int_0^{2\pi} v(x) \cos(kx)\, dx \qquad \forall\ k \geq 0\,,$$

$$b_k^v = \pi^{-1} \int_0^{2\pi} v(x) \sin(kx)\, dx \qquad \forall\ k \geq 1\,.$$

We define for any $v \in H_0^1(0, 2\pi)$ and $s \in (0, 1)$

$$|||v|||_s^2 := \pi \sum_{k=1}^{\infty} \left(\varepsilon^{2-2s} k^2 + k^{2s}\right) \left[(a_k^v)^2 + (b_k^v)^2\right];$$

note that

$$|||v|||_s^2 \leq 2\pi \sum_{1 \leq k \leq \varepsilon^{-1}} k^{2s} \left[(a_k^v)^2 + (b_k^v)^2\right] + \ 2\pi \sum_{k > \varepsilon^{-1}} \varepsilon^{2-2s} k^2 \left[(a_k^v)^2 + (b_k^v)^2\right]. \tag{3.2}$$

The following proposition holds true:

Proposition 3.1. *The continuity estimate* (2.4) *and the inf–sup condition* (2.5) *hold uniformly with respect to* ε *with the choice*

$$\begin{aligned} \|\cdot\|_S &:= |||\cdot|||_s\,, \\ \|\cdot\|_T &:= |||\cdot|||_{1-s}\,, \end{aligned} \tag{3.3}$$

for any $s \in (0,1)$.

Proof. First of all note that we can write the bilinear form a in terms of Fourier coefficients of its arguments:

$$a(w,v) = \pi \sum_{k=1}^{\infty} \varepsilon k^2 \left(a_k^w a_k^v + b_k^w b_k^v\right) + k\left(b_k^w a_k^v - a_k^w b_k^v\right);$$

then, using the Cauchy–Schwartz inequality for series we get

$$\begin{aligned} a(w,v) &= \pi \sum_{k=1}^{\infty} \left(\varepsilon^{1-s} k a_k^w\right)\left(\varepsilon^s k a_k^v\right) + \pi \sum_{k=1}^{\infty} \left(\varepsilon^{1-s} k b_k^w\right)\left(\varepsilon^s k b_k^v\right) \\ &\quad + \pi \sum_{k=1}^{\infty} \left(k^s b_k^w\right)\left(k^{1-s} a_k^v\right) - \pi \sum_{k=1}^{\infty} \left(k^s a_k^w\right)\left(k^{1-s} b_k^v\right) \\ &\le \|w\|_S \|v\|_T \end{aligned}$$

and (2.4) is proved, with $\kappa = 1$ independent of ε.

Now consider a generic $w \in H_0^1(0,2\pi)$ and define $\tilde{w}$ as the function with Fourier coefficients

$$a_k^{\tilde{w}} := k^{2s-1} b_k^w \qquad \forall\ k : 1 \le k \le \varepsilon^{-1}\,, \tag{3.4}$$

$$b_k^{\tilde{w}} := -k^{2s-1} a_k^w \qquad \forall\ k : 1 \le k \le \varepsilon^{-1}\,, \tag{3.5}$$

$$a_k^{\tilde{w}} := \varepsilon^{1-2s} a_k^w \qquad \forall\ k > \varepsilon^{-1}\,, \tag{3.6}$$

$$b_k^{\tilde{w}} := \varepsilon^{1-2s} b_k^w \qquad \forall\ k > \varepsilon^{-1}\,, \tag{3.7}$$

$$a_0^{\tilde{w}} := -2 \sum_{k=1}^{\infty} a_k^{\tilde{w}}\,. \tag{3.8}$$

The last condition (3.8) sets the homogeneous boundary values; this is well-posed because, by (3.6)–(3.7), the function $\tilde{w}$ belongs to $H^1(0,2\pi)$. In particular we have, using (3.2)

$$\|\tilde{w}\|_T^2 \le 2\pi \sum_{1 \le k \le \varepsilon^{-1}} k^{2s}\left[(a_k^w)^2 + (b_k^w)^2\right] + 2\pi \sum_{k > \varepsilon^{-1}} \varepsilon^{2-2s} k^2 \left[(a_k^w)^2 + (b_k^w)^2\right]. \tag{3.9}$$

Moreover

$$a(w,\tilde{w}) = \pi \sum_{1 \le k \le \varepsilon^{-1}} k^{2s}\left[(a_k^w)^2 + (b_k^w)^2\right] + \pi \sum_{k > \varepsilon^{-1}} \varepsilon^{2-2s} k^2 \left[(a_k^w)^2 + (b_k^w)^2\right]. \tag{3.10}$$

Finally, using (3.2), (3.9) and (3.10), we get

$$\sup_{v\in V}\frac{a(w,v)}{\|w\|_S\|v\|_T}\geq\frac{a(w,\tilde{w})}{\|w\|_S\|\tilde{w}\|_T}\geq\frac{1}{2} \tag{3.11}$$

and (2.5) follows, with $\gamma=1/2$. □

This completes the analysis of the behavior of the continuous operator, as indicated in Sec. 2. As a result, we obtain the uniform estimate (2.7). However, to be sure that (2.7) is really useful, we have to guarantee that there is no hidden dependence on ε in $\|f\|_{T'}$ for the source terms f we are considering. This is not obvious and actually it does not hold for all s considered up to now: the following proposition clarifies it, by comparing $\|\cdot\|_{T'}$ with $\|\cdot\|_{L^2}$. In the sequel we shall denote by C any generic constant independent of ε, h, f, u and u_h.

Proposition 3.2. *With the previous notation and $s\in(0,1/2)$ we have*

$$\|f\|_{T'}\leq C\|f\|_{L^2(0,2\pi)},\qquad\forall\ f\in L^2(0,2\pi),\qquad\forall\ \varepsilon>0\,. \tag{3.12}$$

Moreover if $s=1/2$, then

$$\|f\|_{T'}\leq C|\log\varepsilon|^{1/2}\|f\|_{L^2(0,2\pi)},\qquad\forall\ f\in L^2(0,2\pi),\qquad\forall\ \varepsilon:1/2\geq\varepsilon>0\,. \tag{3.13}$$

Proof. This is a simple consequence of the Poincaré-type inequality

$$\|v\|_{L^2(0,2\pi)}\leq\alpha\|v\|_T,\qquad\forall\ v\in H^1_0(0,2\pi) \tag{3.14}$$

with $\alpha=C$ or $\alpha=C|\log\varepsilon|^{1/2}$ depending on the condition $s\leq1/2$ or $s=1/2$. Indeed, using (3.14), we get

$$\begin{aligned}\|f\|_{T'}&=\sup_{v\in V}\frac{{}_{V'}\langle f,v\rangle_V}{\|v\|_T}\leq\alpha\sup_{v\in V}\frac{\int_0^{2\pi}f(x)v(x)\,dx}{\|v\|_T}\\&=\alpha\|f\|_{L^2(0,2\pi)}\,.\end{aligned}$$

In order to prove (3.14), we write the left-hand side in terms of the Fourier coefficients:

$$\begin{aligned}\|v\|^2_{L^2(0,2\pi)}&=\frac{\pi}{2}\,(a^v_0)^2+\pi\sum_{k=1}^{\infty}(a^v_k)^2+(b^v_k)^2\\&=I+II\,;\end{aligned}$$

for the second term we get immediately

$$II\leq C\|v\|_T\,;$$

for the first term, the homogeneous boundary conditions and the Cauchy–Schwartz inequality yield

$$
\begin{aligned}
I &= 2\pi \left(\sum_{k=1}^{\infty} a_k^v \right)^2 \\
&\leq 2\pi \sum_{k=1}^{\infty} \left(\varepsilon^{2s} k^2 + k^{2-2s} \right)^{-1} \sum_{k=1}^{\infty} \left(\varepsilon^{2s} k^2 + k^{2-2s} \right) (a_k^v)^2 \\
&\leq 2\|v\|_T^2 \sum_{k=1}^{\infty} \left(\varepsilon^{2s} k^2 + k^{2-2s} \right)^{-1} ;
\end{aligned}
$$

when $s < 1/2$ we get

$$
\sum_{k=1}^{\infty} \left(\varepsilon^{2s} k^2 + k^{2-2s} \right)^{-1} \leq \sum_{k=1}^{\infty} k^{2s-2} < C ;
$$

while, if $s = 1/2$ and $0 < \varepsilon \leq 1/2$ we get

$$
\begin{aligned}
\sum_{k=1}^{\infty} \left(\varepsilon k^2 + k \right)^{-1} &\leq (\varepsilon + 1)^{-1} + \int_1^{+\infty} \left(\varepsilon t^2 + t \right)^{-1} dt \\
&= (\varepsilon + 1)^{-1} + \log(\varepsilon^{-1}) + \log(\varepsilon + 1) \\
&\leq C \log(\varepsilon^{-1}) ,
\end{aligned}
$$

which concludes the proof. □

3.2. *Numerical evaluation of the discretization*

In Proposition 3.2 we restricted the interesting values of s to $0 < s < 1/2$; in the sequel we consider $s = 1/4$. For what concerns the computation of fractional order scalar products — i.e. the computation of $\mathbf{S}$ and $\mathbf{T}$ in (2.23) and (2.24) — we do it by truncating the series of coefficients (3.2) in order to reach a prescribed relative accuracy of 10^{-6}. Moreover, as one can expect or check by computer, condition (2.20) holds true.

In the first two tests we compare the standard Galerkin scheme and the SUPG scheme with the exponential fitting stabilization (2.22). In the case of the standard Galerkin scheme we distinguish between meshes with an even number of elements and meshes with an odd number of elements: in the first case the scheme becomes singular, and then any simple minded eigenvalue analysis of the related matrix can detect the bad behavior of the method; on the other hand, in the second case it is more difficult to recognize the unstable behavior of the scheme (see also the analysis of Bathe *et al.*[2]).

In Fig. 1 we plot the values of ρ for ε varying from 10^{-5} up to 10^{-1}, with a grid of 100 and 101 elements for the standard Galerkin method ("S.G. even" and "S.G. odd", respectively), and with a grid of 100 elements for SUPG. We can see that the value of ρ is always small for SUPG, near to the optimal value $\rho = 1$, while standard Galerkin on both meshes gives small values of ρ only when the problem is

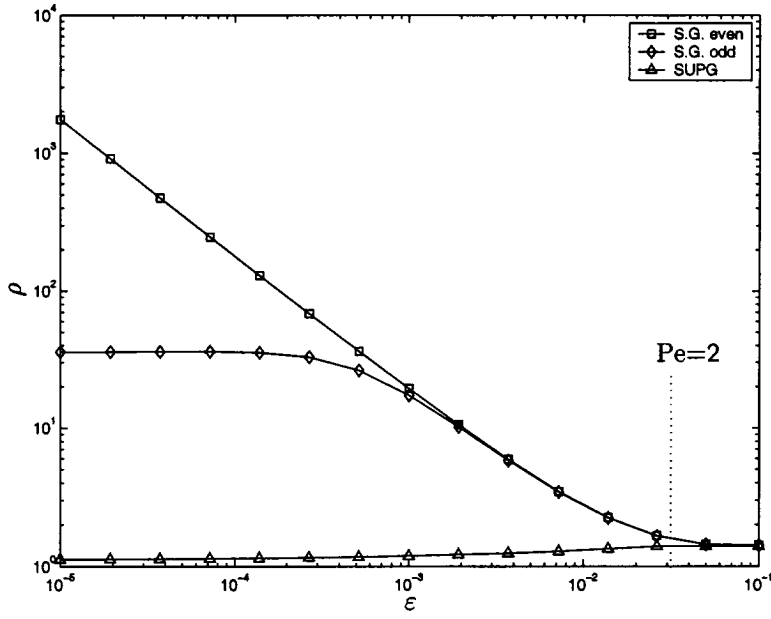

Fig. 1. Plot of ρ for 100–101 elements and ε varying.

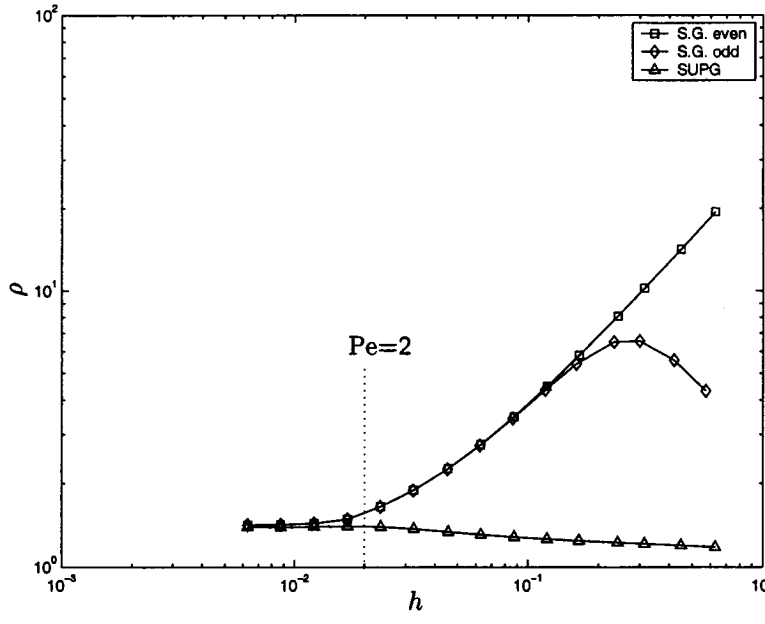

Fig. 2. Plot of ρ for 10 up to 10^3 elements and $\varepsilon = 10^{-2}$.

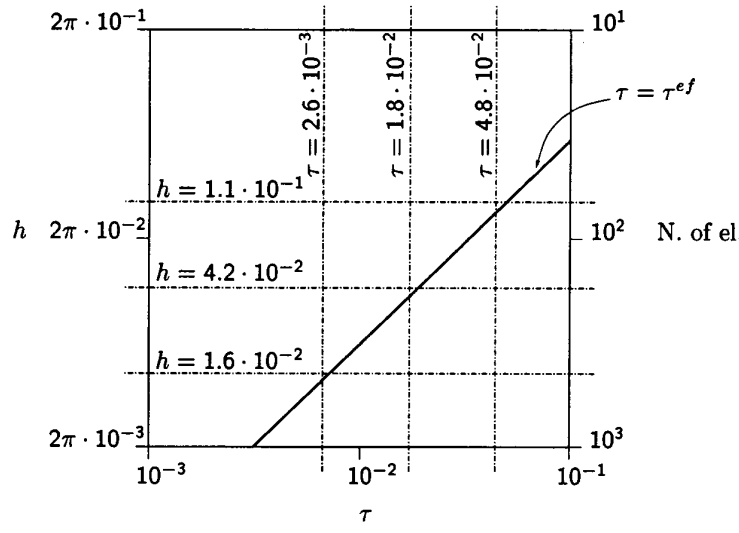

Fig. 3. SUPG: plot of the optimal stabilization $\tau = \tau^{\text{ef}}$ in the (τ, h) domain. This is the same domain for Fig. 4 (and Fig. 11), while significant sections (in dash-dotted lines) are considered in Figs. 5 and 6 (and Figs. 12 and 13).

diffusion-dominated, but not in the advection-dominated case, where the method is actually unstable. We have drawn the edge between the two regions, where the mesh Péclet number $\text{Pe} := h/\varepsilon$ is 2.

In Fig. 2 we compare the methods with $\varepsilon = 10^{-2}$ on different meshes, from 10 up to 10^3 elements. The standard Galerkin method turns out to be not accurate.

Now we focus our attention on the SUPG scheme and in particular on the effect of a change in τ. Figure 3 clarifies the subsequent Figs. 4–6: it shows a (τ, h) region, h ranging in $[2\pi \cdot 10^{-3}, 2\pi \cdot 10^{-1}]$ — namely the number of elements ranges from 10 up to 10^3 — and τ varying in $[10^{-3}, 10^{-1}]$. Then, in Fig. 4, we show the computed ρ for these τ and h, related to the SUPG method with a very small diffusion coefficient $\varepsilon = 10^{-7}$. The valley shaped surface shows a minimum region that represents the suggested optimal value of τ, depending on a given h. Our interest is to compare this minimum region with $\tau = \tau^{\text{ef}}$ — since $\varepsilon \ll h$ we can actually take $\tau^{\text{ef}} \approx h/2$, as shown in Fig. 3 — that gives a sort of theoretical optimality, as we pointed out in the introduction. Figures 5 and 6 are sections of Fig. 4, for $h = 1.1 \cdot 10^{-1}$, $4.2 \cdot 10^{-2}$, $1.6 \cdot 10^{-2}$ and respectively $\tau = 4.8 \cdot 10^{-2}$, $1.8 \cdot 10^{-2}$, $2.6 \cdot 10^{-3}$, as represented also in Fig. 3 by dash-dotted lines. As clearly shown, our procedure recognizes $\tau = \tau^{\text{ef}}$ as the optimal stabilization for SUPG.

4. Weighted Norms

4.1. *Analysis of the continuous operator*

This section presents a second approach for the analysis of (1.2). For very small

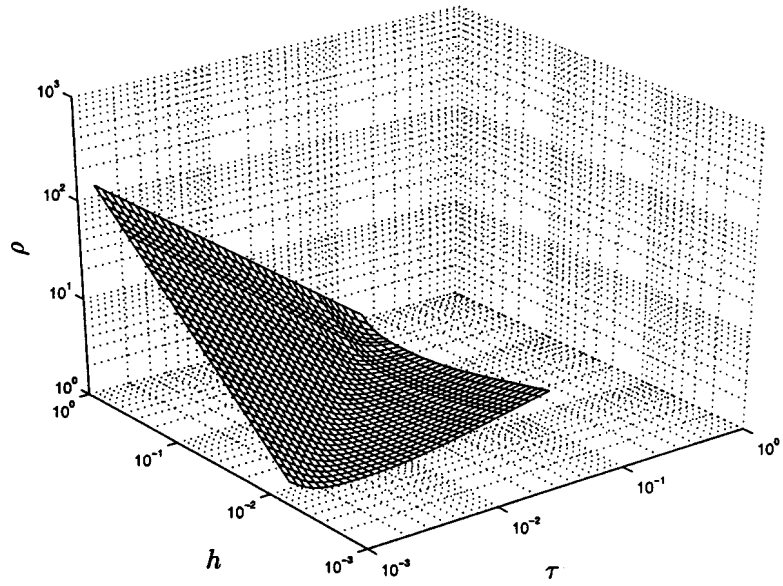

Fig. 4. SUPG: plot of ρ as a function of both τ and h.

values of ε, the actual solution u looks regular, apart from the strong boundary layer. However we know where it occurs: roughly speaking, it affects the last ε-wide part of the domain. This suggests to define $\|\cdot\|_S$ as the H^1_0-norm modified by a weight ψ that damps down the layer part. There is a large literature devoted to error estimates in weighted norms; for what concerns the advection–diffusion operator we refer e.g. to Refs. 9 and 11. The approach followed here is inspired by these works, with variations needed to recover the framework of Theorem 2.1.

Let ψ a positive function on $[0, 2\pi]$ (which, in what follows, will depend on ε and h). We define $\|\cdot\|_\psi$ as

$$\|v\|_\psi := \|v\,\psi^{1/2}\|_{L^2(0,2\pi)}\,;$$

and we define $\|v\|_{\psi^{-1}}$ in a similar way. Then we set

$$\begin{aligned} \|w\|_S^2 &:= 2\|w'\|_\psi^2, && \forall\ w \in H^1_0(0,2\pi), \\ \|v\|_T^2 &:= \varepsilon^2\|v'\|_{\psi^{-1}}^2 + \|v\|_{\psi^{-1}}^2, && \forall\ v \in H^1_0(0,2\pi)\,. \end{aligned} \tag{4.1}$$

The later analysis of the continuous operator requires the three following condition on ψ:

$$|\psi'| \le (2\varepsilon)^{-1}\psi \tag{4.2}$$

$$\int_0^{2\pi} \psi^{-1}(x)\,dx \le C\varepsilon\psi^{-1}(2\pi) \tag{4.3}$$

$$\varepsilon\|v'\|_\psi \le C\|v\|_\psi, \qquad \forall\ v \in V_h\,; \tag{4.4}$$

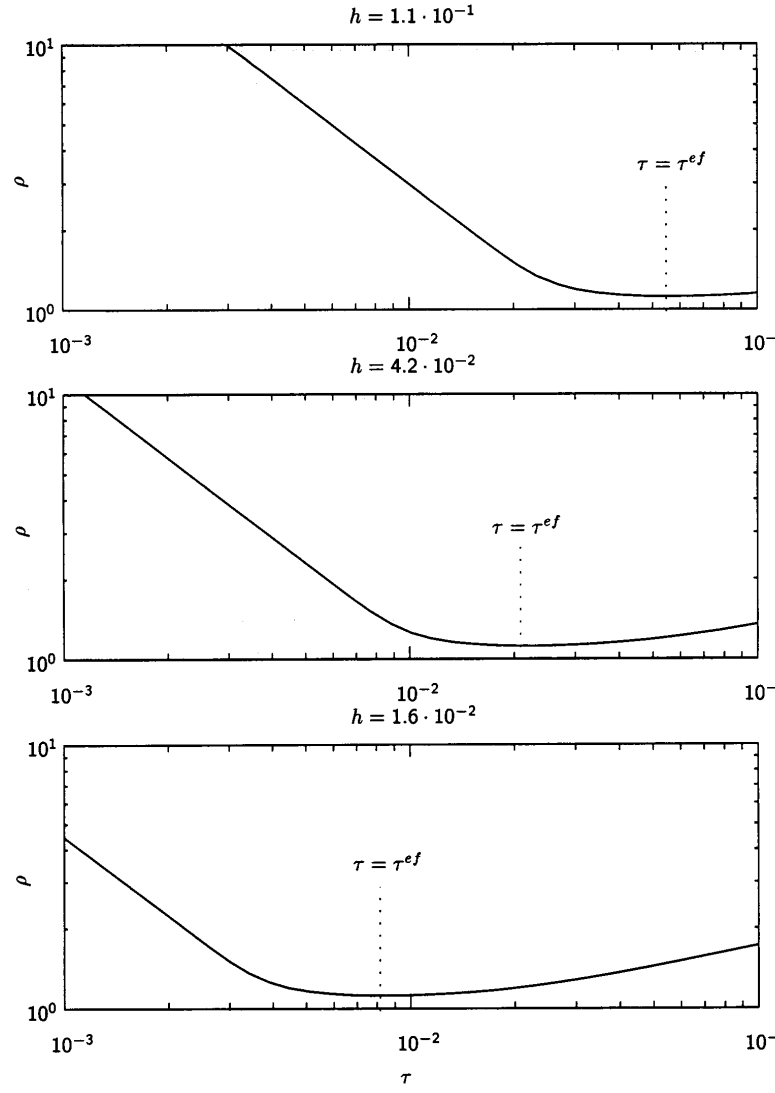

Fig. 5. SUPG: plot of ρ for $h = 1.1 \cdot 10^{-1}$, $4.2 \cdot 10^{-2}$, $1.6 \cdot 10^{-2}$ and τ varying.

moreover, in order to apply our evaluation procedure, we also need that (2.20) holds. If $\mathrm{Pe} \leq 2$ we set (see Fig. 7)

$$\psi(x) := \begin{cases} 1 & \text{if } 0 \leq x \leq \bar{x} := 2\pi + 2\sigma\varepsilon \log(2\varepsilon) \\ e^{(\bar{x}-x)/(2\varepsilon)} & \text{if } \bar{x} < x \leq 2\pi, \end{cases} \qquad (\mathrm{Pe} \leq 2 \text{ case})$$

where the parameter $\sigma > 1$ will be determined later; conditions (2.20), (4.2) and (4.3) follow by simple computation, while (4.4) holds true provided Pe is "not too

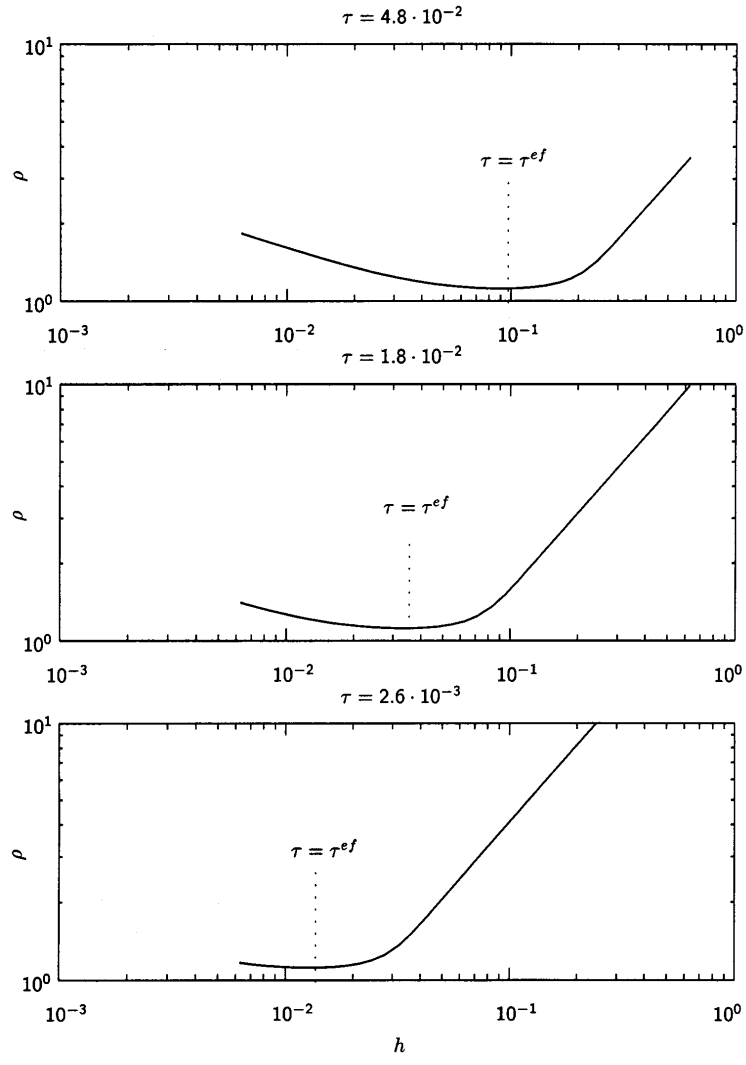

Fig. 6. SUPG: plot of ρ for $\tau = 4.8 \cdot 10^{-2}$, $1.8 \cdot 10^{-2}$, $2.6 \cdot 10^{-3}$ and h varying.

close to zero". For instance, if $\mathrm{Pe} \geq \mathrm{Pe}_0 > 0$, we have

$$\varepsilon \max_{x \in T_i} \psi(x) \leq C(\mathrm{Pe}_0) h \min_{x \in T_i} \psi(x), \qquad \forall\, i = 1, 2, \ldots, 2\pi h^{-1}\,. \tag{4.5}$$

On the other hand, when $\mathrm{Pe} > 2$, condition (2.20) becomes more delicate. It is more convenient to distinguish between the discrete layer, involving the last element, and the continuous layer. For the discrete layer we set

$$\psi_1(x) := \begin{cases} 1 & \text{if } \ 0 \leq x \leq \bar{x}_1 := 2\pi + \sigma h \log(h) \\ e^{(\bar{x}_1 - x)/h} & \text{if } \ \bar{x}_1 < x \leq 2\pi\,, \end{cases}$$

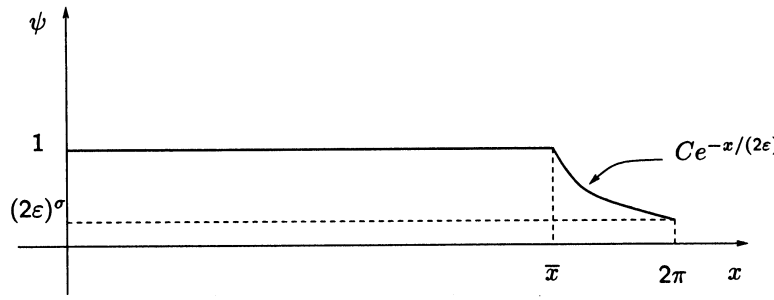

Fig. 7. Qualitative plot of ψ in the case $\mathrm{Pe} \leq 2$.

while for the continuous layer we set

$$\psi_2(x) := \begin{cases} 1 & \text{if } 0 \leq x \leq \bar{x}_2 := 2\pi + \dfrac{2\varepsilon h}{h - 2\varepsilon} \log\left(2\varepsilon h^{-1}\right) \\ e^{\frac{h-2\varepsilon}{2\varepsilon h}(\bar{x}_2 - x)} & \text{if } \bar{x}_2 < x \leq 2\pi\,, \end{cases}$$

and finally

$$\psi := \psi_1\, \psi_2; \qquad (\mathrm{Pe} > 2 \text{ case})$$

in order to clarify this last definition, note that

$$\psi := \begin{cases} 1 & \text{if } 0 \leq x \leq \bar{x}_1 \\ c_1 e^{-x/h} & \text{if } \bar{x}_1 < x \leq \bar{x}_2 \\ c_2 e^{-x/(2\varepsilon)} & \text{if } \bar{x}_2 < x \leq 2\pi\,, \end{cases} \qquad (\mathrm{Pe} > 2 \text{ case})$$

where c_1 and c_2 are proper positive constants (depending on h and ε) that give the continuity in $\bar{x}_1$ and $\bar{x}_2$ (see Fig. 8). In this case ψ also verifies (4.2) and (4.4), because of (4.5). An annoying but direct calculation yields (4.3) and (2.20), whose details have been omitted.

The continuity estimate (2.4) with respect to weighted norms follows easily from the Cauchy–Schwartz inequality, as stated in the following proposition.

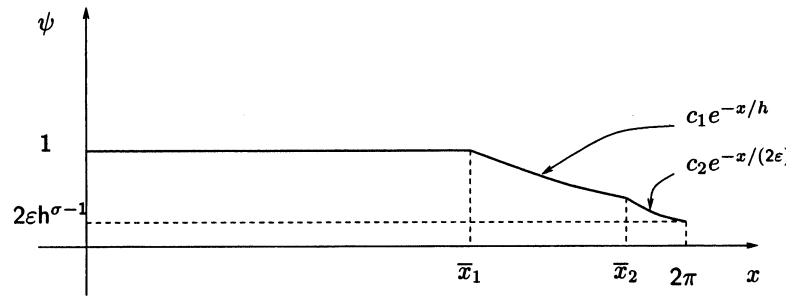

Fig. 8. Qualitative plot of ψ in the case $\mathrm{Pe} > 2$.

Proposition 4.1. *Assuming* (4.1) *we have*

$$|a(w,v)| \leq \|w\|_S \|v\|_T, \qquad \forall\ w,v \in H_0^1(0,2\pi)\,.$$

Moreover, we can prove the following stability estimate:

Lemma 4.1. *Assuming* (4.1)*and* $h \leq 1$ *we have*

$$\|w\|_S \leq C\|\mathcal{L}w\|_\psi, \qquad \forall\ w \in H_0^1(0,2\pi) \cap H^2(0,2\pi). \tag{4.6}$$

Proof. In the diffusion dominated regime, say $\varepsilon \geq 1/4$, (4.6) is a straightforward consequence of the elliptic regularity, since $0 < C \leq \psi \leq 1$. Now we focus our attention on the case $\varepsilon < 1/4$. We have

$$\begin{aligned}\|\mathcal{L}w\|_\psi^2 &= \int_0^{2\pi} (\varepsilon w''(x))^2 \psi(x)\,dx + \int_0^{2\pi} (w'(x))^2 \psi(x) dx \\ &\quad - \int_0^{2\pi} \varepsilon(2w'(x)w''(x))\psi(x)\,dx \geq \|w\|_S^2 - \int_0^{2\pi} \varepsilon((w'(x))^2)'\psi(x)\,dx\,.\end{aligned} \tag{4.7}$$

Integrating by parts, the last addendum yields

$$\begin{aligned}-\int_0^{2\pi} \varepsilon((w'(x))^2)'\psi(x)dx &= \int_0^{2\pi} \varepsilon(w'(x))^2\psi'(x)dx - [\varepsilon(w'(x))^2\psi(x)]_0^{2\pi} \\ &= I + II\,.\end{aligned} \tag{4.8}$$

Using (4.2) we get

$$|I| \leq \frac{1}{2}\|w\|_S^2\,, \tag{4.9}$$

while, by our assumption on ε, we have $\psi(2\pi) \leq 1/2$ whence

$$II \geq \psi(2\pi)\left(2\varepsilon(w'(0))^2 - \varepsilon(w'(2\pi))^2\right)\,. \tag{4.10}$$

Note that, since w assumes homogeneous boundary conditions, we have

$$\varepsilon w'(2\pi) - \varepsilon w'(0) = \int_0^{2\pi} \mathcal{L}w(x)\,dx\,; \tag{4.11}$$

then

$$\begin{aligned}(\varepsilon w'(2\pi))^2 &= \left(\varepsilon w'(0) + \int_0^{2\pi} \mathcal{L}w(x)\,dx\right)^2 \\ &\leq 2(\varepsilon w'(0))^2 + 2\left(\int_0^{2\pi} \mathcal{L}w(x)\,dx\right)^2\end{aligned} \tag{4.12}$$

and, using the Cauchy–Schwartz inequality and (4.3), we obtain

$$\begin{aligned}\left(\int_0^{2\pi} \mathcal{L}w(x)\,dx\right)^2 &\leq \int_0^{2\pi} (\mathcal{L}w(x))^2\psi(x)\,dx \cdot \int_0^{2\pi} \psi^{-1}(x)\,dx \\ &\leq C\epsilon\psi^{-1}(2\pi)\|\mathcal{L}w\|_\psi^2\,;\end{aligned} \tag{4.13}$$

hence (4.12) and (4.13) yield

$$(\varepsilon w'(2\pi))^2 - 2(\varepsilon w'(0))^2 \le C\epsilon\psi^{-1}(2\pi)\|\mathcal{L}w\|_\psi^2\,, \tag{4.14}$$

that is

$$II \ge -C\|\mathcal{L}w\|_\psi^2\,. \tag{4.15}$$

Finally (4.7)–(4.9) and (4.15) give (4.6). □

We can now prove the inf–sup condition (2.5): we restrict ourselves to the case $\mathrm{Pe} \ge 1/2$ and $w \in U_h$ (i.e. $\mathcal{L}w \in \Phi_h \equiv V_h$); as mentioned in Remark 2.1, this is not restrictive for our analysis.

Proposition 4.2. *Assuming* (4.1), $h \le 1$ *and* $\mathrm{Pe} \ge 1/2$, *we have*

$$\inf_{w\in U_h} \sup_{v\in H_0^1} \frac{a(w,v)}{\|w\|_S\|v\|_T} \ge \gamma > 0\,; \tag{4.16}$$

with γ *independent of* ε *and* h.

Proof. Let $t(\cdot,\cdot)$ be the scalar product associated with the norm $\|\cdot\|_T$, i.e.

$$t(v_1,v_2) := \varepsilon^2 \int_0^{2\pi} v_1'(x)v_2'(x)\psi^{-1}(x)\,dx + \int_0^{2\pi} v_1(x)v_2(x)\psi^{-1}(x)\,dx\,,$$

and let $\eta = \eta(w)$ be the solution of the variational problem:

$$t(\eta, w) = a(w,v), \qquad \forall\ v \in V_h\,. \tag{4.17}$$

Then we can express the sup in (4.16) in terms of η:

$$\inf_{w\in U_h} \sup_{v\in V} \frac{a(w,v)}{\|w\|_S\|v\|_T} = \inf_{w\in U_h} \sup_{v\in V} \frac{t(\eta,v)}{\|w\|_S\|v\|_T} = \inf_{w\in U_h} \frac{\|\eta\|_T}{\|w\|_S}\,, \tag{4.18}$$

and we have to prove that

$$\|w\|_S \le C\|\eta\|_T\,. \tag{4.19}$$

We choose $v = \mathcal{L}w\,\psi$ in (4.17). Notice that $w \in U_h$ implies $\mathcal{L}w\,\psi \in H_0^1(0,2\pi)$. Integrating by parts we have

$$\begin{aligned}
\|\mathcal{L}w\|_\psi^2 &= a(w, \mathcal{L}w\,\psi) \\
&= t(\eta, \mathcal{L}w\psi) \\
&= \varepsilon^2 \int_0^{2\pi} \eta'(x)(\mathcal{L}w)'(x)\,dx \\
&\quad + \varepsilon^2 \int_0^{2\pi} \eta'(x)\mathcal{L}w(x)\psi'(x)\psi^{-1}(x)\,dx \\
&\quad + \int_0^{2\pi} \eta(x)\mathcal{L}w(x)\,dx \\
&= I + II + III\,.
\end{aligned}$$

Using the Cauchy–Schwartz inequality and (4.4) we obtain

$$I \leq \|\varepsilon\eta'\|_{\psi^{-1}} \|\varepsilon(\mathcal{L}w)'\|_{\psi} \leq \|\eta\|_T \|\mathcal{L}w\|_{\psi} ;$$

using (4.2) and the Cauchy–Schwartz inequality yields

$$II \leq \frac{\varepsilon}{2} \int_0^{2\pi} |\eta'(x)||\mathcal{L}w| \, dx \leq \|\eta\|_T \|\mathcal{L}w\|_{\psi} ,$$

while, simply using the Cauchy–Schwartz inequality

$$III \leq \|\eta\|_{\psi^{-1}} \|\mathcal{L}w\|_{\psi} \leq \|\eta\|_T \|\mathcal{L}w\|_{\psi} .$$

Hence $\|\mathcal{L}w\|_{\psi} \leq C\|\eta\|_T$, and Lemma 4.1 gives us (4.19). □

4.2. *Numerical evaluation of the discretizations*

The weight ψ defined in the previous section depends on a parameter σ that can influence its effect: the bigger is σ, the stronger is the damping in the layer region, as shown in Figs. 7 and 8. Here we use $\sigma = 5$, but different choices of σ, say $2 \leq \sigma \leq 8$, give similar qualitative results (in the sense that we still clearly recognize the merits and demerits of the methods under analysis).

The results we are showing are the counterpart of the ones of Sec. 3. In Figs. 9 and 10 we compare the standard Galerkin method and the SUPG method for

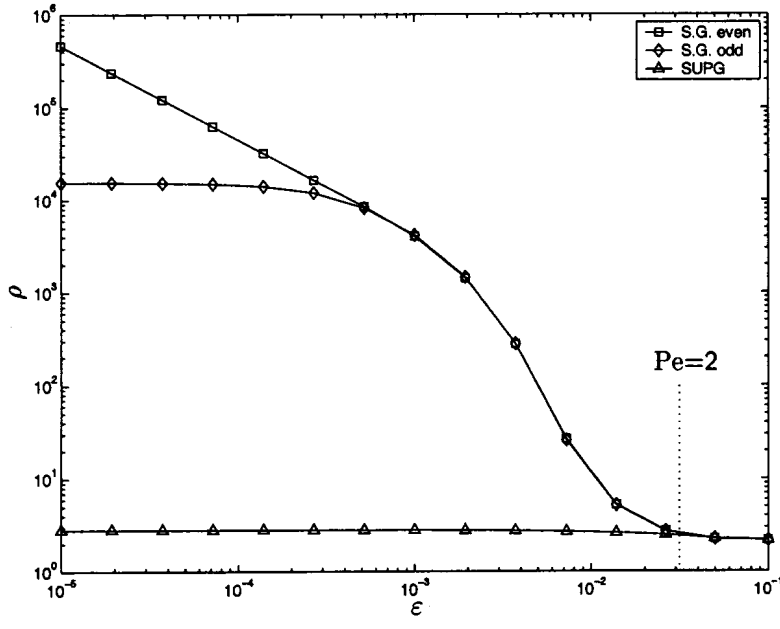

Fig. 9. Plot of ρ for 100–101 elements and ε varying.

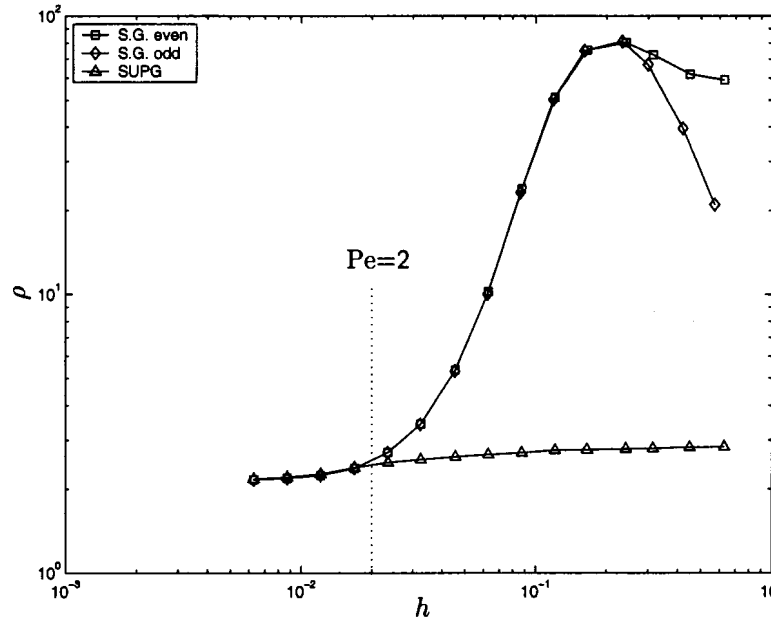

Fig. 10. Plot of ρ for 10 up to 10^3 elements and $\varepsilon = 10^{-2}$.

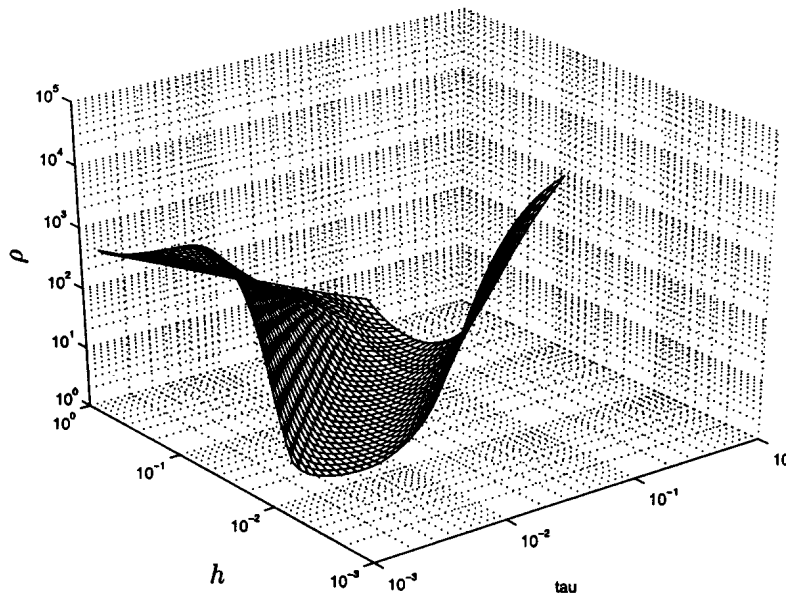

Fig. 11. SUPG: plot of ρ as a function of both τ and h.

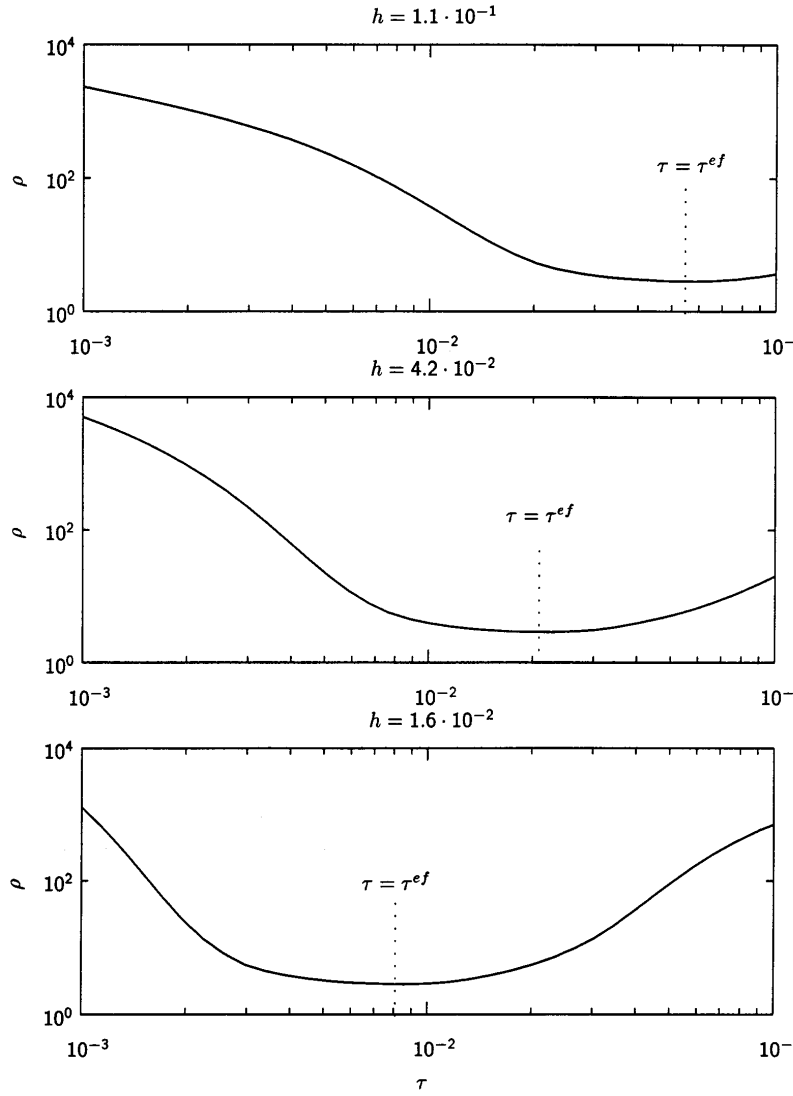

Fig. 12. SUPG: plot of ρ for $h = 1.1 \cdot 10^{-1}$, $4.2 \cdot 10^{-2}$, $1.6 \cdot 10^{-2}$ and τ varying.

different values of ε (from 10^{-5} up to 10^{-1}) and for different number of elements (from 10 up to 10^3) respectively. We can draw the same conclusion, and in particular we see that here there is a larger gap between good and bad methods: this is due to the particular structure of the norms used, that allow a stronger control outside the layer region. Figures 11–13 deal with the effect of τ in SUPG. Keeping in mind Fig. 3, it can be seen in Fig. 11 and in its significant sections plotted as Figs. 12–13 that again our procedure recognizes the optimal behavior of SUPG for $\tau = \tau^{\text{ef}}$.

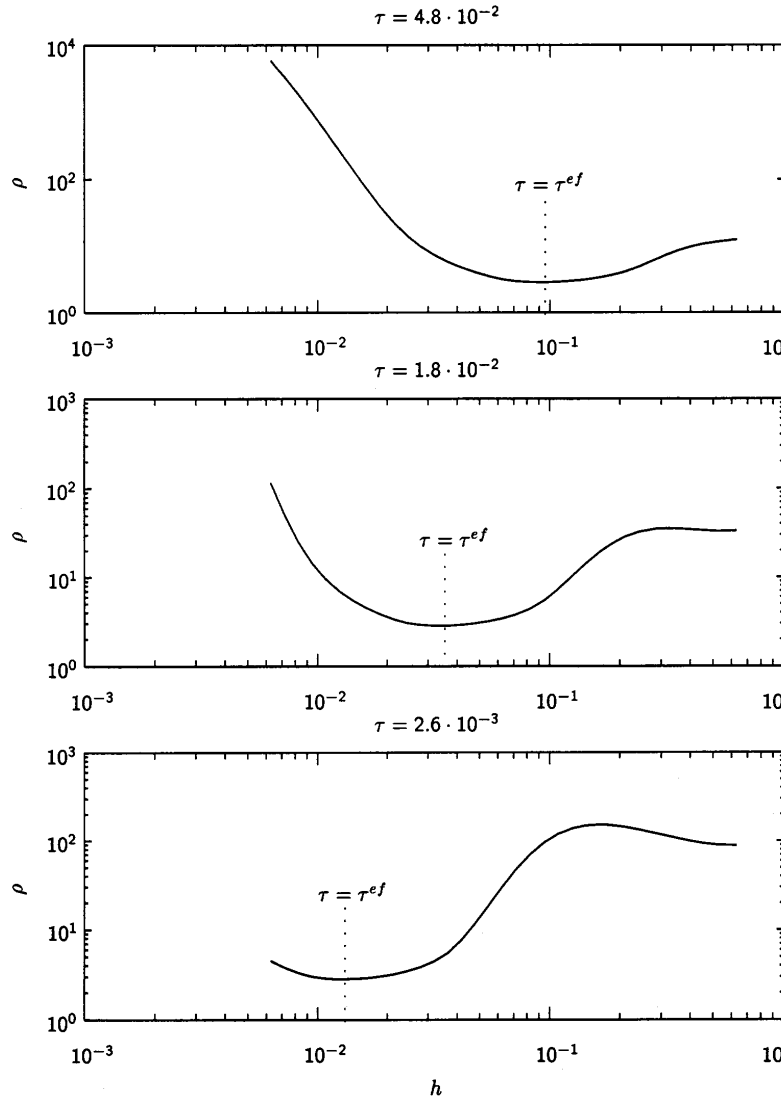

Fig. 13. SUPG: plot of ρ for $\tau = 4.8 \cdot 10^{-2}$, $1.8 \cdot 10^{-2}$, $2.6 \cdot 10^{-3}$ and h varying.

5. Conclusion

In this work we have proposed an automatic procedure to evaluate the accuracy of various finite element discretizations of linear differential problems. We have tested it on the one-dimensional advection–diffusion differential operator, for advection-dominated regimes. The paper is mostly devoted to the analysis of the behavior of this operator: besides being interesting in view of numerical methods, it is also a preparatory step for our procedure. We have proposed two different analyses

of this operator in Secs. 3 and 4. In both cases we have introduced a class of norms, depending on parameters, which brings the advection–diffusion operator into the framework of Theorem 2.1, uniformly in ε. This leads to two different implementations of the procedure.

We point out the main improvements of the present procedure with respect to the previous ones[2,10]:

- after the analysis of the continuous operator, Proposition 2.1 gives a systematic way to obtain the discrete procedure,
- we have focused the crucial points in order to obtain a very sharp evaluation.

As a consequence, our procedure goes well beyond the rough classification (Standard Galerkin = bad, SUPG = good), but is able to distinguish, in SUPG methods, among the different values of the stabilizing parameter τ, and split the optimal one. In this respect, we can say that the present analysis allows a much finer evaluation of the quality of the different methods.

References

1. I. Babuška and A. K. Aziz, *Survey lectures on the mathematical foundations of the finite element method*, in *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*, ed. A. K. Aziz (Academic Press, 1972).
2. K. J. Bathe, F. Brezzi, D. Hendriana and G. Sangalli, *Inf–sup testing of up-wind methods*, *Internat. J. Numer. Methods Engrg.* **48** (2000) 745–760.
3. F. Brezzi, *On the existence, uniqueness and approximation of saddle point problems arising from Lagrangian multipliers*, *R.A.I.R.O. Anal. Numer.* **8** (1974) 129–151.
4. F. Brezzi and A. Russo, *Choosing bubbles for advection–diffusion problems*, *Math. Models Methods Appl. Sci.* **4** (1994) 571–587.
5. D. Chapelle and K. J. Bathe, *The inf–sup test*, *Comput. Structures* **47** (1993) 537–545.
6. L. P. Franca, S. L. Frey and T. J. R. Hughes, *Stabilized finite element methods: I. Application to the advective–diffusive model*, *Comput. Methods Appl. Mech. Engrg.* **95** (1992) 253–276.
7. A. Iosilevich, K. J. Bathe and F. Brezzi, *On evaluating the inf–sup condition for plate bending elements*, *Internat. J. Numer. Methods Engrg.* **40** (1997) 3639–3663.
8. C. Johnson, A. H. Schatz and L. B. Wahlbin, *Crosswind smear and pointwise errors in streamline diffusion finite element methods*, *Math. Comp.* **49** (1987) 25–38.
9. C. Johnson and U. Nävert, *An analysis of some finite element methods for advection–diffusion problems*, in *Analytical and Numerical Approaches to Asymptotic Problems in Analysis* (Proc. Conf., Univ. Nijmegen, Nijmegen, 1980) (North-Holland, 1981), pp. 99–116.
10. G. Sangalli, *Numerical evaluation of finite element methods for convection–diffusion problems*, *Calcolo* **37** (2000) 233–251.
11. G. Sangalli, *Global and local error analysis for the residual-free bubble method applied to advection-dominated problems*, *SIAM J. Numer. Anal.* **38** (2000) 1496–1522.
12. J. Xu and L. Zikatanov, *Some observations on Babuska and Brezzi theories*, Technical Report AM222, the Pennsylvania State University (2000).