Finite element approximation of the convection-diffusion equation: subgrid-scale spaces, local instabilities and anisotropic space-time discretizations

Ramon Codina

Universitat Politècnica de Catalunya, Jordi Girona 1-3, Edifici C1, 08034 Barcelona, Spain, ramon.codina@upc.edu

Summary. The objective of this paper is to give an overview of the finite element approximation of the convection-diffusion equation that we have been developing in our group during the last years, together with some recent methods. We discuss three main aspects, namely, the global stabilization in the convective dominated regime, the treatment of the local instabilities that still remain close to layers when a stabilized formulation is used and the way to deal with transient problems.

The starting point of our formulation is the variational multiscale framework. The main idea is to split the unknown into a finite element component and a remainder that is assumed that the finite element mesh cannot resolve. A closed form expression is then proposed for this remainder, referred to as subgrid-scale. When inserted into the equation for the finite element component, a method with enhanced stability properties is obtained. In our approach, we take the space for the subgrid-scales orthogonal to the finite element space.

Once global instabilities have been overcome, there are still local oscillations near layers due to the lack of monotonicity of the method. Shock capturing techniques are often employed to deal with them. Here, our point of view is that this lack of monotonicity is inherent to the integral as duality pairing intrinsic to the variational formulation of the problem. We claim that if appropriate weighting functions are introduced when computing the integral, giving a reduced weight to layers, the numerical behavior of the method is greatly improved.

The final point we treat is the time integration in time-dependent problems. Most stabilized finite element method require a link between the time step size of classical finite difference schemes in time and the mesh size employed for the spatial discretization. We show that this can be avoided by considering the subgrid-scales as time dependent, and discretizing them in time as well. That allows us to perform a complete numerical analysis which is not restricted by any condition on the time step size, thus permitting anisotropic space-time discretizations.

1 Two-scale approximation of the convection-diffusion-reaction equation

The objective of this section is to summarize the basic stabilized finite element method we use to solve the convection-diffusion-reaction equation (CDRE) in the case in which diffusion is small, that is to say, convective effects are dominant. It is not our intention here neither to describe the details of the problem, which are well known, not to give a fair acknowledgment of the key contributions to design the final method that can be found in the literature. This is why, apart from our own work, only reference to the landmark paper [11] is made. References to other contributions can be consulted in those cited along this work.

Let us start with the problem we are interested in. For the purposes of this section it is enough to consider the stationary CRDE with homogeneous Dirichlet boundary conditions. The problem consists of finding u such that

$$\begin{aligned} \mathcal{L}u := -k\Delta u + \boldsymbol{a}\cdot\nabla u + su = f \qquad & \text{in } \Omega \\ u = 0 \qquad & \text{on } \partial\Omega \end{aligned}$$

where $k > 0$ is the diffusion coefficient, $s \geq 0$ the reaction coefficient, $\boldsymbol{a} \in \mathbb{R}^d$ is the advection coefficient and f a given datum. The problem is posed in the domain $\Omega \subset \mathbb{R}^d$ ($d = 2, 3$). Constant coefficients will be assumed throughout, for the sake of conciseness.

The variational form of the problem can be written as follows: find $u \in V = H_0^1(\Omega)$ such that

$$B(u, v) = \langle f, v\rangle \quad \forall v \in V \tag{1}$$

where:

$$B(u, v) = k(\nabla u, \nabla v) + (\boldsymbol{a}\cdot\nabla u, v) + s(u, v)$$

As usual, $(\cdot,\cdot)$ denotes the L^2 inner product and $\langle\cdot,\cdot\rangle$ the integral of the product of two functions, including the duality pairing.

The conforming Galerkin finite element approximation of the problem is standard. If $V_h \subset V$ is a finite element space to approximate V, it consists of finding $u_h \in V_h$ such that

$$B(u_h, v_h) = \langle f, v_h\rangle \quad \forall v_h \in V_h$$

Again for simplicity, we will consider that the finite element partition associated to V_h is uniform, h being the size of the element domains.

It is well known that this formulation lacks stability when k is small. To justify the method we propose, it is interesting to start trying to elucidate which is the stability it has with some more detail than what is usual. If we take $v_h = u_h$ it is readily seen that

$$B(u_h, u_h) = k\|\nabla u_h\|^2 + s\|u_h\|^2 \tag{2}$$

The question is, what control, if any, can be obtained over the convective term? That is to say, is it possible to have a bound for $\|\boldsymbol{a}\cdot\nabla u_h\|$? To answer this fundamental question, we may obtain an improved stability estimate for the Galerkin method in the form of an inf-sup condition. If we take the test function as $v_{h,0} = \tau P_h(\boldsymbol{a}\cdot\nabla u_h)$, with the parameter τ to be defined and P_h being the L^2 projection onto V_h, we obtain:

$$\begin{aligned} B(u_h, v_{h,0}) \gtrsim\ & \tau\|P_h(\boldsymbol{a}\cdot\nabla u_h)\|^2 \\ & - k\|\nabla u_h\|\frac{C_{\text{inv}}}{h}\tau\|P_h(\boldsymbol{a}\cdot\nabla u_h)\| \\ & - s\|u_h\|\tau\|P_h(\boldsymbol{a}\cdot\nabla u_h)\| \end{aligned}$$

where $\gtrsim$ stands for $\geq$ up to positive constants and C_{inv} is the constant in standard inverse inequalities . If the parameter τ is chosen such that $\tau \leq \min\left\{\frac{h^2}{C_{\text{inv}}^2 k}, \frac{1}{s}\right\}$ then

$$B(u_h, v_{h,0}) \gtrsim \tau\|P_h(\boldsymbol{a}\cdot\nabla u_h)\|^2 - k\|\nabla u_h\|^2 - s\|u_h\|^2$$

The last two terms can be controled, according to (2). It is then easily seen that $B(u_h, v_h) \gtrsim k\|\nabla u_h\|^2 + s\|u_h\|^2 + \tau\|P_h(\boldsymbol{a}\cdot\nabla u_h)\|^2$, with $v_h = u_h + \beta v_{h,0}$ (β sufficiently small), and that $k\|\nabla v_{h,0}\|^2 + s\|v_{h,0}\|^2 + \tau\|P_h(\boldsymbol{a}\cdot\nabla v_{h,0})\|^2 \lesssim \tau\|P_h(\boldsymbol{a}\cdot\nabla u_h)\|^2$, from where an inf-sup condition follows. Therefore, we may conclude that *only control over* $\tau\|P_h^{\perp}(\boldsymbol{a}\cdot\nabla u_h)\|^2$ *is missing*, with $P_h^{\perp} = I - P_h$, the projection orthogonal to the finite element space. This control is, *at least*, what any stabilized method must provide.

Let us describe now the formulation we propose. It is based on the splitting of the unknown u in a component u_h which can be resolved by the finite element space, and a remainder, that will be called subgrid scale (SGS). An approximation for the SGS is required to define a particular numerical formulation. The framework we use in based on [11]. Let $V = V_h \oplus \tilde{V}$, where $\tilde{V}$ is the space for the SGS. Then, problem (1) unfolds into two variational equations: we have to seek $u_h \in V_h$ and $\tilde{u} \in \tilde{V}$ such that

$$\begin{aligned} B(u_h, v_h) + B(\tilde{u}, v_h) &= \langle f, v_h\rangle \quad \forall v_h \in V_h \\ B(u_h, \tilde{v}) + B(\tilde{u}, \tilde{v}) &= \langle f, \tilde{v}\rangle \quad \forall \tilde{v} \in \tilde{V} \end{aligned}$$

Suppose for a moment that $\tilde{V}$ is made of smooth functions (which are anyhow dense in the complement of V_h). Then we may write

$$B(u_h, v_h) + \langle \tilde{u}, \mathcal{L}^* v_h\rangle = \langle f, v_h\rangle \quad \forall v_h \in V_h \tag{3}$$

$$\langle \mathcal{L}u_h, \tilde{v}\rangle + \langle \mathcal{L}\tilde{u}, \tilde{v}\rangle = \langle f, \tilde{v}\rangle \quad \forall \tilde{v} \in \tilde{V} \tag{4}$$

where second derivatives applied to finite element functions have to be understood in the sense of distributions. The problem now can be stated as: *how do*

we model $\tilde{u}$?. At this point is where approximations are required and different methods may be devised according to the approximation chosen.

The first approximation we shall use is that

$$\langle \mathcal{L}v_h, \tilde{v}\rangle \approx \sum_K (\mathcal{L}v_h, \tilde{v})_K \equiv (\mathcal{L}v_h, \tilde{v})_h \tag{5}$$

This essentially means that jumps of derivatives of finite elements functions across edges of the mesh are neglected. We shall stick to this assumption, although it can be relaxed, as explained in [9].

The second approximation, which is definitely the most crucial, is

$$\langle \mathcal{L}\tilde{u}, \tilde{v}\rangle \approx \tau^{-1}(\tilde{u}, \tilde{v}) \quad \text{where} \quad \tau^{-1} = c_1\frac{k}{h^2} + c_2\frac{|\boldsymbol{a}|}{h} + c_3 s \tag{6}$$

where c_1, c_2 and c_3 are numerical parameters. There are many ways to arrive at this expression, which we shall not describe. For an overview, see [5].

Equation (6) can be understood as a *lumping* of the equation for the SGS. This lumping is needed *to make this equation directly solvable*, without the need to introduce additional degrees of freedom into the problem. Both (5) and (6) can be justified from an approximate Fourier analysis requiring $\tau^{-1} \approx \|\mathcal{L}\|$ [7]. Having introduced them, the final problem to be solved is

$$B(u_h, v_h) + (\tilde{u}, \mathcal{L}^* v_h)_h = \langle f, v_h\rangle \quad \forall v_h \in V_h \tag{7}$$

$$(\mathcal{L}u_h, \tilde{v})_h + \tau^{-1}(\tilde{u}, \tilde{v}) = \langle f, \tilde{v}\rangle \quad \forall \tilde{v} \in \tilde{V} \tag{8}$$

which has to be compared with (3)-(4).

At this point we may already check which is the stability of the two scales introduced, namely, u_h and $\tilde{u}$. Using standard inverse inequalities, we have that

$$\begin{aligned}
&B(u_h, u_h) + (\tilde{u}, \mathcal{L}^* u_h)_h + (\mathcal{L}u_h, \tilde{u})_h + \tau^{-1}(\tilde{u}, \tilde{u}) \\
&\quad = k\|\nabla u_h\|^2 + s\|u_h\|^2 + 2(\tilde{u}, -k\Delta u_h + s u_h)_h + \tau^{-1}\|\tilde{u}\|^2 \\
&\quad \gtrsim k\|\nabla u_h\|^2 + s\|u_h\|^2 - 2\left(\tau\frac{k}{h^2}C_{\text{inv}}^2\right)k\|\nabla u_h\|^2 \\
&\qquad - 2(\tau s)s\|u_h\|^2 - \frac{\tau^{-1}}{2}\|\tilde{u}\|^2 + \tau^{-1}\|\tilde{u}\|^2 \\
&\quad \gtrsim k\|\nabla u_h\|^2 + s\|u_h\|^2 + \tau^{-1}\|\tilde{u}\|^2
\end{aligned}$$

where the last step holds for an adequate choice of the constants in (6). We observe that we have the same control on the finite element component as for the Galerkin method *plus additional L^2 control on the SGS*.

The SGS is so far undefined. To choose the subspace $\tilde{V}$ we consider non-conforming approximations, and thus $\tilde{V}$ might not be a subspace of $H_0^1(\Omega)$ (see [9]). If $\tilde{P}$ is the L^2 projection to $\tilde{V}$, we have from (8) that

$$\tilde{u} = \tau\tilde{P}(f - \mathcal{L}u_h)$$

There are two obvious options:

- Choice I:

$$\tilde{V} \subset \mathcal{L}V_h + \text{span}\{f\} \iff \tilde{u} = \tau(f - \mathcal{L}u_h)$$

  In this case, $\tilde{P}$ is the identity when applied to the finite element residual $f - \mathcal{L}u_h$. This option is the most common in the literature. It yields to stable formulations, as we shall see. From the conceptual point of view, the danger it has is that the assumption $V_h \cap \tilde{V} = \{0\}$, crucial to derive the method, may not hold.
- Choice II:

$$\tilde{V} = V_h^\perp \iff \tilde{u} = \tau P_h^\perp(f - \mathcal{L}u_h)$$

  This option was proposed in [6, 7]. In fact, it can be shown that if the SGS are further approximated as

$$\tilde{u} \approx -\tau P_h^\perp(\boldsymbol{a} \cdot \nabla u_h) \tag{9}$$

  the method keeps the order of accuracy. Some care is needed though in the treatment of boundary effects.

Once the two choices have been described, let us write down the final finite element problem to be solved and obtain a simple stability estimate. For choice I the final problem is

$$B(u_h, v_h) + \tau(\mathcal{L}u_h, -\mathcal{L}^* v_h)_h = \langle f, v_h \rangle + \tau(f, -\mathcal{L}^* v_h)_h$$

It is immediately checked that

$$B(u_h, u_h) + \tau(\mathcal{L}u_h, -\mathcal{L}^* u_h)_h \gtrsim k\|\nabla u_h\|^2 + s\|u_h\|^2 + \tau\|\boldsymbol{a} \cdot \nabla u_h\|^2$$

Therefore, this method provides control *over the whole convective term.*

For choice II the finite element problem is

$$B(u_h, v_h) + \tau(P_h^\perp(\boldsymbol{a} \cdot \nabla u_h), P_h^\perp(\boldsymbol{a} \cdot \nabla v_h))_h = \langle f, v_h \rangle$$

and now we have that

$$\begin{aligned} B(u_h, u_h) &+ \tau(P_h^\perp(\boldsymbol{a} \cdot \nabla u_h), P_h^\perp(\boldsymbol{a} \cdot \nabla u_h))_h \\ &\gtrsim k\|\nabla u_h\|^2 + s\|u_h\|^2 + \tau\|P_h^\perp(\boldsymbol{a} \cdot \nabla u_h)\|^2 \end{aligned}$$

Thus, this simple stability estimate shows that the method provides control *only in the component of the convective term orthogonal to the finite element space.* However, *the term $\tau\|P_h^\perp(\boldsymbol{a} \cdot \nabla u_h)\|^2$ is precisely what the Galerkin method lacks.* It is not difficult to foresee that one can in fact obtain optimal stability with choice II.

The results of the numerical analysis of the formulations arising both from choice I and from choice II is summarized next. Let:

$$|||v|||^2 := k\|\nabla v\|^2 + s\|v\|^2 + \tau\|\boldsymbol{a}\cdot\nabla v\|^2$$
$$E(h)^2 := \left(\frac{k}{h^2} + \frac{|\boldsymbol{a}|}{h} + s\right) h^{2(p+1)}|u|_{p+1}^2 \approx \tau^{-1}h^{2(p+1)}|u|_{p+1}^2$$

where $|u|_{p+1}$ is the H^{p+1} seminorm of the exact solution u. If $B_{\rm stab}$ is the bilinear form of *any of the two* stabilized methods introduced, it holds

$$\inf_{u_h\in V_h}\sup_{v_h\in V_h}\frac{B_{\rm stab}(u_h,v_h)}{|||u_h|||\,|||v_h|||} \geq C > 0 \quad \text{Stability}$$
$$|||u-u_h||| \lesssim E(h) \quad \text{Optimal convergence}$$

From these results, there are some remarks to be made:

- The stability and convergence estimates presented are optimal.
- These estimates remain meaningful for all values of the physical parameters, which is the main goal of stabilized finite element methods.
- There is no need to refer to "$h^{p+1/2}$" estimates.

2 Avoiding local instabilities

The methods proposed in the previous section yield stability and convergence *in global* norms. However, local oscillations may still remain in regions where the solution exhibits sharp layers. Even though these oscillations might be considered acceptable in linear problems, in nonlinear situations they may lead to a global failure of iterative schemes. Therefore, eliminating them in linear problems is a required step to extend the formulation to nonlinear equations. Methods aiming to avoid these local oscillations are often termed "shock capturing" or "discontinuity capturing" (DC) techniques.

To start, let us describe the guidelines to design DC methods as presented in [4] and references therein. Suppose that $s = 0$ and let

$$\boldsymbol{a}_\parallel = \frac{\boldsymbol{a}\cdot\nabla u_h}{|\nabla u_h|} \quad \text{if} \quad |\nabla u_h| \neq 0, \quad \boldsymbol{a}_\parallel = \boldsymbol{0} \quad \text{otherwise}$$

The following observations are crucial:

- For regular P_1 elements, the discrete maximum principle (DMP) holds if an artificial diffusion $k_{\rm art}$ is added, $k_{\rm art}$ being of the form

$$k_{\rm art} = \frac{1}{2}\alpha h|\boldsymbol{a}_\parallel|, \quad \alpha \geq C - \frac{1}{{\rm Pe}_\parallel}, \quad {\rm Pe}_\parallel = \frac{|\boldsymbol{a}_\parallel|h}{2k} \tag{10}$$

  where C is a constant that depends on the shape of the elements.
- If the DMP holds, L^∞ stability can be proved.
- If a numerical scheme is *linear* then it is at most first order accurate in L^∞ (a reformulation of Godunov's theorem).

In view of these facts, DC methods may be designed trying to satisfy the DMP, at least in some simple situations, and need to be *nonlinear*.

The first family of DC methods proposed is that in which an artificial diffusion depending on the finite element residual is added to the basic stabilized formulation. The essential idea of these *residual based DC methods* is to design the artificial diffusion in a way similar to (10) but computing $k_{\rm art}$ with

$$\frac{|R(u_h)|}{|\nabla u_h|} \quad \text{instead of} \quad |\boldsymbol{a}_{\|}| = \frac{|\boldsymbol{a}\cdot\nabla u_h|}{|\nabla u_h|}$$

where $R(u_h) = f - \mathcal{L}u_h = f - (-k\Delta u_h + \boldsymbol{a}\cdot\nabla u_h + s u_h)$ ($s \geq 0$ may be considered now). The resulting method is consistent, in the sense that if it is applied to the exact solution u the residual is zero.

The semilinear form of the problem is

$$B_{\rm dc}(u_h, v_h) = B_{\rm stab}(u_h, v_h) + \sum_K (k_{\rm dc}\nabla u_h, \nabla v_h)_K \tag{11}$$

with

$$k_{\rm dc} = \frac{1}{2}\alpha h \frac{|R(u_h)|}{|\nabla u_h|} \tag{12}$$

Several DC methods of this type can be found in the literature (see references in [4]).

A refinement of this approach was proposed in [4]. The idea is that the diffusion introduced by the basic stabilization method can be shown to satisfy the requirements posed by the DMP (in some model cases), but it is only introduced along the streamlines. Therefore, $k_{\rm art}$ needs to be added only in the crosswind direction. This is accomplished by adding a diffusive term with the diffusion tensor

$$\boldsymbol{k}_{\rm dc} = \frac{1}{2}\alpha h \frac{|R(u_h)|}{|\nabla u_h|}\left(\boldsymbol{I} - \frac{1}{|\boldsymbol{a}|^2}\boldsymbol{a}\otimes\boldsymbol{a}\right)$$

to the basic stabilized finite element method, $\boldsymbol{I}$ being the second order identity tensor.

Following the guidelines to design DC methods discussed above, a different possibility to make the method consistent while introducing additional diffusion is to make it proportional to the projection of the gradient orthogonal to the finite element space. Thus, if $k_{\rm art}$ is the diffusion to be added, in order to make it active only in regions of sharp gradients *which cannot be resolved by the finite element mesh*, it can be multiplied by

$$\frac{|P_h^{\perp}(\nabla u_h)|}{|\nabla u_h|} \tag{13}$$

The semilinear form of the resulting problem is again (11), but now with $k_{\rm dc}$ given by

$$k_{\rm dc} = \frac{1}{2}\alpha(|\boldsymbol{a}|h + sh^2)\frac{|P_h^{\perp}(\nabla u_h)|}{|\nabla u_h|}$$

instead of (12). Note that, apart from the factor (13), the artificial diffusion in this method is taken as $\frac{1}{2}\alpha(|\boldsymbol{a}|h + sh^2)$, independent of the finite element solution. As in the residual based DC methods, this diffusion can be introduced only in the crosswind direction. Even though we have not published before this proposal, we have used it routinely in applications requiring a resolution of sharp gradients without local oscillations.

To conclude this section, let us describe another approach in which we have been working even if it has not been published before. The idea is as follows. Consider the problem $\mathcal{L}u = f$. The discrete problem can be formally written as a "projection" of this equation onto the finite element space, using the integral as "inner product". If P_h is this projection, it is well known that it is non-monotone, in the sense that if φ is discontinuous, $P_h(\varphi)$ is oscillating. Therefore, if local oscillations have to be avoided, a natural option seems to be to modify the projection. If, moreover, these oscillations appear in regions of sharp gradients, it seems reasonable *to introduce a weighting function ρ in the integral to lighten the weight of sharp layers*. This intuitive idea has many possible realizations. A possibility that we have successfully checked consists of replacing

$$\langle \mathcal{L}u_h, v_h\rangle = \langle f, v_h\rangle \quad \text{by} \quad \langle \mathcal{L}u_h, \rho v_h\rangle = \langle f, \rho v_h\rangle$$

where $\rho \to \rho_0 \ll 1$ as $|\nabla u_h| \to \infty$. An example of weight that we have tested is $\rho = \rho_0 + (1-\rho_0)\exp\left(-|\nabla u_h|/G\right)$, where G is a reference gradient. It is obvious that the method can be considered of Petrov-Galerkin type, with test function ρv_h.

This approach has the following properties:

- The diffusion term is non-symmetric.
- The partition-of-unity property is lost.
- The numerical performance that we have observed in several examples is excellent.

3 Time dependent problems

Let us move our attention now to time dependent problems. The statement of the initial and boundary problem we are interested in is:

$$\begin{aligned} \partial_t u + \mathcal{L}u &= f \quad \text{in } \Omega, t > 0 \\ u &= 0 \quad \text{on } \partial\Omega, t > 0 \\ u &= u^0 \quad \text{in } \Omega, t = 0 \end{aligned}$$

Our approach consists in extending the scale splitting introduced in Section 1 to this problem. The time dependent counterpart of (3)-(4) is

$$(\partial_t u_h + \partial_t \tilde{u}, v_h) + B(u_h, v_h) + \langle \tilde{u}, \mathcal{L}^* v_h \rangle = \langle f, v_h \rangle \quad \forall v_h \in V_h$$
$$(\partial_t u_h + \partial_t \tilde{u}, \tilde{v}) + \langle \mathcal{L} u_h, \tilde{v} \rangle + \langle \mathcal{L}\tilde{u}, \tilde{v} \rangle = \langle f, \tilde{v} \rangle \quad \forall \tilde{v} \in \tilde{V}$$

The approximations used to arrive at (7)-(8) now lead to

$$(\partial_t u_h + \partial_t \tilde{u}, v_h) + B(u_h, v_h) + (\tilde{u}, \mathcal{L}^* v_h)_h = \langle f, v_h \rangle \quad \forall v_h \in V_h$$
$$(\partial_t u_h + \partial_t \tilde{u}, \tilde{v}) + (\mathcal{L} u_h, \tilde{v})_h + \tau^{-1}(\tilde{u}, \tilde{v}) = \langle f, \tilde{v} \rangle \quad \forall \tilde{v} \in \tilde{V}$$

If the space of SGS is chosen as orthogonal to the finite element space and approximation (9) is used, the problem to be solved becomes

$$(\partial_t u_h, v_h) + B(u_h, v_h) - (\tilde{u}, \boldsymbol{a} \cdot \nabla v_h) = \langle f, v_h \rangle \quad \forall v_h \in V_h \tag{14}$$
$$(\partial_t \tilde{u}, \tilde{v}) + (\boldsymbol{a} \cdot \nabla u_h, \tilde{v}) + \tau^{-1}(\tilde{u}, \tilde{v}) = 0 \quad \forall \tilde{v} \in V_h^{\perp} \tag{15}$$

The important point is that the SGS have been considered time dependent [7, 10]. Their evolution equation can be written as

$$\partial_t \tilde{u} + \tau^{-1} \tilde{u} = -P_h^{\perp}(\boldsymbol{a} \cdot \nabla u_h)$$

If the time derivative of the SGS is neglected, they can be inserted into (14) to obtain a closed problem for the finite element component alone. The full analysis of the resulting formulation can be found in [1, 8].

It is interesting to analyze the dissipative structure of problem (14)-(15). This was done in [12] in the more complex case of the Navier-Stokes equations. Here we will apply the results of the cited reference to the CDRE.

If, for each fixed t, we take $v_h = u_h$ and $\tilde{v} = \tilde{u}$ in (14)-(15) it is readily checked that

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u_h\|^2 + \mathcal{D}_h + \mathcal{T} = \mathcal{P}_h \tag{16}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\|\tilde{u}\|^2 + \tilde{\mathcal{D}} - \mathcal{T} = \tilde{\mathcal{P}} \tag{17}$$

with

$\mathcal{D}_h = k\|\nabla u_h\|^2 + s\|u_h\|^2$	Dissipation of the finite element scale
$\tilde{\mathcal{D}} = \tau^{-1}\|\tilde{u}\|^2$	Dissipation of the SGS
$\mathcal{T} = (\tilde{u}, \mathcal{L}^* u_h)_h = -(\tilde{u}, \boldsymbol{a} \cdot \nabla v_h)$	Energy transfer between scales

These definitions have been introduced thinking of the L^2 norm of the unknown as an energy. In this case, $\mathcal{P}_h$ and $\tilde{\mathcal{P}}$ can be considered the external power applied to the finite element scale and the SGS, respectively. From (16)-(17), with the definition of the different terms introduced above, we may draw an important conclusion. It is observed that the "energy balance" for the finite element component is the same as for the Galerkin method plus the addition of $\mathcal{T}$, which on average can be shown to be positive. In turn, this

additional dissipation is precisely injected with a negative sign in the energy balance equation for the SGS. Therefore, the global energy is conserved, but there is an energy transfer from the "large" scales to the "small" scales. This is the correct dissipative structure for dissipative systems. In particular, it is crucial for the correct modelling of turbulence.

Let us present some stability and convergence results for (14)-(15). If T is the final time of analysis, let us start with stability estimates for $T<\infty$. Suppose that $s=0$ for simplicity. Taking $v_h=u_h$, $\tilde{v}=\tilde{u}$ and integrating on $[0,t']$, $t'\leq T$ in (14)-(15) we obtain

$$\|u_h(t')\|^2+\|\tilde{u}(t')\|^2+\int_0^{t'}k\|\nabla u_h\|^2\mathrm{d}t+\int_0^{t'}\tau^{-1}\|\tilde{u}\|^2\mathrm{d}t \\ \leq\int_0^{t'}\frac{1}{k}\|f\|_{-1}^2\mathrm{d}t+\|u^0\|^2$$

from where

$$\|u_h\|\in L^\infty(0,T),\quad k^{1/2}\|\nabla u_h\|\in L^2(0,T) \\ \|\tilde{u}\|\in L^\infty(0,T),\quad \tau^{-1/2}\|\tilde{u}\|\in L^2(0,T)$$

These results indicate that the stability of (14)-(15) is the same as for the Galerkin method *plus additional stability on the SGS*.

Let us move now to the long term behavior, that is to say $T=\infty$. The results to be presented are proved in [3] for the incompressible Navier-Stokes equations. Taking $v_h=u_h$, $\tilde{v}=\tilde{u}$ and using the classical Gronwall lemma it is found that

$$\|u_h\|\in L^\infty(0,\infty),\quad \|\tilde{u}\|\in L^\infty(0,\infty)$$

and also

$$\limsup_{t\to\infty}\left(\|u_h\|+\|\tilde{u}\|\right)\leq C\frac{|\Omega|^{2/d}}{k}\|f\|_{L^\infty(0,\infty;L^2(\Omega))}$$

from where we conclude that there is a $L^2(\Omega)\oplus L^2(\Omega)$-absorbing set in $V_h\oplus\tilde{V}$, not only in V_h, as for the Galerkin method.

We may also obtain stronger stability estimates using additional regularity assumptions. Using the uniform Gronwall lemma it is found that

$$k^{1/2}\|\nabla u_h\|\in L^\infty(0,\infty),\quad \tau^{-1/2}\|\tilde{u}\|\in L^\infty(0,\infty)$$

and

$$\limsup_{t\to\infty}\left(k\|\nabla u_h\|^2+\tau^{-1}\|\tilde{u}\|^2\right)\leq C\left(a_1+\frac{a_2}{\tilde{t}}\right)\exp(a_3)$$

for certain constants a_1, a_2 and a_3 that behave as k^{-4}. For the Navier-Stokes equations, this leads to the existence of an attractor (see [3]).

Finally, let us present a simple convergence result. Suppose that the time interval is discretized using a uniform partition of size δt. Let us denote with a superscript n the approximation to u and $t^n = n\delta t$. If the backward Euler scheme is used for the time integration, there holds

$$\|u(t^N) - u_h^N\|^2 + \sum_{n=1}^{N} \delta t k \|\nabla u^n - \nabla u_h^n\|^2 + \sum_{n=1}^{N} \delta t \tau \|\boldsymbol{a} \cdot \nabla u^n - \boldsymbol{a} \cdot \nabla u_h^n\|^2 \leq I_h + I_{\delta t}$$

where $I_{\delta t}$ and I_h are optimal interpolation errors in space and time (for a proof for the Stokes problem, see [2]).

From the stability and convergence properties described, and also from the design of the formulation itself, the following properties are particularly relevant:

- No relationship between δt and h is required. *Anisotropic space-time discretizations* are possible.
- No instabilities for small δt can appear.
- τ is independent of δt. "Consistent" behavior is obtained for $t \to \infty$ (the steady state solution does not depend on δt).

These properties do not hold for the most popular stabilized finite element methods for transient problems that can be found in the literature.

4 Conclusions

In this work we have summarized the formulation we have developed during the last years to approximate flow problems and, in particular, the CDRE. The most salient aspect we would like to stress is that in the splitting of the unknown into finite element scales and SGS, the latter have their own "personality" and, in particular, their own variational equation.

We favor the choice of taking the SGS *orthogonal* to the finite element space. This leads to several advantages, in particular for transient problems. In this case, *dynamic* SGS solve inconsistencies encountered in several stabilized formulations (order of space and time discretization, link between h and δt, steady-state dependence on δt, etc.) In general, stability properties of the continuous problem are inherited by the finite element solution *plus the SGS*. Extension to anisotropic meshes is possible, the approach relying on appropriate definitions of τ and, obviously, anisotropic interpolation estimates. In applications, such estimates are usually not feasible. Methods with intrinsic stability are mandatory.

We have also discussed discontinuity capturing techniques, which are required if local instabilities need to be avoided. This is important in nonlinear

problems in which sharp layers my be developed. Classical residual based DC methods have been reviewed, and two new ideas have been proposed.

References

1. S. Badia and R. Codina. Analysis of a stabilized finite element approximation of the transient convection-diffusion equation using an ALE framework. *SIAM Journal on Numerical Analysis*, 44:2159–2197, 2006.
2. S. Badia and R. Codina. On a multiscale approach to the transient Stokes problem. Transient subscales and anisotropic space-time discretization. *Applied Mathematics and Computation*, 207:415–433, 2009.
3. S. Badia, R. Codina, and J.V. Gutiérrez-Santacreu. Long term stability estimates and existence of global attractors in a finite element approximation of the Navier-Stokes equations with numerical sub-grid scale modeling. *SIAM Journal on Numerical Analysis*, 48:1013–1037, 2010.
4. R. Codina. A discontinuity-capturing crosswind-dissipation for the finite element solution of the convection-diffusion equation. *Computer Methods in Applied Mechanics and Engineering*, 110:325–342, 1993.
5. R. Codina. Comparison of some finite element methods for solving the diffusion-convection-reaction equation. *Computer Methods in Applied Mechanics and Engineering*, 156:185–210, 1998.
6. R. Codina. Stabilization of incompressibility and convection through orthogonal sub-scales in finite element methods. *Computer Methods in Applied Mechanics and Engineering*, 190:1579–1599, 2000.
7. R. Codina. Stabilized finite element approximation of transient incompressible flows using orthogonal subscales. *Computer Methods in Applied Mechanics and Engineering*, 191:4295–4321, 2002.
8. R. Codina and J. Blasco. Analysis of a stabilized finite element approximation of the transient convection-diffusion-reaction equation using orthogonal subscales. *Computing and Visualization in Science*, 4:167–174, 2002.
9. R. Codina, J. Principe, and J. Baiges. Subscales on the element boundaries in the variational two-scale finite element method. *Computer Methods in Applied Mechanics and Engineering*, 198:838–852, 2009.
10. R. Codina, J. Principe, O. Guasch, and S. Badia. Time dependent subscales in the stabilized finite element approximation of incompressible flow problems. *Computer Methods in Applied Mechanics and Engineering*, 196:2413–2430, 2007.
11. T.J.R. Hughes. Multiscale phenomena: Green's function, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized formulations. *Computer Methods in Applied Mechanics and Engineering*, 127:387–401, 1995.
12. J. Principe, R. Codina, and F. Henke. The dissipative structure of variational multiscale methods for incompressible flows. *Computer Methods in Applied Mechanics and Engineering*, 199:791–801, 2010.