

Recent results in the treatment of subgrid scales

Franco Brezzi

Résumé. Plusieurs tentatives ont récemment été faites visant à récupérer des informations des échelles sous-grille et à les transférer aux échelles de calcul. Nombre de techniques de stabilisation peuvent aussi être considérées comme partie de ces efforts. Nous discutons ici un cadre dans lequel une partie de ces tentatives peuvent être formulées et analysées.

Abstract. *In recent times, several attempts have been made to recover some information from the subgrid scales and transfer them to the computational scales. Many stabilising techniques can also be considered as part of this effort. We discuss here a framework in which some of these attempts can be set and analysed.*

Key words. *Residual-Free Bubbles, Stabilisation, Subgrid scales, Advection-diffusion, Composite materials, Upscaling, Two-levels.*

AMS subject classification. 65NXX, 65N30, 65F10, 65F30, 65N55

In the numerical simulation of a certain number of problems, there are physical effects that take place on a scale which is much smaller than the smallest one representable on the computational grid, but have a strong impact on the larger scales, and, therefore, cannot be neglected without jeopardizing the overall quality of the final solution.

In other cases, the discretised scheme lacks the necessary stability properties because it does not treat in a proper way the smallest scales allowed by the computational grid. As a consequence, some "smallest scale mode" appears as abnormally amplified in the final numerical results. Most types of numerical instabilities are produced in this way, as the checkerboard pressure mode for nearly incompressible materials, or the fine-grid spurious oscillations in convection-dominated flows. See for instance [21] and the references therein for a classical overview of several types of these and other instabilities of this nature.

In the last decade it has become clear that several attempts to recover stability, in these cases, could be interpreted as a way of improving the simulation of the effects of the smallest scales on the larger ones. By doing that, the small scales can be *seen* by the numerical scheme and therefore be kept under control.

These two situations are quite different, in nature and scale. Nevertheless it is not unreasonable to hope that some techniques that have been developed for dealing with the latter class of phenomena might be adapted to deal with the former one. In this sense, one of the most promising technique seems to be the use of Residual-Free Bubbles (see e.g. [10], [19].) In the following sections, we are going to summarise the general idea behind it, trying to underly its potential and

its limitations. In particular, we shall first present in Section 1 the basic principles of the strategy: *divide and conquer*, *static condensation* and *approximate solution*. In Section 2 we present some examples of toy-problems, related to advection dominated flows, composite materials and viscous incompressible flows, in order to show how the general theory can be applied. We shall see in particular that several other methods, like SUPG, Petrov-Galerkin, upscaling methods, multiscale methods, and others can actually be seen as being variants of the RFB procedure. Finally, in Section 3, we discuss some questions related to the implementation of these methods.

1 The augmented problem and the bubble equation

At a very general level, the Residual-Free Bubbles approach can be summarised as follows. We start with a given problem, that for simplicity we assume to be linear, and in variational form:

$$(1) \quad \begin{cases} \text{find } u \in V \text{ such that :} \\ \mathcal{L}(u, v) = (f, v), \quad \forall v \in V. \end{cases}$$

We assume that we are given a discretised problem:

$$(2) \quad \begin{cases} \text{find } u_h \in V_h \text{ such that :} \\ \mathcal{L}(u_h, v_h) = (f, v_h), \quad \forall v_h \in V_h, \end{cases}$$

where $V_h \subset V$ is a finite element space, corresponding to a given decomposition \mathcal{T}_h of the computational domain. To fix the ideas, we shall use a two-dimensional lexicon, and in our examples we shall always assume that the computational domain is two-dimensional. The basic idea, however, is quite general, and could be applied in any dimension. We suppose, roughly speaking, that \mathcal{T}_h is the finest grid we are ready to afford in the computation, in the sense that we are not ready to solve a final system having much more unknowns than the dimension of V_h .

To the space V and to the decomposition \mathcal{T}_h we associate the space of bubbles $B(V; \mathcal{T}_h)$ defined as

$$(3) \quad B(V; \mathcal{T}_h) := \Pi_K B_V(K)$$

where, for each K in \mathcal{T}_h ,

$$(4) \quad B_V(K) := \{v \mid v \in V, \text{ supp}(v) \subset K\}.$$

Let us see some examples that will help understanding how the spaces $B_V(K)$ change for different choices of V . As a first, and most typical example, consider the case $V = H_0^1(\Omega)$. Then clearly $B_V(K) = H_0^1(K)$ (the usual bubbles). If however $V = H^1(\Omega)$, then we still have $B_V(K) = H_0^1(K)$ for the elements internal to Ω , but if K has one or more edges on $\partial\Omega$, there the elements of $B_V(K)$ are allowed to be different from zero. On the other hand, if we are dealing, for instance, with a problem written in mixed formulation, and

$$(5) \quad V = H(\text{div}; \Omega) \times L^2(\Omega),$$

then the space $B_V(K)$ will have the form $B_V(K) = \Sigma_B(K) \times U_B(K)$ with

$$(6) \quad \Sigma_B(K) := \{\tau \in H(\text{div}; K), \text{ such that } \tau \cdot \mathbf{n} = 0 \text{ on } \partial K \cap \Omega\},$$

where, as usual, \mathbf{n} represents the outward unit normal to ∂K , and

$$(7) \quad U_B(K) := L^2(K).$$

In other cases, as for instance in plate bending problems, the space V is $H_0^2(\Omega)$, and accordingly each $B_V(K) = H_0^2(K)$. And so on.

Problem (2) is now, temporarily and artificially, augmented by considering a new subspace of V :

$$(8) \quad V_A := V_h + B_h,$$

where B_h has the form

$$(9) \quad B_h := \Pi_K B_h(K),$$

and, for every K in \mathcal{T}_h , $B_h(K)$ is a subset (possibly infinite dimensional) of $B_V(K)$. The augmented problem reads now:

$$(10) \quad \begin{cases} \text{find } u_A \in V_A \text{ such that :} \\ \mathcal{L}(u_A, v_A) = (f, v_A), \quad \forall v_A \in V_A. \end{cases}$$

At this level of generality, we cannot do better than *assuming* that (10) has a unique solution for every right-hand side $f \in V'$, although, in general, its solution will not be computable in practice. In principle, problem (10) should be able to take into account all the small scales that *do not cross the boundaries of the elements K* , (in a sense that, as we have seen, depends on the nature of V .) This is a severe limitation, but corresponds to a sort of **divide and conquer** principle that might, in the end, ensure some feasibility to the whole procedure. If necessary, the space V_A could be further augmented, in a sort of iterative, self-adaptive strategy. For instance, out of a previous resolution, we might have hints that, on certain edges of \mathcal{T}_h , there are one or two *fine grid-modes* that should be taken into account. We might then add, for every such edge e , the functions having support in the union of the elements having e as an edge, and having the crucial fine-grid mode as trace on e . For the sake of simplicity, we shall not discuss this option here, apart from a short hint at the hand of the paper.

We now proceed to eliminate, at least formally, the bubble unknowns from problem (10). The technique that we are going to use is well known in the Engineering practice, under the name of **static condensation**. However, here we apply it in a more general, infinite dimensional, case.

The first problem that arises is that (8) might not correspond to a direct sum. To fix the ideas, if the original V_h is made of all piecewise cubic continuous functions, then it already contains one cubic bubble per element, that however might appear in B_h as well. Our choice (roughly speaking) is to keep, as much as possible, all the elements of the intersection $B_h \cap V_h$ in B_h , and reduce, if

necessary, V_h , introducing a proper subspace $V_h^\# \subset V_h$. This rule, however, will not be the best one in every case. Let us see some examples in order to clarify the situation.

Example 1.1 Assume as before that $V = H_0^1(\Omega)$ and V_h is made of piecewise cubic functions. Assume moreover that the space B_h is the biggest one allowed, that is $B_h = B(V; \mathcal{T}_h)$, given by (3) and (4). Then we have several choices. For instance, we might take as $V_h^\#$ the set of functions \tilde{v}_h that are polynomials of degree ≤ 3 on the interelement boundaries, and verify

$$(11) \quad L\tilde{v}_h = 0$$

separately in each element K , where L is the operator associated with the bilinear form \mathcal{L} . On the other hand, we can take as $V_h^\#$ any space of piecewise cubic continuous functions whose value can be determined in a unique way by their traces at the interelement boundaries. For instance, any serendipity cubic element would do; see, as an example, the element described in [12], page 50. Notice that these two choices produce the *same* augmented space V_A , and hence the *same* solution u_A to (10).

Example 1.2 As another example assume now that $V = H(\operatorname{div}; \Omega) \times L^2(\Omega)$ and its approximation $V_h = \Sigma_h \times U_h$ is made by lowest order Raviart-Thomas elements (see for instance [3]). Assume again that $B_h = B(V; \mathcal{T}_h)$. Then we might take $V_h^\#$ to be the space $\Sigma_h \times \{0\}$ (all the piecewise constant displacements are bubbles!). However, in most cases, for reasons that will become clear in a while, it will be more convenient to keep the piecewise constants (that are in U_h) in $V_h^\#$. This implies that we take now $B_h(K) = \Sigma_B(K) \times U_B^0(K)$, where $\Sigma_B(K)$ is given by (6) and

$$(12) \quad U_B^0(K) := L_0^2(K) \equiv \{v \in L^2(K), \text{ such that } \int_K v \, dx = 0\}.$$

In most cases, the use of bubbles in $\Sigma_B(K)$ whose normal component does not vanish on the boundary $\partial\Omega$ will be unnecessary (and in particular it would complicate the notation.) In these cases, we might use instead

$$(13) \quad \Sigma_B(K) := H_0(\operatorname{div}; K) \equiv \{\tau \in H(\operatorname{div}; K), \text{ such that } \tau \cdot \mathbf{n} = 0 \text{ on } \partial K\}.$$

With these choices we will have $V_h^\# = V_h$, and B_h (as given through (13) and (7)) will be the space of all pairs $(\tau, v) \in V$ such that τ has zero normal component at the boundary of each element, and v has zero mean value in each element. The same choice for B_h would be suitable also in the case of higher order Raviart-Thomas spaces (or, say, for BDM spaces; see always [3]), but then V_h should lose all internal degrees of freedom, apart from the piecewise constant scalars. At a more general level we notice, however, that for problems in mixed form the choice of B_h might be delicate, and might have to satisfy some particular requirements. Indeed, for some naive choice of B_h the augmented problem (10) might lose existence and/or uniqueness in V_A .

Example 1.3 A similar approach can be taken for a Stokes problem, where $V = (H_0^1(\Omega))^2 \times L_0^2(\Omega)$, and $L_0^2(\Omega)$ is, as in (12), the space of L^2 functions with zero mean value in Ω . Assume, this time, that V_h is made of piecewise quadratic velocities in $(H_0^1(\Omega))^2$, and piecewise linear discontinuous pressures. In this case one can see that the space $B(V; \mathcal{T}_h)$ as defined in (3)-(4) has the form $B(V; \mathcal{T}_h) = (\Pi_K(H_0^1(K))^2) \times L_0^2(\Omega)$. It will then be convenient to take $B_h =$

$\Pi_K\{(H_0^1(K))^2 \times L_0^2(K)\}$, and reduce the space V_h , taking as $V_h^\#$ the space of quadratic velocities and *constant* pressures. It is easy to check that with this last choice we have a direct sum in (8).

We shall come back later to the merits and drawbacks of the different choices. One way or another, we shall have in the end

$$(14) \quad V_A = V_h^\# \oplus B_h,$$

with a direct sum. Once this point has been, at least partly, clarified, we can now take an easier notation, dropping the # in (14) and *assuming* that in (8) we have a direct sum of subspaces of V . We can then write u_A and v_A , in a unique way, as:

$$(15) \quad u_A = u_h + u_B,$$

and

$$(16) \quad v_A = v_h + v_B,$$

respectively. Inserting (15) into (10), and taking $v_A = v_B$ we obtain the so-called **bubble equation**:

$$(17) \quad \begin{cases} \text{find } u_B \in B_h \text{ such that :} \\ \mathcal{L}(u_B, v_B) = -\mathcal{L}(u_h, v_B) + (f, v_B), \quad \forall v_B \in B_h. \end{cases}$$

The bubble equation (17) will play an important role in the following discussion. We take advantage of the split nature of the space B_h . For every element K in \mathcal{T}_h we define $u_{B,K}$ as the restriction of u_B to the element K . Then we have

$$(18) \quad \mathcal{L}(u_{B,K}, \varphi) = (f - Lu_h, \varphi), \quad \forall \varphi \in B_h(K),$$

where, as in (11), L indicates the operator associated with the bilinear form \mathcal{L} . In certain problems, the existence and uniqueness of the solution of each equation (18) will be immediate. However, for other types of problems (for instance for problems in mixed form) the choice of the spaces $B_h(K)$ has to be made taking into account the unique solvability of (18), that is an important feature in the whole procedure. Here, at the abstract level, we shall make the *assumption* that the problem

$$(19) \quad \begin{cases} \text{find } w_B \in B_h \text{ such that :} \\ \mathcal{L}(w_B, v_B) = (g, v_B), \quad \forall v_B \in B_h. \end{cases}$$

has a unique solution for every right-hand side g , say, in V' . We have then an operator L_B^{-1} from V' into B_h which is *defined* as the solution operator of (19), in the sense that $w_B = L_B^{-1}g$ iff w_B is the solution of (19). It is clear that (19) could also be split in independent problems, similar to (18), one in each element K . Accordingly, L_B^{-1} can be split as a sum of independent operators $L_{B,K}^{-1}$ where $L_{B,K}^{-1}g$ is the restriction $w_{B,K}$ to K of $w_B = L_B^{-1}g$.

In many cases problem (18) can easily be written in strong form as

$$(20) \quad Lu_{B,K} = f - Lu_h \quad \text{in } K,$$

with the associated boundary conditions. Let us stress the fact that, in order to write the problem in the strong form (20), it is usually essential that the space $B_h(K)$ is chosen to be equal to the full bubble space $B_V(K)$ of (4). As far as the boundary conditions are concerned, they are usually *included* in the requirement $u_{B,K} \in B_V(K)$, that is

$$(21) \quad u_{B,K} \in V \quad \text{and} \quad \text{supp}(u_{B,K}) \subset K.$$

For instance, if $B_V(K)$ is equal to $H_0^1(K)$, then the boundary conditions will simply be

$$(22) \quad u_{B,K} = 0 \quad \text{on } \partial K.$$

We can now write the solution of (17) as $u_B = L_B^{-1}(f - Lu_h)$, or, in its split form, as

$$(23) \quad u_{B,K} = L_{B,K}^{-1}(f - Lu_h), \quad \forall K \in \mathcal{T}_h.$$

We then go back to (10) and take $v_A = v_h$; inserting (15) and (23) we obtain:

$$(24) \quad \begin{cases} \text{find } u_h \in V_h \text{ such that :} \\ \mathcal{L}(u_h, v_h) + \sum_K \mathcal{L}(L_{B,K}^{-1}(f - Lu_h), v_h) = (f, v_h), \quad \forall v_h \in V_h. \end{cases}$$

Note that (24) has the same form (and the same number of unknowns) of (2). However, the additional term

$$(25) \quad \mathcal{L}(u_B, v_h) = \mathcal{L}(L_B^{-1}(f - Lu_h), v_h) = \sum_K \mathcal{L}(L_{B,K}^{-1}(f - Lu_h), v_h)$$

takes now into account the effect of *some* small scales (the ones that do not cross the interelement boundaries) onto the scales that are visible on the computational grid. As we mentioned, one can try to improve the situation by adding some fine-grid mode at the interelement boundaries, but we are not going to discuss it now. It is clear that the effect of small scales onto the coarse ones will be reproduced in a better way by taking the space B_h as big as possible, as we have done in our examples so far. It is also clear that if, as in all our examples, the space B_h is infinite dimensional, the bubble equations (18) will be practically unsolvable. Then it will be necessary to compute an **approximate solution** of the bubble equation, that however can be done **in parallel**. Moreover, as we shall see, in a certain number of applications, the structure of the additional term (25) is such that only some averaged quantities will be needed. In these cases, the approximate solution of (18) will not need to be extremely accurate. We shall come back to this point in a moment. For the time being, we point out that, when applied to a certain number of problems, the RFB procedure gives back classical stabilising techniques, as the so-called SUPG methods [11], [16], or variants of them (like in [30]). See for instance [10], [27], [4], [28], [9] for advection dominated flows, and [27], [4], [19], [15], [17] for more general problems (including Stokes, Reissner-Mindlin,

Helmholtz, and others.) Error bounds for these methods can be proved by analysing the stabilised form (24) (see for instance [29], [24], or [31] and the references therein), or, with a more recent approach, by looking directly at the form (10), as done in [7], [9], and [33].

The procedure sketched above follows the classical derivation of [10] and [19]. In particular we point out that the name *Residual-Free Bubbles* comes from the fact that u_A will satisfy the equation

$$(26) \quad \mathcal{L}(u_A, \varphi) = (f, \varphi), \quad \forall \varphi \in B_h(K),$$

which, in many cases, will imply exactly $Lu_A = f$ separately in each $K \in \mathcal{T}_h$ (that is, the residual will be zero in each K).

However, in order to see in a better way the connections with other older and newer procedures, we point out here a different way of looking at problem (10). Indeed, using the unique solvability of (19) and forgetting again the problem of the actual computation of its solution, we could tackle (10) in a different way. We might think to associate, to every function $v_h \in V_h$, a new function \tilde{v}_h solution of

$$(27) \quad \begin{cases} \text{find } \tilde{v}_h \in V_A \text{ such that } \tilde{v}_h - v_h \in B_h \text{ and} \\ \mathcal{L}(\tilde{v}_h, v_B) = 0, \quad \forall v_B \in B_h, \end{cases}$$

that in many cases could also be written as

$$(28) \quad L\tilde{v}_h = 0 \quad \text{in } K, \quad \forall K \in \mathcal{T}_h.$$

Taking all possible v_h 's in V_h , we construct a new space \tilde{V}_h , which is in one-to-one correspondence with V_h (we use here the fact that in (8) we have a direct sum.) As we remarked before, when discussing the different ways of splitting the elements of V_A , we still have $V_A = \tilde{V}_h + B_h$, and we still have a direct sum. We can now write the solution u_A of (10) as $u_A = \tilde{u}_h + w_B$. Taking $v_A \in B_h$ in (10) we have a *new* bubble equation

$$(29) \quad \mathcal{L}(w_B, v_B) = -\mathcal{L}(\tilde{u}_h, v_B) + (f, v_B), \quad \forall v_B \in B_h$$

that however, this time, becomes

$$(30) \quad \mathcal{L}(w_B, v_B) = (f, v_B), \quad \forall v_B \in B_h,$$

and, with some surprise, is *independent of \tilde{u}_h* . Substituting back in (10) with $v_A = \tilde{v}_h$ we have now that \tilde{u}_h is the unique solution of the problem

$$(31) \quad \begin{cases} \text{find } \tilde{u}_h \in \tilde{V}_h \text{ such that :} \\ \mathcal{L}(\tilde{u}_h, \tilde{v}_h) = -\mathcal{L}(w_B, \tilde{v}_h) + (f, \tilde{v}_h), \quad \forall \tilde{v}_h \in \tilde{V}_h. \end{cases}$$

It has to be pointed out that the two formulations (24) and (31) are perfectly equivalent, and $u_h + u_B = \tilde{u}_h + w_B = u_A$, solution of (10). Along these lines, we might also think about another

variant. Indeed, we might *assume* that the adjoint problem of (27) is also uniquely solvable for every right-hand side $g \in V'$. Then, always for every $v_h \in V_h$, we might construct another function, \tilde{v}_h^* , solution of

$$(32) \quad \begin{cases} \text{find } \tilde{v}_h^* \in V_A \text{ such that } \tilde{v}_h^* - v_h \in B_h \text{ and} \\ \mathcal{L}(v_B, \tilde{v}_h^*) = 0, \quad \forall v_B \in B_h. \end{cases}$$

Applying this procedure to every v_h we generate a new space, \tilde{V}_h^* , that in general will be different from \tilde{V}_h , unless the bilinear form \mathcal{L} is symmetric. We have again $V_A = \tilde{V}_h^* + B_h$, always with a direct sum. Therefore in (10) we might think of using the splitting $u_A = \tilde{u}_h + w_B$ for u_A , and the splitting $v_A = \tilde{v}_h^* + v_B$ for v_A , *always without changing the final solution* u_A . An easy computation shows that \tilde{u}_h can also be seen as the unique solution of

$$(33) \quad \begin{cases} \text{find } \tilde{u}_h \in \tilde{V}_h \text{ such that :} \\ \mathcal{L}(\tilde{u}_h, \tilde{v}_h^*) = (f, \tilde{v}_h^*), \quad \forall \tilde{v}_h^* \in \tilde{V}_h^*. \end{cases}$$

On the other hand, w_B will (obviously) still be the solution of (30). We also notice that the solutions u_h and \tilde{u}_h will have the same values at the interelement boundaries. Therefore if, for some reason, we are just interested in the values of the approximate solution on the coarse grid, then u_h and \tilde{u}_h will provide the same information.

We shall see in the next section that these alternative formulations of (10) have, in various applications, strong connections with other procedures that were introduced before and after the Residual-Free Bubbles, following different and independent arguments.

2 Some examples

It is now convenient to see the above (abstract) procedures at work on some particular simple problem.

Example 2.1 Let us consider the classical toy-problem of advection-dominated linear equations. From the physical point of view, we may think to the problem of the passive transport of a scalar diffusive quantity in a fluid whose velocity is known. Let then Ω be, for instance, a convex polygon, ε a positive number (= diffusion coefficient), \mathbf{c} a bounded mapping from Ω to \mathbb{R}^2 (= velocity field) and f , say, an element of $L^2(\Omega)$ (= source term). We consider then the problem of finding u in $H_0^1(\Omega)$ such that:

$$(34) \quad -\varepsilon \Delta u + \mathbf{c} \cdot \nabla u = f \quad \text{in } \Omega.$$

We can set $Lu := -\varepsilon \Delta u + \mathbf{c} \cdot \nabla u$, and

$$(35) \quad \mathcal{L}(u, v) := \varepsilon a(u, v) + c(u, v) \quad \forall u, v \in H_0^1(\Omega),$$

where, in a natural way,

$$(36) \quad a(u, v) := \int_{\Omega} \nabla u \cdot \nabla v \, dx, \quad \text{and} \quad c(u, v) := \int_{\Omega} \mathbf{c} \cdot \nabla u \, v \, dx, \quad \forall u, v \in H_0^1(\Omega).$$

Assume now that we are given a decomposition \mathcal{T}_h of Ω into triangles, and assume moreover that \mathbf{c} and f are piecewise constant on \mathcal{T}_h . We take then V_h to be the space of piecewise linear continuous functions vanishing on $\partial\Omega$, and B_h as in (9) with $B_h(K) = B_V(K) = H_0^1(K)$ for each K . The unique solvability of the augmented problem (10), in this case, is obvious. If we apply the theory of the previous section, the bubble equation (20) becomes, in each triangle K : find $u_{B,K}$ in $H_0^1(K)$ such that:

$$(37) \quad -\varepsilon \Delta u_{B,K} + \mathbf{c} \cdot \nabla u_{B,K} = -(-\varepsilon \Delta u_h + \mathbf{c} \cdot \nabla u_h) + f \quad \text{in } K,$$

where, for a better understanding of more general cases, we kept the term $\varepsilon \Delta u_h$ that is actually zero in each K . We point out that, in this case, the unique solvability of (37) is also obvious. On the other hand, even for the present toy problem, the solution of (37) cannot be computed in practice. However, we want to point out here the **use** that has to be done of it. In particular, it is not difficult to check that, in the present case, we have $a(u_B, v_h) = 0$ for every $u_B \in B_h$ and for every $v_h \in V_h$. Hence the additional term (25) arising in (24) becomes

$$(38) \quad \mathcal{L}(u_B, v_h) = c(u_B, v_h) = \int_{\Omega} \mathbf{c} \cdot \nabla u_B \, v_h \, dx = - \int_{\Omega} u_B \mathbf{c} \cdot \nabla v_h \, dx,$$

with an obvious integration by parts. We also remark that the term $\mathbf{c} \cdot \nabla v_h$ is piecewise constant. Hence we see that **only the mean value** of u_B in each K will be used in the final system (24) for computing u_h . Moreover, still in our assumptions, we observe that the right-hand side of (37) is also constant in K , so that $u_{B,K}$, in each K , can be written as

$$(39) \quad u_{B,K} = b_K R_K,$$

where

$$(40) \quad R_K := -(-\varepsilon \Delta u_h + \mathbf{c} \cdot \nabla u_h) + f$$

is the *residual* in K (taking u_h as approximate solution,) and the bubble b_K is the solution of the scaled problem:

$$(41) \quad \begin{cases} \text{find } b_K \in H_0^1(K) \text{ such that :} \\ -\varepsilon \Delta b_K + \mathbf{c} \cdot \nabla b_K = 1 \quad \text{in } K. \end{cases}$$

A simple computation shows that, inserting (39) in (38), the additional term (25) becomes

$$(42) \quad \mathcal{L}(u_B, v_h) = c(u_B, v_h) = \sum_K \frac{\int_K b_K \, dx}{|K|} \int_K (\mathbf{c} \cdot \nabla u_h - f) \mathbf{c} \cdot \nabla v_h \, dx,$$

where b_K is still the solution of (41). This, as already pointed out in [10] (see also [27], [4],) corresponds to the use of the well known SUPG method (see [11], [16]) with the stabilising parameter chosen as

$$(43) \quad \tau_K = \frac{\int_K b_K dx}{|K|}.$$

Assume now that we proceed with the alternative formulations (31) or (33). In this case, the space \tilde{V}_h will be made of functions \tilde{v}_h that are linear along each interelement boundary, and satisfy $L\tilde{v}_h = 0$ in each K . On the other hand \tilde{V}_h^* will be made of functions \tilde{v}_h^* that are linear along each interelement boundary, and satisfy

$$(44) \quad L^*\tilde{v}_h^* := -\varepsilon\Delta\tilde{v}_h^* - \mathbf{c} \cdot \nabla\tilde{v}_h^* = 0, \quad \forall K \in \mathcal{T}_h.$$

The formulations (31) and, mostly, (33) are clearly in the family of Petrov-Galerkin methods that have been proposed, with several variants, for convection diffusion problems. See for instance [30] and the references therein. We underline once more that the three formulations (24), (31), and (33) are perfectly equivalent, and $u_h + u_B = \tilde{u}_h + w_B = u_A$, solution of (10). However, the computational strategy that they *suggest* is different. Nevertheless, as we shall see in the last section, the computational effort related to these different strategies would also come out to be comparable. The different formulations can also *suggest* different strategies for analysing the method, in order to prove stability and error estimates. In our opinion, the (more recent) strategy of looking directly at the formulation (10), as in [9], [33], is more effective, but the situation might change from one case to another. We point out that the bubble component u_B could also be used for deriving *a posteriori* error estimates (see e.g. [32]).

Example 2.2 We consider now another toy-problem, related to *composite materials*. Let Ω be our computational domain and let $\alpha(x) \geq \alpha_0 > 0$ be a function from Ω into \mathbb{R} (we took for simplicity the scalar case, but in applications α could often be a matrix). Let again f be a given forcing term (like an external load, or a source term) and consider the problem of finding u in $V = H_0^1(\Omega)$ such that:

$$(45) \quad Lu := -\operatorname{div}(\alpha\nabla u) = f \quad \text{in } \Omega.$$

Problem (45) is meant to mimic for instance a linear elasticity problem, or a pressure equation in oil recovery problems, where α (that we assume to be rapidly varying, although not necessarily periodical) represents the properties of a composite material. We can set now

$$(46) \quad \mathcal{L}(u, v) := \int_{\Omega} \alpha \nabla u \cdot \nabla v \, dx, \quad \forall u, v \in H_0^1(\Omega),$$

and write (45) in the usual variational form

$$(47) \quad \mathcal{L}(u, v) = (f, v) \quad \forall v \in V.$$

Assume now that we are given, as in the previous example, a decomposition \mathcal{T}_h of Ω into triangles. We take again V_h as the space of piecewise linear continuous functions vanishing on $\partial\Omega$,

and B_h as in (9) with $B_h(K) = B(K) = H_0^1(K)$ for each K . If we want to apply the theory of the previous section, we see again that existence and uniqueness of the solution of the augmented problem (10) are trivially proved. Moreover the bubble equation (20) becomes, in each triangle K : find $u_{B,K}$ in $H_0^1(K)$ such that:

$$(48) \quad -\operatorname{div}(\alpha \nabla u_{B,K}) = \operatorname{div}(\alpha \nabla u_h) + f \quad \text{in } K.$$

In this case as well, the unique solvability of (48) is obvious. It is also clear that, at the practical level, equation (48) is (as all the other ones so far) unsolvable. However, as in the previous example, we can check the structure of the additional term (25), which is now

$$(49) \quad \mathcal{L}(u_B, v_h) = \int_{\Omega} \alpha \nabla u_B \cdot \nabla v_h \, dx,$$

with $v_h \in V_h$. Although the “bubble stress” $-\alpha \nabla u_B$ will surely have a complicated structure in each triangle K (if we think that α is rapidly varying) it will again be true that only its **mean value** (actually, the mean value of each component) is needed in the computation of (49), since ∇v_h is constant within each element.

We could also follow, even in the present example, the alternative formulation (31), that now would coincide with (33) as the problem is symmetric. In this case, $\tilde{V}_h \equiv \tilde{V}_h^*$ will be the space of functions \tilde{v}_h that are linear on the interelement boundaries and satisfy the homogeneous equation $\operatorname{div}(\alpha \nabla \tilde{v}_h) = 0$ separately in each element K . Again, with some attention, the computational aspects of the two approaches can be made very similar. It is interesting to see that in this case the approach (31) is quite close to the strategy which has been developed for instance in the multilevel approach of [25], [26], although using a quite independent derivation.

Example 2.3 We consider now the same problem of the previous example, but, this time, writing problem (45) in mixed form. We introduce therefore the “stress field”

$$(50) \quad \sigma = -\alpha \nabla u,$$

and the spaces $\Sigma = H(\operatorname{div}; \Omega)$ and $U = L^2(\Omega)$. We also set

$$(51) \quad a(\sigma, \tau) = \int_{\Omega} \alpha^{-1} \sigma \cdot \tau \, dx, \quad \forall \sigma, \tau \in \Sigma, \quad \text{and} \quad b(\tau, v) = \int v \operatorname{div} \tau \, dx, \quad \forall \tau \in \Sigma, \forall v \in U.$$

The mixed formulation of (45) reads now

$$(52) \quad \left\{ \begin{array}{l} \text{find } (\sigma, u) \in V \equiv \Sigma \times U \text{ such that :} \\ a(\sigma, \tau) - b(\tau, u) = 0, \quad \forall \tau \in \Sigma, \\ b(\sigma, v) = (f, v), \quad \forall v \in U. \end{array} \right.$$

Assume that we have a triangulation \mathcal{T}_h as before. We can take, as in Example 1.2, $V_h = \Sigma_h \times U_h$, where Σ_h is the space of lowest order Raviart-Thomas vectors and U_h the space of piecewise constant scalars. Assume now that B_h is chosen as in Example 1.2, in particular discarding the “boundary bubbles”. We recall that this means that each $B_h(K)$ has the form $B_h(K) = \Sigma_B(K) \times$

$U_B^0(K)$ with $\Sigma_B(K)$ given by (13) and U_B^0 by (12). We introduce some additional notation. We set $\Sigma_B := \Pi_K \Sigma_B(K)$ and $U_B^0 := \Pi_K U_B^0(K)$. Then $\Sigma_A := \Sigma_h + \Sigma_B$ and $U_A := U_h + U_B^0$. Finally $V_A := \Sigma_A \times U_A$. The augmented problem (10) becomes now

$$(53) \quad \begin{cases} \text{find } (\sigma, u) \in V_A \text{ such that :} \\ a(\sigma_h + \sigma_B, \tau_A) - b(\tau_A, u_h + u_B) = 0, \quad \forall \tau_A \in \Sigma_A, \\ b(\sigma_h + \sigma_B, v_A) = (f, v_A), \quad \forall v_A \in U_A. \end{cases}$$

To prove existence and uniqueness of the solution of (53) is an exercise, using the general theory of mixed methods and the properties of Raviart-Thomas spaces (see e.g. [3].) Essentially, you have to show that you can construct a continuous lifting of the divergence operator from $L^2(\Omega)$ back into Σ_A . For this, given a $v \in L^2(\Omega)$ you first take its projection \bar{v} into the space of piecewise constants, and you take a $\bar{\tau}_h \in \Sigma_h$ such that $\text{div} \bar{\tau}_h = \bar{v}$. This is possible since Raviart-Thomas spaces satisfy the *inf-sup* condition. Then, in every K , you can find a $\tau_{B,K} \in \Sigma_B(K)$ such that $\text{div} \tau_{B,K} = v - \bar{v}$. This is possible since $v - \bar{v}$ has zero mean value in each K . Clearly $\text{div}(\bar{\tau}_h + \tau_{B,K}) = v$.

We consider now, in each K , the local bubble equation (18), that is

$$(54) \quad \begin{cases} \text{find } (\sigma_B, u_B) \in \Sigma_B(K) \times U_B^0(K) \text{ such that :} \\ a(\sigma_B, \tau_B) - b(\tau_B, u_B) = -a(\sigma_h, \tau_B) + b(\tau_B, u_h) = -a(\sigma_h, \tau_B), \quad \forall \tau_B \in \Sigma_B(K), \\ b(\sigma_B, v_B) = -b(\sigma_h, v_B) + (f, v_B) = (f, v_B), \quad \forall v_B \in U_B^0(K). \end{cases}$$

We notice that u_h disappears from the first equation of (54) since $b(\tau_B, u_h) = 0$ for all τ_B having vanishing normal component on each ∂K , and σ_h disappears from the second equation since $b(\sigma_h, v_B) = 0$ for all v_B having zero mean value in each K . To prove existence and uniqueness of the solution of (54) is also very easy. In particular, for doing this, we make again use of the fact that the divergence operator is surjective (and has a bounded lifting) from $\Sigma_B(K)$ onto $U_B^0(K)$. This would not have been true, had we chosen $U_B(K) = L^2(K)$ as scalars for the bubble spaces, instead of $U_B^0(K)$. Whence the importance of having kept the piecewise constant scalars into V_h , as suggested in Example 1.2.

With some manipulations, one can check that the strong form of the bubble equation (54) is

$$(55) \quad \begin{cases} -\text{div}(\alpha \nabla u_B) = f_0 + \text{div} \sigma_h & \text{in } K, \\ -\alpha \nabla u_B \cdot \mathbf{n} = \sigma_h \cdot \mathbf{n} & \text{on } \partial K, \end{cases}$$

where f_0 is the difference between f and its mean value on K . We see in this case that (55) does not have exactly the general form (20), because our choice of $B_h(K)$ does *not* coincide with the full bubble space $B_V(K)$. Indeed, the piecewise constant scalars have been kept in V_h (with very good reasons.) Existence and uniqueness of (55) can clearly be deduced from those of (54), but they can also be checked directly. Indeed, existence is ensured by the compatibility of the two right-hand sides, and uniqueness is ensured by the requirement that u_B has zero mean value in K . The bubble stress σ_B can then be deduced from $\sigma_B + \sigma_h = -\alpha \nabla u_B$. Notice that if f is piecewise constant then $f_0 = 0$. If moreover, say, $\alpha = 1$, then for our case (lowest order Raviart-Thomas

elements) the solution of (55) would be the (unique) function u_B having $\nabla u_B = -\sigma_h$ and zero mean value in K . In particular σ_B would be zero. This case would be particularly uninteresting, since the additional term (25) would then vanish. All the procedure would just produce a sort of a post-processing giving $\sigma_A = \sigma_h$ and $u_A = u_h + u_B$ where the scalars would have slightly better convergence properties. We discard this curiosity, and we go back to the more interesting case of a general coefficient α . In particular we note that the additional term (25) is now

$$(56) \quad \mathcal{L}((\sigma_B, u_B), (\tau_h, v_h)) = a(\sigma_B, \tau_h) - b(\tau_h, u_B) + b(\sigma_B, v_h) = a(\sigma_B, \tau_h)$$

where we made use once more of Gauss theorem in each K . We note again that in (56) only the “stress part” σ_B of the solution of the bubble equation will be needed, and, most important, only its integral against a Raviart-Thomas vector. As an alternative, one might also compute the additional term by using in each K the formula

$$(57) \quad a_K(\sigma_B, \tau_h) = - \int_K \nabla u_{B,K} \cdot \tau_h \, dx = - \int_{\partial K} u_{B,K} \tau_h \cdot \mathbf{n} \, ds,$$

with obvious meaning for the bilinear form a_K ; then only the **mean value** of u_B on each edge would be used.

We are now interested in seeing, in the present case, what happens if one follows the alternative path (27)-(31). The procedure of Section 1 suggests that to every (τ_h, v_h) in $\Sigma_h \times U_h$ we associate a pair $(\tilde{\tau}_h, \tilde{v}_h) \in \Sigma_A \times U_A$ defined as the solution of the problem

$$(58) \quad \begin{cases} \text{find } (\tilde{\tau}_h, \tilde{v}_h) \in \Sigma_A \times U_A \text{ such that } (\tilde{\tau}_h - \tau_h) \in \Sigma_B, (\tilde{v}_h - v_h) \in U_B^0 \text{ and} \\ a(\tilde{\tau}_h, \tau_B) - b(\tau_B, \tilde{v}_h) = 0, \quad \forall \tau_B \in \Sigma_B, \\ b(\tilde{\tau}_h, v_B) = 0, \quad \forall v_B \in U_B^0. \end{cases}$$

Notice that (58) corresponds to solving, in each K , the following Neumann problem (similar to (55)) for $\tilde{v}_0 := \tilde{v}_h - v_h$

$$(59) \quad \begin{cases} -\text{div}(\alpha \nabla \tilde{v}_0) = \text{div} \tau_h & \text{in } K, \\ -\alpha \nabla \tilde{v}_0 \cdot \mathbf{n} = \tau_h \cdot \mathbf{n} & \text{on } \partial K, \end{cases}$$

and then taking $\tilde{\tau}_h = -\alpha \nabla \tilde{v}_0$ and $\tilde{v}_h = \tilde{v}_0 + v_h$. Existence and uniqueness of the solution of (59) can be checked as we did for (55). We remark in particular that, starting from a pair of the type $(0, v_h)$, we would obtain $\tilde{\tau}_h = 0$ and $\tilde{v}_h = v_h$. On the other hand, starting from a pair $(\tau_h, 0)$ both $\tilde{\tau}_h$ and \tilde{v}_h will, in general, be different from zero.

Our space \tilde{V}_h should now be made of the pairs $(\tilde{\tau}_h, \tilde{v}_h)$ that can be constructed by (58). This however does not have an “elementary” structure of product space, as it can be easily seen from the above discussion. This might complicate, at least, the formalism. In order to have a simple product space, it is then more convenient to consider just all $\tilde{\tau}_h$ ’s, and generate with them the space $\tilde{\Sigma}_h$, that clearly has the same dimension of Σ_h . We also have $\Sigma_A = \tilde{\Sigma}_h + \Sigma_B$ as a direct sum.

Using this last splitting (and leaving $U_A = U_h + U_B^0$ unchanged) we have now the new bubble equation

$$(60) \quad \begin{cases} \text{find } (\chi_B, z_B) \in \Sigma_B \times U_B^0 \text{ such that :} \\ a(\chi_B, \tau_B) - b(\tau_B, z_B) = -a(\tilde{\sigma}_h, \tau_B), \quad \forall \tau_B \in \Sigma_B, \\ b(\chi_B, v_B) = (f, v_B), \quad \forall v_B \in U_B^0, \end{cases}$$

where the terms $b(u_h, \tau_B)$ and $b(\tilde{\sigma}_h, v_B)$ have been neglected being equal to zero. Using the first equation of (58) we see that there exists a $\tilde{w}_B \in U_B^0$ such that $a(\tilde{\sigma}_h, \tau_B) = b(\tau_B, \tilde{w}_B)$ for every $\tau_B \in \Sigma_B$. Hence the pair $(\chi_B, z_B + \tilde{w}_B)$ solves the problem

$$(61) \quad \begin{cases} \text{find } (\chi_B, w_B) \in \Sigma_B \times U_B^0 \text{ such that :} \\ a(\chi_B, \tau_B) - b(\tau_B, w_B) = 0, \quad \forall \tau_B \in \Sigma_B, \\ b(\chi_B, v_B) = (f, v_B), \quad \forall v_B \in U_B^0, \end{cases}$$

which is now independent of $(\tilde{\sigma}_h, u_h)$. Hence, problem (53) can be solved by computing first the solution (χ_B, w_B) of (61), and then solving

$$(62) \quad \begin{cases} \text{find } (\tilde{\sigma}_h, u_h) \in \tilde{\Sigma}_h \times U_h \text{ such that :} \\ a(\tilde{\sigma}_h, \tilde{\tau}_h) - b(\tilde{\tau}_h, u_h) = -a(\chi_B, \tilde{\tau}_h), \quad \forall \tilde{\tau}_h \in \tilde{\Sigma}_h, \\ b(\tilde{\sigma}_h, v_h) = (f, v_h), \quad \forall v_h \in U_h. \end{cases}$$

We notice that the bubble term does not disappear from (62) because we did not take the true space \tilde{V}_h (in order to deal with a simple product space.) As we shall see in the next section, the computational aspects related to different strategies are however more alike than one might think at first sight. It is interesting to see that a quite similar approach, with a different derivation, is used in the upscaling technique of [1], [2].

Example 2.4 To conclude, we consider now an example related to the Stokes problem for viscous incompressible fluids. We shall deal again with a toy-problem, namely

$$(63) \quad \begin{cases} \text{find } (u, p) \in (H_0^1(\Omega))^2 \times L_0^2(\Omega) \text{ such that :} \\ -\Delta u + \nabla p = f \quad \text{in } \Omega, \\ \text{div } u = 0 \quad \text{in } \Omega, \end{cases}$$

where f is a “forcing term”, for instance in $(L^2(\Omega))^2$. Our discussion will be very short since, by now, the general idea should already be sufficiently clear. Introducing the spaces $\mathcal{V} = (H_0^1(\Omega))^2$ and $Q = L_0^2(\Omega)$ and the usual bilinear forms

$$(64) \quad a(u, v) = \int_{\Omega} \nabla u : \nabla v \, dx, \quad \forall u, v \in \mathcal{V}, \quad \text{and} \quad b(v, q) = \int_{\Omega} q \, \text{div } u \, dx, \quad \forall v \in \mathcal{V}, \forall q \in Q,$$

we get the usual variational formulation

$$(65) \quad \begin{cases} \text{find } (u, p) \in V \equiv \mathcal{V} \times Q \text{ such that :} \\ a(u, v) - b(v, p) = (f, v), \quad \forall v \in \mathcal{V}, \\ b(u, q) = 0, \quad \forall q \in Q. \end{cases}$$

Assume that the choices for V_h and B_h are made as in Example 1.3. We recall that this means that in $V_h (= \mathcal{V}_h \times Q_h)$ we have piecewise quadratic continuous velocities and piecewise constant pressures, and $B_h (= \mathcal{V}_B \times Q_B)$ is given by $\Pi_K(H_0^1(K))^2 \times \Pi_K L_0^2(K)$.

The existence and uniqueness of the solution to (10), in the present case, is also very easy. The lifting of the divergence operator can be constructed with a procedure that mimics the one of the previous example, just using piecewise quadratic velocities (to take care of the piecewise constant part of the pressure) instead of Raviart-Thomas vectors (see always, e.g., [3]). The bubble equation will then be, in each K ,

$$(66) \quad \begin{cases} \text{find } (u_B, p_B) \in (H_0^1(K))^2 \times L_0^2(K) \text{ such that :} \\ a(u_B, v_B) - b(v_B, p_B) = -a(u_h, v_B) + b(v_B, p_B) + (f, v_B), \quad \forall v_B \in (H_0^1(K))^2, \\ b(u_B, q_B) = -b(u_h, q_B), \quad \forall q_B \in L_0^2(K). \end{cases}$$

As in the previous example, existence and uniqueness of (66) are easily proven, making profit of the choice of $L_0^2(K)$ as local pressure space for bubbles, instead of the whole $L^2(K)$. The strong form of (66) would be, in each K :

$$(67) \quad \begin{cases} \text{find } (u_{B,K}, p_{B,K}) \in (H_0^1(K))^2 \times L_0^2(K) \text{ such that :} \\ -\Delta u_{B,K} + \nabla p_{B,K} = f + \Delta u_h \text{ in } \Omega \\ \text{div } u_{B,K} = -(\text{div } u_h)_0 \text{ in } \Omega, \end{cases}$$

where $(\text{div } u_h)_0$ is the difference between $\text{div } u_h$ and its mean value in K . Again we see that we do not get the form (20), having kept the piecewise constant pressures in Q_h . The present version of (24) will then be

$$(68) \quad \begin{cases} \text{find } (u_h, p_h) \in V_h \equiv \mathcal{V}_h \times Q_h \text{ such that :} \\ a(u_h, v_h) - b(v_h, p_h) + a(u_B, v_h) - b(p_B, v_h) = (f, v_h), \quad \forall v_h \in \mathcal{V}_h, \\ b(u_h, q_h) = 0, \quad \forall q_h \in Q_h. \end{cases}$$

Writing the second equation of (68) we took advantage of the fact that $(\text{div } u_B, q_h)$ vanishes for all piecewise constant q_h . Integrating by parts we can see that the computation of the additional term (25) requires now the mean value of u_B in each K and the two first order moments of p_B , always in each K .

It might be interesting to notice that some bad initial choices of V_h are *unredeemable* with the standard RFB approach. For instance, starting with piecewise linear continuous velocities and

constant pressures, the augmented problem (10) would be ill-posed, *no matter which subspace* $B_h \subset B(V; \mathcal{T}_h)$ we choose. Indeed, taking in B_h the biggest possible choice for velocities (that is $\Pi_K(H_0^1(K))^2$) and the smallest possible choice for pressures (that is $\{0\}$) the *inf-sup* condition will still fail in the augmented spaces. Actually, the divergences of velocities in $\Pi_K(H_0^1(K))^2$ will always have zero mean value in each K , and they cannot help in controlling piecewise constants. The situation could be improved only by using macro-bubbles, having support in more than one element (see e.g. [20]), but we shall not discuss it here.

Many aspects of the above discussion extend from the toy-problems of this section to more general situations. As we have seen, the main difficulty is that, in general, (17) cannot be solved explicitly, so that the additional term (25) cannot be computed exactly. We saw however in our examples that, in order to have a reasonable approximation of (25), we will not need a very accurate solution of (17), as only some *averaged quantities* of the bubble component are needed. This is true in a wider variety of cases. Indeed, as we already observed, it is always only *the effect of the small scales on the larger ones* that needs to be simulated. Even at a fully general level one can still notice that in (25) the bubble term $L_{B,K}^{-1}(f - Lu_h)$ is tested against v_h , which belongs to the coarse space.

In the next section we are going to see some more practical aspects of the above strategy.

3 Some hints on the implementation

We shall present now a brief discussion on various problems related with the implementation of the above strategies. To start with, a point that we did not stress so far is that, in order to perform the *static condensation*, one has to be able to substitute u_B (or actually, in practice, its approximation) *as a function of the unknown* u_h in (24). We shall see now with some more detail how this can be done.

As a first step we identify, in each K , the smallest linear space that contains all possible residuals, namely

$$(69) \quad R_{K,h} := \text{span}\{f|_K, Lv_h|_K, v_h \in V_h\}.$$

In many cases, $f|_K$ can be approximated, without major loss of information, by means of elements of the space $(LV_h)|_K$, that can therefore be used in place of $R_{K,h}$. Otherwise, the dimension of $R_{K,h}$ will be equal to the dimension of $(LV_h)|_K$ plus one.

The second step is then to choose, for each K , a basis $\{g_K^i\}_{i=1,\dots,N_K}$ of the space $R_{K,h}$. Clearly, N_K denotes the dimension of such space. Then, for each $i = 1, \dots, N_K$, we seek an approximate solution of the local problem:

$$(70) \quad \begin{cases} \text{find } \psi_K^i \in B_h(K) \text{ such that :} \\ \mathcal{L}(\psi_K^i, v) = (g_K^i, v), \quad \forall v \in B_h(K). \end{cases}$$

Now, in each K , we can express every possible right-hand side of (17) as a linear combination of the functions $\{g_K^i\}_{i=1,\dots,N_K}$. In particular, assume that $\{v_h^j\}_{j=1,\dots,N}$ is the basis that

we chose for V_h . Clearly, N denotes the dimension of V_h . Then for each K and for each v_h^j we express the restriction to K of Lv_h^j as a linear combination of the g_K^i . Then the corresponding (approximate) solution $L_{B,K}^{-1}(-Lv_h^j)$ can be obtained as a linear combination of the ψ_K^i . Similarly, the contribution $L_{B,K}^{-1}(f)$ to the right-hand side can also be expressed as a combination of the ψ_K^i . The desired substitution in (24) can then be done during the assembling procedure of the final stiffness matrix.

In various applications, N_K will be rather small. Referring to the examples of the previous section, and assuming that f , in each K , could be approximated by elements of $(LV_h)|_K$ without major losses in accuracy, we have, for the examples of the previous section, $N_K = 1$ in the first example, $N_K = 3$ in Examples 2.2 and 2.3, and $N_K = 4$ in the last example. At a general level we might say that, for problems with constant coefficients *and* without zero order terms (as in Examples 2.1 and 2.4,) N_K will often be smaller (and sometimes much smaller) than the number of degrees of freedom of V_h in K . For instance, in Example 2.4 the number of degrees of freedom of V_h in K is 13 (twelve velocities and one pressure,) while, as we have seen, $N_K = 4$. Notice that the original number of degrees of freedom in V_h was 15 (twelve velocities and three pressures) and went down to 13 when we left in V_h only the piecewise constant part of the pressure.

Let us consider now the alternative formulations (31) and (33). We claim that the (approximate) resolution of the N_K problems (70) is still an essential ingredient for computing the stiffness matrix of (31). Clearly, the adjoint problems will also be needed in order to use (33).

For the sake of simplicity, we concentrate on (31). We have to compute a basis for \tilde{V}_h . An obvious choice is to start from the basis $\{v_h^j\}$ of V_h . Then we associate to every v_h^j an element \tilde{v}_h^j in \tilde{V}_h as in (27). To do that, we write \tilde{v}_h^j as $v_h^j + v_0^j$, with $v_0^j \in B_h$. Then, in each K , v_0^j will be the solution of

$$(71) \quad \mathcal{L}(v_0^j, v_B) = -\mathcal{L}(v_h^j, v_B) = -(Lv_h^j, v_B), \quad \forall v_B \in B_h(K).$$

It is clear that, in each K , every v_0^j can be expressed as a linear combination of the N_K solutions ψ_K^i of (70). If we keep the assumption that, in each K , the right-hand side f is approximated by elements of $(LV_h)|_K$, then the solution w_B of (30) can also be computed, element by element, as a linear combination of the ψ_K^i .

Finally, let us compare the effort for the actual computation of the stiffness matrices of (24) and of (31). In many applications, the computation of $\mathcal{L}(v_h^j, v_h^k)$ ($j, k = 1, \dots, N$) would be much simpler than the direct computation of $\mathcal{L}(\tilde{v}_h^j, \tilde{v}_h^k)$. However, we note that, using (27), we have

$$(72) \quad \mathcal{L}(\tilde{v}_h^j, \tilde{v}_h^k) = \mathcal{L}(\tilde{v}_h^j, v_h^k + v_0^k) = \mathcal{L}(\tilde{v}_h^j, v_h^k) = \mathcal{L}(v_h^j, v_h^k) + \mathcal{L}(v_0^j, v_h^k).$$

We notice that the first term in the right-hand side of (72) is the same that would be computed in (24), and the second term have the same nature as the right-hand sides of (70). Indeed,

$$(73) \quad \mathcal{L}(v_0^j, v_h^k) = (v_0^j, L^* v_h^k).$$

Finally, the term containing w_B (which disappears if \mathcal{L} is symmetric) can be treated as

$$(74) \quad \mathcal{L}(w_B, \tilde{v}_h^k) = \mathcal{L}(w_B, v_h^k) + (f, v_0^k) = (w_B, L^* v_h^k) + (f, v_0^k).$$

We also observe that all the terms in (73) and (74) will require the computation of the integral of the product of a bubble times an element of $R_{K,h}^*$, that we define as

$$(75) \quad R_{K,h}^* := \text{span}\{f|_K, L^*v_h|_K, v_h \in V_h\}.$$

This is the same that has to be done in the RFB approach, when computing the additional term (25). We can conclude that the computational effort required by the two strategies is comparable (after all, they are the *same* method, just written differently.) Similar considerations hold for the formulation (33). We omit them, as they are a straightforward extension of the above discussion. Other variants, as the one discussed in Example 2.3, can be treated with similar arguments.

All these matters might deserve however a deeper investigation, in each particular case, in order to assess the most convenient strategy that has to be adopted. For instance, in cases like Example 2.2, where essentially N_K equals the dimension of $V_h|_K$, the easiest way to tackle (31) seems to solve directly the problems

$$(76) \quad \begin{cases} \text{find } \tilde{\psi}_K^i \in V_{A|K} \text{ such that } (\tilde{\psi}_K^i - v_h^i) \in B_h(K), \text{ and} \\ \mathcal{L}(\tilde{\psi}_K^i, v) = 0, \quad \forall v \in B_h(K), \end{cases}$$

and then to write the \tilde{v}_h^j 's as combinations of the $\tilde{\psi}_K^i$'s. And so on.

Let us consider now the problem of solving (70) or (76). The most general and widely applicable strategy in order to obtain an approximate solution of, say, (70) consists, obviously, in the use of a Galerkin approximation, corresponding to a suitable choice of $B_K^* \subset B_h(K)$ (warning: in spite of the notation, this has no particular relationship with the adjoint problem). Then one can solve N_K problems of the type:

$$(77) \quad \begin{cases} \text{find } \psi_K^{*,i} \in B_K^* \text{ such that :} \\ \mathcal{L}(\psi_K^{*,i}, v) = (g_K^i, v), \quad \forall v \in B_K^*. \end{cases}$$

The technique has been successfully applied to advection dominated flows, also with the use of suitably refined subgrid meshes (of Shishkin type) near the outflow boundary of each K . See e.g. [18]. For an application of this technique to the Helmholtz equation see for instance [15], [17]. This, essentially, is also what has been done for computing the solutions of (76), in problems related to composite materials, in [1], [2], [25], [26], although, as we have said, the derivation there was done in a framework independent of the Residual-Free Bubbles approach.

In more special classes of problems one might also think of using some special *tricks*. For instance, for convection-dominated flows with a very small diffusion coefficient, we have seen that (70) reduces essentially to solve (41). Assuming that the convective coefficient is piecewise constant, one can solve, instead of (41), the (limit) purely convective problem (corresponding to $\varepsilon = 0$) that can be dealt with by hand. See e.g. [10], [5].

Another possible choice is to take B_K^* with a very small dimension (sometimes even one-dimensional) but choosing it in a very suitable way, which depends on the particular form of the

equation in each K . See for instance [8]. The use of Shishkin subgrid meshes for advection dominated problems, as in [18], could also be seen as being part of this strategy.

On the other hand, for singularly perturbed problems where some artificial viscosity (or similar regularisation) is usually employed, the idea of using in (77) only one (or very few) degrees of freedom with a kind of *subgrid viscosity* (or subgrid regularisation,) as in [22], [23], is surely appealing for its simplicity and rather wide range of applicability. However, as shown in [6], the choice of the actual value for the subgrid artificial viscosity appears to be crucial, and requires deeper investigations.

At a more general level the **effects on error estimates** (both *a priori* and *a posteriori*) of taking an approximate solution for the bubble equation deserve a better attention, and should be the object of a more accurate analysis.

The application of the paradigm “divide and conquer/ static condensation/ approximate solution” to some nonlinear problems is currently under investigation. The obvious choice would be to apply it to the various linearised problems in an iterative procedure, but in particular cases the structure of the nonlinearity might suggest a better strategy.

Finally, we point out that, in some cases, the procedure of augmenting the space V_h can be done in *two steps*. Indeed, together with the addition of suitable bubble spaces in each K , we might think of adding some *edge functions* where convenient. For instance, just to make an academic example, one can decide that on a particular edge e the addition of a function with a special shape could be beneficial for the overall quality of the approximation. This could be a rapidly oscillating one, with a given wavelength, or a function with a “bump” (or a “step”) at a given location, and so on. Typically, in practice, this could only be done in some iterative way, using information from the subgrid solutions at previous iterations or time steps. One way or another, *assume* that you know that a certain shape φ^* is needed at a certain edge e^* . Then you might think of *increasing first* the space V_h by adding a function having value φ^* on the edge e^* , and having support in the union of the (one or two) elements having e^* as an edge. The value of such a function inside the element(s) could profitably be defined as the solution of the homogeneous equation (11). You are equipped for that, as you are ready to solve subgrid problems related to the operator L . In general, there will not be just *one* edge and *one* function that is needed, but we might expect that they are not too many. After this first enrichment of V_h , the bubble procedure can be started. It is clear that, after addition and elimination of the bubbles, the total number of equations will be equal to the original number of boundary degrees of freedom in V_h *plus* the ones that were added during the first enrichment (these will not go away in the static condensation process). This procedure, on one hand, has to be used with some care. Indeed, we started by assuming that V_h corresponded roughly to the finest grid we were ready to afford. Clearly, the number of additional degrees of freedom has to be limited to the essential. On the other hand this procedure can provide a partial healing to the very undesirable (but, somehow, necessary) effect of *impermeability of the interelement boundaries to the small scales* that is inherent to the *divide and conquer* strategy. It is clear that much more investigation and experimentation are needed in order to design effective strategies for doing all that.

As a last point, the possible use of “nonconforming bubbles” is surely worth investigating. A nonconforming bubble would be a function having support in a single element K , whose extension by zero outside K does not belong to the space V . In some cases, the boundary values of these

nonconforming bubbles might have zero mean value (or zero moments up to some order k) at each interelement edge (see for instance [14]). In other cases, this will have to be imposed by means of some interelement multiplier or similar trick, in order to compensate for the variational crime. This approach has been recently applied with success in [13], which we refer to for additional information. It is reasonable to assume that the idea might be extended to more general situations, and it is surely quite appealing.

Acknowledgements

This research was supported in part by grants of M.U.R.S.T. cofin 99.

References

- [1] T. Arbogast, “Numerical subgrid upscaling of two-phase flow in porous media,” in “Multiphase flows and transport in porous media: State of the art”, (Z. Chen, R.E. Ewing, and Z.-C. Shi eds.), Lecture Notes in Physics, Springer, Berlin, 2000.
- [2] T. Arbogast, S.E. Minkoff, and P.T. Keenan, “An operator-based approach to upscaling the pressure equation,” in: Computational Methods in Water Resources XII, v.1, V.N. Burganos et als., eds., Computational Mechanics Publications, Southampton, U.K., 1998.
- [3] F. Brezzi, M. Fortin, “Mixed and Hybrid Finite Element Methods,” Springer Verlag, New York, Springer Series in Computational Mathematics **15**, 1991.
- [4] F. Brezzi, L.P. Franca, T.J.R. Hughes, and A. Russo, “ $b = \int g$,” *Comput. Methods Appl. Mech. Engrg.* **145**, 329-339 (1997).
- [5] F. Brezzi, L.P. Franca, and A. Russo, “Further considerations on residual-free bubbles for advective-diffusive equations,” *Comput. Methods Appl. Mech. Engrg.* **166**, 25-33 (1998).
- [6] F. Brezzi, P. Houston, L.D. Marini, and E. Süli, “Modeling subgrid viscosity for advection-diffusion problems,” to appear in *Comput. Methods Appl. Mech. Engrg.*
- [7] F. Brezzi, T.J.R. Hughes, L.D. Marini, A. Russo, and E. Süli, “A priori error analysis of a finite element method with residual-free bubbles for advection-dominated equations,” *SIAM J. Num. Anal.* **36**, 1933-1948 (1999)
- [8] F. Brezzi, D. Marini, and A. Russo, “Applications of pseudo residual-free bubbles to the stabilization of convection-diffusion problems,” *Comput. Methods Appl. Mech. Engrg.* **166**, 51-63 (1998).
- [9] F. Brezzi, D. Marini, and E. Süli, “Residual-free bubbles for advection-diffusion problems: the general error analysis,” *Numer. Math.* **85**, 31-47 (2000).
- [10] F. Brezzi, A. Russo, “Choosing bubbles for advection-diffusion problems,” *Math. Mod. and Meth. in Appl. Sci.* **4**, 571-587 (1994).
- [11] A.N. Brooks, T.J.R. Hughes, “Streamline Upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations,” *Comput. Methods Appl. Mech. Engrg.* **32**, 199-259 (1982).
- [12] Ph.G. Ciarlet, “The finite element method for elliptic problems,” North-Holland, 1978.
- [13] C. Farhat, A. Macedo, and M. Lesoinne, “A two-level domain decomposition method for the iterative solution of high frequency exterior Helmholtz problems,” *Numer. Math.* **85**, 283-308 (2000)
- [14] M. Fortin, M. Soulie, “A non-conforming piecewise quadratic finite element on triangles,” *Int. J. Num. Meth. in Eng.* **19**, 505-520 (1983).

- [15] L.P. Franca, C. Farhat, A.P. Macedo and M. Lesoinne, “Residual-Free Bubbles for the Helmholtz Equation,” *Int. J. Num. Meth. in Eng.* **40**, 4003-4009 (1997).
- [16] L.P. Franca, S.L. Frey and T.J.R. Hughes, “Stabilized finite element methods: I. Applications to advective-diffusive model,” *Comput. Methods Appl. Mech. Engrg.* **95**, 253-276 (1992).
- [17] L.P. Franca, A.P. Macedo, “A Two-Level Finite Element Method and its Application to the Helmholtz Equation,” *Int. J. Num. Meth. in Eng.* **43**, 23-32 (1998).
- [18] L.P. Franca, A. Nesliturk and M. Stynes, “On the Stability of Residual-Free Bubbles for Convection-Diffusion Problems and Their Approximation by a Two-Level Finite Element Method,” *Comput. Methods Appl. Mech. Engrg.* **166**, 35-49 (1998).
- [19] L.P. Franca, A. Russo, “Deriving upwinding, mass lumping and selective reduced integration by residual-free bubbles,” *Appl. Math. Lett.* **9**, 83-88 (1996).
- [20] L.P. Franca, A. Russo, “Approximation of the Stokes problem by residual-free macro bubbles,” *East-West J. Numer. Math.* **4**, 265-278 (1996).
- [21] D.F. Griffiths, A.R. Mitchell, “Spurious behaviour and nonlinear instability in discretised partial differential equations,” In: *The dynamics of numerics and the numerics of dynamics. Inst. Math. Appl. Conf. Ser., New Ser.* **34**, 215-242 (1992).
- [22] J.L. Guermond, “Stabilization of Galerkin approximations of transport equations by subgrid modeling,” Submitted to *Math. Mod. Num. Anal.*
- [23] J.L. Guermond, “Stabilisation par viscosité de sous-maille pour l’approximation de Galerkin des opérateurs monotones,” *C. R. Acad. Sci. Paris, Série I* **328**, 617-622 (1999).
- [24] P. Hansbo, C. Johnson, “Streamline diffusion finite element methods for fluid flow,” *von Karman Institute Lectures*, 1995.
- [25] T.Y. Hou, X.H. Wu, “A multiscale finite element method for elliptic problems in composite materials and porous media,” *J. Comput. Phys.* **134**, 169-189 (1997).
- [26] T.Y. Hou, X.H. Wu, and Z. Cai, “Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients,” *Math. of Comp.* **68**, 913-943 (1999).
- [27] T.J.R. Hughes, “Multiscale phenomena: Green’s functions, the Dirichlet to Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods,” *Comput. Methods Appl. Mech. Engrg.* **127**, 387-401 (1995).
- [28] T.J.R. Hughes, G. Feijoo, L. Mazzei, and Jean-Baptiste Quincy, “The variational multiscale method – A paradigm for computational mechanics,” *Comput. Methods Appl. Mech. Engrg.* **166**, (1998).
- [29] C. Johnson, U. Nävert, and J. Pitkäranta, “Finite element methods for linear hyperbolic problems,” *Comput. Methods Appl. Mech. Engrg.* **45**, 285-312 (1984).
- [30] A.R. Mitchell, D.F. Griffiths, “Generalised Galerkin methods for second order equations with significant first derivative terms,” In: *Proc. bienn. Conf., Dundee 1977, Lect. Notes Math* **630**, 90-104 (1978).
- [31] H.-G. Roos, M. Stynes, and L. Tobiska, “Numerical methods for singularly perturbed differential equations: convection diffusion and flow problems,” Springer-Verlag, 1996.
- [32] A. Russo, “A posteriori error estimators via bubble functions,” *Math. Models Methods Appl. Sci.* **6**, 353-360 (1997).

- [33] G. Sangalli, “Global and local error analysis for the Residual Free Bubble method applied to advection-dominated problems,” Submitted to Numer. Math.

Istituto di Analisi Numerica del CNR and Dipartimento di Matematica
Università di Pavia
Via Ferrata 1
I-27100 Pavia, Italy

mailto: brezzi@ian.pv.cnr.it