

BUBBLE STABILIZATION OF DISCONTINUOUS GALERKIN METHODS

PAOLA ANTONIETTI, FRANCO BREZZI, AND L. DONATELLA MARINI

ABSTRACT. We analyze the stabilizing effect of the introduction of suitable bubble functions in DG formulations for linear second order elliptic problems, working, for the sake of simplicity, on Laplace operator. In particular we analyze the nonsymmetric formulation of Baumann-Oden on rather general decompositions, and we show that the piecewise linear discontinuous approximation, without jump stabilization, can be used if suitable bubbles are added to the local spaces (

1. INTRODUCTION

Most commonly used DG methods need the addition of suitable stabilizing terms in order to provide good convergence properties. The typical stabilizing procedure consists in the introduction of *penalty terms* that penalize the jumps across neighboring elements. Sometimes, in hyperbolic or in convection dominated problems, one can also use *upwind techniques*, consisting in replacing the *average* $((u^+ + u^-)/2)$ on an internal edge with the upwind value (that is, u^+ or u^- , according with the direction of the “wind”). This however, in most cases, can be seen again as a jump stabilization ([18], [15], [11]).

Another possible way of stabilizing DG methods consists in the addition of suitable terms (this time, internal to each element) of the so-called *Hughes-Franca* type: in general, the integral of the original equation (or one of the original equations), written in strong form *inside each element* in terms of the finite element unknowns (= trial functions), multiplied by a similar expression acting on the test functions. The most famous stabilization of this type, for standard Galerkin methods, is surely the SUPG stabilization of convection dominated equations [13]. A typical problem, in these cases, is the choice of the proper *stabilization coefficient* to be put in front of the stabilizing term.

In a recent paper (see [6]) we pointed out that, in DG methods, the jumps are themselves to be regarded as “equations”, so that jump stabilizations (and hence *upwind*) could be regarded as Hughes-Franca stabilizations as well. And, indeed, the optimal choice of the coefficient in a jump-stabilization term is still a subject that might need a further investigation.

In standard Galerkin methods (for instance in Stokes problem or in advection-diffusion problems) one of the possible ways of stabilizing an unstable formulation is to add one or more *bubble function* per element. We recall that a bubble function is, by definition, a function whose support is contained in a single element. The bubble stabilization, in its turn, can also be seen as a Hughes-Franca stabilization after eliminating the bubbles by static condensation. This has the effect of shifting the problem: from the choice the optimal coefficient to the choice of the optimal

shape of the bubble (see e.g. [5], [2]). This last problem can however be solved, in some cases, with the use of Residual Free Bubbles (see [12], [16]), or Pseudo Residual Free Bubbles (see [9], [10]).

When using a discontinuous method the addition of bubble functions does not mean much, as all the basis functions already have support in a single element (hence, in a sense, they are *all*, already, bubbles). We could therefore consider that for DG methods adding bubbles is just the same as augmenting the finite element space, in an arbitrary way. For instance, in two dimensions, shifting from linear discontinuous elements to quadratic discontinuous elements could be seen as adding three bubbles per element (corresponding to x^2 , y^2 , and xy). The same is obviously true for any other increase of the local polynomial degree.

The problem whether the addition of bubbles could provide some additional stability for DG methods has therefore a rather academic nature. However, it is intellectually tackling to check whether and when a suitable (and possibly minimal) increase in the finite element space can turn an unstable formulation into a stable one. And, possibly, any discovery in this direction can provide some additional understanding of the underlying nature of DG methods.

Here we consider as a model (toy) problem the Poisson problem in a polygonal domain, and we address our attention to the so-called Baumann-Oden DG formulation (see [3], [4], [19], [20], and many other papers). In particular we consider the (unstable) choice of piecewise linear discontinuous elements. We already know that, for triangular elements, the use of piecewise quadratic elements (always for the Baumann-Oden formulation) *is* indeed stable ([20]). Hence we know already that, in some sense, adding three bubbles per element can stabilize the problem for triangular elements. In a previous paper, [8], we proved that for triangular elements the addition of a single quadratic bubble can stabilize the Baumann-Oden formulation. A similar result for the *Interior Penalty* formulation has recently been obtained by Burman-Stamm [14].

Here we will complete the result on the Baumann-Oden formulation, showing that the addition of $n - 2$ suitable bubbles per element can stabilize the Baumann-Oden formulation for a decomposition in polygonal elements with n edges.

The practical impact of our investigation is surely questionable, although the possibility of avoiding the jump stabilization for linear elements is surely appealing, as it leads to a more “natural” choice of the interelement fluxes. Moreover we believe that our analysis provide a better understanding of some basic aspects and mechanisms related to DG methods, that might be of some help in designing new future methods. And, as such, it might interest several curious scientists.

An outline of the paper is as follows. In the next section we recall some notation on DG methods, and the Baumann-Oden formulation for Poisson problem. Then we introduce, in an abstract form, the bubble stabilization: under suitable assumptions on the local bubble spaces we prove the stability of the augmented formulation, and optimal error estimates. For triangular decompositions, the properties showed and analyzed in [8] easily imply our abstract assumptions. Hence, in the subsequent section we show how to construct local spaces that work for (non-degenerate) quadrilaterals. Finally, in the last section, we present some numerical experiments.

2. THE MODEL PROBLEM AND THE BAUMANN-ODEN METHOD

Let Ω be a convex polygonal domain, with boundary $\partial\Omega$. For every f , say, in $L^2(\Omega)$ we consider the model problem:

$$(1) \quad -\Delta u = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega.$$

It is well known that problem (1) has a unique solution, that moreover belongs to $H^2(\Omega) \cap H_0^1(\Omega)$.

Let $\{\mathcal{K}_h\}_h$ be a sequence of *compatible* decompositions of Ω into polygons K . Here, “compatible” means that the intersection of the closure of two different polygons is either empty, or a common edge, or a common vertex. For every polygon K we will denote by n_K the number of its edges and by h_K its diameter. Moreover, for every edge e we will denote by $|e|$ its length. We shall also assume that

- There exists an integer $n_* \geq 3$ such that for every h and for every $K \in \mathcal{K}_h$ we have

$$(2) \quad n_K \leq n_*.$$

- There exists a constant $\rho_1 > 0$ such that for every h , for every $K \in \mathcal{K}_h$, and for every edge e of K , we have

$$(3) \quad |e| \geq \rho_1 h_K.$$

- There exists a constant $\rho_2 > 0$ such that for every h , for every $K \in \mathcal{K}_h$, and for every edge e of K , the biggest disk that is tangent to e and is contained in K has radius

$$(4) \quad \rho \geq \rho_2 h_K.$$

We consider first the (infinite dimensional) space $V(\mathcal{K}_h)$ defined as

$$(5) \quad V(\mathcal{K}_h) = \{v \in L^2(\Omega) \text{ such that } v|_K \in H^2(K) \forall K \in \mathcal{K}_h\}.$$

Elements $v \in V(\mathcal{K}_h)$ will, in general, be discontinuous when passing from one element to a neighboring one. As usual in DG methods we have therefore to introduce boundary operators as *averages* and *jumps*. As we shall deal also with vector-valued functions which are smooth in each polygon but discontinuous from one polygon to another, we shall introduce these boundary operators for scalar and for vector-valued functions. Following [1] we set (see Figure 1):

FIGURE 1. Two neighboring polygons and their normals.

$$\begin{aligned} \{v\} &= \frac{v^+ + v^-}{2}; & \llbracket v \rrbracket &= v^+ \mathbf{n}^+ + v^- \mathbf{n}^- & \text{for all internal edges} \\ \{\boldsymbol{\tau}\} &= \frac{\boldsymbol{\tau}^+ + \boldsymbol{\tau}^-}{2}; & \llbracket \boldsymbol{\tau} \rrbracket &= \boldsymbol{\tau}^+ \cdot \mathbf{n}^+ + \boldsymbol{\tau}^- \cdot \mathbf{n}^- & \text{for all internal edges} \end{aligned}$$

On the boundary edges we define $\llbracket v \rrbracket = v \mathbf{n}$; $\{\boldsymbol{\tau}\} = \boldsymbol{\tau}$.

We introduce now some further notation. For functions in $V(\mathcal{K}_h)$ we first introduce the elementwise gradient ∇_h , and then for u and v in $V(\mathcal{K}_h)$ we set

$$(\nabla_h u, \nabla_h v) := \sum_{K \in \mathcal{K}_h} \int_K \nabla u \cdot \nabla v dx, \quad \langle \{\nabla_h u\}, \llbracket v \rrbracket \rangle := \sum_{e \in \mathcal{E}_h} \int_e \{\nabla_h u\} \cdot \llbracket v \rrbracket ds,$$

where, here and in all the sequel, \mathcal{E}_h denotes the set of all the edges of the decomposition \mathcal{K}_h . Setting, for u and v in $V(\mathcal{K}_h)$

$$(6) \quad a(u, v) := (\nabla_h u, \nabla_h v) - \langle \{\nabla_h u\}, [v] \rangle + \langle \{\nabla_h v\}, [u] \rangle,$$

the Baumann-Oden “continuous” formulation of (1) is now

$$(7) \quad \begin{cases} \text{Find } u \in V(\mathcal{K}_h) \text{ such that:} \\ a(u, v) = (f, v) \quad \forall v \in V(\mathcal{K}_h). \end{cases}$$

In $V(\mathcal{K}_h)$ we define the jump seminorm

$$(8) \quad \|v\|_j^2 = \sum_{e \in \mathcal{E}_h} \frac{1}{|e|} \int_e |[v]|^2 ds,$$

and the norm

$$(9) \quad \|v\|_{V(\mathcal{K}_h)}^2 := \sum_{K \in \mathcal{K}_h} (\|\nabla_h v\|_{0,K}^2 + h_K^2 |v|_{2,K}^2) + \|v\|_j^2.$$

We recall now the following useful result, which follows easily from a well known result of Agmon (see, e.g., [1]):

$$(10) \quad \forall K, \forall e \in \partial K, \forall v \in H^1(K) : \int_e v^2 ds \leq C_a (h_K^{-1} \|v\|_{0,K}^2 + h_K |v|_{1,K}^2),$$

with C_a only depending on the constant ρ_2 in (4). Hence we have

$$(11) \quad \begin{aligned} |\langle \{\tau\}, [v] \rangle| &= \sum_{e \in \mathcal{E}_h} \int_e \{\tau\} \cdot [v] ds \\ &\leq C \left[\sum_K (\|\tau\|_{0,K}^2 + h_K^2 |\tau|_{1,K}^2) \right]^{1/2} \left[\sum_{e \in \mathcal{E}_h} h_e^{-1} \int_e |[v]|^2 ds \right]^{1/2}, \end{aligned}$$

for all τ that are in $(H^1(K))^2$ for every K , and for all $v \in V(\mathcal{K}_h)$. From (11) and (9) we easily deduce the following proposition.

Proposition 2.1. *There exists a constant C_{cont} , depending only on ρ_1 and ρ_2 in (3)-(4), such that*

$$(12) \quad a(u, v) \leq C_{cont} \|u\|_{V(\mathcal{K}_h)} \|v\|_{V(\mathcal{K}_h)} \quad \forall u, v \in V(\mathcal{K}_h).$$

3. APPROXIMATION AND ABSTRACT ERROR ESTIMATES

For every element K we choose now a finite dimensional polynomial space $V(K)$ and we define

$$(13) \quad \Sigma(K) := \nabla(V(K))$$

On the choice of the spaces $V(K)$ we make the following assumptions.

- $V(K)$ contains all the polynomials of degree ≤ 1 , that is

$$(14) \quad V(K) \supset P_1.$$

- There exists an integer $N > 0$ such that for all h and for all $K \in \mathcal{K}_h$ we have

$$(15) \quad V(K) \subset P_N$$

where P_N is the space of polynomials of degree $\leq N$.

- There exists a constant $\gamma_1 > 0$ such that: for all h , for all $K \in \mathcal{K}_h$, and for all function $g \in L^2(\partial K)$, constant on each edge of ∂K , there exists $\ell(v) \in \Sigma(K)$ such that

$$(16) \quad \|\ell(v)\|_{0,K}^2 \leq \gamma_1 h_K^{-1} \|g\|_{0,\partial K}^2,$$

$$(17) \quad \int_e \frac{\partial \ell(v)}{\partial \mathbf{n}_K} - g \, ds = 0 \quad \text{for all edge } e \in \partial K \quad \forall K \in \mathcal{K}_h,$$

where \mathbf{n}_K is the outward normal to ∂K . We then extend our spaces to the whole Ω setting

$$V_h := \prod_K V(K), \quad \Sigma_h := \prod_K \Sigma(K).$$

The discrete problem is then:

$$(18) \quad \begin{cases} \text{Find } u_h \in V_h \text{ such that, } \forall v \in V_h : \\ (\nabla_h u_h, \nabla_h v)_- < \{\nabla_h u_h\}, [v] > + < \{\nabla_h v\}, [u_h] > = (f, v), \end{cases}$$

that, using (6), can also be written

$$(19) \quad a(u_h, v) = (f, v) \quad \forall v \in V_h.$$

In the finite element space V_h we introduce the usual DG norm

$$(20) \quad \|v\|^2 = |v|_{1,h}^2 + \|v\|_j^2.$$

We note immediately that, using (15) and our assumptions on the decomposition (2)-(4), with a simple use of the inverse inequality, we have that on V_h the DG norm (20) is equivalent to the norm (9) originally introduced in $V(\mathcal{K}_h)$. In particular we have

$$(21) \quad \|v_h\|_{V(\mathcal{K}_h)} \leq C_{inv} \|v_h\| \leq C_{inv} \|v_h\|_{V(\mathcal{K}_h)} \quad v_h \in V_h,$$

where C_{inv} depends only on n_* , ρ_1 , ρ_2 , and N . In a similar way (11) could be simplified to

$$(22) \quad | < \{\boldsymbol{\tau}\}, [v] > | \leq C_s \|\boldsymbol{\tau}\|_0 \|v\|_j \quad \forall \boldsymbol{\tau} \in \Sigma_h \quad \forall v \in V_h.$$

Hence we immediately have the following result.

Proposition 3.1. *There exists a constant C_{cont} , depending only on n_* , ρ_1 , ρ_2 , and N such that*

$$(23) \quad a(u_h, v_h) \leq C_{cont} \|u_h\| \|v_h\| \quad \forall u_h, v_h \in V_h. \quad \square$$

Our main task will be now to prove *stability* of the bilinear form $a(u, v)$ in the DG norm (20). This however will not be done by showing *ellipticity* of the bilinear form a , but rather by proving that there exists a mapping $S : V_h \rightarrow V_h$ such that

$$(24) \quad \sup_v \frac{a(u, v)}{\|v\|} \geq \frac{a(u, S(u))}{\|S(u)\|} \geq \kappa \|u\| \quad \forall u \in V_h$$

for a suitable constant κ depending only on the minimum angle of \mathcal{K}_h . The target (24) will be reached by constructing an operator S which is *bounded*

$$(25) \quad \|S(u)\| \leq \kappa_2 \|u\|,$$

and *bounding*

$$(26) \quad a(u, S(u)) \geq \kappa_1 \|u\|^2,$$

so that (24) will follow with $\kappa = \kappa_1/\kappa_2$. The construction of the operator S will be done in several steps.

To start with, for every element K and every $\boldsymbol{\tau} \in \Sigma(K)$ we define its *potential* $p(\boldsymbol{\tau})$ by

$$\nabla p(\boldsymbol{\tau}) = \boldsymbol{\tau} \quad \text{and} \quad \int_K p(\boldsymbol{\tau}) = 0.$$

Note that p is *one-to-one* from $\Sigma(K)$ to the subset of $V(K)$ of functions having zero mean value on K .

We then extend the above definitions globally, defining $p : \Sigma_h \rightarrow V_h$ in the (obvious) element by element way, and we note that every $v \in V_h$ can be split in a *unique* way as

$$(27) \quad v = v_0 + v_1 \quad \text{with} \quad v_0 = \text{piecewise constant} \quad \text{and} \quad v_1 = p(\nabla v).$$

We shall now prove the boundedness of the p operator.

Proposition 3.2. *There exists a constant C_p , depending only on n_* , ρ_1 , ρ_2 , and N , such that*

$$(28) \quad \|p(\boldsymbol{\tau})\| \leq C_p \|\boldsymbol{\tau}\|_{0,\Omega}$$

for all $\boldsymbol{\tau} \in \Sigma_h$.

Proof. We first note that, since $p(\boldsymbol{\tau})$ has zero mean value in each K , we have

$$\forall K, \forall \boldsymbol{\tau} \in \Sigma(K) : \|p(\boldsymbol{\tau})\|_{0,K}^2 \leq C_1 h_K^2 |p(\boldsymbol{\tau})|_{1,h}^2 = C_1 h_K^2 \|\boldsymbol{\tau}\|_0^2.$$

Hence, using again the Agmon inequality (10) we deduce that

$$\begin{aligned} \forall \boldsymbol{\tau} \in \Sigma_h : \|p(\boldsymbol{\tau})\|^2 &= \|\boldsymbol{\tau}\|_0^2 + \|p(\boldsymbol{\tau})\|_j^2 \\ &\leq \|\boldsymbol{\tau}\|_0^2 + C C_a \sum_K \frac{1}{h_K} (h_K^{-1} \|p(\boldsymbol{\tau})\|_{0,K}^2 + h_K |p(\boldsymbol{\tau})|_{1,K}^2) \\ &\leq (1 + C C_a C_1 + C C_a) \|\boldsymbol{\tau}\|_0^2, \end{aligned}$$

and the result (28) follows immediately. \square

Next, we construct a mapping L from the space of piecewise constant scalars to the space Σ_h . For v_0 piecewise constant, we define first on \mathcal{E}_h the vector valued function:

$$(29) \quad \mathbf{g}(v_0) = \llbracket v_0 \rrbracket.$$

Then, element by element, we construct $L(v_0)$ setting

$$(30) \quad \boldsymbol{\tau} = L(v_0) \quad \text{iff} \quad \int_e (\boldsymbol{\tau} - \mathbf{g}) \cdot \mathbf{n}_e \, ds = 0 \quad \forall \text{ edge } e \in \mathcal{E}_h,$$

\mathbf{n}_e being one of the two normal directions to e . The existence of $L(v_0)$ is provided by the existence of the operator ℓ in (16)-(17). The following two properties of the map L will play an important role in our analysis. The first is immediate, but important, and we state it as a proposition.

Proposition *Let L be the operator defined in (30). Then for every piecewise constant v_0 we have*

$$(31) \quad \sum_{e \in \mathcal{E}_h} \int_e L(v_0) \cdot \llbracket v_0 \rrbracket \, ds = \|v_0\|^2.$$

Proof. Equality (31) follows immediately from the definitions (30) (of L) and (20) (of the DG norm), taking into account that for a piecewise constant v_0 we have $|v_0|_{1,h} = 0$. \square

The second property expresses the continuity (uniform in h) of the mapping $v_0 \rightarrow p(L(v_0))$.

Proposition 3.3. *Let L be the operator defined in (30). Then there exists a constant γ , depending only on n_* , ρ_1 , ρ_2 , N , and γ_1 , such that*

$$(32) \quad \|p(L(v_0))\|_{0,\Omega}^2 \leq \gamma^2 \|v_0\|^2$$

for every piecewise constant v_0 .

Proof. We first note that, from the property (17) of the local operator ℓ , we have immediately that for all piecewise constants v_0

$$(33) \quad \|L(v_0)\|_{0,\Omega}^2 \leq C_i^2 \|v_0\|^2 \equiv C_i^2 \|v_0\|_j^2,$$

with a constant C_i depending only on γ_1 . Then the required (32) follows easily using (28):

$$(34) \quad \|p(L(v_0))\|_{0,\Omega}^2 \leq C_p^2 \|L(v_0)\|_{0,\Omega}^2 \leq C_p^2 C_i^2 \|v_0\|^2 =: \gamma^2 \|v_0\|^2.$$

\square

We can now start our proof of (25) and (26). The first step will be the following proposition.

Proposition 3.4. *There exists a constant κ_0 , depending only on n_* , ρ_1 , ρ_2 , N , and γ_1 , such that*

$$(35) \quad a(u, p(L(u_0))) \geq \|u_0\|^2 - \kappa_0 |u|_{1,h} \|u_0\|$$

for all $u \in V_h$, where u_0 is obtained from u through the splitting (27).

Proof. For $u \in V_h$, with $u = u_0 + u_1 = u_0 + p(\nabla_h u)$ as in (27), we have

$$\begin{aligned} a(u, p(L(u_0))) &= (\nabla_h u, L(u_0)) - \langle \{\nabla_h u\}, \llbracket p(L(u_0)) \rrbracket \rangle + \langle \{L(u_0)\}, \llbracket u \rrbracket \rangle \\ &= (\nabla_h u, L(u_0)) - \langle \{\nabla_h u\}, \llbracket p(L(u_0)) \rrbracket \rangle + \langle \{L(u_0)\}, \llbracket p(\nabla_h u) \rrbracket \rangle + \langle \{L(u_0)\}, \llbracket u_0 \rrbracket \rangle \\ &= (\nabla_h u, L(u_0)) - \langle \{\nabla_h u\}, \llbracket p(L(u_0)) \rrbracket \rangle + \langle \{L(u_0)\}, \llbracket p(\nabla_h u) \rrbracket \rangle + \|u_0\|^2. \end{aligned}$$

Using this, Cauchy-Schwarz inequality, (11), (32)-(33), and finally (28) we then have

$$\begin{aligned} \|u_0\|^2 &= a(u, p(L(u_0))) - (\nabla_h u, L(u_0)) + \langle \{\nabla_h u\}, \llbracket p(L(u_0)) \rrbracket \rangle - \langle \{L(u_0)\}, \llbracket p(\nabla_h u) \rrbracket \rangle \\ &\leq a(u, p(L(u_0))) + |u|_{1,h} \|L(u_0)\|_{0,\Omega} + C_s |u|_{1,h} \|p(L(u_0))\| + C_s \|L(u_0)\|_{0,\Omega} \|p(\nabla_h u)\| \\ &\leq a(u, p(L(u_0))) + |u|_{1,h} C_i \|u_0\| + C_s |u|_{1,h} \gamma \|u_0\| + C_s C_i \|u_0\| C_1 |u|_{1,h} \\ &= a(u, p(L(u_0))) + \kappa_0 |u|_{1,h} \|u_0\|, \end{aligned}$$

which is inequality (35). \square

The operator S will then be constructed as

$$(36) \quad S(u) := u + \alpha p(L(u_0))$$

for a suitable choice of α . It is clear that S , constructed as in (36), will be bounded. In particular we shall have

$$(37) \quad \begin{aligned} \|u + \alpha p(L(u_0))\|^2 &\leq 2(\|u\|^2 + \|\alpha p(L(u_0))\|^2) \leq 2(\|u\|^2 + \gamma^2 \alpha^2 \|u_0\|^2) \\ &\leq \kappa_2^2 \|u\|^2, \end{aligned}$$

that is precisely the *boundedness* property (25), with κ_2 depending only on α (still to be chosen) and on the usual parameters n_* , ρ_1 , ρ_2 , N , and γ_1 . Let us see that the *bounding* property (26) is also verified, for α small enough. Indeed, we remark first from the definition (6) of the bilinear form a that for all $v \in V_h$

$$(38) \quad a(v, v) = |v|_{1,h}^2$$

which is indeed the nicest feature of the Baumann-Oden formulation, compared with other DG formulations. Then we choose $\alpha := 2/(1 + \kappa_0^2)$, where κ_0 is given in (35), and we have

$$(39) \quad \begin{aligned} a(u, u + \alpha p(L(u_0))) &= |u|_{1,h}^2 + a(u, \alpha p(L(u_0))) \\ &\geq |u|_{1,h}^2 + \alpha(\|u_0\|^2 - \kappa_0 |u|_{1,h} \|u_0\|) \\ &= \frac{\alpha}{2}(|u|_{1,h}^2 + \|u_0\|^2) + (1 - \frac{\alpha}{2})|u|_{1,h}^2 + \frac{\alpha}{2}\|u_0\|^2 - \alpha\kappa_0 |u|_{1,h} \|u_0\| \\ &= \frac{1}{1 + \kappa_0^2}(|u|_{1,h}^2 + \|u_0\|^2) + \frac{1}{1 + \kappa_0^2}(\kappa_0 |u|_{1,h} - \|u_0\|)^2 \\ &\geq \frac{1}{1 + \kappa_0^2}(|u|_{1,h}^2 + \|u_0\|^2). \end{aligned}$$

On the other hand, using (27) and then (28) we have

$$(40) \quad \|u\|^2 \leq 2(\|u_0\|^2 + \|p(\nabla u)\|^2) \leq 2(\|u_0\|^2 + C_p^2 |u|_{1,h}^2),$$

which combined with (39) gives

$$(41) \quad a(u, u + \alpha p(L(u_0))) \geq \kappa_1 \|u\|^2,$$

with κ_1 depending only on κ_0 and C_p , that is (26).

We summarize the result in the following theorem

Theorem 3.5. *There exists a constant κ , depending only on n_* , ρ_1 , ρ_2 , N , and γ_1 , such that: for every $u_h \in V_h$, different from zero, there exists a $v_h (= S(u_h))$ in V_h , different from zero, such that*

$$(42) \quad a(u_h, v_h) \geq \kappa \|u_h\| \|v_h\|.$$

Now it is classical to deduce the error estimate.

Theorem 3.6. *In the above assumptions, for every $f \in L^2(\Omega)$ the discrete problem (18) has a unique solution u_h . Moreover the distance between u_h and the solution u of (1) can be estimated as*

$$(43) \quad \|u - u_h\|_{V(\mathcal{K}_h)} \leq C h |u|_{2,\Omega}.$$

where C is a constant depending only on n_* , ρ_1 , ρ_2 , N , and γ_1 .

Proof. The proof is now classical. We start by defining u_I , in each K , as the $L^2(K)$ projection of u onto the space of polynomials of degree ≤ 1 . The function u_I will belong to V_h thanks to (14). With usual arguments we have immediately that, for each K ,

$$(44) \quad |u - u_I|_{r,K} \leq C_I h_K^{2-r} |u|_2 \quad (r = 0, 1, 2),$$

with a constant C_I depending only on n_* , ρ_1 , ρ_2 , N , and γ_1 . Then we use (21), then (42), then Galerkin orthogonality, then (12), then again (21) to obtain

$$(45) \quad \begin{aligned} \|u_h - u_I\|_{V(\mathcal{K}_h)} &\leq C_{inv} \|u_h - u_I\| \leq \frac{C_{inv} a(u_h - u_I, S(u_h - u_I))}{\kappa \|S(u_h - u_I)\|} \\ &= \frac{C_{inv} a(u - u_I, S(u_h - u_I))}{\kappa \|S(u_h - u_I)\|} \\ &\leq \frac{C_{inv} C_{cont} \|u - u_I\|_{V(\mathcal{K}_h)} \|S(u_h - u_I)\|_{V(\mathcal{K}_h)}}{\kappa \|S(u_h - u_I)\|} \\ &\leq \frac{C_{inv} C_{cont} \|u - u_I\|_{V(\mathcal{K}_h)} C_{inv} \|S(u_h - u_I)\|}{\kappa \|S(u_h - u_I)\|} \\ &= \frac{C_{inv}^2 C_{cont}}{\kappa} \|u - u_I\|_{V(\mathcal{K}_h)} \end{aligned}$$

and the result follows using (44). \square

Remark 3.7. *We point out that both (31) and (32) are global properties. We proved them using (15)-(17), that are instead local (ad hoc) properties, in each K . It is clear that proving global properties starting from our element by element construction is easy, and this is the main motivation for our choice. On the other hand we could have asked that the space V_h satisfy some convenient global properties (but not necessarily (16), (17)). And it is quite possible (and also likely, in our opinion) that (31) and (32) (and hence stability) could hold as well for local choices of $V(K)$ that do not satisfy (15)-(17). For instance, taking on a regular grid of rectangles $V(K) = Q_1 \oplus x^2 + y^2$ does not satisfy (16)-(17) but the numerical results are still quite good, as we shall see in the last Section.*

4. APPLICATIONS

Form the above theory it is clear that, once we are given a decomposition \mathcal{K}_h satisfying (2)-(4), the only problem that is left is to find, for every K , a suitable space $V(K)$ satisfying (14)-(15) (this is easy), and such that an operator ℓ with the properties (16)-(17) exists. This obviously depends on the geometry of K .

As already discussed in [8], when K is a triangle (with the usual *shape regularity properties* implied by (2)-(4)) one can choose

$$V(K) := \{v \mid v = a + bx + cy + d(x^2 + y^2)\},$$

implying

$$\Sigma(K) := \nabla(V(K)) = RT_0(K),$$

where $RT_0(K)$ denotes the lowest order Raviart-Thomas space over the element K .

This, as analyzed in more details in in [8], easily implies the existence of the lifting operator ℓ satisfying (16)-(17). Indeed the values of the normal component on each edge is the most commonly used set of degrees of freedom for the lowest order Raviart-Thomas element.

Hence we turn our attention to quadrilaterals, whose use is much less straightforward than that of triangles. We denote by e^1, e^2, e^3, e^4 the edges of the quadrilateral K , and by \mathbf{t}^i and \mathbf{n}^i the unit tangent and normal vector to the edge e^i ($i = 1, 4$) in the clockwise ordering of the edges. While ordering the edges, we also choose e^1 and e^2 in such a way that

$$(46) \quad \nabla(\mathbf{t}^1 \cdot \mathbf{x}) \cdot \mathbf{n}^3|_{e^3} \quad \text{and} \quad \nabla(\mathbf{t}^2 \cdot \mathbf{x}) \cdot \mathbf{n}^4|_{e^4} \quad \text{have the same sign.}$$

(See figure 2).

FIGURE 2. The quadrilateral and the vectors \mathbf{t}^i .

Let $m_{31}(x, y) = 0$ be the equation of the straight line connecting the midpoints of e^3 and e^1 , and let $m_{42}(x, y) = 0$ be the equation of the straight line connecting the midpoints of e^4 and e^2 respectively. For future use, we shall also assume that $m_{31}(A) = m_{42}(A) = 1$, where A is the vertex of K common to e^1 and e^2 . We then set

$$(47) \quad \varphi^0(\mathbf{x}) \equiv 1, \quad \varphi^1(\mathbf{x}) \equiv \mathbf{t}^1 \cdot \mathbf{x}, \quad \varphi^2(\mathbf{x}) \equiv -\mathbf{t}^2 \cdot \mathbf{x}, \quad \varphi^3(\mathbf{x}) \equiv m_{31}^2, \quad \varphi^4(\mathbf{x}) \equiv m_{41}^2$$

and we take

$$(48) \quad V(K) := \text{span}\{\varphi^0, \varphi^1, \varphi^2, \varphi^3, \varphi^4\} \equiv \text{span}\{1, \mathbf{t}^1 \cdot \mathbf{x}, -\mathbf{t}^2 \cdot \mathbf{x}, m_{31}^2, m_{41}^2\}.$$

We have then the following proposition.

Proposition 4.1. *Let $V(K)$ be defined as in (48) and assume that K satisfies the assumptions (3)-(4). In particular γ_1 depends only on ρ_1 and ρ_2 .*

Proof. The existence of an operator ℓ satisfying (17) will be proved if we show that

$$(49) \quad \sum_{i=1}^4 \alpha_i \int_{\partial K} \nabla \varphi^i \cdot \mathbf{n}^i = 0 \Leftrightarrow \alpha_i \equiv 0 \forall i.$$

Then, after (49) is proven, assumptions (3)-(4) and a usual scaling argument give easily (16) with γ_1 depending only on ρ_1 and ρ_2 .

In order to prove (49) we form the 4×4 matrix M with coefficients

$$(50) \quad M_{ij} = \int_{e^j} \nabla \varphi^i \cdot \mathbf{n}^j, \quad i, j = 1, 4,$$

and prove that $\det(M) \neq 0$. We first notice that $\nabla \varphi^3$ is a vector valued polynomial of degree 1, and that both components vanish identically on the line $m_{31} = 0$; hence by the midpoint quadrature formula we have

$$(51) \quad \int_{e^1} \nabla \varphi^3 \cdot \mathbf{n}^1 = \int_{e^3} \nabla \varphi^3 \cdot \mathbf{n}^3 = 0.$$

A similar argument shows that

$$(52) \quad \int_{e^2} \nabla \varphi^4 \cdot \mathbf{n}^2 = \int_{e^4} \nabla \varphi^4 \cdot \mathbf{n}^4 = 0.$$

Moreover $\nabla \varphi^3 \cdot \mathbf{n}^2$ is positive on e^2 and $\nabla \varphi^3 \cdot \mathbf{n}^4$ is positive on e^3 , and so are their integrals on e^2 and on e^3 respectively. We denote them by

$$(53) \quad \int_{e^2} \nabla \varphi^3 \cdot \mathbf{n}^2 =: \gamma_2 > 0, \quad \int_{e^4} \nabla \varphi^3 \cdot \mathbf{n}^4 =: \gamma_4 > 0.$$

Similarly, always referring to Figure2:

$$(54) \quad \int_{e^1} \nabla \varphi^4 \cdot \mathbf{n}^1 = \delta_1 > 0, \quad \int_{e^3} \nabla \varphi^4 \cdot \mathbf{n}^3 = \delta_3 > 0.$$

Next, for φ^1 and φ^2 we obviously have

$$(55) \quad \int_{e^1} \nabla \varphi^1 \cdot \mathbf{n}^1 = 0, \quad \int_{e^2} \nabla \varphi^2 \cdot \mathbf{n}^2 = 0,$$

and, referring to Figure 2,

$$(56) \quad \begin{aligned} \int_{e^2} \nabla \varphi^1 \cdot \mathbf{n}^2 &= \alpha_2 > 0, & \int_{e^4} \nabla \varphi^1 \cdot \mathbf{n}^4 &= -\alpha_4 < 0, \\ \int_{e^1} \nabla \varphi^2 \cdot \mathbf{n}^1 &= \beta_1 > 0, & \int_{e^3} \nabla \varphi^2 \cdot \mathbf{n}^3 &= -\beta_3 < 0. \end{aligned}$$

Finally, by Gauss theorem

$$(57) \quad \int_{e^3} \nabla \varphi^1 \cdot \mathbf{n}^3 = -\alpha_2 + \alpha_4, \quad \int_{e^4} \nabla \varphi^2 \cdot \mathbf{n}^4 = -\beta_1 + \beta_3.$$

and using (46)

$$(58) \quad (-\alpha_2 + \alpha_4)(-\beta_1 + \beta_3) < 0.$$

Collecting (51)–(58) we obtain the matrix

$$(59) \quad M = \begin{bmatrix} 0 & \alpha_2 & \alpha_4 - \alpha_2 & -\alpha_4 \\ \beta_1 & 0 & -\beta_3 & \beta_3 - \beta_1 \\ 0 & \gamma_2 & 0 & \gamma_4 \\ \delta_1 & 0 & \delta_3 & 0 \end{bmatrix}$$

As both δ_1 and γ_2 are different from zero, the determinant $\det(M)$ is different from zero if and only if the determinant of

$$(60) \quad M^* = \begin{bmatrix} 0 & \alpha_2 & \alpha_4 - \alpha_2 & -\alpha_4 \\ \beta_1 & 0 & -\beta_3 & \beta_3 - \beta_1 \\ 0 & 1 & 0 & \gamma_4/\gamma_2 \\ 1 & 0 & \delta_3/\delta_1 & 0 \end{bmatrix}$$

is different from zero. In its turn, $\det(M^*)$ is different from zero if, and only if, the determinant of

$$(61) \quad M^{**} = \begin{bmatrix} 0 & 0 & \alpha_4 - \alpha_2 & \alpha_4 \alpha_2 \gamma_4 / \gamma_2 \\ 0 & 0 & \beta_3 \beta_1 \delta_3 / \delta_1 & \beta_3 - \beta_1 \\ 0 & 1 & 0 & \gamma_4 / \gamma_2 \\ 1 & 0 & \delta_3 / \delta_1 & 0 \end{bmatrix}$$

is different from zero. Finally using (58) and the (positive) sign of α_2 , α_4 , β_1 , β_3 , γ_2 , γ_4 , δ_1 , δ_3 we have

$$(62) \quad \det(M^{**}) = (\alpha_4 - \alpha_2)(\beta_3 - \beta_1) - \frac{\alpha_4 \alpha_2 \gamma_4 \beta_3 \beta_1 \delta_3}{\gamma_2 \delta_1} < 0$$

and (49) follows. \square

5. NUMERICAL RESULTS

We took $\Omega =]0, 1[^2$, and f and the boundary conditions such that the exact solution was, in all cases

$$(63) \quad u(x, y) = e^{xy}(x - 1)(y - 1).$$

We tested three different types of sequences decompositions into quadrilaterals, that we denote by *squares*, *quadrilaterals* and *squares (asymptotically parallelograms)*. An example of each of the first two sequences (squares and quadrilaterals) is shown in Figure 3, together with two examples of the last sequence (asymptotically parallelograms), one coarser and one finer, to show that in this sequence the elements become more and more regular as the mesh becomes finer.

For each mesh we tested seven different approximations: for the first two, ($NIPG - Q_1$ and $NIPG - P_1$, respectively, the method (19) is stabilized by adding the usual jump penalty to the bilinear form a . For the other five approximations there is no jump penalty (hence we are doing *pure Baumann-Oden*) and the difference among the methods is given by the choice of the local spaces $V(K)$, which are taken, respectively, as follows. For the reference square:

for $NIPG - Q_1$	$V(K) := span\{1, x, y, xy\}$
for $NIPG - P_1$	$V(K) := span\{1, x, y\}$
for $B - O - Q_2$	$V(K) := span\{1, x, y, xy, x^2, y^2, x^2y, xy^2, x^2y^2\}$
for $B - O - P_2$	$V(K) := span\{1, x, y, xy, x^2, y^2\}$
for $B - O - P_1 + Bubble$	$V(K) := span\{1, x, y, m_{13}^2, m_{24}^2\}$
for $B - O - Q_1 + Bubble(Q2)$	$V(K) := span\{1, x, y, xy, b_2\}$
for $B - O - Q_1 + Bubble(Q4)$	$V(K) := span\{1, x, y, xy, b_4\}$

where m_{13} and m_{24} have been described in the previous section, $b_2 \dots$:-) I don't know what it is because the first author did not explain it properly :-), and b_4 is the product of the equations of the four edges.

For each sequence of grids

REFERENCES

- [1] D.N. Arnold, F. Brezzi, B. Cockburn, and L.D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.* (2002) **39**, 1749–1779.
- [2] C. Baiocchi, F. Brezzi, and L.P. Franca, Virtual bubbles and GaLS. *Comput. Methods Appl. Mech. Engrg.* (1993), **105**, 125–141.
- [3] C. E. Baumann and J. T. Oden, A discontinuous *hp* finite element method for the Navier-Stokes equations, 10th. International Conference on Finite Element in Fluids, 1998.
- [4] C. E. Baumann and J. T. Oden, A discontinuous *hp* finite element method for convection-diffusion problems. *Comput. Methods Appl. Mech. Engrg.* (1999) **175**, 311–341.
- [5] F. Brezzi, M.O. Bristeau, L.P. Franca, M. Mallet, and G. Rogé, A relationship between stabilized finite element methods and the Galerkin method with bubble functions. *Comput. Methods Appl. Mech. Engrg.* (1992), **96**, 117–129.
- [6] F. Brezzi, B. Cockburn, L.D. Marini, and E. Süli, Stabilization mechanisms in discontinuous Galerkin finite element methods. *Comput. Methods Appl. Mech. Engrg.* (2006) **195**, No.25-28, 3293-3310.
- [7] F. Brezzi, L. P. Franca, T. J. R. Hughes, and A. Russo, $b = \int g$. *Comput. Methods Appl. Mech. Engrg.* (1997), **142**, 353–360.

- [8] F. Brezzi, L. D. Marini, Bubble stabilization of discontinuous Galerkin methods, in *W. Fitzgibbon, R. Hoppe, J. Periaux, O. Pironneau, Y. Vassilevski (Eds.): Proceedings of International Conference on the occasion of the 60th birthday of Y.A. Kuznetsov* Institute of Numerical Mathematics of the Russian Academy of Sciences. Moskow, (2006) 25-36
- [9] F. Brezzi, L. D. Marini, and A. Russo, Applications of Pseudo Residual-Free Bubbles to the Stabilization of Convection-Diffusion Problems. *Comput. Methods Appl. Mech. Engrg.* (1998), **166**, 51-63.
- [10] F. Brezzi, L. D. Marini, and A. Russo, On the choice of a Stabilizing Subgrid for Convection-Diffusion Problems. *Comput. Methods Appl. Mech. Engrg.* (2005), **194**, 127-148.
- [11] F. Brezzi, L.D. Marini, and E. Süli, Discontinuous Galerkin methods for first-order hyperbolic problems. *Math. Models Methods Appl. Sci.* (2004) **14**, No.12, 1893-1903.
- [12] F. Brezzi and A. Russo, Choosing bubbles for advection-diffusion problems. *Math. Models Meth. Appl. Sci.* (1994), **4**, 571-587.
- [13] A. N. Brooks and 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.* (1982), **32**, 199-259.
- [14] E. Burman and B. Stamm *To appear*
- [15] B. Cockburn, Discontinuous Galerkin methods, *Zeitschrift fur Angewandte Mathematik und Mechanik* (2003), **83**, No.11, 731-754.
- [16] L. P. Franca and A. Russo, Deriving upwinding, mass lumping and selective reduced integration by RFB. *Appl. Math. Letters* (1996), **9**, 83-88.
- [17] AGGIUNGERE QUALCOSA SU NIPG
- [18] P. Lesaint, and P.-A. Raviart, On a finite element method for solving the neutron transport equation. In: *Mathematical Aspects of Finite Elements in Partial Differential Equations* (Ed. C.A. deBoor), Academic Press, New York, 1974, pp. 89-123.
- [19] J. T. Oden, I. Babuška, and C. E. Baumann, A discontinuous *hp* finite element method for diffusion problems, *J. Comput. Phys.* **146** (1998), 491-519.
- [20] B. Rivière, M. F. Wheeler, and V. Girault, Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. *Comp. Geo.* (1999) **3**, 337-360.

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI PAVIA, ITALY, VIA FERRATA 1, 27100, PAVIA, ITALY

E-mail address: Paola.Antonietti@unipv.it

URL: <http://www-dimat.unipv.it/~antonietti/>

ISTITUTO UNIVERSITARIO DI STUDI SUPERIORI, PAVIA, ITALY AND IMATI-CNR, VIA FERRATA 1, 27100, PAVIA, ITALY

E-mail address: brezzi@imati.cnr.it

URL: <http://www.imati.cnr.it/~brezzi/>

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI PAVIA, ITALY AND IMATI-CNR, VIA FERRATA 1, 27100, PAVIA, ITALY

E-mail address: marini@imati.cnr.it

URL: <http://www.imati.cnr.it/~marini/>