ESAIM: MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

PETR KNOBLOCH
LUTZ TOBISKA

Stabilization methods of bubble type for the Q_1/Q_1-element applied to the incompressible Navier-Stokes equations

ESAIM: Modélisation mathématique et analyse numérique, tome 34, nº 1 (2000), p. 85-107

<http://www.numdam.org/item?id=M2AN_2000__34_1_85_0>

Mathematical Modelling and Numerical Analysis M2AN, Vol. 34, N° 1, 2000, pp. 85–107
Modélisation Mathématique et Analyse Numérique

STABILIZATION METHODS OF BUBBLE TYPE FOR THE Q_1/Q_1–ELEMENT APPLIED TO THE INCOMPRESSIBLE NAVIER–STOKES EQUATIONS

PETR KNOBLOCH[1] AND LUTZ TOBISKA[2]

Abstract. In this paper, a general technique is developed to enlarge the velocity space V_h^1 of the unstable Q_1/Q_1–element by adding spaces V_h^2 such that for the extended pair the Babuška–Brezzi condition is satisfied. Examples of stable elements which can be derived in such a way imply the stability of the well-known Q_2/Q_1–element and the $4Q_1/Q_1$–element. However, our new elements are much more cheaper. In particular, we shall see that more than half of the additional degrees of freedom when switching from the Q_1 to the Q_2 and $4Q_1$, respectively, element are not necessary to stabilize the Q_1/Q_1–element. Moreover, by using the technique of reduced discretizations and eliminating the additional degrees of freedom we show the relationship between enlarging the velocity space and stabilized methods. This relationship has been established for triangular elements but was not known for quadrilateral elements. As a result we derive new stabilized methods for the Stokes and Navier–Stokes equations. Finally, we show how the Brezzi–Pitkäranta stabilization and the SUPG method for the incompressible Navier–Stokes equations can be recovered as special cases of the general approach. In contrast to earlier papers we do not restrict ourselves to linearized versions of the Navier–Stokes equations but deal with the full nonlinear case.

Mathematics Subject Classification. 65N30, 65N12, 76D05.

Received: November 25, 1998. Revised: July 6, 1999.

1. INTRODUCTION

In this paper we introduce a general class of stable finite element spaces suitable for a numerical solution of the Stokes equations

$$-\nu\,\Delta\,\boldsymbol{u} + \nabla p = \boldsymbol{f}\,, \quad \operatorname{div}\boldsymbol{u} = 0 \quad \text{in } \Omega\,, \qquad \boldsymbol{u} = \boldsymbol{0} \ \text{ on } \partial\Omega\,, \tag{1}$$

the Navier–Stokes equations

$$-\nu\,\Delta\,\boldsymbol{u} + (\nabla\boldsymbol{u})\boldsymbol{u} + \nabla p = \boldsymbol{f}\,, \quad \operatorname{div}\boldsymbol{u} = 0 \quad \text{in } \Omega\,, \qquad \boldsymbol{u} = \boldsymbol{0} \ \text{ on } \partial\Omega \tag{2}$$

or other problems describing incompressible materials. In the equations (1) and (2), $\boldsymbol{u}$ is the velocity and p is the pressure in a linear viscous fluid contained in a bounded domain $\Omega \subset \mathbb{R}^2$ with a polygonal boundary $\partial\Omega$.

Keywords and phrases. Babuška–Brezzi condition, stabilization, Stokes equations, Navier–Stokes equations.

[1] Institute of Numerical Mathematics, Faculty of Mathematics and Physics, Charles University, Malostranské náměstí 25, 118 00 Praha 1, Czech Republic. e-mail: `knobloch@karlin.mff.cuni.cz`

[2] Institute of Analysis and Numerics, Otto von Guericke University, PF 4120, 39016 Magdeburg, Germany.
e-mail: `tobiska@mathematik.uni-magdeburg.de`

The parameter $\nu > 0$ is the kinematic viscosity and $\boldsymbol{f}$ is an external body force, *e.g.* the gravity. Denoting

$$a(\boldsymbol{u},\boldsymbol{v}) = \int_\Omega \nabla \boldsymbol{u} \cdot \nabla \boldsymbol{v} \,\mathrm{d}x\,, \qquad n(\boldsymbol{u},\boldsymbol{w},\boldsymbol{v}) = \int_\Omega \boldsymbol{v}\cdot(\nabla \boldsymbol{w})\boldsymbol{u}\,\mathrm{d}x\,,$$
$$b(\boldsymbol{v},p) = -\int_\Omega p \operatorname{div} \boldsymbol{v}\,\mathrm{d}x\,,$$

the usual weak formulation of (1) reads: Given $\nu > 0$ and $\boldsymbol{f} \in H^{-1}(\Omega)^2$, find $\boldsymbol{u} \in H_0^1(\Omega)^2$ and $p \in L_0^2(\Omega)$ such that

$$\nu\, a(\boldsymbol{u},\boldsymbol{v}) + b(\boldsymbol{v},p) - b(\boldsymbol{u},q) = \langle \boldsymbol{f},\boldsymbol{v}\rangle \qquad \forall\ \boldsymbol{v} \in H_0^1(\Omega)^2,\ q \in L_0^2(\Omega)\,, \tag{3}$$

where $L_0^2(\Omega)$ consists of $L^2(\Omega)$ functions having zero mean value on Ω. It can be shown that this problem has a unique solution (*cf.* [14], p. 80, Theorem 5.1). The weak formulation of (2) is given by

$$\nu\, a(\boldsymbol{u},\boldsymbol{v}) + n(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}) + b(\boldsymbol{v},p) - b(\boldsymbol{u},q) = \langle \boldsymbol{f},\boldsymbol{v}\rangle \quad \forall\ \boldsymbol{v} \in H_0^1(\Omega)^2,\ q \in L_0^2(\Omega)\,. \tag{4}$$

The problem (4) has a solution which is unique if ν is sufficiently large and/or $\boldsymbol{f}$ is sufficiently small (*cf.* [14], pp. 291 and 292).

A standard Galerkin finite element discretization of (3) reads: Find $\boldsymbol{u}_h \in \mathrm{V}_h$ and $p_h \in \mathrm{Q}_h$ satisfying

$$\nu\, a(\boldsymbol{u}_h,\boldsymbol{v}_h) + b(\boldsymbol{v}_h,p_h) - b(\boldsymbol{u}_h,q_h) = \langle \boldsymbol{f},\boldsymbol{v}_h\rangle \qquad \forall\ \boldsymbol{v}_h \in \mathrm{V}_h,\ q_h \in \mathrm{Q}_h\,, \tag{5}$$

where $\mathrm{V}_h \subset H_0^1(\Omega)^2$ and $\mathrm{Q}_h \subset L_0^2(\Omega)$ are some finite element spaces defined using a triangulation $\mathcal{T}_h$ of Ω. In this paper, we shall consider only triangulations consisting of quadrilaterals T (cf. Sect. 2) and we shall use the spaces

$$\mathrm{V}_h^1 = \{\boldsymbol{v} \in H_0^1(\Omega)^2;\ \boldsymbol{v}\circ F_T \in Q_1(\widehat{T})^2 \ \ \forall\, T \in \mathcal{T}_h\}\,,$$
$$\mathrm{Q}_h = \{q \in H^1(\Omega)\cap L_0^2(\Omega);\ q\circ F_T \in Q_1(\widehat{T}) \ \ \forall\, T \in \mathcal{T}_h\}$$

for approximating the velocity and the pressure, respectively. Here, $Q_1(\widehat{T})$ is the space of bilinear functions defined on the reference square $\widehat{T}$ and $F_T \in Q_1(\widehat{T})^2$ is a one–to–one mapping which maps $\widehat{T}$ onto T. It is well known that this pair of spaces does not satisfy the Babuška–Brezzi condition

$$\exists\ \beta > 0\ : \qquad \sup_{\boldsymbol{v}_h \in \mathrm{V}_h^1\setminus\{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}_h,q_h)}{|\boldsymbol{v}_h|_{1,\Omega}} \geq \beta\, \|q_h\|_{0,\Omega} \quad \forall\ q_h \in \mathrm{Q}_h,\ h > 0\,, \tag{6}$$

which often causes that the problem (5) with $\mathrm{V}_h = \mathrm{V}_h^1$ is not solvable or that its solution contains spurious oscillations. One way to suppress these oscillations and to assure the solvability is to add some extra terms to the discretization (5) (*cf. e.g.* [7, 9, 15, 19]). Another way is to enlarge the space V_h^1 by a space V_h^2 so that the Babuška–Brezzi condition is satisfied. Here we shall first consider the second possibility and construct a general class of spaces V_h^2 assuring the fulfilment of the Babuška–Brezzi condition. Then we shall show that, for suitable spaces V_h^2, the V_h^1–component of $\boldsymbol{u}_h$ and the function p_h are solutions of the stabilized methods of [9, 15].

In case of the mini element [1], which is defined by enriching continuous piecewise linear functions by cubic bubble functions, the close relation to the stabilized methods of [9, 15] was already discussed in [3, 18]. Similar results for a convection–diffusion equation were obtained in [6]. In an abstract framework, the equivalence between Galerkin methods with bubble functions and stabilized methods was investigated for linear problems in [2]. For the linearized incompressible Navier–Stokes equations, the relation between a Galerkin method with the mini element and the streamline upwind Petrov–Galerkin method (SUPG) was studied in [12]. In [20], this relation was investigated for residual–free bubbles and it was shown for the triangular P_1/P_1–element that also the correct stabilization parameters in both the diffusion–dominated and the convection–dominated regimes can be recovered. However, generally, *e.g.* for the Q_1/Q_1–element considered here, a stabilization using residual-free bubbles is not equivalent to the SUPG method (*cf.* [8]). Finally, it was also shown that bubble functions can help to design new stabilized methods (*cf. e.g.* [11,13]).

There is a lot of further papers devoted to investigations of discretizations stabilized using bubble functions, but the most of them are restricted to *triangular* elements and to *linear* problems. In this paper, we deal with quadrilateral elements and, in addition, we consider more general functions than bubble functions. Apart from investigating the relations to some well–known stabilized methods, we shall also derive, eliminating a suitable space V_h^2 from the discretization, a new type of stabilization which can be applied to both the Stokes and the Navier–Stokes equations. In addition, we shall establish a discretization of the Navier–Stokes equations which is, after elimination of a suitable space V_h^2, equivalent to the SUPG method studied for the linearized Navier–Stokes equation in [12] and in the full nonlinear case in [21].

The space V_h^2 added to V_h^1 to satisfy the Babuška–Brezzi condition will be defined in a general way as

$$\mathrm{V}_h^2 = \operatorname{span}\{\varphi_h^i \, \boldsymbol{t}_h^i\}_{i=1}^{N_h} ,$$

where $\varphi_h^i \in H_0^1(\Omega)$ and $\boldsymbol{t}_h^i \in \mathbb{R}^2$ are some suitable functions and vectors, respectively. The proof of the Babuška–Brezzi condition for the spaces $\mathrm{V}_h \equiv \mathrm{V}_h^1 \oplus \mathrm{V}_h^2$ and Q_h, which uses some ideas of [4] and a modification of the Verfürth trick [22], requires that the functions φ_h^i have localized supports and that, for any φ_h^i, there exists a point $A_h^i \in \overline{\Omega}$ such that

$$\int_\Omega \frac{\partial q_h}{\partial \boldsymbol{t}_h^i}\, \varphi_h^i \,\mathrm{d}x = \frac{\partial q_h}{\partial \boldsymbol{t}_h^i}(A_h^i) \int_\Omega \varphi_h^i \,\mathrm{d}x \qquad \forall\, q_h \in \mathrm{Q}_h\,,\ i = 1,\dots,N_h\,, \tag{7}$$

$$|q_h|_{1,T}^2 \le C\, h_T^2 \sum_{\substack{i=1,\\ A_h^i \in T}}^{N_h} \left| \frac{\partial q_h}{\partial \boldsymbol{t}_h^i}(A_h^i) \right|^2 \qquad \forall\, q_h \in \mathrm{Q}_h,\ T \in \mathcal{T}_h\,. \tag{8}$$

We shall give explicit examples of spaces V_h^2 such that (7) and (8) are fulfilled.

If A_h^i lies on an edge E of the triangulation $\mathcal{T}_h$, the corresponding function φ_h^i can be associated with E and we require that $\boldsymbol{t}_h^i$ is tangent to E. In other words, vector functions associated with edges used to stabilize the Q_1/Q_1–element are tangent to the edges. This is not the case for a stabilization of finite elements with discontinuous pressure like the quadrilateral Q_1/P_0–element or the triangular P_1/P_0–element, where vector functions orthogonal to the edges are used (see [5,10]).

The plan of the paper is as follows. In Section 2, we introduce some notations and summarize the assumptions on the triangulations and the functions φ_h^i needed for proving the Babuška–Brezzi condition in Section 3. In Section 4, we give some examples of the functions φ_h^i and construct proper subspaces of the stable Q_2/Q_1–element and the stable $4Q_1/Q_1$–element which satisfy the Babuška–Brezzi condition. Further, in this section, we also recover the stability of the Q_1–bubble/Q_1–element by Mons and Rogé [17]. We investigate discretizations obtained from (5) by eliminating the V_h^2–component of $\boldsymbol{u}_h$ in Section 5 and discuss the general framework between this technique and stabilized schemes. Particularly, we derive a new type of stabilization in Section 6 and show the equivalence to the stabilized methods of [9,15] in Sections 7 and 8. Finally, in Section 9, we show that, for a modified discretization of the incompressible Navier–Stokes equations and a suitable choice of the space V_h^2, the V_h^1–component of $\boldsymbol{u}_h$ and the function p_h are solutions of the SUPG method analyzed in [12,21].

2. Assumptions and notations

We assume that we are given a family $\{\mathcal{T}_h\}$ of triangulations of the domain Ω parametrized by a positive parameter $h \to 0$ and having the following properties. Each triangulation $\mathcal{T}_h$ consists of a finite number of closed convex quadrilaterals T (which will be often called elements in the following) such that $h_T \equiv \mathrm{diam}(T) \le h$, $\overline{\Omega} = \bigcup_{T\in\mathcal{T}_h} T$ and any two different elements T_1, $T_2 \in \mathcal{T}_h$ are either disjoint or possess either a common vertex or a common edge. In order to prevent the elements from degenerating when h tends to zero, we assume that any triangle $\widetilde{T}$, the vertices of which are three vertices of an element $T \in \mathcal{T}_h$, satisfies

$$\frac{h_{\widetilde{T}}}{\varrho_{\widetilde{T}}} \le C_1 \,, \tag{9}$$

where

$$h_{\widetilde{T}} = \mathrm{diam}(\widetilde{T}) \equiv \sup_{x,y\in\widetilde{T}} |x-y| \,, \qquad \varrho_{\widetilde{T}} = \sup_{B\subset\widetilde{T} \text{ is a circle}} \mathrm{diam}(B)$$

and the constant C_1 is independent of h.

We introduce a reference Cartesian coordinate system with axes $\widehat{x}_1$, $\widehat{x}_2$ and we define a reference element $\widehat{T} = [0,1]^2$. For any $T \in \mathcal{T}_h$, we denote by $F_T = (F_{T1}, F_{T2})$ a fixed one–to–one mapping $F_T \in Q_1(\widehat{T})^2$ which maps $\widehat{T}$ onto T. Such a mapping always exists and the assumption (9) guarantees that

$$|F_T|_{1,\infty,\widehat{T}} \le C\, h_T \,, \qquad |F_T^{-1}|_{1,\infty,T} \le C\, h_T^{-1} \qquad \forall\, T \in \mathcal{T}_h \,, \tag{10}$$

where the constant C depends only on C_1. Thus, we have

$$C\, h_T \, \|v \circ F_T\|_{0,\widehat{T}} \le \|v\|_{0,T} \le \widetilde{C}\, h_T \, \|v \circ F_T\|_{0,\widehat{T}} \qquad \forall\, v \in L^2(T),\ T \in \mathcal{T}_h \,, \tag{11}$$

$$C\, |v \circ F_T|_{1,\widehat{T}} \le |v|_{1,T} \le \widetilde{C}\, |v \circ F_T|_{1,\widehat{T}} \qquad \forall\, v \in H^1(T),\ T \in \mathcal{T}_h \,. \tag{12}$$

We shall use the notation $J_T(\widehat{x}) = \mathrm{D}F_T/\mathrm{D}\widehat{x}(\widehat{x})$ for the Jacobi matrix of F_T.

In the following, we formulate general assumptions which are essential for the construction of the supplementary space V_h^2. Later, in Sections 4, 6, 7 and 9, we shall show how these assumptions can be satisfied in special cases.

We suppose that we are given functions $\{\widehat{\varphi}^\alpha\}_{\alpha\in\mathcal{P}} \subset H^1(\widehat{T})$ (where $\mathcal{P}$ is some parameter set which is usually finite) such that, for any $\alpha \in \mathcal{P}$, the function $\widehat{\varphi}^\alpha$ vanishes on at least three edges of $\widehat{T}$ and there exists a point $\widehat{A}^\alpha \in \widehat{T}$ different from the vertices of $\widehat{T}$ satisfying

$$\int_{\widehat{T}} \widehat{q}\, \widehat{\varphi}^\alpha \,\mathrm{d}\widehat{x} = \widehat{q}(\widehat{A}^\alpha) \int_{\widehat{T}} \widehat{\varphi}^\alpha \,\mathrm{d}\widehat{x} \qquad \forall\, \widehat{q} \in Q_1(\widehat{T}) \,. \tag{13}$$

Further, for any $\alpha \in \mathcal{P}$, we introduce a unit vector $\widehat{\boldsymbol{t}}^\alpha = (\widehat{t}_1^\alpha, \widehat{t}_2^\alpha)$ and we denote $\widehat{\boldsymbol{n}}^\alpha = (\widehat{t}_2^\alpha, \widehat{t}_1^\alpha)$. If $\widehat{A}^\alpha \in \partial\widehat{T}$, we require that $\widehat{\boldsymbol{t}}^\alpha$ coincides with the direction of the edge of $\widehat{T}$ containing $\widehat{A}^\alpha$. We admit $\widehat{\varphi}^\alpha = \widehat{\varphi}^\beta$ for $\alpha \ne \beta$ in order to be able to use the same function $\widehat{\varphi}^\alpha$ with two different directions $\widehat{\boldsymbol{t}}^\alpha$. For formal reasons, we also admit $\widehat{\varphi}^\alpha = 0$. In this case, (13) is automatically satisfied for any point $\widehat{A}^\alpha$.

Now, using the mappings F_T, we transform the functions $\widehat{\varphi}^\alpha$ onto elements T of a triangulation $\mathcal{T}_h$ and introduce finite element functions $\varphi_h^i \in H_0^1(\Omega) \setminus \{0\}$, $i = 1, \dots, N_h$, having their supports always in one or two elements. Precisely, we assume that, for any $i \in \{1, \dots, N_h\}$, either

$$\exists\ T \in \mathcal{T}_h,\ \alpha \in \mathcal{P}: \quad F_T(\widehat{A}^\alpha) \in \partial\Omega \cup \mathrm{int}\, T, \quad \varphi_h^i|_T = \widehat{\varphi}^\alpha \circ F_T^{-1}, \quad \varphi_h^i|_{\Omega\setminus T} = 0 \tag{14}$$

or

$$\exists\ T, T' \in \mathcal{T}_h,\ \alpha, \alpha' \in \mathcal{P}: \quad T \cap T' = \{\text{edge}\}, \quad F_T(\widehat{A}^\alpha) = F_{T'}(\widehat{A}^{\alpha'}),$$
$$\varphi_h^i|_T = \widehat{\varphi}^\alpha \circ F_T^{-1}, \quad \varphi_h^i|_{T'} = \widehat{\varphi}^{\alpha'} \circ F_{T'}^{-1}, \quad \varphi_h^i|_{\Omega\setminus(T\cup T')} = 0 \,. \tag{15}$$

The assumption $F_T(\widehat{A}^\alpha) = F_{T'}(\widehat{A}^{\alpha'})$ in (15) implies that $\widehat{A}^\alpha$, $\widehat{A}^{\alpha'} \in \partial\widehat{T}$. Note that the function φ_h^i defined by (15) may vanish on one of the elements T, T'. In both cases (14) and (15), we set

$$A_h^i = F_T(\widehat{A}^\alpha)\,, \qquad \boldsymbol{t}_h^i = J_T(\widehat{A}^\alpha)\,\widehat{\boldsymbol{t}}^\alpha / |J_T(\widehat{A}^\alpha)\,\widehat{\boldsymbol{t}}^\alpha|$$

and we denote

$$\begin{array}{llll} & \widehat{A}_T^i = \widehat{A}^\alpha\,, & \widehat{\boldsymbol{t}}_T^i = \widehat{\boldsymbol{t}}^\alpha\,, & \widehat{\boldsymbol{n}}_T^i = \widehat{\boldsymbol{n}}^\alpha \\ (\text{and} & \widehat{A}_{T'}^i = \widehat{A}^{\alpha'}\,, & \widehat{\boldsymbol{t}}_{T'}^i = \widehat{\boldsymbol{t}}^{\alpha'}\,, & \widehat{\boldsymbol{n}}_{T'}^i = \widehat{\boldsymbol{n}}^{\alpha'}\,). \end{array}$$

In the case of (15), we then also have $A_h^i = F_{T'}(\widehat{A}^{\alpha'})$ and

$$\boldsymbol{t}_h^i = J_{T'}(\widehat{A}^{\alpha'})\,\widehat{\boldsymbol{t}}^{\alpha'} / |J_{T'}(\widehat{A}^{\alpha'})\,\widehat{\boldsymbol{t}}^{\alpha'}| \qquad \text{or} \qquad \boldsymbol{t}_h^i = -J_{T'}(\widehat{A}^{\alpha'})\,\widehat{\boldsymbol{t}}^{\alpha'} / |J_{T'}(\widehat{A}^{\alpha'})\,\widehat{\boldsymbol{t}}^{\alpha'}|\,.$$

We suppose that the functions $\{\varphi_h^i\,\boldsymbol{t}_h^i\}_{i=1}^{N_h}$ are linearly independent and that

$$\operatorname{card}\Big\{\, i \in \{1,\dots,N_h\}\,;\; A_h^i \in T \,\Big\} \le C_2 \qquad \forall\, T \in \mathcal{T}_h\,,$$

where the constant C_2 is independent of h. The support of any function φ_h^i is contained in the union of the elements containing the point A_h^i which will be denoted by P_h^i. Thus, P_h^i consists of one or two elements. Further, we introduce the quantities

$$\gamma_h^i = \frac{|\int_\Omega \varphi_h^i \,\mathrm{d}x|}{|P_h^i|\,|\varphi_h^i|_{1,\Omega}}\,, \quad i = 1,\dots,N_h\,, \qquad\qquad \gamma_h = \min_{i=1,\dots,N_h} \gamma_h^i$$

which influence the magnitude of the constant in the Babuška–Brezzi condition. Defining the functions φ_h^i in a suitable way, the value of γ_h can be made arbitrarily small. However, arbitrarily large values of γ_h cannot be obtained. It can be shown that $\gamma_h \le 2\,C_1$ and, if $\mathcal{T}_h$ consists of rectangles, we even have $\gamma_h \le 1$.

Finally, we introduce an assumption assuring the validity of (8). We assume that, for any $T \in \mathcal{T}_h$, there exist points A_h^i, A_h^j, $A_h^k \in T$ (with $i, j, k \in \{1,\dots,N_h\}$) such that the number

$$\widehat{S}_T^{ijk} = (\widehat{\boldsymbol{n}}_T^i \times \widehat{\boldsymbol{n}}_T^j)\,\widehat{\boldsymbol{n}}_T^k \cdot \widehat{A}_T^k + (\widehat{\boldsymbol{n}}_T^j \times \widehat{\boldsymbol{n}}_T^k)\,\widehat{\boldsymbol{n}}_T^i \cdot \widehat{A}_T^i + (\widehat{\boldsymbol{n}}_T^k \times \widehat{\boldsymbol{n}}_T^i)\,\widehat{\boldsymbol{n}}_T^j \cdot \widehat{A}_T^j \tag{16}$$

satisfies

$$|\widehat{S}_T^{ijk}| \ge C_3 > 0\,, \tag{17}$$

where the constant C_3 is independent of T and h, the vector product $\boldsymbol{a} \times \boldsymbol{b}$ is defined as $a_1\,b_2 - b_1\,a_2$ and $\widehat{\boldsymbol{n}}_T^i \cdot \widehat{A}_T^i$ is defined as $\widehat{\boldsymbol{n}}_T^i \cdot (\widehat{A}_T^i - \boldsymbol{0})$.

Remark 1. Using the identity $(\widehat{\boldsymbol{n}}_T^i \times \widehat{\boldsymbol{n}}_T^j)\,\widehat{\boldsymbol{n}}_T^k + (\widehat{\boldsymbol{n}}_T^j \times \widehat{\boldsymbol{n}}_T^k)\,\widehat{\boldsymbol{n}}_T^i + (\widehat{\boldsymbol{n}}_T^k \times \widehat{\boldsymbol{n}}_T^i)\,\widehat{\boldsymbol{n}}_T^j = 0$, we have

$$\widehat{S}_T^{ijk} = (\widehat{\boldsymbol{n}}_T^i \times \widehat{\boldsymbol{n}}_T^j)\,\widehat{\boldsymbol{n}}_T^k \cdot (\widehat{A}_T^k - \widehat{A}_T^j) + (\widehat{\boldsymbol{n}}_T^j \times \widehat{\boldsymbol{n}}_T^k)\,\widehat{\boldsymbol{n}}_T^i \cdot (\widehat{A}_T^i - \widehat{A}_T^j)\,.$$

If, particularly, $\widehat{\boldsymbol{n}}_T^i$ and $\widehat{\boldsymbol{n}}_T^j$ are equal and orthogonal to $\widehat{\boldsymbol{n}}_T^k$, then

$$|\widehat{S}_T^{ijk}| = |\widehat{\boldsymbol{n}}_T^i \cdot (\widehat{A}_T^i - \widehat{A}_T^j)|\,,$$

which illustrates the meaning of (17).

Remark 2. If $\mathcal{T}_h$ consists of parallelograms, it is sufficient for proving the Babuška–Brezzi condition to assume

$$\widehat{\boldsymbol{n}}^\alpha \cdot \int_{\widehat{T}} (\widehat{x} - \widehat{A}^\alpha)\, \widehat{\varphi}^\alpha(\widehat{x})\, \mathrm{d}\widehat{x} = 0 \tag{18}$$

instead of (13) (*cf.* Remark 6 in Sect. 3). Functions satisfying the property (18) are easier to construct than those ones satisfying (13).

Remark 3. Let $\widehat{\varphi}^\alpha$ be given by a formula which is invariant to which vertex of $\widehat{T}$ is chosen as the origin of the coordinate system $\widehat{x}_1$, $\widehat{x}_2$ (with axes in the directions of edges of $\widehat{T}$). Let $\{\widehat{q}^i\}_{i=1}^4$ be a basis of $Q_1(\widehat{T})$ consisting of bilinear functions equal to 0 in three vertices of $\widehat{T}$ and equal to 1 in the remaining vertex. Then $\int_{\widehat{T}} \widehat{\varphi}^\alpha\, \widehat{q}^i\, \mathrm{d}\widehat{x} = \int_{\widehat{T}} \widehat{\varphi}^\alpha\, \widehat{q}^1\, \mathrm{d}\widehat{x}$ for $i = 2, 3, 4$ and since $\sum_{i=1}^4 \widehat{q}^i = 1$, we infer that $\int_{\widehat{T}} \widehat{\varphi}^\alpha\, \widehat{q}^i\, \mathrm{d}\widehat{x} = \frac{1}{4} \int_{\widehat{T}} \widehat{\varphi}^\alpha\, \mathrm{d}\widehat{x}$, $i = 1, \dots, 4$. Any $\widehat{q} \in Q_1(\widehat{T})$ can be written as $\widehat{q} = \sum_{i=1}^4 \alpha^i\, \widehat{q}^i$ and hence $\int_{\widehat{T}} \widehat{\varphi}^\alpha\, \widehat{q}\, \mathrm{d}\widehat{x} = \frac{1}{4} \sum_{i=1}^4 \alpha^i \int_{\widehat{T}} \widehat{\varphi}^\alpha\, \mathrm{d}\widehat{x} = \widehat{q}(C_{\widehat{T}}) \int_{\widehat{T}} \widehat{\varphi}^\alpha\, \mathrm{d}\widehat{x}$, where $C_{\widehat{T}} = (\frac{1}{2}, \frac{1}{2})$ is the barycentre of $\widehat{T}$. Thus, (13) holds with $\widehat{A}^\alpha = C_{\widehat{T}}$. An example of such an invariant function $\widehat{\varphi}^\alpha$ is the biquadratic function

$$\widehat{\varphi}^\alpha(\widehat{x}) = \widehat{x}_1\,(1 - \widehat{x}_1)\, \widehat{x}_2\,(1 - \widehat{x}_2)\,.$$

Remark 4. It is not necessary to construct invariant functions $\widehat{\varphi}^\alpha$ to satisfy (13). An example of a non-invariant function satisfying the relation (13) is the biquadratic function

$$\widehat{\varphi}^\alpha(\widehat{x}) = \widehat{x}_1\,(1 - \widehat{x}_1)\,(1 - \widehat{x}_2)\,(1/2 - \widehat{x}_2)\,,$$

for which $\widehat{A}^\alpha = (1/2, 0)$.

Remark 5. If A_h^i lies on an edge of some element of the triangulation $\mathcal{T}_h$, then $\boldsymbol{t}_h^i$ is a unit vector in the direction of this edge. Therefore, the derivative $\dfrac{\partial q_h}{\partial \boldsymbol{t}_h^i}(A_h^i)$ is well defined for any $q_h \in \mathrm{Q}_h$ and any $i \in \{1, \dots, N_h\}$. That is essential for our proceeding in the following section.

3. Proof of the Babuška–Brezzi condition

In this section, we prove that, under the assumptions made in Section 2, the spaces $\mathrm{V}_h \equiv \mathrm{V}_h^1 \oplus \mathrm{V}_h^2$ and Q_h satisfy the Babuška–Brezzi condition with a constant proportional to γ_h. First, in Lemmas 1 and 2, we prove the validity of (7) and (8). Then, in Lemma 3, we establish a Babuška–Brezzi condition with a 'wrong' norm of q_h and, finally, in Theorem 1, we prove the desired Babuška–Brezzi condition applying the modified Verfürth trick.

Lemma 1. *We have*

$$\int_\Omega \frac{\partial q_h}{\partial \boldsymbol{t}_h^i}\, \varphi_h^i\, \mathrm{d}x = \frac{\partial q_h}{\partial \boldsymbol{t}_h^i}(A_h^i) \int_\Omega \varphi_h^i\, \mathrm{d}x \qquad \forall\, q_h \in \mathrm{Q}_h\,,\ i \in \{1, \dots, N_h\}\,. \tag{19}$$

Proof. Consider any $T \subset P_h^i$ and $q_h \in \mathrm{Q}_h$ and set $\widehat{q}_T = q_h \circ F_T$. Since

$$(\nabla q_h)(F_T(\widehat{x})) = J_T(\widehat{x})^{-\mathsf{T}}\, \widehat{\nabla} \widehat{q}_T(\widehat{x}) \qquad \forall\, \widehat{x} \in \widehat{T}\,, \tag{20}$$

where $\widehat{\nabla} = (\partial/\partial\widehat{x}_1, \partial/\partial\widehat{x}_2)^{\mathsf{T}}$, we have

$$\int_T \varphi_h^i\, \boldsymbol{t}_h^i \cdot \nabla q_h\, \mathrm{d}x = \int_{\widehat{T}} \widehat{\varphi}^\alpha\, \boldsymbol{t}_h^i \cdot J_T^{-\mathsf{T}}\, \widehat{\nabla} \widehat{q}_T\, |\det J_T|\, \mathrm{d}\widehat{x}\,. \tag{21}$$

It is easy to verify that

$$J_T^{-1} = \frac{1}{\det J_T} \begin{pmatrix} \dfrac{\partial F_{T2}}{\partial \widehat{x}_2} & -\dfrac{\partial F_{T1}}{\partial \widehat{x}_2} \\ -\dfrac{\partial F_{T2}}{\partial \widehat{x}_1} & \dfrac{\partial F_{T1}}{\partial \widehat{x}_1} \end{pmatrix}.$$

Since the $\widehat{x}_1$–derivative of a function from $Q_1(\widehat{T})$ is a linear function of $\widehat{x}_2$ which does not depend on $\widehat{x}_1$ (and similarly for the $\widehat{x}_2$–derivative), we infer that $(\det J_T)\, J_T^{-\mathsf{T}}\, \widehat{\nabla}\widehat{q}_T \in Q_1(\widehat{T})$. Using the fact that $\det J_T \neq 0$ on $\widehat{T}$, it follows from (21) and (13) that

$$\int_T \varphi_h^i\, \boldsymbol{t}_h^i \cdot \nabla q_h \,\mathrm{d}x = \boldsymbol{t}_h^i \cdot (J_T^{-\mathsf{T}}\, \widehat{\nabla}\widehat{q}_T \,|\det J_T|)(\widehat{A}^\alpha) \int_{\widehat{T}} \widehat{\varphi}^\alpha \,\mathrm{d}\widehat{x}.$$

Applying (13) with $\widehat{q} = |\det J_T|$ and using (20), we get

$$\int_T \frac{\partial q_h}{\partial \boldsymbol{t}_h^i}\, \varphi_h^i \,\mathrm{d}x = \boldsymbol{t}_h^i \cdot (\nabla q_h)(A_h^i) \int_{\widehat{T}} \widehat{\varphi}^\alpha \,|\det J_T|\,\mathrm{d}\widehat{x} = \frac{\partial q_h}{\partial \boldsymbol{t}_h^i}(A_h^i) \int_T \varphi_h^i \,\mathrm{d}x\,.$$

□

Remark 6. If $\mathcal{T}_h$ consists of parallelograms, then $J_T = \text{const.}$ and it follows from (21) that

$$\int_T \varphi_h^i\, \boldsymbol{t}_h^i \cdot \nabla q_h \,\mathrm{d}x = \int_{\widehat{T}} \widehat{\varphi}^\alpha\, \widehat{\boldsymbol{t}}^\alpha \cdot \widehat{\nabla}\widehat{q}_T \,\mathrm{d}\widehat{x}\,|\det J_T|/|J_T\,\widehat{\boldsymbol{t}}^\alpha|\,,$$

where we assume that $\boldsymbol{t}_h^i = J_T\,\widehat{\boldsymbol{t}}^\alpha/|J_T\,\widehat{\boldsymbol{t}}^\alpha|$ (if $\boldsymbol{t}_h^i = -J_T\,\widehat{\boldsymbol{t}}^\alpha/|J_T\,\widehat{\boldsymbol{t}}^\alpha|$, we can proceed analogously). Denoting $\widehat{q}_T(\widehat{x}) = \xi_0 + \xi_1\,\widehat{x}_1 + \xi_2\,\widehat{x}_2 + \xi_3\,\widehat{x}_1\,\widehat{x}_2$ and $\widehat{A}^\alpha = (\widehat{a}_1, \widehat{a}_2)$, we have for $\widehat{x} \in \widehat{T}$

$$\frac{\partial \widehat{q}_T}{\partial \widehat{x}_1}(\widehat{x}) = \frac{\partial \widehat{q}_T}{\partial \widehat{x}_1}(\widehat{A}^\alpha) + \xi_3\,(\widehat{x}_2 - \widehat{a}_2)\,, \qquad \frac{\partial \widehat{q}_T}{\partial \widehat{x}_2}(\widehat{x}) = \frac{\partial \widehat{q}_T}{\partial \widehat{x}_2}(\widehat{A}^\alpha) + \xi_3\,(\widehat{x}_1 - \widehat{a}_1)\,.$$

Hence

$$\widehat{\boldsymbol{t}}^\alpha \cdot \widehat{\nabla}\widehat{q}_T(\widehat{x}) = \widehat{\boldsymbol{t}}^\alpha \cdot \widehat{\nabla}\widehat{q}_T(\widehat{A}^\alpha) + \xi_3\,\widehat{\boldsymbol{n}}^\alpha \cdot (\widehat{x} - \widehat{A}^\alpha)$$

and we see that, for proving (19), it suffices to assume (18) instead of (13).

Lemma 2. *There exists a constant C_4 independent of h such that, for any $T \in \mathcal{T}_h$, we have*

$$|q_h|_{1,T}^2 \le C_4\, h_T^2 \sum_{\substack{i=1,\\ A_h^i \in T}}^{N_h} \left| \frac{\partial q_h}{\partial \boldsymbol{t}_h^i}(A_h^i) \right|^2 \qquad \forall\, q_h \in \mathrm{Q}_h\,. \tag{22}$$

Proof. Let $\widehat{A}^i \in \widehat{T}$, $\widehat{\boldsymbol{n}}^i \in \mathbb{R}^2$, $|\widehat{\boldsymbol{n}}^i| = 1$, $i = 1, 2, 3$, be points and vectors satisfying (17) and let again $\widehat{\boldsymbol{t}}^i = (\widehat{n}_2^i, \widehat{n}_1^i)$, $i = 1, 2, 3$. Let us set for $\widehat{x} \in \widehat{T}$ and $i = 1, 2, 3$

$$\widehat{q}^i(\widehat{x}) = \{\widehat{\boldsymbol{n}}^{i+1}\,(\widehat{\boldsymbol{n}}^{i-1} \cdot \widehat{A}^{i-1}) - \widehat{\boldsymbol{n}}^{i-1}\,(\widehat{\boldsymbol{n}}^{i+1} \cdot \widehat{A}^{i+1})\} \cdot (-\widehat{x}_1, \widehat{x}_2) + (\widehat{\boldsymbol{n}}^{i+1} \times \widehat{\boldsymbol{n}}^{i-1})\,\widehat{x}_1\,\widehat{x}_2\,,$$

with the convention that $i-1 \equiv 3$ for $i=1$ and $i+1 \equiv 1$ for $i=3$. For any $\widehat{\boldsymbol{t}} = (\widehat{t}_1, \widehat{t}_2)$ and $\widehat{\boldsymbol{n}} = (\widehat{t}_2, \widehat{t}_1)$, we have

$$\frac{\partial \widehat{q}^i}{\partial \widehat{\boldsymbol{t}}}(\widehat{x}) = \{(\widehat{\boldsymbol{n}} \times \widehat{\boldsymbol{n}}^{i+1})(\widehat{\boldsymbol{n}}^{i-1} \cdot \widehat{A}^{i-1}) - (\widehat{\boldsymbol{n}} \times \widehat{\boldsymbol{n}}^{i-1})(\widehat{\boldsymbol{n}}^{i+1} \cdot \widehat{A}^{i+1})\} + (\widehat{\boldsymbol{n}}^{i+1} \times \widehat{\boldsymbol{n}}^{i-1})\, \widehat{\boldsymbol{n}} \cdot \widehat{x}$$

so that

$$\frac{\partial \widehat{q}^i}{\partial \widehat{\boldsymbol{t}}^j}(\widehat{A}^j) = \widehat{S}^{1\,2\,3}\, \delta_{ij}\,, \qquad i,j = 1,2,3\,,$$

where $\widehat{S}^{1\,2\,3}$ is defined by (16), $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$. Thus, for any $\widehat{q} \in Q_1(\widehat{T})$, the function

$$\widehat{p}(\widehat{x}) = \widehat{q}(0) + \frac{1}{\widehat{S}^{1\,2\,3}} \sum_{i=1}^{3} \frac{\partial \widehat{q}}{\partial \widehat{\boldsymbol{t}}^i}(\widehat{A}^i)\, \widehat{q}^i(\widehat{x})\,, \qquad \widehat{x} \in \widehat{T}\,,$$

satisfies

$$\frac{\partial \widehat{p}}{\partial \widehat{\boldsymbol{t}}^i}(\widehat{A}^i) = \frac{\partial \widehat{q}}{\partial \widehat{\boldsymbol{t}}^i}(\widehat{A}^i)\,, \qquad i = 1,2,3\,. \tag{23}$$

Let us show that $\widehat{p} = \widehat{q}$. We denote $\widehat{p}(\widehat{x}) - \widehat{q}(\widehat{x}) = \xi_1\, \widehat{x}_1 + \xi_2\, \widehat{x}_2 + \xi_3\, \widehat{x}_1\, \widehat{x}_2$. Then it follows from (23) that

$$\xi_1\, \widehat{n}_2^i + \xi_2\, \widehat{n}_1^i + \xi_3\, \widehat{A}^i \cdot \widehat{\boldsymbol{n}}^i = 0\,, \qquad i = 1,2,3\,, \tag{24}$$

which implies that

$$\xi_2\, (\widehat{\boldsymbol{n}}^i \times \widehat{\boldsymbol{n}}^3) + \xi_3\, \{(\widehat{A}^i \cdot \widehat{\boldsymbol{n}}^i)\, \widehat{n}_2^3 - (\widehat{A}^3 \cdot \widehat{\boldsymbol{n}}^3)\, \widehat{n}_2^i\} = 0\,, \qquad i = 1,2\,.$$

Subtracting the second equation multiplied by $\widehat{\boldsymbol{n}}^1 \times \widehat{\boldsymbol{n}}^3$ from the first equation multiplied by $\widehat{\boldsymbol{n}}^2 \times \widehat{\boldsymbol{n}}^3$, we infer that $\xi_3\, \widehat{n}_2^3\, \widehat{S}^{1\,2\,3} = 0$. Analogously we obtain from (24)

$$\xi_1\, (\widehat{\boldsymbol{n}}^3 \times \widehat{\boldsymbol{n}}^i) + \xi_3\, \{(\widehat{A}^i \cdot \widehat{\boldsymbol{n}}^i)\, \widehat{n}_1^3 - (\widehat{A}^3 \cdot \widehat{\boldsymbol{n}}^3)\, \widehat{n}_1^i\} = 0\,, \qquad i = 1,2\,,$$

and $\xi_3\, \widehat{n}_1^3\, \widehat{S}^{1\,2\,3} = 0$. Thus, $\xi_3 = 0$ and it follows from (24) that

$$\xi_1\, (\widehat{\boldsymbol{n}}^i \times \widehat{\boldsymbol{n}}^j) = 0\,, \qquad \xi_2\, (\widehat{\boldsymbol{n}}^i \times \widehat{\boldsymbol{n}}^j) = 0\,, \qquad i,j = 1,2,3\,,$$

which gives $\xi_1 = \xi_2 = 0$ in view of (17). Therefore, $\widehat{p} = \widehat{q}$. Since $|\widehat{x}| \leq \sqrt{2}$ for any $\widehat{x} \in \widehat{T}$, we have $\left|\frac{\partial \widehat{q}^i}{\partial \widehat{\boldsymbol{t}}}(\widehat{x})\right| \leq 3\sqrt{2}$ for any $\widehat{\boldsymbol{t}} \in \mathbb{R}^2$ with $|\widehat{\boldsymbol{t}}| = 1$. Hence, applying (17), we get

$$|\widehat{q}|_{1,\widehat{T}}^2 = |\widehat{p}|_{1,\widehat{T}}^2 \leq \frac{108}{C_3^2} \sum_{i=1}^{3} \left|\frac{\partial \widehat{q}}{\partial \widehat{\boldsymbol{t}}^i}(\widehat{A}^i)\right|^2,$$

which implies (22) in view of (12), (20) and (10). $\square$

Lemma 3. *We have*

$$\sup_{\boldsymbol{v}_h \in \mathrm{V}_h^2 \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}_h, q_h)}{|\boldsymbol{v}_h|_{1,\Omega}} \geq \frac{\gamma_h}{2\, C_1\, \sqrt{C_2\, C_4}} \sqrt{\sum_{T \in \mathcal{T}_h} h_T^2\, |q_h|_{1,T}^2} \qquad \forall\, q_h \in \mathrm{Q}_h\,. \tag{25}$$

Proof. Consider any $\boldsymbol{v}_h \in \mathrm{V}_h^2$. Then $\boldsymbol{v}_h = \sum_{i=1}^{N_h} \alpha^i \varphi_h^i \boldsymbol{t}_h^i$ for some real numbers α^i. We can assume that $\int_\Omega \varphi_h^i \,\mathrm{d}x \neq 0$ for $i = 1, \dots, N_h$ since otherwise $\gamma_h = 0$ and (25) holds. Thus, we have because of the definitions of C_2 and γ_h

$$|\boldsymbol{v}_h|_{1,\Omega}^2 = \sum_{T\in\mathcal{T}_h} |\boldsymbol{v}_h|_{1,T}^2 \le C_2 \sum_{T\in\mathcal{T}_h} \sum_{\substack{i=1,\\ A_h^i \in T}}^{N_h} |\alpha^i|^2 \, |\varphi_h^i|_{1,T}^2 \le \frac{C_2}{\gamma_h^2} \sum_{i=1}^{N_h} |\alpha^i|^2 \, \frac{\left|\int_\Omega \varphi_h^i \,\mathrm{d}x\right|^2}{|P_h^i|^2} \,. \tag{26}$$

Applying Lemma 1, we obtain for any $q_h \in \mathrm{Q}_h$

$$b(\boldsymbol{v}_h, q_h) = \int_\Omega \boldsymbol{v}_h \cdot \nabla q_h \,\mathrm{d}x = \sum_{i=1}^{N_h} \alpha^i \int_\Omega \frac{\partial q_h}{\partial \boldsymbol{t}_h^i} \varphi_h^i \,\mathrm{d}x = \sum_{i=1}^{N_h} \alpha^i \, \frac{\partial q_h}{\partial \boldsymbol{t}_h^i}(A_h^i) \int_\Omega \varphi_h^i \,\mathrm{d}x \,.$$

Choosing

$$\alpha^i = \frac{|P_h^i|^2}{\int_\Omega \varphi_h^i \,\mathrm{d}x} \, \frac{\partial q_h}{\partial \boldsymbol{t}_h^i}(A_h^i) \,,$$

it follows that

$$b(\boldsymbol{v}_h, q_h) = \sum_{i=1}^{N_h} |\alpha^i|^2 \, \frac{\left|\int_\Omega \varphi_h^i \,\mathrm{d}x\right|^2}{|P_h^i|^2} = \left[\sum_{i=1}^{N_h} |\alpha^i|^2 \, \frac{\left|\int_\Omega \varphi_h^i \,\mathrm{d}x\right|^2}{|P_h^i|^2}\right]^{\frac{1}{2}} \left[\sum_{i=1}^{N_h} |P_h^i|^2 \left|\frac{\partial q_h}{\partial \boldsymbol{t}_h^i}(A_h^i)\right|^2\right]^{\frac{1}{2}} ,$$

which implies (25) owing to (26), (9) and (22). □

Theorem 1. *There exists a constant $C_5 > 0$ independent of h such that the spaces $\mathrm{V}_h = \mathrm{V}_h^1 \oplus \mathrm{V}_h^2$ and Q_h satisfy the Babuška–Brezzi condition*

$$\sup_{\boldsymbol{v}_h \in \mathrm{V}_h \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}_h, q_h)}{|\boldsymbol{v}_h|_{1,\Omega}} \ge C_5 \, \gamma_h \, \|q_h\|_{0,\Omega} \qquad \forall \, q_h \in \mathrm{Q}_h \,. \tag{27}$$

Proof. Applying the modified Verfürth trick presented in [7], pp. 255–256, we obtain

$$\sup_{\boldsymbol{v}_h \in \mathrm{V}_h \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}_h, q_h)}{|\boldsymbol{v}_h|_{1,\Omega}} \ge C \, \|q_h\|_{0,\Omega} - \sqrt{\sum_{T\in\mathcal{T}_h} h_T^2 \, |q_h|_{1,T}^2} \qquad \forall \, q_h \in \mathrm{Q}_h \tag{28}$$

with a constant $C > 0$ independent of h and the theorem follows from Lemma 3 and the bound $\gamma_h \le 2\,C_1$. For completeness, we recall the main arguments leading to (28). Since the spaces $H_0^1(\Omega)^2$ and $L_0^2(\Omega)$ satisfy the inf–sup condition (6), we have

$$\beta \, \|q_h\|_{0,\Omega} \le \sup_{\boldsymbol{v} \in H_0^1(\Omega)^2 \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}, q_h)}{|\boldsymbol{v}|_{1,\Omega}} \qquad \forall \, q_h \in \mathrm{Q}_h \,. \tag{29}$$

Using an operator $\pi_h : H_0^1(\Omega)^2 \to \mathrm{V}_h$ satisfying

$$|\pi_h \boldsymbol{v}|_{1,\Omega} \le \widetilde{C} \, |\boldsymbol{v}|_{1,\Omega} \,, \quad \sqrt{\sum_{T\in\mathcal{T}_h} h_T^{-2} \, \|\boldsymbol{v} - \pi_h \boldsymbol{v}\|_{0,T}^2} \le \widetilde{C} \, |\boldsymbol{v}|_{1,\Omega} \qquad \forall \, \boldsymbol{v} \in H_0^1(\Omega)^2 \,,$$

where $\widetilde{C}$ is independent of h, we get for any $\boldsymbol{v} \in H_0^1(\Omega)^2$ and $q_h \in \mathrm{Q}_h$

$$b(\boldsymbol{v} - \pi_h \boldsymbol{v}, q_h) = \int_\Omega (\boldsymbol{v} - \pi_h \boldsymbol{v}) \cdot \nabla q_h \, \mathrm{d}x \leq \widetilde{C}\, |\boldsymbol{v}|_{1,\Omega} \sqrt{\sum_{T \in \mathcal{T}_h} h_T^2 \, |q_h|_{1,T}^2}\,,$$

$$b(\pi_h \boldsymbol{v}, q_h) \leq \widetilde{C}\, |\boldsymbol{v}|_{1,\Omega} \frac{|b(\pi_h \boldsymbol{v}, q_h)|}{|\pi_h \boldsymbol{v}|_{1,\Omega}} \leq \widetilde{C}\, |\boldsymbol{v}|_{1,\Omega} \sup_{\boldsymbol{v}_h \in \mathrm{V}_h \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}_h, q_h)}{|\boldsymbol{v}_h|_{1,\Omega}}\,.$$

Substituting the sum of these two relations into (29), we derive (28). □

Remark 7. Let $\int_{\widehat{T}} \widehat{\varphi}^\alpha \, \mathrm{d}\widehat{x} \geq 0$ for any $\alpha \in \mathcal{P}$. Consider any $T \in \mathcal{T}_h$ and $i \in \{1, \dots, N_h\}$ and let $\varphi_h^i|_T = \widehat{\varphi}^\alpha \circ F_T^{-1}$ for some $\alpha \in \mathcal{P}$. Then, according to (13),

$$\int_T \varphi_h^i \, \mathrm{d}x = \int_{\widehat{T}} \widehat{\varphi}^\alpha \, |\det J_T| \, \mathrm{d}\widehat{x} = |\det J_T(\widehat{A}^\alpha)| \int_{\widehat{T}} \widehat{\varphi}^\alpha \, \mathrm{d}\widehat{x}$$

and hence, in view of (10) and (9),

$$\int_\Omega \varphi_h^i \, \mathrm{d}x \geq C\, |P_h^i| \inf_{\alpha \in \mathcal{P},\, \widehat{\varphi}^\alpha \neq 0} \int_{\widehat{T}} \widehat{\varphi}^\alpha \, \mathrm{d}\widehat{x}\,, \qquad i = 1, \dots, N_h\,,$$

where $C > 0$ is a constant independent of h. Finally, applying (12), we get

$$\gamma_h \geq C \, \frac{\inf\limits_{\alpha \in \mathcal{P},\, \widehat{\varphi}^\alpha \neq 0} \int_{\widehat{T}} \widehat{\varphi}^\alpha \, \mathrm{d}\widehat{x}}{\sup\limits_{\alpha \in \mathcal{P}} |\widehat{\varphi}^\alpha|_{1,\widehat{T}}}\,,$$

where $C > 0$ is again independent of h. Thus, if $\mathcal{P}$ is finite and $\int_{\widehat{T}} \widehat{\varphi}^\alpha \, \mathrm{d}\widehat{x} > 0$ for any $\widehat{\varphi}^\alpha \neq 0$, then $\gamma_h \geq C > 0$ with C independent of h.

4. Examples of stable elements implying the stability of known elements

In this section, we derive several explicit examples of supplementary spaces V_h^2 such that the pair $\mathrm{V}_h = \mathrm{V}_h^1 \oplus \mathrm{V}_h^2$, Q_h satisfies the Babuška–Brezzi condition. All examples which will be discussed in this section can be divided into three classes. The first class is characterized by the fact that $\mathrm{V}_h^1 \oplus \mathrm{V}_h^2$ is a proper subspace of the space

$$\mathrm{V}_h^{Q_2} = \{\boldsymbol{v} \in H_0^1(\Omega)^2;\ \boldsymbol{v} \circ F_T \in Q_2(\widehat{T})^2 \ \ \forall\, T \in \mathcal{T}_h\}\,,$$

where $Q_2(\widehat{T})$ is the space of biquadratic functions defined on $\widehat{T}$. In this way, we get an alternative proof for the stability of the Q_2/Q_1–element. The second class will be constructed such that $\mathrm{V}_h^1 \oplus \mathrm{V}_h^2$ is a proper subspace of the space

$$\mathrm{V}_h^{4Q_1} = \{\boldsymbol{v} \in H_0^1(\Omega)^2;\ \boldsymbol{v} \circ F_T \in Q_1(\widehat{T})^2 \ \ \forall\, T \in \mathcal{T}_{h/2}\}\,,$$

where $\mathcal{T}_{h/2}$ is a triangulation obtained from $\mathcal{T}_h$ by dividing each quadrilateral $T \in \mathcal{T}_h$ into four quadrilaterals connecting the midpoints of opposite edges of T. This also represents a new proof for the stability of the $4\,Q_1/Q_1$–element. The basic feature of the last class is that the space V_h^2 consists of bubble functions, *i.e.*, any function from V_h^2 vanishes on all edges of the triangulation $\mathcal{T}_h$. Here we recover the stability of the Q_1–bubble/Q_1–element [17] by our general approach.

We start with the first class and introduce functions $\widehat{\varphi}^\alpha$ and points $\widehat{A}^\alpha$ satisfying the relation (13). As we have already seen in Remark 4, the biquadratic function

$$\widehat{\varphi}^1(\widehat{x}) = \widehat{x}_1\,(1-\widehat{x}_1)\,(1-\widehat{x}_2)\,(1/2-\widehat{x}_2)$$

satisfies (13) with $\widehat{A}^1 = (1/2, 0)$. Analogously we define the biquadratic functions $\widehat{\varphi}^2$, $\widehat{\varphi}^3$ and $\widehat{\varphi}^4$ satisfying (13) with $\widehat{A}^\alpha$ equal to $(1, 1/2)$, $(1/2, 1)$ and $(0, 1/2)$, respectively. Further, we define the functions

$$\widehat{\varphi}^5(\widehat{x}) = \widehat{\varphi}^6(\widehat{x}) = \widehat{x}_1\,(1-\widehat{x}_1)\,\widehat{x}_2\,(1-\widehat{x}_2)$$

which satisfy (13) with $\widehat{A}^5 = \widehat{A}^6 = (1/2, 1/2)$ (*cf.* Remark 3). Finally, we introduce the vectors $\widehat{\boldsymbol{t}}^1 = \widehat{\boldsymbol{t}}^3 = \widehat{\boldsymbol{t}}^5 = (1,0)$ and $\widehat{\boldsymbol{t}}^2 = \widehat{\boldsymbol{t}}^4 = \widehat{\boldsymbol{t}}^6 = (0,1)$. Now, we follow the lines of Section 2 and construct the functions φ_h^i and vectors $\boldsymbol{t}_h^i$ generating the space V_h^2. Each φ_h^i is constructed using either the functions $\widehat{\varphi}^1, \dots, \widehat{\varphi}^4$ or the functions $\widehat{\varphi}^5$, $\widehat{\varphi}^6$. In the former case, the point A_h^i is the midpoint of an inner edge of $\mathcal{T}_h$, $\boldsymbol{t}_h^i$ is a unit vector in the direction of this edge and $\operatorname{supp}\varphi_h^i$ consists of the two elements containing A_h^i. In the latter case, the point $A_h^i \equiv F_T(1/2, 1/2)$ lies in the interior of an element T, $\boldsymbol{t}_h^i$ is a unit vector in the direction determined by the midpoints of two opposite edges of T and $\operatorname{supp}\varphi_h^i = T$.

The last assumption for satisfying the Babuška–Brezzi condition is the relation (17). According to Remark 1, this relation is satisfied if, for any element T, there exist three points A_h^i, A_h^j, $A_h^k \in T$ such that the corresponding vectors $\widehat{\boldsymbol{t}}_T^i$, $\widehat{\boldsymbol{t}}_T^j$ and $\widehat{\boldsymbol{t}}_T^k$ are not all equal. For example, it is sufficient if, for a given element T, there exist three functions φ_h^i with $A_h^i \in \partial T$ or if there exist one function φ_h^i with $A_h^i \in \partial T$ and two functions φ_h^i with $A_h^i \in \operatorname{int} T$. Thus, using the above defined functions φ_h^i and vectors $\boldsymbol{t}_h^i$, we are able to define various spaces V_h^2 such that the spaces $\mathrm{V}_h = \mathrm{V}_h^1 \oplus \mathrm{V}_h^2$ and Q_h satisfy the Babuška–Brezzi condition (27). Moreover, in all these cases, the Babuška–Brezzi condition holds uniformly with respect to h (*cf.* Remark 7). As a simple consequence, we also see that, owing to

$$\mathrm{V}_h = \mathrm{V}_h^1 \oplus \mathrm{V}_h^2 \subset \mathrm{V}_h^{Q_2},$$

the Babuška–Brezzi condition is also satisfied for the spaces $\mathrm{V}_h^{Q_2}$, Q_h. The remarkable aspect of the new class of elements described above is that the Q_2/Q_1–element remains stable if more than one half of the basis functions from the velocity space are dropped. Particularly, the functions φ_h^i defined using $\widehat{\varphi}^5$ and $\widehat{\varphi}^6$ are needed for the validity of the Babuška–Brezzi condition only on those elements which have two or three edges on $\partial\Omega$.

We now derive the second class of elements implying the stability of the $4\,Q_1/Q_1$–element. The construction of V_h^2 is similar to the class above, however, we have to use piecewise bilinear functions instead of biquadratic functions $\widehat{\varphi}^\alpha$. First, we introduce the functions

$$\widehat{\varphi}(\widehat{x}_1) = \begin{cases} \widehat{x}_1 & \text{for } \widehat{x}_1 \in [0, \frac{1}{2}], \\ 1-\widehat{x}_1 & \text{for } \widehat{x}_1 \in [\frac{1}{2}, 1], \end{cases} \qquad \widehat{\psi}(\widehat{x}_2) = \begin{cases} 3-7\,\widehat{x}_2 & \text{for } \widehat{x}_2 \in [0, \frac{1}{2}], \\ \widehat{x}_2 - 1 & \text{for } \widehat{x}_2 \in [\frac{1}{2}, 1]. \end{cases}$$

Setting $\widehat{\varphi}^1(\widehat{x}) = \widehat{\varphi}(\widehat{x}_1)\,\widehat{\psi}(\widehat{x}_2)$, we obtain a function which is piecewise bilinear with respect to a subdivision of $\widehat{T}$ into four equal squares and which satisfies (13) with $\widehat{A}^1 = (\frac{1}{2}, 0)$. Analogously we define the functions $\widehat{\varphi}^2$, $\widehat{\varphi}^3$ and $\widehat{\varphi}^4$ satisfying (13) with the same points as in the biquadratic case. The functions $\widehat{\varphi}^5 = \widehat{\varphi}^6$ are now piecewise bilinear functions which vanish on the boundary of $\widehat{T}$ and are equal 1 in the point $(1/2, 1/2)$. According to Remark 3, the functions $\widehat{\varphi}^5$, $\widehat{\varphi}^6$ satisfy (13) with $\widehat{A}^5 = \widehat{A}^6 = (1/2, 1/2)$. Now we can proceed in the same way as in the biquadratic case and construct various spaces V_h^2 which guarantee the fulfilment of the Babuška–Brezzi condition. The assumption (17) can be satisfied as in the case of the first class. In all possible cases, we have $\mathrm{V}_h^1 \oplus \mathrm{V}_h^2 \subset \mathrm{V}_h^{4Q_1}$ and hence we particularly infer that the $4\,Q_1/Q_1$–element satisfies the Babuška–Brezzi condition with a constant independent of h. Again, the $4\,Q_1/Q_1$–element remains stable if more than one half of the basis functions from the velocity space are dropped.

As an example of the third class mentioned at the beginning of this section, we shall investigate the Q_1–bubble/Q_1–element by Mons and Rogé [17]. To describe the space V_h^2, we divide the reference element $\widehat{T}$ into the triangles $\widehat{T}_1$, $\widehat{T}_2$ having the vertices $(0,0)$, $(1,0)$, $(0,1)$ and $(1,0)$, $(1,1)$, $(0,1)$, respectively. Denoting

$$\widehat{\psi}^1(\widehat{x}) = \begin{cases} \widehat{x}_1\,\widehat{x}_2\,(1-\widehat{x}_1-\widehat{x}_2) & \text{for } \widehat{x}\in\widehat{T}_1\,, \\ 0 & \text{for } \widehat{x}\in\widehat{T}_2\,, \end{cases}$$

$$\widehat{\psi}^2(\widehat{x}) = \begin{cases} 0 & \text{for } \widehat{x}\in\widehat{T}_1\,, \\ (1-\widehat{x}_1)\,(1-\widehat{x}_2)\,(\widehat{x}_1+\widehat{x}_2-1) & \text{for } \widehat{x}\in\widehat{T}_2\,, \end{cases}$$

we have

$$\mathrm{V}_h^2 = \{\boldsymbol{v}\in H_0^1(\Omega)^2;\ \boldsymbol{v}\circ F_T\in[\mathrm{span}\{\widehat{\psi}^1,\widehat{\psi}^2\}]^2\quad \forall\, T\in\mathcal{T}_h\}\,.$$

We want to show that the stability of the Q_1–bubble/Q_1–element follows from our general theory. For this, we cannot use the functions $\widehat{\psi}^1$, $\widehat{\psi}^2$ since they do not satisfy (13) for any points $\widehat{A}^1$, $\widehat{A}^2$. Therefore, we introduce new basis functions

$$\widehat{\varphi}^1 = \mu\,\widehat{\psi}^1+\widehat{\psi}^2\,,\qquad \widehat{\varphi}^2 = \widehat{\psi}^1+\mu\,\widehat{\psi}^2\,,\qquad \mu = \frac{5+\sqrt{21}}{2}\,,$$

for which (13) holds with

$$\widehat{A}^1 = \left(\frac{21-\sqrt{21}}{42},\frac{21-\sqrt{21}}{42}\right),\qquad \widehat{A}^2 = \left(\frac{21+\sqrt{21}}{42},\frac{21+\sqrt{21}}{42}\right),$$

respectively. Further, we set $\widehat{\varphi}^3=\widehat{\varphi}^1$, $\widehat{\varphi}^4=\widehat{\varphi}^2$, $\widehat{A}^3=\widehat{A}^1$, $\widehat{A}^4=\widehat{A}^2$, $\widehat{\boldsymbol{t}}^1=\widehat{\boldsymbol{t}}^2=(1,0)$, $\widehat{\boldsymbol{t}}^3=\widehat{\boldsymbol{t}}^4=(0,1)$. Defining functions φ_h^i and vectors $\boldsymbol{t}_h^i$, $i=1,\dots,N_h$ ($N_h = 4\,\mathrm{card}\,\mathcal{T}_h$), as in Section 2, we deduce that

$$\mathrm{V}_h^2 = \mathrm{span}\{\varphi_h^i\,\boldsymbol{t}_h^i\}_{i=1}^{N_h}\,.$$

It is easy to check that all the assumptions made in Section 2 are fulfilled and hence it follows from Theorem 1 and Remark 7 that the Q_1–bubble/Q_1–element satisfies the Babuška–Brezzi condition with a constant independent of h.

5. General relationship between enlarging the velocity space and stabilizing the continuity equation

It is well known that a standard Galerkin finite element discretization of the Stokes or Navier–Stokes equations with the spaces V_h^1 and Q_h cannot be used because of failing the Babuška–Brezzi condition (6). We have already seen in Section 4 that, enlarging the velocity space by V_h^2, the stability for the spaces $\mathrm{V}_h=\mathrm{V}_h^1\oplus\mathrm{V}_h^2$, Q_h can be achieved. An alternative way for stabilizing a Galerkin finite element discretization using V_h^1, Q_h consists in adding some terms to the continuity equation

$$b(\boldsymbol{u}_h,q_h) = 0\qquad \forall\, q_h\in\mathrm{Q}_h\,.$$

Here, we shall show that this technique is in some sense equivalent to eliminating the degrees of freedom of the corresponding supplementary space V_h^2 from the discrete problem formulated for the spaces $\mathrm{V}_h = \mathrm{V}_h^1 \oplus \mathrm{V}_h^2$, Q_h. In the following, we shall confine ourselves to functions φ_h^i and vectors $\boldsymbol{t}_h^i$ satisfying

$$a(\varphi_h^i\, \boldsymbol{t}_h^i, \varphi_h^j\, \boldsymbol{t}_h^j) = 0 \qquad \forall\, i \neq j\,,\ \ i,j \in \{1,\dots,N_h\}\,. \tag{30}$$

Examples of such functions and vectors will be given in the following sections.

We start with a reduced discretization of the Stokes equations given by:

Find $\boldsymbol{u}_h^1 \in \mathrm{V}_h^1$, $\boldsymbol{u}_h^2 \in \mathrm{V}_h^2$ and $p_h \in \mathrm{Q}_h$ satisfying

$$\nu\, a(\boldsymbol{u}_h^1, \boldsymbol{v}_h^1) \qquad\qquad + b(\boldsymbol{v}_h^1, p_h) = \langle \boldsymbol{f}, \boldsymbol{v}_h^1\rangle \quad \forall\, \boldsymbol{v}_h^1 \in \mathrm{V}_h^1\,, \tag{31}$$

$$\nu\, a(\boldsymbol{u}_h^2, \boldsymbol{v}_h^2) + b(\boldsymbol{v}_h^2, p_h) = 0 \quad \forall\, \boldsymbol{v}_h^2 \in \mathrm{V}_h^2\,, \tag{32}$$

$$b(\boldsymbol{u}_h^1, q_h) + b(\boldsymbol{u}_h^2, q_h) = 0 \quad \forall\, q_h \in \mathrm{Q}_h\,. \tag{33}$$

This problem was obtained from the discretization (5) with $\mathrm{V}_h = \mathrm{V}_h^1 \oplus \mathrm{V}_h^2$ by dropping the terms $\nu\, a(\boldsymbol{u}_h^2, \boldsymbol{v}_h^1)$, $\nu\, a(\boldsymbol{u}_h^1, \boldsymbol{v}_h^2)$ and $\langle \boldsymbol{f}, \boldsymbol{v}_h^2\rangle$. In [16] it has been shown that, for $\boldsymbol{f} \in L^2(\Omega)^2$, the solution $\boldsymbol{u}_h = \boldsymbol{u}_h^1 + \boldsymbol{u}_h^2$, p_h of (31)–(33) has asymptotically the same rate of convergence as the solution of the original problem (5). Note that, in the special case when the triangulation $\mathcal{T}_h$ consists of rectangles only and the functions from V_h^2 vanish on all edges of $\mathcal{T}_h$, the terms $a(\boldsymbol{u}_h^2, \boldsymbol{v}_h^1)$ and $a(\boldsymbol{u}_h^1, \boldsymbol{v}_h^2)$ vanish identically. Owing to (30), the elimination of $\boldsymbol{u}_h^2$ by means of (32) becomes simple. Indeed, using the basis representation $\boldsymbol{u}_h^2 = \sum_{j=1}^{N_h} \alpha^j\, \varphi_h^j\, \boldsymbol{t}_h^j$ and setting $\boldsymbol{v}_h^2 = \varphi_h^i\, \boldsymbol{t}_h^i$ in (32), we get

$$\alpha^i\, \nu\, a(\varphi_h^i\, \boldsymbol{t}_h^i, \varphi_h^i\, \boldsymbol{t}_h^i) + b(\varphi_h^i\, \boldsymbol{t}_h^i, p_h) = 0\,.$$

Since $a(\varphi_h^i\, \boldsymbol{t}_h^i, \varphi_h^i\, \boldsymbol{t}_h^i) = |\varphi_h^i|_{1,\Omega}^2 \neq 0$, we can eliminate $\boldsymbol{u}_h^2$ from (31)–(33). Lemma 1 implies that

$$b(\varphi_h^i\, \boldsymbol{t}_h^i, q_h) = \frac{\partial q_h}{\partial \boldsymbol{t}_h^i}(A_h^i) \int_\Omega \varphi_h^i \,\mathrm{d}x \qquad \forall\, q_h \in \mathrm{Q}_h,\ i \in \{1,\dots,N_h\}$$

and hence we obtain a stabilized Q_1/Q_1–discretization of the Stokes equations in the form:

Find $\boldsymbol{u}_h^1 \in \mathrm{V}_h^1$ and $p_h \in \mathrm{Q}_h$ satisfying

$$\nu\, a(\boldsymbol{u}_h^1, \boldsymbol{v}_h^1) + b(\boldsymbol{v}_h^1, p_h) = \langle \boldsymbol{f}, \boldsymbol{v}_h^1\rangle \quad \forall\, \boldsymbol{v}_h^1 \in \mathrm{V}_h^1\,, \tag{34}$$

$$b(\boldsymbol{u}_h^1, q_h) - c_h(p_h, q_h) = 0 \quad \forall\, q_h \in \mathrm{Q}_h\,, \tag{35}$$

where the stabilizing term is given by

$$c_h(p_h, q_h) = \sum_{i=1}^{N_h} \frac{\partial p_h}{\partial \boldsymbol{t}_h^i}(A_h^i)\, \frac{\partial q_h}{\partial \boldsymbol{t}_h^i}(A_h^i)\, \frac{\left|\int_\Omega \varphi_h^i \,\mathrm{d}x\right|^2}{\nu\, |\varphi_h^i|_{1,\Omega}^2}\,. \tag{36}$$

It is easy to show (*cf.* [16]) that $\boldsymbol{u}_h^1$ converges to the solution $\boldsymbol{u}$ of (3) with the same rate as $\boldsymbol{u}_h = \boldsymbol{u}_h^1 + \boldsymbol{u}_h^2$. That means that any stabilized discretization of the Stokes equations having the form (34),(35) with $c_h(p_h, q_h)$ given by (36) possesses optimal approximation properties.

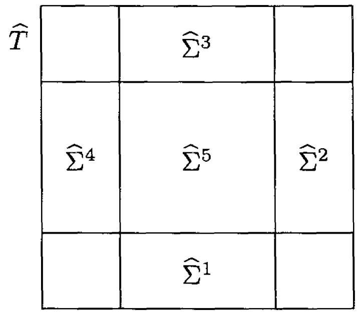

FIGURE 1. Supports of the functions $\{\widehat{\varphi}^i\}_{i=1}^5$.

Similarly as for the Stokes equations we can also proceed for the Navier–Stokes equations. We start with the reduced discretization:

Find $\boldsymbol{u}_h^1 \in \mathrm{V}_h^1$, $\boldsymbol{u}_h^2 \in \mathrm{V}_h^2$ and $p_h \in \mathrm{Q}_h$ satisfying

$$\nu\, a(\boldsymbol{u}_h^1, \boldsymbol{v}_h^1) + n(\boldsymbol{u}_h^1, \boldsymbol{u}_h^1, \boldsymbol{v}_h^1) + b(\boldsymbol{v}_h^1, p_h) = \langle \boldsymbol{f}, \boldsymbol{v}_h^1 \rangle \qquad \forall\, \boldsymbol{v}_h^1 \in \mathrm{V}_h^1\,, \tag{37}$$

$$\nu\, a(\boldsymbol{u}_h^2, \boldsymbol{v}_h^2) + b(\boldsymbol{v}_h^2, p_h) = 0 \qquad \forall\, \boldsymbol{v}_h^2 \in \mathrm{V}_h^2\,, \tag{38}$$

$$b(\boldsymbol{u}_h^1, q_h) + b(\boldsymbol{u}_h^2, q_h) = 0 \qquad \forall\, q_h \in \mathrm{Q}_h\,, \tag{39}$$

the solution of which has asymptotically, for $\boldsymbol{f} \in L^2(\Omega)^2$, the same convergence rate as the solution of the standard Galerkin finite element discretization of the Navier–Stokes equations (*cf.* [16]). Assuming (30) and eliminating $\boldsymbol{u}_h^2$ from (37)–(39), we arrive at the stabilized discretization:

Find $\boldsymbol{u}_h^1 \in \mathrm{V}_h^1$ and $p_h \in \mathrm{Q}_h$ satisfying

$$\nu\, a(\boldsymbol{u}_h^1, \boldsymbol{v}_h^1) + n(\boldsymbol{u}_h^1, \boldsymbol{u}_h^1, \boldsymbol{v}_h^1) + b(\boldsymbol{v}_h^1, p_h) = \langle \boldsymbol{f}, \boldsymbol{v}_h^1 \rangle \qquad \forall\, \boldsymbol{v}_h^1 \in \mathrm{V}_h^1\,, \tag{40}$$

$$b(\boldsymbol{u}_h^1, q_h) - c_h(p_h, q_h) = 0 \qquad \forall\, q_h \in \mathrm{Q}_h\,, \tag{41}$$

where the stabilizing term $c_h(p_h, q_h)$ is given by (36). Again, it follows that this stabilized discretization has optimal approximation properties.

Remark 8. Since the matrix $\{a(\varphi_h^i\, \boldsymbol{t}_h^i, \varphi_h^j\, \boldsymbol{t}_h^j)\}_{i,j=1}^{N_h}$ is regular, the assumption (30) is not necessary for transforming the problem (31)–(33) (resp. (37), (39)) into the form (34)–(35) (resp. (40), (41)) with some stabilizing term $c_h(p_h, q_h)$. However, the stabilizing term then generally cannot be written in a compact form like (36).

6. NEW STABILIZATION TERMS

In this section, we discuss some choices of φ_h^i and $\boldsymbol{t}_h^i$ satisfying (30) and leading to new stabilization terms. A sufficient condition for (30) is

$$(\operatorname{int}\operatorname{supp}\varphi_h^i) \cap (\operatorname{int}\operatorname{supp}\varphi_h^j) = \emptyset \quad \text{or} \quad \boldsymbol{t}_h^i \cdot \boldsymbol{t}_h^j = 0 \qquad \forall\, i \neq j\,,\ \ i, j \in \{1, \dots, N_h\}\,. \tag{42}$$

The functions φ_h^i introduced in Section 4 do not satisfy this condition but we can easily modify them so that (42) holds. We define the sets $\widehat{\Sigma}^1, \dots, \widehat{\Sigma}^5 \subset \widehat{T}$ as depicted in Figure 1 (the set $\widehat{\Sigma}^5$ is a square with the vertices $(1/4, 1/4)$, $(3/4, 1/4)$, $(3/4, 3/4)$, $(1/4, 3/4)$) and we transform the biquadratic functions $\widehat{\varphi}^1, \dots, \widehat{\varphi}^5$ from Section 4 onto $\widehat{\Sigma}^1, \dots, \widehat{\Sigma}^5$, respectively. For simplicity, we denote the transformed functions again $\widehat{\varphi}^1, \dots, \widehat{\varphi}^5$ and we set $\widehat{\varphi}^6 = \widehat{\varphi}^5$. We remark that the sets $\widehat{\Sigma}^i$ could be defined in many other ways (it is only important that their interiors are disjoint) and we could also use various other functions $\widehat{\varphi}^i$ (*e.g.*, we could transform the piecewise bilinear functions from Section 4 onto the sets $\widehat{\Sigma}^i$). However, to fix ideas, we shall now consider only the biquadratic functions defined on the sets from Figure 1. Thus, for example, the function $\widehat{\varphi}^1$ satisfies

$$\widehat{\varphi}^1(\widehat{x}) = (\widehat{x}_1 - \frac{1}{4})\,(\frac{3}{4} - \widehat{x}_1)\,(\frac{1}{4} - \widehat{x}_2)\,(\frac{1}{8} - \widehat{x}_2) \qquad \text{for } \widehat{x} \in \widehat{\Sigma}^1$$

and vanishes in $\widehat{T} \setminus \widehat{\Sigma}^1$. The points $\widehat{A}^1, \dots, \widehat{A}^6$ and the vectors $\widehat{\boldsymbol{t}}^1, \dots, \widehat{\boldsymbol{t}}^6$ remain the same as in Section 4, *i.e.*,

$$\widehat{A}^1 = (\frac{1}{2}, 0)\,, \quad \widehat{A}^2 = (1, \frac{1}{2})\,, \quad \widehat{A}^3 = (\frac{1}{2}, 1)\,, \quad \widehat{A}^4 = (0, \frac{1}{2})\,, \quad \widehat{A}^5 = \widehat{A}^6 = (\frac{1}{2}, \frac{1}{2})$$

and

$$\widehat{\boldsymbol{t}}^1 = \widehat{\boldsymbol{t}}^3 = \widehat{\boldsymbol{t}}^5 = (1,0)\,, \qquad \widehat{\boldsymbol{t}}^2 = \widehat{\boldsymbol{t}}^4 = \widehat{\boldsymbol{t}}^6 = (0,1)\,.$$

The construction of a space V_h^2 is analogous as in Section 4. Thus, using the functions $\widehat{\varphi}^1, \dots, \widehat{\varphi}^4$, we construct functions φ_h^i with points A_h^i lying on inner edges and, using the functions $\widehat{\varphi}^5$ and $\widehat{\varphi}^6$, we construct functions φ_h^i with A_h^i lying in the interiors of elements. However, using the functions $\widehat{\varphi}^1, \dots, \widehat{\varphi}^6$, we are not able to construct φ_h^i with A_h^i lying on a boundary edge. Therefore, for $i = 1, \dots, 4$, we further introduce functions $\widehat{\varphi}^{i*} \in H_0^1(\widehat{T})$ with $\operatorname{supp} \widehat{\varphi}^{i*} \subset \widehat{\Sigma}^i$ satisfying (13) with $\widehat{A}^i$. The function $\widehat{\varphi}^{1*}$ is given in $\widehat{\Sigma}^1$ by

$$\widehat{\varphi}^{1*}(\widehat{x}) = (\widehat{x}_1 - \frac{1}{4})\,(\frac{3}{4} - \widehat{x}_1)\,\widehat{x}_2\,(\frac{1}{4} - \widehat{x}_2)\,(\frac{3}{20} - \widehat{x}_2)$$

and the functions $\widehat{\varphi}^{2*}$, $\widehat{\varphi}^{3*}$, $\widehat{\varphi}^{4*}$ are defined analogously. Now, for any edge lying on $\partial\Omega$, we introduce a function φ_h^i with A_h^i belonging to this edge. For clarity, we shall not use the functions $\widehat{\varphi}^{1*}, \dots, \widehat{\varphi}^{4*}$ for constructing functions φ_h^i with $A_h^i \in \Omega$. Now, if we choose a subset of the functions $\{\varphi_h^i\}$ so that the assumption (17) is satisfied and define the space V_h^2 as the linear hull of this subset, then the Babuška–Brezzi condition (27) will hold. The question which functions φ_h^i should be chosen in order to have (17) was already discussed in Section 4. We recall that, for example, the assumption (17) is satisfied if, for any element T, we have three functions φ_h^i with A_h^i on edges of T. As we know, it is also sufficient if, for a given element, we have two functions φ_h^i with $A_h^i \in \operatorname{int} T$ and one function φ_h^i with A_h^i on an edge of T. However, there is an important difference between these two possibilities: A set of functions φ_h^i with A_h^i on edges of the triangulation satisfies the condition (42) (the supports of these φ_h^i are disjoint) whereas if the points A_h^i also lie in the interiors of elements, (42) generally does not hold. The reason is that functions φ_h^i, φ_h^j defined by $\varphi_h^i|_T = \widehat{\varphi}^5 \circ F_T^{-1}$, $\varphi_h^j|_T = \widehat{\varphi}^6 \circ F_T^{-1}$ for some element T have the same supports and the corresponding vectors $\boldsymbol{t}_h^i$, $\boldsymbol{t}_h^j$ are not orthogonal unless T is a rectangle.

We denote by $\mathcal{E}_h$ the set of all edges E of the triangulation $\mathcal{T}_h$, by C_E the midpoint of each edge E and by $\boldsymbol{t}_E$ a unit vector in the direction of E. Further, for each edge E, we have a function φ_h^i with $A_h^i = C_E$ and we denote $\varphi_E = \varphi_h^i$. Finally, for any $T \in \mathcal{T}_h$, we set $\varphi_T = \widehat{\varphi}^5 \circ F_T^{-1}$ and we denote by C_T the barycentre of T. As we know from the previous paragraph, the space $\mathrm{V}_h^2 = \operatorname{span}\{\varphi_E\, \boldsymbol{t}_E\}_{E \in \mathcal{E}_h}$ guarantees the validity of (27) with $\gamma_h \geq C > 0$. For this choice of V_h^2, the stabilizing term (36) can be rewritten into

$$c_h(p_h, q_h) = \sum_{E \in \mathcal{E}_h} \frac{\delta_E\, h_E^4}{\nu}\, \frac{\partial p_h}{\partial \boldsymbol{t}_E}(C_E)\, \frac{\partial q_h}{\partial \boldsymbol{t}_E}(C_E)\,, \tag{43}$$

where h_E denotes the length of the edge E and the parameter

$$\delta_E = \frac{\left|\int_\Omega \varphi_E \,\mathrm{d}x\right|^2}{h_E^4\, |\varphi_E|_{1,\Omega}^2}$$

is bounded from below and from above by positive constants independent of h (*cf.* Remark 7).

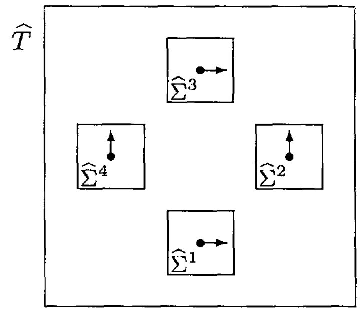

FIGURE 2. Supports of the functions $\{\widehat{\varphi}^i\}_{i=1}^4$ and directions of the vectors $\{\widehat{t}^i\}_{i=1}^4$.

If we also use the functions φ_T and the triangulation $\mathcal{T}_h$ consists of rectangles, then (36) can be rewritten into

$$c_h(p_h, q_h) = \sum_{E\in\mathcal{E}_h} \frac{\delta_E\, h_E^4}{\nu} \frac{\partial p_h}{\partial \boldsymbol{t}_E}(C_E)\, \frac{\partial q_h}{\partial \boldsymbol{t}_E}(C_E) + \sum_{T\in\mathcal{T}_h} \frac{\delta_T\, h_T^4}{\nu} \nabla p_h(C_T)\cdot\nabla q_h(C_T)\,, \tag{44}$$

where δ_E is the same as above and

$$\delta_T = \frac{\left|\int_\Omega \varphi_T \,\mathrm{d}x\right|^2}{h_T^4\, |\varphi_T|_{1,\Omega}^2}\,.$$

Again, δ_T is bounded from below and from above by positive constants independent of h. We recall that, in the sums of (43) and (44), it is sufficient to consider only those terms which assure that the assumption (17) is satisfied. For instance, if some $T \in \mathcal{T}_h$ is present in the second sum of (44), we need only one edge $E \subset T$ in the first sum of (44). Note also that the number of entries in each row of the matrix corresponding to (43) is equal to one plus the number of edges containing the vertex associated with the given row. Thus, for a uniform triangulation, the matrix corresponding to (43) has only five entries in each row like the usual five point star for the discretization of the Laplacian. The matrix corresponding to (44) has typically nine entries per row.

We have seen in Section 5 that the stabilized finite element discretizations (34), (35) and (40), (41) of the Stokes and Navier–Stokes equations, respectively, possess optimal approximation properties provided that the stabilizing term $c_h(p_h, q_h)$ can be written in the form (36). The results of this section show that, particularly, the mentioned discretizations have optimal approximation properties if $c_h(p_h, q_h)$ is defined by (43) or (44) with some suitable parameters δ_E and δ_T, which is a new result.

7. Recovering of the Brezzi–Pitkäranta stabilization

Eliminating the space V_h^2 from a discretization, we can not only derive new stabilized discretizations as in the previous section, but we can also obtain some existing ones. That often provides a deeper insight into their behaviour. Here, we show that, choosing the functions φ_h^i and vectors $\boldsymbol{t}_h^i$ in (36) in a suitable way, we can recover a stabilization introduced and studied by Brezzi and Pitkäranta [9]. For this, we introduce functions $\widehat{\varphi}^1, \dots, \widehat{\varphi}^4 \in H_0^1(\widehat{T})$ having their supports in the sets $\widehat{\Sigma}^1, \dots, \widehat{\Sigma}^4$ depicted in Figure 2. The sets $\widehat{\Sigma}^i$ are squares with side length $\frac{3-\sqrt{3}}{6}$ and barycentres in the points $\widehat{A}^1 = (\frac{1}{2}, \frac{3-\sqrt{3}}{6})$, $\widehat{A}^2 = (\frac{3+\sqrt{3}}{6}, \frac{1}{2})$, $\widehat{A}^3 = (\frac{1}{2}, \frac{3+\sqrt{3}}{6})$ and

$\widehat{A}^4 = (\frac{3-\sqrt{3}}{6}, \frac{1}{2})$, respectively. The choice of the points $\widehat{A}^i$ assures that

$$\int_{\widehat{T}} \widehat{g}\,\mathrm{d}\widehat{x} = \frac{1}{2}\,[\widehat{g}(\widehat{A}^2) + \widehat{g}(\widehat{A}^4)] \quad \text{for} \quad \widehat{g}(\widehat{x}) = (\xi_0 + \xi_1\,\widehat{x}_1)\,\widehat{q}(\widehat{x})\,, \tag{45}$$

$$\int_{\widehat{T}} \widehat{h}\,\mathrm{d}\widehat{x} = \frac{1}{2}\,[\widehat{h}(\widehat{A}^1) + \widehat{h}(\widehat{A}^3)] \quad \text{for} \quad \widehat{h}(\widehat{x}) = (\xi_0 + \xi_1\,\widehat{x}_2)\,\widehat{q}(\widehat{x})\,, \tag{46}$$

where $\widehat{q} \in Q_1(\widehat{T})$ and ξ_0, ξ_1 are arbitrary real numbers. Each function $\widehat{\varphi}^i$ is biquadratic in $\widehat{\Sigma}^i$, vanishes on the boundary of $\widehat{\Sigma}^i$ and is equal to 1 in $\widehat{A}^i$. The corresponding vectors $\widehat{\boldsymbol{t}}^i$ are depicted in Figure 2 and are defined by $\widehat{\boldsymbol{t}}^1 = \widehat{\boldsymbol{t}}^3 = (1,0)$ and $\widehat{\boldsymbol{t}}^2 = \widehat{\boldsymbol{t}}^4 = (0,1)$. Now, for any $T \in \mathcal{T}_h$ and $i \in \{1,\dots,4\}$, we set

$$\varphi_T^i|_T = \widehat{\varphi}^i \circ F_T^{-1}\,, \quad \varphi_T^i|_{\Omega\setminus T} = 0\,, \quad A_T^i = F_T(\widehat{A}^i)\,, \quad \boldsymbol{t}_T^i = J_T(\widehat{A}^i)\,\widehat{\boldsymbol{t}}^i / |J_T(\widehat{A}^i)\,\widehat{\boldsymbol{t}}^i|\,.$$

The functions φ_T^i and vectors $\boldsymbol{t}_T^i$ satisfy all the assumptions made in Section 2 and the space $\mathrm{V}_h^2 = \mathrm{span}\{\varphi_T^i\,\boldsymbol{t}_T^i\}_{T\in\mathcal{T}_h, i=1,\dots,4}$ guarantees the validity of the Babuška–Brezzi condition for the spaces $\mathrm{V}_h^1 \oplus \mathrm{V}_h^2$ and Q_h. The condition (42) is clearly satisfied and therefore, the discretizations (34), (35) and (40), (41) are stable for the stabilizing term

$$c_h(p_h, q_h) = \sum_{T\in\mathcal{T}_h} \sum_{i=1}^{4} \frac{\partial p_h}{\partial \boldsymbol{t}_T^i}(A_T^i)\,\frac{\partial q_h}{\partial \boldsymbol{t}_T^i}(A_T^i)\,\frac{\left|\int_\Omega \varphi_T^i\,\mathrm{d}x\right|^2}{\nu\,|\varphi_T^i|_{1,\Omega}^2}\,.$$

Denoting $\widehat{p}_T = p_h \circ F_T$, $\widehat{q}_T = q_h \circ F_T$ for any $T \in \mathcal{T}_h$, it follows using (13) and (20) that

$$c_h(p_h, q_h) = \sum_{T\in\mathcal{T}_h} \sum_{i=1}^{4} \frac{\partial \widehat{p}_T}{\partial \widehat{\boldsymbol{t}}^i}(\widehat{A}^i)\,\frac{\partial \widehat{q}_T}{\partial \widehat{\boldsymbol{t}}^i}(\widehat{A}^i)\,\frac{|\det J_T(\widehat{A}^i)|^2}{|J_T(\widehat{A}^i)\,\widehat{\boldsymbol{t}}^i|^2}\,\frac{|\int_{\widehat{T}} \widehat{\varphi}^i\,\mathrm{d}\widehat{x}|^2}{\nu\,\int_{\widehat{T}} |J_T^{-\mathsf{T}}\,\widehat{\nabla}\widehat{\varphi}^i|^2\,|\det J_T|\,\mathrm{d}\widehat{x}}\,.$$

If the triangulation $\mathcal{T}_h$ consists of parallelograms, then J_T is constant for any $T \in \mathcal{T}_h$ and we obtain

$$\begin{aligned} c_h(p_h, q_h) = \sum_{T\in\mathcal{T}_h} \Bigg[&\left(\frac{\partial \widehat{p}_T}{\partial \widehat{x}_1}(\widehat{A}^1)\,\frac{\partial \widehat{q}_T}{\partial \widehat{x}_1}(\widehat{A}^1) + \frac{\partial \widehat{p}_T}{\partial \widehat{x}_1}(\widehat{A}^3)\,\frac{\partial \widehat{q}_T}{\partial \widehat{x}_1}(\widehat{A}^3) \right) \frac{1}{|J_T\,\widehat{\boldsymbol{t}}^1|^2} \\ &+ \left(\frac{\partial \widehat{p}_T}{\partial \widehat{x}_2}(\widehat{A}^2)\,\frac{\partial \widehat{q}_T}{\partial \widehat{x}_2}(\widehat{A}^2) + \frac{\partial \widehat{p}_T}{\partial \widehat{x}_2}(\widehat{A}^4)\,\frac{\partial \widehat{q}_T}{\partial \widehat{x}_2}(\widehat{A}^4) \right) \frac{1}{|J_T\,\widehat{\boldsymbol{t}}^2|^2} \Bigg] \frac{|\det J_T|\,|\int_{\widehat{T}} \widehat{\varphi}^1\,\mathrm{d}\widehat{x}|^2}{\nu\,\int_{\widehat{T}} |J_T^{-\mathsf{T}}\,\widehat{\nabla}\widehat{\varphi}^1|^2\,\mathrm{d}\widehat{x}}\,. \end{aligned}$$

Since $\partial\widehat{p}_T/\partial\widehat{x}_k$ is a linear function independent of $\widehat{x}_k$, $k = 1,2$, we infer applying (45) and (46) that the terms in the square brackets are equal to

$$2\int_{\widehat{T}} \frac{\partial \widehat{p}_T}{\partial \widehat{x}_1}\,\frac{\partial \widehat{q}_T}{\partial \widehat{x}_1}\,\frac{1}{|J_T\,\widehat{\boldsymbol{t}}^1|^2} + \frac{\partial \widehat{p}_T}{\partial \widehat{x}_2}\,\frac{\partial \widehat{q}_T}{\partial \widehat{x}_2}\,\frac{1}{|J_T\,\widehat{\boldsymbol{t}}^2|^2}\,\mathrm{d}\widehat{x}\,.$$

Applying (20), we obtain

$$c_h(p_h, q_h) = \sum_{T\in\mathcal{T}_h} \frac{\gamma_T\,h_T^2}{\nu} \int_T \frac{\partial p_h}{\partial \boldsymbol{t}_T^1}\,\frac{\partial q_h}{\partial \boldsymbol{t}_T^1} + \frac{\partial p_h}{\partial \boldsymbol{t}_T^2}\,\frac{\partial q_h}{\partial \boldsymbol{t}_T^2}\,\mathrm{d}x\,,$$

where the parameter

$$\gamma_T = \frac{2\,|\int_\Omega \varphi_T^1\,\mathrm{d}x|^2}{h_T^2\,|T|\,|\varphi_T^1|_{1,\Omega}^2}$$

is bounded from below and from above by positive constants independent of h. The vectors t_T^1, t_T^2 are unit vectors in the directions of the edges of the parallelogram T. Thus, if the triangulation $\mathcal{T}_h$ consists of rectangles, we obtain

$$c_h(p_h, q_h) = \sum_{T\in\mathcal{T}_h} \frac{\gamma_T\, h_T^2}{\nu} \int_T \nabla p_h \cdot \nabla q_h \,\mathrm{d}x\,,$$

which is the stabilization introduced by Brezzi and Pitkäranta [9] for stabilizing a discretization of the Stokes equations.

8. Consistent stabilized discretizations of the Stokes equations

A drawback of the stabilizations discussed up to now is that they are not consistent. First of all, the consistency error comes from the dropped right–hand side in (32), resp. in (38). Let us consider the Stokes equations (the Navier–Stokes equations will be treated in the next section) and let us replace the equation (32) in the reduced discretization (31)–(33) by

$$\nu\, a(\boldsymbol{u}_h^2, \boldsymbol{v}_h^2) + b(\boldsymbol{v}_h^2, p_h) = \langle \boldsymbol{f}, \boldsymbol{v}_h^2\rangle \qquad \forall\, \boldsymbol{v}_h^2 \in \mathrm{V}_h^2\,. \tag{47}$$

The resulting discrete problem (31), (47), (33) has a solution which converges to the solution of (3) with the same rate as the solution of (31)–(33) (cf. [16]). If the triangulation $\mathcal{T}_h$ consists of rectangles, the equation (47) can be written as $\nu\, a(\boldsymbol{u}_h^2, \boldsymbol{v}_h^2) = r_h(\boldsymbol{u}_h^1, p_h, \boldsymbol{v}_h^2)$ with

$$r_h(\boldsymbol{w}, q, \boldsymbol{v}) = \langle \boldsymbol{f}, \boldsymbol{v}\rangle - b(\boldsymbol{v}, q) + \nu \sum_{T\in\mathcal{T}_h} \int_T \boldsymbol{v}\cdot\Delta\,\boldsymbol{w}\,\mathrm{d}x\,.$$

A solution of (3) with $\boldsymbol{u} \in H^2(\Omega)^2$ satisfies $r_h(\boldsymbol{u}, p, \boldsymbol{v}) = 0$ for any $\boldsymbol{v} \in H_0^1(\Omega)^2$ and hence $\boldsymbol{u}_h^1 = \boldsymbol{u}$, $p_h = p$ solves the new discrete problem (31), (47), (33). In this sense, the new discrete problem is consistent. If the elements of $\mathcal{T}_h$ are not rectangular, the discretization is not consistent any more, but if they are nearly rectangular, we can hope that the consistency error is small.

Similarly as at the beginning of this section, we can eliminate $\boldsymbol{u}_h^2$ from the discretization (31), (47), (33) and obtain a stabilized Q_1/Q_1–discretization of the Stokes equations. This discretization now reads:

Find $\boldsymbol{u}_h^1 \in \mathrm{V}_h^1$ and $p_h \in \mathrm{Q}_h$ satisfying

$$\begin{aligned} \nu\, a(\boldsymbol{u}_h^1, \boldsymbol{v}_h^1) + b(\boldsymbol{v}_h^1, p_h) &= \langle \boldsymbol{f}, \boldsymbol{v}_h^1\rangle & \forall\, \boldsymbol{v}_h^1 \in \mathrm{V}_h^1\,,\\ b(\boldsymbol{u}_h^1, q_h) - c_h(p_h, q_h) &= l_h(q_h) & \forall\, q_h \in \mathrm{Q}_h\,, \end{aligned}$$

where $c_h(p_h, q_h)$ is defined by (36) and

$$l_h(q_h) = -\sum_{i=1}^{N_h} \frac{\partial q_h}{\partial \boldsymbol{t}_h^i}(A_h^i)\, \langle \boldsymbol{f}, \varphi_h^i\, \boldsymbol{t}_h^i\rangle\, \frac{\int_\Omega \varphi_h^i\,\mathrm{d}x}{\nu\, |\varphi_h^i|_{1,\Omega}^2}\,.$$

The particular formulas for $l_h(q_h)$ corresponding to (43) or (44) can be introduced in a straightforward way. We only derive a formula for $l_h(q_h)$ in case of the functions φ_T^i.

Let $\mathcal{T}_h$ consists of parallelograms and let

$$\boldsymbol{f} \in \{\boldsymbol{v} \in L^2(\Omega)^2;\ \boldsymbol{v}\circ F_T \in Q_1(\widehat{T})^2\ \ \forall\, T \in \mathcal{T}_h\}\,.$$

Denoting $\widehat{\boldsymbol{f}}_T = \boldsymbol{f} \circ F_T$, $\widehat{q}_T = q_h \circ F_T$ for any $T \in \mathcal{T}_h$, we infer using (13) and (20) that

$$l_h(q_h) = -\sum_{T\in\mathcal{T}_h} \sum_{i=1}^{4} \frac{\gamma_T\, h_T^2}{2\,\nu} \frac{\partial \widehat{q}_T}{\partial \widehat{\boldsymbol{t}}^i}(\widehat{A}^i)\, \widehat{\boldsymbol{f}}_T(\widehat{A}^i) \cdot J_T \widehat{\boldsymbol{t}}^i \, \frac{|\det J_T|}{|J_T \widehat{\boldsymbol{t}}^i|^2}\,.$$

Applying (45), (46) and (20), we obtain

$$l_h(q_h) = -\sum_{T\in\mathcal{T}_h} \frac{\gamma_T\, h_T^2}{\nu} \int_T \frac{\partial q_h}{\partial \boldsymbol{t}_T^1}\, \boldsymbol{f} \cdot \boldsymbol{t}_T^1 + \frac{\partial q_h}{\partial \boldsymbol{t}_T^2}\, \boldsymbol{f} \cdot \boldsymbol{t}_T^2 \, \mathrm{d}x$$

and hence, if the triangulation $\mathcal{T}_h$ consists of rectangles, the stabilized continuity equation reads

$$b(\boldsymbol{u}_h^1, q_h) + \sum_{T\in\mathcal{T}_h} \frac{\gamma_T\, h_T^2}{\nu} \int_T (\boldsymbol{f} - \nabla p_h) \cdot \nabla q_h \, \mathrm{d}x = 0\,. \tag{48}$$

This stabilization is identical with the Petrov–Galerkin formulation of the Stokes equations introduced in [15]. Increasing the number of the bubble functions, we can derive this equation also for $\boldsymbol{f}$ being generated by higher degree polynomials defined on the reference element. That will be also seen in the next section.

9. Recovering of the SUPG method

The aim of this section is to show that, eliminating a sufficiently rich space V_h^2 from a modified reduced discretization of the Navier–Stokes equations, we can obtain the streamline upwind/Petrov–Galerkin (SUPG) method of [12] analyzed for arbitrary combinations of approximation spaces for the velocity and pressure in [21]. This equivalence will be established without linearizing the convective term, unlike other papers investigating the relationship between Galerkin methods with bubble functions and the SUPG method.

We confine ourselves to triangulations consisting of rectangles and, similarly as in Section 5, we again start with a reduced discretization of the Navier–Stokes equations. In contrast with (37)–(39), we now drop only the terms $\nu\, a(\boldsymbol{u}_h^2, \boldsymbol{v}_h^1)$, $\nu\, a(\boldsymbol{u}_h^1, \boldsymbol{v}_h^2)$, $n(\boldsymbol{u}_h^1, \boldsymbol{u}_h^2, \boldsymbol{v}_h^2)$ and $n(\boldsymbol{u}_h^2, \boldsymbol{u}_h, \boldsymbol{v}_h)$. Then we replace the term $n(\boldsymbol{u}_h^1, \boldsymbol{u}_h^2, \boldsymbol{v}_h^1)$ by the term $-n(\boldsymbol{u}_h^1, \boldsymbol{v}_h^1, \boldsymbol{u}_h^2)$. The last modification is motivated by the fact that

$$n(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}) = -n(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) - \int_\Omega (\boldsymbol{v} \cdot \boldsymbol{w})\, \mathrm{div}\, \boldsymbol{u} \, \mathrm{d}x \qquad \forall\ \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in H_0^1(\Omega)^2\,.$$

Thus, we consider the following modified discretization of the Navier–Stokes equations:

Find $\boldsymbol{u}_h^1 \in \mathrm{V}_h^1$, $\boldsymbol{u}_h^2 \in \mathrm{V}_h^2$ and $p_h \in \mathrm{Q}_h$ satisfying

$$\nu\, a(\boldsymbol{u}_h^1, \boldsymbol{v}_h^1) + n(\boldsymbol{u}_h^1, \boldsymbol{u}_h^1, \boldsymbol{v}_h^1) - n(\boldsymbol{u}_h^1, \boldsymbol{v}_h^1, \boldsymbol{u}_h^2) + b(\boldsymbol{v}_h^1, p_h) = \langle \boldsymbol{f}, \boldsymbol{v}_h^1 \rangle\,, \tag{49}$$

$$\nu\, a(\boldsymbol{u}_h^2, \boldsymbol{v}_h^2) + n(\boldsymbol{u}_h^1, \boldsymbol{u}_h^1, \boldsymbol{v}_h^2) + b(\boldsymbol{v}_h^2, p_h) = \langle \boldsymbol{f}, \boldsymbol{v}_h^2 \rangle\,, \tag{50}$$

$$b(\boldsymbol{u}_h^1, q_h) + b(\boldsymbol{u}_h^2, q_h) = 0 \tag{51}$$

for any $\boldsymbol{v}_h^1 \in \mathrm{V}_h^1$, $\boldsymbol{v}_h^2 \in \mathrm{V}_h^2$ and $q_h \in \mathrm{Q}_h$. Using the techniques of [16], it is possible to prove the same convergence results for (49)–(51) as we have for the standard Galerkin finite element discretization of the Navier–Stokes equations. Since $\Delta(\boldsymbol{u}_h^1|_T) = 0$ for any $T \in \mathcal{T}_h$ and $\boldsymbol{u}_h^2$ can be considered as a stabilization device only, the above discrete problem for $\boldsymbol{u}_h^1$, p_h is consistent (*cf.* the previous section). The term $-n(\boldsymbol{u}_h^1, \boldsymbol{v}_h^1, \boldsymbol{u}_h^2)$ introduces an influence of the space V_h^2 into the momentum balance (49). We shall see that this influence corresponds to a stabilization of the convective term $n(\boldsymbol{u}_h^1, \boldsymbol{u}_h^1, \boldsymbol{v}_h^1)$ analogous to the SUPG effect.

In view of the presence of the convective terms in (49) and (50), we shall need a more accurate integration formula than (45) and (46). Therefore, we have to increase the number of the points $\widehat{A}^i$ and hence of the bubble

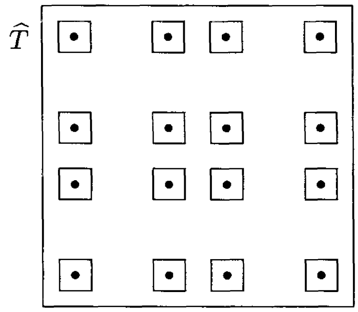

FIGURE 3. Supports of the functions $\{\widehat{\varphi}^i\}_{i=1}^{16}$.

functions $\widehat{\varphi}^i$. We introduce 16 points $\widehat{A}^i \in \widehat{T}$, $i = 1, \dots, 16$, depicted in Figure 3 whose coordinates are all possible combinations of the values $\frac{1}{2}(1+\xi)$, $\frac{1}{2}(1-\xi)$, $\frac{1}{2}(1+\eta)$, $\frac{1}{2}(1-\eta)$, where

$$\xi = \sqrt{\frac{1}{3}\left(1+\frac{2}{\sqrt{5}}\right)}, \qquad \eta = \sqrt{\frac{1}{3}\left(1-\frac{2}{\sqrt{5}}\right)}.$$

Then

$$\int_{\widehat{T}} \widehat{q}\,\mathrm{d}\widehat{x} = \frac{1}{16}\sum_{i=1}^{16} \widehat{q}(\widehat{A}^i) \qquad \forall\, \widehat{q} \in Q_5(\widehat{T}), \tag{52}$$

where $Q_5(\widehat{T})$ is the space of polynomials of degrees less than or equal to 5 in each variable. To define the functions $\widehat{\varphi}^i$ corresponding to the points $\widehat{A}^i$, we first prove the following lemma.

Lemma 4. *Let $\widehat{A} \in \widehat{T}$ and $\widehat{\psi} \in H_0^1(\widehat{T}) \setminus \{0\}$ with $\widehat{\psi} \geq 0$ on $\widehat{T}$ be given. Then there exists $\widehat{p} \in Q_2(\widehat{T})$ such that the function $\widehat{\varphi} = \widehat{p}\,\widehat{\psi}$ satisfies $\widehat{\varphi} \in H_0^1(\widehat{T}) \setminus L_0^2(\widehat{T})$ and*

$$\int_{\widehat{T}} \widehat{q}\,\widehat{\varphi}\,\mathrm{d}\widehat{x} = \widehat{q}(\widehat{A}) \int_{\widehat{T}} \widehat{\varphi}\,\mathrm{d}\widehat{x} \qquad \forall\, \widehat{q} \in Q_2(\widehat{T}). \tag{53}$$

Proof. Let us denote $((\widehat{u},\widehat{v})) = \int_{\widehat{T}} \widehat{\psi}\,\widehat{u}\,\widehat{v}\,\mathrm{d}\widehat{x}$ and $\widehat{\mathrm{M}} = \{\widehat{q} \in Q_2(\widehat{T})\,;\ \widehat{q}(\widehat{A}) = 0\}$. Then $((\cdot,\cdot))$ is an inner product on $Q_2(\widehat{T})$ and $\widehat{\mathrm{M}}$ is a linear subspace of $Q_2(\widehat{T})$ with $\dim \widehat{\mathrm{M}} = \dim Q_2(\widehat{T}) - 1$. Thus, the orthogonal complement $\widehat{\mathrm{M}}^\perp$ of $\widehat{\mathrm{M}}$ in $Q_2(\widehat{T})$ with respect to $((\cdot,\cdot))$ is a one-dimensional space and hence there exists $\widehat{p} \in \widehat{\mathrm{M}}^\perp \setminus \{0\}$ such that $((\widehat{p},\widehat{q})) = 0 \ \ \forall\, \widehat{q} \in \widehat{\mathrm{M}}$. That means that the function $\widehat{\varphi} = \widehat{p}\,\widehat{\psi}$ satisfies $\int_{\widehat{T}} \widehat{q}\,\widehat{\varphi}\,\mathrm{d}\widehat{x} = 0 \ \ \forall\, \widehat{q} \in \widehat{\mathrm{M}}$. Since $\widehat{q} - \widehat{q}(\widehat{A}) \in \widehat{\mathrm{M}}$ for any $\widehat{q} \in Q_2(\widehat{T})$, (53) holds. Let us assume that $\widehat{\varphi} \in L_0^2(\widehat{T})$. Then $\int_{\widehat{T}} \widehat{q}\,\widehat{\varphi}\,\mathrm{d}\widehat{x} = 0 \ \ \forall\, \widehat{q} \in Q_2(\widehat{T})$, *i.e.*, $((\widehat{p},\widehat{q})) = 0 \ \ \forall\, \widehat{q} \in Q_2(\widehat{T})$, which is in contradiction with $\widehat{p} \neq 0$. □

For each point $\widehat{A}^i$, $i = 1, \dots, 16$, we define a square with a side length 0.1 and a barycentre in $\widehat{A}^i$ (*cf.* Fig. 3). Transforming the function $\widehat{\varphi}$ from Lemma 4 (for $\widehat{A} = (1/2, 1/2)$ and some fixed function $\widehat{\psi}$) onto the squares around the points $\widehat{A}^i$, we obtain functions $\widehat{\varphi}^i \in H_0^1(\widehat{T})$, $i = 1, \dots, 16$, satisfying

$$\int_{\widehat{T}} \widehat{q}\,\widehat{\varphi}^i\,\mathrm{d}\widehat{x} = \widehat{q}(\widehat{A}^i) \int_{\widehat{T}} \widehat{\varphi}^i\,\mathrm{d}\widehat{x} \qquad \forall\, \widehat{q} \in Q_2(\widehat{T}),\ i \in \{1, \dots, 16\}.$$

Further, we denote $\widehat{\varphi}^{i+16} = \widehat{\varphi}^i$, $\widehat{A}^{i+16} = \widehat{A}^i$, $\widehat{\boldsymbol{t}}^i = (1,0)$ and $\widehat{\boldsymbol{t}}^{i+16} = (0,1)$, $i = 1,\dots,16$. Finally, we again set for any $T \in \mathcal{T}_h$ and $i \in \{1,\dots,32\}$

$$\varphi_T^i|_T = \widehat{\varphi}^i \circ F_T^{-1}, \quad \varphi_T^i|_{\Omega\setminus T} = 0, \quad A_T^i = F_T(\widehat{A}^i), \quad \boldsymbol{t}_T^i = J_T \widehat{\boldsymbol{t}}^i / |J_T \widehat{\boldsymbol{t}}^i|.$$

(Note that $J_T = \text{const.}$ for each $T \in \mathcal{T}_h$.) Since $\mathcal{T}_h$ consists of rectangles, we have

$$\int_T q\, \varphi_T^i \,\mathrm{d}x = q(A_T^i) \int_T \varphi_T^i \,\mathrm{d}x \qquad \forall\, q \in Q_2(T),\ T \in \mathcal{T}_h,\ i \in \{1,\dots,32\}. \tag{54}$$

The space $\mathrm{V}_h^2 = \mathrm{span}\{\varphi_T^i\, \boldsymbol{t}_T^i\}_{T\in\mathcal{T}_h, i=1,\dots,32}$ guarantees the validity of the Babuška–Brezzi condition (27) for the spaces $\mathrm{V}_h^1 \oplus \mathrm{V}_h^2$ and Q_h with $\gamma_h \geq C > 0$. Of course, much less bubble functions would be enough to get a stable element (it would be sufficient to have three functions φ_T^i in each element T) but the supplementary space $\mathrm{V}_h^2 = \mathrm{span}\{\varphi_T^i\, \boldsymbol{t}_T^i\}_{T\in\mathcal{T}_h, i=1,\dots,32}$ enables us to show a relationship to the SUPG method.

Now, let us eliminate the function

$$\boldsymbol{u}_h^2 = \sum_{T\in\mathcal{T}_h} \sum_{i=1}^{32} \alpha_T^i\, \varphi_T^i\, \boldsymbol{t}_T^i$$

from the discrete problem (49)–(51). Since the basis functions of V_h^2 are orthogonal with respect to the bilinear form $a(\cdot,\cdot)$, the elimination of $\boldsymbol{u}_h^2$ is again easy. We shall assume that

$$\boldsymbol{f} \in \{\boldsymbol{v} \in L^2(\Omega)^2;\ \boldsymbol{v}|_T \in Q_2(T)^2 \ \ \forall\, T \in \mathcal{T}_h\}$$

and we shall employ that the spaces V_h^1 and Q_h now consist of piecewise bilinear functions and that $\boldsymbol{t}_T^i \cdot \boldsymbol{t}_T^{i+16} = 0$ for any $T \in \mathcal{T}_h$ and $i \in \{1,\dots,16\}$. Applying (54) and the fact that

$$(\boldsymbol{f} - (\nabla \boldsymbol{u}_h^1)\boldsymbol{u}_h^1 - \nabla p_h)|_T \in Q_2(T)^2 \qquad \forall\, T \in \mathcal{T}_h,$$

we obtain from (50)

$$\alpha_T^i\, \nu\, |\varphi_T^i|_{1,\Omega}^2 = \boldsymbol{t}_T^i \cdot \Big(\boldsymbol{f} - (\nabla \boldsymbol{u}_h^1)\boldsymbol{u}_h^1 - \nabla p_h\Big)(A_T^i) \int_\Omega \varphi_T^i \,\mathrm{d}x.$$

Using (54) and the fact that each function φ_T^i is only a shifted function φ_T^1, we further infer that

$$n(\boldsymbol{u}_h^1, \boldsymbol{v}_h^1, \boldsymbol{u}_h^2) = \sum_{T\in\mathcal{T}_h} \sum_{i=1}^{16} \Big((\boldsymbol{f} - (\nabla \boldsymbol{u}_h^1)\boldsymbol{u}_h^1 - \nabla p_h) \cdot (\nabla \boldsymbol{v}_h^1)\boldsymbol{u}_h^1\Big)(A_T^i)\, \frac{|\int_\Omega \varphi_T^1 \,\mathrm{d}x|^2}{\nu\, |\varphi_T^1|_{1,\Omega}^2}.$$

Since the quadrature rule (52) is exact for polynomials from $Q_5(\widehat{T})$, we finally get

$$n(\boldsymbol{u}_h^1, \boldsymbol{v}_h^1, \boldsymbol{u}_h^2) = \sum_{T\in\mathcal{T}_h} \frac{\gamma_T\, h_T^2}{\nu} \int_T (\boldsymbol{f} - (\nabla \boldsymbol{u}_h^1)\boldsymbol{u}_h^1 - \nabla p_h) \cdot (\nabla \boldsymbol{v}_h^1)\boldsymbol{u}_h^1 \,\mathrm{d}x,$$

where

$$\gamma_T = \frac{16\, |\int_\Omega \varphi_T^1 \,\mathrm{d}x|^2}{h_T^2\, |T|\, |\varphi_T^1|_{1,\Omega}^2}$$

is again bounded from below and from above by positive constants independent of h. Analogously we obtain

$$b(\boldsymbol{u}_h^2, q_h) = \sum_{T\in\mathcal{T}_h} \frac{\gamma_T\, h_T^2}{\nu} \int_T (\boldsymbol{f} - (\nabla \boldsymbol{u}_h^1)\boldsymbol{u}_h^1 - \nabla p_h)\cdot \nabla q_h \,\mathrm{d}x\,.$$

Thus, the discrete problem (49)–(51) can be equivalently written in the form:

Find $\boldsymbol{u}_h^1 \in \mathrm{V}_h^1$ and $p_h \in \mathrm{Q}_h$ satisfying

$$\begin{aligned}&\nu\, a(\boldsymbol{u}_h^1, \boldsymbol{v}_h^1) + n(\boldsymbol{u}_h^1, \boldsymbol{u}_h^1, \boldsymbol{v}_h^1) + b(\boldsymbol{v}_h^1, p_h) - b(\boldsymbol{u}_h^1, q_h)\\ &\quad = \langle \boldsymbol{f}, \boldsymbol{v}_h^1\rangle + \sum_{T\in\mathcal{T}_h} \frac{\gamma_T\, h_T^2}{\nu}\int_T (\boldsymbol{f} - (\nabla \boldsymbol{u}_h^1)\boldsymbol{u}_h^1 - \nabla p_h)\cdot((\nabla \boldsymbol{v}_h^1)\boldsymbol{u}_h^1 + \nabla q_h)\,\mathrm{d}x \end{aligned} \tag{55}$$

for any $\boldsymbol{v}_h^1 \in \mathrm{V}_h^1$ and $q_h \in \mathrm{Q}_h$. This form of the discrete problem (49), (51) is identical with the SUPG method of [12, 21] in the diffusion–dominated case, *i.e.*, for small values of the element Reynolds number $\mathrm{Re}_T = |\boldsymbol{u}_h^1|_{1,\infty,T}\, h_T/\nu$. The stabilized continuity equation now reads

$$b(\boldsymbol{u}_h^1, q_h) + \sum_{T\in\mathcal{T}_h} \frac{\gamma_T\, h_T^2}{\nu} \int_T (\boldsymbol{f} - (\nabla \boldsymbol{u}_h^1)\boldsymbol{u}_h^1 - \nabla p_h)\cdot \nabla q_h \,\mathrm{d}x = 0\,,$$

which is a generalization of (48) to the nonlinear case.

In the convection–dominated case, *i.e.*, for large values of Re_T, the factor in front of the integral in (55) is usually chosen proportional to $h_T/|\boldsymbol{u}_h^1|_{1,\infty,T}$. Thus, for large values of Re_T, we should have $\gamma_T \sim 1/\mathrm{Re}_T$. That can be always fulfilled since

$$\gamma_T \le \frac{1}{625\, C^2}\,[\mathrm{diam}(\mathrm{supp}\,\widehat{\psi})]^4\,,$$

where $\widehat{\psi}$ is the function from Lemma 4 and C is the same constant as in (12). For each element T, we can use another function $\widehat{\psi}$ and obtain the correct value of γ_T (the parameter set $\mathcal{P}$ introduced in Section 2 is then generally infinite). However, if $\gamma_T \sim 1/\mathrm{Re}_T$, then the parameter γ_h from the Babuška–Brezzi condition (27) behaves like

$$\gamma_h \sim \frac{1}{\max\limits_{T\in\mathcal{T}_h} \sqrt{\mathrm{Re}_T}}\,,$$

which means that, for large values of Re_T, the SUPG method is equivalent to the problem (49)–(51) with spaces $\mathrm{V}_h = \mathrm{V}_h^1 \oplus \mathrm{V}_h^2$ and Q_h satisfying the Babuška–Brezzi condition (27) with a small parameter γ_h. Although this dependency has been not focussed in [20], a careful inspection shows that also for the P_1/P_1–element enlarged by residual–free bubbles the constant in the Babuška–Brezzi condition behaves in the convection–dominated case like $O(1/\sqrt{\mathrm{Re}})$. Therefore, it seems to be more convenient to stabilize the continuity equation and the convective term separately with different parameters γ_T. This discretization corresponds to a modified discretization of the Navier–Stokes equations with a velocity space $\mathrm{V}_h = \mathrm{V}_h^1 \oplus \mathrm{V}_h^2 \oplus \mathrm{V}_h^3$, where the supplementary space V_h^2 guarantees the fulfilment of the Babuška–Brezzi condition and V_h^3 gives additional stability in the convection–dominated case (like in the SUPG approach).

The authors would like to thank the DAAD for supporting this research (grant KZ A/98/18415). The research was also partly supported by the Czech Grant Agency under the grant No. 201/96/0313.

References

[1] D.N. Arnold, F. Brezzi and M. Fortin, A stable finite element for the Stokes equations. *Calcolo* **21** (1984) 337–344.

[2] C. Baiocchi, F. Brezzi and L.P. Franca, Virtual bubbles and Galerkin-least-squares type methods (Ga.L.S.). *Comput. Methods Appl. Mech. Eng.* **105** (1993) 125–141.

[3] R.E. Bank and B.D Welfert, A comparison between the mini-element and the Petrov-Galerkin formulations for the generalized Stokes problem. *Comput. Methods Appl. Mech. Eng.* **83** (1990) 61–68.

[4] M. Bercovier and O. Pironneau, Error estimates for finite element method solution of the Stokes problem in the primitive variables. *Numer. Math.* **33** (1979) 211–224.

[5] C. Bernardi and G. Raugel, Analysis of some finite elements for the Stokes problem. *Math. Comput.* **44** (1985) 71–79.

[6] 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. Eng.* **96** (1992) 117–129.

[7] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods.* Springer–Verlag, New York (1991).

[8] F. Brezzi, L.P. Franca and A. Russo, Further considerations on residual–free bubbles for advective–diffusive equations. Technical Report UCD/CCM 113, Univerity of Colorado at Denver, Center for Computational Mathematics (1997).

[9] F. Brezzi and J. Pitkäranta, On the stabilization of finite element approximations of the Stokes equations, in Efficient Solutions of Elliptic Systems, W. Hackbusch Ed., *Notes on Numerical Fluid Mechanics* **10** Vieweg–Verlag Braunschweig (1984) 11–19.

[10] M. Fortin, Old and new finite elements for incompressible flows. *Int. J. Numer. Methods Fluids* **1** (1981) 347–364.

[11] L.P. Franca and C. Farhat, Bubble functions prompt unusual stabilized finite element methods. *Comput. Methods Appl. Mech. Eng.* **123** (1995) 299–308.

[12] L.P. Franca and S.L. Frey, Stabilized finite element methods. II: The incompressible Navier–Stokes equations. *Comput. Methods Appl. Mech. Eng.* **99** (1992) 209–233.

[13] L.P. Franca and A. Russo, Approximation of the Stokes problem by residual–free macro bubbles, *East–West J. Numer. Math.* **4** (1996) 265–278.

[14] V. Girault and P.–A. Raviart, Finite Element Methods for Navier–Stokes Equations, Springer–Verlag, Berlin (1986).

[15] T.J. Hughes, L.P. Franca and M. Balestra, A new finite element formulation for computational fluid dynamics. V: Circumventing the Babuška–Brezzi condition: A stable Petrov–Galerkin formulation of the Stokes problem accommodating equal–order interpolations. *Comput. Methods Appl. Mech. Eng.* **59** (1986) 85–99.

[16] P. Knobloch, Reduced finite element discretizations of the Stokes and Navier–Stokes equations. *J. Math. Fluid Mechanics* (to appear).

[17] P. Mons and G. Rogé, L'élément Q_1–bulle/Q_1. *Math. Mod. Numer. Anal.* **26** (1992) 507–521.

[18] R. Pierre, Simple C^0 approximations for the computation of incompressible flows. *Comput. Methods Appl. Mech. Eng.* **68** (1988) 205–227.

[19] T.C. Rebollo, A term by term stabilization algorithm for finite element solution of incompressible flow problems, *Numer. Math.* **79** (1998) 283–319.

[20] A. Russo, Bubble stabilization of finite element methods for the linearized incompressible Navier–Stokes equations. Comput. *Methods Appl. Mech. Eng.* **132** (1996) 335–343.

[21] L. Tobiska and R. Verfürth, Analysis of a streamline diffusion finite element method for the Stokes and Navier–Stokes equations. *SIAM J. Numer. Anal.* **33** (1996) 107–127.

[22] R. Verfürth, Error estimates for a mixed finite element approximation of the Stokes equations. *RAIRO Anal. Numér.* **18** (1984) 175–182.