STABILITY OF FINITE ELEMENT MIXED INTERPOLATIONS FOR CONTACT PROBLEMS

KLAUS-JÜRGEN BATHE AND FRANCO BREZZI

Abstract. We consider the formulation of contact problems using a Lagrange multiplier to enforce the contact no-penetration constraint. The finite element discretization of the formulation must satisfy stability conditions which include an *inf-sup* condition. To identify which finite element interpolations in the contact constraint lead to stable (and optimal) numerical solutions we focus on the finite element discretization and solution of a " simple" model problem. While a simple problem to avoid the need for technicalities, the analysis of the finite element discretizations to solve the problem gives valuable insight and allows quite general conclusions on the use of different interpolation schemes.

1. Introduction

While contact problems are already being solved for some time, and many finite element programs offer contact analysis capabilities that are being used daily in production and research applications, efforts to reach more effective solution schemes are still intense [6]. One reason is the multitude of different kind of contact problems that are encountered, which can involve large relative motions, frictional forces, and static or dynamic conditions. Problems involving contact between bodies are, for example, analysed in mechanical designs of seals, in soil-structure interactions, in the analyses of bridges, in metal forming simulations, and in automobile crash and crush analyses [6].

Another reason for the continued research on contact solution procedures is simply the fact that a generally applicable, always effective, optimal and in practice easy-to-use finite element contact solution scheme is still not available.

To reach such a solution scheme, a number of requirements need to be fulfilled [4]. These include that the contact constraints can be satisfied for arbitrary geometries of the contacting bodies and for arbitrary analyis conditions, that a Jacobian of the contact constraints should be available, that an effective use without user to-be-adjusted factors should be possible, and most importantly, that certain fundamental mathematical conditions be satisfied by the contact solution algorithm.

In this paper we focus on the fundamental mathematical conditions that must be fulfilled by an effective contact solution procedure. We assume that a Lagrange multiplier mixed formulation for the solution of the contact problems is used and we study the stability of various finite element contact discretizations that can be employed.

1991 *Mathematics Subject Classification.* 65N30, 65N15, 65N50.

Key words and phrases. Contact problems, elasticity, inf-sup condition.

In an earlier contribution, we proposed a new contact solution procedure that satisfies the patch condition and the mixed formulation stability conditions [1]. This contact solution algorithm shows optimal convergence in the solution, which means that as the mesh is refined the errors *at the contact interface* diminish with the optimal rate. The contact algorithm was proposed and analysed for stability using the *numerical inf-sup* test, which is an appropriate test to perform when an analytical evaluation of the *inf-sup* condition is not available. The objective of the present paper is to present an analytical study of the stability of various contact discretizations considered in ref. [1]. The study gives very valuable insight in the solution scheme and confirms the numerical results published earlier.

While we are interested in contact solution algorithms that are general in applications, see section 2, for the purpose of the mathematical analysis carried out here we can consider a "simple model problem". We present this problem in section 3. The solution of the problem encompasses the fundamental mathematical difficulties encountered in the solution of geometrically more complex contact problems but avoids certain technicalities, the discussion of which would not add to the fundamental understanding of the stability of the solution scheme. For the same reasons, we do not discuss *regularity results* for the solution of our model problem, that in any case would be difficult to extend to more general situations. In section 4, we then introduce the solution scheme and the finite element discretizations considered and discuss abstract error norms. The importance of the *inf-sup* condition is clearly demonstrated. In section 5, we present stable finite element spaces which can be recommended for general analysis use. Finally, in section 6 we give the conclusions of this investigation.

Throughout the paper, the usual notation for Sobolev spaces and for their norms $||\cdot||$ and seminorms $|\cdot|$ is used; see for instance [17], [11].

2. The generic contact problem considered

Figure 1 shows the generic contact problem considered. We show here two flexible bodies, fixed on the boundary Γ_D , and subjected to forces that bring the bodies into contact over the area Γ_C. Only two bodies and only one contact area are shown, but the same principles discussed below are applicable when there are more bodies in contact with many contact areas. Of course, in general, the area(s) of contact are unknown and must be solved for as part of the overall solution of the problem.

Since we are focussing on the fundamental requirements for stability of the contact solution procedure we assume conditions of zero friction and small displacements. The results that we will derive will of course also be used when these conditions no longer hold.

The basic contact conditions are that

$$\lambda \geq 0, \quad g \geq 0, \quad g \cdot \lambda = 0 \tag{2.1}$$

where λ is the contact normal traction between the bodies (positive for compression), and g is the gap between the bodies. The gap is measured based on the original geometries and the displacements of the bodies.

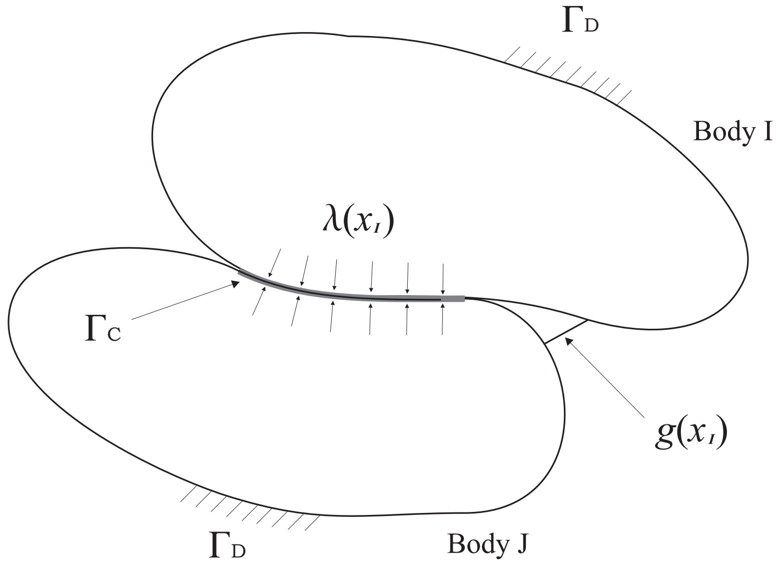

FIGURE 1. Two bodies in contact

Let x_I and x_J be the position vectors of material particles on the surfaces S_I and S_J respectively, see Fig. 2. For a given $x_I \in S_I$ let x_J^* on S_J be defined by

$$||x_I - x_J^*||_2 = \min_{x_J \in S_J} ||x_I - x_J||_2. \tag{2.2}$$

Then the gap (or gap function) between the bodies at x_I is given by

$$g(x_I) = (x_I - x_J^*) \cdot n^* \tag{2.3}$$

where n^* is the unit normal vector on S_J (outward from body J) at the material particle with position vector x_J^*.

The third condition in (2.1) is the complementary condition which stipulates that the contact force is zero if the gap is larger than zero, and vice versa.

The solution of the problem therefore requires that the conditions of equilibrium and compatibility, and the constitutive relations be fulfilled for each differential element of the bodies, subject to the boundary conditions, and that the contact conditions in (2.1) be satisfied. For a complex problem, in finite element analysis, the principle of virtual displacements is generally used with the contact conditions imposed as a constraint.

Figure 3 shows generic finite element discretizations of the bodies (shown here in two-dimensional actions). We note that as long as there is no contact the solution is obtained as in usual linear elastic finite element analysis, using for example the displacement-based finite element procedures. However, when contact is established, that is, the gap is closed anywhere along the surfaces of the bodies, an additional normal contact traction is developed along the contact area and the magnitude of the traction depends on the loading, the geometry, boundary conditions and elastic constants of the bodies.

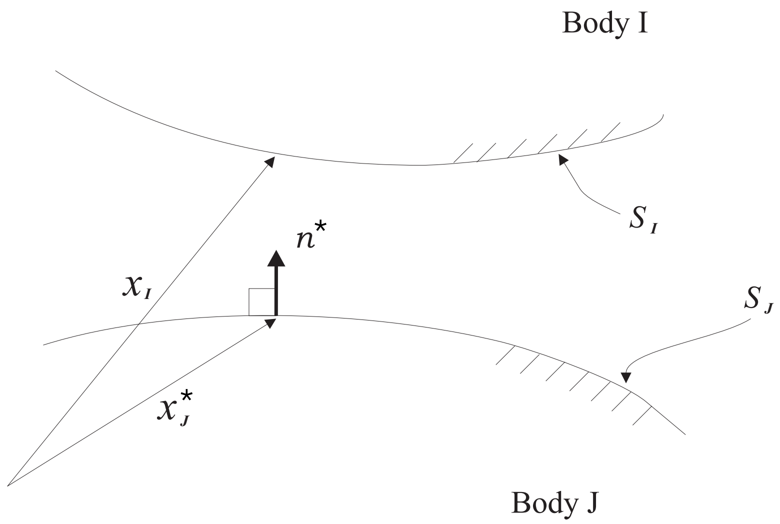

FIGURE 2. Geometry used to calculate the gap

Various finite element approaches can be used to solve the contact problem. To develop the basic principle of virtual displacements (weak formulation) subject to the contact constraints (2.1) we can proceed as follows.

Let V be the Hilbert space of displacements v of the bodies, and $K_V \subset V$ be the non-empty closed subset satisfying $g(v) \geq 0$; under reasonable geometric assumptions, K_V turns out to be convex. Let finally f be an element of V'; then the functional $J(v)$ is given by [13, 6, 4, 1, 3, 12, 16]

$$J(v) := \frac{1}{2}a(v,v) - (f,v) \tag{2.4}$$

where $a : V \times V \to \mathbb{R}$ is the bilinear form of the elasticity problem considered and the solution u is given as the minimizing argument of J over K_V, that is

$$J(u) = \inf_{v \in K_V} J(v). \tag{2.5}$$

The solution of (2.5) can also be obtained as the solution u of the variational inequality [3, 12, 16]

$$\begin{cases} \text{find } u \in K_V \text{ such that :} \\ a(u, u-v) \leq (f, u-v) \quad \forall v \in K_V. \end{cases} \tag{2.6}$$

This is the basic principle of virtual displacements, where we note that in this inequality the only variables are the displacements of the bodies. However, in practice, to reach an effective solution algorithm for complex problems, it is expedient to introduce the contact traction λ as an additional unknown for the solution of the problem. The resulting solution procedure is then a *mixed* finite element method based on the unknown displacements and contact traction (a Lagrange multiplier), which is closely related to penalty methods, perturbed Lagrangian and augmented Lagrangian techniques [6, 4]. The basic step in this

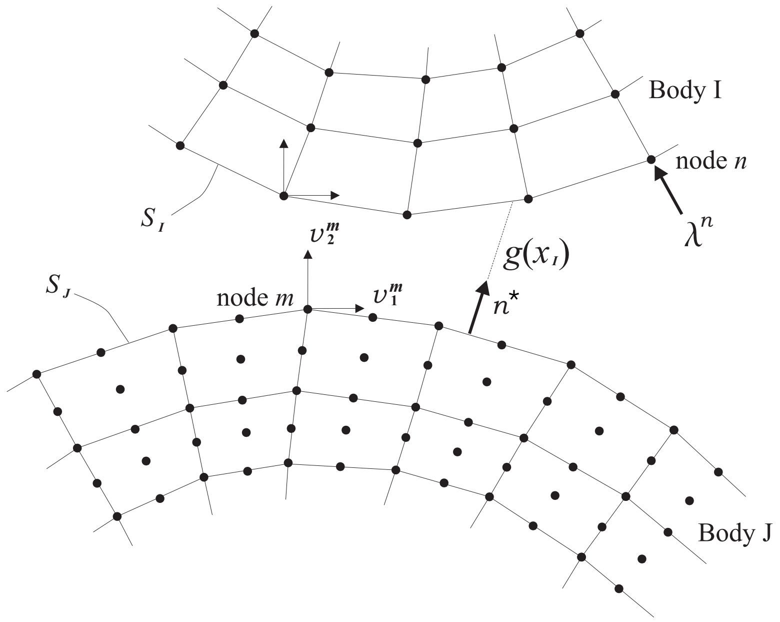

FIGURE 3. Discretizations of bodies in contact region; nodal-point displacements (and nodal Lagrange multiplier if used) are shown.

mixed finite element method is to assume the appropriate interpolation for the contact traction/Lagrange multiplier λ for a given displacement interpolation. The pair of interpolations must satisfy the stability conditions and ideally correspond to an optimal solution scheme.

The crucial stability condition to be satisfied in the selection of the interpolations is the *inf-sup* condition for the problem formulation [9, 8, 5], and we address the difficulties to satisfy this condition in the next section. We do not wish to claim that the results obtained below are all new, but present this exposition also in order to show how the mathematical analysis can be performed in a rather simple and elucidating manner.

3. The model mathematical problem

To simplify the notation and to avoid technicalities, we consider now a "simple model" problem. The discussion of this problem is, in our opinion, very valuable to clarify the difficulties related to the satisfaction of the *inf-sup* condition, and to obtain very useful results (that have quite general applicability).

Figure 4 shows the problem considered. Two adjacent pretensioned membranes are fixed on three of their edges and are free to displace into the x_3-direction (only) on the adjoining edge. The membranes are transversely loaded. Clearly, unless only a specific loading is allowed, a gap will tend to open along the common boundary Γ of the membranes. The physical requirement for the problem is that along the common boundary the transverse displacement of the top membrane must be greater than or equal to the transverse displacement of the bottom membrane. Hence, we have a contact problem.

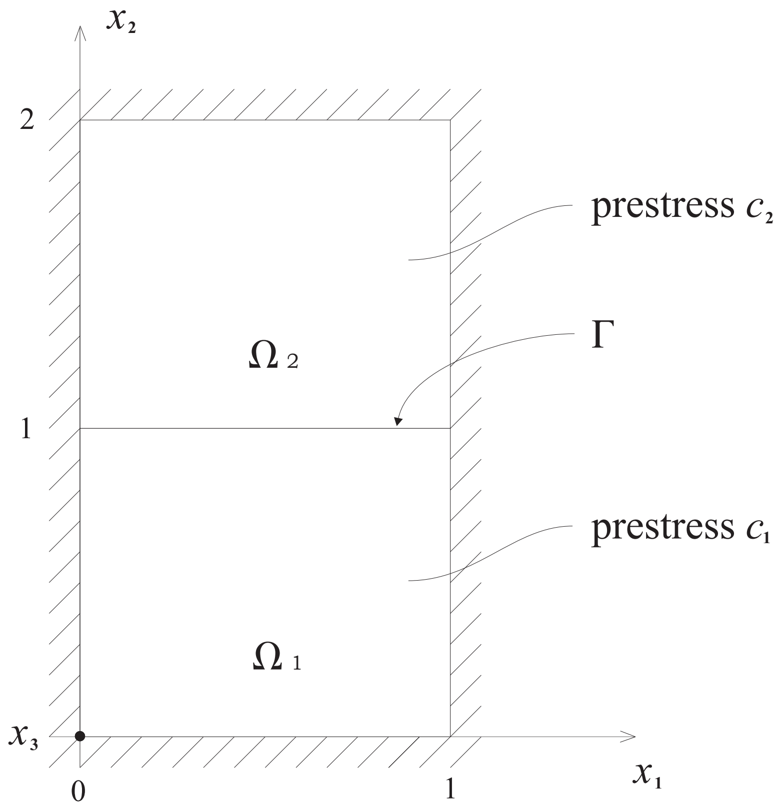

FIGURE 4. Model problem considered: two pretensioned membranes with displacement into the x_3-direction, tranverse loading into x_3-direction is $f(x_1, x_2)$.

Let us now mathematically formulate the problem considered. In Fig. 4 we consider two rectangular domains, Ω_1 and Ω_2, with $\Omega_1 =]0,1[\times]0,1[$ and $\Omega_2 =]0,1[\times]1,2[$ and denote by Γ the common part of the two boundaries, that is $\Gamma :=]0,1[\times\{1\}$. For the analysis we set, for $i = 1, 2$,

$$V_i := \{v \in H^1(\Omega_i),\ v = 0 \text{ on } \partial\Omega_i \setminus \Gamma\} \tag{3.1}$$

and, for u_i, v_i in V_i,

$$a_i(u_i, v_i) := \int_{\Omega_i} c_i \nabla u_i \cdot \nabla v_i \,\mathrm{d}x \tag{3.2}$$

where clearly $x = (x_1, x_2)$ and c_1, c_2 are positive constants (representing the prestress in the two membranes). We also set

$$V := V_1 \times V_2 \qquad \Omega :=]0,1[\times]0,2[. \tag{3.3}$$

Elements of V will be denoted by $u = (u_1, u_2)$ and $v = (v_1, v_2)$. For u and v in V we set

$$a(u, v) := a_1(u_1, v_1) + a_2(u_2, v_2) \tag{3.4}$$

and

$$g(v) := (v_2)_{|\Gamma} - (v_1)_{|\Gamma}. \tag{3.5}$$

and we consider the closed convex subset of V defined by

$$K_V := \{v \in V,\ g(v) \geq 0\}. \tag{3.6}$$

It is clear that on Γ our functions $v \in K_V$ behave like the normal components of the displacements on the contact surface of the previous section.

Our problem is now to find the minimizing argument u in K_V of (see (2.4) to (2.6))

$$J(v) := \frac{1}{2}a(v,v) - (f,v) \tag{3.7}$$

where f is a given (load) function in $L^2(\Omega)$ and $(\,.\,,\,.\,)$ denotes as usual the $L^2(\Omega)$-inner product. The solution is obtained by solving

$$\left\{\begin{array}{l} \text{find } u \in K_V \text{ such that :} \\ a(u, u-v) \leq (f, u-v) \quad \forall v \in K_V. \end{array}\right. \tag{3.8}$$

As mentioned in the previous section, our aim is to impose the condition $v \in K_V$ by means of a suitable Lagrange multiplier on Γ. For this we define the space

$$M := (H^{1/2}_{00}(\Gamma))' \tag{3.9}$$

and the convex cone

$$K_\Lambda := \{\mu \in M,\ \mu \geq 0\}. \tag{3.10}$$

We also define the continuous bilinear form b on $V \times M$

$$b(v,\mu) :=< g(v), \mu > \tag{3.11}$$

where $<\ .\,,\,.\ >$ denotes the duality pairing between $H^{1/2}_{00}(\Gamma)$ and its dual space M, and we consider the *mixed variational inequality*

$$\left\{\begin{array}{l} \text{find } (u,\lambda) \in V \times K_\Lambda \text{ such that :} \\ a(u,v) - b(v,\lambda) = (f,v) \quad \forall v \in V \\ b(u, \mu - \lambda)\ \geq\ 0 \quad \forall \mu \in K_\Lambda. \end{array}\right. \tag{3.12}$$

It is easy to check that (3.12) also has a unique solution (u,λ), where u coincides with the solution of (3.8) and

$$\lambda = c_1\big(\frac{\partial u_1}{\partial n_1}\big)_{|\Gamma} = -c_2\big(\frac{\partial u_2}{\partial n_2}\big)_{|\Gamma} = c_1\big(\frac{\partial u_1}{\partial x_2}\big)_{|\Gamma} = c_2\big(\frac{\partial u_2}{\partial x_2}\big)_{|\Gamma}. \tag{3.13}$$

The existence and uniquenes of the solution of (3.12) can be deduced, for instance, as an application of [10], or as a particular case of the more general result in [2]. We note that, in particular, by taking $\mu = 0$ and then $\mu = 2\lambda$ in the second equation of (3.12) we get

$$b(u,\lambda) = 0, \tag{3.14}$$

which will be used later on.

Remark 3.1. *We explicitly point out that the choice (3.9) for the space M of Lagrange multipliers is essential in order to have the well-posedness of (3.12). Indeed this is the choice which ensures that the continuous* inf-sup *condition holds: there exists a $\beta_c > 0$ such that*

$$\sup_{v \in V \setminus \{0\}} \frac{b(v,\mu)}{||v||_V} \geq \beta_c ||\mu||_M \quad \forall \mu \in M. \tag{3.15}$$

Other choices for M, as for instance $M = L^2(\Gamma)$ used in [15], *will not satisfy (3.15) and can result in nonoptimal estimates for the discretized problems.*

4. Discretization and abstract error estimates

If V_h and M_h are finite dimensional subspaces of V and M, respectively, and K_h a closed convex cone in M_h, we can consider the discrete counterpart of (3.12):

$$\left\{ \begin{array}{l} \text{find } (u_h, \lambda_h) \in V_h \times K_h \text{ such that :} \\ a(u_h, v_h) - b(v_h, \lambda_h) = (f, v_h) \quad \forall v_h \in V_h \\ b(u_h, \mu_h - \lambda_h) \geq 0 \quad \forall \mu_h \in K_h. \end{array} \right. \tag{4.1}$$

Existence and uniqueness of the solution of (4.1) follow rather easily, by the arguments in [10], provided we have, for all $\mu_h \in M_h$ with $\mu_h \neq 0$:

$$\sup_{v_h \in V_h \setminus \{0\}} \frac{b(v_h,\mu_h)}{||v_h||_V} > 0. \tag{4.2}$$

With classical arguments, assuming that

$$K_h \subset K_\Lambda \tag{4.3}$$

we have then, for $u_I \in V_h$ and $\lambda_I \in K_h$:

$$\begin{aligned} ||u_h - u||_V^2 &= a(u_h - u, u_h - u) \\ &= a(u_h - u, u_I - u) + a(u_h - u, u_h - u_I) \\ &= I + b(u_h - u_I, \lambda_h - \lambda) \\ &= I + b(u_h - u_I, \lambda_I - \lambda) + b(u_h - u_I, \lambda_h - \lambda_I) \\ &= I + II + b(u_h - u_I, \lambda_h - \lambda_I) \\ &\leq I + II - b(u_I, \lambda_h - \lambda_I) \\ &= I + II + b(u - u_I, \lambda_h - \lambda_I) - b(u, \lambda_h - \lambda_I) \\ &= I + II + III + b(u, \lambda_I - \lambda_h) \\ &\leq I + II + III + b(u, \lambda_I - \lambda), \end{aligned} \tag{4.4}$$

where we used the ellipticity of a, additions and subtractions, the first equation of (3.12) combined with the first equation of (4.1), and, in the fifth-to-sixth line, we used the second equation of (4.1); finally, in the last line, we used the (positive) sign of $b(u, \lambda_h)$, and (3.14).

The pieces I and II in (4.4) are then easily estimated by the Cauchy-Schwarz inequality, trace theorems, and usual interpolation estimates. However, in order to estimate III we need an estimate for $\lambda_h - \lambda_I$. For this we need a stronger form of (4.2), that is the usual *inf-sup* condition: there exists a $\beta > 0$, independent of h, such that

$$\sup_{v_h \in V_h \setminus \{0\}} \frac{b(v_h, \mu_h)}{||v_h||_V} \geq \beta\, ||\mu_h||_M \qquad \forall \mu_h \in M_h. \tag{4.5}$$

Using (4.5) we immediately obtain the existence of a $w_h \in V_h$, with $||w_h||_V = 1$, such that

$$\begin{aligned} \frac{\beta}{2}||\lambda_I - \lambda_h|| &\leq b(w_h, \lambda_I - \lambda_h) = b(w_h, \lambda_I - \lambda) + b(w_h, \lambda - \lambda_h) \\ &\leq C||\lambda_I - \lambda||_M + a(u - u_h, w_h) \leq C(||\lambda_I - \lambda||_M + ||u - u_h||_V), \end{aligned} \tag{4.6}$$

where, here and in what follows, C is a constant independent of u and h, possibly having different values at different occurrences. From (4.4) and (4.6) we have then easily

$$\begin{aligned} ||u - u_h||_V^2 &\leq C\{||u - u_h||_V\, ||u - u_I||_V + ||u_h - u_I||_V\, ||\lambda - \lambda_I||_M \\ &+||u - u_I||_V\, (||\lambda - \lambda_I||_M + ||u - u_h||_V)\} + b(u, \lambda_I - \lambda). \end{aligned} \tag{4.7}$$

Using the triangle inequality $||u_I - u_h||_V \leq ||u_I - u||_V + ||u - u_h||_V$ we then have from (4.7)

$$\begin{aligned} ||u - u_h||_V^2 \leq\ & C\{||u - u_h||_V\, (||u - u_I||_V + ||\lambda - \lambda_I||_M) \\ &+||u - u_I||_V\, ||\lambda - \lambda_I||_M\} + b(u, \lambda_I - \lambda). \end{aligned} \tag{4.8}$$

As usual, (4.8) can then be combined with (4.6) in order to have an estimate on $||\lambda - \lambda_h||_M$. For each particular choice of V_h, M_h, and K_h, the first term of (4.8) will then be estimated by usual interpolation errors, while the last term will be estimated, on a case by case basis, using the available regularity and possibly (3.14). Just to give an idea we point out that, if λ_I is chosen as the $L^2(\Gamma)$- projection of λ onto M_h, then we can define g_I^M as the $L^2(\Gamma)$-projection of $g(u)$ onto M_h and obtain

$$b(u, \lambda_I - \lambda) = \int_\Gamma (g(u) - g_I^M)(\lambda_I - \lambda)\, \mathrm{d}x_1 \tag{4.9}$$

which reduces the whole estimate (4.8) to a classical interpolation error.

It is interesting to note that a different estimate can also be derived, assuming that we can easily obtain a good approximation u_I of u in K_V. In this case, we can use, for u_I in $K_V \cap V_h$ and $\lambda_I \in K_h$, the following estimate

$$\begin{aligned} ||u_h - u||_V^2 &= a(u_h - u, u_h - u) \\ &= a(u_h - u, u_I - u) + a(u_h - u, u_h - u_I) \\ &= I + b(u_h - u_I, \lambda_h - \lambda) \\ &= I + b(u_h - u_I, \lambda_I - \lambda) + b(u_h - u_I, \lambda_h - \lambda_I) \\ &= I + II + b(u_h - u_I, \lambda_h - \lambda_I) \\ &\leq I + II - b(u_I, \lambda_h - \lambda_I) \\ &\leq I + II + b(u_I, \lambda_I) \\ &= I + II + b(u_I - u, \lambda_I) + b(u, \lambda_I - \lambda), \end{aligned} \tag{4.10}$$

(having used, to obtain the last line, the positive sign of $b(u_I, \lambda_h)$ that was not used in (4.4)). The estimate (4.10) immediately gives (using again the triangle inequality)

$$\begin{aligned} ||u-u_h||_V^2 &\leq C\{||u-u_h||_V\,(||u-u_I||_V + ||\lambda-\lambda_I||_M) \\ &+||u-u_I||_V\,||\lambda-\lambda_I||_M\} + b(u_I-u,\lambda_I) + b(u,\lambda_I-\lambda). \end{aligned} \tag{4.11}$$

It is clear however that (4.11) can only provide an estimate for the error $||u-u_h||_V$ (although the estimate for $||u-u_I||_V$ is more difficult now) but the error $||\lambda-\lambda_h||_M$ cannot be estimated without the *inf-sup* condition (4.5). Most importantly, without having at least (4.2) we cannot even ensure the uniqueness of the solution of the discrete problem (4.1).

Considering our next steps of analysis, it is not within the scope of this paper to study the error estimates, in terms of powers of h and of the regularity of u, that can be obtained from (4.8) or (4.11). Indeed, the results obtained would also not be applicable for the solution of the practical problems we have in mind, which are much more complicated than our model problem, see sections 1 and 2. Instead, we will focus on the *stability* of various possible approximations, and in particular on the *inf-sup* condition (4.5). It is quite reasonable to expect that the stability results obtained for our simple model problem will in fact hold for much more complex problems, and in particular the contact problems considered in sections 1 and 2 .

Hence in the next section we consider several possible choices for V_h and M_h, and check whether the *inf-sup* condition is satisfied.

5. Examples of stable finite element spaces

Assume now that we are given, for each $i = 1, 2$, a decomposition $\mathcal{T}_h^i$ of Ω_i. The two decompositions are not supposed to be identical on Γ. For each $i = 1, 2$ and for each integer $k \geq 1$ we consider the spaces

$$V_{hi}^k := \{v \in V_i, \quad v_{|T} \in P_k(T) \quad \forall T \in \mathcal{T}_{hi}\} \tag{5.1}$$

and then the space

$$V_h^k := V_{h1}^k \times V_{h2}^k \tag{5.2}$$

where if the superscript k is not given, any $k \geq 1$ is considered. Finally, we assume that we are given a decomposition $\mathcal{G}_h$ of Γ. For any integer $s \geq 0$ and $r = 0, 1$ we consider the space

$$\mathcal{M}_s^r := \{\mu_h \in H^r(\Gamma), \quad \mu_{h|\mathcal{I}} \in P_s \quad \forall \mathcal{I} \in \mathcal{G}_h\} \tag{5.3}$$

and the closed convex cone

$$\mathcal{K}_s^r := \{\mu_h \in \mathcal{M}_s^r, \quad \mu_h \geq 0 \text{ on } \Gamma\}. \tag{5.4}$$

In general, we use $s = 0$ or $s = 1$. For a larger s, the condition $\mu_h \geq 0$ is difficult to enforce in a finite element code. In these cases, we might just require that $\mu_h \geq 0$ at the nodes, but the abstract estimates of the previous section must then be adjusted since condition (4.3) will not hold. We also consider the space $\mathcal{M}_{0,s}^r$ and the cone $\mathcal{K}_{0,s}^r$ defined as

$$\mathcal{M}_{0,s}^r = \{\mu_h \in \mathcal{M}_s^r,\ \mu_h(0) = \mu_h(1) = 0\}, \tag{5.5}$$

$$\mathcal{K}^r_{0,s} = \{\mu_h \in \mathcal{K}^r_s,\ \mu_h(0) = \mu_h(1) = 0\}. \tag{5.6}$$

As we have seen in the last section, two inequalities are at the basis of the error estimate (4.8): the ellipticity of the bilinear form $a(u, v)$ – that for our problem is always satisfied for every choice of the discrete space V_h – and the *inf-sup* condition – that, on the contrary, will heavily depend on the choices of the spaces V_h and M_h. We are going to identify, in what follows, three families of choices for which the *inf-sup* condition is satisfied. The proof, for each family, will be based on the so-called Fortin trick [14] recalled in the following theorem.

Theorem 1 Let V and M be Hilbert spaces, and let b be a bilinear continuous form on $V \times M$ such that the continuous *inf-sup* condition (3.15) is satisfied. Assume that we are given a family of subspaces $V_h \subset V$ and $M_h \subset M$, where h is a parameter spanning the interval $]0, h_0]$. We assume that, for each h, we are given a linear operator Π_h from V to V_h with the following properties:

$$b(v - \Pi_h v, \mu_h) = 0 \quad \forall \mu_h \in M_h \tag{5.7}$$

and there exists a constant C_F, independent of h, such that

$$||\Pi_h v||_V \leq C_F\, ||v||_V \quad \forall v \in V. \tag{5.8}$$

Then the discrete *inf-sup* condition

$$\sup_{v_h \in V_h \setminus \{0\}} \frac{b(v_h, \mu_h)}{||v_h||_V} \geq \beta ||\mu_h||_M \quad \forall \mu_h \in M_h. \tag{5.9}$$

holds with $\beta = \beta_c / C_F$. $\diamond$

The proof can be found in [9, 8, 14].

Of course, (3.15) holds for our problem. We are now going to consider particular choices of spaces V_h and M_h. It is intuitively clear that, for a given decomposition $\mathcal{G}_h$ of Γ, the *worst possible scenario* is obtained when the two decompositions $\mathcal{T}_{h1}$ and $\mathcal{T}_{h2}$, restricted to Γ, coincide. This indeed minimizes the dimension of the space spanned by all possible differences between v_{h1} and v_{h2} on Γ. If our aim is to prove the *inf-sup* condition, it is clear that a sufficient condition would be to have it satisfied when the supremum in (5.9) is just taken only on the pairs $v_h = (0, v_{h2})$. On the other hand, if the two spaces $(V_{h1})_{|\Gamma}$ and $(V_{h2})_{|\Gamma}$ coincide, there is no gain in taking the supremum on all v_h's rather than just on the ones of the type $(0, v_{h2})$. Hence the condition

$$\sup_{v_h \in V_{h2} \setminus \{0\}} \frac{\int_\Gamma v_h\, \mu_h\, \mathrm{d}x_1}{||v_h||_{V_2}} \geq \beta ||\mu_h||_M \quad \forall \mu_h \in M_h. \tag{5.10}$$

is always *sufficient* for having (5.9), and becomes *necessary* when $(V_{h1})_{|\Gamma}$ and $(V_{h2})_{|\Gamma}$ coincide. In what follows, we are therefore going to see whether a given choice of the spaces V_h and M_h satisfies (5.10). The following lemma is an immediate consequence of Theorem 1 and of the above discussion.

Lemma 1 Assume that the continuous version of (5.10) holds, namely, there exists a $\beta_c > 0$ such that

$$\sup_{v \in V_2 \setminus \{0\}} \frac{\int_\Gamma v\, \mu\, \mathrm{d}x_1}{||v||_{V_2}} \geq \beta_c ||\mu||_M \quad \forall \mu \in M. \tag{5.11}$$

Assume moreover that for each $h \in]0, h_0]$ there exists a linear operator Π_h from V_2 into V_{h2}^k satisfying

$$\int_\Gamma (v - \Pi_h v)\mu_h \, \mathrm{d}x_1 = 0 \quad \forall \mu_h \in M_h, \tag{5.12}$$

and there exists a constant C_F, independent of h such that

$$||\Pi_h v||_{V_2} \leq C_F \, ||v||_{V_2} \quad \forall v \in V_2. \tag{5.13}$$

Then (5.10) holds, and therefore (4.5) also holds. $\diamond$

To avoid technicalities, we assume that the decomposition cast on Γ by $\mathcal{T}_{h2}$ *coincides* with $\mathcal{G}_h$. This is a rather particular case, but not unrealistic. Finally, always in order to simplify the exposition, we assume that the decomposition $\mathcal{T}_{h2}$ is quasi-uniform. Under this assumption it is rather easy to check (see e.g. [19]) that for every $v_h \in V_{h2}^k$ we can find a $\tilde{v}_h \in V_{h2}^k$ such that

$$\tilde{v}_h = v_h \text{ on } \Gamma, \tag{5.14}$$

and

$$||\tilde{v}_h||_{V_2} \leq C \, ||v_h||_{H^{1/2}_{00}(\Gamma)}. \tag{5.15}$$

Under the above assumptions on the decompositions we have therefore the following theorem.

Theorem 2 Let W_h be the space of the traces of V_{h2}^k on Γ, and assume that we are given, for each h, an operator π_h from $H^{1/2}_{00}(\Gamma)$ into W_h with the following properties:

$$\int_\Gamma (w - \pi_h w)\mu_h \, \mathrm{d}x_1 = 0 \quad \forall \mu_h \in M_h, \tag{5.16}$$

$$||\pi_h w||_{H^{1/2}_{00}(\Gamma)} \leq C_\Gamma \, ||w||_{H^{1/2}_{00}(\Gamma)} \quad \forall w \in H^{1/2}_{00}(\Gamma), \tag{5.17}$$

where C_Γ is a constant independent of h and v. Then an operator Π_h satisfying (5.12) and (5.13) exists (and hence the *inf-sup* condition (5.10) holds.) $\diamond$

Proof. Given $v \in V_2$ we consider $w := v_{|\Gamma}$ and $w_h := \pi_h w$. We then lift w_h, in an arbitrary way, to an element $v_h \in V_{h2}^k$ such that $v_h = w_h$ on Γ. Then we define $\Pi_h v$ as $\widetilde{v_h}$. Using properties (5.14) and (5.16) we immediately get (5.12). Then (5.17) and (5.15) easily give (5.13). $\diamond$

Remark 5.1. *It is quite intuitive that the discrete* inf-sup *condition (5.10) should depend only on the space M_h and on the space W_h of the traces of functions in V_{h2}, even if norms over Ω_2 are involved. The role of Theorem 2 is indeed to reduce the proof of (5.10) to a property (the existence of a suitable π_h) that depends only on M_h and W_h.*

In what follows we will keep $V_h = V_h^k$ fixed, and we consider three possible choices for the corresponding M_h. We notice that, with the above choice for V_h, the space of traces W_h will also be fixed, equal to

$$W_h = \{w_h \in C^0([0,1], \text{ such that } w_{h|\mathcal{I}} \in P_k \quad \forall \mathcal{I} \in \mathcal{G}_h, \text{ and } w_h(0) = w_h(1) = 0\}. \tag{5.18}$$

The first choice corresponds to having as space of multipliers *the same space* that is spanned by the traces of V_h (including the zero boundary conditions at the endpoints of Γ).

Theorem 3 Assume that $V_h := V_h^k$ and $M_h := \mathcal{M}_{0,k}^1$. Then an operator π_h satisfying (5.16) and (5.17) exists (and hence the *inf-sup* condition is satisfied.) $\diamond$

Proof. We note that, in this case, we have $W_h = M_h$. Then we can define $\pi_h w$ as the $L^2(\Gamma)$-projection of w on $W_h = M_h$. Property (5.16) is clearly verified. It is also obvious that

$$||\pi_h w||_{L^2(\Gamma)} \le ||w||_{L^2(\Gamma)} \quad \forall w \in L^2(\Gamma). \tag{5.19}$$

By usual approximation properties we also have

$$||\pi_h w - w||_{L^2(\Gamma)} \le C\, h\, ||w||_{H_0^1(\Gamma)} \quad \forall w \in H_0^1(\Gamma). \tag{5.20}$$

From the inverse inequality, the triangle inequality, and usual interpolation estimates we then have

$$\begin{aligned} ||\pi_h w - w_I||_{H_0^1(\Gamma)} &\le C\, h^{-1}\, ||\pi_h w - w_I||_{L^2(\Gamma)} \\ &\le C\, h^{-1}\, (||\pi_h w - w||_{L^2(\Gamma)} + ||w - w_I||_{L^2(\Gamma)}) \\ &\le C\, ||w||_{H_0^1(\Gamma)} \quad \forall w \in H_0^1(\Gamma), \end{aligned} \tag{5.21}$$

where w_I is the usual interpolant of w.

From (5.21) we immediately obtain

$$||\pi_h w||_{H_0^1(\Gamma)} \le ||\pi_h w - w_I||_{H_0^1(\Gamma)} + ||w_I||_{H_0^1(\Gamma)} \le C\, ||w||_{H_0^1(\Gamma)} \quad \forall w \in H_0^1(\Gamma). \tag{5.22}$$

Interpolating between (5.19) and (5.22) we obtain (5.17). $\diamond$

Remark 5.2. *It is very easy to see that, taking a space of multipliers made of continuous piecewise P_k functions that* do not *vanish at the endpoints, the* inf-sup *condition will not hold. Indeed, the number of degrees of freedom for the space M_h of multipliers would be, in the case of N intervals, equal to $N \times k + 1$, which is bigger than $N \times k - 1$, the dimension of W_h. Hence the* inf-sup *condition cannot hold.*

Before proceeding, we select a useful property that comes out immediately from the proof of Theorem 3.

Lemma 2 If $\mathcal{G}_h$ is quasi-uniform, then there exists a linear operator π_h^1 from $H_{00}^{1/2}(\Gamma)$ into the space of piecewise linear functions on $\mathcal{G}_h$, and two constants C_1 and C^1, independent of h, such that

$$||\pi_h^1 w||_{r,\Gamma} \le C_1 |||w||_{r,\Gamma} \quad r = 0, 1, \tag{5.23}$$

and

$$||\pi_h^1 w - w||_{0,\Gamma} \le C^1\, h\, ||w||_{1,\Gamma} \quad \forall w \in H_0^1(\Gamma). \qquad \diamond \tag{5.24}$$

The proof, as we said, is a byproduct of the proof of Theorem 3, by taking $k = 1$. ◇

We shall consider in a while two other possible choices of finite element spaces. For them, the construction of the operator π_h is made in two steps (see for instance [9]). The strategy is to use the operator π_h^1 of Lemma 2, and look for another operator π_h^2 with the properties

$$\int_\Gamma (\pi_h^2 w - w)\mu_h \,\mathrm{d}x_1 = 0 \quad \forall \mu_h \in M_h, \tag{5.25}$$

and

$$||\pi_h^2 w||_{t,\Gamma} \leq C_2\, h^{-t}\, ||w||_{0,\Gamma} \quad t = 0, 1, \tag{5.26}$$

and then to define π_h as

$$\pi_h := \pi_h^1 - \pi_h^2(\pi_h^1 - I). \tag{5.27}$$

It is then clear that from (5.27) and (5.25) we have

$$\begin{aligned} \textstyle\int_\Gamma (\pi_h w - w)\mu_h \,\mathrm{d}x_1 &= \textstyle\int_\Gamma (\pi_h^1 w - \pi_h^2(\pi_h^1 - I)w - w)\mu_h \,\mathrm{d}x_1 \\ &= \textstyle\int_\Gamma ((\pi_h^1 w - w) - \pi_h^2(\pi_h^1 w - w))\mu_h \,\mathrm{d}x_1 = 0, \end{aligned} \tag{5.28}$$

for all $\mu_h \in M_h$, that is (5.16); moreover, using (5.27), then (5.23) with $r = 0$, and finally (5.26) with $t = 0$ we have, for all $w \in L^2(\Gamma)$:

$$||\pi_h w||_{0,\Gamma} \leq C_1 ||w||_{0,\Gamma} + C_2\, ||\pi_h^1 w - w||_{0,\Gamma} \leq (C_1 + C_2(1 + C_1))\, ||w||_{0,\Gamma}. \tag{5.29}$$

On the other hand, using (5.27), then (5.23) with $r = 1$, then (5.26) with $t = 1$, and finally (5.24) we obtain, for all $w \in H_0^1(\Gamma)$:

$$||\pi_h w||_{1,\Gamma} \leq C_1 ||w||_{1,\Gamma} + C_2\, h^{-1}\, ||\pi_h^1 w - w||_{0,\Gamma} \leq (C_1 + C_2\, C^1)\, ||w||_{1,\Gamma}. \tag{5.30}$$

Equations (5.29) and (5.30) then easily imply, by interpolation,

$$||\pi_h w||_{H_{00}^{1/2}(\Gamma)} \leq C\, ||w||_{H_{00}^{1/2}(\Gamma)} \tag{5.31}$$

that is (5.17).

We summarize the above results in the following lemma.

Lemma 3 Let W_h be the space of traces of V_h^2, and let the decomposition $\mathcal{T}_h^2$ be quasi-uniform. Assume that we can construct an operator π_h^2, from $H_{00}^{1/2}(\Gamma)$ into W_h, with the properties (5.25) and (5.26); then the *inf-sup* condition (5.10) is satisfied (and hence (4.5) also holds). ◇

Proof. The proof follows from the above construction of π_h and Theorem 2. ◇

Remark 5.3. *The role of Lemma 3, as we shall see, is to reduce the verification of the assumptions of Theorem 2 (and hence the proof of the* inf-sup *condition) to the construction of a* local *operator π_h^2. This was not possible for the first choice of multipliers (considered in Theorem 3), but will be possible for the next two choices.*

We now continue with our study of the different choices of spaces. The second case that we consider corresponds to using a space of multipliers which are discontinuous and have local degree $k-2$, with $k \geq 2$, if k is the local degree of V_h. This is discussed in the following theorem.

Theorem 4 Assume that $V_h := V_h^k$ with $k \geq 2$, and $M_h := \mathcal{M}_{k-2}^0$. Then an operator π_h^2 satisfying (5.25) and (5.26) exists (and hence the *inf-sup* condition holds by Lemma 3). $\diamond$

Proof. To construct π_h^2 we can easily proceed with an element by element (actually, interval by interval) argument. For each $\mathcal{I} \in \mathcal{G}_h$ and for each w in, say, $L^2(\Gamma)$ we define $\pi_h^2 w$ as the polynomial of degree k (in $\mathcal{I}$), vanishing at the endpoints of $\mathcal{I}$, and satisfying

$$\int_{\mathcal{I}} (\pi_h^2 w - w) p_{k-2} \, \mathrm{d}x_1 = 0 \quad \forall p_{k-2} \in P_{k-2}. \tag{5.32}$$

It is rather easy to check that properties (5.25) and (5.26) hold true. $\diamond$

Remark 5.4. *It is easy to see that, by taking M_h to be the space of discontinuous piecewise P_{k-1} functions, the* inf-sup *condition will not hold. Indeed, as in Remark 5.2, the dimensional count gives $N \times k$ as dimension of M_h, while W_h has dimension $N \times k - 1$.*

The third case that we consider corresponds to using a space of multipliers which are continuous, do not necessarily vanish at the endpoints of Γ, and have locally one degree less than the degree used in V_h. Comparing with our first case, we see that here (for the same V_h) the space for Lagrange multipliers has, in general, a much smaller dimension. Indeed, with the same notation of Remarks 5.2 and 5.4, the dimension of M_h is $N \times k - 1$ for the first case, and $N \times (k-1) + 1$ for this last case. In view of the previous result (that the *inf-sup* condition holds for case 1) we expect case 3, reasonably, to work as well. However, the new space of multipliers is not a subspace of the previous one, and an independent proof is therefore necessary. Consider the following theorem.

Theorem 5 Assume that $V_h := V_h^k$ with $k \geq 2$, and $M_h := \mathcal{M}_{k-1}^1$. Then an operator π_h^2 satisfying (5.25) and (5.26) exists (and hence the *inf-sup* condition holds by Lemma 3). $\diamond$

Proof. To define π_h^2 we shall use a *macro-element* technique. In order to avoid the technicalities related with the use of macro-elements, we shall detail the proof only in the case when the mesh on Γ has an *even* number of intervals. It should be clear however that the result holds in general. Having an even number of elements, we can take non-overlapping macro-elements $\mathcal{J}$ made of pairs of adjacent elements. In the usual application of the macro-element technique (see [9] or [18]) the macro-elements overlap. Our case is simpler. In each macroelement $\mathcal{J}$, and for each w, say, in $L^2(\Gamma)$, we contruct π_h^2 as the element of W_h having support in $\mathcal{J}$ and such that

$$\int_{\mathcal{J}} (w - \pi_h^2 w) p_{k-1} \, \mathrm{d}x_1 = 0, \tag{5.33}$$

for all p_{k-1} continuous on $\mathcal{J}$ and polynomial of degree $\leq k-1$ in each of the two elements $\mathcal{I}$ of $\mathcal{G}_h$ contained in $\mathcal{J}$. The system (5.33) has $2k-1$ unknowns (the dimension of continuous locally P_k functions, on a mesh of two elements, vanishing at the endpoints) and $2k-1$ equations (the dimension of continuos locally P_{k-1} functions, on a mesh of two elements, with no conditions at the endpoints). It is easy to check that (5.33) has a unique solution, and that (5.25) and (5.26) hold true. $\diamond$

Remark 5.5. *The analysis of the previous three cases, together with Remarks 5.2 and 5.4, can often help in deciding whether other possible choices are viable or not. For instance, it is obvious that if we start from a case where the* inf-sup *condition holds and we increase W_h or decrease M_h, then the* inf-sup *condition will still hold. On the other hand, if we start from a case where the* inf-sup *condition does not hold, and we decrease W_h or we increase M_h then the* inf-sup *condition will still fail to hold.*

6. Conclusions

We have considered the solution of general contact problems for which a mixed finite element interpolation is used. The solution approach involves a Lagrange multiplier to interpolate the unknown normal contact tractions (in addition to the usual interpolations of the displacements for the bodies). The mixed formulation needs to satisfy stability requirements and the objective of this paper was to give insight into these requirements and give specific results as to what Lagrange multiplier interpolation is appropriate, and efffective, with a specific displacement interpolation.

While these results were derived by considering a simple model problem (in order to avoid certain technicalities), valuable insight was gained and there is no reason why the results should not be generally applicable. The analytical results reported in the paper confirm also earlier obtained conclusions based on numerical tests [1].

References

[1] N. El-Abbasi and K.J. Bathe, *Stability and Patch Test Performance of Contact Discretizations,* submitted to Computers & Structures.

[2] C. Baiocchi, G. Buttazzo, F. Gastaldi and F. Tomarelli, *General Existence Theorems for Unilateral Problems in Continuum Mechanics*, Arch. Rational Mech. Anal. **100** (1988) 149-189.

[3] C. Baiocchi and A. Capelo, *Variational and Quasivariational Inequalities, J. Wiley and Sons, 1984.*

[4] K.J. Bathe, *Finite Element Procedures, Prentice Hall, 1996.*

[5] K.J. Bathe, *The Inf-Sup Condition and its Evaluation for Mixed Finite Element Methods,* Computers & Structures **79** (2001), 243-252 and 971.

[6] K.J. Bathe, ed., *Computational Fluid and Solid Mechanics, Elsevier, 2001.*

[7] H. Brezis, *Problèmes unilateraux* J. Math. Pures Appl. **51** (1972) 1-168.

[8] F. Brezzi and K.J. Bathe, *A Discourse on the Stability Conditions for Mixed Finite Element Formulations,* Computer Methods in Applied Mechanics and Engineering **82** (1990), 27-57.

[9] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods, Springer-Verlag, 1991.*

[10] F. Brezzi, W.W. Hager and P.-A. Raviart, *Error Estimates for the Finite Element Solution of Variational Inequalities. Part II Mixed Methods,* Numer. Math. **31** (1978), 1-16.

[11] Ph.G. Ciarlet, *The Finite Element Method for Elliptic Problems, North Holland, 1978.*

[12] G. Duvaut and J.L. Lions, *Inequalities in Mechanics and Physics, Springer-Verlag, 1976.*

[13] G. Fichera, *Unilateral Constraints in Elasticity*, in *Actes du Congrès International des Mathématiciens (Nice, 1970)*, **3**, Gauthier-Villars, (1971), 79–84

[14] M. Fortin, *An Analysis of Convergence of the Mixed Finite Element Method,* RAIRO Analyse Numérique **11** (1997), 341-354.

[15] N. Kikuchi and J.T. Oden, *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods, SIAM Studies in Applied mathematics,* **8**, *1988.*

[16] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications, Academic Press, 1980.*

[17] J.L. Lions and E. Magenes, *Nonhomogeneous Boundary Value Problems and Applications, vol 1, Springer, 1972.*

[18] R. Stenberg, *A Family of Mixed Finite Elements for the Elasticity Problem*, Numer. Math **53** (1988), 513-538.

[19] O. Widlund, *An Extension Theorem for Finite Element Spaces with three Applications*, in Numerical Techniques in Continuum Mechanics (W. Hackbush and K. Witsch Eds.), Vieweg & Sohn, 1987.

Franco Brezzi, Dipartimento di Matematica and I.A.N.-C.N.R., Via Ferrata 1, 27100 Pavia, Italy

E-mail address: brezzi@dragon.ian.pv.cnr.it

K.J. Bathe, Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 massachusetts Avenue, Cambridge, MA 02139, USA

E-mail address: kjb@mit.edu