The p-version of the finite element method
for constraint boundary conditions

I. Babuška
Institute for Physical Science and Technology
University of Maryland
College Park, Maryland 20742

and

Manil Suri
Department of Mathematics
University of Maryland Baltimore County
Catonsville, Maryland 21228

Technical Note BN-1064

April 1987
The paper addresses the implementation of general constraint boundary conditions for a system of equations by the p-version of the finite element method. By constraint boundary conditions we mean conditions where some relation between the components is prescribed at the boundary. Optimal error bounds are proven.
The p-version of the finite element method for constraint boundary conditions

by

I. Babuška
Institute for Physical Science and Technology
University of Maryland
College Park, Maryland 20742

and

Manil Suri
Department of Mathematics
University of Maryland Baltimore County
Catonsville, Maryland 21228

1Partially supported by ONR Contract N00014-85-K-0169.

2Research partially supported by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR- 85-0322.
The p-version of the finite element method for constraint boundary conditions

Abstract. The paper addresses the implementation of general constraint boundary conditions for a system of equations by the p-version of the finite element method. By constraint boundary conditions we mean conditions where some relation between the components is prescribed at the boundary. Optimal error bounds are proven. Keywords: convergence, Sobolev spaces.
1. Introduction.

There is a large variety of boundary conditions for systems of differential equations of elliptic type. Some physically natural conditions may be formulated by a variational approach through constraint conditions. For example, the two dimensional elasticity problem can be formulated as the minimization of a quadratic functional $F(u)$, $u = (u_1, u_2)$ over a set $H$ satisfying

$$(H^1_0(\Omega))^2 \subset H \subset (H^1(\Omega))^2.$$ 

Selections of $H$ then characterize the boundary conditions.

Obviously the choice $H = (H^1_0(\Omega))^2$ induces the (essential) Dirichlet conditions, i.e., the displacement is given on $\partial \Omega$, while $H = (H^1(\Omega))^2$ induces the (natural) Neumann conditions, i.e., the tractions are prescribed on $\partial \Omega$. In addition to these classical conditions other types are important in applications. One of these conditions is characterized by

$$(1.1) \quad H = \{(u_1, u_2) \in (H^1(\Omega))^2 | u_1 \varphi_1(s) + u_2 \varphi_2(s) = 0 \text{ on } \partial \Omega\}$$

where $\varphi_1$ and $\varphi_2$ are given functions defined on $\partial \Omega$. These conditions are in the most simple case the symmetry conditions and in general traction free constraints at the boundary.

So far we have only mentioned homogeneous boundary conditions. Nonhomogeneous conditions are defined in the usual way, when the minimization of $F$ is over a hyperplane $H_v = \{u + v | u \in H, v \in (H^1(\Omega))^2\}$.

The constraint boundary condition we mentioned above is a type of essential condition. Hence when solving such problems by the finite element method in general and by the $p$ or $h$-$p$ versions
in particular, we face the problem of implementing the nonhomogeneous boundary conditions (which are outside the finite element space).

The p and h-p versions are recent developments, where p, the degree of the elements used is not fixed but is increasing. This is in contrast to the classical h-version, where the degree p is kept fixed. The first commercial programs available are PROBE (Noetic Tech., St. Louis) and FIESTA (ISMES, Bergamo, Italy).

The implementation of Dirichlet boundary conditions for the p-version of the finite element method has been addressed by us in [2] and [4]. A general survey on the state of the art of the p and h-p versions may be found in [1].

In this paper we will address the implementation of the constraint conditions (1.1) in a simplified setting (to avoid notational difficulties). Section 2 deals with preliminaries and notation. In Section 3 we formulate an abstract approach and based on it prove that the suggested finite element formulation of the constraint boundary condition leads to the optimal rate of convergence of the p-version. Section 4 addresses some implementational aspects.
2. The basic notation and preliminaries.

2.1. The Sobolev spaces

Let $\mathbb{R}^2$ be the two dimensional Euclidean space, $x = (x_1, x_2) \in \mathbb{R}^2$. Let $\Omega \subset \mathbb{R}^2$ be a bounded Lipschitzian domain with the boundary $\Gamma = \partial \Omega$. We will assume that $\Gamma$ is a Jordan curve, $\Gamma = \bigcup_{i=1}^{m} \Gamma_i$ where $\Gamma_i$ are smooth open arcs with parametric description

$$\Gamma_i = \{(x_1, x_2) | x_1 = x_{1,i}(\xi), x_2 = x_{2,i}(\xi), |\xi| < 1\},$$

$i = 1, \ldots, m$.

Denoting $I = (-1,1), \Gamma_i$ is obviously the image of $I$ by the mapping $F_i = (x_{1,i}, x_{2,i}),$ i.e., $\Gamma_i = F_i(I)$. If $u(s)$ is defined on $\Gamma_i$ then by $U(\xi) = u(F_i(\xi))$ we denote its transform on $I$.

The ends of $\Gamma_i$ will be called vertices and denoted by $A_i = (x_{1,i}(-1), x_{2,i}(-1))$, $B_i = (x_{1,i}(1), x_{2,i}(1))$. We will further assume that $B_i = A_{i+1}$, $B_m = A_1$, $i = 1, \ldots, m$. By this, the orientation of $\Gamma_i$ is established. In general we will denote the vertices by $A_i (= B_{i-1}), i = 1, \ldots, m$. The scheme of the domain and the pertinent notation is shown in Figure 2.1.

![Figure 2.1. Scheme of the domain and notation.](image-url)
Remark 2.1. We assumed that the domain $\Omega$ is simply connected. This assumption has been made only for notational simplicity.

Remark 2.2. We assumed that the domain is Lipschitzian. Once more, our results are valid (with proper modification) in the case when, for example, some arcs coincide (as in the case of the slit domain).

Remark 2.3. We have assumed that the arcs $\Gamma_i$ are sufficiently smooth. For the sake of simplicity we assume that they are $C^\infty$ arcs (i.e., the functions $x_{ij}$, $i = 1, \ldots, m$, $j = 1, 2$ are $C^\infty$ functions).

By $H^k(\Omega)$, $k \geq 0$ integer we denote the usual Sobolev space of functions with square integrable derivatives on $\Omega$. The norm will be denoted by $\| \cdot \|_{H^k(\Omega)}$. If $\ell < q < \ell + 1$, $\ell \geq 0$ integer, then we define $H^q(\Omega) = (H^\ell(\Omega), H^{\ell+1}(\Omega))_\theta$, $\theta = q - \ell$ where by $(\cdot, \cdot)_\theta$ we denote the usual interpolated space using the $K$-method (see [5]). The scalar product $(\cdot, \cdot)_{H^q(\Omega)}$ and the norm $\| \cdot \|_{H^q(\Omega)}$ are defined accordingly.

By $C^k(\bar{\Omega})$, $k \geq 0$ integer, we denote the space of all functions with $k$ continuous derivatives on $\bar{\Omega}$. It is possible to show that $H^k(\Omega) \hookrightarrow C^0(\bar{\Omega})$ for $k > 1$, where by $\hookrightarrow$ we denote continuous imbedding. On the other hand, $H^1(\Omega) \subset C^0(\bar{\Omega})$.

For $I = (-1,1)$, $H^k(I)$, $k \geq 0$ is defined analogously as before. If $k > 1/2$ then $H^k(I) \hookrightarrow C^0(\bar{I})$ but $H^k(I) \subset C^0(\bar{I})$ for $k \leq 1/2$.

So far we have defined $H^k(I)$, $k \geq 0$. We will also be interested in $H^k(I)$, $k < 0$. We define for $k \geq 0$
\[ \| u \|_{H^{-k}(I)} = \sup_{v \neq 0} \frac{\int_{-1}^{1} uv \, dx}{\| v \|_{H^k(I)}}. \]

(Let us remark that sometimes (see e.g. [5]) our space \( H^{-k}(I) \) is denoted by \( (H^k(I))' \), whereas \( H^{-k}(I) \) is used to denote the dual space of \( H^k_0(I) \)).

If \( u \) is defined on \( I_1 \), then we define
\[
H^k(I_1) = \{ u | u(F_1(\xi)) = U(\xi) \in H^k(I) \}
\]
\[
\| u \|_{H^q(I_1)} = \| u \|_{H^k(I)}
\]

So far we have considered only scalar functions on \(\Omega\) and \(I\).

The spaces of vector functions are defined by Cartesian products,
\[ 2H^k(\Omega) = (H^k(\Omega))^2. \]

Let now
\[
Q = \{(x_1, x_2) | |x_1| < 1, |x_2| < 1 \}
\]
\[
I_1^Q = \{(x_1, x_2) | |x_1| < 1, x_2 = -1 \}
\]

\(Q\) will be called the standard square and \(I_1^Q, i = 1, 2, 3, 4\) its sides (\(I_1^Q, i = 2, 3, 4\) are defined analogously to \(I_1^Q\) in an obvious way). Let
\[
T = \{(x_1, x_2) | |x_1| < 1, 0 < x_2 < (1 + x_1)^3 \text{ for } x_1 < 0, \]
\[
0 < x_2 < (1 - x_1)^3 \text{ for } x_1 > 0 \}
\]
\[
I_1^T = \{(x_1, x_2) | |x_1| < 1, x_2 = 0 \}.
\]

\(T\) will be called the standard triangle and \(I_1^T, i = 1, 2, 3\) its sides.
Let us remark that the sides of $T$ and $Q$ are each of length 2. Later we will often not distinguish between $\gamma_1$ and $I$.

We now define

\[ p^2_p(Q) = \{ u \mid u \text{ is a polynomial of degree } \leq p \text{ in each variable } x_1 \text{ and } x_2 \text{ over } Q \}. \]

\[ p^1_p(T) = \{ u \mid u \text{ is a polynomial of (total) degree } \leq p \text{ on } T \}. \]

\[ p_p(I) = \{ u \mid u \text{ is a polynomial of degree } \leq p \text{ on } I \}. \]

We have then

**Lemma 2.1.** Let $v \in H_0^1(\gamma_1^Q) \cap p_p(\gamma_1^Q)$ (respectively $v \in H_0^1(\gamma_1^T) \cap p_p(\gamma_1^T)$) such that

\[ \|v\|_{H^t(\gamma_1^Q)} \leq p^{-(1-t)} A, \quad t = 0,1 \]
respectively

\[ \|v\|_{H^1(\gamma_1)} \leq p^{-(1-t)}A, \quad t = 0, 1. \]

Then there exists \( u \in \mathcal{P}_p^2(Q) \) (respectively \( \mathcal{P}_p^1(T) \)) such that

\[ u|_{\gamma_1} = v \quad \text{(respectively} \quad u|_{T} = v), \quad u|_{\partial Q - \gamma_1} = 0 \quad \text{(respectively} \quad u|_{\partial T - \gamma_1} = 0) \]

and

\[ \|u\|_{H^1(Q)} \leq C_p^{-1/2}A \]

respectively

\[ \|u\|_{H^1(T)} \leq C_p^{-1/2}A. \]

For the proof see [2] or [3].

2.2. The model problem

Let

\[ 2H^1_0(\Omega) \subset \mathcal{X}(\Omega) \subset 2H^1(\Omega) \]

where \( \mathcal{X}(\Omega) \) is closed in \( 2H^1(\Omega) \). \( \mathcal{X}(\Omega) \) will be called the constraint space. Assume that there is given a continuous bilinear form \( B(u,v) \) on \( 2H^1(\Omega) \times 2H^1(\Omega) \), \( u = (u_1, u_2), v = (v_1, v_2) \) such that

\[ (2.1) \quad B(u,u) \geq \gamma \|u\|_{2H^1(\Omega)}^2, \quad \gamma > 0 \quad \text{for any} \quad u \in 2H^1(\Omega). \]

Then obviously for any \( G_1 \in (\mathcal{X}(\Omega))' \), there is a unique \( u_0 \in \mathcal{X}(\Omega) \) such that

\[ B(u_0,v) = G_1(v) \]

holds for any \( v \in \mathcal{X}(\Omega) \). We also have

\[ \|u_0\|_{2H^1(\Omega)} \leq C_1 G_1 \| (2H^1(\Omega))' \]
Denote \( \mathcal{X}_\rho(\Omega) = (u \in H^1(\Omega), u-\rho \in \mathcal{X}(\Omega)) \). \( \mathcal{X}_\rho(\Omega) \) will be called the \( \rho \)-hyperplane. Then our model problem is given by:

Find \( u_0 \in \mathcal{X}_\rho(\Omega) \) such that

\[
(2.2) \quad B(u_0, v) = G_1(v), \quad \forall \ v \in \mathcal{X}(\Omega).
\]

We have then

\[
(2.3) \quad \|u_0\|_{H^1(\Omega)} \leq C[\|\rho\|_{H^1(\Omega)} + \|G_1\|_{H^1(\Omega)}].
\]

If \( \rho = 0 \) then we will speak about a homogeneous constraint problem while for \( \rho \neq 0 \) we will speak about a nonhomogeneous constraint problem. We call these constraint problems because \( \mathcal{X}(\Omega) = H^1(\Omega) \).

There are many constraint problems in applications. We will consider the one when

\[
\mathcal{X}(\Omega) = \{(u_1, u_2) \in H^1(\Omega) \mid \sum_{\ell=1}^{2} a^{(j)}_{k,\ell} u_\ell|_{\Gamma_j} = 0, \ k = 1, 2, \ j = 1, \ldots, m\}
\]

where \( a^{(j)} = (a^{(j)}_{k,\ell}) \) are matrices of smooth functions on \( \Gamma_j \) (say \( C^\infty(\Gamma_j) \)). Additional assumptions on \( (a^{(j)}_{k,\ell}) \) will be imposed later.

Obviously when \( a^{(j)}_{k,k} = 1, \ a^{(j)}_{k,\ell} = 0 \) for \( k \neq \ell \) we get Dirichlet boundary conditions (in general we get Dirichlet conditions when \( a^{(j)} \) have rank 2 for all \( x \in \Gamma_j \)). If \( (a^{(j)}_{k,\ell}) = 0 \) there is no constraint and we have the Neumann problem.

If \( a^{(j)} \) has rank 1 then we can write the constraint on \( \Gamma_j \) as

\[
a^{(j)}_{1,1} u_1 + a^{(j)}_{1,2} u_2 = 0
\]

which will be written in the form
(2.4a) \[ \alpha(j)u_1 + \beta(j)u_2 = 0. \]

Obviously if \( \rho = (\rho_1, \rho_2) \) then the nonhomogeneous constraint problem is characterized on \( \Gamma_j \) by

(2.4b) \[ \alpha(j)u_1 + \beta(j)u_2 = \alpha(j)\rho_1 + \beta(j)\rho_2. \]

Problems of this type are common, for example, in the theory of elasticity. For simplicity of the exposition and notation we will restrict ourselves to the model problem where

\[
B(u,v) = \int_\Omega \left( \frac{\partial u_1}{\partial x_1} \frac{\partial v_1}{\partial x_1} + \frac{\partial u_1}{\partial x_2} \frac{\partial v_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1} \frac{\partial v_2}{\partial x_1} + \frac{\partial u_2}{\partial x_2} \frac{\partial v_2}{\partial x_2} 
+ u_1v_1 + u_2v_2 \right) dx_1 dx_2
\]

and

\[
G_1(v) = \int_\Omega (f_1v_1 + f_2v_2) dx_1 dx_2, \quad f = (f_1, f_2) \in H^0(\Omega).
\]

Although we restrict ourselves to this special case, our results hold in general, e.g., for elasticity problems, etc.

We will assume that \( a_{k,\ell}^{(j)} \in C^0(\overline{\Gamma_j}) \). In practice we have the nonhomogeneous constraint problem defined so that

1) If \( \{a_{k,\ell}^{(j)}\} \) has rank 2 on \( \overline{\Gamma_j} \) then the constraint is

\[ \sum_{\ell=1}^2 a_{k,\ell}^{(j)} u_\ell = g_k^{(j)}, \quad k = 1, 2 \]

where \( (g_1^{(j)}, g_2^{(j)}) \) are defined on \( \overline{\Gamma_j} \). Hence obviously we can transform the above constraint equation to

\[
\begin{align*}
u_1 &= g_1^{(j)} \\
u_2 &= g_2^{(j)}
\end{align*}
\]

(2.6)
Because $a_{k, \ell}^{(j)}$ are assumed to be smooth, we see that $\{g_{i}^{(j)}\}, i = 1,2$ have the same smoothness as $\{g_{i}^{(j)}\}$.

ii) If $a_{k, \ell}^{(j)}$ has rank 1 then the constraint equation is

$$a^{(j)} u_1 + 3^{(j)} u_2 = g^{(j)}.$$  

We add the condition

$$a^{(j)} + 3^{(j)} = 0 \text{ on } F_j.$$  

This enables us to transform the constraint equation to

$$a^{(j)} u_1 + 3^{(j)} u_2 = \tilde{g}^{(j)} \text{ with } a^{(j)} + 3^{(j)} = 1.$$  

As formulated above, $\tilde{g}^{(j)}$ are defined separately on each $F_j$. We will assume that $\tilde{g}^{(j)}$ satisfy consistency conditions, namely that there exists $\rho = (\rho_1, \rho_2) \in H^1(\Omega)$ such that

$$\tilde{g}^{(j)} = \rho_1 |F_j, \tilde{g}^{(j)} = \rho_2 |F_j,$$

respectively

$$a^{(j)} \rho_1 |F_j + 3^{(j)} \rho_2 |F_j = \tilde{g}^{(j)}.$$  

These conditions have to be imposed especially at the vertices of $\Omega$.

The sides $F_j$ where the constraint (2.6) is imposed will be called total constraint sides, while $F_j$ where the constraint (2.7) is imposed will be called partial constraint sides. We will enumerate the total constraint sides as $F_{1j}, j = 1, \ldots, m_1$ and the partial constraint sides as $F_{1j}, j = m_1+1, \ldots, m$.

2.3. The p-version of the finite element method

Assume that the domain $\Omega$ has been partitioned into a finite
number of subdomains $Q_i$, i.e., $\bar{\Omega} = \bigcup_{i=1}^{n} \bar{Q}_i$. We shall assume that $Q_i$ is the curvilinear quadrilateral

$$Q_i = \gamma_i(Q)$$

or curvilinear triangle

$$Q_i = \gamma_i(T)$$

where $Q$ and $T$ are the standard square and triangle, respectively. The domains $Q_i$ will be called elements. We will assume that $\gamma_i^{-1}$ is a smooth one to one mapping of $Q_i$ onto $Q$, respectively $T$. It is obvious what the vertices and sides of $Q_i$ correspond to. If $\gamma$ is a side of $Q_i$ then $\gamma_i$ induces mapping $F_i$ of $I$ onto $\gamma$ (realizing that all the sides of the standard square and triangle have the same length as $I$).

We shall assume the following about the partition and the mappings $F_i$:

1. If $\bar{Q}_i \cap \bar{Q}_j = R_{ij}, R_{ij} \neq \emptyset$ then $R_{ij}$ is either a common vertex or a side of both $Q_i$ and $Q_j$.

2. If $R_{ij} = \gamma_i, j$ then we will assume that the mappings $F$ of $I$ onto $\gamma_i, j$ induced by the mappings $\gamma_i$ and $\gamma_j$ are identical. We denote $F$ by $F_{ij}$. This implies the following. Let $A, B$ the vertices of $Q_i$ and $Q_j$ be the end points of $\gamma_{ij}$.

Assume that $(a_1, b_1)$ and $(a_2, b_2)$ are the end points of the sides $\gamma_{i}$ or $\gamma_{j}$ such that $\gamma_i(a_1) = \gamma_j(a_2) = A, \gamma_i(b_1) = \gamma_j(b_2) = B$. Then if $C \in \gamma_{ij}$, and $C = \gamma_i(c_1) = \gamma_j(c_2)$, $\overline{a_1c_1} = \overline{a_2c_2}$ and $\overline{c_1b_1} = \overline{c_2b_2}$.

Since we assumed that $\gamma_i$ are smooth mappings, the vertices of $\Omega$ necessarily have to coincide with some of the vertices of $Q_i$. We will further assume that for any $F_j$ there is an element
$Q_1$ such that one of its sides coincides with $\Gamma_j$. This assumption is made without any loss of generality.

Denote now

$$1_{\mathcal{P}}(Q) = \{ u \in H^1(Q) \mid u|_{Q_1}(\tau_j(\xi)) \in \mathcal{P}_p(Q) \text{ if } Q_1 \text{ is a quadrilateral and } u|_{Q_1}(\tau_j(\xi)) \in \mathcal{P}_p(I) \text{ if } Q_1 \text{ is a triangle} \}.$$

$$2_{\mathcal{P}}(\mathcal{P}) = (1_{\mathcal{P}}(\mathcal{P}))^2.$$

$$2_{\mathcal{P}}(\mathcal{P}) = \{ u \mid u_j|_{\Gamma_1}((\mathcal{P}_1(\xi))) \in \mathcal{P}_p(I), j = 1, 2 \}.$$

Here $\xi = \Gamma_j$ or any side of an element. Let us define the constraint space $\mathcal{P}(Q) = 2_{\mathcal{P}}(\mathcal{P})$ as follows.

1) If $\Gamma_j$ is a total constraint side with end points $A_j, A_{j+1}$ and $u \in \mathcal{P}(Q)$, then $u(A_j) = u(A_{j+1}) = 0, i = 1, 2$, and

$$\int_{\Gamma_j} u_i \nu_i ds = 0, \text{ for all } u_i \in 1_{\mathcal{P}}(\Gamma_j), i = 1, 2.$$

2) If $\Gamma_j$ is a partial constraint side then

$$(\alpha^{(j)}u_1 + \beta^{(j)}u_2)(A_k) = 0, k = j, j+1$$

and

$$\int_{\Gamma_j} (\alpha^{(j)}u_1 + \beta^{(j)}u_2) \nu ds = 0, \text{ for all } u \in 2_{\mathcal{P}}(\Gamma_j).$$

The $\rho$ hyperplane $\mathcal{P}(Q)$ is defined analogously. Let $g$ be defined in terms of $\rho$ by (2.8). Then on $\Gamma_j$ we impose

$$u_i(A_k) = g_i^{(j)}(A_k), i = 1, 2, k = j, j+1$$

respectively.
\[(a^{(j)}u_1 + b^{(j)}u_2)(A_k) = g^{(j)}(A_k), \quad k = j, j+1\]

and

\[(2.9a) \quad \int_{\Gamma_j} u_j v_1 ds = \int_{\Gamma_j} g^{(j)} v_1 ds, \quad i = 1, 2\]

respectively

\[(2.9b) \quad \int_{\Gamma_j} (a^{(j)}u_1 + b^{(j)}u_2) v ds = \int_{\Gamma_j} g^{(j)} v ds.\]

The p-version is then defined analogously as before. Find \(u_p \in X_p^p(\Omega)\) such that

\[(2.10) \quad B(u_p, v) = G_1(v) \quad \forall \ v \in X_p^p(\Omega).\]

**Remark.** Constraints of the type considered are typical in elasticity theory. Here \(u_1\) and \(u_2\) are the displacements in the directions \(x_1\) and \(x_2\) respectively. Assume now that the displacement is constrained in the normal direction only (and is friction free in the tangential direction). Then on the boundary, we obtain the partial constraint \(u_1 \cos \phi + u_2 \sin \phi = 0\) where \(\phi\) is the angle of the outer normal with the axis \(x_1\).
3. The convergence of the p-version of the finite element method.

3.1. An abstract result

We will first describe an abstract framework which will be the basis for the forthcoming analysis.

Let \( X \) and \( W \) be Hilbert spaces and \( X_p \subset X, W_p \subset W, p = 1,2,... \) be one parameter families of finite-dimensional subspaces. \( X_p \subset X, W_p \subset W \) will denote corresponding families of hyperplanes such that \((u-v) \in X_p\) whenever \( u,v \in X_p \) and \((\phi-v) \in W_p\) whenever \( \phi, v \in W_p \).

Let \( a(u,v), u,v \in X \) be a continuous bilinear form on \( X \times X \) and \( b(v,\phi) \) be a continuous bilinear form on \( X \times W \) such that
\[
b(v,\phi) \leq C\|v\|_X\|\phi\|_W.
\]

Let \( u_0 \in X, \phi_0 \in W \) and \( u_p \in X_p, \phi_p \in W_p \) be such that
\[
(a) \quad a(u_0,v) + b(v,\phi_0) = F_1(v) \forall v \in X_p \\
(b) \quad b(u_0,v) = F_2(v) \forall v \in W_p
\]
and
\[
(a) \quad a(u_p,v) + b(v,\phi_p) = F_1(v) \forall v \in X_p \\
(b) \quad b(u_p,v) = F_2(v) \forall v \in W_p.
\]

Define \( Z_p = (v \in X_p, b(v,\phi) = 0 \forall \phi \in W_p) \subset X_p \). Then we have

**Theorem 3.1.** Let \( a(u,u) \geq \gamma\|u\|_X^2, \gamma > 0, \) for any \( u \in Z_p \). Then
\[
(b) \quad \inf_{u_0-W_p} u_0 - w_p x + \inf_{\phi_0-W_p} \phi_0 - \phi p w \\
b(u_0-w_p,\phi) = 0 \forall \phi \in W_p \forall w \in W_p
\]
\[
u_p - w_p \in \hat{X}_p
\]
(i.e., \( u_p-w_p \in Z_p \))

15
Proof. for arbitrary \( w_p \in X_p \) and \( y_p \in W_p \) we have

\[
(3.4a) \quad a(u_p - w_p, v) + b(v, \varphi_0 - y_p) = a(u_0 - w_p, v) + b(v, \varphi_0 - y_p) \quad \forall v \in X_p
\]

\[
(3.4b) \quad b(u_p - w_p, v) = b(u_0 - w_p, v) \quad \forall v \in W_p
\]

For \( v \in Z_p \) (3.4a) yields

\[
(3.5) \quad a(u_p - w_p, v) = a(u_0 - w_p, v) + b(v, \varphi_0 - y_p)
\]

Suppose now that \( w_p \) is such that

\[
b(w_p, v) = b(u_0, v) \quad \forall v \in W_p.
\]

Then by (3.4b)

\[
b(u_p - w_p, v) = 0 \quad \forall v \in W_p
\]

and hence

\[
u_p - w_p \in Z_p.
\]

Now using \( v = u_p - w_p \) in (3.5) we get

\[
a(u_p - w_p, u_p - w_p) \leq C\left[\|u_0 - w_p\|_X + \|u_p - w_p\|_X + \|\varphi_0 - y_p\|_W\right]\|u_p - w_p\|_X
\]

and hence by coercivity of \( a(\cdot, \cdot) \) on \( Z_p \),

\[
\|u_p - w_p\|_X \leq C\left[\|u_0 - w_p\|_X + \|\varphi_0 - y_p\|_W\right]\|u_p - w_p\|_X
\]

and hence also

\[
\|u_0 - w_p\|_X \leq C\left[\|u_0 - w_p\|_X + \|\varphi_0 - y_p\|_W\right]\|u_0 - w_p\|_X
\]

from which (3.3) follows.

3.2. The convergence of the \( p \)-version

Let \( u_0 = (u_{0,1}, u_{0,2}) \in H^1(\Omega) \) be the solution of our constrained problem (2.2), (2.5) and \( u_p = (u_{p,1}, u_{p,2}) \in \mathbb{P}_p(\Omega) \) be the approximation given by (2.10).
We will assume that $u_0 \in H^k(\Omega)$, $k > 3/2$. Hence $\frac{\partial u_0}{\partial n} \in H^0(\Gamma_j^i)$, $i = 1, 2, \ldots, m$. Let the constraints on $\Gamma_j$ be as in section 2.2, with $\Gamma_j^i$, $j = 1, \ldots, m_1$ being the total constraint sides and $\Gamma_j^j$, $j = m_1 + 1, \ldots, m$ the partial constraint sides. Then it can be verified that for any $v \in H^1(\Omega)$,

$$B(u_0, v) - \sum_{j=1}^{m_1} \int_{\Gamma_j^i} \left( \frac{\partial u_0, 1}{\partial n} v_1 + \frac{\partial u_0, 2}{\partial n} v_2 \right) ds$$

(3.6)

$$- \sum_{j=m_1+1}^{m} \int_{\Gamma_j^j} (\alpha_j v_1 + \beta_j v_2) \left( \frac{\partial u_0, 1}{\partial n} + \frac{\partial u_0, 2}{\partial n} \right) ds = G_1(v)$$

where we have assumed $\alpha_j + \beta_j = 1$. Moreover, for $v, v_1, v_2$ and $\rho$ as in (2.6) - (2.9),

$$\sum_{j=1}^{m_1} \int_{\Gamma_j^i} (u_p, 1 v_1 + u_p, 2 v_2) ds + \sum_{j=m_1+1}^{m} \int_{\Gamma_j^j} (\alpha_j u_p, 1 + \beta_j u_p, 2) ds$$

(3.7)

$$= \sum_{j=1}^{m_1} \int_{\Gamma_j^i} (u_0, 1 v_1 + u_0, 2 v_2) ds + \sum_{j=m_1+1}^{m} \int_{\Gamma_j^j} (\alpha_j u_0, 1 + \beta_j u_0, 2) ds$$

$$= \sum_{j=1}^{m_1} \int_{\Gamma_j^i} (\rho_1 v_1 + \rho_2 v_2) ds + \sum_{j=m_1+1}^{m} \int_{\Gamma_j^j} (\alpha_j \rho_1 + \beta_j \rho_2) ds$$

and

(3.8a) $u_{p, k}(A_j) = u_{0, k}(A_j) = \rho_{k}(A_j)$, $j = \ell, \ell + 1$, $k = 1, 2$.

$\ell = 1, \ldots, m_1$
\[(a^\ell u_{p,1})(A_j) + (b^\ell u_{p,2})(A_j) = (a^\ell u_{0,1})(A_j) + (b^\ell u_{0,2})(A_j)\]

\[(3.8b)\]

\[j = \ell, \ell+1, \ell = i_{m_1+1}, \ldots, i_m.\]

(We remark that \(u_{0,1}(A_j), i = 1, 2\) has meaning because we assumed that \(u_0 \in H^k(\Omega), k > 3/2\).)

We now define

\[X = H^1(\Omega), \quad X = \mathcal{H}^1(\Omega)\]

and for any \(\delta = (\delta_1, \delta_2) \in X\),

\[X_p, \delta = (u = (u_1, u_2) \in H_p(\Omega), u_k(A_j) = \delta_k(A_j), k = 1, 2\]

\[j = \ell, \ell+1, \ell = i_1, \ldots, i_{m_1}, (a^\ell u_1)(A_j) + (b^\ell u_2)(A_j)\]

\[= (a^\ell \delta_1)(A_j) + (b^\ell \delta_2)(A_j), j = \ell, \ell+1, \ell = i_{m_1+1}, \ldots, i_m.\]

We then take in our abstract framework

\[X_p = X_p, \rho, \quad \bar{X}_p = X_p, 0\]

where \(\rho\) satisfies (2.6) - (2.8). Moreover, let

\[W = \prod_{j=1}^{i_{m_1}} H^{-1/2}(\Gamma_{1j}), \quad \prod_{j=i_{m_1}+1}^{i_m} H^{-1/2}(\Gamma_{1j})\]

with the norm

\[\|v\|_W = \left[ \sum_{j=1}^{i_{m_1}} \|v_j\|^2_{H^{-1/2}(\Gamma_{1j})} + \sum_{j=i_{m_1}+1}^{i_m} \|v_j\|^2_{H^{-1/2}(\Gamma_{1j})} \right]^{1/2}\]

We see then that \(\phi_0 \in W\) where (see (3.6))
\[ \phi_{0,j} = \left( \begin{array}{c} \frac{\partial u_{0,1}}{\partial n} |_{\Gamma_{i,j}} \\ \frac{\partial u_{0,2}}{\partial n} |_{\Gamma_{i,j}} \end{array} \right), \quad j = 1, \ldots, m_1 \]

\[ = \left( \begin{array}{c} \frac{\partial u_{0,1}}{\partial n} + \frac{\partial u_{0,2}}{\partial n} \end{array} \right) |_{\Gamma_{i,j}}, \quad j = m_1 + 1, \ldots, m. \]

Define \( W_p = \prod_{j=1}^m W_p \subset W \) where

\[ j W_p = \tilde{W}_p = \prod_{j=1}^m W_p \subset W \]

\[ j W_q = \mathbb{H}_1 \left( \psi_{i,j} \right) \cap \mathbb{H}_p-2 \left( \Gamma_j \right), \quad j = 1, \ldots, m_1 \]

\[ = \mathbb{H}_1 \left( \psi_{i,j} \right) \cap \mathbb{H}_p-2 \left( \Gamma_j \right), \quad j = m_1 + 1, \ldots, m. \]

Let \( a(\cdot, \cdot) \) and \( b(\cdot, \cdot) \) be bilinear forms defined respectively on \( X \times X \) and \( X \times W \) by

\[ a(u, v) = B(u, v) \]

\[ b(u, v) = \sum_{j=1}^{m_1} \int_{\Gamma_{i,j}} (u_1 \psi_1 + u_2 \psi_2) ds + \sum_{j=m_1+1}^m \int_{\Gamma_{i,j}} (a_j u_1 + b_j u_2) \psi ds. \]

It may be seen that the right hand side of (3.7) defines a linear functional \( G_2 \) on \( W_p \). Then (3.6) - (3.8) show that \((u_0, \phi_0)\) satisfy (3.1) with \( F_k = G_k \), \( k = 1, 2 \). Moreover, if we can find a unique pair \((u_p, \phi_p)\) satisfying (3.2), then \( u_p \) will be precisely our finite element solution satisfying (2.10). We will now verify that the mixed method defined above satisfies the assumptions of Theorem 3.1. This in turn will lead to the existence and uniqueness of the solution \((u_p, \phi_p)\) of (3.2) and an estimate of the rate of convergence of \( u_p \) to \( u \).

Obviously, \( a(u, v) \) satisfies the desired continuity and
coercivity conditions. For \( j = 1, 2, \ldots, m \) we have

\[
\left| \int_{\Gamma_{ij}} ur \, ds \right| \leq C u_{H^{1/2}(\Gamma_{ij})} h^{-1/2}(\Gamma_{ij})
\]

from which the continuity of \( b(\cdot, \cdot) \) may be deduced. Hence Theorem 3.1 is applicable. Let us now estimate \( \inf_{r \in W_p} \varphi_0 - r \).

First, let \( m_1 + 1 \leq j \leq m \). We assumed that \( o_{0,j} \in H^{k-3/2}(\Gamma_{ij}) \), \( k > 3/2 \). Hence \( o_{0,j}(F_{ij}(t)) = \mathcal{G}(t) \in H^{k-3/2}(I) \). Let \( \sigma = \psi_{p-2}^{(I)} \) be such that

\[
(3.9) \quad \int_I \sigma \rho \, dt = \int_I \psi \rho \, dt, \quad \forall \rho \in \psi_{p-2}^{(I)}.
\]

Then, with \( q = \mathcal{G} - \sigma \), we have

\[
(3.10) \quad q \sigma \in H^0(I) \Rightarrow \psi_{p-2}^{(I)} H^{k-3/2}(I).
\]

Now, for arbitrary \( v \in H^1(I) \), we have by (3.9)

\[
\frac{\int_I q v \, dt}{H^1(I)} = \frac{\int_I q(v - \sigma_1) \, dt}{H^1(I)} = \frac{q H^0(I)}{v H^1(I)} \leq \psi_{p-2}^{(I)} q H^0(I)
\]

where \( \sigma_1 \) is a polynomial of degree \( p - 2 \) satisfying

\[
\psi_{p-2}^{(I)} H^0(I) \Rightarrow \psi_{p-2}^{(I)} H^1(I).
\]

This yields

\[
(3.11) \quad q \mathcal{G}^{-1} H^{-1}(I) \Rightarrow \psi_{p-2}^{(I)} H^{-1}(I) \Rightarrow \psi_{p-2}^{(I)} H^{k-3/2}(I).
\]

Interpolating (3.10), (3.11) and using the fact that \( F_{ij} \) is a smooth mapping, we obtain
$$\inf_{\chi_p \in \mathcal{W}_p} \| \phi_0, j \|_{X_p} - \chi_p \|_{H^{-1/2}(\Omega)} \leq C_p^{-(k-1)} \| \phi_0, j \|_{H^{k-3/2}(\Omega)}.$$ 

We get similar estimates for $\Gamma_{ij}$, $j = 1, \ldots, m_1$, so that

$$\inf_{\chi_p \in \mathcal{W}_p} \| \phi_0, j \|_{X_p} - \chi_p \|_{W} \leq C_p^{-(k-1)} \| u_0 \|_{H^{k}(\Omega)}.$$ 

We now estimate $\inf \| u_0 - w \|_{X}$. Using the results from [2], there exist $z_i \in \mathcal{P}_p(\Omega), i = 1, 2$ such that

\begin{align*}
(3.12a) & \quad \| u_0, i - z_i \|_{H^t(\Omega)} \leq C_p^{-(k-t)} \| u_0, i \|_{H^t(\Omega)}, i = 1, 2, t = 0, 1 \\
(3.12b) & \quad u_0, i(N) = z_i(N) \text{ for each node } N \text{ of the mesh} \\
(3.12c) & \quad \| u_0, i - z_i \|_{H^t(\Gamma_j)} \leq C_p^{-(k-1/2-t)} \| u_0, i \|_{H^t(\Omega)}, t = 0, 1, i = 1, 2, j = 1, 2, \ldots, m
\end{align*}

Let $m_1 + 1 \leq j \leq m$. Let us denote $\kappa = \alpha_j u_0, i - z_i + \beta_j u_0, z_i$. Then we have $\kappa(A_j) = \kappa(A_{j+1}) = 0$. Let $\widetilde{\kappa}(\xi) = \kappa(F(\xi))$ and let $\tau(F(\xi)) = \tilde{\tau}(\xi) \in \mathcal{P}_p(I)$ satisfy $\tau(\omega) = 0$ and

$$\int_I \tilde{\kappa} v \, dx = \int_I \tilde{\tau} v \, dx, \quad \forall v \in \mathcal{P}_{p-2}(I).$$

Because of (3.12b) we can write

$$\int_I \tilde{\kappa} \omega' \, dx = \int_I \tilde{\tau} \omega' \, dx, \quad \forall \omega \in \mathcal{P}_p(I), \omega(\omega) = 0$$

and hence by Lemma 3.2 of [2]

$$\| \kappa - \tilde{\tau} \|_{H^t(I)} \leq C_p^{-(k-1/2-t)} \| u_0 \|_{H^t(\Omega)}, t = 0, 1.$$ 

Using (3.12c), this gives
Now using Lemma 2.1 it follows that there is a \( w \in \mathcal{P}_p(\Omega) \) such that \( w = 0 \) on \( \partial \Omega \) where \( \tilde{\Omega} \) is the element with the side \( \Gamma_{i_j}, \ w = \tau \) on \( \Gamma_{i_j}, \ w = 0 \) on \( \partial \tilde{\Omega} - \Gamma_{i_j} \) and

\[
\|w\|_{H^1(\Omega)} \leq C_p \|u_0\|_{H^k(\Omega)}.
\]

(3.14)

Letting \( w_{ij} = (w, w) \in \mathcal{P}_p(\Omega) \), we see that \( w_{ij} \) will satisfy (3.14) with \( H^1(\Omega) \) replaced by \( H^2(\Omega) \).

Let now \( w_p = z + w_{ij} = (z_1, z_2) + (w, w) \). Using (3.13b) and the fact that \( w(A_\ell) = 0 \), we obtain

\[
(\alpha^1_J(u_0, 1 - w_p, 1) + \beta^1_J(u_0, 2 - w_p, 2))(A_\ell) = 0, \ \ell = i, i+1.
\]

Moreover, for \( \psi \in \mathcal{P}_{p-2}(\Gamma_{i_j}) \),

\[
\int_{\Gamma_{i_j}} (\alpha^1_J(u_0, 1 - w_p, 1) + \beta^1_J(u_0, 2 - w_p, 2))\psi \, ds
\]

\[
= \int_{\Gamma_{i_j}} (\kappa - (\alpha^1_J + \beta^1_J)w)\psi \, ds = \int_{\Gamma_{i_j}} (\kappa-w)\psi \, ds = 0
\]

where we have used \( \alpha^1_J + \beta^1_J = 1 \). We may construct \( w_{ij} \) as above for all partial constraint sides. An analogous construction can be carried out for total constraint sides as well. Then if

\[
w_p = z + \sum_{j=1}^{m} w_{ij}
\]

we see that
\[ u_p - w_p \in \tilde{X}_p \]

\[ b(u_0 - w_p, \psi) = 0, \forall \psi \in W_p \]

and

\[ \|u_0 - w_p\|_{X} \leq \|u_0 - z\|_{X} + \sum_{j=1}^{m} \|w_j\|_{X} \leq C_{p}^{-1} (k-1) \|u_0\|_{2H^k(\Omega)}. \]

This provides a bound for the first term in the right hand side of (3.3). Hence we have proven

**Theorem 3.2.** Let \( u_0 \in H^{k}(\Omega), k > 3/2 \). Then

\[ \|u_0 - u_p\|_{L^2(\Omega)} \leq C_{p} (k-1) \|u_0\|_{2H^k(\Omega)} \]

where \( u_0 \) is the exact solution and \( u_p \) is the finite element solution of the constrained problem, provided that \( u_0 \) and \( u_p \) exist.

The next theorem deals with the question of existence and uniqueness of \( (u_0, \phi_0) \) and \( (u_p, \phi_p) \).

**Theorem 3.3.** The (exact) solution \( (u_0, \phi_0) \) of the constrained problem exists. The finite element solution \( (u_p, \phi_p) \) exists and is unique.

**Proof.** In Section 2.2 we have shown that \( u_0 \) exists and hence \( (u_0, \phi_0) \) exists, too. The finite element solution \( (u_p, \phi_p) \) is determined by the solution of a linear system of equations with square matrix. Hence the existence follows from the uniqueness.

Assume therefore that there is a solution \( (u_p, \phi_p) \) of the trivial problem. Obviously \( \tilde{u} = \tilde{\phi} = 0 \) is also a solution of this problem.

Hence \( u_p = 0 \) because of Theorem 3.1. We have to show therefore that
\[
\int_I (\alpha v_1 + \beta v_2) \varphi_p \, d\zeta = 0
\]
implies \( \varphi_p = 0 \). Because \( \alpha + \beta = 1 \) we also have \( \int_I v \varphi_p \, d\zeta = 0 \) for all \( v \in \mathcal{P}_p(I) \cap H_0^1(I) \) while \( \varphi_p \in \mathcal{P}_{p-2}(I) \). This leads to \( \varphi_p = 0 \) which leads to the desired result.

Remark. We have dealt only with a model problem. It is obvious that the theorem holds in general, as for example, for the theory of elasticity.
4. **Some aspects of implementation.**

Here we will make some comments about the implementation in
the framework of the code PROBE* (see [6]). The shape functions
are defined as usual on the standard square or triangle. There are
three types:

a) the model shape functions which are linear on every side
   of $Q$, respectively $T$;

b) the side functions which are zero at the vertices of $Q$,
   respectively $T$ and on $\gamma$ are of the form
   $$\xi_j = \int_{-1}^{1} \xi_j(\xi) dx, \ j = 1, 2, \ldots$$
   where $\xi_j$ is the Legendre polynomial of degree $j$. $\xi_j$
   is then a polynomial of degree $j + 1$;

c) The internal shape functions which are zero on $\partial Q$
   (respectively $\partial T$).

The stiffness matrices are first computed in the standard way
without constraints. Then the constraints are imposed at the ver-
tices $A_j$. This only involves the amplitudes for the nodal shape
functions. Then the conditions (2.9 a,b) only involve amplitudes
for the side shape functions. The functions $v$ in (2.9 a,b) are
computed as derivatives of the Legendre polynomials from the usual
recurrence formula and the integration is made using numerical
quadrature.

The condition (2.9a) is especially simple because $u_i =
\sum c_i \xi_i$. Integrating by parts and exploiting orthogonality of the

---

*The code PROBE is the code of Noetic Tech., St. Louis.
Legendre polynomials we get the amplitudes for the side shape functions on the total constraint sides directly.
References


The Laboratory for Numerical Analysis is an integral part of the Institute for Physical Science and Technology of the University of Maryland, under the general administration of the Director, Institute for Physical Science and Technology. It has the following goals:

- To conduct research in the mathematical theory and computational implementation of numerical analysis and related topics, with emphasis on the numerical treatment of linear and nonlinear differential equations and problems in linear and nonlinear algebra.

- To help bridge gaps between computational directions in engineering, physics, etc., and those in the mathematical community.

- To provide a limited consulting service in all areas of numerical mathematics to the University as a whole, and also to government agencies and industries in the State of Maryland and the Washington Metropolitan area.

- To assist with the education of numerical analysts, especially at the postdoctoral level, in conjunction with the Interdisciplinary Applied Mathematics Program and the programs of the Mathematics and Computer Science Departments. This includes active collaboration with government agencies such as the National Bureau of Standards.

- To be an international center of study and research for foreign students in numerical mathematics who are supported by foreign governments or exchange agencies (Fulbright, etc.)

Further information may be obtained from Professor I. Babuška, Chairman, Laboratory for Numerical Analysis, Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742.
END
9-87
DTIC