

The p -Version of the Finite Element Method for Constraint Boundary Conditions

By I. Babuška* and Manil Suri**

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.

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

$$(H_0^1(\Omega))^2 \subset H \subset (H^1(\Omega))^2.$$

Selections of H then characterize the boundary conditions.

Obviously, the choice $H = (H_0^1(\Omega))^2$ 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

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

where φ_1 and φ_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 $H_v = \{u + v \mid u \in H, v \in (H^1(\Omega))^2\}$.

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).

Received May 13, 1987.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 65N30, 41A10, 41A25.

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

**Research partially supported by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-85-0322.

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. Basic Notation and Preliminaries.

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

$$\Gamma_i = \{(x_1, x_2) \mid x_1 = x_{i,1}(\xi), x_2 = x_{i,2}(\xi), |\xi| < 1\}, \quad i = 1, \dots, m.$$

Denoting $I = (-1, 1)$, Γ_i is obviously the image of I by the mapping

$$F_i = \{x_{i,1}, x_{i,2}\},$$

i.e., $\Gamma_i = F_i(I)$. If $u(s)$ is defined on Γ_i , then by $U(\xi) = u(F_i(\xi))$ we denote its transform on I . The ends of Γ_i will be called *vertices* and denoted by $A_i = (x_{i,1}(-1), x_{i,2}(-1))$, $B_i = (x_{i,1}(1), x_{i,2}(1))$. We will further assume that $B_i = A_{i+1}$, $B_m = A_1$, $i = 1, \dots, m$. By this, the orientation of Γ_i is established. In general we will denote the vertices by $A_i (= B_{i-1})$, $i = 1, \dots, m$. The scheme of the domain and the pertinent notation is shown in Figure 2.1.

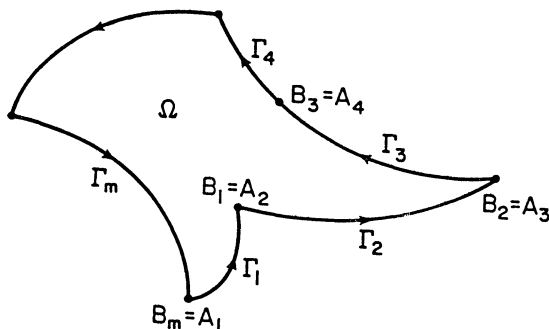


FIGURE 2.1

Scheme of the domain and notation.

Remark 2.1. We assumed that the domain Ω 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 Γ_i are sufficiently smooth. For the sake of simplicity we assume that they are C^∞ arcs (i.e., the functions $x_{i,j}$, $i = 1, \dots, m$, $j = 1, 2$, are C^∞ functions).

By $H^k(\Omega)$, $k \geq 0$ integer, we denote the usual Sobolev space of functions with square integrable derivatives on Ω . The norm will be denoted by $\|\cdot\|_{H^k(\Omega)}$. If $l < q < l + 1$, $l \geq 0$ integer, then we define $H^q(\Omega) = (H^l(\Omega), H^{l+1}(\Omega))_\theta$, $\theta = q - l$, where by $(\cdot, \cdot)_\theta$ we denote the usual interpolated space using the K -method (see [5]). The scalar product $(\cdot, \cdot)_{H^q}$ 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, $H^1(\Omega) \not\subset C^0(\bar{\Omega})$.

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 $H^k(I) \not\subset C^0(\bar{I})$ 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 $k \geq 0$

$$\|u\|_{H^{-k}(I)} = \sup_{v \in H^k(I), v \neq 0} \frac{\int_{-1}^1 uv \, dx}{\|v\|_{H^k(I)}}.$$

(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 $H_0^k(I)$.)

If u is defined on Γ_i , then we define

$$H^k(\Gamma_i) = \{u \mid u(F_i(\xi)) = U(\xi) \in H^k(I)\},$$

$$\|u\|_{H^q(\Gamma_i)} = \|U\|_{H^k(I)}.$$

So far we have considered only scalar functions on Ω and I . The spaces of vector functions are defined by Cartesian products, ${}^2H^k(\Omega) = (H^k(\Omega))^2$.

Now let

$$Q = \{(x_1, x_2) \mid |x_1| < 1, |x_2| < 1\},$$

$$\gamma_1^Q = \{(x_1, x_2) \mid |x_1| < 1, x_2 = -1\}.$$

Q will be called the standard square and γ_i^Q , $i = 1, 2, 3, 4$, its sides (γ_i^Q , $i = 2, 3, 4$, are defined analogously to γ_1^Q in an obvious way). Let

$$T = \{(x_1, x_2) \mid |x_1| < 1, 0 < x_2 < (1 + x_1)\sqrt{3} \text{ for } x_1 < 0,$$

$$0 < x_2 < (1 - x_1)\sqrt{3} \text{ for } x_1 > 0\},$$

$$\gamma_1^T = \{(x_1, x_2) \mid |x_1| < 1, x_2 = 0\}.$$

T will be called the standard triangle and γ_i^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 γ_i and I .

We now define

$$\mathcal{P}_p^2(Q) = \{u \mid u \text{ is a polynomial of degree } \leq p$$

$$\text{in each variable } x_1 \text{ and } x_2, \text{ over } Q\},$$

$$\mathcal{P}_p^1(T) = \{u \mid u \text{ is a polynomial of (total) degree } \leq p \text{ on } T\},$$

$$\mathcal{P}_p(I) = \{u \mid u \text{ is a polynomial of degree } \leq p \text{ on } I\}.$$

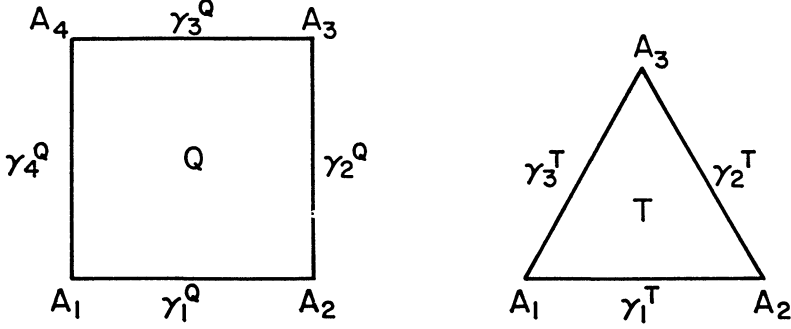


FIGURE 2.2

The scheme of the standard square and standard triangle.

We then have

LEMMA 2.1. Let $v \in H_0^1(\gamma_1^Q) \cap \mathcal{P}_p(\gamma_1^Q)$ (respectively $v \in H_0^1(\gamma_1^T) \cap \mathcal{P}_p(\gamma_1^T)$) such that

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

respectively

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

Then there exists $u \in \mathcal{P}_p^2(Q)$ (respectively $\mathcal{P}_p^1(T)$) such that $u|_{\gamma_1^Q} = v$ (respectively $u|_{\gamma_1^T} = v$), $u|_{\partial Q - \gamma_1^Q} = 0$ (respectively $u|_{\partial T - \gamma_1^T} = 0$) and

$$\|u\|_{H^1(Q)} \leq Cp^{-1/2} A,$$

respectively

$$\|u\|_{H^1(T)} \leq Cp^{-1/2} A.$$

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

2.2. The Model Problem. Let

$${}^2H_0^1(\Omega) \subset \mathcal{H}(\Omega) \subset {}^2H^1(\Omega),$$

where $\mathcal{H}(\Omega) \not\subset {}^2H^1(\Omega)$ is closed in ${}^2H^1(\Omega)$. $\mathcal{H}(\Omega)$ 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

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

Then, obviously, for any $G_1({}^2H^1(\Omega))'$ there is a unique $u_0 \in \mathcal{H}(\Omega)$ such that

$$B(u_0, v) = G_1(v)$$

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

$$\|u_0\|_{{}^2H^1(\Omega)} \leq C \|G_1\|_{({}^2H^1(\Omega))'}.$$

Denote $\mathcal{H}_\rho(\Omega) = \{u \in {}^2H^1(\Omega), u - \rho \in \mathcal{H}(\Omega)\}$. $\mathcal{H}_\rho(\Omega)$ will be called the ρ -hyperplane. Then our model problem is given by: Find $u_0 \in \mathcal{H}_\rho(\Omega)$ such that

$$(2.2) \quad B(u_0, v) = G_1(v) \quad \forall v \in \mathcal{H}(\Omega).$$

We then have

$$(2.3) \quad \|u_0\|_{{}^2H^1(\Omega)} \leq C (\|\rho\|_{{}^2H^1(\Omega)} + \|G_1\|_{({}^2H^1(\Omega))'}).$$

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 $\mathcal{H}(\Omega) \neq {}^2H^1(\Omega)$.

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

$$\mathcal{H}(\Omega) = \left\{ (u_1, u_2) \in {}^2H^1(\Omega) \mid \sum_{l=1}^2 \alpha_{k,l}^{(j)} u_l \Big|_{\Gamma_j} = 0, \quad k = 1, 2, \quad j = 1, \dots, m \right\},$$

where $\alpha^j = \{\alpha_{k,l}^{(j)}\}$ are matrices of smooth functions on $\bar{\Gamma}_j$ (say $C^\infty(\bar{\Gamma}_j)$). Additional assumptions on $\{\alpha_{k,l}^{(j)}\}$ will be imposed later.

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

If α^j has rank 1, then we can write the constraint on Γ_j as

$$\alpha_{1,1}^{(j)} u_1 + \alpha_{1,2}^{(j)} u_2 = 0,$$

which will be written in the form

$$(2.4a) \quad \alpha^{(j)} u_1 + \beta^{(j)} u_2 = 0.$$

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

$$(2.4b) \quad \alpha^{(j)} u_1 + \beta^{(j)} u_2 = \alpha^{(j)} \rho_1 + \beta^{(j)} \rho_2.$$

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

$$(2.5) \quad B(u, v) = \int_Q \left(\frac{\partial u_1}{\partial x_1} \frac{\partial v_1}{\partial x_1} + \frac{\partial u_1}{\partial x_1} \frac{\partial v_1}{\partial x_2} + \frac{\partial u_2}{\partial x_2} \frac{\partial v_2}{\partial x_1} + \frac{\partial u_2}{\partial x_2} \frac{\partial v_2}{\partial x_2} + u_1 v_1 + u_2 v_2 \right) dx_1 dx_2$$

and

$$G_1(v) = \int_\Omega (f_1 v_1 + f_2 v_2) dx_1 dx_2, \quad f = (f_1, f_2) \in {}^2H^0(\Omega).$$

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

We will assume that $\alpha_{k,l}^{(j)} \in C^\infty(\bar{\Gamma}_j)$. In practice, we have the nonhomogeneous constraint problem defined so that

(i) If $\{\alpha_{k,l}^{(j)}\}$ has rank 2 on $\bar{\Gamma}_j$ then the constraint is

$$\sum_{l=1}^2 \alpha_{k,l}^{(j)} u_l = g_k^{(j)}, \quad k = 1, 2,$$

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

$$(2.6) \quad u_1 = \tilde{g}_1^{(j)}, \quad u_2 = \tilde{g}_2^{(j)}.$$

Because $\alpha_{k,i}^{(j)}$ are assumed to be smooth, we see that $\{\tilde{g}_i^{(j)}\}$, $i = 1, 2$, have the same smoothness as $\{g_i^{(j)}\}$.

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

$$\alpha^{(j)}u_1 + \beta^{(j)}u_2 = g^{(j)}.$$

We add the condition

$$\alpha^{(j)} + \beta^{(j)} \neq 0 \quad \text{on } \bar{\Gamma}_j.$$

This enables us to transform the constraint equation to

$$(2.7) \quad \alpha^{(j)}u_1 + \beta^{(j)}u_2 = \tilde{g}^{(j)} \quad \text{with } \alpha^{(j)} + \beta^{(j)} = 1.$$

As formulated above, $\tilde{g}^{(j)}$ are defined separately on each $\bar{\Gamma}_j$. We will assume that $\tilde{g}^{(j)}$ satisfy consistency conditions, namely that there exists $\rho = (\rho_1, \rho_2) \in {}^2H^1(\Omega)$ such that

$$(2.8a) \quad \tilde{g}_1^{(j)} = \rho_1|_{\bar{\Gamma}_j}, \quad \tilde{g}_2^{(j)} = \rho_2|_{\bar{\Gamma}_j},$$

respectively

$$(2.8b) \quad \alpha^{(j)}\rho_1|_{\bar{\Gamma}_j} + \beta^{(j)}\rho_2|_{\bar{\Gamma}_j} = \tilde{g}^{(j)}.$$

These conditions have to be imposed especially at the vertices of Ω .

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

2.3. The p -Version of the Finite Element Method. Assume that the domain Ω has been partitioned into a finite number of subdomains Ω_i , i.e., $\bar{\Omega} = \bigcup_{i=1}^n \bar{\Omega}_i$. We shall assume that Ω_i is the curvilinear quadrilateral

$$\Omega_i = \mathcal{F}_i(Q)$$

or curvilinear triangle

$$\Omega_i = \mathcal{F}_i(T),$$

where Q and T are the standard square and triangle, respectively. The domains Ω_i will be called elements. We will assume that \mathcal{F}_i^{-1} is a smooth one-to-one mapping of Ω_i onto Q , respectively T . It is obvious what the vertices and sides of Ω_i correspond to. If γ is a side of Ω_i , then \mathcal{F}_i induces mapping F_i of I onto γ (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 :

(α) If $\bar{\Omega}_i \cap \bar{\Omega}_j = R_{i,j} \neq \emptyset$, then $R_{i,j}$ is either a common vertex or a side of both Ω_i and Ω_j .

(β) If $R_{i,j} = \gamma_{i,j}$, then we will assume that the mappings F of I onto $\gamma_{i,j}$ induced by the mappings \mathcal{F}_i and \mathcal{F}_j are identical. We denote F by $F_{i,j}$. This implies the following: Let A, B , the vertices of Ω_i and Ω_j , be the end points of $\gamma_{i,j}$. Assume that (a_1, b_1) and (a_2, b_2) are the end points of the sides γ_k^Q or γ_l^T such that $\mathcal{F}_i(a_1) = \mathcal{F}_j(a_2) = A$, $\mathcal{F}_i(b_1) = \mathcal{F}_j(b_2) = B$. Then, if $C \in \gamma_{i,j}$ and $C = \mathcal{F}_i(c_1) = \mathcal{F}_j(c_2)$, $\overline{a_1c_1} = \overline{a_2c_2}$ and $\overline{c_1b_1} = \overline{c_2b_2}$.

Since we assumed that \mathcal{F}_i are smooth mappings, the vertices of Ω necessarily have to coincide with some of the vertices of Ω_i . We will further assume that for any Γ_j there is an element Ω_i such that one of its sides coincides with Γ_j . This assumption is made without any loss of generality.

Denote now

$${}^1\mathcal{P}_p(\Omega) = \left\{ u \in H^1(\Omega) \left| \begin{array}{l} u|_{\Omega_i}(\mathcal{F}_i(\xi)) \in \mathcal{P}_p^2(Q) \text{ if } \Omega_i \text{ is a quadrilateral} \\ \text{and } u|_{\Omega_i}(\mathcal{F}_i(\xi)) \in \mathcal{P}_p^1(T) \text{ if } \Omega_i \text{ is a triangle} \end{array} \right. \right\},$$

$${}^2\mathcal{P}_p(\Omega) = ({}^1\mathcal{P}_p(\Omega))^2,$$

$${}^2\mathcal{P}_p(\gamma) = \left\{ u \left| u_j|_{\gamma}(F_i(\xi)) \in \mathcal{P}_p(I), j = 1, 2 \right. \right\}.$$

Here, $\gamma = \Gamma_i$ or any side of an element. Let us define the constraint space $\mathcal{K} \mathcal{P}_p(\Omega) \subset {}^2\mathcal{P}_p(\Omega)$ as follows:

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

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

(ii) If Γ_j is a partial constraint side, then

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

and

$$\int_{\Gamma_j} (\alpha^{(j)} u_1 + \beta^{(j)} u_2) \psi ds = 0 \quad \text{for all } \psi \in {}^2\mathcal{P}_{p-2}(\Gamma_j).$$

The ρ -hyperplane $\mathcal{K}_\rho \mathcal{P}_p(\Omega)$ is defined analogously. Let g be defined in terms of ρ by (2.8). Then on Γ_j we impose

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

respectively

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

and

$$(2.9a) \quad \int_{\Gamma_j} u_i \psi_i ds = \int_{\Gamma_j} g_i^{(j)} \psi_i ds, \quad i = 1, 2,$$

respectively

$$(2.9b) \quad \int_{\Gamma_j} (\alpha^{(j)} u_1 + \beta^{(j)} u_2) \psi ds = \int_{\Gamma_j} g^{(j)} \psi ds.$$

The p -version is then defined analogously as before: Find $u_p \in \mathcal{K}_\rho \mathcal{P}_p(\Omega)$ such that

$$(2.10) \quad B(u_p, v) = G_1(v) \quad \forall v \in \mathcal{K} \mathcal{P}_p(\Omega).$$

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 $u_1 \cos \varphi + u_2 \sin \varphi = 0$, where φ 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 $\tilde{X}_p \subset X$, $\tilde{W}_p \subset W$, $p = 1, 2, \dots$, be one-parameter families of finite-dimensional subspaces. $X_p \subset X$, $W_p \subset W$ will denote corresponding families of hyperplanes such that $(u - v) \in \tilde{X}_p$ whenever $u, v \in X_p$, and $(\varphi - \psi) \in \tilde{W}_p$ whenever $\varphi, \psi \in W_p$.

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

Let $u_0 \in X$, $\varphi_0 \in W$ and $u_p \in X_p$, $\varphi_p \in W_p$ be such that

$$(3.1a) \quad a(u_0, v) + b(v, \varphi_0) = F_1(v) \quad \forall v \in \tilde{X}_p,$$

$$(3.1b) \quad b(u_0, \psi) = F_2(\psi) \quad \forall \psi \in \tilde{W}_p$$

and

$$(3.2a) \quad a(u_p, v) + b(v, \varphi_p) = F_1(v) \quad \forall v \in \tilde{X}_p,$$

$$(3.2b) \quad b(u_p, \psi) = F_2(\psi) \quad \forall \psi \in \tilde{W}_p.$$

Define $Z_p = \{v \in \tilde{X}_p, b(v, \psi) = 0 \quad \forall \psi \in \tilde{W}_p\} \subset \tilde{X}_p$. Then we have

THEOREM 3.1. *Let $a(u, u) \geq \gamma\|u\|_X^2$, $\gamma > 0$, for any $u \in Z_p$. Then*

$$(3.3) \quad \|u_0 - u_p\|_X \leq C \left[\begin{array}{c} \inf_{\substack{b(u_0 - w_p, \psi) = 0 \quad \forall \psi \in \tilde{W}_p \\ u_p - w_p \in \tilde{X}_p \\ (\text{i.e., } u_p - w_p \in Z_p)}} \|u_0 - w_p\|_X + \inf_{\chi_p \in W_p} \|\varphi_0 - \chi_p\|_W \end{array} \right].$$

Proof. For arbitrary $w_p \in X_p$ and $\chi_p \in W_p$ we have

$$(3.4a) \quad \begin{aligned} & a(u_p - w_p, v) + b(v, \varphi_p - \chi_p) \\ & = a(u_0 - w_p, v) + b(v, \varphi_0 - \chi_p) \quad \forall v \in \tilde{X}_p, \end{aligned}$$

$$(3.4b) \quad b(u_p - w_p, \psi) = b(u_0 - w_p, \psi) \quad \forall \psi \in \tilde{W}_p.$$

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

$$(3.5) \quad a(u_p - w_p, v) = a(u_0 - w_p, v) + b(v, \varphi_0 - \chi_p).$$

Suppose now that w_p is such that

$$b(w_p, \psi) = b(u_0, \psi) \quad \forall \psi \in \tilde{W}_p.$$

Then, by (3.4b),

$$b(u_p - w_p, \psi) = 0 \quad \forall \psi \in \tilde{W}_p$$

and hence

$$u_p - w_p \in Z_p.$$

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

$$a(u_p - w_p, u_p - w_p) \leq C[\|u_0 - w_p\|_X\|u_p - w_p\|_X + \|\varphi_0 - \chi_p\|_W\|u_p - w_p\|_X]$$

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

$$\|u_p - w_p\|_X \leq C[\|u_0 - w_p\|_X + \|\varphi_0 - \chi_p\|_W]$$

and hence also

$$\|u_0 - w_p\|_X \leq C[\|u_0 - w_p\|_X + \|\varphi_0 - \chi_p\|_W],$$

from which (3.3) follows. \square

3.2. *The Convergence of the p -Version.* Let $u_0 = (u_{0,1}, u_{0,2}) \in {}^2H^1(\Omega)$ be the solution of our constrained problem (2.2), (2.5) and $u_p = (u_{p,1}, u_{p,2}) \in {}^2\mathcal{P}_p(\Omega)$ be the approximation given by (2.10).

We will assume that $u_0 \in {}^2H^k(\Omega)$, $k > 3/2$. Hence $\partial u_0 / \partial n \in {}^2H^0(\Gamma_i)$, $i = 1, 2, \dots, m$. Let the constraints on Γ_j be as in Subsection 2.2, with Γ_{i_j} , $j = 1, \dots, m_1$, being the total constraint sides and Γ_{i_j} , $j = m_1 + 1, \dots, m$, the partial constraint sides. Then it can be verified that for any $v \in {}^2H^1(\Omega)$,

$$(3.6) \quad \begin{aligned} B(u_0, v) - \sum_{j=1}^{m_1} \int_{\Gamma_{i_j}} \left(\frac{\partial u_{0,1}}{\partial n} v_1 + \frac{\partial u_{0,2}}{\partial n} v_2 \right) ds \\ - \sum_{j=m_1+1}^m \int_{\Gamma_{i_j}} (\alpha^{i_j} v_1 + \beta^{i_j} v_2) \left(\frac{\partial u_{0,1}}{\partial n} + \frac{\partial u_{0,2}}{\partial n} \right) ds = G_1(v), \end{aligned}$$

where we have assumed $\alpha^{i_j} + \beta^{i_j} = 1$. Moreover, for ψ, ψ_1, ψ_2 and ρ as in (2.8)–(2.9),

$$(3.7) \quad \begin{aligned} & \sum_{j=1}^{m_1} \int_{\Gamma_{i_j}} (u_{p,1} \psi_1 + u_{p,2} \psi_2) ds + \sum_{j=m_1+1}^m \int_{\Gamma_{i_j}} (\alpha^{i_j} u_{p,1} + \beta^{i_j} u_{p,2}) \psi ds \\ &= \sum_{j=1}^{m_1} \int_{\Gamma_{i_j}} (u_{0,1} \psi_1 + u_{0,2} \psi_2) ds + \sum_{j=m_1+1}^m \int_{\Gamma_{i_j}} (\alpha^{i_j} u_{0,1} + \beta^{i_j} u_{0,2}) \psi ds \\ &= \sum_{j=1}^{m_1} \int_{\Gamma_{i_j}} (\rho_1 \psi_1 + \rho_2 \psi_2) ds + \sum_{j=m_1+1}^m \int_{\Gamma_{i_j}} (\alpha^{i_j} \rho_1 + \beta^{i_j} \rho_2) \psi ds \end{aligned}$$

and

$$(3.8a) \quad u_{p,k}(A_j) = u_{0,k}(A_j) = \rho_k(A_j), \quad j = l, l+1, k = 1, 2, \\ l = i_1, \dots, i_{m_1},$$

$$(3.8b) \quad (\alpha^l u_{p,1})(A_j) + (\beta^l u_{p,2})(A_j) = (\alpha^l u_{0,1})(A_j) + (\beta^l u_{0,2})(A_j) \\ = (\alpha^l \rho_1)(A_j) + (\beta^l \rho_2)(A_j), \quad j = l, l+1, l = i_{m_1+1}, \dots, i_m.$$

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

We now define

$$X = {}^2H^1(\Omega), \quad \|\cdot\|_X = \|\cdot\|_{{}^2H^1(\Omega)}$$

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

$$\begin{aligned} X_{p,\delta} = \{u = (u_1, u_2) \in {}^2\mathcal{P}_p(\Omega), u_k(A_j) = \delta_k(A_j), k = 1, 2, \\ j = l, l+1, l = i_1, \dots, i_{m_1}, (\alpha^l u_1)(A_j) + (\beta^l u_2)(A_j) \\ = (\alpha^l \delta_1)(A_j) + (\beta^l \delta_2)(A_j), j = l, l+1, l = i_{m_1+1}, \dots, i_m\}. \end{aligned}$$

We then take in our abstract framework

$$X_p = X_{p,\rho}, \quad \tilde{X}_p = X_{p,0},$$

where ρ satisfies (2.6)–(2.8). Moreover, let

$$W = \prod_{j=1}^{m_1} {}^2H^{-1/2}(\Gamma_{i_j}) \times \prod_{j=m_1+1}^m {}^1H^{-1/2}(\Gamma_{i_j})$$

with the norm

$$\|\psi\|_W = \left[\sum_{j=1}^{m_1} \|\psi_j\|_{{}^2H^{-1/2}(\Gamma_{i_j})}^2 + \sum_{j=m_1+1}^m \|\psi_j\|_{{}^1H^{-1/2}(\Gamma_{i_j})}^2 \right]^{1/2}.$$

We then see that $\varphi_0 \in W$, where (see (3.6))

$$\begin{aligned} \varphi_{0,j} &= - \left(\frac{\partial u_{0,1}}{\partial n} \Big|_{\Gamma_{i_j}}, \frac{\partial u_{0,2}}{\partial n} \Big|_{\Gamma_{i_j}} \right), \quad j = 1, \dots, m_1 \\ &= - \left(\frac{\partial u_{0,1}}{\partial n} + \frac{\partial u_{0,2}}{\partial n} \right) \Big|_{\Gamma_{i_j}}, \quad j = m_1 + 1, \dots, m. \end{aligned}$$

Define $W_p = \tilde{W}_p = \prod_{j=1}^m {}_jW_p \subset W$, where

$$\begin{aligned} {}_jW_p &= {}^2H^{-1/2}(\Gamma_{i_j}) \cap {}^2\mathcal{P}_{p-2}(\Gamma_j), \quad j = 1, \dots, m_1 \\ &= {}^1H^{-1/2}(\Gamma_{i_j}) \cap {}^1\mathcal{P}_{p-2}(\Gamma_j), \quad j = m_1 + 1, \dots, m. \end{aligned}$$

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

$$\begin{aligned} a(u, v) &= B(u, v), \\ b(u, \psi) &= \sum_{j=1}^{m_1} \int_{\Gamma_{i_j}} (u_1 \psi_1 + u_2 \psi_2) ds + \sum_{j=m_1+1}^m \int_{\Gamma_{i_j}} (\alpha^{i_j} u_1 + \beta^{i_j} u_2) \psi ds. \end{aligned}$$

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, φ_0) satisfy (3.1) with $F_k = G_k$, $k = 1, 2$. Moreover, if we can find a unique pair (u_p, φ_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, φ_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 $j = 1, 2, \dots, m$ we have

$$\left| \int_{\Gamma_{i_j}} u \psi ds \right| \leq C \|u\|_{H^{1/2}(\Gamma_{i_j})} \|\psi\|_{H^{-1/2}(\Gamma_{i_j})},$$

from which the continuity of $b(\cdot, \cdot)$ may be deduced. Hence Theorem 3.1 is applicable. Let us now estimate $\inf_{\chi \in W_p} \|\varphi_0 - \chi\|_W$. First, let $m_1 + 1 \leq j \leq m$. We assumed that $\varphi_{0,j} \in H^{k-3/2}(\Gamma_{i_j})$, $k > 3/2$. Hence, $\varphi_{0,j}(F_{i_j}(\xi)) = \varphi(\xi) \in H^{k-3/2}(I)$. Let $\sigma \in {}^1\mathcal{P}_{p-2}(I)$ be such that

$$(3.9) \quad \int_I \sigma \rho d\xi = \int_I \varphi \rho d\xi \quad \forall \rho \in {}^1\mathcal{P}_{p-2}(I).$$

Then, with $q = \varphi - \sigma$, we have

$$(3.10) \quad \|q\|_{H^0(I)} \leq Cp^{-(k-3/2)} \|\varphi\|_{H^{k-3/2}(I)}.$$

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

$$\frac{\int_I qv \, d\xi}{\|v\|_{H^1(I)}} = \frac{\int_I q(v - \sigma_1) \, d\xi}{\|v\|_{H^1(I)}} \leq \frac{\|q\|_{H^0(I)} \|v - \sigma_1\|_{H^0(I)}}{\|v\|_{H^1(I)}} \leq Cp^{-1} \|q\|_{H^0(I)},$$

where σ_1 is a polynomial of degree $p - 2$ satisfying

$$\|v - \sigma_1\|_{H^0(I)} \leq Cp^{-1} \|v\|_{H^1(I)}.$$

This yields

$$(3.11) \quad \|q\|_{H^{-1}(I)} \leq Cp^{-(k-1/2)} \|\varphi\|_{H^{k-3/2}(I)}.$$

Interpolating (3.10), (3.11) and using the fact that F_{i_j} is a smooth mapping, we obtain

$$\inf_{\chi_p \in \mathcal{W}_p} \|\varphi_{0,j} - \chi_p\|_{H^{-1/2}(\Gamma_{i_j})} \leq Cp^{-(k-1)} \|\varphi_{0,j}\|_{H^{k-3/2}(\Gamma_{i_j})}.$$

We get similar estimates for Γ_{i_j} , $j = 1, \dots, m_1$, so that

$$(3.12) \quad \inf_{\chi_p \in \mathcal{W}_p} \|\varphi_0 - \chi_p\|_W \leq Cp^{-(k-1)} \|u_0\|_{2H^k(\Omega)}.$$

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

$$(3.13a) \quad \|u_{0,i} - z_i\|_{H^t(\Omega)} \leq Cp^{-(k-t)} \|u_{0,i}\|_{H^k(\Omega)}, \quad i = 1, 2, t = 0, 1,$$

$$(3.13b) \quad u_{0,i}(N) = z_i(N) \quad \text{for each node } N \text{ of the mesh,}$$

$$(3.13c) \quad \|u_{0,i} - z_i\|_{H^t(\Gamma_j)} \leq Cp^{-(k-1/2-t)} \|u_{0,i}\|_{H^k(\Omega)}, \quad t = 0, 1, i = 1, 2, \\ j = 1, 2, \dots, m.$$

Let $m_1 + 1 \leq j \leq m$. Let us denote $\kappa = \alpha^{i_j}(u_{0,1} - z_1) + \beta^{i_j}(u_{0,2} - z_2)$. Then we have $\kappa(A_{i_j}) = \kappa(A_{i_j+1}) = 0$. Let $\tilde{\kappa}(\xi) = \kappa(F(\xi))$ and let $\tau(F(\xi)) = \tilde{\tau}(\xi) \in {}^1\mathcal{P}_p(I)$ satisfy $\tilde{\tau}(\pm 1) = 0$ and

$$\int_I \tilde{\tau}v \, d\xi = \int_I \tilde{\kappa}v \, d\xi \quad \forall v \in {}^1\mathcal{P}_{p-2}(I).$$

Because of (3.13b) we can write

$$\int_I \tilde{\tau}'\omega' \, d\xi = \int_I \tilde{\kappa}'\omega' \, d\xi \quad \forall \omega \in {}^1\mathcal{P}_p(I), \quad \omega(\pm 1) = 0,$$

and hence by Lemma 3.2 of [2]

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

Using (3.13c), this gives

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

Now using Lemma 2.1, it follows that there is a $w \in {}^1\mathcal{P}_p(\Omega)$ such that $w = 0$ on $\Omega - \tilde{\Omega}$, where $\tilde{\Omega}$ is the element with the side Γ_{i_j} , $w = \tau$ on Γ_{i_j} , $w = 0$ on $\partial\tilde{\Omega} - \Gamma_{i_j}$ and

$$(3.14) \quad \|w\|_{H^1(\Omega)} \leq Cp^{-(k-1)} \|u_0\|_{2H^k(\Omega)}.$$

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

Let now $w_p = z + w_{i_j} = (z_1, z_2) + (w, w)$. Using (3.13b) and the fact that $w(A_l) = 0$, we obtain

$$(\alpha^{i_j}(u_{0,1} - w_{p,1}) + \beta^{i_j}(u_{0,2} - w_{p,2}))(A_l) = 0, \quad l = i_j, i_j + 1.$$

Moreover, for $\psi \in {}^1\mathcal{P}_{p-2}(\Gamma_{i_j})$,

$$\begin{aligned} & \int_{\Gamma_{i_j}} (\alpha^{i_j}(u_{0,1} - w_{p,1}) + \beta^{i_j}(u_{0,2} - w_{p,2}))\psi \, ds \\ &= \int_{\Gamma_{i_j}} (\kappa - (\alpha^{i_j} + \beta^{i_j})w)\psi \, ds = \int_{\Gamma_{i_j}} (\kappa - w)\psi \, ds = 0, \end{aligned}$$

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

$$w_p = z + \sum_{j=1}^m w_{i_j},$$

we see that

$$\begin{aligned} u_p - w_p &\in \tilde{X}_p, \\ b(u_0 - w_p, \psi) &= 0 \quad \forall \psi \in W_p \end{aligned}$$

and

$$\|u_0 - w_p\|_X \leq \|u_0 - z\|_X + \sum_{j=1}^m \|w_{i_j}\|_X \leq Cp^{-(k-1)}\|u_0\|_{2H^k(\Omega)}.$$

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

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

$$\|u_0 - u_p\|_{2H^1(\Omega)} \leq Cp^{-(k-1)}\|u_0\|_{2H^k(\Omega)},$$

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, φ_0) and (u_p, φ_p) .

THEOREM 3.3. *The (exact) solution (u_0, φ_0) of the constrained problem exists. The finite element solution (u_p, φ_p) exists and is unique.*

Proof. In Subsection 2.2 we have shown that u_0 exists and hence (u_0, φ_0) exists, too. The finite element solution (u_p, φ_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, φ_p) of the trivial problem. Obviously $\bar{u} = \bar{\varphi} = 0$ is also a solution of this problem. Hence $u_p = 0$ because of Theorem 3.1. We have to show, therefore, that

$$\int_I (\alpha v_1 + \beta v_2)\varphi_p \, d\xi = 0$$

implies $\varphi_p = 0$. Because $\alpha + \beta = 1$ we also have $\int_I v \varphi_p d\xi = 0$ for all $v \in {}^1\mathcal{P}_p(I) \cap H_0^1(I)$ while $\varphi_p \in {}^1\mathcal{P}_{p-2}(I)$. This leads to $\varphi_p = 0$, which leads to the desired result. \square

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 nodal 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 γ are of the form

$$\xi_j = \int_{-1}^x l_j(\xi) d\xi, \quad j = 1, 2, \dots,$$

where l_j is the Legendre polynomial of degree j . ξ_j is then a polynomial of degree $j + 1$;

(c) The internal shape functions which are zero on ∂Q (respectively ∂T).

The stiffness matrices are first computed in the standard way without constraints. Then the constraints are imposed at the vertices A_j . This only involves the amplitudes for the nodal shape functions. Then the conditions (2.9a,b) only involve amplitudes for the side shape functions. The functions ψ in (2.9a,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 Legendre polynomials, we get the amplitudes for the side shape functions on the total constraint sides directly.

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

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

1. I. BABUŠKA, *The p and h - p Versions of the Finite Element Method, The State of the Art*, Technical Note BN-1156, Institute for Physical Science and Technology, University of Maryland, 1986.

2. I. BABUŠKA & M. SURI, "The optimal convergence rate of the p -version of the finite element method," *SIAM J. Numer. Anal.*, v. 24, 1987, pp. 750-776.

3. I. BABUŠKA & M. SURI, "The h - p version of the finite element method with quasiuniform meshes," *RAIRO Modél. Math. Anal. Numér.*, v. 21, 1987, pp. 199-238.

4. I. BABUŠKA & M. SURI, *The Treatment of Nonhomogeneous Dirichlet Boundary Conditions by the p -Version of the Finite Element Method*, Technical Note BN-1063, Institute for Physical Science and Technology, University of Maryland, 1987.

5. J. L. LIONS & E. MAGENES, *Non-Homogeneous Boundary Value Problems and Applications-I*, Springer-Verlag, New York, Heidelberg, Berlin, 1972.

6. B. A. SZABÓ, *PROBE: Theoretical Manual*, Noetic Technologies Corporation, St. Louis, Missouri, 1985.

*** The code PROBE is the code of Noetic Tech., St Louis.