MATHEMATICS OF COMPUTATION
VOLUME 44, NUMBER 169
JANUARY 1985, PAGES 53-70

Least Squares Methods for Elliptic Systems*

By A. K. Aziz, R. B. Kellogg and A. B. Stephens

Abstract. A weighted least squares method is given for the numerical solution of elliptic partial differential equations of Agmon-Douglis-Nirenberg type and an error analysis is provided. Some examples are given.

1. Introduction. The use of least squares methods for the approximate solution of equations dates back at least to Gauss. The modern theory of least squares methods in the numerical solution of elliptic boundary value problems starts, in 1970, with the papers of Bramble and Schatz [5], [6]. This work uses a finite-dimensional space S of approximating functions, similar to the spaces used in finite-element methods. The approximate solution is defined to be the minimizer of a least squares functional that is a weighted sum of the least squares residual in the differential equation and the least squares residual in the boundary condition. The paper [5] has an historical importance for the following reason. It appeared during the time when numerical analysts were shifting attention from finite-difference methods to finite-element methods, and it provided, for the first time, a family of approximation methods for the solution of the Dirichlet problem whose order of accuracy could be made arbitrarily large. The paper [6] provided an extension to an elliptic equation of order $2m$, and [3] gave important simplifications in the analysis. The principal advantages of the method are that one need not satisfy exactly the Dirichlet boundary conditions, and that the mathematical analysis dictates, in a natural way, the relative weights that are given to the boundary and interior terms in the least squares functional. Also, the method provides, in a quasioptimal sense, as good a solution as can be expected from the space S. On the other hand, the method requires that S consist of functions which are smooth enough to lie in the domain of the elliptic operator. Also, the method seems to produce matrices with large condition number.

For various reasons, it is of interest to extend the theory of least squares methods to include elliptic systems. First, if a second-order elliptic equation is written as a first-order system, it would seem (and this is borne out by our analysis) that the smoothness requirements for the spaces of approximating functions would be reduced, thus eliminating one of the disadvantages of the method. A second motivation for extending the least squares method to elliptic systems is that elliptic systems occur frequently in applications. An example of an elliptic system is the system of equations for Stokes flow. For this system, the least squares method does not require the space of approximating vector fields to be incompressible. Instead,

Received June 24, 1983.
1980 *Mathematics Subject Classification*. Primary 65N30, 35J55.
* This work is supported by the NSWC Independent Research Fund.

0025-5718/85 $1.00 + $.25 per page

the incompressibility condition is considered as one of the equations in the system, and the analysis provides, in a natural way, weights to put on the residual in the incompressibility equation. The difficulties associated with finding approximating spaces of incompressible vector fields are well-known; the least squares method provides an alternate way of treating these difficulties. Finally, it is desirable to extend the least squares method to elliptic systems to close the gap in the theory of the method.

Some work in least squares methods for elliptic systems has appeared in the literature. In [9], [11], a least squares method is formulated for the first-order system in three unknowns that is associated with a single second-order elliptic equation in the plane. The system is discussed more in Section 5. A theory of least squares methods for elliptic systems of Petrovsky type is developed in [15], and quasioptimal error estimates are obtained for the approximate solution. Petrovsky systems are an important subclass of the class of elliptic systems, in which the different equations and unknowns appearing in the system have taken the same "differentiability order". In the least squares method, for these systems developed in [15], the residual for each of the differential equations in the elliptic system receives the same weight in the least squares functional. Finally, least squares methods have recently been applied to fluid flow problems of mixed type, and to problems whose solutions contain singularities [10].

In this paper there is developed a least squares method for the approximate solution of elliptic boundary value problems of Agmon-Douglis-Nirenberg type (ADN). The method involves the minimization of a least squares functional that consists of a weighted sum of the residuals occurring in the equations and the boundary conditions of the system. The weights occurring in the least squares functional are determined by the indices that enter into the definition of an ADN boundary value problem. A quasioptimal error estimate is obtained for the approximate solution generated by the method. The method reduces to the method of [5], if the system is a single equation and to the method of [15] if the system is an elliptic system of Petrovsky type. Our error analysis assumes that the boundary value problem is uniquely solvable, and that the usual a priori estimate for the solution in terms of the data holds over a range of negative regularity indices (see (2.7)). The verification of this assumption for solvable elliptic boundary value problems seems to involve technical difficulties concerning the ellipticity of the adjoint boundary value problem (see, e.g., [13]). Therefore, we have made the required inequality a hypothesis of our theorem, and we have verified this inequality in a number of examples of particular interest.

Section 2 sets the notation and presents the salient facts concerning ADN systems. Section 3 formulates the least squares method, and Section 4 gives the error analysis of the method. Section 5 shows how the method applies to several elliptic systems occurring in practice. This section concludes with a "nonconforming" version of the method. The error analysis for this version has not been done. Finally, Section 6 contains an estimate for the condition number of the matrix associated with a least squares method.

2. The Boundary Value Problem. Let $\Omega \subset R^n$ be a bounded domain with a smooth boundary Γ. We are concerned with elliptic systems of Agmon-Douglis-Nirenberg

(ADN) type. These are linear systems of N partial differential equations in N unknowns, which we write

$$\sum_{j=1}^{N} L_{ij}(x, D)u_j(x) = f_i(x), \qquad x \in \Omega, 1 \leqslant i \leqslant N. \tag{2.1}$$

Here $L_{ij}(x, D)$ is a polynomial in $D = (D_1, \ldots, D_n)$, $D_i = \partial/\partial x_i$, with coefficients which depend smoothly on x. We shall suppose that there are integers s_i, the "equation indices", and t_j, the "unknown indices", such that

$$\begin{cases} L_{ij}(x, D) \equiv 0 \quad \text{if } s_i + t_j < 0, \\ \lambda_{ij} \equiv \deg L_{ij}(x, D) \leqslant s_i + t_j, \\ s_i \leqslant 0, \quad \sum_{1}^{N} (s_i + t_j) = 2m > 0. \end{cases} \tag{2.2}$$

Together with the system (2.1), we consider m boundary conditions which we write

$$\sum_{j=1}^{N} B_{kj}(x, D)u_j(x) = g_k(x), \qquad x \in \Gamma, 1 \leqslant k \leqslant m, \tag{2.3}$$

where $B_{kj}(x, D)$ is a polynomial in D. We shall suppose that there are integers r_k the "boundary condition indices", such that

$$\begin{cases} B_{kj}(x, D) \equiv 0 \quad \text{if } r_k + t_j < 0, \\ \beta_{kj} = \deg B_{kj}(x, D) \leqslant r_k + t_j. \end{cases} \tag{2.4}$$

In addition to (2.2), (2.4) we require that the operators appearing in (2.1), (2.3) satisfy the ellipticity condition, the supplementary condition, and the complementary boundary condition, as specified in [1]. We shall not state these conditions here as they are somewhat complicated and are not explicitly needed in the sequel. What we shall need in the sequel, and will state explicitly in Theorem 2.1, are the a priori estimates associated with these operators. These a priori estimates follow from the above three conditions, and in fact, are known to be equivalent to them [1].

We require some Hilbert-Sobolev spaces on Ω and Γ. We let $C^\infty(\Omega)$ denote the functions on $\overline{\Omega}$ which are restrictions of functions on R^n all of whose derivatives exist, and we recall that $C^\infty(\overline{\Omega})$ is dense in $H^s(\Omega)$. For $s \in R$, let $H^s(\Omega)$ denote the usual Sobolev space of functions on Ω, with norm $\|u\|_s$ and inner product $(u, v)_s$. For $s > 0$ an integer, $\|u\|_s^2 = \sum_{|\alpha| \leqslant s} \|D^\alpha u\|_0^2$. For $s > 0$ not an integer, $H^s(\Omega)$ is defined by interpolation. For $s > 0$ we define

$$\|u\|_{-s} = \sup_{v \in C^\infty(\overline{\Omega})} \frac{(u, v)_0}{\|v\|_s},$$

and we define $H^{-s}(\Omega)$ to be the closure of functions in $C^\infty(\overline{\Omega})$ with respect to this norm. The spaces $H^s(\Gamma)$, $s \in R$, with norm $|u|_s$ and inner product $\langle u, v \rangle_s$, are defined in a similar way. If $s = 0$ we drop the subscripts. We recall that the families $H^s(\Omega)$ and $H^s(\Gamma)$, $-\infty < s < \infty$, each form an interpolating family of Hilbert spaces. The two families are connected by the trace inequality: if $s > \frac{1}{2}$ and $u \in H^s(\Omega)$, then the restriction of u to Γ has a meaning and this restriction, which

we also denote by u, satisfies

$$|u|_{s-1/2} \leqslant c\|u\|_s. \tag{2.5}$$

With these spaces we now state

THEOREM 2.1. *If the problem* (2.1), (2.3) *satisfies the ellipticity, supplementary, and covering conditions, and if* $l \geqslant 0$, *there is a* $c > 0$ *such that if* $u_j \in H^{l+t_j}(\Omega)$, $1 \leqslant j \leqslant N$, *then*

$$\sum_1^N \|u_j\|_{l+t_j} \leqslant c\sum_1^N \|f_i\|_{l-s_i} + c\sum_1^N |g_k|_{l-r_k-1/2} + c\sum_1^N \|u_j\|. \tag{2.6}$$

The proof of Theorem 2.1 is contained in [1]. We shall require some additional hypotheses concerning the problem (2.1), (2.3). The first condition is that the problem has a unique solution for all smooth data f_i and g_k. This condition enables the L_2-norms of u_j on the right side of (2.6) to be eliminated. The second condition is that the modified form of (2.6) be valid for $l < 0$. The verification of this condition for general ADN systems seems to involve technical difficulties concerning the existence of an adjoint elliptic boundary value problem [13]. We will verify the modified inequality in a number of examples of particular interest. Summarizing our additional hypotheses, in addition to the unique solvability of (2.1), (2.3), we shall assume that for each real l there is a $c > 0$ such that if $\{u_j\}$ are a collection of smooth functions on Ω, and if $\{f_i\}$ and $\{g_k\}$ are defined by (2.1) and (2.3), then

$$\sum_j \|u_j\|_{l+t_j} \leqslant c\sum_i \|f_i\|_{l-s_i} + c\sum_k |g_k|_{l-r_k-1/2}. \tag{2.7}$$

We give some examples of elliptic systems in R^2 to illustrate the ideas. First, let $N = 1$, $m = 1$, and consider the single elliptic equation

$$L_1 u \equiv -\Delta u + u = f \quad \text{in } \Omega \tag{2.8a}$$

with the single boundary condition

$$B_{11}u \equiv u = g \quad \text{on } \Gamma. \tag{2.8b}$$

We define indices

$$\lambda_{11} = 2, \quad \beta_{11} = 0, \quad t_1 = 2, \quad s_1 = 0, \quad r_1 = -2. \tag{2.8c}$$

With this choice of indices, (2.2) is satisfied. It is known that the problem (2.8a,b) has a unique solution u for each $f \in H^l(\Omega)$, $g \in H^{l+3/2}(\Gamma)$, $l \geqslant 0$, and that u satisfies

$$\|u\|_{l+2} \leqslant c\|f\|_l + c|g|_{l+3/2}. \tag{2.9}$$

We now verify this fact for $l < 0$. Let $\alpha \leqslant 0$, and let $u \in C^\infty(\bar{\Omega})$. Since $C^\infty(\bar{\Omega})$ is dense in $H^\alpha(\Omega)$, there is an $h \in C^\infty(\bar{\Omega})$ such that

$$\|u\|_\alpha = \sup_\psi \frac{(u, \psi)}{\|\psi\|_{-\alpha}} \leqslant 2\frac{(u, h)}{\|h\|_{-\alpha}}.$$

Let ϕ be the solution of the problem $L_1\phi = h$ in Ω with $B_{11}\phi = 0$ on Γ. Then from Green's second identity,

$$\begin{aligned}(u, h) &= (L_1 u, \phi) - \left\langle u, \frac{\partial \phi}{\partial n} \right\rangle \\ &\leqslant \|L_1 u\|_{\alpha-2}\|\phi\|_{-\alpha+2} + |u|_{\alpha-1/2}\left|\frac{\partial \phi}{\partial n}\right|_{-\alpha+1/2}.\end{aligned}$$

Using (2.5), and (2.8) with u replaced by h, we get

$$(u, h) \leqslant c\|L_1 u\|_{\alpha-2}\|h\|_{-\alpha} + c\|B_{11}u\|_{\alpha-1/2}\|h\|_{-\alpha}.$$

Hence, we obtain (2.9) with $l = \alpha - 2 \leqslant -2$. From (2.9) we see that the map T: $\{f, g\} \to u$ is a bounded operator from $H^l(\Omega) \times H^{l+3/2}(\Gamma) \to H^{l+2}(\Omega)$, for $l \geqslant 0$, and for $l \leqslant -2$. By interpolation, we find that T is a bounded operator in the intermediate range of l, so (2.9) holds for each $l \in (-\infty, \infty)$.

For the next example we set $N = 3$, $m = 1$, and we consider in R^2 the first-order system

$$\text{(2.10a)} \qquad \begin{cases} L_1 u \equiv u_{1,x} - u_2 = f_1 \\ L_2 u \equiv u_{1,y} - u_3 = f_2 \\ L_3 u \equiv -u_1 + u_{2,x} + u_{3,y} = f_3 \end{cases} \quad \text{in } \Omega,$$

with the single boundary condition

$$\text{(2.10b)} \qquad B_1 u \equiv u_1 = g \quad \text{on } \Gamma.$$

We define the indices of the problem by the equations

$$\text{(2.10c)} \qquad [\lambda_{ij}] = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \qquad [\beta_{1j}] = [0,0,0],$$

$$[s_i] = [-1, -1, 0], \quad [t_j] = [2, 1, 1], \quad r_1 = -2.$$

With this choice of indices, (2.2) is satisfied. It is known that (2.10a, b) satisfies all the conditions of an ADN elliptic boundary value problem. Note that (2.10a) gives $-\Delta u_1 + u_1 = f_3 + f_{1,x} + f_{2,y}$. The problem (2.10) is, basically, the problem (2.8), written as a first-order system. Also, from (2.9) we obtain, if $u_j \in C^\infty(\bar{\Omega}), j = 1, 2, 3$,

$$\|u_1\|_{l+2} \leqslant c \sum_j \|f_i\|_{l-s_i} + c|g|_{l+3/2}.$$

From the first two equations of (2.10a) we then obtain

$$\|u_j\|_{l+1} \leqslant c \sum_i \|f_i\|_{l-s_i} + c|g|_{l+3/2}, \qquad j = 2, 3,$$

so (2.7) has been verified in this case. Alternately, it is possible to assume (2.7) for $l \geqslant 0$ and prove (2.7) for $l \leqslant -2$, by using Green's identities and by introducing an auxiliary boundary value problem to estimate the negative norms.

For our third example, we set $N = 3$, $m = 2$, and consider the Stokes system in R^2,

$$\text{(2.11a)} \qquad \begin{cases} L_1 u \equiv -\Delta u_1 + u_{3x} = f_1 \\ L_2 u \equiv -\Delta u_2 + u_{3y} = f_2 \\ L_3 u \equiv u_{1x} + u_{2y} = f_3 \end{cases} \quad \text{in } \Omega,$$

with the boundary conditions

$$\text{(2.11b)} \qquad \begin{cases} B_1 u \equiv u_1 = g_1 \\ B_2 u \equiv u_2 = g_2 \end{cases} \quad \text{on } \Gamma.$$

The variables (u_1, u_2) represent a velocity field, and u_3 represents pressure. If $f_3 = 0$, (2.11a) are the equations of motion of a steady-state incompressible flow, in which the inertial terms have been neglected. The indices of the problem are

$$[\lambda_{ij}] = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \qquad [\beta_{kj}] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{2.11c}$$

$$[s_i] = [0,0,-1], \quad [t_j] = [2,2,1], \quad [r_k] = [-2,-2].$$

It is known [14] that if $\{f_i\}$ and $\{g_k\}$ are smooth functions which satisfy the compatibility condition

$$\int_\Gamma \underline{n} \cdot \underline{g}\, d\Gamma = \iint_\Omega f_3\, dx\, dy, \tag{2.12}$$

then the problem (2.11) has a unique solution $\{u_j\}$ which satisfies

$$\int_\Omega u_3\, dx\, dy = 0. \tag{2.13}$$

Furthermore, this solution satisfies the a priori inequality, for $l \geqslant 0$,

$$\begin{aligned} &\|u_1\|_{l+2} + \|u_2\|_{l+2} + \|u_3\|_{l+1} \\ &\quad \leqslant c\left\{\|f_1\|_l + \|f_2\|_l + \|f_3\|_{l+1} + |g_1|_{l+3/2} + |g_2|_{l+3/2}\right\}. \end{aligned} \tag{2.14}$$

The inequality (2.14) is the same as the inequality (2.7) for the problem (2.11).

For the Stokes problem, it is convenient to use Sobolev spaces of functions with mean 0. Let $\lambda(\phi) = (\phi, 1)$. For s a nonnegative integer, $|\lambda(\phi)| \leqslant c\|\phi\|_s$, and by interpolation, this inequality holds for all $s \geqslant 0$. Since $|(\phi,1)| \leqslant \|1\|_s \|\phi\|_{-s}$, the inequality also holds for $s < 0$. Hence $\lambda(\phi)$ is a bounded linear functional on $H^s(\Omega)$ for all real s. Let

$$\dot{H}^s(\Omega) = \{\phi \in H^s(\Omega) : \lambda(\phi) = 0\}.$$

Then $\dot{H}^s(\Omega)$ is a closed subspace of $H^s(\Omega)$ of codimension one, the collection of Hilbert spaces $\{\dot{H}^s(\Omega)\}$ forms an interpolating family, and for $s < 0$ and $\phi \in \dot{H}^s(\Omega)$,

$$\|\phi\|_s = \sup\left\{\frac{(\phi, \psi)}{\|\psi\|_{-s}} : \psi \in \dot{H}^{-s}(\Omega)\right\}. \tag{2.15}$$

(See [12].)

We now prove the inequality (2.14) for $l < 0$. Let $\alpha \leqslant 0$, and let $u_1, u_2, u_3 \in C^\infty(\bar{\Omega})$, with $(u_3, 1) = 0$. Pick $h_1, h_2, h_3 \in C^\infty(\bar{\Omega})$, with $(h_3, 1) = 0$, and let $\underline{\phi} = (\phi_1, \phi_2, -\phi_3)$ be the unique solution of the Stokes problem

$$\begin{aligned} &L_i \underline{\phi} = h_i, \qquad i = 1,2,3, \\ &(\phi_3, 1) = 0, \\ &B_k \underline{\phi} = 0, \qquad k = 1,2. \end{aligned}$$

Setting $I = (u_1, h_1) + (u_2, h_2) - (u_3, h_3)$, a computation gives

$$I = \sum_1^2 (f_j, \phi_j) - (f_3, \phi_3) + \sum_1^2 \left[-\left\langle g_j, \frac{\partial \phi_j}{\partial n} \right\rangle + \langle n_j g_j, \phi_3 \rangle \right].$$

We apply (2.14) with $l = -\alpha$, and with $\{u_j\}$ replaced by $\{\phi_j\}$, to obtain

$$\begin{aligned} I &\leqslant \sum_1^2 \|f_j\|_{\alpha-2}\|\phi_j\|_{-\alpha+2} + \|f_3\|_{\alpha-1}\|\phi_3\|_{-\alpha+1} \\ &\quad + c\sum_1^2 \left[|g_j|_{\alpha-1/2}\|\phi_j\|_{-\alpha+2} + |g_j|_{\alpha-1/2}\|\phi_3\|_{-\alpha+1}\right] \\ &\leqslant c\left[\sum_1^2 \|f_j\|_{\alpha-2} + \|f_3\|_{\alpha-1} + \sum_1^2 |g_j|_{\alpha-1/2}\right] \\ &\quad \cdot \left[\sum_1^2 \|h_j\|_{-\alpha} + \|h_3\|_{-\alpha+1}\right]. \end{aligned}$$

To use this inequality, we set $h_2 = h_3 = 0$, and we obtain

$$\|u_1\|_\alpha \leqslant 2\frac{(u_1, h_1)}{\|h_1\|_{-\alpha}} \leqslant c\left[\sum_1^2 \|f_j\|_{\alpha-2} + \|f_3\|_{\alpha-1}\right].$$

Let h_1 be chosen so that

$$\|u_1\|_\alpha \leqslant 2\frac{(u_1, h_1)}{\|h_1\|_{-\alpha}}.$$

The inequality then gives

$$\tfrac{1}{2}\|u_1\|_\alpha\|h_1\|_{-\alpha} \leqslant I \leqslant c\left[\sum_1^2 \|f_j\|_{\alpha-2} + \|f_3\|_{\alpha-1} + \sum_1^2 |g_j|_{\alpha-1/2}\right]\cdot\|h_1\|_{-\alpha}.$$

Dividing both sides by $\|h_1\|_{-\alpha}$ gives an inequality for $\|u_1\|_\alpha$. Inequalities for $\|u_2\|_\alpha$ and $\|u_3\|_{\alpha-1}$ are obtained in the same way. This proves (2.14) with $l = \alpha - 2 \leqslant -2$. The inequality for $l \in [-2, 0]$ then follows by interpolation.

3. The Least Squares Method. In this section we define our least squares method, and we discuss some requirements that are needed by our subspaces. We consider an elliptic boundary value problem, (2.1), (2.3), with the associated collections of indices. We define

$$\begin{aligned} \bar{\mu} &= \max\{s_i, r_k + \tfrac{1}{2}: 1 \leqslant i \leqslant N, 1 \leqslant k \leqslant m\}, \\ \underline{\mu} &= \min\{s_i, r_k + \tfrac{1}{2}: 1 \leqslant i \leqslant N, 1 \leqslant k \leqslant m\}, \\ \alpha_j &= \text{smallest integer} \geqslant \{\lambda_{ij}, \beta_{kj} + \tfrac{1}{2}, 1 \leqslant i \leqslant N, 1 \leqslant k \leqslant m\}. \end{aligned}$$

We will use finite-dimensional subspaces S_h of functions to approximate our solution. The parameter h, which represents a mesh spacing, is used to indicate the approximation property of S_h. Let α and β be integers with $\alpha \leqslant \beta$. We say that S_h *approximates optimally* with respect to (β, α) if $S_h \subset H^\alpha(\Omega)$, and if, for each $u \in H^\beta(\Omega)$, there is a $v \in S_h$ such that

$$\sum_{i=s}^{\alpha} h^i\|u - v\|_i \leqslant ch^\beta\|u\|_\beta, \tag{3.1}$$

Here, s is an integer (positive or negative) which is $\leqslant \alpha - 1$. From a theorem of Bramble and Scott [7], s may be chosen as small as desired, if the boundary Γ is smooth enough.

Since we are dealing with systems, we must consider collections of subspaces. Let $S_{hj} \subset H^{\sigma_j}(\Omega)$, and let $\underline{S}_h = S_{h1} \times \cdots \times S_{hN}$. Let $\underline{y} = (y_j)$, $\underline{z} = (z_j)$, with $y_j, z_j \in H^{\alpha_j}(\Omega)$. We define the bilinear form

$$(\underline{y}, \underline{z})_A = \sum_{i=1}^{N} h^{2s_i} \left(\sum_{j=1}^{N} L_{ij} y_j, \sum_{j=1}^{N} L_{ij} z_j \right) + \sum_{k=1}^{m} h^{2r_k+1} \left\langle \sum_{j=1}^{N} B_{kj} y_j, \sum_{j=1}^{N} B_{kj} z_j \right\rangle, \tag{3.2}$$

and we let $\|\underline{y}\|_A$ denote the corresponding norm. If $\underline{u} = (u_j)$, $u_j \in H^{\alpha_j}(\Omega)$, is the solution of (2.1), (2.3), we define the *least squares* approximation to be the function $\underline{u}_h \in \underline{S}_h$ which satisfies

$$\|\underline{u} - \underline{u}_h\|_A \leqslant \|\underline{u} - \underline{w}\|_A, \qquad \underline{w} \in \underline{S}_h. \tag{3.3}$$

An equivalent formulation of the least squares problem is: $\underline{u}_h$ is that function in $\underline{S}_h$ which minimizes the expression, for $\underline{v} \in \underline{S}_h$,

$$\sum_{i=1}^{N} h^{2s_i} \left\| \sum_j L_{ij} v_j - f_i \right\|^2 + \sum_{k=1}^{m} h^{2r_k+1} \left| \sum_k B_{kj} v_j - g_k \right|^2. \tag{3.4}$$

The expression (3.4) is a weighted L_2-norm of the residual. As we will see, the weights have come from the theory of ADN systems outlined in Section 2. The calculation of $\underline{u}_h$ requires the solution of a symmetric, positive definite linear system of equations. The coefficients of the linear system involve the bilinear form (3.2) applied to basis elements of $\underline{S}_h$. The right-hand side of the linear system involves the right-hand sides, f_i, g_k, of the system (2.1), (2.3).

It is easily seen that if $\underline{e} = \underline{u} - \underline{u}_h$ is the error in the least squares solution, then $\underline{e}$ satisfies the orthogonality property

$$(\underline{e}, \underline{w})_A = 0, \qquad \underline{w} \in \underline{S}_h. \tag{3.5}$$

This formula, which will be used in our error estimates, serves to characterize the least squares approximation.

4. Error Estimates. In this section we state and prove our main result, an optimal error estimate for our least squares approximation. Throughout the section, we suppose that the elliptic boundary value problem is uniquely solvable, for smooth right-hand sides, and that (2.7) holds for all real s. The functions $\{u_j\}$ solve (2.1), (2.3) and, for some $\mu \geqslant \bar{\mu}$, $u_j \in H^{\mu+t_j}(\Omega)$, $j = 1, \ldots, N$. We first prove two lemmas.

LEMMA 4.1. *Suppose S_{hj} approximates optimally with respect to $(\mu + t_j, \alpha_j)$, $j = 1, \ldots, N$. Then*

$$\|\underline{e}\|_A \leqslant ch^{\mu} \sum_{j=1}^{N} \|u_j\|_{\mu+t_j}.$$

Proof. For any $\underline{v} \in \underline{S}_h$ it follows that

$$\|\underline{e}\|_A = \|\underline{u} - \underline{u}_h\|_A \leqslant \|\underline{u} - \underline{v}\|_A \leqslant \sum_{i=1}^{N} h^{s_i} \left\| \sum_{j=1}^{N} L_{ij}(u_j - v_j) \right\| + \sum_{k=1}^{m} h^{r_k+1/2} \left| \sum_{j=1}^{N} B_{kj}(u_j - v_j) \right|,$$

so

$$\|\underline{e}\|_A \leqslant c\sum_{i,j} h^{s_i}\|u_j - v_j\|_{\lambda_{ij}} + c\sum_{k,j} h^{r_k+1/2} \sum_{|\beta|\leqslant\beta_{kj}} \left|D^\beta(u_j - v_j)\right|. \tag{4.1}$$

Using the approximation properties of S_{hj}, and using the fact that $\lambda_{ij} \leqslant s_i + t_j$, we choose the v_j so that

$$\|u_j - v_j\|_{\lambda_{ij}} \leqslant ch^{\mu - s_i}\|u_j\|_{\mu + t_j}. \tag{4.2}$$

Using the inequality $|z| \leqslant c(\varepsilon\|z\|_1 + \varepsilon^{-1}\|z\|)$, (see [8], [12]) with $\varepsilon = h^{1/2}$, and recalling that $\beta_{kj} \leqslant r_k + t_j$, we get

$$\begin{aligned}\sum_{|\beta|\leqslant\beta_{kj}} \left|D^\beta(u_j - v_j)\right| &\leqslant ch^{1/2}\|u_j - v_j\|_{\beta_{kj+1}} + ch^{-1/2}\|u_j - v_j\|_{\beta_{kj}} \\ &\leqslant ch^{1/2}\|u_j - v_j\|_{r_k+t_j+1} + ch^{-1/2}\|u_j - v_j\|_{r_k+t_j} \\ &\leqslant ch^{\mu - r_k - 1/2}\|u_j\|_{\mu+t_j},\end{aligned} \tag{4.3}$$

where, in the last step, we have again used the approximation property of v_j. Using (4.2) and (4.3), we obtain the result.

LEMMA 4.2. *Assume $\mu' \geqslant \bar{\mu}$, and assume that S_{hj} approximates optimally with respect to $(\mu' + t_j, \alpha_j)$. Then*

$$\left\|\sum_j^N L_{ij}e_j\right\|_{s_i - \mu'} \leqslant ch^{\mu' - 2s_i}\|\underline{e}\|_A, \qquad i = 1,\ldots,N,$$

and

$$\left\|\sum_{j=1}^N B_{kj}e_j\right\|_{r_k + 1/2 - \mu'} \leqslant ch^{\mu' - 2r_{k-1}}\|\underline{e}\|_A, \qquad k = 1,\ldots,m.$$

Proof. We consider the elliptic boundary value problem

$$\sum_j L_{ij}v_j = \tilde{f}_i \quad \text{in } \Omega, \qquad \sum_j B_{k,j}v_j = \tilde{g}_k \quad \text{on } \Gamma,$$

with $\tilde{f} \in C^\infty(\bar{\Omega})$, $\tilde{g} \in C^\infty(\Gamma)$ the space of infinitely differentiable functions on $\bar{\Omega}$, and Γ, respectively. Let $\underline{v} = (v_j)$ denote the unique solution to the above problem and let $\underline{v}_h$ be the corresponding least squares approximation to $\underline{v}$. Then using (3.5), Lemma 4.1 and (2.7), we have

$$\begin{aligned}(\underline{e}, \underline{v})_A &= \sum_i h^{2s_i}\left(\sum_j L_{ij}e_j, \tilde{f}_i\right) + \sum_k h^{2r_k+1}\left\langle \sum_j B_{kj}e_j, \tilde{g}_k\right\rangle \\ &= (\underline{e}, \underline{v} - \underline{v}_h)_A \leqslant \|\underline{e}\|_A\|\underline{v} - \underline{v}_h\|_A \\ &\leqslant ch^{\mu'}\|\underline{e}\|_A\left[\sum_i \|\tilde{f}_i\|_{\mu' - s_i} + \sum_k |\tilde{g}_k|_{\mu' - r_k - 1/2}\right].\end{aligned}$$

Then for a given i_0, we choose $g_k = 0$, $k = 1, \ldots, m$, and $\tilde{f}_l = 0$, $l \neq i_0$, so that

$$\left\| \sum_j L_{i_0 j} e_j \right\|_{s_{i_0} - \mu'} = \sup_{\tilde{f}_{i_0} \in C^\infty(\Omega)} \frac{\left(\sum_j L_{i_0 j} e_j, \tilde{f}_{i_0} \right)}{\|\tilde{f}_{i_0}\|_{\mu' - s_{i_0}}}$$

$$= h^{-2s_{i_0}} \sup_{f_{i_0} \in C^\infty(\bar{\Omega})} \frac{(\underline{e}, \underline{v})_A}{\|f_{i_0}\|_{\mu' + s_{i_0}}} \leqslant c h^{\mu - 2s_{i_0}} \|\underline{e}\|_A.$$

In a similar manner we obtain the estimate for $|\sum_j B_{kj} e_j|_{r_{k+1/2} - \mu'}$. This completes the proof.

We now state and prove our main result.

THEOREM 4.1. *Let Ω be a bounded domain with smooth boundary Γ. We consider the elliptic system* (2.1) *with covering boundary conditions* (2.3) *and assume that this boundary value problem has a unique solution $\underline{u}$ which satisfies* (2.7) *for all smooth f_i and g_k. Let $\mu \geqslant \bar{\mu}$, $\nu \leqslant \underline{\mu}$ and $\delta \geqslant \max(2\bar{\mu} - \nu, \mu)$. Assume that the subspaces approximate optimally with respect to $(\delta + t_j, \alpha_j)$; then*

$$\sum_{j=1}^N \|u_j - u_{jh}\|_{\nu + t_j} \leqslant c h^{\mu - \nu} \sum_{j=1}^N \|u_j\|_{\mu + t_j}. \tag{4.4}$$

Proof. We have by (2.7)

$$\sum_j \|e_j\|_{\nu + t_j} \leqslant c \sum_i \left\| \sum_j L_{ij} e_j \right\|_{\nu - s_i} + c \sum_k \left| \sum_j B_{kj} e_j \right|_{\nu - r_k - 1/2}. \tag{4.5}$$

Let $w_i = \sum_j L_{ij} e_j$, and recall that $s_i - \delta \leqslant \nu - s_i \leqslant 0$. If $s_i - \delta < 0$, interpolation of the identity operator gives

$$\|w_i\|_{\nu - s_i} \leqslant c \|w_i\|_{s_i - \delta}^{\theta} \|w_i\|^{1 - \theta}, \qquad \theta = \frac{s_i - \nu}{\delta - s_i}.$$

If $s_i = \delta$, then $\nu = s_i$ and the inequality holds with $\theta = 0$, for example. From (3.2), $\|w_i\| \leqslant c h^{-s_i} \|\underline{e}\|_A$, and from Lemma 4.2,

$$\|w_i\|_{s_i - \delta} \leqslant c h^{\delta - 2s_i} \|\underline{e}\|.$$

Hence, using Lemma 4.1,

$$\|w_i\|_{\nu - s_i} \leqslant c h^{-\nu} \|\underline{e}\|_A \leqslant c h^{\mu - \nu} \sum_j \|u_j\|_{\mu + t_j}. \tag{4.6}$$

The boundary terms are treated in a similar fashion. Let $y_k = \sum_j B_{kj} e_j$, and recall that $r_k - \delta + \frac{1}{2} \leqslant \nu - r_k - \frac{1}{2} \leqslant 0$. If $\delta > r_k + \frac{1}{2}$,

$$|y_k|_{\nu - r_k - 1/2} \leqslant c |y_k|_{r_k - \delta + 1/2}^{\theta} |y_k|^{1 - \theta}, \qquad \theta = \frac{r_k + \frac{1}{2} - \nu}{\delta - r_k - \frac{1}{2}}.$$

If $\delta = r_k + \frac{1}{2}$, then $\nu = r_k + \frac{1}{2}$ and the inequality holds with $\theta = 0$. From (3.2),

$$|y_k| \leqslant c h^{-r_k - 1/2} \|\underline{e}\|_A,$$

and from Lemma 4.2,

$$|y_k|_{r_k - \delta + 1/2} \leqslant c h^{\delta - 2r_k - 1} \|e\|_A.$$

Hence, using Lemma 4.1,

$$|y_k|_{\nu - r_k - 1/2} \leqslant ch^{-\nu}\|\underline{e}\|_A \leqslant ch^{\mu-\nu}\sum_j \|u_j\|_{\mu+t_j}. \tag{4.7}$$

Using the estimates (4.6) and (4.7) in (4.5) finishes the proof.

With additional hypotheses on the subspace S_{hj}, we obtain error estimates in higher norms. Specifically, we assume that S_{hj} satisfies the "inverse" assumption

$$\|v_j\|_q \leqslant ch^{\nu+t_j-q}\|v_j\|_{\nu+t_j} \quad \text{for any } v_j \in S_{hj} \text{ where } \nu + t_j \leqslant q \leqslant \alpha_j. \tag{4.8}$$

Then we have the following corollary to Theorem 4.1.

COROLLARY 4.1. *Suppose, in addition to the hypotheses of Theorem* 4.1, *that each* S_{hj} *satisfies* (4.8). *Then for* $\underline{\mu} \leqslant \gamma_j \leqslant \alpha_j - t_j$,

$$\|u_j - u_{hj}\|_{\gamma_j+t_j} \leqslant ch^{\mu-\gamma_j}\sum_j \|u_j\|_{\mu+t_j}.$$

Proof. We have for $v_{jh} \in S_j$,

$$\|u_j - u_{jh}\|_{\gamma_j+t_j} \leqslant \|u_j - v_{jh}\|_{\gamma_j+t_j} + \|u_{jh} - v_{jh}\|_{\gamma_j+t_j}.$$

Using the approximation property (3.1), we have

$$\|u_j - v_{jh}\|_{\gamma_j+t_j} \leqslant ch^{\mu-\gamma_j}\|u_j\|_{\mu+t_j},$$

and from (4.8),

$$\|u_{jh} - v_{jh}\|_{\gamma_j+t_j} \leqslant ch^{\nu-\gamma_j}\|u_{jh} - v_{jh}\|_{\nu+t_j}.$$

Now

$$\|u_{jh} - v_{jh}\|_{\nu+t_j} \leqslant \|u_j - v_{jh}\|_{\nu+t_j} + \|u_j - v_{jh}\|_{\nu+t_j} \leqslant ch^{\mu-\nu}\|u_j\|_{\mu+t_j}$$

by Theorem 4.1 and (3.1). This completes the proof.

Theorem 4.1 provides a quasi-optimal error estimate, in the sense that the approximate solution for each component function, u_j, has accuracy of order $O(h^{\mu-\nu})$, and no greater order of accuracy could be expected, considering that (i) the error in u_j is measured in $H^{\nu-t_j}(\Omega)$, and (ii), the solution component, u_j, is assumed to lie in $H^{\mu+t_j}(\Omega)$. The regularity requirements for the solution fit naturally into the theory of ADN systems. The particular powers of h that appear in the minimizing functional, $\|u - u_h\|_A$, are critical for the success of the method. It is important to obtain results of the type of Theorem 4.1 with spaces S_{hj} that are as simple as possible. If we impose a, perhaps strange, inverse assumption on the S_{hj}, we can obtain the same conclusion, with reduced approximation hypotheses required of the S_{hj}. The inverse assumption that we need is that there is a $c > 0$ such that for each $\underline{z} \in \underline{S}_h$,

$$\left|\sum_j B_{kj}z_j\right|_l \leqslant ch^{-l}\left|\sum_j B_{kj}z_j\right|, \qquad 0 \leqslant l \leqslant \min_j\left\{\alpha_j - \beta_{kj} - \tfrac{1}{2}\right\}, 1 \leqslant k \leqslant m. \tag{4.9}$$

The reduced approximation hypothesis is expressed in terms of a larger possible value for ν. The result is as follows.

THEOREM 4.2. *Let Ω be a bounded domain with smooth boundary Γ. Suppose that the problem* (2.1), (2.3) *has, for any smooth f_i and g_k, a unique solution $\underline{u}$ which satisfies* (2.7). *Let $\mu \geqslant \bar{\mu}$, $\nu \leqslant \min\{s_i, \alpha_j - t_j, i, j = 1, \ldots, N\}$, $\delta > \max(2\bar{\mu} - \nu, \mu)$. Assume that the subspaces approximate optimally with respect to $(\delta + t_j, \alpha_j)$, and that* (4.9) *holds for each $\underline{z} \in \underline{S}_h$. Then the error estimate* (4.4) *holds.*

Proof. Proceeding as in the proof of Theorem 4.1, we arrive at (4.5). The estimates of the terms on the right side of (4.5) are exactly the same, except for the boundary norm in the case $\nu - r_k - \frac{1}{2} > 0$. Suppose that $\nu - r_k - \frac{1}{2} > 0$, pick a good approximation $\underline{v} \in \underline{S}_h$ to $\underline{u}$, and write $e_j = u_j - v_j + v_j - u_{hj}$. Using the triangle inequality,

$$\left|\sum_j B_{kj} e_j\right|_{\nu - r_k - 1/2} \leqslant \left|\sum_j B_{kj}(u_j - v_j)\right|_{\nu - r_k - 1/2} + \left|\sum_j B_{kj}(v_j - u_{hj})\right|_{\nu - r_k - 1/2}$$
$$\equiv \mathrm{I} + \mathrm{II}.$$

Since $\beta_{kj} \leqslant r_k + t_j$, $0 < \nu - r_k - \frac{1}{2} + \beta_{kj} \leqslant \nu + t_j - \frac{1}{2}$. Hence, we may use the trace inequality and the approximation property to obtain

$$\begin{aligned}
\mathrm{I} &\leqslant c \sum_{j,k} \sum_{|\beta| \leqslant \beta_{kj}} \left|D^\beta (u_j - v_j)\right|_{\nu - r_k - 1/2} \\
&\leqslant c \sum_{j,k} \|u_j - v_j\|_{\nu + \beta_{kj} - r_k} \leqslant c \sum_j \|u_j - v_j\|_{\nu + t_j} \\
&\leqslant c h^{\mu - \nu} \sum_j \|u_j\|_{\mu + t_j}.
\end{aligned}$$

To bound II, since $\underline{v} \in \underline{u}_h \in \underline{S}_h$, there is the possibility of using (4.9). Since $\nu \leqslant \alpha_j - t_j$, $\nu - r_k - \frac{1}{2} \leqslant \alpha_j - t_j - r_k - \frac{1}{2} \leqslant \alpha_j - \beta_{kj} - \frac{1}{2}$. Hence we may apply (4.9) to obtain

$$\begin{aligned}
\mathrm{II} &\leqslant c h^{-\nu + r_k + 1/2} \left|\sum_j B_{kj}(v_j - u_{hj})\right| \leqslant c h^{-\nu} \|\underline{v} - \underline{u}_h\|_A \\
&\leqslant c h^{-\nu} \left(\|\underline{u} - \underline{v}\|_A + \|\underline{u} - \underline{u}_h\|_A\right) \leqslant c h^{-\nu} \|\underline{u} - \underline{u}_h\|_A.
\end{aligned}$$

From Lemma 4.1, we get the desired bound for II, and the proof is finished.

5. Examples. In this section, we apply the least squares method to the elliptic boundary value problems discussed in Section 2. In each case, we list the hypotheses on the subspaces that are required for our theorems, and state the error estimates provided by the theorems.

The first example concerns the problem (2.8). From (2.8c) we find that $\bar{\mu} = 0$, $\underline{\mu} = -\frac{3}{2}$, $\alpha_1 = 2$. Hence, we require that $S_{h1} \subset H^2(\Omega)$. If $\mu \geqslant 0$, $\nu \leqslant -2$, and $\delta \geqslant \max\{\mu, -\nu\}$, and if S_{h1} approximates optimally with respect to $(\delta + 2, 2)$, then the error e_1 in the least squares method satisfies

$$\|e_1\|_{\nu+2} \leqslant c h^{\mu - \nu} \|u_1\|_{\mu+2}. \tag{5.1}$$

If S_{hj} also satisfies the inverse assumption (4.4), we obtain in addition the estimate

$$\|e_1\|_{\gamma+2} \leqslant c h^{\mu - \gamma} \|u_1\|_{\mu+2}, \qquad \nu \leqslant \gamma \leqslant \mu. \tag{5.2}$$

In particular, if we choose $\nu = -2$, $\mu = 2$, then our requirement is that S_{h1} approximate optimally with respect to (4, 2), and (5.1) is $\|e_1\|_0 \leqslant ch^4\|e\|_4$. This can be achieved, for example, by letting S_{h1} be a space of bicubic splines on a uniform mesh. These results are identical with the results obtained in [5], [3].

The second example concerns the boundary value problem (2.10). From (2.10c), we find that $\bar{\mu} = 0$, $\underline{\mu} = -\frac{3}{2}$, $[\alpha_j] = [1, 1, 1]$. Hence we require that $S_{hj} \subset H^1(\Omega)$, $j = 1, 2, 3$. This regularity requirement on the subspaces is less stringent than the requirement of the first example, and is a reason for preferring the reformulation of (2.8) as the first-order system, (2.10), when using a least squares method. If $\mu \geqslant 0$, $\nu \leqslant -2$, and $\delta \geqslant \max\{\mu, -\nu\}$, and if S_{hj} approximates optimally with respect to $(t_j + \delta, 1)$, $j = 1, 2, 3$, then the error e_j in the least squares method satisfies

$$\sum_j \|e_j\|_{\nu + t_j} \leqslant ch^{\mu - \nu} \sum_j \|u_j\|_{\mu + t_j}. \tag{5.3}$$

If $\underline{S}_h$ also satisfies the inverse assumption (4.4), we obtain in addition the estimate

$$\sum \|e_j\|_{\gamma + t_j} \leqslant ch^{\mu - \gamma} \sum_j \|u_j\|_{\mu + t_j}, \qquad \nu \leqslant \gamma \leqslant \mu. \tag{5.4}$$

In particular, if we choose $\nu = -2$, $\mu = 2$, then $\delta = 2$ and our requirement is that S_{hj} approximates optimally with respect to $(t_j + 2, 1)$, and (5.3) implies that

$$\|e_1\| \leqslant ch^4\big(\|u_1\|_4 + \|u_2\|_3 + \|u_3\|_3\big). \tag{5.5a}$$

With suitable inverse assumptions, we also obtain, from (5.4),

$$\|e_1\|_1 \leqslant ch^3\big(\|u_1\|_4 + \|u_2\|_3 + \|u_3\|_3\big). \tag{5.5b}$$

These can be achieved, for example, by letting S_{h1} be a space of continuous, piecewise bicubic polynomials on a uniform mesh, and by letting S_{h2} and S_{h3} be collections of continuous, piecewise biquadratic polynomials on the same mesh.

It is of interest to apply Theorem 4.2 to this example. In this case, the inverse assumption becomes

$$|z|_l \leqslant ch^{-l}|z|, \qquad 0 \leqslant l \leqslant \tfrac{1}{2}, z \in S_{h1}. \tag{5.6}$$

Suppose that S_{h1} satisfies (5.6). We may then use Theorem 4.2 with $\nu \leqslant -1$. Choosing $\nu = -1$, $\mu = 0$, so $\delta = 1$, and assuming that S_{h1} approximates optimally with respect to (3, 1), and S_{h2} and S_{h3} approximate optimally with respect to (2, 1), we obtain the error estimate

$$\|e_1\|_1 + \|e_2\| + \|e_3\| \leqslant ch\big[\|u_1\|_2 + \|u_2\|_1 + \|u_3\|_1\big].$$

This can be achieved, for example, by letting S_{h1} be a space of continuous, piecewise biquadratic polynomials on a uniform mesh, and by letting S_{h2} and S_{h3} be continuous, piecewise bilinear polynomials on the same mesh.

A problem similar to (2.10) has been treated by Jesperson [11] using a least squares method that only contains a weight on the boundary integral. In Jesperson's problem, $f_1 = f_2 = 0$, and the term $-u_1$ is removed from $L_3\underline{u}$. These changes do not affect the ellipticity indices, or our analysis of the problem. The method of Jesperson consists in minimizing, for $\underline{v} \in \underline{S}_h$, the expression

$$\|u_{1,x} - u_2\|^2 + \|u_{1,y} - u_3\|^2 + \|u_{2,x} + u_{3,y} - f_3\|^2 + h^{-1}|u_1 - g|^2. \tag{5.8}$$

The proof of Lemma 4.3 in [11] is incorrect, but this lemma is a special case of the inequality (2.7) for the problem (2.10). In contrast, for this problem, the functional (3.4) becomes

$$(5.9)\quad h^{-2}\|u_{1,x}-u_2\|^2+h^{-2}\|u_{1,y}-u_3\|^2+\|u_{2,x}+u_{3,y}-f_3\|^2+h^{-3}|u_1-g|^2.$$

To describe some typical results that are obtained in [11], let $\underline{\tilde{u}}\in\underline{S}_h$ be the approximate solution that is obtained by minimizing (5.8), and let $\underline{\tilde{e}}=\underline{u}-\underline{\tilde{u}}_h$ be the resulting error. Suppose first that S_{hj}, $j=1,2,3$, is a space of continuous piecewise bilinear functions on a uniform mesh of size h. If $u_1\in H^2(\Omega)$, one has $\|\tilde{e}_1\|\leqslant ch\|u_1\|_2$. If $u_1\in H^3(\Omega)$, one has $\|\tilde{e}_1\|\leqslant ch^2\|u_1\|_3$, and, if the subspaces also satisfy (5.6), it is shown that $\|\tilde{e}_1\|_1\leqslant ch\|u_1\|_3$. Next, suppose that S_{h1} is a space of continuous, piecewise biquadratic functions on a uniform mesh, while S_{h2} and S_{h3} are piecewise bilinear functions on the same mesh. In this case, if $u_1\in H^2(\Omega)$, the result $\|\tilde{e}_1\|\leqslant ch^2\|u_1\|_3$ is obtained, and, if the subspaces satisfy (5.6), $\|\tilde{e}_1\|_1\leqslant ch\|u_1\|_3$. If $u_1\in H^4(\Omega)$, then one has the estimates $\|\tilde{e}_1\|\leqslant ch^3\|u_1\|_4$, and $\|\tilde{e}_1\|_1\leqslant ch^2\|u_1\|_4$, where the latter inequality also assumes (5.6). Comparing the error estimates (5.5) and (5.7) with these estimates, it seems difficult to draw general conclusions. However, it seems that the functional (5.9) provides error estimates that utilize more fully the regularity of the solution, while the functional (5.8) allows the use of simpler spaces of test functions. Perhaps further analysis, as well as numerical studies, would be needed to decide the relative merits of the two least squares methods.

The final example concerns the Stokes problem, (2.11). From (2.11c) we find that $\bar{\mu}=0$, $\underline{\mu}=-1$, $[\alpha_j]=[2,2,1]$. Hence, we require that $S_{hj}\subset H^2(\Omega)$, $j=1,2$, and $S_{h3}\subset H^1(\Omega)$. The A-norm in this case is defined by

$$\|\underline{u}\|_A^2=\|-\Delta u_1+u_{3,x}\|^2+\|-\Delta u_2+u_{3,y}\|^2+h^{-2}\|u_{1,x}+u_{2,y}\|^2$$
$$+h^{-3}\left[|u_1|^2+|u_2|^2\right].$$

The analysis of the method is complicated by the compatibility condition (2.12) that is needed for the solvability of the problem. To handle this difficulty we follow the approach of Wendland [15] and modify the system of equations. For this, let z be a smooth function defined on Ω and such that $(z,1)\neq 0$. The function $z(x)\equiv 1$ will suffice. We consider, instead of (2.11a), the system

$$(5.10)\qquad \begin{cases} -\Delta u_1+u_{3,x}=f_1,\\ -\Delta u_2+u_{3,y}=f_2,\\ u_{1,x}+u_{2,y}+\alpha z=f_3. \end{cases}$$

The modified problem (5.10), (2.11b) is to be solved for the unknown function $\underline{u}$ and the unknown number α. If $\underline{f}$ and $\underline{g}$ are smooth functions, the problem (5.10), (2.11b) has a solution $\{\underline{u},\alpha\}$. The function u_3 is specified up to an additive constant. If u_3 is chosen so that $(u_3,1)=0$, the solution $\{\underline{u},\alpha\}$ satisfies, for all real l, the inequality

$$(5.11)\qquad \|u_1\|_{l+2}+\|u_2\|_{l+2}+\|u_3\|_{l+1}+|\alpha|$$
$$\leqslant C\left[\|f_1\|_l+\|f_2\|_l+\|f_3\|_{l+1}+|g_1|_{l+3/2}+|g_2|_{l+3/2}\right].$$

To see this, define $\alpha = [(f_3, 1) - \langle \underline{n} \cdot \underline{g}, 1\rangle]/(z, 1)$ and replace f_3 by $f_3 - \alpha z$. The new problem then becomes the Stokes equations (2.11b) with a modified right-hand side. The new right-hand side satisfies the compatibility condition (2.12), so from the theory of the Stokes problem (2.11), there is a solution $\underline{u}$ which satisfies $(u_3, 1) = 0$, and which also satisfies (2.14) for all real l. The estimate

$$|\alpha| \leqslant C\left[\|f_3\|_{l+1} + |g_1|_{l+3/2} + |g_2|_{l+3/2}\right],$$

and the estimate (5.11), follows from this.

To formulate and analyze a least squares method for the system (5.10), we define a bilinear form on pairs of triples $\{\underline{u}, \alpha\}$ and $\{\underline{v}, \beta\}$ by the expression

$$\begin{aligned}\|\{\underline{u}, \alpha\}, \{\underline{v}, \beta\}\|_A^2 = {} & (-\Delta u_1 + u_{3,x}, -\Delta v_1 + v_{3,x}) \\ & + (-\Delta u_2 + u_{3,y}, -\Delta v_2 + v_{3,y}) \\ & + h^{-2}(u_{1,x} + u_{2,y} + \alpha z, v_{1,x} + v_{2,y} + \beta z) \\ & + h^{-3}\{\langle u_1, v_1\rangle + \langle u_2, v_2\rangle\}.\end{aligned}$$

We remark that the corresponding quadratic form, which we write $\|\{\underline{u}, \alpha\}\|_A^2$, defines a seminorm, since if $u_3 \equiv 1$, and $u_1 = u_2 \equiv 0$, $\alpha = 0$, then $\|\{\underline{u}, \alpha\}\|_A^2 = 0$. The corresponding least squares approximation, $\{\underline{u}_h, \alpha_h\}$, is not unique, since an arbitrary constant can be added to the third component of $\underline{u}_h$. Nevertheless, virtually the same argument that is used to prove Theorem 4.1 leads to a proof of the following theorem

THEOREM. *Let $\mu \geqslant 0$, $\nu \leqslant -1$, $\delta \geqslant \max(-\nu, \mu)$ and let S_{h_j} approximate optimally with respect to $(\delta + t_j, \alpha_j)$. Let $\{\underline{u}, \alpha\}$ be a solution with $(u_3, 1) = (u_{h3}, 1)$. Then*

$$|\alpha - \alpha_h| + \sum_j \|u_j - u_{jh}\|_{\nu + t_j} \leqslant ch^{\mu - \nu} \sum_j \|u_j\|_{\mu + t_j}.$$

As a final illustration of the least squares methodology, we formulate a nonconforming least squares methods. Since we prove nothing about the method and give no numerical results, its value is a matter of conjecture.

Nonconforming finite element methods have been used to avoid the regularity requirements on the subspaces, especially for higher-order problems (see, e.g., [2]). We formulate a nonconforming least squares method for the problem (2.8). To motivate our method, let Γ_0 be a smooth closed curve in Ω, dividing Ω into two subdomains Ω_1 and Ω_2. The boundary of Ω_1 is Γ_0, the boundary of Ω_2 is $\Gamma_0 \cup \Gamma$. Using these subdomains, the problem (2.8) may be given a different formulation as follows. We seek functions u_1 and u_2, defined in Ω_1 and Ω_2, such that

$$-\Delta u_\kappa + u_\kappa = f \quad \text{in } \Omega_\kappa, \kappa = 1, 2, \tag{5.12a}$$

$$u_1(x) - u_2(x) = 0, \qquad x \in \Gamma_0, \tag{5.12b}$$

$$\frac{\partial u_1}{\partial n}(x) - \frac{\partial u_2}{\partial n}(x) = 0, \qquad x \in \Gamma_0, \tag{5.12c}$$

$$u_2(x) = g(x), \qquad x \in \Gamma_0. \tag{5.12d}$$

In these equations, n denotes the unit normal on Γ_0, pointing from Ω_1 into Ω_2. It may be shown that there is a unique solution pair, u_1, u_2, of (5.12), and the solution

is $u_\kappa = u$ restricted to Ω_κ, $\kappa = 1, 2$, where u is the solution of (2.8). The equations (5.12b, c) serve as boundary conditions for the problem, with indices $r = -2$ and $r = -1$, respectively. (5.12) is a problem of "interface" or "transmission" type, and the theory of elliptic boundary value problems may be extended to include this type of problem.

We shall use (5.12) to motivate our nonconforming method; however, we shall consider a situation in which the curves Γ_0 are not smooth. For this, let there be given a uniform mesh of size h on Ω. The mesh divides Ω into a number of subdomains Ω_κ; each Ω_κ is either a mesh rectangle or the intersection of a mesh rectangle with the domain Ω. The mesh lines consist of a collection of line segments $\Gamma_{\kappa\lambda}$. Each line segment, $\Gamma_{\kappa\lambda}$, is the common boundary of two subdomains Ω_κ and Ω_λ. We choose a unit normal $n_{\kappa\lambda}$ on each line segment $\Gamma_{\kappa\lambda}$, and if $x \in \Gamma_{\kappa\lambda}$, we let

$$u(x \pm) = \lim_{\varepsilon \to 0} u(x \pm \varepsilon n_{\kappa\lambda}).$$

Let S_h be a collection of piecewise polynomials on the mesh; no continuity conditions are required for the functions in S_h. Our nonconforming least squares method is to minimize, for $v \in S_h$, the quantity

$$\sum_\kappa \iint [\Delta v - f]^2 + \sum_{\kappa,\lambda} \int_{\Gamma_{\kappa\lambda}} \left\{ h^{-3}[v(x+) - v(x-)]^2 + h^{-1}\left[\frac{\partial v}{\partial n}(x+) - \frac{\partial v}{\partial n}(x-)\right]^2 \right\} + h^{-3}\int_\Gamma [v - g]^2.$$

The weights appearing in the integral over $\Gamma_{\kappa\lambda}$ are dictated by the values of r associated with the boundary conditions (5.21b, c). It would be of interest to give an error analysis for this method.

6. Condition Number. If the Dirichlet problem (2.8) is solved numerically using Galerkin's method with typical finite-element matrices on a uniform mesh of size h, the resulting stiffness matrix has condition number $O(h^{-2})$. If the same problem is solved using the weighted least squares method of [5], the condition number of the associated matrix is $O(h^{-4})$. This results in extra difficulties in obtaining an accurate solution of the linear system. These considerations led Bramble and Nitsche [4] to formulate a modified least squares method with a reduced condition number. The least squares method in [11] also has a reduced condition number. Here, we give an upper bound for the condition number of our linear system.

For the condition number bounds, we require some assumptions on the subspaces S_{hj}. First, we suppose that there is a set of basis functions, $\phi_{j\kappa}$, of S_{hj}, such that

$$eh^n \sum_\kappa a_\kappa^2 \leqslant \left\| \sum_\kappa a_\kappa \phi_{\kappa j} \right\| \leqslant Eh^n \sum_\kappa a_\kappa^2. \tag{6.1}$$

The positive constants e and E are independent of h. This inequality enables us to estimate the L_2-norm of a function $v_j \in S_{hj}$ in terms of the coefficients in the expansion $v_j = \sum a_\kappa \phi_{j\kappa}$. Secondly, we require the inverse assumption

$$\|v_j\|_l \leqslant Dh^{-l}\|v_j\|, \qquad v_j \in S_{hj}, 0 \leqslant l \leqslant \alpha_j, \tag{6.2}$$

where the constant D is independent of h. Notice that if (6.2) and the upper inequality of (6.1) are combined, we obtain for any numbers a_κ the inequality

$$\sum_{|\beta|\leqslant l}\sum_{\kappa,\lambda} a_\kappa a_\lambda\left(D^\alpha\phi_{j\kappa}, D^\alpha\phi_{j\lambda}\right) \leqslant D^2Eh^{2n-2l}\sum_\kappa a_\kappa^2. \tag{6.3}$$

The matrix problem arising from our least squares method depends on the basis functions chosen for the subspace $\underline{S}_h$. If S_{hj} has dimension d_j, then $\underline{S}_h$ has dimension $d = \Sigma d_j$. To describe d linearly independent functions in $\underline{S}_h$, let l and λ be given with $1 \leqslant l \leqslant N$, $1 \leqslant \lambda \leqslant d_l$. Let $\underline{\psi}^{(l,\lambda)}$ be defined in terms of its component functions $\psi_j^{(l,\lambda)}$ by $\psi_j^{(l,\lambda)} = \phi_{l\lambda}$ for $j = l$, and $\psi_j^{(l,\lambda)} = 0$, for $j \neq l$. The d functions $\underline{\psi}^{(l,\lambda)}$ form a basis for $\underline{S}_h$. The least squares method for the problem (2.1), (2.3), consists in minimizing the expression (3.4) over all $\underline{v} \in \underline{S}_{hj}$. The solution of this problem is given by the solution of an associated linear system of equations. The matrix A of the linear system is of order d; a typical matrix entry is provided by the quantity $(\underline{\psi}^{(k,\kappa)}, \underline{\psi}^{(l,\lambda)})_A$.

The condition number of A is defined to be $\operatorname{cond} A = \|A\| \cdot \|A^{-1}\|$, where the norm is any matrix norm. Choosing the matrix norm arising from the Euclidean vector norm, we find that $\operatorname{cond} A = \lambda_{\max}/\lambda_{\min}$, where $\lambda_{\max}$ and $\lambda_{\min}$ are the largest and smallest eigenvalues of A. Since A is positive definite, these eigenvalues are positive and may be estimated by the Rayleigh quotient $Q = \underline{a}^TA\underline{a}/\underline{a}^T\underline{a}$, where $\underline{a} = (a_{l\lambda}) \in R^d$. For this, if $\underline{a}$ is given, let $\underline{v} = \Sigma a_{l\lambda}\underline{\psi}^{(l,\lambda)} \in \underline{S}_h$. Then $\underline{a}^TA\underline{a} = (\underline{v}, \underline{v})_A$, so

$$Q = (\underline{v}, \underline{v})_A/\Sigma a_{l\lambda}^2. \tag{6.4}$$

We require upper and lower estimates for the numerator. For the upper estimate, we use (3.2), (6.3), and the inequality $|z|^2 \leqslant c(h\|z\|_1^2 + h^{-1}\|z\|^2)$ to obtain

$$\begin{aligned}\|\underline{v}\|_A^2 &\leqslant c\sum_{i,j} h^{2s_i}\|v_j\|_{\lambda_{ij}}^2 + c\sum_{k,j}\sum_{|\beta|\leqslant\beta_{kj}} h^{2r_k+1}\left|D^\beta v_j\right|_{\beta_{kj}}^2 \\ &\leqslant c\sum h^{n+2s_i-2\lambda_{ij}}a_{\kappa j}^2 + c\sum\left\{h^{2r_k+2}\|v_j\|_{\beta_{kj}+1}^2 + h^{2r_k}\|v_j\|_{\beta_{kj}}^2\right\} \\ &\leqslant ch^n\left\{\max_{i,j,k}\left[h^{2s_i-2\lambda_{ij}}, h^{2r_k-2\beta_{kj}}\right]\right\}\sum a_{l\lambda}^2.\end{aligned}$$

Let $\gamma = \max[2\lambda_{ij} - 2s_i, 2\beta_{kj} - 2r_k]$. Since $\lambda_{ij} \leqslant s_i + t_j$, $\beta_{kj} \leqslant r_k + t_j$, and since equality holds for some of these indices, we have $\gamma = 2\max t_j$. The above inequalities then give

$$\|\underline{v}\|_A^2 \leqslant ch^{n-\gamma}\sum a_{l\lambda}^2 \tag{6.5}$$

For the lower estimate, we require a further inverse assumption. We suppose that

$$\left\|\sum_j L_{ij}v_j\right\|_{-s_i} \leqslant ch^{s_i}\left\|\sum_j L_{ij}v_j\right\|, \tag{6.6}$$

$$\left|\sum_j B_{kj}v_j\right|_{-r_k-1/2} \leqslant ch^{r_k+1/2}\left|\sum_j B_{kj}v_j\right|. \tag{6.7}$$

To understand these assumptions, recall that $s_i \leqslant 0$, $r_k + \frac{1}{2} \leqslant 0$. If the differential operators L_{ij} and B_{kj} all have constant coefficients, then the quantities $\Sigma_j L_{ij}v_j$ and

$\sum_j B_{kj} v_j$ are piecewise polynomials, and the inequalities (6.6) and (6.7) are not unreasonable. With these inequalities, (3.2), and (2.7) with $l = 0$, we have

$$\|\underline{v}\|_A^2 \geqslant c\sum_i \left\| \sum_j L_{ij} v_j \right\|_{-s_i}^2 + c\sum_k \left| \sum_j B_{kj} v_j \right|_{-r_k - 1/2}^2 \geqslant c\sum_j \|v_j\|_{t_j}^2 \geqslant c\sum_j \|v_j\|^2.$$

Using (6.1), we then obtain

$$\|\underline{v}\|_A^2 \geqslant ch^n \sum a_{l\lambda}^2. \tag{6.8}$$

The inequalities (6.5) and (6.8) give upper and lower bounds for $\lambda_{\max}$ and $\lambda_{\min}$. Using these bounds, we obtain

$$\text{cond } A \leqslant ch^{-\gamma}. \tag{6.9}$$

In the case of the model problem (2.10), we find that $\gamma = 4$, so cond $A \leqslant ch^{-4}$.

Applied Mathematics Branch (R44)
Naval Surface Weapon Center
Silver Spring, Maryland 20910

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

Institute for Physical Science and Technology
University of Maryland
College Park, Maryland 20740

1. S. AGMON, A. DOUGLIS & L. NIRENBERG, "Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions II," *Comm. Pure Appl. Math.*, v. 17, 1964, pp. 35–92.

2. I. BABUŠKA, J. T. ODEN & J. K. LEE, "Mixed-hybrid finite element approximations of second-order boundary value problems," *Comput. Methods Appl. Mech. Engrg.*, v. 11, 1977, pp. 175–206.

3. G. A. BAKER, "Simplified proofs of error estimates for the least squares method for Dirichlet's problem," *Math. Comp.*, v. 27, 1973, pp. 229–235.

4. J. H. BRAMBLE & J. A. NITSCHE, "A generalized Ritz-least-squares method for Dirichlet problems," *SIAM J. Numer. Anal.*, v. 10, 1973, pp. 81–93.

5. J. H. BRAMBLE & A. H. SCHATZ, "Rayleigh-Ritz-Galerkin-methods for Dirichlet's problem using subspaces without boundary conditions," *Comm. Pure Appl. Math.*, v. 23, 1970, pp. 653–675.

6. J. H. BRAMBLE & A. H. SCHATZ, "Least squares for $2m$th order elliptic boundary-value problems," *Math. Comp.*, v. 25, 1971, pp. 1–32.

7. J. H. BRAMBLE & R. SCOTT, "Simultaneous approximation in scales of Banach spaces," *Math. Comp.*, v. 32, 1978, pp. 947–954.

8. J. H. BRAMBLE & V. THOMÉE, "Pointwise bound for discrete Green's functions," *SIAM J. Numer. Anal.*, v. 6, 1969, pp. 583–590.

9. G. J. FIX, M. D. GUNZBURGER, & R. A. NICOLAIDES, "On finite element methods of the least squares type," *Comput. Math. Appl.*, v. 5, 1979, pp. 87–98.

10. G. J. FIX & E. STEPHAN, *Finite Element Methods of the Least Squares Type for Regions With Corners*, Report No. 81–41, December 16, 1981, Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, Virginia 23665.

11. D. C. JESPERSON, "A least squares decomposition method for solving elliptic equations," *Math. Comp.*, v. 31, 1977, pp. 873–880.

12. J. L. LIONS & E. MAGENES, *Non-Homogeneous Boundary Value Problems and Applications*, Vol. 1, Springer, Berlin, 1972.

13. J. ROĬTBERG & Z. ŠEFTEL, "A theorem about the complete set of isomorphisms for systems elliptic in the sense of Douglis and Nirenberg," *Ukrain. Mat. Zh.*, 1975, pp. 447–450.

14. R. TEMAM, *Navier-Stokes Equations*, North-Holland, Amsterdam, New York, 1977.

15. W. L. WENDLAND, *Elliptic Systems in the Plane*, Pitman, London, 1979.