JOURNAL OF APPROXIMATION THEORY 45, 69–84 (1985)

What Is the Complexity of Elliptic Systems?

ARTHUR G. WERSCHULZ*

Division of Science and Mathematics, Fordham Univerity/College at Lincoln Center, New York, New York 10023 and
Department of Computer Science, Columbia University, New York, New York 10027, U.S.A.

Communicated by Oved Shisha

Received October 23, 1984

This paper deals with the optimal solution of the Petrovsky-elliptic system $lu = f$, where l is of homogeneous order t and $f \in H^r(\Omega)$. Of particular interest is the strength of finite element information (FEI) of degree k, as well as the quality of the finite element method (FEM) using this information. We show that the FEM is quasi-optimal iff $k \geqslant r + t - 1$. Suppose this inequality is violated; is the lack of optimality in the FEM due to the information that it uses, or is it because the FEM makes inefficient use of its information? We show that the latter is the case. The FEI is always quasi-optimal information. That is, the spline algorithm using FEI is always a quasi-optimal algorithm. In addition, we show that the asymptotic penalty for using the FEM when k is too small (rather than the spline algorithm which uses the same finite element information as the FEI is unbounded. © 1985 Academic Press, Inc.

1. INTRODUCTION

This paper is a theoretical study of the optimal solution of systems of linear partial differential equations which are elliptic in the sense of Petrovsky [1, 12, 15]. A number of examples of such problems are described in [15]; these include the Cauchy–Riemann equations for Poisson's equation in the plane, as well as problems of fluid flow and elasticity. (The concept of elliptic system is defined in Section 2.)

Since one of the most commonly used methods for solving such problems is the finite element method (FEM), see [2–5, 11, 15], we wish to determine conditions under which the FEM is *quasi-optimal* (i.e., optimal to within a constant factor).

In order to make the notion of optimality more precise, we use the information-centered approach of [13]. The main idea is that an algorithm for

* This research was supported in part by the National Science Foundation under Grant MCS-8303111.

0021-9045/85 $3.00

solving this problem can only use information of finite cardinality (see Section 3 for definitions of these terms). Hence, there is inherent uncertainty when attempting to solve these infinite-dimensional problems using information of finite cardinality. From this, we are able to determine tight bounds on the nth *minimal error* (i.e., the minimal error among all algorithms using information of cardinality at most n).

In Section 4, we show that the FEM is quasi-optimal if and only if

$$k \geqslant r + t - 1, \tag{1.1}$$

where k is the degree of the finite element subspace, t is the order of the elliptic system, and the problem elements f are (a priori) uniformly bounded in the $H^r(\Omega)$-norm (so that r measures the regularity of the class of problem elements). Thus, the degree of the FEM must increase with the regularity of the class of problem elements, if the FEM is to remain quasi-optimal.

Suppose the inequality (1.1) is violated. Is the non-optimality of the FEM inherent in the finite element information (FEI) it uses, or is it due to the fact that it uses the FEI inefficiently? We show that the latter is the case; regardless of whether (1.1) holds, FEI is quasi-optimal information. That is, the "spline algorithm" using the FEI is quasi-optimal.

In Section 5, we discuss the ε-complexity of the problem, i.e., the complexity of finding approximations which differ by at most ε from the true solution. The FEM is a quasi-optimal-complexity algorithm iff (1.1) holds; if (1.1) is violated, the asymptotic penalty for using the FEM is unbounded. However, the spline algorithm using the FEI (which, again, is the same information that is used by the FEM) is *always* a quasi-optimal-complexity algorithm, regardless of whether (1.1) holds.

2. The Elliptic Boundary-Value Probem

In this section, we define (homogeneous) ellipticity, in the sense of Petrovsky. We quote "shift theorems," which allow a priori estimation of derivatives of the solution in terms of the derivatives of the data. We use standard notations for ($\mathbb{R}^N$-valued) Sobolev spaces, inner product, etc., found in [7] (but extended to include functions whose values are in $\mathbb{R}^N$). Fractional- and negative-order Sobolev spaces are defined via Hilbert space interpolation and duality, respectively (see [4, 6, 11] for details). Since for simplicity we only deal with real systems, we use the notation of [1] when describing ellipticity, even though the shift theorems are taken from [12]. For purposes of exposition, we assume that the coefficients of the system and the boundary of the region over which the problem is to be solved are C^∞.

Let $\Omega \subseteq \mathbb{R}^N$ be a bounded C^∞ region. Define the differential operator

$$l(x, \partial) = [l_{ij}(x, \partial)]_{1 \leqslant i,j \leqslant n},$$

with ∂_l denoting the partial derivative in the lth direction, where (using the standard multi-index notation found in, e.g., [7]) we set

$$l_{ij}(x, \xi) = \sum_{|\mu| \leqslant t} a_\mu^{ij}(x)\, \xi^\mu;$$

here the coefficients $a_\mu^{ij} \in C^\infty(\bar{\Omega})$ and t is a non-negative integer. Let

$$l_{ij}^0(x, \xi) = \sum_{|\mu| = t} a_\mu^{ij}(x)\, \xi^\mu$$

denote the *principal part* of l_{ij}. We assume that l is *elliptic*, i.e.,

$$L(x, \xi) := \det[l_{ij}^0(x, \xi)] \neq 0 \qquad \forall x \in \bar{\Omega}, \forall \text{ non-zero } \xi \in \mathbb{R}^N.$$

We now wish to specify a boundary operator. For $x \in \partial\Omega$, let v_x and τ_x denote unit normalized tangent vectors to $\partial\Omega$ at x, and set

$$L_x(\eta) = L(x, \tau_x + \eta v_x) \qquad \forall \eta \in \mathbb{C}.$$

L_x is a polynomial of degree Nt in the complex variable η, which (by ellipticity) has no real roots; since the coefficients of L_x are real, there is a non-negative integer m such that $Nt = \deg L_x = 2m$. Hence we may factor

$$L_x(\eta) = L_x^+(\eta)\, L_x^-(\eta),$$

where the zeros of L_x^+ (respectively, of L_x^-) have positive (respectively, negative) real part, and $\deg L_x^+ = \deg L_x^- = m$. Then we define a boundary operator

$$b(x, \partial) = [b_{ij}(x, \partial)]_{1 \leqslant i \leqslant m, 1 \leqslant j \leqslant N}$$

by

$$b_{ij}(x, \xi) = \sum_{|\mu| \leqslant r_i} b_\mu^{ij}(x)\, \xi^\mu,$$

where $r_1, ..., r_n$ are positive integers and the coefficients b_μ^{ij} are infinitely differentiable.

Let the *principal part* b_{ij}^0 of b_{ij} be defined by

$$b_{ij}^0(x, \xi) = \sum_{|\mu| = r_i} b_\mu^{ij}(x)\, \xi^\mu.$$

Let $L^{jk}(x, \xi)$ denote the cofactor of $l^0_{jk}(x, \xi)$ in the matrix $[l^0_{rs}(x, \xi)]_{1 \leqslant r,s \leqslant N}$. For $x \in \partial\Omega$ and complex η, let

$$C_x(\eta) = [c_{ij}(x, \eta)]_{1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n},$$

with

$$c_{ij}(x, \eta) = \sum_{k=1}^{N} b^0_{ik}(x, \tau_x + \eta\nu_x)\, L^{jk}(x, \tau_x + \eta\nu_x).$$

The boundary operator b is *complementary* to l if the row vectors of the matrix C_x, considered as polynomials in the *complex* variable η, are linearly independent relative to the modulus of $L^+_x(\eta)$.

We say that l and b are *elliptic* on $\bar{\Omega}$ if l is elliptic and b is complementary to l. For $s \geqslant 0$, let $H^s(\partial)$ denote the completion (with respect to the Sobolev norm $\|\cdot\|_s$) of the set of infinitely differentiable functions u such that $bu = 0$ on $\partial\Omega$. We then have the following "shift theorem," taken from [12]:

LEMMA 2.1. *If l and b are elliptic on $\bar{\Omega}$, then for any $r \geqslant 0$, there exists $\sigma \geqslant 1$ such that*

$$\sigma^{-1}\|lu\|_r \leqslant \|u\|_{r+t} \leqslant \sigma\|lu\|_r \qquad \forall u \in H^{r+t}(\partial).$$

In order to proceed, we must consider the *formal adjoint* l^+ of l given by

$$l^+(x, \partial) = [l^+_{ij}(x, \partial)]_{1 \leqslant i,j \leqslant N}$$

with

$$l^+_{ij}(x, \partial)\, u_j(x) = \sum_{|\mu| \leqslant t} \partial^\mu(a^{ji}_\mu(x)\, u_j(x)).$$

Integrating by parts, one may define an adjoint boundary operator b^+ such that

$$(lu, v)_0 = (u, l^+v)_0 \qquad \forall u \in H^t(\partial), \forall v \in H^t(\partial)^+,$$

where for $s \geqslant 0$, $H^s(\partial)^+$ denotes the $\|\cdot\|_s$-completion of the set of infinitely differential functions v such that $b^+v = 0$ on $\partial\Omega$.

In the remainder of this paper, we assume that l and b are elliptic on $\bar{\Omega}$, as well as l^+ and b^+. (Roĭtberg and Šeftel [12] give a *normality condition* on b such that ellipticity of l and b on $\bar{\Omega}$ implies that of l^+ and b^+.)

We then have the following result from [12]:

LEMMA 2.2. *Let $r \geqslant 0$. There is a constant $\sigma \geqslant 1$ such that the following hold*:

(i) *For any* $f \in H^r(\Omega)$, *there exists* $u \in H^{r+t}(\partial)$ *such that*

$$lu = f \text{ in } \Omega, \qquad bu = 0 \text{ on } \partial\Omega,$$

with

$$\sigma^{-1}\|f\|_r \leqslant \|u\|_{r+t} \leqslant \sigma\|f\|_r.$$

(ii) *For any* $g \in H^r(\Omega)$, *there exists* $v \in H^{r+t}(\partial)^+$ *such that*

$$l^+v = g \text{ in } \Omega, \qquad b^+v = 0 \text{ on } \partial\Omega,$$

with

$$\sigma^{-1}\|g\|_r \leqslant \|v\|_{r+t} \leqslant \sigma\|g\|_r.$$

We are now finally ready to state the problem to be studied in this paper. Given $r \geqslant 0$, define a *solution operator*

$$S\colon H^r(\Omega) \to H^t(\partial)$$

by letting $u = Sf$ satisfy

$$lu = f \text{ in } \Omega, \qquad bu = 0 \text{ on } \partial\Omega.$$

Using Lemma 2.2, we see that S is a bounded injection with range $H^{r+t}(\partial) \subseteq H^t(\partial)$. By the Rellich–Kondrasov theorem [7, p. 114], S is an isomorphism or compact, according to whether $r = 0$ or $r > 0$.

3. Information and Algorithms

In this section, we recall results from [13] concerning optimal algorithms and information, as applied to the problem of solving an elliptic system.

Recall that we are trying to approximate Sf for arbitrary $f \in H^r(\Omega)$, where $S\colon H^r(\Omega) \to H^t(\partial)$ is the solution operator defined above and $r \geqslant 0$. Most methods for solving this problem use a finite number of linear functionals on f when approximating Sf. For instance, such methods may evaluate f at a finite number of points in Ω, or the inner product of f with a finite number of predetermined functions. In fact, even when a closed form expression for f is available, most methods do not explicitly use this expression; they only use the values of a finite number of linear functionals at f. Hence, we assume that we only know the values of a finite number of linear functionals for each problem element f. That is, we are given *information* $\mathcal{N}$ of *cardinality* $n = \mathrm{card}(\mathcal{N})$, which is a linear surjection

$$\mathcal{N}\colon H^r(\Omega) \to \mathbb{R}^n.$$

Such information $\mathcal{N}$ is then used by an *algorithm* φ, which is a mapping $\varphi\colon \mathbb{R}^n \to H^t(\partial)$; the class of such algorithms using $\mathcal{N}$ is denoted $\Phi(\mathcal{N})$. Note that we allow *any* mapping to be an algorithm.

Given information $\mathcal{N}$ and an algorithm $\varphi \in \Phi(\mathcal{N})$, the quality of the approximations produced by φ is measured by its *error*

$$e(\varphi) = \sup_{f \in F} \|Sf - \varphi(\mathcal{N}f)\|_s,$$

where the set F of *problem elements* is taken to be the unit ball of $H^r(\Omega)$

$$F = BH^r(\Omega) := \{f \in H^r(\Omega) : \|f\|_r \leqslant 1\}$$

and $0 \leqslant s \leqslant t$. (In what follows, BH will always denote the unit ball of a Hilbert space H.)

We are interested in algorithms using given information whose error is as small as possible. Let

$$e(\mathcal{N}) = \inf\{e(\varphi) : \varphi \in \Phi(\mathcal{N})\}$$

denote the *optimal error* of algorithms using $\mathcal{N}$. An algorithm $\varphi^* \in \Phi(\mathcal{N})$ is an *optimal error algorithm* using $\mathcal{N}$ if

$$e(\varphi^*) = e(\mathcal{N}).$$

Expressions for the optimal error and an optimal error algorithm are given by the following result from [13, Chap. 4]:

LEMMA 3.1. (i) *The optimal error is given by*

$$e(\mathcal{N}) = \sup\{\|Sh\|_s : h \in F \cap \ker \mathcal{N}\}.$$

(ii) *Let*

$$\mathcal{N}f = [\lambda_1(f) \ldots \lambda_n(f)]^T \qquad \forall f \in H^r(\Omega),$$

where $\lambda_1, \ldots, \lambda_n\colon H^r(\Omega) \to \mathbb{R}$ *are linearly independent bounded linear functionals. Let* $\{f_1, \ldots, f_n\}$ *be a basis for the orthogonal complement* $(\ker \mathcal{N})^\perp$ *of* $\ker \mathcal{N}$ *in* $H^r(\Omega)$ *such that* $\lambda_i(f_j) = \delta_{ij}$. *Then the spline algorithm*

$$\varphi^s(\mathcal{N}f) = \sum_{j=1}^{n} \lambda_j(f)\, Sf_j$$

is an optimal error algorithm using $\mathcal{N}$.

Note that although we allow *any* mapping to be an algorithm, a *linear* optimal error algorithm always exists.

Now that we know how to find an optimal error algorithm for any information, we now seek *optimal* information of given cardinality. Let

$$e(n) = \inf\{e(\mathcal{N}) : \text{card}\, \mathcal{N} \leqslant n\}$$

denote the nth *minimal error*. Information $\mathcal{N}_n^*$ of cardinality at most n is said to be nth *optimal information* if

$$e(\mathcal{N}_n^*) = e(n).$$

An algorithm φ_n^* using information of cardinality at most n for which

$$e(\varphi_n^*) = e(n)$$

is said to be an nth *minimal error algorithm.*

We now determine nth minimal error, optimal information, and a minimal error algorithm. Recall that for a balanced convex subset X of a Hilbert space H, the (Kolmogorov) *n-width* of X in H is given by

$$d_n(X, H) = \inf_{H_n} \sup_{x \in X} \inf_{h \in H_n} \|x - h\|_H,$$

the infimum being over all subspaces H_n of H whose dimension does not exceed n. We then have the following result from [13, Chaps. 2 and 3]:

LEMMA 3.2. (i) *The nth minimal error is given by*

$$e(n) = d_n(SF, H^s(\partial)).$$

(ii) *If $r + t = s$ (which can happen if and only if $r = 0$ and $s = t$), then there exists $\varepsilon_0 > 0$ such that*

$$\lim_{n \to \infty} e(n) = \varepsilon_0.$$

(iii) *If $r + t > s$, let $E\colon H^t(\partial) \to H^s(\partial)$ be the inclusion operator, so that ES is compact. Let $\{e_j\}_{j=1}^{\infty}$ be an orthonormal basis of $H^r(\Omega)$ consisting of eigenvectors of $K = (ES)^*(ES)$, with*

$$Ke_j = \lambda_j e_j$$

$$\lambda_1 \geqslant \lambda_2 \geqslant \dots > 0 \qquad \text{with } \lim_{j \to \infty} \lambda_j = 0.$$

Then

$$e(n) = \sqrt{\lambda_{n+1}},$$

the information

$$\mathcal{N}_n^* f = [(f, e_1)_r \cdots (f, e_n)_r]^T \qquad \forall f \in H^r(\Omega)$$

is nth optimal information, and

$$\varphi_n^*(\mathcal{N}_n^* f) = \sum_{j=1}^{n} (f, e_j)_r \, Se_j \qquad \forall f \in H^r(\Omega)$$

is an nth minimal error algorithm.

The first statement in this lemma gives the nth minimal error as a Kolmogorov n-width. The second implies that there is *no* algorithm whose error is less than ε_0 if $r+t=s$. The third tells us that if $r+t>s$, then $\lim_{n\to\infty} e(n)=0$.

Although we have explicit formulas for optimal information and algorithms, as well as minimal error algorithms, these may be difficult to determine in practice, since they require knowledge of S at the eigenvectors of K. For this reason, we will be willing to settle for *quasi-optimality* [14], i.e., optimality to within a constant which is independent of the cardinality of the information; *quasi-minimal error algorithms* are defined analogously. As a benchmark for establishing quasi-optimality, we now establish an estimate of $e(n)$ using techniques of [16]. The result is phrased in terms of Knuth's big-theta notation [10]:

THEOREM 3.1. *$e(n)=\Theta(n^{-(r+t-s)/N})$ as $n\to\infty$.*

Proof. For $\theta>0$, let

$$X(\theta)=\theta BH^{r+t}(\partial)=\{u\in H^{r+t}(\partial): \|u\|_{r+t}\leqslant\theta\}.$$

Lemma 2.1 yields

$$X(\sigma^{-1})\subseteq SF\subseteq X(\sigma).$$

Since for any $\theta>0$,

$$d_n(X(\theta), H^s(\partial))=\theta\, d_n(X(1), H^s(\partial)),$$

the first statement in Lemma 3.2 yields that

$$\sigma^{-1}\leqslant\frac{e(n)}{d_n(BH^{r+t}(\partial), H^s(\partial))}\leqslant\sigma.$$

Using [2, Theorem 2.5.1] and the results of [8], we have

$$d_n(BH^{r+t}(\partial), H^s(\partial))=\Theta(d_n(BH_0^1(\Omega), L_2(\Omega))^{r+t-s})=\Theta(n^{-(r+t-s)/N}),$$

completing the proof. ∎

4. Optimality of Finite Elements for Elliptic Systems

In this section, we define the (least-squares) finite element information (FEI) of degree k and the (least-squares) finite-element method (FEM) using FEI. We show that the FEM is a quasi-minimal error algorithm iff $k \geqslant r+t-1$, while the FEI is always quasi-optimal information. We use the notation and terminology of [4, 7, 11].

Let k be a non-negative integer. Let $\mathscr{T}_n$ be a triangulation of Ω and let $\mathscr{V}_n$ be an n-dimensional subspace of $H^t(\partial)$ consisting of functions which are piecewise polynomial of degree k with respect to the triangulation $\mathscr{T}_n$. (Of course, there is a problem in that such functions cannot in general satisfy the boundary conditions; this may be handled by using curved elements [8] or isoparametric elements [7] on the boundary, or by using the techniques found in [5, 15].) We assume that the family $\{\mathscr{T}_n\}_{n=1}^{\infty}$ is *quasi-uniform* [11, p. 272].

In what follows, we assume that

$$k \geqslant 2t-1-s. \tag{4.1}$$

See [10, Remark 4.1] for further discussion.

We recall the definition of the least-squares finite element method [5] as applied to systems [2, 3, 15]. Let $f \in H^r(\Omega)$. For each positive integer n, we seek an approximation $u_n \in \mathscr{V}_n$ to u such that

$$\|f-lu_n\|_0 = \min\{\|f-lv_n\|_0 : v_n \in \mathscr{V}_n\},$$

i.e., $u_n \in \mathscr{V}_n$ satisfies

$$(lu_n, lv_n)_0 = (f, lv_n)_0 \qquad \forall v_n \in \mathscr{V}_n.$$

Letting $\{w_1,..., w_n\}$ denote a basis for $\mathscr{V}_n$, define the (least-squares) *finite element information* (FEI) $\mathscr{N}_n$ by

$$\mathscr{N}_n f = [(f, lw_1)_0 \dots (f, lw_n)_0]^T \qquad \forall f \in H^r(\Omega).$$

Then the (least-squares) *finite element method* (FEM) $\varphi_n \in \Phi(\mathscr{N}_n)$ is given by

$$\varphi_n(\mathscr{N}_n f) = u_n.$$

Since the basis functions are linearly independent and l is injective, it is easy to see that φ_n is a well-defined linear algorithm using $\mathscr{N}_n$.

We now compute the error of the FEM.

Theorem 4.1. *Let*

$$\mu = \min(k+1-t, r).$$

Then

$$e(\varphi_n) = \Theta(n^{-(\mu+t-s)/N}) \qquad as \quad n \to \infty,$$

and so $\{\varphi_n\}_{n=1}^{\infty}$ *is a sequence of quasi-minimal error algorithms iff*

$$k \geqslant r + t - 1. \tag{4.2}$$

Proof. We first show the lower bound for the error. If (4.2) holds, then $\mu = r$, and so Theorem 3.1 yields

$$e(\varphi_n) \geqslant e(n) = \Theta(n^{-(\mu+t-s)/N}) \qquad \text{as} \quad n \to \infty.$$

We now suppose (4.2) does not hold, so that $\mu = k + 1 - t$. Using an N-dimensional version of the proof of [16, Theorem 5.2] there exists a non-zero function $u^* \in H^{r+t}(\partial)$, a positive constant C, and a positive integer n_0, such that

$$\inf_{v_n \in \mathcal{V}_n} \|u^* - v_n\|_s \geqslant C n^{-(\mu+t-s)/N} \qquad \forall n \geqslant n_0.$$

Since u^* is non-zero, lu^* is also non-zero. Let $f^* = lu^*/\|lu^*\|_r$. Then $\|f^*\|_r = 1$, so that $f^* \in F$. Since φ_n is linear with range $\mathcal{V}_n$, the previous estimate yields that

$$e(\varphi_n) \geqslant \|Sf^* - \varphi_n(\mathcal{N}_n f^*)\|_s = \frac{1}{\|lu^*\|_r} \|u^* - \varphi_n(\mathcal{N}_n lu^*)\|_s$$

$$\geqslant \frac{1}{\|lu^*\|_r} \inf_{v_n \in \mathcal{V}_n} \|u^* - v_n\|_s \geqslant \frac{C}{\|lu^*\|_r} n^{-(\mu+t-s)/N},$$

completing the proof of the lower bound.

We now establish the upper bound. Let $f \in F$. By (4.1) and (4.2), there exists $C > 0$, independent of f, such that (setting $u = Sf$)

$$\|u - u_n\|_s \leqslant C n^{-(\mu+t-s)/N} \|u\|_{r+t} \qquad \forall n \geqslant 1.$$

(See [15, Chap. 8] for the case $t = 1$, and the references cited therein for the case of arbitrary t.) Hence Lemma 2.2 yields

$$\|Sf - \varphi_n(\mathcal{N}_n f)\|_s = \|u - u_n\|_s \leqslant C n^{-(\mu+t-s)/N} \|u\|_{r+t} \leqslant C\sigma n^{-(\mu+t-s)/N} \|f\|_r.$$

Since $f \in F$ is arbitrary, we have

$$e(\varphi_n) \leqslant C\sigma n^{-(\mu+t-s)/N},$$

completing the proof of the first part of the theorem.

The remainder of the theorem now follows from the first part and from Theorem 3.1. ▌

Hence the FEM is (roughly) a minimal error algorithm iff (4.2) holds. Suppose (4.2) is violated. We show that the non-optimality of the FEM is due to the fact that it uses the FEI inefficiently, rather than being inherent in the FEI itself.

We first establish two intermediate results.

LEMMA 4.1. *There exists $\sigma \geqslant 1$ such that*

$$\|lw\|_{-r} \leqslant \sigma \|w\|_{t-r} \qquad \forall w \in H^t(\partial).$$

Proof. If $r=0$, this follows from Lemma 2.1. Once the result is shown for $r \geqslant t$, it then holds for $0<r<t$ by Hilbert space interpolation [6] of the results for the cases $r=0$ and $r=t$. So, we assume $r \geqslant t$ without loss of generality. Let $w \in H^t(\partial)$. For any $v \in C_0^\infty(\Omega)$, we may use Lemma 2.1 (with r replaced by the non-negative real number $r-t$) to see that

$$|(lw, v)|_0 = |(w, l^+ v)_0| \leqslant \|w\|_{t-r} \|l^+ v\|_{r-t} \leqslant \sigma \|w\|_{t-r} \|v\|_r.$$

Hence

$$\|lw\|_{-r} = \sup \left\{ \frac{|(lw, v)_0|}{\|v\|_r} : v \in C_0^\infty(\Omega),\ v \neq 0 \right\} \leqslant \sigma \|w\|_{t-r},$$

as required. ▌

LEMMA 4.2. *For $g \in C_0^\infty(\Omega)$, let $v \in C^\infty(\Omega)$ be the solution of*

$$l^+ v = g \ \text{in } \Omega, \qquad b^+ v = 0 \ \text{on } \partial\Omega. \tag{4.3}$$

Then there is a constant $\sigma \geqslant 1$, independent of g and w, such that

$$\|v\|_{t-s} \leqslant \sigma \|g\|_{-s}.$$

Proof. By (ii) of Lemma 2.2 (with $r=0$), we find

$$\|v\|_t \leqslant \sigma \|g\|_0. \tag{4.4}$$

We next claim that

$$\|v\|_0 \leqslant \sigma \|g\|_{-t}; \tag{4.5}$$

indeed, (i) of Lemma 2.2 yields

$$\sigma^{-1} \|v\|_0 \|Sv\|_t \leqslant \|v\|_0^2 = |(lSv, v)_0| = |(Sv, g)_0| \leqslant \|Sv\|_t \|g\|_{-t},$$

which implies (4.5). The result now follows by Hilbert space interpolation of (4.4) and (4.5). ∎

We now show that FEI is quasi-optimal, regardless of whether (4.2) holds. Let φ_n^s denote the spline algorithm using the FEI N_n^* (see Lemma 3.1).

THEOREM 4.2. $e(\varphi_n^s) = e(N_n^*) = \Theta(n^{-(r+t-s)/N})$ as $n \to \infty$.

Proof. The first equaity follows from Lemma 3.1. We now establish the second. For the lower bound, note that card $N_n^* = n$, and so

$$e(N_n^*) \geqslant e(n) = \Theta(n^{-(r+t-s)/N}) \qquad \text{as} \quad n \to \infty.$$

We now estabish the upper bound. Let $z \in F \cap \ker N_n^*$, so that

$$(z, lv_n)_0 = 0 \qquad \forall v_n \in \mathcal{S}_n'$$

and

$$\|z\|_r \leqslant 1.$$

Let $g \in C_0^\infty(\Omega)$ be non-zero, and choose $v \in C^\infty(\Omega)$ satisfying (4.3). Then for any $v_n \in \mathcal{S}_n'$, we have

$$\begin{aligned} |(Sz, g)_0| &= |(Sz, l^{-1}v)_0| = |(z, v)_0| = |(z, l(Sv - v_n))_0| \\ &\leqslant \|l(Sv - v_n)\|_{-r} \leqslant \sigma \|Sv - v_n\|_{t-r} \end{aligned}$$

by Lemma 4.1. Since (4.1) holds, standard approximation-theoretic results [4, 7] imply that there exists a positive constant C (independent of z, g, v, and n) and $v_n \in \mathcal{S}_n'$ such that

$$\|Sv - v_n\|_{t-r} \leqslant Cn^{-(r+t-s)/N} \|Sv\|_{2t-s}.$$

But (i) of Lemma 2.2 and Lemma 4.2 imply that

$$\|Sv\|_{2t-s} \leqslant \sigma \|v\|_{t-s} \leqslant \sigma^2 \|g\|_{-s}.$$

Combining the three previous inequalities, we see that there is (another) positive constant C (independent of z, g, and n) such that

$$|(Sz, g)_0| \leqslant Cn^{-(r+t-s)/N} \|g\|_{-s}.$$

Since g is an arbitrary element of C_0^∞, we have

$$\|Sz\|_s = \sup\left\{\frac{|(Sz, g)_0|}{\|g\|_{-s}} : g \in C_0^\infty(\Omega),\ g \neq 0\right\} \leqslant Cn^{-(r+t-s)/N}.$$

Taking the supremum over all $z \in F \cap \ker \mathcal{N}_n'$, we have

$$e(\mathcal{N}_n') \leqslant C n^{-(r+t-s)/N},$$

completing the proof of the theorem. ▮

5. Complexity Analysis

In this section, we discuss the complexity of finding ε-approximations to the solution of the elliptic system, as well as the penalty for using the FEM when $k < t - 1 + r$.

Let $\varepsilon > 0$. An algorithm $\varphi \in \Phi(\mathcal{N})$ produces an *ε-approximation* if

$$e(\varphi) \leqslant \varepsilon.$$

The *complexity*, comp(φ), of an *algorithm* $\varphi \in \Phi(\mathcal{N})$ is defined via the model of computation discussed in [13, Chap. 5]. (Informally, we assume that any linear functional can be evaluated with finite cost c_1, and that the cost of an arithmetic operation is unity.) It then turns out that if $\mathcal{N}$ has cardinality n, then

$$\operatorname{comp}(\varphi) \geqslant nc_1 + n - 1 \qquad \forall \varphi \in \Phi(\mathcal{N}), \tag{5.1}$$

while if φ is linear, then

$$\operatorname{comp}(\varphi) \leqslant nc_1 + 2n - 1; \tag{5.2}$$

see [13, Chap. 5, Section 2] for details. We then define, for $\varepsilon > 0$, the *ε-complexity* of the *problem* to be

$$\operatorname{COMP}(\varepsilon) = \inf\{\operatorname{comp}(\varphi) \colon e(\varphi) \leqslant \varepsilon\}.$$

If φ^* is an algorithm for which

$$e(\varphi^*) \leqslant \varepsilon \qquad \text{and} \qquad \operatorname{comp}(\varphi^*) = \operatorname{COMP}(\varepsilon),$$

then φ^* is said to be an *optimal complexity algorithm* for ε-approximation of the problem.

Remark 5.1. Note the distinction between *algorithmic complexity*, which is the cost of using a *particular algorithm* to solve the problem to within a tolerance of ε, and *problem complexity*, which is the *inherent* cost of solving the problem to within ε. ▮

Remark 5.2. Not surprisingly, it is difficult to determine optimal complexity algorithms. We will generally be willing to settle for optimality to within a constant factor, independent of ε. Hence, we say that a family $\{\varphi_\varepsilon^*\}_{\varepsilon>0}$ of algorithms has *quasi-minimal complexity* for the problem if

$$e(\varphi_\varepsilon^*) \leqslant \varepsilon \qquad \text{for all sufficiently small } \varepsilon > 0$$

and

$$\mathrm{comp}(\varphi_\varepsilon^*) = \Theta(\mathrm{COMP}(\varepsilon)) \qquad \text{as} \quad \varepsilon \to 0. \quad \blacksquare$$

Recall that φ_n denotes the finite element method of degree k using the finite element information $\mathcal{N}_n$ based on the finite element subspace $\mathcal{V}_n$, and that φ_n^s denotes the spline algorithm using this information. We let

$$\mathrm{FEM}(\varepsilon) := \inf\{\mathrm{comp}(\varphi_n) \colon e(\varphi_n) \leqslant \varepsilon\}$$

denote the algorithmic complexity of the FEM, and let

$$\mathrm{SPLINE}(\varepsilon) := \inf\{\mathrm{comp}(\varphi_n^s) \colon e(\varphi_n^s) \leqslant \varepsilon\}$$

denote the algorithmic complexity of the spline algorithm using the FEI. Using the results of Section 4, (5.1), and (5.2), we have

THEOREM 5.1. *The problem complexity is*

$$\mathrm{COMP}(\varepsilon) = \Theta(\varepsilon^{-N/(r+t-s)}) \qquad \textit{as} \quad \varepsilon \to 0.$$

The algorithmic complexity of the spline algorithm is

$$\mathrm{SPLINE}(\varepsilon) = \Theta(\varepsilon^{-N/(r+t-s)}) \qquad \textit{as} \quad \varepsilon \to 0.$$

The algorithmic complexity of the finite element method is

$$\mathrm{FEM}(\varepsilon) = \Theta(\varepsilon^{-N/(\mu+t-s)}) \qquad \textit{as} \quad \varepsilon \to 0,$$

where $\mu = \min(k+1-t, r)$.

Hence, we may draw the following conclusions:

THEOREM 5.2. (i) *The spline algorithm using the FEI is quasi-optimal.*

(ii) *The FEM is quasi-optimal iff* $k \geqslant t+1-r$.

(iii) *Let*

$$\mathrm{pen}(\varepsilon) = \frac{\mathrm{FEM}(\varepsilon)}{\mathrm{COMP}(\varepsilon)}$$

denote the penalty for using the FEM instead of a quasi-optimal algorithm using the same information. If $k < t-1+r$, *then*

$$\mathrm{pen}(\varepsilon) = \Theta(1/\varepsilon^{\lambda N}) \qquad \textit{as} \quad \varepsilon \to 0,$$

where

$$\lambda = \frac{1}{k+1-s} - \frac{1}{r+t-s} = \frac{r-\mu}{(k+1-s)(r+t-s)} > 0,$$

and so

$$\lim_{\varepsilon \to 0} \text{pen}(\varepsilon) = \infty.$$

Thus there is an infinite asymptotic penalty (as $\varepsilon \to 0$) for using the FEM when $k < t-1+r$, rater than the spline algorithm which uses the same information as does the FEM.

Remark 5.3. One of the assumptions in the model of computation used in [13] is that computation of *any* linear functional is allowed, and has finite cost c_1. This holds if *pre-conditioning* is allowed. That is, given an algorithm, any computations which are independent of the problem element f may be done in *advance*, and their cost is not counted when determining the complexity of that algorithm. In particular, this means that when measuring the complexity of the FEM, we do not count the cost of factoring the coefficient matrix which appears when the algorithm is reduced to the solution of a linear system of equations. (This is because the coefficient matrix is independent of the problem element f.) In many situations, this is not a realistic assumption. In such cases, the FEM is no longer quasi-optimal from the viewpoint of minimizing complexity (even when $k \geqslant t-1+r$). It is perhaps possible that multi-grid techniques may be used to transform the FEM into a method which has quasi-optimal complexity in situations where pre-conditioning is not allowed. However, no matter what model of computation is used, the quasi-minimal error properties described in Section 4 still hold, since they are independent of any particular model of computation. ∎

ACKNOWLEDGMENT

I thank Professor J. F. Traub (Columbia University) for his helpful remarks.

REFERENCES

1. S. AGMON, A. DOUGLIS, AND L. NIRENBERG, Estimates near the boundary for higher-order elliptic partial differential equations satisfying general boundary conditions, *Comm. Pure Appl. Math.* **12** (1959), 623–727.
2. A. K. AZIZ AND S. LEVENTHAL, Finite element approximation for first-order systems, *SIAM J. Numer. Anal.* **15** (1978), 1103–1111.

3. A. K. Aziz and A. G. Werschulz, On the numerical solution of Helmholtz's equation by the finite element method, *SIAM J. Numer. Anal.* **17** (1980), 681–686.
4. I. Babuška and A. K. Aziz, Survey lectures on the mathematical foundations of the finite element method, *in* "The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations," (A. K. Aziz, Ed.), pp. 3–359, Academic Press, New York, 1972.
5. J. H. Bramble and A. H. Schatz, Least squares methods for $2m$th order elliptic boundary-value problems, *Math. Comp.* **25** (1971), 1–32.
6. P. L. Butzer and H. Berens, "Semi-groups of Operators and Approximation," Springer-Verlag, Berlin, 1967.
7. P. G. Ciarlet, "The Finite Element Method for Eliptic Problems," North-Holland, Amsterdam, 1976.
8. P. G. Ciarlet and P. A. Raviart, Interpolation theory over curved elements, *Comput. Methods Appl. Mech. Engrg.* **1** (1972), 217–249.
9. J. W. Jerome, Asymptotic estimates of the L_2 n-width, *J. Math. Anal. Appl.* **22** (1968), 449–464.
10. D. E. Knuth, Big omicron and big omega and big thta, *SIGACT News*, Association for Computing Machinery, April, 1976.
11. J. T. Oden and J. N. Reddy, "An Introduction to the Mathematical Theory of Finite Elements," Wiley–Interscience, New York, 1976.
12. J. Roĭtberg and Z. Šeftel, A theorem on homeomorphisms for elliptic systems and its applications, *Math. USSR-Sb.* **7** (1969), 439–465.
13. J. F. Traub and H. Woźniakowski, "A General Theory of Optimal Algorithms," Academic Press, New York, 1980.
14. L. B. Wahlbin, Quasi-optimality of the $\dot{H}^1$ projection into finite element spaces, *in* "Lectures on the Numerical Solution of Partial Differential Equations: Proceedings of the Special Year in Numerical Analysis" (I. Babuška, T.-P. Liu, and J. Osborn, Eds.), Lecture Notes 20, pp. 420–434, Department of Mathematics, University of Maryland, College Park, Md., 1981.
15. W. Wendland, "Elliptic Systems in the Plane," Pitman, London, 1979.
16. A. G. Werschulz, "Complexity of Indefinite Elliptic Problems," *Math. Comp.*, to appear.

Printed in Belgium