Complexity of Indefinite Elliptic Problems

By Arthur G. Werschulz*

Abstract. This paper deals with the approximate solution of a linear regularly-elliptic $2m$th-order boundary-value problem $Lu = f$, with $f \in H^r(\Omega)$ for $r \geq -m$. Suppose that the problem is indefinite, i.e., the variational form of the problem involves a weakly-coercive bilinear form. Of particular interest is the quality of the finite element method (FEM) of degree k using n inner products of f. The error of the approximation is measured in the Sobolev l-norm ($0 \leq l \leq m$); we assume that $k \geq 2m - 1 - l$. We assume that an *a priori* bound is known for either the Sobolev r-norm or for the Sobolev r-seminorm of f. We first consider the normed case. We find that the FEM has minimal error if and only if $k \geq 2m-1+r$. Regardless of the values of k and r, there exists a linear combination (called the spline algorithm) of the inner products used by the FEM which *does* have minimal error. For the seminormed case, we give a very restrictive condition which is necessary and sufficient for the error of the FEM to have a bound which is independent of f. When this condition holds, we find that the FEM has minimal error if and only if $k \geq 2m - 1 + r$. However, we once again find that the spline algorithm (using the same inner products as does the FEM) has minimal error, no matter what values k and r have and regardless of whether the FEM has uniformly bounded error. We also show that the inner products used by the FEM is the best set of linear functionals to use.

1. Introduction. This paper is a theoretical study of the approximate solution of the elliptic boundary-value problem $Lu = f$. We assume the problem is linear regularly-elliptic of $2m$th order with homogeneous boundary conditions, and that f is defined on a region $\Omega \subset \mathbf{R}^N$. The smoothness of f will be given by the assumption that $f \in H^r(\Omega)$, where $r \geq -m$. (See Section 2.) We want to approximate the solution u by using the values of n linear functionals of f, a typical example being the values of f at n points in Ω. (See Section 3.)

We address two problems:

(i) Given the values of n linear functionals of f, how may they be combined so as to approximate the solution of the elliptic problem with the smallest possible error?

(ii) What is the best set of n linear functionals to use?

This paper is an extension of [15], where this problem was considered under the following conditions:

(i) The elliptic problem $Lu = f$ is *definite*, i.e., its variational form involves a coercive bilinear form.

(ii) The error of the approximate solution is measured by the energy norm, which is equivalent to the Sobolev m-norm.

(iii) An *a priori* bound is given on the Sobolev r-norm of f.

In [15], we studied the finite element method (FEM) of degree k, which uses n inner products $(f, s_1), \ldots, (f, s_n)$, where $\{ s_1, \ldots, s_n \}$ is a basis for a finite element subspace of degree k and $(\cdot, \cdot)$ is the $L_2(\Omega)$-inner product. (See Section 4.) Here, we must assume $k \geq m$; otherwise, the FEM is not well-defined [15, Lemma 4.1]. The main result of [15] is that the FEM has minimal error (to within a constant) if and only if $k \geq 2m - 1 + r$. We also prove that the set $\{ (f, s_1), \ldots, (f, s_n) \}$ of inner products is always (to within a constant) the best choice of linear functionals. This holds for $\{ s_1, \ldots, s_n \}$ a basis for a finite element subspace of degree k, for any value of $k \geq m$ and any value of r.

How crucial are the conditions in [15]?

Condition (i) disallows problems such as the *Helmholtz problem*: given $f \in H^r(\Omega)$, find $u \colon \overline{\Omega} \to \mathbf{R}$ such that

$$
\begin{aligned}
\Delta u + \lambda u &= f \quad \text{in } \Omega, \\
u &= 0 \quad \text{on } \partial\Omega,
\end{aligned}
$$

(1.1)

where λ is not an eigenvalue of $-\Delta$. If λ is bigger than the smallest eigenvalue of $-\Delta$, this problem does not yield a coercive bilinear form.

Condition (ii) is that the energy norm was used. Although this norm is equivalent to the Sobolev m-norm, the constant which measures this equivalence may be so large that a very good energy-norm solution may not be sufficiently accurate in the m-norm. (This would appear to be the situation in boundary-layer problems.) Moreover, it is sometimes of greater interest to use other norms, such as the $L_2(\Omega)$-norm for measuring displacement error.

Condition (iii) is needed to make the error of the approximate solution finite, and is implicitly used as a standard assumption for partial differential equations and finite element methods [1], [3], [8]. However, instead of assuming that the Sobolev r-norm of f is bounded, one may only wish to assume an *a priori* bound on the Sobolev r-seminorm of f.

In this paper, the results of [15] are extended by weakening conditions (i), (ii), and (iii) above. We assume:

(i)$'$ The elliptic problem $Lu = f$ is *indefinite*: its variational form involves a weakly coercive bilinear form.

(ii)$'$ Error is measured in the Sobolev l-norm, where $0 \leq l \leq m$.

(iii)$'$ Both the *normed* and *seminormed cases* are considered, i.e., an *a priori* bound is given on the Sobolev r-norm or r-seminorm of $f \in H^r(\Omega)$ (respectively), the seminormed case making sense if and only if r is a nonnegative integer.

Again, we will be interested in whether the FEM of degree k has minimal error, and whether the inner products used by the FEM form the best set of linear functionals. We assume in this paper that

(1.2) $k \geq 2m - 1 - l$;

see Section 3 for further discussion.

It turns out that in the normed case, replacing (i) by (i)$'$ causes almost no difficulty, while replacing (ii) by (ii)$'$ may be done via a variant of the Aubin-Nitsche duality argument [5, pp. 136–139]. Hence, in the normed case, the main results of [15] still hold; see Section 4 for details.

The situation is different for the seminormed case. As an example, consider the problems

(1.3) $-u'' + u = f$ in $(0, 1)$, $u(0) = u(1) = 0$,

and

(1.4) $-u'' + u = f$ in $(0, 1)$, $u'(0) = u'(1) = 0$,

where

(1.5) $f \in H^1(0, 1)$ and $\int_0^1 \left(f'(x) \right)^2 dx \leq 1$.

For the problem (1.3), the error of the FEM is not uniformly bounded over all f satisfying (1.5), no matter what the value of k is. Despite this, there exists a linear combination of the inner products used by the FEM whose error *is* uniformly bounded over all such f. This linear combination is called the spline algorithm; see Section 3 for further discussion. In fact, the error of the spline algorithm goes to zero, independently of f, as the number of inner products goes to infinity. Moreover, there is no better choice of linear functionals for the problem (1.3) than the inner products which are used by the FEM.

For the problem (1.4), the situation is quite different. The FEM always has finite error and, when $k \geq 2$, its error is almost minimal. Once again, the inner products used by the FEM are (to within a constant factor) the best linear functionals, regardless of the value of $k \geq 1$.

We discuss the general seminormed case in detail in Section 5. We show that the FEM has finite error if and only if $P_{r-1}(\Omega) \subseteq LS_n$, where $P_{r-1}(\Omega)$ is the restriction to Ω of the space of polynomials of degree at most $r - 1$, and S_n is the finite element subspace of dimension n. This condition is very restrictive; see Remark 6.1. When this condition holds, we find that the error of the FEM is (to within a constant factor) minimal if and only if $k \geq 2m - 1 + r$. Regardless of whether the error of the FEM is minimal, there is a linear combination (called the spline algorithm) of the inner products used by the FEM whose error is minimal. Moreover, the inner products used by the FEM are (to within a constant factor) the best linear functionals, regardless of the value of $k \geq m$, and regardless of whether $P_{r-1}(\Omega) \subseteq LS_n$.

In Section 6, we discuss the complexity (i.e., the minimal cost) of obtaining approximate solutions whose error is less than ε. In the normed case, the FEM produces an approximation with minimal cost if and only if $k \geq 2m - 1 + r$. If, however, $k < 2m - 1 + r$, the penalty for using the FEM (instead of the spline algorithm, which always produces an approximation with minimal cost) is unbounded as $\varepsilon \to 0$, for both the normed and seminormed cases. Since this is an asymptotic measure, we also wish to know whether there is a penalty for using the FEM for fixed ε of moderate size when $k < 2m - 1 + r$. We consider a simple model problem, and show that the complexity of the spline algorithm is less than that of the FEM whenever $\varepsilon < \varepsilon_0$, where $\varepsilon_0 \doteq 0.227$.

Finally, in Section 7, we summarize our work, and point out some possible extensions and open questions.

2. The Variational Boundary-Value Problem. In what follows, we use the standard notation for Sobolev spaces, inner products, and norms, for multi-indices, etc., found in [5]. Fractional- and negative-order Sobolev spaces are defined by

Hilbert-space interpolation and duality, respectively; see [3, Chapter 2] and [8, Chapter 4] for details.

Let $\Omega \subset \mathbf{R}^N$ be a bounded, simply connected, C^∞ region. Define the uniformly strongly elliptic operator

$$Lv = \sum_{|\alpha|,|\beta| \leq m} (-1)^{|\alpha|} D^\alpha (a_{\alpha\beta} D^\beta v),$$

with real coefficients $a_{\alpha\beta} \in C^\infty(\overline{\Omega})$ such that $a_{\alpha\beta} = a_{\beta\alpha}$. In order to have appropriate boundary conditions, let

$$B_j v = \sum_{|\alpha| \leq q_j} b_{j\alpha} D^\alpha v \qquad (0 \leq j \leq m - 1),$$

where $b_{j\alpha} \in C^\infty(\partial\Omega)$ are real-valued and

$$0 \leq q_0 \leq \cdots \leq q_{m-1} \leq 2m - 1.$$

We assume that $\{ B_j \}_{j=0}^{m-1}$ is a normal family of operators which covers L on $\partial\Omega$. To make the boundary-value problem be selfadjoint, we let

$$m^* = \min\{ j \colon q_j \geq m \},$$

and require that

$$\{ q_j \}_{j=0}^{m^*-1} \cup \{ 2m - 1 - q_j \}_{j=m^*}^{m-1} = \{ 0, \ldots, m - 1 \}.$$

(See [3, Chapter 3] and [8, Chapter 5] for further definitions and illustrative examples.) We are interested in solving the *elliptic boundary-value problem*:

Given $f \in H^r(\Omega)$, where $r \geq -m$, find $u \colon \overline{\Omega} \to \mathbf{R}$ such that

(2.1)
$$Lu = f \quad \text{in } \Omega,$$
$$B_j u = 0 \quad \text{on } \partial\Omega \qquad (0 \leq j \leq m - 1).$$

Let

$$H_E^m(\Omega) = \{ v \in H^m(\Omega) \colon B_j v = 0 \, (0 \leq j \leq m^* - 1) \}$$

denote the space of $H^m(\Omega)$-functions satisfying the essential boundary conditions. We define a symmetric, continuous bilinear form B on $H_E^m(\Omega)$ by

$$B(v, w) = \sum_{|\alpha|,|\beta| \leq m} \int_\Omega a_{\alpha\beta} D^\alpha v D^\beta w \, .$$

In [15], we assumed that B was $H_E^m(\Omega)$-*coercive* [8], i.e., that there exists a positive constant γ such that

$$B(v, v) \geq \gamma \|v\|_m^2 \qquad \forall v \in H_E^m(\Omega).$$

When $m = 0$, the conditions on L yield that B is $L_2(\Omega)$-coercive. However, for $m \geq 1$, there exist elliptic boundary-value problems which do not yield a bilinear form that is $H_E^m(\Omega)$-coercive (such as the Dirichlet problem for the Helmholtz equation).

In this paper, we assume instead that B is *weakly $H_E^m(\Omega)$-coercive* [8, p. 310]. Since B is symmetric, this means that there exists a positive constant γ such that:

For any nonzero $v \in H_E^m(\Omega)$, there exists nonzero $w \in H_E^m(\Omega)$ such that

(2.2) $|B(v, w)| \geq \gamma \|v\|_m \|w\|_m \, .$

The following lemma gives a condition which is sufficient to establish weak coercivity. The result appears to be well-known; its proof for arbitrary m is a straightforward modification of the proof for the case $m = 1$ which is found in [3, Chapter 5].

LEMMA 2.1. *Let $m \geq 1$. Suppose that*
(i) *the only solution of* (2.1) *with $f = 0$ is $u = 0$, and*
(ii) *B is $[H_E^m(\Omega), L_2(\Omega)]$-coercive* [8, p. 301]; *that is, there exist $\gamma_0 > 0$ and $\gamma_1 \geq 0$ such that*

$$(2.3) \qquad |B(v,v)| \geq \gamma_0 \|v\|_m^2 - \gamma_1 \|v\|_0^2 \quad \forall v \in H_E^m(\Omega).$$

Then B is weakly $H_E^m(\Omega)$-coercive.   □

Remark 2.1. Suppose that B_j is the jth normal derivative operator ($0 \leq j \leq m - 1$), so that (2.1) is a Dirichlet problem and $H_E^m(\Omega) = H_0^m(\Omega)$. Then (2.3) is Gårding's inequality [1], which follows immediately from the conditions on L. Hence, B is weakly $H_0^m(\Omega)$-coercive provided that (i) holds in Lemma 2.1. For example, the Helmholtz problem (1.1) is weakly $H_0^1(\Omega)$-coercive if λ is not an eigenvalue of $-\Delta$.   □

We now define the *variational boundary problem* as follows. Let $r \geq -m$. We wish to solve the problem:

Given $f \in H^r(\Omega)$, find $u = Sf \in H_E^m(\Omega)$ such that

$$(2.4) \qquad B(u,v) = (f,v)_0 = \int_\Omega fv \quad \forall v \in H_E^m(\Omega).$$

From the Generalized Lax-Milgram Theorem [3, Theorem 5.2.1], S is a Hilbert space isomorphism of $H^{-m}(\Omega)$ onto $H_E^m(\Omega)$, and so $S \colon H^r(\Omega) \to H_E^m(\Omega)$ is a bounded linear injection. Since B is only assumed to be weakly coercive (i.e., we do not know that $B(v,v) \geq \gamma \|v\|_m^2$ holds), the problem (2.4) is said to be *indefinite*.

It is useful to recall the "shift theorem" ([3, Chapter 3], [8, Chapter 5]), which states that since $f \in H^r(\Omega)$, we have $Sf \in H_E^m(\Omega) \cap H^{2m+r}(\Omega)$, and that there exists a positive constant σ, which is independent of f, such that

$$(2.5) \qquad \sigma^{-1} \|Sf\|_{2m+r} \leq \|f\|_r \leq \sigma \|Sf\|_{2m+r}.$$

If $r > N/2$, then the shift theorem, Sobolev's embedding theorem, and an m-fold integration by parts yield that $u = Sf$ is the solution to (2.1).

3. Finite Element Methods and Information. In this section, we discuss the finite element method (FEM). We describe the finite element information (FEI) which the FEM uses, and recall some quasi-optimality results from [15] for the case where assumptions (i), (ii), and (iii) from Section 1 hold.

We let $\{ S_n \}_{n=1}^\infty$ be a regular family of *finite element subspaces* of *degree k*. That is, S_n is an n-dimensional subspace of $H_E^m(\Omega)$ consisting of piecewise polynomials of degree k over a triangulation $\mathcal{T}_n$ of Ω, where $\{ \mathcal{T}_n \}_{n=1}^\infty$ is regular [8, p. 132]. Of course, since Ω is C^∞, we must make an additional assumption about boundary elements to guarantee that $S_n \subseteq H_E^m(\Omega)$ in the situation where (2.1) is not a Neumann problem. (For instance, we may use curved elements as in [6].)

Remark 3.1. As indicated in (1.2) of the introduction, we assume that $k \geq 2m - 1 - l$ in this paper. This technical assumption is needed for the proofs of some upper bounds which appear below. However, there are a number of situations where this holds automatically:

(i) If $\max\{ 0, m-1 \} \leq l \leq m$, then (1.2) holds by [15, Lemma 4.1]. In particular, (1.2) holds when $m = 0$ (where we have $l = 0$) or when $m = 1$ (in which case $l = 0$ or $l = 1$).

(ii) If $N \geq 2$ and triangular elements are used, the results of [16] and the fact that $l \geq 0$ show that (1.2) holds.

Hence, the only possibility that (1.2) will fail is when $N = 1$ or when rectangular elements are used. However, it *is* possible for $k = m$ in either of these settings. Thus (1.2) can indeed fail to hold when $N = 1$ or when rectangular elements are used. □

The *finite element method* (FEM) using $\{ S_n \}_{n=1}^{\infty}$ is defined as follows.

Let $f \in H^r(\Omega)$. Let $\{ s_1, \ldots, s_n \}$ be a basis for S_n. Given n inner products

$$\mathcal{N}_n f = \begin{bmatrix} (f, s_1)_0 \\ \vdots \\ (f, s_n)_0 \end{bmatrix},$$

choose $u_n \in S_n$ such that

$$(3.1) \qquad B(u_n, s_i) = (f, s_i)_0 \qquad (1 \leq i \leq n).$$

The FEM yields u_n which depends only on $\mathcal{N}_n f$; we write $u_n = \varphi_n(\mathcal{N}_n f)$. We call $\mathcal{N}_n$ *finite element information* (FEI) of *cardinality* n.

We discuss some properties of the FEM. It is well-known (see, e.g., [3, Chapter 6] for the case $m = 1$, the general case being similar) that if $\{ \mathcal{T}_n \}_{n=1}^{\infty}$ is quasi-uniform [8, p. 272], then B is weakly coercive on $\{ S_n \}_{n=1}^{\infty}$ in the sense of [8, Theorem 8.1], and hence there exists a unique solution $u_n \in S_n$ to (3.1). Moreover, in this quasi-uniform case, we may use (1.2) and [8, Theorems 8.2 and 8.6] to see that there exists a positive constant C (independent of f, u, n, and u_n) such that

$$(3.2) \qquad \|Sf - \varphi_n(\mathcal{N}_n f)\|_m = \|u - u_n\|_m \leq C \inf_{s \in S_n} \|u - s\|_m$$

and

$$(3.3) \qquad \|Sf - \varphi_n(\mathcal{N}_n f)\|_l = \|u - u_n\|_l \leq C n^{-(\mu+m-l)/N} \|f\|_r \qquad \forall f \in H^r(\Omega),$$

where

$$(3.4) \qquad \mu = \min\{k + 1 - m, \, m + r\}.$$

We now relate the results of [15], where we assumed that (i), (ii), and (iii) from the Introduction held. These results, as well as those in the remainder of this paper, will be described using the Ω- and Θ-notations, as well as the more standard O-notation: for functions f and g, we write

$$f = \Omega(g) \quad \text{iff } g = O(f)$$

and

$$f = \Theta(g) \quad \text{iff } f = O(g) \text{ and } g = O(f).$$

Define the *error* of the FEM as

$$(3.5) \qquad e_m(\varphi_n, \mathcal{F}) = \sup_{f \in \mathcal{F}} \|Sf - \varphi_n(\mathcal{N}_n f)\|_m,$$

where $\mathcal{F}$ is the unit ball of $H^r(\Omega)$. Then

$$(3.6) \qquad e_m(\varphi_n, \mathcal{F}) = \Omega(n^{-\mu/N}) \quad \text{as } n \to \infty,$$

where μ is as defined above. Furthermore, in the quasi-uniform case, the bound in (3.6) can be achieved, i.e.,

$$(3.7) \qquad e_m(\varphi_n, \mathcal{F}) = \Theta(n^{-\mu/N}) \quad \text{as } n \to \infty.$$

We next ask whether there is any better way of combining the inner products in $\mathcal{N}_n f$, so as to minimize the error. Let

$$(3.8) \qquad e_m(\mathcal{N}_n, \mathcal{F}) = \inf_{\varphi} \sup_{f \in \mathcal{F}} \| Sf - \varphi(\mathcal{N}_n f) \|_m.$$

Here φ is any mapping (possibly nonlinear) which uses the inner products in $\mathcal{N}_n f$ to approximate Sf. We say that φ is an *algorithm* using $\mathcal{N}_n$. Thus $e_m(\mathcal{N}_n, \mathcal{F})$ measures the minimal error among all algorithms using $\mathcal{N}_n$. We then have

$$(3.9) \qquad e_m(\mathcal{N}_n, \mathcal{F}) = \Omega(n^{-(m+r)/N}) \quad \text{as } n \to \infty.$$

Furthermore, in the quasi-uniform case, the bound in (3.9) can be achieved, i.e.,

$$(3.10) \qquad e_m(\mathcal{N}_n, \mathcal{F}) = \Theta(n^{-(m+r)/N}) \quad \text{as } n \to \infty.$$

From this, it follows that the FEM is quasi-optimal if and only if $k \geq 2m - 1 + r$. By *quasi-optimal* [13], we mean that the FEM has error which (to within a constant factor) is minimal.

We finally ask whether there is any better set of inner products to use, so as to minimize the error. Let

$$(3.11) \qquad e_m(n, \mathcal{F}) = \inf_{\mathcal{N}} e_m(\mathcal{N}, \mathcal{F}),$$

the minimum being over all information $\mathcal{N}$ consisting of n linear functionals. We then have

$$(3.12) \qquad e_m(n, \mathcal{F}) = \Theta(n^{-(m+r)/N}) \quad \text{as } n \to \infty.$$

From this it follows that the finite element information $\mathcal{N}_n$ is quasi-optimal.

4. Analysis of the Normed Case. In the next two sections, we extend the results of [15] to the case where the bilinear form B is weakly coercive and where error is measured in the norm $\| \cdot \|_l$ (where $0 \leq l \leq m$ and (1.2) holds). In this (as in the previous) section, we let $\mathcal{F}$ denote the unit ball in $H^r(\Omega)$. Moreover, we let $e_l(\varphi_n, \mathcal{F})$, $e_l(\mathcal{N}_n, \mathcal{F})$, and $e_l(n, \mathcal{F})$ be as in the previous section, except that now "m" is replaced by "l" in (3.5), (3.8), and (3.11).

We first determine the nth minimal error $e_l(n, \mathcal{F})$.

THEOREM 4.1.

$$e_l(n, \mathcal{F}) = \Theta(n^{-(r+2m-l)/N}) \quad \text{as } n \to \infty.$$

Proof. From [11], we have

$$(4.1) \qquad e_l(n, \mathcal{F}) = d_n(S\mathcal{F}, H_E^l(\Omega)).$$

(Here, and in what follows, we let d_n denote the Kolmogorov n-width. In addition, a subscript E attached to a Sobolev space indicates that those essential boundary conditions which are well-defined for that space are satisfied.) For any $\theta > 0$, let

$$(4.2) \qquad X(\theta) = \{\, u \in H_E^{r+2m}(\Omega) \colon \|u\|_{r+2m} \leq \theta \,\} = \theta B H_E^{r+2m}(\Omega).$$

(Here, and in what follows, BX denotes the unit ball of the space X.) Then the shift theorem (2.5) yields

$$(4.3) \qquad X(\sigma^{-1}) \subseteq S\mathcal{F} \subseteq X(\sigma),$$

and so (4.1) and (4.3) yield

$$(4.4) \qquad d_n\left(X(\sigma^{-1}), H_E^l(\Omega) \right) \leq e_l(n, \mathcal{F}) \leq d_n\left(X(\sigma), H_E^l(\Omega) \right).$$

But (4.2) implies that

$$(4.5) \qquad d_n\left(X(\theta)\right) = \theta d_n\left(BH_E^{r+2m}(\Omega), H_E^l(\Omega)\right) \quad \forall \theta > 0,$$

so that (4.4) and (4.5) yield

$$(4.6) \qquad \sigma^{-1} d_n\left(BH_E^{r+2m}(\Omega), H_E^l(\Omega)\right) \le e_l(n, \mathcal{F}) \le \sigma d_n\left(BH_E^{r+2m}(\Omega), H_E^l(\Omega)\right).$$

From [3, Theorem 2.5.1], we find

$$(4.7) \qquad d_n\left(BH_E^{r+2m}(\Omega), H_E^l(\Omega)\right) = \Theta\left(d_n\left(BH_E^{1/2}(\Omega), L_2(\Omega)\right)^{2(r+2m-l)}\right).$$

But $H_E^{1/2}(\Omega) = H^{1/2}(\Omega) = H_0^{1/2}(\Omega)$, and so

$$(4.8) \qquad \begin{aligned} d_n\left(BH_E^{1/2}(\Omega), L_2(\Omega)\right) &= d_n\left(BH_0^{1/2}(\Omega), L_2(\Omega)\right) \\ &= \Theta\left(d_n\left(BH_0^1(\Omega), L_2(\Omega)\right)^{1/2}\right), \end{aligned}$$

the last by another application of [3, Theorem 2.5.1]. Hence (4.6), (4.7), and (4.8) yield

$$(4.9) \qquad e_l(n, \mathcal{F}) = \Theta\left(d_n\left(BH_0^1(\Omega), L_2(\Omega)\right)^{r+2m-l}\right).$$

But the results of [7] yield that

$$(4.10) \qquad d_n\left(BH_0^1(\Omega), L_2(\Omega)\right) = \Theta(n^{-1/N}).$$

The theorem follows immediately from (4.9) and (4.10). $\square$

We want to determine $e_l(\varphi_n, \mathcal{F})$, the $H^l(\Omega)$-error of the FEM. Recall from (3.4) that $\mu = \min\{k+1-m, m+r\}$.

THEOREM 4.2. (i) $e_l(\varphi_n, \mathcal{F}) = \Omega(n^{-(\mu+m-l)/N})$ as $n \to \infty$.
(ii) If $\{\mathcal{T}_n\}_{n=1}^\infty$ is quasi-uniform, then

$$e_l(\varphi_n, \mathcal{F}) = \Theta(n^{-(\mu+m-l)/N}) \quad as \ n \to \infty.$$

Proof. First note that Theorem 4.1 yields

$$(4.11) \qquad e_l(\varphi_n, \mathcal{F}) \ge e_l(n, \mathcal{F}) = \Theta(n^{-(r+2m-l)/N}).$$

It remains to show that

$$(4.12) \qquad e_l(\varphi_n, \mathcal{F}) = \Omega(n^{-(k+1-l)/N}),$$

since (4.11) and (4.12) yield (i), while (i) and the usual estimate (3.3) yield (ii).

The proof of (4.12) is very similar to that of [15, (4.17)]. Let Ω^0 be the interior of a hypercube whose closure is contained in Ω, and let

$$\mathcal{T}_n^0 = \{\, K \in \mathcal{T}_n \colon K \in \overline{\Omega^0}\,\}.$$

Choose $u \in H_E^{r+2m}(\Omega)$ such that

$$u(x) = \frac{1}{(k+1)!} x_1^{k+1} \quad \forall x \in \overline{\Omega^0}.$$

In what follows, we define (for any region $K \in \mathbf{R}^n$) $|\cdot|_{l,K}$ to be the usual seminorm [5, (3.1.2)] for nonnegative integers l, while for nonintegral values of $l \ge 0$, we define $|\cdot|_{l,K}$ by the Sloboditskiǐ technique [8, p. 96], i.e.,

$$|v|_{l,K}^2 = \sum_{|\alpha|=\lfloor l \rfloor} \int_K \int_K \frac{(D^\alpha v(x) - D^\alpha v(\xi))^2}{|x-\xi|^{N+2(l-\lfloor l \rfloor)}} \, d\xi \, dx.$$

In any case, we have $\| \cdot \|_{l,K} \geq | \cdot |_{l,K}$. We write $P_k(K)$ to denote the space of polynomials of degree up through k over K.

Let $K \in \mathcal{T}_n^0$. We claim that there is a positive constant C_1, independent of K and n, such that

$$(4.13) \qquad \inf_{s \in P_k(K)} |u - s|_{l,K}^2 \geq C_1^2 \operatorname{vol}(K)^{2(k+1-l)/N+1} .$$

To show (4.13) is true, note that K is the affine image of a "reference element" $\hat{K}$ which is independent of K. It is straightforward to check that the functionals

$$\hat{v} \mapsto |\hat{v}|_{k+1,\hat{K}} \quad \text{and} \quad \hat{v} \mapsto \inf_{\hat{s} \in P_k(\hat{K})} |\hat{v} - \hat{s}|_{l,\hat{K}}$$

are seminorms on $P_{k+1}(\hat{K})$. Since $l \leq m \leq k$ (the last by [15, Lemma 4.1]), they have the same kernel $P_k(\hat{K})$. Since $P_{k+1}(\hat{K})$ is finite-dimensional, there is a positive constant $\hat{C}_1 = \hat{C}_1(k, m, \hat{K})$ such that

$$(4.14) \qquad \inf_{\hat{s} \in P_k(\hat{K})} |\hat{v} - \hat{s}|_{l,\hat{K}} \geq \hat{C}_1 |\hat{v}|_{k+1,\hat{K}} \quad \forall \hat{v} \in P_{k+1}(\hat{K}).$$

As in [15], we may then use [5, Theorem 3.1.2] and (4.14) to conclude that (4.13) holds.

Let

$$\Omega_n = \operatorname{int} \bigcup \{ K \colon K \in \mathcal{T}_n^0 \}.$$

Letting $\# \mathcal{T}_n^0$ denote the number of elements in the set $\mathcal{T}_n^0$, we may then use (4.13) to see that

$$(4.15) \qquad \begin{aligned} \inf_{s \in S_n} |u - s|_l^2 &\geq \sum_{K \in \mathcal{T}_n^0} \inf_{s \in P_k(K)} |u - s|_{l,K}^2 \\ &\geq C_1^2 \sum_{K \in \mathcal{T}_n^0} \operatorname{vol}(K)^{2(k+1-l)/N+1} \\ &\geq C_1^2 \frac{\operatorname{vol}(\overline{\Omega_n})^{2(k+1-l)/N+1}}{(\# \mathcal{T}_n^0)^{2(k+1-l)/N}}, \end{aligned}$$

the last because

$$\sum_{K \in \mathcal{T}_n^0} \operatorname{vol}(K) = \operatorname{vol}(\overline{\Omega_n}).$$

Since $\Omega_n \subset \Omega^0$ and $\lim_{n \to \infty} \operatorname{vol}(\overline{\Omega_n}) = \operatorname{vol}(\overline{\Omega^0})$, there is an index n_0 such that

$$(4.16) \qquad \operatorname{vol}(\overline{\Omega_n}) \geq \tfrac{1}{2} \operatorname{vol}(\overline{\Omega^0}) \quad \forall n \geq n_0.$$

Hence (4.15) and (4.16) yield that there is a positive constant C_2, independent of n, such that

$$(4.17) \qquad \inf_{s \in S_n} |u - s|_l \geq C_2 (\# \mathcal{T}_n^0)^{-(k+1-l)/N} \quad \forall n \geq n_0.$$

But [15, (4.14)] shows that $\# \mathcal{T}_n^0 = O(n)$, and so there exists a positive constant C_3, independent of n, such that

$$(4.18) \qquad \inf_{s \in S_n} |u - s|_l \geq C_3 n^{-(k+1-l)/N} \quad \forall n \geq n_0.$$

Now let $f = Lu$. Then f is a nonzero element of $H^r(\Omega)$, since L is injective and u is a nonzero element of $H_E^{r+2m}(\Omega)$. Let

$$f^* = \frac{1}{\|f\|_r} f,$$

so that

$$Sf^* = \frac{1}{\|f\|_r} u.$$

Then $\varphi_n(\mathcal{N}_n f^*) \in S_n$ and $\|\cdot\|_l \geq |\cdot|_l$ yield

$$\|Sf^* - \varphi_n(\mathcal{N}_n f)\|_l \geq |Sf^* - \varphi_n(\mathcal{N}_n f)|_l \geq \inf_{s \in S_n} |Sf^* - s|_l$$

(4.19)
$$\geq \frac{1}{\|f\|_r} \inf_{s \in S_n} |u - s|_l$$

the latter since S_n is a subspace of $H_E^m(\Omega)$. Letting

$$C = \frac{C_3}{\|f\|_r}$$

(which is a positive constant, independent of n), we see that (4.18) and (4.19) imply

(4.20) $$\|Sf^* - \varphi_n(\mathcal{N}_n f)\|_l \geq C n^{-(k+1-l)/N} \quad \forall n \geq n_0.$$

Since $f^* \in \mathcal{F}$, the desired conclusion (4.12) follows immediately from (4.20) and the definition (3.5) of the error of the FEM.   □

We now show that (in the quasi-uniform case) the FEM is a quasi-optimal method using FEI if and only if $k \geq 2m-1+r$; however, FEI is *always* quasi-optimal information.

THEOREM 4.3. (i) $e_l(\mathcal{N}_n, \mathcal{F}) = \Omega(n^{-(r+2m-l)/N})$ as $n \to \infty$.
(ii) *If* $\{\mathcal{T}_n\}_{n=1}^{\infty}$ *is quasi-uniform, then*

$$e_l(\mathcal{N}_n, \mathcal{F}) = \Theta(n^{-(r+2m-l)/N}) \quad as \ n \to \infty.$$

Proof. Using Theorem 4.1, we have

$$e_l(\mathcal{N}_n, \mathcal{F}) \geq e_l(n, \mathcal{F}) = \Theta(n^{-(r+2m-l)/N})$$

establishing (i). To establish (ii), we let $z \in \mathcal{F} \cap \ker \mathcal{N}_n$, so that

(4.21) $$(z, s)_0 = 0 \quad \forall s \in S_n \text{ and } \|z\|_r \leq 1.$$

Let $g \in H^{-l}(\Omega)$. Then $l \leq m$ yields that $g \in H^{-m}(\Omega)$. Symmetry of B and (4.21) yield

$$(Sz, g)_0 = B(Sz, Sg) = (z, Sg)_0 = (z, Sg - s)_0 \quad \forall s \in S_n,$$

and so

(4.22) $$|(Sz, g)_0| \leq \|z\|_r \|Sg - s\|_{-r} \quad \forall s \in S_n.$$

By [3, Theorem 4.1.1], there exists a positive constant C_1, independent of n, as well as an $s \in S_n$, such that

(4.23) $$\|Sg - s\|_{-r} \leq C_1 n^{-\lambda/N} \|Sg\|_{2m-l} \leq C_1 \sigma n^{-\lambda/N} \|g\|_{-l}.$$

Here, λ is given by

(4.24) $$\lambda = \min\{r + 2m - l, k + 1 + r\} = r + 2m - l$$

(the latter due to (1.2)), while σ in the right-hand inequality comes from the shift theorem (2.5). Hence, (4.21)–(4.24) yield

(4.25) $$\frac{|(Sz, g)_0|}{\|g\|_{-l}} \leq C_1 \sigma n^{-(r+2m-l)/N}.$$

Since $g \in H^{-l}(\Omega)$ is arbitrary, (4.25) implies

$$(4.26) \qquad \|Sz\|_l = \sup_{g \in H^{-l}(\Omega)} \frac{|(Sz,g)_0|}{\|g\|_{-l}} \leq C_1 \sigma n^{-(r+2m-l)/N}.$$

Since $z \in \mathcal{F} \cap \ker \mathcal{N}_n$ is arbitrary, (4.26) and [11, Chapters 2 and 3] yield that

$$(4.27) \qquad e_l(\mathcal{N}_n, \mathcal{F}) = \sup_{z \in \mathcal{F} \cap \ker \mathcal{N}_n} \|Sz\|_l \leq C_1 \sigma n^{-(r+2m-l)/N}.$$

The estimate (ii) follows immediately from (i) and (4.27). $\quad\square$

Hence, in the quasi-uniform case, FEI is always quasi-optimal information, while the FEM is a quasi-optimal algorithm using FEI if and only if $k \geq 2m - 1 + r$. It is known that for any information, there always exists a linear combination (of the linear functionals making up that information) whose error is minimal among *all* algorithms using that information; this algorithm is called the *spline algorithm* [11, Chapter 4]. In particular, this implies that the spline algorithm using FEI is always quasi-optimal among *all* algorithms using *any* information, while the FEM is quasi-optimal among all algorithms using any information if and only if $k \geq 2m - 1 + r$.

5. Analysis of the Seminormed Case.
In the previous section, we extended the results of [15] to the case where the bilinear form B is weakly coercive and error is measured in the norm $\|\cdot\|_l$ for $0 \leq l \leq m$. In this section, we weaken one further condition of [15]. Rather than assume that $\mathcal{F} = BH^r(\Omega)$, the unit ball of $H^r(\Omega)$, we assume in this section that $\mathcal{F} = \mathcal{B}H^r(\Omega)$, the set of all functions f in $H^r(\Omega)$ for which $|f|_r \leq 1$.

In this section, $e_l(\varphi_n, \mathcal{F})$, $e_l(\mathcal{N}_n, \mathcal{F})$, and $e_l(n, \mathcal{F})$ will be defined as in the previous section, except that now $\mathcal{F} = \mathcal{B}H^r(\Omega)$. We assume that r is a nonnegative integer. Let $P_{r-1}(\Omega)$ denote the space of all polynomials in N variables whose degree does not exceed $r - 1$, treated as a space of functions defined on Ω. We also let

$$n^* = \dim P_{r-1}(\Omega) = \binom{N+r-1}{r-1}$$

denote the dimension of this space of polynomials.

We first obtain an estimate of the nth minimal error.

THEOREM 5.1. (i) *If* $n < n^*$, *then* $e_l(n, \mathcal{F}) = \infty$.
(ii) $e_l(n, \mathcal{F}) = \Theta(n^{-(r+2m-l)/N})$ *as* $n \to \infty$.

Proof. (i) follows immediately from [11, Theorem 2.3.2]. To establish (ii), let

$$\hat{H}^r(\Omega) = H^r(\Omega)/P_{r-1}(\Omega)$$

under the quotient norm. Then [14, Lemma 5.3] yields

$$(5.1) \qquad e_l(n, \mathcal{F}) = O\left(e_l\left(n - n^*, B\hat{H}^r(\Omega)\right)\right) = O\left(e_l\left(n - n^*, BH^r(\Omega)\right)\right),$$

where (as in the previous section) "B" denotes the unit ball and the second step is because $B\hat{H}^r(\Omega) \subseteq BH^r(\Omega)$. But Theorem 4.1 yields

$$(5.2) \qquad e_l\left(n - n^*, BH^r(\Omega)\right) = \Theta\left((n - n^*)^{-(r+2m-l)/N}\right) = \Theta(n^{-(r+2m-l)/N}).$$

Hence, (5.1) and (5.2) yield

$$(5.3) \qquad e_l(n, \mathcal{F}) = O(n^{-(r+2m-l)/N}).$$

On the other hand, $BH^r(\Omega) \subseteq \mathcal{B}H^r(\Omega) = \mathcal{F}$ yields

$$(5.4) \qquad e_l(n, \mathcal{F}) \geq e_l(n, BH^r(\Omega)) = \Theta(n^{-(r+2m-l)/N}).$$

The estimate (ii) then follows from (5.3) and (5.4). $\quad\square$

We next investigate the error of the FEM. We show that either the FEM has infinite error, or the estimates of Theorem 4.2 hold. Once again, recall from (3.4) that $\mu = \min\{ k+1-m, m+r \}$.

THEOREM 5.2. (i) If $SP_{r-1}(\Omega) \not\subseteq S_n$, then $e_l(\varphi_n, \mathcal{F}) = \infty$.
(ii) If $SP_{r-1}(\Omega) \subseteq S_n$ for all sufficiently large n, then

$$e_l(\varphi_n, \mathcal{F}) = \Omega(n^{-(\mu+m-l)/N}) \quad as \ n \to \infty,$$

where the "Ω" may be changed to "Θ" when $\{ T_n \}_{n=1}^{\infty}$ is quasi-uniform.

Proof. (i) Let $SP_{r-1}(\Omega) \not\subseteq S_n$. Since the range of φ_n is S_n, there exists $f \in P_{r-1}(\Omega)$ such that $\varphi_n(\mathcal{N}_n f) \neq Sf$. Since φ_n is linear and $P_{r-1}(\Omega) = \ker |\cdot|_r$, we may use [11, Lemma 3.2.2] to see that $e_l(\varphi_n, \mathcal{F}) = \infty$.

(ii) Let $SP_{r-1}(\Omega) \subseteq S_n$. Since $BH^r(\Omega) \subseteq \mathcal{F}$, Theorem 4.2(i) yields

$$e_l(\varphi_n, \mathcal{F}) \geq e_l(\varphi_n, BH^r(\Omega)) = \Theta(n^{-(\mu+m-l)/N}).$$

Suppose now, in addition, that $\{ T_n \}_{n=1}^{\infty}$ is quasi-uniform. We need only show that

$$(5.5) \qquad e_l(\varphi_n, \mathcal{F}) = O(n^{-(\mu+m-l)/N}) \quad as \ n \to \infty.$$

Let $f \in \mathcal{F}$. Then there exists a unique choice of

$$f_1 \in P_{r-1}(\Omega) \quad and \quad f_2 \in \hat{H}^r(\Omega) = H^r(\Omega)/P_{r-1}(\Omega)$$

such that

$$(5.6) \qquad\qquad f = f_1 + f_2.$$

Recall [5, Theorem 3.1.1] that there is a positive constant C_1 such that

$$(5.7) \qquad\qquad \|\cdot\|_r \leq C_1 |\cdot|_r \quad on \ \hat{H}^r(\Omega).$$

Since $f_1 \in P_{r-1}(\Omega)$, and $f \in \mathcal{F}$, we have

$$\|f_2\|_r \leq C_1 |f_2|_r = C_1 |f|_r \leq C_1,$$

so that (3.3) yields that there is a positive constant C_2 (independent of n and f) such that

$$(5.8) \qquad \|Sf_2 - \varphi_n(\mathcal{N}_n f_2)\|_l \leq Cn^{-(\mu+m-l)/N}\|f_2\|_r \leq C_2 n^{-(\mu+m-l)/N}.$$

Now $f_1 \in P_{r-1}(\Omega)$ and $SP_{r-1}(\Omega) \subseteq S_n$, so that $Sf_1 \in S_n$. Since $l \leq m$, we may use (3.2) to see that

$$(5.9) \qquad \|Sf_1 - \varphi_n(\mathcal{N}_n f_1)\|_l \leq \|Sf_1 - \varphi_n(\mathcal{N}_n f_1)\|_m \leq C \inf_{s \in S_n} \|Sf_1 - s\|_m = 0.$$

Since S, φ_n, and $\mathcal{N}_n$ are linear, the estimates (5.6), (5.8), and (5.9) yield that

$$\|Sf - \varphi_n(\mathcal{N}_n f)\|_l \leq C_2 n^{-(\mu+m-l)/N}.$$

Since $f \in \mathcal{F}$ is arbitrary, this yields (5.5). $\quad\square$

Remark 5.1. We illustrate the different possibilities in Theorem 5.2 by considering the model problems (1.3) and (1.4), where we have $r = 1$ and $m = 1$. Hence we define two operators $S_1 \colon H^1(0,1) \to H_0^1(0,1)$ and $S_2 \colon H^1(0,1) \to H^1(0,1)$ by

$$u = S_1 f \quad \text{satisfies} \quad \begin{array}{l} -u'' + u = f \quad \text{in } (0,1) \\ u(0) = u(1) = 0 \end{array} \quad \forall f \in H^1(0,1),$$

and

$$u = S_2 f \quad \text{satisfies} \quad \begin{array}{l} -u'' + u = f \quad \text{in } (0,1) \\ u'(0) = u'(1) = 0 \end{array} \quad \forall f \in H^1(0,1).$$

(Note that S_1 and S_2 differ only in their boundary conditions.) We claim that the FEM has infinite error for S_1, but has finite error for S_2. Again, keep in mind that these problems are being solved for all $f \in H^1(0,1)$ such that $|f|_1 \leq 1$.

To see that the FEM has infinite error for S_1, note that $S_1(P_0(0,1))$ is spanned by the solution of the problem

$$-z'' + z = 1 \quad \text{in } (0,1), \qquad z(0) = z(1) = 0,$$

the solution of which is

$$z(x) = 1 - \left(\frac{e-1}{e^2 - 1} \right) e^x - \left(\frac{e^2 - e}{e^2 - 1} \right) e^{-x}.$$

Since z is not a piecewise polynomial, we find that $S_1(P_1(0,1)) \not\subseteq S_n$, no matter how big the degree k of the subspace S_n or the dimension n of S_n are. Hence, the FEM has infinite error for the problem S_1.

We now consider the problem S_2. We find that $S_2(P_0(0,1)) = P_0(0,1)$, since the only solution to

$$-z'' + z = 1 \quad \text{in } (0,1), \qquad z'(0) = z'(1) = 0$$

is

$$z(x) \equiv 1.$$

Since [15, Lemma 4.1] yields that $k \geq 1$, and there are no essential boundary conditions for this problem, we have $S_2(P_0(0,1)) \subseteq S_n$ for all $n \geq 1$ and any choice of k. This shows that the FEM has finite error for the problem S_2. In fact, the error of the FEM for this problem using S_n (in the quasi-uniform case) is $\Theta(n^{-(\chi - l)})$, where $\chi = \min\{ k + 1, 3 \}$. Comparing Theorems 5.1 and 5.2, we see that the FEM is quasi-optimal for this problem if and only if $k \geq 2$. $\square$

Remark 5.2. Since $S_n \subset H_E^m(\Omega)$, the condition $SP_{r-1}(\Omega) \subseteq S_n$ is equivalent to the condition $P_{r-1}(\Omega) \subseteq LS_n$. For situations in which the explicit form of the solution is unknown (i.e., most cases arising in practice), it will generally be easier to verify whether $P_{r-1}(\Omega) \subseteq LS_n$ than whether $SP_{r-1}(\Omega) \subseteq S_n$. $\square$

Remark 5.3. The condition $SP_{r-1}(\Omega) \subseteq S_n$ is very restrictive, since it is not generally the case that the solution u of the problem $Lu = f$ (with f polynomial) is a piecewise polynomial satisfying the boundary conditions. (For example, we saw in Remark 5.1 that the solution $u = S_1 f$ of (1.3) with $f \equiv 1$ involves exponential functions.) It would be extremely unlikely to have $SP_{r-1}(\Omega) \subseteq S_n$ in most situations, especially when the boundary of Ω is complicated or the coefficients $a_{\alpha\beta}$ of the L are nonpolynomial. Hence, we see that the FEM has finite error for the seminormed case only under exceptional circumstances. $\square$

Hence, there are situations in which the FEM has infinite error, no matter how big k and n are. Is this a feature of the FEM itself, or is it a feature of the

information which the FEM uses? In the remainder of this section, we show that the fault lies with the FEM rather than with the FEI. In fact, we will show that FEI is quasi-optimal information.

In order to do this, we first establish

LEMMA 5.1. *There exists a positive integer* n_0 *and a positive constant* C *such that for any* $n \geq n_0$,
$$\|z\|_r \leq C|z|_r \quad \forall z \in \ker \mathcal{N}_n .$$

Proof. If $r = 0$, this is immediate.

Suppose now that $r \geq 1$. If the conclusion is false, then there is a subsequence $\{ z_{n_i} \in \ker \mathcal{N}_{n_i} \}_{i=1}^{\infty}$ such that
$$(5.10) \qquad \|z_{n_i}\|_r = 1 \quad \text{and} \quad \lim_{i \to \infty} |z_{n_i}|_r = 0 .$$

Following the proof of [8, Theorem 3.1.1], the Rellich-Kondrašov compactness theorem yields that there exists $z \in P_{r-1}(\Omega)$ and a subsequence, which we again denote $\{ z_{n_i} \in \ker \mathcal{N}_{n_i} \}_{i=1}^{\infty}$, such that
$$(5.11) \qquad \lim_{i \to \infty} z_{n_i} = z \quad \text{in } H^r(\Omega) \text{ and thus in } L_2(\Omega).$$

Hence, we see that
$$(5.12) \qquad \|z\|_r = 1.$$

We claim that $z = 0$, contradicting (5.12). Indeed, let $\varepsilon > 0$. Using denseness of $C_0^{\infty}(\Omega)$ in $L_2(\Omega)$, there exists $w \in C_0^{\infty}(\Omega)$ such that
$$(5.13) \qquad \|z - w\|_0 < \tfrac{1}{3}\varepsilon.$$

Since $C_0^{\infty}(\Omega) \subseteq H_E^m(\Omega) \cap H^1(\Omega)$, the standard results (found in, e.g., [8, Chapter 6]) yield that there is a positive constant C_1 (independent of z and w) such that for any $j > 0$, there exists $w_j \in S_j$ for which
$$\|w - w_j\|_0 \leq C_1 j^{-1/N} |w|_1 .$$

Hence, there is an index $i_0(\varepsilon)$ such that for any $i \geq i_0(\varepsilon)$, there exists $w_{n_i} \in S_{n_i}$ satisfying
$$(5.14) \qquad \|w - w_{n_i}\|_0 \leq \tfrac{1}{3}\varepsilon.$$

From (5.11) and $\| \cdot \|_0 \leq \| \cdot \|_r$, there is an index $i_1(\varepsilon)$ such that for any $i \geq i_1(\varepsilon)$, there exists $z_{n_i} \in \ker \mathcal{N}_{n_i}$ for which
$$(5.15) \qquad \|z - z_{n_i}\|_0 \leq \tfrac{1}{3}\varepsilon.$$

Let $i_2(\varepsilon) = \max\{ i_0(\varepsilon), i_1(\varepsilon) \}$. Then (5.13)–(5.15) and the triangle inequality yield
$$(5.16) \qquad \|z_{n_i} - w_{n_i}\|_0 \leq \varepsilon \quad \forall i \geq i_2(\varepsilon).$$

But $z_{n_i} \in \ker \mathcal{N}_{n_i} = H^r(\Omega) \cap (L_2(\Omega)/S_{n_i})$ and $w_{n_i} \in S_{n_i}$. Hence $(z_{n_i}, w_{n_i})_0 = 0$, which, when combined with (5.16), yields
$$\|z_{n_i}\|_0^2 \leq \|z_{n_i}\|_0^2 - 2(z_{n_i}, w_{n_i})_0 + \|w_{n_i}\|_0^2 = \|z_{n_i} - w_{n_i}\|_0^2 \leq \varepsilon^2.$$

Thus for any $\varepsilon > 0$, there is an index $i_2(\varepsilon)$ for which
$$\|z_{n_i}\|_0 \leq \varepsilon \quad \forall i \geq i_2(\varepsilon).$$

Hence
$$(5.17) \qquad \lim_{i \to \infty} z_{n_i} = 0 \quad \text{in } L_2(\Omega).$$

From (5.11) and (5.17), we have $z = 0$, the desired contradiction. $\quad\square$

We are now ready to show that the FEI is always quasi-optimal information for the seminormed case.

THEOREM 5.3. (i) $e_l(\mathcal{N}_n, \mathcal{F}) = \Omega(n^{-(r+2m-l)/N})$ as $n \to \infty$.
(ii) If $\{ \mathcal{T}_n \}_{n=1}^{\infty}$ is quasi-uniform, then

$$e_l(\mathcal{N}_n, \mathcal{F}) = \Theta(n^{-(r+2m-l)/N}) \quad as\ n \to \infty.$$

Proof. Using Theorem 5.1, we have

$$e_l(\mathcal{N}_n, \mathcal{F}) \geq e_l(n, \mathcal{F}) = \Theta(n^{-(r+2m-l)/N}),$$

establishing (i). To prove (ii), let $z \in \mathcal{F} \cap \ker \mathcal{N}_n$. For any $g \in H^{-l}(\Omega)$, we use (4.22) to see that

$$(5.18) \qquad |(Sz, g)_0| \leq \|z\|_r \inf_{s \in S_n} \|Sg - s\|_{-r}.$$

By (4.23) and (4.24), there is a positive constant C_1 (independent of n, g, and z) such that

$$(5.19) \qquad \inf_{s \in S_n} \|Sg - s\|_{-r} \leq C_1 n^{-(r+2m-l)/N} \|g\|_{-l}.$$

Since $z \in \mathcal{F}$, Lemma 5.1 implies that there is a positive constant C_2 (independent of n and z) such that

$$(5.20) \qquad \|z\|_r \leq C_2 |z|_r \leq C_2.$$

Hence, (5.18)–(5.20) yield the existence of a positive constant C (independent of n, z, and g) such that

$$(5.21) \qquad \frac{|(Sz, g)_0|}{\|g\|_{-l}} \leq C n^{-(r+2m-l)/N}.$$

Taking the supremum in (5.21) over all nonzero $g \in H^{-l}(\Omega)$, we find

$$(5.22) \qquad \|Sz\|_l \leq C n^{-(r+2m-l)/N}.$$

Since $z \in \mathcal{F} \cap \ker \mathcal{N}_n$ is arbitrary, (5.22) and [11, Chapters 2 and 3] yield

$$e_l(\mathcal{N}_n, \mathcal{F}) \leq C n^{-(r+2m-l)/N},$$

which, along with (i), establishes (ii). □

Hence when the grid sequence is quasi-uniform, FEI is always quasi-optimal information for the seminormed case. This means that the spline algorithm [11, Chapter 4] using FEI is always quasi-optimal among *all* algorithms using *any* information. This is to be contrasted with the fact that in the seminormed case, the error of the FEM need not even be finite; however, when the FEM has finite error, it is quasi-optimal among all algorithms using any information if and only if $k \geq 2m - 1 + r$.

6. Complexity Analysis. In this section, we discuss the complexity (minimal cost) of finding ε-approximations to the solution of the variational boundary-value problem, as well as the penalty for using the FEM when $k < 2m - 1 + r$.

Let $\varepsilon > 0$. An algorithm φ produces an ε-*approximation* to the problem $(S, \mathcal{F})$ in the $H^l(\Omega)$-norm if $e_l(\varphi, \mathcal{F}) \leq \varepsilon$. The *complexity* comp$(\varphi)$ of an *algorithm* φ is defined via the model of computation discussed in [11, Chapter 5]. Informally, we assume that any linear functional required by φ can be evaluated with finite cost c_1, and that the cost of an arithmetic operation is unity.

Recall that φ_n denotes the FEM of degree k using the FEI $\mathcal{N}_n$ based on the finite element subspace S_n. We assume that $\{ \mathcal{T}_n \}_{n=1}^{\infty}$ is quasi-uniform, where $\mathcal{T}_n$

is the triangulation of Ω upon which S_n is based. Since the FEI $\mathcal{N}_n$ contains n inner products, we find that

$$\operatorname{comp}(\varphi_n) \geq nc_1.$$

Throughout the remainder of this paper, we assume that

(6.1) $$\operatorname{comp}(\varphi_n) = \Theta(n) \quad \text{as } n \to \infty.$$

Remark 6.1. The assumption (6.1) is reasonable in either of two cases. In the first case, we actually assume the existence of an algorithm which can solve the linear system generated by the FEM, whose number of operations is linear in n, the size of the linear system. This condition holds in a number of special cases— finding such linear-time algorithms is still an open problem for the general case.

Alternatively, one may wish to make an assumption of *preconditioning*. That is, we assume that any computation which is independent of the right-hand side f is done in advance, and not counted when determining the number of operations required when approximating the solution u to the problem $Lu = f$. To make this more precise, recall that the FEM is a linear algorithm, i.e., it produces a linear approximation to the exact solution $u(x)$ having the form $\sum_{i=1}^{n} (f, s_i)_0 g_i(x)$, where $g_1, \ldots, g_n \in S_n$. Since $g_1, \ldots g_n$ are independent of f, they may be determined in advance. (This precomputation may be especially efficient in the case where the problem $Lu = f$ is to be solved for many different right-hand sides f.) Hence, computing the value of the FEM at any point in Ω requires at most n multiplications and $n - 1$ additions, once the n inner products $(f, s_1), \ldots, (f, s_n)$ have been evaluated. Thus (6.1) holds. $\square$

Let

$$\mathrm{FEM}_l(\varepsilon, \mathcal{F}) = \inf\{ \operatorname{comp}(\varphi_n) \colon n \text{ is an index such that } e_l(\varphi_n, \mathcal{F}) \leq \varepsilon \}$$

denote the cost of solving the problem in the $H^l(\Omega)$-norm using the FEM. From the results of Sections 4 and 5, along with (6.1), we find

THEOREM 6.1. (i) $\mathrm{FEM}_l\left(\varepsilon, \mathcal{B}H^r(\Omega)\right) = \Theta(\varepsilon^{-N/(\mu+m+l)})$ *as* $\varepsilon \to 0$, *where* $\mu = \min\{k + 1 - m, m + r\}$.

(ii) *If there exists no integer* $n \geq 0$ *for which* $SP_{r-1}(\Omega) \subseteq S_n$, *then*

$$\mathrm{FEM}_l\left(\varepsilon, \mathcal{B}H^r(\Omega)\right) = \infty \quad \forall \varepsilon \geq 0.$$

(iii) *If there exists an integer* $n_0 \geq 0$ *such that* $SP_{r-1}(\Omega) \subseteq S_n$ *for* $n \geq n_0$, *then*

$$\mathrm{FEM}_l\left(\varepsilon, \mathcal{B}H^r(\Omega)\right) = \Theta(\varepsilon^{-N/(\mu+m+l)}) \quad \text{as } \varepsilon \to 0. \quad \square$$

As mentioned previously in Sections 4 and 5, for any information, there exists a linear algorithm (the *spline algorithm*) having minimal error. Let φ_n^s denote the spline algorithm using the FEI $\mathcal{N}_n$. If we agree once more to accept the idea of preconditioning as discussed in Remark 6.1, we find

(6.2) $$\operatorname{comp}(\varphi_n^s) = \Theta(n) \quad \text{as } n \to \infty.$$

We now let

$$\mathrm{SPLINE}_l(\varepsilon, \mathcal{F}) = \inf\{ \operatorname{comp}(\varphi_n^s) \colon n \text{ is an index such that } e_l(\varphi_n^s) \leq \varepsilon \}$$

denote the cost of solving the problem using the spline algorithm. Using (6.2), the minimal error properties of the spline algorithm, and the results of Sections 4 and 5,

we find

THEOREM 6.2. *In both the normed case $\mathcal{F} = BH^r(\Omega)$ and the seminormed case $\mathcal{F} = \mathcal{B}H^r(\Omega)$, we have*

$$\text{SPLINE}_l(\varepsilon, \mathcal{F}) = \Theta(\varepsilon^{-N/(r+2m-l)}) \quad as \ \varepsilon \to 0. \quad \square$$

We now wish to determine the minimal cost of solving the problem. Let

$$\text{COMP}_l(\varepsilon, \mathcal{F}) = \inf\{\,\text{comp}(\varphi)\colon \varphi \text{ is an algorithm for which } e_l(\varphi, \mathcal{F}) \le \varepsilon\,\}$$

denote the *problem complexity*, i.e., the inherent cost of solving the problem with error not exceeding ε. Using the results of the two previous sections, we then find

THEOREM 6.3. *In both the normed case $\mathcal{F} = BH^r(\Omega)$ and the seminormed case $\mathcal{F} = \mathcal{B}H^r(\Omega)$, we have*

$$\text{COMP}_l(\varepsilon, \mathcal{F}) = \Theta(\varepsilon^{-N/(r+2m-l)}) \quad as \ \varepsilon \to 0. \quad \square$$

Hence, we may draw the following conclusions:

COROLLARY 6.1. (i) *The spline algorithm using the FEI is always quasi-optimal, for both the normed and seminormed cases.*

(ii) *The FEM is quasi-optimal for the normed case if and only if $k \ge 2m - 1 + r$. If $k < 2m - 1 + r$, then*

$$\frac{\text{FEM}_l\left(\varepsilon, BH^r(\Omega)\right)}{\text{COMP}_l\left(\varepsilon, BH^r(\Omega)\right)} = \Theta\left(\left(\frac{1}{\varepsilon}\right)^{\lambda N}\right) \quad as \ \varepsilon \to 0,$$

where

(6.3) $$\lambda = \frac{1}{k+1-l} - \frac{1}{r+2m-l} > 0,$$

so that

$$\lim_{\varepsilon \to 0} \frac{\text{FEM}_l\left(\varepsilon, BH^r(\Omega)\right)}{\text{COMP}_l\left(\varepsilon, BH^r(\Omega)\right)} = \infty.$$

(iii) *The FEM is quasi-optimal for the seminormed case if and only if $k \ge 2m - 1 + r$ and $SP_{r-1}(\Omega) \subseteq S_n$ for all sufficiently large n. If $k < 2m - 1 + r$ and $SP_{r-1}(\Omega) \subseteq S_n$ for all sufficiently large n, then*

$$\frac{\text{FEM}_l\left(\varepsilon, \mathcal{B}H^r(\Omega)\right)}{\text{COMP}_l\left(\varepsilon, \mathcal{B}H^r(\Omega)\right)} = \Theta\left(\left(\frac{1}{\varepsilon}\right)^{\lambda N}\right) \quad as \ \varepsilon \to 0,$$

where λ is given by (6.3). In the case that $SP_{r-1}(\Omega) \not\subseteq S_n$ for all $n \ge 0$, we have

$$\text{FEM}_l\left(\varepsilon, \mathcal{B}H^r(\Omega)\right) = \infty \quad \forall \varepsilon \ge 0. \quad \square$$

Hence when k is too small for a given value of r, there is an infinite asymptotic penalty for using the FEM instead of the spline algorithm. Corollary 6.1 implies that there is an $\varepsilon_0 > 0$ such that

(6.4) $$\text{SPLINE}_l(\varepsilon, \mathcal{F}) < \text{FEM}_l(\varepsilon, \mathcal{F}) \quad \text{for } 0 < \varepsilon < \varepsilon_0.$$

What is the value of ε_0? If ε_0 is unreasonably small, it may turn out that it is more reasonable to use the FEM for "practical" values of ε. We determine the value of ε_0 for a model problem in

Example 6.1. Let $N = 1$, $\Omega = (0, \pi)$, $m = 1$, $r = 1$, $H^1_E(\Omega) = H^1_0(0, \pi)$, and consider the bilinear form $B\colon H^1_0(0, \pi) \times H^1_0(0, \pi) \to \mathbf{R}$ defined by

$$B(v, w) = \int_0^\pi v'w' \quad \forall v, w \in H^1_0(0, \pi).$$

Hence for $f \in H^1(0, \pi)$, $u = Sf$ is the variational solution to the problem

$$-u'' = f \quad \text{in } (0, \pi), \qquad u(0) = u(\pi) = 0.$$

We choose S_n to be the n-dimensional subspace of $H_0^1(0, \pi)$ consisting of piecewise linear polynomials with nodes at $x_j = j\pi/(n+1)$ for $0 \le j \le n+1$, so that $k = 1$.

We wish to determine ε_0 such that (6.4) holds with $l = 1$ and $\mathcal{F} = BH^1(0, \pi)$. This is similar to [15, Example 6.1]. The only difference is that in [15], we measured error by the energy norm (which is the H^1-seminorm $|\cdot|_1$), while here we use the H^1-norm $\|\cdot\|_1$.

Using the Poincaré inequality

$$\|\cdot\|_0 \le |\cdot|_1 \quad \text{on } H_0^1(0, \pi)$$

(see [9]), we have

$$|\cdot|_1 \le \|\cdot\|_1 \le \sqrt{2}\,|\cdot|_1 \quad \text{on } H_0^1(0, \pi).$$

Hence [15, (6.25) and (6.32)] and this inequality imply

$$(6.5) \qquad e_1(\varphi_n, \mathcal{F}) \ge \frac{\pi}{\sqrt{12}\,(n+1)} \doteq \frac{0.90689968}{n+1}$$

and

$$(6.6) \qquad e_1(\varphi_n^s, \mathcal{F}) \le \frac{\sqrt{2}}{(n+1)^2} \doteq \frac{1.4142136}{(n+1)^2},$$

where (as before) φ_n denotes the FEM and φ_n^s denotes the spline algorithm using FEI.

To proceed further, we require sharper versions of (6.1) and (6.2). First, note that since any algorithm which actually makes use of all of the inner products of $\mathcal{N}_n$ must perform at least $n - 1$ arithmetic operations, we find (in particular) that

$$(6.7) \qquad \operatorname{comp}(\varphi_n) \ge nc_1 + n - 1.$$

Next, we accept the idea of preconditioning. In this case, we find that linearity of the spline algorithm implies that it may be evaluated at any point in Ω using at most n multiplications and $n - 1$ additions, once the inner products of $\mathcal{N}_n$ have been evaluated. Hence

$$(6.8) \qquad \operatorname{comp}(\varphi_n^s) \le nc_1 + 2n - 1.$$

Using (6.5) and (6.7), we find that

$$\operatorname{FEM}_1(\varepsilon, \mathcal{F}) \ge (c_1 + 1)\left(\frac{\pi}{\sqrt{12}}\varepsilon^{-1} - 1\right) - 1,$$

while (6.6) and (6.8) yield

$$\operatorname{SPLINE}_1(\varepsilon, \mathcal{F}) \le (c_1 + 2)(\sqrt[4]{2}\varepsilon^{-1/2} - 1) - 1.$$

From these two inequalities, it follows that

$$(6.9) \qquad \operatorname{SPLINE}_1(\varepsilon, \mathcal{F}) < \operatorname{FEM}_1(\varepsilon, \mathcal{F}) \quad \text{if } 0 < \varepsilon < \varepsilon_0,$$

where ε_0 is the smallest positive solution ε of

$$(c_1 + 1)\left(\frac{\pi}{\sqrt{12}}\varepsilon^{-1} - 1\right) = (c_1 + 2)(\sqrt[4]{2}\varepsilon^{-1/2} - 1).$$

Some algebra yields

$$\varepsilon_0 = \varepsilon_0(c_1) = \left(\sqrt[4]{2}\left(\frac{1}{2}c_1 + 1\right) - \sqrt{\sqrt{2}\left(\frac{1}{2}c_1 + 1\right)^2 - \frac{\pi}{\sqrt{12}}(c_1 + 1)} \right)^2.$$

We now examine the value of $\varepsilon_0(c_1)$ under various assumptions on the cost c_1 of evaluating a linear functional, noting that $\varepsilon_0(c_1)$ increases with (nonnegative) c_1. Clearly, $c_1 \geq 0$, so that

$$\varepsilon_0(c_0) \geq \varepsilon_0(0) = \left(\sqrt[4]{2} - \left(\sqrt{2} - \frac{\pi}{\sqrt{12}}\right)^{1/2} \right)^2 \doteq 0.22747884.$$

This tells us that (6.9) holds for all ε less than (roughly) 0.227. Next, we assume that $c_1 \geq 1$, i.e., that evaluating a linear functional is at least as hard as an arithmetic operation. (It would be hard to imagine otherwise.) In this case,

$$\varepsilon_0(c_0) \geq \varepsilon_0(1) = \left(\frac{3}{2}\sqrt[4]{2} - \left(\frac{9}{4}\sqrt{2} - \frac{\pi}{\sqrt{3}}\right)^{1/2} \right)^2 \doteq 0.37714081.$$

Finally, it is reasonable to assume that c_1 is very large, i.e., that evaluating a linear functional is *much* harder than an arithmetic operation [11, p. 85]. One may check that

$$\lim_{c_1 \to \infty} \varepsilon_0(c_1) = \frac{1}{24}\pi^2\sqrt{2} \doteq 0.58157202,$$

giving an estimate of $\varepsilon_0(c_1)$ for large values of c_1.   □

Based on this example, it seems reasonable to conjecture that (6.4) generally holds for "reasonable" values of ε_0. However, see the discussion at the end of [15, Section 6] for some comments regarding this conjecture.

7. Summary, Extensions, and Open Questions. We have shown that FEI of degree $k \geq m$ is always quasi-optimal information for indefinite linear elliptic problems $Lu = f$ under the following conditions:
   (i) Error is measured in the Sobolev l-norm, where $0 \leq l \leq m$.
   (ii) Either $\|f\|_r \leq 1$ (where $r \geq -m$) or $|f|_r \leq 1$ (where r is a nonnegative integer).
   (iii) $k \geq 2m - 1 - l$.
However, the FEM is not always quasi-optimal among all algorithms using FEI. In the normed case $\|f\|_r \leq 1$, the FEM is quasi-optimal if and only if $k \geq 2m - 1 + r$. In the seminormed case $|f|_r \leq 1$, the FEM has finite error if and only if the finite element subspace contains $SP_{r-1}(\Omega)$; if this holds, then the FEM is quasi-optimal if and only if $k \geq 2m - 1 + r$. In the case where $k < 2m - 1 + r$, the asymptotic penalty for using the FEM is infinite.

What happens when we try to weaken the assumptions above?

One natural weakening of (i) is to allow l to satisfy the inequality $l \leq m$. The proofs of Theorems 5.2(i) and 6.2(ii) (the lower bound for the FEM) do not hold, since there seems to be no natural definition of the Sobolev l-seminorm for negative l. In the case where $\{T_n\}_{n=1}^{\infty}$ is uniform and $H_E^m(\Omega) = H^m(\Omega)$ (i.e., a Neumann problem), the results in [10] show that these results do hold for negative l in this special situation. We conjecture that this is true in general, i.e., the lower bounds for the FEM given in these theorems hold for any $l \leq m$. However, the other results in this paper do hold for any l such that $l \leq m$, provided (iii) still holds.

A further natural weakening of (i) is to allow the error to be measured in non-Hilbert norms. For example, one may wish to consider sup-norm estimates, since it is often more important to be able to get pointwise estimates on the error than mean-square error estimates.

Condition (ii) may be weakened in a number of ways. Rather than use the norm (or seminorm) over $H^r(\Omega)$, we may use $\|\cdot\|_{r,p}$ or $|\cdot|_{r,p}$, the norm or seminorm in $W^{r,p}(\Omega)$. Alternatively, we may decide to use the norm in more complicated spaces, such as the Besov space $B_p^{r,s}(\Omega)$ (see [4]). If we let $\mathcal{F}$ be the set of elements in one of these spaces whose norm (or seminorm) is bounded by unity, is FEI still quasi-optimal information? When is the FEM quasi-optimal?

What happens when (iii) no longer holds? In this case, the bounds that may be established using the techniques of this paper are no longer tight. Although (iii) holds for $l \geq 0$ in most cases of practical interest, it is important to find out what happens when (iii) no longer holds, as is the case when rectangular elements are used or when error is measured in negative norms.

We finally mention that all of the results of this paper hold for a *worst-case setting*, in which the error is determined by the maximum value (over all elements of $\mathcal{F}$) of the difference between the exact solution and the approximate solution. This is by no means the only possible model. As an example, Trojan [12] considers an *asymptotic setting*, in which a fixed (but arbitrary) element $f \in \mathcal{F}$ is chosen, a sequence of functionals of f is evaluated, and these values are used to produce an approximation of the solution u of the problem $Lu = f$. Other possible settings include an average setting and an asymptotic-average setting. How close is the FEM to being optimal in such settings?

Division of Science and Mathematics
Fordham University/College at Lincoln Center
113 West 60th Street
New York, New York 10023

Department of Computer Science
Columbia University
New York, New York 10027

1. S. AGMON, *Lectures on Elliptic Boundary Value Problems*, Van Nostrand, Princeton, NJ, 1965.

2. P. M. ANSELONE & P. J. LAURENT, "A general method for the construction of interpolating or smoothing spline functions," *Numer. Math.*, v. 12, 1968, pp. 66–82.

3. I. BABUŠKA & A. K. AZIZ, "Survey lectures on the mathematical foundations of the finite element method," *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations* (A. K. Aziz, ed.), Academic Press, New York, 1972, pp. 3–359.

4. P. L. BUTZER & H. BERENS, *Semi-groups of Operators and Approximation*, Springer-Verlag, Berlin, 1967.

5. P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

6. P. G. CIARLET & P. A. RAVIART, "Interpolation theory over curved elements," *Comput. Methods Appl. Mech. Engrg.*, v. 1, 1972, pp. 217–249.

7. J. W. JEROME, "Asymptotic estimates of the L_2 n-width," *J. Math. Anal. Appl.*, v. 22, 1968, pp. 449–464.

8. J. T. ODEN & J. N. REDDY, *An Introduction to the Mathematical Theory of Finite Elements*, Wiley-Interscience, New York, 1976.

9. M. SCHULTZ, *Spline Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1973.

10. G. STRANG & G. FIX, "A Fourier analysis of the finite element variational method," *Constructive Aspects of Functional Analysis, Part II*, C.I.M.E., Rome, 1973.

11. J. F. TRAUB & H. WOŹNIAKOWSKI, *A General Theory of Optimal Algorithms*, Academic Press, New York, 1980.

12. J. M. TROJAN, "Asymptotic model for linear problems," (In preparation).

13. L. B. WAHLBIN, "Quasi-optimality of the H_0^1 projection into finite element spaces," *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.), Dept. of Math., Univ. of Maryland, College Park, MD, Lecture Notes #20, 1981.

14. A. G. WERSCHULZ, "Does increased regularity lower complexity?," *Math. Comp.*, v. 42, 1984, pp. 66–93.

15. A. G. WERSCHULZ, "Finite element methods are not always optimal," (Submitted for publication).

16. A. ŽENIŠEK, "Hermite interpolation on simplexes in the finite element method," *Proceedings EquaDiff 3*, J. E. Purkyně University, Brno, 1972, pp. 271–277.