What is the complexity of solution-restricted operator equations?

Technical Report CUCS-009-95

Arthur G. Werschulz

Department of Computer and Information Sciences, Fordham University
Fordham College at Lincoln Center
New York, NY 10023
and
Department of Computer Science
Columbia University
New York, NY 10027

May 25, 1995

Abstract. We study the worst case complexity of operator equations $Lu = f$, where $L\colon G \to X$ is a bounded linear injection, G is a Hilbert space, and X is a normed linear space. Past work on the complexity of such problems has generally assumed that the class F of problem elements f to be the unit ball of X. However, there are many problems for which this choice of F yields unsatisfactory results. Mixed elliptic-hyperbolic problems are one example, the difficulty being that our technical tools are not strong enoguh to give good complexity bounds. Ill-posed problems are another example, because we know that the complexity of computing finite-error approximations is infinite if F is a ball in X. In this paper, we pursue another idea. Rather than directly restrict the class F of problem elements f, we will consider problems that are solution-restricted, i.e., we restrict the class U of solution elements u. In particular, we assume that U is the unit ball of a Hilbert space W continuously embedded in G.

The main idea is that our problem can be reduced to the standard *approximation problem* of approximating the embedding of W into G. This allows us to characterize optimal information and algorithms for our problem. Then, we consider specific applications. The first application we consider is any problem for which G and W are standard Sobolev Hilbert spaces; we call this the "standard problem" since it includes many problems of practical interest. We show that finite element information and generalized Galerkin methods are nearly optimal for standard problems. We then look at elliptic boundary-value problems, Fredholm integral equations of the second kind, the Tricomi problem (a mixed hyperbolic-elliptic problem arising in the study of transonic flow), the inverse finite Laplace transform, and the backwards heat equation. (Note that with the exception of the backwards heat equation, all of these are standard problems. Moreover, the inverse finite Laplace transform and the backwards heat equation are ill-posed problems.) We determine the problem complexity and derive nearly optimal algorithms for all these problems.

This research was supported in part by the National Science Foundation under Grant CCR-91-01149.

1. Introduction

Operator equationsd $Lu = f$ are among the most important problems of applied mathematics. We are interested in the ε-complexity of such problems; that is, we want to find the minimal cost of computing ε-accurate approximations and to find algorithms that yield such approximations with (nearly) minimal cost. Past work on the worst case complexity of such problems has generally assumed that the class F of problem elements f is a unit ball. As we shall see, this assumption is not strong enough for many classes of problems. In this paper, we assume instead that the class U of solution elements u is a unit ball. Such problems are called "solution-restricted" problems. We study the complexity of solution-restricted operator equations $Lu = f$ in this paper.

More precisely, suppose that $L\colon G \to X$ is a bounded linear injection, where G is a Hilbert space and X is a normed linear spaces. Let $F \subset X$ be a fixed class of problem elements f. We wish to solve $Lu = f$ for $f \in F$, our only knowledge of any f being the values of a finite set of linear functionals of f, each evaluation having cost c. Since we only have partial information about f, we can only calculate approximations of $u = L^{-1}f$. Our goal is to calculate ε-approximations (i.e., approximations with error at most ε) with minimal cost. This is, of course, a problem of information-based complexity (IBC); see [22] for further discussion. We will solve our problem in a worst case setting, so that the error and cost of an algorithm are given by their maximum values over all f.

Researchers in IBC have been most successful in getting good complexity bounds for operator equation problems (for F a ball in X) whenever the solution operator L^{-1} has been bounded. One such class of problems is the solution of elliptic operator equations $Lu = f$, for which we have found a wealth of complexity-related results. For instance, one popular method for elliptic problems is the finite element method (FEM); we have been able to find conditions that are necessary and sufficient for the FEM to be a nearly optimal error or nearly optimal complexity algorithm. Many of these results use a "shift theorem," which says that if L is elliptic of order $2m$, then f has smoothness r (i.e., r derivatives in some sense) iff u has smoothness $r + 2m$. Similarly, since there is a shift theorem (with $m = 0$) for Fredholm integral equations of the second kind, we have been successful in proving results about the complexity of second-kind Fredholm problems, as well as characterizing nearly optimal FEMs for such problems. An exhaustive treatment of this subject may be found in [28, Chapters 5 and 6].

Unfortunately, there has been far less success to date in dealing with operator equations for which there is no shift theorem. For example, consider the Tricomi problem. This is a mixed elliptic-hyperbolic problem arising in the study of transonic flow across an airfoil, see [7, Chapter X] and [13]. Since there is no shift theorem for the Tricomi problem,[1] we have not been able to use these techniques to obtain sharp complexity bounds for the Tricomi problem.

Things are even worse when we consider the solution of ill-posed problems $Lu = f$, in which the solution element u does not depend continuously on the problem element f, e.g., when L is compact. We have a strong negative result on the worst case complexity of ill-posed problems when F is a ball in X, namely, that there exists no algorithm having

[1] The best global smoothness result we know for solutions of the Tricomi problem is that of [2], which only proves first-order smoothness for a solution of a second-order problem, so there is a smoothness gap.

finite error (see [26]). Hence the complexity of computing an ε-approximation is infinite, no matter how large we choose ε to be. This means we cannot solve ill-posed problems in a worst case setting. If we need to solve such problems, we must go to a different setting, such as the average case setting. For further discussion, see [23], [27], [28].

Summarizing, we see that if the class F is a ball in X, good worst case complexity bounds for operator equations $Lu = f$ have so far eluded us, except for L enjoying special properties. If we wish to solve such problems in the worst case setting for L not satisfying these properties, we need to look at other classes F.

This paper uses an idea of Tikhonov [20] that is often used in the solution of ill-posed problems (see also [15] for a fuller development, as well as the discussion in [19]). Instead of solving the problem $Lu = f$ under the a priori assumption that f belongs to a known set F (typically a ball in X), we assume that u belongs to a known set U. Thus for such a *solution-restricted* operator equation, we restrict the solution elements instead of the problem elements.[2]

The main point of this paper is that we can often get good complexity results for solution-restricted problems, simply because much of the work can be rephrased in terms of the well-studied problem of approximating an identity embedding. We now outline our main results.

In Section 2, we formally describe the problem to be solved. The class U will be the unit ball of a Hilbert space W that is continuously embedded in G. Part of our problem description includes a discussion of the class Λ of permissible information operations. We will be mainly interested in two classes Λ. The first is the usual class Λ^* of continuous linear functionals over X, which has been well-studied in previous work on IBC. The second class of permissible information operations is the class Λ^F of linear functionals on F that are bounded on F. This new class of information operations is quite natural for our problem, since it is the largest class of linear functionals defined on F.

In Section 3, we show that a solution-restricted operator equation can be reduced to the *approximation problem* of approximating the embedding of W into G. Since the approximation problem is a well-studied problem of information-based complexity, we can easily adapt known results for the approximation problem to our problem. As a result, we quickly determine optimal algorithms and information for our problem. Moreover, we find that the minimal radii are the same for the classes Λ^* and Λ^F. This means that in principle, there is no need to consider Λ^F. However, our introduction of Λ^F is no mere artifice. Indeed, it is often easier to first consider the problem for $\Lambda = \Lambda^F$, and then to approximate Λ^F-information by Λ^*-information, rather than to directly consider Λ^*-information at the start. Moreover, there are situations for which Λ^F-information is more natural than Λ^*-information.

In Section 4, we use these results to study several important problems. The first problem we study is called the "standard problem," which is any problem for which G and W are standard Sobolev Hilbert spaces $H^q_{\mathrm{bd}}(\Omega)$ and $H^r_{\mathrm{bd}}(\Omega)$ (the "bd" meaning that certain homogeneous boundary conditions may be imposed), with $q < r$ and $\Omega \subset \mathbb{R}^d$. Note that since L and X are not specified, this problem is really a meta-problem, which includes many

[2] Of course, since our class F of problem elements is now given by $F = L^{-1}(U)$, we do have an implicit definition of F. However, F need not be a ball in X.

important examples. Also note that there is no requirement that L have a bounded inverse, so that the standard problem covers both well-posed and ill-posed problems. Hence we are solving any linear operator equation $Lu = f$, with error being measured in the $H^q_{\mathrm{bd}}(\Omega)$-norm, subject to the constraint that $\|u\|_{H^r_{\mathrm{bd}}(\Omega)} \le 1$. Although the standard problem is quite general, we are able to develop complexity bounds and to find optimal algorithms and information for the standard problem. We find that the nth minimal radius is $\Theta(n^{-(r-q)/d})$ and the ε-complexity is $c\cdot\Theta\big((1/\varepsilon)^{d/(r-q)}\big)$. The optimal information is given by eigenvectors of E^*E, where E is the embedding of $H^r_{\mathrm{bd}}(\Omega)$ into $H^q_{\mathrm{bd}}(\Omega)$. We show that the eigenproblem for E^*E can be expressed as a generalized eigenproblem for a partial differential equation. Unfortunately, this eigenproblem usually does not have closed-form solutions. Hence, we need to find more accessible nearly optimal information. We are able to show that finite element information is nearly optimal, and that generalized Galerkin methods turn out to be nearly optimal algorithms.

In the remainder of Section 4, we analyze several important specific applications. The first class of applications is elliptic boundary-value problems. We consider a $2m$th-order problem, with error measured in the energy norm, which is equivalent to the Sobolev $\|\cdot\|_{H^m(\Omega)}$-norm. This is then set as a standard problem with $G = H^m_0(\Omega)$ and $W = H^r(\Omega)$. The nth minimal error is $\Theta(n^{-(r-m)/d})$, and the ε-complexity is $c\cdot\Theta\big((1/\varepsilon)^{d/(r-m)}\big)$. For $\Lambda = \Lambda^F$, we find that the usual finite element method is nearly optimal, the proof not requiring a shift theorem. This means that we can handle elliptic problems that do not admit a shift theorem.[3] For the sake of completeness, we develop the results for $\Lambda = \Lambda^*$, although the method that results would probably not be used in practice.

We next look at Fredholm integral equations of the second kind. This is set up as a standard problem with $G = L_2(\Omega)$ and $W = H^r(\Omega)$. We find that the nth minimal error is $\Theta(n^{-r/d})$ and the ε-complexity is $c\cdot\Theta\big((1/\varepsilon)^{d/r}\big)$. Finite element information is nearly optimal, and we exhibit nearly optimal generalized Galerkin methods.

Next, we look at the Tricomi problem. Once again, we set this up as a standard problem with $G = L_2(\Omega)$ and $W = H^r(\Omega)$. Our results are essentially the same as for the Fredholm problem of the second kind. The nth minimal error is $\Theta(n^{-r/d})$ and the ε-complexity is $c\cdot\Theta\big((1/\varepsilon)^{d/r}\big)$. We find that finite element information and generalized Galerkin methods are nearly optimal.

Our next application is the inverse finite Laplace transform, which is a Fredholm integral equation of the first kind and hence is ill-posed. This particular problem arises in the study of "measurement of the distribution of an absorbing gas (such as ozone in the earth's atmosphere) from the spectrum of scattered light", see [21, pp. 12–13]. This may be set up as a standard problem with $G = L_2(\Omega)$ and $W = H^r(\Omega)$. We find that the nth minimal error is $\Theta(n^{-r/d})$ and the ε-complexity is $c\cdot\Theta\big((1/\varepsilon)^{d/r}\big)$. Once again, finite element information is nearly optimal, as are generalized Galerkin methods.

To show that our techniques are not limited to standard problems, we close this paper by studying a problem that is not a standard problem. Our final application is the heat equation running backwards in time, with final data f. That is, we want to know what the

[3] This means that our analysis can be easily extended to include problems for which the coefficients or the boundary of Ω are not very smooth, or problems having boundary conditions of mixed Dirichlet-Neumann type.

temperature distribution was at some time $t = -t_0$ in the past (where $t_0 > 0$), where the temperature distribution at time $t = 0$ is given by the function f. Suppose that both the final data and solution are measured in the L_2-sense. Then the backwards heat equation is ill-posed, said ill-posedness being related to the second law of thermodynamics, see [12]. To express this as a solution-restricted problem, we let $t_1 > t_0$, and solve the problem under the condition that the solution at time $-t_1$ must have finite L_2-norm, see also [10] for further discussion. We then find that the nth minimal error is $\exp\big(-\pi^2(n+1)^2(t_1 - t_0)\big)$, so that the ε-complexity is $c \cdot \Theta(\sqrt{\ln(1/\varepsilon)})$. Moreover, if we truncate the usual series representation for the solution of the heat equation, we get an optimal algorithm. For further discussion of the complexity of this problem, see [27].

We close this introduction by noting that one of our main assumptions has been that the calss Λ of permissible information operations has been either Λ^F or Λ^*. There is a third class that one could investigate. Suppose that X is a function space on some domain $\Omega \subset \mathbb{R}^d$, as is generally the case in most applications. Then we could study the class Λ^{std} of standard information, i.e., evaluations of f (or some of its derivatives) at points in Ω. At this time, we only have results for standard information under (what we feel are) unnecessarily-restrictive conditions on L, conditions that do not apply (for example) when L is compact, i.e., when the original problem is ill-posed. We hope that we will be able to coherently deal with standard information for solution-restricted operator equations in future work.

2. Problem description

Let G be a Hilbert space, and let X be a normed linear space. We assume that both G and X are infinite-dimensional. Let $L\colon G \to X$ be a bounded linear injection, whose range D is dense in X. We define a solution operator $S\colon D \subseteq X \to G$ as

$$u = Sf \iff Lu = f \qquad \text{for } u \in G \text{ and } f \in D. \tag{2.1}$$

Note that if L does not have a closed range, then S is only densely defined.

Remark: Note that we restrict our attention to the Hilbert case. We do this for two reasons. The first is for ease of exposition, while the second is that all the examples we consider are instances of the Hilbert case. We note in passing that many of the results contained here also hold for the case of general normed linear spaces. □

Let F be a balanced, convex subset of D, and let Λ be a class of continuous linear functionals on F (more precisely, a class of functionals whose extensions to the linear hull of F are continuous linear functionals). Throughout this paper, we will be especially interested in the following classes Λ:

(1) $\Lambda = \Lambda^*$, the class of all continuous linear functionals defined over X. This is a "standard" choice of Λ for many problems arising in information-based complexity; see [22].
(2) $\Lambda = \Lambda^F$, the class of all linear functionals on F that are bounded on F. Note that $\Lambda^* \subseteq \Lambda^F$, with equality when F is a ball in X. The strict inclusion $\Lambda^* \subset \Lambda^F$ holds, e.g., when F is a compact subset of X. Our interest in using Λ^F instead of Λ^*

is that we can use our a priori knowledge that $f \in F$ to allow us to expand the possible choice of information functionals from those that must be defined over all of X to those that need be defined only over F, which is a (possibly small) subset of X.

Our abstract setting is the usual one of information-based complexity, see, e.g., [22] and [28]. We wish to compute approximate values of Sf for $f \in F$, given a finite number of information values $\lambda(f)$ for some elements $\lambda \in \Lambda$. However, our point of departure will be to assume that there is a subset $U \subseteq G$ such that $F = L(U)$. Thus

$$f \in F \iff u = Sf \in U.$$

(Thus, F is a subset of D.) Note that we are now restricting our solution elements u rather than our problem elements f. For this reason, we will call our problem a *solution-restricted* operator equation.

In particular, we will assume in this paper that U is the unit ball of a Hilbert space W that is continuously embedded in G. That is,

$$U = \{ Ew : \|w\|_W \leq 1 \}, \tag{2.2}$$

where the embedding mapping $E\colon W \to G$, defined as $Ew = w$ for $w \in W$, is continuous and dense. To simplify the exposition in what follows, we assume without essential loss of generality that $\|E\| \leq 1$. This assumption holds for many, but not all, cases of practical interest.[4] Since $Ew = w$ for $w \in W$, we will often write w instead of Ew whenever this will simplify the presentation, provided that no confusion will result.

We close this section by recalling some standard terminology from [22]. Let $\mathbb{N}$ denote the nonnegative integers. For any $n \in \mathbb{N}$, we say that N is *information from Λ of cardinality n* if there exist linearly independent linear functionals $\lambda_1, \ldots, \lambda_n \in \Lambda$ such that

$$Nf = [\lambda_1(f), \ldots, \lambda_n(f)] \qquad \forall f \in F. \tag{2.3}$$

We let Λ_n denote the class of information from Λ whose cardinality is at most n. An *algorithm* ϕ using N is a mapping $\phi\colon N(F) \to G$, its error being given by

$$e(\phi, N) = \sup_{f \in F} \|Sf - \phi(Nf)\|_G.$$

The *radius of information* N is given as

$$r(N) = \inf_{\phi} e(\phi, N),$$

the infimum being over all algorithms ϕ using N. An algorithm ϕ_N for which $e(\phi_N, N) = r(N)$ is an *optimal error algorithm* using N.

[4] For example, this assumption holds for the embedding $H^r(\Omega) \hookrightarrow H^q(\Omega)$ of Sobolev spaces ($q < r$), but it does not hold for the embeddings given by either the general statement of Sobolev's embedding theorem or by the Rellich-Kondrasov theorems (see, e.g., [6, pg. 114] for statements).

The nth *minimal radius* in Λ is defined to be

$$r(n,\Lambda)=\inf_{N\in\Lambda_n} r(N),$$

and information $N_n^*\in\Lambda_n$ is nth *optimal information* if $r(N_n^*)=r(n,\Lambda)$. (Note that since F is balanced and convex, we are only considering nonadaptive information of fixed cardinality; see [22, pp. 57–60].) An optimal error algorithm using nth optimal information is said to be an nth *minimal error algorithm.*

We turn these error-related concepts into complexity-related concepts by introducing a model of computation. Our model is the standard one of [22], see also [28, pg. 27]. For any $\lambda\in\Lambda$ and any $f\in F$, the cost of evaluating $\lambda(f)$ is c, and the cost of basic combinatory operations is 1. Typically, $c\gg 1$. Then the cost of an algorithm ϕ using information N is given as

$$\operatorname{cost}(\phi,N)=\sup_{f\in F}\operatorname{cost}(\phi,N,f),$$

where for $f\in F$, we let $\operatorname{cost}(\phi,N,f)$ denote the cost of computing $\phi(Nf)$ under this model of computation. As an example, suppose that ϕ is a *linear* algorithm using N given by (2.3), i.e., there exist $g_1,\dots,g_n\in G$ such that $\phi(Nf)=\sum_{j=1}^n\lambda_j(f)g_j$ for $f\in F$; then $\operatorname{cost}(\phi,N)\le(c+2)n-1$.

Finally, we say that for $\varepsilon>0$, the ε-*complexity* in the class Λ is

$$\operatorname{comp}(\varepsilon,\Lambda)=\inf\{\,\operatorname{cost}(\phi,N): e(\phi,N)\le\varepsilon\,\},$$

the infimum being over all algorithms ϕ using information N of finite cardinality from Λ, and an algorithm ϕ_ε using information N_ε of finite cardinality from Λ for which

$$\operatorname{cost}(\phi_\varepsilon,N_\varepsilon)=\operatorname{comp}(\varepsilon,\Lambda)$$

is an (ε,Λ)-*optimal complexity algorithm.* Since we cannot often find optimal complexity algorithms, we will usually be happy to settle for *nearly optimal complexity algorithms*, which compute ε-approximations with cost at most a constant multiple of the ε-complexity, said constant being independent of ε. One way of doing this involves the (ε,Λ)-*cardinality number*

$$m(\varepsilon,\Lambda)=\inf\{\,n: r(n,\Lambda)\le\varepsilon\,\}.$$

Suppose that $N\in\Lambda_n$ is information of cardinality $n=m(\varepsilon,\Lambda)$ for which $r(N)=\varepsilon$, and that ϕ is a linear optimal error algorithm using N. Then

$$c\,m(\varepsilon,\Lambda)\le\operatorname{comp}(\varepsilon,\Lambda)\le\operatorname{cost}(\phi,N)\le(c+2)m(\varepsilon,\Lambda)-1,$$

and ϕ is an (ε,Λ)-nearly optimal complexity algorithm for our problem.

3. Reduction to the approximation problem

Our main strategy for finding optimal algorithms for solution-restricted operator is to reduce such problems to the *approximation problem* of approximating the embedding $E\colon W \to G$.

Let Λ be a class of linear functionals on F, and let N be information from Λ of finite cardinality. Suppose that ϕ is an algorithm using N. For any $f \in F$, we have

$$\|Sf - \phi(Nf)\|_G = \|Ew - \phi(N_L w)\|_G, \tag{3.1}$$

where $Ew = Sf$ and

$$N_L = N \circ LE. \tag{3.2}$$

Note that we are being somewhat cautious in explicitly using the embedding operator E, so we can keep closer track of the spaces in which various elements may be found. Also note that N and N_L are information operators on the spaces X and W, respectively.

Equation (3.1) tells us that the error of any algorithm for our problem equals the error of the same algorithm (using different information) for the approximation problem. In what follows, we will let e^{app}, r^{app}, m^{app}, and $\mathrm{comp}^{\mathrm{app}}$ denote the error, radius, cardinality number, and complexity for the approximation problem.

Using [22, Chapter 3], we then find the following:

Theorem 3.1. *Let Λ be a class of linear functionals on F.*

(1) *Let N be information from Λ of finite cardinality.*

(a) *For any algorithm ϕ using N, we have*

$$e(\phi, N) = e^{\mathrm{app}}(\phi, N_L),$$

where N_L is defined by (3.2).

(b) *An algorithm ϕ is an optimal error algorithm using N for our problem iff ϕ is an optimal error algorithm using N_L for the approximation problem, and*

$$r(N) = r^{\mathrm{app}}(N_L).$$

(2) *For any nonnegative integer n,*

$$r(n, \Lambda) = r^{\mathrm{app}}(n, \Lambda_L),$$

where

$$\Lambda_L = E^* L^* \Lambda.$$

Hence,

$$m(\varepsilon, \Lambda) = m^{\mathrm{app}}(\varepsilon, \Lambda_L).$$

(3) *An (ε, Λ_L)-(nearly) optimal complexity algorithm for the approximation problem is an (ε, Λ)-(nearly) optimal complexity algorithm for our problem, and*

$$\mathrm{comp}(\varepsilon, \Lambda) = \mathrm{comp}^{\mathrm{app}}(\varepsilon, \Lambda_L).$$

In particular, if $N_L \in (\Lambda_L)_n$ is information of cardinality $n = m^{\mathrm{app}}(\varepsilon, \Lambda_L)$ for which $r(N_L) = \varepsilon$ and ϕ is a linear optimal error algorithm using N, then

$$c\, m(\varepsilon, \Lambda) \le \mathrm{comp}(\varepsilon, \Lambda) \le \mathrm{cost}(\phi, N) \le (c+2) m(\varepsilon, \Lambda) - 1,$$

and so ϕ is a nearly optimal complexity algorithm for our problem. $\square$

The results in Theorem 3.1 essentially hold even in the case where the spaces G and W are not Hilbert spaces. However, we are dealing only with the Hilbert case in this paper. Let us use this fact to derive specifc formulas for a linear optimal error algorithm.

Let N be information of cardinality n. Then there exist linearly independent $\lambda_1, \dots, \lambda_n \in X^*$ such that

$$Nf = [\lambda_1(f), \dots, \lambda_n(f)] \qquad \forall f \in F. \tag{3.3}$$

Let

$$w_j = E^* L^* \lambda_j \qquad (1 \le j \le n). \tag{3.4}$$

Since $f_1, \dots, f_n$ are linearly independent and $E^* L^*$ is an injection, we see that $w_1, \dots, w_n$ are linearly independent. Define

$$\mathscr{S}_N = \operatorname{span}\{w_1, \dots, w_n\}.$$

For $f \in F$, let $u_N \in \mathscr{S}_N$ satisfy

$$\lambda_i(Lu_N) = \lambda_i(f) \qquad (1 \le i \le n). \tag{3.5}$$

Note that this is a generalized Galerkin method, i.e., a Galerkin method with different spaces of test and trial functions. Clearly the dependence of u_N on f is only through the information Nf, so we may write

$$u_N = \phi_N(Nf) \qquad \forall f \in F. \tag{3.6}$$

Lemma 3.1. *Let N be given by (3.3).*

(1) *The linear algorithm ϕ_N given by (3.6) is an optimal error algorithm for N.*
(2) *If $\{w_1, \dots, w_n\}$ is a W-orthonormal set, then*

$$\phi_N(Nf) = \sum_{j=1}^{n} \lambda_j(f) w_j = \sum_{j=1}^{n} \langle Sf, w_j \rangle_W w_j \qquad \forall f \in F.$$

Proof: Using (3.4), we see that

$$\lambda_i(Lu_N) = \langle u_N, w_i \rangle_W \qquad \text{and} \qquad \lambda_i(f) = \langle Sf, w_i \rangle_W.$$

Hence $u_N \in \mathscr{S}_N$ satisfies (3.5) iff it satisfies

$$\langle u_N, w_i \rangle_W = \langle Sf, w_i \rangle_W \qquad (1 \le i \le n). \tag{3.7}.$$

Now use [22, Theorem 5.5.3]. □

As mentioned in Section 2, we will deal with the classes $\Lambda = \Lambda^F$ and $\Lambda = \Lambda^*$ of permissible information operations. In the remainder of this section, we show that the nth minimal radii are the same for these classes, being given by the (same) Gelfand n-width. We also construct nth optimal (or nearly optimal) information for these classes.

First, we prove the following

LEMMA 3.2. *Let $N \in \Lambda_n^F$. For any $\delta > 0$, there exists information $N_\delta \in \Lambda_n^*$ such that*

$$r(N_\delta) \le (1+\delta)\, r(N).$$

PROOF: Since $N \in \Lambda_n^F$, there exist linearly independent $\lambda_1, \dots, \lambda_k \in \Lambda^F$ (with $k \le n$) such that

$$N = [\lambda_1, \dots, \lambda_k].$$

Since each functional $Sf \mapsto \lambda_i(f)$ is a bounded linear functional on W, there exist linearly independent $w_1, \dots, w_k \in W$ such that $\lambda_i(f) = \langle Sf, w_i \rangle_W$ for $f \in W$. Without loss of generality, we may assume that $w_1, \dots, w_k$ are orthonormal. Now let $\delta > 0$, and define

$$\rho = \frac{\delta\, r(N)}{1 + r(N)}.$$

Since $LE\colon W \to X$ is an injection, we see that $E^*L^*\colon X^* \to W$ is dense, and so there exist $\lambda_{1,\delta}, \dots, \lambda_{k,\delta} \in X^*$ such that

$$\|w_i - E^*L^*\lambda_{i,\delta}\|_W \le \frac{\rho}{k}.$$

We then define information $N_\delta \in \Lambda_n^*$ as

$$N_\delta = [\lambda_{1,\delta}, \dots, \lambda_{k,\delta}].$$

Recall that the radius of information for the Hilbert case is given by the simplified expression

$$r(N_\delta) = \sup_{\substack{w \in \ker(N_\delta)_L \\ \|w\|_W \le 1}} \|w\|_G = \sup_{\substack{Lw \in \ker N_\delta \\ \|w\|_W \le 1}} \|w\|_G. \tag{3.8),}$$

see [22, pg. 80]. Let $Lw \in \ker N$ with $\|w\|_W \le 1$. For $1 \le i \le k$, we have

$$(E^*L^*\lambda_{i,\delta})(w) = \lambda_{i,\delta}(Lw) = 0$$

and so

$$|\langle w_i, w \rangle_W| \le \|w_i - E^*L^*\lambda_{i,\delta}\|_{W^*} \|w\|_W \le \frac{\rho}{k}.$$

Let

$$w_0 = w - \sum_{j=1}^{k} \langle w_j, w \rangle_W w_j, \tag{3.9}$$

Note that

$$\lambda_i(Lw_0) = E^*L^*\lambda_i(w_0) = \langle w_i, w_0 \rangle_W = 0 \qquad (1 \le i \le k),$$

and so $Lw_0 \in \ker N$. Moreover

$$\|w_0\|_W \le \|w\|_W + \sum_{j=1}^{k} |\langle w_j, w \rangle_W|\, \|w_j\|_W \le 1 + \rho.$$

It now follows that

$$\|w_0\|_G \le \sup_{\substack{Lv\in\ker N\\ \|v\|_W\le 1+\rho}} \|v\|_G = (1+\rho)\, r(N).$$

From (3.9) and the definition of ρ, it now follows that

$$\|w\|_G \le \|w_0\|_G + \sum_{j=1}^{k} |\langle w_j, w\rangle_W|\, \|w_j\|_G \le (1+\rho)\, r(N) + \rho = (1+\delta)\, r(N).$$

Since w is an arbitrary element in the W-unit ball such that $Lw \in \ker N$, the lemma follows immediately from (3.8). □

Recall that the *Gelfand n-width* of the embedding E is defined as

$$d^n\big(E(W), G\big) = \inf_{W^n\in\mathscr{W}^n} \sup_{\substack{w\in W^n\\ \|w\|_W\le 1}} \|w\|_G,$$

where $\mathscr{W}^n$ is the class of W-subspaces whose codimension is at most n. From [18, Corollary II.7.3], there always exists an nth optimal Gelfand subspace of E, i.e., a subspace $W^n \in \mathscr{W}^n$ at which this infimum is attained. Without loss of generality, we may assume that

$$W^n = \ker[\langle\cdot, w_1\rangle_W, \dots, \langle\cdot, w_n\rangle_W], \tag{3.10}$$

where $\{w_1, \dots, w_n\}$ is a W-orthonormal set.

Theorem 3.2.

(1) *For any $n \in \mathbb{N}$,*

$$r(n, \Lambda^F) = r(n, \Lambda^*) = d^n\big(E(W), G\big).$$

(2) *Suppose that $\Lambda = \Lambda^F$. Let $w_1, \dots, w_n$ be as in (3.10). Define information $N_n \in \Lambda^F$ by*

$$N_n f = [\langle Sf, w_1\rangle_W, \dots, \langle Sf, w_n\rangle_W] \qquad \forall f \in F.$$

Then

$$r(N_n) = r(n, \Lambda^*) = d^n\big(E(W), G\big),$$

i.e., $N_n \in \Lambda_n^F$ is nth optimal information.

(3) *Suppose that $\Lambda = \Lambda^*$. For $\delta > 0$, let*

$$\rho = \frac{2\,\delta\, d^n\big(E(W), G\big)}{1 + 2\,\delta\, d^n\big(E(W), G\big)},$$

and choose $\lambda_{1,\delta}, \dots, \lambda_{n,\delta} \in X^*$ such that

$$\|w_i - E^* L^* \lambda_{i,\delta}\|_W \le \frac{\rho}{n}.$$

Define information $N_{n,\delta} \in \Lambda_n^*$ by

$$N_{n,\delta} = [\lambda_{1,\delta}, \dots, \lambda_{n,\delta}].$$

Then

$$r(N_{n,\delta}) \le (1+\delta)\, r(n, \Lambda^*),$$

so that $N_{n,\delta} \in \Lambda^*$ is nth nearly optimal information.

PROOF: Parts (1) and (2) follows immediately from Lemma 2 and the observation that if $\Lambda = \Lambda_F$, then $\Lambda_L = \Lambda^*$. We need only prove part (3). For $\delta > 0$, we have

$$\rho = \frac{\delta\, r(N_n)}{1 + \delta\, r(N_n)}$$

by part (1). Using the notation of Lemma 3.2, we see that for $N = N_n$, we have $N_\delta = N_{n,\delta}$, and so

$$r(N_{n,\delta}) \le (1+\delta)\, r(N_n) = (1+\delta)\, r(n, \Lambda^*) = (1+\delta)\, d^n\big(E(W), G\big),$$

as required. □

Suppose for a moment that E is not a compact embedding. Then $\lim_{n\to\infty} d^n\big(E(W), G\big)$ is strictly positive (see [18, Proposition II.7.4]). This means that the nth minimal radius is bounded away from zero for the classes Λ^F and Λ^*. In short, if E is not compact, then the problem is not *convergent*, i.e., we cannot get arbitrarily good approximations at finite cost.

Hence, in the remainder of this paper, we shall assume that W is compactly embedded in G. This implies that the space W has an orthonormal basis consisting of eigenvectors of E^*E. Thus there exist $\gamma_1 \ge \gamma_2 \ge \cdots > 0$ with $\lim_{n\to\infty} \gamma_j = 0$, and a complete orthonormal basis $\{z_j\}_{j=1}^{\infty}$ for W such that

$$E^*E z_j = \gamma_j^2 z_j \qquad \text{for } j = 1, 2, \ldots.$$

We then find that for any $n \in \mathbb{N}$,

$$d^n\big(E(W), G\big) = \gamma_{n+1}, \tag{3.11}$$

see [18, Theorem IV.2.2].

We first look at optimal information and algorithms for the class Λ^F. Let

$$N_n f = [\langle Sf, z_1\rangle_W, \ldots, \langle Sf, z_n\rangle_W] \qquad \forall f \in F, \tag{3.12}$$

where $z_1, \ldots, z_n$ are the eigenvectors of E^*E corresponding to the n largest eigenvalues of E^*E. Clearly we have $N_n \in \Lambda_n^F$. Define an algorithm ϕ_n using N_n as

$$\phi_n(N_n f) = \sum_{j=1}^{n} \langle Sf, z_j\rangle_W z_j \qquad \forall f \in F. \tag{3.13}$$

THEOREM 3.3.

(1) *For any $n \in \mathbb{N}$, we have*

$$e(\phi_n, N_n) = r(N_n) = r(n, \Lambda^F) = d^n\big(E(W), G\big) = \gamma_{n+1}.$$

Hence N_n is nth optimal information and ϕ_n is an nth minimal error algorithm.

(2) *Let $\varepsilon > 0$. Then the ε-cardinality number for Λ^F is given by*

$$m(\varepsilon, \Lambda^F) = \inf\{\, \textit{integers } n \ge 0 : \gamma_{n+1} \le \varepsilon \,\}.$$

Moreover, let $N_n \in \Lambda_n^F$ be the information given by (3.12) and let ϕ_n be the algorithm given by (3.13), with $n = m(\varepsilon, \Lambda)$. Then

$$c\, m(\varepsilon, \Lambda^F) \le \mathrm{comp}(\varepsilon, \Lambda^F) \le \mathrm{cost}(\phi_n, N_n) \le (c+2) m(\varepsilon, \Lambda^F) - 1.$$

Hence for $c \gg 1$, the algorithm ϕ_n using information N_n is a nearly optimal complexity algorithm.

PROOF: Immediate from (3.11), along with Theorems 3.1 and 3.2, as well as Lemma 3.1.□

We now consider optimal information and algorithms for the class Λ^*. Let $\delta > 0$, and choose $\lambda_{1,\delta}, \ldots, \lambda_{n,\delta} \in X^*$ such that

$$\|z_j - L^*\lambda_{j,\delta}\|_G \leq \frac{\delta\gamma_{n+1}}{n}, \tag{3.14}$$

where we recall that $z_1, \ldots, z_n$ are the eigenvectors of E^*E corresponding to the n largest eigenvalues of E^*E. Our information $N_{n,\delta} \in \Lambda_n^*$ is defined as

$$N_{n,\delta} f = [\lambda_{1,\delta}(f), \ldots, \lambda_{n,\delta}(f)] \qquad \forall f \in F. \tag{3.15}$$

Define an algorithm $\phi_{n,\delta}$ using $N_{n,\delta}$ as

$$\phi_{n,\delta}(N_{n,\delta} f) = \sum_{j=1}^{n} \lambda_{j,\delta}(f) z_j \qquad \forall f \in F. \tag{3.16}$$

We then have

THEOREM 3.4.

(1) *For any $n \in \mathbb{N}$ and any $\delta > 0$, we have*

$$e(\phi_{n,\delta}, N_{n,\delta}) \leq (1+\delta)\gamma_{n+1}.$$

Hence $N_{n,\delta}$ is nth nearly optimal information and $\phi_{n,\delta}$ is an nth nearly minimal error algorithm.

(2) *Let $\varepsilon > 0$. Then the ε-cardinality number for Λ^* is given by*

$$m(\varepsilon, \Lambda^*) = \inf\{\, \text{integers } n \geq 0 : \gamma_{n+1} \leq \varepsilon \,\}.$$

Suppose that γ_n is strictly monotonically decreasing with n. Let $N_{n,\delta} \in \Lambda_n^F$ be the information given by (3.15) and let $\phi_{n,\delta}$ be the algorithm given by (3.16), with $n = m(\varepsilon, \Lambda^*) + 1$ and δ satisfying

$$0 < \delta < \frac{\gamma_n}{\gamma_{n+1}} - 1.$$

Then

$$c\, m(\varepsilon, \Lambda^F) \leq \operatorname{comp}(\varepsilon, \Lambda^*) \leq \operatorname{cost}(\phi_{n,\delta}, N_{n,\delta}) \leq (c+2)\big(m(\varepsilon, \Lambda^*) + 1\big) - 1.$$

Hence for $c \gg 1$, the algorithm $\phi_{n,\delta}$ using information $N_{n,\delta}$ is a nearly optimal complexity algorithm.

PROOF: Note that part (2) follows from Theorem 3.1 and part (1). So we need only prove part (1). To do this, let $n \in \mathbb{N}$, and let $\delta > 0$. Choose $f \in F$. For any j, we have

$$|\langle Sf, z_j \rangle_W - \lambda_{j,\delta}(f)| = |\langle Sf, z_j - L^*\lambda_{j,\delta} \rangle_G| \leq \|Sf\|_G \|z_j - L^*\lambda_{j,\delta}\|_G \leq \frac{\delta\gamma_{n+1}}{n}.$$

Recalling the definition of N_n and ϕ_n from Theorem 3.3, we then have

$$\|\phi_n(N_n f) - \phi_{n,\delta}(N_{n,\delta} f)\|_G \leq \sum_{j=1}^{n} |\langle Sf, z_j \rangle_W - \lambda_{j,\delta}(f)| \|z_j\|_G \leq \delta\gamma_{n+1}.$$

So

$$\|Sf - \phi_{n,\delta}(N_{n,\delta} f)\|_G \leq \|Sf - \phi_n(N_n f)\|_G + \|\phi_n(N_n f) - \phi_{n,\delta}(N_{n,\delta} f)\|_G \leq (1+\delta)\gamma_{n+1}.$$

Taking the supremum over all such f, we find the desired bound on $e(\phi_{n,\delta}, N_{n,\delta})$. Now that we have proved this bound, the rest of the result follows from (3.11) and Theorem 3.2, as well as Lemma 3.1. □

Remark: This result tells us that $\phi_{n,\delta}$ is an nth nearly minimal error algorithm in Λ^*. Of course, this implies that $\phi_{n,\delta}$ is an nearly optimal error algorithm using the nth nearly optimal information $N_{n,\delta}$. Note that we decided to use the algorithm $\phi_{n,\delta}$, instead of the algorithm $\phi_{N_{n,\delta}}$, i.e., the optimal error algorithm using $N_{n,\delta}$ given by Lemma 3.1. We did this because the algorithm $\phi_{n,\delta}$ is simpler than the optimal error algorithm $\phi_{N_{n,\delta}}$, since the latter would have required either an orthogonalization of the vectors $E^*L^*\lambda_{1,\delta}, \ldots, E^*L^*\lambda_{n,\delta}$ or the solution of an $n \times n$ linear system. Note that the increased error when using $\phi_{n,\delta}$ instead of $\phi_{N_{n,\delta}}$ is small compared to the error of $\phi_{n,\delta}$ and to the nth minimal radius. All things considered, it appears better to use a simpler nearly optimal error algorithm than a more complicated optimal error algorithm in this situation. □

Remark: The strict monotonicity assumption in Theorem 3.4 is not necessary, but only used to simplify the statement of the theorem. A more general (and more complicated) statement is possible for the case where E^*E has multiple eigenvalues, provided that the multiplicity of the eigenvalues does not increase super-exponentially. Since we will not need such a result, we will not pursue this further. □

4. Applications

In this section, we apply the previous results to several problems. We first look at a common situation, namely, a problem in which the spaces G and W are the Sobolev spaces $H^q_{\text{bd}}(\Omega)$ and $H^r_{\text{bd}}(\Omega)$, respectively. Here, $q < r$, and the "bd" indicates that the spaces may satisfy certain homogeneous boundary conditions. Any such problem will be called a "standard problem." We develop detailed results for standard problems. Once we have these results, we can use them to study specific instances of the standard problem. In particular, we will consider elliptic boundary-value problems, Fredholm integral equations of the second kind, mixed elliptic-hyperbolic problems, and the inverse finite Laplace transform. For all these problems, we find the problem complexity and derive nearly optimal algorithms. In particular, we discuss the optimality of Galerkin algorithms using finite element information. We then conclude by looking at the heat equation running backwards in time, an application that is not an instance of the standard problem. Suppose we know an a priori L_2-bound on the solution of the backwards heat equation at time $t = -t_1$ and that we want to solve the equation at time $t = -t_0$, where $0 < t_0 < t_1$. The problem element is "final data", i.e., the solution at time $t = 0$. We then find that by truncating the standard series representation of the solution, we get an optimal algorithm.

Note that two of our applications (the inverse finite Laplace transform and the backwards heat equation) are ill-posed problems. Hence we see that the techniques of this paper are powerful enough for us to determine that the ε-complexity of ill-posed solution-restricted problems is finite.

In what follows we use the standard terminology and notation for multi-indices, as well as Sobolev spaces, norms, and inner products. For details, consult any standard reference on elliptic boundary-value problems and finite element methods, such as [1], [3], [7, Chapter IV], [9], or [16].

The letter C will denote a generic constant whose value may change from one place to the next. All O-, Θ-, Ω-, and $\sim$-estimates will be independent of n or ε, depending on the context.

Finally, we note that in the specific applications we consider here, the space X is also a Sobolev Hilbert space, along with G and W. Hence, an information operation from Λ^* can be represented as an inner product over X. We shall do this consistently, without further comment.

4.1. The standard problem.

Our first problem is really a meta-problem, since it includes many important practical problems as particular instances. Let $\Omega \subset \mathbb{R}^d$ be a sufficiently smooth simply-connected region, which is bounded. Given q and r with $q < r$, we let $H^q_{\text{bd}}(\Omega)$ and $H^r_{\text{bd}}(\Omega)$ be closed subspaces of $H^q(\Omega)$ and $H^r(\Omega)$, respectively, for which $C^\infty_0(\Omega) \subset H^r_{\text{bd}}(\Omega)$ and $H^r_{\text{bd}}(\Omega) \hookrightarrow H^q_{\text{bd}}(\Omega)$. That is, functions in these spaces may satisfy homogeneous boundary conditions, and the conditions satisfied by $H^q_{\text{bd}}(\Omega)$-functions are also satisfied by functions in $H^r_{\text{bd}}(\Omega)$.

We now consider any problem for which $G = H^q_{\text{bd}}(\Omega)$ and $W = H^r_{\text{bd}}(\Omega)$, so that the mapping $E\colon H^r_{\text{bd}}(\Omega) \to H^q_{\text{bd}}(\Omega)$ is the usual inclusion embedding and F is the unit ball of $H^r_{\text{bd}}(\Omega)$. Since this kind of problem will be useful in later applications, we will call it the *standard problem*. Although we have specified neither the space X nor the mapping L, we can still discover much about such a problem. In particular, we show that optimal information in Λ^F is given by the solution of a generalized elliptic eigenproblem whenever $q, r \in \mathbb{N}$. We then show that for any r and q, the nth minimal radius is proportional to $n^{-(r-q)/d}$, so that the ε-complexity is proportional to $(1/\varepsilon)^{d/(r-q)}$. Since it is not generally possible to find a closed-form solution of this eigenproblem, we need to consider nearly optimal that is easier to obtain. We show that finite element information of degree k is nearly optimal if $k \geq r-1$. Having dealt with the case $\Lambda = \Lambda^F$, we then prove analogous results for the case $\Lambda = \Lambda^*$.

We first show that the eigenvectors and eigenvectors of E^*E are solutions of a generalized elliptic eigenproblem when $q, r \in \mathbb{N}$. For any $s \in \mathbb{N}$, we define the partial differential operator P_s as

$$P_s v = \sum_{|\alpha| \le s} (-1)^{|\alpha|} D^{2\alpha} v$$

From [1, Theorem 10.2], we see that for $s \le j \le 2s-1$, there exist partial differential operators $\Psi_{s,j}$ of exact order j, such that the Green's formula

$$\langle z, w \rangle_{H^s(\Omega)} = \langle P_s z, w \rangle_{L_2(\Omega)} + \sum_{j=0}^{s} \langle \Psi_{s,2s-1-j}(z), \partial_\nu^j w \rangle_{L_2(\partial\Omega)} \qquad \forall\, z, w \in C^\infty(\Omega) \quad (4.1.1)$$

holds, with ∂_ν denoting the outward-pointing normal derivative on $\partial\Omega$. We then have

LEMMA 4.1.1. *Let E be the embedding of $H^r_{\mathrm{bd}}(\Omega)$ into $H^q_{\mathrm{bd}}(\Omega)$. For $0 \le j \le r-1$, let*

$$\Gamma_{\gamma,j} = \begin{cases} \Psi_{r,2r-1-j} - \gamma^{-2}\Psi_{q,2q-1-j} & \text{for } 0 \le j \le q-1, \\ \Psi_{r,2r-1-j} & \text{for } q \le j \le r-1. \end{cases}$$

Then $E^*Ez = \gamma^2 z$ iff z and γ are a solution of the generalized elliptic eigenproblem

$$P_r z = \gamma^{-2} P_q z \qquad \text{in } \Omega, \tag{4.1.2}$$

subject to the homogeneous boundary conditions

$$\Gamma_{\gamma,j}(z)\Big|_{\partial\Omega} = 0 \qquad (0 \le j \le r-1). \tag{4.1.3}$$

PROOF: : We have $E^*Ez = \gamma^2 z$ iff

$$\gamma^2 \langle z, w \rangle_{H^r_{\mathrm{bd}}(\Omega)} = \langle z, w \rangle_{H^q_{\mathrm{bd}}(\Omega)}, \qquad \forall\, w \in C^\infty(\Omega). \tag{4.1.4}$$

Using (4.1.1) and (4.1.4), we see that $E^*Ez = \gamma^2 z$ iff

$$\langle [P_r - \gamma^{-2} P_q] z, w \rangle_{L_2(\Omega)} = \sum_{j=0}^{r-1} \langle \Gamma_{\gamma,j}(z), \partial_\nu^j w \rangle_{L_2(\partial\Omega)} \qquad \forall w \in C^\infty(\Omega).$$

Choosing w with compact support, we find that

$$\langle [P_r - \gamma^{-2} P_q] z, w \rangle_{L_2(\Omega)} = 0 \qquad \forall w \in C_0^\infty(\Omega).$$

Hence z and γ are weak solutions of the eigenproblem (4.1.2). Since Ω is smooth, we see that weak and smooth solutions of this eigenproblem coincide. Moreover, by appropriate choices of w, it follows that the boundary conditions (4.1.3) are satisfied. □

Note that (4.1.2)–(4.1.3) is a generalized eigenproblem, akin to the generalized eigenproblems $Az = \gamma Bz$ studied in computational linear algebra. It reduces to a standard elliptic eigenproblem whenever $q = 0$, i.e., whenever $G = L_2(\Omega)$.

We illustrate this construction by two one-dimensional cases.

Example ($q = 0$): Suppose that $d = 1$ and $\Omega = I = [0,1]$, with $q = 0$. Then the eigenfunctions z_j and eigenvalues γ_j of E^*E are solutions of the eigenproblem

$$\sum_{i=0}^{q}(-1)^i z^{(2i)}(s) = \gamma^{-2} y(s) \qquad \forall\, s \in [0,1],$$
$$\sum_{i=0}^{j}(-1)^i z^{(q-j+2i)}(0) = \sum_{i=0}^{j}(-1)^i z^{(q-j+2i)}(1) = 0 \qquad (0 \le j \le q-1).$$

When $q = 0$ and $r = 1$, we easily find that

$$z_j(s) = \begin{cases} 1 & \text{if } j = 1 \\ \sqrt{2}\cos\big((j-1)\pi s\big) & \text{if } j \ge 2 \end{cases}$$

and that

$$\gamma_j = \frac{1}{\sqrt{1 + \pi^2 (j-1)^2}}.$$

Unfortunately, the exact solution of the eigenproblem is unknown for arbitrary r. □

Example ($q = 1$, $r = 2$): Again suppose that $d = 1$ and $\Omega = I = [0,1]$, but now with $q = 1$ and $r = 2$. Then the eigenfunctions z_j and eigenvalues γ_j of E^*E are solutions of the generalized eigenproblem

$$z - z'' + z'''' = \gamma^2 (z - z''), \tag{4.1.5}$$

subject to the boundary conditions

$$\begin{gathered} z''(0) = z''(1) = 0, \\ z'''(0) + (\gamma^{-2} - 1) z'(0) = z'''(1) + (\gamma^{-2} - 1) z'(1) = 0. \end{gathered} \tag{4.1.6}$$

Using Mathematica, we quickly found that a basis for the solution space of (4.1.5) is given by $\{\psi_{\pm,\pm} = \exp(\pm x\sqrt{\beta \pm i\alpha})\}$, where $\alpha = \frac{1}{2}\sqrt{2\gamma^4 + 4\gamma^2 - 6}$ and $\beta = \frac{1}{2}\sqrt{2}(1 - \gamma^2)$, and so

$$z_j = C_1\psi_{+,+} + C_2\psi_{-,+} + C_3\psi_{+,-} + C_4\psi_{-,-}.$$

However, we were unsuccessful in finding values for the weights $C_1, \ldots, C_4$ and for γ such that the boundary conditions (4.1.6) hold. □

4.1.1. Results for the case $\Lambda = \Lambda^F$.

We now suppose that $\Lambda = \Lambda^F$. Using the solution to the eigenproblem in Lemma 4.1.1, we can now find nth optimal information in Λ^F, along with an nth minimal error algorithm. Moreover, we can get a tight bound on the nth minimal radius of information in Λ^F.

THEOREM 4.1.1.1. *Let z_j and γ_j be the jth eigenfunction and eigenvalue satisfying (4.1.2), subject to the boundary conditions (4.1.3). Let*

$$N_n f = [\langle Sf, z_1\rangle_{H^r_{\mathrm{bd}}(\Omega)}, \dots, \langle Sf, z_n\rangle_{H^r_{\mathrm{bd}}(\Omega)}] \qquad \forall f \in F$$

and

$$\phi_n(N_n f) = \sum_{j=1}^{n} \langle Sf, z_j\rangle_{H^r_{\mathrm{bd}}(\Omega)} z_j \qquad \forall f \in F.$$

Then N_n is nth optimal information in Λ^F and ϕ_n is an nth minimal error algorithm, with

$$e(\phi_n, N_n) = r(n, \Lambda^F) = \gamma_{n+1} = \Theta(n^{-(r-q)/d}).$$

PROOF: Using Lemma 4.1.1, along with Theorem 3.3, we immediately have

$$e(\phi_n, N_n) = r(n, \Lambda^F) = \gamma_{n+1} = d^n\left(E\big(H^r_{\mathrm{bd}}(\Omega)\big), H^q_{\mathrm{bd}}(\Omega)\right).$$

Following the proof of [28, Theorem 5.4.1], we see that

$$d^n\left(E\big(H^r_{\mathrm{bd}}(\Omega)\big), H^q_{\mathrm{bd}}(\Omega)\right) = \Theta(n^{-(r-q)/d}). \qquad \square$$

Thus the nth minimal radius is proportional to $n^{-(r-q)/d}$. We now consider the ε-complexity of our standard problem in the class Λ^F. Using Theorems 3.3 and 4.1.1.1, we have

THEOREM 4.1.1.2. *Let $\varepsilon > 0$. Then*

$$\mathrm{comp}(\varepsilon, \Lambda^F) = c \cdot \Theta\big((1/\varepsilon)^{d/(r-q)}\big).$$

Moreover, let N_n and ϕ_n be the information and algorithm given by Theorem 4.1.1.1. Then

$$c\, m(\varepsilon, \Lambda^F) \le \mathrm{comp}(\varepsilon, \Lambda^F) \le \mathrm{cost}(\phi_n, N_n) \le (c+2)m(\varepsilon, \Lambda^F) - 1.$$

Hence for $c \gg 1$, the algorithm ϕ_n using information N_n is a nearly optimal complexity algorithm. $\square$

Since we cannot usually find a closed form solution to the eigenproblem (4.1.2)–(4.1.3), we need to find other kinds of information that will be nearly as good. To do this, we will look at finite element information.

Let $\mathscr{S}_{n,k}$ be an n-dimensional *finite element subspace* of $H^q_{\mathrm{bd}}(\Omega)$, whose degree is k. That is, there is a triangulation $\mathscr{T}_n$ of Ω such that

$$\mathscr{S}_{n,k} = \left\{\, s \in H^q_{\mathrm{bd}}(\Omega) : s\big|_K \in P_k(K)\ \forall K \in \mathscr{T}_n \,\right\},$$
$$\dim \mathscr{S}_{n,k} = n.$$

(Here, $P_k(K)$ is the space of polynomials of degree at most k, considered as functions defined over the region K.) We will further require that $\{\mathscr{T}_n\}_{n=1}^{\infty}$ be a *quasi-uniform* family of triangulations of Ω. This means that

$$\limsup_{n\to\infty} \sup_{K\in\mathscr{T}_n} \frac{h_K}{\rho_K} < \infty,$$

where

$$h_K = \operatorname{diam} K$$

and

$$\rho_K = \sup\{\, \operatorname{diam} B : \text{spheres } B \text{ containing } K \,\}$$

for any $n \in \mathbb{N}$ and any $K \in \mathscr{T}_n$. (See [3], [7, Chapter XII], [16], or [28, Chapter 5 and Appendix] for further properties of finite element spaces.)

We can now define our information. For any $n \in \mathbb{N}$, let

$$N_{n,k}f = [\langle Sf, g_1\rangle_{H^q_{\mathrm{bd}}(\Omega)}, \dots, \langle Sf, g_n\rangle_{H^q_{\mathrm{bd}}(\Omega)}] \qquad \forall f \in F, \tag{4.1.1.1}$$

where $\{g_1, \dots, g_n\}$ is a basis for $\mathscr{S}_{n,k}$. Of course, $N_{n,k} \in \Lambda^F_n$. We say that $N_{n,k}$ is *finite element information* (FEI).

Now consider the following algorithm: For $f \in F$, let $u_{n,k} \in \mathscr{S}_{n,k}$ satisfy

$$\langle u_{n,k}, g_i\rangle_{H^q_{\mathrm{bd}}(\Omega)} = \langle Sf, g_i\rangle_{H^q_{\mathrm{bd}}(\Omega)}. \qquad (1 \le i \le n) \tag{4.1.1.2}$$

Since $u_{n,k}$ depends on f only through the information $N_{n,k}f$, we may write

$$u_{n,k} = \phi_{n,k}(N_{n,k}f).$$

Clearly, $\phi_{n,k}$ is a Galerkin method using test and trial space $\mathscr{S}_{n,k}$. Note that we can determine the coefficients $a = [\alpha_1, \dots, \alpha_n]$ of $u_{n,k}$ (with respect to the basis $g_1, \dots, g_n$) by solving a linear system $Ka = b$, where $\kappa_{j,j} = \langle g_j, g_i\rangle_{H^q_{\mathrm{bd}}(\Omega)}$ and $\beta_i = \langle Sf, g_j\rangle_{H^q_{\mathrm{bd}}(\Omega)}$.

Remark: Note that we refer to $\phi_{n,k}$ as a Galerkin method using the FEI $N_{n,k}$. However, $\phi_{n,k}$ is not a "finite element method" in the usual sense of the term, since the information used is of the form $\langle Sf, g_i\rangle_{H^q_{\mathrm{bd}}(\Omega)}$ instead of the form $\lambda_i(f)$ used by the finite element method. $\square$

Our main error estimate for the Galerkin method using FEI is

Theorem 4.1.1.3. *Let $k \ge r-1$. Then for any n, we have*

$$e(\phi_{n,k}, N_{n,k}) = \Theta(n^{-(r-q)/d}) = \Theta\big(r(n, \Lambda^F)\big).$$

Thus if $k \ge r-1$, then the information $N_{n,k}$ is nth nearly optimal information in Λ^F, and the algorithm $\phi_{n,k}$ is an nth nearly minimal error algorithm.

PROOF: Let $f \in F$. Then $\phi_{n,k}(N_{n,k}f)$ is the $H^q_{\mathrm{bd}}(\Omega)$-projection of Sf onto $\mathscr{S}_{n,k}$. By [28, Lemma 5.4.3], there is a positive constant C, independent of n and f, such that

$$\|Sf - \phi_{n,k}(N_{n,k}f)\|_{H^q_{\mathrm{bd}}(\Omega)} = \inf_{s \in \mathscr{S}_{n,k}} \|Sf - s\|_{H^q_{\mathrm{bd}}(\Omega)} \le Cn^{-(\min\{k+1,r\}-q)/d} \|Sf\|_{H^r_{\mathrm{bd}}(\Omega)}.$$

Since $k+1 \ge r$, we thus find that

$$e(\phi_{n,k}, N_{n,k}) = O(n^{-(r-q)/d}).$$

Using Theorem 4.1.1.1, we have

$$e(\phi_{n,k}, N_{n,k}) \ge r(n, \Lambda^F) = \Theta(n^{-(r-q)/d}),$$

completing the proof of the theorem. □

Remark: Note that we do not claim that $\phi_{n,k}$ is an optimal error algorithm using $N_{n,k}$. If we use the prescription of Lemma 3.1, we would construct an optimal error algorithm $\phi_{N_{n,k}}$ using $N_{n,k}$ as follows. For $f \in F$, let $u_{N_{n,k}} \in E^*\mathscr{S}_{n,k}$ satisfy

$$\langle u_{N_{n,k}}, E^*g_i\rangle_{H^r_{\mathrm{bd}}(\Omega)} = \langle Sf, E^*g_i\rangle_{H^r_{\mathrm{bd}}(\Omega)} \qquad (1 \le i \le n).$$

Since $u_{N_{n,k}}$ depends on f only through the information $N_{n,k}f$, we may write $u_{N_{n,k}} = \phi_{N_{n,k}}(N_{n,k}f)$. Then $\phi_{N_{n,k}}$ is an optimal error algorithm using $N_{n,k}$. □

Remark: It is possible to describe somewhat different finite element information such that the resulting Galerkin method (using the same space of test and trial functions) is an optimal error algorithm using this new FEI. To do this, let $\hat{\mathscr{S}}_{n,k}$ be an n-dimensional finite element subspace $H^r_{\mathrm{bd}}(\Omega)$, whose degree is k, with the spaces $\hat{\mathscr{S}}_{n,k}$ being based on a quasi-uniform family of triangulations of Ω. Then our FEI has the form

$$\hat{N}_{n,k}f = [\langle Sf, w_1\rangle_{H^r_{\mathrm{bd}}(\Omega)}, \ldots, \langle Sf, w_n\rangle_{H^r_{\mathrm{bd}}(\Omega)}] \qquad \forall f \in F,$$

where $\{w_1, \ldots, w_n\}$ is a basis for $\hat{\mathscr{S}}_{n,k}$.

Now consider the following algorithm: For $f \in F$, let $\hat{u}_{n,k} \in \hat{\mathscr{S}}_{n,k}$ satisfy

$$\langle \hat{u}_{n,k}, w_i\rangle_{H^r_{\mathrm{bd}}(\Omega)} = \langle Sf, w_i\rangle_{H^r_{\mathrm{bd}}(\Omega)}.$$

Since $\hat{u}_{n,k}$ depends on f only through the information $N_{n,k}f$, we may write

$$\hat{u}_{n,k} = \hat{\phi}_{n,k}(\hat{N}_{n,k}f).$$

Clearly, $\hat{\phi}_{n,k}$ is a Galerkin method using test and trial space $\hat{\mathscr{S}}_{n,k}$. Moreover, it is possible to show that

$$e(\hat{\phi}_{n,k}, \hat{N}_{n,k}) = r(\hat{N}_{n,k}) = \Theta(n^{-(r-q)/d}).$$

(The proof is slightly more involved than that of Theorem 4.1.1.3, requiring the use of the duality estimate in [3, Theorem 2.3.1].) Note that since $\hat{\mathscr{S}}_{n,k}$ is a subspace of $H^r_{\mathrm{bd}}(\Omega)$, we automatically have $k \geq r-1$ (see [28, Lemma 5.4.2]).

It might seem that since $\hat{\phi}_{n,k}$ is an optimal error algorithm, one would prefer using the algorithm $\hat{\phi}_{n,k}$ instead of the algorithm $\phi_{n,k}$. However, note that if we use $\hat{\phi}_{n,k}$, then we need to construct a finite element space $\hat{\mathscr{S}}_{n,k}$ that is a subspace of $H^r_{\mathrm{bd}}(\Omega)$. In practical situations, this would require more initial precomputation than constructing a finite element space $\mathscr{S}_{n,k}$ that is a subspace of $H^q_{\mathrm{bd}}(\Omega)$. Moreover, the errors of $\phi_{n,k}$ and $\hat{\phi}_{n,k}$ are roughly the same. Hence any gain that might be realized in using $\hat{\phi}_{n,k}$ will be offset by the loss involved in the additional precomputation. For this reason, we prefer to use $\phi_{n,k}$. □

We now determine the cost of using our Galerkin method to compute ε-approximations for our standard problem. Let us denote this cost by

$$\mathrm{cost}^{\mathrm{Gal}}(\varepsilon, \Lambda^F) = \min\{\, \mathrm{cost}(\phi_{n,k}, N_{n,k}) : e(\phi_{n,k}, N_{n,k}) \leq \varepsilon\,\}.$$

We then have

THEOREM 4.1.1.4. *Let $\varepsilon > 0$. Then*

$$\mathrm{cost}^{\mathrm{Gal}}(\varepsilon, \Lambda^F) = c \cdot \Theta\big((1/\varepsilon)^{d/(r-q)}\big).$$

Hence the Galerkin method using FEI is a nearly optimal complexity algorithm when $\Lambda = \Lambda^F$.

PROOF: From Theorem 4.1.1.3, we see that $e(\phi_{n,k}, N_{n,k}) \leq \varepsilon$ holds if and only if we choose $n = \Theta\big((1/\varepsilon)^{d/(r-q)}\big)$. Now use Theorem 4.1.1.2. □

4.1.2. Results for the case $\Lambda = \Lambda^*$.

Now, we suppose that $\Lambda = \Lambda^*$, i.e., continuous linear functionals are permissible information. We can find optimal information, minimal error algorithms, and optimal complexity algorithms for the case $\Lambda = \Lambda^*$ by using the results for the case $\Lambda = \Lambda^F$, along with Theorem 3.4.

We first look at information based on the eigenvectors of E^*E.

THEOREM 4.1.2.1. *Let z_j and γ_j be the jth eigenfunction and eigenvalue satisfying (4.1.2), subject to the boundary conditions (4.1.3). For any $n \in \mathbb{N}$ and any $\delta > 0$, choose $\lambda_{1,\delta}, \ldots, \lambda_{n,\delta} \in X^*$ such that*

$$\|z_j - L^*\lambda_{j,\delta}\|_{H^q_{\mathrm{bd}}(\Omega)} \leq \frac{\delta\gamma_{n+1}}{n}.$$

Let

$$N_{n,\delta} f = [\lambda_{1,\delta}(f), \ldots, \lambda_{n,\delta}(f)] \qquad \forall f \in F,$$

and

$$\phi_{n,\delta}(N_{n,\delta} f) = \sum_{j=1}^{n} \lambda_{j,\delta}(f) z_j \qquad \forall f \in F.$$

Then

$$e(\phi_{n,\delta}, N_{n,\delta}) \le (1+\delta)\gamma_{n+1} = \Theta(n^{-(r-q)/d})$$

and

$$r(n, \Lambda^*) = r(n, \Lambda^F) = \gamma_{n+1} = \Theta(n^{-(r-q)/d}).$$

Hence $N_{n,\delta}$ is nth nearly optimal information in Λ^* and $\phi_{n,\delta}$ is an nth nearly minimal error algorithm.

PROOF: Immediate from Theorems 3.4 and 4.1.1.1. □

Having determined nearly optimal information and nearly minimal algorithms for the case $\Lambda = \Lambda^*$, we can now determine the problem complexity and find nearly optimal complexity algorithms.

THEOREM 4.1.2.2.

(1) *For any $\varepsilon > 0$, we have*

$$\mathrm{comp}(\varepsilon, \Lambda^*) = c \cdot \Omega\big((1/\varepsilon)^{d/(r-q)}\big).$$

(2) *Suppose that the eigenvalues γ_n are strictly monotonically decreasing in n. Then for any $\varepsilon > 0$, we have*

$$\mathrm{comp}(\varepsilon, \Lambda^*) = c \cdot \Theta\big((1/\varepsilon)^{d/(r-q)}\big).$$

Let $N_{n,\delta} \in \Lambda_n^*$ be the information given by (3.15) and let $\phi_{n,\delta}$ be the algorithm given by (3.16), with $n = m(\varepsilon, \Lambda^*) + 1$ and δ satisfying

$$0 < \delta < \frac{\gamma_n}{\gamma_{n+1}} - 1 \sim \left(\frac{n+1}{n}\right)^{(r-q)/d} - 1.$$

Then

$$c\, m(\varepsilon, \Lambda^*) \le \mathrm{comp}(\varepsilon, \Lambda^*) \le \mathrm{cost}(\phi_{n,\delta}, N_{n,\delta}) \le (c+2)\big(m(\varepsilon, \Lambda^*) + 1\big) - 1.$$

Hence for $c \gg 1$, the algorithm $\phi_{n,\delta}$ using information $N_{n,\delta}$ is a nearly optimal complexity algorithm. □

PROOF: This follows immediately from Theorems 3.4 and 4.1.2.1. □

Remark: Note that the assumption that the γ_n are strictly monotonically decreasing in n generally will not hold unless $d = 1$. For example, suppose that $q = 0$ and $r = 0$; then the γ_n are eigenvalues of the operator $(1-\Delta)^{-1}$, where Δ is the Laplacian. This operator has multiple eigenvalues, unless $d = 1$, and so Theorem 4.1.2.2 cannot be immediately applied. However, it is possible to modify this theorem to account for the case of multiple eigenvalues. Since we will not need this modified version of this theorem in what follows, we do not feel that it is important to state the modified theorem. □

As in the previous section, the optimal information and algorithm of Theorem 4.1.2.2 are usually not available in closed form. Hence, we need to look at alternative nearly optimal information and algorithms. In particular, we now consider modified Galerkin methods for computing ε-approximations to our problem. As in the previous section, we let $\mathscr{S}_{n,k}$ denote an n-dimensional finite element subspace of $H^q_{\mathrm{bd}}(\Omega)$, whose degree is k, with the corresponding family $\{\mathscr{T}_n\}_{n=1}^{\infty}$ of triangulations being quasi-uniform. We let $\{g_1,\ldots,g_n\}$ again be a basis for $\mathscr{S}_{n,k}$. However, we require that this basis satisfy the condition that there exist a positive constant C, independent of n, such that

$$\left[\sum_{j=1}^{n}\alpha_j^2\right]^{1/2} \le Cn^{1/2}\left\|\sum_{j=1}^{n}\alpha_j g_j\right\|_{L_2(\Omega)} \qquad \forall\,\alpha_1,\ldots,\alpha_n\in\mathbb{R}. \tag{4.1.2.1}$$

This condition is satisfied by the usual finite element basis functions having small support; see [5] and [28, pp. 216 ff.] for details. Since the range of L^* is dense in X^*, there exist $\lambda_1,\ldots,\lambda_n\in X^*$ such that

$$\|g_i - L^*\lambda_i\|_{H^q_{\mathrm{bd}}(\Omega)} \le Cn^{-((r-q)/d+1)}. \tag{4.1.2.2}$$

We can now define our information and algorithm. For any $n\in\mathbb{N}$, let

$$\tilde{N}_{n,k} = [\lambda_1(f),\ldots,\lambda_n(f)] \qquad \forall f\in F.$$

Note that $\tilde{N}_{n,k}\in\Lambda^*_n$ is an approximation of the finite element information $N_{n,k}\in\Lambda^F_n$ defined in the previous section, and so we refer to $N_{n,k}$ as *modified finite element information*, or modified FEI, for short. Then for $f\in F$, we seek $\tilde{u}_{n,k}\in\mathscr{S}_{n,k}$ satisfying

$$\langle\tilde{u}_{n,k}, g_i\rangle_{H^q_{\mathrm{bd}}(\Omega)} = \lambda_i(f). \qquad (1\le i\le n) \tag{4.1.2.3}$$

Since $\tilde{u}_{n,k}$ depends on f only through the information $\tilde{N}_{n,k}f$, we may write

$$\tilde{u}_{n,k} = \tilde{\phi}_{n,k}(\tilde{N}_{n,k}f).$$

We call $\tilde{\phi}_{n,k}$ an *modified Galerkin method* using modified FEI.

Remark: It is easy to see that the Galerkin and modified Galerkin methods may be reduced to the solution of $n\times n$ linear systems whose solution gives the (respective) coefficients of $u_{n,k}$ or $\tilde{u}_{n,k}$ with respect to the basis $\{g_1,\ldots,g_n\}$. The only difference between the Galerkin and modified Galerkin methods is that the latter method uses $\lambda_i(f)$ to approximate the $\langle Sf, g_i\rangle_{H^q_{\mathrm{bd}}(\Omega)}$ appearing in the definition of the former method. However, the same coefficient matrix is used for both algorithms. □

Our main error estimate is

THEOREM 4.1.2.3. *Let $k\ge r-1$. Then for any $n\in\mathbb{N}$, we have*

$$e(\tilde{\phi}_{n,k},\tilde{N}_{n,k}) = \Theta(n^{-(r-q)/d}) = \Theta\big(r(n,\Lambda^*)\big).$$

Thus if $k\ge r-1$, then the information $\tilde{N}_{n,k}$ is nth nearly optimal information in Λ^*, and the algorithm $\tilde{\phi}_{n,k}$ is an nth nearly minimal error algorithm.

PROOF: Let $n \in \mathbb{N}$. For $f \in F$, let $u_{n,k} = \phi_{n,k}(N_{n,k}f)$ and let $\tilde{u}_{n,k} = \tilde{\phi}_{n,k}(\tilde{N}_{n,k}f)$. Writing $e_{n,k} = u_{n,k} - \tilde{u}_{n,k} = \sum_{j=1}^{n} \varepsilon_j g_j$, we may use the discrete Cauchy-Schwarz inequality to see that

$$\|e_{n,k}\|^2_{H^q_{\mathrm{bd}}(\Omega)} = \sum_{j=1}^{n} \varepsilon_j \langle e_{n,k}, g_j \rangle_{H^q_{\mathrm{bd}}(\Omega)} \le \left[\sum_{j=1}^{n} \varepsilon_j^2\right]^{1/2} \left[\sum_{j=1}^{n} \langle e_{n,k}, g_j \rangle^2_{H^q_{\mathrm{bd}}(\Omega)}\right]^{1/2}. \tag{4.1.2.4}$$

By (4.1.2.1), there is a positive constant C such that

$$\left[\sum_{j=1}^{n} \varepsilon_j^2\right]^{1/2} \le C n^{1/2} \|e_{n,k}\|_{L_2(\Omega)} \le C n^{1/2} \|e_{n,k}\|_{H^q_{\mathrm{bd}}(\Omega)}. \tag{4.1.2.5}$$

Moreover, we may use (4.1.1.2), (4.1.2.3), and (4.1.2.2) to see that

$$\begin{aligned} |\langle e_{n,k}, g_j \rangle_{H^q_{\mathrm{bd}}(\Omega)}| &= |\langle Sf, g_j \rangle_{H^q_{\mathrm{bd}}(\Omega)} - \lambda_j(f)| = |\langle Sf, g_j - L^* \lambda_j \rangle_{H^q_{\mathrm{bd}}(\Omega)}| \\ &\le \|Sf\|_{H^q_{\mathrm{bd}}(\Omega)} \|g_j - L^*\lambda_j\|_{H^q_{\mathrm{bd}}(\Omega)} \le C n^{-((r-q)/d+1)}. \end{aligned} \tag{4.1.2.6}$$

Using (4.1.2.5) and (4.1.2.6) in (4.1.2.4), we find that

$$\|e_{n,k}\|_{H^q_{\mathrm{bd}}(\Omega)} \le C n^{-(r-q)/d}$$

for some positive constant C. Hence

$$\|Sf - \tilde{\phi}_{n,k}(\tilde{N}_{n,k}f)\|_{H^q(\Omega)} \le \|Sf - \phi_{n,k}(N_{n,k}f)\|_{H^q_{\mathrm{bd}}(\Omega)} + \|e_{n,k}\|_{H^q_{\mathrm{bd}}(\Omega)} \le C n^{-(r-q)/d}$$

for a positive constant C. Since $f \in F$ is arbitrary, we find that

$$e(\tilde{\phi}_{n,k}, \tilde{N}_{n,k}) = O(n^{-(r-q)/d}).$$

Using Theorem 4.1.2.1, we have

$$e(\phi_{n,k}, N_{n,k}) \ge r(n, \Lambda^*) = \Theta(n^{-(r-q)/d}),$$

completing the proof of the theorem. □

We now determine the ε-complexity of the standard problem in the class Λ^*, as well as the cost of using our modified Galerkin method to compute ε-approximations. Let us denote this cost by

$$\mathrm{cost}^{\text{mod-Gal}}(\varepsilon, \Lambda^F) = \min\{\, \mathrm{cost}(\tilde{\phi}_{n,k}, \tilde{N}_{n,k}) : e(\tilde{\phi}_{n,k}, \tilde{N}_{n,k}) \le \varepsilon \,\}.$$

We then have

THEOREM 4.1.2.4. *Let $\varepsilon > 0$. Then*

$$\mathrm{comp}(\varepsilon, \Lambda^*) = c \cdot \Theta\big((1/\varepsilon)^{d/(r-q)}\big)$$

and

$$\mathrm{cost}^{\text{mod-Gal}}(\varepsilon, \Lambda^*) = c \cdot \Theta\big((1/\varepsilon)^{d/(r-q)}\big).$$

Hence the modified Galerkin method using modified FEI is a nearly optimal complexity algorithm when $\Lambda = \Lambda^*$.

PROOF: From Theorem 4.1.2.3, we see that $e(\tilde{\phi}_{n,k}, \tilde{N}_{n,k}) \le \varepsilon$ holds iff $n = \Theta\big((1/\varepsilon)^{d/(r-q)}\big)$. Thus

$$\mathrm{comp}(\varepsilon, \Lambda^*) = c \cdot O\big((1/\varepsilon)^{d/(r-q)}\big)$$

Now use Theorem 4.1.2.2. □

4.2. Elliptic boundary-value problems.

We now consider the complexity of solution-restricted elliptic boundary-value problems. For the sake of brevity, we provide neither definitions of standard vocabulary nor lists of standard results in the study of elliptic problems. The interested reader should consult any standard reference (such as [1], [3], [7, Chapter V], [9], or [16]) for more details. Previous results on the complexity of elliptic problems may be be found in [28, Chapter 5], and the references found therein. To simplify the exposition, we will follow the approach taken in [25]. However, more complicated problems could have been handled as well.

Let $\Omega \subset \mathbb{R}^d$ be a smooth, bounded, simply-connected region. Recall that $H_0^m(\Omega)$ is the space of all $H^m(\Omega)$-functions whose normal derivatives of order less than m vanish at the boundary of Ω. Let

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

be a symmetric $H_0^m(\Omega)$-coercive bilinear form. Then B is an inner product on $H_0^m(\Omega)$, yielding an *energy norm* $\|\cdot\|_B$ defined by

$$\|v\|_B = \sqrt{B(v,v)} \qquad \forall\, v \in H_0^m(\Omega)$$

that is equivalent to the usual norm $\|\cdot\|$ on $H_0^m(\Omega)$. Note that if we let

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

then L is a uniformly strongly elliptic operator of order $2m$. Moreover, we have

$$B(v, w) = \langle Lv, w \rangle_{L_2(\Omega)} \qquad \forall\, v, w \in H_0^m(\Omega),$$

where (as usual) the $L_2(\Omega)$-inner product may be interpreted as the duality pairing of $H_0^m(\Omega)$ and its dual space $H^{-m}(\Omega)$.

We now show how our problem may be expressed in terms of the general framework that we have developed. Let $r \ge m$, and let $W = H^r_{\text{bd}}(\Omega) = \{\, v \in H^r(\Omega) : v \in H^m_0(\Omega) \,\}$. We then take $G = H^m_0(\Omega)$ and $X = H^{-m}(\Omega)$. The class F of problem elements is now the unit ball of $H^r(\Omega)$. Note that the Lax-Milgram lemma implies that the range of L is $H^{-m}(\Omega)$. So $L\colon H^m_0(\Omega) \to H^{-m}(\Omega)$ is a bijection. Hence $S = L^{-1}\colon H^{-m}(\Omega) \to H^m_0(\Omega)$.

Hence our solution operator $S\colon H^{-m}(\Omega) \to H^m_0(\Omega)$ is given as

$$B(Sf, v) = \langle f, v \rangle_{L_2(\Omega)} \qquad \forall f \in H^{-m}(\Omega), v \in H^m_0(\Omega). \tag{4.2.1}$$

Again, the Lax-Milgram lemma tells us that S is well-defined. For any $f \in H^{-m}(\Omega)$, we have that $u = Sf$ is the variational solution of

$$\begin{aligned} Lu &= f \quad \text{in } \Omega \\ \partial^j_\nu u &= 0 \quad \text{on } \partial\Omega \qquad (0 \le j \le m-1), \end{aligned}$$

a $2m$th-order *elliptic boundary problem* satisfying homogeneous Dirichlet boundary conditions.

Note that our problem (4.2.1) is a standard problem (with $q = m$), and so we can apply the results in Section 4.1. We first look at the case $\Lambda = \Lambda^F$.

For any $n \in \mathbb{N}$, let $\mathscr{S}_{n,k}$ be an n-dimensional finite element subspace of $H^m_0(\Omega)$ having degree k. Let $\{g_1, \ldots, g_n\}$ be a basis for $\mathscr{S}_{n,k}$. Define finite element information $N_{n,k}$ by

$$N_{n,k}f = [\langle f, g_1 \rangle_{L_2(\Omega)}, \ldots, \langle f, g_n \rangle_{L_2(\Omega)}].$$

For $f \in F$, let $u_{n,k} = \phi_{n,k}(N_{n,k}f)$ be the Galerkin method given by

$$B(u_{n,k}, g_i) = \langle f, g_i \rangle_{L_2(\Omega)} \qquad (1 \le i \le n). \tag{4.2.2}$$

Note that the Galerkin algorithm $\phi_{n,k}$ is the standard finite element method (FEM).

Theorem 4.2.1. *Let $\Lambda = \Lambda^F$.*

(1) *For any $n \in \mathbb{N}$, we have*

$$r(n, \Lambda^F) = \Theta(n^{-(r-m)/d}).$$

(2) *Let $k \ge r-1$. For any $n \in \mathbb{N}$, we have*

$$e(\phi_{n,k}, N_{n,k}) = \Theta(n^{-(r-m)/d}),$$

so that $\phi_{n,k}$ is an nth nearly minimal error algorithm, and $N_{n,k}$ is nth nearly optimal information.

(3) *For any $\varepsilon > 0$, we have*

$$\operatorname{comp}(\varepsilon, \Lambda^F) = c \cdot \Theta\big((1/\varepsilon)^{d/(r-m)}\big).$$

(4) *Let $k \ge r-1$. For any $\varepsilon > 0$, let $n = \Theta\big((1/\varepsilon)^{d/(r-m)}\big)$. Then*

$$\operatorname{cost}^{\text{Gal}}(\varepsilon, \Lambda^F) = c \cdot \Theta\big((1/\varepsilon)^{d/(r-m)}\big),$$

and so the FEM $\phi_{n,k}$ using FEI $N_{n,k}$ is a nearly optimal complexity algorithm.

PROOF: Recall that $B(\cdot,\cdot)$ is the inner product on $H_0^m(\Omega)$. But $B(Sf, g_i) = \langle f, g_i\rangle_{L_2(\Omega)}$ by (4.2.1). Hence $N_{n,k}$ and $\phi_{n,k}$ as defined in the statement of the theorem are the same as in (4.1.1.1) and (4.1.1.2). The result now follows from Theorems 4.1.1.3 and 4.1.1.4. □

So we see that the standard finite element method of degree $k \geq r-1$ is nearly optimal for a $2m$th-order elliptic problem, if the solution elements are constrained to lie in the unit ball of $H^r(\Omega)$. This result should not be too surprising, given the known results about optimality of FEMs for elliptic problems, see [28, Chapter 5] and the results cited therein. The novelty in this result lies in the fact that we did not need to use a shift theorem (i.e., a result saying that if f has $r-2m$ derivatives, then Sf has r derivatives) to prove the optimality of the FEM.

We now look at the case $\Lambda = \Lambda^*$, mainly for the sake of completeness. It will turn out that FEI for this problem (which is a priori only Λ^F-information) is really continuous linear information.

Rather than use the standard $H^{-m}(\Omega)$ inner product on $H^{-m}(\Omega)$, it will be more convenient to consider $H^{-m}(\Omega)$ as the dual of $H_0^m(\Omega)$ under the energy norm, i.e., we use the norm

$$\|v\|_{B'} = \sup_{w\in H_0^m(\Omega)} \frac{\langle v, w\rangle_{L_2(\Omega)}}{\|w\|_B} \qquad \forall\, v \in H^{-m}(\Omega). \tag{4.2.3}$$

We let $\langle\cdot,\cdot\rangle_{B'}$ denote the corresponding inner product. Since the energy norm $\|\cdot\|_B$ is equivalent to the usual $H_0^m(\Omega)$-norm, it follows that this norm $\|\cdot\|_{B'}$ is equivalent to the usual $H^{-m}(\Omega)$-norm.

LEMMA 4.2.1.

(1) *For any $v, w \in H^{-m}(\Omega)$, we have*

$$\langle v, w\rangle_{B'} = B(Sv, Sw).$$

(2) $L^* = S$.

PROOF: From (4.2.1) and (4.2.3), we find

$$\|v\|_{B'} = \sup_{v\in H_0^m(\Omega)} \frac{\langle v, w\rangle_{L_2(\Omega)}}{\|w\|_B} = \sup_{v\in H_0^m(\Omega)} \frac{B(Sv, w)}{\|w\|_B} = \|Sv\|_B,$$

the last holding because $B(\cdot,\cdot)$ is the inner product corresponding to the norm $\|\cdot\|_B$. Part (1) now follows immediately. To prove part (2), note that we must have

$$\langle Lv, w\rangle_{B'} = B(v, L^*w) \qquad \forall\, v \in H_0^m(\Omega), w \in H^{-m}(\Omega).$$

But

$$\langle Lv, w\rangle_{B'} = B(SLv, Sw) = B(v, Sw) \qquad \forall\, v \in H_0^m(\Omega), w \in H^{-m}(\Omega).$$

Comparing these last two, we see that $L^* = S$. □

Now suppose that $N_{n,k}$ is FEI as in the statement of Theorem 4.2.1. If for $1 \le i \le n$, we let $v_i = Lg_i$, then we have $L^* v_i = Sv_i = g_i$ by Lemma 4.2.1. Thus we find that

$$\langle f, g_i\rangle_{L_2(\Omega)} = \langle f, v_i\rangle_{B'} \qquad (1 \le i \le n).$$

We now see that

$$N_{n,k} f = [\langle f, v_1\rangle_{B'}, \ldots, \langle f, v_n\rangle_{B'}].$$

For $f \in F$, it now follows that $u_{n,k} = \phi_{n,k}(N_{n,k} f)$ satisfies

$$B(u_{n,k}, g_i) = \langle f, v_i\rangle_{B'} \qquad (1 \le i \le n). \tag{4.2.4}$$

So the information $N_{n,k}$ defined in Theorem 4.2.1 is continuous linear information. Theorem 4.2.1 now tells us that the FEM is a nearly optimal error algorithm and a nearly optimal complexity algorithm using continuous linear information. Of course, Theorems 4.1.1.3 and 4.1.1.4 tell us that $r(n, \Lambda^*) = r(n, \Lambda^F) = \Theta(n^{-(r-m)/d})$ and $\mathrm{comp}(\varepsilon, \Lambda^*) = c \cdot \Theta\big((1/\varepsilon)^{d/(r-m)}\big)$.

Hence (4.2.2) and (4.2.4) are two formulations of the FEM algorithm using FEI. The first formulation shows that this FEI is information from Λ^F, whereas the second shows that it actually is information from Λ^*. However, the B'-inner product is more complicated than the $L_2(\Omega)$-inner product. So in practice, we would probably rather use (4.2.2) than (4.2.4).

4.3. Fredholm problems of the second kind.

We now look at the complexity of solution-restricted Fredholm integral equations of the second kind. For previous work on the complexity of Fredholm problems of the second kind, see [24] and [28, Section 6.3].

Let $\Omega \subset \mathbb{R}^d$ be a smooth, bounded, simply-connected region. Suppose that $k\colon \Omega\times\Omega \to \mathbb{R}$ satisfies

$$\int_\Omega \int_\Omega |k(x, y)|^2\, dx\, dy < \infty.$$

Then the operator $K\colon L_2(\Omega) \to L_2(\Omega)$, defined as

$$(Kv)(x) = \int_\Omega k(x, y)v(y)\, dy \qquad \forall\, v \in L_2(\Omega),$$

is an integral operator with a Hilbert-Schmidt kernel, and is thus compact. Assume that 1 is not an eigenvalue of $(K^*K)^{1/2}$. For $f\colon \Omega \to \mathbb{R}$, we wish to find (approximations to) $u\colon \Omega \to \mathbb{R}$ satisfying

$$u(x) - \int_\Omega k(x, y)u(y)\, dy = f(x) \qquad \forall\, x \in \Omega, \tag{4.3.1}$$

a *Fredholm integral equation* of the second kind.

We formally describe our problem by taking $X = G = L_2(\Omega)$ and $W = H^r(\Omega)$ for some $r \ge 0$. Hence the class F of problem elements is the unit ball of $H^r(\Omega)$. Let $L = I - K$. Then L is a bounded bijection of G onto X, and so $S = L^{-1}$ is a bounded bijection of X

onto G. Note that our problem (4.3.1) is a standard problem (with $q = 0$), and so we can apply the results in Section 4.1.

We first look at the case $\Lambda = \Lambda^F$. For any $n \in \mathbb{N}$, let $\mathscr{S}_{n,k}$ be an n-dimensional finite element subspace of $L_2(\Omega)$ having degree k. Let $\{g_1, \ldots, g_n\}$ be a basis for $\mathscr{S}_{n,k}$. Define finite element information $N_{n,k}$ by

$$N_{n,k}f = [\langle Sf, g_1\rangle_{L_2(\Omega)}, \ldots, \langle Sf, g_n\rangle_{L_2(\Omega)}].$$

For $f \in F$, let $u_{n,k} = \phi_{n,k}(N_{n,k}f)$ be the Galerkin method given by

$$\langle u_{n,k}, g_i\rangle_{L_2(\Omega)} = \langle Sf, g_i\rangle_{L_2(\Omega)} \qquad (1 \le i \le n). \tag{4.3.2}$$

We then have

THEOREM 4.3.1. *Let $\Lambda = \Lambda^F$.*

(1) *For any $n \in \mathbb{N}$, we have*

$$r(n, \Lambda^F) = \Theta(n^{-r/d}).$$

(2) *Let $k \ge r - 1$. For any $n \in \mathbb{N}$, we have*

$$e(\phi_{n,k}, N_{n,k}) = \Theta(n^{-r/d}),$$

so that $\phi_{n,k}$ is an nth nearly minimal error algorithm, and $N_{n,k}$ is nth nearly optimal information.

(3) *For any $\varepsilon > 0$, we have*

$$\mathrm{comp}(\varepsilon, \Lambda^F) = c \cdot \Theta\big((1/\varepsilon)^{d/r}\big).$$

(4) *Let $k \ge r - 1$. For any $\varepsilon > 0$, let $n = \Theta\big((1/\varepsilon)^{d/r}\big)$. Then*

$$\mathrm{cost}^{\mathrm{Gal}}(\varepsilon, \Lambda^F) = c \cdot \Theta\big((1/\varepsilon)^{d/r}\big),$$

and so the Galerkin method $\phi_{n,k}$ using FEI $N_{n,k}$ is a nearly optimal complexity algorithm for the Fredholm problem of the second kind.

PROOF: Immediate from Theorems 4.1.1.3 and 4.1.1.4. □

Remark: The algorithm $\phi_{n,k}$ defined by (4.3.2) is not a finite element method, a marked contrast with what happened when we were looking at elliptic boundary-value problems. The only possibility we would have for using the approach in the previous section, in which the Galerkin method $\phi_{n,k}$ turned out to be an FEM, would be to assume that K is symmetric and $\|K\| < 1$. If this were the case, then $B(v, w) = \langle Lv, w\rangle_{L_2(\Omega)}$ would be an inner product on $L_2(\Omega)$, with the norm $\|\cdot\|_B = \sqrt{B(\cdot,\cdot)}$ being equivalent to the usual $L_2(\Omega)$-norm. Under these hypotheses, it then turns out that the Galerkin method (4.1.1.2) is the standard FEM with test and trial spaces $\mathscr{S}_{n,k}$. However, we often need to solve our problem $(I-K)u=f$ for non-symmetric K or $\|K\| \not< 1$, so this approach is not generally applicable. $\square$

Note that linear functionals $\langle Sf, g_i\rangle_{L_2(\Omega)}$ are hard to directly evaluate, because of their dependence on Sf. However, since $S\colon X \to G$ is a bounded bijection, we can write

$$\langle Sf, g_i\rangle_{L_2(\Omega)} = \langle f, S^* g_i\rangle_{L_2(\Omega)},$$

where $S^* = (L^*)^{-1} = (I-K^*)^{-1}$. In principle, we can then consider the computation of $S^*g_1, \dots, S^*g_n$ as precomputation. We now see that

$$N_{n,k}f = [\langle f, S^*g_1\rangle_{L_2(\Omega)}, \dots, \langle f, S^*g_n\rangle_{L_2(\Omega)}].$$

For $f \in F$, it now follows that $u_{n,k} = \phi_{n,k}(N_{n,k}f)$ satisfies

$$\langle u_{n,k}, g_i\rangle_{L_2(\Omega)} = \langle f, S^*g_i\rangle_{L_2(\Omega)} \qquad (1 \le i \le n). \tag{4.3.3}$$

So the information $N_{n,k}$ defined in Theorem 4.3.1 is continuous linear information. Moreover, the Galerkin algorithm (4.3.3) is a nearly optimal error algorithm and a nearly optimal complexity algorithm using continuous linear information. Of course, Theorems 4.1.1.3 and 4.1.1.4 tell us that $r(n,\Lambda^*) = r(n,\Lambda^F) = \Theta(n^{-r/d})$ and $\mathrm{comp}(\varepsilon,\Lambda^*) = c\cdot\Theta\big((1/\varepsilon)^{d/r}\big)$.

Hence (4.3.2) and (4.3.3) are two formulations of the Galerkin algorithm using FEI. The first formulation shows that this FEI is information from Λ^F, whereas the second shows that it actually is information from Λ^*. Moreover, the first formulation uses functionals $\langle Sf, g_i\rangle_{L_2(\Omega)}$, whereas the second uses $\langle f, S^*g_i\rangle_{L_2(\Omega)}$. Since the formulation (4.3.2) is simpler than the formulation (4.3.3), as well as making it clear that we are using continuous linear information. So in practice, we would probably rather use (4.3.3) than (4.3.2).

Unfortunately, there is still one difficulty with the Galerkin method, even if we decide to use the simpler formulation (4.3.2). Even though $S^*g_1, \dots, S^*g_n$ are well-defined, they may not be easy to calculate. For this reason, we will look at modified Galerkin methods for the solution-restricted Fredholm problem of the second kind.

Once again, we let $\mathscr{S}_{n,k}$ be a finite element subspace of $L_2(\Omega)$, of dimension n and degree k. Let $\{g_1, \dots, g_n\}$ be a basis of $\mathscr{S}_{n,k}$ satisfying (4.1.2.1). Next, we let $\tilde r \in \{1, \dots, r\}$, and let $\mathscr{S}_{\tilde n,\tilde k}$ be a finite element subspace of $H^{\tilde r}(\Omega)$, of dimension $\tilde n$ and degree $\tilde k$. Of course, we must have $\tilde k \ge \tilde r$, since $\mathscr{S}_{\tilde n,\tilde k} \subset H^{\tilde r}(\Omega)$.

For any $n \in \mathbb{N}$, let

$$\tilde n = \Theta(n^{(r-\tilde r)/\tilde r + d/(2\tilde r)}). \tag{4.3.4}$$

For $1 \le i \le n$, let $v_i \in \mathscr{S}_{\tilde n,\tilde k}$ satisfy

$$\langle L^* v_i, s\rangle_{L_2(\Omega)} = \langle g_i, s\rangle_{L_2(\Omega)}. \tag{4.3.5}$$

We then have

LEMMA 4.3.1. *For $1 \le i \le n$, we have*

$$\|g_i - L^* v_i\|_{L_2(\Omega)} \le C n^{-(r/d+1)}.$$

PROOF: Let $1 \le i \le n$. We first note that v_i is an $\mathscr{S}_{\tilde{n},\tilde{k}}$-finite element approximation (in the usual sense) of $S^* g_i$. Using [28, Theorem 6.3.3.2] and (4.3.5), it follows that

$$\|g_i - L^* v_i\|_{L_2(\Omega)} \le \|L^*\| \, \|S^* g_i - v_i\|_{L_2(\Omega)} \le C \tilde{n}^{-\tilde{r}/d} \|S^* g_i\|_{H^{\tilde{r}}(\Omega)} \le C \|S^*\| \, \tilde{n}^{-\tilde{r}/d} \|g_i\|_{H^{\tilde{r}}(\Omega)}.$$

(Here $\|L^*\|$ and $\|S^*\|$ respectively denote the $L_2(\Omega)$ and $H^{\tilde{r}}(\Omega)$ operator norms. See [28, Theorem 6.3.1.1] for a proof that $\|S^*\|$ is finite.) Recall that the spaces $\mathscr{S}_{n,k}$ are based on a quasi-uniform family of triangulations. For $0 \le j \le r$, we may use [28, Lemma A.2.3.4] to see that the inverse inequalities

$$|g_i|_{H^j(\Omega)} \le C n^{(\tilde{r}-j)/d} \|g_i\|_{L_2(\Omega)}$$

hold, where

$$|v|_{H^j(\Omega)} = \left[\sum_{|\alpha|=j} \|D^\alpha v\|^2_{L_2(\Omega)} \right]^{1/2} \qquad \forall\, v \in H^j(\Omega)$$

is the $H^j(\Omega)$-seminorm. So

$$\|g_i\|_{H^{\tilde{r}}(\Omega)} = \left[\sum_{j=0}^{\tilde{r}} |g_i|^2_{H^j(\Omega)} \right]^{1/2} \le C n^{\tilde{r}/d} \|g_i\|_{L_2(\Omega)} \le C n^{\tilde{r}/d - 1/2}$$

the latter since g_i is a bounded function whose support has volume $\Theta(n^{-1})$. Combining these results, we see that

$$\|L^* v_i\|_{L_2(\Omega)} \le C \tilde{n}^{-\tilde{r}/d} n^{\tilde{r}/d - 1/2}.$$

Using this inequality along with (4.3.4), the lemma follows. □

We can now define our information and algorithm. For any $n \in \mathbb{N}$, let

$$\tilde{N}_{n,k} = [\langle f, v_1 \rangle_{L_2(\Omega)}, \ldots, \langle f, v_n \rangle_{L_2(\Omega)}] \qquad \forall f \in F.$$

Then for $f \in F$, we seek $\tilde{u}_{n,k} = \tilde{\phi}_{n,k}(\tilde{N}_{n,k} f)$ satisfying

$$\langle \tilde{u}_{n,k}, g_i \rangle_{L_2(\Omega)} = \langle f, v_i \rangle_{L_2(\Omega)} \qquad (1 \le i \le n).$$

We then have

THEOREM 4.3.2. *Let $\Lambda = \Lambda^*$.*

(1) *For any $n \in \mathbb{N}$, we have*

$$r(n, \Lambda^*) = \Theta(n^{-r/d}).$$

(2) *Let $k \geq r - 1$. For any $n \in \mathbb{N}$, we have*

$$e(\tilde{\phi}_{n,k}, \tilde{N}_{n,k}) = \Theta(n^{-r/d}),$$

so that $\tilde{\phi}_{n,k}$ is an nth nearly minimal error algorithm, and $\tilde{N}_{n,k}$ is nth nearly optimal information.

(3) *For any $\varepsilon > 0$, we have*

$$\mathrm{comp}(\varepsilon, \Lambda^*) = c \cdot \Theta\big((1/\varepsilon)^{d/r}\big).$$

(4) *Let $k \geq r - 1$. For any $\varepsilon > 0$, let $n = \Theta\big((1/\varepsilon)^{d/r}\big)$. Then*

$$\mathrm{cost}^{\text{mod-Gal}}(\varepsilon, \Lambda^*) = c \cdot \Theta\big((1/\varepsilon)^{d/r}\big),$$

and so the modified Galerkin method $\tilde{\phi}_{n,k}$ using modified FEI $\tilde{N}_{n,k}$ is a nearly optimal complexity algorithm for the Fredholm problem of the second kind.

PROOF: Using Lemma 4.3.1, we see that inequality (4.1.2.2) holds. Now we can apply Theorems 4.1.2.3 and 4.1.2.4. □

We close this section by discussing the optimal choice of $\tilde{r}$ in (4.3.4), as well as the amount of preprocessing required in computing $v_1, \ldots, v_n$.

Clearly, $v_1, \ldots, v_n$ are independent of any problem element f, and so their calculation may be be considered precomputation. However, in practice, we would like to compute them as cheaply as possible. There are (at least) two conflicting reasons why it may be difficult or expensive to calculate $v_1, \ldots, v_n$ satisfying (4.3.5) with $\tilde{n}$ given by (4.3.4). The first is that we need $\mathscr{S}_{\tilde{n},\tilde{k}}$ to be a subspace of $H^{\tilde{r}}(\Omega)$. On the one hand, we want to simplify the task of designing the basis functions of the finite element space $\mathscr{S}_{\tilde{n},\tilde{k}}$ over a reference element; this tells us that we should choose $\tilde{r}$ as small as possible, i.e., $\tilde{r} = 1$. On the other hand, we want to minimize amount of work required to calculate $v_1, \ldots, v_n$ once we have designed these basis functions from (4.3.4), this criterion tells to choose $\tilde{r}$ as large as possible, i.e., $\tilde{r} = r$. The question is now one of which criterion to use. To solve this conundrum, we note that we only design the reference element basis functions once, independent of n, whereas the calculation of $v_1, \ldots, v_n$ depends on n. We are probably willing to expend the extra effort involved in designing the basis functions (which only needs to be done once), thereby saving cost arising in the calculation of $v_1, \ldots, v_n$ for various n. In other words, we feel that it would be preferable to choose $\tilde{r} = r$. Thus we choose a finite element subspace $\mathscr{S}_{\tilde{n},\tilde{k}}$ of $H^r(\Omega)$, where $\tilde{k}$ is the degree of the subspace and the dimension $\tilde{n}$ of the subspace satisfies $\tilde{n} = n^{d/(2r)}$.

We now discuss the cost of computing $v_1, \dots, v_n$. Assuming we choose $\tilde{r} = r$, the previous analysis implies that the cost of computing $v_1, \dots, v_n$ is $\Theta(n^{d/(2r)+1})$. Since this cost grows faster than n, on-the-fly calculation of $v_1, \dots, v_n$ will dwarf the remainder of the calculation of $\tilde{\phi}_{n,k}(\tilde{N}_{n,k} f)$ for a problem element f. However, it is possible to precompute $v_1, \dots, v_n$, since they are independent of any f. If we decide to compute ε-approximations for many $f \in F$, with a fixed value of ε, then we may consider the cost of this precomputation as an overhead whose cost we can ignore.

4.4. Mixed elliptic-hyperbolic problems.

In this section, we look at the complexity of a solution-restricted Tricomi problem. This is a simple mixed hyperbolic-elliptic problem, which arises in the study of two-dimensional transonic flow across an airfoil; see [7, Chapter X] and [13] for further discussion.

Let Γ_0 be a simple curve in the region $y > 0$ of the two-dimensional (x, y)-plane, intersecting the x-axis only at the points $A(-1, 0)$ and $B(1, 0)$. Let Γ_1 and Γ_2 be given by

$$x = \begin{cases} -1 + \frac{2}{3}(-y)^{3/2} & (x, y) \in \Gamma_1, \\ 1 - \frac{2}{3}(-y)^{3/2} & (x, y) \in \Gamma_2. \end{cases}$$

Note that Γ_1 and Γ_2 intersect at the point $C(0, y_C)$, where $y_C = -(\frac{3}{2})^{2/3} \doteq -1.31037$. Let $\Omega \subset \mathbb{R}^2$ be the region whose boundary is $\Gamma_0 \cup \Gamma_1 \cup \Gamma_2$.

Define a partial differential operator L by

$$Lu = yu_{xx} + u_{yy}.$$

L is called the *Tricomi operator*. Note that

(1) L is elliptic (but not strongly elliptic) in the elliptic region $\Omega_E = \{ (x, y) \in \Omega : y > 0 \}$. This corresponds to subsonic flow.
(2) L is parabolic on the parabolic line $J = \{ (x, y) \in \Omega : y = 0 \}$. This corresponds to sonic flow.
(3) L is hyperbolic in the hyperbolic region $\Omega_E = \{ (x, y) \in \Omega : y < 0 \}$. This corresponds to supersonic flow.

We also note that Γ_1 and Γ_2 are the characteristic curves of L, respectively emanating from A and B.

Let $f \colon \Omega \to \mathbb{R}$. The *Tricomi problem* is to find a function $u \colon \overline{\Omega} \to \mathbb{R}$ satisfying

$$\begin{aligned} Lu &= f \qquad \text{in } \Omega, \\ u &= 0 \qquad \text{on } \Gamma_0 \cup \Gamma_1. \end{aligned} \tag{4.4.1}$$

Since the domain Ω is divided by the sonic line into elliptic and hyperbolic regions, the Tricomi problem is an example of a mixed elliptic-hyperbolic problem.

Remark: Note that we only prescribe boundary data for the Tricomi problem on part of $\partial\Omega$. The simple explanation for this is that since the Tricomi problem is hyperbolic in Ω_H, prescribing boundary data on one of the characteristic lines that bounds Ω_H is sufficient for solvability on $\overline{\Omega}_H$, whereas requiring the solution to satisfy given boundary values on two characteristic lines would overly constrain the problem in Ω_H. Note that in principle, once we have solved the problem in $\overline{\Omega}_H$, we could solve the remaining degenerate elliptic problem in $\overline{\Omega}_E$. $\square$

We now show how to express (4.4.1) as a standard problem. First, we let

$$H^2(\mathrm{bd}) = \{ v \in H^2(\Omega) : v = 0 \text{ on } \Gamma_0 \cup \Gamma_1 \}$$

and

$$H^2(\mathrm{bd})^+ = \{ v \in H^2(\Omega) : v = 0 \text{ on } \Gamma_0 \cup \Gamma_2 \}.$$

(The notations bd and bd^+ respectively refer to boundary conditions and adjoint boundary conditions; see [4, pg. 71].) Then

$$\langle Lv, w\rangle_{L_2(\Omega)} = \langle v, Lw\rangle_{L_2(\Omega)} \qquad \forall\, v \in H^2(\mathrm{bd}), w \in H^2(\mathrm{bd})^+, \tag{4.4.2}$$

see [4, pg. 308]. Let $H^{-2}(\mathrm{bd})$ denote the dual space of $H^2(\mathrm{bd})$, which is a Hilbert space under the norm

$$\|v\|_H^{-2}(\mathrm{bd}) = \sup_{w\in H^2(\mathrm{bd})^+} \frac{\langle v, w\rangle_{L_2(\Omega)}}{\|w\|_{H^2(\Omega)}}.$$

We can then define a linear transformation $L\colon L_2(\Omega) \to H^{-2}(\mathrm{bd})$ by letting

$$Lv = \langle v, L\cdot\rangle_{L_2(\Omega)} \qquad \forall\, v \in L_2(\Omega).$$

Using (4.4.2), we see that there is a positive constant C for which

$$\|Lv\|_H^{-2}(\mathrm{bd}) \le C\|v\|_{L_2(\Omega)} \qquad \forall\, v \in L_2(\Omega).$$

Thus the bounded linear transformation $L\colon L_2(\Omega) \to H^{-2}(\mathrm{bd})$ is a weak extension of the original Tricomi operator L. Using [4, Theorem IV.3.1], we find that L is a dense injection.

Hence, we see that the problem (4.4.1) may be expressed as a standard problem, with $X = H^{-2}(\mathrm{bd})$, $G = L_2(\Omega)$, and $W = H^r(\Omega)$ for some $r \ge 0$. Note that $q = 0$. Once again, the class F of problem elements is the unit ball of $H^r(\Omega)$. Since $L\colon L_2(\Omega) \to H^{-2}(\mathrm{bd})$ is defined weakly, we must also define $S = L^{-1}$ weakly. That is, for $f \in F$, we require that $u = Sf$ satisfy

$$\langle u, Lv\rangle_{L_2(\Omega)} = \langle f, v\rangle_{L_2(\Omega)} \qquad \forall\, v \in H^2(\mathrm{bd})^+.$$

Note that this weak definition of S allows us to avoid the issue of boundary values for $L_2(\Omega)$-functions.

We now apply the results of Section 4.1. First, we look at the case $\Lambda = \Lambda^F$. For any $n \in \mathbb{N}$, we once again let $\mathscr{S}_{n,k}$ be an n-dimensional finite element subspace of $L_2(\Omega)$ having degree k. Let $\{g_1, \dots, g_n\}$ be a basis for $\mathscr{S}_{n,k}$. Define finite element information $N_{n,k}$ by

$$N_{n,k}f = [\langle Sf, g_1\rangle_{L_2(\Omega)}, \dots, \langle Sf, g_n\rangle_{L_2(\Omega)}].$$

For $f \in F$, let $u_{n,k} = \phi_{n,k}(N_{n,k}f) \in \mathscr{S}_{n,k}$ be the Galerkin approximation given by

$$\langle u_{n,k}, g_i\rangle_{L_2(\Omega)} = \langle Sf, g_i\rangle_{L_2(\Omega)} \qquad (1 \le i \le n). \tag{4.4.3}$$

We then have

THEOREM 4.4.1. *Let* $\Lambda = \Lambda^F$.

(1) *For any* $n \in \mathbb{N}$, *we have*

$$r(n, \Lambda^F) = \Theta(n^{-r/2}).$$

(2) *Let* $k \geq r-1$. *For any* $n \in \mathbb{N}$, *we have*

$$e(\phi_{n,k}, N_{n,k}) = \Theta(n^{-r/2}),$$

so that $\phi_{n,k}$ *is an* n*th nearly minimal error algorithm, and* $N_{n,k}$ *is* n*th nearly optimal information.*

(3) *For any* $\varepsilon > 0$, *we have*

$$\mathrm{comp}(\varepsilon, \Lambda^F) = c \cdot \Theta\big((1/\varepsilon)^{2/r}\big).$$

(4) *Let* $k \geq r-1$. *For any* $\varepsilon > 0$, *let* $n = \Theta\big((1/\varepsilon)^{2/r}\big)$. *Then*

$$\mathrm{cost}^{\mathrm{Gal}}(\varepsilon, \Lambda^F) = c \cdot \Theta\big((1/\varepsilon)^{2/r}\big),$$

and so the Galerkin method $\phi_{n,k}$ *using FEI* $N_{n,k}$ *is a nearly optimal complexity algorithm for the Tricomi problem.*

PROOF: Immediate from Theorems 4.1.1.3 and 4.1.1.4. □

Once again, we see that linear functionals $\langle Sf, g_i\rangle_{L_2(\Omega)}$ are hard to directly evaluate, because of their apparent dependence on Sf. To overcome this difficulty, we describe an auxilliary adjoint problem. Even though (4.4.2) tells us that L is formally self-adjoint, we will now write L^* for the adjoint operator.[5] For $g\colon \Omega \to \mathbb{R}$, we wish to find a function $v\colon \overline{\Omega} \to \mathbb{R}$ satisfying

$$\begin{aligned} L^* v &= g \qquad \text{in } \Omega, \\ v &= 0 \qquad \text{on } \Gamma_0 \cup \Gamma_2. \end{aligned} \tag{4.4.4}$$

This problem is called the *adjoint Tricomi problem.* We need to find the proper weak formulation of this adjoint problem. To do this, let $H^{-2}(\mathrm{bd})^+$ denote the dual space of $H^2(\mathrm{bd})^+$, which is a Hilbert space under the norm

$$\|v\|_H^{-2}(\mathrm{bd})^+ = \sup_{w \in H^2(\mathrm{bd})} \frac{\langle v, w\rangle_{L_2(\Omega)}}{\|w\|_{H^2(\Omega)}}.$$

Define a linear transformation $L^*\colon L_2(\Omega) \to H^{-2}(\mathrm{bd})^+$ by letting

$$L^* v = \langle L\cdot, v\rangle_{L_2(\Omega)} \qquad \forall\, v \in L_2(\Omega).$$

[5] We do this because L and L^*, when considered as operators with domain $L_2(\Omega)$, have different codomains.

Using (4.4.2), we see that there is a positive constant C for which

$$\|L^* v\|_H^{-2}(\mathrm{bd})^+ \le C\|v\|_{L_2(\Omega)} \qquad \forall\, v \in L_2(\Omega).$$

Thus the bounded linear transformation $L^*\colon L_2(\Omega) \to H^{-2}(\mathrm{bd})^+$ is a weak extension of the adjoint Tricomi operator L^*. From [4, Theorem IV.3.1], it follows that L^* is a dense injection.

Hence, we have defined L^* weakly. This means that we can define $S^* = (L^*)^{-1}$ weakly. For $g \in F$, we require that $v = S^* g$ satisfy

$$\langle Lu, v\rangle_{L_2(\Omega)} = \langle u, g\rangle_{L_2(\Omega)} \qquad \forall\, u \in H^2(\mathrm{bd}).$$

This is the weak formulation of the adjoint Tricomi problem.

Using (4.4.2), we find that

$$\langle Sf, g\rangle_{L_2(\Omega)} = \langle f, S^* g\rangle_{L_2(\Omega)} \qquad \forall f, g \in L_2(\Omega).$$

In particular, we have

$$\langle Sf, g_i\rangle_{L_2(\Omega)} = \langle f, S^* g_i\rangle_{L_2(\Omega)} \qquad (1 \le i \le n).$$

This means that we can once again consider the computation of $S^* g_1, \ldots, S^* g_n$ as precomputation. Hence we have

$$N_{n,k} f = [\langle f, S^* g_1\rangle_{L_2(\Omega)}, \ldots, \langle f, S^* g_n\rangle_{L_2(\Omega)}].$$

For $f \in F$, it now follows that $u_{n,k} = \phi_{n,k}(N_{n,k} f)$ satisfies

$$\langle u_{n,k}, g_i\rangle_{L_2(\Omega)} = \langle f, S^* g_i\rangle_{L_2(\Omega)} \qquad (1 \le i \le n). \tag{4.4.5}$$

So the information $N_{n,k}$ defined in (4.4.3) is continuous linear information. Moreover, the Galerkin algorithm (4.4.5) is a nearly optimal error algorithm and a nearly optimal complexity algorithm using continuous linear information. Of course, Theorems 4.1.1.3 and 4.1.1.4 tell us that $r(n, \Lambda^*) = r(n, \Lambda^F) = \Theta(n^{-r/2})$ and $\mathrm{comp}(\varepsilon, \Lambda^*) = c \cdot \Theta\big((1/\varepsilon)^{2/r}\big)$.

We now see that (4.4.3) and (4.4.5) are two formulations of the Galerkin algorithm using FEI, the first showing that this FEI is information from Λ^F and the second showing that it actually is from Λ^*. Moreover, the first formulation uses functionals $\langle Sf, g_i\rangle_{L_2(\Omega)}$, whereas the second uses $\langle f, S^* g_i\rangle_{L_2(\Omega)}$. As was the case in the previous section, we prefer to use (4.4.5) instead of (4.4.3).

As in the previous section we see that $S^* g_1, \ldots, S^* g_n$ may be hard to calculate, even though they are well-defined. Hence we need to once again consider modified Galerkin methods for our problem.

Of course, the only issue that we need to resolve is how to calculate $v_1, \ldots, v_n$ such that $L^* v_i$ is sufficiently close to g_i for $1 \le i \le n$. One idea is to let $\mathscr{S}_{\tilde{n},k+2}$ be an $\tilde{n}$-dimensional finite element subspace of $L_2(\Omega)$ having degree $k+2$, where

$$\tilde{n} = \Theta(n^{r+2}). \tag{4.4.6}$$

Suppose that we choose v_i solving the least squares problem

$$\|g_i - L^* v_i\|_{L_2(\Omega)} = \inf_{s \in \mathscr{S}_{\tilde{n},k+2}} \|g_i - L^* s\|_{L_2(\Omega)},$$

i.e., $v_i \in \mathscr{S}_{\tilde{n},k+2}$ satisfies

$$\langle L v_i, L s\rangle_{L_2(\Omega)} = \langle g_i, L s\rangle_{L_2(\Omega)} \qquad \forall\, s \in \mathscr{S}_{\tilde{n},k+2}.$$

We are now ready to define our information and algorithm. For any $n \in \mathbb{N}$, let

$$\tilde{N}_{n,k} = [\langle f, v_1\rangle_{L_2(\Omega)}, \ldots, \langle f, v_n\rangle_{L_2(\Omega)}] \qquad \forall f \in F.$$

Then for $f \in F$, we seek $\tilde{u}_{n,k} = \tilde{\phi}_{n,k}(\tilde{N}_{n,k} f)$ satisfying

$$\langle \tilde{u}_{n,k}, g_i\rangle_{L_2(\Omega)} = \langle f, v_i\rangle_{L_2(\Omega)} \qquad (1 \le i \le n).$$

We then have

THEOREM 4.4.2. *Let $\Lambda = \Lambda^*$.*

(1) *For any $n \in \mathbb{N}$, we have*

$$r(n, \Lambda^*) = \Theta(n^{-r/2}).$$

(2) *Let $k \ge r - 1$. For any $n \in \mathbb{N}$, we have*

$$e(\tilde{\phi}_{n,k}, \tilde{N}_{n,k}) = \Theta(n^{-r/2}),$$

so that $\tilde{\phi}_{n,k}$ is an nth nearly minimal error algorithm, and $\tilde{N}_{n,k}$ is nth nearly optimal information.

(3) *For any $\varepsilon > 0$, we have*

$$\mathrm{comp}(\varepsilon, \Lambda^*) = c \cdot \Theta\big((1/\varepsilon)^{2/r}\big).$$

(4) *Let $k \ge r - 1$. For any $\varepsilon > 0$, let $n = \Theta\big((1/\varepsilon)^{2/r}\big)$. Then*

$$\mathrm{cost}^{\text{mod-Gal}}(\varepsilon, \Lambda^*) = c \cdot \Theta\big((1/\varepsilon)^{2/r}\big),$$

and so the modified Galerkin method $\tilde{\phi}_{n,k}$ using modified FEI $\tilde{N}_{n,k}$ is a nearly optimal complexity algorithm for the Tricomi problem.

PROOF: Since $Lv \in P_k$ for $v \in P_{k+2}$, one can exactly solve the problem $v_i = S^* g_i$, except possibly on the boundary elements. But these boundary elements have area $\Theta(\tilde{n}^{-1/2}) = \Theta(n^{-(r/2+1)})$. Hence $\|g_i - L^* v_i\|_{L_2(\Omega)} \leq C n^{-(r/2+1)}$ for $1 \leq i \leq n$. Now use Theorems 4.1.2.3 and 4.1.2.4. □

We once again note that on-the-fly computation of $v_1, \ldots, v_n$ may be expensive. Indeed, from (4.4.6), we see that $v_1, \ldots, v_n$ may be calculated with cost $\Theta(n^{r+3})$, which of course greatly outweighs the cost of calculating $\tilde{u}_{n,k}$. However, we once again point out that $v_1, \ldots, v_n$ are independent of any $f \in F$. So, if we precompute $v_1, \ldots, v_n$ and if we do not charge for this precomputation (since it is independent of any problem element), we can ignore the cost of the precomputation.

Remark: Note that we used a very weak error bound to show that (4.4.6) implies that $v_1, \ldots, v_n$ are sufficiently accurate. This was motivated by our lack of a shift theorem for the Tricomi problem, so that we cannot assume enough global smoothness in $S^* g_i$ (where g_i is piecewise polynomial) to use the error estimates in 2. It is quite possible that our estimate is overly-pessimistic, and that we can find sufficiently accurate piecewise polynomial approximations to $S^* g_i$ using fewer degrees of freedom. □

4.5. Inverse finite Laplace transform.

In this section, we look at the complexity of a solution-restricted inverse Laplace transform. This is an example of a Fredholm integral equation of the first kind, and is thus an ill-posed problem. This problem arises in remote sensing problems of geomathematics; see [21] for discussion and further examples.

Without loss of generality, we assume that our functions are defined over the unit interval $I = [0, 1]$. Define an operator $L\colon L_2(I) \to L_2(I)$ as

$$(Lu)(s) = \int_0^1 e^{-st} u(t)\, dt \qquad (0 \leq s \leq 1)$$

for $u \in L_2(I)$. Thus Lu is the finite Laplace transform of u. We are interested in the *inverse finite Laplace transform problem*: for $f \in L_2(I)$, find $u \in L_2(I)$ such that $Lu = f$, i.e., such that

$$\int_0^1 e^{-st} u(t)\, dt = f(s) \qquad (0 \leq s \leq 1). \tag{4.5.1}$$

From [14], we see that $L\colon L_2(I) \to L_2(I)$ is an injection. Since L is self-adjoint as an operator on $L_2(I)$, we see that the range of L is dense in $L_2(I)$. Hence the solution operator $S = L^{-1}$ is densely defined in $L_2(I)$. Thus we can express our problem (4.5.1) as a standard problem if we choose $G = X = L_2(I)$ and $W = H^r(I)$ for some $r \geq 0$.

Note that L is compact. Thus the problem of finding $u \in L_2(I)$ satisfying $Lu = f$ is ill-posed for $f \in L_2(I)$.

Remark: The problem of inverting L is in fact *very* ill-posed. From [28, pp. 198-199], we find that $L\colon L_2(I) \to H^r(I)$ is a compact injection with dense range for any $r \geq 0$. Thus the (densely defined) solution operator $S = L^{-1}\colon H^r(I) \to L_2(I)$ is unbounded, no matter how big we choose r to be. Hence the results of [26] imply that if we choose our problem elements to be the unit ball of a Hilbert Sobolev space $H^r(I)$, then the error of any finite-cost algorithm is infinite, no matter how large we choose r to be. Simply stated, this means that restricting the problem elements for (4.5.1) will not work. This explains why we are interested in a solution-restricted version of this problem. □

Having expressed (4.5.1) as a solution-restricted problem, we can now use the results of Section 4.1. First, we look at the case $\Lambda = \Lambda^F$. For any $n \in \mathbb{N}$, let $\mathscr{S}_{n,k}$ be an n-dimensional finite element subspace of $L_2(I)$ having degree k. Let $\{g_1, \dots, g_n\}$ be a basis for $\mathscr{S}_{n,k}$. Define finite element information $N_{n,k}$ by

$$N_{n,k}f = [\langle Sf, g_1\rangle_{L_2(I)}, \dots, \langle Sf, g_n\rangle_{L_2(I)}].$$

For $f \in F$, let $u_{n,k} = \phi_{n,k}(N_{n,k}f)$ be the Galerkin method given by

$$\langle u_{n,k}, g_i\rangle_{L_2(I)} = \langle Sf, g_i\rangle_{L_2(I)} \qquad (1 \le i \le n). \tag{4.5.2}$$

We then have

THEOREM 4.5.1. *Let $\Lambda = \Lambda^F$.*

(1) *For any $n \in \mathbb{N}$, we have*
$$r(n, \Lambda^F) = \Theta(n^{-r}).$$
(2) *Let $k \ge r - 1$. For any $n \in \mathbb{N}$, we have*
$$e(\phi_{n,k}, N_{n,k}) = \Theta(n^{-r}),$$
so that $\phi_{n,k}$ is an nth nearly minimal error algorithm, and $N_{n,k}$ is nth nearly optimal information.

(3) *For any $\varepsilon > 0$, we have*
$$\mathrm{comp}(\varepsilon, \Lambda^F) = c \cdot \Theta\big((1/\varepsilon)^{1/r}\big).$$
(4) *Let $k \ge r - 1$. For any $\varepsilon > 0$, let $n = \Theta\big((1/\varepsilon)^{1/r}\big)$. Then*
$$\mathrm{cost}^{\mathrm{Gal}}(\varepsilon, \Lambda^F) = c \cdot \Theta\big((1/\varepsilon)^{1/r}\big),$$
and so the Galerkin method $\phi_{n,k}$ using FEI $N_{n,k}$ is a nearly optimal complexity algorithm for the inverse finite Laplace transform problem.

PROOF: Immediate from Theorems 4.1.1.3 and 4.1.1.4. □

As with problems that we have considered previously, we prefer to avoid using the functionals $\langle Sf, g_i\rangle_{L_2(I)}$ making up the finite element information $N_{n,k}$. We would rather use functionals of the form $\langle f, v_i\rangle_{L_2(I)}$. In the discussion that follows, we will take advantage of the self-adjointness of L as an operator on $L_2(I)$, which implies that S is self-adjoint in $L_2(I)$. Thus we write L and S instead of L^* and S^* in what follows.

Note that we cannot rewrite the functionals $\langle Sf, g_i\rangle_{L_2(I)}$ in the form $\langle f, Sg_i\rangle_{L_2(I)}$. This is because functions belonging to the domain of S (i.e., the range of L) are infinitely differentiable, while the functions g_i are only piecewise smooth. So the nearly optimal Galerkin method using FEI for the inverse finite Laplace transform, which (of course) uses Λ^F-information, cannot be cosmetically rewritten as a method using Λ^*-information. This is an important difference between this problem and the other problems that we have studied.

This being the case, we now let $\Lambda = \Lambda^*$, and look for nearly optimal Λ^*-information, using our general results in Section 4.1.2. The main idea is to choose $v_1, \dots, v_n \in X$ such that $Lv_1, \dots, Lv_n$ are sufficiently good approximations of $g_1, \dots, g_n$. To do this, we use ideas based on those in [28, pp. 225–226].

We assume that the finite element space $\mathscr{S}_{n,k}$ is defined over a uniform partition of I. Suppose that $\mathscr{S}_{n,k} \subset H^p(I)$ for some positive integer p, so that $k \geq p$. Recall that $\{g_1, \dots, g_n\}$ is a standard basis for $\mathscr{S}_{n,k}$, having small supports. Our task is to find functions $\{v_1, \dots, v_n\}$ such that (4.1.2.2) holds.

To do this, we let R and m be parameters depending on n, p, and r. We will give criteria for choosing R and m later. For any $j \in \mathbb{N}$, we let $G_j(x) = P_j(2x-1)$, where P_j is the usual Legendre polynomial of degree j. We write $G_j(x) = \sum_{l=0}^{j} \beta_{j,l} x^l$. Next, for any $l \in \mathbb{N}$, we let

$$w_l = \begin{cases} (-1)^j \binom{l}{j} R^{l+1} & \text{if } t \in [j/R, (j+1)/R) \text{ for some } j \in \{0, \dots l\}, \\ 0 & \text{if } t \in [(l+1)/R, 1] \end{cases}$$

We then let

$$v_i = \sum_{j=0}^{m} (2j+1) \langle g_i, G_j \rangle_{L_2(I)} \sum_{l=0}^{j} \beta_{j,l} w_l$$

for $1 \leq i \leq n$. Then we have

LEMMA 4.5.1. *For any n and k, there exist R and m such that*

$$\|Lv_i - g_i\|_{L_2(I)} \leq Cn^{-(r+1)} \qquad (1 \leq i \leq n). \tag{4.5.3}$$

PROOF: Let $i \in \{1, \dots, n\}$. Since the support of g_i has length $\Theta(n^{-1})$, there exists a positive constant A such that

$$\|g_i\|_{L_2(I)} \leq An^{-1/2}$$

Since $\mathscr{S}_{n,k}$ is defined over a uniform partition of I, we know that the inverse inequality [28, Lemma A.2.3.4] holds, and so there exists $B > 0$ such that

$$\|g_i^{(p)}\|_{L_2(I)} \leq BA^{-1} n^p \|g_i\|_{L_2(I)} \leq Bn^{p+1/2}.$$

We claim that for (4.5.3) to hold, it suffices to choose

$$R \geq \frac{A}{\sqrt{3}C} n^{r+1/2} \qquad \text{and} \qquad m \geq \frac{e}{4} \left(\frac{2B}{C} \right)^{1/p} n^{(r+p+1/2)/p}. \tag{4.5.4}$$

Indeed, let

$$g_{i,m} = \sum_{j=0}^{m} (2j+1)\langle g_i, G_j\rangle_{L_2(I)} G_j$$

be the mth Legendre series approximation of g_i. From [8, Theorem 1.3.2], we have

$$\|g_i - g_{i,m}\|_{L_2(I)} \le \left(\frac{e}{4m}\right)^p \|g_i^{(p)}\|_{L_2(I)}.$$

Following [28, pp. 225–226], we find that

$$\|Lv_i - g_i\|_{L_2(I)} \le \frac{A}{2R\sqrt{3}} \cdot \frac{1}{n^{1/2}} + B\left(\frac{e}{4m}\right)^p n^{p-1/2}.$$

Using (4.5.4), we now find that (4.5.3) holds, as claimed. □

We are now ready to define our information and algorithm. For any $n \in \mathbb{N}$, let $v_1, \dots, v_n$ be as defined in Lemma 4.5.1. Define information $\tilde{N}_{n,k}$ as

$$\tilde{N}_{n,k} = [\langle f, v_1\rangle_{L_2(\Omega)}, \dots, \langle f, v_n\rangle_{L_2(\Omega)}] \qquad \forall f \in F.$$

The modified Galerkin algorithm $\tilde{\phi}_{n,k}$ is as follows: For $f \in F$, we seek $\tilde{u}_{n,k} = \tilde{\phi}_{n,k}(\tilde{N}_{n,k} f)$ satisfying

$$\langle \tilde{u}_{n,k}, g_i\rangle_{L_2(\Omega)} = \langle f, v_i\rangle_{L_2(\Omega)} \qquad (1 \le i \le n).$$

We then have

Theorem 4.5.2. *Let $\Lambda = \Lambda^*$.*

(1) *For any $n \in \mathbb{N}$, we have*

$$r(n, \Lambda^*) = \Theta(n^{-r}).$$

(2) *Let $k \ge r-1$. For any $n \in \mathbb{N}$, we have*

$$e(\tilde{\phi}_{n,k}, \tilde{N}_{n,k}) = \Theta(n^{-r}),$$

so that $\tilde{\phi}_{n,k}$ is an nth nearly minimal error algorithm, and $\tilde{N}_{n,k}$ is nth nearly optimal information.

(3) *For any $\varepsilon > 0$, we have*

$$\mathrm{comp}(\varepsilon, \Lambda^*) = c \cdot \Theta\big((1/\varepsilon)^{1/r}\big).$$

(4) *Let $k \ge r-1$. For any $\varepsilon > 0$, let $n = \Theta\big((1/\varepsilon)^{1/r}\big)$. Then*

$$\mathrm{cost}^{\text{mod-Gal}}(\varepsilon, \Lambda^*) = c \cdot \Theta\big((1/\varepsilon)^{1/r}\big),$$

and so the modified Galerkin method $\tilde{\phi}_{n,k}$ using modified FEI $\tilde{N}_{n,k}$ is a nearly optimal complexity algorithm for the inverse finite Laplace transform problem.

PROOF: Immediate from Theorems 4.1.2.3 and 4.1.2.4, along with Lemma 4.5.1. □

We briefly discuss the choice of p. Clearly the larger we make p, the smaller m needs to be. The optimal choice is then to make $p = k$. Since the best choice for k is to let $k = r - 1$, we see that the best choice for p is $p = r - 1$. It then follows that $m = \Theta(n^\theta)$, where $\theta = 2 + 3/(2r - 2)$. Since m grows faster than n^2, the cost of computing $v_1, \ldots, v_n$ on-the-fly grows faster than n^3, and is therefore impractical. However, since $v_1, \ldots, v_n$ are independent of any f, they may be precomputed. If we wish to compute ε-approximations for many $f \in F$, with a fixed value of ε, then we can safely ignore the cost of this precomputation.

4.6. The backwards heat equation.

In our final application, we look at the complexity of the heat equation running backwards in time. This is one of the most famous classical examples of an ill-posed problem. Further discussion and references may be found in [10], [11], [12], [17], and [19].

Let $I = [0, 1]$. For any positive integer j, let

$$s_j(x) = \sqrt{2} \sin j\pi x \qquad \forall\, x \in I. \tag{4.6.1}$$

Then $\{s_j\}_{j=1}^\infty$ is an orthonormal basis for $L_2(I)$. For $t \in \mathbb{R}$, we define a positive-definite self-adjoint operator H_t in $L_2(I)$ by letting

$$H_t f = \sum_{j=1}^\infty e^{-\pi^2 j^2 t} \langle f, s_j \rangle_{L_2(I)} s_j \qquad \forall f \in L_2(I). \tag{4.6.2}$$

Note that

(1) If $t > 0$, then H_t is compact.
(2) H_0 is the identity map.
(3) If $t < 0$, then H_t is a densely-defined unbounded operator.
(4) We have the semigroup properties

$$\begin{aligned} H_t H_\tau &= H_{t+\tau} \qquad && \forall\, t, \tau \in \mathbb{R}, \\ H_{-t} &= H_t^{-1} && \forall\, t \in \mathbb{R}. \end{aligned}$$

(5) For any $t \in \mathbb{R}$, let $u(\cdot, t) = H_t f$. Then $u(x, t)$ satisfies

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} \qquad && \text{in } I \times \mathbb{R}, \\ u(\cdot, 0) &= f && \text{in } I, \\ u(0, \cdot) &= u(1, \cdot) = 0 && \text{in } \mathbb{R}. \end{aligned}$$

Thus $H_t f$ is the solution of the heat equation at time t, with u at time $t = 0$ being given by f. We will call these conditions at time $t = 0$ *initial conditions* if $t > 0$ and *final conditions* if $t < 0$.

We wish to solve this problem at time $t = -t_0$, where $t_0 > 0$. Since the elapsed time is negative, we call this problem a *backwards heat equation.* The operator H_{-t_0} is a unbounded in $L_2(I)$, and so this problem is ill-posed. We shall formulate our problem as a solution-restricted operator equation. Our restriction will be to assume that the solution is $L_2(I)$-bounded at some time $t = -t_1$ in the past, where $t_0 < t_1$. It is well-known that the ill-posed backwards heat equation becomes well-posed under this hypothesis, see [10]. Note that the approach taken here is similar to that in [27].

To express this problem as a solution-restricted operator equation, we let $X = G = L_2(I)$ and define $L\colon G \to X$ be defined as $L = H_{t_0}$. Then $Lu = f$ iff f is the solution of a heat equation at time $t = t_0$, with initial conditions given by u at time $t = 0$. From the previous comments, we see that L is a compact injection with dense range. Our solution operator is now $S = L^{-1}\colon G \to X$. From the semigroup properties, we see that $S = H_{-t_0}$, i.e., $u = Sf$ is the solution of a heat equation at time $t = -t_0$, with final conditions given by f at time $t = 0$. Since S is an unbounded operator, we are trying to solve an ill-posed problem. Finally, we let $W = H_{t_1-t_0}\big(L_2(I)\big)$, which is a Hilbert space under the norm

$$\|w\|_W = \|H_{-(t_1-t_0)}w\|_{L_2(I)} \qquad \forall\, w \in W.$$

Since $t_1 > t_0$, we see that $\|\cdot\|_{L_2(I)} \le \|\cdot\|_W$, and so the identity embedding $E\colon W \to L_2(I)$ is continuous, with $\|E\| \le 1$. Hence, the spaces X, G, and W, along with the operator $L\colon G \to X$, define a solution-restricted operator equation. It is straightforward to check that the class F of problem elements is the set of all $f \in L_2(I)$ for which $H_{-t_1}f$ belongs to the unit ball of $L_2(I)$. Hence this solution-restricted problem is the solution of the backwards heat equation at time $t = -t_0$ with prescribed final data at time $t = 0$, under the constraint of a known bound on the solution at the earlier time $t = -t_1$.

We will use the results contained in Theorem 3.3. This requires us to find an orthonormal basis for W consisting of eigenvectors of E^*E, as well as the corresponding eigenvalues $\gamma_1^2 \ge \gamma_2^2 \ge \cdots > 0$. For $j = 1, 2, \dots$, we will let

$$z_j = H_{t_1-t_0}s_j = e^{-\pi^2 j^2 (t_1-t_0)}s_j. \tag{4.6.3}$$

Then $\{z_j\}_{j=1}^\infty$ is an orthonormal basis for W.

Lemma 4.6.1. *For any positive integer j, we have*

$$E^*Ez_j = \gamma_j^2 z_j,$$

where

$$\gamma_j = e^{-\pi^2 j^2 (t_1-t_0)} \tag{4.6.4}$$

and z_j is given by (4.6.3).

Proof: We first claim that $E^*E = H_{2(t_1-t_0)}$. Indeed, let $v, w \in W$. Since $H_{-(t_1-t_0)}$ is self-adjoint on $L_2(I)$, we have

$$\begin{aligned}\langle v, w\rangle_{L_2(I)} &= \langle Ev, Ew\rangle_{L_2(I)} = \langle v, E^*Ew\rangle_W \\ &= \langle H_{-(t_1-t_0)}v, H_{-(t_1-t_0)}E^*Ew\rangle_{L_2(I)} = \langle v, H_{-2(t_1-t_0)}E^*Ew\rangle_{L_2(I)}.\end{aligned}$$

Thus $H_{-2(t_1-t_0)}E^*E$ is the identity operator on W, and so $E^*E = H^{-1}_{-2(t_1-t_0)} = H_{2(t_1-t_0)}$, as claimed.

Next, we claim that for any $w \in W$ and any index j, we have

$$\langle w, z_j \rangle_W z_j = \langle w, s_j \rangle_{L_2(I)} s_j. \tag{4.6.5}$$

Indeed, since $H_{-(t_1-t_0)}s_j$ is a scalar multiple of s_j and $H_{-(t_1-t_0)}$ is self-adjoint, we have

$$\begin{aligned}\langle w, z_j \rangle_W z_j &= \langle H_{-(t_1-t_0)}w, H_{-(t_1-t_0)}z_j \rangle_{L_2(I)} z_j = \langle H_{-(t_1-t_0)}w, s_j \rangle_{L_2(I)} H_{t_1-t_0} s_j \\ &= \langle w, H_{-(t_1-t_0)} 2s_j \rangle_{L_2(I)} H_{t_1-t_0} s_j = \langle w, s_j \rangle_{L_2(I)} s_j,\end{aligned}$$

as claimed.

Now since $E^*E = H_{2(t_1-t_0)}$, we may use (4.6.2) and (4.6.5) to find that

$$E^*Ew = H_{2(t_1-t_0)}w = \sum_{j=1}^{\infty} e^{-2\pi^2 j^2 (t_1-t_0)} \langle w, s_j \rangle_{L_2(I)} s_j = \sum_{j=1}^{\infty} e^{-2\pi^2 j^2 (t_1-t_0)} \langle w, z_j \rangle_W z_j.$$

The result follows immediately. $\square$

We are now ready to apply the results in Theorem 3.3. First, we consider the case $\Lambda = \Lambda^F$. For any $n \in \mathbb{N}$, we define information N_n as

$$N_n f = [\langle Sf, z_1 \rangle_W, \ldots, \langle Sf, z_n \rangle_W],$$

where $z_1, \ldots, z_n$ are given by (4.6.3), and an algorithm ϕ_n using N_n as

$$\phi_n(N_n f) = \sum_{j=1}^{n} \langle Sf, z_j \rangle_W z_j \qquad \forall f \in F.$$

We then have

THEOREM 4.6.1. *Let $\Lambda = \Lambda^F$.*

(1) *For any $n \in \mathbb{N}$, we have*

$$r(n, \Lambda^F) = e(\phi_n, N_n) = e^{-\pi^2 (n+1)(t_1-t_0)}.$$

Hence N_n is nth optimal information in Λ^F, and ϕ_n is an nth minimal error algorithm.

(2) *For any $\varepsilon > 0$, the ε-cardinality number*

$$m(\varepsilon, \Lambda^F) = \inf \{ \text{integers } n \geq 0 : \gamma_{n+1} \leq \varepsilon \}$$

is given by

$$m(\varepsilon, \Lambda^F) = \max \left\{ \left\lceil \frac{1}{\pi\sqrt{t_1 - t_0}} \sqrt{\ln \frac{1}{\varepsilon}} \right\rceil - 1, 0 \right\}.$$

(3) *For $\varepsilon > 0$, let $n = m(\varepsilon, \Lambda^F)$. Then*

$$c\, m(\varepsilon, \Lambda^F) \leq \text{comp}(\varepsilon, \Lambda^F) \leq \text{cost}(\phi_n, N_n) \leq (c+2) m(\varepsilon, \Lambda^F) - 1.$$

Hence for $c \gg 1$, the algorithm ϕ_n using information N_n is a nearly optimal complexity algorithm.

PROOF: Immediate from Lemma 4.6.1 and Theorem 3.4. $\square$

Hence we see that $\mathrm{comp}(\varepsilon, \Lambda^F) = c \cdot \Theta\big(\sqrt{\ln(1/\varepsilon)}\big)$, giving the complexity of the solution-restricted backwards heat equation when $\Lambda = \Lambda^F$.

We now look at the case $\Lambda = \Lambda^*$. For any $n \in \mathbb{N}$, we define information $\tilde{N}_n \in \Lambda_n^*$ as

$$\tilde{N}_n f = [\langle f, s_1\rangle_{L_2(I)}, \dots, \langle f, s_n\rangle_{L_2(I)}],$$

where $s_1, \dots, s_n$ are given by (4.6.1), and an algorithm $\tilde{\phi}_n$ using $\tilde{N}_n$ as

$$\tilde{\phi}_n(\tilde{N}_n f) = \sum_{j=1}^{n} e^{\pi^2 j^2 t_0} \langle f, s_j\rangle_{L_2(I)} s_j \qquad \forall f \in F.$$

Note that the algorithm $\tilde{\phi}_n$ is the n-term truncation of the standard series representation H_{-t_0} of the solution to the heat equation at time $t = -t_0$.

We then have

THEOREM 4.6.2. *Let $\Lambda = \Lambda^*$.*

(1) *For any $n \in \mathbb{N}$, we have*

$$r(n, \Lambda^*) = e(\phi_n, N_n) = e^{-\pi^2 (n+1)(t_1 - t_0)}.$$

Hence N_n is nth optimal information in Λ^*, and ϕ_n is an nth minimal error algorithm.

(2) *For any $\varepsilon > 0$, the ε-cardinality number*

$$m(\varepsilon, \Lambda^*) = \inf\{\, \text{integers } n \ge 0 : \gamma_{n+1} \le \varepsilon \,\}$$

is given by

$$m(\varepsilon, \Lambda^*) = \max\left\{ \left\lceil \frac{1}{\pi\sqrt{t_1 - t_0}} \sqrt{\ln \frac{1}{\varepsilon}} \right\rceil - 1, 0 \right\}.$$

(3) *For $\varepsilon > 0$, let $n = m(\varepsilon, \Lambda^*)$. Then*

$$c\, m(\varepsilon, \Lambda^*) \le \mathrm{comp}(\varepsilon, \Lambda^*) \le \mathrm{cost}(\phi_n, N_n) \le (c+2) m(\varepsilon, \Lambda^*) - 1.$$

Hence for $c \gg 1$, the algorithm ϕ_n using information N_n is a nearly optimal complexity algorithm.

PROOF: Let $j \in \{1, \dots, n\}$. Using (4.6.5) and the self-adjointness of $S = H_{-t_0}$, we find

$$\langle Sf, z_j\rangle_W z_j = \langle H_{-t_0} f, s_j\rangle_{L_2(I)} s_j = \langle f, H_{-t_0} s_j\rangle_{L_2(I)} s_j = e^{\pi^2 j^2 t_0} \langle f, s_j\rangle_{L_2(I)} s_j.$$

Thus $\tilde{\phi}_n(\tilde{N} f) = \phi_n(N_n f)$ for any $f \in F$. The theorem now follows immediately from Theorem 4.6.1. $\square$

Hence, we have shown that the optimal Λ^F-information is actually Λ^*-information, and that we get a minimal error algorithm by truncating the standard series representation for the solution of the backwards heat equation. Moreover, $\mathrm{comp}(\varepsilon, \Lambda^*) = c \cdot \Theta\big(\sqrt{\ln(1/\varepsilon)}\big)$, giving the complexity of the solution-restricted backwards heat equation when $\Lambda = \Lambda^*$.

Bibliography

1. Agmon, S., *Lectures on elliptic boundary value problems*, Van Nostrand, Princeton, 1965.
2. Aziz, A. K., Schneider, M., Werschulz, A., *A finite element method for the Tricomi problem*, Numerische Mathematik **35** (1980), 13–20.
3. Babuška, I. and Aziz, A. K., *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. Berezanskiĭ, Yu. M., *Expansion in Eigenvectors of Self-Adjoint Operators*, Vol. 17, Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, 1968.
5. Bramble, J. H. and Nitsche, J., *A generalized Ritz-least-squares method for Dirichlet problems*, SIAM Journal on Numerical Analysis **10** (1973), 81–93.
6. Ciarlet, P. G., *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
7. Dautrary, R. and Lions, J.-L., *Mathematical Analysis and Numerical Methods for Science and Technology Vol 3: Functional and Variational Methods*, Springer, Berlin, 1988;Vol. 4: Integral Equations and Numerical Methods Springer, Berlin, 1990.
8. Gui, W. Z., *Hierarchical elements, local mappings, and the h-p version of the finite element method (I)*, Journal of Computational Mathematics **6** (1988), 54–67.
9. Hormander, L., *Analysis of linear partial differential operators*, Springer-Verlag, Berlin, 1983.
10. John, F., *Continuous dependence on data for solutions of partial differential equations with a prescribed bound*, Comm. Pure Appl. Math. **13** (1960), 551–585.
11. Lavrentjev, M. M., *On the Cauchy problem for the Laplace equation*, (in Russian), Dokl. Akad. Nauk SSSR **127** (1959), 205–206.
12. Lavrent'ev, M. M., Romanov, V. G., Shishatskiĭ, S. P., *Ill-Posed Problems of Mathematical Physics and Analysis*, Vol. 64, Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, 1986.
13. Manwell, A. R., *The Tricomi Equation With Applications to the Theory of Plane Transonic Flow*, Vol. 35, Research Notes in Mathematics, Pitman, Boston, 1979.
14. Marti, J. T., *Numerical solution of Fujita's equation*, Improperly posed problems and their numerical treatment, (G. Hämmerlin and K.-H. Hoffmann, eds.), Birkhäuser-Verlag, Basel, 1983, pp. 179–187.
15. Morozov, V. A., *Methods for Solving Incorrectly Posed Problems*, Springer-Verlag, New York, 1984.
16. Oden, J. T. and Reddy, J. N., *An introduction to the mathematical theory of finite elements*, Wiley-Interscience, New York, 1976.
17. Nedelkov, I. P., *Improper problems in computational physics*, Computer Physics Communications (1972), 157–164.
18. Pinkus, A., *n-Widths in Approximation Theory*, Springer-Verlag, Berlin, 1985.
19. Seidman, T. I., *Computational approaches to ill-posed problems: general considerations*, Proceedings of the 1976 Conference on Information Science and Systems (F. M. Davidson and G. M.Masson, eds.), Johns Hopkins University Press, Baltimore, 1978, pp. 179–182.
20. Tikhonov, A. N., *On the solution stability of inverse problems*, Doklady Akademii Nauk SSSR **5** (1943), 195–198. (Russian)
21. Twomey, S., *Introduction to the mathematics of inversion in remote sensing and indirect measurement*, Developments in Geomathematics **3**, Elsevier Scientific Publ., Amsterdam, 1977.
22. Traub, J. F., Wasilkowski, G. W., and Woźniakowski, H., *Information-Based Complexity*, Academic Press, New York, 1988.
23. Traub, J. F. and Werschulz, A. G., *Linear ill-posed problems are solvable on the average for all Gaussian measures*, Mathematical Intelligencer **16** (1993), 42–48.
24. Werschulz, A. G., *What is the complexity of the Fredholm problem of the second kind?*, Journal of Integral Equations **9** (1985), 213–241.
25. ———, *Finite elements are not always optimal*, Advances in Applied Mathematics **8** (1987), 354–375.
26. ———, *What is the complexity of ill-posed problems?*, Numerical Functional Analysis and Optimization **9** (1987), 945–967.
27. ———, *Ill-posed problems in various settings: an information-based survey*, Proceedings of the 1990 American Control Conference, IEEE Press, Piscataway, NJ, 1990, pp. 380–385.

28. ______, *The Computational Complexity of Differential and Integral Equations: An Information-Based Approach*, Oxford University Press, Oxford, 1991.