The Discrete Galerkin Method
for Integral Equations

By Kendall Atkinson* and Alex Bogomolny

Abstract. A general theory is given for discretized versions of the Galerkin method for solving Fredholm integral equations of the second kind. The discretized Galerkin method is obtained from using numerical integration to evaluate the integrals occurring in the Galerkin method. The theoretical framework that is given parallels that of the regular Galerkin method, including the error analysis of the superconvergence of the iterated Galerkin and discrete Galerkin solutions. In some cases, the iterated discrete Galerkin solution is shown to coincide with the Nyström solution with the same numerical integration method. The paper concludes with applications to finite element Galerkin methods.

1. Introduction. Consider the numerical solution of the Fredholm integral equation of the second kind,

$$(1.1) \qquad x(s) - \int_D K(s,t)x(t)\, d\sigma(t) = y(s), \qquad s \in D.$$

In this paper we will define and analyze the discrete Galerkin method for the numerical solution of (1.1). The Galerkin method is a well-known procedure for the approximate solution of this equation (e.g., see [5, p. 62]); and the discrete Galerkin method results when the integrations of the Galerkin method are evaluated numerically. Before giving a more precise definition of the discrete Galerkin method, we review results for the Galerkin method.

In Eq. (1.1), the region D is to be a closed subset of $\mathbf{R}^m$, some $m \geq 1$; and the dimension of D can be less than m, for example, if D is a surface in $\mathbf{R}^3$. For the discrete Galerkin method, we will assume that $K(s,t)$ is continuous for $s,\ t \in D$, although that is not necessary for the discussion of Galerkin's method given below. The equation (1.1) is written symbolically as

$$(1.2) \qquad\qquad (I - \mathcal{K})x = y,$$

with the integral operator assumed to be a compact operator from $L^2(D)$ to $L^2(D)$ and from $L^\infty(D)$ to $C(D)$. Further, it is assumed that (1.1) is uniquely solvable in $\mathcal{X}$ for all $y \in \mathcal{X}$, for both $\mathcal{X} = C(D)$ and $\mathcal{X} = L^2(D)$. Additional assumptions on D and $K(s,t)$ will be given as they are needed in the applications presented later in the paper. Generally $y \in C(D)$, and this will imply $x \in C(D)$.

The Galerkin method for solving (1.2) is defined as follows. Let $\{S_h: h > 0\}$ denote a sequence of finite-dimensional approximating subspaces of both $L^2(D)$ and $C(D)$, and let P_h denote the orthogonal projection of $L^2(D)$ onto S_h. Then to approximate (1.2), solve

$$(1.3) \qquad (I - P_h\mathscr{K})x_h = P_h y, \qquad x_h \in L^2(D);$$

or equivalently, pick $x_h \in S_h$ such that

$$(1.4) \qquad ((I - \mathscr{K})x_h, \psi) = (y, \psi), \quad \text{all } \psi \in S_h,$$

using the inner product of $L^2(D)$. This is called the Galerkin method for (1.2), relative to the subspace S_h. In addition, define the iterated Galerkin solution by

$$(1.5) \qquad x_h^* = y + \mathscr{K}x_h.$$

The error analysis of x_h and x_h^* is well known, both in $L^2(D)$ and $C(D)$. For a simple error analysis of x_h in $L^2(D)$, see [5, p. 62]; for more general error analyses of x_h and x_h^*, in both $L^2(D)$ and $C(D)$, see [12], [13], [15], [28], [29].

The error analysis of x_h and x_h^*, whether in $L^2(D)$ or $C(D)$, usually depends on showing that

$$(1.6) \qquad \|\mathscr{K} - P_h\mathscr{K}\| \to 0 \quad \text{as } h \to 0$$

with the norm dependent on which space $\mathscr{X} = L^2(D)$ or $C(D)$ is being considered. If

$$(1.7) \qquad P_h x \to x \quad \text{as } h \to 0, \text{ all } x \in \mathscr{X},$$

then (1.6) follows from the compactness of $\mathscr{K}$ on $\mathscr{X}$; and generally we will be assuming (1.7) for both choices of $\mathscr{X}$. From (1.6) and the assumed existence of $(I - \mathscr{K})^{-1}$ on $\mathscr{X}$, it follows that $(I - P_h\mathscr{K})^{-1}$ exists and is uniformly bounded for all sufficiently small values of h. Furthermore,

$$(1.8) \qquad x - x_h = (I - P_h\mathscr{K})^{-1}(x - P_h x),$$

$$(1.9) \qquad \|x - x_h\| \leqslant \|(I - P_h\mathscr{K})^{-1}\|\|x - P_h x\|.$$

Together with (1.7), this shows convergence, along with the rate of convergence. The value of $\|x - P_h x\|$ will depend on both S_h and the smoothness of the unknown solution x.

For the iterated Galerkin solution x_h^*, it is straightforward to show that

$$(1.10) \qquad (I - \mathscr{K}P_h)x_h^* = y$$

and that

$$(1.11) \qquad P_h x_h^* = x_h.$$

Using an error analysis for x_h^*, one can also be given for x_h as follows:

$$(1.12) \qquad \begin{aligned} x - x_h &= x - P_h x_h^* = [x - P_h x] + P_h[x - x_h^*], \\ \|x - x_h\| &\leqslant \|x - P_h x\| + \|P_h\|\|x - x_h^*\|. \end{aligned}$$

The analysis of (1.10) in $L^2(D)$ hinges on the fact that with the corresponding operator norm,

$$(1.13) \qquad \|\mathscr{K} - \mathscr{K}P_h\| = \|\mathscr{K}^* - P_h\mathscr{K}^*\| \to 0 \quad \text{as } h \to 0.$$

The compactness of $\mathscr{K}$ implies the same for $\mathscr{K}^*$, and then the assumption (1.7) implies $\|\mathscr{K}^* - P_h\mathscr{K}^*\| \to 0$. Derivation of (1.13) employs the fact that P_h is an orthogonal projection in $L^2(D)$. When P_h is not orthogonal, as on $\mathscr{X} = C(D)$, one

uses ad hoc methods to establish superconvergence; for example, see [15]. It is the purpose of this note to introduce a formalism that allows the derivation of similar estimates for a wide class of discrete Galerkin methods.

Using (1.13), a straightforward stability and convergence analysis can be given for (1.10). In fact, (1.10) is a degenerate kernel method if the application of (1.10) to (1.1) is examined in more detail. The analysis of (1.10) generalizes to $C(D)$. For this, see [12], [13], [15], [28], [29].

For (1.10),

$$(1.14) \qquad x - x_h^* = (I - \mathscr{K}P_h)^{-1}\mathscr{K}(I - P_h)x.$$

Since $I - P_h$ is a projection, $(I - P_h) = (I - P_h)^2$ and

$$(1.15) \qquad \|x - x_h^*\| \leqslant \|(I - \mathscr{K}P_h)^{-1}\| \|\mathscr{K}(I - P_h)\| \|(I - P_h)x\|.$$

Using (1.13) for the case $\mathscr{X} = L^2(D)$, (1.15) shows that x_h^* converges to x more rapidly than does x_h. A similar result can be shown in the space $C(D)$; see the papers cited above. This more rapid convergence is called 'superconvergence'.

In practice, the integrals in (1.4) and (1.5) are not computed exactly, which leads to the discrete and discrete iterated Galerkin methods, respectively. We shall introduce these discrete methods in the next section, initially posing them in a matrix algebra framework. It is shown that in many important cases, the iterated discrete Galerkin method is exactly the same as the direct application of the numerical integration method to (1.1), yielding a Nyström method. Section 3 contains some applications and implications of this result. In Sections 4, 5, and 6 we introduce a functional analysis framework for discrete Galerkin methods, a framework that is a 'discrete analogue' of the analyses given above for the regular Galerkin method. Unlike the previous analyses of the discrete Galerkin method by [12, Chapter 3], [23], [30], our approach yields convergence results for the discrete methods directly, without referring to the convergence of their continuous analogues. Albeit more restrictive, our approach yields more general results where applicable. For reasons of space, Section 6 is placed separately at the end of the issue, in photo-offset form. Section 7 contains applications of this method to several classes of problems.

This paper is devoted to the nonhomogeneous equation (1.1), but the results generalize to the numerical solution of the associated eigenvalue problem. The approximating operators in Section 5 are shown to be a collectively compact family. This means that the standard analyses of the numerical eigenvalue problem can be used, for example, [1], [3], [4], [14, Chapters 5–7], and [25]; and these results can be combined with the techniques of this paper to give an analysis of the discrete Galerkin method for solving the eigenvalue problem. Future papers will discuss the eigenvalue problem, iterative variants to solve the associated linear systems, and applications to Galerkin methods for nonlinear equations.

Discrete Galerkin methods for boundary value problems have also been analyzed previously. An important early paper is [20]; and an analysis of a discrete least-squares method is given in [2]. Although their results are related to those given here, our schema is more general and is not as restricted in the properties of the operators, the approximating subspaces, and the numerical integration scheme.

2. The Discrete Galerkin Method. The Galerkin equation (1.3)–(1.4) is solved by reducing it to an equivalent finite linear system. To this end, let $\{\varphi_1, \dots, \varphi_N\}$ be a basis of S_h, with $N = N_h$ the dimension of S_h. Assume

$$(2.1) \qquad x_h(s) = \sum_{j=1}^{N} \alpha_j \varphi_j(s), \qquad s \in D.$$

Then (1.4) is equivalent to solving for $\{\alpha_j\}$ in the linear system

$$(2.2) \qquad \sum_{j=1}^{N} \alpha_j \left\{ (\varphi_j, \varphi_i) - (\mathcal{K}\varphi_j, \varphi_i) \right\} = (y, \varphi_i), \qquad i = 1, \dots, N.$$

We will assume that this is easily solvable, although in practice the size of N may necessitate an iterative method of solution. The conditioning of this linear system is examined in [5, p. 79] and [32]. The iterated Galerkin solution x_h^* is obtained by substituting (2.1) into the definition (1.5).

To solve (2.2) and (1.5) in practice, usually we must numerically evaluate the integrals that occur in these formulas. Thus, introduce the numerical integration scheme

$$(2.3) \qquad \int_D f(t)\, d\sigma(t) \doteq \sum_{k=1}^{R} w_{kR} f(t_{kR}),$$

with all $t_{kR} \in D$ and all $w_{kR} \neq 0$. Here $R = R_h$ is the number of node points; and we assume that the numerical integration scheme converges as $R \to \infty$ ($h \to 0$), for all $f \in C(D)$. Ordinarily, the weights and nodes will be written simply as w_k and t_k, with the dependence on R (or h) understood implicitly. In all cases when using (2.3) in this paper, we will assume that

$$(2.4) \qquad\qquad [H1] \qquad R_h \geqslant N_h, \, h > 0.$$

This will be needed for reasons that will become apparent later.

Use (2.3) to approximate all of the integrals in (2.2), applying it twice to evaluate the iterated integrals of $(\mathcal{K}\varphi_j, \varphi_i)$. Let

$$(2.5) \qquad z_h(s) = \sum_{j=1}^{N} \beta_j \varphi_j(s), \qquad s \in D,$$

with $\beta_1, \dots, \beta_N$ determined from the linear system

$$(2.6) \qquad \sum_{j=1}^{N} \beta_j \left\{ \sum_{k=1}^{R} w_k \varphi_j(t_k) \varphi_i(t_k) - \sum_{k=1}^{R} \sum_{l=1}^{R} w_k w_l K(t_k, t_l) \varphi_j(t_l) \varphi_i(t_k) \right\}$$
$$= \sum_{k=1}^{R} w_k y(t_k) \varphi_i(t_k), \qquad i = 1, \dots, N.$$

Note that the coefficients (φ_i, φ_j) are also being approximated, as this is necessary for some of our later applications. In the earlier works [12], [30], these coefficients were assumed to be evaluated exactly; and in some cases, our work will imply the earlier results. To complete our approximation of the Galerkin method, define the

discrete iterated Galerkin solution by

$$(2.7) \qquad z_h^*(s) = y(s) + \sum_{k=1}^{R} w_k K(s, t_k) z_h(t_k), \qquad s \in D.$$

To simplify the analysis of z_h and z_h^*, introduce the following matrix notation. Let Φ be the matrix of order $N \times R$ with

$$(2.8) \qquad \Phi_{ik} = \varphi_i(t_k), \qquad i = 1, \ldots, N, \ k = 1, \ldots, R.$$

Also, define

$$(2.9) \qquad \begin{aligned} \tilde{K} &= [K(t_k, t_l)], \qquad k, l = 1, \ldots, R, \\ W &= \text{diagonal}[w_1, \ldots, w_R], \end{aligned}$$

$$(2.10) \qquad \begin{aligned} \mathbf{y} &= [y(t_1), \ldots, y(t_R)]^T, \qquad \boldsymbol{\beta} = [\beta_1, \ldots, \beta_N]^T, \\ \mathbf{z}_h &= [z_h(t_1), \ldots, z_h(t_R)]^T, \qquad \mathbf{z}_h^* = [z_h^*(t_1), \ldots, z_h^*(t_R)]^T. \end{aligned}$$

The linear system (2.6) can be written as

$$(2.11) \qquad \Phi W \Phi^T \boldsymbol{\beta} - \Phi W \tilde{K} W \Phi^T \boldsymbol{\beta} = \Phi W \mathbf{y},$$

and

$$(2.12) \qquad \mathbf{z}_h = \Phi^T \boldsymbol{\beta}.$$

The iterated solution z_h^* satisfies

$$(2.13) \qquad \mathbf{z}_h^* = \mathbf{y} + \tilde{K} W \mathbf{z}_h.$$

With this, we have

LEMMA 2.1. $\Phi W \mathbf{z}_h = \Phi W \mathbf{z}_h^*$.

Proof. Multiply (2.13) by ΦW, yielding

$$\Phi W \mathbf{z}_h^* = \Phi W \mathbf{y} + \Phi W \tilde{K} W \mathbf{z}_h.$$

Combining (2.12) and (2.11),

$$\Phi W \mathbf{z}_h = \Phi W \mathbf{y} + \Phi W \tilde{K} W \mathbf{z}_h.$$

This shows the desired result. $\square$

For the remainder of the paper, we will assume

$$(2.14) \qquad [\text{H2}] \qquad \text{Rank}(\Phi) = N, \qquad h > 0.$$

The lemma leads directly to our first important result.

THEOREM 2.2. *Let $N = R$, and suppose that the system (2.11) has a solution z_h. Then the iterate z_h^* is a solution to the Nyström method for solving (1.1), with (2.3) as the numerical integration method:*

$$(2.15) \qquad z_h^*(s) - \sum_{k=1}^{R} w_k K(s, t_k) z_h^*(t_k) = y(s), \qquad s \in D.$$

Proof. From the assumption $N = R$, Φ is a nonsingular square matrix. Then from Lemma 2.1, $\mathbf{z}_h = \mathbf{z}_h^*$. Combine this with the definition (2.7) to obtain (2.15), as desired. $\square$

This theorem has consequences that will be explored in the following section. We first say something further about the matrix Φ. For $R \geqslant N$, the matrix

$$(2.16) \qquad\qquad G_h = \Phi W \Phi^T$$

is the discretization of the Gram matrix of the original Galerkin system (2.2),

$$(2.17) \qquad\qquad \Gamma_h = \left[(\varphi_i, \varphi_j) \right].$$

The discrete Gram matrix G_h is not necessarily an approximation of Γ_h, as will be shown in the next section with an example, where S_h is a space of periodic piecewise-linear functions. This means that Eq. (2.6) need not be an approximation to (2.2) in order to assure convergence of z_h (or z_h^*) to x. The theory developed in Sections 4 and 5 to explain this convergence differs from the customary approach via perturbation theory, as given in [12], [30]. In certain cases $G_h = \Gamma_h$, and then our discrete Galerkin method (2.6) becomes the same as that analyzed earlier. Such cases are considered in Sections 3 and 7.

To aid in further understanding the meaning of the matrix Φ, we give the following results.

LEMMA 2.3. *Let $N = R$. Then* [H2] *is true if and only if for every set of data* $\{ b_1, \ldots, b_N \}$, *there exists a unique element* $\varphi \in S_h$ *with*

$$(2.18) \qquad\qquad \varphi(t_i) = b_i, \qquad i = 1, \ldots, N.$$

In fact, we have that the mapping $f \in C(D) \to \varphi \in S_h$, with $\varphi(s)$ interpolating $f(s)$ at the nodes $\{ t_j \}$, is a bounded linear projection operator on $C(D)$.

Proof. This is straightforward and well known. Also, this interpolation property is well known to be true for many pairs S_h, $\{ t_j \}$, and thus [H2] is easily checked by considering the equivalent interpolation problem (2.18). □

Remark. Let $N = R$. Then

$$(2.19) \qquad\qquad \Phi W \Phi^T = I \quad \text{if and only if } \Phi^T \Phi W = I.$$

Proof. Again, the proof is straightforward and we omit it. □

3. Applications of the Discrete Galerkin-Nyström Method. In this section we consider only the case $N_h = R_h$, $h > 0$, with the resultant Theorem 2.2. When the discrete Galerkin solution z_h exists, the iterated discrete Galerkin solution z_h^* is simply a solution of the Nyström method for solving (1.1), with (2.3) as the integration method. Therefore, the well-developed convergence theory for the Nyström method can be used to show the convergence of the discrete Galerkin method, even in the heretofore disregarded event of the Gram matrix Γ_h not being computed exactly.

The Nyström approximation method is

$$(3.1) \qquad\qquad \zeta_h(s) - \sum_{l=1}^{N} w_l K(s, t_l) \zeta_h(t_l) = y(s), \qquad s \in D.$$

With the assumptions that (i) Eq. (1.1) is uniquely solvable on $C(D)$, (ii) $K(s, t)$ is continuous for s, $t \in D$, and (iii) the numerical integration method (2.3) is convergent on $C(D)$, it is well known that (3.1) is uniquely solvable in $C(D)$ for all sufficiently small h, say $h \leqslant h_0$. For a development of this theory, see [5, p. 88].

From the error analysis of (3.1),

$$(3.2) \qquad x - \zeta_h = (I - \mathscr{K}_h)^{-1}(\mathscr{K}x - \mathscr{K}_h x), \qquad h \leqslant h_0,$$

where $\mathscr{K}_h$ denotes the numerical integral operator in (3.1). Because of Theorem 2.2, we will henceforth identify ζ_h with z_h^*.

The Nyström method also gives a justification of the discrete Galerkin method when $N = R$. To show that the system in (2.11) is nonsingular, write it as

$$(3.3) \qquad \Phi W \Phi^T - \Phi W \tilde{K} W \Phi^T = \Phi W [I - \tilde{K} W] \Phi^T.$$

The matrix $I - \tilde{K} W$ arises when solving (3.1), and it is known that

$$(3.4) \qquad \left\| (I - \tilde{K}W)^{-1} \right\| \leqslant \left\| (I - \mathscr{K}_h)^{-1} \right\|, \qquad h \leqslant h_0.$$

The matrix norm on the left side is the matrix row norm; see [5, p. 105]. Since Φ, W, and Φ^T are also nonsingular, this shows that the system (3.1) is nonsingular; and the condition number of the matrix on the left side of (3.3) can be found from information on Φ, G_h, and $I - \tilde{K}W$.

The result (3.2) says that the choice of S_h is not important in the rate of convergence of z_h^* to x, provided that [H2] is satisfied, along with $N = R$. Also, the Gram matrix Γ_h does not have to be evaluated exactly in order to obtain convergence.

Application 1. Let $n > 0$, $D = [a, b]$, $h = (b - a)/n$, $s_j = a + jh$ for $0 \leqslant j \leqslant n$. Let $S_h = \mathscr{P}_{r-1}$, the set of functions that are polynomials of degree $< r$ on each subinterval $[s_{j-1}, s_j]$, $j = 1, \ldots, n$. The functions in $\mathscr{P}_{r-1}$ need not be continuous at the node points s_j. The dimension of S_h is

$$(3.5) \qquad N_h = rn.$$

On each subinterval $[s_{j-1}, s_j]$, let an integration scheme be given:

$$(3.6) \qquad \int_{s_{j-1}}^{s_j} f(t)\, dt \doteq \sum_{k=1}^{r} \nu_{kj} f(\tau_{kj}),$$

with $s_{j-1} \leqslant \tau_{1j} < \cdots < \tau_{rj} \leqslant s_j$. Assume that (3.6) has degree of precision d on each subinterval $[s_{j-1}, s_j]$. The formula (3.6) leads naturally to an integration formula over all of $[a, b]$, and the number of integration nodes will be $R = rn$. (Note: Because of the discontinuous nature of the functions in S_h, we can allow $\tau_{1j} = s_{j-1}$ and $\tau_{rj} = s_j$, while still considering τ_{rj} and $\tau_{1, j+1}$ as distinct nodes.)

The space S_h is not contained in $C[a, b]$, so the analysis will be extended to allow functions which can be considered as continuous on each subinterval $[s_{j-1}, s_j]$. This can be done in several ways, as is pointed out in [9]. We will use a formulation using $L^\infty[a, b]$, with the point functional evaluations defined as in Section 2 of the cited paper. With this, we have the following theorem.

THEOREM 3.1. *Assume* (1.1) *is uniquely solvable on* $C[a, b]$. *Further assume that* $x \in C^{d+1}[a, b]$ *and that* $K(s, t)$ *is* $d + 1$ *times differentiable with respect to* t. *Then the iterated discrete Galerkin solution* z_h^* *satisfies*

$$(3.7) \qquad \left\| x - z_h^* \right\|_\infty = O(h^{d+1}).$$

If $d \geqslant 2r - 2$, *then* $\Gamma_h = G_h$. *For a maximal order of convergence, choose* (3.6) *to be Gauss-Legendre quadrature on* $[s_{j-1}, s_j]$, $j = 1, \ldots, n$. *Then* $d = 2r - 1$ *and*

$$(3.8) \qquad \left\| x - z_h^* \right\|_\infty = O(h^{2r}).$$

Proof. These results are a straightforward use of (3.2), together with easily derived error bounds for the numerical integration in (3.6). The fact that $G_h = \Gamma_h$ when $d \geqslant 2r - 2$ follows from the fact that the integrand in

$$(\varphi_l, \varphi_k) = \sum_{j=1}^{n} \int_{s_{j-1}}^{s_j} \varphi_l(t) \varphi_k(t)\, dt$$

is a polynomial of degree $\leqslant 2r - 2$ on each subinterval. The result (3.8) is the same as that given in [12, Theorem 3.6]. $\square$

Application 2. Let D be a piecewise-smooth surface in $\mathbf{R}^3$, with

(3.9) $$D = D_1 \cup \cdots \cup D_J.$$

For each D_j, assume the existence of a smooth mapping

(3.10) $$F_j \colon \hat{D}_j \to D_j, \qquad j = 1, \ldots, J,$$

where $\hat{D}_j$ is a polygonal region in the plane. For each such region $\hat{D}_j$, let $\{\hat{\Delta}_{k,j}\}$ be a triangulation of $\hat{D}_j$, and let $F_j(\hat{\Delta}_{k,j}) = \Delta_{k,j}$ define a corresponding triangulation of D_j. Collecting together all these triangulations of $D_1, \ldots, D_J$ one has a triangulation $\{\Delta_1, \ldots, \Delta_n\}$ *of the surface* D.

To have a standard means of defining approximations and numerical integration on $\{\Delta_k\}$, we introduce an alternative way of defining Δ_k. For each $\hat{\Delta}_k$, let $\hat{v}_{1,k}$, $\hat{v}_{2,k}$, and $\hat{v}_{3,k}$ be its three vertices. Also let $\hat{e}$ be the unit simplex in the plane:

(3.11) $$\hat{e} = \{(\hat{s}, \hat{t}) \mid \hat{s}, \hat{t}, 1 - \hat{s} - \hat{t} \geqslant 0\}.$$

Define $m_k \colon \hat{e} \overset{1-1}{\underset{\text{onto}}{\to}} \Delta_k$ by

(3.12) $$m_k(\hat{s}, \hat{t}) = F_j(\hat{u}\hat{v}_{1,k} + \hat{t}\hat{v}_{2,k} + \hat{s}\hat{v}_{3,k}), \qquad (\hat{s}, \hat{t}) \in \hat{e}, \; \hat{u} = 1 - \hat{s} - \hat{t},$$

where $\hat{\Delta}_k \subset \hat{D}_j$. We define S_h as the set of functions that are images under m_k of polynomials of degree $\leqslant r$ on $\hat{e}$, for $k = 1, \ldots, n$. Again the space S_h will not be contained in the continuous functions, and the analysis must be extended to $L^\infty(D)$ in order to carry out an error analysis.

For numerical integration, let

(3.13) $$\int_{\hat{e}} f(\hat{s}, \hat{t})\, d\sigma \doteq \sum_{j=1}^{\mathscr{J}} \hat{w}_j f(\hat{s}_j, \hat{t}_j).$$

Then for integrals over Δ_k, use

(3.14)
$$\int_{\Delta_k} f(Q)\, dS(Q) = \int_{\hat{e}} f(m_k(\hat{s}, \hat{t})) |D_s m_k \times D_t m_k|\, d\sigma$$
$$\doteq \sum_{j=1}^{\mathscr{J}} w_{kj} f(m_k(\hat{s}_j, \hat{t}_j))$$

with $w_{kj} = \hat{w}_j |D_s m_k \times D_t m_k(\hat{s}_j, \hat{t}_j)|$. D_s denotes the partial derivative with respect to s, and similarly for D_t. For more information on both the triangulation of D and the numerical integration (3.14), see [7], [8].

The dimension of S_h is

(3.15) $$N = n(r + 1)(r + 2)/2;$$

and the total number of integration nodes is

(3.16) $$R = n\mathscr{J}.$$

Thus to have $N = R$, we must choose

$$(3.17) \qquad \mathscr{I} = \frac{(r+1)(r+2)}{2}.$$

Subject to this, we choose (3.13) to maximize the degree of precision d. For formulas chosen from [24], [31], we have the possibilities shown in Table 1.

Assuming (i) an integration rule of degree of precision d, (ii) sufficient smoothness of the functions $K(s, t)$ and $x(t)$ on each D_j, and (iii) sufficient smoothness of each D_j (by means of the mappings F_j), we obtain from (3.2) that

$$(3.18) \qquad \|x - z_h^*\|_\infty = O(h^{d+1}),$$

where h denotes the maximum of the diameters of the triangles $\hat{\Delta}_j$ making up the triangulation of $\hat{D}_1, \ldots, \hat{D}_J$. This is a relatively straightforward argument, and we refer the reader to [7].

TABLE 1

Maximal degree integration rules

r	0	1	2	3
$\mathscr{I}$	1	3	6	10
d	1	2	4	6

We note that, in general, $G_h \neq \Gamma_h$, because of the presence of the Jacobian $|D_s m_k \times D_t m_k|$ in the integrand of (3.14). Nonetheless, what is referred to as superconvergence is still attained. With approximations in S_h of degree $\leqslant r$, one would ordinarily have

$$(3.19) \qquad \|x - x_h\|_\infty = O(h^{r+1})$$

for the Galerkin approximation and

$$(3.20) \qquad \|x - x_h^*\|_\infty = O(h^{2r+2})$$

for the iterated Galerkin approximation. According to Table 1 and (3.18), we do not quite attain this order of superconvergence with z_h^*, although there is an improvement over (3.19). More accurate numerical integration will be needed to replicate (3.20), and this will be returned to in Section 7.

Application 3. We consider two different approximating families S_h while using the same integration rule, thus arriving at the same iterated discrete Galerkin method. Let D be a smooth simple closed curve in the plane. Then $C(D)$ can be replaced by $C_p[0, 2\pi]$, the space of continuous periodic functions on $[0, 2\pi]$. As our first choice of S_h, use $S_h = \mathscr{T}_n$, the set of trigonometric polynomials of degree $\leqslant n$,

$$\varphi(s) = a_0 + \sum_{j=1}^{n} \left[a_j \cos(js) + b_j \sin(js) \right].$$

Then the dimension of S_h is $N = 2n + 1$. For the second choice of S_h, let $S_h = \mathscr{L}_h$, the set of continuous piecewise-linear periodic functions on $[0, 2\pi]$, with each $\varphi \in \mathscr{L}_h$ linear on $[t_{j-1}, t_j]$, $j = 1, \ldots, 2n + 1$,

$$(3.21) \qquad t_j = jh, \qquad 0 \leqslant j \leqslant 2n, \qquad h = 2\pi/(2n + 1),$$

$$t_{2n+1} = t_0, \qquad t_{2n+2} = t_1.$$

The dimension of $\mathscr{L}_h$ is also $N = 2n + 1$.

For the integration rule, use the trapezoidal rule with the nodes in (3.21). Then the number of distinct nodes is only $R = 2n + 1$ because of the periodicity of the functions being integrated. This is a natural integration rule to use when $S_h = \mathcal{T}_n$, because the integration rule then has degree of precision $d = 2n$ and $G_h = \Gamma_h$. This would seem to preserve the rapid convergence of the Galerkin method with $S_h = \mathcal{T}_h$. For the space $S_h = \mathcal{L}_h$, however, we will not have $G_h = \Gamma_h$. For a basis of $\mathcal{L}_h$, use the standard 'hat' functions:

$$\varphi_j(s) = \begin{cases} (s - t_{j-1})/h, & t_{j-1} \leqslant s \leqslant t_j, \\ (t_{j+1} - s)/h, & t_j \leqslant s \leqslant t_{j+1}, \; j = 1, \ldots, 2n + 1, \\ 0, & \text{otherwise.} \end{cases}$$

Then Γ_h is almost tridiagonal (it is circulant); and $\Phi = I$, $G_h = hI$. G_h is not a good approximation of Γ_h.

According to the standard theory of Galerkin's method, one would expect

$$\|x - x_h\|_\infty = O(h^2), \qquad \|x - x_h^*\|_\infty = O(h^4)$$

with $S_h = \mathcal{L}_h$. But according to Theorem 2.2 and formula (3.2), we need only consider the integration error in examining the error in z_h^*, and it will coincide with the result using the more sophisticated approximation from $\mathcal{T}_n$. It is well known that for sufficiently smooth periodic integrands, the trapezoidal rule converges very rapidly; see [6, p. 253] and [16, p. 314]. Thus $\|x - z_h^*\|_\infty$ will converge very rapidly, regardless of which underlying space S_h is being used. When going back to z_h, however, this greater speed will be lost for approximants from $\mathcal{L}_h$, except at the node points (since $z_h(t_i) = z_h^*(t_i)$).

4. The Discrete Orthogonal Projection. We will introduce a discrete analogue to the orthogonal projection P_h of $L^2(D)$ onto S_h. Using this discrete projection, we will give an error analysis in Section 5 for the discrete Galerkin method when $R \geqslant N$. In this section the discrete projection is defined and its properties are examined. Examples with important subspaces S_h will be given in Sections 6 and 7.

Using the numerical integration method (2.3), define

$$(4.1) \qquad (f, g)_h = \sum_{k=1}^{R} w_k f(t_k) g(t_k), \qquad f, g \in C(D),$$

$$(4.2) \qquad \|f\|_h = \sqrt{(f, f)_h}, \qquad f \in C(D).$$

The latter is only a seminorm on $C(D)$, but we will henceforth assume it is a norm on S_h:

$$(4.3) \qquad [\text{H3}] \qquad f \in S_h \quad \text{and} \quad \|f\|_h = 0 \quad \text{implies } f = 0.$$

We also assume

$$(4.4) \qquad [\text{H4}] \qquad w_j > 0, \qquad j = 1, \ldots, R.$$

With [H4], it is usually straightforward to prove [H3]. In particular, we have the following easily proven result.

LEMMA 4.1. *Let* [H4] *hold, and, in addition, assume that functions from* S_h *possess the interpolating property* (2.18) *with respect to a subset of* N *of the integration nodes used in* (4.1). *Then* [H3] *holds.* □

With the definitions (4.1) and (4.2) and using [H4], it is easy to show that

$$(4.5) \qquad |(f,g)_h| \leq \|f\|_h \|g\|_h, \qquad f, g \in C(D),$$

$$(4.6) \qquad \|f\|_h \leq c\|f\|_\infty, \qquad c = \left[\sum_1^R w_k \right]^{1/2}, \qquad f \in C(D).$$

Since the integration rule is convergent on $C(D)$, the constant c can be bounded independent of h.

Define the projection operator $Q_h: C(D) \to S_h$ by

$$(4.7) \qquad (Q_h f, \psi)_h = (f, \psi)_h \quad \text{all } \psi \in S_h.$$

To see that such a $Q_h f$ is uniquely defined, let $S_h = \text{Span}\{\varphi_1, \ldots, \varphi_N\}$ and write

$$(4.8) \qquad Q_h f(s) = \varphi(s) = \sum_{j=1}^N \alpha_j \varphi_j(s), \qquad s \in D.$$

To satisfy (4.7), it is necessary and sufficient that $\alpha_1, \ldots, \alpha_N$ satisfy the linear system

$$(4.9) \qquad \sum_{j=1}^N \alpha_j (\varphi_j, \varphi_i)_h = (f, \varphi_i)_h, \qquad i = 1, \ldots, N.$$

The coefficient matrix is a Gram matrix relative to the inner product (4.1), and it is also the matrix G_h of (2.16). To show G_h is nonsingular, we use the standard arguments to show it is positive definite, with the aid of [H3]. Thus (4.9) is uniquely solvable and $Q_h f$ is uniquely defined. The linearity of (4.9), relative to f, will also show Q_h is linear; and the uniqueness of $Q_h f$ shows Q_h is a projection, i.e., $Q_h^2 = Q_h$.

As a first indication of the usefulness of Q_h, we give the following result.

LEMMA 4.2. *Let z_h and z_h^* be the discrete Galerkin and discrete iterated Galerkin solutions, from (2.5)–(2.7), assuming they exist. Then*

$$(4.10) \qquad z_h = Q_h z_h^*,$$

in analogy with (1.11) for the original Galerkin method.

Proof. Recall Lemma 2.1, that $\Phi W z_h = \Phi W z_h^*$.

Equating corresponding elements,

$$\sum_{k=1}^R w_k \varphi_i(t_k) z_h(t_k) = \sum_{k=1}^R w_k \varphi_i(t_k) z_h^*(t_k), \qquad i = 1, \ldots, N,$$

$$(z_h, \varphi_i)_h = (z_h^*, \varphi_i)_h, \qquad i = 1, \ldots, N.$$

Since $\{\varphi_1, \ldots, \varphi_N\}$ is a basis of S_h, this says

$$(z_h, \psi)_h = (z_h^*, \psi)_h \quad \text{all } \psi \in S_h.$$

From $z_h \in S_h$, the uniqueness of $Q_h z_h^*$ combined with this latter result proves (4.10). $\square$

Some additional properties of Q_h are given in the following

LEMMA 4.3. (a) *Q_h is selfadjoint on $C(D)$, relative to the discrete inner product (4.1):*

$$(4.11) \qquad (Q_h f, g)_h = (f, Q_h g)_h, \qquad f, g \in C(D).$$

(b)

$$(4.12) \qquad \|Q_h f\|_h \leqslant \|f\|_h, \qquad f \in C(D).$$

(c) *If the family $\{Q_h: h > 0\}$ is uniformly bounded on $C(D)$, using the usual operator norm induced by the uniform norm $\|\cdot\|_\infty$, then*

$$(4.13) \qquad \|f - Q_h f\|_\infty \leqslant c \cdot \underset{\varphi \in S_h}{\text{minimum}} \|f - \varphi\|_\infty,$$

with c independent of h and f.

(d) *When $N = R$, the projection Q_h is the interpolating projection from $C(D)$ onto S_h, i.e.,*

$$(4.14) \qquad (Q_h f)(t_i) = f(t_i), \qquad i = 1, \ldots, N.$$

Proof. (a) From (4.7) with $\psi = Q_h g$,

$$(Q_h f, Q_h g)_h = (f, Q_h g)_h.$$

Similarly,

$$(Q_h g, Q_h f)_h = (g, Q_h f)_h.$$

Using the symmetry of the inner product, result (4.11) is proven.

(b)

$$\|Q_h f\|_h^2 = (Q_h f, Q_h f)_h$$
$$= (f, Q_h f)_h, \quad \text{using } \psi = Q_h f \text{ in (4.7)}$$
$$\leqslant \|f\|_h \|Q_h f\|_h, \quad \text{using (4.5).}$$

If $Q_h f \neq 0$, cancel $\|Q_h f\|_h$ to prove (4.12). If $Q_h f = 0$, then (4.12) is trivially true.

(c) Let

$$(4.15) \qquad q = \underset{h > 0}{\text{Supremum}} \|Q_h\| < \infty.$$

Then for any $\varphi \in S_h$,

$$(4.16) \qquad f - Q_h f = f - \varphi + Q_h \varphi - Q_h f = (I - Q_h)(f - \varphi),$$

$$(4.17) \qquad \|f - Q_h f\|_\infty \leqslant (1 + q)\|f - \varphi\|_\infty.$$

Let $c = 1 + q$, and form the minimum of the right side over S_h to prove (4.13).

(d) The system (4.9), defining $Q_h f$, can be written as

$$G_h \alpha = \Phi W \mathbf{f}, \qquad \mathbf{f} = [f(t_1), \ldots, f(t_N)]^T.$$

Using (2.16) for G_h,

$$\Phi W \Phi^T \alpha = \Phi W \mathbf{f}.$$

Since Φ and W are nonsingular, we have

$$\Phi^T \alpha = \mathbf{f}.$$

But from (4.8), with $s = t_i$, we have

$$Q_h f(t_i) = (\Phi^T \alpha)_i = (\mathbf{f})_i = f(t_i), \qquad i = 1, \ldots, N,$$

and thus the result is proven. □

Part (c) shows that $Q_h f \to f$ in $C(D)$ if (i) the family $\{Q_h: h > 0\}$ is uniformly bounded on $C(D)$, and (ii) the family $\{S_h: h > 0\}$ approximates all elements of $C(D)$. Conversely, if $Q_h f \to f$ for all $f \in C(D)$, then $\{Q_h\}$ is uniformly bounded

by the principle of uniform boundedness. We examine next the general problem of bounding the family $\{Q_h\}$ for some important approximating families $\{S_h\}$.

In general, we need to bound $Q_h f$ in (4.8), subject to $\alpha_1, \ldots, \alpha_N$ satisfying (4.9). The choice of the basis $\{\varphi_1, \ldots, \varphi_N\}$ of S_h is at our disposal. Since the coefficient matrix of (4.8) is G_h, a bound for Q_h can be produced by using

$$
\|\alpha\|_\infty \leqslant \|G_h^{-1}\| \underset{1 \leqslant i \leqslant N}{\mathrm{Max}} |(f, \varphi_i)_h|
$$

(4.18)

$$
\leqslant \|G_h^{-1}\| \|f\|_\infty \underset{1 \leqslant i \leqslant N}{\mathrm{Max}} \sum_{k=1}^{N} w_k |\varphi_i(t_k)|
$$

with $\alpha = [\alpha_1, \ldots, \alpha_N]^T$. The matrix norm used in $\|G_h^{-1}\|$ is the row norm. Combining (4.8) and (4.18),

(4.19)
$$
\|Q_h f\|_\infty \leqslant c \|f\|_\infty,
$$

(4.20)
$$
c = \|G_h^{-1}\| \left[\underset{s \in D}{\mathrm{Max}} \sum_{i=1}^{N} |\varphi_i(s)| \right] \left[\underset{1 \leqslant i \leqslant N}{\mathrm{Max}} \sum_{k=1}^{R} w_k |\varphi_i(t_k)| \right].
$$

Thus the problem of bounding Q_h is reduced to that of bounding G_h^{-1}, at least in the case when the underlying subspaces S_h are of finite element type. For then,

$$
\underset{s \in D}{\mathrm{Max}} \sum_{i=1}^{N} |\varphi_i(s)|
$$

can be reasonably assumed to be bounded independent of h; while for the last term in (4.20), an estimate

$$
\underset{1 \leqslant i \leqslant N}{\mathrm{Max}} \sum_{k=1}^{R} w_k |\varphi_i(t_k)| \leqslant c h^\delta
$$

could be established.

Bounds on $\|\Gamma_h^{-1}\|$ were studied in [10], [18] in case of spline spaces S_h, and in [17], [19] for finite element spaces over subregions in $\mathbf{R}^n$. The possibility of defining the discrete inner product and the corresponding projection was mentioned in the latter paper; and the discrete projection onto spline spaces in conjunction with collocation methods at Gaussian points was used in [21], [22]. Unlike the latter two papers and aforementioned works [12], [30], we do not require $G_h = \Gamma_h$ in order to establish convergence $\|x - z_h^*\| \to 0$.

Anticipating the superconvergence results for the discrete iterate z_h^*, we next show that the discrete Galerkin solution z_h inherits certain superconvergence properties of z_h^*.

LEMMA 4.4. *Let z_h and z_h^* be defined as in Lemma 4.2. Then*

(4.21)
$$
\|Q_h x - z_h\|_h \leqslant c \|x - z_h^*\|_\infty,
$$

where c is the constant in (4.6).

Proof. Using consecutively (4.10), (4.12), and (4.6), one has

$$
\|Q_h x - z_h\|_h = \|Q_h(x - z_h^*)\|_h \leqslant \|x - z_h^*\|_h \leqslant c \|x - z_h^*\|_\infty. \quad \square
$$

Assuming (4.19), we state a stronger result.

LEMMA 4.5. *Assume* (4.19) *holds, i.e.,* $\|Q_h f\|_\infty \leqslant c\|f\|_\infty$ *for all* $f \in C(D)$ *with* c *independent of* h. *Then*

$$(4.22) \qquad\qquad \|Q_h x - z_h\|_\infty \leqslant c\|x - z_h^*\|_\infty.$$

Proof. In (4.19), set $f = x - z_h^*$; and then use

$$\|Q_h x - z_h\|_\infty = \|Q_h(x - z_h^*)\|_\infty \leqslant c\|x - z_h^*\|_\infty. \qquad \Box$$

With $N = R$, Q_h is just the interpolating projection. Thus (4.22) extends the result of Richter [26], on the superconvergence of $(x - x_h)(t_i)$ for collocation at Gaussian points, to the case of the discrete Galerkin method.

Thus we see that the convergence of the discrete Galerkin method hinges on availability of the bound (4.19). This will be discussed in Sections 6 and 7. But first we introduce the formalism for handling the fully discretized Galerkin method.

5. General Error Analysis. Recall the numerical integral operator $\mathcal{K}_h$ based on the integration rule (2.3),

$$(5.1) \qquad\qquad (\mathcal{K}_h f)(s) = \sum_{j=1}^{R} w_j K(s, t_j) f(t_j), \qquad s \in D.$$

Using the notion of the discrete inner product (4.1), we also have

$$(5.2) \qquad\qquad (\mathcal{K}_h f)(s) = (k_s, f)_h,$$

where $k_s = K(s, \cdot)$.

LEMMA 5.1. *The discrete Galerkin method* (2.5)–(2.6) *can be written equivalently as*

$$(5.3) \qquad\qquad (I - Q_h \mathcal{K}_h) z_h = Q_h y, \qquad z_h \in C(D).$$

Moreover, the discrete iterated Galerkin solution satisfies

$$(5.4) \qquad\qquad (I - \mathcal{K}_h Q_h) z_h^* = y.$$

Proof. From (5.3), $z_h = Q_h(y + \mathcal{K}_h z_h) \in S_h$. Using this, rewrite (5.3) as

$$Q_h[(I - \mathcal{K}_h) z_h - y] = 0, \qquad z_h \in S_h.$$

This is equivalent to saying that $z_h \in S_h$ and that it satisfies

$$((I - \mathcal{K}_h) z_h - y, \psi)_h = 0 \quad \text{all } \psi \in S_h.$$

Choose a basis $\{\varphi_1, \ldots, \varphi_N\}$ of S_h, and successively set $\psi = \varphi_i$, $i = 1, \ldots, N$. This leads directly to the formulation (2.5)–(2.6), showing the equivalence with the formula (5.3).

As to z_h^*, the definition (2.7) is equivalent to

$$(5.5) \qquad\qquad z_h^* = y + \mathcal{K}_h z_h,$$

from which $Q_h z_h^* = Q_h y + Q_h \mathcal{K}_h z_h = z_h$, supplying another proof of Lemma 4.2. Now replacing z_h with $Q_h z_h^*$ in (5.5), we prove (5.4). $\Box$

Solvability of the two equations (5.3) and (5.4) follows easily from the standard theory of collectively compact families of operators, [1], [5].

THEOREM 5.2. *Assume the integral equation* $(I - \mathcal{K})x = y$ *is uniquely solvable for every* $y \in C(D)$, *and assume* $K(s, t)$ *is continuous for* s, $t \in D$. *Let*

$$(5.6) \qquad\qquad \{S_h \subset C(D): h > 0\}$$

be a sequence of approximating subspaces with the property that for each $f \in C(D)$,

$$(5.7) \qquad \rho_h(f) = \min_{\varphi \in S_h} \| f - \varphi \|_\infty \to 0 \quad as\ h \to 0.$$

Finally, assume that the discrete projections $\{ Q_h: h > 0 \}$ are uniformly bounded,

$$(5.8) \qquad q \equiv \sup_{h > 0} \| Q_h \| < \infty.$$

Then $\{ Q_h \mathcal{X}_h \}$ and $\{ \mathcal{X}_h Q_h \}$ each form a pointwise convergent and collectively compact family of operators. Hence, for all sufficiently small h, the operators $I - Q_h \mathcal{X}_h$ and $I - \mathcal{X}_h Q_h$ are invertible on $C(D)$ (for, say, $h \leqslant h_0$), and

$$(5.9) \qquad \left\| (I - Q_h \mathcal{X}_h)^{-1} \right\| \leqslant B, \qquad \left\| (I - \mathcal{X}_h Q_h)^{-1} \right\| \leqslant B < \infty.$$

Moreover,

$$(5.10) \qquad (I - Q_h \mathcal{X}_h)(x - z_h) = (I - Q_h)(y + \mathcal{X}_h x) + (\mathcal{X} - \mathcal{X}_h)x$$

and

$$(5.11) \qquad (I - \mathcal{X}_h Q_h)(x - z_h^*) = \mathcal{X}x - \mathcal{X}_h Q_h x.$$

Thus both $\| x - z_h \|_\infty$ and $\| x - z_h^* \|$ tend to 0 as $h \to 0$.

Proof. We concentrate on the equation (5.4). Equation (5.3) can be dealt with in a similar manner. We show that $\{ \mathcal{X}_h Q_h: h > 0 \}$ is a pointwise convergent and collectively compact family of operators on $C(D)$. Then the general theory of collectively compact operators can be used to complete the proof (e.g., see [1] or [5, p. 96]).

From (5.7), (5.8), and (4.17),

$$(5.12) \qquad \| f - Q_h f \|_\infty \leqslant (1 + q)\rho_h(f),$$

proving that $Q_h f \to f$ as $h \to 0$, for all $f \in C(D)$. Then

$$(5.13) \qquad \mathcal{X}f - \mathcal{X}_h Q_h f = [\mathcal{X}f - \mathcal{X}_h f] + \mathcal{X}_h[f - Q_h f].$$

It is straightforward to show that $\| \mathcal{X}f - \mathcal{X}_h f \|_\infty \to 0$ (see [5, p. 90]); and then the principle of uniform boundedness implies $\{ \mathcal{X}_h \}$ is uniformly bounded. Combining the latter with (5.12), the last term in (5.13) also goes to zero, completing the proof that $\mathcal{X}_h Q_h f \to \mathcal{X}f$.

The collective compactness of $\{ \mathcal{X}_h Q_h: h > 0 \}$ will follow by standard arguments. Combined with the pointwise convergence of $\{ \mathcal{X}_h Q_h \}$, the remaining results follow from known theory [1], [5]. $\square$

We now give another proof of Theorem 2.2, within this new framework.

COROLLARY 5.3. *When $N = R$, the discrete iterated Galerkin solution results from a Nyström approximation of the integral equation (1.1).*

Proof. From Lemma 4.3(d), we have $\mathcal{X}_h Q_h f(s) = \mathcal{X}_h f(s)$ for all $f \in C(D)$. Thus formula (5.4) for z_h^* reduces to

$$(I - \mathcal{X}_h)z_h^* = y,$$

the Nyström method. $\square$

To obtain superconvergence results for z_h^* to x in the original Galerkin method, examine the error term in (5.11). This reduces to examining the error

$$\mathcal{X}x(s) - \mathcal{X}_h Q_h x(s).$$

LEMMA 5.4. *Let* $k_s(t) = K(s, t)$, s, $t \in D$. *Then*

$$(5.14) \quad \mathscr{K}x(s) - \mathscr{K}_h Q_h x(s) = [\mathscr{K}x(s) - \mathscr{K}_h x(s)] + ((I - Q_h)k_s, (I - Q_h)x)_h$$

and

$$(5.15) \quad \begin{aligned} |\mathscr{K}x(s) - \mathscr{K}_h Q_h x(s)| &\leqslant |\mathscr{K}x(s) - \mathscr{K}_h x(s)| \\ &\quad + \|(I - Q_h)k_s\|_{h,1} \cdot \|(I - Q_h)x\|_{h,\infty}. \end{aligned}$$

The seminorms are

$$(5.16) \quad \|f\|_{h,1} = \sum_{j=1}^{R} w_j |f(t_j)|, \qquad \|f\|_{h,\infty} = \underset{1 \leqslant j \leqslant R}{\text{Maximum}} |f(t_j)|.$$

Proof. Write

$$\mathscr{K}x(s) - \mathscr{K}_h Q_h x(s) = [\mathscr{K}x(s) - \mathscr{K}_h x(s)] + \mathscr{K}_h(I - Q_h)x(s).$$

Then note that from (5.2),

$$\begin{aligned} \mathscr{K}_h(I - Q_h)x(s) &= (k_s, (I - Q_h)x)_h = (k_s, (I - Q_h)^2 x)_h \\ &= ((I - Q_h)k_s, (I - Q_h)x)_h. \end{aligned}$$

The last two steps used the facts that $I - Q_h$ is a projection and that it is selfadjoint from Lemma 4.3(a). (5.15) follows easily from (5.14). □

This shows the quantities that must be examined in order to obtain convergence results for z_h^*, namely, (1) $\mathscr{K}x - \mathscr{K}_h x$, (2) $x - Q_h x$, and (3) $(I - Q_h)k_s$. The first is simply the numerical integration error. The second and third are the errors in using the projection of $C(D)$ into S_h, and these are examined for particular cases in the next two sections.

When S_h contains only piecewise-continuous functions, the above arguments can still be used. Note that if $S_h \subset L^\infty(D)$, then $\mathscr{K}: L^\infty(D) \to C(D)$ implies that $\mathscr{K}_h Q_h: C(D) \to C(D)$, and z_h^* will still be a continuous function. The discrete solution z_h will not be continuous, but the following bound in (5.17) is still valid.

Note also that for the difference $x - z_h$ one has

$$(5.17) \quad \|x - z_h\|_\infty = \|x - Q_h z_h^*\|_\infty \leqslant \|x - Q_h x\|_\infty + \|Q_h\| \|x - z_h^*\|_\infty,$$

which provides a way other than (5.10) to estimate $\|x - z_h\|_\infty$. As the first application of the above lemma, consider the simple case of a smooth kernel $K(s, t)$ and a smooth right-hand side y. Assume also that for smooth f,

$$\inf_{\varphi \in S_h} \|f - \varphi\|_\infty = O(h^l);$$

and assume the integration formula used is of degree of precision d. Then under the assumption that $\{Q_h\}$ is uniformly bounded,

$$(5.18) \quad \|x - z_h\|_\infty = O(h^{\min(l, d+1)}),$$

$$(5.19) \quad \|x - z_h^*\|_\infty = O(h^{\min(2l, d+1)}).$$

(5.19) actually coincides with a result by Chandler [12], while (5.18) improves the corresponding result of his in the event we are concerned with, i.e., when the same integration formula is used in defining Q_h and $\mathscr{K}_h$.

While we are on the case of smooth kernels, let us use (5.14) to make a remark that parallels one in [28]. Namely, the rate of convergence observed in (5.19) remains the same for all derivatives of $x - z_h^*$, i.e.,

$$(5.20) \qquad \left\| D^\alpha \left(x - z_h^* \right) \right\|_\infty = O\left(h^{\min(2l, d+1)} \right),$$

where α is a multi-index. This follows from (i) the fact that the uniform boundedness of $(I - \mathscr{K}_h Q_h)^{-1}$ in $C(D)$, $h \leqslant h_\infty$, implies their uniform boundedness in $C^k(D)$, the space of k times continuously differentiable functions, and (ii) the easy formula

$$(5.21) \qquad \left\| D^\alpha \mathscr{K}_h (I - Q_h) x \right\|_\infty = \max_s \left((I - Q_h) k_s^\alpha, (I - Q_h) x \right)_h,$$

where $k_s^\alpha = D_s^\alpha K(s, \cdot)$.

6. Bounds on G_n^{-1}. See the Supplements section at the end of this issue.

7. Applications. In this section, we derive the bound in (4.19)–(4.20), using the results of the preceding section. To define the projection Q_h, we have to specify the spaces S_h and the numerical quadratures (2.3).

We consider finite element spaces as described in Section 6. Starting with an integration formula over the reference element,

$$(7.1) \qquad \int_{\hat{e}} \hat{\varphi}(\hat{t}) \, d\hat{t} \doteq \sum_{j=1}^{\mathscr{J}} \hat{w}_j \hat{\varphi}(\hat{t}_j),$$

we define the composite formula over D,

$$(7.2) \qquad \int_D \varphi(t) \, dt \doteq \sum_{k=1}^{E} \sum_{j=1}^{\mathscr{J}} w_{jk} \varphi(t_{jk}),$$

where $t_{jk} = F_k(\hat{t}_j)$, and $w_{jk} = \hat{w}_j |DF_k(\hat{t}_j)|$, $k = 1, \ldots, E$, $j = 1, \ldots, \mathscr{J}$, in accordance with (3.13)–(3.14).

Using (7.2), we define the scalar product $(\cdot, \cdot)_h$ as in (4.1) and (6.2), with

$$(7.3) \qquad (\varphi, \psi)_{e_k} = \sum_{j=1}^{\mathscr{J}} w_{jk} \varphi(t_{jk}) \psi(t_{jk}).$$

This clearly satisfies (6.10) and (6.11), if the scalar product $(\cdot, \cdot)\hat{}$ is defined by means of

$$(7.4) \qquad (\hat{\varphi}, \hat{\psi})\hat{} = \sum_{j=1}^{\mathscr{J}} \hat{w}_j \hat{\varphi}(\hat{t}_j) \hat{\psi}(\hat{t}_j).$$

To state a general result, we have to impose two additional conditions, in addition to (a)–(j) of Section 6.

(k) F_k is continuously differentiable and
$$\max_{\hat{t} \in \hat{e}} |DF_k(\hat{t})| \leqslant \kappa m(e_k), \qquad k = 1, \ldots, E.$$

In case F_k is an affine mapping as in (6.11), this follows from $|DF_k| = m(e_k)/m(\hat{e})$.

(l) $\qquad$ Let $M_j = \{ k : C_j \supset e_k \}$. Then $|M_j| \leqslant M$, $j = 1, \ldots, N$.

This is easily verified for all commonly used finite element spaces.

We can now state the following general result.

THEOREM 7.1. *Let spaces S_h be defined as in* (a)–(l), *where for definiteness, we set* $h = \max(d(e_k))$, $k = 1, \ldots, E$. *Then with* $(\cdot, \cdot)_h$ *defined as above and Q_h as in* (4.7),

$$(7.5) \qquad\qquad \|Q_h\| \leqslant C,$$

independent of h.

Proof. We have to estimate the three terms in (4.20). First, using (j), we get

$$\operatorname*{Max}_{s \in D} \sum_{i=1}^{N} |\varphi_i(s)| = \operatorname*{Max}_{k=1,\ldots,E} \operatorname*{Max}_{s \in e_k} \sum_{i=1}^{N} |\varphi_i(s)|$$

$$= \operatorname*{Max}_{k=1,\ldots,E} \operatorname*{Max}_{s \in e_k} \sum_{i \in I_k} |\varphi_i(s)|$$

$$(7.6)$$

$$\leqslant M \cdot \operatorname*{Max}_{k=1,\ldots,E} \operatorname*{Max}_{s \in e_k} \operatorname*{Max}_{i \in I_k} |\varphi_i(s)|$$

$$\leqslant M \cdot \operatorname*{Max}_{j=1,\ldots,\nu} \operatorname*{Max}_{\hat{s} \in \hat{e}} |\hat{\varphi}_j(\hat{s})| \leqslant C \cdot M,$$

for a constant C independent of h.

For all $i = 1, \ldots, N$ and $k \in M_i$, denote the progenitor of φ_i over e_k as $\hat{\varphi}_{i_k}$:

$$\hat{\varphi}_{i_k} = \varphi_i|_{e_k} \circ F_k$$

as in (6.14). Using this notation, we next estimate the last factor in (4.20). Thus, for all $i = 1, \ldots, N$, we see that

$$\sum_{k=1}^{R} w_k |\varphi_i(t_k)| = \sum_{k=1}^{E} \sum_{j=1}^{\mathscr{J}} w_{jk} |\varphi_i(t_{jk})|$$

$$= \sum_{k \in M_i} \sum_{j=1}^{\mathscr{J}} w_{jk} |\hat{\varphi}_{i_k}(\hat{t}_j)|, \quad \text{from (j)}$$

$$(7.7)$$

$$\leqslant \kappa \sum_{k \in M_i} \sum_{j=1}^{\mathscr{J}} \hat{w}_j m(e_k) |\hat{\varphi}_{i_k}(\hat{t}_j)|, \quad \text{from (k)}$$

$$\leqslant \kappa \cdot \operatorname*{Max}_{l=1,\ldots,\nu} \sum_{k \in M_i} \sum_{j=1}^{\mathscr{J}} |\hat{\varphi}_l(\hat{t}_j)| \cdot \operatorname*{Max}_{k=1,\ldots,E} m(e_k)$$

$$\leqslant \kappa \cdot M \cdot ch^{\bar{n}} = Ch^{\bar{n}},$$

where $\bar{n} = n$ or $\bar{n} = n - 1$ as in (h), and C is independent of h.

Finally, applying Lemma 6.3, (k), and the same bound on $m(e_k)$ as in (7.7), we obtain through Theorem 6.1 the estimate

$$(7.8) \qquad\qquad \|G_h^{-1}\| \leqslant Ch^{-\bar{n}}$$

with C independent of h.

Combining (4.20), (7.6)–(7.8), we finish the proof. $\square$

Examples. Consider the unit simplex $\hat{e}$, as in (3.11). Let $\hat{v}_1$, $\hat{v}_2$, $\hat{v}_3$ be its vertices while $\hat{v}_{12}$, $\hat{v}_{23}$, $\hat{v}_{13}$ are the midpoints of its sides. First we study the element $\{\hat{e}, \hat{\Sigma}^0, \mathscr{P}_1\}$, where $\mathscr{P}_1$ is the set of polynomials of degree less than 2, and $\hat{\Sigma}^0 = \{\delta_{\hat{v}_1}, \delta_{\hat{v}_2}, \delta_{\hat{v}_3}\}$. The scalar products are defined with

$$(7.9) \qquad\qquad (\hat{\varphi}, \hat{\psi})^{\hat{}} = \frac{1}{6} \sum_{i=1}^{3} \hat{\varphi}(\hat{v}_i) \hat{\psi}(\hat{v}_i).$$

The space S_h consists of piecewise-linear continuous functions, and the discrete iterated Galerkin method reduces to the Nyström method with nodes at vertices of the triangulation.

If one chooses instead $\hat{\Sigma}^0 = \{\delta_{\hat{v}_{12}}, \delta_{\hat{v}_{23}}, \delta_{\hat{v}_{13}}\}$ and replaces (7.9) with

$$(7.10) \qquad (\hat{\varphi}, \hat{\psi})^{\hat{}} = \frac{1}{6} \sum_{\substack{i,\,j=1 \\ i \neq j}}^{3} \hat{\varphi}(\hat{v}_{ij}) \hat{\psi}(\hat{v}_{ij}),$$

then S_h will consist of piecewise-linear functions, continuous at the midpoints of the sides of the triangular elements. The discrete iterated Galerkin method coincides with the Nyström method with nodes at these points. Again, although (7.10) is exact for the finite elements concerned, Q_h does not reduce to the customary orthogonal projection P_h.

It immediately follows from the general theory that, taking operators $\mathscr{K}$ with a smooth kernel, we have $\|x - z_h^*\|_\infty \leqslant ch^2$ in the first case, while $\|x - z_h^*\|_\infty \leqslant ch^3$ in the second case, since the degree of precision of the quadrature in (7.9) is 1, while that of the quadrature in (7.10) is 2. Note that $\|x - x_h^*\|_\infty \leqslant ch^4$, from which we see that, though the discrete iterate z_h^* converges faster than z_h (or for that matter x_h), it does not attain the full superconvergence of x_h^*. As we see, to achieve this, it does not suffice to choose an integration formula for which the integral in (7.4) is evaluated exactly. To obtain the convergence rate $O(h^4)$ for z_h^*, an integration formula with degree of precision $d \geqslant 3$ is needed. From [24], there exists a formula with $d = 3$ that has 4 nodes.

As the last example of finite element spaces, we consider the piecewise-smooth surface D from (3.9). The triangulation is defined by (3.10)–(3.12), while the scalar products $(\cdot, \cdot)_{e_i}$ are defined by means of the integration (3.13)–(3.14). The kernel of the integral operator $\mathscr{K}$ is assumed to be smooth on each of the patches D_j. The same holds for the solution x, provided the right-hand side y is sufficiently smooth.

Theorem 7.1 applies directly; thus we have that (4.19) holds. Convergence estimates easily follow from (5.14), (5.17). The difference $(\mathscr{K} - \mathscr{K}_h)x(s)$ is bounded as in [7], [8]; we have as before

$$\|x - z_h\|_\infty \leqslant c\left(h^{d+1} + h^{r+1}\right) \quad \text{and} \quad \|x - z_h^*\|_\infty \leqslant c\left(h^{d+1} + h^{2(r+1)}\right).$$

To attain the rate of convergence of the exact Galerkin iterate x_h^*, which is $h^{2(r+1)}$, one has to choose an integration formula with

$$(7.11) \qquad d + 1 \geqslant 2(r + 1).$$

Finally, note again that (7.11) does not imply $Q_h = P_h$.

Lastly, we would like to remark on the applicability of the developed theory to the spline spaces. The crucial assumption (j) holds for the spline spaces $S(r, k)$ with $2k \leqslant r - 1$, where

$$S(r, k) = \left\{ f \in C^k[a, b] : f|_{[t_{i-1}, t_i]} \in \mathscr{P}_r \right\}.$$

Here, Δ: $a = t_0 < \cdots < t_m = b$ is a partition of a finite interval $[a, b]$, $\mathscr{P}_r$ is the set of polynomials of degree less than r. When $2k = r - 1$ one has a space of Hermite splines; with $2k < r - 1$ the splines are less continuous than in the latter case.

Thus we can formulate the following

COROLLARY 7.2. *Assume for* $S_h = S(r, k)$ *with* $2k \leqslant r - 1$ *that* (a)–(k) *hold. Then with* $(\cdot, \cdot)_h$ *defined by* (6.2) *and* Q_h *by* (4.7),

$$\|Q_h\| \leqslant C,$$

independent of h. □

Example. Let K be an operator with a smooth kernel. Then using the above-defined spline spaces S_h of Corollary 7.2, we have two estimates for the exact Galerkin and iterated Galerkin methods:

(7.12)
$$\|x - x_h\|_\infty \leqslant ch^r$$

and

(7.13)
$$\|x - x_h^*\|_\infty \leqslant ch^{2r},$$

provided the right-hand side $y \in C'[a, b]$.

If the quadrature in (7.1) has degree of precision d, then from (5.14) we immediately have

(7.14)
$$\|x - z_h^*\|_\infty \leqslant c(h^{d+1} + h^{2r});$$

and combined with (5.17),

(7.15)
$$\|x - z_h\|_\infty \leqslant c(h^{d+1} + h^r).$$

Then using numerical integration maintains the accuracy of the Galerkin method if $d \geqslant r - 1$. To preserve the accuracy of the iterate, we need $d \geqslant 2r - 1$; although with $d > r - 1$, the discrete iterate will exhibit superconvergence. Compare with [12].

Remark. Consider more general spline spaces $S_h = S(r, k)$, with a greater order of smoothness. Assume that the numerical integration (7.1) over the reference element has degree of precision $d \geqslant 2r - 2$. As a basis for S_h, use the normalized B-splines. Then $G_h = \Gamma_h$, because the numerical integration of all inner products is exact,

$$(\varphi, \psi)_h = (\varphi, \psi), \qquad \varphi, \psi \in S_h.$$

Thus, although $Q_h \neq P_h$, the results of de Boor [10] can be used to bound $\|G_h^{-1}\|$, giving the result (7.8) with $\bar{n} = 1$. Also, since B-splines are nonnegative and constitute a partition of unity, (7.6) is verified directly, with

$$\sum_{i=1}^{N} |\varphi_i(s)| = 1, \qquad a \leqslant s \leqslant b.$$

The assumption (l) holds easily, and this implies (7.7). Thus we also have for the general spline spaces $S_h = S(r, k)$ that

$$\|Q_h\| \leqslant C,$$

independent of h, provided the degree of precision d of (7.1) satisfies $d \geqslant 2r - 2$. With this, the inequalities (7.12)–(7.15) also apply to $S_h = S(r, k)$. In (7.15), we need $d \geqslant 2r - 1$ in order to recover the full speed of convergence of the iterated Galerkin method. This was noted previously in [12].

Mathematics Department
University of Iowa
Iowa City, Iowa 52242

1. P. M. ANSELONE, *Collectively Compact Operator Approximation Theory*, Prentice-Hall, Englewood Cliffs, N. J., 1971.

2. U. ASCHER, "Discrete least squares approximations for ordinary differential equations," *SIAM J. Numer. Anal.*, v. 15, 1978, pp. 478–496.

3. K. ATKINSON, "The numerical solution of the eigenvalue problem for compact integral operators," *Trans. Amer. Math. Soc.*, v. 129, 1967, pp. 458–465.

4. K. ATKINSON, "Convergence rates for approximate eigenvalues of compact integral operators," *SIAM J. Numer. Anal.*, v. 12, 1975, pp. 213–222.

5. K. ATKINSON, *A Survey of Numerical Methods for the Solution of Fredholm Integral Equations of the Second Kind*, SIAM, Philadelphia, Pa., 1976.

6. K. ATKINSON, *An Introduction to Numerical Analysis*, Wiley, New York, 1978.

7. K. ATKINSON, "Piecewise polynomial collocation for integral equations on surfaces in three dimensions," *J. Integral Equations*, v. 9, 1985, pp. 25–48.

8. K. ATKINSON, "Solving integral equations on surfaces in space," in *Constructive Methods for the Practical Treatment of Integral Equations* (G. Hämmerlin and K.-H. Hoffmann, eds.), Birkhäuser Verlag, Basel, 1985, pp. 20–43.

9. K. ATKINSON, I. GRAHAM & I. SLOAN, "Piecewise continuous collocation for integral equations," *SIAM J. Numer. Anal.*, v. 20, 1983, pp. 172–186.

10. C. DE BOOR, "A bound on the L_∞-norm of L_2-approximation by splines in terms of a global mesh ratio," *Math. Comp.*, v. 30, 1976, pp. 765–771.

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

12. G. A. CHANDLER, *Superconvergence of Numerical Solutions to Second Kind Integral Equations*, Ph. D. thesis, Australian National University, Canberra, 1979.

13. G. A. CHANDLER, "Superconvergence for second kind integral equations," in *The Application and Numerical Solution of Integral Equations* (R. S. Anderssen, F. de Hoog, and M. Lukas, eds.), Sijthoff & Noordhoff, Groningen, 1980, pp. 103–118.

14. F. CHATELIN, *Spectral Approximation of Linear Operators*, Academic Press, New York, 1984.

15. F. CHATELIN & R. LEBBAR, "Superconvergence results for the iterated projection method applied to a Fredholm integral equation of the second kind and the corresponding eigenvalue problem," *J. Integral Equations*, v. 6, 1984, pp. 71–91.

16. P. J. DAVIS & P. RABINOWITZ, *Methods of Numerical Integration*, 2nd ed., Academic Press, New York, 1984.

17. J. DESCLOUX, "On finite element matrices," *SIAM. J. Numer. Anal.*, v. 9, 1972, pp. 260–265.

18. J. DOUGLAS JR., TODD DUPONT & LARS WAHLBIN, "Optimal L_∞ error estimates for Galerkin approximations to solutions of two-point boundary value problems," *Math. Comp.*, v. 29, 1975, pp. 475–483.

19. J. DOUGLAS JR., TODD DUPONT & LARS WAHLBIN, "The stability in L^q of the L^2-projection into finite element function spaces," *Numer. Math.*, v. 23, 1975, pp. 193–197.

20. R. J. HERBOLD, M. H. SCHULTZ & R. S. VARGA, "The effect of quadrature errors in the numerical solution of boundary value problems by variational techniques," *Aequationes Math.*, v. 3, 1969, pp. 247–270.

21. G. C. HSIAO, P. KOPP & W. L. WENDLAND, "A Galerkin collocation method for some integral equations of the first kind," *Computing*, v. 25, 1980, pp. 89–130.

22. G. C. HSIAO & W. L. WENDLAND, "The Aubin-Nitsche lemma for integral equations," *J. Integral Equations*, v. 3, 1981, pp. 299–315.

23. S. JOE, *The Numerical Solution of Second Kind Fredholm Integral Equations*, Ph. D. thesis, Univ. of New South Wales, Sydney, Australia, 1985.

24. J. N. LYNESS & D. C. JESPERSEN, "Moderate degree symmetric quadrature rules for the triangle," *J. Inst. Math. Appl.*, v. 15, 1975, pp. 19–32.

25. J. E. OSBORN, "Spectral approximation for compact operators," *Math. Comp.*, v. 29, 1975, pp. 712–725.

26. G. R. RICHTER, "Superconvergence of piecewise polynomial Galerkin approximations for Fredholm integral equations of the second kind," *Numer. Math.*, v. 31, 1978, pp. 63–70.

27. L. L. Schumaker, *Spline Functions: Basic Theory*, Wiley, New York, 1981.

28. I. H. Sloan, "Superconvergence and the Galerkin method for integral equations," in *Treatment of Integral Equations by Numerical Methods* (C. T. H. Baker and G. F. Miller, eds.), Academic Press, London, 1982, pp. 197–208.

29. I. H. Sloan & Vidar Thomée, "Superconvergence of the Galerkin iterates for integral equations of the second kind," *J. Integral Equations*, v. 9, 1985, pp. 1–23.

30. A. Spence & K. S. Thomas,"On superconvergence properties of Galerkin's method for compact operator equations," *IMA J. Numer. Anal.*, v. 3, 1983, pp. 253–271.

31. A. H. Stroud, *Approximate Calculation of Multiple Integrals*, Prentice-Hall, Englewood Cliffs, N. J., 1971.

32. R. Whitley, "The stability of finite rank methods with applications to integral equations," *SIAM J. Numer. Anal.*, v. 23, 1986, pp. 118–134.

Added in proof. Results related to the material of Section 2 can be found in the following reference: L. M. Delves & J. Mohamed, *Computational Methods for Integral Equations*, Cambridge University Press, 1985, pp. 208–209.

Supplement to
The Discrete Galerkin Method
for Integral Equations

By Kendall Atkinson and Alex Bogomolny

6. **Bound on** G_h^{-1}. In this section we adopt the approach of [17] in order to bound $\|G_h^{-1}\|$ in (4.20). This approach is quite general, and the reader may wish to refer ahead to Section 7, which contains more specific examples of subspaces S_h.

We distinguish between two cases:

(i) D is a region in $\mathbb{R}^n$,

(ii) D is a surface in $\mathbb{R}^n$.

In the first case, the modifications needed of Descloux's presentation are just notational; while in the second, they are minor in every respect. Thus we introduce the following notations and assumptions.

(a) $D \subset \mathbb{R}^n$ is either a region or a piecewise smooth hypersurface in $\mathbb{R}^n$. In the latter case, the surface D is assumed to be Lipschitz; and the general schema for D is that of application 2 in Section 3.

(b) m is either the Lebesgue measure in $\mathbb{R}^n$ or the measure over D induced by the latter.

(c) The closure $\bar{D}$ of D is subdivided into the union of closed sets e_i, $i = 1, \cdots, E_h = E$, each one of which is the closure of its interior in the natural topology of D. h is a characteristic parameter of the division. In the case D is a piecewise smooth surface, the interior of each e_i belongs to one and only one of the smooth sections that make up D.

(d) $m(e_i) > 0$, $i = 1, \cdots, E$;
$m(e_i \cap e_j) = 0$, $i, j = 1, \cdots, E$; $i \neq j$.

(e) $S_h = \text{span}(\varphi_1, \cdots, \varphi_N)$, $N = N_h$; and for all $i = 1, \cdots, N$. $C_i = \text{supp } \varphi_i$ is a union of a finite number of e_i's.

(f) For all i, there exists j such that $C_j \supset e_i$.

(g) Let $I_i = \{j : C_j \supset e_i\}$; and let $|Z|$ denote the number of elements in a finite set Z. Then

$$|I_i| \le M, \quad i = 1, \cdots, E$$

independent of h.

(h) For $Y \subset \mathbb{R}^n$, let $d(Y)$ denote the diameter of Y. Then with $\bar{n} = n$ or $\bar{n} = n-1$, depending on whether D is a region or a surface,

$$\bar{d}^n(C_i)/m(e_j) \le \tau, \quad i = 1, \cdots, N, \quad j = 1, \cdots, E.$$

and

$$\bar{d}^n(e_i)/\bar{d}^n(e_j) \le \delta, \quad i, j = 1, \cdots, E.$$

both independent of h.

(i) On $S_h|_{e_i}$, let there be given a scalar product $(\cdot, \cdot)_{e_i}$. Assume also that for all $x \in \mathbb{R}^n$ and $i = 1, \cdots, E$.

$$\alpha \sum_{j \in I_i} x_j^2 \le \sum_{j,k \in I_i} (\varphi_j, \varphi_k)_{e_i} x_j x_k \le \omega \sum_{j \in I_i} x_j^2 \quad (6.1)$$

with $\alpha > 0$.

With this, we state the following theorem.

Theorem 6.1. Assume (a)-(i) hold. Define

$$(\cdot, \cdot)_h = \sum_{i=1}^{E} (\cdot, \cdot)_{e_i}. \quad (6.2)$$

Then $(\cdot, \cdot)_h$ is an inner product on S_h; and for the Gram matrix $G_h = ((\varphi_i, \varphi_j)_h)$, one has

$$\|G_h^{-1}\| \le C \alpha^{-1}. \quad (6.3)$$

where C does not depend on h.

Proof. The proof follows that of Theorem 1 in [17] very closely. For some $f \in S_h$, suppose $(f,f)_h = \sum_{i=1}^{E} (f,f)_{e_i} = 0$.

This implies $(f,f)_{e_i} = 0$, $i = 1, \cdots, E$. and it is obvious that $f|_{e_i} = 0$, $i = 1, \cdots, E$. By (e). supp $f \subset \bigcup_{j=1}^{N} \text{supp } \varphi_j = \bigcup_{i=1}^{E} e_i$. by (f) and (e). Therefore, $f \equiv 0$.

Now in the case of D a subregion of $\mathbb{R}^n$,

$$C = s^{-1}(M c_n p^n \tau)^{\frac{1}{2}}. \quad (6.4)$$

where $0 < s < 1$, p is the smallest integer satisfying

$$\mu^{-2} M c_n \gamma n (1-\mu)^{p-1} p^{n-1} \le (1-s)^2, \qquad (6.5)$$

$\mu = \alpha\omega^{-1} \le 1$, and c_n is the volume of the unit ball in $\mathbb{R}^n$.

The relations (6.4)–(6.5) are found in [17]. These can be modified in different ways to the case of D being a surface in $\mathbb{R}^n$. The most immediate modification is as follows.

Let τ be a constant with the property

$$m(D \cap B_n(r)) \le \tau \, area(S_n(r)) \qquad (6.6)$$

for all balls $B_n(r)$ of radius r and the corresponding spheres $S_n(r)$, about all points of D. Let $\bar{c}_n$ denote the area of the unit sphere. Then

$$C = s^{-1}(\tau M \bar{c}_n p^{n-1} \gamma)^{1/2}, \qquad (6.7)$$

where $0 < s < 1$ and p is the smallest integer satisfying

$$\tau\mu^{-2} M \bar{c}_n \gamma (1-\mu)^{p-1} p^{n-1} \le (1-s)^2. \qquad (6.8)$$

The latter inequality (6.8) can be improved by making a more stringent assumption than (6.6) on the behavior of D. (6.6) in turn evidently holds for surfaces that are Lipschitz and piecewise smooth. ∎

Using Theorem 6.1, the problem of bounding $\|G_h^{-1}\|$ is thus reduced to evaluating (6.1). To see the connection with the previously discussed bound (4.19)–(4.20), note that one possible choice of the inner products $(\cdot,\cdot)_{e_i}$ in (6.1) is through the use of integration formulas on e_i ([17] uses exact integration). Therefore, the inner product (4.1) based on (4.4) and the integration (2.3) fits well into the above framework.

Observe also, that the set I_i that appears in (6.1) admits a very simple interpretation, as indicated by the following

Lemma 6.2. Assume (a)–(i) hold. Then the set $S_{e_i}^0 = \{\varphi_k|_{e_i} : k \in I_i\}$ forms a basis in $S_h|_{e_i} = \{\varphi|_{e_i} : \varphi \in S_h\}$.

Proof. Since $\alpha > 0$, it follows from (6.1) that functions from $S_{e_i}^0$ are linearly independent.

On the other hand, if $\psi \in S_h|_{e_i}$, then there exists $\varphi \in S_h$ such that $\psi = \varphi|_{e_i}$. By (e), $\varphi = \sum_{k=1}^N c_k\varphi_k$, hence

$$\psi = \varphi|_{e_i} = \sum_{k=1}^N c_k\varphi_k|_{e_i} = \sum_{k\in I_i} c_k\varphi_k|_{e_i}.$$ ∎

Construction of S_h. Now let us be more specific about the construction of the finite-dimensional spaces S_h. As in

[11, pp. 88-90], we assume there is given a reference element $\hat{e}$, a finite-dimensional space $\hat{S}$ of functions over $\hat{e}$, and for each e_k, $k = 1, \cdots, E$, a 1-1 transformation F_k of $\hat{e}$ onto e_k. We shall write $s \in \hat{e}$ and $s = F_k(\hat{s}) \in e_k$. Set

$$S_{e_k} = \{\varphi : \varphi(s) = \hat{\varphi}(\hat{s}) \text{ for some } \hat{\varphi} \in \hat{S}\}. \tag{6.9}$$

Finally, suppose on $\hat{S}$ there is defined an inner product $(\cdot, \cdot)^{\hat{}}$. Then it induces inner products $(\cdot, \cdot)_{e_k}$ on S_{e_k} by means of

$$(\varphi(s), \psi(s))_{e_k} = (|DF_k|^{\frac{1}{2}} \hat{\varphi}, |DF_k|^{\frac{1}{2}} \hat{\psi})^{\hat{}}. \tag{6.10}$$

where DF_k is the Jacobian matrix of F_k.

In the case when each F_k is an affine mapping,

$$s = F_k \hat{s} = B_k \hat{s} + b_k, \tag{6.11}$$

where $DF_k = \det B_k$ and $|DF_k| = m(e_k)/m(\hat{e})$. Hence (6.10) becomes

$$(\varphi, \psi)_{e_k} = |\det B_k|(\hat{\varphi}, \hat{\psi})^{\hat{}} = \frac{m(e_k)}{m(\hat{e})}(\hat{\varphi}, \hat{\psi})^{\hat{}}. \tag{6.12}$$

Spaces S_h we consider all possess the property

$$S_h\big|_{e_k} = S_{e_k}, \quad k = 1, \cdots, E. \tag{6.13}$$

It is not difficult to see that all finite element spaces do satisfy (6.13). The proof for spaces of spline functions is much less immediate; see [27, Thm. 4.18].

In addition, we require existence of a basis $\hat{S}^0$ of $\hat{S}$ with the property

(j) for all $k = 1, \cdots, E$, $i = 1, \cdots, N$ and $j \in I_i$ there exists $\hat{\varphi} \in \hat{S}^0$ such that $\hat{\varphi} = \varphi_i\big|_{e_k} \circ F_k$ (6.14)

and write $\hat{S}^0 = \{\hat{\varphi}_i : i = 1, \cdots, \nu\}$. By Lemma 6.2 the function $\hat{\varphi}$ in (j) is unique.

The condition (j) is what singles out finite element spaces, and most of the spline spaces fail to satisfy it. For some that do, see Corollary 7.2.

Assuming (j), (6.1) is rewritten as

$$\alpha \sum_{i=1}^{\nu} X_i^2 \leq \sum_{i,j=1}^{\nu} (|DF_k|^{\frac{1}{2}} \hat{\varphi}_i, |DF_k|^{\frac{1}{2}} \hat{\varphi}) X_i X_j \leq \omega \sum_{i=1}^{\nu} X_i^2. \tag{6.15}$$

If (6.11) holds, (6.15) simplifies to

$$\alpha \sum_{i=1}^{\nu} X_i^2 \leq |\det B_k| \sum_{i,j=1}^{\nu} (\hat{\varphi}_i, \hat{\varphi}_j) X_i X_j \leq \omega \sum_{i=1}^{\nu} X_i^2. \tag{6.16}$$

In this case, if α_0 and ω_0 are lower and upper bounds for

the eigenvalues of $((\hat\varphi_i, \hat\varphi_j)^\wedge)_{i,j=1,\ldots,\nu}$, then $\alpha \le |\det(B_k)|\alpha_0$ and $\omega \ge |\det(B_k)|\omega_0$, $k = 1,\ldots,E$.

In the more general case, assume

$$(|DF_k|^{\frac{1}{2}}\hat\varphi, |DF_k|^{\frac{1}{2}}\hat\varphi)^\wedge = (\hat\varphi, |DF_k|\hat\psi)^\wedge \tag{6.17}$$

for all $\hat\varphi, \hat\psi \in \hat S$ and also that

$$(DF_k)(\hat s) = \det(B_k)(1+O_k(h,\hat s)) \tag{6.18}$$

for some matrix B_k.

Then write

$$A = ((\hat\varphi_i, \hat\varphi_j)^\wedge)$$

and

$$\Delta = [(\hat\varphi_i, O_k(h,s)\hat\varphi_j)^\wedge)] = [O_k(h,i,j)]. \tag{6.19}$$

Lemma 6.3. *With the above notations, assume*

$$|O_k(h,i,j)| \le O(h), \quad i,j = 1,\ldots,\nu; \quad k = 1,\ldots,E. \tag{6.20}$$

Then in (6.15) one can choose

$$\alpha = \frac{\alpha_0}{2}\,\underset{1\le k\le E}{\text{Min}}\,|\det(B_k)|. \tag{6.21}$$

provided h is sufficiently small.

Proof. The proof uses a standard result from perturbation theory; see [6, p. 505]. If λ is an eigenvalue of $A+\Delta$ then

$$\min_{1\le j\le\nu}|\lambda-\lambda_j| \le \|\Delta\|_2,$$

where λ_i are eigenvalues of A. Then, since $\|\Delta\|_2 \le \|\Delta\|_F \le O(h)$, where $\|\cdot\|_F$ is the Frobenius matrix norm, the smallest eigenvalue λ of $A+\Delta$ satisfies

$$\lambda \ge \alpha_0 - |O(h)| > \tfrac{1}{2}\alpha_0.$$

provided h is sufficiently small. ∎