Improved convergence rates for some discrete Galerkin methods
M.A. Golberg
2025 University Circle, Las Vegas, NV 89119-6051, USA

Abstract

We show how to obtain almost optimal convergence rates for discrete Galerkin methods for a number of Fredholm and singular integral equations. These rates improve on previous ones we have obtained for one-dimensional Fredholm equations and are new for singular and multidimensional Fredholm equations.

1 Introduction

The reformulation of boundary value problems leads to a variety of integral equations. These include one-dimensional Fredholm equations [1], Cauchy singular equations [2], hypersingular equations [3] and multidimensional Fredholm equations [4]. Traditionally in the BEM literature these equations are solved numerically by piecewise polynomial collocation methods [5]. However, there are many situations where global approximations based on polynomials for one-dimensional equations and spherical harmonics for surface integral equations are natural [1, 5]. Moreover, recent work shows that Galerkin methods may be more efficient than collocation in these situations, because the Galerkin matrices are ‘almost sparse’ while the collocation matrices are not. In addition, for surface equations, collocation methods based on spherical harmonics are not available [5]. For this situation Galerkin methods may be the method of choice.

In recent years there has been intensive analysis of Galerkin methods using piecewise-polynomials. In particular, much work has been done on the effects of numerical integration on their convergence [6]. More recently, we and others have studied the convergence of globally-based discrete Galerkin methods [1, 2, 7]. Here, using an argument analogous to that of Joe [7] for piecewise-polynomial approximations [6], we improve on our previous results.
for one-dimensional Fredholm and singular equations with smooth kernels [1, 2] and extend these to some equations with singular kernels, particularly the generalized airfoil equation (GAF), hypersingular equations and also to Fredholm equations on the sphere.

In Section 2 we review Galerkin’s method and in Section 3 we introduce the discretized versions resulting from the numerical approximation of the integral transforms and inner products. We also state some new error bounds for the quadratures we use and these are then applied to establish convergence rates for the algorithms we discuss. We close with some directions for future research.

2 Galerkin’s Method

We consider the numerical solution of the following integral equations by Galerkin’s method:

\[ u(x) = \int_{-1}^{1} k(x, t)u(t)dt + f(x), \quad \text{Fredholm equation} \]  \hspace{1cm} (1)

\[- \frac{1}{\pi} \int_{-1}^{1} \frac{w(t)u(t)dt}{t-x} + \frac{1}{\pi} \int_{-1}^{1} w(t)a(x-t) \log |x-t| u(t)dt \]

\[ = \int_{-1}^{1} w(t)k(x, t)u(t)dt + f(x), \quad w(t) = (1-t)^{\frac{1}{2}}(1+t)^{-\frac{1}{2}}, \]  \hspace{1cm} (2)

[generalized airfoil equation]

\[ \frac{1}{\pi} \frac{d}{dx} \int_{-1}^{1} \frac{(1-t^2)^{\frac{3}{2}} u(t)}{t-x} = \int_{-1}^{1} (1-t^2)^{1/2} k(x, t)u(t)dt + f(x), \]  \hspace{1cm} (3)

[hypersingular equation],

and

\[ u(P) = \int_{D} k(P, Q)u(Q)ds + f(P), \quad P \in D, \]  \hspace{1cm} (4)

where \( D = \{ x^2 + y^2 + z^2 = 1 \} \) is the unit sphere in \( R^3 \).

For convenience of exposition we rewrite (1)-(4) in operator form as

\[ Hu + Wu = Ku + f \]  \hspace{1cm} (5)

where \( H = I \), the identity operator, in (1) and (4) and \( H \) and \( W \) are defined by the left hand sides in (2)-(3) \( [W = 0 \text{ in (1), (3) - (4)}] \).

For appropriately defined Hilbert spaces \( (X,Y) \) [to be defined] we assume that \( H : X \to Y \) is a bounded, invertible operator and \( (K,W) : X \to Y \) are compact. We also assume that (1)-(4) are uniquely solvable for each \( f \in Y \).
To solve (1)-(4) by Galerkin's method, let \( \{ \varphi_k \}_{k=0}^\infty \) be a complete orthogonal basis for \( X \) and approximate \( u \) by

\[
u_N = \sum_{k=0}^{N} a_k \varphi_k,
\]

where the coefficients \( \{ a_k \} \) are determined by setting the residual \( r_N = Hu_N + Wu_N - Ku_N - f \) orthogonal to \( \{ \psi_j \}_{j=0}^N \) in \( Y \). This gives the set of \( N + 1 \) linear equations

\[
\sum_{k=0}^{N} a_k \langle (H + W) \varphi_k, \psi_j \rangle_Y = \sum_{k=0}^{N} a_k \langle K \varphi_k, \psi_j \rangle_Y + \langle f, \psi_j \rangle_Y, \quad j = 0, 1, \ldots, N,
\]

where \( \langle \cdot, \cdot \rangle_Y \) is the inner product on \( Y \).

Although there are many possibilities for \( \{ \varphi_k \} \) and \( \{ \psi_j \} \), here we use the following choices. For (1) we let \( X = Y = L_2 \), the space of real square-integrable functions on \([-1, 1]\) and choose \( \varphi_k = \psi_k, \ k \geq 0, \) the normalized Legendre polynomials, with \( \deg \varphi_j = j \) and \( N = n \), the degree of \( \varphi_n \). In this case \( \langle \varphi_k, \varphi_j \rangle_X = \delta_{jk} \) so that (7) becomes

\[
a_j = \sum_{k=0}^{n} a_k \langle K \varphi_k, \varphi_j \rangle_X + \langle f, \varphi_j \rangle_X, \quad j = 0, 1, \ldots, n.
\]

For (2) we let \( X = L_w \), the space of real functions square integrable with respect to \( w \), \( Y = L_{1/w} \) and

\[
\varphi_k = \sin \left( (k + 1/2) t \right) / \sin \left( t/2 \right), \quad \psi_k = \cos \left( (k + 1/2) t \right) / \cos \left( t/2 \right), \quad k \geq 0,
\]

are orthogonal bases for \( L_w \) and \( L_{1/w} \) respectively. To simplify the analysis in this case, we assume that \( a(x) = a/\pi. \) Then it is known that [8]

\[
H \varphi_k = -\psi_k, \quad k \geq 0,
\]

and

\[
W \varphi_0 = (1/2 - \log 2) \psi_0 + \psi_1, W \varphi_k = \frac{1}{2} \left[ \psi_{k+1} / k + 1 - \psi_k / (k + 1) - \psi_{k-1} / k \right],
\]

\[
k \geq 1.
\]

Using (10)-(11) and the orthogonality of \( \{ \psi_j \} \), one can compute \( \langle w \varphi_k, \psi_j \rangle_Y \) analytically. It follows from this and the results of [9] that \( H + W \) is invertible if \( a \) is real.

For (3) \( X = L_p, \rho = (1 - t^2)^{1/2}, \) and \( Y \) is the closed subspace \( L_1(\rho) \) of \( X \) given by \( u \in L_\rho \) such that

\[
\sum_{k=0}^{\infty} (k + 1)^2 \langle u, \varphi_k \rangle_X^2 < \infty
\]
where
\[ \psi_k = \varphi_k = \frac{2}{\pi} U_k, \quad U_k = \sin \left[(k + 1) t\right]/\sin t, \quad k \geq 0, \quad (13) \]
the Chebyshev polynomials of the second kind. In this case we use the \( L_p \)
inner product in (7). Since \( \mathcal{H}\varphi_k = - (k + 1) \varphi_k \) [3] (7) simplifies to
\[ -(j + 1) a_j = \sum_{k=0}^{n} \langle K \varphi_k, \varphi_j \rangle_X + \langle f, \varphi_j \rangle_X, \quad 0 \leq j \leq n. \quad (14) \]

Last, for (4) let \( X = Y = L_2(D) \) and let \( \{ \varphi_k \} = \{ \psi_j \} \) be the normalized spherical harmonics on \( D \). We order these in such a way that \( N = (n + 1)^2 \)
and \( \{ \varphi_0, \ldots, \varphi_N \} \) is a basis for the spherical polynomials of degree \( \leq n \).
Since \( \mathcal{H} = I \) in this case, (7) becomes
\[ a_j = \sum_{k=0}^{N} a_k \langle K \varphi_k, \varphi_j \rangle_X + \langle f, \varphi_j \rangle_X, \quad 0 \leq j \leq n. \quad (15) \]

### 2.1 Convergence of Galerkin’s method

When the inner products and integral transforms in (7) are computed exactly, the convergence of Galerkin’s method for (1) has been studied in [1],
for (2) in [8], for (3) in [3] and for (4) in [4]. For this let \( Y_N = \text{span} \{ \psi_j \} \),
and let \( \mathbb{C} \) be the operator of orthogonal projection onto \( Y_N \). Then \( u_N \)
satisfies the operator equation
\[ \mathbb{C} (\mathcal{H} u_N + W u_N - \mathcal{K} u_N) = \mathbb{C} f. \quad (16) \]
For (1), (3) and (4) (16) is equivalent to
\[ \mathcal{H} u_N - \mathbb{C} K u_N = \mathbb{C} f, \quad (17) \]
while for (2) we have
\[ \mathcal{H} u_N + P_N W u_N = P_N K u_N + P_N f. \quad (18) \]

Since \( \mathcal{H} \) is invertible, \( K \) and \( W \) are compact and \( P_N g \to g \quad \forall g \in Y \), it follows
from standard arguments that \( u_N \to u \) in \( Y \), and for all \( N \) sufficiently large that [2]
\[ \| u - u_N \|_X \leq c \| \mathcal{H} u - P_N \mathcal{H} u \|_Y. \quad (19) \]
With further smoothness assumptions on \( k \) and \( f \) we can obtain optimal rates of convergence for \( u_N \). In (2), (3) and (4) we assume that \( k \) and \( f \) are \( C^r \), \( r > 1 \), so that \( \mathcal{H} u = K u + f \) is \( C^r \) as well. Furthermore, in (2) and (3)
using \( \| P_N \| = 1 \) and Jackson’s theorem, \( \| \mathcal{H} u - P_N \mathcal{H} u \|_Y \leq c n^{-r} \) so that
\( \| u - u_N \|_X \leq c n^{-r} \). For (4) we use a theorem of Ragozin [10] and \( \| P_N \| = 1 \)
to show that \( \| u - P_N u \|_X \leq c n^{-r} \) giving \( \| u - u_N \|_X \leq c n^{-r} \) as well.
For the GAF we use the argument in [2] or [10] to show that \( Hu \in C^r \), since \( Hu \) satisfies a first order differential equation with \( C^r \) data. Hence, it follows from (19) and Jackson’s theorem again that \( \| u - u_N \|_X \leq cn^{-r} \).

3 Discrete Galerkin methods

For practical implementation of Galerkin’s method it is usually necessary to approximate the inner products and integral transforms in (7) by numerical integration. To describe this, denote the domains in (1)-(4) generically by \( D \). Then integrals of the form \( \int_D w(t)g(t)dt \) are approximated by

\[
\int_D w(t)g(t)dt \approx Q_L(g) = \sum_{\ell=1}^{L} w_{\ell} g(t_{\ell})
\]

and integrals of the form \( \int_D \sigma(t)g(t)dt \) by

\[
\int_D \sigma(t)g(t)dt \approx Q_S(g) = \sum_{m=1}^{M} \sigma_{m} g(t_{m}),
\]

where (i) \( w_{\ell} > 0, 1 \leq \ell \leq L \), (ii) \( \sigma_{m} > 0, 1 \leq m \leq M \), and in (1)-(3) \( Q_L \) and \( Q_S \) are exact for all polynomials of degree \( \leq 2n \) and for (4) \( Q_L \) is exact for all spherical polynomials of degree \( \leq 2n \).

In (8) we use (20) with \( w = 1 \), \( \langle f, \varphi_j \rangle_X = Q_L(f \varphi_j) \) and \( \langle K \varphi_k, \varphi_j \rangle_X = Q_L \times Q_L(k \varphi_k \varphi_j) \), where \( Q_L \times Q_L \) is a product rule on \( D \times D \). For the GAF we use (20) and (21) with \( \sigma = 1/w \), \( \langle f, \psi_j \rangle_Y = Q_S (f \psi_j) \) and \( \langle K \varphi_k, \psi_j \rangle_Y = Q_L \times Q_S (k \varphi_k \psi_j) \). In (14) we use \( w = \sigma = \rho \) and \( \langle f, \varphi_j \rangle_X = Q_L(f \varphi_j) \) and \( \langle K \varphi_k, \varphi_j \rangle_X = Q_L \times Q_L(k \varphi_k \varphi_j) \) while in (15) with \( w = \sigma = 1 \), \( \langle f, \varphi_j \rangle_X = Q_L(f \varphi_j) \) and \( \langle K \varphi_k, \varphi_j \rangle_X = Q_L \times Q_L(k \varphi_k \varphi_j) \).

Using (20)-(21),

\[
K_N u(x) = \sum_{\ell=1}^{L} w_{\ell} K(x, t_{\ell}) u(t_{\ell}),
\]

and

\[
\pi_N u(x) = \sum_{j=0}^{N} Q_S (w \psi_j) \psi_j
\]

the discrete Galerkin approximation \( v_N \) to \( u \) satisfies [1, 2, 6]

\[
Hu_N + P_N W v_N = \pi_N K_N u_N + \pi_N f.
\]

There are several approaches to analyzing the convergence of \( v_N \). Here we use the theory of discrete projection methods. Although this method sometimes gives slightly less sharp convergence rates than a more direct analysis of (24), it is more elementary and more general.
Boundary Elements XVII

Hence, rewrite (24) as

$$H v_N = -P_N W v_N + P_N k v_N + R_N v_N + P_N f$$  \hspace{1cm} (25)$$

where

$$R_N v_N = -(P_N K v_N - \tau_N K N v_N), \quad r_N = -(P_N f - \pi_N f).$$  \hspace{1cm} (26)$$

Some straightforward manipulation shows that

$$R_N v_N = -\sum_{j=0}^{N} E_j (k v_N \psi_j) \psi_j, \quad r_N = -\sum_{j=0}^{N} e_j (f \psi_j) \psi_j,$$  \hspace{1cm} (27)$$

where \( E_j (k v_N \psi_j) \) and \( e_j (f \psi_j) \) are the integration errors in approximating \( \langle K v_N, \psi_j \rangle_Y \) by \( Q_L \times Q_S (k v_N \psi_j) \) and \( \langle f, \psi_j \rangle_Y \) by \( Q_S (f \psi_j) \) respectively.

Letting \( X_N = \text{span} \{ \varphi_k \}^{N}_{k=0} \)

$$\| R_N \|_X = \text{lub} \{ \| R_N w_N \|_Y, w_N \in X_N, \| w_N \|_X = 1 \},$$  \hspace{1cm} (28)$$

it follows from Theorem 1 of [12] that if \( \| R_N \|_X \to 0 \) and \( \| v_N \|_Y \to 0 \), \( N \to \infty \), that \( u_N \to u \) and

$$\| u - v_N \|_X \leq c (\| u - u_N \|_X + \| R_N \|_X + \| v_N \|_Y).$$  \hspace{1cm} (29)$$

From Section 2 \( \| u - u_N \|_X \leq cn^{-\tau} \) if \( k \) and \( f \) are \( C^r \). Thus it suffices to bound \( \| R_N \|_X \) and \( \| r_N \|_Y \) in order to bound \( \| u - v_N \|_X \).

To do this we proceed differently than in [1, 2] to directly bound \( E_j \) rather than \( E_j (k \varphi_k \psi_j) \) as in [1, 2]. This leads to improved convergence rates compared to those in [1, 2] and gives new ones for (2) and (3). Since the proofs are rather lengthy, we merely state the necessary results and leave the details to [11].

**Theorem 3.1** Let \( w(x) \geq 0 \) be a non-negative integrable weight function and let \( \{ \varphi_k \} \) be the orthogonal polynomials associated with \( w \). Let \( X_n = \text{span} \{ \varphi_k \}^{N}_{k=0} \) and let \( v_n \in X_n \) with \( \| v_n \|_X = 1 \). Consider the integral

$$I_n = \int_{-1}^{1} w(t) f(t) v_n(t) dt,$$

where \( f \in C^r, r \geq 1 \). Suppose \( I_n \) is approximated by \( Q (f v_n) \) where \( Q \) satisfies the conditions in Section 2. Then the error

$$|c| \leq c n^{-\tau}$$  \hspace{1cm} (30)$$

where \( c \) depends only on \( f \).

**Theorem 3.2** Let \( \rho \geq 0, \sigma \geq 0 \) be non-negative weight functions on \([-1, 1]\) and let \( \{ \varphi_k \} \) and \( \{ \psi_k \} \) be the orthogonal polynomials associated with \( \rho \) and \( \sigma \) respectively. Let \( k(x,t) \in C^r, r \geq 1 \), and consider the integrals

$$I_n = \int_{-1}^{1} \int_{-1}^{1} \rho(x) \sigma(t) k(x,t) \zeta_n(x) v_n(t) dx dt,$$  \hspace{1cm} (31)$$
where \( v_n \in \text{span} \{ \psi_k \}_{k=0}^{n} \) and \( z_n \in \text{span} \{ \psi_k \}_{k=0}^{n} \) with \( \| v_n \|_X = \| z_n \|_V = 1 \). Then the error \( E = I_n - Q_L \times Q_S (ktv_n, z_n) \) satisfies
\[
|E| \leq cn^{-r}
\]
where \( c \) depends only on \( k \).

**Theorem 3.3** Let \( D \) be the unit sphere and \( f \in C^r(D) \), \( r \geq 1 \), let \( Q_L \) be an integration rule satisfying the conditions in Section 2. Let \( w_N \in X_N \) with \( \| w_N \|_X = 1 \). Then the integration error \( e = (f, w_N)_X - Q_L (f w_N) \) satisfies
\[
|e| \leq cn^{-r}
\]
where \( c \) depends only on \( f \).

**Theorem 3.4** Let \( k(P, Q) \in C^r(D \times D) \) and let \( (w_N, z_N) \in X_N \) with \( \| w_N \|_X = \| z_N \|_X = 1 \). Then the integration error \( E = (Kw_N, z_N)_X - Q_L \times Q_L (k w_N z_N) \) satisfies
\[
|E| \leq cn^{-r}
\]
where \( c \) depends only on \( k \).

### 3.1 Convergence of the discrete Galerkin method

Using the estimate (29) and Theorems 3.1 - 3.4 the convergence of the discrete Galerkin methods proposed in Section 2 follows immediately if \( (f, k) \in C^r, r > 1 \). For (1)-(3) using Theorems 3.1 - 3.2 it follows that \( \| R_n \|_n \leq cn^{-r+1/2} \) and \( \| r_n \|_V \leq cn^{-r+1/2} \). Using this in (29) shows that \( \| u - u_n \|_X \leq cn^{-r+1/2}, r \geq 1 \).

For (4) it follows from Theorems 3.3 - 3.4 that \( \| R_n \|_N \leq Cn^{-r+1} \) and \( \| r_N \|_X \leq cn^{-r+1} \). Hence \( u_N \to u \) if \( r > 1 \) and \( \| u - u_N \|_X \leq cn^{-r+1} \). For (1) the \( O\left(n^{-r+1/2}\right) \) convergence rate improves on the \( O\left(n^{-r+1}\right) \) rate given in [1, 2], while the rate for (2)-(3) is new. That for (4) is the same as given in [7], but the proof here is considerably less technical. Similar proofs can be given for other discrete Galerkin methods such as those in [2].

### 4 Discussion

We have shown how to obtain almost optimal convergence rates for discrete Galerkin methods for a number of Fredholm and singular integral equations. These rates improve on those given previously for (1) and are new for (2) - (4). Although collocation methods have traditionally been the method of choice for solving boundary integral equations with piecewise polynomials, such methods may not be available for global approximations. Moreover, for some problems, such as those which occur in aerodynamics
Boundary Elements XVII

[2, 8], Galerkin’s method allows the exploitation of the superconvergence properties of Galerkin’s method to accurately compute linear functionals of the solution, which collocation cannot do. We expect to elaborate on this in future work.

References


