Spectral Equivalence and Proper Clusters for Matrices from the Boundary Element Method

Eugene E. Tyrtyshnikov* Raymond H. Chan†

ABSTRACT

The Galerkin matrices A_n from applications of the boundary element method to integral equations of the first kind usually need to be preconditioned. In the Laplace equation context, we highlight a family of preconditioners C_n that simultaneously enjoy two important properties: (a) A_n and C_n are spectrally equivalent, and (b) the eigenvalues of $C_n^{-1}A_n$ have a proper cluster at unity. In the Helmholtz equation context, we prove the spectral equivalence for the so-called second Galerkin matrices and that the eigenvalues of $C_n^{-1}A_n$ still have a proper cluster at unity. We then show that some circulant integral approximate operator (CIAO) preconditioners belong to this family, including the well-known optimal CIAO. Consequently, if we use the preconditioned conjugate gradients to solve the problems, the number of iterations for a prescribed accuracy does not depend on n, and, what is more, the convergence rate is superlinear.

Keywords: Boundary Element Method, Spectral Equivalence, Circulant Integral Operator.

*Institute of Numerical Mathematics, Russian Academy of Sciences, Gubkina 8, Moscow 117333, Russia. Research was carried out during the visit to the Institute of Mathematical Sciences of the Chinese University of Hong Kong. Supported in part by the RFBR Grants 97-01-00155, 99-01-00017 and Volkswagen-Stiftung.

†Department of Mathematics, Chinese University of Hong Kong, Shatin, Hong Kong. Research supported in part by Hong Kong Research Grants Council Grant No. CUHK 4207/97P and CUHK DAG Grant No. 2060143.

1 Introduction

In this paper, we study preconditioners for the moment matrices $A_n \in \mathbb{C}^{n\times n}$ in the Galerkin method applied to some integral equations of the first kind. Since the spectral condition number of A_n may (and does) grow with n, we use as preconditioners some other matrices $C_n \in \mathbb{C}^{n\times n}$ that are easier to invert. The purpose here is to make the condition number of $C_n^{-1}A_n$ bounded or, if that can not be acheived, at least to slow down its growth with n and *anyway to improve the behavior of the eigenvalues*. The first case can be achieved if A_n and C_n are *spectrally equivalent*:

Definition 1 *Consider two sequences of Hermitian positive definite matrices A_n and C_n and assume that all the eigenvalues λ of $C_n^{-1}A_n$ satisfy $c_1 \le \lambda \le c_2$ with positive c_1 and c_2 independent of n. Then A_n and C_n are said to be* spectrally equivalent.

In this case, the number of iterations in the preconditioned conjugate gradients (PCG) depends only on the desired reduction of the residual and not upon n. This is the well-known linear convergence result of PCG, see for instance [2]. However, we may expect even a faster convergence for sufficiently large n if the eigenvalues behave properly:

Definition 2 *Denote by $\gamma_n(\varepsilon)$ the number of those eigenvalues of A_n that lie at a distance farther than ε from $z \in \mathbb{C}$. The eigenvalues of A_n are said to have a* proper cluster *at z if $\gamma_n(\varepsilon) \le c(\varepsilon)$ for all n and for all $\varepsilon > 0$. If two matrices A_n and C_n are nonsingular and the eigenvalues of $C_n^{-1}A_n$ have a proper cluster at unity, then we say that A_n and C_n are* properly equivalent.

In the Hermitian case, "properly equivalent" is what accounts for the so-called *superlinear* convergence of PCG.

Definition 3 *An iterative method converges* superlinearly *for A_n if, for any $0 < q < 1$ and for all sufficiently large n the residual at the ith iteration is bounded above by cq^i (c may depend on q, n, the right-hand side, and the initial vector).*

In the Hermitian case, if C_n is a properly equivalent preconditioner for A_n, then PCG converge superlinearly, see for instance [2, 3]. In this sense, C_n may be called an optimal preconditioner for A_n; yet, since the word "optimal" is overused, we will call C_n a *superlinear preconditioner* (see [13]). If at least one of the matrices A_n or C_n has the Frobenius norm growing not faster than $o(\sqrt{n})$, then, it follows from [15] that the eigenvalues of C_n and A_n are *equally distributed* (in the spirit of G. Weyl's and G. Szegö's approach [10] extended in [15]).

Why superlinear preconditioners make PCG converge superlinearly is explained, for example, in [2] and in [3]. The latter paper considered the case of A_n being Toeplitz matrices with a positive symbol from the Wiener class and C_n being G. Strang's or T. Chan's circulants. It is proved therein that the eigenvalues of $C_n^{-1}A_n$ have a proper cluster at unity and how the superlinear convergence follows from that. The crucial notion of cluster has then been extended and studied in [15, 16].

In the non-Hermitian case, we can apply PCG to solve the normal equations $A_n^*A_n\mathbf{x} = A_n^*\mathbf{b}$. Then, we are interested to have $A_n^*A_n$ and $C_n^*C_n$ spectrally equivalent. It is equal to the claim that the singular values σ of $A_nC_n^{-1}$ all satisfy $c_1 \le \sigma \le c_2$ with some positive c_1 and c_2. In addition, if $C_n^*C_n$ and $A_n^*A_n$ are properly equivalent, then superlinear convergence comes as a fringe benefit.

In this paper, as basic examples we consider two typical single layer potential equations related to the Dirichlet boundary value problems for the Laplace ($\Delta w = 0$) and Helmholtz ($(\Delta + \kappa^2)w = 0$) equations in two dimensions. The first one is (some call it Symm's equation [12])

$$-\frac{1}{2\pi}\int_{\partial\Omega} \log|x-y|\, U(y)\, ds(y) \;=\; F(x), \quad x \in \partial\Omega, \tag{1}$$

and the second is

$$\frac{\mathbf{i}}{4}\int_{\partial\Omega} H_0^{(1)}(\kappa|x-y|)\, U(y)\, ds(y) \;=\; F(x), \quad x \in \partial\Omega, \tag{2}$$

here $ds(y)$ is the arclength element. For simplicity, we assume for the time being that $\partial\Omega$ is an infinitely smooth closed curve cutting the plane into the interior domain Ω and the exterior Ω^c (our results in later sections cover boundaries with a finite smoothness). The kernels for the equations are the fundamental solutions for the Laplace and Helmholtz equations in two dimensions. In particular, $H_0^{(1)}$ in (2) is the Hankel function of order 0 of the first kind. Recall that

$$H_0^{(1)}(z) \sim \frac{2\mathbf{i}}{\pi}\log z \quad \text{as} \quad z \to 0, \tag{3}$$

and so the two kernels have the same type of singularity at $x = y$. Theory and applications of these equations are well-known [7]; nonetheless we will recollect the crucial facts when they are needed.

The Galerkin matrices A_n are Hermitian for the equation (1) while non-Hermitian (yet symmetric) for (2). In our analysis, we rely heavily on the results in the theory of elliptic boundary integral equations which are well appreciated nowadays and mostly due to Hsiao and Wendland [11]; cf. [7]. We always *assume* that the integral operators in question have a trivial null-space. For (1) that means the diameter of the contour is less than 1 and,

if it is not the case, it can be easily remedied by a contraction [7]. As far as (2) is concerned, the assumption amounts to that some (resonant for Ω) values of κ are prohibited.

For the equation (1), we find that if C_n are the Galerkin matrices for *any other* smooth contour of diameter less than 1, then A_n and C_n are spectrally equivalent. It implies that C_n can correspond to any convenient contour for which C_n is easily invertible. The best choice might be a circle, in which case C_n are circulant matrices, see [3, 14, 16]. It was shown in [4] that the optimal circulant integral approximate operator (CIAO) also leads to preconditioners spectrally equivalent to A_n (some previous papers on CIAOs for the Wiener–Hopf equations are [5, 9]). Note that the circle-based preconditioner, though not optimal, in the sense of [6] and [4], is also a CIAO. In light of our results, the performance of the optimal CIAO preconditioner for (1) is the same as that of the circle-based CIAO (and of any other contour-based preconditioner). We will illustrate this by numerical results in §6.

In the case of (2), we prove the spectral equivalence for the so-called second Galerkin matrices (defined in (6)) which are related to the normal equations of (2).

Furthermore, for both equations, we prove that the Galerkin matrices A_n and C_n corresponding to *any two* smooth contours are properly equivalent. It is therefore natural to use the circle-based CIAOs which yield easily invertible matrices. It is a superlinear preconditioner for A_n, and equally good compared with any other contour-based preconditioner (at least asymptotically). Even we may choose C_n to be the matrices coming from (1) to precondition A_n from (2).

All in all, for equations (1) and (2), we have pretty simple preconditioners which are both spectrally equivalent and properly equivalent (provide a proper cluster at unity). In general, these two properties do not imply each other but in our case we have as much as both. The proof is of a certain generic nature, so without significant changes we can cover other kinds of equations.

The outline of the paper is as follows. In Sections 2 and 3, we consider spectral equivalent preconditioners for matrices from (1) and (2) respectively. In Section 4, we consider the clustering properties of these preconditioners. In Section 5, we show that the CIAO preconditioners have the properties we studied in Sections 2–5. Numerical results are given in Section 6 to illustrate our theory.

2 Spectral equivalence: Hermitian case

Given two sequences of Hermitian positive definite matrices A_n and C_n, both of order n, we wish to know whether they are spectrally equivalent. Here is an obvious observation that can help.

Lemma 2.1 *Two sequences of Hermitian positive definite matrices A_n and C_n are spectrally equivalent if and only if there are constants $0 < c_1 \le c_2$ such that for any n, there exists a vector norm $||\cdot||_*$ on $\mathbb{C}^n$ for which*

$$c_1 \, ||\mathbf{x}||_*^2 \le \mathbf{x}^* A_n \mathbf{x} \le c_2 \, ||\mathbf{x}||_*^2$$

and

$$c_1 \, ||\mathbf{x}||_*^2 \le \mathbf{x}^* C_n \mathbf{x} \le c_2 \, ||\mathbf{x}||_*^2$$

hold for any $\mathbf{x} \in \mathbb{C}^n$.

This lemma may shed some light on why we are so anxious in matrix analysis to study norms in a general setting. Still, where does this norm $||\cdot||_*$ come from?

Assume that $\mathcal{A}: \; H \to H'$ is a continuously invertible linear operator on a pair of Hilbert spaces H and H' which are dual with respect to a bounded Hermitian bilinear form (f, u), where $f \in H'$ and $u \in H$. Consider a sequence of finite sets $\{u_{in}\}_{i=1}^n$ of linearly independent (basis) functions $u_{in} \in H$ and the corresponding Galerkin matrices

$$A_n \equiv [\mathcal{A}]_n = [a_{ij}], \quad a_{ij} = (\mathcal{A}u_{jn}, u_{in}), \quad 1 \le i, j \le n. \tag{4}$$

Assume additionally that

$$(\mathcal{A}u, v) = \overline{(\mathcal{A}v, u)}, \quad \forall \, u, v \in H$$

and, for $0 < c_1 < c_2$,

$$c_1 \, ||u||_H^2 \le (\mathcal{A}u, u) \le c_2 \, ||u||_H^2, \quad \forall u \in H.$$

In this case, we say that $\mathcal{A}$ is a *Hermitian positive definite* operator. Now, the vector norm on $\mathbb{C}^n$ can be defined as follows:

$$||\mathbf{x}||_* = ||u||_H, \quad u = \sum_{i=1}^n x_i \, u_{in}, \quad \mathbf{x} = [x_1, \ldots, x_n]^T.$$

It is easy to check that $\mathbf{x}^* A_n \mathbf{x} = (\mathcal{A}u, u)$. Hence, we are immediately led to the following.

Lemma 2.2 *Suppose that $\mathcal{A}$ and $\mathcal{B}$ are Hermitian positive definite operators on the same pair of Hilbert spaces. Then the Galerkin matrices $A_n = [\mathcal{A}]_n$ and $B_n = [\mathcal{B}]_n$ are spectrally equivalent.*

Now, let $\mathcal{A}$ be the integral operator from (1) or (2). Since any curve is dealt with through some parametrization, functions on $\partial\Omega$ can be identified with those on $[0, 2\pi]$ (or with 2π-periodic functions on the real axis). For example, we can use the parametrization by arclength and tailor the parameter by mapping it linearly onto $[0, 2\pi]$. Thus without loss of generality, let $\partial\Omega = \{\gamma(t),\ 0 \le t \le 2\pi\}$. We demand γ to be a diffeomorphism, which implies that $\gamma'(t)$ never vanishes, and, sometimes, γ may be required to have some higher order derivatives. (However, some results hold for Lipschitz boundaries [8].) With the parametrization, (1) and (2) can be casted in the form

$$\int_0^{2\pi} a(\gamma(\tau), \gamma(t))\ u(t)\, dt \ = \ f(\tau),$$

with different kernels $a(x, y)$. Obviously,

$$u(t) = U(\gamma(t))\, |\gamma'(t)|, \quad f(\tau) = F(\gamma(\tau)).$$

It is exactly how we proceed in practice. Also, in view of Lemma 2.2, we are now able (and keen) to consider different kernels and contours with exactly the same spaces for $u(t)$ and $f(\tau)$.

Denote by H^s the Sobolev space of order $s \in \mathbb{R}$. Recall that it can be defined as the completion of infinitely smooth functions $u(t)$ with respect to the norm

$$||u||_s = \left(\sum_{j=-\infty}^{\infty} |\hat{u}_j|^2\ (1 + |j|)^{2s} \right)^{1/2}$$

expressed through the Fourier coefficients $\hat{u}_j$ of u. This norm is naturally related to the scalar product

$$(u, v)_s = \sum_{j=-\infty}^{\infty} \hat{u}_j\ \overline{\hat{v}_j}\ (1 + |j|)^{2s},$$

where

$$u(t) = \sum_{j=-\infty}^{\infty} \hat{u}_j e^{\mathbf{i}jt}, \quad v(t) = \sum_{j=-\infty}^{\infty} \hat{v}_j e^{\mathbf{i}jt}.$$

Thus, H^s is a Hilbert space (cf. [12]).

Let $a(x, y)$ be either the kernel function of the integral in (1) or (2). We gather from the theory of pseudodifferential equations that for any $\gamma \in C^\infty$ the corresponding operator $\mathcal{A}$ can be viewed as a continuous operator from H^s to H^{s+1} for any $s \in \mathbb{R}$ (as a pseudodifferential operator of order -1) [7]. For finitely smooth γ (and even for Lipschitz boundaries [8]), it is still a continuous operator from $H^{-1/2}$ to $H^{1/2}$. In line with the above, we take

$H = H^{-1/2}$, $H' = H^{1/2}$, and the Hermitian bilinear form is defined by the Lebesgue integral

$$(f, u) = \int_0^{2\pi} f(t)\,\overline{u(t)}\,dt \tag{5}$$

(it is easy to see that $f\overline{u} \in L_2 = H^0$).

In the case of (1), the fundamental result of Hsiao and Wendland [11] can be put in the following form.

Theorem 2.1 (Hsiao and Wendlandt [11]) *Let $a(x, y) = -\frac{1}{2\pi}\,\log|x - y|$, and assume that the boundary γ is sufficiently smooth and of diameter less than 1. Then $\mathcal{A}$ is a Hermitian positive definite operator from $H^{-1/2}$ to $H^{1/2}$ with respect to the Hermitian bilinear form defined by (5).*

Combining the above theorem with the previous lemmas, we immediately come up with the following result.

Theorem 2.2 *Let γ_1 and γ_2 be two contours satisfying the hypotheses of Theorem 2.1 with the corresponding logarithmic-kernel operators being $\mathcal{A}_1$ and $\mathcal{A}_2$. Then the Galerkin matrices $[\mathcal{A}_1]_n$ and $[\mathcal{A}_2]_n$ are spectrally equivalent.*

Remark. In agreement with the results reported in [8], the spectral equivalence still holds for Lipschitz γ. However, in the presence of corners, we may need to use some graded meshes or nonuniform finite elements. Consequently, some nice structures of the Galerkin matrices, even for circular contours, might vanish.

3 Spectral equivalence: non-Hermitian case

We get on to the equation (2) (yet everything applies to the equation (1), too). Since

$$c_1\,||u||^2_{-1/2} \le ||\mathcal{A}u||^2_{1/2} \le c_2\,||u||^2_{-1/2}, \quad \forall\, u \in H^{-1/2},$$

the matrices $[(\mathcal{A}u_{jn}, \mathcal{A}u_{in})_{1/2}]$ for different contours are spectrally equivalent. But these matrices never appear in practice. Of some real-life interest are matrices of the form

$$[\mathcal{A}^*\mathcal{A}]_n \equiv [a_{ij}], \quad a_{ij} = (\mathcal{A}u_{jn}, \mathcal{A}u_{in}) = (\mathcal{A}u_{jn}, \mathcal{A}u_{in})_0, \quad 0 \le i, j \le n, \tag{6}$$

which we call the *second Galerkin matrices*. They are related to the normal equations of $[\mathcal{A}]_n\mathbf{x} = \mathbf{b}$, see (8) and (9) below.

If $\gamma \in C^\infty$, then the inequalities

$$c_1 \, ||u||_s^2 \le ||\mathcal{A}u||_{s+1}^2 \le c_2 \, ||u||_s^2$$

hold for any $s \in \mathbb{R}$ (the constants depend on s) [7, 12]. The choice $s = -1$ gives

$$c_1 \, ||u||_{-1}^2 \le (\mathcal{A}u, \mathcal{A}u)_0 \le c_2 \, ||u||_{-1}^2, \tag{7}$$

and, therefore, the second Galerkin matrices for different $\gamma \in C^\infty$ are spectrally equivalent. (Note that the norm $||\cdot||_{-1}$ is somewhat irrelevant to the problem, for all functions u of physical interest lie at least in $H^{-1/2}$. This norm is still what has made things trivial.) We formulate the result as follows.

Theorem 3.1 *Let $a(x,y) = -\frac{1}{2\pi} \log|x-y|$ or $\frac{\mathbf{i}}{4} H_0^{(1)}(\kappa|x-y|)$. Assume that γ_1 and γ_2 are sufficiently smooth such that the corresponding operators $\mathcal{A}_1$ and $\mathcal{A}_2$ satisfy (7). Then the second Galerkin matrices $[\mathcal{A}_1^*\mathcal{A}_1]_n$ and $[\mathcal{A}_2^*\mathcal{A}_2]_n$ are spectrally equivalent.*

Remark. For finitely smooth γ and even for Lipschitz boundaries the inequalities (7) are still valid [8] (for Lipschitz domains, the Hilbert scale boils down to $|s| \le 1$). Consequently, the above premises are fulfilled.

In practice, if we use step functions on a uniform grid with stepsize h and compute the outer integrals in the quantities $(\mathcal{A}u_{jn}, u_{in})$ and $(\mathcal{A}u_{jn}, \mathcal{A}u_{in})$ by the rectangular rule using the same nodes, then

$$\underline{[\mathcal{A}^*\mathcal{A}]_n} = h^{-1} \, \underline{[\mathcal{A}]_n^*} \, \underline{[\mathcal{A}]_n}, \tag{8}$$

where computed matrices are underlined. If the stepsize is not constant, say, h_j for the jth function, then

$$\underline{[\mathcal{A}^*\mathcal{A}]_n} = \underline{[\mathcal{A}]_n^*} \, \mathrm{diag}\,\{h_j^{-1}\} \, \underline{[\mathcal{A}]_n}. \tag{9}$$

In a bit less straightforward way, this observation can be extended also to other quadrature rules and different trial functions.

4 Proper clusters

In this section, we consider conditions under which two Galerkin matrices are properly equivalent. Let H and H' be Banach spaces which are dual with respect to a bounded Hermitian bilinear form (f,u), $f \in H'$ and $u \in H$, and consider a continuous linear operator $\mathcal{A}$ from H to H'.

The Galerkin method can be viewed as the projection method wherein the solution $u \in H$ to the equation

$$\mathcal{A}u = f$$

is approximated by the solutions $u_n \in \mathfrak{U}_n \equiv \operatorname{span}\{u_{1n}, \ldots, u_{nn}\}$ to the finite-dimensional equations

$$\mathcal{A}_n u_n = f_n, \quad \mathcal{A}_n = \mathcal{Q}_n \mathcal{A} \mathcal{P}_n, \quad f_n = \mathcal{Q}_n f.$$

Here $\mathcal{P}_n = \mathcal{P}_n^2$ is a projector in H onto $\mathcal{P}_n H = \mathfrak{U}_n \subset H$, and $\mathcal{Q}_n = \mathcal{Q}_n^2$ is the dual projector acting on H'. These projectors can be expressed explicitly as follows. Let $f_{jn} \in H'$, $1 \le j \le n$, be the dual system for $u_{in} \in H$, $1 \le i \le n$. This means, by definition, that

$$(f_{jn}, u_{in}) = \begin{cases} 1, & i = j, \\ 0, & i \ne j. \end{cases}$$

Then (cf. [16])

$$\mathcal{P}_n\, u = \sum_{j=1}^{n} u_{jn}\, \overline{(f_{jn}, u)}, \qquad \mathcal{Q}_n\, f = \sum_{i=1}^{n} f_{in}\, (f, u_{in}).$$

In practice, one never cares about these projectors $\mathcal{P}_n$ and $\mathcal{Q}_n$, because the Galerkin matrices are entirely determined by $\mathfrak{U}_n$. If $u_n = \sum\limits_{j=1}^{n} x_{jn}\, u_{jn}$, then the equation $\mathcal{A}_n u_n = f_n$ takes the form

$$\sum_{i=1}^{n} f_{in} \sum_{j=1}^{n} (\mathcal{A} u_{jn}, u_{in})\, x_{jn} = \sum_{i=1}^{n} f_{in}\, (f, u_{in}).$$

Since f_{in} are linearly independent, we equate the coefficients at f_{in} and arrive at the familiar Galerkin system

$$A_n\, \mathbf{x} = \mathbf{b}, \quad \mathbf{x} = \begin{bmatrix} x_{1n} \\ \vdots \\ x_{nn} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} (f, u_{1n}) \\ \vdots \\ (f, u_{nn}) \end{bmatrix},$$

where A_n is the Galerkin matrix defined by (4).

Next two lemmas relate the Galerkin operators with the Galerkin matrices.

Lemma 4.1 *Let $\mathcal{A}$ and $\mathcal{C}$ be linear operators from H to H' defined on $\mathfrak{U}_n$, and A_n and C_n be the corresponding Galerkin matrices. Consider the Galerkin operators $\mathcal{A}_n = \mathcal{Q}_n \mathcal{A} \mathcal{P}_n$ and $\mathcal{C}_n = \mathcal{Q}_n \mathcal{C} \mathcal{P}_n$. Then*

$$\mathcal{A}_n u = \lambda\, \mathcal{C}_n u, \quad u \ne 0, \tag{10}$$

if and only if

$$A_n \mathbf{x} = \lambda\, C_n \mathbf{x}, \quad \mathbf{x} \ne 0, \tag{11}$$

where u and $\mathbf{x} = [x_1, \ldots, x_n]^T$ are related by the formula $u = \sum\limits_{j=1}^{n} x_j u_{jn}$.

Proof. It is easy to verify that (10) is equivalent to

$$\sum_{i=1}^{n} f_{in} \sum_{j=1}^{n} (\mathcal{A}u_{jn}, u_{in})\, x_j = \lambda \sum_{i=1}^{n} f_{in} \sum_{j=1}^{n} (\mathcal{C}u_{jn}, u_{in})\, x_j,$$

which, thanks to the linear independence of f_{in}, is equivalent to (11). □

Lemma 4.2 *Under the hypotheses of Lemma 4.1, assume that the operator $\mathcal{C}_n$ has a trivial null-space in $\mathcal{P}_n H$. Then the matrix C_n is nonsingular, and $C_n^{-1} A_n$ is the matrix representation of $\mathcal{C}_n^{-1}\mathcal{A}_n$ in the basis $u_{1n}, \ldots, u_{nn}$. The same matrix also represents $\mathcal{A}_n\mathcal{C}_n^{-1}$ in the basis $\mathcal{C}_n u_{1n}, \ldots, \mathcal{C}_n u_{nn}$.*

Proof. Let $M = [m_{ij}]$ represent $\mathcal{C}_n^{-1}\mathcal{A}_n$ in the basis $u_{1n}, \ldots, u_{nn}$. Then

$$\mathcal{C}_n^{-1}\mathcal{A}_n u_{jn} = \sum_{l=1}^{n} m_{lj}\, u_{ln}.$$

After multiplying both sides by $\mathcal{C}_n$ and taking scalar products with u_{in}, we obtain

$$(\mathcal{A}_n u_{jn}, u_{in}) = \sum_{l=1}^{n} m_{lj}\, (\mathcal{C}_n u_{ln}, u_{in}),$$

which gives the first assertion. The second is proved similarly. □

From now on, by $\mathcal{A}_n \to \mathcal{A}$ we mean that $\mathcal{A}_n u \to \mathcal{A}u$ for any $u \in H$. Denote by $\mathcal{I}$ and $\mathcal{I}'$ the identity operators on H and H' respectively. If $\mathcal{C}_n = \mathcal{Q}_n\mathcal{C}\mathcal{P}_n$ has only a trivial null-space in $\mathcal{P}_n H$, then its restriction $\mathcal{C}_n : \mathcal{P}_n H \to \mathcal{Q}_n H'$ is a one-to-one mapping. Let $\mathcal{C}_n^{-1} : \mathcal{Q}_n H' \to \mathcal{P}_n H$ denote the corresponding inverse operator.

Lemma 4.3 *Assume that $\mathcal{P}_n \to \mathcal{I}$ and $\mathcal{Q}_n \to \mathcal{I}'$. Let $\mathcal{C}$ be a continuously invertible operator from H to H', and assume that the Galerkin method is stable for $\mathcal{C}$ in the sense that, for some $c > 0$, there is a compact operator $\mathcal{T}$ from H to H' such that*

$$\mathrm{Re}\,((\mathcal{C} - \mathcal{T})u, u) \geq c||u||_H, \quad \forall u \in \mathfrak{U}_n.$$

Then $\mathcal{C}_n^{-1} : \mathcal{Q}_n H' \to \mathcal{P}_n H$ exists for all sufficiently large n and

$$\mathcal{C}_n^{-1}\, \mathcal{Q}_n \to \mathcal{C}^{-1}.$$

Behind the thesis is the well-known formula "approximation and stability $\Rightarrow$ convergence" in conjunction with the convergence-preserving property of compact perturbations. The proof can be found, for example, in [16]. Also, note that if $\mathcal{T} = 0$ then $\mathcal{C}_n$ and hence C_n are invertible for all n.

Now assume that $\mathcal{A}$ and $\mathcal{C}$ are continuously invertible linear operators from H to H' and, in addition, the difference

$$\mathcal{K} = \mathcal{A} - \mathcal{C}$$

is a compact operator from H to H'. Then

$$\mathcal{A}\mathcal{C}^{-1} = \mathcal{I} + \mathcal{K}\mathcal{C}^{-1},$$

where $\mathcal{K}\mathcal{C}^{-1}$ is a compact operator from H' to H'. From the operator theory, its spectrum contains zero and for any $\varepsilon > 0$ there are only finitely many nonzero spectral points, each being an eigenvalue of finite algebraic multiplicity.

For the corresponding Galerkin matrices for $\mathcal{K}$, $\mathcal{A}$, and $\mathcal{C}$ we obtain, obviously,

$$K_n = A_n - C_n$$

and, if C_n is nonsingular,

$$C_n^{-1}\, A_n = I + C_n^{-1}\, K_n.$$

To proceed further, we need to know whether and how the eigenvalues of $C_n^{-1}\, K_n$ approximate the spectrum of the compact operator $\mathcal{K}\mathcal{C}^{-1}$. Owing to Lemma 4.2, the question reduces to the approximation properties of the nonzero eigenvalues of $\mathcal{K}_n\, \mathcal{C}_n^{-1}\, \mathcal{Q}_n$.

Theorem 4.1 *Let $\mathcal{K} : H \to H'$ be compact and $\mathcal{C} : H \to H'$ obey the hypotheses of Lemma 4.3. Then for sufficiently large n the Galerkin matrices C_n are nonsingular, and the eigenvalues of $C_n^{-1}K_n$ have a proper cluster at zero.*

Proof. We are interested in counting only the nonzero eigenvalues of $C_n^{-1}K_n$. From Lemma 4.2, we know that they coincide (together with algebraic multiplicities) with those of $\mathcal{K}_n\mathcal{C}_n^{-1}\mathcal{Q}_n$. Owing to Lemma 4.3, $\mathcal{C}_n^{-1}\,\mathcal{Q}_n \to \mathcal{C}^{-1}$, and, as a consequence,

$$\mathcal{T}_n \equiv \mathcal{K}_n\, \mathcal{C}_n^{-1}\, \mathcal{Q}_n \;\to\; \mathcal{T} \equiv \mathcal{K}\, \mathcal{C}^{-1}.$$

Let us show that $\mathcal{T}_n$, $n \in \mathbb{N}$, are *collectively compact* operators, i.e. those for which the set $\{\mathcal{T}_n u \;:\; n \in \mathbb{N},\; ||u||_H \le 1\}$ is totally bounded. Note that $\mathcal{K}_n$ are collectively compact. First, since $\mathcal{P}_n \to \mathcal{I}$, $\mathcal{Q}_n \to \mathcal{I}'$, and $\mathcal{K}$ is compact, we can prove that $\mathcal{Q}_n\mathcal{K}$ converges to $\mathcal{K}$ not only pointwise but also in norm, and any convergent-in-norm sequence of bounded operators is collectively compact [1]. Second, if a collectively compact sequence is postmultiplied by a uniformly bounded sequence of operators, then the resulting sequence is also collectively compact [1]. It remains to note that, by

the Banach–Steinhaus theorem, the operators $\mathcal{Q}_n$, $\mathcal{P}_n$, and $\mathcal{C}_n^{-1}\,\mathcal{Q}_n$ must be uniformly bounded in norm (due to their pointwise convergence to $\mathcal{C}^{-1}$).

Thus, $\mathcal{T}_n$ are collectively compact, $\mathcal{T}_n \to \mathcal{T}$, and therefore $\mathcal{T}$ is a compact operator. With that much we are aware [1] that if $z \neq 0$ is an eigenvalue of algebraic multiplicity α for $\mathcal{T}$, then there is an open set $O(z) \ni z$ that contains exactly α (with multiplicities) eigenvalues of $\mathcal{T}_n$ for all sufficiently large n, and these eigenvalues converge to z as $n \to \infty$. This obviously completes the proof. □

Corollary. *Under the premises of Theorem 4.1, the eigenvalues of* $I + C_n^{-1}K_n$ *have a proper cluster at unity.*

We are now in the position to apply Theorem 4.1 to the integral operators from equations (1) and (2).

Theorem 4.2 *Let* $a(x,y) = -\frac{1}{2\pi}\log|x-y|$ *or* $\frac{\mathbf{i}}{4}H_0^{(1)}(\kappa|x-y|)$, *and assume that* γ_1 *and* γ_2 *are two boundaries belonging to* $H^{2+\delta}$ *for some* $\delta > 0$. *Let the corresponding operators* $\mathcal{A}_1$ *and* $\mathcal{A}_2$ *have a trivial null-space. Then their Galerkin matrices* $[\mathcal{A}_1]_n$ *and* $[\mathcal{A}_2]_n$ *are nonsingular for sufficiently large* n, *and* $[\mathcal{A}_2]_n^{-1}\,[\mathcal{A}_1]_n$ *have a proper eigenvalue cluster at unity.*

Proof. The operators in question enjoy the assumptions of Lemma 4.3 with $H = H^{-1/2}$ and $H' = H^{1/2}$ [7]. In line with Theorem 4.1, it suffices to prove that the difference $\mathcal{K} = \mathcal{A}_1 - \mathcal{A}_2$ is a compact operator from $H^{-1/2}$ to $H^{1/2}$. Let the kernel function of $\mathcal{K}$ be $k(\tau,t)$.

If γ_1 and γ_2 belong to C^∞, the assertion follows straightforwardly from the observation that

$$k(\tau,t) \equiv \log\left|\frac{\gamma_1(\tau)-\gamma_1(t)}{\gamma_2(\tau)-\gamma_2(t)}\right| \in C^\infty([0,2\pi]\times[0,2\pi]);$$

and in fact, this infinitely smooth kernel generates a so-called *smoothing operator*. For Hankel kernel, we use (3) to get the same conclusion. If γ_1 and γ_2 are relegated to $H^{2+\delta}$ with some $\delta > 0$, then

$$k(\tau,t) \in H^{1+\delta}([0,2\pi]\times[0,2\pi]).$$

In this case, by a direct calculation we can show that if $u \in H^{-1/2}$ then $\mathcal{K}u \in H^{1/2+\varepsilon}$ with some $\varepsilon > 0$. Since the imbedding $H^{1/2+\varepsilon} \subset H^{1/2}$ is compact (for example, see [7]), $\mathcal{K}$ is a compact operator from $H^{-1/2}$ to $H^{1/2}$, and this completes the proof. □

We do not claim that the smoothness assumptions imposed on γ can not be mellowed. Nevertheless, *some* smoothness assumptions are certainly needed.

5 CIAO preconditioners

For a circle, the integral operators in (1) and (2) are of the form

$$(\mathcal{C}u)(\tau) = \int_0^{2\pi} c(\tau - t)u(t)\,dt$$

with a 2π-periodic kernel function $c(\cdot)$. We call $\mathcal{C}$ a *circulant integral operator* [4, 5, 9]. Circulant integral operator appears more often implicitly than explicitly: it usually comes in as a result of the approximation of more general operators. Consider a general operator

$$(\mathcal{A}u)(\tau) = \int_0^{2\pi} a(\tau, t)u(t)\,dt,$$

where $a(\tau, t)$ is 2π-periodic both in τ and t, we may choose a function

$$c(\tau - t) \;\approx\; a(\tau, t) \tag{12}$$

and regard $\mathcal{C}$ as a *circulant integral approximate operator* (CIAO) for $\mathcal{A}$.

The choice of the CIAO depends on how we understand (12). Some recent constructions [4, 5, 9] were inspired by T. Chan's idea of optimal circulant preconditioners [6] (developed further in [14]). A matrix $C = [c_{ij}]$, $0 \le i, j \le n-1$, is called a circulant matrix if c_{ij} is constant along any wrapped diagonal $i - j = k \,(\mathrm{mod}\, n)$. The optimal circulant for a matrix $A = [a_{ij}]$ is the minimizer of $||A - C||_F$. It can be shown that the diagonals c_j of the optimal circulant are given by

$$c_j = \frac{1}{n} \sum_{l=0}^{n-1} a_{j+l\,l},$$

where the values of a_{ij} are viewed as n-periodical in both indices [6].

This suggests that we consider

$$c(\phi) = \frac{1}{2\pi} \int_0^{2\pi} a(\phi + \theta, \theta)\,d\theta, \tag{13}$$

as the kernel for the approximate operator. The corresponding CIAO is called *optimal* in [9, 5]. If C is the optimal circulant preconditioner for a matrix A, then [14]

$$(Ae_j, e_j) = (Ce_j, e_j)$$

for all columns e_j of the Fourier matrix. In the optimal CIAO case, quite the same arises in the form [4]

$$(\mathcal{A}e_j, e_j) = (\mathcal{C}e_j, e_j) \tag{14}$$

for any $e_j(t) = e^{\mathbf{i}jt}$, $j = 0, \pm 1, \ldots$. It makes it easy to disclose that the optimal CIAO $\mathcal{C}$ minimizes the Hilbert–Schmidt norm (assume that $a, c \in L_2$)

$$|||\mathcal{A} - \mathcal{C}||| \equiv \int_0^{2\pi} \int_0^{2\pi} |a(\tau, t) - c(\tau - t)|^2 \, d\tau \, dt,$$

which imparts a variational sense to the name given.

In [4], the optimal CIAO was used as a preconditioner for the Galerkin matrices from (1). The main theoretical result was the spectral equivalence. In fact, it was proved that $\mathcal{A}$ and $\mathcal{C}$ are both Hermitian positive definite operators from $H^{-1/2}$ to $H^{1/2}$ (cf. Lemmas 2.1 and 2.2 above). Now we can add to this that $\mathcal{C}$ is also a superlinear preconditioner for $\mathcal{A}$.

Theorem 5.1 *Let $a(x, y) = -\frac{1}{2\pi} \log|x - y|$ or $\frac{\mathbf{i}}{4} H_0^{(1)}(\kappa|x - y|)$, and assume that $\gamma \in H^{2+\delta}$ for some $\delta > 0$, and the corresponding operator $\mathcal{A}$ has a trivial null-space in $H^{-1/2}$. Let $\mathcal{C}$ denote the optimal CIAO for $\mathcal{A}$, and assume that $\mathcal{C}$ has a trivial null-space in $H^{-1/2}$. Then the Galerkin matrices $[\mathcal{A}]_n$ and $[\mathcal{C}]_n$ are nonsingular for sufficiently large n, and $[\mathcal{C}]_n^{-1}$ $[\mathcal{A}]_n$ have a proper eigenvalue cluster at unity.*

Proof. We need to ascertain first that $\mathcal{C}$ enjoys the hypotheses of Lemma 4.3. To this end, we can exploit (14) and easily adopt the proof from [4] that $\mathcal{C}$ is positive definite in the logarithmic kernel case. Take into account that

$$k(\tau, t) = \log \left| \frac{\gamma(\tau) - \gamma(t)}{e^{\mathbf{i}\tau} - e^{\mathbf{i}t}} \right| \in H^{1+\delta}([0, 2\pi] \times [0, 2\pi]), \quad \delta > 0.$$

Then, the optimal CIAO kernel is the sum of the two terms: the optimal CIAO kernel for $a(\tau, t) - k(\tau, t)$ and that for $k(\tau, t)$, respectively. Since the operator with kernel $a(\tau, t) - k(\tau, t)$ is already an integral circulant operator, its optimal CIAO is the operator itself. The kernel of the optimal CIAO for $k(\tau, t)$ inherits the smoothness of $k(\tau, t)$ (see (13)), and we can use the arguments from the proof of Theorem 4.2. For the Hankel kernel, we use (3) to get the same conclusion. □

Remark. In the logarithmic kernel case, $\mathcal{C}$ inherits the positive definiteness from $\mathcal{A}$, which is the case when the diameter of γ is less than 1.

For the equation (2), there is also a possibility of building up the optimal CIAO for $\mathcal{A}^*\mathcal{A}$. For simplicity, let $\gamma \in C^\infty$. Then $\mathcal{A}^*\mathcal{A}$ can be viewed as a Hermitian positive definite operator from H^{-1} to H^0. Its kernel is

$$\tilde{a}(\tau, t) = \int_0^{2\pi} a(\tau, \theta) \, \overline{a(t, \theta)} \, d\theta.$$

Let $\tilde{\mathcal{C}}$ denote the corresponding optimal CIAO. Based on (14) and the method of [4], it is easy to check that it takes the positive definiteness from $\mathcal{A}^*\mathcal{A}$. Then, owing to Lemma 2.2, matrices $[\mathcal{A}^*\mathcal{A}]_n$ and $[\tilde{\mathcal{C}}]_n$ are spectrally equivalent.

Now, it can be argued that one might find the optimal CIAO preconditioners less beneficial than some simpler CIAO. For integral equations on closed curves, the CIAO generated from a cicular contour may be a better choice as it is easier to construct and invert. Usually it does not count much if it gives one or two iterations more than the optimal CIAO. The circle-based CIAO in fact solves some approximation problem of the form (12): it makes the difference $a(\tau, t) - c(\tau - t)$ *smoother*. Moreover, when γ is infinitely smooth, it makes it infinitely smooth too.

It is remarkable that the performance of the circle-based CIAO does not depend on the contour, or in other words, on a particular problem from a particular class. It is entirely determined by the class – in the integral equation context, by the kernel singularity, or more precisely, by the principal symbol of the integral operator. The conclusion is that as long as the boundary has a certain degree of smoothness, the singularity is what matters, not the geometry of the contour. The latter is somehow captured by the optimal CIAO, though with just a minor effect. In the next section these theoretical findings are illustrated by numerical experiments.

6 Numerical illustrations

Consider the Hankel-kernel equation (2) on an elliptic boundary with half-axes $a = 0.4$, $b = 0.05$. Let A_n denote the Galerkin matrices for the step functions on a uniform grid. To illustrate the above theory, we consider the following preconditioners C_n:

- the Galerkin matrices corresponding to the same equation for the ellipse with half-axes $a = 0.05$ and $b = 0.4$ (this ridiculuos choice is not for practice, just to check the clustering property);
- the Galerkin (circulant) matrices corresponding to the same equation for the circle of radius 0.1;
- the Galerkin (circulant) matrices corresponding to the logarithmic kernel equation (1) for the circle of radius 0.1.

In Figure 1, we can see that the eigenvalues and singular values of the matrices $A_n C_n^{-1}$ have a cluster at unity in all three cases ($n = 256$).

Next we consider the performance of the preconditioned conjugate gradients applied to the normal equations with the coefficient matrices $A_n^* A_n$ preconditioned by $C_n^* C_n$. As C_n, we take the optimal circulant matrices and

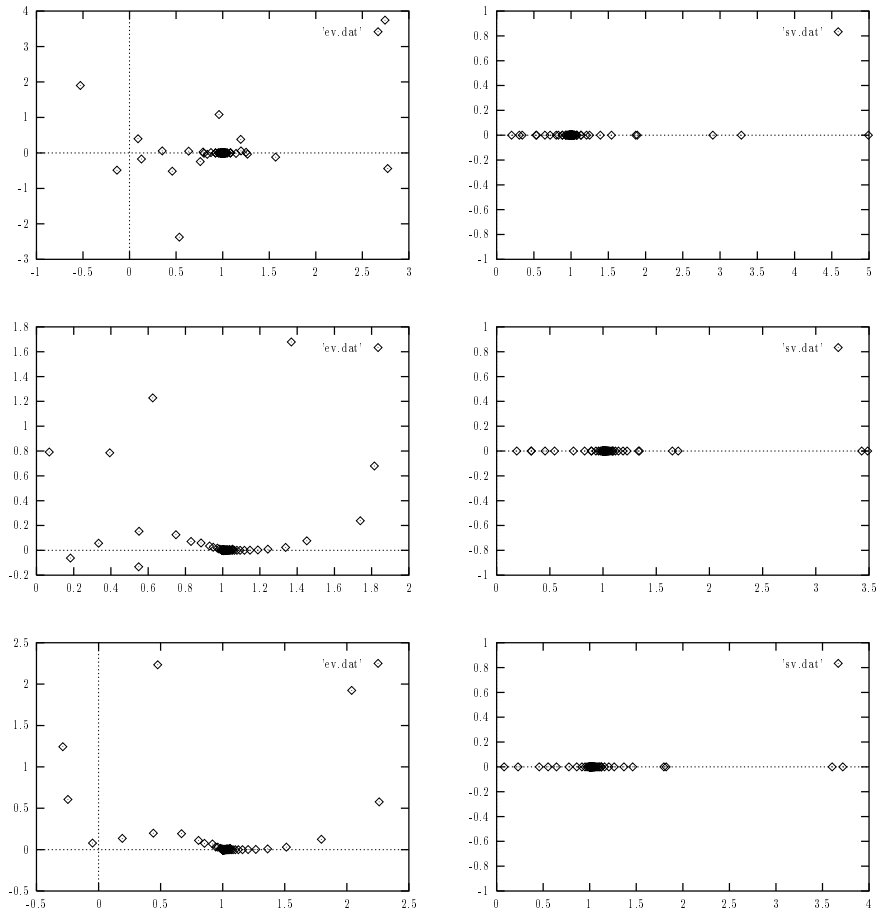

Figure 1: Eigenvalues and singular values for $A_n C_n^{-1}$.

the circle-based CIAO matrices for the circle of radius 0.1. The stopping criterion is when the residual becomes less than ε times the initial residual where we take $\varepsilon = 10^{-9}$. The right-hand side function is $f(t) = |\sin t|^{3/2}$. The Hankel kernel is taken with $\kappa = 15$. The contour is the ellipse with $a = 0.8$ and $b = 0.05$. We also depict the approximate relative errors computed by comparing the solutions in the L_2 norm for two neighboring grids. The numbers of iterations required for convergence are given in Table 1.

Matrix size	64	128	256	512	1024	2048
No preconditioner	33	63	114	182	266	359
Optimal circulant preconditioner	10	11	11	11	12	11
Circle-based preconditioner	11	12	12	12	12	12
Relative error	0.68%	0.30%	0.15%	0.07%	0.04%	

Table 1: The preconditioned conjugate gradients for normal equations in the Hankel-kernel case.

References

[1] P. Anselone, *Collectively Compact Operator Approximation Theory and Applications to Integral Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1971.

[2] O. Axelsson and G. Lidskog, On the rate of convergence of the preconditioned conjugate gradient method, *Numer. Math.* 48: 499–523 (1986).

[3] R. Chan and G. Strang, Toeplitz equations by conjugate gradients with circulant preconditioners, *SIAM J. Sci. Stat. Comput.* 10: 104–119 (1989).

[4] R. Chan, H. Sun, and W. Ng, Circulant preconditioners for ill-conditioned boundary integral equations from potential equations, *Int. J. Numer. Meth. Engng.* 43: 1505–1521 (1998).

[5] R. Chan, X. Jin, and M. Ng, Circulant integral operators as preconditioners for Wiener–Hopf equations, *Integr. Equat. Oper. Th.* 21: 12–23 (1995).

[6] T. Chan, An optimal circulant preconditioner for Toeplitz systems, *SIAM J. Sci. Stat. Comput.* 9: 766–771 (1988).

[7] G. Chen and J. Zhou, *Boundary Element Methods*, Academic Press, London, 1992.

[8] M. Costabel, Boundary integral equations on Lipschitz domains: elementary results, *SIAM J. Math. Anal.* 19 (3): 613–625 (1988).

[9] I. Gohberg, M. Hanke, and I. Koltracht, Fast preconditioned conjugate gradient algorithms for Wiener–Hopf equations, *SIAM J. Numer. Anal.* 31: 429–443 (1994).

[10] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*, University of California Press, Berkeley, CA, 1958.

[11] G. Hsiao and W. Wendland, A finite element method for some integral equations of the first kind, *I. Math. Anal. Appl.* 58: 449–481 (1977).

[12] S. Prossdorf and B. Silberman, *Numerical Analysis for Integral and Related Operator Equations*, Birkhauser Verlag, Basel, 1991.

[13] S. Serra Capizzano and E. Tyrtyshnikov, Any circulant-like preconditioner for multilevel matrices is not superlinear, *SIAM J. Matrix Anal. Appl.*, to appear.

[14] E. Tyrtyshnikov, Optimal and superoptimal circulant preconditioners, *SIAM J. Matrix Anal. Appl.* 12(2): 459–473 (1992).

[15] E. Tyrtyshnikov, A unifying approach to some old and new theorems on distribution and clustering, *Linear Algebra Appl.* 232: 1–43 (1996).

[16] E. Tyrtyshnikov, *A Brief Introduction to Numerical Analysis*, Birkhauser, Boston, 1997.