Preconditioned Conjugate Gradient Methods for Integral Equations of the Second Kind Defined on the Half-Line*

Raymond H. Chan† Fu-Rong Lin‡

April 19, 2001

Abstract

We consider solving integral equations of the second kind defined on the half-line $[0, \infty)$ by the preconditioned conjugate gradient method. Convergence is known to be slow due to the non-compactness of the associated integral operator. In this paper, we construct two different circulant integral operators to be used as preconditioners for the method to speed up its convergence rate. We prove that if the given integral operator is close to a convolution-type integral operator, then the preconditioned systems will have spectrum clustered around 1 and hence the preconditioned conjugate gradient method will converge superlinearly. Numerical examples are given to illustrate the fast convergence.

Abbreviated Title: Preconditioners for Integral Equations.

Keywords: Preconditioned conjugate gradient methods, integral equations, circulant integral operators.

AMS(MOS) Subject Classifications. 45E10, 65R20.

*Research supported in part by HKRGC grant no. 221600070.

†Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong

‡Department of Mathematics, University of Hong Kong, Pokfulam Road, Hong Kong

1 Introduction

In this paper, we study numerical solutions to integral equations of the second kind defined on the half-line. More precisely, we consider the equation

$$y(t) + \int_0^\infty a(t,s)y(s)ds = g(t), \quad 0 \le t < \infty \tag{1}$$

where $g(t)$ is a given function in $L_2[0,\infty)$ and the kernel function $a(s,t)$ is in $L_2(\mathbb{R}^2)$. One way of solving (1) is by the projection method [3] where the solution $y(t)$ is approximated by the solution $y_\tau(t)$ of the finite-section equation

$$y_\tau(t) + \int_0^\tau a(t,s)y_\tau(t) = g(t), \quad 0 \le t \le \tau. \tag{2}$$

It is shown in [3] that

$$\lim_{\tau\to\infty} ||y_\tau - y||_{L_p[0,\tau)} = 0, \quad 1 \le p < \infty.$$

The finite-section equation (2) can be solved numerically by either direct or iterative methods. For a fixed τ, the finite-section operator A_τ defined by

$$(A_\tau x)(t) = \begin{cases} \int_0^\tau a(t,s)x(s)ds, & 0 \le t \le \tau, \\ 0, & t > \tau. \end{cases} \tag{3}$$

is a compact operator. Therefore, the spectrum of the operator $I + A_\tau$ is clustered around 1 and hence solving (2) by iterative methods such as the conjugate gradient (CG) method will be less expensive than direct methods. However, as $\tau \to \infty$, the spectrum of A_τ becomes dense in that of A, where A is defined as

$$Ax(t) = \int_0^\infty a(t,s)x(s)ds, \quad 0 \le t < \infty,$$

and hence the convergence rate of the CG method will deteriorate, see the numerical results in §5.

One way of speeding up the convergence rate of the CG method is to apply a preconditioner to (2). Thus instead of solving (2), we solve the preconditioned equation

$$(I + H_\tau)^{-1}(I + A_\tau)y_\tau(t) = (I + H_\tau)^{-1}g(t). \tag{4}$$

We will call the operator H_τ a preconditioner for the operator A_τ. A good preconditioner H_τ is an operator that is close to A_τ in some norm and yet the operator equation

$$(I + H_\tau)x(t) = f(t) \tag{5}$$

is easier to solve than (2) for arbitrary function $f \in L_2[0,\tau]$. A class of candidates is the class of operators of the form

$$H_\tau x(t) = \int_0^\tau h_\tau(t-s)x(s)ds, \quad 0 \le t \le \tau,$$

where the kernel functions h_τ are periodic in $[0,\tau]$. They are called *circulant integral operators* in [5]. The eigenfunctions and eigenvalues of the operator H_τ are given by

$$u_m(t) = \frac{1}{\sqrt{\tau}} e^{2\pi imt/\tau}, \quad m \in \mathbb{Z}, \tag{6}$$

and

$$\lambda_m = \sqrt{\tau}(h_\tau, u_m)_\tau = \sqrt{\tau}\int_0^\tau h_\tau(t)\bar{u}_m(t)dt, \quad m \in \mathbb{Z}. \tag{7}$$

Therefore, (5) can be solved effeciently by using the Fourier transforms.

The convergence rate of solving the preconditioned system (4) with CG method depends on how close the operator $(I + H_\tau)$ is to the operator $(I + A_\tau)$, see Axelsson and Barker [1, p.28]. Therefore, a natural idea is to find the circulant integral operator H_τ that minimizes the difference $A_\tau - H_\tau$ in some norm over all circulant integral operators. In this paper, we will consider the minimization in the Hilbert-Schmidt norm $|||\cdot|||$. We will construct two different kinds of circulant integral preconditioners for A_τ. The first one minimizes $|||A_\tau - H_\tau|||$ and the second one minimizes $|||I - (I + H_\tau)^{-1}(I + A_\tau)|||$. Following the terminologies used in the study of Toeplitz matrices, we call the first minimizer the *optimal preconditioner* and denote it by $\mathcal{P}(A_\tau)$, see [2]. The circulant operator that minimizes the second one will be called the *super-optimal* preconditioner, see [6].

We will prove some of the properties of the operator $\mathcal{P}$. In particular, we will show that for self-adjoint operators A_τ,

$$\inf_{||x||_2=1}(A_\tau x, x)_\tau \le \inf_{||x||_2=1}(\mathcal{P}(A_\tau)x, x)_\tau \le \sup_{||x||_2=1}(\mathcal{P}(A_\tau)x, x)_\tau \le \sup_{||x||_2=1}(A_\tau x, x)_\tau,$$

where

$$(a, b)_\tau \equiv \int_0^\tau a(t)\bar{b}(t)dt.$$

Thus if A_τ is a positive operator, then so is $\mathcal{P}(A_\tau)$. We also show that the operator norms of $\mathcal{P}$ derived from the 2-norm and the Hilbert-Schmidt norm are both equal to 1. We then show that the super-optimal preconditioners are good preconditioners for integral equations with convolution kernels because the spectra of the preconditioned operators will be clustered around 1 for sufficiently large τ. As a corollary, we prove that the preconditioned conjugate gradient (PCG) method will converge superlinearly for integral operators A that are close to convolution-type operators in the Hilbert-Schmidt norm.

The outline of the paper is as follows. In §2, we construct the optimal circulant integral preconditioners $\mathcal{P}(A_\tau)$ for integral operators A_τ and study some of its spectral properties. In §3, we construct the super-optimal circulant integral preconditioners. The convergence analysis of the preconditioned operators for convolution-type operators and for general integral operators are discussed in §4. Finally numerical results are given in §5.

2 Optimal Circulant Integral Operator

In this section, we discuss some of the properties of the optimal circulant integral operator $\mathcal{P}(A_\tau)$ for integral operator A_τ given in (3). The preconditioner $\mathcal{P}(A_\tau)$ is defined to be the circulant integral operator that minimizes the Hilbert-Schmidt norm

$$|||A_\tau - H_\tau|||^2 \equiv \int_0^\tau \int_0^\tau (a(t,s) - h_\tau(t-s))(\bar{a}(t,s) - \bar{h}_\tau(t-s))dsdt \qquad (8)$$

over all circulant integral operators H_τ. We first give the expression of the kernel function of $\mathcal{P}(A_\tau)$.

Lemma 1 *Let $a(\cdot,\cdot) \in L_2([0,\tau]^2)$. Then the kernel function of $\mathcal{P}(A_\tau)$ is given by*

$$p_{A_\tau}(s) = \frac{1}{\tau}\int_{\tau-s}^{\tau} a(v+s-\tau, v)dv + \frac{1}{\tau}\int_0^{\tau-s} a(v+s,v)dv. \qquad (9)$$

In terms of Fourier expansions,

$$p_{A_\tau}(t-s) = \sum_{m=-\infty}^{\infty} (A_\tau u_m, u_m)_\tau u_m(t)\bar{u}_m(s), \quad 0 \le s,t \le \tau. \qquad (10)$$

Proof: Since $a(\cdot,\cdot) \in L_2([0,\tau]^2)$, we can write, by using Fourier expansions

$$a(t,s) = \sum_{m,n=-\infty}^{\infty} \nu_{m,n} u_m(t)\bar{u}_n(s), \quad 0 \le s,t \le \tau, \qquad (11)$$

where $u_m(t)$ is given in (6) and

$$\nu_{m,n} \equiv \int_0^\tau \int_0^\tau a(t,s)u_n(s)\bar{u}_m(t)dsdt = (A_\tau u_n, u_m)_\tau, \quad m,n \in \mathbb{Z}. \qquad (12)$$

Let H_τ be a circulant integral operator with kernel function h_τ in $L_2[-\tau,\tau]$. By means of Fourier expansion, we can write

$$h_\tau(t-s) = \sum_{m=-\infty}^{\infty} \lambda_m u_m(t)\bar{u}_m(s), \quad 0 \le s,t \le \tau$$

where λ_m is given in (7). Combining this with (11) and using the orthogonality of u_n, we can rephrase the distance (8) as

$$|||A_\tau - H_\tau|||^2 = \sum_{m=-\infty}^{\infty} |\nu_{m,m} - \lambda_m|^2 + \sum_{\substack{m,n=-\infty \\ m\neq n}}^{\infty} |\nu_{m,n}|^2.$$

Clearly, the expression becomes minimal if and only if $\lambda_m = \nu_{m,m} = (A_\tau u_m, u_m)_\tau$ for all integers m. Thus (10) follows.

To obtain (9), we observe from (12) that for all integer m,

$$\lambda_m = \nu_{m,m} = \int_0^\tau \int_0^\tau a(t,s) u_m(s) \bar{u}_m(t) ds dt = \frac{1}{\sqrt{\tau}} \int_0^\tau \int_0^\tau a(t,s) \bar{u}_m(t-s) ds dt.$$

By using Fubini's theorem and the substitutions $v = t - s$ and $s = s$, we get

$$\begin{aligned}
\lambda_m &= \frac{1}{\sqrt{\tau}} \left\{ \int_{-\tau}^{0} \int_{-s}^{\tau} a(v+s, v) \bar{u}_m(s) dv ds + \int_0^\tau \int_0^{\tau - s} a(v+s, v) \bar{u}_m(s) dv ds \right\} \\
&= \sqrt{\tau} \int_0^\tau \left\{ \frac{1}{\tau} \int_{\tau - s}^{\tau} a(v + s - \tau, v) dv + \frac{1}{\tau} \int_0^{\tau - s} a(v+s, v) dv \right\} \bar{u}_m(s) ds.
\end{aligned}$$

Comparing this with (7), we see that the kernel function p_{A_τ} is given by (9). $\square$

In Lemma 3 below, we study some of the properties of the operator $\mathcal{P}$ which are useful in proving convergence in §4. We first note the following result whose proof is trivial and will be omitted.

Lemma 2 *Let A_τ and B_τ be two integral operators with kernel functions a_τ and b_τ respectively. If $a_\tau(\cdot,\cdot)$ and $b_\tau(\cdot,\cdot)$ are in $L_2([0,\tau]^2)$, then the kernel function $d_\tau(\cdot,\cdot)$ of the composite operator $D_\tau = A_\tau B_\tau$ is also in $L_2([0,\tau]^2)$ and is given by*

$$d_\tau(t,s) = \int_0^\tau a_\tau(t,w) b_\tau(w,s) dw. \qquad (13)$$

Moreover, we have $||D_\tau||_2 \leq ||A_\tau||_2 ||B_\tau||_2$.

Lemma 3 *The following properties of the operator $\mathcal{P}$ hold.*

(i) $\mathcal{P}$ is a linear projection operator, i.e., $\mathcal{P}(\mathcal{P}(A_\tau)) = \mathcal{P}(A_\tau)$.

(ii) Let H_τ be any circulant integral operator. Then we have

$$\mathcal{P}(H_\tau A_\tau) = H_\tau \mathcal{P}(A_\tau) = \mathcal{P}(A_\tau) H_\tau = \mathcal{P}(A_\tau H_\tau).$$

(iii) $||\mathcal{P}||_2 = |||\mathcal{P}||| = 1$ *where* $||\cdot||_2$ *and* $|||\cdot|||$ *are the operator norms of* $\mathcal{P}$ *derived from the 2-norm and the Hilbert-Schmidt norm respectively.*

(iv) If A_τ *is self-adjoint, i.e.,* $\bar{a}(t,s) = a(s,t)$*, then so is* $\mathcal{P}(A_\tau)$*; and we have*

$$\inf_{||x||_2=1}(A_\tau x, x)_\tau \le \inf_{||x||_2=1}(\mathcal{P}(A_\tau)x, x)_\tau \le \sup_{||x||_2=1}(\mathcal{P}(A_\tau)x, x)_\tau \le \sup_{||x||_2=1}(A_\tau x, x)_\tau.$$

In particular, if A_τ *is positive, i.e.* $(A_\tau x, x)_\tau \ge 0$ *for all* $x \in L_2[0,\tau]$*, then* $\mathcal{P}(A_\tau)$ *is also positive.*

Proof: The proof of (i) is obvious and will be omitted. For (ii), we first prove that the operators $\mathcal{P}(H_\tau A_\tau)$ and $H_\tau \mathcal{P}(A_\tau)$ have the same kernel function. Let the kernel functions of A_τ and H_τ be

$$a_\tau(t,s) = \sum_{m,n=-\infty}^{\infty} \nu_{m,n} u_m(t)\bar{u}_n(s), \ \ 0 \le s,t \le \tau$$

and

$$h_\tau(t-s) = \sum_{m=-\infty}^{\infty} \lambda_m u_m(t)\bar{u}_n(s), \ \ 0 \le s,t \le \tau$$

respectively. By (13) the kernel function of $H_\tau A_\tau$ at the point (t,s) is given by

$$\int_0^\tau \sum_{m=-\infty}^{\infty} \lambda_m u_m(t)\bar{u}_m(w) \sum_{m,n=-\infty}^{\infty} \nu_{m,n} u_m(w)\bar{u}_n(s) ds = \sum_{m,n=-\infty}^{\infty} \lambda_m \nu_{m,n} u_m(t)\bar{u}_n(s).$$

By Lemma 1, the kernel function of $\mathcal{P}(H_\tau A_\tau)$ at the point $(t-s)$ is therefore given by

$$\sum_{m=-\infty}^{\infty} \lambda_m \nu_{m,m} u_m(t)\bar{u}_m(s).$$

It is easy to check that $H_\tau \mathcal{P}(A_\tau)$ has the same kernel function. Thus the first equality in (ii) holds. The third equality in (ii) can be proved likewise. The second equality in (ii) follows from the fact that the composite of two circulant integral operators with the same period is commutative, see for instance [4, p.181].

To prove (iii), we first note that for an arbitrary function $x(s) = \sum_{m=-\infty}^{\infty} \alpha_m u_m(s)$ in $L_2[0,\tau]$, we have by (10)

$$||\mathcal{P}(A_\tau)x(t)||_2^2 = \sum_{m=-\infty}^{\infty} |\alpha_m|^2 |\nu_{m,m}|^2 \le \sup_{m\in\mathbb{Z}} |\nu_{m,m}|^2 ||x||_2^2,$$

where $\nu_{m,m}$ are defined in (12). Thus $||\mathcal{P}(A_\tau)||_2 \le \sup_{m\in\mathbb{Z}} |\nu_{m,m}|$. However, for each integer m,

$$||\mathcal{P}(A_\tau)||_2 \ge ||\mathcal{P}(A_\tau)u_m||_2 = |\nu_{m,m}|,$$

it follows that $||\mathcal{P}(A_\tau)||_2 = \sup_{m\in\mathbb{Z}} |\nu_{m,m}|$. On the other hand, since u_n are orthonormal, we see that for all $n \in \mathbb{Z}$,

$$||A_\tau||_2^2 \ge ||A_\tau u_n||_2^2 = ||\sum_{m=-\infty}^{\infty} \nu_{m,n} u_m(t)||_2^2 = \sum_{m=-\infty}^{\infty} |\nu_{m,n}|^2.$$

Hence,

$$||A_\tau||_2^2 \ge \sup_{n\in\mathbb{Z}} \sum_{m=-\infty}^{\infty} |\nu_{m,n}|^2 \ge \sup_{m\in\mathbb{Z}} |\nu_{m,m}|^2 = ||\mathcal{P}(A_\tau)||_2^2.$$

Since when A_τ is circulant, $||A_\tau||_2 = ||\mathcal{P}(A_\tau)||_2$, it follows that $||\mathcal{P}||_2 = 1$. That $|||\mathcal{P}||| = 1$ can be proved similarly. In fact, by (11) and (10), we see that

$$|||A_\tau|||^2 = \sum_{m,n=-\infty}^{\infty} |\nu_{m,n}|^2 \ge \sum_{m=-\infty}^{\infty} |\nu_{m,m}|^2 = |||\mathcal{P}(A_\tau)|||^2.$$

Finally, we prove (iv). It is clear that if A_τ is self-adjoint, then $\mathcal{P}(A_\tau)$ is also self-adjoint. By (11) and (12), we see that

$$\inf_{||x||_2=1} (A_\tau x, x)_\tau \le \inf_{m\in\mathbb{Z}} (A_\tau u_m, u_m)_\tau.$$

However, by (10),

$$\inf_{m\in\mathbb{Z}} (A_\tau u_m, u_m)_\tau = \inf_{m\in\mathbb{Z}} (\mathcal{P}(A_\tau) u_m, u_m)_\tau = \inf_{||x||_2=1} (\mathcal{P}(A_\tau) x, x)_\tau.$$

The inequality for the supremum in (iv) can be proved likewise. $\square$

3 Super-Optimal Integral Preconditioners

In this section, we consider another type of circulant integral preconditioners which are obtained by minimizing the Hilbert-Schmidt norm

$$|||I - (I + H_\tau)^{-1}(I + A_\tau)||| \tag{14}$$

over all circulant integral operators H_τ such that $(I+H_\tau)^{-1}$ exists. The reason we consider this preconditioner is that in the PCG method (cf (4)), we want the preconditioned operator $(I+H_\tau)^{-1}(I+A_\tau)$ to be as close to the identity operator I as possible. If the minimum of (14) is attained by $I+C_\tau$, then we call C_τ the *super-optimal* circulant integral preconditioner for A_τ. In order to find the kernel function for C_τ, we first characterize the inverse of operators of the form $I+H_\tau$.

Lemma 4 *Let H_τ be a circulant integral operator with kernel function h_τ and eigenvalues λ_m given in (7). If $I+H_\tau$ is invertible, then its inverse is given by*

$$(I+H_\tau)^{-1} = I - K_\tau$$

where K_τ is also a circulant integral operator with kernel function

$$k_\tau(t-s) = \sum_{n=-\infty}^{\infty} \left(\frac{\lambda_n}{1+\lambda_n}\right) u_n(t)\bar{u}_n(s). \tag{15}$$

Proof: Since $I+H_\tau$ is invertible, $1+\lambda_n \neq 0$ for all integers n. Moreover, since h_τ is in $L_2[-\tau,\tau]$, $\sum_{n=-\infty}^{\infty}|\lambda_n|^2 < \infty$ and therefore $|1+\lambda_n| \geq 1/2$ for all $|n|$ sufficiently large. In particular, the function k_τ defined in (15) is a τ-periodic function in $L_2[-\tau,\tau]$. By (13), the kernel function of $H_\tau K_\tau$ at the point $(t-s)$ is given by

$$\int_0^\tau \sum_{m=-\infty}^{\infty} \lambda_m u_m(t)\bar{u}_m(w) \sum_{n=-\infty}^{\infty}\left(\frac{\lambda_n}{1+\lambda_n}\right) u_n(w)\bar{u}_n(s)dw = \sum_{n=-\infty}^{\infty}\left(\frac{\lambda_n^2}{1+\lambda_n}\right) u_n(t)\bar{u}_n(s).$$

From this, it is straightforward to check that the kernel function of

$$H_\tau - K_\tau - H_\tau K_\tau = (I+H_\tau)(I-K_\tau) - I$$

is the zero function. Thus $(I+H_\tau)^{-1} = (I-K_\tau)$. $\square$

In view of this Lemma, the problem of minimizing the norm (14) becomes the problem of minimizing $|||I-(I-K_\tau)(I+A_\tau)|||$ over all circulant integral operator K_τ. In this formulation, it is easy to find the super-optimal circulant preconditioner C_τ for A_τ.

Lemma 5 *Let $a(\cdot,\cdot) \in L_2([0,\tau]^2)$ be such that $I+\mathcal{P}(A_\tau)$ is invertible. Let C_τ be the super-optimal circulant integral operator for A_τ. Then $I+C_\tau$ is invertible and the kernel function c_τ of C_τ is given by*

$$c_\tau(t-s) = \sum_{m=-\infty}^{\infty}\left(\frac{\zeta_m+\nu_{m,m}}{1+\bar{\nu}_{m,m}}\right) u_m(t)\bar{u}_m(s), \tag{16}$$

where $\nu_{m,n} = (A_\tau u_n, u_m)_\tau$ and $\zeta_m = \sum_{n=-\infty}^{\infty}|\nu_{m,n}|^2$.

Proof: Let H_τ be any circulant integral operator such that $I + H_\tau$ is invertible. Denote the kernel function of H_τ by

$$h_\tau(t-s) = \sum_{m=-\infty}^{\infty} \lambda_m u_m(t)\bar{u}_m(s).$$

By Lemma 4, $(I + H_\tau)^{-1} = I - K_\tau$ where the kernel function of K_τ is given in (15). Thus

$$|||I - (I + H_\tau)^{-1}(I + A_\tau)||| = |||K_\tau + K_\tau A_\tau - A_\tau|||.$$

By (13) and (15), the kernel function of $K_\tau + K_\tau A_\tau - A_\tau$ at the point (t, s) is given by

$$\sum_{m,n=-\infty}^{\infty} \left\{ \frac{\delta_{m,n}\lambda_m}{1+\lambda_m} + \frac{\lambda_m \nu_{m,n}}{1+\lambda_m} - \nu_{m,n} \right\} u_m(t)\bar{u}_n(s) = \sum_{m,n=-\infty}^{\infty} \left(\frac{\delta_{m,n}\lambda_m - \nu_{m,n}}{1+\lambda_m} \right) u_m(t)\bar{u}_n(s),$$

where $\delta_{m,n}$ denotes the Kronecker symbol. By the definition of the Hilbert-Schmidt norm,

$$|||I - (I + H_\tau)^{-1}(I + A_\tau)|||^2 = \sum_{m,n=-\infty}^{\infty} |\frac{\delta_{m,n}\lambda_m - \nu_{m,n}}{1+\lambda_m}|^2.$$

It is clear that the above expression is minimized if and only if the term

$$\frac{|\lambda_m|^2 - \lambda_m \bar{\nu}_{m,m} - \bar{\lambda}_m \nu_{m,m} + \zeta_m}{|1+\lambda_m|^2}$$

is minimized for all integers m. However, by differentiating this quotient with respect to the real and imaginary parts of λ_m, we see that the minimum will be obtained if we set

$$\lambda_m = \frac{\zeta_m + \nu_{m,m}}{1 + \bar{\nu}_{m,m}}$$

for all m. Hence (16) follows. We note that by (10), the assumption that $I + \mathcal{P}(A_\tau)$ is invertible implies that the denominator $1 + \bar{\nu}_{m,m}$ in the above expression is nonzero for all m. Moreover, since

$$1 + \lambda_m = \frac{\sum_{n=-\infty, n\neq m}^{\infty} |\nu_{m,n}|^2 + |1 + \bar{\nu}_{m,m}|^2}{1 + \bar{\nu}_{m,m}} \neq 0,$$

we see that $I + C_\tau$ is invertible. $\square$

4 Convergence Analysis

In this section we consider the convergence rate of the optimal and super-optimal circulant integral preconditioners for solving integral equations of the second kind. We begin with equations having convolution kernel first. In this case, the convergence analysis for the optimal circulant integral preconditioners has already been studied.

Lemma 6 (Gohberg, Hanke and Koltracht [5]) *Let A be a self-adjoint, positive convolution-type integral operator with kernel function $a(\cdot) \in L_1(\mathbb{R})$. Let $\mathcal{P}(A_\tau)$ be the optimal circulant integral operator of A_τ. Then for each $\epsilon > 0$ there is a positive integer ρ and a $\tau^* > 0$ such that for each $\tau \geq \tau^*$, there exists a decomposition*

$$A_\tau - \mathcal{P}(A_\tau) = Q_\tau + R_\tau, \tag{17}$$

with self-adjoint operators Q_τ and R_τ satisfying $||Q_\tau||_2 \leq \epsilon$ and rank $R_\tau \leq \rho$. *Moreover, the spectrum of*

$$(I + \mathcal{P}(A_\tau))^{-1/2}(I + A_\tau)(I + \mathcal{P}(A_\tau))^{-1/2}$$

has at most ρ eigenvalues outside interval $(1-\epsilon, 1+\epsilon)$.

This lemma basically states that the spectrum of the preconditioned operator is clustered around 1. Hence using standard theory of the PCG method, see for instance [1, p.28], we can conclude that the method with the optimal preconditioner converges superlinearly. We now prove a similar result for the super-optimal preconditioner.

Theorem 1 *Let A be a self-adjoint, positive convolution-type integral operator with kernel function $a(\cdot) \in L_1(\mathbb{R}) \cap L_2(\mathbb{R})$. Let C_τ be the super-optimal circulant integral preconditioner for A_τ. Then for each $\epsilon > 0$, there is a positive integer ρ and a $\tau^* > 0$ such that for each $\tau > \tau^*$, there exists a decomposition*

$$A_\tau - C_\tau = S_\tau + T_\tau, \tag{18}$$

where S_τ and T_τ are self-adjoint operators satisfying $\|S_\tau\|_2 \leq \epsilon$ and rank $T_\tau \leq \rho$. *Moreover, the spectrum of $(I + C_\tau)^{-1/2}(I + A_\tau)(I + C_\tau)^{-1/2}$ has at most ρ eigenvalues outside interval $(1-\epsilon, 1+\epsilon)$.*

By comparing (17) and (18), we see that Theorem 1 can be easily proved if we can show that

$$\lim_{\tau \to \infty} ||\mathcal{P}(A_\tau) - C_\tau||_2 = 0.$$

The next Lemma and Corollary are devoted to proving this limit.

Lemma 7 *Let A be a self-adjoint convolution-type integral operator with kernel function $a \in L_1(\mathbb{R}) \cap L_2(\mathbb{R})$. Then*

$$\lim_{\tau\to\infty} ||\mathcal{P}({A_\tau}^2) - \mathcal{P}(A_\tau)^2||_2 = 0. \tag{19}$$

Proof: Since $a \in L_1(\mathbb{R})$, for each $\epsilon > 0$, there is a $\tau_\epsilon > 0$ such that $\int_{\tau_\epsilon}^{\infty} |a(s)|ds < \epsilon$. Define

$$\tau^* \equiv \max\{\tau_\epsilon ||a||_1/\epsilon, 36\tau_\epsilon^2 ||a||_2^2/\epsilon, 2\tau_\epsilon\}. \tag{20}$$

For each $\tau > \tau^*$, we decompose the difference $A_\tau - \mathcal{P}(A_\tau)$ as

$$A_\tau - \mathcal{P}(A_\tau) = E_\tau + F_\tau \tag{21}$$

where E_τ and F_τ are self-adjoint operators with kernel functions

$$e_\tau(s) = \begin{cases} a(s) - p_{A_\tau}(s), & |s| \le \tau - \tau_\epsilon \\ 0, & |s| > \tau - \tau_\epsilon \end{cases} \tag{22}$$

and

$$f_\tau(s) = \begin{cases} a(s) - p_{A_\tau}(s), & \tau - \tau_\epsilon \le |s| \le \tau \\ 0, & \text{otherwise} \end{cases} \tag{23}$$

respectively. Using the decomposition (21) and Lemma 3(ii), we then have

$$\begin{aligned} \mathcal{P}({A_\tau}^2) - \mathcal{P}(A_\tau)^2 &= \mathcal{P}[(A_\tau - \mathcal{P}(A_\tau))(A_\tau - \mathcal{P}(A_\tau))] \\ &= \mathcal{P}[(E_\tau + F_\tau)^2] \\ &= \mathcal{P}({E_\tau}^2 + E_\tau F_\tau + F_\tau E_\tau) + \mathcal{P}({F_\tau}^2). \end{aligned}$$

Therefore by Lemma 3(iii), we then have

$$||\mathcal{P}({A_\tau}^2) - \mathcal{P}(A_\tau)^2||_2 \le ||{E_\tau}^2 + E_\tau F_\tau + F_\tau E_\tau||_2 + ||\mathcal{P}({F_\tau}^2)||_2. \tag{24}$$

We now estimate the 2-norm of the two terms on the right hand side of (24).

For the first term, we need estimates of $||E_\tau||_2$ and $||F_\tau||_2$. From (22), (20) and (9), we get

$$\begin{aligned} ||E_\tau||_2 &\le ||e_\tau||_1 = 2\int_0^{\tau-\tau_\epsilon} \frac{s}{\tau}|a(s-\tau) - a(s)|ds \\ &\le 2\left\{\int_0^{\tau-\tau_\epsilon} |a(s-\tau)|ds + \int_0^{\tau_\epsilon} \frac{\tau_\epsilon}{\tau}|a(s)|ds + \int_{\tau_\epsilon}^{\tau-\tau_\epsilon} |a(s)|ds\right\} \\ &\le 2\left\{\int_{-\tau}^{-\tau_\epsilon} |a(v)|dv + \frac{\tau_\epsilon}{\tau}||a||_1 + \epsilon\right\} < 6\epsilon. \end{aligned}$$

From (21) and Lemma 3(iii), we see that

$$||F_\tau||_2 \le ||\mathcal{P}(A_\tau)||_2 + ||A_\tau||_2 + ||E_\tau||_2 \le 2||A_\tau||_2 + ||E_\tau||_2 \le 2||a||_1 + 6\epsilon.$$

Thus for the first term in the right hand side of (24), we have

$$||{E_\tau}^2 + E_\tau F_\tau + F_\tau E_\tau||_2 \le 36\epsilon^2 + 12\epsilon(2||a||_1 + 6\epsilon) = 108\epsilon^2 + 24||a||_1\epsilon. \qquad (25)$$

Next we estimate second term in the right hand side of (24). By (13), the kernel function of ${F_\tau}^2$ is given by

$$\hat{f}_\tau(s,t) = \int_0^\tau f_\tau(s-w) f_\tau(w-t) dw$$

and by (9), the kernel function of $\mathcal{P}({F_\tau}^2)$ is

$$p_{{F_\tau}^2}(s) = \frac{1}{\tau}\int_{\tau-s}^{\tau} \hat{f}_\tau(v+s-\tau, v)dv + \frac{1}{\tau}\int_0^{\tau-s} \hat{f}_\tau(v+s, v)dv.$$

Using the definition of f_τ in (23), we can check that

$$\begin{cases} |\hat{f}_\tau(s,t)| \le ||f_\tau||_2^2, & \tau - \tau_\epsilon \le s,t \le \tau \text{ or } 0 \le s,t \le \tau_\epsilon, \\ |\hat{f}_\tau(s,t)| = 0, & \text{otherwise.} \end{cases}$$

Using this, it is straightforward to check that

$$\begin{cases} p_{{F_\tau}^2}(s) = 0, & \tau_\epsilon \le s \le \tau - \tau_\epsilon, \\ |p_{{F_\tau}^2}(s)| \le 2\tau_\epsilon ||f_\tau||_2^2/\tau, & 0 \le s \le \tau_\epsilon \text{ or } \tau - \tau_\epsilon \le s \le \tau. \end{cases}$$

Therefore, it follows that for all $\tau > \tau^*$

$$||p_{{F_\tau}^2}||_1 = \int_0^{\tau_\epsilon} |p_{{F_\tau}^2}(s)|ds + \int_{\tau-\tau_\epsilon}^{\tau} |p_{{F_\tau}^2}(s)|ds \le \frac{4||f_\tau||_2^2 \tau_\epsilon^2}{\tau}.$$

We now claim that $||f_\tau||_2^2 \le 9||a||_2^2$. If this is true, then by our choice of τ^* in (20), $||\mathcal{P}({F_\tau}^2)||_2 \le ||p_{{F_\tau}^2}||_1 < \epsilon$. Putting this result and (25) back into (24), our Lemma follows.

Thus it remains to prove that $||f_\tau||_2^2 \le 9||a||_2^2$. But by (23),

$$||f_\tau||_2^2 \le ||p_{A_\tau}||_2^2 + ||a_\tau||_2^2 + 2||a_\tau||_2 ||p_{A_\tau}||_2$$

and by (9),

$$|p_{A_\tau}(t)| = |\frac{\tau - t}{\tau} a(t) + \frac{t}{\tau} a(t-\tau)| \le |a(t)| + |a(t-\tau)|, \quad 0 \le t \le \tau.$$

Using Schwarz's inequality, we have $||p_{A_\tau}||_2^2 \le 4||a||_2^2$. Hence $||f_\tau||_2^2 \le 9||a||_2^2$. $\square$

Corollary 1 *Let A be a self-adjoint convolution-type integral operator with kernel function $a \in L_1(\mathbb{R}) \cap L_2(\mathbb{R})$. Let $\nu_{m,n} = (Au_m, u_n)_\tau$. Then*

$$\lim_{\tau \to \infty} \sup_{m \in \mathbb{Z}} \sum_{\substack{n=-\infty \\ n \neq m}}^{\infty} |\nu_{m,n}|^2 = 0. \tag{26}$$

Proof: By (11) and (13), the kernel function of ${A_\tau}^2$ at the point $(t-s)$ is given by

$$\sum_{m,k=-\infty}^{\infty} (\sum_{n=-\infty}^{\infty} \nu_{m,n}\nu_{n,k}) u_m(t)\bar{u}_k(s).$$

Therefore by (10), the kernel function $\mathcal{P}({A_\tau}^2)$ at the point $(t-s)$ is given by

$$\sum_{m=-\infty}^{\infty} (\sum_{n=-\infty}^{\infty} \nu_{m,n}\nu_{n,m}) u_m(t)\bar{u}_m(s) = \sum_{m=-\infty}^{\infty} \left(\sum_{n=-\infty}^{\infty} |\nu_{m,n}|^2 \right) u_m(t)\bar{u}_m(s).$$

However by (10) and (13) again, the kernel function of $\mathcal{P}(A_\tau)^2$ at the point $(t-s)$ is given by

$$\sum_{m=-\infty}^{\infty} |\nu_{m,m}|^2 u_m(t)\bar{u}_m(s).$$

Thus the kernel function of $\mathcal{P}({A_\tau}^2) - \mathcal{P}(A_\tau)^2$ at the point $(t-s)$ is given by

$$\sum_{m=-\infty}^{\infty} \left(\sum_{\substack{n=-\infty \\ n \neq m}}^{\infty} |\nu_{m,n}|^2 \right) u_m(t)\bar{u}_m(s).$$

Hence (26) follows from (19). $\square$

Now we are ready to prove Theorem 1.

Proof of Theorem 1. By Lemma 6, $I + \mathcal{P}(A_\tau)$ is invertible. Hence by Lemma 5, the super-optimal preconditioner C_τ of A_τ exists. Using (10) and (16) and noting that $\nu_{m,m} = (A_\tau u_m, u_m)_\tau$ is real, we see that the kernel function of $C_\tau - \mathcal{P}(A_\tau)$ is given by

$$\sum_{m=-\infty}^{\infty} \left(\frac{\zeta_m - \nu_{m,m}^2}{1 + \nu_{m,m}} \right) u_m(t)\bar{u}_m(s) = \sum_{m=-\infty}^{\infty} \left(\frac{1}{1 + \nu_{m,m}} \sum_{\substack{n=-\infty \\ n \neq m}}^{\infty} |\nu_{m,n}|^2 \right) u_m(t)\bar{u}_m(s),$$

Since A is a positive operator, $1+\nu_{m,m} \geq 1$ for all m. Hence (26) implies that

$$\lim_{\tau\to\infty} ||C_\tau - \mathcal{P}(A_\tau)||_2 = \lim_{\tau\to\infty} \sup_m \left(\frac{1}{1+\nu_{m,m}} \sum_{\substack{n=-\infty \\ n\neq m}}^{\infty} |\nu_{m,n}|^2 \right) = 0.$$

By combining this result with (17), equation (18) follows with $S_\tau = Q_\tau + \mathcal{P}(A_\tau) - C_\tau$ and $T_\tau = R_\tau$.

Finally we prove the clustering of the spectrum of the preconditioned operator. By (16), the eigenvalues of the operator $I + C_\tau$ are equal to

$$\frac{\zeta_m + 2\nu_{m,m} + 1}{1+\nu_{m,m}} \geq 1.$$

Thus $I + C_\tau$ is a positive operator with $||(I+C_\tau)^{-1}||_2 \leq 1$. Using (18), we therefore have

$$\begin{aligned} & (I+C_\tau)^{-1/2}(I+A_\tau)(I+C_\tau)^{-1/2} - I \\ = \; & (I+C_\tau)^{-1/2}T_\tau(I+C_\tau)^{-1/2} + (I+C_\tau)^{-1/2}S_\tau(I+C_\tau)^{-1/2} = \tilde{T}_\tau + \tilde{S}_\tau, \end{aligned}$$

where clearly we have rank $\tilde{T}_\tau =$ rank $T_\tau \leq \rho$ and

$$||\tilde{S}_\tau||_2 \leq ||S_\tau||_2 ||(I+C_\tau)^{-1/2}||_2^2 \leq \epsilon.$$

Thus, by the min-max theorem [4, p.123], Theorem 1 follows. □

In the following, we extend the convergence result to operators that are close to convolution operators in the Hilbert-Schmidt norm. We first note that if E is an operator with finite Hilbert-Schmidt norm, then E_τ, the restriction operator of E onto $[0,\tau)$, can be decomposed as the sum of a small norm operator and a low rank operator.

Lemma 8 *Let E be a self-adjoint operator with kernel function $e(s,t)$. If*

$$\int_0^\infty \int_0^\infty |e(t,s)|^2 dt ds \leq M < \infty \tag{27}$$

for some constant M, then for each given $\epsilon > 0$, at most M/ϵ^2 eigenvalues of E_τ outside the interval $(-\epsilon, \epsilon)$.

Proof: The Lemma follows easily by noting that the eigenvalues λ_n of E_τ satisfies

$$\sum_n |\lambda_n|^2 \leq \int_0^\tau \int_0^\tau |e(t,s)|^2 dt ds \leq M,$$

see [7, p.32] for instance. □

Combining this together with Lemma 6 or Theorem 1, we have our main theorem.

Theorem 2 *Let B be a self-adjoint integral operator with kernel function $b(t,s) = a(t-s) + e(t,s)$ where $a(\cdot) \in L_1(\mathbb{R}) \cap L_2(\mathbb{R})$ and $e(s,t)$ satisfies (27). Let A_τ be the operator on $[0,\tau)$ with kernel function $a(\cdot)$ and D_τ be the optimal (or super-optimal) preconditioner for A_τ. Then for each $\epsilon > 0$, there is a positive integer ρ and a $\tau^* > 0$ such that for each $\tau > \tau^*$, there exists a decomposition*

$$B_\tau - D_\tau = S_\tau + T_\tau,$$

where S_τ and T_τ are self-adjoint operators satisfying $\|S_\tau\|_2 \leq \epsilon$ and $\operatorname{rank} T_\tau \leq \rho$. Moreover, the spectrum of $(I + D_\tau)^{-1/2}(I + B_\tau)(I + D_\tau)^{-1/2}$ has at most ρ eigenvalues outside interval $(1-\epsilon, 1+\epsilon)$.

Thus if $I + B_\tau$ is preconditioned by $I + D_\tau$, we expect fast convergence.

5 Numerical Results

In this section, we test the convergence performance of the optimal and super-optimal preconditioners for solving integral equations of the second kind. In the tests, the operators are all discretized by the rectangular quadrature rule. The rule using N points will yield N-by-N matrices. Random vectors are used as initial guesses and are kept the same for all preconditioners. The stopping criterion is $||\mathbf{r}_k||_2/||\mathbf{r}_0||_2 < 10^{-7}$, where $\mathbf{r}_k$ is the residual vector of preconditioned conjugate gradient method after k iterations.

We first test the preformance on integral equations with convolution kernels. Two kernel functions were tested and they are:

1. $a_1(t) = \dfrac{\sigma}{1 + e^{|t|}}$.
2. $a_2(t) = \dfrac{\sigma}{1 + |t|^{1.01}}$.

We note that in real applications, σ^{-1} is the regularization parameter used and is usually small. In all our experiments, we set $\sigma^{-1} = 0.01$. We note that the discretization matrices formed from $I + A_\tau$ and $I + \mathcal{P}(A_\tau)$ are Toeplitz and circulant matrices respectively. Tables 1 and 2 gives the numbers of iterations required for convergence for different preconditioners. In the tables, O, S and I denote that the optimal, super-optimal and no preconditioner is used respectively. We see from the tables that the two preconditioners perform almost the same when τ is large and their performances are much better than that of no preconditioning.

Table 1. Number of iterations for test function $a_1(t)$.

N	τ=16			τ=64			τ=256			τ=1024		
	O	S	I	O	S	I	O	S	I	O	S	I
1024	7	11	38	6	7	65	6	6	59	5	5	23
2048	8	11	40	6	8	66	6	6	71	5	5	40
4096	7	11	39	6	7	66	6	6	75	6	6	59
8192	8	11	42	7	8	68	5	7	75	6	6	70

Table 2. The number of iterations for test function $a_2(t)$.

N	τ=16			τ=64			τ=256			τ=1024		
	O	S	I	O	S	I	O	S	I	O	S	I
1024	7	10	48	7	8	77	7	7	66	7	7	31
2048	7	11	52	7	9	85	8	8	96	8	8	52
4096	7	10	48	7	8	79	8	8	109	8	8	83
8192	8	10	54	8	9	88	9	9	123	8	8	122

Next we test our algorithms for general integral equations. We tried the following two kernel functions:

1. $a_3(s,t) = \sigma(\frac{1}{1+(s-t)^2} + e^{-\sqrt{t^2+s^2}})$.

2. $a_4(s,t) = \dfrac{\sigma}{1+(s-t)^2(1+(1+t+s)^{-1})}$.

Again σ^{-1} was set to 0.01 in our experiments. We remark that with $a(t) = \sigma/(1+t^2)$, the function $a_3(s,t)$ satisifies the assumptions in Theorem 2 while $a_4(s,t)$ does not. In the tests, we used the optimal preconditioner for the operator A_τ as the preconditioner in both cases. The convergence results are listed in Table 3. We see that the numbers of iterations of PCG method are smaller than that of the CG method considerably. We also emphasize that if we look at the numbers of iterations for fixed mesh-size, i.e. τ/N is fixed, we see that the numbers are increasing rapidly with increasing τ for the non-preconditioned systems. This indiciates that the finite-section equation (2) is less well-conditioned as $\tau \to \infty$. However, the numbers stay basically unchanged for the preconditioned one.

Table 3. The numbers of iterations for $a_3(s,t)$ and $a_4(s,t)$.

	$a_3(s,t)$								$a_4(s,t)$							
	τ=16		τ=32		τ=64		τ=128		τ=16		τ=32		τ=64		τ=128	
N	O	I	O	I	O	I	O	I	O	I	O	I	O	I	O	I
64	8	45	8	49	*	*	*	*	9	45	8	47	*	*	*	*
128	10	42	8	66	7	61	*	*	9	45	9	66	8	63	*	*
256	10	42	8	65	8	85	7	70	10	45	10	66	9	83	8	69
512	10	40	9	63	8	85	8	97	10	44	9	66	10	84	8	97

References

[1] O. Axelsson and V. Barker, *Finite Element Solution of Boundary Value Problems*, Academic Press, Orlando, 1984.

[2] T. Chan, *An Optimal Circulant Preconditioner for Toeplitz Systems*, SIAM J. Sci. Statist. Comput., 9 (1988), 766–771.

[3] I. Gohberg and I. Fel'dman, *Convolution Equations and Projection Methods for Their Solutions,* Transl. Math. Monographs, 41, Amer. Math. Soc., Providence. RI, 1974.

[4] I. Gohberg and S. Goldberg, *Basic Operator Theory,* Birkhauser, Boston, 1981.

[5] I. Gohberg, M. Hanke and I. Koltracht, *Fast Preconditioned Conjugate Gradient Algorithms for Wiener-Hopf Integral Equations*, to appear.

[6] E. Tyrtyshnikov, *Optimal and Super-optimal Circulant Preconditioners,* SIAM Matrix Anal. Appl., 13 (1992), 459–473.

[7] P. Zabreyko, A. Koshelev, M. Krasnosel'skii, S. Mikhlin, L. Rakovshchik and V. Stet'senko, *Integral Equations – a Reference Text*, Noordhoff International Publishing, Leyden, 1968.