THE SPECTRA OF SUPER–OPTIMAL CIRCULANT PRECONDITIONED TOEPLITZ SYSTEMS

Raymond H. Chan, Xiao–Qing Jin and Man–Chung Yeung

Department of Mathematics

University of Hong Kong

Hong Kong

April 1990

Revised September 1990

Abstract. We study the solutions of Hermitian positive definite Toeplitz systems $A_n x = b$ by the preconditioned conjugate gradient method. The preconditioner, called the "super-optimal" preconditioner, is the circulant matrix T_n that minimizes $\|I - C_n^{-1} A_n\|_F$ over all circulant matrices C_n. The convergence rate is known to be governed by the distribution of the eigenvalues of $T_n^{-1} A_n$. For n-by-n Toeplitz matrix A_n with entries being Fourier coefficients of a positive function in the Wiener class, we find the asymptotic behaviour of the eigenvalues of the preconditioned matrix $T_n^{-1} A_n$ as n increases and prove that they are clustered around one.

Abbreviated Title. Super-optimal Circulant Preconditioners

Key Words. Toeplitz matrix, super-optimal preconditioner, circulant matrix, preconditioned conjugate gradient method

AMS(MOS) Subject Classifications. 65F10, 65F15

§1 Introduction.

In this paper, we discuss the solutions to a class of Hermitian positive definite Toeplitz systems $A_n x = b$ by the preconditioned conjugate gradient method. Strang in [7] proposed using preconditioned conjugate gradient method with circulant preconditioners for solving symmetric positive definite Toeplitz systems. For n-by-n Toeplitz matrix A_n, the number of operations per iteration is of order $O(n \log n)$, as circulant systems can be solved efficiently by the Fast Fourier Transform. R. Chan and Strang [1] then considered using a circulant preconditioner S_n, obtained by copying the central diagonals of A_n and bringing them around to complete the circulant. In that paper, they proved that if the underlying generating function f, the Fourier coefficients of which give the entries of A_n, is a positive function in the Wiener class, then for n sufficiently large, S_n and S_n^{-1} are uniformly bounded in the ℓ_2 norm and the eigenvalues of the preconditioned matrix $S_n^{-1} A_n$ are clustered around one.

T. Chan in [5] proposed another circulant matrix C_n that is obtained by averaging the corresponding diagonals of A_n with the diagonals of A_n being extended to length n by a wrap-around. He proved that such C_n minimizes $\|C_n - A_n\|_F$ over all circulant matrices, where $\|\cdot\|_F$ means the Frobenius norm. It was then shown in R. Chan [2] that $\lim_{n\to\infty} \|C_n - S_n\|_2 = 0$ and hence the spectrum of $C_n^{-1} A_n$ is also clustered around one for sufficiently large n. These results are generalized to Hermitian positive definite Toeplitz systems in R. Chan [3] and a more precise convergence rate of these methods is given there.

Recently, Tyrtyshnikov [8] proposed another circulant preconditioner T_n that minimizes $\|I - T_n^{-1} A_n\|_F$ over all non-singular circulant matrices. In that paper, C_n and T_n are called *optimal* and *super-optimal* preconditioners respectively and it is proved that if A_n is positive definite, then so are C_n and T_n.

In this paper, we will prove that if the generating function f is a positive function in the Wiener class, i.e. its Fourier coefficients are absolutely summable, then T_n and T_n^{-1} are uniformly bounded in the ℓ_2 norm. We also show that $\lim_{n\to\infty} \|T_n - C_n\|_2 = 0$. Hence we can conclude that the eigenvalues of the preconditioned matrix $T_n^{-1}A_n$ are clustered around one. Therefore, if the conjugate gradient method is applied to solve this preconditioned system, we can expect the method to have the same convergence rate as the preconditioned systems $S_n^{-1}A_n$ and $C_n^{-1}A_n$. The outline of the paper is as follows. In §2, we discuss some of the properties of the optimal preconditioner C_n that will be needed later. The spectral results are given in §3 and numerical results are reported in §4.

§2 Properties of the Circulant Operator.

Let $(\mathcal{M}_{n\times n}, \|\cdot\|_1)$ denote the Banach algebra of all n-by-n matrices over the complex field equipped with the ℓ_1 norm $\|\cdot\|_1$. We now define an operator c from $\mathcal{M}_{n\times n}$ into the subalgebra of all n-by-n circulant matrices. For any $A_n = (a_{ij})$ in $\mathcal{M}_{n\times n}$, let

$$c(A_n) \equiv \sum_{j=0}^{n-1} \left(\frac{1}{n} \sum_{p-q\equiv j \pmod n} a_{pq} \right) Q^j. \tag{1}$$

Here Q is the n-by-n circulant matrix

$$Q = \begin{bmatrix} 0 & & & & 1 \\ 1 & 0 & & 0 & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ 0 & & & 1 & 0 \end{bmatrix}. \tag{2}$$

We remark that the coefficients of Q^j in (1) can also be written as

$$\frac{1}{n} \sum_{p-q\equiv j \pmod n} a_{pq} = \frac{1}{n} \operatorname{tr}\,(Q^{-j}A), \tag{3}$$

where $\operatorname{tr}(\cdot)$ means the trace. From (3), it can be checked easily that if $A_n = (a_{i-j})$ is a Toeplitz matrix in $\mathcal{M}_{n\times n}$, then $c(A_n)$ is the circulant preconditioner given in T. Chan [5]. More precisely, the entries $c_{ij} = c_{i-j}$ of $c(A_n)$ are given by

$$c_k = \frac{1}{n}\{k a_{k-n} + (n-k)a_k\}\ , \quad k = 0, \cdots n-1. \tag{4}$$

We will call c the *circulant operator.* In the following, we will investigate some of its properties that will be needed later on.

Theorem 1. *The circulant operator c is a bounded linear operator in the Banach algebra $(\mathcal{M}_{n\times n}, \|\cdot\|_1)$. Moreover, for all $n > 0$,*

$$\|c\| = \sup_{\|A_n\|_1=1} \|c(A_n)\|_1 = 1.$$

Proof. It is clear from the definition (1) that c is linear. To prove that $\|c\| = 1$, we first note that if $A_n = I$, the identity matrix, then $\|c(A_n)\|_1 = \|I\|_1 = 1$. For general A_n, we have

$$\begin{aligned}\|c(A_n)\|_1 &= \sum_{j=0}^{n-1}\left|\frac{1}{n}\sum_{p-q\equiv j \pmod n} a_{pq}\right| \le \frac{1}{n}\sum_{j=0}^{n-1}\sum_{p-q\equiv j \pmod n} |a_{pq}| \\ &= \frac{1}{n}\sum_{i=0}^{n-1}\sum_{k=0}^{n-1}|a_{ik}| \le \frac{1}{n}\cdot n\cdot\|A_n\|_1 = 1.\end{aligned}$$

Hence $\|c\| = 1$ for all n. □

Theorem 2. *Let $A_n \in \mathcal{M}_{n\times n}$ and C_n be any n-by-n circulant matrix in $\mathcal{M}_{n\times n}$. Then*

$$c(C_nA_n) = C_n\cdot c(A_n)\ ,$$

$$c(A_nC_n) = c(A_n)\cdot C_n\ .$$

Proof. Let $C_n = \sum_{j=0}^{n-1} c_jQ^j$ where Q is given by (2). We will prove that $c(C_nA_n) = C_n\cdot c(A_n)$. Since $c(C_nA_n)$ and $C_n\cdot c(A_n)$ are circulant, it suffices to show that their first columns are the same. For $\ell = 0,\cdots,n-1$, by (3), the $(\ell,0)$-th entry of $c(C_nA_n)$ is given by

$$\begin{aligned}[c(C_nA_n)]_{\ell 0} &= \frac{1}{n}\ \mathrm{tr}\ \left(C_nA_nQ^{-\ell}\right) = \frac{1}{n}\ \mathrm{tr}\ \left(\sum_{j=0}^{n-1} c_jQ^jA_nQ^{-\ell}\right) \\ &= \frac{1}{n}\sum_{j=0}^{n-1} c_j\ \mathrm{tr}\ \left(Q^jA_nQ^{-\ell}\right) = \frac{1}{n}\sum_{j=0}^{n-1} c_j\ \mathrm{tr}\ \left(Q^{j-\ell}A_n\right)\ .\end{aligned}$$

To compute the first column of $C_n\cdot c(A_n)$, we first note that by (3),

$$C_n\cdot c(A_n) = \left(\sum_{j=0}^{n-1} c_jQ^j\right)\left(\sum_{k=0}^{n-1}\frac{1}{n}\ \mathrm{tr}\ (Q^{-k}A_n)Q^k\right) = \frac{1}{n}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1} c_j\ \mathrm{tr}\ (Q^{-k}A_n)Q^{j+k}.$$

Hence if e_i denotes the i-th unit vector and $*$ denotes the complex conjugate, then we have

$$[C_n \cdot c(A_n)]_{\ell 0} = e_\ell^*[C_n \cdot c(A_n)]e_0 = \frac{1}{n}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1} c_j \ \mathrm{tr}\ \left(Q^{-k}A_n\right)\delta_{\ell,(j+k)\ (\mathrm{mod}\ n)}.$$

Since $\ell \equiv j+k \pmod n$ implies that $Q^{-k} = Q^{j-\ell}$, we see that

$$[C_n \cdot c(A_n)]_{\ell 0} = \frac{1}{n}\sum_{j=0}^{n-1} c_j \ \mathrm{tr}\ \left(Q^{j-\ell}A_n\right) = [c(C_nA_n)]_{\ell 0},$$

for all $0 \le \ell < n$. Thus $C_n \cdot c(A_n) = c(C_nA_n)$. By similar arguments, we can prove that $c(A_nC_n) = c(A_n)\cdot C_n$. □

Theorems 3 and 4 below are just generalization of Theorems 3.1 and 4.1 in Tyrtyshnikov [8] from the real scalar field to the complex field and their proofs are given in R. Chan, Jin and Yeung [4].

Theorem 3. *If A_n is Hermitian and positive definite, then $c(A_n)$ is Hermitian and positive definite. Moreover we have,*

$$0 < \lambda_{\min}(A_n) \le \lambda_{\min}(c(A_n)) \le \lambda_{\max}(c(A_n)) \le \lambda_{\max}(A_n),$$

where $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ denote the smallest and the largest eigenvalues respectively.

Theorem 4. *Let $A_n \in \mathcal{M}_{n\times n}$ be such that $x^*A_nx > 0$ for all nonzero x. Let T_n^{-1} be the super-optimal circulant preconditioner for A_n, defined by*

$$\|I - T_n^{-1}A_n\|_F = \min \|I - C_n^{-1}A_n\|_F \ ,$$

where the minimum is taken over all n-by-n nonsingular circulant matrices C_n in $\mathcal{M}_{n\times n}$. Then

$$T_n^{-1} = [c(A_nA_n^*)]^{-1}c(A_n^*) \ . \tag{5}$$

§3 Spectra of the Preconditioned Systems.

Let us assume that the Hermitian Toeplitz matrices A_n are finite sections of a fixed simply infinite matrix A_∞, see R. Chan and Strang [1]. Thus the (i,j)-th entries of A_n and A_∞ are a_{i-j}, with $a_k = \overline{a}_{-k}$ for all k. We associate to A_∞ the real-valued generating function

$$f(\theta) = \sum_{-\infty}^{\infty} a_k e^{-ik\theta}$$

defined on $[0, 2\pi)$. We will assume that f is a positive function in the Wiener class, i.e.

$$\sum_{k=-\infty}^{\infty} |a_k| = M < \infty \ . \qquad (6)$$

It easily follows that, see Grenander and Szegö [6], A_n are Hermitian positive definite matrices for all n. Moreover, if $0 < f_{\min} < f < f_{\max} < \infty$, then

$$\sigma(A_n) \subset [f_{\min}, f_{\max}] \ . \qquad (7)$$

For Hermitian A_n, we have, by (5),

$$T_n^{-1} = [c(A_n A_n^*)]^{-1} c(A_n^*) = [c(A_n^2)]^{-1} c(A_n).$$

Hence

$$\begin{aligned} T_n^{-1} A_n &= I + T_n^{-1}[A_n - c(A_n)] + T_n^{-1}[c(A_n) - T_n] \\ &= I + T_n^{-1}[A_n - c(A_n)] + [c(A_n^2)]^{-1}[c(A_n)^2 - c(A_n^2)] \ . \end{aligned} \qquad (8)$$

In the following, we will show that under the assumptions mentioned above, T_n^{-1} and $[c(A_n^2)]^{-1}$ are uniformly bounded in the ℓ_2 norm and that $\lim_{n\to\infty} \|c(A_n)^2 - c(A_n^2)\|_2 = 0$. By recalling that the spectrum of $A_n - c(A_n)$ is clustered around zero, see R. Chan [3], we can then conclude that the spectrum of $T_n^{-1} A_n$ is clustered around one.

Lemma 1. *Let the generating function f of A_n be a positive function in the Wiener class. Then for n sufficiently large, $c(A_n)$, $c(A_n^2)$, T_n and their inverses are all Hermitian, positive definite and uniformly bounded in the ℓ_2 norm.*

Proof. Since f is positive and in the Weiner class, A_n is Hermitian and positive definite for all n and (7) holds. Hence $\sigma\,(A_n^2) \subset \left[f_{\min}^2\,, f_{\max}^2\right]$. By Theorem 3, therefore $c(A_n)$, $c(A_n^2)$ and T_n are all Hermitian and positive definite. Moreover, we also have $\sigma\,(c(A_n)) \subset [f_{\min}, f_{\max}]$ and $\sigma\,(c(A_n^2)) \subset [f_{\min}^2\,,\ f_{\max}^2]$. Finally, we have

$$\|T_n^{-1}\|_2 = \|c(A_n^2)^{-1}c(A_n)\|_2 \le \|c(A_n^2)^{-1}\|_2\|c(A_n)\|_2 \le \frac{f_{\max}}{f_{\min}^2}\ ,$$

and

$$\|T_n\|_2 = \|c(A_n)^{-1}c(A_n^2)\|_2 \le \|c(A_n)^{-1}\|_2\|c(A_n^2)\|_2 \le \frac{f_{\max}^2}{f_{\min}}\ . \quad \square$$

In the following, instead of proving $\|c(A_n)^2 - c(A_n^2)\|_2$ tends to zero, we will prove the stronger result that $\|c(A_n)^2 - c(A_n^2)\|_1$ tends to zero. For simplicity, we will assume that $n = 2m+1$ is odd. The case for even n can be proved similarly. Let S_n be the circulant preconditioner defined in R. Chan and Strang [1]. More precisely, the entries $s_{ij} = s_{i-j}$ of S_n are given by

$$s_k = \begin{cases} a_k & 0 \le k \le m, \\ a_{k-n} & m < k < n. \end{cases} \tag{9}$$

Then by Theorem 3 and the fact that S_n is circulant, we have

$$\begin{aligned} \|c(A_n)^2 - c(A_n^2)\|_1 &= \|c(A_n)^2 - c(A_n^2) - c(A_n)S_n + c(A_n)S_n\|_1 \\ &\le \|c(A_n)\big(c(A_n) - S_n\big)\|_1 + \|c(A_nS_n) - c(A_n^2)\|_1, \end{aligned}$$

for all n. Hence by Theorem 1,

$$\begin{aligned} \|c(A_n)^2 - c(A_n^2)\|_1 &\le \|c(A_n)\|_1\|c(A_n) - S_n\|_1 + \|c(A_nS_n - A_n^2)\|_1 \\ &\le \|A_n\|_1\,\|c(A_n) - S_n\|_1 + \|c\big(A_n(S_n - A_n)\big)\|_1. \end{aligned} \tag{10}$$

The following Lemma shows that the first term in (10) tends to zero as n tends to infinity.

Lemma 2. *Let the generating function f of A_n be a real-valued function in the Wiener class. Let S_n be the circulant preconditioner defined by (9), then*

$$\lim_{n\to\infty} \|S_n - c(A_n)\|_1 = 0\ .$$

Proof. By (4) and (9), it is clear that $B_n \equiv S_n - c(A_n)$ is circulant and its entries $b_{ij} = b_{i-j}$ are given by

$$b_k = \begin{cases} \frac{k}{n}(a_k - a_{k-n}) & 0 \le k \le m, \\ \frac{n-k}{n}(a_{k-n} - a_k) & m < k < n. \end{cases}$$

Since B_n is circulant,

$$\|B_n\|_1 = \sum_{k=0}^{n-1} |b_k| \le 2\sum_{k=1}^{m} \frac{k}{n}|a_k| + 2\sum_{k=m+1}^{n-1} |a_k|.$$

Now all for $\varepsilon > 0$, we can always find an $N_1 > 0$ and an $N_2 > 2N_1$, such that

$$\sum_{k=N_1}^{\infty} |a_k| < \varepsilon \quad \text{and} \quad \frac{1}{N_2}\sum_{k=1}^{N_1} k|a_k| < \varepsilon.$$

Thus for all $n > N_2$,

$$\|B_n\|_1 < \frac{2}{N_2}\sum_{k=1}^{N_1} k|a_k| + 2\sum_{k=N_1+1}^{m} |a_k| + 2\sum_{k=m+1}^{\infty} |a_k| < 6\varepsilon \ . \quad \square$$

Clearly (6) implies that $\|A_n\|_1 \le M$ for all $n > 0$. Hence Lemma 2 implies that the first term of (10) tends to zero as n tends to infinity. The following Lemma shows that the second term in (10) also tends to zero as n tends to infinity.

Lemma 3. *Let the generating function f of A_n be a real-valued function in the Wiener class. Then*

$$\lim_{n\to\infty} \|c\big(A_n(S_n - A_n)\big)\|_1 = 0.$$

Proof. For all $\varepsilon > 0$, by (6), we see that there exists an $N > 0$ such that

$$\sum_{|k|\ge N} |a_k| < \varepsilon \ .$$

For all $n > \max(7N, N^2/\varepsilon)$, we partition $A_n = A_n^{(N)} + A_n^{(n-N)}$ as follows. Let the matrix $A_n^{(N)}$ be the matrix obtained by copying the $2N-1$ central diagonals of A_n and setting other entries to zero. In block form, $A_n^{(N)}$ is given by

$$A_n^{(N)} = \begin{bmatrix} A_N & B_N^* & 0 & 0 & 0 \\ B_N & A_N & \dagger & 0 & 0 \\ 0 & \dagger & \dagger & \dagger & 0 \\ 0 & 0 & \dagger & A_N & B_N^* \\ 0 & 0 & 0 & B_N & A_N \end{bmatrix}, \tag{11}$$

where † denotes nonzero block, A_N is the N-by-N principal submatrix of A_n and B_N is an N-by-N upper triangular Toeplitz matrix given by

$$B_N = \begin{bmatrix} 0 & a_{N-1} & a_{N-2} & \cdots & a_1 \\ & 0 & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & a_{N-2} \\ & \text{\Large 0} & & \ddots & a_{N-1} \\ & & & & 0 \end{bmatrix} . \tag{12}$$

Define $A_n^{(n-N)} = A_n - A_n^{(N)}$. Correspondingly, we partition $S_n - A_n$ as follows. Let $X_n^{(N)}$ be the matrix obtained by setting the $n-2N-1$ central diagonals of $S_n - A_n$ to zero. More precisely, $X_n^{(N)}$ has the following block form:

$$X_n^{(N)} = \begin{bmatrix} 0_N & 0 & 0 & 0 & Z_N^* \\ 0 & 0_N & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0_N & 0 \\ Z_N & 0 & 0 & 0 & 0_N \end{bmatrix} , \tag{13}$$

where 0_N is the N-by-N zero matrix and Z_N is an N-by-N lower triangular Toeplitz matrix given by

$$Z_N = \begin{bmatrix} a_{-N} - a_{n-N} & & & \text{\Large 0} \\ \vdots & \ddots & & \\ \vdots & \ddots & \ddots & \\ a_{-1} - a_{n-1} & \cdots & \cdots & a_{-N} - a_{n-N} \end{bmatrix} . \tag{14}$$

Let $Y_n^{(N)} = S_n - A_n - X_n^{(N)}$. Then by Theorem 1,

$$\begin{aligned} \|c\big(A_n(S_n - A_n)\big)\|_1 &= \|c\big(A_n(X_n^{(N)} + Y_n^{(N)})\big)\|_1 \le \|c\big(A_n X_n^{(N)}\big)\|_1 + \|c\big(A_n Y_n^{(N)}\big)\|_1 \\ &\le \|c\big(A_n^{(N)} X_n^{(N)}\big)\|_1 + \|c\big(A_n^{(n-N)} X_n^{(N)}\big)\|_1 + \|A_n\|_1 \, \|Y_n^{(N)}\|_1 \\ &\le \|c\big(A_n^{(N)} X_n^{(N)}\big)\|_1 + \|A_n^{(n-N)}\|_1 \|X_n^{(N)}\|_1 + \|A_n\|_1 \, \|Y_n^{(N)}\|_1 . \end{aligned} \tag{15}$$

Clearly by (6), $\|A_n\|_1$ and $\|X_n^{(N)}\|_1$ both are bounded above by M. Moreover,

$$\|Y_n^{(N)}\|_1 = \sum_{k=m+1}^{n-N-1} |a_{k-n} - a_k| \le \sum_{k=N+1}^{n-N-1} |a_k| < \varepsilon ,$$

and

$$\|A_n^{(n-N)}\|_1 \le \sum_{|k|\ge N} |a_k| < \varepsilon .$$

Thus (15) becomes

$$\|c\big(A_n(S_n - A_n)\big)\|_1 \le \|c\big(A_n^{(N)} X_n^{(N)}\big)\|_1 + 2M\varepsilon. \tag{16}$$

It remains to estimate $\|c\big(A_n^{(N)} X_n^{(N)}\big)\|_1$.

By (11) and (13), we have

$$A_n^{(N)} X_n^{(N)} = \begin{bmatrix} 0_N & 0 & 0 & 0 & A_N Z_N^* \\ 0 & 0_N & 0 & 0 & B_N Z_N^* \\ 0 & 0 & 0 & 0 & 0 \\ B_N^* Z_N & 0 & 0 & 0_N & 0 \\ A_N Z_N & 0 & 0 & 0 & 0_N \end{bmatrix}. \tag{17}$$

We note that the absolute values of all the elements in $A_n^{(N)} X_n^{(N)}$ are bounded by M^2. In fact, since

$$\|A_N\|_1 \le \sum_{|j|<N-1} |a_j| \le M,$$
$$\|B_N\|_1 \le \sum_{j=1}^{N-1} |a_j| \le M,$$
$$\|Z_N\|_1 \le \sum_{j=1}^{n-1} |a_j| \le M,$$

we see that for any $0 \le i, j < N$,

$$|(A_N Z_N^*)_{ij}| \le \sum_{k=0}^{N-1} |(A_N)_{ik}||(Z_N^*)_{kj}| \le \|A_N\|_1 \|Z_N\|_1 \le M^2 \ .$$

Similar arguments show that for any $0 \le i, j < n$,

$$|(A_n^{(N)} X_n^{(N)})_{ij}| \le M^2 \ . \tag{18}$$

From (17), it is clear that if $n > 7N$, the first column of the circulant matrix $c\big(A_n^{(N)} X_n^{(N)}\big)$ is of the form

$$(0, d_1, \cdots, d_{3N-2}, 0, \cdots, 0, d_{n-3N+2}, \cdots, d_{n-1})^*.$$

By (18), the absolute values of $d_1, \cdots, d_{3N-2}$ and $d_{n-3N+2}, \cdots, d_{n-1}$ are all bounded by NM^2/n. Therefore,

$$\|c(A_n^{(N)} X_n^{(N)})\|_1 \le 2(3N-2)\frac{NM^2}{n} \le \frac{6N^2M^2}{n} < 6M^2\varepsilon \ ,$$

for all $n > N^2/\varepsilon$. Thus by (16), we see that

$$\|c\big(A_n(S_n - A_n)\big)\|_1 \le (6M^2 + 2M)\varepsilon$$

for all $n > \max(7N, N^2/\varepsilon)$. $\square$

By using (10), Lemmas 2 and 3 and the fact that for Hermitian matrices, $\|\cdot\|_2 \le \|\cdot\|_1$, we have the following immediate corollary.

Corollary. *Let the generating function f of A_n be a real-valued function in the Wiener class. Then*

$$\lim_{n\to\infty} \|c(A_n)^2 - c(A_n^2)\|_1 = \lim_{n\to\infty} \|c(A_n)^2 - c(A_n^2)\|_2 = 0.$$

Now we state our main theorem.

Theorem 5. *Let the generating function f of A_n be a positive function in the Wiener class, then the spectrum of $T_n^{-1}A_n - I$ is clustered around zero. More precisely, for all $\varepsilon > 0$, there exist $N_1, N_2 > 0$, such that, for all $n > N_1$, at most N_2 eigenvalues of $T_n^{-1}A_n - I$ have absolute value larger than ε.*

Proof. By Lemma 1 and the fact that any two circulant matrices commute, we see that $T_n^{-\frac{1}{2}}$ is well-defined and is given by

$$T_n^{-\frac{1}{2}} = [c(A_n^2)]^{-\frac{1}{2}}[c(A_n)]^{\frac{1}{2}}. \tag{19}$$

Clearly the spectra of $T_n^{-1}A_n$ and $T_n^{-\frac{1}{2}}A_nT_n^{-\frac{1}{2}}$ are the same. By (8) and (19), we have

$$T_n^{-\frac{1}{2}}A_nT_n^{-\frac{1}{2}} - I = T_n^{-\frac{1}{2}}\big(A_n - c(A_n)\big)T_n^{-\frac{1}{2}} + [c(A_n^2)]^{-\frac{1}{2}}\big(c(A_n)^2 - c(A_n^2)\big)[c(A_n^2)]^{-\frac{1}{2}}. \tag{20}$$

By Lemma 1 and the above Corollary, the ℓ_2 norm of the last term in (20) tends to zero as n tends to infinity. Thus the Theorem follows by noting that the spectrum of $A_n - c(A_n)$ is clustered around zero, see R. Chan [3]. $\square$

Hence, if the conjugate gradient method is applied to solve the preconditioned system $T_n^{-1}A_n$, we can expect the convergence rate to be the same as the preconditioned systems $S_n^{-1}A_n$ and $C_n^{-1}A_n$. This is consistent with the numerical results we have in the next section. We note that a more precise bound on the convergence rate of these systems can be found in R. Chan [3]. The bound depends on the differentiablity of the generating function f.

We remark that although these preconditioned systems converge at the same rate for large n, the costs of generating the corresponding preconditioners are not the same. Clearly S_n defined by (9) can be generated at no cost, C_n defined by (4) can be computed in about $3n/2$ operations and Tyrtyshnikov in [8] proposed an algorithm of finding T_n in $9n\log n + O(n)$ operations. An algorithm of finding T_n in $6n\log n + O(n)$ operations can be found in R. Chan, Jin and Yeung [4].

§4 Numerical Results.

To test the convergence rates of the preconditioners, we have applied the preconditioned conjugate gradient method to $A_n x = b$ with

$$a_k = \begin{cases} \dfrac{1+\sqrt{-1}}{(1+k)^{1.1}} & k > 0, \\ 2 & k = 0, \\ \bar{a}_{-k} & k < 0. \end{cases}$$

The underlying generating function f is given by

$$f(\theta) = 2\sum_{k=0}^{\infty} \frac{\sin(k\theta) + \cos(k\theta)}{(1+k)^{1.1}}.$$

Clearly f is in the Wiener class. The spectra of A_n, $S_n^{-1}A_n$, $C_n^{-1}A_n$ and $T_n^{-1}A_n$ for $n = 32$ are represented in Figure 1. Table 1 shows the number of iterations required to make $||r_q||_2/||r_0||_2 < 10^{-7}$, where r_q is the residual vector after q iterations. The right hand side b is the vector of all ones and the zero vector is our initial guess. We see that as n increases, the number of iterations increases like $O(\log n)$ for the original matrix A_n, while it stays almost the same for the preconditioned matrices. Moreover, all preconditioned systems converge at the same rate for large n.

TABLE 1

Number of iterations for different systems.

n	A_n	$S_n^{-1}A_n$	$C_n^{-1}A_n$	$T_n^{-1}A_n$
16	13	8	7	7
32	15	8	6	7
64	18	7	7	7
128	19	7	7	7
256	21	7	7	7

References

[1] R. Chan and G. Strang, *Toeplitz Equations by Conjugate Gradients with Circulant Preconditioner*, SIAM J. Sci. Statist. Comput., 10 (1989), pp. 104–119.

[2] R. Chan, *The Spectrum of a Family of Circulant Preconditioned Toeplitz Systems*, SIAM J. Numer. Anal., 26 (1989), pp. 503–506.

[3] R. Chan, *Circulant Preconditioners for Hermitian Toeplitz Systems*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 542–550.

[4] R. Chan, X. Jin and M. Yeung, *The Circulant Operator in the Banach Algebra of Matrices*, Linear Algebra Appl., to appear.

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

[6] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*. Second edition, Chelsea, New York, 1984.

[7] G. Strang, *A Proposal for Toeplitz Matrix Calculations*, Stud. Appl. Math., 74 (1986), pp. 171–176.

[8] E. Tyrtyshnikov, *Optimal and Super-optimal Circulant Preconditioners*. Tech. Report, Dept. of Numerical Math., USSR Academy of Science, 1990.