Circulant Preconditioners Constructed From Kernels

Raymond H. Chan and Man-Chung Yeung
Department of Mathematics
University of Hong Kong
Pokfulam Road
Hong Kong

August 1991

Abstract

We consider circulant preconditioners for Hermitian Toeplitz systems from the view point of function theory. We show that some well-known circulant preconditioners can be derived from convoluting the generating function f of the Toeplitz matrix with famous kernels like the Dirichlet and the Fejér kernels. Several circulant preconditioners are then constructed using this approach. Finally, we prove that if the convolution product converges to f uniformly, then the circulant preconditioned Toeplitz systems will have clustered spectrum.

Abbreviated Title. Circulant Preconditioners from Kernels.

Key Words. Toeplitz matrix, circulant matrix, preconditioned conjugate gradient method, generating function, kernel.

AMS(MOS) Subject Classifications. 65F10.

1 Introduction

An n-by-n matrix $A_n = (a_{i,j})$ is said to be Toeplitz if $a_{i,j} = a_{i-j}$, i.e. A_n is constant along its diagonals. Toeplitz systems arise in a variety of applications, especially in signal processing and control theory. Existing direct methods for dealing with them include the Levinson-Trench-Zohar $O(n^2)$ algorithms [21], and a variety of $O(n \log^2 n)$ algorithms such as the one by Ammar and Gragg [1]. The stability properties of these direct methods for symmetric positive definite matrices are discussed in Bunch [2].

An n-by-n matrix C_n is said to be circulant if it is Toeplitz and its diagonals c_j satisfy $c_{n-j} = c_{-j}$ for $0 < j \le n-1$. Circulant matrices can always be diagonalized by Fourier matrix, i.e.

$$C_n = F_n \Lambda_n F_n^*, \tag{1}$$

where Λ_n is diagonal and F_n is the n-by-n Fourier matrix with (j,k)th entry given by

$$[F_n]_{jk} = \frac{1}{\sqrt{n}} e^{\frac{-2\pi ijk}{n}}, \quad 0 \le j,k < n, \tag{2}$$

see Davis [12]. Hence for any vector y, $C_n^{-1}y = F_n \Lambda_n^{-1} F_n^* y$ can be computed by the Fast Fourier Transform in $O(n \log n)$ operations.

Strang [19] first proposed using the preconditioned conjugate gradient method with circulant preconditioners C_n for solving positive definite Toeplitz systems. Instead of solving $A_n x = b$, we solve the preconditioned system $C_n^{-1} A_n x = C_n^{-1} b$ by the conjugate gradient method with C_n being a circulant matrix. The number of operations per iteration in the preconditioned conjugate gradient method depends mainly on the work of computing the matrix-vector multiplication $C_n^{-1} A_n y$, see for instance Golub and van Loan [13]. As remarked above, $C_n^{-1} y$ can be computed in $O(n \log n)$ operation by the Fast Fourier Transform. For $A_n y$, it can also be computed by the Fast Fourier Transform by first embedding A_n into a $2n$-by-$2n$ circulant matrix. Thus computing $A_n y$ requires $O(2n \log(2n))$ operations. It follows that the total operations per iteration is of order $O(n \log n)$.

In order to compete with direct methods, the circulant matrix C_n should be chosen such that the conjugate gradient method converges sufficiently fast when applied to the preconditioned system $C_n^{-1} A_n x = C_n^{-1} b$. It is well-known that the convergence rate of the method depends on the spectrum of

$C_n^{-1}A_n$. The more clustered the eigenvalues are, the faster the convergence. Specifically, we want $C_n^{-1}A_n$ to be of the form $I_n + U_n + W_n$ where I_n is the identity matrix, U_n is a matrix of low rank and W_n is a matrix of small ℓ_2 norm.

Several circulant preconditioners have been proposed and analysed, see for instance, Chan and Strang [3], T. Chan [10], Chan [4, 5], Tyrtyshinkov [22], Ku and Kuo [17], Chan, Jin and Yeung [6], Huckle [16] and Chan and Jin [8]. It has been shown in these papers that if the diagonals a_j of the Toeplitz matrix A_n are Fourier coefficients of a positive function f in the Wiener class, then the spectrum of the preconditioned system $C_n^{-1}A_n$ will be clustered around one for large n. It follows that the preconditioned conjugate gradient methods, when applied to the preconditioned system, converges superlinearly for large n. More precisely, we have for all $\epsilon > 0$, there exists a constant $c(\epsilon) > 0$ such that the error vector e_q of the preconditioned conjugate gradient method at the qth iteration satisfies

$$||e_q|| \le c(\epsilon)\epsilon^q ||e_0|| \tag{3}$$

when n is sufficiently large. Here $||x||^2 = x^* C_n^{-1/2} A_n C_n^{-1/2} x$. Hence the number of iterations required for convergence is independent of the size of the matrix A_n when n is large. In particular, the system $A_n x = b$ can be solved in $O(n \log n)$ operations.

Recently, Chan and Yeung [7] extended the above superlinearly convergence result from the Wiener class of functions to the class of 2π-periodic continuous functions for the T. Chan's [10] preconditioner. One of the aims of this paper is to find other circulant preconditioners that has the same superlinear convergence property. Our approach is to consider circulant preconditioners as convolution products of the generating function f with some kernels. We show that most of the known circulant preconditioners can be derived easily by this approach. In particular, we see that the T. Chan's preconditioner is obtained by convoluting f with the Fejèr kernel while the Strang's preconditioner is obtained by convoluting f with the Dirichlet kernel. We also prove that if the convolution product converges to f uniformly, then the corresponding circulant preconditioner will have the clustering and superlinear convergence properties mentioned above. Several circulant preconditioners possessing these properties are then constructed by using known kernels in function theory and signal processing.

The outline of the paper is as follows. In §2, we express some of the circulant preconditioners mentioned above in terms of convolution of kernels. Using the idea, we design in §3 some other circulant preconditioners by considering kernels from function theory and signal processing. In §4, we prove that if the kernel K_n is such that the convolution product $K_n * f$ converges uniformly to f, then the circulant preconditioners so constructed will have clustering and superlinearly convergence properties. Numerical examples and concluding remarks are given in §5 and §6 respectively.

2 The Kernels of Some Circulant Preconditioners

Let $\mathcal{C}_{2\pi}$ be the Banach space of all 2π-periodic continuous real-valued functions defined on the real line $\mathbf{R}$ and equipped with the supremum norm $||\cdot||_\infty$. For all $f \in \mathcal{C}_{2\pi}$, let

$$a_k[f] = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)e^{-ik\theta}d\theta, \quad k = 0, \pm 1, \pm 2, \cdots$$

be the Fourier coefficients of f. For simplicity, we will write $a_k[f]$ as a_k. Since f is real-valued,

$$a_{-k} = \bar{a}_k, \quad k = 0, \pm 1, \pm 2, \cdots.$$

Let $A_n[f]$ be the n-by-n Hermitian Toeplitz matrix with the (j,l)th entry given by a_{j-l}. The function f is called the generating function of the matrices $A_n[f]$. The jth partial sum of f is defined as

$$s_j[f](\theta) \equiv \sum_{k=-j}^{j} a_k e^{ik\theta}, \qquad \forall \theta \in \mathbf{R}. \tag{4}$$

Given the matrix equation $A_n[f]x = b$, we consider the preconditioned system $C_n^{-1}A_n[f]x = C_n^{-1}b$ for some circulant matrix C_n. In order to have fast convergence rate, the circulant preconditioner C_n should be chosen such that $C_n^{-1}A_n[f]$ has spectrum clustered around one. Before we analyze few known circulant preconditioners that have this property, let us first emphasize the relationship between the first column of a circulant matrix and its eigenvalues.

By (1) and (2), if e_1 and 1_n denote the first unit vector and the vector of all ones respectively, then we have

$$\sqrt{n}C_n e_1 = F_n \Lambda_n 1_n. \tag{5}$$

Hence if

$$C_n e_1 = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{pmatrix}, \tag{6}$$

then the eigenvalues of C_n are given by

$$\lambda_j(C_n) = (\Lambda_n)_{jj} = \sum_{k=0}^{n-1} c_k \zeta_j^k, \quad 0 \le j < n, \tag{7}$$

where

$$\zeta_j \equiv e^{2\pi i j/n} \qquad 0 \le j < n.$$

Conversely, if the eigenvalues of C_n are given by the right hand side of (7), then the first column of C_n is given by (6). Notice that we have

$$\zeta_j^{n-k} = \zeta_j^{-k} = \bar{\zeta}_j^k, \quad 0 \le j, k < n. \tag{8}$$

1) Strang's preconditioner $S_n[f]$.

Given $A_n[f]$, the corresponding Strang's preconditioner $S_n[f]$ is defined to be the circulant matrix that copies the central diagonals of $A_n[f]$ and reflects them around to complete the circulant, see Strang [19]. More precisely, the kth entry in the first column of $S_n[f]$ is given by

$$(S_n[f])_{k0} = \begin{cases} a_k & 0 \le k \le m, \\ a_{k-n} & m < k < n. \end{cases}$$

Here we assume for simplicity that $n = 2m + 1$. If $n = 2m$, we define $(S_n[f])_{m0} = 0$.

We remark that the eigenvalues of $S_n[f]$ are given by the partial sum $s_m[f]$ of f at evenly-spaced points in $[0, 2\pi]$. In fact, by (7) and (8), we see that the eigenvalues of $S_n[f]$ are equal to

$$\begin{aligned}
\lambda_j(S_n[f]) &= \sum_{k=0}^{m} a_k \zeta_j^k + \sum_{k=m+1}^{n-1} a_{k-n} \zeta_j^k \\
&= \sum_{k=0}^{m} a_k \zeta_j^k + \sum_{k=1}^{m} a_{-k} \zeta_j^{-k} \\
&= \sum_{k=0}^{m} a_k \zeta_j^k + \sum_{k=1}^{m} \bar{a}_k \bar{\zeta}_j^k \\
&= s_m[f](\frac{2\pi j}{n}), \quad 0 \le j < n.
\end{aligned}$$

¿From Fourier analysis, see Zygmund [25, p.49] for instance, the partial sum $s_m[f]$ of f is given by the convolution of f with the Dirichlet kernel $\hat{D}_m$, i.e.

$$s_m[f](\theta) = \frac{1}{2\pi} \int_0^{2\pi} f(\phi) \hat{D}_m(\theta - \phi) d\phi \equiv (f * \hat{D}_m)(\theta), \tag{9}$$

where

$$\hat{D}_k(\theta) = \frac{\sin(k + \frac{1}{2})\theta}{\sin(\frac{1}{2}\theta)}, \quad k = 0, 1, \cdots.$$

Thus the eigenvalues of $S_n[f]$ can also be expressed as

$$\lambda_j(S_n[f]) = (f * \hat{D}_m)(\frac{2\pi j}{n}), \quad 0 \le j < n.$$

2) T. Chan's preconditioner $T_n[f]$.

Given $A_n[f]$, the corresponding T. Chan's preconditioner $T_n[f]$ is defined to be the circulant matrix with diagonals that are arithmetic average of the diagonals of $A_n[f]$ (extended to length n by wrap-around when necessary), see T. Chan [10]. More precisely, the entries in the first column of $T_n[f]$ are given by

$$(T_n[f])_{k0} = \frac{1}{n}\{(n-k)a_k + k\bar{a}_{n-k}\}, \quad 0 \le k < n.$$

By (7) and (8) again, the eigenvalues of $T_n[f]$ are given by

$$\lambda_j(T_n[f]) = \sum_{k=0}^{n-1} \frac{n-k}{n} a_k \zeta_j^k + \sum_{k=1}^{n-1} \frac{k}{n} \bar{a}_{n-k} \zeta_j^k$$

$$
\begin{aligned}
&= \sum_{k=0}^{n-1} \frac{n-k}{n} a_k \zeta_j^k + \sum_{k=1}^{n-1} \frac{n-k}{n} \bar{a}_k \bar{\zeta}_j^k \\
&= \frac{1}{n} \sum_{k=1}^{n-1} (n-k)\{a_k \zeta_j^k + \bar{a}_k \bar{\zeta}_j^k\} + a_0 \\
&= \frac{1}{n} \sum_{k=-(n-1)}^{n-1} (n-|k|) a_k \zeta_j^k, \quad 0 \le j < n.
\end{aligned}
$$

We note that this is a Cesàro summation process of order 1 for the Fourier series of f, see Zygmund [25, p.76]. Using the definition of partial sum and after some rearrangements of the terms, we get

$$
\lambda_j(T_n[f]) = \frac{1}{n} \sum_{k=0}^{n-1} s_k[f](\frac{2\pi j}{n}), \quad 0 \le j < n.
$$

Thus the eigenvalues of $T_n[f]$ are just the values of the arithmetic mean of the first n partial sums of f at $2\pi j/n$. It is well-known that this arithmetic mean is given by the convolution of f with the Fejér kernel $\hat{F}_n$, i.e.

$$
\frac{1}{n} \sum_{k=0}^{n-1} s_k[f](\theta) = \frac{1}{2\pi} \int_0^{2\pi} f(\phi) \hat{F}_n(\theta - \phi) d\phi \equiv (f * \hat{F}_n)(\theta), \tag{10}
$$

where

$$
\hat{F}_k(\theta) = \frac{1}{k} \left\{ \frac{\sin(\frac{k}{2}\theta)}{\sin(\frac{1}{2}\theta)} \right\}^2, \quad k = 1, 2, \cdots,
$$

see Zygmund [25, p.88]. Thus the eigenvalues of $T_n[f]$ can be expressed as

$$
\lambda_j(T_n[f]) = (f * \hat{F}_n)(\frac{2\pi j}{n}), \quad 0 \le j < n. \tag{11}
$$

3) R. Chan's preconditioner $R_n[f]$.

Given $A_n[f]$, the R. Chan's circulant preconditioner $R_n[f]$ has the first column given by

$$
(R_n[f])_{k0} = \begin{cases} a_0 & k = 0, \\ a_k + \bar{a}_{n-k} & 0 < k < n, \end{cases}
$$

see R. Chan [5]. Thus the eigenvalues of $R_n[f]$ are given by

$$\begin{aligned}
\lambda_j(R_n[f]) &= a_0 + \sum_{k=1}^{n-1}\{a_k + \bar{a}_{n-k}\}\zeta_j^k \\
&= a_0 + \sum_{k=1}^{n-1} a_k\zeta_j^k + \sum_{k=1}^{n-1} \bar{a}_k\bar{\zeta}_j^k \\
&= s_{n-1}[f](\frac{2\pi j}{n}), \quad 0 \le j < n.
\end{aligned}$$

Using (9), we then have

$$\lambda_j(R_n[f]) = (f * \hat{D}_{n-1})(\frac{2\pi j}{n}), \quad 0 \le j < n. \tag{12}$$

4) Huckle's preconditioner $H_n[f]$.

Given any $0 < p < n$, the Huckle's circulant preconditioner $H_n^p[f]$, see Huckle [16], is defined to be the circulant matrix with eigenvalues given by

$$\lambda_j(H_n^p[f]) = \frac{1}{p}\sum_{k=-p}^{p}(p - |k|)a_k\zeta_j^k, \quad 0 \le j < n,$$

which is also a Cesàro summation process of order 1 for the Fourier series of f. In fact, using (10) and after some simplifications, we have

$$\lambda_j(H_n^p[f]) = \frac{1}{p}\sum_{k=0}^{p-1} s_k[f](\frac{2\pi j}{n}) = (f * \hat{F}_p)(\frac{2\pi j}{n}), \quad 0 \le j < n.$$

5) Ku and Kuo's preconditioner $K_n[f]$.

One of the preconditioners proposed in Ku and Kuo [17] is the skew-circulant matrix $K_n[f]$ which, using our notations, can be defined as

$$K_n[f] = 2 \cdot A_n[f] - R_n[f].$$

Notice that if Θ_n is the n-by-n diagonal matrix with $(\Theta_n)_{jj} = e^{\pi ij/n}$ for $0 \le j < n$, then $\Theta_n^* K_n[f]\Theta_n$ is a circulant matrix. In fact, this property holds for any skew-circulant matrix, see Davis [12].

By (7) and (8) again, it is then straighforward to verify that

$$\begin{aligned}\lambda_j(K_n[f]) &= \lambda_j(\Theta_n^* K_n[f]\Theta_n) \\ &= s_{n-1}[f](\frac{2\pi j}{n} - \frac{\pi}{n}) \\ &= (f * \hat{D}_{n-1})(\frac{2\pi j}{n} - \frac{\pi}{n}), \quad 0 \le j < n.\end{aligned}$$

Comparing this with (12), we see that the eigenvalues of $K_n[f]$ and the eigenvalues of $R_n[f]$ are just the values of $f * \hat{D}_{n-1}$ sampled at different points in $[0, 2\pi]$.

3 Circulant Preconditioners from Kernels

In this section, we apply the idea explored in §2 to design other circulant preconditioners from kernels that are commonly used in function theory and signal processing. These kernels are listed in Table 1, see Hamming [15], Natanson [18, p.58] and Walker [23, p.88].

Kernel	$\hat{C}_n(\theta)$
Modified Dirichlet	$\frac{1}{2}\{\hat{D}_{n-1}(\theta) + \hat{D}_{n-2}(\theta)\}$
De la Vallée Poussin	$2\hat{F}_{2\lfloor n/2\rfloor}(\theta) - \hat{F}_{\lfloor n/2\rfloor}(\theta)$
von Hann	$\frac{1}{4}\{\hat{D}_{n-1}(\theta - \frac{\pi}{n}) + 2\hat{D}_{n-1}(\theta) + \hat{D}_{n-1}(\theta + \frac{\pi}{n})\}$
Hamming	$0.23\{\hat{D}_{n-1}(\theta - \frac{\pi}{n}) + \hat{D}_{n-1}(\theta + \frac{\pi}{n})\} + 0.54\hat{D}_{n-1}(\theta)$
Bernstein	$\frac{1}{2}\{\hat{D}_{n-1}(\theta) + \hat{D}_{n-1}(\theta + \frac{\pi}{n})\}$

Table 1. Some kernels and their definitions.

Given a kernel $\hat{C}_n(\theta)$ defined on $[0, 2\pi]$, we let $C_n[f]$ to be the circulant matrix with eigenvalues given by

$$\lambda_j(C_n[f]) = [f * \hat{C}_n](\frac{2\pi j}{n}), \quad 0 \le j < n. \tag{13}$$

The first column of $C_n[f]$ can be obtained by using (5).

Let us illustrate the construction process by using the De la Vallée Poussin's kernel which is defined as

$$\hat{C}_n(\theta) = 2\hat{F}_{2m}(\theta) - \hat{F}_m(\theta),$$

where $\hat{F}_k$ is the Fejér kernel and $m = \lfloor n/2 \rfloor$. For simplicity, let us consider the case where $n = 2m$. Then by (10), (8) and (4), we have

$$\begin{aligned}
& [f * \hat{C}_n](\frac{2\pi j}{n}) \\
= \; & 2(f * \hat{F}_{2m})(\frac{2\pi j}{n}) - (f * \hat{F}_m)(\frac{2\pi j}{n}) \\
= \; & 2\frac{s_0[f] + \cdots + s_{2m-1}[f]}{2m}(\frac{2\pi j}{n}) - \frac{s_0[f] + \cdots + s_{m-1}[f]}{m}(\frac{2\pi j}{n}) \\
= \; & \frac{1}{m}\{s_m[f] + \cdots + s_{2m-1}[f]\}(\frac{2\pi j}{n}) \\
= \; & s_m[f](\frac{2\pi j}{n}) + \frac{2}{n}\left\{ \sum_{k=m+1}^{2m-1} (n-k) a_k \zeta_j^k + \sum_{k=m+1}^{2m-1} (n-k) \bar{a}_k \bar{\zeta}_j^k \right\} \\
= \; & \sum_{k=0}^{m} (a_k + \frac{2k}{n}\bar{a}_{n-k})\zeta_j^k + \sum_{k=m+1}^{2m-1} (\frac{2(n-k)}{n} a_k + \bar{a}_{n-k})\zeta_j^k.
\end{aligned}$$

Hence the first column of $C_n[f]$ is given by

$$(C_n[f])_{k0} = \begin{cases} a_k + \frac{2k}{n}\bar{a}_{n-k} & 0 \le k \le m, \\ \frac{2(n-k)}{n} a_k + \bar{a}_{n-k} & m < k < n. \end{cases}$$

Table 2 lists the first column of the circulant preconditioners for the kernels in Table 1. The main diagonal of $C_n[f]$, i.e. $(C_n[f])_{jj}$, is equal to a_0, hence is omitted from the table.

Kernel	$(C_n[f])_{k0},\ \ 1 \le k < n$
Modified Dirichlet	$\begin{cases} a_1 + \frac{1}{2}\bar{a}_{n-1} & k=1 \\ a_k + \bar{a}_{n-k} & 2 \le k \le n-2 \\ \frac{1}{2}a_{n-1} + \bar{a}_1 & k = n-1 \end{cases}$
De la Vallée Poussin	$\begin{cases} a_k + \frac{k}{m}\bar{a}_{2m-k} & 1 \le k \le m, \\ \frac{2m-k}{m}a_k + \bar{a}_{2m-k} & m < k < 2m, \quad m = \lfloor n/2 \rfloor \\ 0 & k = 2m. \end{cases}$
von Hann	$\cos^2(\frac{\pi k}{2n})a_k + \cos^2(\frac{\pi(n-k)}{2n})\bar{a}_{n-k}$
Hamming	$(0.54 + 0.46\cos\frac{\pi k}{n})a_k + (0.54 + 0.46\cos\frac{\pi(n-k)}{n})\bar{a}_{n-k}$
Bernstein	$\frac{1}{2}\{(1+\exp(\frac{i\pi k}{n}))a_k + (1-\exp(\frac{i\pi k}{n}))\bar{a}_{n-k}\}$

Table 2. The first column of the circulant preconditioner.

4 Clustering of the Eigenvalues

In this section, we discuss the convergence property of the circulant preconditioned systems for those circulant preconditioners derived from kernels. We prove that if the kernel $\hat{C}_n$ is such that the convolution product of $\hat{C}_n * f$ tends to the generating function f uniformly, then the corresponding preconditioned system $C_n^{-1}A_n[f]$ will have clustered spectrum. ¿From Fourier analysis, see Zygmund [25, p.89], we know that for the Fejér kernel $\hat{F}_n$, $f * \hat{F}_n$ tends to f uniformly on $[0, 2\pi]$ for all f in $\mathcal{C}_{2\pi}$. Hence the T. Chan's circulant preconditioned system $T_n[f]^{-1}A_n[f]$ should have clustered spectrum for all f in $\mathcal{C}_{2\pi}$. This result was proved in R. Chan and Yeung [7] and we restate it here as the following Lemma.

Lemma 1 *Let $f \in \mathcal{C}_{2\pi}$. Then for all $\epsilon > 0$, there exist positive integers N and M such that for all $n > N$, at most M eigenvalues of $A_n[f] - T_n[f]$ have absolute value greater than ϵ.*

This Lemma basically states that the spectrum of $A_n[f] - T_n[f]$ is clustered around zero. Using this Lemma, we can easily get the same result for other circulant preconditioners derived from kernels.

Lemma 2 *Let $f \in \mathcal{C}_{2\pi}$. Let $\hat{C}_n$ be a kernel such that $\hat{C}_n * f$ tends to f uniformly on $[0, 2\pi]$. If $C_n[f]$ is the circulant matrix with eigenvalues given by (13), then for all $\epsilon > 0$, there exist positive integers N and M such that for all $n > N$, at most M eigenvalues of $A_n[f] - C_n[f]$ have absolute value greater than ϵ.*

Proof: We first rewrite $A_n[f] - C_n[f]$ as

$$A_n[f] - C_n[f] = \{A_n[f] - T_n[f]\} + \{T_n[f] - C_n[f]\},$$

where $T_n[f]$ is the T. Chan's circulant preconditioner. In view of Lemma 1, it suffices to show that

$$\lim_{n\to\infty} ||T_n[f] - C_n[f]||_2 = 0. \tag{14}$$

Since $T_n[f]$ and $C_n[f]$ are both circulant matrices and hence can be diagonalized by the same Fourier matrix F_n, we see that (14) is equivalent to

$$\lim_{n\to\infty} \max_{0\le j<n} |\lambda_j(T_n[f]) - \lambda_j(C_n[f])| = 0. \tag{15}$$

However, by (11) and (13), we have

$$\begin{aligned}
\max_{0\le j<n} |\lambda_j(T_n[f]) - \lambda_j(C_n[f])| &= \max_{0\le j<n} |(\hat{F}_n * f)(\frac{2\pi j}{n}) - (\hat{C}_n * f)(\frac{2\pi j}{n})| \\
&\le ||\hat{F}_n * f - \hat{C}_n * f||_\infty \\
&\le ||\hat{F}_n * f - f||_\infty + ||f - \hat{C}_n * f||_\infty.
\end{aligned}$$

Since $\hat{F}_n * f$ and $\hat{C}_n * f$ both converge to f uniformly, (15) follows. $\square$

Next we will show that if f is positive, then $C_n[f]$ is positive definite and uniformly invertible for large n.

Lemma 3 *Let $f \in \mathcal{C}_{2\pi}$ with minimum value $f_{\min} > 0$. Let $\hat{C}_n$ be a kernel such that $\hat{C}_n * f$ tends to f uniformly on $[0, 2\pi]$. If $C_n[f]$ is the circulant matrix with eigenvalues given by (13), then for all n sufficiently large, we have*

$$\lambda_j(C_n[f]) \geq \frac{1}{2} f_{\min} > 0, \quad 0 \leq j < n.$$

Proof: Since $\hat{C}_n * f$ converges to f uniformly and $f_{\min} > 0$, there exists an $N > 0$, such that for all $n > N$,

$$|[f - f * \hat{C}_n](\frac{2\pi j}{n})| \leq ||f - f * \hat{C}_n||_\infty \leq \frac{1}{2} f_{\min}, \quad 0 \leq j < n.$$

Thus by (13), we have

$$\begin{aligned} \lambda_j(C_n[f]) &= [f * \hat{C}_n - f](\frac{2\pi j}{n}) + f(\frac{2\pi j}{n}) \\ &\geq f_{\min} - [f - f * \hat{C}_n](\frac{2\pi j}{n}) \\ &\geq \frac{1}{2} f_{\min}, \quad 0 \leq j < n. \quad \square \end{aligned}$$

Combining Lemmas 2 and 3, we have our main theorem, namely that the spectrum of $C_n^{-1}[f]A_n[f]$ is clustered around one.

Theorem 1 *Let $f \in \mathcal{C}_{2\pi}$ be positive. Let $\hat{C}_n$ be a kernel such that $\hat{C}_n * f$ tends to f uniformly on $[0, 2\pi]$. If $C_n[f]$ is the circulant matrix with eigenvalues given by*

$$\lambda_j(C_n[f]) = [\hat{C}_n * f](\frac{2\pi j}{n}), \quad 0 \leq j < n,$$

then for all $\epsilon > 0$, there exist positive integers N and M such that for all $n > N$, at most M eigenvalues of $I_n - C_n^{-1}[f]A_n[f]$ have absolute value greater than ϵ.

Proof: We just note that

$$I_n - C_n^{-1}[f]A_n[f] = C_n^{-1}[f](C_n[f] - A_n[f]). \quad \square$$

It follows easily from Theorem 1 that the conjugate gradient method, when applied to the preconditioned system $C_n^{-1}A_n[f]$, converges superlinearly, see Chan and Strang [3] for a proof. Thus the number of iterations required to achieve a fixed accuracy remains bounded as the matrix order n is increased. Recall that in each iteration, the work is of order $O(n \log n)$, therefore, the total work of solving the equation $A_n[f]x = b$ to a given accuracy is also of order $O(n \log n)$.

5 Numerical Results

In this section, we test the convergence rate of the preconditioned systems with generating functions given by the Hardy-Littlewood series:

$$H_\alpha(\theta) = \sum_{k=1}^{\infty}\left(\frac{e^{ik\log k}}{k^\alpha}e^{ik\theta} + \frac{e^{-ik\log k}}{k^\alpha}e^{-ik\theta}\right),$$

see Zygmund [25, p.197]. It converges uniformly to a function in $\mathcal{C}_{2\pi}$ when $\alpha > 0.5$. In the examples below, we investigate the convergence rate of the preconditioned systems for $\alpha = 1.0$ and 0.5. We note that for $\alpha = 0.5$, $H_{0.5}$ is not even a function in $L^2[0, 2\pi]$ for its Fourier coefficients are not in ℓ_2.

We remark that in general, $H_\alpha(\theta)$ is not a positive function in $[0, 2\pi]$. In fact, we find numerically that when $n = 512$, the minimum values of the partial sum $s_n[H_\alpha](\theta)$ are approximately equal to -4.146 and -6.492 for $\alpha = 1.0$ and 0.5 respectively. Thus, in the experiments, we choose the functions $H_1(\theta) + 4.2$ and $H_{0.5}(\theta) + 6.5$ as our generating functions.

Eight different circulant preconditioners are tested. Tables 3 and 4 below shows the number of iterations required to make $||r_q||_2/||r_0||_2 \le 10^{-7}$, where r_q is the residual vector after q iterations. The right hand size b is the vector of all ones and the zero vector is our initial guess. The computations are done by using double precision arithmetic on a Vax 6420.

We see that as n increases, the number of iterations increases 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. We note that the convergence rate depends on the degree of smoothness of the generating function. For the Strang's preconditioner, this dependence has been proved in Chan and Yeung [9]. Specifically, we have proved that for generating function which is either Lipschitz of order $0 < \nu \leq 1$ or has a continuous νth order derivative, $\nu \geq 1$, then the estimate in (3) becomes

$$||e_{2q}|| \leq \prod_{k=2}^{q} \left(\frac{c \log k}{k^{\nu}} \right)^2 ||e_0||.$$

Finally, we notice that for n small, some of the preconditioners may have negative eigenvalues (Cf. Lemma 3). However, it is interesting to note that the preconditioned conjugate gradient method still converges in these cases.

Preconditioner Used	n					
	16	32	64	128	256	512
No	13	18	27	43	51	58
Strang	8*	9	9	9	9	9
T. Chan	8	10	11	11	10	9
R. Chan	8	10	9	9	9	9
Modified Dirichlet	8*	10	9	9	9	9
De la Vallée Poussin	9	9	9	9	9	9
von Hann	8	9	9	9	9	9
Bernstein	9	10	10	9	9	9
Hamming	8	9	9	9	9	9

Table 3. Number of Iterations when $f(\theta) = H_1(\theta) + 4.2$.
* Preconditioner has negative eigenvalues.

Preconditioner Used	n					
	16	32	64	128	256	512
No	12	18	29	44	66	67
Strang	9*	11	16*	16	16	15
T. Chan	8	12	13	14	15	14
R. Chan	10*	12	14	16	17	15
Modified Dirichlet	9*	12	14	16	16	15
De la Vallée Poussin	8*	11	14	15	16	15
von Hann	8	11	12	13	15	15
Bernstein	9	12	14	14	16	15
Hamming	8	11	12	13	15	15

Table 4. Number of Iterations for $f(\theta) = H_{0.5}(\theta) + 6.5$
* Preconditioner has negative eigenvalues.

6 Concluding Remarks

In this paper, we introduce a new method of finding good circulant preconditioners for Toeplitz systems. We see from the derivation that these circulant preconditioners are designed so that their eigenvalues approximate the values of f at $2\pi j/n$, $0 \le j < n$. Thus if $f(2\pi j/n)$ can be computed efficiently, then the circulant preconditioners with eigenvalues given by $f(2\pi j/n)$ is certainly a good choice. Its corresponding kernel is just the Dirac delta function. For f that are trigonometric polynomials of degree $n-1$ say, as is in the case of finite Toeplitz systems, then this circulant preconditioner reduces to the R. Chan's preconditioner, which has eigenvalues given by $s_{n-1}[f](2\pi j/n) = f(2\pi j/n)$.

References

[1] G. Ammar and W. Gragg, *Superfast Solution of Real Positive Definite Toeplitz Systems*, SIAM J. Matrix Appl. V9 (1988), pp. 61–76.

[2] J. Bunch, *Stability of Methods for Solving Toeplitz Systems of Equations*, SIAM J. Sci. Statist. Comput., V6 (1985), pp. 349–364.

[3] R. Chan and G. Strang, *Toeplitz Equations by Conjugate Gradients with Circulant Preconditioner*, SIAM J. Sci. Statist. Comput., V10 (1989), pp. 104–119.

[4] R. Chan, *The Spectrum of a Family of Circulant Precondtioned Toe-plitz Systems*, SIAM J. Numer. Anal., V26 (1989), pp. 503-506.

[5] R. Chan, *Circulant Preconditioners for Hermitian Toeplitz Systems*, SIAM J. Matrix Anal. Appl., V10 (1989), pp. 542–550.

[6] R. Chan, X. Jin and M. Yeung, *The Spectra of Super-optimal Circulant Preconditioned Systems*, SIAM J. Numer. Anal., V28 (1991), pp. 871-879.

[7] R. Chan and M. Yeung, *Circulant Preconditioners for Toeplitz Matrices with Positive Continuous Generating Functions*, Math. Comp., to appear.

[8] R. Chan and X. Jin, *Circulant and Skew-Circulant Preconditioners for Skew-Hermitian Type Toeplitz Systems*, BIT, to appear.

[9] R. Chan and M. Yeung, *Jackson's Theorem and the Circulant Preconditioned Toeplitz Systems*, J. Approx. Theory, to appear.

[10] T. Chan, *An Optimal Circulant Preconditioner for Toeplitz Systems*, SIAM J. Sci. Statist. Comput., V9 (1988), pp. 766–771.

[11] E. Cheney, *Introduction to Approximation Theory*, McGraw-Hill Book Co., New York, 1966.

[12] P. Davis, *Circulant Matrices*, John Wiley & Sons, Inc., New York, 1979.

[13] G. Golub and C. van Loan, *Matrix Computations*, The Johns Hopkins University Press, Maryland, 1983.

[14] U. Grenander and G. Szegö, *Toeplitz Forms and their Applications*, 2nd Ed., Chelsea Pub. Co., New York, 1984.

[15] R. Hamming, *Digital Filters*, 3rd Ed., Prentice Hall, New Jersey, 1989.

[16] T. Huckle, *Circulant and Skew-circulant Matrices for Solving Toeplitz Matrix Problems*, SIAM J. Matrix Anal. Appl., to appear.

[17] T. Ku and C. Kuo, *Design and Analysis of Toeplitz Preconditioners*, IEEE Trans. Acoust. Speech Signal Process., to appear.

[18] I. Natanson, *Theory of Functions of a Real Variable*, Vol II, Frederick Ungar Publishing Co., New York, 1967.

[19] G. Strang, *A Proposal for Toeplitz Matrix Calculations*, Stud. Appl. Math., V74 (1986), pp. 171–176.

[20] L. Trefethen, *Approximation Theory and Numerical Linear Algebra*, Algorithms for Approximation II, M. Cox and J. Mason, eds, 1989.

[21] W. Trench, *An Algorithm for the Inversion of Finite Toeplitz Matrices*, SIAM J. Appl. Math., V12 (1964), pp. 515–522.

[22] E. Tyrtyshnikov, *Optimal and Super-optimal Circulant Preconditioners*, SIAM J. Matrix Anal. Appl., to appear.

[23] J. Walker, *Fourier Analysis*, Oxford University Press, Oxford, 1988.

[24] J. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.

[25] A. Zygmund, *Trigonometric Series*, Vol I, Cambridge University Press, Cambridge, 1959.