Toeplitz Preconditioners for Toeplitz Systems with Nonnegative Generating Functions

Raymond H. Chan
Department of Mathematics
University of Hong Kong
Hong Kong

August 1989
Revised September 1990

Abstract. The preconditioned conjugate gradient method is employed to solve Toeplitz systems $T_n x = b$ where the generating functions of the n-by-n Toeplitz matrices T_n are continuous nonnegative periodic functions defined in $[-\pi, \pi]$. The preconditioners C_n are band Toeplitz matrices with band-widths independent of n. We prove that the spectra of $C_n^{-1} T_n$ are uniformly bounded by constants independent of n. In particular, we show that the solutions of $T_n x = b$ can be obtained in $O(n \log n)$ operations.

Abbreviated Title. Band Toeplitz Preconditioners

Key words. Toeplitz matrix, band matrix, generating function, preconditioned conjugate gradient method

AMS(MOS) subject classifications. 65F10,65F15

1. Introduction.

In this paper we discuss the solutions to a class of Hermitian Toeplitz systems $T_n x = b$ by the preconditioned conjugate gradient method. Direct methods that are based on the Levinson recursion formula are in constant use; see for instance, Levinson [11] and Trench [13]. For an n-by-n Toeplitz matrix T_n, these methods require $O(n^2)$ operations. Faster algorithms that require $O(n \log^2 n)$ operations have been developed, see Bitmead and Anderson [2], Brent, Gustavson and Yun [3] and Ammar and Gragg [1]. The stability properties of these direct methods for symmetric positive definite matrices are discussed in Bunch [4].

In [12], Strang proposed using the preconditioned conjugate gradient method with circulant preconditioners for solving symmetric positive definite Toeplitz systems. 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. Chan and Strang [5] then considered using a circulant preconditioner S_n that is obtained by copying the central diagonals of T_n and bringing them around to complete the circulant. In that paper, we proved that if the underlying generating function f, the Fourier coefficients of which give the diagonals of T_n, is an even and 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 that the eigenvalues of the preconditioned matrix $S_n^{-1}T_n$ cluster around one. In particular, we showed that the conjugate gradient method converges superlinearly.

These results are generalized to the case of Hermitian positive definite Toeplitz systems in Chan [7]. We showed that the foregoing conclusions are still valid if the generating function f is a real-valued positive function in the Wiener class. We also gave an estimate of the convergence rate of the method in terms of the degree of smoothness of the function f. Moreover, we proved that S_n is "optimal" in the sense that it minimizes the l_1 norm $||S_n - T_n||_1$ over the space of all circulant matrices. We remark that the optimal circulant preconditioner B_n that minimizes the Frobenius norm is given in Chan [8]. The spectral analysis of $B_n^{-1}T_n$ is studied in Chan [6], where we showed that $\lim_{n\to\infty} ||B_n^{-1}T_n - S_n^{-1}T_n||_2 = 0$. Hence the two optimal preconditioners behave more or less the same for large n. Thus in the following, we will use S_n for the comparison with our band preconditioners.

In this paper, we will consider nonnegative continuous periodic generating functions defined on $[-\pi, \pi]$. A typical example is given by the 1-dimensional discrete Laplacian:

$$\text{tridiag}[-1, 2, -1]. \tag{1}$$

Its generating function is $2 - 2\cos\theta$, which has a zero at $\theta = 0$. We note that the eigenvalues of T_n are given by

$$\lambda_j(T_n) = 4\sin^2(\frac{\pi j}{2n+2}), \quad j = 1, 2, \cdots, n, \tag{2}$$

hence T_n is nonsingular for all n. However, the circulant preconditioner S_n is inapplicable here because it is singular. Instead of finding other possible nonsingular circulant preconditioners, we resort to using band Toeplitz matrices as preconditioners. We will show that if the global minimum of f is attained at finitely many points, and f is sufficiently smooth around these points, then there exists a band Toeplitz preconditioner C_n, with band-width independent of n, such that the condition number $\kappa(C_n^{-1}T_n)$ is uniformly bounded.

The outline of the rest of the paper is as follows. In §2, we introduce our preconditioners and some of their properties. In §3, we consider the case where the generating function f has a unique global minimum. The case where the minimun is attained at finitely many points is given in §4. In §5, we evaluate the computational cost and storage requirement in solving $T_n x = b$ by the preconditioned conjugate gradient method using our band Toeplitz matrix as preconditioner. In §6, we consider the possibility of extending these results to cases where either f is zero in an interval or the minimum of f is less than zero. Numerical results are given in §7.

Before we begin our discussion, let us recall some of the properties of Toeplitz forms. Let f be a real-valued periodic function in $L_1[-\pi, \pi]$ with

$$m = \text{ess inf } f \quad \text{and} \quad M = \text{ess sup } f.$$

Thus $M \geq f(\theta) \geq m$ almost everywhere. We denote by $T_n[f]$ the n-by-n Toeplitz matrix with entries $t_{i,j} = t_{j-i}$, where for all integers k,

$$t_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{-ik\theta}d\theta.$$

The function f is called the generating function of $T_n[f]$. Since f is real-valued, $t_{-k} = \bar{t}_k$, and $T_n[f]$ is Hermitian. We will order its eigenvalues in ascending order, i.e.,

$$\lambda_1^{(n)} \leq \lambda_2^{(n)} \leq \cdots \leq \lambda_n^{(n)}.$$

We will drop the superscript (n) if the order of the matrix is clear from the context. Notice that for any n-vector $u = (u_1, u_2, \cdots, u_n)^*$,

$$\begin{aligned} u^* T_n[f] u &= \sum_{j,k=1}^{n} \bar{u}_j (T_n[f])_{j,k} u_k = \sum_{j,k=1}^{n} \bar{u}_j t_{k-j} u_k \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{j,k=1}^{n} \bar{u}_j u_k e^{-i(k-j)\theta} f(\theta) d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |\sum_{j=1}^{n} u_j e^{-ij\theta}|^2 f(\theta) d\theta. \end{aligned} \tag{3}$$

It follows immediately that for all $n > 0$,

$$m \leq \lambda_i^{(n)} \leq M, \quad i = 1, 2, \cdots, n. \tag{4}$$

In particular, if $M = m$, then $T_n[f] = mI_n$ for all n. In the following, we will therefore assume that $M > m$. We note that we then have a stronger result than (4).

Lemma 1. *If $m < M$, then for all $n > 0$,*

$$m < \lambda_i^{(n)} < M, \quad i = 1, \cdots, n. \tag{5}$$

In particular, if $m \geq 0$, then $T_n[f]$ are positive definite for all n.

Proof: By contradiction, let us assume that $\lambda_1^{(n)} = m$ with corresponding eigenvector u. Then by (3), we have

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |\sum_{j=1}^{n} u_j e^{-ij\theta}|^2 (f(\theta) - m) d\theta = u^* T_n[f] u - m u^* u = 0.$$

Since the integrand is nonnegative almost everywhere and the integral is zero, the integrand must be zero almost everywhere. Therefore,

$$|\sum_{j=1}^{n} u_j e^{-ij\theta}| = 0$$

on the set $\{\theta | f(\theta) > m\}$ which has positive measure. However this implies that the complex polynomial $\sum_{j=1}^{n} u_j z^j = 0$ has more than n roots on the unit circle, which is clearly impossible. Thus $\lambda_1^{(n)} > m$. Similarly, we can show that $\lambda_n^{(n)} < M$. $\Box$

In the following, we will mainly consider continuous periodic functions in $[-\pi, \pi]$. Suppose that θ_0 is a zero of $f(\theta) - m$. We say that θ_0 is of *order* ν if ν is the smallest positive integer such that $f^{(\nu+1)}(\theta)$ is continuous in a neighborhood of θ_0 and $f^{(\nu)}(\theta_0) \neq 0$. By Taylor's theorem,

$$f(\theta) = m + \frac{f^{(\nu)}(\theta_0)}{\nu!}(\theta - \theta_0)^\nu + O((\theta - \theta_0)^{\nu+1}),$$

for all θ in that neighborhood. We note that $f^{(\nu)}(\theta_0) > 0$ and ν must be even. We let $l = \nu/2$.

2. Properties of the Preconditioners.

For all $l \geq 1$, we define

$$a_l(\theta) = (2 - 2\cos\theta)^l = (2\sin(\frac{\theta}{2}))^{2l},$$

which has a unique zero of order $2l$ at $\theta = 0$. Let

$$A_n[l] = T_n[a_l(\theta)].$$

We note that $A_n[1]$ is the discrete Laplacian given by (1) with eigenvalues given by (2). $A_n[l]$ will be used as our preconditioners in subsequent sections. It is therefore necessary that the diagonals of $A_n[l]$ can be found easily. We remark that

$$2 - 2\cos\theta = -\frac{1}{z}(1 - z)^2 = -(\frac{1}{z} + 2 - z), \tag{6}$$

where $z = e^{i\theta}$. Hence by the binomial theorem,

$$(2 - 2\cos\theta)^l = \sum_{k=-l}^{l} a_k^{(l)} z^k, \tag{7}$$

where

$$a_j^{(l)} = a_{-j}^{(l)} = (-1)^j \begin{pmatrix} 2l \\ l+j \end{pmatrix},$$

are the binomial coefficients of $(-1)^l(1-z)^{2l}$. Hence the diagonals of $A_n[l]$ can be obtained easily from the Pascal triangle. From (7), it is clear that $A_n[l]$ is a symmetric band Toeplitz matrix of band-width $(2l+1)$. We first investigate the spectrum of $A_n[l]$.

Theorem 1. *For all $l \geq 1$,*

$$a_0^{(l)} = \frac{(2l)!}{(l!)^2} \leq \lambda_n(A_n[l]) < 4^l, \tag{8}$$

and

$$0 < \lambda_1(A_n[l]) \leq \left(\frac{(2l+1)\pi}{n+1} \right)^{2l}, \tag{9}$$

for all $n > 0$. In particular, $A_n[l]$ are nonsingular for all n and

$$\kappa(A_n[l]) \geq \frac{(2l)!}{(l!)^2} \left(\frac{n+1}{(2l+1)\pi} \right)^{2l} = O(n^{2l}).$$

Proof: The right hand inequality of (8) and the left hand inequality in (9) follow directly from (5). To prove the left hand inequality in (8), we simply choose $u = (1, 0, \cdots, 0)^*$. Then $a_0^{(l)} = u^* A_n[l] u \leq \lambda_n$. To obtain an upper bound for λ_1, we write

$$A_n[l] = (A_n[1])^l + H_n[l].$$

We note that both $A_n[l]$ and $(A_n[1])^l$ are discrete approximation of the 1-dimensional operator $(-1)^{l-1} d^{(2l-2)}/d\theta^{2l-2}$ with center differencing scheme. The only difference in them is the handling of the boundary conditions. Thus the $(l+1)$-th up to the $(n-l)$-th rows of $A_n[l]$ and $(A_n[1])^l$ are the same. In fact, for any $l < j \leq n - l$, the j-th entry of $A_n[l]u$ for any vector u is given by

$$\begin{aligned} (A_n[l]u)_j &= \sum_{k=1}^{n} a_{k-j}^{(l)} u_k = \sum_{k=-l}^{l} (-1)^k \begin{pmatrix} 2l \\ l+k \end{pmatrix} u_{k+j} \\ &= \sum_{k=-l}^{l} (-1)^k \begin{pmatrix} 2l \\ l+k \end{pmatrix} B^k u_j = \sum_{k=-l}^{l} a_k^{(l)} B^k u_j, \end{aligned}$$

where B is the shift operator: $B^k u_j = u_{k+j}$. By (6) and (7), we see that

$$(A_n[l]u)_j = (-B + 2 - B^{-1})^l u_j,$$

which is equal to the j-th entry of $(A_n[1])^l u$. Thus the j-th rows of $A_n[l]$ and $(A_n[1])^l$ are the same for all $l < j \le n - l$. By symmetry, we see that the $(l+1)$-th up to the $(n-l)$-th columns of $A_n[l]$ and $(A_n[1])^l$ are the same too. Hence $H_n[l]$ is nonzero only in the first l-by-l and the last l-by-l principal blocks. In particular, $H_n[l]$ is a matrix of rank at most $2l$. Hence $\lambda_{n-2l}(H_n[l]) = 0$. By Cauchy interlace theorem, see Wilkinson [14], we have

$$\lambda_1(A_n[l]) \le \lambda_{2l+1}((A_n[1])^l) + \lambda_{n-2l}(H_n[l]) = \lambda_{2l+1}((A_n[1])^l).$$

Using the formula for $\lambda_j(A_n[1])$ in (2), we get

$$\lambda_1(A_n[l]) \le (\lambda_{2l+1}(A_n[1]))^l = 4^l \sin^{2l}\left(\frac{(2l+1)\pi}{2n+2}\right) \le \left(\frac{(2l+1)\pi}{n+1}\right)^{2l}. \quad \Box$$

3. Generating Functions with a Unique Global Minimum.

In this section, we will consider generating functions f which are continuous nonnegative periodic functions in $[-\pi, \pi]$ with a unique global minimum point at θ_0. We first note that we can assume without loss of generality that $\theta_0 = 0$.

Lemma 2. *Let $\tilde{f}(\theta) = f(\theta + \theta_0)$. Then for all $n > 0$,*

$$T_n[\tilde{f}] = D_n^* T_n[f] D_n,$$

where

$$D_n = \text{diag}(1, e^{-i\theta_0}, \cdots, e^{-i(n-1)\theta_0}).$$

In particular, the spectra of $T_n[\tilde{f}]$ and $T_n[f]$ are the same.

Proof: Let $\tilde{T}_n = D_n^* T_n[f] D_n$. Then $\tilde{T}_n$ is a Toeplitz matrix with entries

$$\begin{aligned} [\tilde{T}_n]_{j,k} &= e^{ij\theta_0} T_n[f]_{j,k} e^{-ik\theta_0} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-k)(\theta-\theta_0)} f(\theta) d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-k)\theta} f(\theta + \theta_0) d\theta. \end{aligned}$$

Thus $\tilde{T}_n = T_n[\tilde{f}]$. $\square$

We note that $\tilde{f}$ now has the minimum at $\theta_0 = 0$. With regard to the linear system $T_n x = b$, we can solve $\tilde{T}_n \tilde{x} = \tilde{b}$ with $\tilde{b} = D_n^* b$ instead. Then x will be given by $D_n \tilde{x}$. Thus in the following, we assume that f has the minimum at $\theta_0 = 0$. We then have our main theorem.

Theorem 2. *Suppose that $f(\theta) - m$ has a unique zero at $\theta = 0$ with order equals to $2l$. Define for all $n > 0$,*

$$C_n \equiv A_n[l] + mI_n = T_n[(2 - 2\cos\theta)^l + m].$$

Then $\kappa(C_n^{-1} T_n[f])$ is uniformly bounded for all $n > 0$.

Proof: Define

$$F(\theta) = \frac{f(\theta)}{(2 - 2\cos\theta)^l + m}.$$

Then clearly F is continuous and positive for all $\theta \neq 0$. Since

$$\lim_{\theta \to 0} F(\theta) = \begin{cases} 1 & \text{if } m > 0, \\ \dfrac{f^{(2l)}(0)}{(2l)!} & \text{if } m = 0, \end{cases}$$

is positive, F is a continuous positive function in $[-\pi, \pi]$. Hence there exist constants $b_1, b_2 > 0$, such that $b_1 \leq F(\theta) \leq b_2$ for all θ in $[-\pi, \pi]$. Using (3), we then have

$$b_1 \leq \frac{u^* T_n[f] u}{u^* C_n u} \leq b_2 \tag{10}$$

for any n-vector u. Therefore $\kappa(C_n^{-1} T_n[f]) \leq b_2/b_1$, which is independent of n. $\square$

Thus the preconditioner C_n is spectrally equivalent to $T_n[f]$. We note that this theorem also gives an estimate of the condition number of $T_n[f]$.

Corollary. *Suppose that $f(\theta) - m$ has a unique zero at $\theta = 0$ with order equals to $2l$. Then for all $n > 0$, we have*

$$\lambda_1(T_n[f]) \leq b_3 m + b_4 n^{-2l}, \tag{11}$$

and

$$\kappa(T_n[f]) \geq \frac{b_5}{b_6 + mn^{2l}} n^{2l}, \tag{12}$$

where $\{b_i\}_{i=3}^6$ *are constants independent of* n*.*

Proof: By (9) and (10), we have

$$\lambda_1(T_n[f]) \leq b_2 \cdot \lambda_1(C_n) \leq b_2 \left(m + \left(\frac{(2l+1)\pi}{n+1} \right)^{2l} \right).$$

Since

$$\lambda_n(T_n[f]) \geq (T_n[f])_{1,1} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) d\theta,$$

which is a constant independent of n, we get the bound in (12). □

We remark that if $m = 0$, our preconditioner has improved the condition number from $\kappa(T_n[f]) = O(n^{2l})$ to $\kappa(C_n^{-1} T_n[f]) = O(1)$. In the case when $m > 0$, the condition number may still be improved, see, for instance, the numerical results in §7. However, we emphasize that the spectrum of $C_n^{-1} T_n[f]$ in general will not be clustered around 1 although they are uniformly bounded.

4. Generating Functions with Multiple Minimum Points.

Let f be a continuous, nonnegative periodic function defined in $[-\pi, \pi]$ with global minimum m attained at $\{\theta_i\}_{i=1}^k$. Let the order of θ_i be $2l_i$ and we order them such that $l_1 \leq \cdots \leq l_k$. Let $l = \sum_{i=1}^k l_i$. We define

$$a(\theta) = \prod_{i=1}^k [2 - 2\cos(\theta - \theta_i)]^{l_i}.$$

The matrix $T_n[a(\theta) + m] = T_n[a(\theta)] + mI_n$ will be used as our preconditioner for $T_n[f]$. To compute the diagonals a_j of $T_n[a(\theta)]$, we note that

$$a(\theta) = \prod_{j=1}^k [(1 - e^{i(\theta - \theta_j)})(1 - e^{-i(\theta - \theta_j)})]^{l_j}$$

$$
\begin{aligned}
&= \prod_{j=1}^{k}[(1-ze^{-i\theta_j})(1-z^{-1}e^{i\theta_j})]^{l_j} \\
&= \frac{(-1)^l}{z^l}\prod_{j=1}^{k}(e^{i\theta_j}-2z+z^2e^{-i\theta_j})^{l_j} \qquad (13)\\
&= \sum_{j=-l}^{l} a_j z^j,
\end{aligned}
$$

where $z = e^{i\theta}$. Thus the diagonals a_j can be obtained by expanding the product in (13). Notice that $T_n[a(\theta)]$ is a Toeplitz matrix of band-width equals to $(2l+1)$. By repeating the arguments in Theorem 2, we have

Theorem 3. *There exist constants $c_1, c_2 > 0$, such that*

$$
c_1 \le \frac{f(\theta)}{a(\theta)+m} \le c_2, \quad \forall\, \theta \in [-\pi, \pi].
$$

In particular, if we let $C_n \equiv T_n[a(\theta)] + mI_n$, then $\kappa(C_n^{-1}T_n[f]) \le c_2/c_1$ for all $n > 0$. □

Thus the preconditioner C_n is also spectrally equivalent to $T_n[f]$. Next we try to estimate the condition number of the original matrix $T_n[f]$.

Corollary. *We have, for all $n > 0$,*

$$
\lambda_1(T_n[f]) \le m + c_3 n^{-2l_k},
$$

and

$$
\kappa(T_n[f]) \ge c_4 m + c_5 n^{2l_k},
$$

where $\{c_i\}_{i=3}^{5}$ are constants independent of n.

Proof: To compute an upper bound for the smallest eigenvalue, we let, for all $j = 1, \cdots, k$,

$$
\tilde{f}_j(\theta) = \frac{f(\theta)-m}{\prod_{\substack{i=1\\ i\ne j}}^{k}[2-2\cos(\theta-\theta_i)]^{l_i}}, \quad \forall\, \theta \in [-\pi, \pi].
$$

Then $\tilde{f}_j$ has a unique zero at θ_j with order equals to $2l_j$. By (11), we have,

$$\lambda_1(T_n[\tilde{f}_j]) \le b_4 n^{-2l_j}.$$

Since

$$0 \le f(\theta) - m = \tilde{f}_j(\theta) \cdot \prod_{\substack{i=1 \\ i \ne j}}^{k} [2 - 2\cos(\theta - \theta_i)]^{l_i} \le \tilde{f}_j(\theta) \cdot 4^l,$$

for all $\theta \in [-\pi, \pi]$, we have

$$\lambda_1(T_n[f]) - m = \lambda_1(T_n[f - m]) \le 4^l \lambda_1(T_n[\tilde{f}_j]) \le 4^l b_4 n^{-2l_j},$$

for all $j = 1, \cdots, k$. Thus

$$\lambda_1(T_n[f]) \le m + 4^l b_4 n^{-2l_k}.$$

Using the fact that $\lambda_n(T_n[f]) \ge (T_n[f])_{1,1}$, which is a constant independent of n, we get the required bound for the condition number. $\square$

Thus for $m = 0$, the condition number of the original system is of order $O(n^{2l_k})$, where we recall that l_k is the largest order of all the zeros of f. However, the condition number of the preconditioned system is still of order $O(1)$.

5. Computational Cost and Storage Requirement.

We now consider the cost of solving $T_n[f]x = b$ by using the preconditioned conjugate gradient method with the band Toeplitz matrix C_n as preconditioner. For a discussion on preconditioned conjugate gradient method and band matrix solvers, see Golub and van Loan [9].

It is known that the cost per iteration in the preconditioned conjugate gradient method is about $5n$ operations plus the cost of computing $T_n y$ and $C_n^{-1} d$ for some vectors y and d. By operation, we mean one complex multiplication together with one addition. The matrix-vector multiplication $T_n y$ can be done by the Fast Fourier Transform by first embedding T_n into a $2n$-by-$2n$ circulant matrix, see Strang [12]. The cost is about $2n\log(2n) + 2n$

operations. The vector $C_n^{-1}d$ can be found by using any band matrix solver. The cost of factorizing C_n is about $\frac{1}{2}l^2n$ operations, and then each solve requires an extra $(2l+1)n$ operations. Hence the cost per iteration is about $n(2\log(2n)+2l+8)$ operations, which is of order $O(n\log n)$, as l is independent of n.

The number of iterations required for convergence will depend on the condition number of the preconditioned system. It is well-known that the number of iterations required to attain a given tolerance ϵ is bounded by

$$\frac{1}{2}\sqrt{\kappa(C_n^{-1}T_n)}\log(\frac{1}{\epsilon})+1, \tag{14}$$

which in our case is uniformly bounded. Hence the overall work required to attain the given tolerance is given by

$$n(2\log(2n)+2l+8)\cdot\{\frac{1}{2}\sqrt{\kappa(C_n^{-1}T_n)}\log(\frac{1}{\epsilon})+1\}+\frac{1}{2}l^2n = O(n\log n).$$

As for the storage, we need five n-vectors in the conjugate gradient method. The diagonals of T_n will require another n-vector, and finally, we need an n-by-$(l+1)$ matrix to hold the factors of the preconditoner C_n. Thus the overall storage requirement is about $(7+l)n$.

6. More General Generating Functions.

We now investigate the possibility of extending our method to more general generating functions. We first consider generating function f which is zero in a sub-interval of $[-\pi,\pi]$. Without loss of generality, let us assume that f is zero in $(-\tau,\tau)$, $\tau>0$.

Theorem 4. *If $f(\theta)=0$ in $(-\tau,\tau)$, $\tau>0$, then for all integers l, there exist a constant b, which depends on f and l only, such that for all $n>0$,*

$$\lambda_1(T_n[f]) \le bn^{-2l}.$$

In particular,

$$\kappa(T_n[f]) \ge cn^{2l},$$

for some constant c independent of n.

Proof: For all $l \geq 1$, we let

$$f_l(\theta) = M \left(\frac{\theta}{\tau}\right)^{2l},$$

where $M = \text{ess sup } f$. Clearly, f_l has a zero of order $2l$ at $\theta = 0$. Since $f_l(\theta) \geq f(\theta)$ for all θ in $[-\pi, \pi]$, by (11),

$$\lambda_1(T_n[f]) \leq \lambda_1(T_n[f_l]) \leq bn^{-2l}. \quad \square$$

Thus the smallest eigenvalue of $T_n[f]$ goes to zero faster than any fixed power of n^{-1}, and its condition number goes to infinity faster than any fixed power of n. Hence our band matrix $A_n[l]$, with the smallest eigenvalue going to zero at the rate of n^{-2l}, will not be a good preconditioner in this case.

Next we consider the case where $m < 0$. We then have

$$\lim_{n \to \infty} \lambda_1(T_n[f]) = m < 0, \tag{15}$$

see Grenander and Szegö [10]. Hence for n sufficiently large, the Toeplitz matrix $T_n[f]$ is non-definite and may even be singular. The conjugate gradient method may be divergent in this case. We remark that this is also the case where the direct method may be unstable, see Bunch [4].

7. Numerical Results and Concluding Remarks.

Let us begin by investigating the spectra of the preconditioned systems. We first consider $f(\theta) = \theta^4$, which has a zero of fourth order at $\theta = 0$ and $M \approx 97.4$. Let $C_n = A_n[2]$. The spectra of $T_n[f]$ and $C_n^{-1}T_n[f]$ for $n = 32$ are given in Figure 1. We remark that the circulant matrix S_{32} has a negative eigenvalue, hence is non-definite, and cannot be used as a preconditioner. The condition number of $T_{32}[f]$ is found to be about $2.24 * 10^5$ whereas the condition number of $C_{32}^{-1}T_{32}[f]$ is about 5.56. We note that in this case, we can actually compute an upper bound for $\kappa(C_n^{-1}T_n[f])$. In fact, since

$$(2 - 2\cos\theta)^4 = 16\sin^4(\frac{\theta}{2}) \leq \theta^4$$

for all $\theta \in [-\pi, \pi]$, $\lambda_1(C_n^{-1}T_n[f]) \geq 1$ for all $n > 0$. Using the fact that

$$\sin^2(\frac{\theta}{2}) \geq \frac{\theta^2}{\pi^2},$$

which holds for all $\theta \in [-\pi, \pi]$, we have $\lambda_n(C_n^{-1}T_n[f]) \leq \pi^4/16$ for all $n > 0$. Thus for all $n > 0$,

$$\kappa(C_n^{-1}T_n[f]) \leq \frac{\pi^4}{16} \approx 6.09\ .$$

Now we apply our method to solve the linear system $T_n[\theta^4]x = b$. The right hand side b is chosen to be the vector of all ones and the zero vector is our initial guess. Computations are done in 8-byte arithmetic. Table 1 shows the numbers of iterations required to make $||r_q||_2/||r_0||_2 < 10^{-7}$, where r_q is the residual vector after q iterations. We see that for the original system $T_n[\theta^4]$, the number of iterations grows like $O(n^2)$, as expected from (14), while for the preconditioned system, it approaches a constant.

n	T_n	$C_n^{-1}T_n$
16	9	8
32	27	15
64	98	20
128	377	24
256	1692	27
512	7457	29

Table 1. Number of Iterations for $f(\theta) = \theta^4$.

Next we consider the function $f(\theta) = \theta^4 + 1 \geq 1$. The condition number of $T_n[f]$ is bounded above by $M \approx 98.4$. By (15) and the fact that

$$\lim_{n\to\infty} \lambda_n(T_n[f]) = M,$$

see Grenander and Szegö [10], we see that $\kappa(T_n[f])$ is actually approaching M as n tends to infinity. However, we can easily show that with the band preconditioner $C_n = A_n[2] + I_n$, the condition number of the preconditioned matrix is still bounded above by $\pi^4/16 \approx 6.09$. Thus preconditioning in this case will also improve the condition number. We remark that since

$f(\theta) = \theta^4 + 1$ is a positive function in the Wiener class, we may use the circulant preconditioner S_n as an alternative. In Figure 2, we compare the resulting spectra of the preconditioned systems for $n = 32$. We see that the circulant preconditioned system has a highly clustered spectrum, while the band Toeplitz preconditioned system gives a smaller condition number. The numbers of iterations required for solving the linear system $T_n[\theta^4 + 1]x = b$ with these preconditioners are given in Table 2. We see that the circulant preconditioner performs much better than the band preconditioner.

n	T_n	$C_n^{-1}T_n$	$S_n^{-1}T_n$
16	8	8	6
32	19	12	5
64	35	15	5
128	54	17	5
256	66	17	5
512	70	17	5

Table 2. Number of Iterations for $f(\theta) = \theta^4 + 1$.

We finally consider an example where f is zero in a sub-interval of $[-\pi, \pi]$. We let

$$f(\theta) = \begin{cases} 0 & |\theta| < \pi/2, \\ \dfrac{2}{\pi}|\theta| - 1 & |\theta| \geq \pi/2. \end{cases}$$

We found that $\lambda_1(T_4) = 1.9 * 10^{-3}$, $\lambda_1(T_8) = 2.2 * 10^{-6}$ and $\lambda_1(T_{16}) = 2.3 * 10^{-12}$. Thus the matrices $T_n[f]$ are very ill-conditioned.

We close our paper by remarking that the results we obtained here are for generating functions f that are continuous periodic and nonnegative in $[-\pi, \pi]$. Moreover the results are true for all n, they are not asymptotic results. In contrast, the results for the circulant preconditioners are proved under the assumptions that f is in the Wiener class and that n is sufficiently large, see for instance, Chan [7]. However, we emphasize that the circulant preconditioners work pretty well even for small n in practice, as was demonstrated by the second example above. Our band preconditioners will be a good complementary alternative in cases where the circulant preconditioners become singular or non-definite.

References

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

2. Bitmead, R., and Anderson, B., *Asymptotically Fast Solution of Toeplitz and Related Systems of Equations*, Lin. Alg. Appl., V34 (1980), pp. 103-116.

3. Brent, R., Gustavson, F. and Yun, D., *Fast Solution of Toeplitz Systems of Equations and Computations of Pade Approximations*, J. Algorithms, V1 (1980), pp. 259-295.

4. Bunch, J., *Stability of Methods for Solving Toeplitz Systems of Equations*, SIAM J. Sci. Stat. Comp., V6 (1985), pp. 349-364.

5. Chan, R. and Strang, G., *Toeplitz Equations by Conjugate Gradients with Circulant Preconditioner*, SIAM J. Sci. Stat. Comp., V10 (1989), pp. 104-119.

6. Chan, R. *The Spectrum of a Family of Circulant Preconditioned Toeplitz Systems*, SIAM J. Num. Anal., V26 (1989), pp. 503-506.

7. Chan, R. *Circulant Preconditioners for Hermitian Toeplitz Systems*, SIAM J. Matrix Anal. Appl. V10 (1989), pp. 542-550.

8. Chan, T., *An Optimal Circulant Preconditioner for Toeplitz Systems*, SIAM J. Sci. Stat. Comp., V9 (1988), pp. 766-771.

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

10. Grenander, U., and Szegö, G., *Toeplitz Forms and Their Applications*, 2nd Ed., Chelsea Pub. Co., New York, 1984.

11. Levinson, N., *The Wiener rms (root-mean-square) Error Criterion in Filter Design and Prediction*, J. Math. Phys., V25 (1947), pp. 261-278.

12. Strang, G., *A Proposal for Toeplitz Matrix Calculations*, Studies in App. Math., V74 (1986), pp. 171-176.

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

14. Wilkinson, J., *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.