SIAM J. MATRIX ANAL. APPL.
Vol. 16, No. 2, pp. 628–645, April 1995

019

FAST TRANSFORM BASED PRECONDITIONERS FOR TOEPLITZ EQUATIONS*

E. BOMAN† AND I. KOLTRACHT‡

Abstract. We present a new preconditioner for $n \times n$ symmetric, positive definite Toeplitz systems. This preconditioner is an element of the n–dimensional vector space of matrices that are diagonalized by the discrete sine transform. Conditions are given for which the preconditioner is positive definite and for which the preconditioned system has asymptotically clustered eigenvalues. The diagonal form of the preconditioner can be calculated in $O(n\log(n))$ operations if $n = 2^k - 1$. Thus only n additional parameters need be stored. Moreover, complex arithmetic is not needed. To use the preconditioner effectively, we develop a new technique for computing a fast convolution using the discrete sine transform (also requiring only real arithmetic). The results of numerical experimentation with this preconditioner are presented. Our preconditioner is comparable, and in some cases superior, to the standard circulant preconditioner of Tony Chan. Possible generalizations for other fast transforms are also indicated.

Key words. Toeplitz matrix, conjugate gradient algorithm, preconditioner, fast sine transform

AMS subject classifications. 65F10, 65T20, 65Y20, 65F99

0. Introduction. The preconditioned conjugate gradient algorithm (PCGA) is traditionally used to solve sparse systems of linear equations; see [6] or [13], for example. In recent years, starting with the proposal of Strang [14], this algorithm has also been used for computations with dense structured matrices, in particular for symmetric, positive definite systems of equations

$$Ax = b, \tag{1}$$

where the coefficient array is Toeplitz (see [2] and [3], for example.) Toeplitz matrices are constant along the diagonals $A = (a_{i-j})_{i,j=0}^{n-1}$. It is assumed that A is Toeplitz throughout this paper.

The successful application of the PCGA to (1) relies on the existence of a good preconditioner; that is, a matrix P such that the following properties are satisfied.

Property 1. The spectrum of $P^{-1}A$ is clustered.

Property 2. P is positive definite.

Property 3. The complexity of computing P is comparable to the complexity of computing Ax.

Property 4. The complexity of solving $Pz = b$, for arbitrary b, is comparable to the complexity of computing Ax.

Since Strang's proposal several investigators have been successful in choosing preconditioners for Toeplitz problems from among the set of circulants. Circulant matrices form a subspace of the vector space of Toeplitz matrices and may be characterized as those matrices that are diagonal in the orthogonal basis defined by the columns of the discrete Fourier transform (DFT); see [4]. Thus if P is circulant and F is the DFT then there is a diagonal matrix Λ such that $P = F\Lambda F^*$. Since Λ can

* Received by the editors August 23, 1993; accepted for publication (in revised form) by R. J. Plemmons, March 29, 1994.

† United Technologies Research Center, East Hartford, Connecticut 06119 (boman@math.uconn.edu).

‡ University of Connecticut, U-9, Storrs, Connecticut 06269 (koltrach@uconnvm.edu) This author's work was supported in part by National Science Foundation grants DMS-9007030 and DMS-9306357.

be computed in $O(n\log(n))$ operations (see [4]), this factorization allows one to solve $Pz = b$ for arbitrary $b \in R^n$ in $O(n\log(n))$ operations provided $n = 2^k$ for some integer k. Since the computation of a matrix–vector product where the matrix is Toeplitz can be performed in at most $O(n\log(n))$ flops one arrives at an $O(n\log(n))$ algorithm for solving Toeplitz systems, provided that the spectrum of $P^{-1}A$ is clustered as explained in the next section. This is faster than more traditional $O(n^2)$ algorithms like Levinson's algorithm ([10]), for example.

Strang built a circulant preconditioner for Toeplitz problems in [14] by copying the central diagonals of A around to complete the circulant. That is, if $A = \left(a_{|i-j|}\right)_{i\,j=1}^{n}$ then Strang's preconditioner is given by

$$p_{i\,j} = \begin{cases} a_{|i-j|} & \text{if } |i-j| \leq \frac{n}{2}, \\ a_{n-|i-j|} & \text{if } |i-j| > \frac{n}{2}. \end{cases}$$

In [3], Tony Chan built a circulant preconditioner by taking

$$P = \operatorname*{argmin}_{M\,\mathrm{circulant}} \|M - A\|_F,$$

where F denotes the Frobenius norm, and in [15] Tyrtyshnikov built a circulant preconditioner by taking

$$P^{-1} = \operatorname*{argmin}_{M\,\mathrm{circulant}} \|I - MA\|_F.$$

In [1], Raymond Chan showed that as n grows, all of the above preconditioners give similar asymptotic clustering of the spectrum of $P^{-1}A$ when A is a finite section of a singly infinite Toeplitz matrix associated with a Weiner class function.

In this paper we study preconditioners based on the fast sine transform (FST), called S_1-diagonal preconditioners, and compare them with Tony Chan's circulant preconditioner on some common Toeplitz problems. We also indicate generalizations to other fast transforms. S_1-diagonal preconditioners can be implemented quickly and do not require complex arithmetic (as do the circulant preconditioners listed above). They are banded when the Toeplitz matrix is banded and, in this case, perform better than the circulant preconditioner in our numerical experiments.

This paper is organized as follows. Section 1 is a brief outline of the PCGA. Section 2 develops the underlying theory that supports the generation of new preconditioners for Toeplitz equations, based on a given fast transform, and shows how the circulant matrices fit into this broader scheme.

In §3 we show how to apply a symmetric Toeplitz matrix to an arbitrary vector by embedding the matrix into an S_1-diagonal matrix.

Section 4 develops a new preconditioner for banded, symmetric Toeplitz systems. This new preconditioner is based on the discrete sine transform in the same way that circulants are based on the DFT. When A is banded, the spectrum of the product $P^{-1}A$ will not only be clustered, but all eigenvalues are equal to 1 except for a few outliers. The number of the outlying eigenvalues depends only on the bandwidth of A.

Under certain conditions this preconditioner also works for full Toeplitz systems. Asymptotic clustering of the eigenvalues of $P^{-1}A$ when A is full is shown in §4.2.

The results of numerical experimentation with the new preconditioners are presented in §5, and the Appendix indicates how preconditioners based on some other fast transforms may be generated.

1. The preconditioned conjugate gradient algorithm (PCGA). The PCGA for solving (1), as given in [6], follows here.

ALGORITHM 1. Set $x_0 = 0$, and $r_0 = b$. Then for $k = 1$, to convergence repeat the following.

(i) If $r_{k-1} = 0$ set $x = x_{k-1}$ and stop.
(ii) Otherwise,
1. Solve $Pz_{k-1} = r_{k-1}$ (This is the preconditioning step.)
2. Set $\beta_k = z_{k-1}^T r_{k-1} / z_{k-2}^T r_{k-2}, \quad \beta_1 \equiv 0$
3. Set $p_k = z_{k-1} + \beta_k p_{k-1}, \quad p_1 \equiv z_0$
4. Set $\alpha_k = z_{k-1}^T r_{k-1} / p_k^T A p_k$
5. Set $x_k = x_{k-1} + \alpha_k p_k$
6. Set $r_k = r_{k-1} - \alpha_k A p_k$

Applying the PCGA to (1) is equivalent to applying the conjugate gradient algorithm to $\tilde{A}\tilde{x} = \tilde{b}$, where $\tilde{A} = P^{-1/2}AP^{-1/2}$, $\tilde{x} = P^{1/2}x$, and $\tilde{b} = P^{-1/2}b$ (see [6]). The important feature of this algorithm is that if it is performed in exact arithmetic, it will converge to the correct solution in $k \leq n$ iterations where k is the number of distinct eigenvalues of $\tilde{A}$. This is because after the jth iteration of the PCGA,

$$\left(\tilde{A}x_j - \tilde{b}\right)^T \tilde{A}^{-1} \left(\tilde{A}x_j - \tilde{b}\right),$$

has minimal norm over all vectors spanned by the Krylov vectors: $\left\{\tilde{b}, \tilde{A}\tilde{b}, \tilde{A}^2\tilde{b}, \ldots, \tilde{A}^j\tilde{b}\right\}$ (see [12]).

Clearly the choice of the preconditioner is all important. A preconditioner that reduces the number of distinct eigenvalues of $P^{-1/2}AP^{-1/2}$ also reduces the number of iterations required for convergence. We remark that the spectra of $P^{-1/2}AP^{-1/2}$ and $P^{-1}A$ are equal since the two matrices are similar. We will work with the latter matrix hereafter.

In practice it is rarely possible to actually reduce the number of distinct eigenvalues. However, Jennings [8] shows that if the eigenvalues of $P^{-1}A$ are "clustered" — if the eigenvalues occur in $q < n$ clusters — the convergence characteristics of the PCGA are almost as good as if there were only q distinct eigenvalues.

2. T–diagonal matrices.

DEFINITION 1. *Let T be an arbitrary, nonsingular matrix. A matrix M is T–diagonal if $T^{-1}MT$ is diagonal.*

DEFINITION 2. *For a given T, denote by D_T the n-dimensional vector space of all T–diagonal matrices.*

Much of the recent work on preconditioners for Toeplitz systems is focused on circulant matrices that are F–diagonal, where F denotes the DFT,

$$F = \frac{1}{\sqrt{n}} \left(\exp\left(\frac{jk2\pi i}{n} \right) \right)_{j,k=0}^{n-1}.$$

The key to the use of circulant matrices as preconditioners for Toeplitz systems of equations is the following well-known property (see [4]). If C is circulant and c_1 is the first row of C, then $\sqrt{n}c_1F$ is a row-vector whose elements are the eigenvalues of C.

This is a special case of the next proposition. It shows that if T and $M \in D_T$ are given, it is not necessary to compute $T^{-1}MT$ to produce the diagonal form (eigenvalues) of M.

DEFINITION 3. *Let* $x = (x_1, \ldots, x_n)$. *Then* $\Delta(x) \equiv \operatorname{diag}\{x_1, \ldots, x_n\}$.

Thus if $x_i \neq 0$, $i = 1, \ldots, n$, then $[\Delta(x)]^{-1} \equiv \operatorname{diag}\left\{\frac{1}{x_1}, \ldots, \frac{1}{x_n}\right\}$.

Some care must be taken here. We will need to refer to Δ^{-1}, which is the inverse of Δ as an operator *not* the inverse of $\operatorname{diag}\{x_1, \ldots, x_n\}$, thus $\Delta^{-1}(\operatorname{diag}\{x_1, \ldots, x_n\}) = (x_1, \ldots, x_n)$.

PROPOSITION 2.1. *Let T be a nonsingular matrix that has at least one row τ_k, which is everywhere nonzero. Let μ_k be the corresponding row of M. Then*

$$T^{-1}MT = \Delta\left(\mu_k T \left[\Delta(\tau_k)\right]^{-1}\right).$$

Proof. Since $M \in D_T$ there is a diagonal matrix, Λ such that $MT = T\Lambda$. Thus

$$\mu_k T = \tau_k \Lambda = \Delta^{-1}(\Lambda)\, \Delta(\tau_k).$$

Therefore $\Delta\left(\mu_k T \left[\Delta(\tau_k)\right]^{-1}\right) = \Lambda$. □

Observe that the computation of Λ is dominated by the product $\mu_k T$ that generally is $O(n^2)$. If T admits a fast calculation, for example if T is the DFT and $n = 2^k$, then $\mu_k T$ can be computed in $O(n\log(n))$ operations.

If M is circulant and T is the DFT, then take $k = 1$ and observe that $\tau_1 = \frac{1}{\sqrt{n}}(1, \ldots, 1)$ to recover the previously mentioned property of circulant matrices.

2.1. The diagonal space of the discrete sine transform, (DST1).

Let S_1 denote the first discrete sine transform (there are at least two; see [16])

$$S_1 = \sqrt{\frac{2}{n+1}}\left(\sin\left(\frac{ij\pi}{n+1}\right)\right)_{i,j=1}^{n}.$$

In this section a basis for D_{S_1} is explicitly displayed.

It is well known that $S_1^2 = I$, $S_1 = S_1^T$ and that S_1 can be applied to a vector in $O(n\log(n))$ flops as long as $n = 2^k - 1$ for some integer k. This will be assumed hereafter. For a detailed exposition of the properties of the DST1, see [16].

Recall that the n-dimensional vector space of $n \times n$ circulant matrices is spanned by the set $\{C_p\}_{p=0}^{n-1}$, where C_1 is defined by the following relation. If $x = (x_1, \ldots, x_n)^T$ then $C_1 x \equiv (x_n, x_1, \ldots, x_{n-1})^T$ and $C_p = C_1^p$. Similarly, a basis for D_{S_1} is given by choosing ζ so that

$$\zeta(i,j) = \begin{cases} 1 & \text{if } |i-j| = 1, \\ 0 & \text{otherwise,} \end{cases}$$

(see ζ_1 displayed below). A basis is then given by $\{I, \zeta^1, \ldots, \zeta^{n-1}\}$ just like the circulant basis. Since nothing that follows depends on this basis we simply remark that it exists.

For our purposes we have found it more convenient to use the basis $\{\zeta_p\}_{p=0}^{n-1}$ where

$$\zeta_p(i,j) = \begin{cases} 1 & \text{if } |i-j| = p, \\ -1 & \text{if } i+j = p, \\ -1 & \text{if } i+j = 2(n+1) - p, \\ 0 & \text{otherwise.} \end{cases}$$

To display the structure of the ζ_p's more clearly, we now display the complete basis for the 5×5 case. Of course, ζ_0 is the identity. The other four basis matrices are given below.

$$\zeta_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \qquad \zeta_2 = \begin{pmatrix} -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix},$$

$$\zeta_3 = \begin{pmatrix} 0 & -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 & 0 \end{pmatrix}, \qquad \zeta_4 = \begin{pmatrix} 0 & 0 & -1 & 0 & 1 \\ 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 \\ 1 & 0 & -1 & 0 & 0 \end{pmatrix}.$$

That this is in fact a basis is the subject of the next lemma. That it is more convenient is taken up in §4. The dimension of ζ_p will always be clear from context.

LEMMA 2.2. *Let $n \in Z^+$ be given. Then $\{\zeta_p\}_{p=0}^{n-1}$ is a basis for D_{S_1}. Moreover if $p > 0$ then the spectrum of ζ_p is*

$$\left\{2\cos\left(\frac{p\pi k}{n+1}\right)\right\}_{k=1}^{n}.$$

Proof. Clearly $\{\zeta_p\}_{p=0}^{n-1}$ is a linearly independent set and it is large enough to span D_{S_1}. All that remains is to show that for all $p = 0, \ldots, n-1$, ζ_p is in D_{S_1} and that the spectrum is as given. Denote by s_k the kth column of S_1. Calculating the components of $\zeta_p s_k = (\nu_1, \ldots, \nu_n)^T$ directly will show that every column of S_1 is an eigenvector of ζ_p and, incidentally, gives the spectrum of ζ_p, which will complete the proof.

There are five cases: (i) $i < p$, (ii) $i = p$, (iii) $p < i < n - p + 1$, (iv) $i = n - p + 1$, and (v) $i > n - p + 1$. If $i < p$ then

$$\nu_i = \sqrt{\frac{2}{n+1}}\left[-\sin\left(\frac{(p-i)\pi k}{n+1}\right) + \sin\left(\frac{(p+i)\pi k}{n+1}\right)\right],$$

which by an elementary trigonometric identity is

$$\nu_i = 2\cos\left(\frac{p\pi k}{n+1}\right)\sqrt{\frac{2}{n+1}}\sin\left(\frac{i\pi k}{n+1}\right).$$

The other four cases are similar. $\square$

Remark. D_{S_1} is clearly a subspace of the space of Toeplitz plus Hankel matrices. It may prove useful to precondition such matrices with a preconditioner drawn from D_{S_1}, but we have not done so here.

3. Efficient matrix–vector multiplication. Before proceeding to develop our preconditioner, we first indicate how the diagonal space D_{S_1} may be used to apply a symmetric Toeplitz matrix $A \in R^{n\times n}$ to a vector $x \in R^n$ in $O(n\log(n))$ operations. Step 4 of the PCGA requires this product, which is commonly performed by embedding A into a larger circulant matrix, C, and x into a a larger vector, $\bar{x}$. Applying C

to $\bar{x}$ requires only $O(n\log(n))$ operations and the desired product Ax emerges from the nature of the embeddings. The difficulty of this method is that the efficient application of C to $\bar{x}$ requires the use of the DFT thus requiring complex arithmetic. But if A and x are both real their product will also be real and we would like to avoid the added computational and storage burden of using complex arithmetic to compute real results.

Our method is similar to the method just outlined in that we embed the Toeplitz matrix in an S_1–diagonal matrix. No complex arithmetic is required at any step.

The method is best demonstrated by example. Let

$$A = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 \\ a_1 & a_0 & a_1 & a_2 \\ a_2 & a_1 & a_0 & a_1 \\ a_3 & a_2 & a_1 & a_0 \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

We seek a matrix $\hat{A} \in D_{S_1}$ such that A is a submatrix of $\hat{A}$. Clearly,

$$\hat{A} = \begin{pmatrix} (a_0-a_2) & (a_1-a_3) & a_2 & a_3 & 0 & 0 \\ (a_1-a_3) & \mathbf{a_0} & \mathbf{a_1} & \mathbf{a_2} & \mathbf{a_3} & 0 \\ a_2 & \mathbf{a_1} & \mathbf{a_0} & \mathbf{a_1} & \mathbf{a_2} & a_3 \\ a_3 & \mathbf{a_2} & \mathbf{a_1} & \mathbf{a_0} & \mathbf{a_1} & a_2 \\ 0 & \mathbf{a_3} & \mathbf{a_2} & \mathbf{a_1} & \mathbf{a_0} & (a_1-a_3) \\ 0 & 0 & a_3 & a_2 & (a_1-a_3) & (a_0-a_2) \end{pmatrix}$$

is such a matrix since $\hat{A} = \sum_{i=0}^{3} a_i \zeta_i$. The embedding of A is indicated with boldface.

Similarly, embed x in $\hat{x}$

$$\hat{x} = \begin{pmatrix} 0 \\ x_0 \\ x_1 \\ x_2 \\ x_3 \\ 0 \end{pmatrix}, \quad \text{so that } \hat{A}\hat{x} = \begin{pmatrix} * \\ Ax \\ * \end{pmatrix}.$$

Since by construction $\hat{A}$ is in D_{S_1}, there is a diagonal matrix Λ such that $\hat{A} = S_1 \Lambda S_1$. Thus the product $\hat{A}\hat{x}$ can be computed in $O(n\log(n))$ operations once Λ is known. Proposition 2.1 gives an algorithm for computing Λ in $O(n\log(n))$ operations. Thus the entire computation is again $O(n\log(n))$.

In general an $n \times n$ real Toeplitz matrix A can be embedded in the S_1–diagonal matrix,

$$\hat{A} = \begin{pmatrix} A_1 - H_1 & A_2 - H_2 & 0 \\ (A_2 - H_2)^T & A & J(A_2 - H_2)^T J \\ 0 & J(A_2 - H_2)J & J(A_1 - H_1)J \end{pmatrix},$$

where

$$J \equiv \begin{pmatrix} 0 & \cdots & 1 \\ \vdots & ⋰ & \vdots \\ 1 & \cdots & 0 \end{pmatrix}$$

and A_1, H_1, A_2, and H_2 are defined as follows. Take $k = (n-3)/2$ if n is odd and $k = (n-4)/2$ if n is even. Then

$$A_1 = \begin{pmatrix} a_0 & \cdots & a_k \\ \vdots & \ddots & \vdots \\ a_k & \cdots & a_0 \end{pmatrix}, \qquad H_1 = \begin{pmatrix} a_2 & \cdots & a_{k+2} \\ \vdots & \iddots & \vdots \\ a_{k+2} & \cdots & a_{n-1} \end{pmatrix},$$

$$A_2 = \begin{pmatrix} a_{k+1} & \cdots & a_{n-1} & 0 & \cdots & 0 \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ a_1 & \cdots & a_{k+1} & \cdots & a_{n-1} & 0 \end{pmatrix},$$

and

$$H_2 = \begin{pmatrix} a_{k+3} & \cdots & a_{n-1} & 0 & \cdots & 0 \\ \vdots & \iddots & \iddots & & \iddots & \vdots \\ a_{n-1} & 0 & & \iddots & & 0 \\ 0 & \cdots & 0 & \cdots & \cdots & 0 \end{pmatrix},$$

Finally, it is not necessary to compute all of $\hat{A}$. By Proposition 2.1 all that is required to calculate the diagonal form (eigenvalues) of $\hat{A}$ is its first row.

4. S_1–diagonal preconditioners.

4.1. Banded matrices. The goal of this section is to construct an S_1–diagonal preconditioner, P_b, which approximates the banded Toeplitz matrix,

$$A = \begin{pmatrix} a_0 & a_1 & \cdots & a_b & & & & 0 \\ a_1 & a_0 & \cdots & & a_b & & & \\ \vdots & & \cdot & & & \ddots & & \\ a_b & & & \cdot & & & a_b & \\ & a_b & & & \cdot & & & a_b \\ & & \ddots & & & \cdot & & \vdots \\ & & & a_b & & \cdots & a_0 & a_1 \\ 0 & & & & a_b & \cdots & a_1 & a_0 \end{pmatrix} \in \mathcal{R}^{n \times n}$$

with a banded matrix from D_{S_1} and to give conditions on A for which it meets the criteria for a good preconditioner as set forth in the first section.

Recall that $\zeta_p = T_p + H_p$, where T_p is Toeplitz and H_p is Hankel. Thus any element of D_{S_1} will be Toeplitz plus Hankel. We wish to choose P_b in D_{S_1} such that its Toeplitz part is A. In view of Lemma 2.2 it is clear how the preconditioner is to be constructed. Clearly,

$$P_b(n) \equiv P_b = \sum_{i=0}^{b} a_i \zeta_i \tag{2}$$

is the desired preconditioner. This is why the basis $\{\zeta_p\}_{p=0}^{n-1}$ is convenient, the Toeplitz part of the preconditioner is exactly A. An example is instructive. Let

$$A = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & 0 & 0 & 0 & 0 & 0 \\ a_1 & a_0 & a_1 & a_2 & a_3 & 0 & 0 & 0 & 0 \\ a_2 & a_1 & a_0 & a_1 & a_2 & a_3 & 0 & 0 & 0 \\ a_3 & a_2 & a_1 & a_0 & a_1 & a_2 & a_3 & 0 & 0 \\ 0 & a_3 & a_2 & a_1 & a_0 & a_1 & a_2 & a_3 & 0 \\ 0 & 0 & a_3 & a_2 & a_1 & a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & 0 & a_3 & a_2 & a_1 & a_0 & a_1 & a_2 \\ 0 & 0 & 0 & 0 & a_3 & a_2 & a_1 & a_0 & a_1 \\ 0 & 0 & 0 & 0 & 0 & a_3 & a_2 & a_1 & a_0 \end{pmatrix}.$$

Then $b = 3$ and P_3 is given by

$$\begin{aligned} P_3 &= a_0 I + a_1\zeta_1 + a_2\zeta_2 + a_3\zeta_3 \\ &= A - \begin{pmatrix} a_2 & a_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_3 & a_2 \end{pmatrix}. \end{aligned}$$

In general, as in the above example, it is clear that $A = P_b + H$ where, in block form, H is

$$H = \begin{pmatrix} G & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \hat{G} \end{pmatrix}, \tag{3}$$

and G and $\hat{G}$ are given by

$$G = \begin{pmatrix} a_2 & a_3 & \dots & a_b \\ a_3 & & \cdot^{\cdot^{\cdot}} & 0 \\ \vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ a_b & 0 & \dots & 0 \end{pmatrix}, \qquad \hat{G} = \begin{pmatrix} 0 & \dots & 0 & a_b \\ \vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ 0 & \cdot^{\cdot^{\cdot}} & & a_3 \\ a_b & \dots & a_3 & a_2 \end{pmatrix}.$$

Note that the blocks in (3) are not necessarily the same size. In particular, G and $\hat{G}$ are $(b-1) \times (b-1)$ and the central block of zeros is $(n - 2(b-1)) \times (n - 2(b-1))$.

First observe that P_n satisfies Property 1. Since H has at most $2(b-1)$ nonempty columns $\operatorname{rank}(H) \le 2(b-1)$. Thus if P_b is nonsingular, which we address below, then $P_b^{-1}A = I + P_b^{-1}H$ and all but $2(b-1)$ of the eigenvalues are equal to 1. Thus $P^{-1}A$ has at most $2b-1$ distinct eigenvalues. Therefore P_b satisfies Property 1 if $b \ll n$, which is true for instance when n is increasing while b remains fixed. An identical argument and conclusion are possible for Strang's circulant preconditioner. Indeed, our preconditioner is the S_1–diagonal analogue of Strang's preconditioner.

Regarding Property 4, notice that the complexity of computing Ax is $O(bn)$, while solving $P_n z_{k-1} = r_{k-1}$ is $O(n\log(n))$. (Recall that this is the preconditioning step of the PCGA.) Although asymptotically solving $P_n z_{k-1} = r_{k-1}$ is slower than computing Ax, this does not appear to be a drawback for the problems we consider in the framework of this paper. In fact, in all of our numerical work we have $bn \geq n\log(n)$. Note also that when A is nonbanded (see §4.2), Ax is computed in $O(n\log(n))$ operations so that in this case Property 4 is satisfied. Moreover, as seen in §3 Ax can be computed using only real arithmetic.

Property 3 is satisfied in the banded case by computing the diagonal form of P_b via (2). Lemma 2.2 gives the spectrum of each of the ζ_p's so that $\zeta_p = S_1\Lambda_p S_1$ for some known Λ_p. Thus the diagonal form of the preconditioner is given by $S_1 P_b S_1 = \sum_{i=0}^{b} a_i\Lambda_i$. This formula requires $O(bn)$ operations to compute.

The following method is $O(n\log(n))$ regardless of the bandwidth of A. Let Z be the shift operator, i.e., $(x_1, \dots, x_n)\, Z \equiv (x_2, \dots, x_n, 0)$. In view of Proposition 2.1 only the first row of P_b is required to calculate the spectrum of P_b. Denoting by ρ and τ the first rows of P_b and A, ρ is computed as follows. From $P_b = \sum_{i=0}^{b} a_i\zeta_i$,

$$\rho = a_0 \begin{pmatrix} 1 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}^T + a_1 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}^T + a_2 \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}^T + \cdots + a_b \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}^T = \tau\left(I - Z^2\right).$$

Thus ρ may be computed in $O(n)$ operations via the formula, $\rho = \tau\left(I - Z^2\right)$. Applying Proposition 2.1 proves the following lemma.

LEMMA 4.1. *Let $\Lambda = S_1 P_b S_1$ be the diagonal form of P_b and let σ denote the first row of S_1. Then*

$$\Lambda = \Delta\left(\tau\left(I - Z^2\right) S_1 \left[\Delta(\sigma)\right]^{-1}\right). \tag{4}$$

Proof. From Proposition 2.1 if P_b is an element of D_{S_1}, then

$$S_1 P_b S_1 = \Delta\left(\rho S_1 \left[\Delta(\sigma)\right]^{-1}\right).$$

(Recall that $S_1 = S_1^{-1}$.) Observing that $\rho = \tau\left(I - Z^2\right)$ completes the proof. □

It is now clear that the preconditioner $P_b(n)$ given in (2) satisfies Properties 1, 3, and 4 given in the Introduction.

Only Property 2 remains to be shown. The next theorem gives conditions on A for which P_b is positive definite.

THEOREM 4.2. *If A is an $n \times n$ section of an infinite, positive definite, Toeplitz matrix, $A(\infty)$ with bandwidth $2b+1 < n$ then P_b, is also positive definite.*

Proof. Since $A(\infty)$ is positive definite, $A(\infty) = U(\infty)U(\infty)^T$ where $U(\infty)$ is an upper triangular Toeplitz matrix (see e.g., [5, proof of Thm. 1.1]),

$$U(\infty) = \begin{pmatrix} c_0 & c_1 & \cdots & c_b & 0 & \cdots & \cdots \\ 0 & c_0 & \cdots & c_{b-1} & c_b & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

and hence

$$A = UU^T, \tag{5}$$

where

$$U = \begin{pmatrix} c_0 & c_1 & \cdots & c_b & 0 & \cdots & \cdots & 0 \\ 0 & c_0 & \cdots & c_{b-1} & c_b & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & c_0 & c_1 & \cdots & c_b \end{pmatrix} \in R^{n\times(n+b)}.$$

Remark. All of the zeros of the polynomial $c_0 z^b + \cdots + c_b$ are in the open unit disk (see [5]).

For our purposes it is more convenient to write U in the following block form:

$$U = \begin{pmatrix} \xi & U_1 & 0 \\ 0 & U_2 & 0 \\ 0 & U_3 & \beta \end{pmatrix},$$

where

$$\xi = \begin{pmatrix} c_0 & \cdots & c_{b-2} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_0 \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} c_b & \cdots & 0 \\ \vdots & \ddots & \vdots \\ c_2 & \cdots & c_b \end{pmatrix}.$$

Let $A[a:b, c:d]$ be the submatrix of A consisting of the intersection of rows a through b and columns c through d.

As before,

$$P_b = A - H \text{ and } H = \begin{pmatrix} G & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \hat{G} \end{pmatrix}.$$

Note that since $2b < n+1$,

$$A(b+1:2b-1, 1:b-1) = \begin{pmatrix} a_b & a_{b-1} & \cdots & a_2 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & a_{b-1} \\ 0 & \cdots & 0 & a_b \end{pmatrix} = GJ.$$

Recall from §2.1 that J is the anti-identity. From the block representation above it is clear that $\xi\beta^T = A(b+1:2b-1, 1:b-1)$ so that $G = \xi\beta^T J$. Similarly $\hat{G} = \beta\xi^T J$.

Observe next that $P_b = \hat{U}\hat{L}$ where

$$\hat{U} = \begin{pmatrix} I & U_1 & 0 \\ 0 & U_2 & 0 \\ 0 & U_3 & I \end{pmatrix} \quad \text{and} \quad \hat{L} = \begin{pmatrix} (\xi\xi^T - \xi\beta^T J) & 0 & 0 \\ U_1^T & U_2^T & U_3^T \\ 0 & 0 & (\beta\beta^T - \beta\xi^T J) \end{pmatrix},$$

as may be seen by mulitplying $\hat{U}$ and $\hat{L}$ directly and comparing the result with $UU^T = A$.

With these definitions in place the theorem can now be proved. Since A and H are both symmetric and $P_b = A - H$, we need only show that all of the eigenvalues of P_b are positive.

Suppose λ is an eigenvalue of P_b with corresponding nonzero eigenvector $x = (x_1, \ldots, x_n)^T$. Then at least one of $y \equiv x \pm Jx$ is also a nonzero eigenvector of P_b with the same eigenvalue. (This follows from the identity $JP_b = P_bJ$.) Therefore, we may asssume without loss of generality that y is an eigenvector corresponding to λ such that $Jy = \pm y$ and $\|y\| = 1$.

Partition y as

$$y = \begin{pmatrix} y_1 \\ \hat{y} \\ y_2 \end{pmatrix},$$

where y_1 and y_2 are of length $b-1$. Thus $y_1 = \pm Jy_2$. Moreover ξy_1 and βy_2 are defined.

Next observe that

$$y^T \hat{U} = \left(y_1^T, z^T, y_2^T\right),$$

$$\hat{L}y = \begin{pmatrix} (\xi\xi^T - \xi\beta^T J)y_1 \\ z \\ (\beta\beta^T - \beta\xi^T J)y_2 \end{pmatrix},$$

and that

$$y^T U = \left(y_1^T \xi, z^T, y_2^T \beta\right),$$

$$U^T y = \begin{pmatrix} \xi^T y_1 \\ z \\ \beta^T y_2 \end{pmatrix}. \tag{6}$$

Thus

$$\begin{aligned}
\lambda &= y^T P_b y \\
&= \begin{pmatrix} y_1^T, & \hat{y}^T, & y_2^T \end{pmatrix} \hat{U}\hat{L} \begin{pmatrix} y_1 \\ \hat{y} \\ y_2 \end{pmatrix} \\
&= \begin{pmatrix} y_1^T, & z^T, & y_2^T \end{pmatrix} \begin{pmatrix} (\xi\xi^T - \xi\beta^T J)y_1 \\ z \\ (\beta\beta^T - \beta\xi^T J)y_2 \end{pmatrix} \\
&= y_1^T(\xi\xi^T - \xi\beta^T J)y_1 + y_2^T(\beta\beta^T - \beta\xi^T J)y_2 + z^T z \\
&= y_1^T\xi\xi^T y_1 - y_1^T\xi\beta^T J y_1 + y_2^T\beta\beta^T y_2 - y_2^T\beta\xi^T J y_2 + z^T z \\
&= y_1^T\xi\xi^T y_1 - y_1^T\xi\beta^T J y_1 + y_1^T J\beta\beta^T J y_1 - y_1^T J\beta\xi^T y_1 + z^T z \\
&= y_1^T(\xi - J\beta)(\xi^T - \beta^T J)y_1 + z^T z \\
&= \left\|(\xi^T - \beta^T J)y_1\right\|_2^2 + \|z\|_2^2 \\
&\geq 0.
\end{aligned}$$

We now show that this last inequality is actually strict since $z \neq 0$. The proof is by contradiction. Suppose $z = 0$.

Defining P and Q as

$$P = \begin{pmatrix} c_0 & & 0 \\ \vdots & \ddots & \\ c_b & \cdots & c_0 \end{pmatrix}, \qquad Q = \begin{pmatrix} c_b & & 0 \\ \vdots & \ddots & \\ c_0 & \cdots & c_b \end{pmatrix},$$

and repartitioning y as

$$y = \begin{pmatrix} \omega_1 \\ v \\ \omega_2 \end{pmatrix},$$

where $P\omega_1$ is defined and $\omega_1 = \pm J\omega_2$, we can write

$$Ay = UU^T y = \begin{pmatrix} P^T P\omega_1 \\ 0 \\ QQ^T\omega_2 \end{pmatrix}. \tag{7}$$

Observe that neither c_0 nor c_b are equal to zero since $c_0 c_b = a_b$ and by assumption $a_b \neq 0$. Therefore both $P^T P$ and QQ^T are nonsingular. Moreover since $JA = AJ$ it follows that $P^T P\omega_1 = \pm JQQ^T\omega_2$. In the case of a negative sign, we have $\omega_1^T P^T P\omega_1 = -\omega_2^T QQ^T\omega_2$. The left-hand side of this equality is nonnegative and the right-hand side is nonpositive. Thus $\omega_1 = \omega_2 = 0$, which implies by (7) that $y = 0$ as well. This contradicts the assumption that y is a nonzero eigenvector of A.

In the case of a positive sign, write

$$JP^T PJ\omega_2 = QQ^T\omega_2.$$

Since $JP^T PJ = JP^T JJPJ = PP^T$, we have

$$\begin{pmatrix} c_0 & & 0 \\ \vdots & \ddots & \\ c_b & \cdots & c_0 \end{pmatrix}\begin{pmatrix} c_0 & \cdots & c_b \\ & \ddots & \vdots \\ 0 & & c_0 \end{pmatrix}\omega_2 = \begin{pmatrix} c_b & & 0 \\ \vdots & \ddots & \\ c_0 & \cdots & c_b \end{pmatrix}\begin{pmatrix} c_b & \cdots & c_0 \\ & \ddots & \vdots \\ 0 & & c_b \end{pmatrix}\omega_2. \tag{8}$$

Since $z = 0$ we can rewrite (6) in more detail as

$$\begin{pmatrix} c_0 & & & & 0 \\ c_1 & \cdot & & & \\ \vdots & \cdot & \cdot & & \\ c_b & \cdot & \cdot & \cdot & \\ & \cdot & \cdot & \cdot & c_0 \\ & & \cdot & \cdot & c_1 \\ & & & \cdot & \vdots \\ 0 & & & & c_b \end{pmatrix}\begin{pmatrix} \omega_1 \\ v \\ \omega_2 \end{pmatrix} = \begin{pmatrix} \phi_0 \\ \vdots \\ \phi_{b-2} \\ 0 \\ \vdots \\ 0 \\ \psi_0 \\ \vdots \\ \psi_{b-2} \end{pmatrix}.$$

Inspecting row $b+1$ counting from below, we see that

$$(c_b, \ldots, c_0)\omega_2 = 0. \tag{9}$$

In combination with (8), this implies that

$$\begin{pmatrix} c_0 & & 0 \\ \vdots & \ddots & \\ c_b & \cdots & c_0 \end{pmatrix} \begin{pmatrix} c_0 & \cdots & c_b \\ & \ddots & \vdots \\ 0 & & c_0 \end{pmatrix} \omega_2 = \begin{pmatrix} 0 & & & 0 \\ 0 & c_b & & \\ \vdots & \vdots & \ddots & \\ 0 & c_1 & \cdots & c_b \end{pmatrix} \begin{pmatrix} 0 & & \cdots & 0 \\ 0 & c_b & \cdots & c_1 \\ & & \ddots & \vdots \\ 0 & & & c_b \end{pmatrix} \omega_2,$$

which in turn is equal to

$$\begin{pmatrix} 0 & & & 0 \\ c_b & \ddots & & \\ \vdots & \ddots & \ddots & \\ c_1 & \cdots & c_b & 0 \end{pmatrix} \begin{pmatrix} 0 & c_b & \cdots & c_1 \\ & \ddots & \ddots & \vdots \\ & & \ddots & c_b \\ 0 & & & 0 \end{pmatrix} \omega_2.$$

Thus we have derived the identity

$$\left[\begin{pmatrix} c_0 & & & 0 \\ \vdots & \ddots & & \\ \vdots & & \ddots & \\ c_b & \cdots & \cdots & c_0 \end{pmatrix} \begin{pmatrix} c_0 & \cdots & \cdots & c_b \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ 0 & & & c_0 \end{pmatrix} - \begin{pmatrix} 0 & & & 0 \\ c_b & \ddots & & \\ \vdots & \ddots & \ddots & \\ c_1 & \cdots & c_b & 0 \end{pmatrix} \begin{pmatrix} 0 & c_b & \cdots & c_1 \\ & \ddots & \ddots & \vdots \\ & & \ddots & c_b \\ 0 & & & 0 \end{pmatrix} \right] \omega_2 = 0.$$

It is known (see, for example, Theorem 3 in [9]) that if all zeros of the polynomial $c_0 z^b + \cdots + c_b$ are in the open unit disk (as they are in this case; see the remark at the beginning of the proof), then the vector $(c_0, \ldots, c_b)^T$ is the first column of the inverse of some positive definite Toeplitz matrix, say Φ. Then by the Gohberg–Semencul formula ([5, p. 86])

$$c_0 \Phi^{-1} = \begin{pmatrix} c_0 & & & 0 \\ \vdots & \ddots & & \\ \vdots & & \ddots & \\ c_b & \cdots & \cdots & c_0 \end{pmatrix} \begin{pmatrix} c_0 & \cdots & \cdots & c_b \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ 0 & & & c_0 \end{pmatrix} - \begin{pmatrix} 0 & & & 0 \\ c_b & \ddots & & \\ \vdots & & \ddots & \\ c_1 & \cdots & c_b & 0 \end{pmatrix} \begin{pmatrix} 0 & c_b & \cdots & c_1 \\ & \ddots & \ddots & \vdots \\ & & \ddots & c_b \\ 0 & & & 0 \end{pmatrix}.$$

Since Φ^{-1} is positive definite whenever Φ is we conclude that $\omega_2 = 0$. This contradiction completes the proof. □

Next we show that the inverses of P_b are in fact uniformly bounded or, equivalently, that the minimal eigenvalues of P_b are bounded away from zero independent of n. Let $\lambda_{\min}(n)$ be the minimal eigenvalue of P_b with corresponding eigenvector

$$y(n) = \begin{pmatrix} y_1(n) \\ \hat{y}(n) \\ \pm J y_1(n) \end{pmatrix} = \begin{pmatrix} \omega_1(n) \\ v(n) \\ \pm J\omega_1(n) \end{pmatrix}$$

partitioned as before so that $\xi y_1(n)$ and $P\omega_1(n)$ are defined and normalized so that $\|y(n)\| = 1$. The proof is similar to the last part of the proof of Theorem 4.2, and we use the same notation except that we use n to indicate the order of the involved matrix.

COROLLARY 4.3. *For $n > 2b+1$, there is a constant $\kappa > 0$ such that $\lambda_{\min}(n) > \kappa$.*

Proof. If $\liminf \lambda_{\min}(n) > 0$ there is nothing to prove so without loss of generality assume $\lim_{n\to\infty} \lambda_{\min}(n) = 0$ and hence $\lim_{n\to\infty} \|z(n)\| = 0$. This implies that (9) becomes $\lim_{n\to\infty}(c_b, \ldots, c_0)\omega_2(n) = 0$. It follows as in the proof of Theorem 4.2 that

$$(JPP^TJ \pm QQ^T)\omega_2(n) \to 0,$$

since P and Q are defined independent of n. In the case of a positive sign the matrix is positive definite and fixed. Hence as $n \to \infty$, $\omega_2(n) \to 0$.

In the case of a negative sign, as n increases without bound, it follows that

$$\left[\begin{pmatrix} c_0 & & & 0 \\ \vdots & \ddots & & \\ \vdots & & \ddots & \\ c_b & \cdots & \cdots & c_0 \end{pmatrix} \begin{pmatrix} c_0 & \cdots & \cdots & c_b \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ 0 & & & c_0 \end{pmatrix} \right.$$

$$\left. - \begin{pmatrix} 0 & & & 0 \\ c_b & \ddots & & \\ \vdots & \ddots & \ddots & \\ c_1 & \cdots & c_b & 0 \end{pmatrix} \begin{pmatrix} 0 & c_b & \cdots & c_1 \\ & \ddots & \ddots & \vdots \\ & & \ddots & c_b \\ 0 & & & 0 \end{pmatrix}\right] \omega_2(n) \to 0,$$

as $n \to \infty$ so that once again $\omega_2(n) \to 0$.

In both cases it follows from (7) that $A_n y(n) \to 0$. Since A is positive definite it follows that $\|A^{-1}(n)\| > \|A^{-1}\|$ for all n. Therefore $y(n) \to 0$ as $n \to \infty$. This contradiction completes the proof. □

4.2. Nonbanded matrices. If $A = \left(a_{|i-j|}\right)_{i\,j=1}^n$ is not banded, then it is natural to use

$$P_n(n) = P_n = \sum_{i=0}^{n} a_i \zeta_i \tag{10}$$

as a preconditioner instead of (2). Note that for fixed n, the cost of computing the spectrum of $P_b(n)$ via (4) is fixed regardless of the value of b. In particular the spectrum of $P_n(n)$ may be computed at the same cost. In general, $P_n(n)$ may not satisfy the first two properties. (This is also true of Strang's circulant preconditioner [14].) If however, the Toeplitz matrix $A = A(n)$ is a finite $n \times n$ section of a singly infinite positive definite Toeplitz matrix $A(\infty) = \left(a_{|i-j|}\right)_{i\,j=1}^{\infty}$ such that $\sum_{i=-\infty}^{\infty} |a_i| < \infty$ (in other words $A(\infty)$ is generated by the Wiener class function $f(\theta) = \sum_{k=-\infty}^{\infty} a_k e^{ik\theta}$), then for n sufficiently large these two properities will be satisfied as the following standard argument shows.

Since $\sum_{i=-\infty}^{\infty} |a_i| < \infty$, then given $\varepsilon > 0$, there is an integer $b > 0$ such that $\sum_{|i|>b} |a_i| < \varepsilon$. If we write $A(\infty) = A_b(\infty) + A_\varepsilon(\infty)$, where $A_b(\infty)$ has bandwidth $2b+1$, then clearly

$$\|A_\varepsilon(\infty)\| \leq \sum_{|i|>b} |a_i| < 2\varepsilon.$$

If ε is chosen such that

$$\Delta \equiv \inf \frac{x^T A(\infty) x}{x^T x} \geq 2\varepsilon > 0,$$

then clearly $A_b(\infty)$ is also positive definite, and hence for n large enough $P_b(n)$ is positive definite too, by Theorem 4.2. By Corollary 4.3 we can further constrain ε such that $\lambda_{\min}(P_b(n)) > 2\varepsilon$ for all sufficiently large n.

Writing $P(n) = P_b(n) + P_\varepsilon(n)$ we see that for such ε, $P(n)$ is positive definite whenever $P_b(n)$ is positive definite. Indeed, it is clear that $\|\zeta_p\| \leq 2$ for any p and hence

$$\|P_\varepsilon(n)\| \leq 2 \sum_{i>b} |a_i| = \sum_{|i|>b} |a_i| < 2\varepsilon.$$

Since $A(n) - P(n) = [A_b(n) - P_b(n)] + [A_\varepsilon(n) - P_\varepsilon(n)]$ we get

$$I - A^{-1}(n)P(n) = A^{-1}(n)\left[A_b(n) - P_b(n)\right] + A^{-1}(n)\left[A_\varepsilon(n) - P_\varepsilon(n)\right].$$

Therefore $\left\|A^{-1}(n)\left[A_\varepsilon(n) - P_\varepsilon(n)\right]\right\| \leq 4\varepsilon/\Delta$. Defining $\tilde{\varepsilon}$ to be $4\varepsilon/\Delta$ we see that the the interval $(1-\tilde{\varepsilon}, 1+\tilde{\varepsilon})$ contains the spectrum of $A^{-1}(n)P(n)$ except possibly for $2b+1$ outliers. Therefore the interval $\left(1-\tilde{\varepsilon}+O(\tilde{\varepsilon}^2), 1+\tilde{\varepsilon}+O(\tilde{\varepsilon}^2)\right)$ contains spectrum of $P^{-1}(n)A(n)$ except possibly for $2b+1$ outliers and hence is asymptotically clustered.

We remark that if the problem is poorly conditioned in the sense that $1/\Delta = \left\|A^{-1}(\infty)\right\|_2$ is large, then a large b may be needed to obtain a satisfactory clustering of the eigenvalues of $P^{-1}(n)A(n)$.

We also remark that if A is positive definite itself but not necessarily a section of an infinite matrix, then a positive definite S_1−diagonal preconditioner can be built as follows. Let $D = \Delta(S_1 A S_1)$ and let $P = S_1 D S_1$. Then clearly P is positive definite. It is also the nearest element of D_{S_1} to A in Frobenius norm. This P is the analogue of the optimal circulant preconditioner of T. Chan [3]. We remark that the determination of P via the computation of $S_1 A S_1$ is prohibitively expensive as it requires $O(n^2 \log(n))$ flops. A fast $O(n \log(n))$ method for computing P will be suggested elsewhere.

5. Numerical results. To test the S_1−diagonal preconditioner, we have implemented the PCGA on the Connection Machine 200 at United Technologies Research Center in East Hartford, Connecticut. This computer was configured with 16,384 bit serial processors, 512 floating point processors, and a Vax 6320 front end.

The PCGA was used to solve $Ax = b$ for each of the following matrices and $b = (1, \ldots, 1)^T$.

Matrix 1. $A = \left[\frac{1}{(|i-j|+1)^{1.1}}\right]_{i,j=1}^{n}$.
Matrix 2. $A = \left[\frac{1}{|i-j|+1}\right]_{i,j=1}^{n}$.
Matrix 3. Bandwidth = 41, (see text).
Matrix 4. Bandwidth = 201, (see text).
Matrix 5. $A = \left[\frac{\cos(j)}{j+1}\right]_{i,j=1}^{n}$.
Matrix 6. $A = \left[\frac{1}{(|i-j|+1)^2}\right]_{i,j=1}^{n}$.
Matrix 7. $A = \left[\frac{1}{2^{|i-j|}}\right]_{i,j=1}^{n}$.

Except for Matrix 3 and Matrix 4, most of these matrices have appeared in the literature previously (see [14], [1], and [3]). These were generated in the following manner. It is well known that for $\rho_i \in R$, $|\rho_i| < 1$, $i = 1, \ldots, k$, $k < n-1$ the coefficients of $\prod_{i=1}^{k} (1 - z\rho_i)(1 - z^{-1}\rho_i)$ define for any n an $n \times n$ positive definite, symmetric, Toeplitz matrix with bandwidth $2k + 1$.

Table 1 shows the convergence results for Matrices 1–4. N is the problem size; I, C, and S, represent, respectively, no preconditioning, Tony Chan's circulant preconditioner, and the S_1–diagonal preconditioner. The body of the table gives the iteration count for each matrix and each preconditioner.

TABLE 1

	Matrix 1			Matrix 2			Matrix 3			Matrix 4		
N	I	C	S	I	C	S	I	C	S	I	C	S
255	19	5	5	21	5	5	255	47	9	147	10	7
511	20	5	5	22	5	5	511	37	8	169	9	7
1023	21	5	5	23	5	5	837	29	9	190	9	7
2047	22	5	5	24	6	5	860	21	9	204	9	7
4095	22	6	5	25	6	5	874	17	9	206	9	7
8191	22	6	5	25	6	6	876	16	10	204	9	7

The cases $k = 20$ and $k = 100$, $\rho_i = -d + i\Delta$, $i = 1, \ldots, k$ where $\Delta = \frac{2d}{k}$ and $d = 0.75$ were then used to generate matrices 3 and 4, respectively.

The stopping criterion $\|r_i\|_\infty < 10^{-7}$ was used in all cases, except that the algorithm was terminated if the number of iterations ever exceeded the order of the matrix.

Table 2 shows the convergence results for Matrices 5–7. The format is the same as in Table 1.

TABLE 2

	Matrix 5			Matrix 6			Matrix 7		
N	I	C	S	I	C	S	I	C	S
1023	21	7	7	11	4	4	16	3	3
2047	23	7	7	11	4	4	16	3	3
4095	23	7	7	10	4	4	15	3	3
8191	24	7	7	10	4	4	15	3	3
16383	25	7	7	10	4	4	14	3	3
32767	25	7	6	9	4	4	14	3	3

Since no efficient implementation of the FST was available it was performed via the fast Fourier transform (FFT) as in [11]. The FFT from the Connection Machine Scientific Software Library was used. If an efficient FST were available the algorithm could be faster overall since the FST can be faster than the FFT; see [7]. Moreover, when the coefficient matrix of the problem is real, using the FST eliminates the need for complex arithmetic. Thus storage can also be reduced.

Note that for the nonbanded problems the S_1−diagonal preconditioner is competitive with the circulant preconditioner, while for the banded problems it is clearly superior.

6. Appendix. Other diagonal spaces. This section displays bases for the diagonal spaces of some common fast transforms. All proofs are very similar to the proof of Lemma 2.2 and are therefore omitted.

Denote by S_2, C_1, and C_2 the second discrete sine transform and the two discrete cosine transforms defined in [16]. That is,

$$S_2 \equiv \left(\sin\left(\frac{i(2j-1)\pi}{2n}\right)\right)_{i,j=1}^{n},$$

$$C_2 \equiv \left(\cos\left(\frac{i(2j+1)\pi}{2n}\right)\right)_{i,j=0}^{n-1}, \quad \text{where } k_i = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } i = 1, \\ 1 & \text{otherwise.} \end{cases}$$

$$C_1 \equiv (c_{ij})_{i,j=0}^{n}, \text{ where } c_{ij} \equiv \begin{cases} \frac{1}{2} & \text{if } j = 0, \\ \frac{(-1)^i}{2} & \text{if } j = n, \\ \cos\left(\frac{i(2j+1)\pi}{2n}\right) & \text{otherwise.} \end{cases}$$

LEMMA 6.1. *Let $n \in Z^+$ be given. Then $\{\eta_p\}_{p=0}^{n-1}$ is a basis for D_{S_2}, where η_0 is the identity and*

$$\eta_p(i,j) = \begin{cases} 1 & \text{if } |i-j| = p \text{ and } i < n, \\ -1 & \text{if } i+j = p+1, \\ 1 & \text{if } 2n-(i+j) = p \text{ and } i < n, \\ 2 & \text{if } i = n \text{ and } j = n-p, \\ 0 & \text{otherwise.} \end{cases}$$

The spectrum of η_p, $p > 0$ is $\left\{ 2\cos\left(\frac{(2k-1)p\pi}{2n}\right)\right\}_{k=1.}^{n}$

LEMMA 6.2. *Let $n \in Z^+$ be given. Then $\{\xi_p\}_{p=1}^{n}$ is a basis for D_{C_1} where ξ_0 is the identity and*

$$\xi_p(i,j) = \begin{cases} 2 & \text{if } i = 1 \text{ and } j = p+1, \\ 2 & \text{if } i = n \text{ and } j = n-p, \\ 1 & \text{if } |i-j| = p,\ i \neq 1,\ i \neq n, \\ 1 & \text{if } i+j = p+2,\ j \neq p+1, \\ 1 & \text{if } i+j = 2n-p,\ j \neq n-p, \\ 0 & \text{otherwise.} \end{cases}$$

The spectrum of ξ_p $p > 0$ is $\left\{2\cos\left(\frac{kp\pi}{n-1}\right)\right\}_{k=0.}^{n-1}$

LEMMA 6.3. *Let $n \in Z^+$ be given. Then $\{\xi_p\}_{p=1}^n$ is a basis for $D_{C_2^T}$ where χ_0 is the identity and*

$$\chi_p(i,j) = \begin{cases} 1 & \text{if } |i-j| = p, \\ 1 & \text{if } i+j = p+1, \\ 1 & \text{if } i+j = 2n-p+1, \\ 0 & \text{otherwise.} \end{cases}$$

The spectrum of χ_p is $\left\{ 2\cos\left(\frac{kp\pi}{n}\right)\right\}_{k=1}^n$.

Of course, bases abound for each of the diagonal spaces above but for constructing preconditioners for Toeplitz matrices the bases given have certain advantages. Principally, if we take $P_b = \sum_{i=0}^{n-1} a_i\beta_i$ where $\{\beta_i\}$ is one of $\{\eta_i\}$, $\{\xi_i\}$, or $\{\chi_i\}$ then

$$A - P_b = \begin{pmatrix} G & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \hat{G} \end{pmatrix},$$

as in (3). Thus most of the eigenvalues of $P^{-1}A$ are equal to one and the number of outlying eigenvalues will depend linearly on b.

Also as with the S_1–diagonal preconditioner of §4 all of these transforms are real so complex arithmetic can be avoided by using them rather than the DFT.

Acknowledgments. We would like to thank the referees for many helpful comments and suggestions.

REFERENCES

[1] R. CHAN, *The spectrum of a family of circulant preconditioned Toeplitz systems*, SIAM J. Numer. Anal., 26 (1989), pp. 503–506.

[2] R. CHAN, X. JIN, AND M.-C. YEUNG, *The circulant operator in the Banach algebra of matrices*, Linear Algebra Appl., 149 (1991), pp. 41–53.

[3] T. F. CHAN, *An optimal circulant preconditioner for Toeplitz systems*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 766–771.

[4] P. DAVIS, *Circulant Matrices*, John Wiley and Sons, New York, 1979.

[5] I. GOHBERG AND I. FEL̀DMAN, *Convolution Equations and Projection Methods for their Solution*, American Mathematical Society, Providence, RI, 1974. Translated from Russian by F. M. Goldware.

[6] G. GOLUB AND C. VAN LOAN, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, 1983.

[7] A. K. JAIN, *Fundamentals of Digital Image Processing*, System Sciences Series, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[8] A. JENNINGS, *Influence of eigenvalue spectrum on the convergence rate of the conjugate gradient method*, J. Instit. Math. Appl., 20 (1977), pp. 61–72.

[9] I. KOLTRACHT AND M. NEUMANN, *On the inverse M-matrix problem for real symmetric positive-definite Toeplitz matrices*, SIAM J. Matrix Anal. Appl., 12 (1991), pp. 310–320.

[10] N. LEVINSON, *The Weiner rms error criterion in filter design and prediction*, J. Math. Phys., 25 (1947), pp. 261–278.

[11] W. H. PRESS, B. P. FLANNERY, S. A. TEUKOLSKY, AND W. T. VETTERLING, *Numerical Recipes: The Art of Scientific Computing (Fortran Version)*, Cambridge University Press, Cambridge, MA, 1989.

[12] J. STOER AND R. BULIRSCH, *Introduction to Numerical Analysis*, Springer-Verlag, New York, 1980. Translated from German by R. Bartels, W. Gautschi, and C. Witzgall.

[13] G. STRANG, *Introduction to Applied Mathematics*, Wellesly–Cambridge Press, Wellesly, MA, 1986.

[14] ———, *A proposal for Toeplitz matrix calculations*, Stud. Appl. Math., 74 (1986), pp. 171–176.

[15] E. E. TYRTYSHNIKOV, *Optimal and super-optimal circulant preconditioners*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 459–473.

[16] C. VAN LOAN, *Computational Frameworks for the Fast Fourier Transform*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.