Displacement structure approach to discrete-trigonometric-transform based preconditioners of G.Strang type and of T.Chan type *

Thomas Kailath and Vadim Olshevsky

Information System Laboratory,
Stanford University,
Stanford CA 94305 – 4055, U.S.A.;

E-mail : olshevsk@rascals.stanford.edu

Abstract

In this paper a *displacement structure* technique is used to design a class of new *preconditioners* for the *conjugate gradient method* applied to the solution of large Toeplitz linear equations. Explicit formulas are suggested for the G.Strang-type and for the T.Chan-type preconditioners belonging to any of 8 classes of matrices diagonalized by the corresponding *discrete cosine or sine transforms*. Under the standard Wiener class assumption the *clustering property* is established for all of these preconditioners, guaranteeing a rapid convergence of the preconditioned conjugate gradient method. The formulas for the G.Strang-type preconditioners have another important application : they suggest a wide variety of new $O(m \log m)$ algorithms for multiplication of a Toeplitz matrix by a vector, based on any of the 8 DCT's and DST's.

Recently *transformations* of Toeplitz matrices to Vandermonde-like or Cauchy-like matrices have been found to be useful in developing accurate *direct* methods for Toeplitz linear equations. Here it is suggested to further extend the range of the transformation approach by exploring it for *iterative* methods; this technique allowed us to reduce the complexity of each iteration of the preconditioned conjugate gradient method to 4 discrete transforms per iteration.

1 Introduction

1.1. PCGM for Toeplitz linear equations. We consider the solution of a large linear system of equations $A_m x = b$ whose coefficient matrix A_m is a $m \times m$ leading submatrix of a single-infinite

*This work was supported in part by ARO contract DAAH04-96-1-0176, NSF contract CCR-962811 and also by DARPA contract F49620-95-1-0525. The views and conclusions contained in these documents are those of the author(s) and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the National Science Foundation, the Advanced Research Projects Agency, the Army Research office, or the U.S. Government.

symmetric Toeplitz matrix of the form

$$A = \left[\ a_{|i-j|}\ \right] = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 & \cdots \\ a_1 & a_0 & a_1 & a_2 & \ddots \\ a_2 & a_1 & a_0 & a_1 & \ddots \\ a_3 & a_2 & a_1 & a_0 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}, \tag{1.1}$$

usually associated with the corresponding *generating function* $a(z) = \sum_{k=-\infty}^{\infty} a_{|k|} z^k$. There are special *fast* Toeplitz solvers (for example, the classical Schur and Levinson algorithms) all taking advantage of the structure (1.1) to compute the solution in $O(m^2)$ (or even $O(m \log^2 m)$) operations, thus achieving a favorably efficiency as compared to $O(m^3)$ of structure-ignoring Gaussian elimination.

Along with the above *direct methods*, the *preconditioned conjugate gradient method* (PCGM) for solving Toeplitz linear systems has garnered much attention. This is a well-known *iterative* procedure, for which the main computational burden of each iteration is to compute the product of the coefficient matrix by a vector. In the Toeplitz case it requires $O(m \log m)$ operations per iteration. The number of iterations depends upon the clustering of the spectrum of the A_m, and if the latter has $m - s$ eigenvalues clustered around 1, then PCGM will converge in only s iterations, see, e.g., [GL89]. Unfortunately, classical results on the eigenvalue distribution of Toeplitz matrices (see, e.g., [GS84]) indicate that we cannot expect, in general, any clustering. This disadvantage motivated G.Strang to propose to apply the algorithm to a *preconditioned system*

$$P^{-1}Ax = P^{-1}b. \tag{1.2}$$

where the *preconditioner* P should satisfy the following three requirements.

Property 1. *The complexity of the construction of P should be small, not exceeding $O(m \log m)$ operations.*

Property 2. *A linear system with P should be solved in $O(m \log m)$ operations.*

Property 3. *The spectrum of $P^{-1}A_m$ should be clustered around 1, more precisely the following holds :*

- *For any $\varepsilon > 0$ there exist integers N and s such that for any $m > N$, at most s eigenvalues of $P^{-1}A$ lie outside the interval $[1 - \varepsilon, 1 + \varepsilon]$.*

If a preconditioner satisfying the above properties 1-3 can be constructed, then the complexity of the PCGM will be reduced to only $O(m \log m)$ operations.

1.2. Circulant preconditioners. The first (now well-known) proposed preconditioners of G.Strang [S86] and of T.Chan [C88] were *circulant* matrices, defined respectively by

$$S(A_m) = \text{ circ } (a_0, a_1, a_2, \ldots, a_2, a_1),$$

$$C(A_m) = \text{ circ } (a_0,\ a_1 + \frac{(a_{m-1} - a_1)}{m},\ a_2 + \frac{2(a_{m-2} - a_2)}{m},\ \ldots,\ a_{m-1} + \frac{(m-1)(a_1 - a_{m-1})}{m}).$$

Here circ (r) denotes a circulant matrix specified by its first row r. For these two preconditioners the first property holds by their construction, and since circulant matrices are diagonalized by the discrete Fourier transform (DFT) matrix $\mathcal{F}$, the second property is also immediately satisfied.

Moreover for the case when the generating function is a function from the Wiener class, positive on the unit circle, the 3rd property for $S(A_m)$ and $C(A_m)$ was established in [C89], [CS89] and in [CY92], resp.

Following the Strang proposal [S86] many other circulant preconditioners were designed, including those of R.Chan, E.Tyrtyshnikov, T.Ku and C.Kuo, T.Huckle, and others.

1.3. Noncirculant preconditioners. Part of the motivation for circulant preconditioners stems from the fact that such matrices are diagonalized by the DFT matrix which has fast algorithms for its computation. However, there are other relatives of the DFT with fast $O(m \log m)$ algorithms, for example various versions of discrete cosine and sine transforms. This motivated D.Bini and F.Di Benedetto to propose in [BB90] *non-circulant* analogs of $S(A_m)$ and $C(A_m)$ belonging to the so-called τ-class, introduced earlier in [BC83]. The point is that preconditioners from τ-class have the property 2, because they are all diagonalized by the discrete sine transform I (DST-I). Moreover D.Bini and F.Di Benedetto established the crucial property 3 for their preconditioners. Later DST-I based preconditioners were discussed by several authors.

2 Main results

2.1. G.Strang-type and T. Chan-type preconditioners. In this paper we continue the work started in [S86], [C88], [BB90] and give a systematic account of G.Strang-type and T.Chan-type preconditioners belonging to classes of matrices diagonalized by other discrete trigonometric transforms (we consider 4 DCT's and 4 DST's). For all 8 cases explicit formulas for such preconditioners are obtained, and the above properties 1-3 are established. Moreover, many other useful properties that hold for the classical G.Strang and T.Chan circulant preconditioners are carried over to their discrete-trigonometric-transform analogs.

2.2. Discrete-trigonometric-transform based multiplication of a Toeplitz matrix by a vector. All the computations related to the new preconditioners can be done in real arithmetic. To fully exploit this advanageous property in PCGM one has to suggest discrete cosine/sine transform based algorithm for multiplication of a Toeplitz matrix by a vector. There are two standard FFT based algorithms, one is based on the embedding of a Toeplitz matrix into a circulant, and the second uses a decomposition of a Toeplitz matrix into a sum of a circulant and skew-circulant. It turns out that the new formulas for the G.Strang-type preconditioners lead to a wide variety of generalizations for each of the above two methods, in which (complex) FFT is replaced by any of the 8 considered fast trigonometric transforms.

2.3. Transformations. Recently displacement structure technique has been exploited [P90], [GO94a] to transform different classes of structured matrices (i.e., Toeplitz, Toeplitz-plus-Hankel, Cauchy, Vandermonde , Chebyshev-Vandermonde) to each other. Then this idea has been found to be useful to devise accurate *direct* Toeplitz solvers [He95], [GKO95], [KO95a] among others. Here it is suggested to apply the above approach to *iterative* methods and to transform (1.2) to

$$(T_Q P^{-1} T_Q^T)(T_Q A_m)x = T_Q P^{-1} b,$$

Here T_Q is the corresponding discrete cosine or sine transform matrix, so the transformed preconditioner $(T_Q P^{-1} T_Q^T)$ is now a diagonal matrix. This allows us to save 2 discrete transforms per iteration, because the transformed Toeplitz matrix, $(T_Q A_m)$ has the so-called Chebyshev-Vandermonde-like structure, which allows multiplication with vector with exactly the same complexity 4 discrete-trigonometric transforms, as the initial Toeplitz matrix A_m.

3 A proposal : ∇_{H_Q}-kernel preconditioner

3.1. Displacement structure. We shall design discrete trigonometric-transform based preconditioners in Sec. 5, using an interpretation next. The displacement structure approach initiated by [KKM79] is based on introducing in a linear space of all $m \times m$ matrices a suitable displacement operator $\nabla(\cdot) : \mathbf{R}^{m\times m} \to \mathbf{R}^{m\times m}$ of the form

$$\nabla(R) = R - FRF^T, \qquad \text{or} \qquad \nabla(R) = F^T R - RF. \tag{3.1}$$

A matrix R is said to have ∇-displacement structure, if it is mapped to a low-rank matrix $\nabla(R)$. Since a low-rank matrix can be described by a small number of parameters, a representation of a matrix by its image $\nabla(R)$ often leads to interesting results, and is useful for the design of many fast algorithms. This approach has been found to be useful for studying many different patterns of structure (for example, Toeplitz, Vandermonde, Cauchy, etc.) by specifying for each of them an appropriate displacement operator. For example, *Toeplitz-like matrices* are defined as having displacement structure with respect to the choice

$$\nabla_{Z_1}(R) = R - Z_1 R Z_1^T, \tag{3.2}$$

where $Z_1 = \text{ circ } (0, \ldots, 0, 1)$. The motivation for the above definition can be inferred from the easily verified fact that for any Toeplitz matrix A_m the rank of $\nabla_{Z_1}(A_m)$ does not exceed 2. Although the latter definition of Toeplitz-like matrices was used by several authors [AG90], [GO94a], it is slightly different from the standard one,

$$\nabla_{Z_0}(R) = R - Z_0 R Z_0^T$$

where Z_0 is the lower shift matrix. The crucial difference is that ∇_{Z_1} clearly has a *nontrivial kernel*, so the image $\nabla_{Z_1}(R)$ no longer contains all the information on R. Such matrices R have been called *partially reconstructible* in [KO95a], and systematically studied there. In the Toeplitz-like case Ker ∇_{Z_1} coincides with the subspace of all circulant matrices in $\mathbf{R}^{m\times m}$, so we can observe that the G.Strang and T.Chan preconditioners are both chosen from Ker ∇_{Z_1}.

3.2. A proposal : ∇_{H_Q}-kernel preconditioner. The above displacement operator ∇_{Z_1} is not the only one associated with the class of Toeplitz matrices. We propose to apply the above interpretation, and develop the analogs of G.Strang and T.Chan preconditioners in the kernels of several other related displacement operators of the form

$$\nabla_{H_Q}(R) = H_Q^T R - R H_Q. \tag{3.3}$$

Specifically Toeplitz matrices have displacement rank 4 for all 8 choces for H_Q listed in the second column of Table 1 below. In each of these 8 cases the kernel of the corresponding displacement operator ∇_{H_Q} coincides with the subspace of $\mathbf{R}^{n\times n}$ of all matrices diagonalized by the discrete trigonometric transforms listed in the first column of Table 1, and whose definitions are displayed in Table 2.

Table 1. Matrices H_Q for displacement operator

DCT-I	$H_Q = \ \mathrm{tridiag}\left[\begin{array}{ccccccccccc} & \frac{1}{\sqrt{2}} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \frac{1}{\sqrt{2}} & \\ 0 & & 0 & & 0 & \cdots & 0 & & 0 & & 0 \\ & \frac{1}{\sqrt{2}} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \frac{1}{\sqrt{2}} & \end{array}\right]$
DCT-II	$H_Q = \ \mathrm{tridiag}\left[\begin{array}{ccccccccccc} & \frac{1}{2} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \frac{1}{2} & \\ \frac{1}{2} & & 0 & & 0 & \cdots & 0 & & 0 & & \frac{1}{2} \\ & \frac{1}{2} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \frac{1}{2} & \end{array}\right]$
DCT-III	$H_Q = \ \mathrm{tridiag}\left[\begin{array}{ccccccccccc} & \frac{1}{\sqrt{2}} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \frac{1}{2} & \\ 0 & & 0 & & 0 & \cdots & 0 & & 0 & & 0 \\ & \frac{1}{\sqrt{2}} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \frac{1}{2} & \end{array}\right]$
DCT-IV	$H_Q = \ \mathrm{tridiag}\left[\begin{array}{ccccccccccc} & \frac{1}{2} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \frac{1}{2} & \\ \frac{1}{2} & & 0 & & 0 & \cdots & 0 & & 0 & & -\frac{1}{2} \\ & \frac{1}{2} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \frac{1}{2} & \end{array}\right]$
DST-I	$H_Q = \ \mathrm{tridiag}\left[\begin{array}{ccccccccccc} & \frac{1}{2} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \frac{1}{2} & \\ 0 & & 0 & & 0 & \cdots & 0 & & 0 & & 0 \\ & \frac{1}{2} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \frac{1}{2} & \end{array}\right]$
DST-II	$H_Q = \ \mathrm{tridiag}\left[\begin{array}{ccccccccccc} & \frac{1}{2} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \frac{1}{2} & \\ -\frac{1}{2} & & 0 & & 0 & \cdots & 0 & & 0 & & -\frac{1}{2} \\ & \frac{1}{2} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \frac{1}{2} & \end{array}\right]$
DST-III	$H_Q = \ \mathrm{tridiag}\left[\begin{array}{ccccccccccc} & \frac{1}{2} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \sqrt{\frac{1}{2}} & \\ 0 & & 0 & & 0 & \cdots & 0 & & 0 & & 0 \\ & \frac{1}{2} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \sqrt{\frac{1}{2}} & \end{array}\right]$
DST-IV	$H_Q = \ \mathrm{tridiag}\left[\begin{array}{ccccccccccc} & \frac{1}{2} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \frac{1}{2} & \\ -\frac{1}{2} & & 0 & & 0 & \cdots & 0 & & 0 & & \frac{1}{2} \\ & \frac{1}{2} & & \frac{1}{2} & & \cdots & & \frac{1}{2} & & \frac{1}{2} & \end{array}\right]$

Here we have to explain the designation H_Q. This is a Jacoby (i.e., real symmetric tridiagonal) matrix, and by Q we denote the associated system of orthonormal polynomials, see, e.g., [KO96] for details.

Table 2. Discrete trigonometric transform matrices T_Q.

	Discrete transform	Inverse transform
DCT-I	$C_N^I = \quad \sqrt{\frac{2}{N-1}}\left[\ \eta_k\eta_{N-1-k}\eta_j\eta_{N-1-j}\cos\frac{kj\pi}{N-1}\ \right]_{k,j=0}^{N-1}$	$[C_N^I]^{-1} = [C_N^I]^T = C_N^I$
DCT-II	$C_N^{II} = \quad \sqrt{\frac{2}{N}}\left[\ \eta_k\cos\frac{k(2j+1)\pi}{2N}\ \right]_{k,j=0}^{N-1}$	$[C_N^{II}]^{-1} = [C_N^{II}]^T = C_N^{III}$
DCT-III	$C_N^{III} = \quad \sqrt{\frac{2}{N}}\left[\ \eta_j\cos\frac{(2k+1)j\pi}{2N}\ \right]_{k,j=0}^{N-1}$	$[C_N^{III}]^{-1} = [C_N^{III}]^T = C_N^{II}$
DCT-IV	$C_N^{IV} = \quad \sqrt{\frac{2}{N}}\left[\ \cos\frac{(2k+1)(2j+1)\pi}{4N}\ \right]_{k,j=0}^{N-1}$	$[C_N^{IV}]^{-1} = [C_N^{IV}]^T = C_N^{IV}$
DST-I	$S_N^I = \quad \sqrt{\frac{2}{N+1}}\left[\ \sin\frac{kj}{N+1}\pi\ \right]_{k,j=1}^{N}$	$[S_N^I]^{-1} = [S_N^I]^T = S_N^I$
DST-II	$S_N^{II} = \quad \sqrt{\frac{2}{N}}\left[\ \eta_k\sin\frac{k(2j-1)}{2N}\pi\ \right]_{k,j=1}^{N}$	$[S_N^{II}]^{-1} = [S_N^{II}]^T = S_N^{III}$
DST-III	$S_N^{III} = \quad \sqrt{\frac{2}{N}}\left[\ \eta_j\sin\frac{(2k-1)j}{2N}\pi\ \right]_{k,j=1}^{N}$	$[S_N^{III}]^{-1} = [S_N^{III}]^T = S_N^{II}$
DST-IV	$S_N^{IV} = \quad \sqrt{\frac{2}{N}}\left[\ \sin\frac{(2k-1)(2j-1)}{4N}\pi\ \right]_{k,j=1}^{N}$	$[S_N^{IV}]^{-1} = [S_N^{IV}]^T = S_N^{IV}$

4 G.Strang-type and T.Chan-type preconditioners

The following statement shows that for the purposed of fast $O(m \log m)$ computations it is advantageous to obtain the description of new preconditioners is terms of their first columns.

Proposition 4.1 *Let H_Q be one of the matrices in Table 1, $R = \ Ker\ \nabla_{H_Q}$ (so that R is diagonalized by the corresponding discrete cosine/sine transform matrix T_Q defined in Table 2). Then*

$$T_Q R T_Q^T = \mathrm{diag}(\lambda_1, \dots, \lambda_n), \tag{4.1}$$

where

$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = W_Q^{-1} T_Q \begin{bmatrix} r_0 \\ \vdots \\ r_{n-1} \end{bmatrix}, \tag{4.2}$$

where $\begin{bmatrix} r_0 & \cdots & r_{m-1} \end{bmatrix}^T$ is the first column of R, and W_Q is the corresponding weight matrix defined in Table 3 below.

Table 3.Weight matrices W_Q.

DCT-I	$W_Q = \sqrt{\frac{2}{N-1}}\mathrm{diag}(\frac{1}{\sqrt{2}}, 1, \dots, 1, \frac{1}{\sqrt{2}})$
DCT-II	$W_Q = \sqrt{\frac{2}{N}}\mathrm{diag}(\frac{1}{\sqrt{2}}, \cos(\frac{\pi}{2N}), \dots, \cos(\frac{(N-1)\pi}{2N}))$
DCT-III	$W_Q = \sqrt{\frac{2}{N}} \cdot I$
DCT-IV	$W_Q = \sqrt{\frac{2}{N}}\mathrm{diag}(\cos(\frac{\pi}{4N}), \cos(3\frac{\pi}{4N}) \dots, \cos(\frac{(2N-1)\pi}{4N}))$
DST-I	$W_Q = \sqrt{\frac{2}{N+1}}\mathrm{diag}(\sin(\frac{\pi}{N+1}), \dots, \sin(\frac{N\pi}{N+1}))$
DST-II	$W_Q = \sqrt{\frac{2}{N}}\mathrm{diag}(\sin(\frac{\pi}{2N}), \dots, \sin(\frac{(N-1)\pi}{2N}), \frac{1}{\sqrt{2}}\sin(\frac{\pi}{2}))$
DST-III	$W_Q = \sqrt{\frac{2}{N}}\mathrm{diag}(\sin(\frac{\pi}{2N}), \sin(\frac{3\pi}{2N}), \dots, \frac{1}{\sqrt{2}}\sin(\frac{(2N-1)\pi}{2N}))$
DST-IV	$W_Q = \sqrt{\frac{2}{N}}\mathrm{diag}(\sin(\frac{\pi}{4N}), \sin(\frac{3\pi}{4N}), \dots, \sin(\frac{(2N-1)\pi}{4N}))$

Now we are ready to present formulas for the first columns of G.Strang-type and T.Chan-type preconditioners, denoted by $S_Q(A_m)$ and $C(A_m)$, respectively. In fact, the first columns of these preconditioners are described by the entries of the first column of A_m via

$$\begin{bmatrix} r_0 \\ \vdots \\ r_{m-1} \end{bmatrix} = R_Q \cdot \begin{bmatrix} a_0 \\ \vdots \\ a_{m-1} \end{bmatrix}, \qquad \begin{bmatrix} r_0 \\ \vdots \\ r_{m-1} \end{bmatrix} = G_Q \cdot \begin{bmatrix} a_0 \\ \vdots \\ a_{m-1} \end{bmatrix}, \tag{4.3}$$

respectively, where the matrices G_Q and R_Q are listed in Tables 4 and 5 below.

Table 4. Definition of $S_Q(A_m)$. The matrix R_Q in (4.3).

DCT-I	$\begin{bmatrix} \sqrt{2} & & & & & \\ & 2 & & & & \\ & & 2 & & & \\ & & & \ddots & & \\ & & & & 2 & \\ & & & & & 2\sqrt{2} \end{bmatrix}$	DST-I	$\begin{bmatrix} 1 & 0 & -1 & & & & \\ & 1 & 0 & -1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & -1 & \\ & & & & 1 & 0 & -1 \\ & & & & & 1 & 0 \\ & & & & & & 1 \end{bmatrix}$

Table 4.Continuation. Definition of $S_Q(A_m)$. The matrix R_Q in (4.3).

DCT-II	$\begin{bmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & 1 \\ & & & & 1 \quad 1 \end{bmatrix}$	DST-II	$\begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ & & & & 1 \quad -1 \end{bmatrix}$
DCT-III	$\begin{bmatrix} \sqrt{2} & & & & & \\ & 2 & & & & \\ & & 2 & & & \\ & & & \ddots & & \\ & & & & 2 & \\ & & & & & 2 \end{bmatrix}$	DST-III	$\begin{bmatrix} 1 & 0 & -1 & & & & \\ & 1 & 0 & -1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & -1 & \\ & & & & 1 & 0 & -1 \\ & & & & & 1 & 0 \\ & & & & & & \sqrt{2} \end{bmatrix}$
DCT-IV	$\begin{bmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & 1 \\ & & & & 1 \quad 1 \end{bmatrix}$	DST-IV	$\begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ & & & & 1 \quad -1 \end{bmatrix}$

Table 5. Definition of the T.Chan-type preconditioner $C_Q(A_m)$. Matrix G_Q for (4.3).

DCT-I	$G = \frac{1}{(m-1)^2} \cdot$ diag $\{\sqrt{2}, 2, 2, \ldots 2, \sqrt{2}\}(D+E+L+U)$, with the terms specified by (4.4), (4.5), (4.6), (4.7)

$$D = \text{ diag } ((m-1)^2, \boxed{2\sqrt{2}(m-1)+(m-3)(m-3)}, \boxed{2\sqrt{2}(m-1)+(m-3)(m-4)}, \ldots \tag{4.4}$$

$$\ldots, \boxed{2\sqrt{2}(m-1)+2(m-3)}, \boxed{2\sqrt{2}(m-1)+(m-3)}, \boxed{2\sqrt{2}(m-1)}, (2m-3)),$$

(a recursion for the $2, 3, \ldots, m-2, m-1$ entries is apparent.)

$$E = -2\sqrt{2} \text{ toeplitz } (\begin{bmatrix} 1, & 0, & 1, & 0, & 1, & 0, & \ldots \end{bmatrix}) \cdot \text{ diag } (\begin{bmatrix} 0, & 1, & 1, & \ldots, 1, & 0 \end{bmatrix}) \tag{4.5}$$

Here we follow the MATLAB notations, where toeplitz (c, r) denotes the Toeplitz matrix with the first column c and the first row r. toeplitz (c) denotes the symmetric Toeplitz matrix with the first column c.

$$L = \text{ toeplitz } (\begin{bmatrix} 0, & 0, & 1, & 0, & 1, & 0, & 1, & 0, & \ldots \end{bmatrix}, \begin{bmatrix} 0, & 0, & 0, & \ldots \end{bmatrix}) \times \tag{4.6}$$

$$\text{diag } (\begin{bmatrix} 0, & 2 \cdot 2, & 2 \cdot 3, & \ldots, & 2 \cdot (m-2), & 0, & 0 \end{bmatrix})$$

$$U = \text{ toeplitz } (\begin{bmatrix} 0, & 0, & 0 & \ldots \end{bmatrix}, \begin{bmatrix} 0, & 0, & 1, & 0, & 1, & 0, & 1, & 0, & \ldots \end{bmatrix}) \times \tag{4.7}$$

$$\text{diag } (\begin{bmatrix} 0, & 0, & -2(m-4), & -2(m-3), & \ldots, & -4, & -2, & 0, & -1 \end{bmatrix}).$$

Table 5. Continuation. Definition of the T.Chan-type preconditioner $C_Q(A_m)$. Matrix G_Q for (4.3).

DCT-II	$\frac{1}{m^2}\begin{bmatrix} m^2 & (m-1)(m-2) & -2(m-2) & \cdots & -4 & -2 \\ 0 & m^2-(m-2) & (m-2)(m-2) & \ddots & -4 & -2 \\ 0 & 2 & m^2-2(m-2) & \ddots & -4 & -2 \\ 0 & 2 & 4 & \ddots & 2(m-2) & -2 \\ 0 & 2 & 4 & \ddots & m^2-(m-2)(m-2) & (m-2) \\ 0 & 2 & 4 & \cdots & 2(m-2) & m^2-(m-1)(m-2) \end{bmatrix}$
DCT-III	$\frac{1}{m}\begin{bmatrix} \sqrt{2}m & & & & & \\ & 2(m+\sqrt{2}-2) & & \mathbf{0} & & \\ & & 2(m+\sqrt{2}-3) & & & \\ & & & \ddots & & \\ & \mathbf{0} & & & 2(\sqrt{2}+1) & \\ & & & & & 2\sqrt{2} \end{bmatrix}$
DCT-IV	$\frac{1}{m}\begin{bmatrix} m & m-1 & & & & \\ & m-1 & m-2 & & \mathbf{0} & \\ & & m-2 & \ddots & & \\ & & & \ddots & 2 & \\ & \mathbf{0} & & & 2 & 1 \\ & & & & & 1 \end{bmatrix}$
DST-I	$\frac{1}{m+1}\begin{bmatrix} m+1 & 0 & -(m-2) & & & & \\ & m+1 & 0 & -(m-3) & & & \\ & & m & \ddots & \ddots & & \\ & & & \ddots & 0 & -2 & \\ & \mathbf{0} & & & 5 & 0 & -1 \\ & & & & & 4 & 0 \\ & & & & & & 3 \end{bmatrix}$
DST-II	$\frac{1}{m^2}\begin{bmatrix} m^2 & -(m-1)(m-2) & -2(m-2) & \cdots & (-1)^{m-2}4 & (-1)^{m-1} \\ 0 & m^2-(m-2) & -(m-2)(m-2) & \ddots & \vdots & \vdots \\ 0 & -2 & m^2-2(m-2) & \ddots & -4 & 2 \\ 0 & 2 & -4 & \ddots & -2(m-2) & -2 \\ 0 & \vdots & \vdots & \ddots & m^2-(m-2)(m-2) & -(m-2) \\ 0 & (-1)^m 2 & (-1)^{m-1}4 & \cdots & -2(m-2) & m^2-(m-1)(m-2) \end{bmatrix}$

Table 5. Continuation. Definition of the T.Chan-type preconditioner $C_Q(A_m)$. Matrix G_Q for (4.3).

DST-III	$\frac{1}{m}\begin{bmatrix} m & 0 & -(-3+\sqrt{2}) & & & & \\ & m-2+\sqrt{2} & 0 & \ddots & & & \\ & & m-3+\sqrt{2} & \ddots & -(2+\sqrt{2}) & & \\ & & & \ddots & 0 & -(1+\sqrt{2}) & \\ & \mathbf{0} & & & 2+\sqrt{2} & 0 & -\sqrt{2} \\ & & & & & 1+\sqrt{2} & 0 \\ & & & & & & 2 \end{bmatrix}$
DST-IV	$\frac{1}{m}\begin{bmatrix} m & -(m-1) & & & & \\ & m-1 & -(m-2) & & \mathbf{0} & \\ & & m-2 & \ddots & & \\ & & & \ddots & -2 & \\ & & \mathbf{0} & & 2 & -1 \\ & & & & & 1 \end{bmatrix}$

The DST-I based G.Strang-type and T.Chan-type preconditioners in Tables 4 and 5 were obtained earlier in [BB90].

5 Terminology: G.Strang-type and T.Chan-type preconditioners

Recall that the classical T.Chan preconditioner $C(A_m)$ is the best Frobenious norm circulant approximant of A_m. Similarly, in all 8 cases we have :

$$\|C_Q(A_m) - A_m\|_F = \min_{R \in \text{ Ker } \nabla_{H_Q}} \|R - A_m\|_F,$$

thus justifying the name T.Chan-type preconditioner for the $C_Q(A_m)$.

Further, the classical G.Strang preconditioner is a "Toeplitz-plus-Toeplitz" matrix,

$$S(A_m) = A_m + T,$$

where the second Toeplitz term can be decomposed into a sum $T = T_{lr} + T_{sn}$ of a low-rank and a small-norm matrices. Similarly, all preconditioners $S_Q(A_m)$ are "Toeplitz-plus-Hankel-plus-border" matrices,

$$S_Q(A_m) = A_m + H + B,$$

where the "border" matrix B is nonzero only in the first and last rows and columns, and the Hankel matrix can be decomposed into a sum, $H = H_{lr} + H_{sn}$, of a low-rank and small-norm matrices.

6 Properties

In this section we assume that the generating function, $a(x)$ is a function from Wiener class,

$$\sum_{k=-\infty}^{\infty} |a_k| < \infty,$$

positive on the unit circle. It is well-known that these conditions imply that A_m is positive definite for all m. Under these conditions all new preconditioners satisfy the 3 properties stated in Introduction, thus guaranteeing the rapid convergence of the PCGM. Moreover, they also have several useful properties, satisfied by the classical G.Strang and T.Chan circulant preconditioners.

- We have Then
$$\lim_{m\to\infty} \|C_Q(A_m) - S_Q(A_m)\|_2 = 0,$$
where $\|\cdot\|_2$ denotes the spectral norm in $\mathbf{R}^{m\times m}$. For the circulant G.Strang and T.Chan preconditioners such a property was established in [C89].

- For any $\varepsilon > 0$ there exist $M > 0$ so that for $m > M$ the spectrum of $S_Q(A_m)$ lies in the interval $[\min_{|z|=1} a(z) - \varepsilon, \max_{|z|=1} a(z) + \varepsilon]$.

- In particular, in all 8 cases $S_Q(A_m)$ is positive definite for sufficiently large m.

- We have
$$\lambda_{min}(A_m) \le \lambda_{min}(C_Q(A_m)) \le \lambda_{max}(C_Q(A_m)) \le \lambda_{max}(A_m). \tag{6.1}$$
For the circulant T.Chan preconditioner such a property was proved in [T92] and [CJY91].

- In particular, in all 8 cases $C_Q(A_m)$ are positive definite independently of m.

- The norms $\|S_Q(A_m)\|_2$, $\|S_Q(A_m)^{-1}\|$, $\|C_Q(A_m)\|_2$ and $\|C_Q(A_m)^{-1}\|$ are uniformly bounded independently of m.

Note that these properties for the DST-I based preconditioners were established earlier in [BB90].

7 Discrete-trigonometric-transform based multiplication of a Toeplitz matrix by a vector

There are two well-known methods to multiply a Toeplitz matrix by a vector, the first based on the embedding of A_m into a larger $2m \times 2m$ circulant matrix, and the second based on a representation of A_m as a sum of circulant and a skew-circulant. In both cases the multiplication is reduced to 4 FFT's of the order m. In this and next section we present a wide variety of analogs of these two methods, based on discrete-trigonometric transforms.

To generalize the first, "embedding-into-a-circulant" method, we recall from Sec. 5 that G.Strang-type preconditioners admit a "Toeplitz-plus-Hankel-plus-border" decomposition,

$$S_Q(A_m) = A_m + H + B,$$

with Hankel and "border" components listed in Table 6 below.

Table 6. A Hankel part and a "border" part of $S_Q(A)$.

	H	B
DCT-I	$\begin{bmatrix} a_0 & a_1 & \cdots & a_{m-2} & 2a_{m-1} \\ a_1 & a_2 & \iddots & \iddots & a_{m-2} \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ a_{m-2} & \iddots & \iddots & a_2 & a_1 \\ 2a_{m-1} & a_{m-2} & \cdots & a_1 & a_0 \end{bmatrix}$	$(\sqrt{2}-2) \cdot \left[\begin{array}{c\|ccc\|c} -\frac{a_0}{\sqrt{2}-2} & a_1 & \cdots & a_{m-2} & -\frac{a_{m-1}}{\sqrt{2}-2} \\ \hline a_1 & & & & a_{m-1} \\ \vdots & & 0 & & \vdots \\ a_{m-2} & & & & a_1 \\ \hline -\frac{a_{m-1}}{\sqrt{2}-2} & a_{m-2} & \cdots & a_1 & -\frac{a_0}{\sqrt{2}-2} \end{array}\right]$
DCT-II	$\begin{bmatrix} a_1 & a_2 & \cdots & a_{m-1} & 0 \\ a_2 & a_3 & \iddots & \iddots & a_{m-1} \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ a_{m-1} & \iddots & \iddots & a_3 & a_2 \\ 0 & a_{m-1} & \cdots & a_2 & a_1 \end{bmatrix}$	0
DCT-III	$\begin{bmatrix} a_0 & a_1 & \cdots & a_{m-2} & a_{m-1} \\ a_1 & a_2 & \iddots & \iddots & 0 \\ \vdots & \iddots & \iddots & \iddots & -a_{m-1} \\ a_{m-2} & \iddots & \iddots & \iddots & \vdots \\ a_{m-1} & 0 & -a_{m-1} & \cdots & -a_2 \end{bmatrix}$	$(\sqrt{2}-2) \cdot \left[\begin{array}{c\|cccc} -\frac{a_0}{\sqrt{2}-2} & a_1 & \cdots & a_{m-2} & a_{m-1} \\ \hline a_1 & & & & \\ \vdots & & & 0 & \\ a_{m-2} & & & & \\ a_{m-1} & & & & \end{array}\right]$
DCT-IV	$\begin{bmatrix} a_1 & a_2 & \cdots & a_{m-1} & 0 \\ a_2 & a_3 & \iddots & \iddots & -a_{m-1} \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ a_{m-1} & \iddots & \iddots & -a_3 & -a_2 \\ 0 & -a_{m-1} & \cdots & -a_2 & -a_1 \end{bmatrix}$	0
DST-I	$\begin{bmatrix} -a_2 & \cdots & -a_{m-1} & 0 & 0 \\ \vdots & \iddots & \iddots & \iddots & 0 \\ -a_{m-1} & \iddots & \iddots & \iddots & -a_{m-1} \\ 0 & \iddots & \iddots & \iddots & \vdots \\ 0 & 0 & -a_{m-1} & \cdots & -a_2 \end{bmatrix}$	0
DST-II	$\begin{bmatrix} -a_1 & -a_2 & \cdots & -a_{m-1} & 0 \\ -a_2 & -a_3 & \iddots & \iddots & -a_{m-1} \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ -a_{m-1} & \iddots & \iddots & -a_3 & -a_2 \\ 0 & -a_{m-1} & \cdots & -a_2 & -a_1 \end{bmatrix}$	0
DST-III	$\begin{bmatrix} -a_2 & \cdots & -a_{m-1} & 0 & a_{m-1} \\ \vdots & & \iddots & \iddots & a_{m-2} \\ -a_{m-1} & \iddots & \iddots & \iddots & \vdots \\ 0 & \iddots & \iddots & a_2 & a_1 \\ a_{m-1} & a_{m-2} & \cdots & a_1 & a_0 \end{bmatrix}$	$\frac{-1}{1+\sqrt{2}} \cdot \left[\begin{array}{cccc\|c} & & & & a_{m-1} \\ & & 0 & & a_{m-2} \\ & & & & \vdots \\ & & & & a_1 \\ \hline a_{m-1} & a_{m-2} & \cdots & a_1 & (1+\frac{1}{\sqrt{2}})a_0 \end{array}\right]$
DST-IV	$\begin{bmatrix} -a_1 & -a_2 & \cdots & -a_{m-1} & 0 \\ -a_2 & -a_3 & \iddots & \iddots & a_{m-1} \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ -a_{m-1} & \iddots & \iddots & a_3 & a_2 \\ 0 & a_{m-1} & \cdots & a_2 & a_1 \end{bmatrix}$	0

These formulas allow us an embedding of A_m in a larger $2m \times 2m$ matrix, diagonalized by any of the 8 DCT's and DST's. Specifically,

- First, we embed a $m \times m$ Toeplitz matrix A_m into a larger $2m \times 2m$ Toeplitz matrix $\mathcal{A}_{2m}$ by padding its first column with m zeros.

- Secondly, we construct for $\mathcal{A}_{2m}$ the G.Strang-type preconditioner $S_Q(\mathcal{A}_{2m})$.

Clearly, the banded structure of $\mathcal{A}_{2m}$ implies that A_m is a submatrix of $S_Q(\mathcal{A}_{2m})$, thus suggesting a number of $O(m \log m)$ multiplication algorithms based on any of the 8 DCT's and DST's.

Although this is beyong our needs in the present paper, note that the formulas for the G.Strang-type preconditioners allow us to multiply in the same way Toeplitz-plus-Hankel matrices by a vector. These methods generalize "embedding-into-a-circulant" algorithm. In the next section we suggest (as a by-product) another set of algorithms, generalizing "circulant-plus-skew-circulant" method.

8 Transformation to Chebyshev-Vandermonde-like matrix

12.1. Transformations. Recently the transformation technique of structured matrices from one class to another has been found to be useful for direct Toeplitz solvers. Here it is suggested to use the results of [KO95a] on transformation of *partially reconstructible* matrices and apply this technique to iterative methods, and instead of applying PCGM to the preconditioned system

$$S_Q(A_m)^{-1} A_m x = b, \tag{8.1}$$

we suggest to apply it to the *transformed system*

$$(T_Q S_Q(A_m)^{-1} T_Q^T) \cdot (T_Q A_m) x = T_Q b$$

where the preconditioner is transformed to the diagonal matrix $(T_Q S_Q(A_m)^{-1} T_Q^T)$, and the Toeplitz matrix A_m is transformed into a Chebyshev-Vandermonde-like matrix $T_Q A_m$. Since a diagonal linear system can be solved in m operations, such transformation saves us 2 discrete transforms per iteration, if we can multiply the Chebyshev-Vandermonde-like matrix $T_Q A_m$ by a vector with exactly the same complexity as the initial matrix A_m. In the rest of the section we describe algorithms for this purpose, these new algorithms are based on the formulas, which are counterparts of the well-know decomposition of a Toeplitz matrix into a sum of a circulant and a skew-circulant matrices.

12.2. Discrete transforms II and IV. For example, from the Toeplitz-plus-Hankel-plus-border decompositions in Table 7 it immediately follows that ,

$$A_m = \frac{1}{2}(S_{C2}(A_m) + S_{S2}(A_m)), \qquad A_m = \frac{1}{2}(S_{C4}(A_m) + S_{S4}(A_m)). \tag{8.2}$$

where $S_{C2}(A_m)$ denotes G.Strang-type preconditioner based on DCT-2, with the other preconditioners designated similarly. Since such preconditioners $S_Q(A_m)$ are diagonalized by the corresponding transform matrices T_Q, each of these formulas clearly allow us to multiply A_m by a vector in just 4 discrete trigonometric transforms (with 2 more transforms needed only once to diagonalize $S_Q(A)$). As to our goal, i.e., $T_Q A_m$, we have the following formulas.

Table 12. Decompositions for II and IV transforms.

DCT-II	$T_{C2}A_m = D_{C2}T_{C2} + T_{C2}T_{S2}^T D_{S2}T_{S2}$	DST-II	$T_{S2}A_m = T_{S2}T_{C2}^T D_{C_2}T_{C2} + D_{S2}T_{S_2}$
DCT-IV	$T_{C4}A_m = D_{C4}T_{C4} + T_{C4}T_{S4}^T D_{S4}T_{S4}$	DST-IV	$T_{S4}A_m = T_{S4}T_{C4}^T D_{C_4}T_{C4} + D_{S4}T_{S_4}$

Here

$$D_Q = W_Q T_Q R_Q \cdot \begin{bmatrix} a_0 \\ \vdots \\ a_{m-1} \end{bmatrix} \tag{8.3}$$

where the matrices R_Q are displayed in Table 4, weight matrices W_Q are collected in Table 3, and the definitions of discrete transform matrices T_Q are listed in Table 2. These formulas reduce the complexity of one iteration to 4 real discrete trigonometric transform of the order m, as compared to 6 such transforms of the methods in the previous section.

12.3. Discrete transforms I. Thus for the II and IV transforms the formulas (8.2) seem to be simple because in these cases the Hankel part of the corresponding cosine and sine G.Strang-type preconditioners differ only by the sign (see, e.g., Table 6). For the I and III transforms this is not so, and the reason seem to be that the definitions of the corresponding discrete transforms are not chosen to imply for them the representations of the form (8.2). However, instead of changing the standard definitions (for example, taking care of different $N+1$ and $N-1$ and of the size for the DCT-I, DST-I, DCT-III and DST-III), we show that even with standard definitions in the remaining two cases one can derive not much more involved formulas, also leading to the same efficiency of 4 discrete transforms per iteration.

Indeed, in the case of DCT-I and DST-I we have the following. Let the numbers $\{c_k\}$, $\{s_k\}$ be defined by

$$(I + (Z^T)^2) \begin{bmatrix} f_0 \\ \vdots \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} a_0 \\ \vdots \\ a_{n-1} \end{bmatrix}, \qquad \begin{bmatrix} e_0 \\ \vdots \\ e_{n-1} \end{bmatrix} = (Z^T)^2 \begin{bmatrix} f_0 \\ \vdots \\ f_{n-1} \end{bmatrix}$$

where Z denotes the lower shift matrix. Then clearly

$$A = S_{C1}(E_m) + S_{S1}(F_m) - B_{C1}$$

where $S_{C1}(E_m)$, $S_{S1}(F_m)$ are G.Strang-type preconditioners from Table 7 for Toeplitz matrices E_m and F_m defined by their first columns $[e_k]$ and $[f_k]$, resp. The matrix B_{C1} is the border matrix of F_m defined in the row DCT-I of the same Table. Therefore we have the following formulas.

<u>Table 12.</u> Continuation.

DCT-I	$T_{C1}A_m = D_{C1}T_{C1} + T_{C1}(T_{S1}^T D_{S1} T_{S1} + B_{C1})$
DST-I	$T_{S1}A_m = T_{S1}(T_{C1}^T D_{C1} T_{C1} + B_{C1}) + D_{S1}T_{S1}$

Here all the diagonal matrices are obtained by (8.3) with the replacement of $[a_k]$ by the $[f_k]$ and $[e_k]$, resp. Since B_{C1} is the rank-four matrix, these formulas allow us to compute the product of $T_Q A_m$ by a vector in 4 real trigonometric transforms of the order m. Note that a different formula of this kind was obtained for DST-I in [H95].

12.4. Discrete transforms III. In this case we have

$$A_m = \frac{1}{2}(S_{C3}(A_m) - B_{C3} + Z S_{S3}(A_m) Z^T),$$

$$A_m = \frac{1}{2}(Z^T S_{C3}(A_m) Z + S_{S3}(A_m) - B_{S3}),$$

leading to the formulas in the next Table.

Table 12. Continuation.

DCT-III	$T_{C3}A_m = \frac{1}{2}(D_{C3}T_{C3} + T_{C3}(ZT_{S3}^T D_{S3}T_{S3}Z^T - B_{C3})$
DST-III	$T_{S3}A_m = \frac{1}{2}(T_{S1}(Z^T T_{C3}^T D_{C3}T_{C3}Z - B_{S3}) + D_{S3}T_{S3}$

Again the complexity of one iteration is 4 real discrete trigonometric transforms per iteration.

References

[AG90] Ammar G. and Gader P., *New decompositions of the inverse of a Toeplitz matrix*, in *Signal processing, Scattering and Operator Theory, and Numerical Methods*, Proc. Int. Symp. MTNS-89, vol. III (Kaashoek M.A., van Schuppen J.H. and Ran A.C.M., Eds), 421-428, Birkhauser, Boston, 1990.

[BB90] D.Bini and F. Di Benedetto, *A new preconditioner for the parallel solution of of positive definite Toeplitz systems*, in Proc. 2nd ACM Symp. on Parallel Algorithms and Architectures, Crete, Greece, 1990, 220-223.

[BC83] D. Bini and M.Capovani, *Spectral and computational properties of band symmetric Toeplitz matrices*, Linear Algebra Appl., **52** (1983), 99-126.

[BK95] E.Boman and I.Koltracht, *Fast transform based preconditioners for Toeplitz equations*, SIAM J. on Matrix Analysis and Appl., **16**(1995), 628-645.

[C88] T.Chan, *An optimal circulant preconditioner for Toeplitz systems*, SIAM J. Sci. Stat. Comput., **9**(1988), 766-771.

[C89] R.Chan, *Circulant preconditioners for Hermitioan Toeplitz systems*, SIAM J. Matrix Analysis and Appl., **10**(1989), 542-550.

[CJY91] R.Chan, X.Jin and M.Yeung, *The circulant operator in the Banach algebra of matrices*, Linear Algebra Appl., **149** (1991), 41-53.

[CS89] R.Chan and G.Strang, *Toeplitz equations by conjugate gradients with circulant preconditioner*, SIAM J. Sci. Stat. Comp., **10** (1989), 104-119.

[CY92] R.Chan and M.Yeng, *Circulant preconditioners for Toeplitz matrices with positive continuous generating functions*, Math. Comp., **58**(1992), 233-240.

[GKO95] I.Gohberg, T.Kailath and V.Olshevsky, *Fast Gaussian elimination with partial pivoting for matrices with displacement structure*, Math. of Computation, **64** (1995), 1557-1576.

[GL89] G. Golub and C, Van Loan, *Matrix Computations*, second edition, Johns Hopkins U. P., Baltimore, 1989.

[GO94a] I.Gohberg and V.Olshevsky, *Complexity of multiplication with vectors for structured matrices*, Linear Algebra Appl., **202** (1994), 163 – 192.

[GO94b] I.Gohberg and V.Olshevsky, *Fast state space algorithms for matrix Nehari and Nehari-Takagi interpolation problems*, Integral Equations and Operator Theory, **20**, No. 1, 1994, 44 – 83.

[GS84] U.Grenader and G.Szegö, *Toeplitz forms and their applications*, 2nd ed., Chelsea, New York, 1984. January 1987.

[H95] T.Huckle, *Cauchy matrices and iterative methods for Toeplitz matrices*, in Proc. of SPIE-95, **2563**(1995), 281-292.

[He95] G. Heinig, *Inversion of generalized Cauchy matrices and other classes of structured matrices*, in : Linear Algebra in Signal Processing, IMA volumes in Mathematics and its Applications, vol. **69** (1995), 95 - 114.

[KKM79] T.Kailath, S.Kung and M.Morf, *Displacement ranks of matrices and linear equations*, J. Math. Anal. and Appl., **68** (1979), 395-407.

[KO95a] T.Kailath and V.Olshevsky, *Bunch-Kaufman Pivoting for Partially Reconstructable Cauchy-like Matrices, with Applications to Toeplitz-like Linear Equations and to Boundary Rational Matrix Interpolation Problems*, to appear Linear Algebra and Appl., (1997), Proc. of the 5-th ILAS conference, 1995.

[KO95b] T.Kailath and V.Olshevsky, *Displacement structure approach to Chebyshev-Vandermonde and related matrices*, Integral equations and Operator Theory, **22**, 1995, 65-92.

[KO96] T.Kailath and V.Olshevsky, *Displacement structure approach to discrete transform based preconditioners of T.Chan and G.Strang types*, submitted, 1996.

[P90] V.Pan, *On computations with dense structured matrices*, Math. of Computation, **55**, No. 191 (1990), 179 – 190.

[S86] G.Strang, *A proposal for Toeplitz matrix calculations*, Stud. Appl. Math., **74** (1986), 171-176.

[T92] E.Tyrtyshnikov, *Optimal and superoptimal circulant preconditioners*, SIAM J. on Matrix Analysis Appl., **13**(1992), 459-473.