SIAM J. MATRIX ANAL. APPL.
Vol. 24, No. 3, pp. 660–677

INVERSION OF DISPLACEMENT OPERATORS*

VICTOR Y. PAN† AND XINMAO WANG‡

Abstract. We recall briefly the displacement rank approach to the computations with structured matrices, which we trace back to the seminal paper by Kailath, Kung, and Morf [*J. Math. Anal. Appl.*, 68 (1979), pp. 395–407]. The concluding stage of the computations is the recovery of the output from its compressed representation via the associated displacement operator L. The recovery amounts to the inversion of the operator. That is, one must express a structured matrix M via its image $L(M)$. We show a general method for obtaining such expressions that works for all displacement operators (under only the mildest nonsingularity assumptions) and thus provides the foundation for the displacement rank approach to practical computations with structured matrices. We also apply our techniques to specify the expressions for various important classes of matrices. Besides unified derivation of several known formulae, we obtain some new ones, in particular, for the matrices associated with the tangential Nevanlinna–Pick problems. This enables acceleration of the known solution algorithms. We show several new matrix representations of the problem in the important confluent case. Finally, we substantially improve the known estimates for the norms of the inverse displacement operators, which are critical numerical parameters for computations based on the displacement approach.

Key words. structured matrices, displacement rank, inverse displacement operators, tangential confluent Nevanlinna–Pick problem

AMS subject classifications. 47A57, 47B35, 47N40, 68Q25, 65F05

PII. S089547980238627X

1. Introduction.

1.1. Displacement rank approach to computations with structured matrices.

Structured matrices are omnipresent in computations for communication, sciences, and engineering (see the extensive bibliographies in [KS95], [KS99], and [P01]). Figure 1.1 displays the four most popular classes of structured matrices. They are generalized to various other highly important matrix structures in the *displacement rank approach*, which we trace back to the seminal paper [KKM79]. We next follow [P01] to outline this approach, which treats various matrix structures in a unified way, based on their association with the *displacement operators*, and then focus on its most fundamental stage of the inversion of the displacement operators.

An $n \times n$ structured matrix M can be associated with an appropriate displacement operator L such that $r = \text{rank}(L(M))$ is small, $r \ll n$. The image matrix $L(M)$ is called the *displacement* of M, and r is called the *displacement rank* of M. The n^2 entries of the displacement $L(M)$ can be represented via only $2rn$ parameters. Such a compressed representation of $L(M)$ can be extended to the matrix M by inverting the displacement operator.

Example 1.1 (Cauchy-like matrices; see [HR84], [GO94]). Let $D(\mathbf{s}) = \text{diag}(s_i)_{i=1}^m$, $D(\mathbf{t}) = \text{diag}(t_j)_{j=1}^n$, $\mathbf{s} = (s_i)_{i=1}^m$, and $\mathbf{t} = (t_j)_{j=1}^n$, where all s_i, t_j are distinct. Consider

*Received by the editors April 4, 2001; accepted for publication (in revised form) by S. Van Huffel June 28, 2002; published electronically January 23, 2003.

http://www.siam.org/journals/simax/24-3/38627.html

†Department of Mathematics and Computer Science, Lehman College of CUNY, Bronx, NY 10468 (vpan@lehman.cuny.edu).

‡Ph.D. Program in Mathematics, Graduate School of CUNY, New York, NY 10036 (xwang2@gc.cuny.edu).

Toeplitz matrices $(t_{i-j})_{i,j=1}^{m,n}$	Hankel matrices $(h_{i+j})_{i,j=1}^{m,n}$
Vandermonde matrices $\left(t_i^{j-1}\right)_{i,j=1}^{m,n}$	Cauchy matrices $\left(\frac{1}{s_i-t_j}\right)_{i,j=1}^{m,n}$

FIG. 1.1. *Four classes of structured matrices.*

the linear operator

$$L(M) = D(\mathbf{s})M - MD(\mathbf{t}).$$

Suppose that

$$L(M) = \sum_{k=1}^{l} \mathbf{g}_k \mathbf{h}_k^T = GH^T, \qquad G = (\mathbf{g}_1, \ldots, \mathbf{g}_l), \quad H = (\mathbf{h}_1, \ldots, \mathbf{h}_l). \tag{1.1}$$

It is immediately verified that (1.1) has a unique solution

$$M = \sum_{k=1}^{l} D(\mathbf{g}_k)C(\mathbf{s}, \mathbf{t})D(\mathbf{h}_k), \tag{1.2}$$

where $C = C(\mathbf{s}, \mathbf{t}) = (\frac{1}{s_i-t_j})_{i,j=1}^{m,n}$ is the Cauchy matrix of Figure 1.1. Equation (1.2) reduces multiplication of a matrix M by a vector to multiplication of $C(\mathbf{s}, \mathbf{t})$ by l vectors. If $l \ll n$, M is called a *Cauchy-like* matrix. This class covers the important subclasses of *Loewner* and *Pick* matrices [P01, pp. 9, 96].

Similarly, the classes of Toeplitz, Hankel, and Vandermonde matrices are extended, and many other structured matrices are also compressed via (1.1) for appropriate operators L. This enables the performing of computations with the matrices M in terms of their *displacement generators* G, H by using much smaller amounts of computer memory and much less CPU time than with the general matrices as long as

(a) the ranks of the displacements are kept small and
(b) the desired output (e.g., the solution of a linear system) is easily recovered at the end.

Here is a flowchart of [P01] for this approach: COMPRESS, OPERATE, DECOMPRESS.

The COMPRESS stage consists in choosing a short displacement generator for the input matrix M (e.g., [P92], [P93], by computing the SVD of its displacement $L(M) = U\Sigma^2V^T = GH^T$, $G = U\Sigma$, $H = V\Sigma$). Simple rules for operating with displacements at the OPERATE stage can be found in [P01, section 1.5]. They include expressions for displacements of the products, sums, linear combinations, Schur complements, and blocks of structured matrices. They also include algorithms for the recovery of a shorter generator from a longer one. These expressions and algorithms are stated for *symbolic displacement*, based on (1.1), where the operator L and matrix class M are not specified. Thus the rules and algorithms are *unified* over various classes of structured matrices. Application of these rules to some basic computations with structured matrices (such as the computation of short displacement generators for the inverses or for the bases of the null spaces) yields effective unified algorithms, which are superfast, that is, which run in $O(n \log^d n)$ time for $d \leq 3$, versus the orders of n^3 time in Gaussian elimination and n^2 time in some fast algorithms.

Furthermore, in [P90], the *displacement transformation* was proposed as a means of extending any successful algorithm available for one class of structured matrices to other classes, and sample transformations among the four classes of matrices of Hankel, Toeplitz, Vandermonde, and Cauchy types were displayed. This approach was pushed forward extensively, yielding effective practical algorithms [H95], [GKO95], [KO96], [G98], [G98a]. On the other hand, the DECOMPRESS stage never enjoyed the systematic treatment it deserves, and thus the entire approach hinged on a few ad hoc formulae scattered in [KKM79], [AG91], [GO94], and [BP94]. Particularly underdeveloped was this basic stage for the important applications using rectangular structured matrices (appearing, e.g., in structured least-squares computations) and singular displacement operators (appearing, e.g., in the study of the Nevanlinna–Pick celebrated problems). An important related issue is the estimation of $\|L^{-1}\|$, which is a critical numerical parameter. For example, whenever the solution of a linear system $M\mathbf{x} = \mathbf{b}$ is recovered from the displacement $L(M^{-1})$ computed numerically, the output errors are proportional to $\|L^{-1}\|$. In another example, a structured matrix is inverted numerically by means of Newton's iteration, and the COMPRESS stage is recursively applied in each iterative step [P92], [P01]. The convergence rate of the process and even the convergence itself critically depend on the residual norm $r_i = \|I - MX_i\|$, where X_i is an approximation to M^{-1} implicitly represented by its compressed displacement. The residual norm r_i is proportional to $\|L^{-1}\|$, and so the convergence is faster where $\|L^{-1}\|$ is smaller.

1.2. Our results and organization of the paper. In the present paper, we fill the cited void in a systematic regular way. We specify bilinear expressions of structured matrices via their displacements or, equivalently, invert the linear displacement operators, covering the most popular classes of structured matrices and *almost all known operators* L (see Remark 6.4). We treat the general case of rectangular matrices M and supply general inversion techniques for possibly singular operators L. We first invert them on the orthogonal complement of their null spaces and then extend the inversion to all matrices by using the first or the last row and/or column of M (see Examples 5.1(3), 5.4(2), and 5.6(2) and sections 6.2(ii) and 6.3(ii)). Because of the high importance of the approach, our work should inevitably have substantial practical impact on the computations with structured matrices supplying a solid foundation for them to replace the collection of random ad hoc recipes available so far. Within the limited space of this paper, we point only to some impact on the solution of the Nevanlinna–Pick and Nehari problems in Remarks 5.8 and 6.4, referring the reader to [BGR90], [OP98], [P01], and the bibliographies therein for further information on these problems and the impact. Sections 2–4 cover some simple and/or known auxiliary results. In sections 5 and 6, we derive the desired bilinear expressions. The derivation is elementary and rather straightforward in section 5 (apart from our novel treatment of the case of singular operators) but is more involved in section 6. There we invert the operators associated with a very general class of confluent matrices. The inversion of these operators is a basis for the design of effective algorithms for the confluent tangential Nevanlinna–Pick problem and was a highly important long-standing open issue. The superfast algorithms of [OP98], [OS00], and [P01] for the confluent Nevanlinna–Pick problem as well as their future amendments and improvements rely and must inevitably rely on the inversion of the associated displacement operators. In Remarks 5.9 and 6.4, we comment on the preceding works. In section 7, we briefly comment on the extension of our results to the products and inverses of structured matrices. Finally, in sections 8 and 9, we substantially advance the known results in [P92], [P93], [PRW02], and [PKRC02] on estimating the norms of the inverse displacement operators.

2. Definitions and basic results. Let us begin with some definitions and simple basic results (cf. [P01] for a detailed and systematic exposition of structured matrix computations). We assume computations in an arbitrary field $\mathbb{F}$, which, in particular, covers computations in the fields of complex, real, or rational numbers ($\mathbb{C}$, $\mathbb{R}$, or $\mathbb{Q}$). $M \in \mathbb{F}^{m\times n}$ denotes an $m \times n$ matrix with the entries in the field $\mathbb{F}$.

- $\mathbf{t}^{-1} = (t_i^{-1})_{1\le i\le k}$, where $\mathbf{t} = (t_i)_{1\le i\le k} \in \mathbb{F}^{k\times 1}$. W^T, $\mathbf{v}^T$ are the *transposes* of a matrix W and a vector $\mathbf{v}$, respectively. W^{-T} is the transpose of W^{-1}, that is, $W^{-T} = (W^{-1})^T = (W^T)^{-1}$.
- $(W_1, \dots, W_n)$ is the $1 \times n$ block matrix with the blocks $W_1, \dots, W_n$. $D(\mathbf{v}) = \operatorname{diag}(\mathbf{v})$ is the $n \times n$ *diagonal* matrix, where $\mathbf{v} = (v_i)_{1\le i\le n}$.
- $\mathbf{e}_i$ is the ith coordinate vector, having its ith coordinate 1 and all other coordinates 0, and so $\mathbf{e}_1 = (1, 0, \dots, 0)^T$. $\mathbf{0} = (0, \dots, 0)^T$, $\mathbf{1} = (1, \dots, 1)^T$. $I = I_n = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ is the $n\times n$ *identity matrix.* 0_n is the $n\times n$ *null matrix.* $J = J_n = (\mathbf{e}_n, \dots, \mathbf{e}_1)$ is the $n \times n$ *reflection matrix.*
- $Z_f = \begin{pmatrix} & f \\ I_{n-1} & \end{pmatrix} = (\mathbf{e}_2, \dots, \mathbf{e}_n, f\mathbf{e}_1)$ is the $n \times n$ *unit f-circulant* matrix. $Z = Z_0$ is the $n \times n$ unit lower triangular Toeplitz matrix. For a vector $\mathbf{v} = (v_1, \dots, v_m)^T$, write

$$Z_{f,m,n}(\mathbf{v}) = (z_{i,j})_{\substack{1\le i\le m\\ 1\le j\le n}},$$

$$z_{i,j} = \begin{cases} v_{i-j+1} & \text{if } i \ge j, \\ f^k v_{m-l} & \text{if } j-i-1 = km+l, 0 \le l \le m-1, k \ge 0. \end{cases}$$

  $Z_{f,m,n}(\mathbf{v})$ is the $m \times n$ f-circulant matrix with the first column $\mathbf{v}$. $Z_f(\mathbf{v}) = \sum_{i=1}^m v_i Z_f^{i-1} = Z_{f,m,m}(\mathbf{v})$. $Z(\mathbf{v}) = Z_0(\mathbf{v})$.
- $V_{m,n}(\mathbf{x}) = (x_i^{j-1})_{\substack{1\le i\le m\\ 1\le j\le n}}$ is the $m \times n$ Vandermonde matrix defined by its second column vector $\mathbf{x} = (x_i)_{1\le i\le m}$. $V(\mathbf{x}) = V_{m,m}(\mathbf{x})$.
- ω_n is a primitive nth root of 1 (that is, $\omega_n^n = 1$, $\omega_n^s \ne 1$, $s = 1, 2, \dots, n-1$); e.g., $\omega_n = e^{2\pi\sqrt{-1}/n}$ in the complex number field $\mathbb{C}$. $\mathbf{w}_n = (\omega_n^{i-1})_{1\le i\le n}$ is the vector of all nth roots of 1.
- $\Omega_n = (\omega_n^{(i-1)(j-1)})_{\substack{1\le i\le n\\ 1\le j\le n}}$ is the $n\times n$ matrix of the *discrete Fourier transform (DFT).* The DFT of a vector $\mathbf{v}$ of dimension n is the vector $DFT(\mathbf{v}) = \Omega_n\mathbf{v}$.
- $\lceil x\rceil$ and $\lfloor x\rfloor$ denote two integers closest to a real x such that $\lfloor x\rfloor \le x \le \lceil x\rceil$.
- For any matrix A, let $\sigma_i(A)$ be its ith largest singular value if $i \le \operatorname{rank}(A)$, and let $\sigma_i(A) = 0$ if $i > \operatorname{rank}(A)$. For any $n\times n$ matrix A, let $\operatorname{spectrum}(A) = \{\lambda_1(A), \dots, \lambda_n(A)\}$ be the set of all of the eigenvalues of A. (We repeat m times any eigenvalue having algebraic multiplicity m.)

The following simple results can be easily verified.

THEOREM 2.1. *$J^2 = I$, $J\mathbf{v} = (v_{n+1-i})_{1\le i\le n}$, $JD(\mathbf{v})J = D(J\mathbf{v})$ for any vector $\mathbf{v} = (v_i)_{1\le i\le n}$.*

THEOREM 2.2. *For the $n \times n$ matrix Z_e and any scalar e, we have $Z_e^n = eI$, $Z_e^T = JZ_eJ$. For $e \ne 0$, we have $Z_e^{-1} = Z_{1/e}^T$.*

THEOREM 2.3 (see [CPW74]). *For the $n \times n$ matrix Z_e and scalar $e \ne 0$, we have $Z_e = V^{-1}\operatorname{diag}(\omega_n^i)_{i=0}^{n-1} V$, where $V = V(\mathbf{t}) = (\omega_n^{ij})_{i,j=0}^{n-1}\operatorname{diag}(t^i)_{i=0}^{n-1}$ and t is a primitive nth root of e.*

THEOREM 2.4. *$O(n \log n)$ flops are sufficient to multiply by a vector the matrices V and V^{-1} of Theorem 2.3 as well as the $n \times n$ Vandermonde matrix $V((\mathbf{t}+s\mathbf{1})^{-1})^T$ for any scalar s.*

Proof. Let $\mathbf{v} = (v_i)_i$, $\mathbf{u} = (u_k)_k$, $v(x) = \sum_{1 \le i \le n} v_i x^{i-1}$. Then the vectors $V(\mathbf{u})\mathbf{v} = (v(u_k))_{k=1}^n$ for $\mathbf{u} = \mathbf{t}$, $\mathbf{u} = \mathbf{t}^{-1}$, and $\mathbf{u} = (\mathbf{t} + s\mathbf{1})^{-1}$ can be computed in $O(n \log n)$ flops [P01, p. 29], [PSD70]. □

The rest of this section is the basis for estimating the operator norms $\|L^{-1}\|$ in sections 8–9.

DEFINITION 2.5.

- *$\sigma_1(A) \ge \cdots \ge \sigma_r(A) > 0$ are all of the singular values of a matrix A, $r =$ rank(A).*
- *$\lambda_1(A), \dots, \lambda_n(A)$ are all of the eigenvalues of an $n \times n$ matrix A with $|\lambda_1(A)| \ge \cdots \ge |\lambda_n(A)|$.*
- *$m \mid n$ means that an integer m divides an integer n; $m \nmid n$ means the opposite.*
- *$\mathrm{lcm}(m, n)$ is the* least common multiple *of two positive integers m and n.*

DEFINITION 2.6 (norms of vectors, operators, and matrices).

- *For a vector $\mathbf{x} = (x_i)$, we define its (Euclidean) norm by $\|\mathbf{x}\| = (\sum_i |x_i|^2)^{1/2}$.*
- *For a linear operator L on a normed vector space V, we define the 2-norm by $\|L\| = \sup_{\mathbf{0} \ne \mathbf{x} \in V} \frac{\|L(\mathbf{x})\|}{\|\mathbf{x}\|}$.*
- *Viewing an $m \times n$ matrix $A = (a_{ij}) = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ as an mn-dimensional vector $\overrightarrow{A} = (\mathbf{a}_1^T, \dots, \mathbf{a}_n^T)^T$, we define its Frobenius norm $\|A\|_F = \|\overrightarrow{A}\| = (\sum_{i,j} |a_{i,j}|^2)^{1/2}$. Alternatively, we may view the matrix as a linear operator $L_A : \mathbf{x} \mapsto A\mathbf{x}$ (or $R_A : \mathbf{x} \mapsto \mathbf{x}^T A$) and define its operator norm (2-norm) $\|A\| = \|L_A\|$ (or $\|R_A\|$).*
- *Given a linear operator L on the $m \times n$ matrix space, we restrict L on the matrices A having rank of at most r and define $\|L\|_r = \sup_{\mathrm{rank}(A) \le r} \frac{\|L(A)\|}{\|A\|}$.*

Here are some simple results.

THEOREM 2.7. *For $r \ge 1$ and a linear operator L on the matrix space, we have $\|L\|_{r-1} \le \|L\|_r \le r\|L\|_1$.*

Proof. (1) $\|L\|_{r-1} \le \|L\|_r$ is obvious. (2) For any matrix A, we know from its SVD that $A = A_1 + \cdots + A_r$, where $r = \mathrm{rank}(A)$, each A_i is a rank-1 matrix, and $\|A_i\| = \sigma_i(A)$. Then $\|L(A)\| \le \|L(A_1)\| + \cdots + \|L(A_r)\| \le \|L\|_1(\|A_1\| + \cdots + \|A_r\|) \le r\|L\|_1\|A\|$. □

THEOREM 2.8. *For any matrix A, $\|A\| = \sigma_1(A)$, $\|A\|_F = (\sum_i \sigma_i(A)^2)^{1/2}$. Therefore, $\|A\|_F/\sqrt{\mathrm{rank}(A)} \le \|A\| \le \|A\|_F$. Furthermore, $\|A\| \ge |\lambda_1(A)|$ if A is a square matrix.*

Example 2.9. For $m \times m$ matrix Z_e,

$$\|Z_e^k\| = \begin{cases} |e|^{k/m} & \text{if } m \mid k, \\ |e|^{\lfloor k/m \rfloor} \max(1, |e|) & \text{if } m \nmid k. \end{cases}$$

3. Linear operators of Sylvester and Stein types. Let us associate structured matrices with *displacement linear operators of Sylvester type*, $L = \nabla_{A,B}$,

$$\nabla_{A,B}(M) = AM - MB, \tag{3.1}$$

and *Stein type*, $L = \Delta_{A,B}$,

$$\Delta_{A,B}(M) = M - AMB, \tag{3.2}$$

where A, B are two fixed *operator matrices*. The image $L(M)$ is called the *L-displacement* of a matrix M or just its *displacement*. We consider the general case of rectangular matrices A, B, and M.

Operators of both types are useful; the ∇ operators may be more effective at the OPERATE stage; the Δ operators may be more effective at the DECOMPRESS stage.

THEOREM 3.1. $\nabla_{A,B} = A\Delta_{A^{-1},B}$ *if the matrix A is nonsingular, and* $\nabla_{A,B} = -\Delta_{A,B^{-1}}B$ *if the matrix B is nonsingular.*

DEFINITION 3.2. *A linear operator L is* nonsingular *if the equation $L(M) = 0$ implies that $M = 0$.*

THEOREM 3.3 (cf. [P01, Theorem 4.3.2]). $\nabla_{A,B}$ *is nonsingular if and only if* $\lambda_i(A) \neq \lambda_j(B)$ *for all pairs of eigenvalues* $(\lambda_i(A), \lambda_j(B))$*;* $\Delta_{A,B}$ *is nonsingular if and only if* $\lambda_i(A)\lambda_j(B) \neq 1$ *for all pairs* $(\lambda_i(A), \lambda_j(B))$.

COROLLARY 3.4. *If the operator $\nabla_{A,B}$ is nonsingular, then A or B is nonsingular.*

Let us relate basic operations with matrices to operations with their displacements (cf. [P01]).

THEOREM 3.5. *For a nonsingular matrix M and a pair of operator matrices A and B, we have*

$$\nabla_{B,A}(M^{-1}) = -M^{-1}\nabla_{A,B}(M)M^{-1}.$$

Furthermore,

$$\Delta_{B,A}(M^{-1}) = BM^{-1}\Delta_{A,B}(M)B^{-1}M^{-1} \quad \text{if } B \text{ is nonsingular},$$
$$\Delta_{B,A}(M^{-1}) = M^{-1}A^{-1}\Delta_{A,B}(M)M^{-1}A \quad \text{if } A \text{ is nonsingular.}$$

THEOREM 3.6. *For any triple of matrices (A, B, M) of compatible sizes, we have*

$$\nabla_{A,B}(M^T) = -\nabla_{B^T,A^T}(M)^T, \qquad \Delta_{A,B}(M^T) = \Delta_{B^T,A^T}(M)^T.$$

THEOREM 3.7. *Let $\hat{A} = VAV^{-1}$, $\hat{B} = W^{-1}BW$ for some nonsingular matrices V and W. Then*

$$\nabla_{\hat{A},\hat{B}}(VMW) = V\nabla_{A,B}(M)W, \qquad \Delta_{\hat{A},\hat{B}}(VMW) = V\Delta_{A,B}(M)W.$$

4. Inversion of displacement operators. Our explicit expressions for a matrix M via its displacement rely on the next simple theorem.

THEOREM 4.1 (see [GO92], [W93], [PRW02]). *For any triple of matrices A, B, and M and for all natural numbers k, we have* $M = A^kMB^k + \sum_{i=0}^{k-1} A^i\Delta_{A,B}(M)B^i$.

By combining Theorems 3.1 and 4.1, we obtain the next result.

COROLLARY 4.2. *Given a triple of matrices A, B, and M and a natural number k, we have* $M = A^{-k}MB^k + \sum_{i=0}^{k-1} A^{-i-1}\nabla_{A,B}(M)B^i$ *if A is nonsingular and* $M = A^kMB^{-k} - \sum_{i=0}^{k-1} A^i\nabla_{A,B}(M)B^{-i-1}$ *if B is nonsingular.*

Theorem 4.1 and Corollary 4.2 enable simple expressions of a matrix M via its displacements $\Delta_{A,B}(M)$ and $\nabla_{A,B}(M)$, respectively, provided that $A^k = cI$ and/or $B^k = cI$ for a scalar c.

COROLLARY 4.3. *Under the assumptions of Theorem* 4.1, *we have* $M(I-aB^k) = \sum_{i=0}^{k-1} A^i\Delta_{A,B}(M)B^i$ *if* $A^k = aI$ *and* $(I-bA^k)M = \sum_{i=0}^{k-1} A^i\Delta_{A,B}(M)B^i$ *if* $B^k = bI$.

Definition 4.4. *Hereafter, $W = (\mathbf{w}_1, \ldots, \mathbf{w}_s)$ is a matrix with columns given by vectors $\mathbf{w}_1, \ldots, \mathbf{w}_s$. Suppose, for an $m \times n$ matrix M and a linear operator L, that we have*

$$L(M) = GH^T = \sum_{k=1}^{l} \mathbf{g}_k \mathbf{h}_k^T, \tag{4.1}$$

$G = (\mathbf{g}_1, \ldots, \mathbf{g}_l)$, $H = (\mathbf{h}_1, \ldots, \mathbf{h}_l)$, and l is "small" ($l = O(1)$ or $l \ll \min(m, n)$). Then M is called a structured matrix with an L-generator (G, H) of length l.

Definition 4.5. *For natural numbers m and n, an $m \times m$ matrix P, and an m-dimensional column vector $\mathbf{v}$, we define the $m \times n$* Krylov matrix $K_{m,n}(P, \mathbf{v}) = (\mathbf{v}, P\mathbf{v}, \ldots, P^{n-1}\mathbf{v})$.

Remark 4.6. The Krylov matrix $K_{m,n}(P, \mathbf{v})$ turns into

(a) the $m \times n$ f-circulant matrix $Z_{f,m,n}(\mathbf{v})$ when $P = Z_f$,

(b) $JZ_{f,m,n}(J\mathbf{v})$ when $P = Z_f^T$, and

(c) the product $D(\mathbf{v})V_{m,n}(P\mathbf{1})$ of the diagonal matrix $D(\mathbf{v})$ and the Vandermonde matrix $V_{m,n}(P\mathbf{1})$ when P is a diagonal matrix; $D(\mathbf{v}) = I_m$ when $\mathbf{v} = \mathbf{1}$.

Theorem 4.1 and Corollary 4.2 imply the next results.

Theorem 4.7. *For an operator $L = \Delta_{A,B}$, an $m \times n$ matrix M satisfying* (4.1), *and all natural numbers k, we have $M = A^k M B^k + \sum_{j=1}^{l} K_{m,k}(A, \mathbf{g}_j) K_{n,k}(B^T, \mathbf{h}_j)^T$.*

Theorem 4.8. *For an operator $L = \nabla_{A,B}$, an $m \times n$ matrix M satisfying* (4.1), *and all natural numbers k, we have $M = A^{-k-1} M B^k + A^{-1} \sum_{j=1}^{l} K_{m,k}(A^{-1}, \mathbf{g}_j) \cdot K_{n,k}(B^T, \mathbf{h}_j)^T$ if A is nonsingular and $M = A^k M B^{-k-1} - \sum_{j=1}^{l} K_{m,k}(A, \mathbf{g}_j) \cdot K_{n,k}(B^{-T}, \mathbf{h}_j)^T B^{-1}$ if B is nonsingular.*

5. Bilinear expressions via generators for fundamental matrix structures. In this section, we extend (1.2) to express a matrix $M \in \mathbb{F}^{m \times n}$ via its displacement $L(M)$, where $L = \Delta_{A,B}$ and $L = \nabla_{A,B}$ for some commonly used operator matrices $A \in \mathbb{F}^{m \times m}$, $B \in \mathbb{F}^{n \times n}$. Let $L(M) = GH^T = \sum_{1 \le j \le l} \mathbf{g}_j \mathbf{h}_j^T$, $G = (g_{i,j})_{1 \le i \le m, 1 \le j \le l}$, $H = (h_{i,j})_{1 \le i \le n, 1 \le j \le l}$.

Example 5.1. The Stein-type operators $L = \Delta_{Z_e, Z_f}$ are associated with Hankel-like matrices. Note that $Z_e^m = eI_m$, $Z_f^n = fI_n$. We begin with the special case in which $e = 0$ (the same tool applies to $f = 0$), then supply the expressions in the general nonsingular case, and finally cover all choices of e and f.

1. $e = 0$. Apply Theorem 4.7, take into account Remark 4.6, and obtain that

$$M = \sum_{j=1}^{l} K_{m,m}(Z, \mathbf{g}_j) K_{n,m}(Z_f^T, \mathbf{h}_j)^T = \sum_{j=1}^{l} Z(\mathbf{g}_j) Z_{f,n,m}(J\mathbf{h}_j)^T J.$$

2. Let the operator Δ_{Z_e, Z_f} be nonsingular. Then the matrix $I_n - eZ_f^m$ is nonsingular due to Definition 3.2. Apply Theorem 4.7 for $k = m$, then recall Remark 4.6 and obtain that

$$\begin{aligned} M &= \sum_{j=1}^{l} K_{m,m}(Z_e, \mathbf{g}_j) K_{n,m}(Z_f^T, \mathbf{h}_j)^T (I_n - eZ_f^m)^{-1} \\ &= \sum_{j=1}^{l} Z_e(\mathbf{g}_j) Z_{f,n,m}(J\mathbf{h}_j)^T J (I_n - eZ_f^m)^{-1}. \end{aligned}$$

3. If the operator Δ_{Z_e,Z_f} is singular, then we cannot recover the matrix M solely from its displacement. We need extra information about M. We begin with the matrix equation

$$\Delta_{Z_e,Z}(M) = \Delta_{Z_e,Z_f}(M) + Z_e M \begin{pmatrix} & f \\ 0_{n-1} & \end{pmatrix} = GH^T + fZ_eM\mathbf{e}_1\mathbf{e}_n^T,$$

apply Theorem 4.7 to the operator $\Delta_{Z_e,Z}$, recall Remark 4.6, and deduce that

$$\begin{aligned} M &= \sum_{j=1}^{l} K_{m,n}(Z_e,\mathbf{g}_j)K_{n,n}(Z^T,\mathbf{h}_j)^T + fK_{m,n}(Z_e,Z_eM\mathbf{e}_1)K_{n,n}(Z^T,\mathbf{e}_n)^T \\ &= \sum_{j=1}^{l} Z_{e,m,n}(\mathbf{g}_j)Z(J\mathbf{h}_j)^T J + fZ_{e,m,n}(JZ_eM\mathbf{e}_1)J, \end{aligned}$$

where $(M_{i,1})_{1\le i\le m} = M\mathbf{e}_1$ is the first column of M.

Remark 5.2. In case 2, we involve the matrix $(I_n - eZ_f^m)^{-1} = V_f^{-1}\operatorname{diag}(\frac{1}{1-et_i^m})_{1\le i\le n}V_f$, where $V_f = (t_i^{j-1})_{1\le i\le n, 1\le j\le n}$, $t_1,\dots,t_n$ are all the nth roots of f.

Example 5.3. The case of the Stein-type operators $L = \Delta_{Z_e,Z_f^T}$, $L = \Delta_{Z_e^T,Z_f}$, and $L = \Delta_{Z_e^T,Z_f^T}$, associated with Toeplitz/Hankel-like matrices, can be reduced to Example 5.1 based on Theorem 2.2.

Example 5.4. The Sylvester-type operators $L = \nabla_{Z_e,Z_f}$ are associated with Toeplitz-like matrices.

1. If $e \ne 0$, then, by Theorems 3.1 and 2.2, we have

$$\Delta_{Z_{1/e}^T,Z_f}(M) = Z_{1/e}^T\nabla_{Z_e,Z_f}(M) = (Z_{1/e}^T G)H^T.$$

The latter equation immediately reduces the problem to the case of the Stein-type operators $\Delta_{Z_e^T,Z_f}$ of Example 5.3. The same tool applies to the case $f \ne 0$.

2. If $e = f = 0$, then we have (cf. [BP93], [BP94] for the proof)

$$M = J\sum_{j=1}^{l} Z(JZ^T\mathbf{g}_j)Z_{0,n,m}(J\mathbf{h}_j)^T J + JZ_{0,n,m}(JM^T\mathbf{e}_m)^T J.$$

Example 5.5. Similarly to Example 5.4, we express Hankel-like and Toeplitz-like matrices M associated with the Sylvester-type operators $L = \nabla_{Z_e,Z_f^T}$, $L = \nabla_{Z_e^T,Z_f}$, and $L = \nabla_{Z_e^T,Z_f^T}$.

Example 5.6. The Stein-type operators $L = \Delta_{D(\mathbf{v}),Z_f^T}$ are associated with the matrix structure of Vandermonde type.

1. If the operator $\Delta_{D(\mathbf{v}),Z_f^T}$ is nonsingular, then the matrix $I_m - fD(\mathbf{v})^n$ is nonsingular, and it follows from Theorem 4.7 and Remark 4.6 that

$$\begin{aligned} M &= (I_m - fD(\mathbf{v})^n)^{-1}\sum_{j=1}^{l} K_{m,n}(D(\mathbf{v}),\mathbf{g}_j)K_{n,n}(Z_f,\mathbf{h}_j)^T \\ &= \operatorname{diag}\left(\frac{1}{1-fv_i^n}\right)_{1\le i\le m}\sum_{j=1}^{l} D(\mathbf{g}_j)V_{m,n}(\mathbf{v})Z_f(\mathbf{h}_j)^T. \end{aligned}$$

2. To relax the nonsingularity assumption, observe that

$$\Delta_{D(\mathbf{v}),Z^T}(M) = \Delta_{D(\mathbf{v}),Z_f^T}(M) + D(\mathbf{v})M \begin{pmatrix} & 0_{n-1} \\ f & \end{pmatrix} = GH^T + fD(\mathbf{v})M\mathbf{e}_n\mathbf{e}_1^T.$$

Therefore, we deduce from Theorem 4.7 and Remark 4.6 that

$$M = \sum_{j=1}^{l} K_{m,n}(D(\mathbf{v}), \mathbf{g}_j)K_{n,n}(Z, \mathbf{h}_j)^T + fD(\mathbf{v})K_{m,n}(D(\mathbf{v}), M\mathbf{e}_n)$$
$$= \sum_{j=1}^{l} D(\mathbf{g}_j)V_{m,n}(\mathbf{v})Z(\mathbf{h}_j)^T + fD(\mathbf{v})D(M\mathbf{e}_n)V_{m,n}(\mathbf{v}).$$

Example 5.7. The Stein-type operators $L = \Delta_{D(\mathbf{v}),Z_f}$, $L = \Delta_{Z_e,D(\mathbf{v})}$, and $L = \Delta_{Z_e^T,D(\mathbf{v})}$ and the Sylvester-type operators $L = \nabla_{D(\mathbf{v}),Z_f}$, $L = \nabla_{D(\mathbf{v}),Z_f^T}$, $L = \nabla_{Z_e,D(\mathbf{v})}$, and $L = \nabla_{Z_e^T,D(\mathbf{v})}$ are also associated with structured matrices of Vandermonde type. Theorems 2.2, 3.1, 3.5, and 3.6 enable us to reduce the problems to the case of the operator $\Delta_{D(\mathbf{v}),Z_f^T}$ of Example 5.6.

Remark 5.8. In [P01], a variant of the displacement transformation techniques of [P90] reduced the solution of the tangential Nevanlinna–Pick and matrix Nehari problems to computations with *skew-Hankel-like* matrices associated with the operator $\nabla_{Z_e^T,-Z_f}$ for a pair of scalars e and f, $e \neq f$. Such matrices have no other natural derivation; their hidden structure appears to be hard to exploit, but the inversion formulae of Example 5.3 are immediately extended. More generally, *all presented expressions for the inversion of the operators* $\nabla_{A,B}$ *and* $\Delta_{A,B}$ *can be immediately extended to operators* $\nabla_{aA,bB}$ *and* $\Delta_{aA,bB}$ *for any pair of nonzero scalars* a *and* b. With this extension, the unified divide-and-conquer algorithms of [P99], [P00], and [P01] solve the cited important problems by using $O(N \log^2 N)$ flops (N being the input size) versus $O(N \log^3 N)$ announced in the proceedings paper [OP98] and the order of N^2 in the preceding works. Important superfast algorithms using $O(N \log^d N)$ flops are known for other rational interpolation problems [BVB97], [CJL96], [CJL96a], [KVB99], but different techniques are required for the Nevanlinna–Pick and Nehari problems.

Remark 5.9. The bilinear expressions of this section have been known for square matrices M and the following operators L: $\Delta_{Z^T,Z}$ and Δ_{Z,Z^T} [KKM79], $\nabla_{D(\mathbf{v}),Z^T}$ [GKKL87], $\nabla_{Z_1,Z_{-1}}$ [AG91], $\nabla_{Z,Z}, \nabla_{Z^T,Z^T}, \Delta_{Z,Z}$, and Δ_{Z^T,Z^T} [BP93], [BP94], and $\Delta_{Z_e,Z_{1/e}^T}$ and $\Delta_{D(\mathbf{v}),Z_{1/e}^T}$ for $e \neq 0$ [GO94]. Some other expressions of this section have been announced in [P01] (see the notes to section 4.4 therein) although only for square matrices M and nonsingular operators L.

6. Bilinear expressions via generators for confluent matrices.

Example 6.1. More general operator matrices $A = \lambda I_m + Z_e$ and $B = \mu I_n + Z_f$ (see Remark 6.3) cover the confluent matrices involved in the tangential confluent Nevanlinna–Pick problem [BGR90], [GO94a], [OS00], [P01]. Then $L = \nabla_{A,B} = \nabla_{(\lambda-\mu)I_m+Z_e,Z_f} = \nabla_{Z_e,(\mu-\lambda)I_n+Z_f}$, and we partition our study depending on the values e, f, and $\lambda - \mu$.

6.1. Case 1. $e = f = 0$.

Assume that $\lambda \neq \mu$. Otherwise, see part 2 of Example 5.4. Combine Corollary 4.2, Remark 4.6, and the equation

$$\nabla_{A,B}(M) = ((\lambda - \mu)I_m + Z)M - MZ = GH^T$$

to obtain

$$\begin{aligned} M &= \sum_{j=0}^{n-1} ((\lambda-\mu)I_m + Z)^{-j-1} G H^T Z^j = \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} \binom{-j-1}{i} (\lambda-\mu)^{-j-1-i} Z^i G H^T Z^j \\ &= \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} \frac{(-1)^i (i+j)!}{i!(\lambda-\mu)^{i+j+1} j!} Z^i G H^T Z^j = \sum_{k=1}^{l} K_{m,m}(Z,\mathbf{g}_k)\Theta_0(\lambda-\mu) K_{n,n}(Z^T,\mathbf{h}_k)^T \\ &= \sum_{k=1}^{l} Z(\mathbf{g}_k)\Theta_0(\lambda-\mu) Z(J\mathbf{h}_k)^T J, \end{aligned}$$

$$\begin{aligned} \Theta_0(s) &= \left(\frac{(-1)^{i-1}(i+j-2)!}{(i-1)! s^{i+j-1} (j-1)!}\right)_{i,j=1}^{m,n} = \operatorname{diag}\left(\frac{(-1)^{i-1}}{(i-1)!}\right)_{i=1}^{m} H \operatorname{diag}\left(\frac{1}{(j-1)!}\right)_{j=1}^{n}, \\ H &= \left(\frac{(i+j-2)!}{s^{i+j-1}}\right)_{\substack{1\le i\le m\\ 1\le j\le n}}, \end{aligned}$$

which is a Hankel matrix.

6.2. Case 2. $e \neq 0$, $f = 0$ (similarly if $e = 0$, $f \neq 0$).

(i) $(\mu-\lambda)^m \neq e$. Write $V = V(\mathbf{t})$, and combine the equation

$$\nabla_{A,B}(M) = ((\lambda-\mu)I_m + Z_e)M - MZ = GH^T$$

with Theorem 2.3, Corollary 4.2, and Remark 4.6 to obtain

$$\begin{aligned} M &= \sum_{j=0}^{n-1} ((\lambda-\mu)I_m + Z_e)^{-j-1} G H^T Z^j = \sum_{j=0}^{n-1} V^{-1}((\lambda-\mu)I_m + D)^{-j-1} V G H^T Z^j \\ &= \sum_{k=1}^{l} V^{-1}((\lambda-\mu)I_m + D)^{-1} K_{m,n}(((\lambda-\mu)I_m + D)^{-1}, V\mathbf{g}_k) K_{n,n}(Z^T,\mathbf{h}_k)^T \\ &= \sum_{k=1}^{l} V^{-1} \operatorname{diag}(V\mathbf{g}_k) \left((\tfrac{1}{\lambda-\mu+t_i})^j\right)_{\substack{1\le i\le m\\ 1\le j\le n}} K_{n,n}(Z^T,\mathbf{h}_k)^T \\ &= \sum_{k=1}^{l} \left(V^{-1} \sum_{r=1}^{m} g_{r,k} D^{r-1}\right) \left((\tfrac{1}{\lambda-\mu+t_i})^j\right)_{\substack{1\le i\le m\\ 1\le j\le n}} K_{n,n}(Z^T,\mathbf{h}_k)^T \\ &= \sum_{k=1}^{l} \left(\sum_{r=1}^{m} g_{r,k} Z_e^{r-1} V^{-1}\right) \left((\tfrac{1}{\lambda-\mu+t_i})^j\right)_{\substack{1\le i\le m\\ 1\le j\le n}} K_{n,n}(Z^T,\mathbf{h}_k)^T \\ &= \sum_{k=1}^{l} K_{m,m}(Z_e,\mathbf{g}_k) V^{-1} \left((\tfrac{1}{\lambda-\mu+t_i})^j\right)_{\substack{1\le i\le m\\ 1\le j\le n}} K_{n,n}(Z^T,\mathbf{h}_k)^T \\ &= \sum_{k=1}^{l} K_{m,m}(Z_e,\mathbf{g}_k)\Theta_1(\lambda-\mu) K_{n,n}(Z^T,\mathbf{h}_k)^T = \sum_{k=1}^{l} Z_e(\mathbf{g}_k)\Theta_1(\lambda-\mu) Z(J\mathbf{h}_k)^T J. \end{aligned}$$

Here $\Theta_1(s) = V^{-1}((\frac{1}{s+t_i})^j)_{\substack{1\le i\le m\\ 1\le j\le n}} = \frac{1}{m} V(\mathbf{t}^{-1})^T V_{m,n}((\mathbf{t}+s\mathbf{1})^{-1}) D(\mathbf{t}+s\mathbf{1})^{-1}$, $\mathbf{t} = (t_i)_{1\le i\le m}$ is the vector of all the mth roots of e.

(ii) $(\mu-\lambda)^m = e$, so the operator L is singular. Note that

$$((\lambda-\mu)I_m + Z)M - MZ = \nabla_{A,B}(M) + (Z - Z_e)M = GH^T - e\mathbf{e}_1\mathbf{e}_m^T M.$$

Proceed similarly to Case 1 to obtain that

$$\begin{aligned} M &= \sum_{k=1}^{l} K_{m,m}(Z, \mathbf{g}_k)\Theta_0(\lambda-\mu)K_{n,n}(Z^T, \mathbf{h}_k)^T \; - \; e\Theta_0(\lambda-\mu)K_{n,n}(Z^T, M^T\mathbf{e}_m)^T \\ &= \sum_{k=1}^{l} Z(\mathbf{g}_k)\Theta_0(\lambda-\mu)Z(J\mathbf{h}_k)^T J \; - \; e\Theta_0(\lambda-\mu)Z(JM^T\mathbf{e}_m)^T J. \end{aligned}$$

6.3. Case 3. $ef \neq 0$.

(i) The operator L is nonsingular, so both matrices $I - f((\lambda-\mu)I_m + Z_e)^n$ and $I - e((\mu-\lambda)I_n + Z_f)^m$ are nonsingular. Apply Corollary 4.2 for $k = m$ and obtain that

$$M = Me((\mu-\lambda)I_n + Z_f)^m + \sum_{i=0}^{m-1} Z_e^{-i-1}GH^T((\mu-\lambda)I_n + Z_f)^i.$$

Therefore, we have

$$\begin{aligned} M &= \left(\sum_{i=0}^{m-1} Z_e^{-i-1}GH^T((\mu-\lambda)I_n + Z_f)^i\right)(I_n - e((\mu-\lambda)I_n + Z_f)^m)^{-1} \\ &= \left(\sum_{i=0}^{m-1} Z_e^{-i-1}GH^T \sum_{j=0}^{n-1}\binom{i}{j}(\mu-\lambda)^{i-j}Z_f^j\right)(I_n - e((\mu-\lambda)I_n + Z_f)^m)^{-1} \\ &= \left(\sum_{i=1}^{m}\sum_{j=1}^{n}\binom{i-1}{j-1}(\mu-\lambda)^{i-j}Z_e^{-i}GH^T Z_f^{j-1}\right)(I_n - e((\mu-\lambda)I_n + Z_f)^m)^{-1} \\ &= \left(\sum_{k=1}^{l} K_{m,m}(Z_e^{-1}, Z_e^{-1}\mathbf{g}_k)\Theta_2(\mu-\lambda)K_{n,m}(Z_f^T, \mathbf{h}_k)^T\right)(I_n - e((\mu-\lambda)I_n + Z_f)^m)^{-1}, \end{aligned}$$

where $\Theta_2(s) = (\frac{(i-1)!s^{i-j}}{(j-1)!(i-j)!})_{\substack{1\le i\le m\\ 1\le j\le i}}$ is an $m \times m$ lower triangular matrix, $\Theta_2(s) = \operatorname{diag}((i-1)!)_{1\le i\le m}(\frac{s^{i-j}}{(i-j)!})_{\substack{1\le i\le m\\ 1\le j\le i}}\operatorname{diag}(\frac{1}{(j-1)!})_{1\le j\le m}$. Recall the equation $Z_e^{-1} = Z_{1/e}^T$ of Theorem 2.2, recall Remark 4.6, and rewrite this expression as follows:

$$M = \left(\sum_{k=1}^{l} JZ_{1/e}(JZ_e^{-1}\mathbf{g}_k)\Theta_3(\mu-\lambda)Z_{f,n,m}(J\mathbf{h}_k)^T J\right)(I_n - e((\mu-\lambda)I_n + Z_f)^m)^{-1}.$$

(ii) The operator L is singular. For any 4-tuple (λ, μ, e, f), apply the equation

$$(\lambda I_m + Z_e)M - M(\mu I_n + Z) = \nabla_{A,B}(M) + M(Z_f - Z) = GH^T + fM\mathbf{e}_1\mathbf{e}_n^T,$$

where $Z^n = 0$, and, as in Case 2, deduce from Theorem 4.7 and Remark 4.6 that

$$\begin{aligned} M &= \sum_{k=1}^{l} K_{m,m}(Z_e, \mathbf{g}_k)\Theta_1(\lambda-\mu)K_{n,n}(Z^T, \mathbf{h}_k)^T + fK_{m,m}(Z_e, M\mathbf{e}_1)\Theta_1(\lambda-\mu)J \\ &= \sum_{k=1}^{l} Z_e(\mathbf{g}_k)\Theta_1(\lambda-\mu)Z(J\mathbf{h}_k)^T J + fZ_e(M\mathbf{e}_1)\Theta_1(\lambda-\mu)J. \end{aligned}$$

6.4. Further remarks.

Remark 6.2. For $ef \neq 0$, we have (cf. Theorem 2.3)

$$(I_m - f((\lambda-\mu)I_m + Z_e)^{-n})^{-1} = V_e^{-1}\,\mathrm{diag}\left(\tfrac{(\lambda-\mu+s_i)^n}{(\lambda-\mu+s_i)^n - f}\right)_{1\le i\le m} V_e,$$
$$(I_n - e((\mu-\lambda)I_n + Z_f)^{-m})^{-1} = V_f^{-1}\,\mathrm{diag}\left(\tfrac{(\mu-\lambda+t_i)^m}{(\mu-\lambda+t_i)^m - e}\right)_{1\le i\le n} V_f,$$

where $V_e = (s_i^{j-1})_{\substack{1\le i\le m\\1\le j\le m}}$, $V_f = (t_i^{j-1})_{\substack{1\le i\le n\\1\le j\le n}}$, $s_1,\ldots,s_m$, are all mth roots of e, and $t_1,\ldots,t_n$ are all nth roots of f.

Remark 6.3. To invert the Sylvester-type operators $L = \nabla_{\lambda I_m + Z_e, \mu I_n + Z_f^T}$, $L = \nabla_{\lambda I_m + Z_e^T, \mu I_n + Z_f}$, and $L = \nabla_{\lambda I_m + Z_e^T, \mu I_n + Z_f^T}$, combine Theorem 2.2, Example 6.1, and the equations

$$\nabla_{\lambda I_m + Z_e, \mu I_n + Z_f}(MJ) = \nabla_{\lambda I_m + Z_e, \mu I_n + Z_f^T}(M)J = G(JH)^T,$$
$$\nabla_{\lambda I_m + Z_e, \mu I_n + Z_f}(JM) = J\nabla_{\lambda I_m + Z_e^T, \mu I_n + Z_f}(M) = (JG)H^T,$$
$$\nabla_{\lambda I_m + Z_e, \mu I_n + Z_f}(JMJ) = J\nabla_{\lambda I_m + Z_e^T, \mu I_n + Z_f^T}(M)J = (JG)(JH)^T.$$

Remark 6.4. For a Sylvester-type operator $L = \nabla_{A,B}$ for any pair of A and B, we have $PAP^{-1} = \mathrm{diag}(\lambda_i(A)I_{m_i} + Z)_{1\le i\le p}$, $QBQ^{-1} = \mathrm{diag}(\lambda_j(B)I_{n_j} + Z)_{1\le j\le q}$. Let us express a matrix M via its displacement $L(M) = GH^T$, the matrices P and Q, and the Jordan blocks $A_i = \lambda_i(A)I_{m_i} + Z$, $i = 1,\ldots,p$; $B_j = \lambda_j(B)I_{n_j} + Z$, $j = 1,\ldots,q$, of the operator matrices A and B. (Already for $P = I_m$, $Q = I_n$, this covers the general class of confluent matrices associated with the tangential confluent Nevanlinna–Pick problem [BGR90].) We recover the matrix M from its displacement $L(M) = GH^T$ by applying the following steps:

1. Represent the matrix PMQ^{-1} as a $p\times q$ block matrix with blocks $M_{i,j}$ of size $m_i \times n_j$; represent the matrix PG as a $p \times 1$ block matrix with blocks G_i of size $m_i \times l$; represent the matrix H^TQ^{-1} as a $1\times q$ block matrix with blocks H_j^T of size $l \times n_j$.
2. Replace the matrix equation $\nabla_{A,B}(M) = GH^T$ by the block equations $\nabla_{A_i,B_j}(M_{i,j}) = G_iH_j^T$ for all pairs (i,j), $i = 1,\ldots,p$; $j = 1,\ldots,q$.
3. Express the blocks $M_{i,j}$ from their displacement generators (G_i, H_j) as in Example 6.1.
4. Express the matrix $M = P^{-1}(M_{i,j})_{\substack{1\le i\le p\\1\le j\le q}}Q$.

For $P = I_m$, $Q = I_n$, we arrive at the matrices M defining the tangential confluent Nevanlinna–Pick problem. In this case, extensively studied since [BGR90], stages 1 and 4 are trivialized. In Case 1 of Example 6.1, a distinct expression for M via $L(M)$ is stated in [OS00]. With omitted proofs and restricted to the case of square matrices M, some of our results were announced in [P01] with reference to the present paper (see the notes of section 4.4 therein).

7. Two implications.

(a) The basic structured matrices can be multiplied by vectors in nearly linear time (see [P01]). Our bilinear expressions of structured matrices via their generators enable immediate extension of these algorithms to more general classes of structured matrices. In particular, we multiply the $n \times n$ matrices of Examples 5.1, 5.3–5.5, and 6.1 by a vector by using $O(ln \log n)$ flops. Similarly, we yield the cost bound of $O(ln \log^2 n)$ flops for the $n \times n$ matrices of Examples 5.6 and 5.7.

(b) Theorem 3.5 enables the extension of all of our expressions to the inverse matrix M^{-1} via the products of this matrix with the $2l$ vectors $\mathbf{g}_k$ and $\mathbf{h}_k$, $k = 1, \dots, l$.

8. Lower and upper bounds on the norms of the inverse displacement operators. In this section, we estimate the operator norm $\|L^{-1}\|$ for the operators $L = \Delta_{Z_e,Z_f}$, $L = \nabla_{Z_e,Z_f}$, $L = \Delta_{Z_e,D(\mathbf{v})}$, $L = \nabla_{Z_e,D(\mathbf{v})}$, and $L = \Delta_{D(\mathbf{u}),D(\mathbf{v})}$, $L = \nabla_{D(\mathbf{u}),D(\mathbf{v})}$. All of our proofs and estimates, however, are invariant to interchanging the operator matrices A and B and to transposing any of A and B, so the same estimates are immediately extended to the operators $\Delta_{Z_e^T,Z_f}$, $\nabla_{Z_e^T,Z_f}$, Δ_{Z_e,Z_f^T}, ∇_{Z_e,Z_f^T}, $\Delta_{Z_e^T,Z_f^T}$, $\nabla_{Z_e^T,Z_f^T}$, $\Delta_{Z_e^T,D(\mathbf{v})}$, $\nabla_{Z_e^T,D(\mathbf{v})}$, $\Delta_{D(\mathbf{v}),Z_e}$, $\nabla_{D(\mathbf{v}),Z_e}$, $\Delta_{D(\mathbf{v}),Z_e^T}$, $\nabla_{D(\mathbf{v}),Z_e^T}$, respectively. This covers the operators associated with the matrices of the most popular structures of Toeplitz, Hankel, Vandermonde, and Cauchy types.

THEOREM 8.1. *For any operator norm and any positive integer r, we have*

$$\max_{i,j} |1 - \lambda_i(A)\lambda_j(B)|^{-1} \le \|\Delta_{A,B}^{-1}\|_r \le \sqrt{r}\|(I - B^T \otimes A)^{-1}\|, \tag{8.1}$$

$$\max_{i,j} |\lambda_i(A) - \lambda_j(B)|^{-1} \le \|\nabla_{A,B}^{-1}\|_r \le \sqrt{r}\|(I \otimes A - B^T \otimes I)^{-1}\|, \tag{8.2}$$

where $\otimes$ is the Kronecker product and the lower bounds on $\|\Delta_{A,B}^{-1}\|_r$ and $\|\nabla_{A,B}^{-1}\|_r$ apply to any operator norm.

Proof. (1) Let $\mathbf{g}$ and $\mathbf{h}$ be two eigenvectors of A and B, respectively, such that $A\mathbf{g} = \lambda_i(A)\mathbf{g}$, $B^T\mathbf{h} = \lambda_j(B)\mathbf{h}$. Let $M = \mathbf{g}\mathbf{h}^T$. Then we have $\Delta_{A,B}(M) = (1 - \lambda_i(A)\lambda_j(B))M$, $\nabla_{A,B}(M) = (\lambda_i(A) - \lambda_j(B))M$; that is, M is an eigenvector of $\Delta_{A,B}$ and $\nabla_{A,B}$. This proves the lower bounds in (8.1) and (8.2).

(2) Recall that $\overrightarrow{\Delta_{A,B}(M)} = (I - B^T \otimes A)\overrightarrow{M}$, $\overrightarrow{\nabla_{A,B}(M)} = (I \otimes A - B^T \otimes I)\overrightarrow{M}$. By Theorem 2.8, $\|M\| \le \|M\|_F = \|\overrightarrow{M}\|$, $\|\Delta_{A,B}(M)\| \ge \|\Delta_{A,B}(M)\|_F/\sqrt{r} = \|\overrightarrow{\Delta_{A,B}(M)}\|/\sqrt{r}$, $\|\nabla_{A,B}(M)\| \ge \|\nabla_{A,B}(M)\|_F/\sqrt{r} = \|\overrightarrow{\nabla_{A,B}(M)}\|/\sqrt{r}$. This proves the upper bounds in (8.1) and (8.2). ∎

Our next upper bounds on $\|L^{-1}\|$ rely on the bilinear expressions for M implied by Example 2.9 and Corollary 4.3. Write $\hat{e} = \max(1, |e|)$, $\hat{f} = \max(1, |f|)$.

THEOREM 8.2. *Let $\ell = \operatorname{lcm}(m, n)$. We have*

$$\|\Delta_{Z_e,Z_f}^{-1}\| \le \frac{\hat{e}\hat{f}}{|1 - e^{\ell/m} f^{\ell/n}|} \sum_{k=0}^{\ell-1} |e|^{\lfloor k/m \rfloor} |f|^{\lfloor k/n \rfloor}, \tag{8.3}$$

$$\|\nabla_{Z_e,Z_f}^{-1}\| \le \frac{\hat{e}\hat{f}}{|e^{\ell/m} - f^{\ell/n}|} \sum_{k=0}^{\ell-1} |e|^{\lfloor (\ell-1-k)/m \rfloor} |f|^{\lfloor k/n \rfloor}. \tag{8.4}$$

Proof. (1) Let $\Delta = \Delta_{Z_e,Z_f}(M)$. Then $M = \frac{1}{1-e^{\ell/m}f^{\ell/n}} \sum_{k=0}^{\ell-1} Z_e^k \Delta Z_f^k$. So $\|M\| \le \frac{1}{|1-e^{\ell/m}f^{\ell/n}|} \sum_{k=0}^{\ell-1} \|Z_e^k\|\ \|\Delta\|\ \|Z_f^k\| \le \frac{\|\Delta\|\hat{e}\hat{f}}{|1-e^{\ell/m}f^{\ell/n}|} \sum_{k=0}^{\ell-1} |e|^{\lfloor k/m \rfloor} |f|^{\lfloor k/n \rfloor}$.

(2) Let $\nabla = \nabla_{Z_e,Z_f}(M)$. We may assume $e \ne 0$; otherwise, $\|\nabla_{Z_0,Z_f}^{-1}\| = \lim_{e\to 0} \|\nabla_{Z_e,Z_f}^{-1}\|$. Then $M = \frac{1}{1-f^{\ell/n}/e^{\ell/m}} \sum_{k=0}^{\ell-1} Z_e^{-1-k} \nabla Z_f^k$. So $\|M\| \le \frac{|e|^{\ell/m}}{|e^{\ell/m}-f^{\ell/n}|} \cdot \sum_{k=0}^{\ell-1} \|Z_e^{-1-k}\|\ \|\nabla\|\ \|Z_f^k\| \le \frac{\|\nabla\|\hat{e}\hat{f}}{|e^{\ell/m}-f^{\ell/n}|} \sum_{k=0}^{\ell-1} |e|^{\lfloor (\ell-1-k)/m \rfloor} |f|^{\lfloor k/n \rfloor}$. ∎

Theorems 8.1 (for $A = Z_e$, $B = Z_f$) and 8.2 together imply the following corollary.

COROLLARY 8.3. *Let* $m = n$, *so* $\operatorname{lcm}(m, n) = n$. *Then*

$$|1 - |ef|^{1/n}\omega_{2n}|^{-1} \le \|\Delta^{-1}_{Z_e,Z_f}\| \le \frac{n\hat{e}\hat{f}}{|1 - ef|},$$

$$||e|^{1/n} - |f|^{1/n}\omega_{2n}|^{-1} \le \|\nabla^{-1}_{Z_e,Z_f}\| \le \frac{n\hat{e}\hat{f}}{|e - f|}.$$

Remark 8.4. Comparing the latter lower and upper bounds as $n \to \infty$, we obtain

$$\lim_{n\to\infty} \frac{|1 - ef|}{n\hat{e}\hat{f}\left|1 - |ef|^{1/n}\omega_{2n}\right|} = \begin{cases} \frac{|1-ef|}{\hat{e}\hat{f}\sqrt{\pi^2+\ln^2|ef|}} & \text{if } ef \ne 0, \\ 0 & \text{if } ef = 0, \end{cases}$$

$$\lim_{n\to\infty} \frac{|e - f|}{n\hat{e}\hat{f}\left||e|^{1/n} - |f|^{1/n}\omega_{2n}\right|} = \begin{cases} \frac{|e-f|}{\hat{e}\hat{f}\sqrt{\pi^2+\ln^2|e/f|}} & \text{if } ef \ne 0, \\ 0 & \text{if } ef = 0. \end{cases}$$

Here ln denotes the natural logarithm. Our estimates of Corollary 8.3 are asymptotically tight as $n \to \infty$; that is, the lower and upper bounds differ by a nonzero constant factor, provided $ef \ne 0$ and either $ef \ne 1$ (for Δ_{Z_e,Z_f}) or $e \ne f$ (for ∇_{Z_e,Z_f}).

Let us improve our lower bounds by sampling matrices in the case in which $f = 0$ for any e (similarly, where $e = 0$ for any f). Let $M = Z_e(\mathbf{1})JZ(\mathbf{1})$. Then $\Delta_{Z_e,Z}(M) = \mathbf{1}\,\mathbf{1}^T$. Write $e = x + y\sqrt{-1}$, where x, y are real numbers. We have $\|\Delta_{Z_e,Z}(M)\| = n$ and

$$\begin{aligned} \|M\|^2 &\ge \frac{1}{n}\|\mathbf{1}^T M\|^2 = \frac{1}{n}\sum_{i=1}^{n}\left|in + \frac{i(i-1)}{2}(e-1)\right|^2 \\ &\ge \left(\frac{n^4}{20} - \frac{n^3}{8} + \frac{n^2}{12}\right)(x-1)^2 + \left(\frac{n^4}{4} - \frac{n^3}{3}\right)(x-1) + \frac{n^4}{3} \\ &\ge \frac{n^2}{48}\left(n^2 - \frac{2}{5}n + 25\right). \end{aligned}$$

Therefore, $\|\Delta^{-1}_{Z_e,Z}\|_1 \ge cn$ for some constant $c > 0$. Similarly, we have $\|\nabla^{-1}_{Z_e,Z}\|_1 \ge c'n$ for another constant $c' > 0$. This leads to much tighter bounds than the ones of Theorem 8.1 for $A = Z_e$, $B = Z_0$.

In all cases, we have $\|\nabla^{-1}_{Z_e,Z_f}\|_1 \ge cn$ for all $ef \ne 1$; $\|\Delta^{-1}_{Z_e,Z_f}\|_1 \ge cn$ for all $e \ne f$, where c is a positive constant independent of n.

THEOREM 8.5.

$$\max_j |1 - |e|^{1/m}|v_j|\omega_{2m}|^{-1} \le \|\Delta^{-1}_{Z_e,D(\mathbf{v})}\| \le \hat{e}\sum_{k=0}^{m-1}\max_j\left|\frac{v_j^k}{1 - ev_j^m}\right|, \tag{8.5}$$

$$\max_j ||e|^{1/m} - |v_j|\omega_{2m}|^{-1} \le \|\nabla^{-1}_{Z_e,D(\mathbf{v})}\| \le \hat{e}\sum_{k=0}^{m-1}\max_j\left|\frac{v_j^k}{e - v_j^m}\right|. \tag{8.6}$$

Proof. (1) The lower bounds in (8.5) and (8.6) follow from Theorem 8.1 for $A = Z_e$, $B = D(\mathbf{v})$.

(2) Let $\Delta = \Delta_{Z_e,D(\mathbf{v})}(M)$. Then $M = \sum_{k=0}^{m-1} Z_e^k \Delta D(\mathbf{v})^k (I_n - eD(\mathbf{v})^m)^{-1}$ (see Corollary 4.3). So $\|M\| \leq \sum_{k=0}^{m-1} \|Z_e^k\| \; \|\Delta\| \; \|D(\mathbf{v})^k (I_n - eD(\mathbf{v})^m)^{-1}\| \leq \|\Delta\| \hat{e} \sum_{k=0}^{m-1} \max_j |\frac{v_j^k}{e - v_j^m}|$.

(3) Assume $e \neq 0$; otherwise, $\|\nabla_{Z_0,D(\mathbf{v})}^{-1}\| = \lim_{e\to 0} \|\nabla_{Z_e,D(\mathbf{v})}^{-1}\|$. Let $\nabla = \nabla_{Z_e,D(\mathbf{v})}(M)$. Then $M = \sum_{k=0}^{m-1} Z_e^{-1-k} \nabla D(\mathbf{v})^k (I_n - eD(\mathbf{v})^m)^{-1}$ (see Corollary 4.3). So $\|M\| \leq \sum_{k=0}^{m-1} \|Z_e^{-1-k}\| \; \|\nabla\| \; \|D(\mathbf{v})^k (I_n - eD(\mathbf{v})^m)^{-1}\| \leq \|\nabla\| \hat{e} \sum_{k=0}^{m} \max_j |\frac{v_j^k}{e - v_j^m}|$. ☐

Remark 8.6. Suppose $|v_j| \notin (1-\epsilon, 1+\epsilon)$ for a constant $\epsilon > 0$ and for all j.

(a) If $e \neq 0$, then

$$\lim_{m\to\infty} \sum_{k=0}^{m-1} \max_j \left| \frac{v_j^k}{1 - ev_j^m} \right| < \frac{1}{\epsilon} \max\left(1, \frac{1}{|e|}\right),$$
$$\lim_{m\to\infty} \sum_{k=0}^{m-1} \max_j \left| \frac{v_j^k}{e - v_j^m} \right| < \frac{1}{\epsilon} \max\left(1, \frac{1}{|e|}\right).$$

(b) If $e = 0$, let us improve the lower bound by sampling $M = (v_j^{i-1})_{i=1}^n \mathbf{e}_j^T$. Then $\Delta_{Z,D(\mathbf{v})}(M) = \mathbf{e}_1 \mathbf{e}_j^T$. Write $v = \max_j |v_j|$. Since $\|\Delta_{Z,D(\mathbf{v})}(M)\| = 1$, we have $\sqrt{\frac{v^{2m}-1}{v^2-1}} \leq \|\Delta_{Z,D(\mathbf{v})}^{-1}\|_1 \leq \frac{v^m - 1}{v-1}$. Compare the latter lower and upper bounds as $m \to \infty$ to obtain $\lim_{m\to\infty} \sqrt{\frac{v^{2m}-1}{v^2-1}} / \frac{v^m-1}{v-1} = |\frac{v+1}{v-1}|$. That is, our estimates (8.5), (8.6) are asymptotically tight as $m \to \infty$ for any e provided that $\{v_j\}$ are not clustered around 1.

Theorem 3.1 for $A = D(\mathbf{u})$, $B = D(\mathbf{v})$ implies the next corollary.

COROLLARY 8.7.

$$\frac{1}{\min_{i,j} |1 - u_i v_j|} \leq \|\Delta_{D(\mathbf{u}),D(\mathbf{v})}^{-1}\|_r \leq \frac{\sqrt{r}}{\min_{i,j} |1 - u_i v_j|}, \tag{8.7}$$
$$\frac{1}{\min_{i,j} |u_i - v_j|} \leq \|\nabla_{D(\mathbf{u}),D(\mathbf{v})}^{-1}\|_r \leq \frac{\sqrt{r}}{\min_{i,j} |u_i - v_j|}. \tag{8.8}$$

Remark 8.8. The lower and upper bounds of Corollary 8.7 are within the factor of $\sqrt{r}$ from each other, and r is small for structured matrices.

Next, we extend the upper estimates of Corollary 8.7 for $\|L^{-1}\|_r$ based on the displacement transformation techniques, with the goal of improving our estimates (8.3)–(8.6) when $|e|$ and $|f|$ are not too small or too large.

THEOREM 8.9. *Let $\hat{A} = VAV^{-1}$, $\hat{B} = W^{-1}BW$ for some nonsingular matrices V and W, $C = \|V\| \; \|V^{-1}\| \; \|W\| \; \|W^{-1}\|$. Then*

$$\|\Delta_{A,B}^{-1}\|_r \leq C \|\Delta_{\hat{A},\hat{B}}^{-1}\|_r, \qquad \|\nabla_{A,B}^{-1}\|_r \leq C \|\nabla_{\hat{A},\hat{B}}^{-1}\|_r.$$

Proof. $\Delta_{\hat{A},\hat{B}}(VMW) = V\Delta_{A,B}(M)W$, $\nabla_{\hat{A},\hat{B}}(VMW) = V\nabla_{A,B}(M)W$. ☐

By combining Theorems 8.9 and 2.3, we transform the operators Δ_{Z_e,Z_f}^{-1}, ∇_{Z_e,Z_f}^{-1} and $\Delta_{Z_e,D(\mathbf{v})}^{-1}$, $\nabla_{Z_e,D(\mathbf{v})}^{-1}$ into the operators $\Delta_{D(\mathbf{u}),D(\mathbf{v})}^{-1}$, $\nabla_{D(\mathbf{u}),D(\mathbf{v})}^{-1}$ and then extend the bounds of Corollary 8.7 to the former operators. We arrive at a corollary showing the desired improvement of (8.3)–(8.6).

COROLLARY 8.10. *Suppose* $ef \neq 0$; $e^{\frac{1}{m}}$ *and* $f^{\frac{1}{n}}$ *are any* m*th and* n*th roots of* e *and* f, *respectively. Write* $\tilde{e} = \max(|e|, \frac{1}{|e|})$, $\tilde{f} = \max(|f|, \frac{1}{|f|})$, *and then*

$$\|\Delta^{-1}_{Z_e,Z_f}\|_r \leq \sqrt{r}\tilde{e}^{\frac{m-1}{m}}\tilde{f}^{\frac{n-1}{n}} \max_{i,j} |1 - e^{\frac{1}{m}}\omega^i_m f^{\frac{1}{n}}\omega^j_n|^{-1},$$

$$\|\nabla^{-1}_{Z_e,Z_f}\|_r \leq \sqrt{r}\tilde{e}^{\frac{m-1}{m}}\tilde{f}^{\frac{n-1}{n}} \max_{i,j} |e^{\frac{1}{m}}\omega^i_m - f^{\frac{1}{n}}\omega^j_n|^{-1},$$

$$\|\Delta_{Z_e,D(\mathbf{v})}\|_r \leq \sqrt{r}\tilde{e}^{\frac{m-1}{m}} \max_{i,j} |1 - e^{\frac{1}{m}}\omega^i_m v_j|^{-1},$$

$$\|\nabla_{Z_e,D(\mathbf{v})}\|_r \leq \sqrt{r}\tilde{e}^{\frac{m-1}{m}} \max_{i,j} |e^{\frac{1}{m}}\omega^i_m - v_j|^{-1}.$$

9. Decreasing the norm $\|L^{-1}\|$. Typically, in the DECOMPRESS stage of the displacement rank approach, the numerical problems diminish where $\|L^{-1}\|$ is smaller. In particular, this factor is critical for rapid convergence of Newton's iteration with recursive compression applied to invert a structured matrix [P92], [P01], [PRW02], [PKRC02], [CPVBW02]. It is, therefore, desired to decrease $\|L^{-1}\|$. Surprisingly, this is possible based on the displacement transformation approach proposed in [P90]. According to this approach, a successful algorithm or method of study for a specific class of structured matrices can be extended to other classes of structured matrices via appropriate transformation of the associated displacement operators. At the end of the preceding section, we applied this approach to extend our estimates of Corollary 8.7 from Cauchy-like to Toeplitz/Hankel-like and Vandermonde-like matrices. Let us demonstrate how this works by another example. Suppose we seek the solution of a nonsingular linear system $M\mathbf{x} = \mathbf{b}$, where the Cauchy-like input matrix M is associated with an operator $L_0 = \nabla_{D(\mathbf{s}),D(\mathbf{t})}$, and suppose the norm $\|L_0^{-1}\|$ is too large. Let us solve the problem by using the displacement transformation method. Choose a vector $\mathbf{v} = (a\omega^i_n)^{n-1}_{i=0}$ for a scalar a such that $\|L^{-1}\|$ is small for $L = \nabla_{D(\mathbf{s}),D(\mathbf{v})}$. According to Corollary 8.7, this is the case if the component sets of the vectors $\mathbf{s}$ and $\mathbf{v}$ are well isolated from each other. Solve the linear system $MC(\mathbf{t},\mathbf{v})\mathbf{y} = \mathbf{b}$ whose coefficient matrix is associated with the operator $L = \nabla_{D(\mathbf{s}),D(\mathbf{v})}$ and is typically not worse conditioned than M; finally, recover $\mathbf{x} = C(\mathbf{t},\mathbf{v})\mathbf{y}$. Due to the transition from L_0 to L, the critical stage of the solution can be dramatically simplified (for instance, if the solution is obtained by using Newton's iteration). The above recipe can be immediately extended to the case of Toeplitz-like, Hankel-like, Vandermonde-like, and other structured matrices M based on their well-known simple transformations into Cauchy-like matrices (see Theorems 2.3, 3.7, and 8.9 and [P90]).

Acknowledgments. We thank both referees for many helpful comments and the editor Prof. Dr. Ir. Sabine Van Huffel for ensuring fast processing of the paper and compression of its first draft.

REFERENCES

[AG91] G. AMMAR AND P. GADER, *A variant of the Gohberg–Semencul formula involving circulant matrices*, SIAM J. Matrix Anal. Appl., 12 (1991), pp. 534–540.

[BGR90] J. A. BALL, I. GOHBERG, AND L. RODMAN, *Interpolation of Rational Matrix Functions*, Oper. Theory Adv. Appl. 45, Birkhäuser, Basel, 1990.

[BP93] D. A. BINI AND V. Y. PAN, *Improved parallel computations with Toeplitz-like and Hankel-like matrices*, Linear Algebra Appl., 188/189 (1993), pp. 3–29.

[BP94] D. A. BINI AND V. Y. PAN, *Polynomial and Matrix Computations, Vol.* 1: *Fundamental Algorithms*, Birkhäuser Boston, Boston, 1994.

[BVB97] A. Bultheel and M. Van Barel, *Linear Algebra, Rational Approximation and Orthogonal Polynomials*, Stud. Comput. Math. 6, North–Holland, Amsterdam, 1997.

[CJL96] S. Cabay, A. R. Jones, and G. Labahn, *Computation of numerical Padé–Hermite and simultaneous Padé systems* I: *Near inversion of generalized Sylvester matrices*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 248–267.

[CJL96a] S. Cabay, A. R. Jones, and G. Labahn, *Computation of numerical Padé–Hermite and simultaneous Padé systems* II: *A weakly stable aglorithm*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 268–297.

[CPVBW02] G. Codevico, V. Y. Pan, M. van Barel, and X. Wang, *Iterative Inversion of Structured Matrices*, Tech. report 200214, Ph.D. Program in Computer Science, The Graduate Center of the City University of New York, New York, NY, 2002 (submitted).

[CPW74] R. E. Cline, R. J. Plemmons, and G. Worm, *Generalized inverses of certain Toeplitz matrices*, Linear Algebra Appl., 8 (1974), pp. 25–33.

[GKKL87] I. Gohberg, T. Kailath, I. Koltracht, and P. Lancaster, *Linear complexity parallel algorithms for linear systems of equations with recursive structure*, Linear Algebra Appl., 88/89 (1987), pp. 271–315.

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

[GO92] I. Gohberg and V. Olshevsky, *Circulants, displacements and decompositions of matrices*, Integral Equations Operator Theory, 15 (1992), pp. 730–743.

[GO94] I. Gohberg and V. Olshevsky, *Complexity of multiplication with vectors for structured matrices*, Linear Algebra Appl., 202 (1994), pp. 163–192.

[GO94a] I. Gohberg and V. Olshevsky, *Fast state-space algorithms for matrix Nehari and Nehari–Takagi interpolation problems*, Integral Equations Operator Theory, 20 (1994), pp. 44–83.

[G98] M. Gu, *Stable and efficient algorithms for structured systems of linear equations*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 279–306.

[G98a] M. Gu, *New fast algorithms for structured linear least squares problems*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 244–269.

[H95] G. Heinig, *Inversion of generalized Cauchy matrices and the other classes of structured matrices*, in Linear Algebra for Signal Processing, IMA Vol. Math. Appl. 69, Springer-Verlag, New York, 1995, pp. 95–114.

[HR84] G. Heinig and K. Rost, *Algebraic Methods for Toeplitz-like Matrices and Operators*, Oper. Theory Adv. Appl. 13, Birkhäuser, Basel, 1984.

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

[KO96] T. Kailath and V. Olshevsky, *Displacement structure approach to discrete transform based preconditioners of G. Strang type and of T. Chan type*, Calcolo, 33 (1996), pp. 191–208.

[KS95] T. Kailath and A. H. Sayed, *Displacement structure: Theory and applications*, SIAM Rev., 37 (1995), pp. 297–386.

[KS99] T. Kailath and A. H. Sayed, eds., *Fast Reliable Algorithms for Matrices with Structure*, SIAM, Philadelphia, 1999.

[KVB99] P. Kravanja and M. Van Barel, *Algorithms for solving rational interpolation problems related to fast and superfast solvers for Toeplitz systems*, in Proceedings of Advanced Signal Processing Algorithms, Architecture and Implementation IX (Denver, CO), SPIE Publications, Bellingham, WA, 1999, pp. 359–370.

[OP98] V. Olshevsky and V. Y. Pan, *A unified superfast algorithm for boundary rational tangential interpolation problem*, in Proceedings of the 39th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, Los Alamitos, CA, 1998, pp. 192–201.

[OS00] V. Olshevsky and M. A. Shokrollahi, *A unified superfast algorithm for confluent tangential interpolation problem and for structured matrices*, in Proceedings of the 14th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2000), University of Perpignan, Perpignan, France, 2000.

[P90] V. Y. Pan, *On computations with dense structured matrices*, Math. Comput., 55 (1990), pp. 179–190. Proceedings version in Proceedings of International Symposium on Symbolic and Algebraic Computation (ISSAC '89), ACM, New York, 1989, pp. 34–42.

[P92] V. Y. Pan, *Parallel solution of Toeplitz-like linear systems*, J. Complexity, 8 (1992), pp. 1–21.

[P93] V. Y. Pan, *Decreasing the displacement rank of a matrix*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 118–121.

[P99] V. Y. Pan, *A Unified Superfast Divide-and-Conquer Algorithm for Structured Matrices over Abstract Fields*, MSRI Preprint 1999-033, Mathematical Sciences Research Institute, Berkeley, CA, 1999.

[P00] V. Y. Pan, *Nearly optimal computations with structured matrices*, in Proceedings of the 11th Annual ACM–SIAM Symposium on Discrete Algorithms (SODA 2000), ACM, New York, SIAM, Philadelphia, 2000, pp. 953–962.

[P01] V. Y. Pan, *Structured Matrices and Polynomials: Unified Superfast Algorithms*, Birkhäuser Boston, Boston, Springer-Verlag, New York, 2001.

[PKRC02] V. Y. Pan, M. Kunin, R. Rosholt, and H. Cebecioğlu, *Homotopic Residual Correction Processes*, Tech. report 200215, Ph.D. Program in Computer Science, The Graduate Center of the City University of New York, New York, NY, 2002 (submitted).

[PRW02] V. Y. Pan, Y. Rami, and X. Wang, *Structured matrices and Newton's iteration: Unified approach,* Linear Algebra Appl., 343/344 (2002), pp. 233–265.

[PSD70] P. Penfield, Jr., R. Spencer, and S. Duinker, *Tellegen's Theorem and Electrical Networks*, MIT Press, Cambridge, MA, 1970.

[W93] D. H. Wood, *Product rules for the displacement of nearly-Toeplitz matrices*, Linear Algebra Appl., 188/189 (1993), pp. 641–663.