Topics in the numerical linear algebra of Toeplitz and Hankel matrices

Albrecht Böttcher[*] and **Karla Rost**[**]

Fakultät für Mathematik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany

This is an introduction to some aspects of the numerical linear algebra of large matrices with Toeplitz, Hankel, and Toeplitz-plus-Hankel structures. The concrete topics we have selected are determined by our preferences and by work we have participated in. We give an introduction to the symbol calculus, we touch exact and asymptotic formulas for inverses, including the notion of a Bezoutian, we consider eigenvalues, pseudospectra, eigenvectors, and condition numbers, and we embark on the fast solution of Toeplitz systems.

1 Introduction

Square Toeplitz and Hankel matrices are matrices of the form

$$T_n = (a_{j-k})_{j,k=1}^n = \begin{pmatrix} a_0 & a_{-1} & \dots & a_{-(n-1)} \\ a_1 & a_0 & \dots & a_{-(n-2)} \\ \vdots & \vdots & & \vdots \\ a_{n-1} & a_{n-2} & \dots & a_0 \end{pmatrix}, \tag{1}$$

$$H_n = (a_{j+k-1})_{j,k=1}^n = \begin{pmatrix} a_1 & a_2 & \dots & a_n \\ a_2 & a_3 & \dots & a_{n+1} \\ \vdots & \vdots & & \vdots \\ a_n & a_{n+1} & \dots & a_{2n-1} \end{pmatrix}, \tag{2}$$

respectively. Each of these two types of $n \times n$ matrices depends on $2n - 1$ parameters, which are usually complex numbers. Such matrices are currently emerging in plenty of applications. People are in particular interested in the spectral properties (for example, in the eigenvalues and eigenvectors), in the structure of the inverse, in several kinds of factorizations, and in efficiently solving linear systems with such coefficient matrices. These questions can all be tackled without difficulty if n is of moderate size, but they become challenging if n is large. The purpose of this article is to present some exemplary results on large Toeplitz and Hankel matrices which may serve as modest illustrations of an actually very big business.

[*] Corresponding author: e-mail: aboettch@mathematik.tu-chemnitz.de
[**] e-mail: krost@mathematik.tu-chemnitz.de

2   Symbol calculus

When working with large matrices, it is sometimes no bad idea to pass to infinite matrices. The infinite versions of matrices (1) and (2) are

$$T = (a_{j-k})_{j,k=1}^{\infty} = \begin{pmatrix} a_0 & a_{-1} & a_{-2} & \dots \\ a_1 & a_0 & a_{-1} & \dots \\ a_2 & a_1 & a_0 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}, \tag{3}$$

$$H = (a_{j+k-1})_{j,k=1}^{\infty} = \begin{pmatrix} a_1 & a_2 & a_3 & \dots \\ a_2 & a_3 & a_4 & \dots \\ a_3 & a_4 & a_5 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}. \tag{4}$$

Matrix (3) is determined by a doubly infinite sequence $(a_n)_{n=-\infty}^{\infty}$, while matrix (4) is specified by a simply infinite sequence $(a_n)_{n=1}^{\infty}$. Classic results by Toeplitz and Nehari characterize all sequences (a_n) for which matrices (3) and (4) induce bounded operators on the usual space ℓ^2. This is in particular the case if $\sum |a_n| < \infty$ (which can also be verified straightforwardly).

The basic properties of an infinite Toeplitz or Hankel matrix are encoded in a certain function, the so-called symbol of the matrix.

Let a be a complex-valued function in L^1 on the complex unit circle $\mathbf{T}$. The Fourier coefficients are defined by

$$a_n = \frac{1}{2\pi} \int_0^{2\pi} a(e^{i\theta}) e^{-in\theta} d\theta, \quad n \in \mathbf{Z}. \tag{5}$$

The set of all $a \in L^1(\mathbf{T})$ for which the sequence of the Fourier coefficients belongs to ℓ^1 is called the Wiener algebra and is denoted by W. In other words, a is in W if and only if $\sum_{n=-\infty}^{\infty} |a_n| < \infty$. The Fourier series of a function $a \in W$ is absolutely convergent, and we have

$$a(e^{i\theta}) = \sum_{n=-\infty}^{\infty} a_n e^{in\theta}, \quad e^{i\theta} \in \mathbf{T}. \tag{6}$$

We emphasize that the product of two functions in W is again in W. Wiener's theorem says that if $a \in W$ has no zeros on $\mathbf{T}$, then a^{-1} is also in W.

Consider now the Toeplitz matrix (3) and suppose $\sum_{n=-\infty}^{\infty} |a_n| < \infty$. In this case (6) defines a function a in W. This function is referred to as the symbol of the matrix (3), and we denote matrix (3) by $T(a)$. Notice that we can also look at things from the reverse side: we start with a function $a \in W$, define the sequence $(a_n)_{n=-\infty}^{\infty}$ via (5), and eventually let $T(a)$ stand for the matrix (3).

For a function $a \in W$ with Fourier coefficients (5), we define two infinite Hankel matrices $H(a)$ and $H(\widetilde{a})$ by

$$H(a) = \begin{pmatrix} a_1 & a_2 & a_3 & \dots \\ a_2 & a_3 & \dots & \\ a_3 & \dots & & \\ \dots & & & \end{pmatrix}, \quad H(\widetilde{a}) = \begin{pmatrix} a_{-1} & a_{-2} & a_{-3} & \dots \\ a_{-2} & a_{-3} & \dots & \\ a_{-3} & \dots & & \\ \dots & & & \end{pmatrix}.$$

Clearly, $H(a)$ and $H(\widetilde{a})$ are uniquely determined by a. However, if we are given matrix (4) with $\sum_{n=1}^{\infty}|a_n| < \infty$, then there are infinitely many $a \in W$ such that $H(a)$ (or $H(\widetilde{a})$) is the matrix (4). Thus, the symbol of a Hankel matrix is not unique.

We are now in a position to state a formula for the product of two infinite Toeplitz matrices. We always assume that $a, b \in W$.

Theorem 2.1 *We have* $T(a)T(b) = T(ab) - H(a)H(\widetilde{b})$.

This theorem, which goes back at least to Gohberg and Feldman [16], can be readily verified by inspection. It tells us that the product of two infinite Toeplitz matrices differs from an infinite Toeplitz matrix by the product of two infinite Hankel matrices. We remark that if $a, b \in W$, then $H(a)$ and $H(\widetilde{b})$ are compact operators on ℓ^2. Consequently, $T(a)T(b)$ equals $T(ab)$ modulo compact operators.

We define the operators P_n and W_n on ℓ^2 by

$$P_n : (x_0, x_1, x_2, \ldots) \mapsto (x_0, \ldots, x_{n-1}, 0, \ldots),$$
$$W_n : (x_0, x_1, x_2, \ldots) \mapsto (x_{n-1}, \ldots, x_0, 0, \ldots),$$

and we identify the ranges of P_n and W_n with $\mathbf{C}^n$ in the natural manner. Matrix (1) may then be identified with $P_nT(a)P_n$, the principal $n \times n$ truncation of $T(a)$. We denote it by $T_n(a)$. Clearly, as for fixed n only the Fourier coefficients $a_{-(n-1)}, \ldots, a_{n-1}$ are specified, there are infinitely many $a \in W$ such that $T_n(a)$ equals (1). Here is the analogue of Theorem 2.1 for finite matrices.

Theorem 2.2 *We have*

$$T_n(a)T_n(b) = T_n(ab) - P_nH(a)H(\widetilde{b})P_n - W_nH(\widetilde{a})H(b)W_n.$$

This beautiful formula is due to Widom [50]. Once it has been guessed, it can again be proved by simply computing and comparing the j, k entries of both sides.

3   Formulas for inverses

Throughout what follows we suppose that $a \in W$. When does the Toeplitz matrix $T(a)$ generate an invertible operator on ℓ^2? This question has a nice answer in terms of the symbol a. The counter-clockwise orientation of $\mathbf{T}$ induces an orientation of the continuous closed curve $a(\mathbf{T})$. If a has no zeros on $\mathbf{T}$, then $a(\mathbf{T})$ does not contain the origin and hence has a well-defined winding number $\operatorname{wind} a$ about the origin. The following classic result has emerged from the work of many authors and was probably first explicitly stated by Gohberg [14].

Theorem 3.1 *The matrix $T(a)$ induces an invertible operator on ℓ^2 if and only if the function a has no zeros on $\mathbf{T}$ and* $\operatorname{wind} a = 0$.

The following formula for the inverse $T^{-1}(a) := (T(a))^{-1}$ has its root in the so-called Wiener-Hopf factorization and is in this form due to Mark Krein [39]. We denote by $T(\widetilde{a})$ the transpose of $T(a)$,

$$T(\widetilde{a}) = (a_{k-j})_{j,k=1}^{\infty} = \begin{pmatrix} a_0 & a_1 & a_2 & \ldots \\ a_{-1} & a_0 & a_1 & \ldots \\ a_{-2} & a_{-1} & a_0 & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{pmatrix}, \tag{7}$$

and by $e_0 \in \ell^2$ the sequence $(1, 0, 0, \ldots)$.

Theorem 3.2 *Let $T(a)$ be invertible. Then the solutions $x = (x_j)_{j=0}^\infty$ and $y = (y_j)_{j=0}^\infty$ of the equations $T(a)x = e_0$ and $T(\widetilde{a})y = e_0$ belong to ℓ^1, we have $x_0 \neq 0$, and*

$$
T^{-1}(a) = \frac{1}{x_0}
\begin{pmatrix}
x_0 & & & \\
x_1 & x_0 & & \\
x_2 & x_1 & x_0 & \\
\cdots & \cdots & \cdots & \cdots
\end{pmatrix}
\begin{pmatrix}
y_0 & y_1 & y_2 & \cdots \\
& y_0 & y_1 & \cdots \\
& & y_0 & \cdots \\
& & & \cdots
\end{pmatrix}.
\tag{8}
$$

To formulate the analogue of Theorem 3.2 for $T_n^{-1}(a) := (T_n(a))^{-1}$, we understand by $e_0 \in \mathbf{C}^n$ the vector $(1, 0, \ldots, 0)$ and consider the equations $T_n(a)x = e_0$ and $T_n(\widetilde{a})y = e_0$. Here $T_n(\widetilde{a})$ is the leading principal $n \times n$ submatrix of the matrix (7). Equivalently, $T_n(\widetilde{a})$ is the transpose of $T_n(a)$. The following inversion formula is well known as the Gohberg-Semencul formula [17].

Theorem 3.3 *Suppose $T_n(a)$ is invertible and let $x = (x_j)_{j=0}^{n-1}$ and $y = (y_j)_{j=0}^{n-1}$ be the solutions of the equations $T_n(a)x = e_0$ and $T_n(\widetilde{a})y = e_0$. Put $x_n = y_n := 0$. If $x_0 \neq 0$, then*

$$
\begin{aligned}
T_n^{-1}(a) \;=\; & \frac{1}{x_0}
\begin{pmatrix}
x_0 & & \\
\vdots & \ddots & \\
x_{n-1} & \cdots & x_0
\end{pmatrix}
\begin{pmatrix}
y_0 & \cdots & y_{n-1} \\
& \ddots & \vdots \\
& & y_0
\end{pmatrix} \\
& - \frac{1}{x_0}
\begin{pmatrix}
y_n & & \\
\vdots & \ddots & \\
y_1 & \cdots & y_n
\end{pmatrix}
\begin{pmatrix}
x_n & \cdots & x_1 \\
& \ddots & \vdots \\
& & x_n
\end{pmatrix}.
\end{aligned}
\tag{9}
$$

Notice that x and y in Theorem 3.3 depend stronger on n than notation suggests: actually we have $x = x^{(n)} = (x_j^{(n)})_{j=0}^n$ and $y = y^{(n)} = (y_j^{(n)})_{j=0}^n$. One nevertheless has the feeling that the right-hand side of (9) should in some sense converge to that of (8). This is indeed the case, which was first proved by Gohberg and Feldman [16].

Theorem 3.4 *If $T(a)$ is invertible, then there is an n_0 such that $T_n(a)$ is invertible for all $n \geq n_0$ and such that the first component $x_0^{(n)}$ of the solution of $T_n(a)x^{(n)} = e_0$ is nonzero for all $n \geq n_0$. Moreover, the operators $T_n^{-1}(a)P_n$ converge strongly (= pointwise) to the operator $T^{-1}(a)$ on ℓ^2.*

This theorem tells us that we may replace the infinite system $T(a)f = g$ by the (large but finite) systems $T_n(a)f^{(n)} = P_n g$, which is a projection method. If $T(a)$ is invertible, then $f^{(n)}$ converges in ℓ^2 to f.

Theorem 3.4 implies in particular that $x_0^{(n)} \to x_0$ as $n \to \infty$ if only the operator $T(a)$ is invertible. By Cramer's rule, $x_0^{(n)}$ is the quotient of two Toeplitz determinants: $x_0^{(n)} = \det T_{n-1}(a)/\det T_n(a)$. On the other hand, one can show that $x_0 = 1/\exp(\log a)_0$, where $(\log a)_0$ is the 0th Fourier coefficient of $\log a$ and $\log a$ is a logarithm of a in W (which exists by Theorem 3.1). Thus, Theorem 3.4 contains as a special case the formula

$$
\lim_{n \to \infty} \frac{\det T_n(a)}{\det T_{n-1}(a)} = \exp(\log a)_0,
$$

which is the first Szegö limit theorem (see, for example, [10] and [16] for more details).

Theorem 3.4 is at the heart of the asymptotic analysis of large Toeplitz matrices. When proving something about $T_n(a)$ as $n \to \infty$, one will often sooner or later be forced to have recourse to this theorem. Nowadays there are many different proofs of this key result. These proofs also yield representations for $T_n^{-1}(a)$ that are completely different from (9). We will see in the following how (9) can be used to design fast algorithms for solving the system $T_n(a)f = g$. For asymptotic problems, however, formula (9) is frequently inconvenient. In this connection a better representation is, for instance,

$$T_n^{-1}(a) = T_n(a^{-1}) + P_n K(a) P_n + W_n K(\widetilde{a}) W_n + C_n,$$

where $\|C_n\| \to 0$ and the compact operators $K(a)$ and $K(\widetilde{a})$ are given by

$$K(a) = T^{-1}(a) - T(a^{-1}) = H(a^{-1})H(\widetilde{a})T^{-1}(a),$$
$$K(\widetilde{a}) = T^{-1}(\widetilde{a}) - T(\widetilde{a}^{-1}) = H(\widetilde{a}^{-1})H(a)T^{-1}(\widetilde{a}).$$

We refer to the text [10] for more on this topic.

4 Eigenvalues and pseudospectra

We denote by $\mathrm{sp}\, A$ the spectrum of a matrix or an operator, that is, the set of all $\lambda \in \mathbf{C}$ for which $A - \lambda I$ is not invertible. Notice that for Toeplitz matrices the equalities $T(a) - \lambda I = T(a - \lambda)$ and $T_n(a) - \lambda I_n = T_n(a - \lambda)$ hold. By Theorem 3.1,

$$\mathrm{sp}\, T(a) = a(\mathbf{T}) \cup \{\lambda \in \mathbf{C} \setminus a(\mathbf{T}) : \mathrm{wind}\,(a, \lambda) \neq 0\},$$

where $\mathrm{wind}\,(a, \lambda)$ is the winding number of the curve $a(\mathbf{T})$ about the point λ. It turns out that in the $n \to \infty$ limit the spectra (= sets of eigenvalues) $\mathrm{sp}\, T_n(a)$ are in general in no obvious way related to $\mathrm{sp}\, T(a)$. Figure 1 shows an example.

In the case of banded Toeplitz matrices, the limit of the spectra $\mathrm{sp}\, T_n(a)$ was completely identified by Schmidt and Spitzer [48]. Thus, let us assume that a is a Laurent polynomial,

$$a(e^{i\theta}) = \sum_{k=-r}^{s} a_k e^{ik\theta}.$$

In this case $T(a)$ is banded with bandwidth $r + s + 1$. We define

$$\Lambda(a) = \bigcap_{\varrho > 0} \mathrm{sp}\, T(a_\varrho) \quad \text{where} \quad a_\varrho(e^{i\theta}) = \sum_{k=-r}^{s} a_k \varrho^k e^{ik\theta}.$$

Here is the result of Schmidt and Spitzer.

Theorem 4.1 *The spectra $\mathrm{sp}\, T_n(a)$ converge to $\Lambda(a)$ in the Hausdorff metric as $n \to \infty$. The set $\Lambda(a)$ is a finite union of analytic arcs.*

For dense Toeplitz matrices $T(a)$, the identification of $\mathrm{sp}\, T_n(a)$ for large n remains a true challenge. The set $\Lambda(a) := \limsup \mathrm{sp}\, T_n(a)$ always exists. It is defined as the set of all $\lambda \in \mathbf{C}$ for which there are $n_1 < n_2 < n_3 < \ldots$ and $\lambda_k \in \mathrm{sp}\, T_{n_k}(a)$ such that $\lambda_k \to \lambda$. Let $\mathcal{H}_{\mathbf{C}}$ denote the metric space of all compact subsets of $\mathbf{C}$ with the Hausdorff metric. In [6] it

is shown that the map $C(\mathbf{T}) \to \mathcal{H}_{\mathbf{C}}, a \mapsto \Lambda(a)$ is discontinuous. Thus, small changes in the entries of a large dense Toeplitz matrix may drastically change the spectrum.

In contrast to spectra, pseudospectra of Toeplitz matrices behave as nicely as one could ever expect. For $\varepsilon > 0$, the ε-pseudospectrum of a matrix or an operator A is defined by

$$\mathrm{sp}_\varepsilon A = \{\lambda \in \mathbf{C} : \|(A - \lambda I)^{-1}\| \geq 1/\epsilon\}, \tag{10}$$

with the convention that $\|(A - \lambda I)^{-1}\| = +\infty$ if $\lambda \in \mathrm{sp}\, A$. Throughout this article, $\|\cdot\|$ is the operator norm on ℓ^2 or on $\mathbf{C}^n$ with the ℓ^2 norm. We should mention that $\mathrm{sp}_\varepsilon A$ admits the alternative description

$$\mathrm{sp}_\varepsilon A = \bigcup_{\|E\| \leq \varepsilon} \mathrm{sp}\,(A + E). \tag{11}$$

Formula (11) implies that we can get an idea of $\mathrm{sp}_\varepsilon A$ by superposing the spectra $\mathrm{sp}(A + E_k)$ for sufficiently many randomly chosen $E_1, \ldots, E_N$ with $\|E_k\| \leq \varepsilon$. Figure 2 is an example. For more on pseudospectra we refer the reader to the forthcoming book [49].
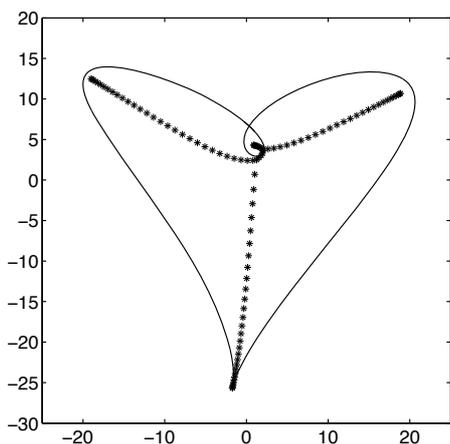
Fig. 1 We see the curve $a(\mathbf{T})$ together with the 100 eigenvalues of $T_{100}(a)$.
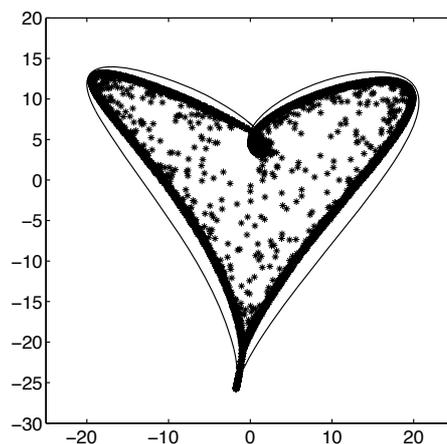
Fig. 2 The picture indicates $\mathrm{sp}_\varepsilon T_n(a)$ for $\varepsilon = 1/100$ and $n = 200$. It shows the superposition of the spectra $\mathrm{sp}\,(T_n(a) + E_k)$ for 50 randomly chosen matrices E_k with $\|E_k\| = \varepsilon$.

Theorem 4.2 *For each $\varepsilon > 0$, the pseudospectra $\mathrm{sp}_\varepsilon T_n(a)$ converge to $\mathrm{sp}_\varepsilon T(a)$ in the Hausdorff metric as $n \to \infty$.*

This theorem goes back to Henry Landau [42] and Reichel and Trefethen [46]. The first clean proof was given in [3]. This proof is based on (10) and the equality

$$\lim_{n \to \infty} \|T_n^{-1}(a - \lambda)\| = \|T^{-1}(a - \lambda)\|. \tag{12}$$

Equality (12) was in turn proved by C^*-algebra techniques (see [4] or [18] for more on this subject). Take, for example, $\lambda = 0$. Then (12) says that $\|T_n^{-1}(a)\|$ converges to $\|T^{-1}(a)\|$

if $T(a)$ is invertible and to $+\infty$ if $T(a)$ is not invertible. This is a significant refinement of Gohberg and Feldman's classic Theorem 3.4, which is essentially equivalent to the statement that $\|T_n^{-1}(a)\|$ remains bounded as $n \to \infty$ if and only if $T(a)$ is invertible.

If a is a Laurent polynomial, then the convergence in (12) is exponentially fast. This implies that $\mathrm{sp}_\varepsilon T_n(a)$ converges to $\mathrm{sp}_\varepsilon T(a)$ extremely rapidly. On the other hand, for non-smooth symbols a the convergence of $\mathrm{sp}_\varepsilon T_n(a)$ to $\mathrm{sp}_\varepsilon T(a)$ may be spectacularly slow, so that it cannot be convincingly seen numerically [5, 7].

The problems for Hankel matrices are often completely different from those for Toeplitz matrices. This happens in particular in connection with eigenvalues and pseudospectra. Let $H_n(a)$ denote the principal $n \times n$ truncation of the infinite Hankel matrix $H(a)$. Since $H(a)$ is compact on ℓ^2, it follows easily that $\mathrm{sp}\, H_n(a) \to \mathrm{sp}\, H(a)$ and $\mathrm{sp}_\varepsilon H_n(a) \to \mathrm{sp}_\varepsilon H(a)$ in the Hausdorff metric as $n \to \infty$.

5   Eigenvectors

Let $T(a)$ be a banded matrix, that is, suppose a is a Laurent polynomial. Fix a point $\lambda \in \mathbf{C}$, and for the sake of simplicity assume that $\lambda \notin a(\mathbf{T})$. Then the winding number $\mathrm{wind}\,(a, \lambda)$ of the curve $a(\mathbf{T})$ about λ is well-defined. Recall that $T(a) - \lambda I = T(a - \lambda)$ and $T_n(a) - \lambda I = T_n(a - \lambda)$. The following theorem, which is essentially due to Gohberg [15] and Mark Krein [39], provides information about the eigenvectors of infinite Toeplitz matrices.

Theorem 5.1 *The point λ is an eigenvalue of the operator $T(a) : \ell^2 \to \ell^2$ if and only if* $\mathrm{wind}\,(a, \lambda) = -m \leq -1$. *In that case*

$$\mathrm{Ker}\, T(a - \lambda) = \mathrm{span}\,\{x, Vx, \ldots, V^{m-1}x\},$$

where $x = (x_j)_{j=0}^\infty$ is the solution of the equation $T(a)x = e_0$ and V is the shift operator $\{\xi_0, \xi_1, \ldots\} \mapsto \{0, \xi_0, \xi_1, \ldots\}$.

Now suppose we have eigenvalues $\lambda_n \in \mathrm{sp}\, T_n(a)$ and $\lambda_n \to \lambda$ as $n \to \infty$. Clearly, λ is necessarily in the limiting set $\Lambda(a)$. How are the eigenvectors of $T_n(a)$ corresponding to λ_n related to the eigenvectors of $T(a)$ for λ? This problem was solved only recently in [9]. First of all, it turns out that everything is fine in the case where $\mathrm{wind}\,(a, \lambda) = -1$.

Theorem 5.2 *If $\mathrm{wind}\,(a, \lambda) = -1$, then there exist a natural number n_0, vectors $x^{(n)} = (x_j^{(n)})_{j=0}^{n-1} \in \mathbf{C}^n$ $(n \geq n_0)$, and a sequence $x \in \ell^2$ such that*

$$\mathrm{Ker}\, T_n(a - \lambda_n) = \mathrm{span}\,\{x^{(n)}\}, \quad x_0^{(n)} = 1,$$
$$\mathrm{Ker}\, T(a - \lambda) = \mathrm{span}\,\{x\},$$
$$x^{(n)} \to x \ \text{ in } \ \ell^2.$$

Things are not so perfect for $\mathrm{wind}\,(a, \lambda) \neq -1$. Let, for example, $a(e^{i\theta}) = e^{-i\theta} - \beta e^{-2i\theta}$. Then $T_n(a)$ has 1 on the first superdiagonal, $-\beta$ on the second superdiagonal, and zeros elsewhere. Obviously, $\mathrm{sp}\, T_n(a) = \{0\}$. It is readily seen that

$$\mathrm{Ker}\, T_n(a) = \{(x_0, 0, \ldots, 0) : x_0 \in \mathbf{C}\}. \tag{13}$$

If $\beta = 1/2$, then $\mathrm{wind}\,(a, 0) = -1$ and

$$\mathrm{Ker}\, T(a) = \{(x_0, 0, \ldots) : x_0 \in \mathbf{C}\}. \tag{14}$$

Clearly, (13) and (14) fit together according to Theorems 5.1 and 5.2. If $\beta = 2$, then wind $(a, 0) = -2$, and it is not difficult to verify that

$$\operatorname{Ker} T(a) = \left\{ \left(x_0, x_1, \frac{x_1}{2}, \frac{x_1}{2^2}, \frac{x_1}{2^3}, \dots \right) : x_0, x_1 \in \mathbf{C} \right\}. \tag{15}$$

Thus, the limit of (13) delivers only a proper part of (15).

For general Laurent polynomials a the situation is as follows. If a certain technical condition is satisfied (which is generically the case), then, for all sufficiently large n, $\operatorname{Ker} T_n(a - \lambda_n) = \operatorname{span}\{x^{(n)}\}$ with vectors $x^{(n)} = (x_j^{(n)})_{j=0}^{n-1} \in \mathbf{C}^n$ satisfying $x_0^{(n)} = 1$. The limits $x_j := \lim_{n \to \infty} x_j^{(n)}$ exist for each $j \geq 1$. If wind $(a, \lambda) \leq -1$, then $(1, x_1, x_2, \dots)$ lies in $\operatorname{Ker} T(a - \lambda)$, while if wind $(a, \lambda) \geq 0$, we have $\operatorname{Ker} T(a - \lambda) = \{0\}$ and $(1, x_1, x_2, \dots) \notin \ell^2$.

6   Condition numbers

Let a be a Laurent polynomial. If $T(a)$ is invertible, then (12) implies that the (spectral) condition numbers $\kappa(T_n(a)) := \|T_n(a)\| \, \|T_n^{-1}(a)\|$ converge to $\|T(a)\| \, \|T^{-1}(a)\|$ and thus to a finite limit. However, if $T(a)$ is not invertible, then $\kappa(T_n(a))$ grows to infinity. The latter happens if either a has zeros on $\mathbf{T}$ or if a has no zeros on $\mathbf{T}$ but nonzero winding number. As the following result (which can already be found in [46]) shows, the latter case is especially treacherous.

Theorem 6.1 *If* wind $a \neq 0$, *then the condition numbers $\kappa(T_n(a))$ grow at least exponentially to infinity, that is, there are constants $C = C(a) > 0$ and $\alpha = \alpha(a) > 0$ such that $Ce^{\alpha n} \leq \kappa(T_n(a)) \leq +\infty$ for all $n \geq 1$.*

The condition number $\kappa(T_n(a), x)$ of $T_n(a)$ at a given nonzero vector $x \in \mathbf{C}^n$ is defined as follows. For $\varepsilon > 0$, let M_ε be the set of all $n \times n$ matrices δA_n satisfying $\|\delta A_n\| \leq \varepsilon \|T_n(a)\|$ and denote by P_ε the set of all vectors $\delta x \in \mathbf{C}^n$ for which there exists a $\delta A_n \in M_\varepsilon$ such that $(T_n(a) + \delta A_n)(x + \delta x) = T_n(a)x$. Then one defines

$$\kappa(T_n(a), x) := \lim_{\varepsilon \to 0} \sup_{\delta x \in P_\varepsilon} \frac{\|\delta x\|}{\varepsilon \|x\|}. \tag{16}$$

It is well known that $\kappa(T_n(a), x)$ does actually not depend on x and is equal to $\kappa(T_n(a))$ (see, for example, [37, 47]).

In practice, we need not consider perturbations of $T_n(a)$ by general matrices. It is rather reasonable to assume that the perturbing matrix is itself Toeplitz. Thus, replace the M_ε in the previous paragraph by the set of all $n \times n$ *Toeplitz* matrices δA_n satisfying $\|\delta A_n\| \leq \varepsilon \|T_n(a)\|$. The limit (16) is then denoted by $\kappa^{\operatorname{Toep}}(T_n(a), x)$ and is an example of a so-called structured condition number [37, 47]. In contrast to $\kappa(T_n(a), x)$, the number $\kappa^{\operatorname{Toep}}(T_n(a), x)$ may depend on x. One expects that $\kappa^{\operatorname{Toep}}(T_n(a), x)$ is in general significantly smaller than $\kappa(T_n(a), x)$, but, curiously, up to now no convincing example in this direction is known. Quite on the contrary, the following result shows that in the Toeplitz case structured condition numbers are rarely better than usual condition numbers and that, moreover, the numerical search for an exception from this rule seems to be a hopeless venture.

Theorem 6.2 *Let $x_0, x_1, \ldots, x_{n-1} \in \mathbf{C}$ be independent random variables whose real and imaginary parts are subject to the standard normal distribution and put $x = (x_j)_{j=0}^{n-1}$. There are universal constants $\delta \in (0, \infty)$ and $n_0 \in \mathbf{N}$ such that*

$$\text{Probability} \left(\frac{\kappa^{\text{Toep}}(T_n(a), x)}{\kappa(T_n(a), x)} \geq \frac{\delta}{n^{3/2}} \right) > \frac{99}{100}$$

for all Laurent polynomials a and all $n \geq n_0$.

This theorem was proved in [8] on the basis of a deep recent theorem by Konyagin and Schlag [38]. Theorems 6.1 and 6.2 reveal that with probability of at least 99 % we have

$$\kappa^{\text{Toep}}(T_n(a), x) \geq \delta \, \frac{\kappa(T_n(a), x)}{n^{3/2}} = \delta \, \frac{\kappa(T_n(a))}{n^{3/2}} \geq C\delta \, \frac{e^{\alpha n}}{n^{3/2}}$$

for all $n \geq n_0$. Clearly, the "structural improvement" $n^{3/2}$ in the denominator is absolutely harmless in comparison with the $e^{\alpha n}$ in the numerator.

7   Superfast algorithms for matrix-vector multiplication

Theorem 3.3 reveals the structure of Toeplitz inverses. After [17], a large number of inversion formulas for Toeplitz and Hankel matrices have been established. Several such representations can be derived as follows. Let $\partial T_n(a)$ denote the $(n-1) \times (n+1)$ Toeplitz matrix that results from $T_n(a)$ by deleting the first row and adding a last column,

$$\partial T_n(a) = (a_{j-k})_{j=1, k=0}^{n-1, n} = \begin{pmatrix} a_1 & a_0 & \ldots & a_{-(n-1)} \\ a_2 & a_1 & \ldots & a_{-(n-2)} \\ \vdots & \vdots & & \vdots \\ a_{n-1} & a_{n-2} & \ldots & a_{-1} \end{pmatrix}.$$

If $T_n(a)$ is invertible, then the kernel $\text{Ker} \, \partial T_n(a) \subset \mathbf{C}^{n+1}$ has the dimension 2. Every basis $\{u, v\}$ of $\text{Ker} \, \partial T_n(a)$ is called a fundamental system for $T_n(a)$.

Theorem 7.1 *If $T_n(a)$ is invertible and $u = (u_j)_{j=0}^n$, $v = (v_j)_{j=0}^n$ constitute a fundamental system for $T_n(a)$, then there exists a nonzero constant c such that*

$$\begin{aligned}
T_n^{-1}(a) &= c \begin{pmatrix} u_0 & & \\ \vdots & \ddots & \\ u_{n-1} & \ldots & u_0 \end{pmatrix} \begin{pmatrix} v_n & \ldots & v_1 \\ & \ddots & \vdots \\ & & v_n \end{pmatrix} \\
&\quad -c \begin{pmatrix} v_0 & & \\ \vdots & \ddots & \\ v_{n-1} & \ldots & v_0 \end{pmatrix} \begin{pmatrix} u_n & \ldots & u_1 \\ & \ddots & \vdots \\ & & u_n \end{pmatrix}.
\end{aligned} \tag{17}$$

For example, if x and q are the solutions of

$$T_n(a)x = e_0, \quad T_n(a)q = (\, -a_{-n} \ -a_{-n+1} \ \ldots \ -a_{-1} \,)^\top, \tag{18}$$

with a_{-n} arbitrarily chosen, then $u = \begin{pmatrix} x \\ 0 \end{pmatrix}$ and $v = \begin{pmatrix} q \\ 1 \end{pmatrix}$ form obviously a fundamental system for $T_n(a)$. The corresponding representation (17) is Heinig's formula. Here $c = 1$ (see [22]).

Beginning with results by Ammar and Gader [1], inversion formulas that involve only diagonal matrices and discrete Fourier transforms (DFT's) have received increasing interest since they speed up matrix-vector multiplication. We confine ourselves to one representation of $T_n^{-1}(a)$ that allows matrix-vector multiplication with only 6 DFT's plus $O(n)$ operations. Let $\omega_0, \ldots, \omega_{2n-1}$ be the $2n$th unit roots, $\omega_j = \exp(\pi \mathrm{i} j/n)$. Notice that $\omega_0, \omega_2, \ldots, \omega_{2n-2}$ are the nth roots of 1 and that $\omega_1, \omega_3, \ldots, \omega_{2n-1}$ are the nth roots of -1. Put

$$F_+ = \frac{1}{\sqrt{n}} \left(\omega_{2j}^k \right)_{j,k=0}^{n-1}, \quad F_- = \frac{1}{\sqrt{n}} \left(\omega_{2j+1}^k \right)_{j,k=0}^{n-1}.$$

The matrix F_+ is the usual DFT and $F_- = F_+ \mathrm{diag}\left(\omega_j\right)_{j=0}^{n-1}$. Both F_+ and F_- are unitary. Finally, for $b \in \mathbf{C}^n$, define diagonal matrices $D_\pm(b)$ the kth diagonal element of which is just the kth component of the vector $F_\pm b$.

Theorem 7.2 *Let $T_n(a)$ be invertible and let $\{u,v\}$ be the fundamental system for $T_n(a)$ given by (18). Then*

$$T_n^{-1}(a) = \frac{1}{2} F_-^* \left[D_-(u) F_- F_+^* D_+(v) - D_-(v) F_- F_+^* D_+(u) \right] F_+. \tag{19}$$

Multiplication of an $n \times n$ triangular Toeplitz matrix by a vector can be realized with 2 DFT's of length $2n$, which is approximately 4 DFT's of length n (see, e.g., [45]). Thus, employing formulas (9) or (17) one can multiply $T_n^{-1}(a)$ by a vector with 12 DFT's of length n plus 8 DFT's for preprocessing. Formula (19) does this job with 6 DFT's plus 4 DFT's for preprocessing. It results that by using FFT the computational complexity for matrix-vector multiplication by an inverse Toeplitz matrix is $O(n \log n)$, which is the reason for the attribute "superfast".

The proof of (19) given in [24] allows important generalizations and is based on the translation of representation (17) into polynomial language, which shows the structure of the inverses of Toeplitz (and Hankel) matrices in a new light. It turns out that these are Bezoutians.

8 Bezoutians

The generating function of a matrix $A = (a_{jk})_{j,k=0}^{n-1}$ or a vector $b = (b_j)_{j=0}^{n-1}$ is defined by

$$A(t,s) = \sum_{j,k=0}^{n-1} a_{jk} t^j s^k, \quad b(t) = \sum_{j=0}^{n-1} b_j t^j.$$

An $n \times n$ matrix B is called a Toeplitz Bezoutian if there exist vectors $u, v \in \mathbf{C}^{n+1}$ such that the generating function of B can be represented in the form

$$B(t,s) = \frac{u(t)(J_{n+1} v)(s) - v(t)(J_{n+1} u)(s)}{1 - ts},$$

where J_n is the $n \times n$ counteridentity (that is, the matrix that turns $(z_1, \ldots, z_n)$ into $(z_n, \ldots, z_1)$). In that case B is denoted by $\mathrm{Bez}(u,v)$.

The concept of a Bezoutian was originally introduced by Hermite in order to solve root location problems for polynomials (see, e.g., [40]). As first observed by Lander [43], actually the following is true.

Theorem 8.1 *An invertible matrix is a Toeplitz Bezoutian if and only if it is the inverse of an invertible Toeplitz matrix.*

With the notion of the Toeplitz Bezoutian, formula (17) can simply be written in the form

$$T_n^{-1}(a) = c \operatorname{Bez}(u, v). \tag{20}$$

Furthermore, representation (9) is nothing but (20) with $c = 1/x_0$, $u = \begin{pmatrix} x \\ 0 \end{pmatrix}$, and $v = \begin{pmatrix} 0 \\ J_n y \end{pmatrix}$.

Analogously one can consider Hankel Bezoutians. They have $t - s$ instead of $1 - ts$ in the denominator and $u(t)v(s) - v(t)u(s)$ in the numerator. A matrix of the form $C_n = T_n(a) + H_n(b)$ is called a Toeplitz-plus-Hankel (T+H) matrix. Surprisingly, the structure of the inverse of an invertible T+H matrix closely resembles a Bezoutian. An $n \times n$ matrix is referred to as a T+H Bezoutian if there are eight vectors $u_i, v_i \in \mathbf{C}^{n+2}$ ($i = 1, 2, 3, 4$) such that

$$B(t, s) = \frac{\sum u_i(t)v_i(s)}{(t - s)(1 - ts)}. \tag{21}$$

The following result was established in [23].

Theorem 8.2 *An invertible matrix is a T+H Bezoutian if and only if it is the inverse of an invertible T+H matrix.*

Now let C_n be an invertible T+H matrix. Then the generating function of C_n^{-1} is of the form (21), and the u_i and v_i can again be found as the solutions of T+H systems with special right-hand sides or as bases of the kernels of appropriately modified T+H matrices.

Multiplying the matrix given by the right-hand side of (21) from the left and right by DFT's, we obtain a Cauchy-like matrix. This matrix can in turn also be represented in terms of DFT's. In the end, we get a matrix representation of (21) that involves (besides permutation and diagonal matrices) only 6 DFT's plus 8 DFT's for preprocessing (see [24]). This improves the so far best result [2], which has 7 DFT's plus 10 DFT's for preprocessing.

When working with real matrices, it is desirable to avoid the complex DFT and thus to have representation formulas which contain discrete Hartley or sine and cosine transforms. For more on this subject we refer to [25] and [26].

9   Fast solution of Toeplitz systems

Let $T_n(a)$ be an invertible Toeplitz matrix. We now discuss the problem of solving the linear system

$$T_n(a)f = b, \tag{22}$$

with computational complexity $O(n^2)$ (recall that Gaussian elimination has the complexity $O(n^3)$). Here $b = (b_j)_{j=0}^{n-1}$ is an arbitrary right-hand side.

A first possibility is to use the results of Sections 3 and 7, which amounts to solving (22) by the superfast matrix-vector multiplication $f = T_n^{-1}(a)b$. The question we are left with is how to compute the parameters in the representations for $T_n^{-1}(a)$. In particular, in order to exploit representation (9) we have to solve the two systems $T_n(a)x^{(n)} = e_0$ and $T_n(\widetilde{a})y^{(n)} = e_0$. Let us instead of these two equations consider the pair $T_n(a)x^{(n)} = e_0$ and $T_n(a)z^{(n)} = e_{n-1}$,

where $e_{n-1} = (0, \ldots, 0, 1) \in \mathbf{C}^n$. This change is unessential because $J_n T_n(\widetilde{a}) J_n = T_n(a)$ and hence $y^{(n)} = J_n z^{(n)}$. Our first problem is to find $x^{(n)}$ and $z^{(n)}$ with at most $O(n^2)$ multiplications.

Theorem 3.3 is based on the assumption that $x_0^{(n)} \neq 0$. By Cramer's rule, this is equivalent to the invertibility of the matrix $T_{n-1}(a)$. Our algorithm will be recursive, and therefore we demand that $T_n(a)$ be strongly nonsingular, which means that the leading principal submatrices $T_1(a), T_2(a), \ldots, T_n(a)$ are all invertible. This is in particular the case if the convex hull of the set $a(\mathbf{T})$ does not contain the origin [10].

We consider the equations $T_k(a)x^{(k)} = e_0$ and $T_k(a)z^{(k)} = e_{k-1}$ for $k = 1, \ldots, n$. We start with the equations $T_1(a)x^{(1)} = e_0$, $T_1(a)z^{(1)} = e_0$, which are simply the equations $a_0 x_0^{(1)} = 1$, $a_0 z_0^{(1)} = 1$. The following theorem shows how the columns $x^{(k+1)}$, $z^{(k+1)}$ can be easily computed once the columns $x^{(k)}$, $z^{(k)}$ are available.

Theorem 9.1 *Suppose $T_n(a)$ is strongly nonsingular. Put*

$$\alpha_k = \begin{pmatrix} a_k \ a_{k-1} \ \ldots \ a_1 \end{pmatrix} x^{(k)}, \quad \beta_k = \begin{pmatrix} a_{-1} \ a_{-2} \ \ldots \ a_{-k} \end{pmatrix} z^{(k)}.$$

Then $1 - \alpha_k \beta_k \neq 0$ and

$$x^{(k+1)} = \frac{1}{1 - \alpha_k \beta_k} \left[\begin{pmatrix} x^{(k)} \\ 0 \end{pmatrix} - \alpha_k \begin{pmatrix} 0 \\ z^{(k)} \end{pmatrix} \right], \tag{23}$$

$$z^{(k+1)} = \frac{1}{1 - \alpha_k \beta_k} \left[\begin{pmatrix} 0 \\ z^{(k)} \end{pmatrix} - \beta_k \begin{pmatrix} x^{(k)} \\ 0 \end{pmatrix} \right], \tag{24}$$

for all $k = 1, 2, \ldots, n-1$.

The recursion (23), (24) is known as the Levinson (or Levinson-Durbin) algorithm. The computational complexity of the algorithm resulting from Theorem 9.1 can easily be shown to be $O(n^2)$.

The Levinson recursion (23), (24) offers another possibility for solving (22) with computational complexity $O(n^2)$. This time we look for a direct recursion formula for the solutions $f^{(k)} = (f_j^{(k)})_{j=0}^{k-1}$ of the truncated systems $T_k(a)f^{(k)} = b^{(k)}$ ($k = 1, 2, \ldots, n$), where $b^{(k)} = (b_j)_{j=0}^{k-1}$.

Theorem 9.2 *Let $T_n(a)$ be strongly nonsingular. For $k = 1, 2, \ldots, n-1$, put*

$$\gamma_k = \begin{pmatrix} a_k \ a_{k-1} \ \ldots \ a_1 \end{pmatrix} f^{(k)}.$$

Then

$$f^{(k+1)} = \begin{pmatrix} f^{(k)} \\ 0 \end{pmatrix} + (b_k - \gamma_k) z^{(k+1)}$$

with $z^{(k+1)}$ from (23), (24).

The Levinson recursion is directly related to a UL-factorization of the inverse $T_n^{-1}(a)$,

$$T_n^{-1}(a) = \quad \blacksquare\!\!\blacktriangledown \quad .$$

With this formula the solution f of system (22) can be computed by 2 matrix-vector multiplications with triangular matrices .

A second kind of a recursive procedure is known as the Schur (or Schur-Barreis) algorithm. It computes recurrently certain residual vectors and is related to an LU-factorization of the matrix itself,

$$T_n(a) = \quad \blacktriangle \quad \blacktriangledown \quad .$$

The solution f is given after solving 2 triangular systems by back substitution.

Of course, one wants to remove the strong nonsingularity from the hypotheses of Theorems 9.1 and 9.2. This can indeed be done by appropriately modifying the algorithms. A modification of the Levinson algorithm that works for Toeplitz matrices with arbitrary rank profile was first provided by Heinig [19] (see also [22]). Fast algorithms for arbitrary invertible T+H matrices were first designed in [21].

Starting with Delsarte and Genin [11], [12], one has developed algorithms that employ not only the basic structure of the matrix (the Toeplitz structure, for instance) but also take advantage of additional symmetries, such as the property of being Hermitian, symmetric, skewsymmetric, centro-symmetric, or centro-skewsymmetric. Clearly, the main objective of such algorithms is a further reduction of the computational amount (see e.g. [41, 44, 20, 27]).

First we observe that these symmetries are also reflected in the inversion formulas. As an example let us consider the case of a centro-symmetric T+H matrix C_n, $C_n = J_n C_n J_n$. The inverse of C_n possesses a surprisingly nice structure. It can be represented as the sum of two special T+H Bezoutians B_i ($i = 1, 2$) with the generating functions

$$B_i(t, s) = \frac{u_i(t)v_i(s) - v_i(t)u_i(s)}{(t - s)(1 - ts)} ,$$

where u_1, v_1 are symmetric and u_2, v_2 are skewsymmetric vectors given by the solution of 4 pure Toeplitz equations (see [29] and [30]).

Moreover, in this case the Levinson-type or Schur-type algorithms lead in a natural way to other kinds of factorizations, called WZ- or ZW-factorizations,

$$C_n^{-1} = \quad \blacktriangleright\blacktriangleleft \quad , \qquad C_n = \quad \blacktriangleleft\blacktriangleright \quad ,$$

respectively. The factors in these representations have symmetries which can be exploited to reduce the number of operations. Such factorizations were also obtained for symmetric, skewsymmetric or Hermitian Toeplitz matrices (see [13, 28, 34]). For more details concerning the T+H case we refer to [30, 31].

Recently algorithms have been designed which do both, they work without any restrictions and take advantage of additional symmetries of the matrices (see [32, 33, 35, 36]).

References

[1] G. Ammar and P. Gader, A variant of the Gohberg-Semencul formula involving circulant matrices, SIAM J. Matrix Analysis Appl. **12**, 534–540 (1991).

[2] E. Bozzo, Algebras of higher dimension for displacement decompositions and computations with Toeplitz-plus-Hankel matrices, Linear Algebra Appl. **188/189**, 3–29 (1993).

[3] A. Böttcher, Pseudospectra and singular values of large convolution operators, J. Integral Equations Appl. **6**, 267–301 (1994).

[4] A. Böttcher, C^*-algebras in numerical analysis. Irish Math. Soc. Bull. **45**, 57–133 (2000).

[5] A. Böttcher, M. Embree, and L. N. Trefethen, Piecewise continuous Toeplitz matrices and operators: slow approach to infinity, SIAM J. Matrix Analysis Appl. **24**, 484–489 (2002).

[6] A. Böttcher and S. Grudsky, Asymptotic spectra of dense Toeplitz matrices are unstable, Numerical Algorithms **33**, 105–112 (2003).

[7] A. Böttcher and S. Grudsky, Toeplitz matrices with slowly growing pseudospectra, in: Factorization, Singular Integral Operators, and Related Topics, edited by S. Samko, A. Lebre, and A. F. dos Santos (Kluwer, Dordrecht, 2003), pp. 43-54.

[8] A. Böttcher and S. Grudsky, Structured condition numbers of large Toeplitz matrices are rarely better than usual condition numbers, Numer. Linear Algebra Appl., to appear.

[9] A. Böttcher, S. Grudsky, and E. Ramírez de Arellano, On the asymptotic behavior of the eigenvectors of large banded Toeplitz matrices, Math. Nachr., to appear.

[10] A. Böttcher and B. Silbermann, Introduction to Large Truncated Toeplitz Matrices (Springer, New York, 1999).

[11] P. Delsarte and Y. Genin, The split Levinson algorithm, IEEE Trans. Acoust. Speech Signal Process. **34**, 470–478 (1986).

[12] P. Delsarte and Y. Genin, On the splitting of classical algorithms in linear prediction theory, IEEE Trans. Acoust. Speech Signal Process. **35**, 645–653 (1987).

[13] C. J. Demeure, Bowtie factors of Toeplitz matrices by means of split algorithms, IEEE Trans. Acoust. Speech Signal Process. **37**, 1601–1603 (1989).

[14] I. Gohberg, On an application of the theory of normed rings to singular integral equations, Uspekhi Mat. Nauk **7**, 149–156 (1952) [Russian].

[15] I. Gohberg, On the number of solutions of homogeneous singular integral equations with continuous coefficients, Dokl. Akad. Nauk SSSR **122**, 327–330 (1958) [Russian].

[16] I. Gohberg and I. A. Feldman, Convolution Equations and Projection Methods for Their Solution (Amer. Math. Soc., Providence, R.I., 1974).

[17] I. Gohberg and A. A. Semencul, The inversion of finite Toeplitz matrices and their continual analogues, Matem. Issled. **7**, 201–223 (1972). [Russian].

[18] R. Hagen, S. Roch, and B. Silbermann, C^*-Algebras and Numerical Analysis (Marcel Dekker, New York, 2001).

[19] G. Heinig, Inversion of Toeplitz and Hankel matrices with singular sections, Wiss. Zeitschrift der TH Karl-Marx-Stadt **25**, 326–333 (1983).

[20] G. Heinig, Chebyshev-Hankel matrices and the splitting approach for centrosymmetric Toeplitz-plus-Hankel matrices, Linear Algebra Appl. **327**, 181–196 (2001).

[21] G. Heinig, Inversion of Toeplitz-plus-Hankel matrices with any rank profile, in: Fast Algorithms for Structured Matrices, edited by V. Olshevsky (AMS series Contemp. Math., vol. 323, 2003), pp. 75–90.

[22] G. Heinig and K. Rost, Algebraic Methods for Toeplitz-Like Matrices and Operators (Birkhäuser, Basel, 1984).

[23] G. Heinig and K. Rost, On the inverses of Toeplitz-plus-Hankel matrices, Linear Algebra Appl. **106**, 39–52 (1988).

[24] G. Heinig and K. Rost, DFT representations of Toeplitz-plus-Hankel Bezoutians with application to fast matrix-vector multiplication, Linear Algebra Appl. **284**, 157–175 (1998).

[25] G. Heinig and K. Rost, Hartley transform representations of inverses of real Toeplitz-plus-Hankel matrices, Numerical Functional Analysis and Optimization **21**, 175–189 (2000).

[26] G. Heinig and K. Rost, Efficient inversion formulas for Toeplitz-plus-Hankel matrices using trigonometric transformations, in: Structured Matrices in Mathematics, Computer Science, and Engineering, edited by V. Olshevsky (AMS series Contemp. Math., vol. 281, 2001), pp. 247–264.

[27] G. Heinig and K. Rost, Centro-symmetric and centro-skewsymmetric Toeplitz matrices and Bezoutians, Linear Algebra Appl. **343/344**, 195–203 (2002).

[28] G. Heinig and K. Rost, Fast algorithms for skewsymmetric Toeplitz matrices, in: Toeplitz Matrices and Singular Integral Equations, edited by A. Böttcher, I. Gohberg, P. Junghanns (Birkhäuser, Basel, OT series, vol. 135, 2002), pp. 193–208.

[29] G. Heinig and K. Rost, Centrosymmetric and centro-skewsymmetric Toeplitz-plus-Hankel matrices and Bezoutians, Linear Algebra Appl. **366**, 257–281 (2003).

[30] G. Heinig and K. Rost, Fast algorithms for centro-symmetric and centro-skewsymmetric Toeplitz-plus-Hankel matrices, Numerical Algorithms **33**, 305–317 (2003).

[31] G. Heinig and K. Rost, New fast algorithms for Toeplitz-plus-Hankel matrices, SIAM J. Matrix Analysis Appl. **25**, 842–857 (2004).

[32] G. Heinig and K. Rost, Split algorithms for skewsymmetric Toeplitz matrices with arbitrary rank profile, Theoretical Computer Science **315**, 453–468 (2004).

[33] G. Heinig and K. Rost, Split algorithms for symmetric Toeplitz matrices with arbitrary rank profile, Numerical Linear Algebra with Applications, to appear.

[34] G. Heinig and K. Rost, Schur-type algorithms for the solution of Hermitian Toeplitz systems via factorization, in: Operator Theory: Advances and Applications, to appear.

[35] G. Heinig and K. Rost, Split algorithms for hermitian Toeplitz matrices with arbitrary rank profile, Linear Algebra Appl., to appear.

[36] G. Heinig and K. Rost, Split algorithms for centrosymmetric Toeplitz-plus-Hankel matrices with arbitrary rank profile, in preparation.

[37] N. J. Higham, Accuracy and Stability of Numerical Algorithms (SIAM, Philadelphia, 1996).

[38] S. V. Konyagin and W. Schlag, Lower bounds for the absolute value of random polynomials on a neighborhood of the unit circle, Trans. Amer. Math. Soc. **351**, 4963–4980 (1999).

[39] M. G. Krein, Integral equations on the half-line with a kernel depending on the difference of the arguments, Uspekhi Mat. Nauk **13**, 3–120 (1958) [Russian].

[40] M. Krein and M. Naimark, The method of symmetric and Hermitian forms in the theory of the separation of the roots of algebraic equations, Linear and Multilinear Algebra **56**, 69–87 (1974).

[41] B. Krishna and H. Krishna, Computationally efficient reduced polynomial based algorithms for hermitian Toeplitz matrices, SIAM J. Appl. Math. **49**, 1275–1282 (1989).

[42] H. Landau, On Szegö's eigenvalue distribution theorem and non-Hermitian kernels, J. Analyse Math. **28**, 335–357 (1975).

[43] F. I. Lander, The Bezoutian and the inversion of Hankel and Toeplitz matrices, Matem. Issled. **9**, 69–87 (1974) [Russian].

[44] A. Melman, A two-step even-odd split Levinson algorithm for Toeplitz systems, Linear Algebra Appl. **338**, 219–237 (2001).

[45] V.Y. Pan, Structured Matrices and Polynomials (Birkhäuser, Boston and Springer, New York, 2001).

[46] L. Reichel and L. N. Trefethen, Eigenvalues and pseudo-eigenvalues of Toeplitz matrices, Linear Algebra Appl. **162/164**, 153–185 (1992).

[47] S. M. Rump, Structured perturbations, Part I: Normwise distances, SIAM J. Matrix Anal. Appl. **25**, 1–30 (2003).

[48] P. Schmidt and F. Spitzer, The Toeplitz matrices of an arbitrary Laurent polynomial, Math. Scand. **8**, 15–38 (1960).

[49] L. N. Trefethen and M. Embree, Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators (Princeton University Press, to appear).

[50] H. Widom, Asymptotic behavior of block Toeplitz matrices and determinants II, Advances in Math. **21**, 1–29 (1976).