Bounds on the Singular Values of Matrices with Displacement Structure

Bernhard Beckermann, Alex Townsend

SIAM REVIEW
Vol. 61, No. 2, pp. 000–000

Bounds on the Singular Values of Matrices with Displacement Structure*

Bernhard Beckermann†
Alex Townsend‡

Abstract. Matrices with displacement structure, such as Pick, Vandermonde, and Hankel matrices, appear in a diverse range of applications. In this paper, we use an extremal problem involving rational functions to derive explicit bounds on the singular values of such matrices. For example, we show that the kth singular value of a real $n \times n$ positive definite Hankel matrix, H_n, is bounded by $C\rho^{-k/\log n}\|H_n\|_2$ with explicitly given constants $C > 0$ and $\rho > 1$, where $\|H_n\|_2$ is the spectral norm. This means that a real $n \times n$ positive definite Hankel matrix can be approximated, up to an accuracy of $\epsilon\|H_n\|_2$ with $0 < \epsilon < 1$, by a rank $\mathcal{O}(\log n \log(1/\epsilon))$ matrix. Analogous results are obtained for Pick, Cauchy, real Vandermonde, Löwner, and certain Krylov matrices.

Key words. singular values, displacement structure, Zolotarev, rational

AMS subject classifications. 15A18, 26C15

DOI. 10.1137/19M1244433

1. Introduction.

Matrices with rapidly decaying singular values frequently appear in computational mathematics. Such matrices are numerically of low rank, and this is exploited in applications such as particle simulations [33], model reduction [2], boundary element methods [35], and matrix completion [20]. However, it can be theoretically challenging to fully explain why low rank techniques are so effective in practice. In this paper, we derive explicit bounds on the singular values of matrices with displacement structure and in doing so justify many of the low rank techniques that are being employed on such matrices.

Let $X \in \mathbb{C}^{m\times n}$ with $m \geq n$, $A \in \mathbb{C}^{m\times m}$, and $B \in \mathbb{C}^{n\times n}$. We say that X has an (A, B)-displacement rank of ν if X satisfies the Sylvester matrix equation given by

$$AX - XB = MN^* \tag{1.1}$$

for some matrices $M \in \mathbb{C}^{m\times\nu}$ and $N \in \mathbb{C}^{n\times\nu}$. Matrices with displacement structure include Toeplitz ($\nu = 2$), Hankel ($\nu = 2$), Cauchy ($\nu = 1$), Krylov ($\nu = 1$), and Vandermonde ($\nu = 1$) matrices, as well as Pick ($\nu = 2$), Sylvester ($\nu = 2$), and Löwner ($\nu = 2$) matrices. Fast algorithms for computing matrix-vector products and for

*Published electronically May 8, 2019. This paper originally appeared in *SIAM Journal on Matrix Analysis and Applications*, Volume 38, Number 4, 2017, pages 1227–1248, under the title "On the Singular Values of Matrices with Displacement Structure."
http://www.siam.org/journals/sirev/61-2/M124443.html
Funding: The work of the first author was supported in part by the Labex CEMPI (ANR-11-LABX-0007-01). The work of the second author was supported by National Science Foundation grant 1818757.

†Laboratoire Paul Painlevé UMR 8524 CNRS, Equipe ANO-EDP, UFR Mathématiques, UST Lille, F-59655 Villeneuve d'Ascq CEDEX, France (Bernhard.Beckermann@univ-lille1.fr).
‡Department of Mathematics, Cornell University, Ithaca, NY 14853 (townsend@cornell.edu).

Table 1.1 *Summary of the bounds proved on the singular values of matrices with displacement structure. For the singular value bounds to be valid for $C_{m,n}$ and L_n, mild "separation conditions" must hold (see section 4). The numbers ρ_j and C_j for $j = 1, \ldots, 6$ are given explicitly in their corresponding sections.*

Matrix class	Notation	Singular value bound	Ref.
Pick	P_n	$\sigma_{1+2k}(P_n) \leq C_1\rho_1^{-k}\|P_n\|_2$	sec. 4.1
Cauchy	$C_{m,n}$	$\sigma_{1+k}(C_{m,n}) \leq C_2\rho_2^{-k}\|C_{m,n}\|_2$	sec. 4.2
Löwner	L_n	$\sigma_{1+2k}(L_n) \leq C_3\rho_3^{-k}\|L_n\|_2$	sec. 4.3
Krylov, Herm. arg.	$K_{m,n}$	$\sigma_{1+2k}(K_{m,n}) \leq C_4\rho_4^{-k/\log n}\|K_{m,n}\|_2$	sec. 5.1
real Vandermonde	$V_{m,n}$	$\sigma_{1+2k}(V_{m,n}) \leq C_5\rho_5^{-k/\log n}\|V_{m,n}\|_2$	sec. 5.1
pos. semidef. Hankel	H_n	$\sigma_{1+2k}(H_n) \leq C_6\rho_6^{-k/\log n}\|H_n\|_2$	sec. 5.2

solving systems of linear equations can be derived for many of these matrices by exploiting (1.1) [36, 42].

In this paper, we use the displacement structure to derive explicit bounds on the singular values of matrices that satisfy (1.1) by using an extremal problem for rational functions from complex approximation theory. In particular, we prove that the following inequality holds (see Theorem 2.1):

$$\sigma_{j+\nu k}(X) \leq Z_k(E,F)\sigma_j(X), \qquad 1 \leq j + \nu k \leq n, \tag{1.2}$$

where $\sigma_1(X), \ldots, \sigma_n(X)$ denote the singular values of X and $Z_k(E,F)$ is the Zolotarev number for complex sets E and F that depend on A and B (see (1.4)). Researchers have previously exploited the connection between the Sylvester matrix equation and Zolotarev numbers for selecting algorithmic parameters in the alternating direction implicit (ADI) method [8, 15, 38], and others have demonstrated that the singular values of matrices satisfying certain Sylvester matrix equations have rapidly decaying singular values [2, 4, 48]. Here, we derive explicit bounds on all the singular values of structured matrices. Table 1.1 summarizes our main singular value bounds.

Not every matrix with displacement structure is numerically of low rank. For example, the identity matrix is a full rank Toeplitz matrix and the exchange matrix[1] is a full rank Hankel matrix. Moreover, we show in Example 5.1 that the inequality in (1.2) is trivial for circulant as well as Toeplitz matrices. The properties of A and B in (1.1) are crucial. If A and B are normal matrices, then one expects X to be numerically of low rank only if the eigenvalues of A and B are well separated (see Theorem 2.1). If A and B are both not normal, then descriptive bounds on the numerical rank of X are more subtle (see Corollary 2.2, [4], and [48, Lemma 1]).

By the Eckart–Young theorem [30, Theorem 2.4.8], singular values measure the distance in the spectral norm from X to the set of matrices of a given rank, i.e.,

$$\sigma_j(X) = \min\left\{\|X - Y\|_2 : Y \in \mathbb{C}^{m\times n},\ \operatorname{rank}(Y) = j-1\right\}.$$

For an $0 < \epsilon < 1$, we say that the *ϵ-rank* of a matrix X is k if k is the smallest integer such that $\sigma_{k+1}(X) \leq \epsilon\|X\|_2$. That is,

$$\operatorname{rank}_\epsilon(X) = \min_{k\geq 0}\left\{k : \sigma_{k+1}(X) \leq \epsilon\|X\|_2\right\}. \tag{1.3}$$

[1] The $n \times n$ exchange matrix X is obtained by reversing the order of the rows of the $n \times n$ identity matrix, i.e., $X_{n-j+1,j} = 1$ for $1 \leq j \leq n$.

Table 1.2 *Summary of the upper bounds proved on the ϵ-rank of matrices with displacement structure. For the bounds above to be valid for $C_{m,n}$ and L_n, mild "separation conditions" must hold (see section 4). The number γ is the absolute value of the cross-ratio of a, b, c, and d; see (3.7). The first three rows show an ϵ-rank of at most $\mathcal{O}(\log\gamma\log(1/\epsilon))$, and the last three rows show an ϵ-rank of at most $\mathcal{O}(\log n\log(1/\epsilon))$.*

Matrix class	Notation	Upper bound on $\mathrm{rank}_\epsilon(X)$	Ref.
Pick	P_n	$2\lceil\log(4b/a)\log(4/\epsilon)/\pi^2\rceil$	sec. 4.1
Cauchy	$C_{m,n}$	$\lceil\log(16\gamma)\log(4/\epsilon)/\pi^2\rceil$	sec. 4.2
Löwner	L_n	$2\lceil\log(16\gamma)\log(4/\epsilon)/\pi^2\rceil$	sec. 4.3
Krylov, Herm. arg.	$K_{m,n}$	$2\lceil 4\log(8\lfloor n/2\rfloor/\pi)\log(4/\epsilon)/\pi^2\rceil+2$	sec. 5.1
real Vandermonde	$V_{m,n}$	$2\lceil 4\log(8\lfloor n/2\rfloor/\pi)\log(4/\epsilon)/\pi^2\rceil+2$	sec. 5.1
pos. semidef. Hankel	H_n	$2\lceil 2\log(8\lfloor n/2\rfloor/\pi)\log(16/\epsilon)/\pi^2\rceil+2$	sec. 5.2

Thus, we may approximate X to a precision of $\epsilon\|X\|_2$ by a rank $k=\mathrm{rank}_\epsilon(X)$ matrix.

An immediate consequence of explicit bounds on the singular values of certain matrices is a bound on the ϵ-rank. Table 1.2 summarizes our main upper bounds on the ϵ-rank of matrices with displacement structure. The form of the inequalities in (1.2) also allows one to use Zolotarev numbers to bound the ϵ-rank of matrices when measured in the Frobenius norm [51, Lemma 5.1], which is a key observation to extending the bounds in this paper to tensors [51].

Zolotarev numbers have already proved useful for deriving tight bounds on the condition number of matrices with displacement structure [5, 6], where the condition number of a rectangular $m\times n$ matrix X is given by $\kappa_2(X)=\sigma_1(X)/\sigma_{\min(m,n)}(X)$. For example, the first author proved that a real $n\times n$ positive definite Hankel matrix, H_n, with $n\geq 3$, is exponentially ill-conditioned [6]. That is,

$$\kappa_2(H_n)\geq\frac{\gamma^{n-1}}{16n},\qquad \gamma\approx 3.210,$$

and that this bound cannot be improved by more than a factor of n times a modest constant. The Hilbert matrix given by $(H_n)_{jk}=1/(j+k-1)$, for $1\leq j,k\leq n$, is the classic example of an exponentially ill-conditioned positive definite Hankel matrix [65, eqn. (3.35)]. Similar exponential ill-conditioning has been shown for certain Krylov matrices and real Vandermonde matrices [6].

This paper extends the application of Zolotarev numbers to deriving bounds on the singular values of matrices with displacement structure—not just the condition number. The bounds we derive are particularly tight for $\sigma_j(X)$, where j is small with respect to n. Improved bounds on $\sigma_j(X)$ when $j/n\to c\in(0,1)$ may be possible with the ideas found in [9]. Nevertheless, our interest here is to justify the application of low rank techniques on matrices with displacement structure by proving that such matrices are often well approximated by low rank matrices. The bounds that we derive are sufficient for this purpose.

For an integer k, let $\mathcal{R}_{k,k}$ denote the set of irreducible rational functions of the form $p(x)/q(x)$, where p and q are polynomials of degree at most k. Given two closed disjoint sets $E,F\subset\mathbb{C}$, the corresponding Zolotarev number, $Z_k(E,F)$, is defined by

$$Z_k(E,F):=\inf_{r\in\mathcal{R}_{k,k}}\frac{\sup_{z\in E}|r(z)|}{\inf_{z\in F}|r(z)|},\tag{1.4}$$

where the infimum is attained for some extremal rational function. As a general rule, the number $Z_k(E,F)$ decreases rapidly to zero with k if E and F are sets that are disjoint and well separated. Zolotarev numbers satisfy several immediate properties: for any sets E and F and integers k, k_1, and k_2, one has $Z_0(E,F)=1$, $Z_k(E,F)=Z_k(F,E)$, $Z_{k+1}(E,F)\leq Z_k(E,F)$, and $Z_{k_1+k_2}(E,F)\leq Z_{k_1}(E,F)Z_{k_2}(E,F)$. They also satisfy $Z_k(E_1,F_1)\leq Z_k(E_2,F_2)$ if $E_1\subseteq E_2$ and $F_1\subseteq F_2$ as well as $Z_k(E,F)=Z_k(T(E),T(F))$, where T is any Möbius transformation [1]. As $k\to\infty$ the value for $Z_k(E,F)$ is known asymptotically to be

$$\lim_{k\to\infty}(Z_k(E,F))^{1/k}=\exp\left(-\frac{1}{\mathrm{cap}(E,F)}\right),$$

where $\mathrm{cap}(E,F)$ is the logarithmic capacity of a condenser with plates E and F; see [31] or [50, Theorem VIII.3.5]. A lower bound is also known [50, Theorem VIII.3.1]:

$$Z_k(E,F)\geq\exp\left(-\frac{k}{\mathrm{cap}(E,F)}\right),\qquad k\geq 0. \tag{1.5}$$

To readers who are not familiar with Zolotarev numbers, it may seem that (1.2) trades a difficult task of directly bounding the singular values of a matrix X with a more abstract task of understanding the behavior of $Z_k(E,F)$. However, Zolotarev numbers have been extensively studied in the literature [1, 31, 66], and for certain sets E and F the extremal rational function is known explicitly [1, section 50] and [53] (see section 3). Our major challenge for bounding singular values is to carefully select sets E and F so that one can use complex analysis and Möbius transformations to convert the associated extremal rational approximation problem in (1.4) into one that has an explicitly known bound.

The paper is structured as follows. In section 2 we prove (1.2), giving us a bound on the singular values of matrices with displacement structure in terms of Zolotarev numbers. In section 3 we derive new sharper bounds on $Z_k([-b,-a],[a,b])$ when $0<a<b<\infty$ by correcting an infinite product formula from Lebedev (see Theorem 3.1 and Corollary 3.2). In section 4 we derive explicit bounds on the singular values of Pick, Cauchy, and Löwner matrices. In section 5 we tackle the challenging task of showing that all real Vandermonde and positive definite Hankel matrices have rapidly decaying singular values and can be approximated, up to an accuracy of $0<\epsilon<1$, by a rank $\mathcal{O}(\log n\log(1/\epsilon))$ matrix. While no additional novel results are included in this updated version of [11], we have added extra commentary on the related literature and reference recent applications.

2. The Singular Values of Matrices with Displacement Structure and Zolotarev Numbers.

Let X be an $m\times n$ matrix with $m\geq n$ that satisfies (1.1). We show that the singular values of X can be bounded from above in terms of Zolotarev numbers. First, we assume that A and B in (1.1) are normal matrices and later remove this assumption in Corollary 2.2. In Theorem 2.1 the spectrum (set of eigenvalues) of A and B is denoted by $\sigma(A)$ and $\sigma(B)$, respectively.[2]

THEOREM 2.1. *Let $A\in\mathbb{C}^{m\times m}$ and $B\in\mathbb{C}^{n\times n}$ be normal matrices with $m\geq n$, and let E and F be complex sets such that $\sigma(A)\subseteq E$ and $\sigma(B)\subseteq F$. Suppose that the matrix $X\in\mathbb{C}^{m\times n}$ satisfies*

$$AX-XB=MN^*,\qquad M\in\mathbb{C}^{m\times\nu},\quad N\in\mathbb{C}^{n\times\nu},$$

[2] The statement of Theorem 2.1 was presented by the first author at the Cortona meeting on Structured Numerical Linear Algebra in 2004 [7] as well as several other locations. Similar statements based on the presentation have appeared in [49, Theorem 2.1.1], [52, Theorem 4], and [14, Theorem 4.2].

where $1 \leq \nu \leq n$ is an integer. Then, for $j \geq 1$, the singular values of X satisfy

$$\sigma_{j+\nu k}(X) \leq Z_k(E,F)\,\sigma_j(X), \qquad 1 \leq j+\nu k \leq n,$$

where $Z_k(E,F)$ is the Zolotarev number in (1.4).

Proof. Let $p(z)$ and $q(z)$ be polynomials of degree at most k. First, we show that

$$\operatorname{rank}(p(A)Xq(B) - q(A)Xp(B)) \leq \nu k, \qquad \nu = \operatorname{rank}(AX - XB). \tag{2.1}$$

Suppose that $p(z) = z^s$ and $q(z) = z^t$, where $k \geq s \geq t$. Then

$$\begin{aligned} p(A)Xq(B) - q(A)Xp(B) &= A^t\left(A^{s-t}X - XB^{s-t}\right)B^t \\ &= \sum_{j=0}^{s-t-1} A^{t+j}(AX - XB)B^{s-1-j} \\ &= \sum_{j=0}^{s-t-1} \left(A^{t+j}M\right)\left(N^*B^{s-1-j}\right). \end{aligned}$$

In the last sum we have terms of the form $(A^\ell M)(N^*B^\wp)$, with $0 \leq \ell, \wp \leq k-1$. By adding together the terms occurring in $p(A)Xq(B) - q(A)Xp(B)$ for general degree k polynomials p and q, we conclude that there exist coefficients $c_{\ell,\wp} \in \mathbb{C}$ such that

$$p(A)Xq(B) - q(A)Xp(B) = \sum_{\ell,\wp=0}^{k-1} c_{\ell,\wp}\left(A^\ell M\right)(N^*B^\wp).$$

This shows that the rank of $p(A)Xq(B) - q(A)Xp(B)$ is bounded above by k times the number of columns of M, proving (2.1).

Now, let $r(z) = p(z)/q(z)$, where p and q are polynomials of degree k so that $r(z)$ is the extremal rational function for the Zolotarev number in (1.4). This means that $p(z)$ and $q(z)$ are not zero on F and E, respectively, so that $p(B)$ and $q(A)$ are invertible matrices. From (2.1) we know that $\Delta = p(A)Xq(B) - q(A)Xp(B)$ has rank at most νk, and hence the matrix

$$Y = -q(A)^{-1}\Delta p(B)^{-1} = X - r(A)Xr(B)^{-1}$$

is of rank at most νk. Let X_j be the best rank $j-1$ approximation to X in $\|\cdot\|_2$, and let $Y_{j-1} = r(A)X_{j-1}r(B)^{-1}$. Since Y_{j-1} is of rank at most $j-1$, $Y + Y_{j-1}$ is of rank at most $j + \nu k - 1$. This implies that

$$\begin{aligned} \sigma_{j+\nu k}(X) &\leq \|X - Y - Y_{j-1}\|_2 \\ &= \left\|r(A)(X - X_{j-1})r(B)^{-1}\right\|_2 \\ &\leq \|r(A)\|_2\left\|r(B)^{-1}\right\|_2\sigma_j(X), \end{aligned}$$

where in the last inequality we used the relation $\sigma_j(X) = \|X - X_{j-1}\|_2$. Finally, since A and B are normal we have $\|r(A)\|_2 = \sup_{z\in\sigma(A)}|r(z)|$ and $\|r(B)^{-1}\|_2 = \sup_{z\in\sigma(B)}|r(z)^{-1}|$. We conclude by the definition of $r(z)$ that

$$\frac{\sigma_{j+\nu k}(X)}{\sigma_j(X)} \leq \sup_{z\in\sigma(A)}|r(z)|\sup_{z\in\sigma(B)}\frac{1}{|r(z)|} = Z_k(\sigma(A),\sigma(B)) \leq Z_k(E,F), \tag{2.2}$$

as required. ∎

Theorem 2.1 shows that if A and B are normal matrices in (1.1), then the singular values decay at least as quickly as $Z_k(\sigma(A),\sigma(B))$ in (1.4). In particular, when $\sigma(A)$ and $\sigma(B)$ are disjoint and well separated we expect $Z_k(\sigma(A),\sigma(B))$ to decay rapidly to zero, and hence so do the singular values of X. The singular value bound in Theorem 2.1 is most useful when $\nu \leq 5$ as it tends to not be tight for large ν. An extension of Theorem 2.1 has derived useful bounds when ν is large, provided that MN^* has rapidly decaying singular values [57]. It is also known that X can have off-diagonal low rank structure if MN^* does [39].

For those readers who are familiar with the ADI method [15], an analogous proof of Theorem 2.1 is to run the ADI method for k steps with shift parameters given by the zeros and poles of the extremal rational function for $Z_k(E,F)$. By doing this, one constructs a rank νk approximant $X_{\nu k}$ for X, which shows that $\sigma_{1+\nu k}(X) \leq \|X - X_{\nu k}\|_2 \leq Z_k(E,F)\sigma_1(X)$. The so-called factored ADI method is a modification of the ADI method that computes $X_{\nu k}$ in low rank form [12, 63] and means that it is computationally possible (and usually efficient) to construct low rank approximants that attain all our singular value bounds in this paper. The connection between Zolotarev numbers and the optimal parameter selection for the ADI method has been previously exploited [38].

For matrices A and B that are not normal, Theorem 2.1 can be extended by using K-spectral sets [3]. Given a matrix A, a complex set E is said to be a K-spectral set for A if the spectrum $\sigma(A)$ of A is contained in E and the inequality $\|r(A)\|_2 \leq K\|r\|_E$ holds for every bounded rational function on E, where $\|r\|_E = \sup_{z\in E}|r(z)|$. Similar extensions have been noted when $B = A^*$ in (1.1) and the sets E and F are taken to be the fields of values[3] for A and B, respectively [4].

We have the following extension of Theorem 2.1.

COROLLARY 2.2. *Suppose that the assumptions of Theorem* 2.1 *hold, except that the matrices A and B are not necessarily normal. Also suppose that E and F are K-spectral sets for A and B for fixed constants $K_A, K_B > 0$, respectively. Then we have $\sigma_{j+\nu k}(X) \leq K_A K_B Z_k(E,F)\,\sigma_j(X)$.*

Proof. It is only the first inequality in (2.2) of the proof of Theorem 2.1 that requires A and B to be normal matrices. When A is not a normal matrix, the equality $\|r(A)\|_2 = \sup_{z\in\sigma(A)}|r(z)|$ may not hold. Instead, we replace it by the K-spectral set bounds given by $\|r(A)\|_2 \leq K_A\|r\|_E$ and $\|r(B)\|_2 \leq K_B\|r\|_F$. Note that since $p(z)$ and $q(x)$ are not zero on F and E, respectively, one can show via the Schur decomposition that $p(B)$ and $q(A)$ are invertible matrices. $\square$

It is worth noting that the bound in Corollary 2.2 is not always descriptive when A and B are nonnormal matrices [4]. There is also the remarkable result in [48, Lemma 1] that shows that for any prescribed eigenvalue distribution and monotonically decaying sequence $s_1 \geq \cdots \geq s_n \geq 0$, there exists a (typically nonnormal) matrix A with the prescribed eigenvalues and a vector $\underline{b}$ such that the singular values of X satisfying $AX + XA^* = -\underline{b}\,\underline{b}^*$ are $\sigma_j(X) = s_j$.

Theorem 2.1 and Corollary 2.2 provide bounds on the singular values of X in terms of Zolotarev numbers. Therefore, to derive analytic bounds on the singular values of matrices with displacement structure, we must now calculate explicit bounds on Zolotarev numbers—a topic that fortunately is extensively studied.

[3]Crouzeix and Palencia have shown that there exists a universal constant $2 \leq K \leq 1+\sqrt{2}$ such that the field of values E of A is a K-spectral set [21, 22].

3. Zolotarev Numbers. In this section, we give explicit lower and upper bounds for Zolotarev numbers for real intervals and disks. Explicit bounds on $Z_k(E,F)$ are known for more general sets E and F, but the bounds tend to not to be tight [27].

3.1. Zolotarev Numbers for Real Symmetric Intervals. Let $0 < a < b < \infty$ and consider the Zolotarev number $Z_k := Z_k([-b,-a],[a,b])$. The sharpest bounds that we are aware of in the literature take the form

$$\rho^{-2k} \leq Z_k \leq 16\,\rho^{-2k}; \tag{3.1}$$

see (1.5) or [31, Theorem 1] for the lower bound and [18, eqn. (2.3)] for the upper bound.[4] There are also bounds obtained directly from an infinite product formula for $\sqrt{Z_k}$ [38, (1.11)]; unfortunately, the original product formula in [38, (1.11)] contains typos and one must be careful, as the typo has been copied elsewhere.

The value of ρ in (3.1) is related to the logarithmic capacity of a condenser with plates $[-b,-a]$ and $[a,b]$:

$$\rho^2 = \exp\left(\frac{1}{\mathrm{cap}([-b,-a],[a,b])}\right), \qquad \rho = \exp\left(\frac{\pi^2}{2\mu(a/b)}\right) > 1, \tag{3.2}$$

where $\mu(\lambda) = \frac{\pi}{2}K(\sqrt{1-\lambda^2})/K(\lambda)$ is the Grötzsch ring function, and K is the complete elliptic integral of the first kind [44, (19.2.8)]:

$$K(\lambda) = \int_0^1 \frac{1}{\sqrt{(1-t^2)(1-\lambda^2 t^2)}}dt, \qquad 0 \leq \lambda \leq 1.$$

The bounds in (3.1) are not asymptotically sharp, and in Corollary 3.2 we show that the constant of 16 in the upper bound can be replaced by 4. For a proof of this sharper upper bound, we first return to the work of Lebedev [38] and derive a corrected infinite product formula for Z_k.

THEOREM 3.1. *Let $k \geq 1$ be an integer and $0 < a < b < \infty$. Then for $Z_k := Z_k([-b,-a],[a,b])$ we have*

$$Z_k = 4\rho^{-2k}\prod_{\tau=1}^{\infty}\frac{(1+\rho^{-8\tau k})^4}{(1+\rho^{4k}\rho^{-8\tau k})^4}, \qquad \rho = \exp\left(\frac{\pi^2}{2\mu(a/b)}\right),$$

where $\mu(\cdot)$ is the Grötzsch ring function.

Proof. We start by establishing a product formula for the inverse of μ that is apparently not widely known. For $\kappa \in (0,1)$ set $q = \exp(-2\mu(\kappa))$. Since $\mu(\kappa) = \frac{\pi}{2}K\left(\sqrt{1-\kappa^2}\right)/K(\kappa)$, we have that $q = \exp(-\pi K\left(\sqrt{1-\kappa^2}\right)/K(\kappa))$, and from [44, (22.2.2)] we obtain

$$\kappa = \left(\frac{\theta_2(0,q)}{\theta_3(0,q)}\right)^2 = 4\sqrt{q}\prod_{\tau=1}^{\infty}\frac{(1+q^{2\tau})^4}{(1+q^{2\tau-1})^4}, \quad q = q(\kappa) = \exp(-2\mu(\kappa))\,. \tag{3.3}$$

Here, $\theta_2(z,q)$ and $\theta_3(z,q)$ are the classical theta functions [44, (20.2.2) and (20.2.3)], and the second equality for κ in (3.3) is derived from the infinite product formula in [44, (20.4.3) and (20.4.4)].

[4] See also [17, eqn. (A1)] and [10, proof of Theorem 6.6] for the related problem of minimal Blaschke products, and see [16, Theorem V.5.5] for how to deal with rational functions with different degree constraints.

In order to deduce an explicit product formula for Z_k, we first note that the value of $2\sqrt{Z_k}/(1+Z_k)$ is extensively reviewed by Akhiezer;[5] see [1, section 51], [1, Tables 1 and 2, p. 150, no. 7 and 8], and [1, Table XXIII]. This value is equal to

$$\frac{2\sqrt{Z_k}}{1+Z_k} = \frac{1-\lambda_k}{1+\lambda_k}, \qquad k\mu(\lambda_k) = \mu(a/b),$$

for some $\lambda_k \in (0,1)$ [1, p. 149]. Here, there is a unique $\lambda_k \in (0,1)$ since the Grötzsch ring function $\mu : [0,1] \to [0,\infty]$ is a strictly decreasing bijection. Next, we recall that Gauss's transformation [1, Table XXI] and Landen's transformation [1, Table XX] are given by

$$\mu\left(\frac{2\sqrt{\lambda}}{1+\lambda}\right) = \frac{\mu(\lambda)}{2}, \qquad \mu\left(\frac{1-\lambda}{1+\lambda}\right) = 2\mu\left(\sqrt{1-\lambda^2}\right), \qquad \lambda \in (0,1), \tag{3.4}$$

from which we conclude that

$$\mu(Z_k) = 2\mu\left(\frac{2\sqrt{Z_k}}{1+Z_k}\right) = 4\mu\left(\sqrt{1-\lambda_k^2}\right) = \frac{\pi^2}{\mu(\lambda_k)} = \frac{\pi^2 k}{\mu(a/b)}. \tag{3.5}$$

Therefore, from (3.5) we have

$$q = q(Z_k) = e^{-2\mu(Z_k)} = \exp\left(-2k\frac{\pi^2}{\mu(a/b)}\right) = \rho^{-4k},$$

where ρ is given in (3.2). The infinite product formula for Z_k follows by setting $\kappa = Z_k$ and $q = \rho^{-4k}$ in (3.3). $\square$

The infinite product in Theorem 3.1 can be estimated by observing that $(1+\rho^{-4k}\rho^{-8\tau k})^2 \leq (1+\rho^{-8\tau k})^2 \leq (1+\rho^{-4k}\rho^{-8\tau k})(1+\rho^{4k}\rho^{-8\tau k})$ for all $\tau \geq 1$. This leads to the following simple upper and lower bounds which are sufficient for the purpose of our paper.

COROLLARY 3.2. *Let $k \geq 1$ be an integer and $0 < a < b < \infty$. Then for $Z_k := Z_k([-b,-a],[a,b])$ we have*

$$\frac{4\rho^{-2k}}{(1+\rho^{-4k})^4} \leq Z_k \leq \frac{4\rho^{-2k}}{(1+\rho^{-4k})^2} \leq 4\rho^{-2k}, \qquad \rho = \exp\left(\frac{\pi^2}{2\mu(a/b)}\right),$$

where $\mu(\cdot)$ is the Grötzsch ring function.

Corollary 3.2 shows that $Z_k \leq 4\rho^{-2k}$ is an asymptotically sharp upper bound in the sense that the geometric decay rate and the constant 4 cannot be improved if one hopes for the bound to hold for all k. However, this does not necessarily imply that our derived singular value inequalities are asymptotically sharp. On the contrary, they are usually not. For asymptotically sharp singular value bounds, we expect that one must consider discrete Zolotarev numbers, i.e., $Z_k(\sigma(A),\sigma(B))$ in Theorem 2.1, which are more subtle to bound and are outside the scope of this paper.

We often prefer the following slightly weaker bound that does not contain the Grötzsch ring function:

$$Z_k([-b,-a],[a,b]) \leq 4\left[\exp\left(\frac{\pi^2}{2\log(4b/a)}\right)\right]^{-2k}, \qquad 0 < a < b < \infty,$$

[5]There is a typo in [1, Tables 1 and 2, p. 150, no. 7 and 8]. There should be no prime on λ_1.

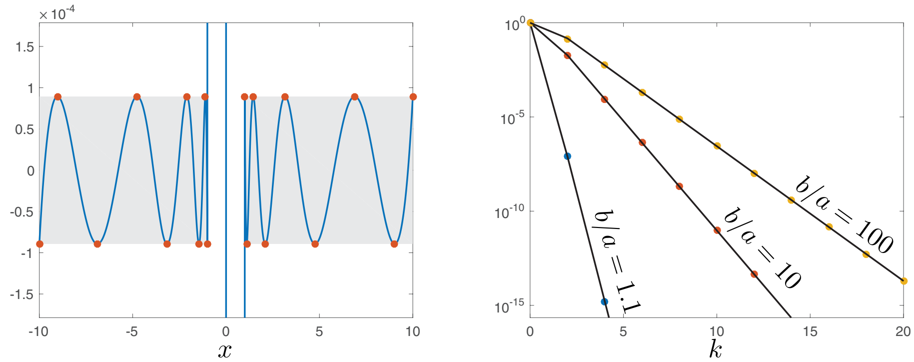

Fig. 3.1 *Zolotarev's rational approximations. Left: The error between the sign function on* $[-10, -1] \cup [1, 10]$ *and its best rational* $\mathcal{R}_{8,8}$ *approximation on the domain* $[-10, 10]$*. The error equioscillates* 9 *times in the interval* $[-10, -1]$ *and* $[1, 10]$ *(see red dots), verifying its optimality* [1, p. 149]*. Right: The upper bound (black line) on* $Z_k([-b, -a], [a, b])$ *(colored dots) in Corollary* 3.2 *for* $0 \leq k \leq 20$*, with* $b/a = 1.1$ *(blue),* 10 *(red),* 100 *(yellow).*

which is obtained by using the bound $\mu(\lambda) \leq \log(2(1 + \sqrt{1 - \lambda^2})/\lambda) \leq \log(4/\lambda)$; see [44, (19.9.5)]. This makes our final bounds on the singular values and the ϵ-rank of matrices with displacement rank more intuitive to those readers who are less familiar with the Grötzsch ring function.

Figure 3.1 (left) shows the error between the sign function on $[-10, -1] \cup [1, 10]$ and its best $\mathcal{R}_{8,8}$ rational approximation, which equioscillates 9 times on $[-10, -1]$ and $[1, 10]$ to confirm its optimality.

3.1.1. Properties of Extremal Rational Functions for $Z_k([-b, -a], [a, b])$. Later, in section 5 we will need to use properties of an extremal rational function for $Z_k = Z_k([-b, -a], [a, b])$, and we prove them now. Zolotarev [66] studied the value Z_k and gave an explicit expression for the extremal function for Z_k (see (1.4)) by showing an equivalence to the problem of best rational approximation of the sign function on $[-b, -a] \cup [a, b]$. We now repeat this to derive the desired properties of the extremal rational function.

Theorem 3.3. *Let $k \geq 1$ be an integer and $0 < a < b < \infty$. There exists an extremal function $R \in \mathcal{R}_{k,k}$ for $Z_k = Z_k([-b, -a], [a, b])$ such that*

(a) *for $z \in [-b, -a]$, we have $-\sqrt{Z_k} \leq R(z) \leq \sqrt{Z_k}$,*
(b) *for $z \in \mathbb{C}$, we have $R(-z) = 1/R(z)$, and*
(c) *for $z \in \mathbb{R}$, we have $|R(iz)| = 1$.*

Proof. We give an explicit expression for an extremal function for Z_k by deriving it from the best rational approximation of the sign function on $[-b, -a] \cup [a, b]$. According to [1, sections 50 and 51, p. 144, line 6] we have

$$\inf_{r \in \mathcal{R}_{k,k}} \sup_{z \in [-b,-a] \cup [a,b]} |\text{sgn}(z) - r(z)| = \frac{2\sqrt{Z_k}}{1 + Z_k}, \qquad \text{sgn}(z) = \begin{cases} 1, & z \in [a, b], \\ -1, & z \in [-b, -a], \end{cases}$$

where the infimum is attained by the rational function [1, section 51, Table 2, no. 7 and 8]

$$\tilde{r}(z) = Mz \frac{\prod_{j=1}^{\lfloor (k-1)/2 \rfloor} z^2 + c_{2j}}{\prod_{j=1}^{\lfloor k/2 \rfloor} z^2 + c_{2j-1}}, \qquad c_j = a^2 \frac{\text{sn}^2(jK(\kappa)/k; \kappa)}{1 - \text{sn}^2(jK(\kappa)/k; \kappa)}. \tag{3.6}$$

Here, M is a real constant selected so that $\mathrm{sgn}(z) - \tilde{r}(z)$ equioscillates on $[-b,-a] \cup [a,b]$, $\kappa = \sqrt{1-(a/b)^2}$, and $\mathrm{sn}(\cdot)$ is the first Jacobian elliptic function.

In order to construct an extremal function for Z_k with the required properties, we observe from (3.6) that M and $c_1, \ldots, c_{k-1}$ are real, and thus

- $\tilde{r}(z)$ is real-valued for $z \in \mathbb{R}$,
- $\tilde{r}(iz)$ is purely imaginary for $z \in \mathbb{R}$, and
- $\tilde{r}(z)$ is an odd function on $\mathbb{R}$, i.e., $\tilde{r}(z) = -\tilde{r}(-z)$ for $z \in \mathbb{R}$.

As a consequence, the rational function given by

$$R(z) = \frac{1 + \frac{1+Z_k}{1-Z_k}\tilde{r}(z)}{1 - \frac{1+Z_k}{1-Z_k}\tilde{r}(z)} \in \mathcal{R}_{k,k}$$

is real-valued for $z \in \mathbb{R}$ with $R(-z) = 1/R(z)$, and of modulus 1 on the imaginary axis. Finally, as $\tilde{r}(z)$ takes values in the interval

$$\left[-1 - \frac{2\sqrt{Z_k}}{1+Z_k}, -1 + \frac{2\sqrt{Z_k}}{1+Z_k}\right] = \left[\frac{-(1+\sqrt{Z_k})^2}{1+Z_k}, \frac{-(1-\sqrt{Z_k})^2}{1+Z_k}\right]$$

for $z \in [-b,-a]$, we have for such z that

$$\frac{1+Z_k}{1-Z_k}\tilde{r}(z) \in \left[-\frac{1+\sqrt{Z_k}}{1-\sqrt{Z_k}}, -\frac{1-\sqrt{Z_k}}{1+\sqrt{Z_k}}\right],$$

implying that $-\sqrt{Z_k} \leq R(z) \leq \sqrt{Z_k}$ for $z \in [-b,-a]$. Hence, using $R(-z) = 1/R(z)$ we have

$$\frac{\sup\limits_{z\in[-b,-a]} |R(z)|}{\inf\limits_{z\in[a,b]} |R(z)|} \leq Z_k = \inf_{r\in\mathcal{R}_{k,k}} \frac{\sup\limits_{z\in[-b,-a]} |r(z)|}{\inf\limits_{z\in[a,b]} |r(z)|},$$

showing that R is extremal for $Z_k([-b,-a],[a,b])$, as required. ∎

Figure 3.1 (right) demonstrates the upper bound in Corollary 3.2 when $b/a = 1.1,\ 10,\ 100$. In section 4 we combine our upper bound on the singular values in Theorem 2.1 with our upper bound on Zolotarev numbers to derive explicit bounds on the singular values of certain Pick, Cauchy, and Löwner matrices.

3.2. Zolotarev Numbers for General Real Intervals. Consider the Zolotarev number $Z_k([a,b],[c,d])$, where either $b < c$ or $d < a$ so that $[a,b] \cap [c,d] = \emptyset$. Since $Z_k(E,F) = Z_k(T(E),T(F))$ for any Möbius transformation $T(z) = (a_1 z + a_2)/(a_3 z + a_4)$, we can find bounds on $Z_k([a,b],[c,d])$ by transplanting $[a,b] \cup [c,d]$ onto symmetric real intervals $[-\alpha,-1] \cup [1,\alpha]$ for some $\alpha > 1$. If $b < c$, then the Möbius transformation satisfies $T(a) = -\alpha$, $T(b) = -1$, $T(c) = 1$, $T(d) = \alpha$. Since T is a Möbius transformation the cross-ratio of the four collinear points a, b, c, and d equals the cross-ratio of $T(a)$, $T(b)$, $T(c)$, and $T(d)$. Hence, if $b < c$ or $d < a$, then we know that α must satisfy

$$\frac{|c-a||d-b|}{|c-b||d-a|} = \frac{(\alpha+1)^2}{4\alpha}.$$

Therefore, by solving the quadratic and noting that $\alpha > 1$ we find that

$$\alpha = -1 + 2\gamma + 2\sqrt{\gamma^2 - \gamma}, \qquad \gamma = \frac{|c-a||d-b|}{|c-b||d-a|}. \tag{3.7}$$

We conclude from Corollary 3.2 that

$$Z_k([a,b],[c,d]) \leq 4\left[\exp\left(\frac{\pi^2}{2\mu(1/\alpha)}\right)\right]^{-2k} \leq 4\left[\exp\left(\frac{\pi^2}{2\log(16\gamma)}\right)\right]^{-2k}. \tag{3.8}$$

Here, the last inequality comes from Gauss's transformation in (3.4) as we have $\mu(1/\alpha) = 2\mu(1/\sqrt{\gamma}) \leq 2\log(4\sqrt{\gamma}) = \log(16\gamma)$.

3.3. Zolotarev Numbers for Disks. For two disjoint closed disks $E, F \in \mathbb{C}$, it is known from a so-called near-circularity characterization of the extremal rational function that [53, p. 123]

$$Z_k(E,F) = \exp\left(-\frac{k}{\operatorname{cap}(E,F)}\right) = Z_1(E,F)^k,$$

where an explicit but technical expression for $\operatorname{cap}(E,F)$ is given in Example VIII.4.2 of [50]. By invariance under Möbius transformations, we see that the same formula holds if one of the sets E or F is either a half plane or the exterior of a disk of the form $\{z \in \mathbb{C} : |z-c| \geq r\}$. We give two examples where we can be a bit more explicit. Provided that $c \in \mathbb{C}$ and $0 < r_1 < r_2$, we have

$$Z_1\left(\{|z-c| \leq r_1\}, \{|z-c| \geq r_2\}\right) = \frac{r_1}{r_2},$$

which easily follows by applying [53, p. 123] to the (extremal) rational function $R(z) = z$. Also, provided that $0 < a < b$, we find that

$$Z_1(-E,E) = \left(\frac{1-\sqrt{a/b}}{1+\sqrt{a/b}}\right)^2, \qquad E = \left\{\left|z - \frac{b+a}{2}\right| \leq \frac{b-a}{2}\right\},$$

where we consider $R = (z+\sqrt{ab})/(z-\sqrt{ab})$. Interesting enough, the same function allows one to show that $Z_1([-b,-a],[a,b]) = Z_1(-E,E)$, even though E is quite a bit larger than $[a,b]$. We also see that the upper bound of $Z_k([-b,-a],[a,b]) \leq Z_1([-b,-a],[a,b])^k$ does not have an asymptotically sharp decay rate since the right-hand side is $Z_k(-E,E)$.

4. The Decay of the Singular Values of Pick, Cauchy, and Löwner Matrices.

In this section we bound the singular values of Pick (see section 4.1), Cauchy (see section 4.2), and Löwner (see section 4.3) matrices. In view of Theorem 2.1 and Corollary 3.2, our first idea is to construct matrices A and B so that the rank of $AX - XB$ is small with the additional hope that $\sigma(A)$ and $\sigma(B)$ are contained in real and disjoint intervals. For the three classes of matrices in this section, this first idea works out under mild "separation conditions." In section 5 the more challenging cases of Krylov, real Vandermonde, and real positive definite Hankel matrices are considered.

4.1. Pick Matrices. An $n \times n$ matrix P_n is called a Pick matrix if there exist a vector $\underline{s} = (s_1, \ldots, s_n)^T \in \mathbb{C}^{n\times 1}$ and a collection of real numbers $x_1 < \cdots < x_n$ from an interval $[a,b]$ with $0 < a < b < \infty$ such that

$$(P_n)_{jk} = \frac{s_j + s_k}{x_j + x_k}, \qquad 1 \leq j,k \leq n. \tag{4.1}$$

All Pick matrices satisfy the following Sylvester matrix equation:

$$D_{\underline{x}} P_n - P_n(-D_{\underline{x}}) = \underline{s}\,\underline{e}^T + \underline{e}\,\underline{s}^T, \qquad D_{\underline{x}} = \operatorname{diag}(x_1, \ldots, x_n), \tag{4.2}$$

where $\underline{e} = (1, \ldots, 1)^T$. Since diagonal matrices are normal matrices and in this case the spectrum of $D_{\underline{x}}$ is contained in $[a, b]$, we have the following bounds on the singular values of P_n.

COROLLARY 4.1. *Let P_n be the $n \times n$ Pick matrix in* (4.1). *Then, for $j \geq 1$, we have*

$$\sigma_{j+2k}(P_n) \leq 4 \left[\exp\left(\frac{\pi^2}{2\mu(a/b)} \right) \right]^{-2k} \sigma_j(P_n), \qquad 1 \leq j + 2k \leq n,$$

where $\mu(\lambda)$ is the Grötzsch ring function (see section 3*). The bound remains valid, but is slightly weakened, if $\mu(a/b)$ is replaced by* $\log(4b/a)$.

Proof. From (4.2), we know that $A = D_{\underline{x}}$, $B = -A$, $\nu = 2$, $E = [a, b]$, and $F = [-b, -a]$ in Theorem 2.1. Therefore, for $j \geq 1$ we have $\sigma_{j+2k}(P_n) \leq Z_k(E, F)\sigma_j(P_n)$, $1 \leq j + 2k \leq n$. The result follows from the upper bound in Corollary 3.2. ◻

There are two important consequences of Corollary 4.1: (1) Pick matrices are ill-conditioned unless b/a is large and/or n is small, and (2) all Pick matrices can be approximated, up to an accuracy of $\epsilon\|P_n\|_2$ with $0 < \epsilon < 1$, by a rank $\mathcal{O}(\log(b/a)\log(1/\epsilon))$ matrix. More precisely, for any Pick matrix in (4.1) we have

$$\kappa_2(P_n) = \frac{\sigma_1(P_n)}{\sigma_n(P_n)} \geq \frac{1}{4} \left[\exp\left(\frac{\pi^2}{2\log(4b/a)} \right) \right]^{2\lceil \frac{n}{2} - 1 \rceil}, \tag{4.3}$$

where for an even integer n we used $\sigma_1(P_n)/\sigma_n(P_n) \geq \sigma_1(P_n)/\sigma_{n-1}(P_n)$. Moreover, by setting k to be the smallest integer so that $\sigma_{1+2k}(P_n) \leq \epsilon\sigma_1(P_n)$, we find the following bound on the ϵ-rank of P_n (see (1.3)):

$$\operatorname{rank}_\epsilon(P_n) \leq 2 \left\lceil \frac{\log(4b/a)\log(4/\epsilon)}{\pi^2} \right\rceil. \tag{4.4}$$

In both (4.3) and (4.4), the bound can be slightly improved by replacing the $\log(4b/a)$ term by $\mu(a/b)$. Previously, bounds on the minimum and maximum singular values of Pick matrices were derived under the additional assumption that P_n is a positive definite matrix [24].

Figure 4.1 (left) demonstrates the bound in Corollary 4.1 on three 100×100 Pick matrices. The black line bounding the singular values has a stepping behavior because of the inequality in Corollary 4.1, which for $j = 1$ only bounds the odd indexed singular values of P_n. To bound $\sigma_{2k}(P_n)$ we use the trivial inequality $\sigma_{2k}(P_n) \leq \sigma_{2k-1}(P_n)$. At this time we can offer little insight into why the singular values of the tested Pick matrices also have a similar stepping behavior.

4.2. Cauchy Matrices. An $m \times n$ matrix $C_{m,n}$ with $m \geq n$ is called a (generalized) Cauchy matrix if there exist vectors $\underline{s} \in \mathbb{C}^{m\times 1}$ and $\underline{t} \in \mathbb{C}^{n \times 1}$, points $x_1 < \cdots < x_m$ on an interval $[a, b]$ with $-\infty < a < b < \infty$, and points $y_1 < \cdots < y_n$ (all distinct from $x_1, \ldots, x_m$) in an interval $[c, d]$ with $-\infty < c < d < \infty$ such that

$$(C_{m,n})_{jk} = \frac{s_j t_k}{x_j - y_k}, \qquad 1 \leq j \leq m, \quad 1 \leq k \leq n. \tag{4.5}$$

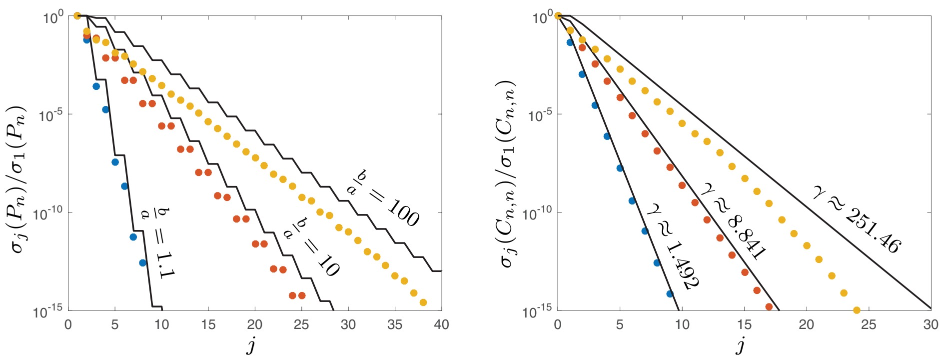

Fig. 4.1 *Left: The scaled singular values of* 100×100 *Pick matrices (colored dots) and the bound in Corollary* 4.1 *(black lines) for* $b/a = 1.1$ *(blue dots),* 10 *(red dots),* 100 *(yellow dots). In* (4.1)*,* $\underline{x}$ *is a vector of equally spaced points in* $[a, b]$ *and* $\underline{s}$ *is a random vector with independent standard Gaussian entries. Right: The scaled singular values of* 100×100 *Cauchy matrices (colored dots) and the bound in Corollary* 4.2 *(black lines) for* $\gamma = 1.1, 10, 100$. *In* (4.5)*,* $\underline{x}$ *is a vector of Chebyshev nodes from* $[-8.5, -2]$ *(blue dots),* $[-100, -3]$ *(red dots), and* $[-101, 2.8]$ *(yellow dots), respectively,* $\underline{y}$ *is a vector of Chebyshev nodes from* $[3, 10]$ *(blue dots),* $[3, 100]$ *(red dots), and* $[3, 100]$ *(yellow dots), respectively, and* $\underline{s}$ *and* $\underline{t}$ *are random vectors with independent standard Gaussian entries. The decay rate depends on the cross-ratio of the endpoints of the intervals.*

Generalized Cauchy matrices satisfy the following Sylvester matrix equation:

$$D_{\underline{x}} C_{m,n} - C_{m,n} D_{\underline{y}} = \underline{s}\,\underline{t}^T, \tag{4.6}$$

where $D_{\underline{x}} = \operatorname{diag}(x_1, \ldots, x_m)$ and $D_{\underline{y}} = \operatorname{diag}(y_1, \ldots, y_n)$.

If we make the further assumption that either $b < c$ or $d < a$ so that the intervals $[a, b]$ and $[c, d]$ are disjoint, then we can bound the singular values of $C_{m,n}$. This "separation condition" is an extra assumption on Cauchy matrices that simplifies the analysis. If the intervals $[a, b]$ and $[c, d]$ overlapped, then one would have to consider discrete Zolotarev numbers to estimate the singular values, and we want to avoid this in this paper. For implicit bounds, i.e., bounds that require computation, on the singular values of Cauchy matrices that continue to hold when the separation condition is violated, see [58, Theorem 3] and [59, (2.34)].

Corollary 4.2. *Let* $C_{m,n}$ *be an* $m \times n$ *Cauchy matrix in* (4.5) *with* $m \geq n$ *and either* $b < c$ *or* $d < a$*. Then*

$$\sigma_{j+k}(C_{m,n}) \leq 4 \left[\exp\left(\frac{\pi^2}{4\mu(1/\sqrt{\gamma})} \right) \right]^{-2k} \sigma_j(C_{m,n}), \qquad 1 \leq j + k \leq n,$$

where $\gamma = |(c - a)(d - b)/((c - b)(d - a))|$ *is the absolute value of the cross-ratio of* a*,* b*,* c*, and* d*. If* $c = -b$ *and* $d = -a$*, then* $2\mu(1/\sqrt{\gamma}) = \mu(a/b)$*. The bound remains valid, but is slightly weakened, if* $4\mu(1/\sqrt{\gamma})$ *is replaced by* $2\log(16\gamma)$*.*

Proof. From (4.6), we know that $A = D_{\underline{x}}$, $B = D_{\underline{y}}$, $\nu = 1$, $E = [a, b]$, and $F = [c, d]$ in Theorem 2.1. Therefore, we have $\sigma_{j+k}(C_{m,n}) \leq Z_k(E, F)\sigma_j(C_{m,n})$ for $1 \leq j + k \leq n$. From (3.8), we conclude that

$$\sigma_{j+k}(C_{m,n}) \leq 4 \left[\exp\left(\frac{\pi^2}{2\mu(1/\alpha)} \right) \right]^{-2k} \sigma_j(C_{m,n}), \qquad 1 \leq j + k \leq n. \qquad \square$$

It is interesting to observe that our bound on the singular values of Cauchy matrices depends on the absolute value of the cross-ratio of a, b, c, and d. Hence, the "separation" of two real intervals $[a, b]$ and $[c, d]$ for the purposes of our singular value estimates is measured in terms of the cross-ratio of a, b, c, and d—not the separation distance $\max(c - b, a - d)$.

Corollary 4.2 shows that the Cauchy matrix in (4.5) (when $b < c$ or $d < a$) has an ϵ-rank of at most

$$\mathrm{rank}_\epsilon(C_{m,n}) \leq \left\lceil \frac{2\mu(1/\sqrt{\gamma})\log(4/\epsilon)}{\pi^2} \right\rceil \leq \left\lceil \frac{\log(16\gamma)\log(4/\epsilon)}{\pi^2} \right\rceil,$$

where γ is the absolute value of the cross-ratio of a, b, c, and d. We expect our bounds to be quite tight when the cross-ratio $\gamma \gg 1$. When the cross-ratio is close to 1, then the discrete nature of the spectra of $D_{\underline{x}}$ and $D_{\underline{y}}$ matters more and our bound may not adequately describe the decay rate of the singular values. When $m \geq n$ and $b < c$ or $d < a$, we have the following lower bound on the condition number of $C_{m,n}$:

$$\kappa_2(C_{m,n}) = \frac{\sigma_1(C_{m,n})}{\sigma_n(C_{m,n})} \geq \frac{1}{4}\left[\exp\left(\frac{\pi^2}{2\log(16\gamma)}\right)\right]^{2(n-1)}, \qquad \gamma = \frac{|c-a||d-b|}{|c-b||d-a|}.$$

From the bounds in subsection 3.3, explicit singular value bounds are also possible to derive when $x_1, \ldots, x_m$ and $y_1, \ldots, y_n$ lie in two disjoint disks.

Remark 4.3. Most of the existing bounds in the literature on the numerical rank of Cauchy matrices are derived from separable approximations of the Cauchy kernel (see [60], [23, section 2.2], and [48, Theorem 1]). For simplicity we only report about the case where $[c, d] = [-b, -a]$ with $0 < a < b$. If there exist $2k$ functions f_j and g_j such that for some $0 < \epsilon < 1$ one has

$$\left| \frac{1}{x-y} - \sum_{j=1}^{k} f_j(x) g_j(y) \right| \leq \epsilon \left| \frac{1}{x-y} \right|, \qquad (x, y) \in [a, b] \times [-b, -a], \tag{4.7}$$

then $\mathrm{rank}_\epsilon(C_{m,n}) \leq k$. If $R \in \mathcal{R}_{k,k}$, then the proof of Theorem 2.1 shows that

$$\frac{1}{x-y}\left(\frac{R(x)}{R(y)} - 1\right) = \sum_{j=1}^{k} f_j(x) g_j(y),$$

where f_j and g_j are rational functions. Thus, (4.7) holds, provided that

$$\max_{x\in[a,b]} |R(x)| \Big/ \min_{y\in[-b,-a]} |R(y)| \leq \epsilon.$$

By sampling the functions f_j and g_j, one can construct a low rank approximant of $C_{m,n}$. Tyrtyshnikov used this observation to construct a skeleton approximation [45] of the Cauchy matrix [60]; see also [32] and [55, Chapter 3] for the hierarchical low rank structure of $C_{m,n}$. In [23, section 2.2] the authors first apply a Möbius transform and then use a polynomial approximation of the Cauchy kernel. Penzl in [48, Theorem 1] obtains a sharper bound in terms of $\sqrt{b/a}$ by constructing a particular rational function $R \in \mathcal{R}_{k,k}$; however, our bound of $4\rho^{2k} \leq \epsilon$ with $\rho \in (0, 1)$ as in Corollary 3.2 is sharper since it is based on an optimal rational function. Another approach investigated by Braess and Hackbusch [19] (see also [16, 18]) is to let $z = x - y \in [2a, 2b]$ and approximate the function $1/z$ by an exponential sum of the form $\sum_{j=1}^{k} \alpha_j \exp(-t_j z)$ for suitable chosen $\alpha_j, t_j \in \mathbb{R}$. A combination of [19, Lemma 2.1 and Theorems 3.3 and 4.1] shows that (4.7) holds with $c\, k\, \rho^{-2k} \leq \epsilon$ with the same ρ as before, and a nonexplicit constant c depending on a/b but not on k. The connection to exponential sum bounds is extensively discussed in [57, section 4.3].

Corollary 4.2 also includes the important Hilbert matrix, i.e., $(H_n)_{jk} = 1/(j+k-1)$ for $1 \leq j, k \leq n$. By setting $x_j = j - 1/2$, $y_j = -k + 1/2$, and $\underline{s} = \underline{r} = (1, \ldots, 1)^T$, the matrix in (4.5) is the Hilbert matrix. In particular, Corollary 4.2 with $[a, b] = [-n + 1/2, -1/2]$ and $[c, d] = [1/2, n - 1/2]$ shows that

$$\sigma_{k+1}(H_n) \leq 4 \left[\exp\left(\frac{\pi^2}{2 \log(8n-4)} \right) \right]^{-2k} \sigma_1(H_n), \qquad 1 \leq k \leq n-1. \tag{4.8}$$

Therefore, the Hilbert matrix can be well approximated by a low rank matrix and has exponentially decaying singular values.[6] In particular, it has an ϵ-rank of at most $\lceil \log(8n-4) \log(4/\epsilon)/\pi^2 \rceil$. Recall that our bounds are most useful for $\sigma_k(H_n)$, where k is relatively small with respect to n. Therefore, (4.8) with $k = n-1$ does not give an asymptotically sharp bound on the condition number of H_n [65, eqn. (3.35)]. The Hilbert matrix is an example of a real positive definite Hankel matrix, and in section 5 we show that bounds similar to (4.8) hold for the singular values of all such matrices.

Figure 4.1 (right) demonstrates the bound in Corollary 4.2 on three $n \times n$ Cauchy matrices, where $n = 100$. In practice, the derived bound is relatively tight for singular values $\sigma_j(C_{m,n})$ when j is small with respect to n.

4.3. Löwner Matrices. An $n \times n$ matrix L_n is called a Löwner matrix if there exist vectors $\underline{r}, \underline{s} \in \mathbb{C}^{n \times 1}$, points $x_1 < \cdots < x_n$ in $[a, b]$ with $-\infty < a < b < \infty$, and points $y_1 < \cdots < y_n$ (all different from $x_1, \ldots, x_n$) in $[c, d]$ with $-\infty < c < d < \infty$ such that

$$(L_n)_{jk} = \frac{r_j - s_k}{x_j - y_k}, \qquad 1 \leq j, k \leq N. \tag{4.9}$$

In the special case when $y_j = -x_j$ and $s_j = -r_j$, a Löwner matrix is a Pick matrix (see section 4.1). Löwner matrices satisfy the Sylvester matrix equation given by

$$D_{\underline{x}} L_n - L_n D_{\underline{y}} = \underline{r}\, \underline{e}^T - \underline{e}\, \underline{s}^T,$$

where $\underline{e} = (1, \ldots, 1)^T$. From Theorem 2.1 we can bound the singular values of L_n, provided that $[a, b]$ and $[c, d]$ are disjoint, i.e., either $b < c$ or $d < a$. We emphasize that the separation condition of the intervals $[a, b]$ and $[c, d]$ is an extra assumption on a Löwner matrix that allows us to proceed with the methodology we have developed.

Corollary 4.4. *Let L_n be an $n \times n$ Löwner matrix in* (4.9) *with $b < c$ or $d < a$. Then, for $j \geq 1$, we have*

$$\sigma_{j+2k}(L_n) \leq 4 \left[\exp\left(\frac{\pi^2}{4\mu(1/\sqrt{\gamma})} \right) \right]^{-2k} \sigma_j(L_n), \qquad 1 \leq j + 2k \leq n,$$

where γ is the absolute value of the cross-ratio of a, b, c, and d (see (3.7)*). If $c = -b$ and $d = -a$, then $2\mu(1/\sqrt{\gamma}) = \mu(a/b)$. The bound remains valid, but is slightly weakened, if $4\mu(1/\sqrt{\gamma})$ is replaced by $2\log(16\gamma)$.*

[6] More generally, skeleton decompositions can be used to show that the Hilbert kernel of $f(x, y) = 1/(x+y)$ on $[a, b] \times [a, b]$ with $0 < a < b < \infty$ has exponentially decaying singular values [45].

Proof. The proof follows the same argument as in the proof of Corollary 4.2, but with $\nu = 2$. ◻

Corollary 4.4 shows that many Löwner matrices can be well approximated by low rank matrices with $\text{rank}_\epsilon(L_n) = \mathcal{O}(\log\gamma\log(1/\epsilon))$ and are exponentially ill-conditioned.

5. The Singular Values of Krylov, Vandermonde, and Hankel Matrices.

The three types of matrices considered in section 4 allowed for direct applications of Theorem 2.1 and Corollary 3.2. In this section, we consider the more challenging tasks of bounding the singular values of Krylov matrices with Hermitian arguments, real Vandermonde matrices, and real positive definite Hankel matrices.

5.1. Krylov and Real Vandermonde Matrices.

An $m \times n$ matrix $K_{m,n}$ with $m \geq n$ is said to be a Krylov matrix if there exist a matrix $A \in \mathbb{C}^{m\times m}$ and a vector $\underline{w} \in \mathbb{C}^{m\times 1}$ such that

$$K_{m,n} = \left[\underline{w} \,\middle|\, A\underline{w} \,\middle|\, \cdots \,\middle|\, A^{n-1}\underline{w} \right]. \tag{5.1}$$

Krylov matrices satisfy the following Sylvester matrix equation:

$$AK_{m,n} - K_{m,n}Q = \underline{s}\,\underline{e}_n^T, \qquad Q = \begin{bmatrix} 0 & & & -1 \\ 1 & & & \\ & \ddots & & \\ & & 1 & 0 \end{bmatrix}, \tag{5.2}$$

where $\underline{s} \in \mathbb{C}^{m\times 1}$ and $\underline{e}_n = (0,\ldots,0,1)^T$. Here, we will only consider $K_{m,n}$ where the matrix A in (5.2) is normal. When $A = D_{\underline{x}}$ and $\underline{w} = (1,\ldots,1)^T$ the Krylov matrix in (5.1) is the Vandermonde matrix $(V_{m,n})_{jk} = x_j^{k-1}$ for $1 \leq j \leq m$ and $1 \leq k \leq n$. This makes the bounds in this section also applicable to Vandermonde matrices. The vector of $\underline{w}$ in (5.1) allows for additional row scaling of $V_{m,n}$. We refer the reader to [41] for a state-of-the-art conditioning analysis of Vandermonde matrices.

We start by considering an example for which A is a unitary matrix and Theorem 2.1 is trivial.

Example 5.1. Let $m = n$ and $A = \widetilde{Q}$ be the $n \times n$ circshift matrix obtained from Q in (5.2) by replacing the -1 entry by 1. Then, for any vector $\underline{s}$ in (5.2), the solution $K_{n,n}$ to the Sylvester matrix equation is a circulant matrix. Moreover, all circulant matrices satisfy (5.2) with $A = \widetilde{Q}$. If $\underline{s}$ is the first or second canonical vector, then $K_{n,n}$ is the identity or circshift matrix, respectively. In both cases, $K_{n,n}$ is an orthogonal matrix and $\sigma_j(K_{n,n}) = 1$ for $1 \leq j \leq n$. Through the link with companion matrices, the eigenvalues of $\widetilde{Q}$ are the nth roots of unity, i.e., the even powers of $\exp(\frac{2\pi i}{2n})$. These eigenvalues perfectly interlace the eigenvalues of Q, which are the odd powers of $\exp(\frac{2\pi i}{2n})$. An application of Theorem 2.1 reveals that

$$1 = \frac{\sigma_{k+1}(I_n)}{\sigma_1(I_n)} \leq Z_k(\sigma(\widetilde{Q}), \sigma(Q)) \leq 1,$$

where I_n is the identity matrix and $\sigma(\widetilde{Q})$ and $\sigma(Q)$ denote the spectrum of $\widetilde{Q}$ and Q, respectively. Hence, we find that $Z_k(\sigma(\widetilde{Q}), \sigma(Q)) = 1$ for $1 \leq k \leq n-1$. As a consequence, for any circulant matrix $K_{n,n}$ our Theorem 2.1 gives the trivial bound that $\sigma_{k+1}(K_{n,n}) \leq \sigma_1(K_{n,n})$ for $1 \leq k \leq n-1$. On the other hand, using the discrete Fourier transform one can construct a circulant matrix with any sequence of singular values $s_1 \geq s_2 \geq \cdots \geq s_n \geq 0$. We conclude that for circulant matrices as well as the larger class of Toeplitz matrices, Theorem 2.1 is not useful.

Similarly, from the Moitra upper bound [41] on the condition number of a rectangular Vandermonde matrix $V_{m,n}$, $m \leq n$, with arbitrary nodes $x_1, \ldots, x_m$ on the unit circle, we get lower bounds for the Zolotarev number $Z_{m-1}(\sigma(D_{\underline{x}}), \sigma(Q))$, provided that the nodes x_j are distributed like n roots of unity.

We now study another class of Krylov matrices where A is Hermitian, and we attempt to use Theorem 2.1 to bound the singular values of $K_{m,n}$. When A is Hermitian, Theorem 2.1 provides a nontrivial bound on the singulars of $K_{m,n}$.

For the analysis that follows, we require that n be an even integer. This is not a loss of generality because of the interlacing theorem for singular values [54]. To see this, let $K_{m,n-1}$ be the $m \times (n-1)$ Krylov matrix obtained from $K_{m,n}$ by removing its last column. If n is odd, then[7]

$$\frac{\sigma_{j+k}(K_{m,n})}{\sigma_j(K_{m,n})} \leq \frac{\sigma_{j+k-1}(K_{m,n-1})}{\sigma_j(K_{m,n-1})}, \qquad 2 \leq j+k \leq n, \tag{5.3}$$

and one can bound $\sigma_{j+k-1}(K_{m,n-1})/\sigma_j(K_{m,n-1})$ instead. From now on in this section we will assume that n is an even integer.

The Sylvester matrix equation in (5.2) contains matrices A and Q, which are both normal matrices. The eigenvalues of A are contained in $\mathbb{R}$ and the eigenvalues of Q are the n (shifted) roots of unity, i.e.,

$$\sigma(Q) = \left\{ z \in \mathbb{C} : z = e^{\frac{2\pi i(j+1/2)}{n}}, 0 \leq j \leq n-1 \right\}.$$

Since n is even, the spectrum of Q and the real line are disjoint. Using Theorem 2.1 we find that for $j \geq 1$ and $1 \leq j+k \leq n$

$$\sigma_{j+k}(K_{m,n}) \leq Z_k(E,F)\sigma_j(K_{m,n}), \qquad E \subseteq \mathbb{R}, \quad F = F_+ \cup F_-,$$

where F_+ and F_- are complex sets defined by

$$F_+ = \{e^{it} : t \in [\tfrac{\pi}{n}, \pi - \tfrac{\pi}{n}]\}, \quad F_- = \{e^{it} : t \in [-\pi + \tfrac{\pi}{n}, -\tfrac{\pi}{n}]\}. \tag{5.4}$$

Figure 5.1 shows the two sets E and F in the complex plane. As $n \to \infty$ the sets F_+ and F_- approach the real line, suggesting that our bound on the singular values must depend on n somehow. Our task is to bound the quantity $Z_k(E, F_+ \cup F_-)$—a Zolotarev number that is not immediately related to one of the form $Z_k([-b,-a],[a,b])$.

The following lemma relates the quantity $Z_{2k}(E, F_+ \cup F_-)$ to the Zolotarev number $Z_k([-1/\ell, -\ell], [\ell, 1/\ell])$ with $\ell = \tan(\pi/(2n))$.

LEMMA 5.2. *Let $k \geq 1$ be an integer and $E \subseteq \mathbb{R}$. Then $Z_{2k+1}(E, F_+ \cup F_-) \leq Z_{2k}(E, F_+ \cup F_-)$ and*

$$Z_{2k}(E, F_+ \cup F_-) \leq \frac{2\sqrt{Z_k}}{1+Z_k}, \quad Z_k := Z_k([-1/\ell, -\ell], [\ell, 1/\ell]),$$

[7] Observe that the singular values of a matrix decrease when removing a column and thus $\sigma_j(K_{m,n-1}) \leq \sigma_j(K_{m,n})$. Let Y be a best rank $j+k-2$ approximation to $K_{m,n-1}$ so that $\sigma_{j+k-1}(K_{m,n-1}) = \|K_{m,n-1} - Y\|_2$ and consider X obtained from Y by concatenating (on the right) the last column of $K_{m,n}$. Then the rank of X is at most $j+k-1$, and hence $\sigma_{j+k}(K_{m,n}) \leq \|K_{m,n} - X\|_2 = \|K_{m,n-1} - Y\|_2 = \sigma_{j+k-1}(K_{m,n-1})$.

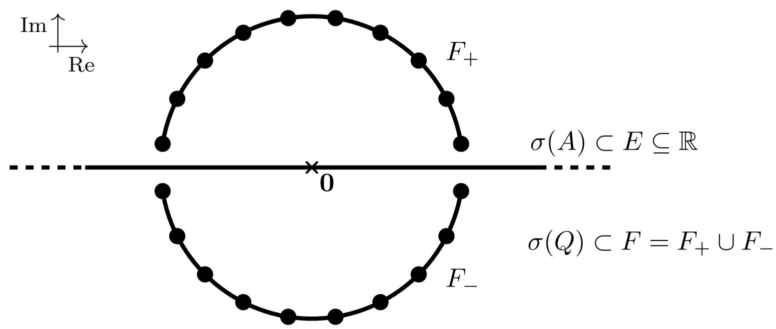

Fig. 5.1 *The sets E and F in the complex plane for the Zolotarev problem (1.4) used to bound the singular values of a 20×20 Krylov matrix with a Hermitian argument. The sets F_+ and F_- are a distance of only $\mathcal{O}(1/n)$ from the real axis, where n is the size of the Krylov matrix, and this causes the $\log n$ dependence in the weaker version of (5.5). The solid black dots denote the spectrum of Q, which is contained in $F_+ \cup F_-$.*

where $\ell = \tan(\pi/(2n))$, the complex sets F_+ and F_- are as defined in (5.4), and n is an even integer.

Proof. Let $R(z) \in \mathcal{R}_{k,k}$ be the extremal function for $Z_k := Z_k([-1/\ell, -\ell], [\ell, 1/\ell])$ characterized in Theorem 3.3, where $\ell = \tan(\pi/(2n))$. Since the Möbius transform given by

$$T(z) = \frac{1}{i}\frac{z-1}{z+1}$$

maps F_+ to $[\ell, 1/\ell]$, F_- to $[-1/\ell, -\ell]$, and $\mathbb{R}$ to $i\mathbb{R}$, we have

$$Z_{2k}(\mathbb{R}, F_+ \cup F_-) = Z_{2k}(i\mathbb{R}, [-1/\ell, -\ell] \cup [\ell, 1/\ell]) = \inf_{r \in \mathcal{R}_{2k,2k}} \frac{\sup_{z\in\mathbb{R}} |r(iz)|}{\inf_{z\in[-1/\ell,-\ell]\cup[\ell,1/\ell]} |r(z)|}.$$

Now, consider the rational function

$$S(z) = \frac{R(z) + 1/R(z)}{2} = \frac{R(z) + R(-z)}{2} \in \mathcal{R}_{2k,2k},$$

where we used the fact that $1/R(z) = R(-z)$ (see Theorem 3.3(b)). Since $|R(iz)| = 1$ for $z \in \mathbb{R}$ (see Theorem 3.3(c)), we have

$$\sup_{z\in\mathbb{R}} |S(iz)| = \sup_{z\in\mathbb{R}} \left| \frac{R(iz) + R(-iz)}{2} \right| \leq 1.$$

Moreover, since $-1 \leq -\sqrt{Z_k} \leq R(z) \leq \sqrt{Z_k} \leq 1$ for $z \in [-1/\ell, -\ell]$ (see Theorem 3.3(a)) and $x \mapsto 2x/(1+x^2)$ is a nondecreasing function on $[-1, 1]$ and $S(-z) = S(z)$, we have

$$\begin{aligned}
\sup_{z\in[-1/\ell,-\ell]\cup[\ell,1/\ell]} \left|\frac{1}{S(z)}\right| &= \sup_{z\in[-1/\ell,-\ell]} \left| \frac{2}{R(z) + 1/R(z)} \right| \\
&= \sup_{z\in[-1/\ell,-\ell]} \left| \frac{2R(z)}{1 + R(z)^2} \right| \\
&\leq \frac{2\sqrt{Z_k}}{1 + Z_k}.
\end{aligned}$$

Therefore, $Z_{2k}(E, F_+ \cup F_-) \leq Z_{2k}(\mathbb{R}, F_+ \cup F_-) \leq 2\sqrt{Z_k}/(1 + Z_k)$ as required. The bound $Z_{2k+1}(E, F_+ \cup F_-) \leq Z_{2k}(E, F_+ \cup F_-)$ trivially holds from the definition of Zolotarev numbers. ∎

By Corollary 3.2 we have the slightly weaker upper bound for $Z_{2k}(E, F_+ \cup F_-)$:

$$Z_{2k}(E, F_+ \cup F_-) \leq 2\sqrt{Z_k([-1/\ell, -\ell], [\ell, 1/\ell])} \leq 4\rho^{-k},$$

where since $\tan x \geq x$ for $0 \leq x \leq \pi/2$, we have

$$\rho = \exp\left(\frac{\pi^2}{2\mu(\tan(\pi/(2n))^2)}\right) \geq \exp\left(\frac{\pi^2}{2\log(4/\tan(\pi/(2n))^2)}\right) \geq \exp\left(\frac{\pi^2}{4\log(4n/\pi)}\right).$$

If n is an even integer, then we can immediately conclude a bound on the singular values from Theorem 2.1. If n is an odd integer, then one must employ (5.3) first.

Corollary 5.3. *The singular values of $K_{m,n}$ can be bounded as follows:*

(5.5)
$$\sigma_{j+2k}(K_{m,n}) \leq 4\left[\exp\left(\frac{\pi^2}{2\mu(\tan(\pi/(4\lfloor n/2\rfloor))^2)}\right)\right]^{-k+[n]_2} \sigma_j(K_{m,n}), \quad 1 \leq j + 2k \leq n,$$

where $\mu(\cdot)$ is the Grötzsch function and $[n]_2 = 1$ if n is odd and is 0 if n is even. The bound above remains valid, but is slightly weakened, if $2\mu(\tan(\pi/(4\lfloor n/2\rfloor))^2)$ is replaced by $4\log(8\lfloor n/2\rfloor/\pi)$.

Figure 5.2 demonstrates the bound on the singular values in (5.5) on $n \times n$ Krylov matrices, where $n = 10, 100, 1000$. The step behavior of the bound is due to the fact that (5.5) only bounds $\sigma_{1+2k}(K_{m,n})$ when n is even, and we use the trivial inequality $\sigma_{2k+2}(K_{m,n}) \leq \sigma_{2k+1}(K_{m,n})$ otherwise. One also observes that the singular values of Krylov matrices with Hermitian arguments can decay at a supergeometric rate; however, the analysis in this paper only realizes a geometric decay. Therefore, (5.5) is only a reasonable bound on $\sigma_j(K_{m,n})$ when j is a small integer with respect to n. If $j/n \to c$ and $c \in (0, 1)$, then improved bounds on $\sigma_j(K_{m,n})$ may be possible by bounding discrete Zolotarev numbers [9]. The bound in (5.5) provides an upper bound on the ϵ-rank of $K_{m,n}$:

$$\text{rank}_\epsilon(K_{m,n}) \leq 2\left\lceil \frac{4\log(8\lfloor n/2\rfloor/\pi)\log(4/\epsilon)}{\pi^2} \right\rceil + 2,$$

which allows for either an odd or an even integer n.

As mentioned before, Vandermonde matrices of size $m \times n$ with real abscissas $\underline{x} \in \mathbb{R}^{m\times 1}$, i.e., $(V_{m,n})_{jk} = x_j^{k-1}$, are also Krylov matrices with $A = D_{\underline{x}}$ and $\underline{w} = (1, \ldots, 1)^T$. Therefore, the bounds in this section also apply to (possibly row-scaled) Vandermonde matrices with real abscissas and show that they have rapidly decaying singular values and are exponentially ill-conditioned, an observation that has been extensively investigated in the literature [6, 28, 46].

5.2. Real Positive Definite Hankel Matrices. An $n \times n$ matrix H_n is a Hankel matrix if the matrix is constant along each antidiagonal, i.e., $(H_n)_{jk} = h_{j+k}$ for $1 \leq j, k \leq n$. Clearly, not all Hankel matrices have decaying singular values. For example, the exchange matrix has repeated singular values of 1. This means that any displacement structure that is satisfied by all Hankel matrices, for example,

$$\text{rank}\left(QX - XQ^T\right) \leq 2,$$

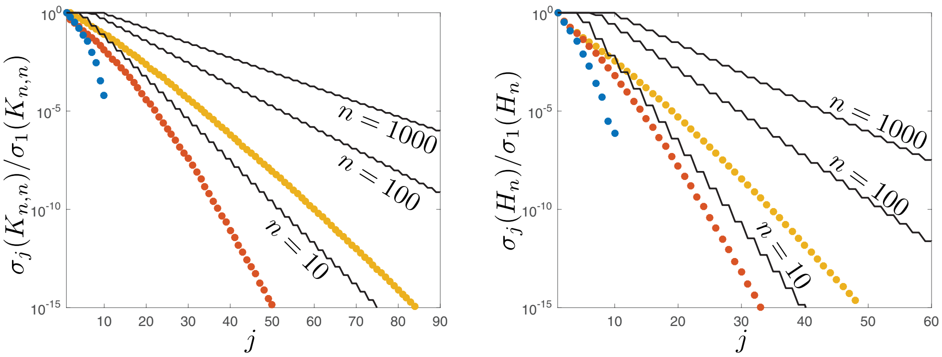

Fig. 5.2 *Left: The singular values of the* $n \times n$ *Krylov matrix (colored dots) compared to the bound in* (5.5) *for* $n = 10$ *(blue dots),* 100 *(red dots),* 1000 *(yellow dots). In* (5.1) *the matrix* A *is a diagonal matrix with entries taken to be equally spaced points in* $[-1, 1]$ *and* $\underline{w}$ *is a random vector with independent Gaussian entries. Right: The singular values of the* $n \times n$ *real positive definite Hankel matrices (colored dots) associated to the measure* $\mu_H(x) = \mathbf{1}|_{-1\leq x\leq 1}$ *compared to the bound in* (5.7) *for* $n = 10$ *(blue dots),* 100 *(red dots),* 1000 *(yellow dots).*

where Q is given in (5.2), does not result in a Zolotarev number that decays. Motivated by the Hilbert matrix in section 4.2, we show that every real and positive definite Hankel matrix has rapidly decaying singular values. Previous work has led to bounds that can be calculated by using a pivoted Cholesky algorithm [2], bounds for very special cases [56], and incomplete attempts [61, 62].

In order to exploit the positive definite structure we recall that the Hamburger moment problem states that a real Hankel matrix is positive semidefinite if and only if it is associated to a nonnegative Borel measure supported on the real line.

Lemma 5.4. *A real* $n \times n$ *Hankel matrix,* H_n*, is positive semidefinite if and only if there exists a nonnegative Borel measure* μ_H *supported on the real line such that*

$$(H_n)_{jk} = \int_{-\infty}^{\infty} x^{j+k-2} \mathrm{d}\mu_H(x), \qquad 1 \leq j, k \leq n. \tag{5.6}$$

Proof. For a proof, see [47, Theorem 7.1]. □

Let H_n be a real positive definite Hankel matrix associated to the nonnegative weight μ_H in (5.6) supported on $\mathbb{R}$. Let $x_1, \ldots, x_n$ and $w_1^2, \ldots, w_n^2$ be the Gauss quadrature nodes and weights associated to μ_H. Then, since a Gauss quadrature is exact for polynomials of degree $2n-1$ or less, we have

$$(H_n)_{jk} = \int_{-\infty}^{\infty} x^{j+k-2} \mathrm{d}\mu_H(x) = \sum_{s=1}^{n} w_s^2 x_s^{j+k-2} = \sum_{s=1}^{n} \left(w_s x_s^{j-1}\right)\left(w_s x_s^{k-1}\right).$$

Therefore, every real positive definite Hankel matrix has a so-called *Fiedler factorization* [25], i.e.,

$$H_n = K_{n,n}^* K_{n,n}, \qquad K_{n,n} = \left[\underline{w} \;\middle|\; D_{\underline{x}}\underline{w} \;\middle|\; \cdots \;\middle|\; D_{\underline{x}}^{n-1}\underline{w}\right],$$

where $K_{n,n}$ is a Krylov matrix with Hermitian argument and $K_{n,n}^*$ is the conjugate transpose of $K_{n,n}$. This means that $\sigma_j(H_n) = \sigma_j(K_{n,n})^2$ for $1 \leq j \leq n$. That is, a bound on the singular values of H_n and the ϵ-rank of H_n directly follows from (5.5).

Corollary 5.5. *Let H_n be an $n \times n$ real positive definite Hankel matrix. Then*

$$\sigma_{j+2k}(H_n) \leq 16 \left[\exp\left(\frac{\pi^2}{4\log(8\lfloor n/2\rfloor/\pi)}\right)\right]^{-2k+2} \sigma_j(H_n), \qquad 1 \leq j+2k \leq n, \tag{5.7}$$

and

$$\operatorname{rank}_\epsilon(H_n) \leq 2\left\lceil \frac{2\log\left(8\lfloor n/2\rfloor/\pi\right)\log\left(16/\epsilon\right)}{\pi^2}\right\rceil + 2,$$

where both bounds allow for n to be an even or odd integer.

We conclude that all real positive definite Hankel matrices have an ϵ-rank of at most $\mathcal{O}(\log n \log(1/\epsilon))$, explaining why low rank techniques are usually advantageous in computational mathematics on such matrices.

Since a real positive semidefinite Hankel matrix can be arbitrarily approximated by a real positive definite Hankel matrix, the results from this section immediately extend to such Hankel matrices.[8] This fact was exploited, but not proved in general, in [56] to derive quasi-optimal complexity fast transforms between orthogonal polynomial bases.

Conclusions. Many special matrices such as Pick, Vandermonde, and Hankel matrices have a displacement structure and satisfy a Sylvester matrix equation. Zolotarev numbers can be used to derive explicit bounds on the singular values of such matrices (see Table 1.1) that help explain the abundance of low rank structures in computational mathematics.

Zolotarev numbers have received renewed attention in the field of numerical linear algebra for computing matrix functions [29, 34], for communication-avoiding eigensolvers [43], for fast Poisson solvers [26, 64], for solving Sylvester equations [48, 49], for solving Riccati equations [37], for bounding exponential sums [18, 19], and for tensor compression [51]. We look forward to the future advances that this renewed interest in Zolotarev numbers will bring.

Appendix A. Infinite Product Formulas for Quantities Related to the Zolotarev Number. Here, we rederive infinite product formulas related to a formula given by Lebedev [38, (1.11)], which unfortunately contained a typo. The typo has been copied several times in the literature.

Lebedev [38] and his successors [40, 45] were not concerned with the Zolotarev problem in (1.4) but instead with the equivalent problem of minimal Blaschke products in the half plane, i.e.,

$$E_k([a,b]) = \min_{z_1,\ldots,z_k\in\mathbb{C}} \max_{z\in[a,b]} \left|\prod_{s=1}^{k} \frac{z-z_s}{z+\overline{z_s}}\right|, \qquad 0 < a < b < \infty. \tag{A.1}$$

In [38, (1.11)], Lebedev presented an infinite product formula for E_k that unfortunately contained typos and resulted in an erroneous lower bound for E_k in [38, (1.12)]. More recently, other erroneous lower bounds have been claimed in [34, (4.1)] for a related problem based on [40, (3.17)].

To correct the situation we first show that with $Z_k := Z_k([-b,-a],[a,b])$ we have

$$\sqrt{Z_k} = E_k([-b,-a]) = E_k([a,b]) = E_k([a/b,1]),$$

[8] For a real positive semidefinite Hankel matrix one may improve our bounds on the singular values of H_n by replacing n by the rank of H_n.

where the last two equalities are immediate from symmetry considerations and scaling. Since any $z_1, \ldots, z_k \in \mathbb{C}$ describes a rational function for $E_k([-b,-a])$ in (A.1), the solution to (A.1) describes a rational function $r(z)$ that is a candidate for the Zolotarev problem in (1.4) and satisfies $1/r(z) = \overline{r(-\overline{z})}$. By construction, we have

$$\max_{z\in[-b,-a]} |r(z)| = E_k([-b,-a]), \qquad \max_{z\in[a,b]} \left|\frac{1}{r(z)}\right| = \max_{z\in[a,b]} |r(-z)| = E_k([-b,-a]),$$

where we used the fact that $[a,b]$ is a real interval. Therefore, $\sqrt{Z_k} \leq E_k([-b,-a])$. Conversely, taking $R(z)$ as in Theorem 3.3 we know from property (c) that $|R(iz)|=1$. Therefore, $R(z)$ satisfies $R(z)\overline{R(-\overline{z})} = 1$ on the imaginary axis and by analytic continuation $R(z)\overline{R(-\overline{z})} = 1$ holds for all $z \in \mathbb{C}$. We conclude that if $p_1, \ldots, p_k \in \mathbb{C}$ are the zeros of $R(z)$, then the poles of $R(z)$ are $-\overline{p}_1, \ldots, -\overline{p}_n$. Thus, $R(z)$ from Theorem 3.3 has the form

$$R(z) = |\alpha| \prod_{j=1}^{k} \frac{z-p_j}{z+\overline{p_j}}, \qquad |\alpha| = 1.$$

This implies that $E_k([-b,-a]) \leq \max_{z\in[-b,-a]} |R(z)| \leq \sqrt{Z_k}$. Here, in the last inequality we have applied property (a). We conclude that $E_k([-b,-a]) = \sqrt{Z_k}$.

Therefore, an infinite product formula for $E_k([\eta,1])$ that corrects [38, (1.11)] is obtained by taking square roots (and setting $a/b = \eta$) in Theorem 3.1. That is, for $0 < \eta < 1$ we have[9]

$$E_k([\eta,1]) = 2\rho^{-k} \prod_{\tau=1}^{\infty} \frac{(1+\rho^{-8\tau k})^2}{(1+\rho^{4k}\rho^{-8\tau k})^2}, \qquad \rho = \exp\left(\frac{\pi^2}{2\mu(\eta)}\right), \tag{A.2}$$

where $\mu(\cdot)$ is the Grötzsch ring function. From (A.2), one obtains upper and lower bounds for $E_k([\eta,1])$ that correct [38, (1.12)], [40, (3.17)], and [45, (15)], namely,

$$\frac{2\rho^{-k}}{(1+\rho^{-4k})^2} \leq E_k([\eta,1]) \leq \frac{2\rho^{-k}}{1+\rho^{-4k}} \leq 2\rho^{-k}, \qquad \rho = \exp\left(\frac{\pi^2}{2\mu(\eta)}\right). \tag{A.3}$$

More refined estimates than (A.3) can be obtained by taking more terms from the infinite product in (A.2).

More recently, the best rational approximation of the sign function on $[-b,-a] \cup [a,b]$ has become important in numerical linear algebra because of a recursive construction of spectral projectors of matrices [34, 43]. In this setting, if $E_{m,n} := E_{m,n}([-b,-a],[a,b])$, then

$$E_{m,n} = \min_{r\in\mathcal{R}_{m,n}} \max_{z\in[-b,-a]\cup[a,b]} |r(z) - \mathrm{sgn}(z)|, \quad \mathrm{sgn}(z) = \begin{cases} 1, & z \in [a,b], \\ -1, & z \in [-b,-a]. \end{cases} \tag{A.4}$$

Unfortunately, lower and upper bounds for $E_{2k,2k} = E_{2k-1,2k}$ are claimed in [34, (4.1)] based on the erroneous infinite product formula in [40, (3.17)] and for $E_{2k+1,2k+1} = E_{2k+1,2k}$ in [43, (3.8)] by incorrectly citing the fundamental work of Gončar [31, (32)].

[9] For those readers who want to make a direct comparison to [38, (1.11)], we know from Theorem 3.1 that the q in [38, (1.11)] satisfies the relationship $q = \rho^{-4k}$.

We believe it is therefore useful to state infinite product formulas for $E_{k,k}$ and the resulting estimates. We recall from the proofs of Theorems 3.1 and 3.3 that we have

$$E_{k,k} = E_{2\lfloor (k-1)/2\rfloor+1,2\lfloor k/2\rfloor} = \frac{2\sqrt{Z_k}}{1+Z_k}, \qquad \mu\left(\frac{2\sqrt{Z_k}}{1+Z_k}\right) = \frac{\mu(Z_k)}{2}.$$

Thus, in the proof of Theorem 3.1 we select $q = \exp(-2\mu(E_{k,k})) = \rho^{-2k}$ and obtain

$$E_{k,k} = 4\rho^{-k}\prod_{\tau=1}^{\infty}\frac{(1+\rho^{-4\tau k})^4}{(1+\rho^{2k}\rho^{-4\tau k})^4}, \qquad \rho = \exp\left(\frac{\pi^2}{2\mu(a/b)}\right). \tag{A.5}$$

Again, this infinite product in (A.5) results in asymptotically tight corrected lower and upper bounds on $E_{k,k}$:

$$\frac{4\rho^{-k}}{(1+\rho^{-2k})^4} \le E_{k,k} \le \frac{4\rho^{-k}}{(1+\rho^{-2k})^2} \le 4\rho^{-k}, \qquad \rho = \exp\left(\frac{\pi^2}{2\mu(a/b)}\right). \tag{A.6}$$

Similarly, more refined estimates than (A.6) can be obtained by taking more terms from the infinite product in (A.5).

In Zolotarev's original work [66], he demonstrated an equivalence between (A.4) and the best rational approximation of $1/\sqrt{z}$ and $\sqrt{z}$ on the interval $[a^2, b^2]$. Since the rational function given in (3.6) can be written as $\tilde{r}(z) = zS(z^2)$ with a rational function S, one can verify by letting $x = z^2$ and $x = a^2b^2/y$ that [1, p. 147]

$$E_{2k,2k} = \min_{r\in\mathcal{R}_{k-1,k}}\ \max_{x\in[a^2,b^2]}\left|\left(\frac{1}{\sqrt{x}} - r(x)\right)\Big/\frac{1}{\sqrt{x}}\right|$$

and

$$E_{2k+1,2k+1} = \min_{r\in\mathcal{R}_{k,k}}\ \max_{x\in[a^2,b^2]}\left|\left(\frac{1}{\sqrt{x}} - r(x)\right)\Big/\frac{1}{\sqrt{x}}\right| = \min_{r\in\mathcal{R}_{k,k}}\ \max_{y\in[a^2,b^2]}\left|\frac{\sqrt{y}-r(y)}{\sqrt{y}}\right|.$$

It is worth noting that these are relative errors as opposed to absolute errors.

Acknowledgments. Many of the inequalities in this paper were numerically verified using RKToolbox [13], and we thank Stefan Güttel for providing support. We thank Yuji Nakatsukasa for carefully reading a draft of this manuscript and pointing us toward [34] and [43]. We also thank Sheehan Olver, Gil Strang, Marcus Webb, and Heather Wilber for discussions. Finally, we thank the referees for carefully reading the paper and providing us with excellent and helpful feedback. The authors are grateful to the SIMAX and *SIAM Review* editors for selecting this work for a SIGEST award. It is an honor to have our work appreciated in this way.

REFERENCES

[1] N. I. Akhieser, *Elements of the Theory of Elliptic Functions*, Transl. Math. Monogr. 79, AMS, Providence RI, 1990. (Cited on pp. 4, 8, 9, 23)

[2] A. C. Antoulas, D. C. Sorensen, and Y. Zhou, *On the decay rate of Hankel singular values and related issues*, Systems Control Lett., 46 (2002), pp. 323–342. (Cited on pp. 1, 2, 20)

[3] C. Badea and B. Beckermann, *Spectral sets*, in Handbook of Linear Algebra, 2nd ed., L. Hogben, ed., Chapman and Hall/CRC, Boca Raton, FL, 2013, Chapter 37. (Cited on p. 6)

[4] J. Baker, M. Embree, and J. Sabino, *Fast singular value decay for Lyapunov solutions with nonnormal coefficients*, SIAM J. Matrix. Anal. Appl., 36 (2015), pp. 656–668, https://doi.org/10.1137/140993867. (Cited on pp. 2, 6)

[5] B. Beckermann, *On the Numerical Condition of Polynomial Bases: Estimates for the Condition Number of Vandermonde, Krylov and Hankel Matrices*, Habilitationsschrift, Universität Hannover, Hannover, Germany, 1996. (Cited on p. 3)

[6] B. Beckermann, *The condition number of real Vandermonde, Krylov and positive definite Hankel matrices*, Numer. Math., 85 (2000), pp. 553–577. (Cited on pp. 3, 19)

[7] B. Beckermann, *Singular Value Estimates for Matrices with Small Displacement Rank*, Presentation at Structured Matrices, Cortona, 2004. (Cited on p. 4)

[8] B. Beckermann, *An error analysis for rational Galerkin projection applied to the Sylvester equation*, SIAM J. Numer. Anal., 49 (2011), pp. 2430–2450, https://doi.org/10.1137/110824590. (Cited on p. 2)

[9] B. Beckermann and A. Gryson, *Extremal rational functions on symmetric discrete sets and superlinear convergence of the ADI method*, Constr. Approx., 32 (2010), pp. 393–428. (Cited on pp. 3, 19)

[10] B. Beckermann and L. Reichel, *Error estimates and evaluation of matrix functions via the Faber transform*, SIAM J. Numer. Anal., 47 (2009), pp. 3849–3883, https://doi.org/10.1137/080741744. (Cited on p. 7)

[11] B. Beckermann and A. Townsend, *On the singular values of matrices with displacement structure*, SIAM J. Matrix Anal. Appl., 38 (2017), pp. 1227–1248, https://doi.org/10.1137/16M1096426. (Cited on p. 4)

[12] P. Benner, R.-C. Li, and N. Truhar, *On the ADI method for Sylvester equations*, J. Comput. Appl. Math., 233 (2009), pp. 1035–1045. (Cited on p. 6)

[13] M. Berljafa and S. Güttel, *A Rational Krylov Toolbox for MATLAB*, MIMS EPrint 2014.56, The University of Manchester, Manchester, UK, 2014. (Cited on p. 23)

[14] D. A. Bini, S. Massei, and L. Robol, *On the decay of the off-diagonal singular values in cyclic reduction*, Linear Algebra Appl., 519 (2017), pp. 27–53. (Cited on p. 4)

[15] G. Birkhoff, R. S. Varga, and D. Young, *Alternating direction implicit methods*, Adv. Comput., 3 (1962), pp. 189–273. (Cited on pp. 2, 6)

[16] D. Braess, *Nonlinear Approximation Theory*, Springer-Verlag, Berlin, Heidelberg, New York, 1986. (Cited on pp. 7, 14)

[17] D. Braess, *Rational approximation of Stieltjes functions by the Carathéodory–Fejér method*, Constr. Approx., 3 (1987), pp. 43–50. (Cited on p. 7)

[18] D. Braess and W. Hackbusch, *Approximation of $1/x$ by exponential sums in* $[1,\infty)$, IMA J. Numer. Anal., 25 (2005), pp. 685–697. (Cited on pp. 7, 14, 21)

[19] D. Braess and W. Hackbusch, *On the efficient computation of high-dimensional integrals and the approximation by exponential sums*, in Multiscale, Nonlinear and Adaptive Approximation, Springer, 2009, pp. 39–74. (Cited on pp. 14, 15, 21)

[20] E. J. Candés and B. Recht, *Exact matrix completion via convex optimization*, Found. Comput. Math., 9 (2009), pp. 717–772. (Cited on p. 1)

[21] M. Crouzeix, *Bounds for analytic functions of matrices*, Integral Equations Operator Theory, 48 (2004), pp. 461–477. (Cited on p. 6)

[22] M. Crouzeix and C. Palencia, *The numerical range is a $(1+\sqrt{2})$-spectral set*, SIAM J. Matrix. Anal. Appl., 38 (2017), pp. 649–655, https://doi.org/10.1137/17M1116672. (Cited on p. 6)

[23] A. Dutt and V. Rokhlin, *Fast Fourier transforms for nonequispaced data,* II, Appl. Comput. Harmon. Anal., 2 (1995), pp. 85–100. (Cited on p. 14)

[24] D. Fasino and V. Olshevsky, *How bad are symmetric Pick matrices?*, in Structured Matrices in Mathematics, Computer Science, and Engineering I, Contemp. Math. 280, AMS, Providence, RI, 2001, pp. 301–312. (Cited on p. 12)

[25] M. Fiedler, *Special Matrices and Their Applications in Numerical Mathematics*, Martinus Nijhoff, Leiden, The Netherlands, 1986. (Cited on p. 20)

[26] D. Fortunato and A. Townsend, *Fast Poisson Solvers for Spectral Methods*, preprint, https://arxiv.org/abs/1710.11259, 2017. (Cited on p. 21)

[27] T. Ganelius, *Rational functions, capacities and approximation*, in Aspects of Contemporary Complex Analysis (Proc. NATO Adv. Study Inst., Univ. Durham, Durham, 1979), Academic Press, London, 1980, pp. 409–414. (Cited on p. 7)

[28] W. Gautschi and G. Inglese, *Lower bounds for the condition number of Vandermonde matrices*, Numer. Math., 52 (1988), pp. 241–250. (Cited on p. 19)

[29] E. S. Gawlik, *Zolotarev Iterations for the Matrix Square Root*, preprint, https://arxiv.org/abs/1804.11000, 2018. (Cited on p. 21)

[30] G. Golub and C. Van Loan, *Matrix Computations*, 4th ed., The Johns Hopkins University Press, Baltimore, MD, 2013. (Cited on p. 2)

[31] A. A. Gončar, *Zolotarev problems connected with rational functions*, Sb. Math., 7 (1969),

pp. 623–635. (Cited on pp. 4, 7, 22)

[32] L. Grasedyck, *Existence of a low rank or $\mathcal{H}$ matrix approximant to the solution of a Sylvester equation*, Numer. Linear Algebra. Appl., 11 (2004), pp. 371–389. (Cited on p. 14)

[33] L. Greengard and V. Rokhlin, *A fast algorithm for particle simulations*, J. Comput. Phys., 73 (1987), pp. 325–348. (Cited on p. 1)

[34] S. Güttel, E. Polizzi, P. T. P. Tang, and G. Viaud, *Zolotarev quadrature rules and load balancing for the FEAST eigensolver*, SIAM J. Sci. Comput., 37 (2015), A2100–A2122, https://doi.org/10.1137/140980090. (Cited on pp. 21, 22, 23)

[35] W. Hackbusch, *A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part* I*: Introduction to $\mathcal{H}$-matrices*, Computing, 62 (1999), pp. 89–108. (Cited on p. 1)

[36] G. Heinig and K. Rost, *Algebraic Methods for Toeplitz-Like Matrices and Operators*, Oper. Theory Adv. Appl. 13, Birkhäuser Verlag, Basel, 1984. (Cited on p. 2)

[37] D. Kressner, S. Massei, and L. Robol, *Low-Rank Updates and a Divide-and-Conquer Method for Linear Matrix Equations*, preprint, https://arxiv.org/abs/1712.04349, 2017. (Cited on p. 21)

[38] V. I. Lebedev, *On a Zolotarev problem in the method of alternating directions*, USSR Comput. Math. Math. Phys., 17 (1977), pp. 58–76. (Cited on pp. 2, 6, 7, 21, 22)

[39] S. Massei, D. Palitta, and L. Robol, *Solving rank-structured Sylvester and Lyapunov equations*, SIAM J. Matrix. Anal. Appl., 39 (2018), pp. 1564–1590, https://doi.org/10.1137/17M1157155. (Cited on p. 6)

[40] A. A. Medovikov and V. I. Lebedev, *Variable time steps optimization of L_ω-stable Crank–Nicolson method*, Russian J. Numer. Anal. Math. Model., 20 (2005), pp. 283–303. (Cited on pp. 21, 22)

[41] A. Moitra, *Super-resolution, extremal functions and the condition number of Vandermonde matrices*, in Proceedings of the 47th Annual ACM Symposium on Theory of Computing, ACM, 2015, pp. 821–830. (Cited on pp. 16, 17)

[42] M. Morf, *Fast Algorithms for Multivariable Systems*, Ph.D. thesis, Stanford University, Stanford, CA, 1974. (Cited on p. 2)

[43] Y. Nakatsukasa and R. W. Freund, *Computing fundamental matrix decompositions accurately via the matrix sign function in two iterations: The power of Zolotarev's functions*, SIAM Rev., 58 (2016), pp. 461–493, https://doi.org/10.1137/140990334. (Cited on pp. 21, 22, 23)

[44] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, *NIST Handbook of Mathematical Functions*, Cambridge University Press, Cambridge, UK, 2010. (Cited on pp. 7, 9)

[45] I. V. Oseledets, *Lower bounds for separable approximations of the Hilbert kernel*, Sb. Math., 198 (2007), pp. 425–432. (Cited on pp. 14, 15, 21, 22)

[46] V. Y. Pan, *How bad are Vandermonde matrices?*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 676–694, https://doi.org/10.1137/15M1030170. (Cited on p. 19)

[47] V. Peller, *Hankel Operators and Their Applications*, Springer, New York, 2012. (Cited on p. 20)

[48] T. Penzl, *Eigenvalue decay bounds for solutions of Lyapunov equations: The symmetric case*, Systems Control Lett., 40 (2000), pp. 130–144. (Cited on pp. 2, 6, 14, 21)

[49] J. Sabino, *Solution of Large-Scale Lyapunov Equations via the Block Modified Smith Method*, Ph.D. thesis, Rice University, Houston, TX, 2006. (Cited on pp. 4, 21)

[50] E. B. Saff and V. Totik, *Logarithmic Potentials with External Fields*, Grundlehren Math. Wiss. 316, Springer, Heidelberg, 1997. (Cited on pp. 4, 11)

[51] T. Shi and A. Townsend, *On the Numerical Ranks of Tensors*, preprint, https://arxiv.org/abs/1812.09576, 2018. (Cited on pp. 3, 21)

[52] V. Simoncini, *Computational methods for linear matrix equations*, SIAM Rev., 58 (2016), pp. 377–441, https://doi.org/10.1137/130912839. (Cited on p. 4)

[53] G. Starke, *Near-circularity for the rational Zolotarev problem in the complex plane*, J. Approx. Theory, 70 (1992), pp. 115–130. (Cited on pp. 4, 11)

[54] R. C. Thompson, *Principal submatrices* IX*: Interlacing inequalities for singular values of submatrices*, Linear Algebra Appl., 5 (1972), pp. 1–12. (Cited on p. 17)

[55] A. Townsend, *Computing with Functions in Two dimensions*, D.Phil. thesis, University of Oxford, Oxford, UK, 2014. (Cited on p. 14)

[56] A. Townsend, M. Webb, and S. Olver, *Fast polynomial transforms based on Toeplitz and Hankel matrices.*, Math. Comp., 87 (2018), pp. 1913–1934. (Cited on pp. 20, 21)

[57] A. Townsend and H. Wilber, *On the singular values of matrices with high displacement rank*, Linear Algebra Appl., 548 (2018), pp. 19–41. (Cited on pp. 6, 15)

[58] N. Truhar, Z. Tomljanović, and R.-C. Li, *Analysis of the solution of the Sylvester equation*

using low-rank ADI with exact shifts, Systems Control Lett., 59 (2010), pp. 248–257. (Cited on p. 13)

[59] N. Truhar and K. Veselić, *Bounds on the trace of a solution to the Lyapunov equation with a general stable matrix*, Systems Control Lett., 56 (2007), pp. 493–503. (Cited on p. 13)

[60] E. E. Tyrtyshnikov, *Mosaic-skeleton approximations*, Calcolo, 33 (1996), pp. 47–57. (Cited on p. 14)

[61] E. E. Tyrtyshnikov, *More on Hankel Matrices*, manuscript, 1998. (Cited on p. 20)

[62] E. E. Tyrtyshnikov, *How bad are Hankel matrices?*, Numer. Math., 67 (1994), pp. 261–269. (Cited on p. 20)

[63] E. L. Wachspress, *Trail to a Lyapunov equation solver*, Comput. Math. Appl., 55 (2008), pp. 1653–1659. (Cited on p. 6)

[64] H. Wilber, A. Townsend, and G. B. Wright, *Computing with functions in spherical and polar geometries* II. *The disk*, SIAM J. Sci. Comput., 39 (2017), C238–C262, https://doi.org/10.1137/16M1070207. (Cited on p. 21)

[65] H. S. Wilf, *Finite Sections of Some Classical Inequalities*, Springer, Heidelberg, 1970. (Cited on pp. 3, 15)

[66] E. I. Zolotarev, *Application of elliptic functions to questions of functions deviating least and most from zero*, Zap. Imp. Akad. Nauk. St. Petersburg, 30 (1877), pp. 1–59. (Cited on pp. 4, 9, 23)