SIAM J. MATRIX ANAL. APPL.
Vol. 30, No. 2, pp. 762–776

PERTURBATION BOUNDS FOR DETERMINANTS AND CHARACTERISTIC POLYNOMIALS*

ILSE C. F. IPSEN† AND RIZWANA REHMAN†

Abstract. We derive absolute perturbation bounds for the coefficients of the characteristic polynomial of a $n \times n$ complex matrix. The bounds consist of elementary symmetric functions of singular values, and suggest that coefficients of normal matrices are better conditioned with regard to absolute perturbations than those of general matrices. When the matrix is Hermitian positive-definite, the bounds can be expressed in terms of the coefficients themselves. We also improve absolute and relative perturbation bounds for determinants. The basis for all bounds is an expansion of the determinant of a perturbed diagonal matrix.

Key words. elementary symmetric functions, singular values, eigenvalues, condition number, determinant

AMS subject classifications. 65F40, 65F15, 65F35, 65Z05, 15A15, 15A18

DOI. 10.1137/070704770

1. Introduction. The characteristic polynomial of a $n \times n$ complex matrix A is defined as

$$\det(\lambda I - A) = \lambda^n + c_1 \lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n,$$

where in particular $c_n = (-1)^n \det(A)$ and $c_1 = -\operatorname{trace}(A)$.

The coefficients of the characteristic polynomial of a complex matrix are of central importance in a quantum physics application. There they supply information about thermodynamic properties of fermionic systems, which arise, for instance, in the study of structure and evolution of neutron stars. These thermodynamic quantities are calculated from partition functions. It turns out that the partition function Z_k for k noninteracting fermions is given by $Z_k = (-1)^k c_k$, where the matrix A is a function of the particle Hamiltonian operator [9]. Partition functions for systems of interacting fermions require repeated calculation of noninteracting partition functions. The matrices A in these problems have fairly small dimension ($n \leq 1000$) and no discernible structure.

In order to assess the stability of numerical methods for computing the characteristic polynomial, though, we first need to know the conditioning of the c_k and their sensitivity to perturbations in the matrix A. To this end, we derive perturbation bounds for absolute normwise perturbations.

Main results. The main idea behind our perturbation bounds is a determinant expansion of a perturbed diagonal matrix (Theorem 2.3). The expansion can be extended to any square matrix via the SVD (Corollary 2.4). The resulting absolute perturbation bounds contain elementary symmetric functions of singular values. Below we present weaker, first-order versions of these bounds.

Let $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$ be the singular values of A, and let $A + E$ be a $n \times n$ complex matrix with $\|E\|_2 < 1$ and characteristic polynomial $\det(\lambda I - (A + E)) =$

*Received by the editors October 8, 2007; accepted for publication (in revised form) by A. Frommer April 3, 2008; published electronically July 2, 2008.
http://www.siam.org/journals/simax/30-2/70477.html

†Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205 (ipsen@ncsu.edu, http://www4.ncsu.edu/~ipsen/, rrehman@unity.ncsu.edu).

$\lambda^n + \tilde{c}_1\lambda^{n-1} + \cdots + \tilde{c}_{n-1}\lambda + \tilde{c}_n$. The extreme coefficients c_1 and c_n have the simplest bounds. The linearity of the trace implies

$$|\tilde{c}_1 - c_1| = |\operatorname{trace}(A+E) - \operatorname{trace}(A)| = |\operatorname{trace}(E)| \leq n\|E\|_2,$$

so that coefficient c_1 is well conditioned with regard to absolute perturbations if the matrix order n is not too large. The determinant satisfies to first order (Remark 2.9)

$$|\tilde{c}_n - c_n| = |\det(A) - \det(A+E)| \leq s_{n-1}\|E\|_2 + \mathcal{O}(\|E\|_2^2),$$

where s_{n-1} is the $(n-1)$st elementary symmetric function in the singular values and has the upper bound $s_{n-1} \leq n\sigma_1 \dots \sigma_{n-1}$.

The remaining coefficients c_k satisfy to first order (Remark 3.4)

$$|\tilde{c}_k - c_k| \leq \binom{n}{k} s_{k-1}^{(k)}\|E\|_2 + \mathcal{O}(\|E\|_2^2), \qquad 2 \leq k \leq n,$$

where $s_{k-1}^{(k)}$ is the $(k-1)$st elementary symmetric function in the k largest singular values, and $s_{k-1}^{(k)} \leq k\sigma_1 \dots \sigma_{k-1}$. However, if the matrix is normal or Hermitian, then the bound improves to (Remark 3.6)

$$|\tilde{c}_k - c_k| \leq (n-k+1)s_{k-1}\|E\|_2 + \mathcal{O}(\|E\|_2^2), \qquad 2 \leq k \leq n,$$

where s_{k-1} is the $(k-1)$st function in all singular values. Also, since A is normal, $\sigma_i = |\lambda_i|$, where λ_i are the eigenvalues. Since the binomial term $\binom{n}{k}$ is reduced to $n-k+1$, the coefficients of a normal matrix are likely to be better conditioned than those of a general matrix.

When the matrix is Hermitian positive-definite, eigenvalues are equal to singular values, and the above bound can be written as (Corollary 3.7)

$$|\tilde{c}_k - c_k| \leq (n-k+1)|c_{k-1}|\|E\|_2 + \mathcal{O}(\|E\|_2^2), \qquad 2 \leq k \leq n.$$

As a result, c_k is well conditioned in the absolute sense if the magnitude of the preceding coefficient $|c_{k-1}|$ is not too large.

Overview. Section 2 deals with determinants. We first derive expansions for determinants (section 2.1), and from them absolute perturbation bounds in terms of elementary symmetric functions of singular values (section 2.2), as well as relative bounds for determinants (section 2.3), and local sensitivity results (section 2.4). Section 3 deals with coefficients c_k of the characteristic polynomial. We derive absolute perturbation bounds for general matrices (section 3.1) and normal matrices (section 3.2), as well as normwise bounds (section 3.3).

Notation. The matrix A is a $n \times n$ complex matrix with singular values $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$, and eigenvalues λ_i, labelled so that $|\lambda_1| \geq \cdots \geq |\lambda_n|$. The two-norm is $\|A\|_2 = \sigma_1$, and A^* is the conjugate transpose of A. The matrix $I = \operatorname{diag}\begin{pmatrix}1 & \dots & 1\end{pmatrix}$ is the identity matrix, with columns e_i, $i \geq 1$. We denote by A_i the principal submatrix of order $n-1$ that is obtained by removing row and column i of A, and by $A_{i_1 \dots i_k}$ the principal submatrix of order $n-k$, obtained by removing rows and columns $i_1 \dots i_k$.

2. Determinants.

We derive expansions and perturbation bounds for determinants. We start with expansions for determinants of perturbed matrices (section 2.1), and from them derive absolute perturbation bounds in terms of elementary symmetric functions of singular values (section 2.2), as well as relative bounds for determinants (section 2.3), and local sensitivity results (section 2.4).

2.1. Expansions. We derive expansions for determinants of perturbed matrices in several steps, by considering perturbations that have only a single nonzero diagonal element (Lemma 2.1), perturbations of diagonal matrices (Theorem 2.3), and at last perturbations of general matrices (Corollary 2.4).

LEMMA 2.1. *Let A be a $n \times n$ complex matrix, α a scalar, and A_i the principal submatrix of order $n-1$ obtained by deleting row and column i of A.*

If $B = A + \alpha e_i e_i^*$, then $\det(B) = \det(A) + \alpha \det(A_i)$, $1 \le i \le n$.

Proof. This follows from a cofactor expansion [8, Theorem 2.3.1] along row i or column i of B. ∎

The above expansion can be used to expand the determinant of a perturbed diagonal matrix. Before deriving this expansion, we motivate its expression on matrices of order 2 and 3.

Example 2.2. If

$$D = \begin{pmatrix} \delta_1 & \\ & \delta_2 \end{pmatrix}, \qquad F = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix},$$

then $\det(D+F) = \det(D) + \det(F) + S_1$, where $S_1 \equiv \delta_1 f_{22} + \delta_2 f_{11}$.

If

$$D = \begin{pmatrix} \delta_1 & & \\ & \delta_2 & \\ & & \delta_3 \end{pmatrix}, \qquad F = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix},$$

then $\det(D+F) = \det(D) + \det(F) + S_1 + S_2$, where

$$S_1 \equiv \delta_1 \det \begin{pmatrix} f_{22} & f_{23} \\ f_{32} & f_{33} \end{pmatrix} + \delta_2 \det \begin{pmatrix} f_{11} & f_{13} \\ f_{31} & f_{33} \end{pmatrix} + \delta_3 \det \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix},$$

and $S_2 \equiv \delta_1\delta_2 f_{33} + \delta_1\delta_3 f_{22} + \delta_2\delta_3 f_{11}$.

These examples illustrate that the expansion of $\det(D+F)$ can be written as a sum, where each term consists of a product of k diagonal elements of D and the determinant of the "complementary" submatrix of order $n-k$ of F.

To derive expansions for diagonal matrices of any order, we denote by $F_{i_1 \ldots i_k}$ the principal submatrix of order $n-k$ obtained by deleting rows and columns $i_1 \ldots i_k$ of the $n \times n$ matrix F.

THEOREM 2.3 (expansion for diagonal matrices). *Let D and F be $n \times n$ complex matrices. If $D = \operatorname{diag}\begin{pmatrix}\delta_1 & \ldots & \delta_n\end{pmatrix}$, then*

$$\det(D+F) = \det(D) + \det(F) + S_1 + \cdots + S_{n-1},$$

where

$$S_k \equiv \sum_{1 \le i_1 < \cdots < i_k \le n} \delta_{i_1} \cdots \delta_{i_k} \det(F_{i_1 \ldots i_k}), \qquad 1 \le k \le n-1.$$

In particular, if $\delta_1 = \cdots = \delta_j = 0$ for some $1 \le j \le n-1$, then

$$\det(D+F) = \det(F) + S_1 + \cdots + S_{n-j},$$

where

$$S_k = \sum_{j+1 \le i_1 < \cdots < i_k \le n} \delta_{i_1} \cdots \delta_{i_k} \det(F_{i_1 \ldots i_k}), \qquad 1 \le k \le n-j.$$

Proof. The proof is by induction over the matrix order n, and Example 2.2 represents the induction basis. Assuming the statement is true for matrices of order $n-1$, we show that it is also true for matrices of order n. Let

$$D^{(j)} \equiv \operatorname{diag}\begin{pmatrix}0 & \dots & 0 & \delta_{j+1} & \dots & \delta_n\end{pmatrix}$$

be a diagonal matrix of order n with j leading zeros. Applying Lemma 2.1 to $A \equiv D^{(1)} + F$ and $B \equiv A + \delta_1 e_1 e_1^*$ gives

$$\det(D+F) = \delta_1 \det(D_1 + F_1) + \det(D^{(1)} + F).$$

We repeat this process on the second summand $\det(D^{(1)} + F)$ to remove the diagonal elements δ_j one by one; $j \geq 2$. To this end, we apply Lemma 2.1 to $A \equiv D^{(j)} + F$ and $B \equiv A + \delta_j e_j e_j^*$, and denote by $(D^{(j)})_j$ the matrix of order $n-1$ obtained by removing row and column j from $D^{(j)}$. This gives

$$\det(D^{(1)} + F) = \sum_{j=2}^{n-1} \delta_j \det\left((D^{(j)})_j + F_j\right) + \delta_n \det(F_n) + \det(F).$$

Putting the above expression into the expansion for $\det(D+F)$ yields

$$\det(D+F) = \det(F) + \delta_1 \det(D_1 + F_1) + \sum_{j=2}^{n-1} \delta_j \det\left((D^{(j)})_j + F_j\right) + \delta_n \det(F_n).$$

Since $D_1 + F_1$ and $(D^{(j)})_j + F_j$ are matrices of order $n-1$, we can apply the induction hypothesis. To take advantage of the fact that the $j-1$ top diagonal elements of $(D^{(j)})_j$ are zero, we define the following sums for matrices of order $n-1$,

$$S_k^{(j)} \equiv \sum_{j+1 \leq i_1 < \cdots < i_k \leq n} \delta_{i_1} \cdots \delta_{i_k} \det(F_{j i_1 \dots i_k}), \qquad 1 \leq j \leq n-1, \quad 1 \leq k \leq n-j,$$

where $F_{j i_1 \dots i_k}$ is the matrix of order $n-k-1$ obtained by removing rows and columns $j, i_1, \dots, i_k$ of F. The induction hypothesis yields

$$\begin{aligned} \det(D_1 + F_1) &= \det(D_1) + \det(F_1) + S_1^{(1)} + \cdots + S_{n-2}^{(1)}, \\ \det\left((D^{(j)})_j + F_j\right) &= \det(F_j) + S_1^{(j)} + \cdots + S_{n-j}^{(j)}, \qquad 2 \leq j \leq n-2, \\ \det\left((D^{(n-1)})_{n-1} + F_{n-1}\right) &= \det(F_{n-1}) + S_1^{(n-1)}. \end{aligned}$$

Now substitute the above expansions into the expression for $\det(D+F)$ and use the fact that $\delta_1 \det(D_1) = \det(D)$, $\sum_{i=1}^n \delta_i \det(F_i) = S_1$, and

$$\sum_{i=1}^{n-j} \delta_i S_j^{(i)} = S_{j+1}, \qquad 1 \leq j \leq n-2. \qquad \square$$

When the leading j diagonal elements of D are zero, then at most $n-j$ of the S_k are nonzero, and within each S_k one needs to account only for the nonzero summands. We now extend Theorem 2.3 to general matrices, by transforming them to diagonal form via the SVD. Let $A = U \Sigma V^*$ be a SVD of A, where $\Sigma = \operatorname{diag}\begin{pmatrix}\sigma_1 & \dots & \sigma_n\end{pmatrix}$ with $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$, and U and V are unitary.

Corollary 2.4 (expansion for general matrices). *Let A and E be $n\times n$ complex matrices, and $F \equiv U^*EV$. Then*

$$\det(A+E) = \det(A) + \det(E) + S_1 + \cdots + S_{n-1},$$

where

$$S_k \equiv \det(UV^*) \sum_{1\le i_1<\cdots<i_k\le n} \sigma_{i_1}\cdots\sigma_{i_k} \det(F_{i_1\ldots i_k}), \qquad 1\le k\le n-1.$$

If $\operatorname{rank}(A) = r$ *for some* $1\le r\le n-1$*, then*

$$\det(A+E) = \det(E) + S_1 + \cdots + S_r,$$

where

$$S_k = \det(UV^*) \sum_{1\le i_1<\cdots<i_k\le r} \sigma_{i_1}\cdots\sigma_{i_k} \det(F_{i_1\ldots i_k}), \qquad 1\le k\le r.$$

Proof. The SVD of A implies $A+E = U(\Sigma+F)V^*$, and Theorem 2.3 implies

$$\det(\Sigma+F) = \det(\Sigma) + \det(F) + \hat{S}_1 + \cdots + \hat{S}_{n-1},$$

where

$$\hat{S}_k \equiv \sum_{1\le i_1<\cdots<i_k\le n} \sigma_{i_1}\cdots\sigma_{i_k} \det(F_{i_1,\ldots,i_k}), \qquad 1\le k\le n-1.$$

With $S_k \equiv \det(UV^*)\hat{S}_k$ we obtain $\det(A+E) = \det(A) + \det(E) + S_1 + \cdots + S_{n-1}$.

Now suppose $\operatorname{rank}(A) = r\le n-1$. Then $n-r$ singular values are zero, so that all products of $r+1$ or more singular values are zero. In particular, $\det(A)=0$. If $\operatorname{rank}(A) = r < n-1$, then $S_{r+1} = \cdots = S_{n-1} = 0$. Moreover, the terms $S_1,\ldots,S_r$ contain only the nonzero singular values $\sigma_1,\ldots,\sigma_r$. ∎

Corollary 2.4 shows that the number of summands in the expansion decreases with the rank of the matrix.

2.2. Absolute perturbation bounds.

We derive absolute perturbation bounds for determinants in terms of elementary symmetric functions of singular values. These bounds give rise to absolute first-order condition numbers. We also derive simpler, but weaker normwise bounds.

To bound the perturbations we need the following inequalities.

Lemma 2.5 (Hadamard's inequality). *If B is a $n\times n$ complex matrix, then*

$$|\det(B)| \le \prod_{i=1}^{n} \|Be_i\|_2 \le \|B\|_2^n.$$

Proof. The first inequality is Hadamard's inequality [6, Corollary 7.8.2]. ∎

The bounds also contain elementary symmetric functions, which are defined as follows [6, Definition 1.2.9].

Definition 2.6 (elementary symmetric functions of singular values). *Let A be a $n\times n$ matrix with singular values $\sigma_1\ge\ldots\ge\sigma_n$. The expressions*

$$s_0 \equiv 1, \qquad s_k \equiv \sum_{1\le i_1<\cdots<i_k\le n} \sigma_{i_1}\cdots\sigma_{i_k}, \qquad 1\le k\le n,$$

are the kth elementary symmetric functions of the singular values of A.

Now we are ready to derive the first perturbation bound for determinants of general matrices.

COROLLARY 2.7 (general matrices). *Let A and E be $n \times n$ complex matrices. Then*

$$|\det(A) - \det(A+E)| \leq \sum_{i=1}^{n} s_{n-i} \|E\|_2^i.$$

If $\mathrm{rank}(A) = r$ *for some* $1 \leq r \leq n-1$, *then*

$$|\det(A+E)| \leq \|E\|_2^{n-r} \sum_{i=0}^{r} s_{r-i} \|E\|_2^i,$$

where the s_j are elementary symmetric functions in the r largest singular values of A, $1 \leq j \leq r$.

The bounds hold with equality for $E = \epsilon UV^*$ with $\epsilon > 0$, where $A = U\Sigma V^*$ is a SVD of A.

Proof. Corollary 2.4 implies $|\det(A) - \det(A+E)| \leq |\det(E)| + |S_1| + \cdots + |S_{n-1}|$. To bound $|S_k|$ use the fact that $|\det(UV^*)| = 1$ and $\sigma_i \geq 0$ to obtain

$$\begin{aligned} |S_k| &\leq \max_{1 \leq i_1 < \cdots < i_k \leq n} |\det(F_{i_1 \ldots i_k})| \sum_{1 \leq i_1 < \cdots < i_k \leq n} \sigma_{i_1} \cdots \sigma_{i_k} \\ &= \max_{1 \leq i_1 < \cdots < i_k \leq n} |\det(F_{i_1 \ldots i_k})| s_k. \end{aligned}$$

Lemma 2.5 implies $|\det(E)| \leq \|E\|_2^n$, and $|\det(F_{i_1 \ldots i_k})| \leq \|F\|_2^{n-k} = \|E\|_2^{n-k}$. Hence $|S_k| \leq s_k \|E\|_2^{n-k}$, $1 \leq k \leq n-1$.

Now suppose $\mathrm{rank}(A) = r$. Then Corollary 2.4 implies

$$|\det(A+E)| \leq |\det(E)| + |S_1| + \cdots + |S_r| \leq \|E\|_2^{n-r} \sum_{i=0}^{r} s_{r-i} \|E\|_2^i,$$

where the terms s_{r-i} contain only nonzero singular values.

If $E = \epsilon UV^*$, then $F = \epsilon I$ and $\det(F_{i_1 \ldots i_k}) = \epsilon^{n-k} = \|E\|_2^{n-k}$, so that $S_k = |S_k| = \|E\|_2^{n-k} s_k$. □

Corollary 2.7 bounds the absolute error in $\det(A+E)$ by elementary symmetric functions of singular values and powers of $\|E\|_2$. Although the bounds for nonsingular and rank-r matrices look different, because the sums start at different indices, they are consistent. If $\mathrm{rank}(A) \leq n-k$ for some $k \geq 1$, then $|\det(A+E)|$ is bounded by a multiple of $\|E\|_2^k$. Hence if $\|E\|_2 < 1$ then determinants of rank-deficient matrices tend to be better conditioned in the absolute sense.

Remark 2.8 (Hermitian positive-definite matrices). In the special case when A is Hermitian positive-definite, singular values are equal to eigenvalues, so that we can write the elementary symmetric functions in terms of the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$. Hence in Corollary 2.7

$$s_k = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \lambda_{i_1} \cdots \lambda_{i_k}, \qquad 1 \leq k \leq n-1.$$

Note that $A+E$ does not have to be Hermitian positive-definite, because no restrictions are placed on E.

Remark 2.9 (first-order absolute condition numbers). Let A be a $n \times n$ complex matrix with $\mathrm{rank}(A) \geq n-1$ and $\|E\|_2 < 1$. Corollary 2.7 implies the first-order bound

$$|\det(A) - \det(A+E)| \leq s_{n-1} \|E\|_2 + \mathcal{O}\left(\|E\|_2^2\right),$$

where $s_{n-1} \leq n\sigma_1 \ldots \sigma_{n-1}$. Hence we can view s_{n-1} or $n\sigma_1 \ldots \sigma_{n-1}$ as first-order condition numbers for absolute perturbations in A.

Example 2.10. The perturbation of a diagonally scaled Jordan block below illustrates that the first-order bound in Remark 2.9 can hold with equality. Let

$$A = \begin{pmatrix} 0 & \alpha_1 & 0 & \ldots & 0 \\ \vdots & 0 & \alpha_2 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 0 & \alpha_{n-1} \\ 0 & 0 & \ldots & \ldots & 0 \end{pmatrix}, \qquad E = \epsilon e_n e_1^*,$$

where $|\epsilon| \leq 1$ and $\alpha_i > 0$, $1 \leq i \leq n-1$. Then $|\det(A+E) - \det(A)| = \alpha_1 \ldots \alpha_{n-1}\epsilon$. Since the singular values of A are 0 and $\alpha_i > 0$, $1 \leq i \leq n-1$, we obtain $|\det(A+E) - \det(A)| = s_{n-1}\|E\|_2$.

Replacing the singular values in Corollary 2.7 by powers of $\|A\|_2$ gives the simpler, but weaker bounds below.

COROLLARY 2.11 (normwise bounds). *Let A and E be $n \times n$ complex matrices. Then*

$$\begin{aligned} |\det(A+E) - \det(A)| &\leq \sum_{i=1}^{n} \binom{n}{i} \|A\|_2^{n-i} \|E\|_2^i \\ &= (\|A\|_2 + \|E\|_2)^n - \|A\|_2^n. \end{aligned}$$

If $\operatorname{rank}(A) = r$ *for some* $1 \leq r \leq n-1$, *then*

$$\begin{aligned} |\det(A+E)| &\leq \|E\|_2^{n-r} \sum_{i=0}^{r} \binom{r}{i} \|A\|_2^{r-i} \|E\|_2^i \\ &= \|E\|_2^{n-r} (\|A\|_2 + \|E\|_2)^r. \end{aligned}$$

Proof. This follows from Corollary 2.7 and $s_{n-i} \leq \binom{n}{n-i}\|A\|_2^{n-i} = \binom{n}{i}\|A\|_2^{n-i}$, $1 \leq i \leq n-1$. □

A bound similar to the one in Corollary 2.11 was already derived in [1, section 20], [2, Problem I.6.11], [3, Theorem 4.7] for any p-norm, by taking Fréchet derivatives of wedge products. Below we give a basic proof from first principles for the two-norm.

THEOREM 2.12 (section 20 in [1], problem I.6.11 in [2], Theorem 4.7 in [3]). *Let A and E be $n \times n$ complex matrices. Then*

$$|\det(A+E) - \det(A)| \leq n\|E\|_2 \max\{\|A\|_2, \|A+E\|_2\}^{n-1}.$$

Proof. We first show the statement for a diagonal matrix. That is, if $D = \operatorname{diag}\begin{pmatrix}\delta_1 & \ldots & \delta_n\end{pmatrix}$ is diagonal, then

$$\det(D+F) = \det(D) + z, \qquad \text{where} \quad |z| \leq n\|F\|_2 \max\{\|D\|_2, \|D+F\|_2\}^{n-1}.$$

The proof is by induction. For $n = 2$

$$D = \begin{pmatrix} \delta_1 & \\ & \delta_2 \end{pmatrix}, \qquad F = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix},$$

and

$$z \equiv \det(D+F) - \det(D) = \delta_1 f_{22} + \det\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & \delta_2 + f_{22} \end{pmatrix}.$$

Lemma 2.5 implies

$$\begin{aligned}|z| &\le \|F\|_2\|D\|_2 + \left\|\begin{pmatrix} f_{11} \\ f_{21}\end{pmatrix}\right\|_2 \left\|\begin{pmatrix} f_{12} \\ \delta_2 + f_{22}\end{pmatrix}\right\|_2 \le \|F\|_2\|D\|_2 + \|F\|_2\|D+F\|_2 \\ &\le 2\|F\|_2 \max\{\|D\|_2, \|D+F\|_2\}.\end{aligned}$$

This completes the induction basis. Assuming the statement is true for matrices of order $n-1$, we show that it is also true for matrices of order n. As in the proof of Theorem 2.3, let $D^{(1)} \equiv \operatorname{diag}\begin{pmatrix}0 & \delta_2 & \dots & \delta_n\end{pmatrix}$ be the matrix obtained from D by replacing δ_1 with 0, and apply Lemma 2.1 to conclude

$$\det(D+F) = \delta_1 \det(D_1 + F_1) + \det(D^{(1)} + F).$$

Since $D_1 + F_1$ is a matrix of order $n-1$, the induction hypothesis applies and gives $\det(D_1 + F_1) = \det(D_1) + z_1$, where

$$\begin{aligned}|z_1| &\le (n-1)\|F_1\|_2 \max\{\|D_1\|_2, \|D_1 + F_1\|_2\}^{n-2} \\ &\le (n-1)\|F\|_2 \max\{\|D\|_2, \|D+F\|_2\}^{n-2}.\end{aligned}$$

Substitute the above expression into the expansion for $\det(D+F)$ to obtain

$$z \equiv \det(D+F) - \det(D) = \delta_1 z_1 + \det(D^{(1)} + F),$$

where $|\delta_1 z_1| \le (n-1)\|F\|_2 \max\{\|D\|_2, \|D+F\|_2\}^{n-1}$. Applying Lemma 2.5 to $\det(D^{(1)} + F)$ yields

$$\det(D^{(1)} + F) \le \|Fe_1\|_2 \prod_{i=2}^{n} \|(D+F)e_i\|_2 \le \|F\|_2 \|D+F\|_2^{n-1}.$$

Therefore we have proved the theorem for diagonal matrices D.

To prove the theorem for general matrices A, let $A = U\Sigma V^*$ be a SVD of A. Then $\det(A+E) = \det(UV^*)\det(\Sigma + F)$, where $F \equiv U^*EV$. Since Σ is diagonal, $\det(\Sigma + F) = \det(\Sigma) + z$, where

$$|z| \le n\|F\|_2 \max\{\|\Sigma\|_2, \|\Sigma + F\|_2\}^{n-1} = n\|E\|_2 \max\{\|A\|_2, \|A+E\|_2\}^{n-1}.$$

Hence $\det(A+E) - \det(A) = \det(UV^*)z$, and the result follows from $|\det(UV^*)| = 1$. $\square$

2.3. Relative perturbation bounds. We derive expansions for relative perturbations of determinants, as well as relative perturbation bounds that improve existing bounds.

THEOREM 2.13 (expansion). *Let A and E be $n \times n$ complex matrices. If A is nonsingular, then*

$$\frac{\det(A+E) - \det(A)}{\det(A)} = \det(A^{-1}E) + S_1 + \cdots + S_{n-1},$$

where

$$S_k \equiv \sum_{1 \le i_1 < \cdots < i_k \le n} \det((A^{-1}E)_{i_1 \dots i_k}), \qquad 1 \le k \le n-1.$$

Proof. Write $\det(A+E) = \det(A)\det(I+A^{-1}E)$ and apply Theorem 2.3 to $\det(I+A^{-1}E)$. ∎

COROLLARY 2.14 (relative perturbation bound). *Let A and E be $n \times n$ complex matrices. If A is nonsingular, then*

$$\frac{|\det(A+E)-\det(A)|}{|\det(A)|} \leq \left(\kappa\frac{\|E\|_2}{\|A\|_2}+1\right)^n - 1,$$

where $\kappa \equiv \|A\|_2\|A^{-1}\|_2$.

Proof. Apply Corollary 2.7 to

$$\frac{|\det(A+E)-\det(A)|}{|\det(A)|} = |\det(I+A^{-1}E)-\det(I)|,$$

and bound $\|A^{-1}E\|_2 \leq \kappa\|E\|_2/\|A\|_2$. ∎

Remark 2.15. Corollary 2.14 is more general and tighter than the following bound from [4, (1.6)], [5, Problem 14.15]:

$$\frac{|\det(A+E)-\det(A)|}{|\det(A)|} \leq \frac{n\kappa\|E\|_2/\|A\|_2}{1-n\kappa\|E\|_2/\|A\|_2},$$

which holds only for $n\kappa\|E\|_2/\|A\|_2 < 1$. This is true because of the following. With $q \equiv \|A^{-1}\|_2\|E\|_2 = \kappa\|E\|_2/\|A\|_2$ we can write the first term in the bound of Corollary 2.14 as

$$(q+1)^n = \sum_{i=0}^{n}\binom{n}{i}q^i \leq \sum_{i=0}^{n} n^i q^i \leq \sum_{i=0}^{\infty}(nq)^i.$$

If $nq < 1$, then $\sum_{i=0}^{\infty}(nq)^i = \frac{1}{1-nq}$, so that

$$(q+1)^n - 1 \leq \frac{1}{1-nq} - 1 = \frac{nq}{1-nq}.$$

This implies for the bound in Corollary 2.14

$$\left(\kappa\frac{\|E\|_2}{\|A\|_2}+1\right)^n - 1 \leq \frac{n\kappa\|E\|_2/\|A\|_2}{1-n\kappa\|E\|_2/\|A\|_2},$$

where the last expression is the bound in [4, inequality (1.6)], [5, Problem 14.15].

2.4. Local sensitivity. We derive a local condition number for determinants from directional derivatives. The directional derivative for $\det(A)$ in the direction E is $\frac{d^k}{dx^k}\det(A+xE)$.

Although we derive the expressions below from the expansion in Theorem 2.3, we could have also used the expression for derivatives of $A(x)$ in [7, equation (6.5.9)].

THEOREM 2.16. *Let A and E be $n \times n$ complex matrices, $F \equiv U^*EV$, and x a real scalar. Then*

$$\det(A+xE) = \sum_{i=1}^{n} S_{n-i}x^i + \det(A),$$

where

$$S_0 \equiv \det(E), \qquad S_k \equiv \det(UV^*)\sum_{1\leq i_1<\cdots<i_k\leq n}\sigma_{i_1}\cdots\sigma_{i_k}\det(F_{i_1\ldots i_k}), \quad 1 \leq k \leq n-1,$$

and

$$\frac{d^k}{dx^k}\det(A+xE)|_{x=0} = k! S_{n-k}, \qquad 1 \le k \le n.$$

Proof. If $D = \operatorname{diag}\begin{pmatrix}\delta_1 & \dots & \delta_n\end{pmatrix}$ is a diagonal matrix, then Theorem 2.3 implies $\det(D+xF) = \det(xF)+\tilde{S}_1+\cdots+\tilde{S}_{n-1}+\det(D)$, where $\det(xF) = x^n\det(F) = x^n S_0$ and

$$\tilde{S}_k = \sum_{1\le i_1<\cdots<i_k\le n} \delta_{i_1}\cdots\delta_{i_k}\det(xF_{i_1\dots i_k}) = x^{n-k}S_k.$$

To derive the expansion for a general matrix, use the SVD as in Corollary 2.4. ◻

The first derivative gives the local condition number of the determinant with regard to small perturbations.

COROLLARY 2.17 (local condition number). *Let A and E be $n\times n$ complex matrices, and x a real scalar. Then*

$$\left|\frac{d}{dx}\det(A+xE)|_{x=0}\right| \le s_{n-1}\|E\|_2, \qquad \text{where} \quad s_{n-1} \le n\sigma_1\dots\sigma_{n-1}.$$

Proof. Theorem 2.16 implies for the first derivative

$$\frac{d}{dx}\det(A+xE)|_{x=0} = \det(UV^*)\sum_{1\le i_1<\cdots<i_{n-1}\le n}\sigma_{i_1}\cdots\sigma_{i_{n-1}}\det(F_{i_1\dots i_{n-1}}),$$

where $F_{i_1\dots i_{n-1}}$ is a diagonal element of F. Lemma 2.5 implies $|\det(F_{i_1\dots i_{n-1}})| \le \|F\|_2 = \|E\|_2$. ◻

Corollary 2.17 shows that the sensitivity of $\det(A)$ to small perturbations in any direction E is determined by s_{n-1} or $n\sigma_1\dots\sigma_{n-1}$. A comparison with Remark 2.9 shows that the local condition number for $\det(A)$ is identical to the first-order condition number.

3. Characteristic polynomial. Based on the determinant results in section 2, we derive absolute perturbation bounds for the coefficients of the characteristic polynomial for general matrices (section 3.1) and normal matrices (section 3.2), as well as simpler, but weaker normwise bounds (section 3.3).

Applying Theorem 2.3 to the characteristic polynomial

$$\det(\lambda I - A) = \lambda^n + c_1\lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n$$

of the $n\times n$ matrix A gives the well-known expressions [6, Theorem 1.2.12]

$$c_{n-k} = (-1)^{n-k}\sum_{1\le i_1<\dots<i_k\le n}\det(A_{i_1\dots i_k}), \quad 0\le k\le n-1,$$

where $A_{i_1\dots i_k}$ is the principal submatrix of order $n-k$ obtained by deleting rows and columns $i_1\dots i_k$ of A. The characteristic polynomial of the perturbed matrix $A+E$ is

$$\det(\lambda I - (A+E)) = \lambda^n + \tilde{c}_1\lambda^{n-1} + \cdots + \tilde{c}_{n-1}\lambda + \tilde{c}_n,$$

where $\tilde{c}_n = (-1)^n\det(A+E)$ and

$$\tilde{c}_{n-k} = (-1)^{n-k}\sum_{1\le i_1<\dots<i_k\le n}\det(A_{i_1\dots i_k}+E_{i_1\dots i_k}), \qquad 1\le k\le n-1.$$

The following example illustrates that products of singular values play an important role in the conditioning of the coefficients c_k.

Example 3.1 (companion matrices). The $n \times n$ matrix

$$A = \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \dots & \alpha_n \\ \eta & 0 & \dots & \dots & 0 \\ 0 & \eta & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \eta & 0 \end{pmatrix}, \qquad \eta > 0,$$

is a multiple of a companion matrix, and let $E = e_1 \begin{pmatrix} \epsilon & \dots & \epsilon \end{pmatrix}$ with $\epsilon > 0$. The respective coefficients of the characteristic polynomials of A and $A+E$ are [5, section 28.6]

$$c_i = \alpha_i \eta^{i-1}, \qquad \tilde{c}_i = (\alpha_i + \epsilon)\eta^{i-1}, \qquad 1 \leq i \leq n.$$

Then $|\tilde{c}_i - c_i| = \epsilon \eta^{i-1}$, $1 \leq i \leq n$. The singular values of A are [5, section 28.6]

$$\sigma_1^2 = \frac{1}{2}\left(\alpha + \sqrt{\alpha^2 - 4|\alpha_n|^2}\right), \qquad \sigma_n^2 = \frac{1}{2}\left(\alpha - \sqrt{\alpha^2 - 4|\alpha_n|^2}\right),$$

where $\alpha \equiv 1 + |\alpha_1|^2 + \cdots + |\alpha_n|^2$, and $\sigma_i = \eta$, $2 \leq i \leq n-1$. Therefore the conditioning of the coefficients c_k is determined by products of singular values.

The products of singular values in our perturbation bounds are expressed in terms of elementary symmetric functions of only the largest singular values of A.

DEFINITION 3.2 (elementary symmetric functions in the largest singular values). *Let A be a $n \times n$ matrix with singular values $\sigma_1 \geq \ldots \geq \sigma_n$. Denote by*

$$s_0^{(k)} \equiv 1, \qquad s_j^{(k)} \equiv \sum_{1 \leq i_1 < \ldots < i_j \leq k} \sigma_{i_1} \ldots \sigma_{i_j}, \qquad 1 \leq j \leq k, \quad 1 \leq k \leq n,$$

where $s_j^{(n)} = s_j$. The expression $s_j^{(k)}$ is the jth elementary symmetric function in the k largest singular values of A.

3.1. General matrices. We use the determinant expansion in Corollary 2.4 to derive perturbation bounds for coefficients c_k of general matrices.

THEOREM 3.3 (general matrices). *Let A and E be $n \times n$ complex matrices. Then*

$$|\tilde{c}_k - c_k| \leq \binom{n}{k} \sum_{i=1}^{k} s_{k-i}^{(k)} \|E\|_2^i, \qquad 1 \leq k \leq n.$$

If $\operatorname{rank}(A) = r$ *for some* $1 \leq r \leq n-1$, *then*

$$|\tilde{c}_k - c_k| \leq \binom{n}{k} \|E\|_2^{k-r} \sum_{i=0}^{r} s_{r-i}^{(k)} \|E\|_2^i, \qquad r+1 \leq k \leq n.$$

Proof. In the perturbed coefficient

$$\tilde{c}_{n-k} = (-1)^{n-k} \sum_{1 \leq i_1 < \ldots < i_k \leq n} \det(A_{i_1 \ldots i_k} + E_{i_1 \ldots i_k}),$$

the matrices $A_{i_1\ldots i_k}+E_{i_1\ldots i_k}$ are of order $n-k$. Fix the indices $i_1,\ldots,i_k$; set $B\equiv A_{i_1\ldots i_k}$ and $F\equiv E_{i_1\ldots i_k}$; and let $\mu_1\geq\ldots\geq\mu_{n-k}$ be the singular values of B. Corollary 2.4 implies $\det(B+F)=\det(B)+\det(F)+S_1+\cdots+S_{n-k-1}$, where

$$S_j=\sum_{1\leq i_1<\ldots<i_j\leq n-k}\mu_{i_1}\ldots\mu_{i_j}\det(F_{i_1\ldots i_j}),\qquad 1\leq j\leq n-k-1.$$

Since B is a submatrix of A, the singular values interlace [6, Theorem 7.3.9], so that $\sigma_j\geq\mu_j$, $1\leq j\leq n-k$. With Lemma 2.5 we obtain $|S_j|\leq s_j^{(n-k)}\|E\|_2^{n-k-j}$. Hence $|S_1|+\cdots+|S_{n-k-1}|\leq\sum_{i=1}^{n-k}s_{n-k-i}^{(n-k)}\|E\|_2^i$. Summing up the terms associated with all $\binom{n}{k}$ submatrices $A_{i_1\ldots i_k}+E_{i_1\ldots i_k}$ gives the desired bound for $|\tilde{c}_{n-k}-c_{n-k}|$.

Now suppose $\operatorname{rank}(A)=r\leq n-1$. Since r singular values are nonzero, the elementary symmetric functions $s_j^{(k)}$ in the k largest singular values remain unchanged for $k\leq r$.

Since $n-r$ singular values are equal to zero, all products of $r+1$ or more singular values are zero. Hence for $k\geq r+1$ we have $s_j^{(k)}=0$ whenever $j\geq r+1$, so that

$$\sum_{i=1}^{k}s_{k-i}^{(k)}\|E\|_2^i=\|E\|_2^{k-r}\sum_{i=0}^{r}s_{r-i}^{(k)}\|E\|_2^i.$$

Moreover, for $j\leq r$ the $s_j^{(k)}$ are functions of the r largest singular values only, so that $s_j^{(k)}=s_j^{(r)}$. Therefore $\sum_{i=0}^{k}s_{k-i}^{(k)}\|E\|_2^i=\|E\|_2^{k-r}\sum_{i=0}^{r}s_{r-i}^{(r)}\|E\|_2^i$, giving the desired bound for $|\tilde{c}_k-c_k|$ when $k\geq r+1$. ▯

For the two extreme coefficients, Theorem 3.3 produces the expected bounds: In the case of $c_n=(-1)^n\det(A)$, the bound coincides with the determinant bound in Corollary 2.7, while for $c_1=-\operatorname{trace}(A)$ we obtain $|\tilde{c}_1-c_1|\leq n\|E\|_2$. Theorem 3.3 shows that the conditioning of c_k with regard to absolute perturbations is determined by the binomial term $\binom{n}{k}$ and the elementary symmetric functions in the k largest singular values. The binomial coefficient is largest for c_k with $k\approx n/2$, because $\binom{n}{n-k}=\binom{n}{k}$, and $\binom{n}{k}$ is monotonically increasing for $k<n/2$. In particular, if n is even, then for $k=n/2$ we have $k\binom{n}{k}\geq k\left(\frac{n}{k}\right)^k=n2^{n/2-1}$.

If $\operatorname{rank}(A)=r\leq n-2$, then the bounds for the coefficients $c_{r+1},\ldots,c_n$ contain higher powers of $\|E\|_2$. Hence if $\|E\|_2<1$, then the coefficients $c_{r+1},\ldots,c_n$ of rank-deficient matrices tend to be better conditioned in the absolute sense.

Remark 3.4 (first-order absolute condition numbers for general matrices). Theorem 3.3 implies for $\|E\|_2<1$ the first-order bound

$$|\tilde{c}_k-c_k|\leq\binom{n}{k}s_{k-1}^{(k)}\|E\|_2+\mathcal{O}(\|E\|_2^2),\qquad 1\leq k\leq n,$$

where $s_{k-1}^{(k)}\leq k\sigma_1\ldots\sigma_{k-1}$. Hence we can view $\binom{n}{k}s_{k-1}^{(k)}$ or $\binom{n}{k}k\sigma_1\ldots\sigma_{k-1}$ as first-order condition numbers for absolute perturbations in the coefficient c_k.

3.2. Normal matrices. We show that for normal matrices, the conditioning of the coefficients improves because the binomial term is smaller, and the elementary symmetric functions depend on all singular values, not just the largest ones. Note that all statements for normal matrices apply in particular to Hermitian matrices.

THEOREM 3.5 (normal matrices). *If the $n\times n$ matrix A is normal, then*

$$|\tilde{c}_k-c_k|\leq\sum_{i=1}^{k}\binom{n-k+i}{i}s_{k-i}\|E\|_2^i,\qquad 1\leq k\leq n.$$

The bound holds with equality if $E = \epsilon I$ with $\epsilon > 0$.

Proof. Since A is normal, it has an eigenvalue decomposition $A = V\Lambda V^*$, where $\Lambda = \operatorname{diag}\begin{pmatrix}\lambda_1 & \cdots & \lambda_n\end{pmatrix}$ is complex diagonal, $|\lambda_1| \geq \ldots \geq |\lambda_n|$, and V is unitary. Set $D \equiv \lambda I - \Lambda$ and $F \equiv -V^*EV$, so that $\det(\lambda I - (A+E)) = \det(D+F)$. Theorem 2.3 implies $\det(D+F) = \det(D) + \det(F) + S_1 + \cdots + S_{n-1}$. Substituting $\det(D) = \lambda^n + \sum_{k=1}^n c_k\lambda^{n-k}$ and $\det(D+F) = \lambda^n + \sum_{k=1}^n \tilde{c}_k\lambda^{n-k}$ in the above expansion gives

$$\sum_{k=1}^{n} (\tilde{c}_k - c_k)\lambda^{n-k} = \det(F) + S_1 + \cdots + S_{n-1}.$$

Thus $\tilde{c}_k - c_k$ is equal to the coefficient of λ^{n-k} on the right-hand side, i.e., in $\det(F) + S_1 + \cdots + S_{n-1}$. Since

$$S_{n-j} \equiv \sum_{1\leq i_1<\cdots<i_{n-j}\leq n} (\lambda - \lambda_{i_1})\cdots(\lambda - \lambda_{i_{n-j}})\det(F_{i_1\ldots i_{n-j}}), \qquad 1 \leq j \leq n-1,$$

has as highest power λ^{n-j}, the term λ^{n-k} can occur only in $S_{n-k}, \ldots, S_{n-1}$. This means $\tilde{c}_k - c_k$ is the sum of the coefficients of λ^{n-k} in $S_{n-1}, \ldots, S_{n-k}$. To bound the coefficient of λ^{n-k} in S_{n-j} in particular, we first bound all coefficients in S_{n-j}.

Observe that S_{n-j} is a sum of $\binom{n}{n-j}$ products $(\lambda - \lambda_{i_1})\cdots(\lambda - \lambda_{i_{n-j}})$. For fixed $i_1, \ldots, i_{n-j}$ we can write the product as

$$(\lambda - \lambda_{i_1})\cdots(\lambda - \lambda_{i_{n-j}}) = \lambda^{n-j} + \gamma_1\lambda^{n-j-1} + \cdots + \gamma_{n-j-1}\lambda + \gamma_{n-j}.$$

The coefficient γ_l is a sum of $\binom{n-j}{l}$ products $\lambda_{j_1}\ldots\lambda_{j_l}$. Hence S_{n-j} contains $\binom{n}{n-j}\binom{n-j}{l}$ such products. Therefore we can bound $|S_{n-j}|$ by a sum of $\binom{n}{n-j}\binom{n-j}{l}$ products $|\lambda_{j_1}|\ldots|\lambda_{j_l}|$. Since A is normal $|\lambda_i| = \sigma_i$, so that these products are also summands of the elementary symmetric function s_l. The sum s_l contains $\binom{n}{l}$ such summands. Therefore the number of occurrences of s_l in the bound for $|S_{n-j}|$ is $\binom{n}{n-j}\binom{n-j}{l}/\binom{n}{l} = \binom{n-l}{j}$.

Now we are ready to return to the coefficient of λ^{n-k} in particular; it is γ_{k-j}. Applying the above counting argument with $l = k-j$ shows that the coefficient of λ^{n-k} in S_{n-j} is bounded by $\binom{n-k+j}{j}s_{k-j}|\det(F_{i_1\ldots i_{n-j}})|$. Lemma 2.5 implies $|\det(F_{i_1\ldots i_{n-j}})| \leq \|F\|_2^j = \|E\|_2^j$. Summing up the contributions from all S_{n-j}, $1 \leq j \leq k$, gives the desired result.

If $E = \epsilon I$, then $F = \epsilon I$ and $\det(F_{i_1\ldots i_k}) = \epsilon^{n-k} = \|E\|_2^{n-k}$. □

Remark 3.6 (first-order absolute condition numbers for normal matrices). If A is normal and $\|E\|_2 < 1$, then Theorem 3.5 implies the first-order bound

$$|\tilde{c}_k - c_k| \leq (n-k+1)s_{k-1}\|E\|_2 + \mathcal{O}(\|E\|_2^2), \qquad 1 \leq k \leq n,$$

where $s_{k-1} \leq k|\lambda_1\ldots\lambda_{k-1}|$. Hence we can view $(n-k+1)s_{k-1}$ or $(n-k+1)k|\lambda_1\ldots\lambda_{k-1}|$ as first-order condition numbers for absolute perturbations in the coefficient c_k.

For Hermitian positive-define matrices, the bound in Theorem 3.5 can be expressed in terms of the coefficients c_k.

COROLLARY 3.7 (Hermitian positive-definite matrices). *If the $n \times n$ matrix A is Hermitian positive-definite, then*

$$|\tilde{c}_k - c_k| \leq \sum_{i=1}^{k} \binom{n-k+i}{i} |c_{k-i}|\|E\|_2^i, \qquad 1 \leq k \leq n.$$

Proof. The coefficients c_k are also elementary symmetric functions in the eigenvalues [6, section 1.2], and the eigenvalue of a Hermitian positive-definite is equal to the singular values. Thus $c_k = (-1)^k s_k$, and the result follows from Theorem 3.5. ∎

To first order, the conditioning of coefficient c_k is determined by the magnitude of the preceding coefficient, $|c_{k-1}|$. As in Corollary 2.8, the matrix $A+E$ in Corollary 3.7 does not have to be Hermitian positive-definite, because E can be arbitrary. Below we illustrate that one cannot use the expression in Corollary 3.7 for indefinite matrices; that is, positive-definiteness of A is crucial for the expression in Corollary 3.7.

Example 3.8. Corollary 3.7 is not valid for indefinite Hermitian matrices and in particular matrices with zero trace.

To see this, let

$$A = \begin{pmatrix} \alpha & \\ & -\alpha \end{pmatrix}, \qquad \tilde{A} = \begin{pmatrix} \alpha - \epsilon & \\ & -\alpha + \epsilon \end{pmatrix},$$

where $\alpha > 0$ and $\epsilon > 0$. The characteristic polynomials are

$$\det(\lambda I - A) = \lambda^2 - \alpha^2, \qquad (\lambda I - (A + E)) = \lambda^2 - (\alpha - \epsilon)^2,$$

so that $\tilde{c}_2 - c_2 = 2\alpha\epsilon - \epsilon^2$. However, $|\tilde{c}_2 - c_2|$ cannot be bounded in terms of c_1, as required by Corollary 3.7, because $c_1 = 0$.

3.3. Normwise bounds. Replacing the singular value products by powers of $\|A\|_2$ gives the following simpler, but weaker bounds.

COROLLARY 3.9 (normwise bounds). *Let A and E be $n \times n$ complex matrices. Then*

$$\begin{aligned} |\tilde{c}_k - c_k| &\le k \binom{n}{k} \sum_{i=1}^{k} \binom{k}{i} \|A\|_2^{k-i} \|E\|_2^i, \\ &= \binom{n}{k} \left((\|A\|_2 + \|E\|_2)^k - \|A\|^k \right), \qquad 1 \le k \le n. \end{aligned}$$

If $\operatorname{rank}(A) = r$ *for some* $1 \le r \le n-1$*, then*

$$\begin{aligned} |\tilde{c}_k - c_k| &\le k \binom{n}{k} \|E\|_2^{k-r} \sum_{i=1}^{r} \binom{k}{i} \|A\|_2^{r-i} \|E\|_2^i, \\ &= \binom{n}{k} \|E\|_2^{k-r} \left((\|A\|_2 + \|E\|_2)^r - \|A\|^r \right), \qquad r+1 \le k \le n. \end{aligned}$$

Proof. This follows from Theorem 3.3 and

$$s_{k-i}^{(k)} \le \binom{k}{k-i} \|A\|_2^{k-i} = \binom{k}{i} \|A\|_2^{k-i}, \qquad 1 \le i \le k-1. \quad ∎$$

A similar bound was was already derived in [1, section 20] and [2, Problem I.6.11] for any p-norm, by taking Fréchet derivatives of wedge products. Below we give a basic proof from first principles for the two-norm.

THEOREM 3.10 (section 20 in [1], problem I.6.11 in [2]). *Let A and E be $n \times n$ complex matrices. Then*

$$|\tilde{c}_k - c_k| \le k \binom{n}{k} \|E\|_2 \max\{\|A\|_2, \|A + E\|_2\}^{k-1}, \qquad 1 \le k \le n.$$

Proof. As in the proof of Theorem 3.3, we use

$$\tilde{c}_{n-k} = (-1)^{n-k} \sum_{1\le i_1<\ldots<i_k\le n} \det(A_{i_1\ldots i_k} + E_{i_1\ldots i_k}).$$

This gives for the absolute error

$$|\tilde{c}_{n-k} - c_{n-k}| \le \sum_{1\le i_1<\ldots<i_k\le n} |\det(A_{i_1\ldots i_k} + E_{i_1\ldots i_k}) - \det(A_{i_1\ldots i_k})|.$$

Theorem 2.12 implies that $|\det(A_{i_1\ldots i_k} + E_{i_1\ldots i_k}) - \det(A_{i_1\ldots i_k})|$ is bounded by

$$(n-k)\|E_{i_1\ldots i_k}\|_2 \max\{\|A_{i_1\ldots i_k}\|_2, \|(A+E)_{i_1\ldots i_k}\|_2\}^{n-k-1}.$$

Bounding the principal submatrices by the norms of the respective matrices and recognizing that the sum contains $\binom{n}{n-k}$ summands yields the desired bound. ∎

Acknowledgment. We thank Dean Lean for many helpful discussions and for bringing this problem to our attention.

REFERENCES

[1] R. Bhatia, *Perturbation Bounds for Matrix Eigenvalues*, Longman Scientific & Technical, New York, 1987.
[2] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.
[3] S. Friedland, *Variation of tensor powers and permanents*, Linear Multilinear Algebra, 12 (1982), pp. 81–98.
[4] S. K. Godunov, A. G. Antonov, O. P. Kiriljuk, and V. I. Kostin, *Guaranteed accuracy in numerical linear algebra*, Math. Appl. 252, Kluwer Academic Publishers, Dordrecht, 1993 (in English).
[5] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, Philadelphia, 2002.
[6] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.
[7] R. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, London, 1991.
[8] R. Horn, C. Johnson, P. Lancaster, and M. Tismenetsky, *The Theory of Matrices, 2 ed.*, Academic Press, Orlando, 1985.
[9] D. Lee and T. Schaefer, *Neutron matter on the lattice with pionless effective field theory*, Phys. Rev. C, 72 (2005), 024006.