Operator Theory:
Advances and Applications, Vol. 199, 203–220

A Priori Estimates on the Structured Conditioning of Cauchy and Vandermonde Matrices

Enrico Bozzo and Dario Fasino

Abstract. We analyze the componentwise and normwise sensitivity of inverses of Cauchy, Vandermonde, and Cauchy-Vandermonde matrices, with respect to relative componentwise perturbations in the nodes defining these matrices. We obtain *a priori*, easily computable upper bounds for these condition numbers. In particular, we improve known estimates for Vandermonde matrices with generic real nodes; we consider in detail Vandermonde matrices with nonnegative or symmetric nodes; and we extend the analysis to the class of complex Cauchy-Vandermonde matrices.

Mathematics Subject Classification (2000). Primary 15A12; Secondary 15A57, 65F35.

Keywords. Condition number, displacement structure, Cauchy matrix, Vandermonde matrix.

1. Introduction

A structured matrix is, in some sense, a matrix whose entries depend on a small set of parameters. Clearly, this dependence is somehow preserved by matrix inversion. The main question addressed here is: How sensitive is the inverse of a structured matrix to perturbations in its parameters?

The answer to the above question is what we call *structured conditioning*. This paper focuses on the structured conditioning of Cauchy, Vandermonde and Cauchy-Vandermonde matrices, which have been one of the main research topics of Georg Heinig, see, e.g., [9, 12, 13]. The main reasons motivating this study are the analysis of the influence of data errors in the solution of linear systems with structured matrices, and the assessment of stability properties of fast algorithms for solving such linear systems. These algorithms act directly on the set of parameters defining a structured matrix rather than on its entries, see, e.g., [3, 4, 5, 9, 10, 12, 15], hence their stability characteristics, which are sometimes surprising, should be examined in the light of suitable refinements of classical condition numbers, taking in due consideration both more accurate measures of perturbations and the structure of the problem considered, as observed, e.g., in [1, 10, 11, 14, 15, 22]. Indeed, suppose that the solution computed by such an algorithm in finite precision is the exact solution of a similarly structured system defined by slightly perturbed parameters. To assess the quality of the computed solution with no further assumptions on the right-hand side, we should make use of some measure of the sensitivity of matrix inversion with respect to perturbation in its parameters. Theoretical bases of this argument can be found in [14], where such a *backward structured error analysis* in introduced rather formally for structured matrices whose dependence on the parameters is linear, and in [1, 15, 22], where a detailed analysis has been carried out for Vandermonde systems, mainly motivated by the stability analysis of the Björk-Pereyra algorithm. Other results in the same streamline can be found in [7, 25], concerning possibly rectangular Cauchy, Vandermonde, Toeplitz and Hankel matrices.

By the way, the results in the above-mentioned papers are mainly aimed at characterizing the structured conditioning in terms of exact (possibly generalized) inverse matrices and solutions of suitable linear systems, hence they are *a posteriori* results, and the resulting expressions may be hard to compute.

The goal of this paper is to obtain *a priori* bounds for the structured conditioning of Cauchy, Vandermonde and other related matrices, that are easily computable right from their parameters, with no involvement of exact (generalized) inverse matrices. Besides to their pervasive occurrence in computations with polynomial and rational functions, Cauchy and Vandermonde matrices play an important role in deriving structural and computational properties of many relevant matrix classes with displacement structure, see, e.g., [9, 13, 17, 18, 19]. For example, they occur as fundamental blocks (together with trigonometric transforms) in decomposition formulas for Toeplitz, Hankel, and related matrices. In this paper, we will pay particular attention to the structured conditioning of the individual columns of their inverses, as they have a special relevance in polynomial and rational interpolation problems. Indeed, let $x_1 \dots x_n$ be pairwise distinct points in the complex plane, considered as parameters, and let $\phi_1(x) \dots \phi_n(x)$ be a fixed set of functions such that the collocation matrix $X \equiv (\phi_j(x_i))$ is nonsingular. In fact, Vandermonde, Cauchy, and Cauchy-Vandermonde matrices arise as collocation matrices when the functions $\phi_j(x)$ are monomials or particular rational functions. Let $X^{-1} \equiv (v_{i,j})$. Then, the functions $\ell_k(x) = \sum_j v_{j,k}\phi_j(x)$ for $1 \le k \le n$ are the Lagrange functions for the interpolation problem defined by the points x_i and the functions $\phi_j(x)$. Hence, a further reason for investigating the structured conditioning of X^{-1} is to give a measure of the sensitivity of the functions $\ell_k(x)$ with respect to perturbations in the interpolation points x_i.

After giving in Section 2 a quick look at Cauchy matrices, we consider Vandermonde matrices in Section 3. There, we will improve the result in [11, Thm. 1] on the mixed structured conditioning of Vandermonde matrices (see Corollary 2), and consider in detail Vandermonde matrices with nonnegative nodes and symmetric nodes. In Section 4 we will extend our analysis to Cauchy-Vandermonde matrices with complex nodes.

1.1. Main definitions and notations

We borrow from [10, 11] the following definitions and notations. Let p and q be positive integers, and let F be a (densely defined) continuous function $F : \mathbb{C}^p \mapsto \mathbb{C}^q$. One usually defines the *normwise condition number* of F in a given point $x \in \mathbb{C}^p$ as

$$\kappa(F,x) = \lim_{\varepsilon \to 0} \sup_{\|x-\tilde{x}\| \le \varepsilon} \frac{\|F(x) - F(\tilde{x})\|}{\|F(x)\|} \frac{\|x\|}{\|x-\tilde{x}\|}.$$

In the limit as the perturbation size tends to zero, this number gives the worst possible magnification of the quantity $\|x-\tilde{x}\|/\|x\|$ in the computation of F. Here and in what follows, $\|\cdot\| = \|\cdot\|_\infty$, unless otherwise noted.

The number $\|x-\tilde{x}\|/\|x\|$ is the *relative normwise distance* between x and $\tilde{x}$. In what follows, we also consider the *relative componentwise distance* between two points $x, \tilde{x} \in \mathbb{C}^p$, defined as

$$\delta(x,\tilde{x}) = \min\{\varepsilon : |x_i - \tilde{x}_i| \le \varepsilon |x_i|, 1 \le i \le p\}. \tag{1}$$

The above definition is of interest in numerical analysis, because it is the most appropriate way to measure errors induced by the finite precision representation of machine numbers. Observe that if an entry of the vector x is zero, the corresponding entry in $\tilde{x}$ must be zero for $\delta(x,\tilde{x})$ be definite: Componentwise relative perturbations do not affect null entries.

Accordingly, we consider the *componentwise condition number* of F in $x \ne 0$ as

$$c(F,x) = \lim_{\varepsilon \to 0} \sup_{\delta(x,\tilde{x}) \le \varepsilon} \frac{\delta(F(x), F(\tilde{x}))}{\delta(x,\tilde{x})}. \tag{2}$$

Throughout this paper, we consider matrix and vector moduli and inequalities as applied componentwise. Hence, we observe that $c(F,x)$ is characterized by the following inequality:

$$|F(x) - F(\tilde{x})| \le |F(x)| c(F,x) \delta(x,\tilde{x}) + o(\delta(x,\tilde{x})). \tag{3}$$

In some cases, it may be also of interest to consider the *mixed condition number*

$$m(F,x) = \lim_{\varepsilon \to 0} \sup_{\delta(x,\tilde{x}) \le \varepsilon} \frac{\|F(x) - F(\tilde{x})\|}{\|F(x)\|} \frac{1}{\delta(x,\tilde{x})}. \tag{4}$$

Since $\|x-\tilde{x}\|/\|x\| \le \delta(x,\tilde{x})$, we have both $m(F,x) \le c(F,x)$ and $m(F,x) \le \kappa(F,x)$, but in general $c(F,x)$ and $\kappa(F,x)$ are unrelated. Moreover, the equivalent definition

$$\delta(x,\tilde{x}) = \sup_{D \text{ diagonal}} \frac{\|D(x-\tilde{x})\|}{\|Dx\|}$$

implies that $m(F,x)$ results by minimizing $\kappa(F,x)$ with respect to all argument normalizations, while $c(F,x) = \sup_D m(DF,x)$ is a “worst case” measure of the sensitivity of F in presence of diagonal scalings in both the parameters and the function values.

If the map F is differentiable, the condition numbers introduced above can be related to its differential as follows: For any n-vector $a = (a_1, \dots, a_n)^T$, let D_a denote the $n \times n$ diagonal matrix whose ith diagonal entry is a_i. Then, if F' denotes the differential of F, we have

$$\kappa(F, x) = \|F'(x)\| \|x\| / \|F(x)\| \tag{5}$$

$$c(F, x) = \|D_{F(x)}^{-1} F'(x) D_x\| \tag{6}$$

$$m(F, x) = \|F'(x) D_x\| / \|F(x)\|. \tag{7}$$

Some further notations are used throughout this paper: Let e_i be the ith standard basis vector, whose order will be made clear from the context, and $\mathbf{1} = (1, \dots, 1)^T$. Let $\mathrm{Vec} : \mathbb{C}^{n \times n} \mapsto \mathbb{C}^{n^2}$ be the operator such that $\mathrm{Vec}(X)$ is the n^2-order vector obtained by stacking downward the columns of X, into one long column. Moreover, for $x = (x_1, \dots, x_p)^T \in \mathbb{C}^p$ and $y = (y_1, \dots, y_q)^T \in \mathbb{C}^q$, let

$$\Delta(x) = \max_{i \neq j} \frac{|x_i|}{|x_i - x_j|}, \qquad \Delta(x, y) = \max_{i,j} \frac{|x_i|}{|x_i - y_j|}. \tag{8}$$

In the above equations, fractions with vanishing denominators assume the value $+\infty$, whatever the numerators are. In particular, $\Delta(x) < +\infty$ if and only if the points $x_1 \dots x_n$ are pairwise distinct, and similarly for $\Delta(x, y)$.

1.2. Basic displacement structured matrices

Cauchy, Vandermonde and Cauchy-Vandermonde matrices are among the best known matrices with a *displacement structure*. Such kind of structure is defined in terms of a *displacement operator*

$$\mathcal{L}_{M,N}(X) = MX - XN,$$

where M, N are two fixed $n \times n$ matrices with disjoint spectra, so that the operator $\mathcal{L}_{M,N}$ is invertible. Throughout this paper, we will deal with matrix spaces having the following form:

$$\mathcal{D}_{M,N} = \{X : \mathrm{rank}(\mathcal{L}_{M,N}(X)) = 1\}.$$

A basic fact that will be used here to reduce the analysis of Vandermonde and Cauchy-Vandemonde matrices to the simpler case of Cauchy matrices is this: If $N = SDS^{-1}$ and $X \in \mathcal{D}_{M,N}$, then $XS \in \mathcal{D}_{M,D}$.

Regarding the connection between displacement structure and conditioning, we recall that in the paper [20] an exponentially growing lower bound is derived for the spectral conditioning of matrices X such that $AX + XA^T = -BB^T$, where B has low rank and all eigenvalues of A have negative real part. By Lyapunov theorem, any such matrix X is symmetric and positive definite; remarkably, the displacement structure induced by the operator $\mathcal{L}_{A,-A^T}$ forces X to be very badly conditioned.

2. Cauchy matrices

Let $x = (x_1, \dots, x_n)^T \in \mathbb{C}^n$ and $y = (y_1, \dots, y_n)^T \in \mathbb{C}^n$ have pairwise distinct entries, with $x_i \neq y_j$ for $1 \leq i, j \leq n$. The *Cauchy matrix* associated with x and y is defined by $C_{x,y} \equiv (1/(x_i - y_j))$. Since the displacement operator $\mathcal{L}_{D_x,D_y}$ is nonsingular, $C_{x,y}$ is the unique solution of the displacement equation $D_x C_{x,y} - C_{x,y} D_y = \mathbf{1}\mathbf{1}^T$. The following explicit formula for the entries of the inverse of $C_{x,y}$ is well known, see, e.g., [10]: For $C_{x,y}^{-1} \equiv (v_{i,j})$ we have

$$v_{i,j} = \frac{\prod_l (y_i - x_l)}{\prod_{l \neq i} (y_i - y_l)} \frac{1}{y_i - x_j} \frac{\prod_l (x_j - y_l)}{\prod_{l \neq j} (x_j - x_l)}. \tag{9}$$

Remark that $v_{i,j} \neq 0$. In [10] the above expression is differentiated with respect to x_i and y_i in order to bound the componentwise conditioning of the inversion of $C_{x,y}$, via (6). In the next result we show a simple bound for the componentwise conditioning of the columns of $C_{x,y}^{-1}$, when only the x node vector is subject to perturbations. This result plays a fundamental role in the subsequent sections. Moreover, as mentioned in the Introduction, it is useful to estimate the sensitivity of a set of rational Lagrange functions with respect to the interpolation nodes. It is apparent from the definition of $C_{x,y}$ that we obtain the corresponding result for y by simply considering the transpose matrix.

Theorem 1. *Let $1 \leq i \leq n$ and $y \in \mathbb{C}^n$ be fixed, with pairwise distinct entries. Let $F_i : \mathbb{C}^n \mapsto \mathbb{C}^n$ be defined as $F_i(x) = C_{x,y}^{-1} e_i$. Then,*

$$c(F_i, x) \leq 2(n-1)\Delta(x) + (2n-1)\Delta(x, y).$$

Proof. From (6) we have

$$c(F_i, x) = \max_{1 \leq j \leq n} \sum_k \left| \frac{x_k}{v_{j,i}} \frac{\partial v_{j,i}}{\partial x_k} \right| = \max_{1 \leq j \leq n} \sum_k \left| x_k \frac{\partial}{\partial x_k} \log |v_{j,i}| \right|, \tag{10}$$

where $v_{j,i}$ is given in (9). For $k \neq i$ we have:

$$\left| x_k \frac{\partial \log |v_{j,i}|}{\partial x_k} \right| = \left| - \frac{x_k}{x_k - x_i} \frac{x_k}{x_k - y_j} \right| \leq \Delta(x) + \Delta(x, y).$$

On the other hand,

$$\left| x_i \frac{\partial \log |v_{j,i}|}{\partial x_i} \right| = \left| \sum_{l \neq i} \frac{x_i}{x_i - x_l} - \sum_l \frac{x_i}{x_i - y_l} \right| \leq (n-1)\Delta(x) + n\Delta(x, y).$$

Plugging the latter inequalities into (10) we arrive at the claim. □

For notational simplicity, in the sequel we use the shorthand

$$\Delta(n, x, y) = 2(n-1)\Delta(x) + (2n-1)\Delta(x, y). \tag{11}$$

In the next corollary we consider the structured conditioning of the matrices in $\mathcal{D}_{D_x,D_y}$. Observe that any matrix in this set can be expressed as $D_1 C_{x,y} D_2$ for some diagonal matrices D_1 and D_2.

Corollary 1. *Let $F : \mathbb{C}^n \mapsto \mathbb{C}^{n^2}$ be defined as $F(x) = \mathrm{Vec}(D_1 C_{x,y}^{-1} D_2)$, where D_1, D_2 are arbitrary diagonal matrices and y is as in the preceding theorem. Then,*

$$c(F, x) \le \Delta(n, x, y),$$

where $\Delta(n, x, y)$ is defined in (11).

Proof. Scaling the entries of $C_{x,y}^{-1}$ by constant factors does not affect their conditioning, as it should be clear from the definitions (1) and (2). Hence we can assume $D_c = D_d = I$. In this case, we observe that

$$c(F, x) = \max_{1 \le i \le n} c(F_i, x) \le \Delta(n, x, y),$$

where F_i is as in the preceding theorem. □

If F is as in the preceding corollary and $D_c = D_d = I$, then $F'(x)$ is the map $z \mapsto \mathrm{Vec}(C_{x,y}^{-1} E(z) C_{x,y}^{-1})$, where $E(z) \equiv (-z_i/(x_i - y_j)^2)$. Hence, in view of (5), the normwise conditioning $\kappa(F, x)$ is essentially driven by the (unstructured) conditioning of $C_{x,y}$, that is, $\|C_{x,y}\| \|C_{x,y}^{-1}\|$. Quite few estimates are currently available for the latter, and only for special vectors x and y. Here we only mention the paper [23], dealing with the spectral conditioning of the Cauchy-Toeplitz matrix $C_T \equiv (1/(a+i-j))$, and [8], for Cauchy matrices $C_{-y,y}$ with positive vector y (note that these two examples fall into the generic case considered in [20] and mentioned in Subsection 1.2). Generally, all these matrices are very ill conditioned, and this fact should be contrasted with the slowly-growing estimate for the componentwise conditioning shown in the above corollary; for example, if y has positive entries we have (see [8])

$$\|C_{-y,y}\|_2 \|C_{-y,y}^{-1}\|_2 > \left(\frac{y_n + y_1}{y_n - y_1} \right)^{2n-2}.$$

Note that also the exponentially ill-conditioned Hilbert matrix $C_H \equiv (1/(i+j-1))$ is a special Cauchy matrix. For that matrix, an $O(n^2)$ bound for its structured conditioning is derived in [10], while its spectral conditioning grows roughly like 34^n [2].

3. Vandermonde matrices

Given a vector $x = (x_1, \ldots, x_n)^T \in \mathbb{C}^n$ with pairwise distinct entries, for any fixed $0 \le \theta < 2\pi$ the Vandermonde matrix $V_x \equiv (x_i^{j-1})$ fulfills the displacement equation

$$\mathcal{L}_{D_x, P_\theta}(V_x) = D_x V_x - V_x P_\theta = \begin{pmatrix} x_1^n - \mathrm{e}^{\mathrm{i}n\theta} \\ \vdots \\ x_n^n - \mathrm{e}^{\mathrm{i}n\theta} \end{pmatrix} e_n^T,$$

where i is the imaginary unit, and

$$P_\theta = \begin{pmatrix} 0 & \cdots & 0 & \mathrm{e}^{\mathrm{i}n\theta} \\ 1 & \ddots & 0 & 0 \\ & \ddots & \ddots & \vdots \\ O & & 1 & 0 \end{pmatrix}.$$

Recall that the spectral decomposition of P_θ is explicitly computable,

$$P_\theta = \mathrm{e}^{\mathrm{i}\theta} D(\theta)\Phi D_\xi \Phi^{-1} D(\theta)^{-1}, \tag{12}$$

where $\xi = (1, \mathrm{e}^{\mathrm{i}2\pi/n}, \ldots, \mathrm{e}^{\mathrm{i}(n-1)2\pi/n})^T$, $\Phi = \sqrt{1/n}\, V_\xi$ is the unitary Fourier matrix of order n, and $D(\theta) = \mathrm{Diag}(1, \mathrm{e}^{-\mathrm{i}\theta}, \ldots, \mathrm{e}^{-\mathrm{i}(n-1)\theta})$.

Only for notational simplicity, in what follows we suppose that the operator $\mathcal{L}_{D_x,P_0}$ is nonsingular, that is, $\Delta(x,\xi) < +\infty$. This hypothesis is of no restriction, and can always be fulfilled by a suitable rotation of the complex plane. Indeed, let $\hat{x} = \mathrm{e}^{-\mathrm{i}\theta}x$; we have $V_x D(\theta) = V_{\hat{x}}$ and, from (12),

$$\begin{aligned} \mathcal{L}_{D_x,P_\theta}(V_x) &= [D_x V_x D(\theta) - \mathrm{e}^{\mathrm{i}\theta} V_x D(\theta) P_0] D(\theta)^{-1} \\ &= \mathrm{e}^{\mathrm{i}\theta}[D_{\hat{x}} V_{\hat{x}} - V_{\hat{x}} P_0] D(\theta)^{-1} \\ &= \mathrm{e}^{\mathrm{i}n\theta} \mathcal{L}_{D_{\hat{x}},P_0}(V_{\hat{x}}). \end{aligned}$$

The last equation uses the fact that $e_n^T D(\theta)^{-1} = \mathrm{e}^{\mathrm{i}(n-1)\theta} e_n^T$. Hence, the discussion of the Vandermonde matrix with node vector x, for a chosen matrix P_θ, can be carried out equivalently by considering the Vandermonde matrix defined by $\hat{x}$ and the displacement equation with the matrix P_0. Since the entries of $V_{\hat{x}}$ and $V_{\hat{x}}^{-1}$ have the same modulus of the entries of V_x and V_x^{-1}, respectively, the rotation $x \mapsto \hat{x}$ leaves unaltered the conditioning properties we are investigating.

In contrast to what happens with Cauchy matrices, inverse Vandermonde matrices can have zero entries, hence the componentwise conditioning of Vandermonde matrices cannot be bounded in general. For this reason, we will consider the mixed conditioning, in the generic case, as in [11, 15, 22]. Furthermore, we will obtain estimates for the componentwise conditioning for particular configurations of the nodes, namely, nonnegative or symmetric nodes.

We will use the following lemma to bridge Cauchy and Vandermode matrices. The statement can be obtained as a consequence of Proposition 3.2 in [9]. For convenience, we provide here a short and self-contained proof.

Lemma 1. *Let $x = (x_1, \ldots, x_n)^T \in \mathbb{C}^n$ and $a = (x_1^n - 1, \ldots, x_n^n - 1)^T$. If D_a is nonsingular, the matrix V_x can be factorized as follows:*

$$V_x = \frac{1}{\sqrt{n}} D_a C_{x,\xi} D_\xi^{-1} \Phi^{-1}.$$

Proof. From (12) we have

$$D_x V_x \Phi - V_x \Phi D_\xi = (D_x V_x - V_x P_0)\Phi = a e_n^T \Phi = \frac{1}{\sqrt{n}} a \mathbf{1}^T D_\xi^{-1}.$$

Since diagonal matrices commute, we have

$$\begin{aligned}\mathcal{L}_{D_x,D_\xi}(D_a^{-1}V_x\Phi D_\xi) &= D_a^{-1}[D_xV_x\Phi - V_x\Phi D_\xi]D_\xi \\ &= \sqrt{1/n}\,D_a^{-1}a\mathbf{1}^TD_\xi^{-1}D_\xi \\ &= \sqrt{1/n}\,\mathbf{1}\mathbf{1}^T.\end{aligned}$$

By hypothesis, the nodes x_i are not roots of unit, hence the operator $\mathcal{L}_{D_x,D_\xi}$ is invertible. Thus $D_a^{-1}V_x\Phi D_\xi = \sqrt{1/n}\,C_{x,\xi}$, whence we obtain the thesis. □

Observe that, if the hypothesis on D_a is false, then the matrix $C_{x,\xi}$ is not even defined. Indeed, the same hypothesis can be restated as $\Delta(x,\xi) < +\infty$.

Lemma 2. *Let $F : \mathbb{C}^n \mapsto \mathbb{C}^n$ be the map $F(v) = \Phi v$. Then, $m(F, v) = \sqrt{n}$.*

Proof. See [6, Corollary 1]. □

Theorem 2. *Let $1 \le i \le n$ be fixed, and let $F_i : \mathbb{C}^n \mapsto \mathbb{C}^n$ be the function defined as $F_i(x) = V_x^{-1}e_i$. Furthermore, let $\xi = (1, \mathrm{e}^{\mathrm{i}2\pi/n}, \ldots, \mathrm{e}^{\mathrm{i}(n-1)2\pi/n})^T$. Then, for any $0 \le \theta < 2\pi$ we have*

$$m(F_i, x) \le \sqrt{n}\left(\Delta(n, \mathrm{e}^{-\mathrm{i}\theta}x, \xi) + \left|\frac{nx_i^n}{x_i^n - \mathrm{e}^{\mathrm{i}n\theta}}\right|\right),$$

where $\Delta(n, \mathrm{e}^{-\mathrm{i}\theta}x, \xi)$ can be obtained from (11).

Proof. In the light of the argument outlined at the beginning of this section, we can restrict the proof to the case $\theta = 0$, since the general case follows by considering the matrix $V_{\hat{x}}$, with $\hat{x} = \mathrm{e}^{-\mathrm{i}\theta}x$. From Lemma 1 we obtain $V_x^{-1} = \sqrt{n}\Phi D_\xi C_{x,\xi}^{-1}D_a^{-1}$, where $a = (x_1^n - 1, \ldots, x_n^n - 1)^T$. Hence, for $1 \le i \le n$,

$$F_i(x) = \frac{\sqrt{n}}{x_i^n - 1}\Phi D_\xi C_{x,\xi}^{-1}e_i.$$

Consider the decomposition $F_i(x) = G(H_i(x))$, where $G(x) = \Phi x$ and $H_i(x) = \sqrt{n}(x_i^n - 1)^{-1}D_\xi C_{x,\xi}^{-1}e_i$. It can be shown that $m(F_i, x) \le m(G, H_i(x))c(H_i, x)$, see [11, p. 692]. Then, from Lemma 2 we obtain $m(F_i, x) \le \sqrt{n}c(H_i, x)$.

Furthermore, for x and $\tilde{x}$ such that $\delta(x, \tilde{x}) = \epsilon$, we have from Theorem 1

$$\begin{aligned}|H_i(x) - H_i(\tilde{x})| &\le \sqrt{n}|x_i^n - 1|^{-1}|C_{x,\xi}^{-1}e_i - C_{\tilde{x},\xi}^{-1}e_i| \\ &\quad + \sqrt{n}|(x_i^n - 1)^{-1} - (\tilde{x}_i^n - 1)^{-1}||C_{\tilde{x},\xi}^{-1}e_i| \\ &\le \epsilon\Delta(n, x, \xi)|H_i(x)| + \epsilon\left|\frac{nx_i^n}{x_i^n - 1}\right||H_i(x)| + o(\epsilon).\end{aligned}$$

From the equivalence of (3) and (2) we have

$$c(H_i, x) \le \Delta(n, x, \xi) + \left|\frac{nx_i^n}{x_i^n - 1}\right|,$$

completing the proof. □

Note that the factor $\sqrt{n}$ in the right-hand side of the thesis of the preceding theorem can be dropped off, if the 2-norm is used instead of the ∞-norm in the definition of $m(F_i, x)$. Indeed, one proves easily that $m_2(G, x) = \|G'(x)D_x\|_2/\|G(x)\|_2 = 1$, and we have the attainable upper bound $m_2(F_i, x) \leq c(H_i, x)$, in the notations of the preceding proof. Hence, in some sense, the Euclidean norm is more appropriate than the ∞-norm to analyze the structured conditioning of V_x.

In the case of real nodes, Gohberg and Koltracht proved in [11] the upper bound

$$m(F_i, x) \leq n^2 \max(n\Delta(x), n + \Delta(x)).$$

We improve this bound in the following corollary.

Corollary 2. *In the notations of Theorem 2, if $x \in \mathbb{R}^n$ we have*

$$m(F_i, x) < \sqrt{n}(2(n-1)\Delta(x) + 2n^2).$$

Proof. Let $\theta = \pi/(2n)$. We have:

$$\Delta(x, \mathrm{e}^{\mathrm{i}\theta}\xi) \leq \frac{1}{\sin(\theta)} < \frac{\pi}{2\theta} = n.$$

Hence from (11)

$$\Delta(n, \mathrm{e}^{-\mathrm{i}\theta}x, \xi) = \Delta(n, x, \mathrm{e}^{\mathrm{i}\theta}\xi) < 2(n-1)\Delta(x) + (2n-1)n.$$

Moreover,

$$\left| \frac{nx_i^n}{x_i^n - \mathrm{e}^{\mathrm{i}n\theta}} \right| = n \left| \frac{x_i^n}{x_i^n - \mathrm{i}} \right| < n,$$

and the claim follows from Theorem 2. □

We consider the complete inverse of V_x in the next corollary, which follows from an argument analogous to that used in Corollary 1.

Corollary 3. *Let $F : \mathbb{C}^n \mapsto \mathbb{C}^{n^2}$ be defined as $F(x) = \mathrm{Vec}(V_x^{-1}D_d)$, where D_d is an arbitrary diagonal matrix whose entries do not depend on x. Then, for any $0 \leq \theta < 2\pi$,*

$$m(F, x) \leq \sqrt{n} \left(\Delta(n, \mathrm{e}^{-\mathrm{i}\theta}x, \xi) + n \max_{1\leq i\leq n} \left| \frac{x_i^n}{x_i^n - \mathrm{e}^{\mathrm{i}n\theta}} \right| \right),$$

where $\Delta(n, \mathrm{e}^{-\mathrm{i}\theta}x, \xi)$ can be obtained from (11).

Proof. Since a constant scaling of the columns of V_x^{-1} does not affect their conditioning, we suppose $D_d = I$. Let F_i be the map introduced in Theorem 2 for $1 \leq i \leq n$. Then, $F(x) = (F_1(x), \ldots, F_n(x))^T$. For arbitrary $x, \tilde{x}$ we have

$$\frac{\|F(x) - F(\tilde{x})\|}{\|F(x)\|} = \frac{\max_i \|F_i(x) - F_i(x)\|}{\max_i \|F_i(x)\|} \leq \max_{1\leq i\leq n} \frac{\|F_i(x) - F_i(x)\|}{\|F_i(x)\|}.$$

Hence, from the definition (4) we obtain $m(F, x) \leq \max_i m(F_i, x)$, and the claim follows from Theorem 2. □

3.1. Vandermonde matrices with nonnegative nodes

If $x_k > 0$ for $1 \le k \le n$ then V_x^{-1} has no zero entries, and we can obtain precise upper bounds for its structured componentwise conditioning. Indeed, let $V_x^{-1} \equiv (v_{i,j})$. Then, it is well known that

$$v_{j,i} = (-1)^{n-j}\sigma_{n-j}^{(n-1)}(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\prod_{l\neq i}(x_i - x_l)^{-1},$$

see, e.g., [11], where $\sigma_i^{(n)}(a_1,\ldots,a_n)$ is the ith *elementary symmetric function* on n variables,

$$\sigma_i^{(n)}(a_1,\ldots,a_n) = \sum_{1\le j_1<\cdots<j_i\le n} a_{j_1}a_{j_2}\cdots a_{j_i}. \tag{13}$$

By definition (13), for any $1 \le k \le n$ we have

$$\begin{aligned}\sigma_i^{(n)}(a_1,\ldots,a_n) &= a_k\sigma_{i-1}^{(n-1)}(a_1,\ldots,a_{k-1},a_{k+1},\ldots,a_n)\\ &\quad + \sigma_i^{(n-1)}(a_1,\ldots,a_{k-1},a_{k+1},\ldots,a_n).\end{aligned}$$

Recall that $\sigma_0^{(n)} \equiv 1$ and $\sigma_i^{(n)} \equiv 0$ for $i > n$ or $i < 0$. Furthermore, we have

$$\sum_{k=1}^{n} a_k\sigma_i^{(n-1)}(a_1,\ldots,a_{k-1},a_{k+1},\ldots,a_n) = (i+1)\sigma_{i+1}^{(n)}(a_1,\ldots,a_n). \tag{14}$$

Actually, we can allow one entry in the vector x, say x_1, to be zero. In this case, the first row of V_x^{-1} is parallel to e_1^T, but no other zeros are introduced in V_x^{-1}; moreover, owing to the definition (1), x_1 is untouched by relative perturbations, hence the zero entries in V_x^{-1} don't vary. In the light of the preceding facts, we obtain the following bound for the componentwise conditioning of the columns of V_x^{-1} with nodes in $\mathbb{R}_+ = \{x \ge 0\}$:

Theorem 3. *Let $F_i : \mathbb{R}_+^n \mapsto \mathbb{C}^n$ be the function defined as $F_i(x) = V_x^{-1}e_i$, for any fixed $1 \le i \le n$. Then*

$$c(F_i, x) \le (n-1)(2\Delta(x) + 1).$$

Proof. From (6) we have

$$c(F_i, x) = \max_{1\le j\le n}\sum_k \left|\frac{x_k}{v_{j,i}}\frac{\partial v_{j,i}}{\partial x_k}\right|.$$

For $k \neq i$ we have:

$$\frac{x_k}{v_{j,i}}\frac{\partial v_{j,i}}{\partial x_k} = \frac{x_k\sigma_{n-j-1}^{(n-2)}(x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_{i-1},x_{i+1},\ldots,x_n)}{\sigma_{n-j}^{(n-1)}(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)} - \frac{x_k}{x_i - x_k}.$$

Moreover,

$$\frac{x_i}{v_{j,i}}\frac{\partial v_{j,i}}{\partial x_i} = x_i\frac{\partial}{\partial x_i}\log|v_{j,i}| = -\sum_{l\neq i}\frac{x_i}{x_i - x_l}.$$

Owing to the positivity of $x_1 \dots x_n$, from (14) we obtain

$$\sum_{k \neq i} \left| \frac{x_k \sigma_{n-j-1}^{(n-1)}(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_{i-1}, x_{i+1}, \dots, x_n)}{\sigma_{n-j}^{(n-1)}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)} \right| = n - j.$$

Hence, we obtain

$$c(F_i, x) \leq \max_{1 \leq j \leq n} |n - j| + \sum_{k \neq i} \left| \frac{x_k}{x_i - x_k} \right| + \left| \sum_{l \neq i} \frac{x_i}{x_i - x_l} \right|,$$

and the claim follows from the triangle inequality and the definitions (8). □

The next corollary follows by an argument analogous to the ones exploited in the proof of Corollary 1 and 3, hence we omit it.

Corollary 4. *Let $F : \mathbb{R}_+^n \mapsto \mathbb{C}^{n^2}$ be defined as $F(x) = \mathrm{Vec}(D_c V_x^{-1} D_d)$, where D_c, D_d are arbitrary diagonal matrices whose entries do not depend on x. Then, $c(F, x) \leq (n-1)(2\Delta(x) + 1)$.*

The preceding results should be contrasted with well-known lower bounds on the conditioning of Vandermonde matrices with real or positive nodes [2, 16, 21, 24], which are exponentially growing functions in the order n. Moreover, we observe that the results in this subsections are trivially extended to the case where the points $x_1 \dots x_n$ belong to a ray in the complex plane, $x_i = |x_i|\omega$, where $\omega = \mathrm{e}^{\mathrm{i}\theta}$, as it is apparent from the factorization $V_x = V_{\omega^{-1}x} \mathrm{Diag}(1, \omega \dots \omega^{n-1})$.

3.2. Vandermonde matrices with symmetric nodes

When the nodes are restricted to be real, in many circumstances they are also symmetrically located with respect to zero. Indeed, symmetric configurations arise naturally when the nodes are Fekete points or zeros of special polynomial sequences (e.g., orthogonal polynomials from symmetric weights), or when one attempts to minimize the (classical) conditioning of Vandermonde matrices, see [16].

For the sake of simplicity, we consider the Vandermonde matrix $V_{(x,-x)}$, of order $2n$, whose nodes are $x_1, \dots, x_n, -x_1, \dots, -x_n$ with $x_i > 0$ (however, consider that the structured conditioning is invariant under permutation of the nodes). Introducing the vector $\hat{x} = (x_1^2, \dots, x_n^2)^T$, we have

$$V_{(x,-x)} = \begin{pmatrix} V_x & D_x^n V_n \\ V_{-x} & D_{-x}^n V_{-n} \end{pmatrix} = \begin{pmatrix} I & I \\ I & -I \end{pmatrix} \begin{pmatrix} V_{\hat{x}} & O \\ O & D_x V_{\hat{x}} \end{pmatrix} \Pi^T,$$

where Π is the perfect shuffle permutation matrix. We obtain

$$\begin{aligned} V_{(x,-x)}^{-1} &= \frac{1}{2} \Pi \begin{pmatrix} V_{\hat{x}}^{-1} & O \\ O & (D_x V_{\hat{x}})^{-1} \end{pmatrix} \begin{pmatrix} I & I \\ I & -I \end{pmatrix} \\ &= \frac{1}{2} \Pi \begin{pmatrix} V_{\hat{x}}^{-1} & (D_x V_{\hat{x}})^{-1} \\ V_{\hat{x}}^{-1} & -(D_x V_{\hat{x}})^{-1} \end{pmatrix}, \end{aligned}$$

and we see from the above decomposition that every entry of the matrix $V_{(x,-x)}^{-1}$ coincides, apart of a constant, with one entry of either $V_{\hat{x}}^{-1}$ or $(D_x V_{\hat{x}})^{-1}$. On the basis of this argument we obtain the following result:

Corollary 5. *Let $F : \mathbb{R}_+^n \mapsto \mathbb{C}^{4n^2}$ be defined as $F(x) = \mathrm{Vec}(D_c V_{(x,-x)}^{-1} D_d)$, where D_c, D_d are diagonal matrices of order $2n$, whose entries do not depend on x. Moreover, for $x = (x_1, \ldots, x_n)^T$, let $\hat{x} = (x_1^2, \ldots, x_n^2)^T$. Then,*

$$c(F, x) \leq 2(n-1)(2\Delta(\hat{x}) + 1) + 1.$$

Proof. We can set $D_c = D_d = I$, without loss in generality. Let $F^{(1)}, F^{(2)} : \mathbb{R}_+^n \mapsto \mathbb{C}^{n^2}$ be defined as $F^{(1)}(x) = \mathrm{Vec}(V_{\hat{x}}^{-1})$, and $F^{(2)}(x) = \mathrm{Vec}((D_x V_{\hat{x}})^{-1})$. By virtue of the preceding argument we have

$$c(F, x) \leq \max\{c(F^{(1)}, x), c(F^{(2)}, x)\}.$$

Our goal reduces to obtain upper bounds for $c(F^{(1)}, x)$ and $c(F^{(2)}, x)$.

First, observe that $F^{(1)}(x) = G(H(x))$, where $G(x) = \mathrm{Vec}(V_x^{-1})$ and $H(x) = \hat{x}$. Hence $c(F^{(1)}, x) \leq c(G, H(x))c(H, x)$, see [11, p. 692]. From Corollary 4 we have $c(G, H(x)) \leq (n-1)(2\Delta(\hat{x})+1)$. Moreover, from (6) we obtain $c(H, x) = 2$. Hence $c(F^{(1)}, x) \leq 2(n-1)(2\Delta(\hat{x}) + 1)$.

In order to estimate the componentwise conditioning of $F^{(2)}$, denote $V_{\hat{x}}^{-1} \equiv (\hat{v}_{i,j})$. Then $(D_x V_{\hat{x}})^{-1} \equiv (\hat{v}_{i,j}/x_j)$. Let fix one particular pair (i, j), and consider the scalar function $F_{i,j}^{(2)}(x) = \hat{v}_{i,j}/x_j$, considering $\hat{v}_{i,j}$ a function of x. Clearly we have $c(F^{(2)}, x) = \max_{i,j} c(F_{i,j}^{(2)}, x)$. Again using (6) we obtain

$$\begin{aligned} c(F_{i,j}^{(2)}, x) &= \sum_k \left| \frac{x_k}{F_{i,j}^{(2)}(x)} \frac{\partial F_{i,j}^{(2)}(x)}{\partial x_k} \right| \\ &\leq \sum_{k \neq j} \left| \frac{x_k}{\hat{v}_{i,j}} \frac{\partial \hat{v}_{i,j}}{\partial x_k} \right| + \left| \frac{x_j^2}{\hat{v}_{i,j}} \frac{\partial}{\partial x_j} \left(\frac{\hat{v}_{i,j}}{x_j} \right) \right| \\ &\leq c(F^{(1)}, x) + 1. \end{aligned}$$

By Corollary 4, we have $c(F^{(2)}, x) \leq 2(n-1)(2\Delta(\hat{x}) + 1) + 1$ and the proof is complete. □

Observe that in the above corollary we have

$$\Delta(\hat{x}) = \max_{i \neq j} \frac{x_i^2}{|x_i^2 - x_j^2|} = \max_{i \neq j} \frac{x_i}{|x_i - x_j|} \frac{x_i}{x_i + x_j} < \Delta(x).$$

4. Cauchy-Vandermonde matrices

Let $x = (x_1, \ldots, x_n)^T \in \mathbb{C}^n$ and $y = (y_1, \ldots, y_k)^T \in \mathbb{C}^k$, where $0 < k < n$, have pairwise distinct entries and $x_i \neq y_j$ for all i, j. The $n \times n$ matrix

$$K_{x,y} = \left(\begin{array}{ccc|cccc} \frac{1}{x_1 - y_1} & \cdots & \frac{1}{x_1 - y_k} & 1 & x_1 & \cdots & x_1^{n-k-1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{x_n - y_1} & \cdots & \frac{1}{x_n - y_k} & 1 & x_n & \cdots & x_n^{n-k-1} \end{array} \right)$$

is called *Cauchy-Vandermonde matrix*. Matrices with the above structure appear in connection with rational interpolation problems for functions with prescribed poles, see [9, 12, 18, 19]. In particular, in [9, 18, 19] special factorizations of the inverse of $K_{x,y}$ are deduced from representation formulas of the rational function solving a particular interpolation problem; such factorizations allows to compute the solution of linear systems with a Cauchy-Vandermonde matrix at the cost of $O(n^2)$ or even $O(n \log^2 n)$ arithmetic operations. Other low-cost algorithms for such systems, based on recursions of Levinson and Schur type, are presented in [12].

For any $0 \leq \theta < 2\pi$, the matrix $K_{x,y}$ fulfills the displacement equation

$$D_x K_{x,y} - K_{x,y} N_\theta = \begin{pmatrix} x_1^{n-k} - \mathrm{e}^{\mathrm{i}(n-k)\theta} \\ \vdots \\ x_n^{n-k} - \mathrm{e}^{\mathrm{i}(n-k)\theta} \end{pmatrix} e_n^T, \qquad N_\theta = \begin{pmatrix} D_y & O \\ e_1 \mathbf{1}^T & P_\theta \end{pmatrix},$$

where the matrix N_θ has square diagonal blocks and P_θ is as in (12). Analogously to the Vandermonde case, we can limit ourselves to consider the case $\theta = 0$. Indeed, if we let $\omega = \mathrm{e}^{-\mathrm{i}\theta}$, we have

$$K_{\omega x, \omega y} = K_{x,y} \Omega, \qquad \Omega = \begin{pmatrix} \omega^{-1} I & O \\ O & \mathrm{Diag}(1, \omega, \ldots, \omega^{n-k-1}) \end{pmatrix}. \tag{15}$$

Since Ω is diagonal and unitary, the entries of $K_{\omega x, \omega y}$ and its inverse have the same modulus of the corresponding entries of $K_{x,y}$ and its inverse, respectively, hence their structured conditioning is the same.

We stress the fact that, differently to the case of Cauchy and Vandermonde matrices, no closed-form formulas are known for the entries of $K_{x,y}^{-1}$, for generic k. Indeed, the inversion formulas presented in [9] involve a polynomial division, and the inversion algorithms introduced in [12, 18, 19] have a recursive character, namely, the entries of $K_{x,y}^{-1}$ are computed according to suitable orderings. Hence, it is not obvious how to study the structured conditioning of $K_{x,y}$ by a straightforward use of the definitions. In the sequel we will exploit a factorization approach analogous to the one already introduced in the preceding section.

Lemma 3. *Let $\xi = (\xi_1, \ldots, \xi_{n-k})^T$, $\xi_j = \mathrm{e}^{\mathrm{i}(j-1)2\pi/(n-k)}$. Furthermore, let $a = (x_1^{n-k} - 1, \ldots, x_n^{n-k} - 1)^T$ and $\hat{y} \in \mathbb{C}^n$, $\hat{y} = (y_1, \ldots, y_k, \xi_1, \ldots, \xi_{n-k})^T$. Let*

$$T = \begin{pmatrix} I & O \\ (n-k)^{-1/2} C_{\xi, y} & \Phi^{-1} \end{pmatrix}, \tag{16}$$

partitioned as N_0, i.e., the Cauchy matrix $C_{\xi,y}$ has order $(n-k)\times k$ and $\Phi = (n-k)^{-1/2}V_\xi$ is the Fourier matrix of order $n-k$. Furthermore, let f be the solution of the linear system $T^T f = e_n$. If $\Delta(x,\hat{y}) < +\infty$ then

$$K_{x,y} = D_a C_{x,\hat{y}} D_f T.$$

Proof. It is a simple task to verify that

$$T^{-1} = \begin{pmatrix} I & O \\ -(n-k)^{-1/2}\Phi C_{\xi,y} & \Phi \end{pmatrix}. \tag{17}$$

Moreover, since $\Phi^{-1}e_1 = (n-k)^{-1/2}\mathbf{1}$ and $\Phi^{-1}P_0\Phi = D_\xi$, we have

$$\begin{aligned} TN_0T^{-1} &= \begin{pmatrix} I & O \\ (n-k)^{-1/2}C_{\xi,y} & \sqrt{n-k}\Phi^{-1} \end{pmatrix} \begin{pmatrix} D_y & O \\ e_1\mathbf{1}^T & P_0 \end{pmatrix} \\ &\quad \times \begin{pmatrix} I & O \\ -(n-k)^{-1/2}\Phi C_{\xi,y} & \Phi \end{pmatrix} \\ &= \begin{pmatrix} D_y & O \\ (n-k)^{-1/2}[C_{\xi,y}D_y - D_\xi C_{\xi,y} + \mathbf{1}\mathbf{1}^T] & \Phi^{-1}P_0\Phi \end{pmatrix} \\ &= D_{\hat{y}}. \end{aligned}$$

Hence $T^{-1}D_{\hat{y}}T = N_0$. Since diagonal matrices commute, we obtain:

$$\begin{aligned} \mathcal{L}_{D_x,D_{\hat{y}}}(D_a^{-1}K_{x,y}T^{-1}) &= D_a^{-1}[D_xK_{x,y} - K_{x,y}N_0]T^{-1} \\ &= D_a^{-1}ae_n^TT^{-1} \\ &= \mathbf{1}\mathbf{1}^TD_f. \end{aligned}$$

The hypothesis stated on x implies both the invertibility of D_a and that of the operator $\mathcal{L}_{D_x,D_{\hat{y}}}$. We obtain $D_a^{-1}K_{x,y}T^{-1} = C_{x,\hat{y}}D_f$, whence thesis follows. □

We remark that $\det(K_{x,y}) \neq 0$ if all entries from x and y are distinct, see, e.g., [9, Thm. 3.1] or [19]. As a consequence, by Binet-Cauchy Theorem we see that, in the hypotheses of the preceding lemma, we have $\det(D_f) \neq 0$, that is, f has no zero entries.

We consider in the following theorem the structured conditioning of the columns of $K_{x,y}^{-1}$ with respect to perturbations in the vector x.

We will use some further notations: For any $x = (x_1,\dots,x_n)^T \in \mathbb{C}^n$, consider the two subvectors

$$x^{(1)} = (x_1,\dots,x_k)^T \in \mathbb{C}^k, \qquad x^{(2)} = (x_{k+1},\dots,x_n)^T \in \mathbb{C}^{n-k},$$

where k is the integer appearing in the definition of the Cauchy-Vandermonde matrix under consideration. Furthermore, introduce the relative distance

$$\hat{\delta}(x,\tilde{x}) = \max\{\delta(x^{(1)},\tilde{x}^{(1)}), \|x^{(2)} - \tilde{x}^{(2)}\|/\|x\|\},$$

and consider the following condition measure, naturally induced by it:

$$\widehat{m}(F,x) = \lim_{\epsilon\to 0} \sup_{\delta(x,\tilde{x})\le\epsilon} \frac{\hat{\delta}(F(x),F(\tilde{x}))}{\delta(x,\tilde{x})}. \tag{18}$$

Remark that $\|x - \tilde{x}\|/\|x\| \leq \hat{\delta}(x, \tilde{x}) \leq \delta(x, \tilde{x})$, and

$$m(F, x) \leq \widehat{m}(F, x). \tag{19}$$

Theorem 4. *Let $1 \leq i \leq n$ be fixed, and let $F_i : \mathbb{C}^n \mapsto \mathbb{C}^n$ be the function defined as $F_i(x) = K_{x,y}^{-1} e_i$. Moreover, let $0 \leq \theta < 2\pi$ be arbitrary and let $\omega = \mathrm{e}^{-\mathrm{i}\theta}$. Then,*

$$\begin{aligned} \widehat{m}(F_i, x) \quad \leq \quad & \left(\frac{\gamma}{\sqrt{n-k}} + \sqrt{n-k} \right) (\gamma + \sqrt{n-k}) \\ & \times \left(\Delta(n, \omega x, \xi) + \left| \frac{(n-k) x_i^{n-k}}{x_i^{n-k} - \mathrm{e}^{\mathrm{i}(n-k)\theta}} \right| \right), \end{aligned}$$

where $\gamma = \|C_{\xi, \omega y}\|$, $\xi = (\xi_1, \ldots, \xi_{n-k})$, $\xi_j = \mathrm{e}^{\mathrm{i}(j-1)2\pi/(n-k)}$ and $\Delta(n, \omega x, \xi)$ can be derived from (11).

Proof. As in Theorem 2, we can reduce the analysis to the case $\theta = 0$ by simply considering the rotated vectors $\hat{x} = \omega x$ and $\hat{y} = \omega y$, by virtue of (15). In the notations of Lemma 3 we have

$$K_{x,y}^{-1} = T^{-1} D_f^{-1} C_{x,\hat{y}}^{-1} D_a^{-1}.$$

Recall that f has no zero entries, hence D_f is invertible. We obtain

$$F_i(x) = \frac{1}{x_i^{n-k} - 1} T^{-1} D_f^{-1} C_{x,\hat{y}}^{-1} e_i.$$

Then, $F_i(x) = G(H_i(x))$, where $G(x) = T^{-1} x$ and

$$H_i(x) = (x_i^{n-k} - 1)^{-1} D_f^{-1} C_{x,\xi}^{-1} e_i.$$

Using (18), it is straightforward to check that $\widehat{m}(F_i, x) \leq \widehat{m}(G, H_i(x)) c(H_i, x)$.

Let $u = G(x)$ and $\tilde{u} = G(\tilde{x})$, for arbitrary $x, \tilde{x}$. In order to bound the quantity $\hat{\delta}(u, \tilde{u})$, observe firstly that $\delta(u^{(1)}, \tilde{u}^{(1)}) = \delta(x^{(1)}, \tilde{x}^{(1)}) \leq \delta(x, \tilde{x})$. Using $\|u\| \geq \|x\|/\|T\|$, we have:

$$\begin{aligned} \frac{\|u^{(2)} - \tilde{u}^{(2)}\|}{\|u\|} \quad & \leq \quad \|T\| \frac{\|(n-k)^{-1/2} \Phi C_{\xi,y}(x^{(1)} - \tilde{x}^{(1)})\| + \|\Phi(x^{(2)} - \tilde{x}^{(2)})\|}{\|x\|} \\ & \leq \quad \|T\| \left(\gamma \frac{\|x^{(1)} - \tilde{x}^{(1)}\|}{\|x\|} + \sqrt{n-k} \frac{\|x^{(2)} - \tilde{x}^{(2)}\|}{\|x\|} \right) \\ & \leq \quad \|T\| (\gamma + \sqrt{n-k}) \delta(x, \tilde{x}). \end{aligned}$$

Since $\|T\| \leq \gamma/\sqrt{n-k} + \sqrt{n-k}$ we have

$$\hat{\delta}(u, \tilde{u}) \leq (\gamma/\sqrt{n-k} + \sqrt{n-k})(\gamma + \sqrt{n-k}) \delta(x, \tilde{x}).$$

From (18) we obtain

$$\widehat{m}(G, x) \leq (\gamma/\sqrt{n-k} + \sqrt{n-k})(\gamma + \sqrt{n-k}).$$

Furthermore, if we let $\delta(x,\tilde{x})=\epsilon$, we have from Theorem 1

$$\begin{aligned}
|H_i(x)-H_i(\tilde{x})| &\le |x_i^{n-k}-1|^{-1}|D_f^{-1}||C_{x,\xi}^{-1}e_i-C_{\tilde{x},\xi}^{-1}e_i| \\
&\quad+\left|\frac{1}{x_i^{n-k}-1}-\frac{1}{\tilde{x}_i^{n-k}-1}\right||D_f^{-1}C_{\tilde{x},\xi}^{-1}e_i| \\
&\le \epsilon|x_i^{n-k}-1|^{-1}|D_f^{-1}C_{x,\xi}^{-1}e_i|\Delta(n,x,\xi) \\
&\quad+\epsilon\frac{(n-k)|x_i^{n-k}|}{|x_i^{n-k}-1|^2}|D_f^{-1}C_{x,\xi}^{-1}e_i|+o(\epsilon) \\
&\le \epsilon\Delta(n,x,\xi)|H_i(x)|+\epsilon\left|\frac{(n-k)x_i^{n-k}}{x_i^{n-k}-1}\right||H_i(x)|+o(\epsilon),
\end{aligned}$$

whence we obtain

$$c(H_i,x)\le\Delta(n,x,\xi)+\left|\frac{(n-k)x_i^{n-k}}{x_i^{n-k}-1}\right|,$$

and the proof is complete. □

Remark that an upper bound for the constant γ appearing in the preceding theorem is obtained as follows:

$$\gamma=\max_{1\le i\le n-k}\sum_{j=1}^{k}\frac{1}{|\xi_i-\omega y_j|}\le\frac{k}{\min_{i,j}|\xi_i-\omega y_j|}=k\Delta(\xi,\omega y).$$

We omit the proof of the following corollary, as it is essentially the same as that of Corollary 3, in the light of (19):

Corollary 6. *Let $F:\mathbb{C}^n\mapsto\mathbb{C}^{n^2}$ be defined as $F(x)=\mathrm{Vec}(K_{x,y}^{-1}D_d)$, where D_d is an arbitrary diagonal matrix whose entries do not depend on x. Furthermore, let $\xi=(\xi_1,\ldots,\xi_{n-k})$, $\xi_j=\mathrm{e}^{\mathrm{i}(j-1)2\pi/(n-k)}$. For any $0\le\theta<2\pi$, let $\omega=\mathrm{e}^{-\mathrm{i}\theta}$. Then,*

$$\begin{aligned}
m(F,x) &\le (\gamma/\sqrt{n-k}+\sqrt{n-k})(\gamma+\sqrt{n-k}) \\
&\quad\times\left(\Delta(n,\omega x,\xi)+(n-k)\max_i\left|\frac{x_i^{n-k}}{x_i^{n-k}-\mathrm{e}^{\mathrm{i}(n-k)\theta}}\right|\right),
\end{aligned}$$

where $\gamma=\|C_{\xi,\omega y}\|$ and $\Delta(n,\omega x,\xi)$ can be obtained from (11).

References

[1] S.G. Bartels and D.J. Higham; The structured sensitivity of Vandermonde-like systems. *Numer. Math.* 62 (1992), 17–33.

[2] B. Beckermann; The condition number of real Vandermonde, Krylov and positive definite Hankel matrices. *Numer. Math.* 85 (2000), 553–577.

[3] T. Bella, Y. Eidelman, I. Gohberg, I. Koltracht and V. Olshevsky; A Björck-Pereyra-type algorithm for Szegö-Vandermonde matrices based on properties of unitary Hessenberg matrices. *Linear Algebra Appl.* 420 (2007), 634–647.

[4] T. Boros, T. Kailath and V. Olshevsky; A fast parallel Björck-Pereyra-type algorithm for solving Cauchy linear equations. *Linear Algebra Appl.* 302/303 (1999), 265–293.

[5] T. Boros, T. Kailath and V. Olshevsky; Pivoting and backward stability of fast algorithms for solving Cauchy linear equations. *Linear Algebra Appl.* 343/344 (2002), 63–99.

[6] E. Bozzo, D. Fasino and O. Menchi; The componentwise conditioning of the DFT. *Calcolo* 39 (2002), 181–187.

[7] F. Cucker and H. Diao; Mixed and componentwise condition numbers for rectangular structured matrices. *Calcolo* 44 (2007), 89–115.

[8] D. Fasino and V. Olshevsky; How bad are symmetric Pick matrices?, in: *Structured Matrices in Operator Theory, Numerical Analysis, Control, Signal and Image Processing* (V. Olshevsky, Ed.) AMS Series on Contemporary Mathematics, 280 (2001), 301–311.

[9] T. Finck, G. Heinig and K. Rost; An inversion formula and fast algorithms for Cauchy-Vandermonde matrices. *Linear Algebra Appl.* 183 (1993), 179–191.

[10] I. Gohberg and I. Koltracht; On the inversion of Cauchy matrices, in: M.A. Kaashoek, J.H. van Schuppen, A.C.M. Ran (Eds.), *Signal processing, scattering and operator theory, and numerical methods* (Proceedings of MTNS-98), pp. 381–392; Birkhäuser, 1990.

[11] I. Gohberg and I. Koltracht; Mixed, componentwise, and structured condition numbers. *SIAM J. Matrix Anal. Appl.* 14 (1993), 688–704.

[12] G. Heinig and K. Rost; Recursive solution of Cauchy-Vandermonde systems of equations. *Linear Algebra Appl.* 218 (1995), 59–72.

[13] G. Heinig and K. Rost; *Representations of inverses of real Toeplitz-plus-Hankel matrices using trigonometric transformations*, in: Large-Scale Scientific Computation of Engineering and Environmental Problems II (M. Griebel, S. Margenov and P. Yalamov, Eds.), Vieweg, 73 (2000) 80–86.

[14] D.J. Higham and N.J. Higham; Backward error and condition of structured linear systems. *SIAM J. Matrix Anal. Appl.* 13 (1992), 162–175.

[15] N.J. Higham; Error analysis of the Björck-Pereyra algorithms for solving Vandermonde systems. *Numer. Math.* 50 (1987), 613–632.

[16] R.-C. Li; Asymptotically optimal lower bounds for the condition number of a real Vandermonde matrix. *SIAM J. Matrix Anal. Appl.* 28 (2006), 829–844.

[17] V. Olshevsky and V. Pan; *Polynomial and rational evaluation and interpolation* (*with structured matrices*). ICALP99 Proceedings, Springer, LNCS 1644 (1999), 585–594.

[18] G. Mühlbach; Interpolation by Cauchy-Vandermonde systems and applications. *J. Comput. Appl. Math.* 122 (2000), 203–222.

[19] G. Mühlbach; On Hermite interpolation by Cauchy-Vandermonde systems: the Lagrange formula, the adjoint and the inverse of a Cauchy-Vandermonde matrix. *J. Comput. Appl. Math.* 67 (1996), 147–159.

[20] T. Penzl; Eigenvalue decay bounds for solutions of Lyapunov equations: the symmetric case. *Systems Control Lett.* 40 (2000), 139–144.

[21] S. Serra Capizzano; An elementary proof of the exponential conditioning of real Vandermonde matrices. *Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat.* 10 (2007), 761–768.

[22] J.-G. Sun; Bounds for the structured backward errors of Vandermonde systems. *SIAM J. Matrix Anal. Appl.* 20 (1999), 45–59.

[23] E.E. Tyrtyshnikov; Singular values of Cauchy-Toeplitz matrices. *Linear Algebra Appl.* 161 (1992), 99–116.

[24] E.E. Tyrtyshnikov; How bad are Hankel matrices? *Numer. Math.* 67 (1994), 261–269.

[25] H. Xiang and Y. Wei; Structured mixed and componentwise condition numbers of some structured matrices. *J. Comput. Appl. Math.* 202 (2007), 217–229.

Enrico Bozzo and Dario Fasino
Dipartimento di Matematica e Informatica
Università di Udine
Via delle Scienze, 208
I-33100 Udine, Italy
e-mail: dario.fasino@dimi.uniud.it
enrico.bozzo@dimi.uniud.it