How (Un)stable Are Vandermonde Systems?

Walter Gautschi Professor, Department of Computer Sciences, Purdue University, West Lafayette, Indiana

ABSTRACT. Results on the condition number of Vandermonde type matrices obtained during the last 25 years are reviewed. Equal emphasis is given to real and complex nodes. Recent work dealing with nodes placed sequentially on circular and elliptic contours in the complex plane receives special attention.

I. INTRODUCTION

Many problems in applied and numerical analysis eventually boil down to solving large systems of linear algebraic equations. Since the matrices and right-hand sides of such systems are typically the result of (sometimes extensive) computations, they are subject to an unavoidable level of noise caused by the rounding errors committed during their generation. It is then a matter of practical concern trying to estimate the effect of such uncertainties upon the solution of the system.

A common answer – and one which we shall adopt in the sequel – concerning any nonsingular system

$$Ax = b, \quad \det A \neq 0, \tag{1.1}$$

is to compute (or estimate) the *condition number*

$$\operatorname{cond} A = \|A\| \cdot \|A^{-1}\| \tag{1.2}$$

of the system, where $\|\cdot\|$ denotes a suitable matrix norm. Norms, for matrices

Work supported, in part, by the National Science Foundation under grant CCR-8704404.

$A = [a_{ij}]$, that will be used here are the ∞ - *norm*,

$$||A||_\infty = \max_i \sum_j |a_{ij}|, \tag{1.3}$$

the *Euclidean* (or spectral) *norm*,

$$||A||_2 = \sqrt{\rho(AA^H)}, \tag{1.4}$$

where $\rho(\cdot)$ denotes spectral radius, and the *Frobenius norm*,

$$||A||_F = \sqrt{\mathrm{tr}(AA^H)} = \sqrt{\sum_{i,j} |a_{ij}|^2}\ . \tag{1.5}$$

If $\varepsilon_x = ||\delta x||/||x||$ is the relative error in the solution x of (1.1), caused by relative errors $\varepsilon_A = ||\delta A||/||A||$, $\varepsilon_b = ||\delta b||/||b||$ in the system, then the condition number in (1.2) indicates how much larger ε_x is compared to ε_A and ε_b, that is, roughly speaking,

$$\varepsilon_x \approx (\mathrm{cond}\, A)(\varepsilon_A + \varepsilon_b). \tag{1.6}$$

It always seemed important to us that the conditioning of matrices be investigated for many special classes of matrices. In this spirit, we began, 25 years ago, to take up the class of Vandermonde matrices. The original motivation came from unpleasant experiences with the computation of Gauss type quadrature rules from the moments of the underlying weight function. The sensitivity of the problem then indeed depends on the condition of certain (confluent) Vandermonde matrices with real nodes. Since then, we have intermittently looked at the conditioning of such matrices, considering not only real, but also complex nodes, and have enlarged the class of matrices by including Vandermonde-like matrices involving polynomial systems other than the system of powers. Here we present a brief survey of results obtained over the years, including also some original material (in Sections IV, V and VI).

To establish terminology and notation, we call a *Vandermonde matrix* a matrix of the form

$$V_n = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_n \\ \cdot & \cdot & \cdots & \cdot \\ z_1^{n-1} & z_2^{n-1} & \cdots & z_n^{n-1} \end{bmatrix}, \quad z_i \in \mathbb{C}, \quad n > 1, \tag{1.7}$$

where z_i are pairwise distinct real or complex numbers called the *nodes*. More generally, a *Vandermonde-like matrix*, with nodes z_i, is a matrix of the form

$$V_n = \begin{bmatrix} p_0(z_1) & p_0(z_2) & \cdots & p_0(z_n) \\ p_1(z_1) & p_1(z_2) & \cdots & p_1(z_n) \\ \cdot & \cdot & \cdots & \cdot \\ p_{n-1}(z_1) & p_{n-1}(z_2) & \cdots & p_{n-1}(z_n) \end{bmatrix}, \tag{1.8}$$

where $\{p_k\}$ is a system of linearly independent polynomials, often with $p_k \in \mathbf{P}_k$, the class of polynomials of degree k. Such matrices (or their transposed) are encountered, for example, when one deals with polynomial interpolation or with interpolatory approximation of linear functionals (see, e.g., [1]), the form (1.7) or (1.8) occurring depending on the choice of basis elements in polynomial spaces. Vandermonde systems with matrix (1.7) are also an important ingredient in Remes' algorithm for constructing best uniform polynomial approximations.

A brief outline of the paper is as follows. Sections II–IV are devoted to ordinary Vandermonde matrices V_n, (1.7). We begin in Section II with some basic inequalities for $\| V_n^{-1} \|_\infty$. These are then applied in Section III to obtain estimates for the condition number $\operatorname{cond}_\infty V_n$ for certain real node configurations. The bottom line here is that *V_n is ill-conditioned when all nodes are real.* Indeed, the condition number, in many cases (and perhaps always), grows exponentially with the order n of the matrix. The scenario changes drastically if one allows complex nodes. The roots of unity, for example, give rise to perfectly conditioned Vandermonde matrices. Other sequences of nodes on the unit circle are studied in Section IV. Some, like the Van der Corput sequence, perform nearly as well, and do so in a linear sequential, rather than triangular, fashion. Others, like "quasi-cyclic" sequences, do much worse. The remaining two sections discuss Vandermonde-like matrices with the polynomials p_k in (1.8) chosen to be orthogonal polynomials. In Section V we consider special real, as well as arbitrary complex nodes, in Section VI nodes placed sequentially on elliptic contours in the complex plane, and a Chebyshev system of polynomials p_k.

II. A BASIC INEQUALITY FOR INVERSES OF VANDERMONDE MATRICES

The inversion of a linear system with the Vandermonde matrix (1.7) as coefficient matrix can be easily described in terms of the elementary Lagrange interpolation polynomials

$$\ell_\lambda(z) = \prod_{\substack{\mu=1 \\ \mu \neq \lambda}}^{n} \frac{z - z_\mu}{z_\lambda - z_\mu}, \quad \lambda = 1,2,\ldots,n, \tag{2.1}$$

associated with the nodes $z_1, z_2, \ldots, z_n$. Indeed, if we expand ℓ_λ in powers of z,

$$\ell_\lambda(z) = \sum_{\mu=1}^{n} u_{\lambda\mu} z^{\mu-1}, \tag{2.2}$$

the inversion is accomplished by multiplying the μth equation by $u_{\lambda\mu}$, $\mu = 1,2,\ldots,n$, and adding up the results. In view of $\ell_\lambda(z_\nu) = \delta_{\lambda\nu}$ (the Kronecker delta), this will express the λth unknown linearly in terms of the right-hand members of the system. Hence,

$$V_n^{-1} = [u_{\lambda\mu}]_{\substack{1 \le \lambda \le n \\ 1 \le \mu \le n}} . \tag{2.3}$$

Combining (2.1) and (2.2) yields

$$u_{\lambda 1} + u_{\lambda 2} z + \cdots + u_{\lambda n} z^{n-1} = \pi_\lambda \prod_{\substack{\mu=1 \\ \mu \ne \lambda}}^{n} (z - z_\mu), \tag{2.4}$$

where

$$\pi_\lambda = \prod_{\mu \ne \lambda} (z_\lambda - z_\mu)^{-1}. \tag{2.5}$$

Therefore, we have the alternative representation

$$u_{\lambda\mu} = (-1)^{\mu-1} \pi_\lambda \sigma_{n-\mu}^{\lambda}(z_1, \ldots, z_{\lambda-1}, z_{\lambda+1}, \ldots, z_n), \tag{2.6}$$

where σ_m^λ denotes the mth elementary symmetric function in the $n-1$ variables z_μ with z_λ removed.

Theorem 2.1. *For arbitrary $z_\nu \in \mathbb{C}$, with $z_\nu \ne z_\mu$ if $\nu \ne \mu$, there holds*

$$\max_\lambda \prod_{\mu \ne \lambda} \frac{\max(1, |z_\mu|)}{|z_\lambda - z_\mu|} < \|V_n^{-1}\|_\infty \le \max_\lambda \prod_{\mu \ne \lambda} \frac{1 + |z_\mu|}{|z_\lambda - z_\mu|}, \tag{2.7}$$

where V_n is the matrix in (1.7). The upper bound is attained if $z_\mu = |z_\mu| e^{i\theta}$, $\mu = 1,2,\ldots,n$, for some fixed $\theta \in \mathbb{R}$.

Proof (Sketch). The upper bound in (2.7), and the statement about equality, follow from (2.6) and from a simple fact (see the Lemma in [7, p.118]) about elementary symmetric functions $\sigma_m = \sigma_m(x_1, \ldots, x_n)$ in n variables, namely that $\sum_{m=0}^{n} |\sigma_m| \le \prod_{\nu=1}^{n} (1 + |x_\nu|)$, with equality precisely if all x_ν lie on the same ray through the origin.

For the lower bound we use the fact that

$$\sum_{\mu=0}^{n} |a_\mu| \ge |a_n| \prod_{\nu=1}^{n} \max(1, |\zeta_\nu|) \tag{2.8}$$

holds for any polynomial $p(z) = a_0 + a_1 z + \cdots + a_n z^n$, $a_n \ne 0$, having the zeros ζ_1,

$\zeta_2, \ldots, \zeta_n$, with equality if and only if $p(z) = a_n z^n$. This is a simple consequence of Jensen's formula (see [10, §2]). Applying (2.8) to the polynomial (of degree $n-1$) in (2.4) then easily yields the lower bound in (2.7); see again [10] for details. □

III. REAL NODES

An important case in which equality holds for the upper bound in (2.7) is when all nodes are nonnegative, $z_\nu = x_\nu \geq 0$. Then, letting

$$p_n(z) = \prod_{\nu=1}^{n} (z - x_\nu) \tag{3.1}$$

denote the node polynomial, we can write (2.7) in the form

$$\|V_n^{-1}\|_\infty = \frac{|p_n(-1)|}{\min\limits_{\nu}\{(1 + x_\nu)|p_n'(x_\nu)|\}} \quad (x_\nu \geq 0). \tag{3.2}$$

The techniques used in the first part of the proof of Theorem 2.1 can be adapted (cf. [8, Theorem 4.3]) to also deal with the case of real nodes located symmetrically with respect to the origin: $x_\nu \in \mathbf{R}$ with $x_\nu + x_{n+1-\nu} = 0$ for $\nu = 1,2, \ldots, n$. In place of (3.2), one obtains

$$\|V_n^{-1}\|_\infty = \frac{|p_n(i)|}{\min\limits_{x_\nu \geq 0}\left\{\dfrac{1 + x_\nu^2}{1 + x_\nu}\,|p_n'(x_\nu)|\right\}} \quad (x_\nu + x_{n+1-\nu} = 0,\ \ x_\nu \in \mathbf{R}). \tag{3.3}$$

Since for given nodes x_ν the norm $\|V_n\|_\infty$ is easily calculated, the results (3.2), (3.3) allow us to evaluate the condition number $\mathrm{cond}_\infty V_n$ exactly in the respective cases. For example, if $|x_\nu| \leq 1$ for all ν, then

$$\mathrm{cond}_\infty V_n = n\|V_n^{-1}\|_\infty \quad (|x_\nu| \leq 1). \tag{3.4}$$

We illustrate (3.2)–(3.4) with a number of examples, ordered in decreasing severity of ill-conditioning.

Example 3.1. Harmonic nodes $x_\nu = 1/\nu$, $\nu = 1,2, \ldots, n$.

Here, an easy calculation gives $|p_n(-1)| = n+1$, and letting $\delta_\nu = (1 + x_\nu)|p_n'(x_\nu)|$, one finds

$$\delta_\nu = \frac{(\nu+1)!(n-\nu)!}{\nu^n n!}, \qquad \nu = 1,2,\ldots,n.$$

There follows

$$\min_\nu \delta_\nu \le \delta_n = \frac{n+1}{n^n}.$$

(Actually, this holds with strict inequality, for $n > 2$, as can be shown by a more detailed analysis). Consequently, by (3.4) and (3.2),

$$\operatorname{cond}_\infty V_n > n^{n+1} \qquad (x_\nu = 1/\nu). \tag{3.5}$$

Note that the condition number in (3.5) grows more rapidly than $n!$, which is far worse than the condition of the notorious Hilbert matrix, which grows ''only'' exponentially!

Example 3.2. Equidistant nodes on $[0,1]$: $x_\nu = \dfrac{\nu-1}{n-1}$, $\nu = 1,2,\ldots,n$.

Defining δ_ν as in the previous example, an elementary computation gives

$$|p_n(-1)| = \frac{(2n-2)!}{(n-1)!(n-1)^{n-1}}, \qquad \delta_\nu = \frac{(n+\nu-2)(\nu-1)!(n-\nu)!}{(n-1)^n}.$$

Putting $\nu = \kappa n$, $0 < \kappa < 1$, and studying $\delta_{\kappa n}$ for $n \to \infty$, reveals that, asymptotically, $\delta_{\kappa n}$ is a minimum when $\kappa = \frac{1}{2}$, and $\delta_{\frac{1}{2}n} \sim \pi en\,(2e)^{-n}$ as $n \to \infty$. Combining this in Eq. (3.2) with the asymptotic expression for $|p_n(-1)|$, obtained by Stirling's formula, and noting that $\|V_n\|_\infty = n$, yields

$$\operatorname{cond}_\infty V_n \sim \frac{\sqrt{2}}{4\pi}\cdot 8^n, \qquad n \to \infty. \tag{3.6}$$

We are now down to exponential growth, but expect that the rate of growth can still be reduced by placing the nodes symmetrically with respect to the origin. This is confirmed in the next example.

Example 3.3. Equidistant nodes on $[-1,1]$, $x_\nu = 1 - \dfrac{2(\nu-1)}{n-1}$, $\nu = 1,2,\ldots,n$.

Here we use (3.3). An asymptotic analysis similar to the one in the previous example, but more involved, shows that [8, Example 6.1]

$$\operatorname{cond}_\infty V_n \sim \frac{1}{\pi}\, e^{-\frac{1}{4}\pi}\, e^{n(\frac{1}{4}\pi + \frac{1}{2}\ln 2)}, \quad n \to \infty. \tag{3.7}$$

Note that the exponential growth rate is now $\exp\left[\frac{1}{4}\pi + \frac{1}{2}\ln 2\right] = 3.1017\ldots,$

compared to 8 in the asymmetric case of Example 3.2.

Can we do better if we take the Chebyshev nodes in (–1,1)?

Example 3.4. Chebyshev nodes $x_\nu = \cos((2\nu - 1)\pi/2n)$, $\nu = 1,2,\ldots, n$. Applying (3.3), one can prove [8, Example 6.2] that

$$\mathrm{cond}_\infty V_n \sim \frac{3^{3/4}}{4}(1+\sqrt{2})^n, \quad n \to \infty. \tag{3.8}$$

Here the growth rate $1+\sqrt{2} = 2.4142\ldots$ is indeed smaller than for equally spaced symmetric nodes, but not by a whole lot.

Seeing the condition of V_n continually improving through the series of examples above, one cannot help wondering whether there is an optimal set of real nodes $x^T = [x_1, x_2, \ldots, x_n]$ (say, with $x_1 > x_2 > \cdots > x_n$), and if so, what they are and what optimal growth rate of $\mathrm{cond}_\infty V_n$ they produce as $n \to \infty$. As far as the existence of the optimum is concerned, the answer is easily seen to be affirmative (cf. [9]). We even conjecture (but have no proof as yet) that the optimal nodes are unique, subject to the above ordering. If this were true, it would follow [9, Theorem 3.1] that the optimal node configuration is symmetric with respect to the origin. Seen in this light, the recent result [14, Theorem 3.1]

$$\mathrm{cond}_\infty V_n > 2^{n/2} \quad (n > 2,\ x_\nu \text{ symmetric}) \tag{3.9}$$

is of interest, since it shows that, accepting the above conjecture, the condition of Vandermonde matrices grows exponentially for *any* set of real nodes. Nevertheless, the growth rate indicated in (3.9) is not believed to be sharp, and the search for the optimal growth rate remains an interesting open problem. There is a result analogous to (3.9) for arbitrary positive nodes, namely [14, Theorem 2.1]

$$\mathrm{cond}_\infty V_n > 2^{n-1} \quad (n > 1,\ x_\nu \geq 0). \tag{3.10}$$

Both bounds in (3.9) and (3.10) can be slightly sharpened (cf. Theorems 3.2 and 2.2, respectively, in [14]).

IV. COMPLEX NODES

The fact that real nodes lead to ill-conditioned Vandermonde matrices is not surprising if one considers that powers constitute, as is well known, a poor basis for polynomial approximation on the real line; see, e.g., the discussion of near linear dependence in [3, pp. 119–120], or of the conditioning of the power basis in [11]. In contrast, replacing the powers by Chebyshev polynomials, and considering the corresponding

Vandermonde-like matrices (1.8), can lead to perfectly conditioned matrices if one chooses the (real) nodes appropriately; cf. Section V below.

Now it so happens that the powers are indeed "Chebyshev polynomials" on any disc in the complex plane centered at the origin, in the sense of deviating least from zero (in the uniform norm on the disc, or on the circumference of the disc) among all monic polynomials of the same degree. Therefore, one expects better conditioning of ordinary Vandermonde matrices if one allows the nodes z_ν to be complex.

If we measure the condition in the Euclidean norm (1.4), and consider for simplicity the *unit* disc, then the nth roots of unity indeed minimize $\mathrm{cond}_2 V_n$; in fact,

$$\mathrm{cond}_2 V_n = 1 \ \text{ if } \ z_\nu = z_\nu^{(n)} = e^{i(\nu-1)2\pi/n}, \quad \nu = 1,2,\ldots, n. \tag{4.1}$$

This is an easy consequence of the orthogonality of trigonometric functions. The roots of unity, therefore, would seem to be an ideal choice for work on the unit disc, if it weren't for the fact that they form a *triangular array* of nodes, i.e., for each n, as indicated in (4.1), there are n distinct nodes $z_\nu^{(n)}$ which change as n is increased by 1. In applications to interpolation and quadrature, this would require that the function to be interpolated, or integrated, be obtained on a two-dimensional set of points. It is an interesting question to ask how well one can do with a *linear array* of nodes on the unit circle.

One naive answer to this is to first note that the set of kth roots of unity, for k even, contains as subset the $(k/2)$th roots of unity. We may therefore generate a linear sequence of nodes by adjoining to the 2^{k-1}th roots of unity every other 2^kth root of unity, going around the circle in the positive direction, and doing this for $k = 1,2,3, \ldots$. More precisely, if

$$2^{k-1} < \nu \le 2^k \tag{4.2}$$

for some $k \ge 1$, then

$$z_1 = 1, \quad z_\nu = e^{2\pi i (2(\nu - 2^{k-1}) - 1)/2^k}. \tag{4.3}$$

We call this sequence, for lack of a better word, the *quasi-cyclic sequence* on the circle. (Such sequences have been used by Eiermann, Niethammer and Varga [4, p. 522] in the context of semiiterative methods for systems of linear algebraic equations.) With this choice of nodes, whenever n is a power of 2, the corresponding Vandermonde matrix V_n is perfectly conditioned, but otherwise, there is a chance, especially for large n, that the condition may deteriorate significantly. The bounds in Theorem 2.1, unfortunately, are too far apart to give much useful information. We therefore computed the condition number for V_n numerically, using, as seems natural on the circle, the Euclidean matrix norm (1.4). The results for $\mathrm{cond}_2 V_n$ are depicted on a logarithmic scale in Figure 4.1 for $3 \le n \le 64$. As expected, the condition number shoots up to considerable heights between two successive (large) powers of 2.

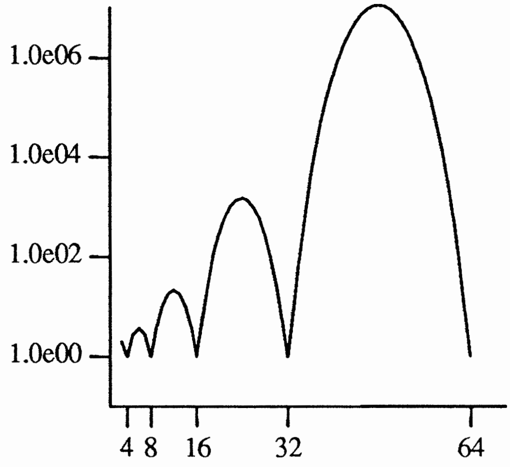

FIG. 4.1 *The condition of Vandermonde matrices* (1.7) *for* n=3(1)64 *with nodes taken from the quasi-cyclic sequence* (4.3).

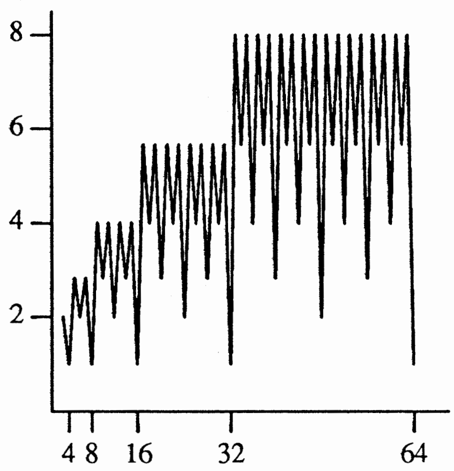

FIG. 4.2 *The condition of Vandermonde matrices* (1.7) *for* n=3(1)64 *with nodes taken from the Van der Corput sequence* (4.6).

How, then, can we avoid these large peaks? More specifically, for an integer n satisfying $2^{k-1} < n < 2^k$, in which order should we adjoin the set of 2^{k-1}th roots of unity by alternate 2^kth roots of unity such that $\max\limits_{2^{k-1} < n < 2^k} \mathrm{cond}_2 V_n$ is minimized? We don't know the solution to this problem, but a good candidate for an optimal (or nearly optimal) node sequence is obtained as follows.

For any integer $\nu \geq 0$, written in binary form

$$\nu = \sum_{j=0}^{\infty} c_j 2^j, \quad \nu_j \in \{0,1\}, \tag{4.4}$$

define the fraction $c_\nu \in [0,1)$ by

$$c_\nu = \sum_{j=0}^{\infty} \nu_j 2^{-j-1}. \tag{4.5}$$

The sequence $\{c_\nu\}_{\nu=0}^{\infty}$ is known as the *Van der Corput sequence.* We then take as nodes

$$z_\nu = e^{2\pi i c_{\nu-1}}, \quad \nu = 1,2,3,\ldots . \tag{4.6}$$

(Such nodes were used to good advantage by Fischer and Reichel [5, p. 228] in connection with the Richardson iteration method; see also Fischer and Reichel [6], Reichel and Opfer [17].) It is easily seen that for $n = 2^{k-1}$ the set $\{z_\nu : \nu = 1,2,\ldots,n\}$ consists of all the nth roots of unity, just like in the quasi-cyclic case. For values of n between 2^{k-1} and 2^k, however, the nodes (4.6) are picked in a zigzag manner from the 2^kth roots of unity, rather than cyclically around the circle as in (4.3). This achieves a more evenly distributed set of nodes, and one can hope that the condition number $\mathrm{cond}_2 V_n$ remains correspondingly smaller. This is indeed confirmed by a computation for $3 \leq n \leq 64$, the results of which are summarized in Figure 4.2. (A similar picture, extended through $n = 148$, has previously been published by Reichel and Opfer [17].) Further computations [15] reveal a rather astonishing pattern for the eigenvalues and eigenspaces of the matrix $V_n V_n^H$ – a Hermitian Toeplitz matrix – which served as an inspiration for the work in [2]. There, it is proved, in particular, that all eigenvalues of $V_n V_n^H$ are powers of 2, the largest, $\lambda_{\max}$, always being equal to $\lambda_{\max} = 2^k$, and the smallest, $\lambda_{\min} = 2^\ell$, where $\ell = 0$ if n is odd, and $0 < \ell \leq k$ otherwise. There follows

$$\mathrm{cond}_2 V_n \leq 2^{k/2} < \sqrt{2n} \quad (2^{k-1} < n \leq 2^k, \ z_\nu \text{ as in (4.6)}), \tag{4.7}$$

with equality on the left holding for every odd n. The various "levels" exhibited in Figure 4.2 thus have heights $2^{k/2}$, $k = 2,3,4,\ldots$. Comparison of Figures 4.2 and 4.1 clearly illustrates the significant improvement achieved by the Van der Corput sequence over the quasi-cyclic sequence. Similar phenomena on ellipses (and also on intervals) will be discussed in Section VI.

V. VANDERMONDE-LIKE MATRICES INVOLVING ORTHOGONAL POLYNOMIALS

As observed at the beginning of Section IV, the choice of orthogonal polynomials as bases in problems of approximation on the real line leads to Vandermonde-like matrices (1.8) which can be expected to have better condition than ordinary Vandermonde matrices. It is the purpose of this section to study the condition of such matrices in the case where

$$p_k(z) = p_k(z\,; d\sigma), \quad k = 0,1,2,\ldots, \tag{5.1}$$

are *orthonormal polynomials* with respect to some (positive) measure $d\sigma$ on the real line,

$$\int_{\mathbb{R}} p_r(x)p_s(x)d\sigma(x) = \delta_{rs} = \begin{cases} 1, & r = s, \\ 0, & r \neq s. \end{cases} \tag{5.2}$$

In most applications, the nodes z_ν are real and contained in the support of $d\sigma$,

$$z_\nu = x_\nu \in \mathbf{R}, \quad x_\nu \in \operatorname{supp} d\sigma. \tag{5.3}$$

A choice that appears particularly natural is that of the zeros of $p_n(\cdot\,; d\sigma)$,

$$x_\nu = x_\nu^{(n)}, \quad p_n(x_\nu\,; d\sigma) = 0, \quad \nu = 1,2,\ldots, n. \tag{5.4}$$

In this case the condition $\operatorname{cond}_F V_n$ of V_n (in the Frobenius norm (1.5)) can be expressed very simply in terms of *Christoffel numbers* $\gamma_\nu = \gamma_\nu^{(n)}(d\sigma)$ belonging to the measure $d\sigma$, i.e., in terms of the weights in the Gauss-Christoffel quadrature formula

$$\int_{\mathbb{R}} f(x)d\sigma(x) = \sum_{\nu=1}^{n} \gamma_\nu^{(n)} f(x_\nu^{(n)}) + R_n(f), \quad R_n(\mathbf{P}_{2n-1}) = 0. \tag{5.5}$$

(As is well known, $\gamma_\nu^{(n)} > 0$ for $\nu = 1,2,\ldots, n$.) Indeed, we have

Theorem 5.1. *The condition of V_n in* (1.8), *where p_k are the orthonormal polynomials* (5.1), (5.2) *and the nodes x_ν given by* (5.4), *equals*

$$\operatorname{cond}_F V_n = \left[\sum_{\nu=1}^{n} \gamma_\nu \sum_{\nu=1}^{n} \frac{1}{\gamma_\nu}\right]^{1/2}, \tag{5.6}$$

where $\gamma_\nu = \gamma_\nu^{(n)}(d\sigma)$ are the Christoffel numbers of $d\sigma$, and the norm used in (5.6) *is* *the Frobenius norm* $\|\cdot\|_F$ *in* (1.5).

The proof rests on the fact that $1/\gamma_\nu^{(n)}$ are the squares of the singular values of V_n, which in turn is a consequence of the discrete orthogonality property of orthogonal polynomials (cf. [12]). Note also that $\mathrm{cond}_F V_n \geq n$ for any set of positive numbers γ_ν.

Our first example is the analogue of the example involving the roots of unity, in the sense that it achieves optimality.

Example 5.1. The Chebyshev measure $d\sigma(x) = (1-x^2)^{-1/2}dx$ on $[-1,1]$.

Here, the Christoffel numbers $\gamma_\nu^{(n)}$ are all equal to π/n. Indeed, the Chebyshev measure is the only measure for which this is true for all n. (There are other measures, however, for which equality of Christoffel numbers holds for selected values of n; see, e.g., [13, §6], [16]). It then follows from (5.6) that $\mathrm{cond}_F V_n = n$, i.e., V_n is optimally conditioned.

Nevertheless, optimality is achieved at a price: a *triangular* array of nodes (just like earlier with roots of unity). We will show in Section VI how one can find a *linear* array of nodes that also produces well-conditioned matrices V_n, (1.8).

Not much worse than the Chebyshev measure are those of Legendre and Chebyshev of the second kind.

Example 5.2. $d\sigma(x) = dx$ and $d\sigma(x) = (1-x^2)^{1/2}dx$ on $[-1,1]$.

Here one computes from (5.6) the following condition numbers for selected values of n:

TABLE 1 *The condition of Vandermonde-like matrices for Legendre and 2nd-kind Chebyshev polynomials*
(Numbers in parentheses indicate decimal exponents.)

n	Legendre	Chebyshev 2nd kind
5	5.362(0)	5.916(0)
10	1.155(1)	1.483(1)
20	2.494(1)	3.924(1)
40	5.367(1)	1.071(2)
80	1.148(2)	2.976(2)

In stark contrast, Laguerre and Hermite polynomials give rise to extremely ill-conditioned matrices V_n, for example, $\mathrm{cond}_F V_{40} = 1.924(30)$ and $3.699(14)$ in the two respective cases. This is due to the presence of very small Christoffel numbers.

If z_ν are arbitrary complex nodes, one can prove a result similar to, but weaker than, (5.6); it involves the Christoffel function, rather than Christoffel numbers. We recall that the *Christoffel function* (for some measure $d\sigma$) is defined by

$$\gamma_n(z_0; d\sigma) = \min_{\substack{p \in \mathbb{P}_{n-1} \\ p(z_0)=1}} \int_{\mathbb{R}} |p(x)|^2 d\sigma(x), \quad z_0 \in \mathbb{C}, \tag{5.7}$$

where the minimum is over all complex polynomials of degree $\le n-1$ taking on the value 1 at z_0. Alternatively,

$$[\gamma_n(z\,;d\sigma)]^{-1} = \sum_{k=0}^{n-1} |p_k(z\,;d\sigma)|^2. \tag{5.7$'$}$$

In place of (5.6) we then have [12]

$$\mathrm{cond}_F V_n \ge \left\{ \sum_{\nu=1}^{n} \gamma_n(z_\nu\,;d\sigma) \sum_{\nu=1}^{n} \frac{1}{\gamma_n(z_\nu\,;d\sigma)} \right\}^{1/2}. \tag{5.8}$$

To prove (5.8), and at the same time give a version of (5.8) involving equality, one must first of all invert the matrix V_n in (1.8). This can be done similarly as in Section II for powers by expanding the fundamental Lagrange polynomial (2.1) not in powers, as in (2.2), but in the orthogonal polynomials $p_k(z) = p_k(z\,;d\sigma)$,

$$\ell_\nu(z) = \sum_{\mu=1}^{n} a_{\nu\mu} p_{\mu-1}(z), \qquad \nu = 1,2,\ldots,n. \tag{5.9}$$

Then as before, one finds

$$V_n^{-1} = A, \qquad A = [a_{\nu\mu}]. \tag{5.10}$$

Now

$$\begin{aligned}
\int_{\mathbb{R}} \sum_{\nu=1}^{n} |\ell_\nu(x)|^2 d\sigma(x) &= \int_{\mathbb{R}} \sum_{\nu} \sum_{\mu} a_{\nu\mu} p_{\mu-1}(x) \sum_{\lambda} \bar{a}_{\nu\lambda} p_{\lambda-1}(x) d\sigma(x) \\
&= \sum_{\nu} \sum_{\mu,\lambda} a_{\nu\mu} \bar{a}_{\nu\lambda} \int_{\mathbb{R}} p_{\mu-1}(x) p_{\lambda-1}(x) d\sigma(x) \\
&= \sum_{\nu,\mu} |a_{\nu\mu}|^2
\end{aligned}$$

on account of the orthonormality of the p_k. Consequently,

$$\| V_n^{-1} \|_F = \left[\int_{\mathbb{R}} \sum_{\nu=1}^{n} |\ell_\nu(x)|^2 d\sigma(x) \right]^{1/2}. \tag{5.11}$$

On the other hand,

$$\| V_n \|_F = \left\{ \sum_{\nu=1}^{n} \sum_{\mu=1}^{n} |p_{\mu-1}(z_\nu)|^2 \right\}^{1/2}, \tag{5.12}$$

which, on account of (5.7′), gives

$$||V_n||_F = \left[\sum_{\nu=1}^{n} \frac{1}{\gamma_n(z_\nu\,;d\sigma)}\right]^{1/2}. \tag{5.13}$$

The assertion (5.8) now follows by multiplying the two expressions in (5.11) and (5.13) and observing that

$$\int_{\mathbb{R}} |\ell_\nu(x)|^2 d\sigma(x) \geq \gamma_n(z_\nu\,;d\sigma)$$

by (5.7), since $\ell_\nu \in \mathbf{P}_{n-1}$ and $\ell_\nu(z_\nu) = 1$.

We note, however, that the product of (5.11) and (5.12) is exactly *equal* to $\mathrm{cond}_F V_n$, and can easily be computed, at least for conventional measures $d\sigma$, the first factor by Gaussian quadrature and the other by recurrence.

Analogous results can be derived, in essentially the same way, for orthogonal polynomials $\{p_k(\cdot\,;d\sigma)\}$ that are *not* normalized (for example, for *monic* polynomials). Letting

$$d_k^2 = \int_{\mathbb{R}} p_k^2(x;d\sigma)d\sigma(x), \quad k = 0,1,2,\ldots, \tag{5.14}$$

and denoting $D = \mathrm{diag}(d_0,d_1,\ldots,d_{n-1})$, the appropriate matrix norm to be used is then

$$||A||_{F,D} = ||D^{-1}AD||_F \tag{5.15}$$

(which clearly satisfies all the axioms of a matrix norm, including submultiplicativity, $||AB||_{F,D} \leq ||A||_{F,D}||B||_{F,D}$). In place of (5.11), one obtains

$$||V_n^{-1}||_{F,D} = \left\{\int_{\mathbb{R}} \sum_{\nu=1}^{n} \frac{1}{d_{\nu-1}^2} |\ell_\nu(x)|^2 d\sigma(x)\right\}^{1/2}, \tag{5.16}$$

and (5.12) must be modified to read

$$||V_n||_{F,D} = \left\{\sum_{\nu=1}^{n} d_{\nu-1}^2 \sum_{\mu=1}^{n} \frac{1}{d_{\mu-1}^2} |p_{\mu-1}(z_\nu)|^2\right\}^{1/2}. \tag{5.17}$$

The condition $\mathrm{cond}_{F,D} V_n$ is then again computable as the product of (5.16) and (5.17), and can be estimated from below by

$$\operatorname{cond}_{F,D} V_n \geq \left[\sum_{\nu=1}^{n} \frac{\gamma_n(z_\nu; d\sigma)}{d_{\nu-1}^2} \sum_{\nu=1}^{n} \frac{d_{\nu-1}^2}{\gamma_n(z_\nu; d\sigma)} \right]^{1/2}. \tag{5.18}$$

For orthonormal polynomials, we have $D = I$, and the results (5.16)–(5.18) reduce to (5.11), (5.12) and (5.8).

VI. VANDERMONDE-LIKE MATRICES INVOLVING CHEBYSHEV POLYNOMIALS ON ELLIPSES

We have noted in Section IV that the powers are "Chebyshev polynomials" (i.e., monic polynomials of minimum uniform norm) on the disc. It is similarly known that the (monic) polynomials

$$p_0(z) = 1, \quad p_k(z) = 2\rho^{k/2} T_k\left(\frac{z}{2\sqrt{\rho}}\right), \quad k = 1,2, \ldots \quad (0 < \rho \leq 1), \tag{6.1}$$

where T_k denotes the Chebyshev polynomial of the first kind, are the "Chebyshev polynomials" on the ellipse $\mathcal{E}_\rho$ with boundary given by

$$\partial\mathcal{E}_\rho = \{z : z = e^{i\theta} + \rho e^{-i\theta},\ 0 \leq \theta \leq 2\pi\} \tag{6.2}$$

if $0 < \rho < 1$, and on the interval $[-2,2]$ (the limit of (6.2) as $\rho \to 1$), when $\rho = 1$; cf. [17]. (The ellipse $\mathcal{E}_\rho$ is scaled so as to have capacity 1.) This suggests the study of Vandermonde-like matrices (1.8) with polynomials p_k given by (6.1) and nodes located on the elliptic contour (6.2), either in quasi-cyclic order, or in the order determined by the Van der Corput sequence. Thus, in the former case, with ν given as in (4.2), and assuming $0 < \rho < 1$,

$$z_1 = 1 + \rho, \quad z_\nu = e^{i\theta_\nu} + \rho e^{-i\theta_\nu}, \quad \theta_\nu = 2\pi(2(\nu - 2^{k-1}) - 1)/2^k, \tag{6.3}$$

and in the latter case,

$$z_\nu = e^{2\pi i c_{\nu-1}} + \rho e^{-2\pi i c_{\nu-1}}, \tag{6.4}$$

where $\{c_\nu\}_{\nu=0}^{\infty}$ is the Van der Corput sequence (4.4), (4.5). In the limit case $\rho=1$, these formulae have to be slightly modified, since we do not want to run back and forth through the interval $[-2,2]$. We then assume, in the quasi-cyclic case,

$$1 + 2^{k-1} < \nu \leq 2^k + 1 \quad (k \geq 1), \tag{6.5}$$

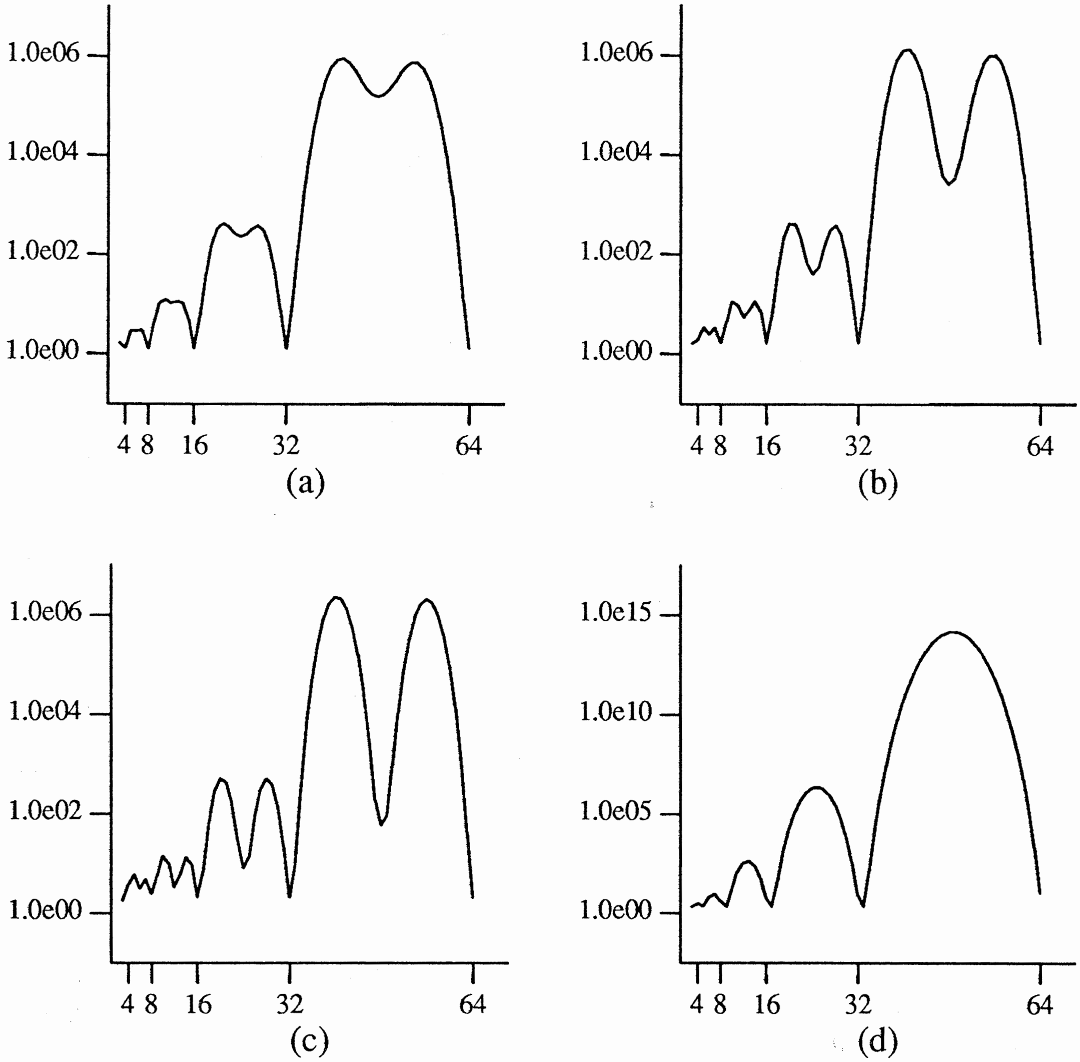

FIG. 6.1 *The condition of Vandermonde-like matrices* (1.8) *for* n=3(1)64 *involving "Chebyshev polynomials" on the ellipse* $\mathcal{E}_\rho$ *and nodes taken from the quasi-cyclic sequence* (6.3) *resp.* (6.3$_1$).
(a) $\rho = .25$ (b) $\rho = .5$ (c) $\rho = .75$ (d) $\rho = 1$.

and define

$$z_1 = -2, \quad z_2 = 2, \quad z_\nu = 2\cos \pi(2(\nu - 2^{k-1}) - 3)/2^k, \quad \nu = 3,4,\ldots\ ; \qquad (6.3_1)$$

in the case of the Van der Corput sequence, we let

$$z_1 = -2, \qquad z_{\nu+1} = 2\cos(\pi c_{\nu-1}), \qquad \nu = 1,2,3,\ldots\ . \qquad (6.4_1)$$

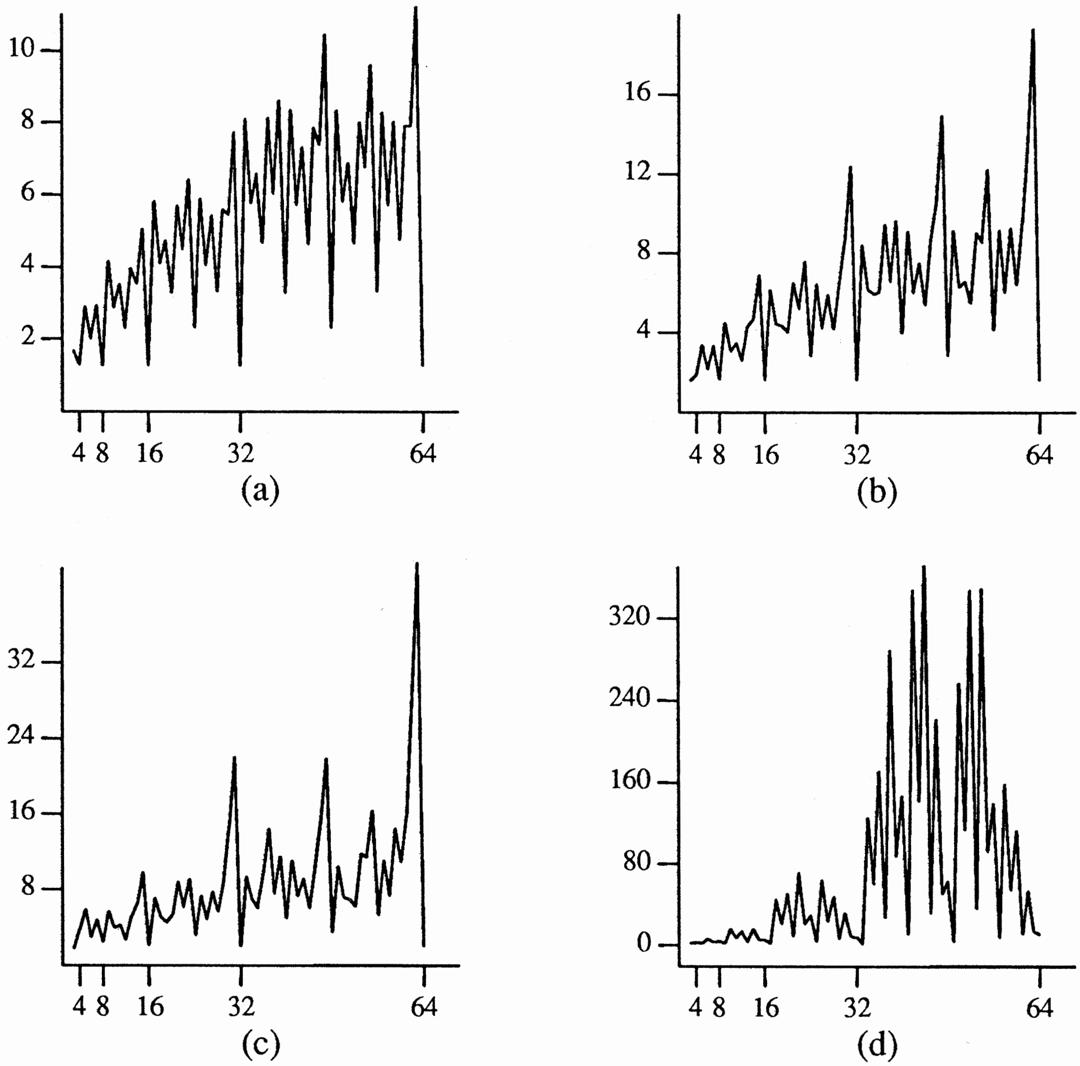

FIG. 6.2 *The condition of Vandermonde-like matrices* (1.8) *for* n=3(1)64 *involving "Chebyshev polynomials" on the ellipse* $\mathcal{E}_\rho$ *and nodes taken from the Van der Corput sequence* (6.4) *resp.* (6.4_1).
(a) $\rho = .25$ (b) $\rho = .5$ (c) $\rho = .75$ (d) $\rho = 1$.

We computed $\text{cond}_2 V_n$ for $3 \le n \le 64$ in the case of ellipses $\mathcal{E}_\rho$ with ρ = .25(.25).75 (for the case ρ=0, see Section IV), and for the segment [–2,2] (i.e., the case ρ=1). The results are shown graphically, on a logarithmic scale, in Figure 6.1 for nodes given by (6.3) [resp. (6.3_1)], and in Figure 6.2, on a linear scale, for the nodes (6.4) [resp. (6.4_1)]. As can be seen, Van der Corput sequences again perform significantly better than quasi-cyclic sequences. The condition, in fact, is known to grow at most polynomially in n, if $0 \le \rho < 1$, and at most like $n^{O(\log n)}$ if $\rho = 1$; cf. [17, Section 3]. In the case of quasi-cyclic sequences, when $0 < \rho < 1$, it is interesting to observe two large peaks between successive powers of 2, in contrast to the cases ρ=0 and ρ=1, which exhibit only one (a surprisingly large one when ρ=1).

ACKNOWLEDGMENT. The author is indebted to Professor Lothar Reichel for useful comments.

REFERENCES

1. Björck, A. and Elfving, T.: Algorithms for confluent Vandermonde systems, *Numer. Math.*, v. 21, 1973, pp. 130–137.

2. Córdova, A., Gautschi, W. and Ruscheweyh, S.: Vandermonde matrices on the circle: spectral properties and conditioning, submitted for publication.

3. Dahlquist, G. and Björck, Å.: *Numerical Methods*, Prentice-Hall, Englewood Cliffs, N.J., 1974.

4. Eiermann, M., Niethammer, W. and Varga, R.S.: A study of semiiterative methods for nonsymmetric systems of linear equations, *Numer. Math.*, v. 47, 1985, pp. 505–533.

5. Fischer, B. and Reichel, L.: A stable Richardson iteration method for complex linear systems, *Numer. Math.*, v. 54, 1988, pp. 225–242.

6. ________ and ________: Newton interpolation in Fejér and Chebyshev points, *Math. Comp.*, v. 53, 1989, to appear.

7. Gautschi, W.: On inverses of Vandermonde and confluent Vandermonde matrices, *Numer. Math.*, v. 4, 1962, pp. 117–123.

8. ________: Norm estimates for inverses of Vandermonde matrices, *Numer. Math.*, v. 23, 1975, pp. 337–347.

9. ________: Optimally conditioned Vandermonde matrices, *Numer. Math.*, v. 24, 1975, pp. 1–12.

10. ________: On inverses of Vandermonde and confluent Vandermonde matrices III, *Numer. Math.*, v. 29, 1978, pp. 445–450.

11. ________: The condition of polynomials in power form, *Math. Comp.*, v. 33, 1979, pp. 343–352.

12. ________: The condition of Vandermonde-like matrices involving orthogonal polynomials, *Linear Algebra Appl.*, v. 52/53, 1983, pp. 293–300.

13. ________: On some orthogonal polynomials of interest in theoretical chemistry, *BIT*, v. 24, 1984, pp. 473–483.

14. ________ and Inglese, G.: Lower bounds for the condition number of Vandermonde matrices, *Numer. Math.*, v. 52, 1988, pp. 241–250.

15. Opfer, G.: personal communication.

16. Peherstorfer, F.: On Gauss quadrature formulas with equal weights, *Numer. Math.*, v. 52, 1988, pp. 317–327.

17. Reichel, L. and Opfer G.: Chebyshev-Vandermonde systems, *Math. Comp.*, to appear.