JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **131**, 341–353 (1988)

Inverses of Generalized Vandermonde Matrices

LUIS VERDE-STAR

Department of Mathematics, Universidad Autónoma Metropolitana, México D.F. CP. 09340, México

Submitted by G.-C. Rota

Received October 15, 1986

1. INTRODUCTION

Vandermonde matrices appear frequently in the theories of approximation and interpolation, linear difference and differential equations, and algebraic coding. In many cases Vandermonde matrices appear as coefficient matrices of systems of equations which must be solved. In other cases the inverse matrix is required. See for example Davis [3], Kalman [8], and Schumaker [12].

It is well known that under certain simple hypothesis confluent Vandermonde matrices are invertible. This fact is often used to show existence and uniqueness of solutions of a great variety of problems, for instance, Davis [3] does this for several interpolation problems.

On the other hand, it seems that there is not much known about the structure and algebraic properties of the inverses of confluent Vandermonde matrices.

In this paper we will prove several matrix equations involving generalized Vandermonde matrices, which give explicit algebraic information about the inverse matrices. For example, Eq. (3.20) says that the inversion of a confluent Vandermonde matrix V is accomplished multiplying V by certain permutation matrices, a triangular change of basis matrix and a block diagonal matrix whose blocks are triangular Toeplitz matrices. All these matrices are completely determined, they have a simple structure and are easily computed.

A convenient way to deal with Vandermonde matrices is through their connection with polynomial interpolation. We define generalized Vandermonde matrices as the transposes of certain evaluation operators on a space of polynomials. Let $z_0, z_1, z_2, \ldots, z_n$ be distinct complex numbers, let $m_0, m_1, m_2, \ldots, m_n$ be positive integers and let $N = \sum m_i - 1$. Denote by E the evaluation operator that assigns to each polynomial q of degree at

0022-247X/88 $3.00

409/131/2-4

most N the vector whose components are the numbers $D^k q(z_i)/k!$, $i = 0, 1, 2, ..., n$, $k = 0, 1, 2, ..., m_i - 1$.

If we consider the standard power basis $1, x, x^2, ..., x^N$ for the vector space Π of polynomials of degree at most N, and the canonical basis for $\mathbb{C}^{N+1}$, then the operator E is represented by the transpose of a classical confluent Vandermonde matrix. If instead of the power basis we consider any other basis for Π then E has as its matrix representation the transpose of what we will call a generalized Vandermonde matrix.

The inverse of the operator E is a Hermite interpolation operator. Therefore the inverses of generalized Vandermonde matrices are very important for the study of interpolation operators.

In [8] Macon and Spitzbart obtained an explicit formula for the entries of the inverse of a nonconfluent Vandermonde matrix in terms of symmetric functions. A generalization of their formula is easily obtained by translating into the language of symmetric functions some of our results.

Bjorck and Pereyra [2] considered confluent Vandermonde systems of linear equations and proposed algorithms for the solution of such systems, essentially based on Newton's divided difference formula for polynomial interpolation.

See Golub and Van Loan [4] for other references on the solution of Vandermonde systems.

In this paper we will use Horner's algorithm and some simple properties of the divided difference of polynomials in order to obtain the inverse of a generalized Vandermonde matrix.

We will show that the inverse of a transpose of a generalized Vandermonde matrix is the product of a generalized Vandermonde matrix by a block diagonal matrix whose blocks are triangular Toeplitz matrices.

Our results yield an algorithm for the inversion of confluent Vandermonde matrices that requires only one division for each interpolation point. The other arithmetical operations (sums and products) are done using Horner's algorithm for the computation of derivatives.

Henrici [5, Section 6.1, Problems 1, 2, 3] mentions the connection between Horner's algorithm and divided differences but he does not use it for the inversion of Vandermonde matrices.

In Section 2 we consider the simple case of nonconfluent Vandermonde matrices associated with Lagrange interpolation.

In Section 3 we deal with confluent Vandermonde matrices (associated with Hermite interpolation) in Theorem 3.1. The general case of generalized Vandermonde matrices with respect to any polynomial basis appears in Corollary 3.3.

Our results may be considered as an initial step in the algebraic study of Vandermonde matrices and their inverses. There is much to be studied about the relations of Vandermonde matrices with groups of matrices, symmetric functions, operators on spaces of polynomials, generating functions, and combinatorial identities. A good setting for such study is Rota's finite operator calculus.

2. Horner's Algorithm and Lagrange Interpolation

In this section we will use the algebraic foundations of Horner's algorithm in order to obtain formulas for the Lagrange interpolation polynomial. Such formulas are equivalent to the inversion of (transposed) nonconfluent Vandermonde matrices, as we will show.

Let N be a fixed nonnegative integer and let Π be the complex vector space of polynomials of degree at most N.

For any polynomial

$$w(x) = \sum_{k=0}^{N+1} b_k x^{N+1-k}$$

with $b_0 \neq 0$, its divided difference, defined by

$$w[x, t] = \frac{w(x) - w(t)}{x - t}, \tag{2.1}$$

is a symmetric polynomial in x and t, of degree N in each variable, which may be expressed in the form

$$w[x, t] = \sum_{k=0}^{N} \sum_{j=0}^{N-k} b_k x^{N-k-j} t^j. \tag{2.2}$$

Changing the order of the sums in (2.2) and using the symmetry of the divided difference we get

$$w[x, t] = \sum_{j=0}^{N} p_{N-j}(x)\, t^j = \sum_{j=0}^{N} p_{N-j}(t)\, x^j, \tag{2.3}$$

where

$$p_j(x) = b_0 x^j + b_1 x^{j-1} + \cdots + b_j, \qquad 0 \leqslant j \leqslant N. \tag{2.4}$$

We will call these polynomials the Horner polynomials of w. Note that they form a basis for Π and satisfy the recurrence formula

$$p_{j+1}(x) = x p_j(x) + b_{j+1}, \qquad 0 \leqslant j \leqslant N. \tag{2.5}$$

Since $p_{N+1} = w$, Eq. (2.5) is just Horner's algorithm for the computation of $w(x)$.

It is clear that $w[x, x] = w'(x)$ for any x and also that $w[x_1, x_2] = 0$ whenever x_1 and x_2 are distinct roots of w. These basic properties of the divided difference allow us to prove the following theorem.

THEOREM 2.1. *Let $z_0, z_1, z_2, ..., z_N$ be distinct real or complex numbers, let $y_0, y_1, y_2, ..., y_N$ be arbitrary numbers and let $P(x)$ be the unique element of Π that satisfies $P(z_j) = y_j$, $0 \leqslant j \leqslant N$, i.e., the Lagrange interpolation polynomial for the given data. Then*

$$P(x) = \sum_{j=0}^{N} \sum_{k=0}^{N} p_{N-j}(z_k) \frac{y_k}{w'(z_k)} x^j, \tag{2.6}$$

and

$$P(x) = \sum_{j=0}^{N} \sum_{k=0}^{N} z_k^{N-j} \frac{y_k}{w'(z_k)} p_j(x), \tag{2.7}$$

where the p_j are the Horner polynomials of

$$w(x) = (x - z_0)(x - z_1)(x - z_2) \cdots (x - z_N).$$

Proof. Let

$$l_k(x) = \frac{w[x, z_k]}{w'(z_k)}, \qquad 0 \leqslant k \leqslant N. \tag{2.8}$$

Then we have

$$l_k(z_j) = \delta_{j,k}, \qquad 0 \leqslant j, k \leqslant N, \tag{2.9}$$

and therefore

$$P(x) = \sum_{k=0}^{N} y_k l_k(x) = \sum_{k=0}^{N} \frac{w[x, z_k]}{w'(z_k)} y_k. \tag{2.10}$$

Substitution of (2.3) in (2.10), with $t = z_k$, gives formulas (2.6) and (2.7).

We introduce next some vector and matrix notation that simplifies the interpretation of Theorem 2.1 in terms of matrices and the generalization to the case of general Hermite interpolation.

We define $\mathbf{s}(x) = (1, x, x^2, ..., x^N)^T$, and whenever we have a sequence $q_0, q_1, q_2, ..., q_N$ of polynomials in Π we write

$$\mathbf{q}(x) = (q_0(x), q_1(x), q_2(x), ..., q_N(x))^T.$$

For each positive integer n let J_n denote the n by n permutation matrix that reverses order, that is,

$$J_n(a_1, a_2, ..., a_n)^T = (a_n, a_{n-1}, ..., a_2, a_1)^T.$$

We will omit the subindex and write just J whenever the order of the matrix is clear from the context.

The notation introduced above allows us to write Eq. (2.3) in the form

$$w[x, t] = \mathbf{s}^{\mathrm{T}}(x)\, J\mathbf{p}(t) = \mathbf{s}^{\mathrm{T}}(t)\, J\mathbf{p}(x) \tag{2.11}$$

and Eq. (2.9) in the form

$$\mathbf{s}^{\mathrm{T}}(z_j)\, J\mathbf{p}(z_k) = \delta_{k,j} w'(z_k), \qquad 0 \leqslant j,\ k \leqslant N. \tag{2.12}$$

Writing this equation in matrix notation we obtain

COROLLARY 2.2.

$$\begin{aligned}(\mathbf{s}(z_0), \mathbf{s}(z_1), \ldots, \mathbf{s}(z_N))^{\mathrm{T}}\, J(\mathbf{p}(z_0), \mathbf{p}(z_1), \ldots, \mathbf{p}(z_N))&\\ = \operatorname{diag}(w'(z_0), w'(z_1), w'(z_2), \ldots, w'(z_N)).&\end{aligned} \tag{2.13}$$

The first matrix in (2.13) is the transpose of the usual Vandermonde matrix V at the points z_i. The matrix $(\mathbf{p}(z_0), \mathbf{p}(z_1), \ldots, \mathbf{p}(z_N))$ is a generalized Vandermonde matrix.

Since the z_i are simple roots of w, from (2.13) we get immediately

$$\begin{aligned}(V^{\mathrm{T}})^{-1} = J(\mathbf{p}(z_0), \mathbf{p}(z_1), \ldots, \mathbf{p}(z_N))&\\ \times \operatorname{diag}(1/w'(z_0), 1/w'(z_1), \ldots, 1/w'(z_N)).&\end{aligned} \tag{2.14}$$

Since the coefficients of w, w' and the Horner polynomials p_j are easily expressed in terms of elementary symmetric functions of some of the z_i, from Eq. (2.14) it is easy to obtain an explicit formula for the inverse of V in terms of symmetric functions of the z_i, which of course coincides with the formula given by Macon and Spitzbart [8].

3. THE MAIN THEOREM

In this section we will obtain a matrix equation that generalizes (2.13) and which holds for confluent Vandermonde matrices. We introduce first some terminology.

The usual differentiation operator is represented by D and for each non-negative integer k we let $d^k = D^k/k!$. The operator d^k acts componentwise on the cartesian products Π^j.

Let $\mathbf{q}$ be a column vector in the space Π^j and let m be a positive integer. We denote by $A(\mathbf{q}, m)$ the j by m matrix whose kth column is the vector $d^k\mathbf{q}$, $0 \leqslant k \leqslant m-1$. That is,

$$A(\mathbf{q}, m) = (\mathbf{q}, d\mathbf{q}, d^2\mathbf{q}, \ldots, d^{m-1}\mathbf{q}). \tag{3.1}$$

If $\mathbf{z}=(z_0, z_1, ..., z_n)$ has distinct complex components, $\mathbf{m}=(m_0, m_1, ..., m_n)$ is a vector of positive integers and $\mathbf{q}$ is in Π^j for some $j \geqslant 1$, the generalized Vandermonde matrix of $\mathbf{q}$ at $\mathbf{z}$ with multiplicity vector $\mathbf{m}$ is the j by $\sum m_i$ block matrix

$$V(\mathbf{q}, \mathbf{z}, \mathbf{m}) = (A(\mathbf{q}(z_0), m_0)\, A(\mathbf{q}(z_1), m_1) \cdots A(\mathbf{q}(z_n), m_n)). \tag{3.2}$$

Whenever $\mathbf{z}$ and $\mathbf{m}$ are kept fixed we write $V(\mathbf{q})$ instead of $V(\mathbf{q}, \mathbf{z}, \mathbf{m})$ and V instead of $V(\mathbf{s})$.

Aitken [1] considers square matrices $V(\mathbf{q})$ and calls them confluent alternant matrices, since their determinants are alternating functions of the z_i. Kalman [8] calls such matrices generalized Vandermonde matrices, as we do.

If A is any rectangular matrix we denote its transpose by A^{T} and we define the reverse of A, which we denote by $A^{\#}$, as the matrix obtained from A reversing the order of the rows and the columns, that is, if A is an n by k matrix then

$$A^{\#} = J_n A J_k. \tag{3.3}$$

It is easy to verify that the reversion of matrices has the following properties:

(i) $(A^{\#})^{\#} = A$ and $(A^{\#})^{\mathrm{T}} = (A^{\mathrm{T}})^{\#}$ for any A.

(ii) $(A_1 A_2)^{\#} = A_1^{\#} A_2^{\#}$ whenever $A_1 A_2$ is defined.

(iii) $(A^{-1})^{\#} = (A^{\#})^{-1}$ for any invertible matrix A.

If $V(\mathbf{q}, \mathbf{z}, \mathbf{m})$ is a generalized Vandermonde matrix we define the associated matrix $V^*(\mathbf{q}, \mathbf{z}, \mathbf{m})$ as

$$V^*(\mathbf{q}, \mathbf{z}, \mathbf{m}) = (A^{\#}(\mathbf{q}(z_0), m_0)\, A^{\#}(\mathbf{q}(z_1), m_1) \cdots A^{\#}(\mathbf{q}(z_n), m_n)). \tag{3.4}$$

Using the permutation matrices J we can write

$$V^*(\mathbf{q}, \mathbf{z}, \mathbf{m}) = JV(\mathbf{q}, \mathbf{z}, \mathbf{m}) \operatorname{diag}(J_{m_0}, J_{m_1}, ..., J_{m_n}). \tag{3.5}$$

The associated matrices V^* will also be called generalized Vandermonde matrices.

A square matrix of the form

$$\begin{pmatrix} a_0 & & & & 0 \\ a_1 & a_0 & & & \\ a_2 & a_1 & a_0 & & \\ \vdots & \vdots & \vdots & & \\ a_k & a_{k-1} \cdots & a_2 & a_1 & a_0 \end{pmatrix} \tag{3.6}$$

will be called (lower triangular) Toeplitz matrix and will be denoted by

$$T(a_0, a_1, ..., a_k).$$

Now we are ready to state our main result.

THEOREM 3.1. *Let* $\mathbf{z} = (z_0, z_1, ..., z_n)$ *be a vector with distinct real or complex components, let* $\mathbf{m} = (m_0, m_1, ..., m_n)$ *be a vector with positive integer components and let* $N = \sum m_i - 1$. *Define*

$$w(x) = (x - z_0)^{m_0} (x - z_1)^{m_1} \cdots (x - z_n)^{m_n}$$

and let $\mathbf{p} = (p_0, p_1, ..., p_N)$ *be the vector of Horner polynomials of* w. *Then we have*

$$V^{\mathrm{T}}(\mathbf{s}, \mathbf{z}, \mathbf{m})\, V^*(\mathbf{p}, \mathbf{z}, \mathbf{m}) = \operatorname{diag}(T_0, T_1, ..., T_n), \tag{3.7}$$

where each block

$$T_j = T(d^{m_j + k} w(z_j);\ k = 0, 1, 2, ..., m_j - 1),\ 0 \leqslant j \leqslant n, \tag{3.8}$$

is an invertible lower triangular Toeplitz matrix.

In order to prove Theorem 3.1 we will need some elementary properties of the divided difference of a polynomial, which we state and prove in the following lemma.

LEMMA 3.2. (i) *If* m, n *and* k *are nonnegative integers such that* $m \geqslant n + k$ *then*

$$\sum_{j=k}^{m-n} \binom{j}{k} \binom{m-j}{n} = \binom{m+1}{n+k+1}. \tag{3.9}$$

(ii) *If* w *is a polynomial,* n, k *are nonnegative integers and* a *is any number then*

$$d_t^k d_x^n w[a, a] = d^{k+n+1} w(a) \tag{3.10}$$

where

$$d_t^k = \frac{1}{k!} \frac{\partial^k}{\partial t^k} \qquad \textit{and} \qquad d_x^n = \frac{1}{n!} \frac{\partial^n}{\partial x^n}$$

(iii) *If* a_1, a_2 *are distinct roots of a polynomial* w *with multiplicities* m_1, m_2, *respectively, then*

$$d_t^k d_x^n w[a_1, a_2] = 0, \qquad 0 \leqslant k < m_2,\ 0 \leqslant n < m_1. \tag{3.11}$$

Proof. (i) Let S be the set of all strictly increasing sequences $(a_1, a_2, ..., a_{n+k+1})$ of positive integers at most equal to $m+1$. Note that the right-hand side of (3.9) is the number of elements in S.

For each j such that $k \leqslant j \leqslant m-n$ let S_j be the set of sequences in S having $a_{k+1} = j+1$. A simple counting argument shows that S_j has $\binom{j}{k}\binom{m-j}{n}$ elements, and since the S_j form a partition of S equation (3.9) holds.

(ii) Let us note first that it is enough to show that (3.10) holds for $w(x) = x^{m+1}$, $m \geqslant 0$. For such w we have $w[x, t] = \sum t^j x^{m-j}$ and hence

$$d_t^k d_x^n w[x, t] = \sum_j \binom{j}{k}\binom{m-j}{n} t^{j-k} x^{m-j-n}. \tag{3.12}$$

Taking $x = t = a$ and using (3.9) we get

$$d_t^k d_x^n w[a, a] = \binom{m+1}{n+k+1} a^{m-n-k} = d^{n+k+1} w(a). \tag{3.13}$$

(iii) Since $w[x, t] = (w(x) - w(t))(x-t)^{-1}$, applying Leibniz rule twice we get

$$\begin{aligned} d_t^k d_x^n w[x, t] = \sum_{j=0}^{n} & (-1)^{n-j} d_x^j w(x)(x-t)^{j-k-n-1} \\ - \sum_{j=0}^{k} & (-1)^n d_t^j w(t)(x-t)^{j-k-n-1}, \qquad n, k \geqslant 0. \end{aligned} \tag{3.14}$$

Taking $x = a_1$, $t = a_2$, $n < m_1$, and $k < m_2$ in (3.14) we get (3.11).

The convolution identity (3.9) appears in Riordan [10, p. 35, Problems 12, 13], where it is proved by means of recurrence relations in two variables. Our combinatorial proof is easier and shows that (3.9) is similar to Vandermonde's convolution formula.

The statements (3.10) and (3.11) about the divided difference $w[x, t]$ hold also when w is a sufficiently differentiable function. For a proof in such case see Verde-Star [13].

Now we are ready to prove Theorem 3.1.

Proof of Theorem 3.1. We will compute the product $V^{\mathrm{T}}(\mathbf{s})\, V^*(\mathbf{p})$ by blocks. Let i and j be indices between 0 and n. The ith block of $V^{\mathrm{T}}(\mathbf{s})$ is the m_i by $N+1$ matrix $A^{\mathrm{T}}(\mathbf{s}(z_i), m_i)$, whose kth row is the vector $d^k\mathbf{s}(z_i)$, $0 \leqslant k \leqslant m_i - 1$. The jth block of $V^*(\mathbf{p})$ is the $N+1$ by m_j matrix $A^*(\mathbf{p}(z_j), m_j)$, whose rth column is the vector $d^{m_j-1-r}\mathbf{p}(z_j)$ in reverse order. Therefore, the entry with indices k, r in the product $A^{T}(\mathbf{s}(z_i), m_i)\ A^{\#}(\mathbf{p}(z_j), m_j)$ is

$$\sum_{l=0}^{N} d^k s_l(z_i)\ d^{m_j-1-r} p_{N-l}(z_j), \tag{3.15}$$

which may be written as

$$d_t^k d_x^{m_j-1-r} \sum_{l=0}^{N} s_l(t)\, p_{N-l}(x)\Big|_{\substack{t=z_i\\ x=z_j}} = d_t^k d^{m_j-1-r} w[z_i, z_j]. \tag{3.16}$$

By Lemma 3.2 the expression in (3.16) vanishes whenever $i \neq j$, and it equals $d^{m_j+k-r} w(z_j)$ when $i=j$.

Since z_j is a root of w of multiplicity m_j, $d^{m_j+k-r} w(z_j)$ is zero for $k<r$ and it is different from zero if $k=r$. In fact, Leibniz rule yields

$$d^{m_j+k} w(z_j) = d^k \left[\prod_{\substack{i=0\\ i\neq j}}^{n} (z-z_i)^{m_i} \right]\Bigg|_{z=z_j}, \qquad k \geq 0. \tag{3.17}$$

In particular, we have

$$d^{m_j} w(z_j) = \prod_{\substack{i=0\\ i\neq j}}^{n} (z_j - z_i)^{m_i} \neq 0. \tag{3.18}$$

This completes the proof of Theorem 3.1.

The coefficients b_k of the polynomial w, which are also the coefficients of the Horner polynomials p_j, play an important role in Theorem 3.1. If we let $B=T(b_0, b_1, \ldots, b_N)$ then B is invertible and $B\mathbf{s}=\mathbf{p}$, that is, B transforms the standard basis of Π into the basis formed by the Horner polynomials of w.

Let us denote by W the block diagonal matrix which appears in the right-hand side of Eq. (3.7). Note that W depends only on $\mathbf{z}$ and $\mathbf{m}$, that is W depends only on the polynomial w, as B does.

COROLLARY 3.3. *With the same hypothesis and notation of Theorem* 3.1 *and with B and W as defined above we have*

(i) $V^{T}(s)\, B^{T} V^{*}(\mathbf{s}) = V^{T}(\mathbf{p})\, V^{*}(\mathbf{s}) = W.$ (3.19)

(ii) $[V^{T}(\mathbf{s})]^{-1} = B^{T} V^{*}(\mathbf{s})\, W^{-1}.$ (3.20)

(iii) *For each ordered basis* $\mathbf{q}=(q_0, q_1, ..., q_N)$ *of* Π *there exists a unique ordered basis* $\mathbf{u}=(u_0, u_1, ..., u_N)$ *of* Π *such that*

$$V^{\mathrm{T}}(\mathbf{q})\, V^*(\mathbf{u})=W. \tag{3.21}$$

(iv) $\det V(\mathbf{s})=\prod_{i<j}(z_j-z_i)^{m_j m_i}$. (3.22)

Proof. Since $B\mathbf{s}=\mathbf{p}$, for every positive integer m we have $A(\mathbf{p}, m)=BA(\mathbf{s}, m)$, and since $B^{\#}=B^{\mathrm{T}}$, we have $V^*(\mathbf{p})=B^{\mathrm{T}}V^*(\mathbf{s})$, which combined with Theorem 3.1 gives (3.19). Part (ii) follows from (i), since W is invertible. For the proof of (iii), define the invertible matrix Q by $Q^{\mathrm{T}}\mathbf{q}=\mathbf{s}$ and let $\mathbf{u}=Q^{\#}\mathbf{p}$. Then $V^{\mathrm{T}}(\mathbf{q})=V^{\mathrm{T}}(\mathbf{s})\,Q^{-1}$ and, since $A(\mathbf{u}, m)=Q^{\#}A(\mathbf{p}, m)$ for every m, we have $A^{\#}(\mathbf{u}, m)=QA^{\#}(\mathbf{p}, m)$ and hence $V^*(\mathbf{u})=QV^*(\mathbf{p})$. Therefore, by Theorem 3.1 we have

$$V^{\mathrm{T}}(\mathbf{q})\, V^*(\mathbf{u})=V^{\mathrm{T}}(\mathbf{s})\,Q^{-1}QV^*(\mathbf{p})=W.$$

For part (iv) note that $\det B=1$ and $V^*(s)$ is obtained by permuting rows and columns of $V(\mathbf{s})$, hence computing determinants in (3.19) we get

$$[\det V(\mathbf{s})]^2\,(-1)^r=\det W, \tag{3.23}$$

where

$$r=\binom{N+1}{2}+\sum_{i=0}^{n}\binom{m_i}{2}.$$

On the other hand,

$$\det W=\prod_{j=0}^{n}\left(\prod_{\substack{i=0\\ i\neq j}}^{n}(z_j-z_i)^{m_i}\right)^{m_j}=(-1)^v\prod_{i<j}(z_j-z_i)^{2m_i m_j}, \tag{3.24}$$

where

$$v=\sum_{i<j}m_i m_j.$$

A simple computation shows that either r and v are both even or they are both odd and hence

$$\det V(\mathbf{s})=\pm\prod_{i<j}(z_j-z_i)^{m_i m_j}. \tag{3.25}$$

Induction on the number of blocks in $V(\mathbf{s})$ shows that the sign in the right-hand side of (3.25) must be plus.

Kalman [8] presents several methods that may be used to get formula (3.22).

4. Final Remarks

Let us note that Theorem 3.1 and Horner's algorithm give a simple algorithm for the inversion of a generalized Vandermonde matrix $V^{\mathrm{T}}(\mathbf{s}, \mathbf{z}, \mathbf{m})$. The algorithm is roughly described as follows:

Step 1. Starting with $\mathbf{z}$ and $\mathbf{m}$ as data compute recursively the coefficients of $w(x) = \sum b_k x^{N+1-k}$.

Step 2. Compute the numbers $d^{m_j+k}w(z_j)$, $0 \leqslant j \leqslant n$, $0 \leqslant k < m_j$ (the entries in the T_j) using Horner's algorithm, saving some of the numbers produced by the intermediate computations (the entries of $V^*(\mathbf{p}, \mathbf{z}, \mathbf{m})$).

Step 3. Compute recusively the entries of the Toeplitz blocks T_j^{-1}.

This algorithm gives the inverse of $V^{\mathrm{T}}(\mathbf{s}, \mathbf{z}, \mathbf{m})$ factored in the form

$$V^*(\mathbf{p}, \mathbf{z}, \mathbf{m}) \operatorname{diag}(T_0^{-1}, T_1^{-1}, ..., T_n^{-1}).$$

Note that Step 1 is Horner's algorithm in reverse, since $w(x)$ is obtained recursively, starting with the constant 1, by successive multiplications by factors of the form $x - z_j$.

The inversion of the Toeplitz block T_j is equivalent to the computation of the first m_j terms of a linearly recurrent sequence.

An alternative algorithm uses Horner's scheme to obtain the elements of T_j^{-1}, which are certain complete symmetric functions. This replaces Step 3 and part of Step 2 in the algorithm described above.

A detailed study of the inversion algorithms will appear elsewhere (Verde-Star [15]).

In Section 2 we have mentioned the connection between interpolation formulas and the inversion of Vandermonde matrices. Corollary 2.2 is obtained immediately translating into matrix notation the interpolation formulas of Theorem 2.1.

Proceeding in the opposite direction we can get interpolation formulas from the matrix equations obtained in Theorem 3.1 and Corollary 3.3. For example, Eq. (3.21) may be considered as a generator of interpolation formulas. In the simplest case ($m_j = 1$ for each j) (3.21) yields the following expression for the Lagrange interpolation polynomial

$$P(x) = \sum_{j=0}^{N} \sum_{k=0}^{N} u_{N-j}(z_k) \frac{y_k}{w'(z_k)} q_j(x), \tag{4.1}$$

where the bases $\mathbf{q}$ and $\mathbf{u}$ are related by

$$w[x, t] = \mathbf{q}^{\mathrm{T}}(x) J\mathbf{u}(t). \tag{4.2}$$

For each choice of the interpolation points z_k and the polynomial basis $\mathbf{q}$, Eq. (4.1) gives an explicit expression for $P(x)$. In this way we can get a variety of formulas using only elementary algebraic computations. See Verde-Star [13].

Another interesting particular case is Taylor's interpolation problem. Here we have only one point z_0 and the confluent Vandermonde matrix is a single square block $A(\mathbf{s}(z_0), m)$. Kalman [7] proved that such matrices form a one parameter group and obtained several combinatorial identities.

This suggests that generalized Vandermonde matrices may be studied in a group theoretical setting, since each confluent Vandermonde matrix is made up of rectangular blocks taken from matrices of the form $A((\mathbf{s}(z_k), N+1)$, which belong to a group of Taylor matrices, and the triangular Toeplitz blocks that appear in W also form an abelian group.

It is possible to express the inverse of a confluent Vandermonde matrix $V(\mathbf{s}, \mathbf{z}, \mathbf{m})$ in terms of symmetric functions of the z_k. In (3.20) the entries in B are elementary symmetric functions of the z_k (with alternating signs) and the entries in W^{-1} are complete symmetric functions. See Aitken [1].

Our matrix equations may be written using Hankel matrices instead of Toeplitz matrices. For example, we can write (3.20) in the form

$$(V^{\mathrm{T}})^{-1} = B^{\mathrm{T}} J V \operatorname{diag}(J_{m_0}, J_{m_1}, ..., J_{m_n})\, W^{-1}. \tag{4.3}$$

Note that $B^{\mathrm{T}}J$ is a triangular Hankel matrix and so is each block $J_{m_i} T_{m_i}^{-1}$. Therefore, (4.3) says that, modulo such Hankel matrices, Vandermonde matrices are orthogonal. This suggests that some kind of duality may be useful to study generalized Vandermonde matrices. See Verde-Star [14].

The main tools used in the present paper are properties of divided differences of polynomials. Such divided differences are symmetric polynomials in several variables; they are connected to combinatorics and generating functions and deserve further study. In this direction the ideas of Joni and Rota [6] are certainly very valuable. Another approach to the study of divided differences, based on determinants, appears in Schumaker's book [12].

The remarks above and some of our results seem to indicate that Vandermonde matrices should be studied as linear operators in spaces of polynomials. This may be done in the context of Rota's finite operator calculus. In this direction the ideas in the papers by Roman [11] and Verde-Star [14] are relevant.

As a closing remark we want to point out that Theorem 3.1 may be a useful tool for the study of regularity of Hermite–Birkhoff interpolation problems.

References

1. A. C. Aitken, "Determinants and Matrices," Oliver & Boyd, Edinburgh, 1956.
2. A. Bjork and V. Pereyra, Solution of Vandermonde systems of equations, *Math. Comput.* **24** (1970), 893–903.
3. P. J. Davis, "Interpolation and Approximation," Dover, New York, 1975.
4. G. H. Golub and C. F. Van Loan, "Matrix Computations," Johns Hopkins Press, Baltimore, MD., 1983.
5. P. Henrici, "Applied and computational Complex Analysis," Wiley, New York, 1974.
6. S. A. Joni and G. C. Rota, Coalgebras and bialgebras in Combinatorics, *Stud. Appl. Math.* **61** (1979), 93–139.
7. D. Kalman, Polynomial translation groups, *Math. Mag.* **56** (1983), 23–25.
8. D. Kalman, The generalized Vandermonde matrix, *Math. Mag.* **57** (1984), 15–21.
9. N. Macon and A. Spitzbart, Inverses of Vandermond matrices, *Amer. Math. Monthly* **65** (1958), 95–100.
10. J. Riordan, "Combinatorial Identities," Wiley, New York, 1968.
11. S. Roman, Polynomials, power series and interpolation, *J. Math. Anal. Appl.* **80** (1981), 333–371.
12. L. L. Schumaker, "Spline Functions: Basic Theory," Wiley-Interscience, New York, 1981.
13. L. Verde-Star, Polynomial Interpolation formulas and inverses of Vandermonde matrices, *Sem. Brasil. Anal.* **16** (1982), 181–192.
14. L. Verde-Star, Dual operators and Lagrange inversion in several variables, *Adv. in Math.* **58** (1985), 89–108.
15. L. Verde-Star, An algorithm for the inversion of generalized Vandermonde matrices, to appear.