The Electronic Journal of Linear Algebra.
A publication of the International Linear Algebra Society.
Volume 1, pp. 18-33, June 1996.
ISSN 1081-3810. http://gauss.technion.ac.il/iic/ela

ELA

SOME PROPERTIES OF THE Q-ADIC VANDERMONDE MATRIX*

VAIDYANATH MANI† AND ROBERT E. HARTWIG†

Abstract. The Vandermonde and confluent Vandermonde matrices are of fundamental significance in matrix theory. A further generalization of the Vandermonde matrix called the q-adic coefficient matrix was introduced in [V. Mani and R. E. Hartwig, *Lin. Algebra Appl.*, to appear]. It was demonstrated there that the q-adic coefficient matrix reduces the Bezout matrix of two polynomials by congruence. This extended the work of Chen, Fuhrman, and Sansigre among others. In this paper, some important properties of the q-adic coefficient matrix are studied. It is shown that the determinant of this matrix is a product of resultants (like the Vandermonde matrix). The Wronskian-like block structure of the q-adic coefficient matrix is also explored using a modified definition of the partial derivative operator.

Key words. Vandermonde, Hermite Interpolation, q-adic expansion.

AMS(MOS) subject classification. Primary 15A15, 15A21, 41A10, 12Y05

1. Introduction. Undoubtedly, the Vandermonde matrix is one of the most important matrices in applied matrix theory. Its generalization to the confluent Vandermonde matrix has numerous applications in engineering. This concept was further generalized to the q-adic Vandermonde matrix in [9], extending the work of Chen [4], Fuhrman [7] and Sansigre [3] among others. In this paper, we investigate some of the properties of the q-adic Vandermonde matrix. In particular, we shall evaluate its determinant and analyze its block structure. The former extends the Vandermonde determinant, while the latter shows a remarkable analogy to the layered Wronskian structure of the confluent Vandermonde matrix.

It is well known that the confluent Vandermonde matrix for the polynomial

$$\psi(x) = \prod_{i=1}^{s} (x - \lambda_i)^{m_i},$$

$deg(\psi) = \sum m_i = n$, can be constructed by forming the Taylor expansion of

* Received by the editors on 5 December 1995. Final manuscript accepted on 29 May 1996. Handling editor: Daniel Hershkowitz.

† North Carolina State University, Raleigh, North Carolina 27695-8205, USA (vmani@iterated.com, hartwig@math.ncsu.edu).

x^t in terms of the powers of $(x-\lambda_i)$, i.e.,

$$\begin{array}{lll} 1 & = & 1 \\ x & = & \lambda_i + 1(x-\lambda_i) \\ x^2 & = & \lambda_i^2 + 2\lambda_i(x-\lambda_i) + 1(x-\lambda_i)^2 \\ \vdots & & \\ x^t & = & \lambda_i^t + t\lambda_i^{t-1}(x-\lambda_i) + \ldots \\ \vdots & & \\ x^{n-1} & = & \sum_{j=1}^{n} \binom{n-1}{j-1} \lambda_i^{n-j}(x-\lambda_i)^j. \end{array} \tag{1}$$

If we set $(x-\lambda_i) = q_i$ and denote the column containing the first m_i coefficients in the above expansion of x^t by $[x^t]_{(q_i,m_i)}$, then

$$[x^t]_{(q_i,m_i)} = \begin{bmatrix} \lambda_i^t \\ t\lambda_i^{t-1} \\ \vdots \\ \binom{t}{m_i-1} \lambda_i^{t-m_i+1} \end{bmatrix}_{m_i \times 1}. \tag{2}$$

In case $t < m_i - 1$, we of course replace the elements involving negative powers of λ_i by zero. Collecting these coordinate columns, we may construct the $n \times m_i$ matrix $\Omega_{m_i}^n(\lambda_i)$ as

$$\Omega_{m_i}^n(\lambda_i)^T = \begin{bmatrix} [1]_{(q_i,m_i)} & [x]_{(q_i,m_i)} & \ldots & [x^{n-1}]_{(q_i,m_i)} \end{bmatrix}.$$

In other words, $\Omega_{m_i}^n(a)$ is made up of first m_i columns of Caratheodory's $n \times n$ binomial matrix $\Omega_n(a) = \left[\binom{k}{j} a^{k-j} \right]_{k,j=0}^{n-1,n-1}$. Lastly, the $n \times n$ confluent Vandermonde is defined by matrix

$$\Omega_\eta = \left[\Omega_{m_1}^n(\lambda_1); \Omega_{m_2}^n(\lambda_2); ..; \Omega_{m_s}^n(\lambda_s) \right]. \tag{3}$$

To extend this concept further, we recall the q-adic expansion of a given polynomial $f(x)$ relative to a monic polynomial $q(x)$. Indeed, the following lemma from [2] is elementary.

LEMMA 1.1. *Let $q(x)$ be a monic polynomial of degree l over $\mathbb{F}$. If $f(x)$ is any polynomial over $\mathbb{F}$, then there exist unique polynomials $r_j(x)$ over $\mathbb{F}$, of degree less than l, such that*

$$f(x) = \sum_{j=0}^{k} r_j(x)[q(x)]^j$$

for some finite non-negative integer k.

Let us now use the above lemma to define the q-adic coordinate column of a polynomial. For an arbitrary polynomial $f(x) = a_0+a_1x+\ldots+a_kx^k$, its standard coordinate column relative to the standard basis $\mathfrak{B}_{ST} = \{1, x, \ldots, x^{n-1}\}$ is defined and denoted by

$$\mathbf{f} = [f]_{(st)} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{bmatrix}.$$

Now for a given pair (q, m) made up of a polynomial $q(x)$ of degree l and a positive integer m, the q-adic expansion of $f(x)$ may be written as

$$f(x) = r_0(x) + r_1(x)q + \ldots + r_{m-1}(x)q^{m-1} + q^m(\star),$$

where $\star$ denotes the (polynomial) quotient after division by q^m. The q-adic coordinate column associated with $f(x)$ for the pair (q, m) is now defined as

$$[f]_{(q,m)} = \begin{bmatrix} \mathbf{r}_0 \\ \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_{m-1} \end{bmatrix}_{lm\times 1},$$

where $\mathbf{r}_i$ is the standard coordinate column for $r_i(x)$. It should be noted that
(i) if the degree of $f(x)$ is larger than $q(x)^m$, we simply ignore the quotient after division by q^m.
(ii) the statements $q^m|f$ and $[f]_{(q,m)} = 0$ are equivalent, i.e.,

$$[f]_{(q,m)} = 0 \;\Leftrightarrow\; q^m|f. \tag{4}$$

(iii) if $deg(f) = k$, then $[f]_{(st)}$ is a $(k+1)\times 1$ column vector.

Throughout this paper, we shall denote the i^{th} unit column vector by $\mathbf{e}_i$.

1.1. The q-adic Vandermonde matrix. We are now ready for the construction of the q-adic Vandermonde matrix. In order to repeat the construction given in (1) over fields which are not algebraically closed, let $\psi(x)$ be a foundation polynomial with the factorization

$$\psi(x) = \prod_{i=1}^{s} q_i^{m_i}$$

with $deg(q_i) = l_i \geq 1$ and $deg(\psi) = \sum_{i=1}^{s} m_il_i = n$. Again, using the q-adic expansion of the standard basis elements x^t, we have

$$x^t = \sum_{j=0}^{m_i-1} r_{ij}^{(t)}(x)q_i(x)^j + q_i(x)^{m_i}(\star).$$

The q_i-adic coordinate column $[x^t]_{(q_i,m_i)}$ may now be formed as

$$(5) \qquad [x^t]_{(q_i,m_i)} = \begin{bmatrix} \mathbf{r}_{i0}^{(t)} \\ \mathbf{r}_{i1}^{(t)} \\ \vdots \\ \mathbf{r}_{i,m_i-1}^{(t)} \end{bmatrix},$$

where $\mathbf{r}_{ij}^{(t)}$ is the standard coordinate column for $r_{ij}^{(t)}(x)$. Lastly, collecting coordinate columns, we set

$$(6) \qquad W_i = \begin{bmatrix} [1]_{(q_i,m_i)} & [x]_{(q_i,m_i)} & \dots & [x^{n-1}]_{(q_i,m_i)} \end{bmatrix}_{l_i m_i \times n}$$

and construct the $n \times n$ q-adic Vandermonde matrix W as

$$W^T = \begin{bmatrix} W_1 \\ W_2 \\ \dots \\ W_s \end{bmatrix}$$

It is clear that when the q_i's are linear polynomials then, (2) reduces to (6) and W^T to Ω^T. We shall as such refer to the matrix W as the q-adic Vandermonde matrix. If W is nonsingular, it solves the q-adic interpolation problem for polynomials as discussed in [9]. Moreover, for any polynomial $f(x)$ with degree less than n and standard coordinate column $[f]_{(st)}$, it is clear from (5) that

$$(7) \qquad W_i[f]_{(st)} = [f]_{(q_i,m_i)},$$

In other words, W_i acts as a *change of basis matrix* from the standard to the q_i-adic coordinates. This is a crucial observation which will be used in the next section.

To analyze W further, we need the relation between the q-adic coordinate columns of $[f]_{(q,m)}$ and $[\lambda^k f]_{(q,m)}$. This relation is best handled via the concept of a polynomial in a hypercompanion matrix which we shall now address.

1.2. Some basic results hypercompanion matrices. If $q(x) = a_0 + a_1 x + \ldots + x^l$ is a monic polynomial of degree l, then the hypercompanion matrix for $q(x)^m$ is the $m \times m$ block matrix

$$(8) \qquad H_{q^m} = \begin{bmatrix} L_q & 0 & . & .. & 0 \\ N & L_q & 0 & .. & 0 \\ . & . & . & . & 0 \\ 0 & 0 & N & L_q & 0 \\ 0 & 0 & 0 & N & L_q \end{bmatrix},$$

where $N = E_{1,l} = \mathbf{e}_1\mathbf{e}_l^T$ and L_q is the companion matrix of $q(x)$ defined by

$$L_q = \begin{bmatrix} 0 & 0 & 0 & \ldots & -a_0 \\ 1 & 0 & 0 & \ldots & -a_1 \\ 0 & 1 & 0 & \ldots & .. \\ . & . & . & \ldots & .. \\ . & . & . & \ldots & -a_{k-2} \\ 0 & 0 & 0 & ..1 & -a_{k-1} \end{bmatrix}.$$

The key result which we need is the following lemma dealing with polynomials of hypercompanion matrices.

LEMMA 1.2. *Let $p(x)$ be a monic polynomial of degree l and let $H_m = H[p(x)^m]$ be the $ml \times ml$ hypercompanion matrix induced by $p(x)$. If $f(x)$ is any polynomial over $\mathbb{F}$ with p-adic expansion*

$$f(x) = \alpha_0(x) + \alpha_1(x)p(x) + \ldots + \alpha_{m-1}(x)p(x)^{m-1} + p(x)^m(\star), \tag{9}$$

and coordinate columns $\boldsymbol{\alpha}_i$, then

$$f(H_m) = \begin{bmatrix} A_0 & 0 & \ldots & 0 \\ A_1 & A_0 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{m-1} & A_{m-2} & \ldots & A_0 \end{bmatrix} = [\ U, \quad JU, \quad \ldots, \quad J^{m-1}U\ ], \tag{10}$$

where $A_0 = \alpha_0(L_p)$, $A_i = \alpha_i(L_p) + \alpha_{i-1}(L_p + N) - \alpha_{i-1}(L_p)$, $i = 1, \ldots, m-1$, $N = E_{1,l}$ and $U = [\mathbf{a}, H\mathbf{a}, \ldots, H^{l-1}\mathbf{a}]$ with $J = J_m(0) \otimes I_l = p(H_m)$ and

$$\mathbf{a} = \begin{bmatrix} \boldsymbol{\alpha}_0 \\ \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_{m-1} \end{bmatrix}.$$

Proof. To calculate a polynomial in the hypercompanion matrix $H_m = H[p(x)^m]$, we proceed in two stages. First suppose that $f(x)$ has degree $deg(f) < deg(p) = l$. Now split H_m as $H_m = I_m \otimes L + J_m(0) \otimes N$, and since

$$NL^kN = 0, \; k = 0, 1, \ldots, l-2, \tag{11}$$

it follows by induction that $(L + N)^k = L^k + Y_k$, $k = 0, 1, \ldots, l$, and that

$$H_m^k = I \otimes L^k + J_m(0) \otimes Y_k, \; k = 0, 1, \ldots, l \tag{12}$$

where $Y_k = L^{k-1}N + L^{k-2}NL + \ldots + NL^{k-1}$ and $Y_0 = 0$.
Next, we define the difference operator for fixed $p(x)$ and $f(x)$ by

$$\Delta f = f(L + N) - f(L).$$

Summing up (12) shows that if $deg(f) < l$, then

$$f(H_m) = I_m \otimes f(L) + J_m(0) \otimes \Delta f. \tag{13}$$

This formula breaks down when $deg(f) \geq l$, since (11) does not hold for values of $k \geq l-2$. But we may proceed in the following way. Suppose $f(x)$ admits the p-adic expansion

$$f(x) = \alpha_0(x) + \alpha_1(x)p(x) + \ldots + \alpha_{m-1}(x)p(x)^{m-1} + p(x)^m(\star).$$

Now recall from [6] that $J = p(H_m) = J_m(0) \otimes I_l$. Clearly, $p(H_m)^k = 0$ for $k \geq l$. Hence by (13), $\alpha_k(H_m)p(H_m)^k$ has the form

$$\alpha_k(H_m)p(H_m)^k = J_m(0)^k \otimes \alpha_k(L) + J_m(0)^{k+1} \otimes \Delta\alpha_k,$$

for $k = 0, 1, \ldots, l-1$. Summing this for $k = 0, 1, \ldots, l-1$, we see that $f(H_m)$ is block Toeplitz of the form given in (10).
For the remaining part it suffices to show that U agrees with the matrix $A = \begin{bmatrix} A_0 \\ A_1 \\ \cdots \\ A_{m-1} \end{bmatrix}$. It is well known that for a polynomial $h(x) = h_0 + h_1x + \ldots + h_{n-1}x^{n-1}$, $h(L) = [\mathbf{h}\ L\mathbf{h}\ \ldots\ L^{n-1}\mathbf{h}]$, where L is any $n \times n$ companion matrix. Moreover, $(\widetilde{L}) = L + N$, $N = E_{1,n}$ is also a companion matrix. From the above two observations, it follows that

$$A_0\mathbf{e}_1 = \alpha_k(L)\mathbf{e}_1 + \alpha_{k-1}(\widetilde{L})\mathbf{e}_1 - \alpha_{k-1}(L)\mathbf{e}_1 = \boldsymbol{\alpha}_k + \boldsymbol{\alpha}_{k-1} - \boldsymbol{\alpha}_{k-1} = \boldsymbol{\alpha}_k.$$

Needless to say, this ensures that $A\mathbf{e}_1 = \mathbf{a} = f(H_m)\mathbf{e}_1$. Consequently, $H_m\mathbf{a} = f(H_m)H_m(\mathbf{e}_1) = f(H_m)\mathbf{e}_2$. Similarly, $H_m^k\mathbf{a} = f(H_m)\mathbf{e}_{k+1}$ for $k = 0, 1, \ldots, l-1$ and hence $U = [\mathbf{a},\ H_m\mathbf{a}\ \ldots\ H_m^{l-1}\mathbf{a}]$, completing the proof. □

An immediate corollary to Lemma 1.2 gives us the required result relating the q-adic coordinate column of $[f]_{(q,m)}$ and $[\lambda^k f]_{(q,m)}$.

Corollary 1.3. *Let* $[f]_{(q,m)} = \mathbf{a}$ *for polynomials* $f(x)$ *and* $q(x)$ *as in Lemma 1.2. Then* $[\lambda^k f]_{(q,m)} = (H_m)^k\mathbf{a}$ *and* $[q^k f]_{(q,m)} = J^k\mathbf{a}$, *where* $H_m = H[q^m]$ *and* $J = J_m(0) \otimes I_l$.

Proof. From Lemma 1.2, we see that for any polynomial $g(x)$, $[g]_{(q,m)} = g(H_m)\mathbf{e}_1$. Consequently, if $g(x) = \lambda^k f(x)$, then

$$[g]_{(q,m)} = g(H_m)\mathbf{e}_1 = (H_m)^k f(H_m)\mathbf{e}_1 = (H_m)^k\mathbf{a}.$$

The remaining result is clear since $q(H_m) = J$. □

Given the special construction of W, it comes as no surprise that W is directly related to the hypercompanion matrix. Indeed, since $[1]_{(q_i,m_i)} = \mathbf{e}_1$, it is clear from Corollary 1.3 that $[\lambda^k]_{(q_i,m-i)} = H_i^k\mathbf{e}_1$, where $H_i = H[q_i^{m_i}]$. Consequently,

$$W_i = \begin{bmatrix} \mathbf{e}_1 & H_i\mathbf{e}_1 & \ldots & H_i^{n-1}\mathbf{e}_1 \end{bmatrix}. \tag{14}$$

1.3. The E matrix. In order to calculate the determinant of W, our strategy shall be to block diagonalize W^T first, and then compute the determinant of the resulting block diagonal matrix. In other words, we wish to find a suitable non singular matrix $E = [E_1, E_2, \ldots, E_s]$ such that $W^T E = diag(D_1, D_2, \ldots, D_s)$, i.e., we want

$$W_i E_j = \begin{cases} 0 & i \neq j \\ D_i & i = j, \end{cases}$$

where W_i is $l_i m_i \times n$ and E_j is $n \times l_j m_j$.

Based on (4) and (7), it suffices to take as the columns of E_j, the standard coordinate columns of suitable polynomials all divisible by $q_i^{m_i}$. In addition, we want E to be such that the blocks $D_i = W_i E_i$ are manageable and $det(E) \neq 0$.

A convenient choice is the chain of polynomials

$$\Pi_j = [x^k q_j(x)^r \psi_j(x)] = \psi_j(x)[\Lambda_j, q_j \Lambda_j, \ldots, q_j^{m_j - 1} \Lambda_j]$$

where $\psi_j = \psi / q_j^{m_j}$, $\Lambda_j = [1, x, \ldots, x^{l_j - 1}]$, $j = 1, \ldots, s$, $r = 0, \ldots, m_j - 1$ and $k = 0, 1, \ldots, l_j - 1$. Clearly, $q_i^{m_i}$ divides every polynomial in Π_j for $i \neq j$. As mentioned earlier, E_j is defined by

$$\Pi_j = \mathfrak{B}_{ST} E_j. \tag{15}$$

From (4), we have

$$W_i [x^k q_j(x)^r \psi_j(x)]_{(st)} = [x^k q_j(x)^r \psi_j(x)]_{(q_i, m_i)} \tag{16}$$

and hence for $j \neq i$, $W_i E_j = 0$. On the other hand when $j = i$, we recall Corollary 1.3, so that (16) reduces to

$$W_i [x^k q_i(x)^r \psi_i(x)]_{(st)} = [\lambda^k q_i^r \psi_i]_{(q_i, m_i)} = J_i^r H_i^k \mathbf{a},$$

where $H_i = H[q_i^{m_i}]$, $J_i = J_{m_i} \otimes I_{l_i}$ and $\mathbf{a} = [\psi_i]_{(q_i, m_i)}$. Now if $[U] = [\mathbf{a}, H_i \mathbf{a}, \ldots, H_i^{l_i - 1} \mathbf{a}]$ then an application of Lemma 1.2 yields

$$W_i E_i = \begin{bmatrix} U & JU & \ldots & J^{m_i - 1} U \end{bmatrix} = \psi_i(H_i).$$

We have thus shown the following result.

LEMMA 1.4. *For the matrices W and E described above*

$$W^T E = diag[\psi_i(H_i)]_{i=1}^s, \tag{17}$$

where $\psi_i = \psi / q_i^{m_i}$ and $H_i = H(q_i^{m_i})$ is the hypercompanion matrix associated with $q_i^{m_i}$.

We close this section with the following remarks.

REMARK 1.5. The blocks E_r in $E = [E_1, E_2, \ldots, E_s]$ may be further described as

$$E_r = \psi_r(L)\left[\left[\begin{array}{c} I_{l_i} \\ 0 \end{array}\right],\ q_i(L)\left[\begin{array}{c} I_{l_i} \\ 0 \end{array}\right],\ \ldots\ldots,\ q_i(L)^{m_i-1}\left[\begin{array}{c} I_{l_i} \\ 0 \end{array}\right]\right]; \tag{18}$$

see (5.4) of [5]. To obtain the above, simply note that

$$\begin{aligned} \psi_r(L)q_r(L)^k\left[\begin{array}{c} I_{l_r} \\ 0 \end{array}\right] &= q_r(L)^k\left[\psi_r, L\psi_r, \ldots, L^{l_r-1}\psi_r\right], \\ &= \left[[q_r^k\psi_r]_{(st)}, [xq_r^k\psi_r]_{(st)}, \ldots, [x^{l_r-1}q_r^k\psi_r]_{(st)}\right], \end{aligned}$$

in which $L = L_\psi$ and $[\psi_r]_{(st)} = \psi_r$ are respectively the companion matrix and the standard coordinate column of $\psi(x)$. Since $n \geq deg(\psi_r) + l_r$, we see that for $k = 0, 1, \ldots, l_r - 1$, $L^k\psi_r$ merely shifts the nonzero entries in ψ_r without losing any of them.

REMARK 1.6. In the next section we shall explicitly calculate the determinant of the E matrix. It will be shown that $det(E)$ is nonzero, and hence E is nonsingular, if and only if the polynomials q_i are pairwise coprime, i.e, $gcd(q_i, q_j) = 1$ for $i \neq j$.

REMARK 1.7. It was shown in [9] that the polynomials Π form a Jacobson chain basis for the space $\mathbb{F}_{n-1}[x]$ of polynomials of degree less than n with respect to a shift operator $\mathcal{S}$. Hence, when the q_i's are pairwise coprime, E is a change of basis matrix from the standard to a chain basis denoted by $\mathfrak{B}_{CH}$. The reader is referred to [9] for more details.

Let us now capitalize on (17) and simultaneously calculate the determinant of E as well as that of W.

2. Determinants. It is well known that if Ω is the confluent Vandermonde matrix for the polynomial $\eta(x) = \prod_{i=1}^{s}(x - \lambda_i)^{m_i}$, then

$$det(\Omega) = \prod_{1 \leq j < i \leq s}(\lambda_i - \lambda_j)^{m_i m_j},$$

where the product is taken over all the roots λ_i . We shall now show that the determinant of the q-adic Vandermonde matrix W induced by $\psi(x) = \prod_{i=1}^{s} q_i^{m_i}$ is given by

$$det(W) = \prod_{1 \leq j < i \leq s} R(q_i, q_j)^{m_i m_j},$$

where $R(q_i, q_j)$ is the resultant of q_i and q_j. This shows that $det(W)$ is non-zero if and only if the q_i's are pairwise coprime.

From (17), we see at once that

$$det[W^T]det[E] = det[W]det[E] = \prod_{i=1}^{s} det[\psi_i(H_i)], \tag{19}$$

and thus in order to calculate $det(W)$, we shall need to evaluate $det(E)$ and $det[\psi_i(H_i)]$, $i = 1, \ldots, s$.

Our strategy shall be to factor E as $E = MT$, in which $det(T) = 1$ and M is a generalized Sylvester matrix. The determinant of M can be calculated with the aid of some results from [5] and this will enable us to compute $det(E)$. We then evaluate $det[\psi_i(H_i)]$ and combine the results to express the $det(W)$ in terms of the factors of the foundation polynomial $\psi(x) = \prod_{i=1}^{s} q_i^{m_i}$.

In the next subsection, we undertake this factorization of the E matrix.

2.1. Resultants and the E matrix. Before we can factor E, we shall need to introduce the Sylvester matrix of two polynomials. We first define the $(r+k) \times r$ striped matrix $S_r(f)$ for a polynomial $f(x) = a_0 + a_1 x + \ldots + a_k x^k$ to be

$$S_r(f) = \begin{bmatrix} a_0 & \ldots & \ldots & 0 \\ a_1 & a_0 & \ldots & \\ \vdots & a_1 & \ddots & \\ a_k & & \ddots & a_0 \\ & a_k & & a_1 \\ & & \ddots & \vdots \\ & 0 & \ldots & a_k \end{bmatrix}, \tag{20}$$

The striped matrices of two polynomials $f_1(x)$ and $f_2(x)$ of degrees n_1 and n_2 respectively can be combined to yield the square Sylvester matrix $S(f_1, f_2)$ given by

$$S(f_1, f_2) = [S_{n_2}(f_1) | S_{n_1}(f_2)]$$

whose determinant $R(f_1, f_2)$ is the resultant of f_1 and f_2. The Sylvester matrix can be constructed only for two polynomials. It is possible to extend this construction to a set of more than two polynomials as follows. A more thorough discussion can be found in [5].

Let $\{f_1(x), f_2(x), \ldots, f_s(x)\}$ be a set of polynomials with $deg(f_i) = n_i$. Moreover, let $f = f_1 f_2 \ldots f_s$, $deg(f) = \sum_{i=1}^{s} n_i = n$ and construct the set of polynomials $\{\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_s\}$ by $\mathcal{F}_i = f/f_i$. Now define the square $n \times n$ generalized Sylvester matrix $M(f_1|f_2 \ldots |f_s)$ to be

$$M = M(f_1|f_2 \ldots |f_s) = [S_{n_1}(\mathcal{F}_1), S_{n_2}(\mathcal{F}_2), \ldots, S_{n_s}(\mathcal{F}_s)]. \tag{21}$$

Clearly

$$M(f_1|f_2) = S(f_2, f_1). \tag{22}$$

The following lemma (corollary 5, page 24, [5]) gives a factorization of the matrix M.

LEMMA 2.1. *For the generalized Sylvester matrix $M = M(f_1|f_2 \dots |f_s)$ described above,*

$$\begin{aligned} M &= M(f_s|f_1 \dots f_{s-1}) \begin{bmatrix} M(f_{s-1}|f_1 \dots f_{s-2}) & 0 \\ 0 & I \end{bmatrix} \\ &\times \dots \begin{bmatrix} M(f_3|f_1 f_2) & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} M(f_2|f_1) & 0 \\ 0 & I \end{bmatrix}. \end{aligned}$$

where I denotes the identity matrix of appropriate size.

Using the above factorization, it is easy to calculate $det(M)$ as follows.

COROLLARY 2.2. *The determinant of the matrix $M = M(f_1|f_2 \dots |f_s)$ defined above is*

$$det(M) = \prod_{1 \le i < j \le s} R(f_i, f_j),$$

where $R(f_i, f_j)$ is the resultant of f_i and f_j.

Proof. The proof easily follows from (22), Lemma 2.1 and the product rule for resultants i.e.

$$R(f_1, f_2 f_3) = R(f_1, f_2) R(f_1, f_3). \quad \square$$

We are now ready to give the desired factorization. Let us first factor the chain E_r given in (18) as follows

$$(23) \; E_r = \psi_r(L) \left[\begin{bmatrix} I_{l_r} \\ 0 \end{bmatrix}, \; q_r(L) \begin{bmatrix} I_{l_r} \\ 0 \end{bmatrix}, \; \dots, \; q_r(L)^{m_r - 1} \begin{bmatrix} I_{l_i} \\ 0 \end{bmatrix} \right] = \hat{\psi}_r(L) T_r,$$

where

$$T_r = \left[\begin{bmatrix} I_{l_r} \\ 0 \end{bmatrix}, \; q_r(L) \begin{bmatrix} I_{l_r} \\ 0 \end{bmatrix}, \; \dots, \; q_r(L)^{m_r - 1} \begin{bmatrix} I_{l_i} \\ 0 \end{bmatrix} \right]$$

is upper triangular with a diagonal of ones and $\widehat{\psi_r}(L_\psi)$ contains the first $m_r l_r$ columns of $\psi_r(L_\psi)$. Since the $deg(\psi_r(x)) = m_r l_r - 1$ and L_ψ is $n \times n$, we may write

$$\hat{\psi}_r(L_\psi) = S_{m_r l_r}(\psi_r),$$

where the striped matrix $S_r(f)$ was defined as in (20). Using (18) and (23), the induced factorization of E becomes

$$E = \begin{bmatrix} \hat{\psi}_1(L_\psi) & \hat{\psi}_2(L_\psi) & \dots & \hat{\psi}_s(L_\psi) \end{bmatrix} diag[T_i]_{i=1}^s = MT.$$

We now observe that

$$\begin{aligned} M &= \begin{bmatrix} \hat{\psi}_1(L_\psi), & \hat{\psi}_2(L_\psi), & \dots, & \hat{\psi}_s(L_\psi) \end{bmatrix} \\ &= [\; S_{m_1 l_1}(\psi_1), \quad \dots, \quad S_{m_s l_s}(\psi_s), \;] \end{aligned}$$

is the generalized Sylvester matrix of the polynomials $\{q_1^{m_1}, q_2^{m_2}, \ldots, q_s^{m_s}\}$ defined in (21). Since T is upper triangular and block diagonal, with a diagonal of ones, $det(T) = 1$ and hence from Corollary 2.1

$$(24) \qquad det(E) = det(M) = \prod_{1 \leq i < j \leq s} R(f_i, f_j) = \prod_{1 \leq i < j \leq s} R(q_i, q_j)^{m_i m_j}.$$

We have thus shown the following result.

LEMMA 2.3. *For the matrix E given in* (15),

$$det(E) = \prod_{1 \leq i < j \leq s} R(q_i, q_j)^{m_i m_j}$$

where $R(q_i, q_j)$ is the resultant of $q_i(x)$ and $q_j(x)$.

In the next subsection, we evaluate $det[\psi_i(H_i)]$ and then calculate $det(W)$ according to (19).

2.2. The determinant of W. Recall from (8) and Lemma 1.2 that $\psi_i(H_i)$ is a $m_i \times m_i$ block upper triangular matrix with diagonal blocks $\psi_i(L_i)$ where $L_i = L_{q_i}$ is a companion matrix. Note further that

$$(25) \qquad \psi_i(L_i) = \prod_{\substack{j=1 \\ j \neq i}}^{s} (q_j(L_i))^{m_j}.$$

It is well known (see [1] for example) that

$$(26) \qquad det[q_j(L_{q_i})] = R(q_i, q_j),$$

where $R(q_i, q_j)$ is the resultant of q_i and q_j. Using (25), (26), and the product rule for determinants, the quantity $det[\psi_i(L_i)]$ can be readily calculated to be

$$det[\psi_i(L_i)] = \prod_{\substack{j=1 \\ k \neq i}}^{s} R(q_i, q_j)^{m_j}.$$

Since $\psi_i(H_i)$ consists of m_i diagonal blocks of the form $\psi_i(L_i)$, we have that

$$det[\psi_i(H_i)] = \prod_{\substack{j=1 \\ k \neq i}}^{s} R(q_i, q_j)^{m_i m_j},$$

and repeating this for $i = 1, \ldots, s$, we arrive at the expression

$$\det\left(diag[\psi_i(H_i)]_{i=1}^{s}\right) = \prod_{i=1}^{s} \prod_{\substack{j=1 \\ j \neq i}}^{s} R(q_i, q_j)^{m_i m_j}. \tag{27}$$

We are now ready to put the pieces together. Let us recall from (17) that

$$W^T E = diag[\psi_i(H_i)]_{i=1}^{s}.$$

and hence taking determinants on both sides of this equation yields

$$det(W^T)det(E) = det(W)det(E) = \prod_{i=1}^{s} det[\psi_i(H_i)]. \tag{28}$$

Substituting from (24) and (27), equation (28) can be rewritten as

$$det(W) \prod_{1 \leq i < j \leq s} R(q_i, q_j)^{m_i m_j} = \prod_{i=1}^{s} \prod_{\substack{j=1 \\ j \neq i}}^{s} R(q_i, q_j)^{m_i m_j}. \tag{29}$$

We now need to consider two cases. First suppose that $gcd(q_i, q_j) = 1$ for $i \neq j$. In this case, $det(E)$ and the right hand side of (29) are both non zero since $R(q_i, q_j) \neq 0$. Hence we may cancel $det(E)$ from both sides of (29) to obtain

$$det(W) = \prod_{1 \leq j < i \leq s} R(q_i, q_j)^{m_i m_j}. \tag{30}$$

Next, assume that $gcd(q_r, q_t) = h(x)$, $deg(h) \geq 1$ for some distinct r, t, $1 \leq t < r \leq s$. We claim that in this case, W is singular.

To show this, consider the polynomial $\widetilde{\psi} = \psi/h$ where $deg(\widetilde{\psi}) \leq n - 1$. Now note that for $1 \leq i \leq s$, $q_i^{m_i} | \widetilde{\psi}$ and thus from (7) and (4), we have

$$W_i \widetilde{\psi} = [\widetilde{\psi}(x)]_{(q_i, m_i)} = 0,$$

i.e, $W\widetilde{\psi} = 0$ and consequently $det(W) = 0$. Needless to say $R(q_r, q_t) = 0$ and hence (30) also holds in this case.

We reiterate the fact that in the above discussion, no assumption has been made about the polynomials q_i being prime or irreducible. Hence (30) holds for every factorization of $\psi(x) = \prod_{i=1}^{s} q_i^{m_i}$ of the polynomial ψ. We have thus proven the following theorem.

THEOREM 2.4. *Let $\psi(x) = \prod_{i=1}^{s} q_i^{m_i}$ be any (not necessarily prime) factorization of a polynomial $\psi(x)$ of degree n. Let W be the q-adic matrix corresponding to this factorization. Then*

$$det(W) = \prod_{1 \le j < i \le s} R(q_i, q_j)^{m_i m_j}.$$

In particular W is nonsingular if and only if $gcd(q_i, q_j) = 1$ for $i \neq j$.

So far we have demonstrated one way in which the q-adic Vandermonde matrix generalizes the Vandermonde matrix. In the next section, we consider the Wronskian structure of the confluent Vandermonde matrix and show that the q-adic Vandermonde matrix has a similar block structure.

3. The Structure of W. We conclude this paper by examining the block structure of the q-adic Vandermonde matrix W. Recall from (3) that the confluent Vandermonde matrix Ω, for $\eta = \prod_{i=1}^{s}(x - \lambda_i)^{m_i}$ has the following structure

$$\Omega_\eta^T = \begin{bmatrix} \Omega_1 \\ \Omega_2 \\ \cdots \\ \Omega_s \end{bmatrix},$$

where

$$\Omega_r = \begin{bmatrix} 1 & \lambda_r & \lambda_r^2 & \dots & \lambda_r^{n-1} \\ 0 & 1 & 2\lambda_r & \dots & (n-1)\lambda_r^{n-2} \\ 0 & 0 & 1 & & \\ 0 & 0 & & \ddots & \vdots \\ 0 & 0 & \dots & & \binom{n-1}{m_r - 1}\lambda_r^{n-m_r} \end{bmatrix}_{m_r \times n}. \tag{31}$$

Clearly, for $1 \le j \le n$ and $1 \le i \le m_r$,

$$[\Omega_r]_{ij} = \begin{cases} \binom{j-1}{i-1}\lambda_r^{j-i} & j \ge i \\ 0 & otherwise \end{cases} \tag{32}$$

We may express the entries in Ω_r via partial differentiation. Indeed if $\mathfrak{D}_r^{(k)}$ is the k^{th} derivative with respect to λ_r, we may denote the quantity on the right hand side of (32) by $(1/i!)\mathfrak{D}_r^{(i)}(\lambda_r^j)$, $j = 0, \dots, n-1$.

Let us now extend the idea of partial differentiation to matrices thereby showing that the q-adic Vandermonde matrix induced by $\psi(x) = \prod_{i=1}^{s} q_i^{m_i}$, has a block structure that is very similar to that of the Wronskian structure exhibited in (31).

Suppose $L_r = L_{q_r}$ of the polynomial $q_r(x) = \alpha_0 + \alpha_1(x) + \ldots + x^{l_r - 1}$. We define the associated partial derivative operator $\mathfrak{D}_r$ acting on an $l_r \times l_r$ matrix by

$$\mathfrak{D}_r(A) = \left[\frac{\partial(a_{ij})}{\partial(-\alpha_0)}\right],$$

i.e., the derivative is taken with respect to the negative of the constant term $(-\alpha_0)$ in $q_r(x)$. For example, $\mathfrak{D}_r(L_r) = N = E_{1,l_r}$ and hence

$$\mathfrak{D}_r(L_r\mathbf{x}) = \mathfrak{D}_r(L_r)\mathbf{x} + L_r(\mathfrak{D}_r\mathbf{x}) = N\mathbf{x} + L_r\mathfrak{D}_r\mathbf{x}.$$

It follows swiftly by induction that

$$\mathfrak{D}^i(L_r\mathbf{x}) = iN_r\mathfrak{D}^{i-1}\mathbf{x} + L_r\mathfrak{D}^i\mathbf{x}. \tag{33}$$

Our essential step is to relate the q-adic coordinate columns to the operator $\mathfrak{D}_r^{(i)}$. In fact, we have the following lemma.

LEMMA 3.1. *If* $[x^j]_{(q_r,m_r)} = \begin{bmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{m_r-1} \end{bmatrix}$, *then* $\mathbf{a}_i = (1/i!)\mathfrak{D}_r^{(i)}(L_r^j\mathbf{e}_1)$ *for* $i = 0, \ldots, m_r - 1$ *and* $j = 0, \ldots, n-1$.

Proof. The proof is by induction on j. The lemma clearly holds for $j = 0$, since $[x^0]_{(q_r,m_r)} = \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{0} \\ \cdots \\ \mathbf{0} \end{bmatrix}_{m_r l_r \times 1}$. Now assume that

$$[x^j]_{(q_r,m_r)} = \begin{bmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{m_r-1} \end{bmatrix}, \text{ and } [x^{j+1}]_{(q_r,m_r)} = \begin{bmatrix} \mathbf{b}_0 \\ \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_{m_r-1} \end{bmatrix},$$

where $\mathbf{a}_i = (1/i!)\mathfrak{D}_r^{(i)}(L_r^j\mathbf{e}_1)$. From Corollary 1.3 with $f = x^j$ we see that $[x^{j+1}]_{(q_r,m_r)} = H_m[x^j]_{(q_r,m_r)}$ which in matrix form becomes

$$\mathbf{b}_0 = L_r\mathbf{a}_0, \ldots, \ \mathbf{b}_i = N\mathbf{a}_{i-1} + L_r\mathbf{a}_i$$

for $i = 1, \ldots, n-1$. Using the induction hypothesis, this gives

$$\mathbf{b}_0 = L_r\mathbf{a}_0 = L_rL_r^j\mathbf{e}_1 = L_r^{j+1}\mathbf{e}_1$$

as well as

$$\mathbf{b}_i = N\left(\frac{1}{(i-1)!}\mathfrak{D}_r^{(i-1)}(L_r^j\mathbf{e}_1)\right) + L_r\left(\frac{1}{i!}\mathfrak{D}_r^i(L_r^j\mathbf{e}_1)\right).$$

Using (33) with $\mathbf{x} = L_r^j \mathbf{e}_1$, we see that (3) reduces to

$$\mathbf{b_i} = \frac{1}{i!}\left(iN\mathfrak{D}_r^{(i-1)}(L_r^j\mathbf{e}_1) + L_r\mathfrak{D}_r^i(L_r^j\mathbf{e}_1)\right) = \frac{1}{i!}\mathfrak{D}_r^{(i)}(L_r^{j+1}\mathbf{e}_1),$$

completing the induction. □

Our final result on the structure of the matrix W is given by the following theorem.

THEOREM 3.2. *The q-adic Vandermonde matrix W induced by $\prod q_i^{m_i}$ has the block form* $W^T = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_s \end{bmatrix}$, *where*

$$W_r = \begin{bmatrix} \mathbf{e}_1 & L_r\mathbf{e}_1 & L_r^2\mathbf{e}_1 & \dots & L_r^{n-1}\mathbf{e}_1 \\ 0 & \mathbf{e}_1 & \mathfrak{D}_r(L_r^2\mathbf{e}_1) & & \mathfrak{D}_r(L_r^{n-1}\mathbf{e}_1) \\ \vdots & 0 & \mathbf{e}_1 & \vdots & \\ & & \dots & & \\ 0 & 0 & \dots & & (1/(m_r-1)!)\mathfrak{D}_r^{(m_r-1)}(L_r^{n-1}\mathbf{e}_1) \end{bmatrix}_{m_r\times n}.$$

In other words, for $i = 0, .., m_r-1$ and $j = 0, .., n-1$, $[W_r]_{ij} = (1/i!)\mathfrak{D}_r^{(i)}(L_r^j\mathbf{e}_1)$.

Proof. We recall from (6) that the columns of W_r are of the form $[x^t]_{(q_r,m_r)}$. The structure of these columns is precisely given by Lemma 3.1 as

$$[x^t]_{(q_r,m_r)} = \begin{bmatrix} L_r^t\mathbf{e}_1 \\ \mathfrak{D}_r(L_r^t\mathbf{e}_1) \\ \vdots \\ (1/(m_r-1)!)\mathfrak{D}_r^{(m_r-1)}(L_r^t\mathbf{e}_1) \end{bmatrix}$$

and hence the proof of the theorem follows. □

REMARK 3.3. It has been pointed out to the authors by an anonymous referee that each of the W_i is a Kalman reachability matrix. Indeed, from equation (14), we have that

$$W_i = \left[\mathbf{e}_1 H_i\mathbf{e}_1 \dots H_i^{n-1}\mathbf{e}_1\right],$$

where $\mathbf{e}_1 = [1, 0, \dots, 0]^T$ and $H_i = H[q_i^{m_i}]$ is a hypercompanion matrix. This structure of this matrix is similar to the one given in Theorem 3.18, page 41 of [8]. The control theoretic aspects of this matrix are still to be explored.

REFERENCES

[1] S. Barnett, *Polynomials and Linear Control Systems*, Marcel Dekker, New York, 1983.
[2] S. Lang, *Algebra*, 3rd edition, Prentice Hall, New York, 1993.
[3] G. Sansigre and M. Alvarez, On Bezoutian Reduction with Vandermonde Matrices, *Linear Algebra and its Applications*, 121:401–408, 1989.
[4] G. Chen and Z. Yang, Bezoutian Representation Via Vandermonde Matrices, *Linear Algebra and its Applications*, 186:37–44, 1993.
[5] R. E. Hartwig, Resultants, Companion Matrices, and Jacobson Chains, *Linear Algebra and its Applications*, 94:1–34, 1987,
[6] R. E. Hartwig, Some Properties of Hypercompanion Matrices, *Industrial Mathematics*, Volume 24, Part 2, 1974.
[7] U. Helmke and P. A. Fuhrman, Bezoutians, *Linear Algebra and its Applications*, 122–124:1039–1097, 1989.
[8] R. E. Kalman, P. L. Falb and M. A. Arbib, *Topics in Mathematical System Theory*, McGraw-Hill, New York, 1969.
[9] V. Mani and R. E. Hartwig, Bezoutian Reduction over Arbitrary Fields, *Linear Algebra and its Applications*, to appear.