

## Symmetric functions and the Vandermonde matrix

Halil Oruç \*, Hakan K. Akmaz

*Department of Mathematics, Dokuz Eylül University*

*Fen Edebiyat Fakültesi, Tinaztepe Kampüsü 35160 Buca İzmir, Turkey*

Received 29 May 2003 ; received in revised form 15 January 2004

This paper is dedicated to fine friend, mentor and esteemed colleague  
George M. Phillips.

---

### Abstract

This work deduces the lower and the upper triangular factors of the inverse of the Vandermonde matrix using symmetric functions and combinatorial identities. The  $L$  and  $U$  matrices are in turn factored as bidiagonal matrices. The elements of the upper triangular matrices in both the Vandermonde matrix and its inverse are obtained recursively. The particular value  $x_i = 1 + q + \dots + q^{i-1}$  in the indeterminates of the Vandermonde matrix is investigated and it leads to  $q$ -binomial and  $q$ -Stirling matrices. It is also shown that  $q$ -Stirling matrices may be obtained from the Pascal matrix.

*AMS classification:* 15A23, 11B65.

*Keywords:* Vandermonde matrix, symmetric functions, triangular and bidiagonal factorization,  $q$ -Stirling numbers.

---

### 1. Introduction

The Vandermonde matrix appears in interpolation problems. Given the values of a function  $f(x)$  at  $n + 1$  distinct points  $x_0, \dots, x_n$ , to find the interpolating polynomial in the form

$$p_n(x) = a_0 + a_1x + \dots + a_nx^n$$

of degree at most  $n$  which assumes  $p_n(x_i) = f(x_i), i = 0, 1, \dots, n$  requires that the linear system

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix} \quad (1.1)$$

has a unique solution. The coefficient matrix denoted by  $V(x_0, \dots, x_n)$  in (1.1) is called the Vandermonde matrix. Traditional methods to solve a linear system of equations

\*Corresponding author E-mail: halil.oruc@deu.edu.tr

$\mathbf{Ax} = \mathbf{b}$  use triangular factorization of the coefficient matrix which is transformed to an easy to solve system  $LU\mathbf{x} = \mathbf{b}$ , where  $L$  is a lower triangular matrix and  $U$  is an upper triangular matrix. Although the system of equations  $V\mathbf{a} = \mathbf{f}$  obtained above to solve the interpolation problem may be transformed to  $LU\mathbf{a} = \mathbf{f}$ , there are several constructive methods. See for example [20, 21]. Instead of using the monomial basis or the Lagrange functions,  $L_0, L_1, \dots, L_n$ , where

$$L_i(x) = \prod_{j \neq i} \left( \frac{x - x_j}{x_i - x_j} \right)$$

the Newton form of the interpolating polynomial is more effective due to the recursive definition. Let  $\lambda_0(x) = 1$  and  $\lambda_i(x) = (x - x_0)(x - x_1) \cdots (x - x_{i-1})$  for  $1 \leq i \leq n$ . Then the interpolating polynomial is

$$p_n(x) = \sum_{i=0}^n f[x_0, x_1, \dots, x_i] \lambda_i(x), \quad (1.2)$$

where the divided differences of  $f(x)$  at  $x_0, x_1, \dots, x_i$  are defined recursively by

$$f[x_0, x_1, \dots, x_i] = \frac{f[x_1, x_2, \dots, x_i] - f[x_0, x_1, \dots, x_{i-1}]}{x_i - x_0}. \quad (1.3)$$

However triangular factorization of the Vandermonde matrix has gained a considerable theoretical interest. The explicit formulas for the triangular factors  $LU$  of  $V$  and their inverses  $U^{-1}L^{-1}$  of  $V^{-1}$  at the integer values are well known [11]. The triangular factorization of Cauchy-Vandermonde matrix and its inverse are given in [6, 14, 15]. The studies [14, 15] investigate bidiagonal factorization of the inverse of the Cauchy-Vandermonde matrix in order to derive a fast algorithm for solving linear systems whose coefficient matrix is of the same type. The symmetric functions have been used in [17] for the explicit formulation of the triangular factors of  $V$ . Then, the bidiagonal products of the triangular factors are obtained to represent  $V$  as a product of bidiagonal matrices and hence prove the total positivity for any positive increasing sequence  $0 < x_0 < x_1 < \dots < x_n$ . Bidiagonal matrices via total positivity are important for curves-surface design purposes.

The Pascal and Stirling matrices arising from the triangular factorization of the Vandermonde matrix  $V$  at the integer values have been worked extensively. The work [4] obtains the explicit formulas for both  $LU$  factors of the Vandermonde matrix and its inverse on the integer nodes. The paper [5] generalizes the results of [4] to the rectangular Vandermonde matrix. Algebraic and combinatorial approach to the Pascal and Stirling matrices have been investigated in [24, 3]. For further investigation via that approach, follow the references there in. In this paper, a generalization is done by taking the  $q$ -integer nodes,  $x_0 = 0, x_i = 1 + q + \dots + q^{i-1} := [i], i = 1, 2, \dots, n$  in the Vandermonde matrix. This is related to the interpolation problem at the geometrically spaced nodes. Thus  $q$ -binomial and  $q$ -Stirling numbers play an important role in decomposing the Vandermonde matrix. Section 2 of this paper outlines some properties of the symmetric

functions and  $q$ -Stirling numbers that will be necessary in this study. In Section 3, the explicit expression for the triangular factors of the inverse of the Vandermonde matrix are derived. In particular the elements of upper triangular matrices  $U$  and  $U^{-1}$  are obtained recursively with the help of symmetric functions in Section 4. We give a different type of recurrence relation for the  $q$ -Stirling numbers leading to connect  $q$ -Stirling and Pascal type matrices. Section 5 yields formulas for the bidiagonal factors of  $L^{-1}$  and  $U^{-1}$ . These formulas cannot be obtained directly by inverting the bidiagonal products of  $L$  and  $U$  from [17] since the inverse of a bidiagonal matrix in general is not bidiagonal.

## 2. Preliminaries

In this study, the main tool in decomposing the Vandermonde matrix and its inverse is the use of the symmetric functions. So, it is helpful to give the definitions of the symmetric functions and some of their properties. The books [13, ch.1] and [23, ch.7] are the guides for this purpose.

**Definition 2.1.** Let  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ , ( $n \geq 1$ ). The  $r^{th}$  ( $1 \leq r \leq n$ ) elementary symmetric function denoted by  $\sigma_r(x_1, \dots, x_n)$ , is the sum of all products of  $r$  distinct variables chosen from  $n$  variables. That is,

$$\sigma_r(\mathbf{x}) = \sum_{1 \leq i_1 < i_2 < \dots < i_r \leq n} x_{i_1} \dots x_{i_r}.$$

**Definition 2.2.** The  $r^{th}$  complete symmetric function  $\tau_r(\mathbf{x})$ , in variables  $x_1, \dots, x_n$  is defined by

$$\tau_r(\mathbf{x}) = \sum_{\lambda_1 + \dots + \lambda_n = r} x_1^{\lambda_1} x_2^{\lambda_2} \dots x_n^{\lambda_n},$$

where the summation extends over all choices of  $\lambda_1, \dots, \lambda_n \in \{0, 1, \dots, r\}$  satisfying the condition.

It is convenient to define  $\sigma_r(\mathbf{x}) = \tau_r(\mathbf{x}) = 0$  for  $r < 0$  and  $\sigma_r(\mathbf{x}) = \tau_r(\mathbf{x}) = 1$  when  $r = 0$ . There are other types of symmetric functions see for example [23, ch.7].

The generating function  $G(x)$  for the elementary symmetric function is

$$G(x) = \prod_{i=1}^n (1 - x_i x) = \sum_{r=0}^n (-1)^r \sigma_r(\mathbf{x}) x^r \tag{2.1}$$

and

$$\frac{1}{G(x)} = \sum_{r=0}^{\infty} \tau_r(\mathbf{x}) x^r \tag{2.2}$$

is the generating function for the complete symmetric function. The function  $G(x)$  may be interpreted as the generating function for the subsets of the set  $\{x_1, \dots, x_n\}$  and  $\sigma_r(\mathbf{x})$  is all  $r$ -element subsets. Similarly,  $1/G(x)$  may be viewed as the generating function of the multiset  $\{\infty.x_1, \dots, \infty.x_n\}$  meaning an element, say  $x_1$ , is repeated infinitely many times. Hence the  $r^{th}$  complete symmetric function  $\tau_r(\mathbf{x})$  represents all  $r$ -element subsets

of the multiset. The recurrence relations of the symmetric functions can be obtained respectively from (2.1) and (2.2) as

$$\sigma_r(x_1, \dots, x_n) = \sigma_r(x_1, \dots, x_{n-1}) + x_n \sigma_{r-1}(x_1, \dots, x_{n-1}) \quad (2.3)$$

and

$$\tau_r(x_1, \dots, x_n) = \tau_r(x_1, \dots, x_{n-1}) + x_n \tau_{r-1}(x_1, \dots, x_{n-1}) \quad (2.4)$$

for  $r \geq 1$  and  $n \geq 2$  number of variables. It follows from  $G(x) \frac{1}{G(x)} = 1$  that

$$\sum_{r=0}^n (-1)^r \sigma_r(\mathbf{x}) \tau_{n-r}(\mathbf{x}) = 0, \quad n \geq 1. \quad (2.5)$$

Note that many combinatorial identities and counting sequences such as binomial coefficients,  $q$ -binomial coefficients, Stirling numbers, their  $q$ -analogues ( $q$ -Stirling numbers) and their recurrence relations, inverse relations can be deduced from the symmetric functions by taking particular values in the indeterminates  $x_1, \dots, x_n$  (cf. [13, 12]). Explicitly, take  $x_i = 1, \forall i$  in order to derive the binomial coefficients  $\sigma_r(1, \dots, 1) = \binom{n}{r}$  and with repetition (allowed)  $\tau_r(1, \dots, 1) = \binom{n+r-1}{r}$ .

The Stirling number of the first kind,  $s(n, k)$  counts the number of ways to arrange  $n$  objects into  $k$  cycles and the Stirling number of the second kind  $S(n, k)$  counts the number of ways to arrange  $n$  objects into  $k$  nonempty subsets. If we set  $x_i = i$  then  $\sigma_r(1, 2, \dots, n) = s(n+1, n+1-r)$  and  $\tau_r(1, 2, \dots, n) = S(n+r, n)$ . The recurrence relations and inverse relations of binomial coefficients and Stirling numbers follow directly from (2.3), (2.4) and (2.5), see [12].

The generating function of restricted partitions may also be obtained from the symmetric functions.

$$\sigma_r(1, q, \dots, q^{n-1}) = q^{r(r-1)/2} \begin{bmatrix} n \\ r \end{bmatrix} \quad \text{and} \quad \tau_r(1, q, \dots, q^{n-1}) = \begin{bmatrix} n+r-1 \\ r \end{bmatrix},$$

where

$$\begin{bmatrix} n \\ r \end{bmatrix} = \begin{cases} \frac{[n][n-1] \cdots [n-r+1]}{[r][r-1] \cdots [1]}, & n \geq r \geq 0, \\ 0, & \text{otherwise,} \end{cases}$$

An extensive treatment of partitions is done in the book [1].

When  $\sigma_r(1, q, \dots, q^{n-1})$  is put in (2.3) we deduce the recurrence relation for restricted partitions

$$\begin{bmatrix} n \\ r \end{bmatrix} = \begin{bmatrix} n-1 \\ r \end{bmatrix} + q^{n-r} \begin{bmatrix} n-1 \\ r-1 \end{bmatrix} \quad (2.6)$$

We recall (see [10]) the  $q$ -analogue of the Stirling numbers of the first kind  $s_q(n, m)$  and of the second kind  $S_q(n, m)$  given by the generating functions

$$\lambda_n(x) = \sum_{m=0}^n s_q(n, m) x^m \quad (2.7)$$

and

$$x^n = \sum_{m=0}^n S_q(n, m) \lambda_m(x), \tag{2.8}$$

where  $x_i = [i]$  for all  $i$  in the polynomial  $\lambda_n(x)$ . When  $q$  is replaced by 1, the  $q$ -Stirling numbers reduce to the classical Stirling numbers. It may be proved by induction on  $n$  that  $s_q(n, m)$  is a polynomial in  $q$  of degree  $\frac{1}{2}(n-1)(n-2) - \frac{1}{2}(m-1)(m-2)$  and  $S_q(n, m)$  is a polynomial in  $q$  of degree  $(m-1)(n-m)$ .

On comparing the coefficients of  $x^m$  in the equation  $\lambda_{n+1}(x) = (x - [n])\lambda_n(x)$  we obtain the recurrence relation

$$s_q(n+1, m) = s_q(n, m-1) - [n]s_q(n, m), \tag{2.9}$$

for  $n \geq 0$  and  $m \geq 1$ . Similarly, from (2.8) we deduce the recurrence relation

$$S_q(n+1, m) = S_q(n, m-1) + [m]S_q(n, m). \tag{2.10}$$

The  $S_q(n, m)$  can be written explicitly in the form

$$S_q(n, m) = \frac{1}{q^{m(m-1)/2}[m]!} \sum_{r=0}^n (-1)^r q^{r(r-1)/2} \begin{bmatrix} n \\ r \end{bmatrix} [n-r]^m$$

and is obtained in [9] in connection with the  $q$ -Bernstein polynomials of monomials. The notation  $[m]!$  means the product  $[m][m-1] \cdots [1]$  and  $[0]! = 1$ .

Substitute  $x_i = [i], i = 1, 2, \dots, n$ . Inductively, one may see that

$$(1-x)(1-[2]x) \cdots (1-[n]x) = \sum_{r=0}^n s_q(n+1, n+1-r)x^r, \tag{2.11}$$

giving  $(-1)^r \sigma_r([1], [2], \dots, [n]) = s_q(n+1, n+1-r)$ , also

$$\frac{1}{(1-x)(1-[2]x) \cdots (1-[n]x)} = \sum_{r=0}^{\infty} S_q(n+r, n)x^r, \tag{2.12}$$

giving  $\tau_r([1], [2], \dots, [n]) = S_q(n+r, n)$ . Now, the recurrence relations for the  $q$ -Stirling numbers, (2.9) and (2.10) may be verified from (2.3) and (2.4) respectively.

### 3. Triangular factors

In this section, we deduce the triangular factors of  $V^{-1}$ . Not only the computation of  $U$  and of  $U^{-1}$  but also the computation of their bidiagonal factors use the recurrence relations of the symmetric functions. The formulas for the elements of the upper triangular matrix  $U$  and for  $U^{-1}$  involve the symmetric functions. Let us introduce the notation

$$\lambda_i^j(x) = (x-x_i)(x-x_{i+1}) \cdots (x-x_j)$$

for nonnegative integers  $j \geq i$  which may be obtained  $G(x)$ . Clearly, the coefficient of  $x^r$  in the expansion of  $\lambda_i^j(x)$  may be expressed in terms of elementary symmetric functions

and equals  $(-1)^r \sigma_{i+j-r}(x_i, \dots, x_j)$ . We write  $(\lambda_i^j)'(x)$  for the derivative of  $\lambda_i^j(x)$  with respect to  $x$ . Then for an integer  $k$ ,

$$(\lambda_i^j)'(x_k) = \prod_{\substack{t=i \\ t \neq k}}^j (x_k - x_t), \quad i \leq k \leq j.$$

The following formulas for the triangular decomposition  $LDU$  of the Vandermonde matrix  $V$  may simply be obtained using [17, Theorem 2.1] from its  $LU$  factors.

**Theorem 3.1.** *Let  $V = (x_i^j)_{i,j=0}^n$  be a Vandermonde matrix such that  $V = LDU$ , where  $L$  is a lower triangular matrix with units on its main diagonal,  $D$  is a diagonal matrix and  $U$  is an upper triangular matrix with units on its main diagonal. Then the elements of  $L$ ,  $D$  and  $U$  satisfy, respectively,*

$$L_{i,j} = \frac{\lambda_0^{j-1}(x_i)}{\lambda_0^{j-1}(x_j)}, \quad 0 \leq j \leq i \leq n, \quad (3.1)$$

$$D_{i,j} = \lambda_0^{i-1}(x_i), \quad i = j, \quad (3.2)$$

$$U_{i,j} = \tau_{j-i}(x_0, \dots, x_i), \quad 0 \leq i \leq j \leq n, \quad (3.3)$$

where  $\lambda_0^{-1}(x)$  denotes 1.

Recently, Phillips [21] has scaled the elements of  $L, U$  in [17, Theorem 2.1] to give  $L^*$  and  $U^*$  whose elements are

$$L_{i,j}^* = \prod_{t=0}^{j-1} (x_i - x_{j-t-1}), \quad 0 \leq j \leq i \leq n, \quad U_{i,j}^* = \tau_{j-i}(x_0, \dots, x_i), \quad 0 \leq i \leq j \leq n. \quad (3.4)$$

Before deriving the formulas of the inverses of the matrices  $L, D$  and  $U$ , we give the following result which is obtained by using (2.3) repeatedly. This will be needed in the proof of a theorem for the triangular factorization of  $V^{-1}$ .

**Lemma 3.1.** *The  $i^{\text{th}}$  iterate of the recurrence relation (2.3) satisfies*

$$\sigma_r(x_1, \dots, x_n) = \sum_{j=0}^i \sigma_{r-j}(x_1, \dots, x_{n-i}) \sigma_j(x_{n-i+1}, \dots, x_n). \quad (3.5)$$

*Proof.* When  $i = 1$  we write  $\sigma_r(x_n)$  in the place of  $\sigma_r(x_n, \dots, x_n)$  in (3.5). Applying (3.5) to  $\sigma_{r-j}(x_1, \dots, x_{n-i})$  and combining the summations appropriately verifies the Lemma.  $\square$

**Theorem 3.2.** *Let  $V^{-1} = U^{-1}D^{-1}L^{-1}$  be the inverse of the Vandermonde matrix.*

Then the  $(i, j)$  element of  $U^{-1}, D^{-1}$  and  $L^{-1}$  satisfy respectively,

$$(U^{-1})_{i,j} = (-1)^{i+j} \sigma_{j-i}(x_0, \dots, x_{j-1}), \quad 0 \leq i \leq j \leq n, \quad (3.6)$$

$$(D^{-1})_{i,j} = \begin{cases} \frac{1}{\lambda_0^{i-1}(x_i)}, & i = j, \\ 0, & \text{otherwise,} \end{cases} \quad (3.7)$$

$$(L^{-1})_{i,j} = \begin{cases} -\frac{(\lambda_0^{i-1})'(x_i)}{(\lambda_0^{i-1})'(x_j)}, & i > j, \\ 1, & i = j, \\ 0, & \text{otherwise.} \end{cases} \quad (3.8)$$

*Proof.* In order to verify (3.6) we need to show that

$$\sum_{k=0}^n U_{i,k}(U^{-1})_{k,j} = \sum_{k=i}^j U_{i,k}(U^{-1})_{k,j} = \delta_{i,j},$$

since  $U$  and  $U^{-1}$  are upper triangular matrices and  $\delta_{i,j}$  denotes the Kronecker delta function. Using (3.3) and (3.6), the last equation equals

$$\sum_{k=i}^j (-1)^{k+j} \tau_{k-i}(x_0, \dots, x_i) \sigma_{j-k}(x_0, \dots, x_{j-1}) = w_{i,j}. \quad (3.9)$$

From the identity (3.5) of the above Lemma 3.1, we may write

$$\sigma_{j-k}(x_0, \dots, x_{j-1}) = \sum_{t=0}^{j-i-1} \sigma_{j-k-t}(x_0, \dots, x_i) \sigma_t(x_{i+1}, \dots, x_{j-1}).$$

Substituting the last formula in (3.9) gives

$$w_{i,j} = \sum_{k=0}^{j-i} (-1)^{i+j+k} \tau_k(x_0, \dots, x_i) \sum_{t=0}^{j-i-1} \sigma_{j-i-k-t}(x_0, \dots, x_i) \sigma_t(x_{i+1}, \dots, x_{j-1}).$$

Rearranging the latter summations we obtain

$$w_{i,j} = \sum_{t=0}^{j-i-1} (-1)^{i+j} \sigma_t(x_{i+1}, \dots, x_{j-1}) \sum_{k=0}^{j-i-t} (-1)^k \tau_k(x_0, \dots, x_i) \sigma_{j-i-t-k}(x_0, \dots, x_i).$$

It follows from (2.5) that the second sum in the last equation vanishes. Thus  $w_{i,j} = 0$  for  $i \neq j$ . It is obvious that (3.6) is equal to 1 for  $i = j$  and hence  $w_{i,j} = \delta_{i,j}$ .

For the proof of (3.8) we will show that  $LL^{-1} = I$ . Since both matrices are lower triangular and from the fact that  $(x_j - x_k)(\lambda_0^{k-1})'(x_j) = (\lambda_0^k)'(x_j)$ , we have

$$\sum_{k=0}^n L_{i,k}(L^{-1})_{k,j} = \sum_{k=j}^i \frac{\lambda_0^{k-1}(x_i)}{(\lambda_0^k)'(x_j)} = w_{i,j}.$$

Multiplying and dividing the last equation by  $\lambda_{k+1}^i(x_j)$  we get

$$\omega_{i,j} = \frac{1}{(\lambda_0^i)'(x_j)} \sum_{k=j}^i \lambda_0^{k-1}(x_i) \lambda_{k+1}^i(x_j). \quad (3.10)$$

Next we show that the sum in (3.10) vanishes for  $i \neq j$ . For this purpose we first prove the following identity by induction on  $m$ .

$$\sum_{k=1}^m \lambda_0^{k+j-2}(x_i) \lambda_{k+j}^i(x_j) = \frac{\lambda_0^{m+j-1}(x_i)}{(x_i - x_j)} \lambda_{m+j}^i(x_j). \quad (3.11)$$

In order to verify the above identity for  $m+1$  we write

$$\sum_{k=1}^{m+1} \lambda_0^{k+j-2}(x_i) \lambda_{k+j}^i(x_j) = \lambda_0^{m+j-1}(x_i) \lambda_{m+j+1}^i(x_j) + \sum_{k=1}^m \lambda_0^{k+j-2}(x_i) \lambda_{k+j}^i(x_j).$$

Using the induction hypothesis (3.11) we obtain

$$\begin{aligned} \sum_{k=1}^{m+1} \lambda_0^{k+j-2}(x_i) \lambda_{k+j}^i(x_j) &= (x_i - x_j) \frac{\lambda_0^{m+j-1}(x_i)}{(x_i - x_j)} \lambda_{m+j+1}^i(x_j) \\ &+ (x_j - x_{m+j}) \frac{\lambda_0^{m+j-1}(x_i)}{(x_i - x_j)} \lambda_{m+j+1}^i(x_j) \\ &= \frac{\lambda_0^{m+j}(x_i)}{(x_i - x_j)} \lambda_{m+j+1}^i(x_j) \end{aligned}$$

and this completes the induction. Substituting  $m = i - j + 1$  in (3.11) causes the sum to vanish since  $\lambda_0^i(x_i) = 0$  on the right side. Thus  $\omega_{i,j} = 0$  for  $i \neq j$ . It is easy to check  $\omega_{i,j} = 1$  for  $i = j$ . This completes the proof.  $\square$

Since deriving the elements of  $U$  and  $U^{-1}$  involves the complete symmetric functions and elementary symmetric functions respectively, and the fact that they satisfy a three term recurrence relation, we deduce the following consequence from the above factorization.

**Theorem 3.3.** For  $1 \leq i \leq j \leq n$ , the elements of  $U$  and  $U^{-1}$  can be obtained recursively in the form

$$U_{i,j} = U_{i-1,j-1} + x_i U_{i,j-1}, \quad (3.12)$$

$$(U^{-1})_{i,j} = (U^{-1})_{i-1,j-1} - x_{j-1} (U^{-1})_{i,j-1}. \quad (3.13)$$

*Proof.* The proof of (3.12) follows from (3.3) using (2.4) and that of (3.13) follows from (3.6) using (2.3).  $\square$

#### 4. Pascal and Stirling matrices

In this section we are concerned with the particular values  $x_i = [i], i = 0, 1, \dots, n$  in the Vandermonde matrix  $V(x_0, x_1, \dots, x_n)$  and its triangular factors. We will see how  $q$ -binomial coefficients and  $q$ -Stirling numbers are related from the Vandermonde matrix and the interpolation problem at geometrically spaced nodes.



**Theorem 4.1.** *The elements of the matrices  $L, D, U$  and  $L^{-1}, D^{-1}, U^{-1}$  of  $V$  and of  $V^{-1}$  on  $q$ -integer nodes  $x_i = [i], i = 0, 1, \dots, n$  are respectively*

$$L_{i,j} = \begin{bmatrix} i \\ j \end{bmatrix}, \tag{4.1}$$

$$D_{i,i} = [i]! \tag{4.2}$$

$$U_{i,j} = S_q(j, i), \tag{4.3}$$

$$(L^{-1})_{i,j} = (-1)^{i-j} q^{(i-j-1)(i-j)/2} \begin{bmatrix} i \\ j \end{bmatrix}, \tag{4.4}$$

$$(D^{-1})_{i,i} = \frac{1}{[i]!} \tag{4.5}$$

$$(U^{-1})_{i,j} = s_q(j, i). \tag{4.6}$$

*Proof.* Substituting  $x_i = [i]$  in (3.1) gives

$$L_{i,j} = \prod_{t=0}^{j-1} \frac{[i] - [t]}{[j] - [t]}$$

and using the relation

$$[i] - [t] = \begin{cases} q^t [i - t], & t \leq i, \\ -q^i [t - i], & t > i, \end{cases} \tag{4.7}$$

in the latter equation yields

$$L_{i,j} = \frac{[i][i-1] \cdots [i-(j-1)]}{[j][j-1] \cdots [1]} = \begin{bmatrix} i \\ j \end{bmatrix}.$$

To verify (4.3) let us take  $f(x) = x^j$  in (1.2) at the  $q$ -integer points  $[0], [1], \dots, [j]$ . Since the interpolation operator reproduces polynomials we have

$$x^j = \sum_{i=0}^j f[[0], [1], \dots, [i]] \lambda_i(x).$$

The relation

$$\tau_{j-i}(x_0, x_1, \dots, x_i) = f[x_0, x_1, \dots, x_i], \quad f(x) = x^j \tag{4.8}$$

between complete symmetric function and divided differences is deduced in [17]. (This identity is well known, see the book [22, p.47].) Thus we see that

$$\tau_{j-i}([0], [1], \dots, [i]) = f[[0], [1], \dots, [i]].$$

Comparing this and (3.3), we obtain the  $q$ -Stirling matrix whose elements are  $q$ -Stirling numbers of the second kind  $U_{i,j} = S_q(j, i)$ . For the proof of (4.4), multiply and divide it by  $([j] - [i])$  and use (4.7). In the next proof we consider the generating function of the  $q$ -Stirling numbers (2.7) and (2.8). In a way they are inverse relations as in (2.5) and the following is deduced

$$\sum_{k=0}^n S_q(i, k) s_q(k, j) = \sum_{k=0}^n s_q(i, k) S_q(k, j) = \delta_{i,j}, \quad 0 \leq i, j \leq n.$$

The relations (4.2) and (4.5) are obvious and this completes the proof.  $\square$

On taking (1.3) and (4.8) into account we deduce the complete symmetric functions in the form

$$\tau_{j-i}(x_0, x_1, \dots, x_i) = \frac{\tau_{j-i+1}(x_1, x_2, \dots, x_i) - \tau_{j-i+1}(x_0, x_1, \dots, x_{i-1})}{x_i - x_0}. \quad (4.9)$$

For further investigation on interpolation at geometrically spaced nodes see [21, 20].

Since the  $q$ -binomial coefficients and the  $q$ -Stirling numbers satisfy a three term recurrence relation, the above elements of the factors of  $V$  may also be evaluated recursively.

**Corollary 4.1.** *The elements of  $L, L^{-1}, U, U^{-1}$  on  $q$ -integer nodes satisfy the following recurrence relations*

$$\begin{aligned} L_{i,j} &= q^{i-j} L_{i-1,j-1} + L_{i-1,j}, \\ (L^{-1})_{i,j} &= q^{i-j} (L^{-1})_{i-1,j-1} - q^{i-j-1} (L^{-1})_{i-1,j}, \\ U_{i,j} &= U_{i-1,j-1} + [i] U_{i,j-1}, \\ (U^{-1})_{i,j} &= (U^{-1})_{i-1,j-1} - [j-1] (U^{-1})_{i,j-1}. \end{aligned}$$

*Proof.* This result is a consequence of the recurrence relations (2.6), (2.9), (2.10) and the formulas in the Theorem 4.1.  $\square$

The Neumann series of  $q$ -Stirling matrices of the first and the second kind are used in [18] in order to prove certain convergence and iteration properties of the  $q$ -Bernstein polynomials.

The above theorem is a generalization of the result of [4], since  $q = 1$  reduces  $q$ -integers to integers and  $q$ -Stirling numbers to Stirling numbers. Gould in [10] defines the  $q$ -Stirling numbers of the first kind, denoted by  $S_1(n, m)$ , as the sum of  $\binom{n}{m}$  possible products of  $m$  distinct factors chosen from the set  $\{[1], [2], \dots, [n]\}$ . This definition should not be confused with  $S_q(n, m)$  at the value  $q = 1$ . Thus,  $S_1(n, m)$  corresponds to the coefficient of  $x^m$  in the expansion of  $(1+x)(1+[2]x) \cdots (1+[n]x)$  and so  $S_1(n, m) = \sigma_m([1], [2], \dots, [n])$ . It is interesting that the matrix whose elements are  $S_1(n, m)$  appeared in [16, 19] as the transition matrix between the Bernstein-Bézier basis and  $q$ -Bernstein basis which are used for geometric design purposes. In [10] the  $q$ -Stirling numbers of the second kind denoted by  $S_2(n, m)$ , is defined by the sum of  $\binom{n+m-1}{m}$  possible products with repetition of at most  $m$  factors chosen from the set  $\{[1], [2], \dots, [n]\}$ . Comparing this definition and the complete symmetric functions we see that  $S_2(n, m) = \tau_m([1], [2], \dots, [n])$ .

Before giving the next theorem we will introduce the Pascal matrix which has been studied quite extensively, see for example [24, 3].

**Definition 4.3.** *The  $n \times n$  Pascal matrix  $P_n$  is defined by*

$$(P_n)_{i,j} = \begin{cases} \binom{i-1}{j-1}, & i \geq j, \\ 0, & \text{otherwise,} \end{cases}$$

where  $1 \leq i, j \leq n$ .

It can be easily derived from (4.4) by taking  $q = 1$  that the inverse of Pascal matrix has the form

$$(P_n^{-1})_{i,j} = \begin{cases} (-1)^{i-j} \binom{i-1}{j-1}, & i \geq j, \\ 0, & \text{otherwise.} \end{cases}$$

**Definition 4.4.** Let  $s_q^n$  and  $S_q^n$  denote respectively the  $n \times n$   $q$ -Stirling matrices of the first and second kind such that

$$(s_q^n)_{i,j} = \begin{cases} s_q(i, j), & i \geq j, \\ 0, & \text{otherwise,} \end{cases} \quad (4.10)$$

and

$$(S_q^n)_{i,j} = \begin{cases} S_q(i, j), & i \geq j, \\ 0, & \text{otherwise.} \end{cases} \quad (4.11)$$

We see from Theorem 4.1 that  $S_q^{n+1} = U^T$  and  $s_q^{n+1} = (U^{-1})^T$  and  $s_1^n, S_1^n$  are the Stirling matrices with  $q$  replaced by 1. Before deriving  $s_q^n$  and  $S_q^n$  from the Pascal matrix we give the following lemma.

**Lemma 4.2.**  $q$ -Stirling numbers of the first and second kind satisfy the following identities

$$s_q(n+1, m+1) = \sum_{k=m}^n (-1)^{k-m} q^{n-k} \binom{k}{m} s_q(n, k), \quad (4.12)$$

$$S_q(n+1, m+1) = \sum_{k=m}^n q^{k-m} \binom{n}{k} S_q(k, m), \quad (4.13)$$

for  $n \geq 1$  and  $0 \leq m \leq n$ .

*Proof.* We only give the proof of (4.12) since (4.13) may be seen in [16] or be proved analogously. The proof is by induction on  $n$ . For  $n = 1$ , (4.12) gives  $s_q(2, 2) = s_q(1, 1) = 1$ . Now assume

$$s_q(n, m+1) = \sum_{k=m}^{n-1} (-1)^{k-m} q^{n-k-1} \binom{k}{m} s_q(n-1, k)$$

is true for any  $n \geq 2$  and all  $m \leq n$ . From (2.9) we obtain

$$\begin{aligned} s_q(n+1, m+1) &= \sum_{k=m-1}^{n-1} (-1)^{k-m+1} q^{n-k-1} \binom{k}{m-1} s_q(n-1, k) \\ &\quad - [n] \sum_{k=m}^{n-1} (-1)^{k-m} q^{n-k-1} \binom{k}{m} s_q(n-1, k). \end{aligned}$$

Write  $[n] = 1 + q[n-1]$  in the last equation, combine the first and second summation using the Pascal identity and then shift the index of summations, and finally use (2.9) to obtain

$$\begin{aligned} s_q(n+1, m+1) &= \binom{n}{m} (-1)^{n-m} + \sum_{k=m}^{n-1} (-1)^{k-m} q^{n-k} \binom{k}{m} s_q(n, k) \\ &= \sum_{k=m}^n (-1)^{k-m} q^{n-k} \binom{k}{m} s_q(n, k), \end{aligned}$$

which completes the induction.  $\square$

Notice that the above lemma evaluates the  $q$ -Stirling numbers recursively in a different way given in Section 2. The relation (4.12) obtains  $s_q(n+1, m+1)$  from the previous row using all  $s_q(n, m), \dots, s_q(n, n)$ . However (4.13) may be considered as the column-wise version of (4.12).

**Theorem 4.2.** *Let  $\oplus$  denote the direct sum of matrices. For the  $n \times n$  Pascal matrix we have*

$$s_q^n = ([1] \oplus \tilde{s}_q^{n-1}) P_n^{-1}, \quad (4.14)$$

$$S_q^n = P_n([1] \oplus \tilde{S}_q^{n-1}), \quad (4.15)$$

where

$$(\tilde{s}_q)_{i,j}^{n-1} = \begin{cases} q^{i-j} s_q(i, j), & i \geq j, \\ 0, & \text{otherwise,} \end{cases}$$

$$(\tilde{S}_q)_{i,j}^{n-1} = \begin{cases} q^{i-j} S_q(i, j), & i \geq j, \\ 0, & \text{otherwise,} \end{cases}$$

and  $[1]$  (here only) is the  $n \times n$  matrix such that the first entry is 1 and the other entries are 0.

*Proof.* Noting that the  $(i, j)$  element of the matrix  $([1] \oplus \tilde{s}_q^{n-1})$  is  $q^{i-j} \tilde{s}_q^{n-1}(i-1, j-1)$  we write the matrix product

$$(([1] \oplus \tilde{s}_q^{n-1}) P_n^{-1})_{i,j} = \sum_{k=j}^i (-1)^{k-j} q^{i-k} \binom{k-1}{j-1} s_q(i-1, k-1).$$

It follows from (4.12) that the sum on the right is equal to  $(s_q^n)_{i,j}$ .

To prove (4.15), we write the product

$$(P_n([1] \oplus \tilde{S}_q^{n-1}))_{i,j} = \sum_{k=j}^i q^{k-j} \binom{i-1}{k-1} S_q(k-1, j-1).$$

From (4.13), we see that the latter equation is identical to  $(S_q^n)_{i,j}$ .  $\square$

This theorem is more general since  $q = 1$  reduces to the result in [24, 3].

## 5. Bidiagonal factors

Bidiagonal (1-banded) factorization of a matrix is directly related to its total positivity. A matrix is called totally positive if all its minors are nonnegative (all square submatrices have a nonnegative determinant). It is shown in [8] that a finite matrix is totally positive if and only if it is a product of bidiagonal matrices with nonnegative elements. Using this result, it is proved in [17] by finding bidiagonal products of the Vandermonde matrix that the matrix  $V(x_0, x_1, \dots, x_n)$  is totally positive whenever  $0 < x_0 < x_1 < \dots < x_n$ . Next

we give bidiagonal products of  $U^{-1}$  and  $L^{-1}$  and hence expressing  $V^{-1}$  as a product of bidiagonal matrices. It is also worth noting that bidiagonal products of  $V^{-1}$  are closely related to the algorithm in [2] (Björck-Pereyra algorithm) in solving the linear system  $V\mathbf{a} = \mathbf{f}$  efficiently. We note that the matrices denoted by  $U^{(k)}$ 's and  $L^{(n)}$ 's are not the same matrices as those in Section 3.

**Theorem 5.1.** *For integers  $n \geq 1$  and distinct real numbers  $x_0, \dots, x_n$  the inverse of the  $(n + 1) \times (n + 1)$  Vandermonde matrix can be factorized into  $2n$  bidiagonal matrices such that*

$$V^{-1} = U^{(n)}U^{(n-1)} \dots U^{(1)}D^{-1}L^{(1)}L^{(2)} \dots L^{(n)}, \quad (5.1)$$

where, for  $1 \leq k \leq n$  the elements of  $U^{(k)}$  and  $L^{(k)}$  are

$$U_{i,j}^{(k)} = \begin{cases} 1, & i = j, \\ -x_{n-k}, & i = j - 1, i \geq n - k, \\ 0, & \text{otherwise,} \end{cases} \quad (5.2)$$

$$L_{i,j}^{(k)} = \begin{cases} 1, & i = j, \\ -\prod_{t=i-n+k}^{i-1} \frac{x_i - x_t}{x_{i-1} - x_{t-1}}, & i = j + 1, i \geq n - k + 1, \\ 0, & \text{else,} \end{cases} \quad (5.3)$$

respectively and  $D^{-1}$  is the same as (3.7).

*Proof.* We use induction on  $n$  and apply a similar notation and block matrices technique as in [17]. When  $n = 1$ , it is easily seen that  $U^{-1}$  and  $L^{-1}$  are  $2 \times 2$  bidiagonal matrices giving  $U^{-1}D^{-1}L^{-1} = V^{-1}$ . We give the rest of the proof in two parts, the factorization of  $U^{-1}$  and  $L^{-1}$ .

First we will show by induction on  $k$  that, for  $1 \leq k \leq n$

$$U^{(k)} \dots U^{(1)} = \begin{bmatrix} I_{n-k} & \mathbf{0} \\ \mathbf{0} & \tilde{U}^{(k)} \end{bmatrix}, \quad (5.4)$$

where each  $\mathbf{0}$  denotes the appropriate zero matrix,  $I_{n-k}$  denotes the  $(n - k) \times (n - k)$  identity matrix,  $\tilde{U}^{(k)}$  is a  $(k + 1) \times (k + 1)$  upper triangular matrix whose elements are

$$\tilde{U}_{i,j}^{(k)} = (-1)^{i+j} \sigma_{j-i}(x_{n-k}, \dots, x_{n-k+j-1}), \quad 0 \leq i \leq j \leq k. \quad (5.5)$$

For  $k = 1$ , we see from (5.2), (5.4) and (5.5) that

$$\begin{bmatrix} I_{n-1} & \mathbf{0} \\ \mathbf{0} & \tilde{U}^{(1)} \end{bmatrix} = U^{(1)}.$$

Assume that (5.4) is true for some  $k \geq 1$ . It is necessary to verify the following identity

$$\begin{bmatrix} I_{n-k-1} & \mathbf{0} \\ \mathbf{0} & \tilde{U}^{(k+1)} \end{bmatrix} = U^{(k+1)} \begin{bmatrix} I_{n-k} & \mathbf{0} \\ \mathbf{0} & \tilde{U}^{(k)} \end{bmatrix}.$$

Let us represent  $U^{(k+1)}$  in block form as

$$U^{(k+1)} = \begin{bmatrix} I_{n-k-1} & \mathbf{0} \\ \mathbf{0} & C^{(k+1)} \end{bmatrix},$$

where  $C^{(k+1)}$  is the  $(k+2) \times (k+2)$  upper bidiagonal matrix defined by

$$C_{i,j}^{(k+1)} = \begin{cases} 1, & i = j, \\ -x_{n-k-1}, & i = j - 1, 0 \leq i \leq k + 1, \\ 0, & \text{otherwise.} \end{cases} \quad (5.6)$$

We also rewrite  $\tilde{U}^{(k)}$  by adding a column and a row to give

$$\hat{U}^{(k)} = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \tilde{U}^{(k)} \end{bmatrix},$$

where  $\mathbf{0}$  is an appropriate zero column vector and

$$\hat{U}_{i,j}^{(k)} = \begin{cases} 1, & i = j, \\ (-1)^{i+j} \sigma_{j-i}(x_{n-k}, \dots, x_{n-k+j-2}), & 1 \leq i \leq j \leq k + 1, \\ 0, & \text{otherwise.} \end{cases} \quad (5.7)$$

Thus

$$\begin{bmatrix} I_{n-k-1} & \mathbf{0} \\ \mathbf{0} & \tilde{U}^{(k+1)} \end{bmatrix} = \begin{bmatrix} I_{n-k-1} & \mathbf{0} \\ \mathbf{0} & C^{(k+1)} \end{bmatrix} \begin{bmatrix} I_{n-k-1} & \mathbf{0} \\ \mathbf{0} & \hat{U}^{(k)} \end{bmatrix},$$

which gives  $\tilde{U}^{(k+1)} = C^{(k+1)} \hat{U}^{(k)}$ . The  $(i, j)$  element of  $C^{(k+1)} \hat{U}^{(k)}$  is, say,

$$N_{i,j} = \sum_{m=0}^{k+1} C_{i,m}^{(k+1)} \hat{U}_{m,j}^{(k)}, \quad 0 \leq i \leq j \leq k + 1.$$

Since  $C^{(k+1)}$  is upper bidiagonal, its only non-zero entries are  $C_{i,i}^{(k+1)}$  and  $C_{i,i+1}^{(k+1)}$ . Thus

$$N_{i,j} = C_{i,i}^{(k+1)} \hat{U}_{i,j}^{(k)} + C_{i,i+1}^{(k+1)} \hat{U}_{i+1,j}^{(k)}.$$

Using (5.6), (5.7), and the recurrence relation (2.3) we obtain

$$\begin{aligned} N_{i,j} &= (-1)^{i+j} (\sigma_{j-i}(x_{n-k}, \dots, x_{n-k+j-2}) + x_{n-k-1} \sigma_{j-i-1}(x_{n-k}, \dots, x_{n-k+j-2})) \\ &= (-1)^{i+j} \sigma_{j-i}(x_{n-k-1}, \dots, x_{n-k+j-2}) = \tilde{U}_{i,j}^{(k+1)}. \end{aligned}$$

Thus when  $k = n$ , we have from (5.4) and (5.5) that

$$U^{(n)} U^{(n-1)} \dots U^{(1)} = \tilde{U}^{(n)} = U^{-1}.$$

One may prove (5.3) in a similar way, but it needs more careful calculation. Briefly, first write the elements in the product  $B = L^{(k)} L^{(k+1)} \dots L^{(n)}$ . Then evaluate  $L^{(k-1)} B$  so that the product  $L^{(k)} \dots L^{(n)}$  gives the desired result when  $k = 1$ .  $\square$

The special structure of the Vandermonde matrix allows us to solve  $n \times n$  linear system  $V\mathbf{a} = \mathbf{f}$  efficiently with computational complexity  $O(n^2)$  obtaining

$$\mathbf{a} = U^{(n-1)}U^{(n-2)} \dots U^{(1)}D^{-1}L^{(1)}L^{(2)} \dots L^{(n-1)}\mathbf{f}$$

since each matrix is bidiagonal. The Björck-Pereyra algorithm in [2] uses Newton's divided difference process and bidiagonal products reaching  $\frac{5}{2}n^2$  flops to solve the system. We discovered that the algorithm is essentially connected to the last expression. The Björck-Pereyra algorithm in [15] is generalized to find a fast solution of a linear system whose coefficient matrix is the Cauchy-Vandermonde matrix. The work [7] compares the numerical properties of the well known fast  $O(n^2)$  Traub and Björck-Pereyra algorithms.

### Acknowledgement

The authors thank the referee for his/her remarks and very careful reading of this paper. This research is partially supported by the grant *03.KB.FEN.035* at DEU.

### References

- [1] G. E. Andrews, *The Theory of Partitions*, Cambridge University Press, Cambridge 1998.
- [2] A. Björck, and V. Pereyra, Solution of Vandermonde systems of equations, *Math. Comp.* **24** (1970), 893-903.
- [3] G.S. Cheon and J.S. Kim, Stirling matrix via Pascal Matrix, *Linear Algebra Appl.* **329** (2001), 49-59.
- [4] A. Eisinberg, G. Franzé, and P. Pugliese, Vandermonde matrices on integer nodes, *Numer. Math.* **80** (1998), 75-85.
- [5] A. Eisinberg, P. Pugliese and N. Salerno, Vandermonde matrix on integer nodes: the rectangular case, *Numer. Math.* **87** (2001), 663-674.
- [6] I. Gohberg, I. Koltracht, Triangular factors of Cauchy and Vandermonde matrices, *Integral Equat. Op. Th.* **26** (1996), 46-59.
- [7] I. Gohberg, V. Olshevsky, The Fast Generalized Parker-Traub Algorithm for Inversion of Vandermonde and Related Matrices, *J. Complexity* **13** (1997), 208-234.
- [8] T.N.T. Goodman, Total positivity and shape of curves, in M. Gasca and C. A. Micchelli (Ed.), *Total positivity and its Applications*, Kluwer Academic Publisher, Dordrecht 1996, pp. 157-186.
- [9] T.N.T. Goodman, H. Oruç and G.M. Phillips, Convexity and generalized Bernstein polynomials, *Proc. Edin. Math. Soci.*, **42** (1999), 179-190.

- [10] H.W. Gould, The  $q$ -Stirling numbers of the first kind and the second kind, *Duke Math. Journal*, **28** (1961), 281-289.
- [11] D.E. Knuth, *The Art of Computer Programming* Vol. 1, Addison-Wesley, 1973.
- [12] J. Konvalin, Generalized-binomial coefficients and the subset-subspace Problem, *Advances in Applied Math.*, **21** (1998), 228-240.
- [13] I. G. MacDonald, *Symmetric functions and Hall polynomials*, Clarendon Press Oxford, 1979.
- [14] J.J. Martínez and J.M. Peña, Factorization of Cauchy-Vandermonde matrices, *Linear Algebra Appl.* **284** (1998), 229-237.
- [15] J.J. Martínez and J.M. Peña, Fast algorithms of Björck-Pereyra type for solving Cauchy-Vandermonde linear systems, *Applied Num. Math.* **26** (1998), 343-352.
- [16] H. Oruç, The generalized Bernstein polynomials and Total positivity, Ph.D. Thesis, University of St Andrews 1998;  
<http://web.deu.edu.tr/halil.oruc/>
- [17] H. Oruç and G.M. Phillips, Explicit factorization of the Vandermonde Matrix, *Linear Algebra Appl.* **315** (2000), 113-123.
- [18] H. Oruç and N. Tuncer, On the convergence and iterates of  $q$ -Bernstein polynomials, *J. Approx. Theory.* **117** (2002), 301-313.
- [19] H. Oruç and G.M. Phillips,  $q$ -Bernstein polynomials and Bézier curves, *J. Comp. Appl. Math.* **151** (2003), 1-12.
- [20] G.M. Phillips, *Two Millenia of Mathematics From Archimedes to Gauss*, Springer Verlag, New York 2000.
- [21] G.M. Phillips, *Interpolation and approximation by polynomials*, Springer Verlag, New York 2003.
- [22] L.L. Schumaker, *Spline functions: Basic Theory*, John Wiley & Sons. Inc. 1981.
- [23] R.P. Stanley, *Enumerative Combinatorics* Vol. 2, Cambridge University Press, Cambridge, Massachusetts 1999.
- [24] Z. Zhizheng, The linear algebra of the generalized Pascal Matrix, *Linear Algebra Appl.* **250** (1997), 51-60.