

# On the Metamorphosis of Vandermonde's Identity

DON RAWLINGS

LAWRENCE SZE

California Polytechnic State University

San Luis Obispo, CA 93407

drawling@calpoly.edu

lsze@calpoly.edu

As motivation for our designs, first consider the identity

$$\sum_{l=0}^n \binom{n}{l} = 2^n.$$

Next, imagine inserting  $a^{n-l}b^l$  in the  $l$ th term of the sum and replacing  $2^n$  by  $(a+b)^n$ . The result is the binomial theorem.

Our objective here is to examine the effect of an analogous "pollination" of Vandermonde's identity

$$\sum_{l=0}^i \binom{j}{l} \binom{n-j}{i-l} = \binom{n}{i}. \quad (1)$$

Guided by our binomial theorem example, we slip four letters (two per binomial coefficient) with exponents that add to  $n$  inside the sum in (1) to obtain

$$\sum_{l=0}^i \binom{j}{l} \binom{n-j}{i-l} a^{n+l-i-j} b^{j-l} c^{i-l} d^l. \quad (2)$$

Although there is no simple formula for this sum to rival the  $(a+b)^n$  of the binomial identity, an interesting metamorphosis ensues, producing a noteworthy matrix identity.

For convenience, we abbreviate the sum in (2) by  $V_{i,j} \binom{a,b}{c,d}$ . The leading role in our story is played by the matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}_n = \left( V_{i,j} \binom{a,b}{c,d} \right)_{0 \leq i,j \leq n}, \quad (3)$$

which we shall refer to as the  $n$ th *Vandermonde matrix* with parameters  $a$ ,  $b$ ,  $c$ , and  $d$ . For  $n = 3$ ,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}_3 = \begin{pmatrix} a^3 & a^2b & ab^2 & b^3 \\ 3a^2c & 2abc + a^2d & b^2c + 2abd & 3b^2d \\ 3ac^2 & bc^2 + 2acd & 2bcd + ad^2 & 3bd^2 \\ c^3 & c^2d & cd^2 & d^3 \end{pmatrix}. \quad (4)$$

At least two cases of the Vandermonde matrix have already achieved some notoriety. When  $c = 0$  and  $a = b = d = 1$ , (3) is upper triangular and contains the first  $n + 1$  rows of Pascal's triangle. For instance,

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}_3 = \left( V_{i,j} \binom{1,1}{0,1} \right)_{0 \leq i,j \leq 3} = \left( \binom{j}{i} \right)_{0 \leq i,j \leq 3} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

When transposed, this case of (3) is aptly referred to as the  $n$ th Pascal matrix. The case of (3) with  $c = 0$  and  $a = d = 1$  has also received some attention. It coincides with the transpose of a generalization of Pascal’s matrix considered first by Call and Velleman [2] and later by Aggarwala and Lamoureux [1].

Getting back to our story, our pollination of the sum in (2) leads to the following result.

**THEOREM 1.** *If  $a, b, \dots, g$  are elements of a field, then*

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}_n \begin{bmatrix} e & f \\ g & h \end{bmatrix}_n = \left[ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} \right]_n. \tag{5}$$

In other words, the product of two Vandermonde matrices is Vandermonde. Moreover, the matrix of parameters for the product miraculously coincides with the product of the underlying two-by-two matrices of parameters!

Before proving (5), we present a sampling of its remarkable implications in the next two sections. In the final section, we briefly describe the context that led us to Theorem 1.

**A sampler of inverses and determinants** The most amusing consequences of Theorem 1 involve inverses and determinants of Vandermonde matrices. For  $ad - bc \neq 0$ , (5) implies that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}_n^{-1} = \begin{bmatrix} \frac{d}{ad - bc} & \frac{-b}{ad - bc} \\ \frac{-c}{ad - bc} & \frac{a}{ad - bc} \end{bmatrix}_n = \frac{1}{(ad - bc)^n} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}_n. \tag{6}$$

As an example of (6), the inverse of the transpose of Pascal’s matrix is readily seen to be

$$\left( \binom{j}{i} \right)_{0 \leq i, j \leq n}^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}_n^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}_n = \left( (-1)^{j-i} \binom{j}{i} \right)_{0 \leq i, j \leq n}.$$

So, for  $n = 3$ ,

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Although the origin of this equality is unclear to us, it is well known and appears in a number of contexts (including inclusion-exclusion [5, p. 67]).

Next, we note that the determinant of the Vandermonde matrix is just a power of the determinant of its underlying two-by-two matrix of parameters.

**COROLLARY 1.**

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}_n = (ad - bc)^{n(n+1)/2}.$$

*Proof.* The result follows directly if  $a, b, c$ , or  $d$  is zero; for instance, if  $c = 0$ , then (3) (the example in (4) is illustrative) is upper triangular with  $i$ th diagonal entry  $a^{n-i}d^i$  for  $0 \leq i \leq n$ . Thus,

$$\det \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}_n = a^n d^0 a^{n-1} d^1 \dots a^0 d^n = (ad - b \cdot 0)^{n(n+1)/2}.$$

So assume that  $a, b, c,$  and  $d$  are all nonzero. As

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}_n = \begin{bmatrix} (ad - bc)/d & b/d \\ 0 & 1 \end{bmatrix}_n \begin{bmatrix} 1 & 0 \\ c & d \end{bmatrix}_n \tag{7}$$

and as the determinant of a product is the product of determinants, we have

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}_n = \left( \frac{ad - bc}{d} \right)^{n(n+1)/2} d^{n(n+1)/2} = (ad - bc)^{n(n+1)/2}. \quad \blacksquare$$

Corollary 1 bears a resemblance to the classical *Vandermonde determinant*

$$\det(x_i^{n-j})_{0 \leq i, j \leq n} = \prod_{0 \leq i < j \leq n} (x_i - x_j).$$

Motivated by the similarities, we tried introducing even more variables, indexing our parameters  $a$  and  $b$  by row. The result in (8) below, which we refer to as the Vandermonde expansion, may be regarded as a distant cousin of the binomial theorem.

For  $\vec{a} = (a_0, a_1, \dots, a_n)$  and  $\vec{b} = (b_0, b_1, \dots, b_n)$ , let

$$\begin{bmatrix} \vec{a} & \vec{b} \\ c & d \end{bmatrix}_n = \left( V_{i,j} \begin{pmatrix} a_i, b_i \\ c, d \end{pmatrix} \right)_{0 \leq i, j \leq n}.$$

Then, a slight variation on our proof of Corollary 1 leads to

$$\det \begin{bmatrix} \vec{a} & \vec{b} \\ c & d \end{bmatrix}_n = (a_0d - b_0c)^n (a_1d - b_1c)^{n-1} \dots (a_nd - b_nc)^0. \tag{8}$$

The key is to observe that the matrix on the extreme right in (7) is independent of  $a$  and  $b$ . So (7) remains true when, in both of the other matrices, the  $a$  and  $b$  in row  $i$  are respectively replaced by  $a_i$  and  $b_i$  for  $0 \leq i \leq n$ .

Although the Vandermonde expansion will never become as popular as the binomial theorem, it contains some striking special cases. When  $n = 3, c = -1, d = 1, \vec{a} = (a, 1, 1, 1),$  and  $\vec{b} = (b, 0, 0, 0),$  (8) reduces to

$$(a + b)^3 = \begin{vmatrix} a^3 & a^2b & ab^2 & b^3 \\ -3 & 1 & 0 & 0 \\ 3 & -2 & 1 & 0 \\ -1 & 1 & -1 & 1 \end{vmatrix}.$$

In the above, note that the usual suspects in the expansion of  $(a + b)^3$  appear across the first row and that a truncated, signed Pascal triangle is contained in the lower left corner. Of course, such determinant formulas for  $(a + b)^k, 0 \leq k \leq n(n + 1)/2,$  may be obtained by simply running monomials through appropriate rows of the signed Pascal triangle. For instance, if  $n = 3, c = -1, d = 1, \vec{a} = (a, a, 1, 1),$  and  $\vec{b} = (b, b, 0, 0),$  then (8) implies

$$(a + b)^5 = \begin{vmatrix} a^3 & a^2b & ab^2 & b^3 \\ -3a^2 & a^2 - 2ab & 2ab - b^2 & 3b^2 \\ 3 & -2 & 1 & 0 \\ -1 & 1 & -1 & 1 \end{vmatrix}.$$

Among the outright curious, the balanced selections  $\vec{a} = (a, 1/2, 1/2, 1/2)$  and  $\vec{b} = (b, 1/2, 1/2, 1/2)$  with  $n = 3, c = -1,$  and  $d = 1$  in (8) give

$$(a + b)^3 = \begin{vmatrix} a^3 & a^2b & ab^2 & b^3 \\ -3/4 & -1/4 & 1/4 & 3/4 \\ 3/2 & -1/2 & -1/2 & 3/2 \\ -1 & 1 & -1 & 1 \end{vmatrix}.$$

**A sampler of scalar identities** Theorem 1 is a cache of binomial identities ranging from the well-known to the exotic. Careful selection of the parameters will in fact reveal Vandermonde’s identity, the binomial theorem, and other results of interest. First, by the definition of matrix multiplication, (5) is equivalent to

$$\sum_{k=0}^n V_{i,k} \begin{pmatrix} a, b \\ c, d \end{pmatrix} V_{k,j} \begin{pmatrix} e, f \\ g, h \end{pmatrix} = V_{i,j} \begin{pmatrix} ae + bg, af + bh \\ ce + dg, cf + dh \end{pmatrix}, \tag{9}$$

where  $j$  is any integer from 0 to  $n$ .

To extract (1) from (9), note that

$$V_{i,k} \begin{pmatrix} 1, 0 \\ 1, 0 \end{pmatrix} = \sum_{l=0}^i \binom{k}{l} \binom{n-k}{i-l} 1^{n+l-i-k} 0^{k-l} 1^{i-l} 0^l = \begin{cases} \binom{n}{i} & \text{if } k = 0, \\ 0 & \text{if } 0 < k \leq n. \end{cases}$$

Similarly,

$$V_{k,j} \begin{pmatrix} 1, 1 \\ 0, 1 \end{pmatrix} = \sum_{l=0}^k \binom{j}{l} \binom{n-j}{k-l} 1^{n+l-k-j} 1^{j-l} 0^{k-l} 1^l = \binom{j}{k}.$$

Thus, we obtain (1):

$$\binom{n}{i} = \sum_{k=0}^n V_{i,k} \begin{pmatrix} 1, 0 \\ 1, 0 \end{pmatrix} V_{k,j} \begin{pmatrix} 1, 1 \\ 1, 0 \end{pmatrix} = V_{i,j} \begin{pmatrix} 1, 1 \\ 1, 1 \end{pmatrix} = \sum_{l=0}^i \binom{j}{l} \binom{n-j}{i-l}.$$

Of course, further setting  $j = 1$  gives the well-known binomial recurrence

$$\binom{n}{i} = \binom{n-1}{i} + \binom{n-1}{i-1}. \tag{10}$$

To see the binomial theorem emerge from (9), observe that

$$V_{0,k} \begin{pmatrix} a, b \\ a, b \end{pmatrix} = a^{n-k} b^k \quad \text{and} \quad V_{k,0} \begin{pmatrix} 1, 1 \\ 1, 1 \end{pmatrix} = \binom{n}{k}.$$

So

$$(a + b)^n = V_{0,0} \begin{pmatrix} a + b, a + b \\ a + b, a + b \end{pmatrix} = \sum_{k=0}^n V_{0,k} \begin{pmatrix} a, b \\ a, b \end{pmatrix} V_{k,0} \begin{pmatrix} 1, 1 \\ 1, 1 \end{pmatrix} = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k.$$

In the realm of the exotic, we note that the special case

$$\sum_{k=0}^n V_{i,k} \begin{pmatrix} 1, -1 \\ 1, 1 \end{pmatrix} V_{k,j} \begin{pmatrix} 1, 1 \\ 2, 1 \end{pmatrix} = V_{i,j} \begin{pmatrix} -1, 0 \\ 3, 2 \end{pmatrix}$$

of (9) translates into

$$\sum_{k=0}^n \sum_{m=0}^i \sum_{l=0}^k (-1)^{k-m} \binom{k}{m} \binom{n-k}{i-m} \binom{j}{l} \binom{n-j}{k-l} 2^{k-l} = (-1)^{n-i} \binom{n-j}{i-j} 2^j 3^{i-j}.$$

Closed formulas for such sums are easily produced. The trick is to select  $a, b, \dots, h$  so that at least one of the parameters on the right in (9) is zero.

Next, as Vandermonde's identity (1) holds when  $j$  is viewed as an indeterminate, it is only natural to ask whether (9) has a similar extension. Annoyingly, the answer in general is no. Recall that for  $j$  an indeterminate and  $l$  a nonnegative integer, the extended binomial coefficient is defined by

$$\binom{j}{l} = \begin{cases} \frac{j(j-1)\cdots(j-l+1)}{l!} & \text{if } l > 0, \\ 1 & \text{if } l = 0. \end{cases} \quad (11)$$

Of course, (11) agrees with the usual binomial coefficient when  $j$  is replaced by an integer greater than or equal to  $l$ . Also, (11) is a polynomial in  $j$  of degree  $l$ . The difficulty in extending (9) is exposed by noting that, if  $j$  is a real number other than  $0, 1, \dots, \text{ or } n$ , then

$$\sum_{k=0}^n V_{i,k} \binom{0,1}{0,1} V_{k,j} \binom{0,0}{0,1} = \binom{n}{i} \binom{j}{n} \neq 0^{n-j} \binom{n}{i} = V_{i,j} \binom{0,1}{0,1}.$$

However, (9) may be extended with some restrictions.

One approach, among many, is to restrict  $a, b, \dots, h$  to the field of real numbers and require that

$$eh = gf \quad \text{and that } e, f, ae + bg \text{ and } af + bh \text{ are all positive.} \quad (12)$$

The proof that (9) holds under these conditions parallels the standard technique for extending (1): First, for  $0 \leq k \leq n$ , note that

$$e^j f^{-j} V_{k,j} \binom{e,f}{g,h} = \sum_{l=0}^k \binom{j}{l} \binom{n-j}{k-l} e^{n+l-k} f^{-l} g^{k-l} h^l$$

is a polynomial in  $j$  of degree at most  $k$ . Now, for  $0 \leq i \leq n$ , define

$$p(j) = e^j f^{-j} V_{i,j} \binom{ae+bg, af+bh}{ce+dg, cf+dh} - e^j f^{-j} \sum_{k=0}^n V_{i,k} \binom{a,b}{c,d} V_{k,j} \binom{e,f}{g,h}.$$

With (12) in mind and a little work, it may be verified that  $p(j)$  is a polynomial in  $j$  of degree at most  $n$ . As (9) implies that  $p(j)$  has at least  $n+1$  roots (namely,  $0, 1, \dots, n$ ),  $p(j)$  must in fact be the zero polynomial. Thus, under the terms of (12), (9) holds for  $j$  an indeterminate.

As a final example, we note that setting  $a = c = e = f = 1, b = d = g = h = 0$ , and  $j = -1$  in our extension of (9) gives the commonplace equality

$$\binom{n}{i} = \sum_{l=0}^i (-1)^l \binom{n+1}{i-l}.$$

**An algebraic proof** There are a number of proofs of Theorem 1. For one, it is possible to extend the usual combinatorial proof of (1). It is not too difficult to see that the sum in (2) may be interpreted as a *weighted* selection of a committee of size  $i$  from a group of people consisting of  $j$  women and  $n-j$  men. Induction will also do the job: With judicious use of (10), it may be verified that both sides of (9) (which, we recall, is equivalent to (5) in Theorem 1) satisfy the recurrence relationship

$$\rho_{i,j}(n) = (af + bh)\rho_{i,j-1}(n-1) + (cf + dh)\rho_{i-1,j-1}(n-1).$$

However, we find neither the combinatorial approach nor the induction argument entirely satisfying. In our opinion, the slickest proof, the only one we present, relies on linear algebra.

We set the stage. Let  $V$  be an  $m$ -dimensional vector space with ordered basis  $\beta = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m)$ . The matrix representation relative to  $\beta$  of a linear operator  $T : V \rightarrow V$  is  $[T]_\beta = (a_{i,j})_{1 \leq i,j \leq m}$ , where  $a_{i,j}$  is the  $i$ th coordinate of  $T(\vec{v}_j)$ , that is, the  $a_{i,j}$  are scalars satisfying

$$T(\vec{v}_j) = \sum_{i=1}^m a_{i,j} \vec{v}_i \quad \text{for } 1 \leq j \leq m.$$

The algebraic key to (5) is the fact that the matrix associated with the composition of linear operators is the product of the matrices of the operators. In other words, if  $S, T : V \rightarrow V$  are linear operators, then (see Friedberg, Insel, and Spence [3, Ch. 2])

$$[S \circ T]_\beta = [S]_\beta [T]_\beta. \tag{13}$$

Identity (13) is a handy tool for establishing properties of matrix multiplication that avoids much of the tedium of indices! It's perfect for our purposes.

*Proof of Theorem 1.* Let  $F$  denote a field,  $a, b, \dots, h$  be elements of  $F$ , and  $F[x, y]$  be the ring of polynomials over  $F$  in the commuting indeterminates  $x$  and  $y$ . For each polynomial  $p(x, y) \in F[x, y]$ , define

$$M(p) = p(ax + cy, bx + dy) \quad \text{and} \quad N(p) = p(ex + gy, fx + hy).$$

By thinking in terms of matrix products, we may express the formulas for  $M$  and  $N$  in the more satisfying forms

$$M(p) = p \left( (x, y) \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right) \quad \text{and} \quad N(p) = p \left( (x, y) \begin{pmatrix} e & f \\ g & h \end{pmatrix} \right). \tag{14}$$

Note that  $M, N : F[x, y] \rightarrow F[x, y]$ . Also, both  $M$  and  $N$  are ring homomorphisms. From (14), the composition of  $M$  with  $N$  applied to a polynomial  $p(x, y) \in F[x, y]$  is seen to be

$$M \circ N(p) = p \left( (x, y) \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} \right).$$

We now turn our attention to the subset

$$H_n = \{a_0x^n + a_1x^{n-1}y + \dots + a_ny^n : a_0, a_1, \dots, a_n \in F\}$$

of  $F[x, y]$ . The nonzero elements in  $H_n$  are just the homogeneous polynomials of degree  $n$  in the indeterminates  $x$  and  $y$ . Note that  $H_n$  is a vector space over  $F$  and that  $\gamma = (x^n, x^{n-1}y, \dots, y^n)$  constitutes an ordered basis of  $H_n$ . Moreover, the restrictions of  $M$  and  $N$  to  $H_n$  are linear operators on  $H_n$ .

To determine the matrix of  $M$  relative to the ordered basis  $\gamma$ , we fish the coefficient of  $x^{n-i}y^i$  out of

$$M(x^{n-j}y^j) = (ax + cy)^{n-j}(bx + dy)^j. \tag{15}$$

For  $0 \leq l \leq i$ , the binomial theorem tells us that the coefficient of

$$x^{n+l-i-j}y^{i-l} \quad \text{in} \quad (ax + cy)^{n-j} \quad \text{is} \quad \binom{n-j}{i-l} a^{n+l-i-j} c^{i-l}$$

and that the coefficient of

$$x^{j-l}y^l \text{ in } (bx + dy)^j \text{ is } \binom{j}{l}b^{j-l}d^l.$$

Noting that  $x^{n-i}y^i = x^{n+l-i-j}y^{i-l}x^{j-l}y^l$  for  $0 \leq l \leq i$ , it is then evident that the coefficient of  $x^{n-i}y^i$  in (15) is none other than

$$\sum_{l=0}^i \binom{j}{l} \binom{n-j}{i-l} a^{n+l-i-j} b^{j-l} c^{i-l} d^l = V_{i,j} \begin{pmatrix} a, b \\ c, d \end{pmatrix}.$$

So the matrix relative to  $\gamma$  of  $M$  restricted to  $H_n$  is

$$[M]_{\gamma} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}_n.$$

As similar considerations lead to

$$[N]_{\gamma} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}_n \quad \text{and} \quad [M \circ N]_{\gamma} = \begin{bmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \begin{pmatrix} e & f \\ g & h \end{pmatrix} \end{bmatrix}_n,$$

(13) delivers the final blow:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}_n \begin{bmatrix} e & f \\ g & h \end{bmatrix}_n = [M]_{\gamma} [N]_{\gamma} = [M \circ N]_{\gamma} = \begin{bmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \begin{pmatrix} e & f \\ g & h \end{pmatrix} \end{bmatrix}_n. \quad \blacksquare$$

For the adventurous, we note that our algebraic proof is readily adapted to deduce identities for *extended* Vandermonde matrices with  $k^2$  parameters for any integer  $k \geq 2$ .

**Concluding remarks** We were led to consider the pollinated sum (2) and to the discovery of Theorem 1 by certain practical considerations. There are many natural contexts in which the elements of a fixed set vary with time between two states. For one, the members of a given population may or may not have a certain contagious disease. From one moment to the next, a healthy individual may become ill and an infected individual may recover. For another, the components of a system of service may either be in or out of service. Again, with each passing moment, an in-service component may fail while a broken component may be repaired and returned to service. It turns out that, under certain probabilistic assumptions, such processes can be modeled as Markov chains. Moreover, the corresponding transition matrices are Vandermonde.

In this context, Theorem 1 is an indispensable tool. It allows us to manipulate (multiply, invert, and diagonalize) Vandermonde matrices at will. As we've demonstrated, such computations miraculously boil down to working with the underlying two-by-two matrices of associated parameters.

## REFERENCES

1. R. Aggarwala and M. P. Lamoureux, Inverting the Pascal matrix plus one, *Amer. Math. Monthly* **109** (2002), 371–377.
2. G. S. Call and D. J. Velleman, Pascal's matrices, *Amer. Math. Monthly* **100** (1993), 372–376.
3. S. Friedberg, A. Insel, and L. Spence, *Linear Algebra*, third edition, Prentice Hall, NJ, 1997.
4. D. Rawlings and L. Sze, On the reliability of an  $n$ -component system, in preparation.
5. R. P. Stanley, *Enumerative Combinatorics* Vol. I, Wadsworth and Brooks/Cole, Pacific Grove, CA, 1986.