Indag. Mathem., N.S., 4 (1), 11-16 March 29, 1993

The q-binomial theorem and spectral symmetry

by Rajendra Bhatia* and Ludwig Elsner

Indian Statistical Institute, New Delhi 110016, India and
Fakultät für Mathematik, Universität Bielefeld, 4800 Bielefeld 1, Germany

Communicated by Prof. R. Tijdeman at the meeting of April 27, 1992

SUMMARY

In various contexts, several mathematicians have discovered a binomial theorem of the following form: Let T_1, T_2 be complex matrices such that $T_2T_1 = qT_1T_2$. Then

$$(T_1+T_2)^n = \sum_{k=0}^{n} \alpha_{n,k}(q)T_1^kT_2^{n-k}$$

and the polynomials $\alpha_{n,k}(q)$ are given explicitly. We describe an application of this result in our work on matrices whose eigenvalues have certain symmetries.

In various contexts, several mathematicians in recent years have discovered a beautiful binomial theorem:

THEOREM 1. *Let T_1, T_2 be (complex) matrices such that*

(1) $$T_2T_1 = qT_1T_2$$

for some complex number q. Then for each positive integer n, we have the binomial expansion

(2) $$(T_1+T_2)^n = \sum_{k=0}^{n} \alpha_{n,k}(q)T_1^kT_2^{n-k},$$

where the coefficient $\alpha_{n,k}(q)$ are polynomials in q satisfying the properties

*The first author thanks SFB 343 at the University of Bielefeld for their support and hospitality during June 1990 when a part of this work was done, the NBHM, India for a travel grant and the DAE, India for a research grant.

$$
(3) \quad \begin{cases} \text{(i)} \quad \alpha_{n+1,k}(q) = \alpha_{n,k}(q) + q^{n+1-k}\alpha_{n,k-1}(q) \quad \textit{for } k = 1,2,\ldots,n; \\ \qquad \alpha_{n,0} = \alpha_{n,n} = 1; \\ \text{(ii)} \quad \alpha_{n,k}(q) = \dfrac{(1-q)(1-q^2)\cdots(1-q^n)}{(1-q)(1-q^2)\cdots(1-q^k)(1-q)(1-q^2)\cdots(1-q^{n-k})}; \\ \text{(iii)} \quad \textit{degree } \alpha_{n,k}(q) = k(n-k). \end{cases}
$$

PROOF. Multiplying (2) by $T_1 + T_2$ and comparing coefficients we get (i). Then (ii) follows from (i) by induction and (iii) as a simple consequence. ■

This theorem is attributed to several authors in the recent monograph [4, p. 28]. It was discovered yet again in [3]. We were led to it in the course of our work on matrices whose eigenvalues have certain symmetries. This application of Theorem 1 is described in this note.

The eigenvalues of the $p \times p$ matrix

$$
(4) \quad \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ t & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}
$$

are the pth roots of t. This symmetric distribution of roots is a very special instance of the following general situation. Let X be a complex matrix of order $n = pr$, having the special form

$$
(5) \qquad X = \begin{bmatrix} 0 & A_1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & A_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & A_{p-1} \\ A_p & 0 & 0 & 0 & \cdots & 0 \end{bmatrix},
$$

where $A_1, \ldots, A_p$ are square matrices of order r. We will call such a matrix a $p \times p$ *block cyclic matrix*. Let S be the *block diagonal* matrix

$$
(6) \qquad S = \operatorname{diag}(I_r, \omega I_r, \ldots, \omega^{p-1} I_r),
$$

where I_r is the identity matrix of order r and ω is the primitive pth root of unity. Then $S^{-1}XS = \omega X$. This implies that the eigenvalues of X are symmetrically distributed in the following sense: they can be enumerated as

$$
(7) \qquad (\lambda_1, \ldots, \lambda_r, \omega\lambda_1, \ldots, \omega\lambda_r, \ldots, \omega^{p-1}\lambda_1, \ldots, \omega^{p-1}\lambda_r).
$$

We call an n-tuple like (7) a *p-Carrollian* n-tuple and we say that the matrix X has a *p-Carrollian* spectrum.

Note that if a scalar matrix is added to (4) then its spectrum is no longer p-Carrollian. However, certain diagonal perturbations do preserve this property. In [2] Choi proved the following interesting proposition. Let

$$(8) \qquad Z = \begin{bmatrix} R & A_1 \\ A_2 & -R \end{bmatrix},$$

where A_1, A_2 and R are $r \times r$ matrices such that R commutes with A_1. Then the spectrum of Z is 2-Carrollian. (Note that such a matrix Z is not necessarily similar to $-Z$.)

Choi has used this in connection with his work on some K-theoretic questions about matrices. For us its interest is in the following interpretation. Write

$$(9) \qquad Y = \begin{bmatrix} R & 0 \\ 0 & -R \end{bmatrix}.$$

Then $Z = X + Y$, where X is a 2×2 block cyclic matrix. We saw above that the spectrum of X is 2-Carrollian and Choi's proposition says that this property is preserved when we perturb X by adding to it the block diagonal matrix Y, provided R commutes with A_1.

It turns out that this phenomenon occurs for all values of p. More precisely, we have:

THEOREM 2. *Let X be a $p \times p$ block cyclic matrix as in* (5) *and let Y be a block diagonal matrix of the form*

$$(10) \qquad Y = \begin{bmatrix} R & 0 & 0 & \cdots & 0 \\ 0 & \omega R & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \omega^{p-1} R \end{bmatrix},$$

where ω is the primitive pth root of unity. Let

$$(11) \qquad Z = X + Y.$$

Suppose R commutes with $A_1, A_2, \ldots, A_{p-1}$. Then the spectrum of Z is p-Carrollian.

In an earlier paper [1] we gave a proof of Theorem 2 based on the block LU-decomposition often used in numerical analysis. Choi had used a very similar idea in his proof [2]. We will now give a proof based on Theorem 1.

PROOF OF THEOREM 2. We will first prove the theorem in a special case: assume that R commutes with all the matrices $A_1, A_2, \ldots, A_p$. In this case the matrices X and Y given by (5) and (10) satisfy the commutation relation

$$(12) \qquad XY = \omega YX.$$

From (3) one sees that

$$\alpha_{p,k}(\omega)=0 \quad \text{for } 1\le k\le p-1.$$

Hence, by Theorem 1

(13) $$(X+Y)^p = X^p + Y^p.$$

Now note that the matrix X^p is block-diagonal and its diagonal entries are $A_{\sigma(1)}A_{\sigma(2)}\cdots A_{\sigma(p)}$, where σ runs over all cyclic permutations of $\{1,2,\dots,p\}$. So, by (13) the matrix Z^p is also block-diagonal. Hence Z^m is block-diagonal for $m=p,2p,3p,\dots$. We claim that

(14) $$\operatorname{tr} Z^m = 0 \quad \text{if } m\ne p,2p,3p,\dots.$$

To see this note that X^k always has zero blocks on its diagonal if k is not an integral multiple of p. Hence $\operatorname{tr} Y^jX^k=0$ for all j and for all $k\ne p,2p,\dots$. If $k=rp$ then X^k is block-diagonal whose diagonal entries are the rth powers of $A_{\sigma(1)}A_{\sigma(2}\cdots A_{\sigma(p)}$, σ running over all cyclic permutations of $\{1,2,\dots,p\}$. So, if j is not an integral multiple of p but k is, then again $\operatorname{tr} Y^jX^k=0$. So, the statement (14) follows from (2).

Now, it is a consequence of Newton's identities connecting elementary symmetric polynomials and sums of powers that if E and F are two $n\times n$ matrices with $\operatorname{tr} E^k=\operatorname{tr} F^k$, $1\le k\le n$, then E and F have the same eigenvalues. See, e.g., [5, p. 44]. Hence, (14) implies that the matrices $Z,\omega Z,\dots,\omega^{p-1}Z$ all have the same eigenvalues. This completes the proof of the special case of the Theorem.

In the general case, when R commutes with $A_1,\dots,A_{p-1}$ but not with A_p, the above proof can be modified as indicated below.

Instead of (12) we now have a relation

(15) $$XY=\omega YX+E,$$

where E is a $p\times p$ block matrix all whose block entries are zero except the one in the bottom left corner, and this entry is $E_{p1}=A_pR-RA_p$.

At the next step, we find that now Z^p is not necessarily block-diagonal. However, it still has a special form: it turns out to be a block lower triangular matrix. To see this note that if $1\le \nu\le p$ and if J is a product of ν matrices each of which is a $p\times p$ block cyclic matrix then J has a special block structure: all blocks of J are zero except those which are on the νth superdiagonal or on the $(p-\nu)$th subdiagonal. (The case $\nu=p$ says that a product of p such matrices is block-diagonal.) Now consider a typical mixed term C in the expansion of $(Y+X)^p$. If X occurs k times as a factor in C, where $1\le k\le p-1$ and if C is not yet of the form $Y^{p-k}X^k$, then we can write

(16) $$C=PXYQ,$$

where P is a product of ν block cyclic matrices. Q is a product of μ block cyclic matrices and $\nu+\mu=k-1$. One application of (15) converts the equation (16) into

$$C=\omega PYXQ+PEQ.$$

The matrix PEQ in its block form is strictly lower triangular, as can be verified. Repeated applications of (15) finally bring C to the form

$$C = \omega^m Y^{p-k} X^k + T, \tag{17}$$

where T is a strictly lower triangular block matrix and m is an integer equal to the number of times a letter Y occurring in the original form (16) is interchanged with a letter X using the rule (15) till we reach the final form (17). Had E been zero our commutation rule (15) would have reduced to (12) and our expansion for $(Y+X)^p$ would have been given by Theorem 1. Hence, we must have

$$Z^p = (Y+X)^p = \sum_{k=0}^{p} \alpha_{p,k}(\omega) Y^k X^{p-k} + S = Y^p + X^p + S, \tag{18}$$

where S is a strictly block lower triangular matrix. In other words Z^p is a block lower triangular matrix as we claimed, and its diagonal entries are $R^p + A_{\sigma(1)} A_{\sigma(2)} \cdots A_{\sigma(p)}$ where σ runs over cyclic permutations of $\{1, 2, \ldots, p\}$.

The statement (14) remains true in this case; the details of its verification are omitted. As before, the general case of the Theorem follows. ■

Finally, we make a few remarks on Theorem 1. The coefficients $\alpha_{n,k}(q)$ are called q-binomial coefficients in [4] and, in a more suggestive notation, are denoted by $[{n \atop k}]_q$. Following the analogy with the usual binomial coefficients one may wonder whether a similar multinomial theorem can be proved. Indeed, one can prove that if $T_1, T_2, \ldots, T_m$ are matrices satisfying the commutation rules

$$T_j T_i = q T_i T_j \quad \text{for } j > i, \tag{19}$$

then we have

$$(T_1 + \cdots + T_m)^n = \sum \begin{bmatrix} n \\ j_1, j_2, \ldots, j_m \end{bmatrix}_q T_1^{j_1} T_2^{j_2} \cdots T_m^{j_m}, \tag{20}$$

where the summation is over all choices of indices $j_1, \ldots, j_m$ such that $j + \cdots + j_m = n$, and the coefficients occurring in the above expansion are defined as

$$\begin{bmatrix} n \\ j_1, j_2, \ldots, j_m \end{bmatrix}_q = \begin{bmatrix} n \\ j_1 \end{bmatrix}_q \begin{bmatrix} n - j_1 \\ j_2 \end{bmatrix}_q \cdots \begin{bmatrix} n - (j_1 + j_2 + \cdots + j_{m-2}) \\ j_{m-1} \end{bmatrix}_q. \tag{21}$$

A little more generally we can prove using Theorem 1 that if $T_1, T_2, \ldots, T_m$ are matrices satisfying the commutation rules

$$T_j T_i = q_i T_i T_j \quad \text{for } j > i;\ i = 1, 2, \ldots, m-1, \tag{22}$$

then we have

$$(T_1 + \cdots + T_m)^n = \sum \alpha_{n; j_1, \ldots, j_m}(q_1, \ldots, q_{m-1}) T_1^{j_1} T_2^{j_2} \cdots T_m^{j_m}, \tag{23}$$

where the summation is over all indices $j_1, \ldots, j_m$ such that $j_1 + \cdots + j_m = n$ and the coefficients are defined as

$$\alpha_{n;\, j_1, j_2, \ldots, j_m}(q_1, \ldots, q_{m-1}) \tag{24}$$
$$= \begin{bmatrix} n \\ j_1 \end{bmatrix}_{q_1} \begin{bmatrix} n-j_1 \\ j_2 \end{bmatrix}_{q_2} \cdots \begin{bmatrix} n-(j_1+j_2+\cdots+j_{m-2}) \\ j_{m-1} \end{bmatrix}_{q_{m-1}}.$$

We wonder whether a neat and simple multinomial theorem can be obtained for matrices $T_1, T_2, \ldots, T_m$ which obey a more general commutation rule

$$T_j T_i = q_{ij} T_i T_j \quad \text{for } j > i;\ 1 \le i \le m-1. \tag{25}$$

Since the q-binomial theorem and the q-binomial coefficients turn up in diverse problems in combinatorics, number theory, probability, geometry, analysis and physics [4, p. 29], there are likely to be uses for a similar multinomial theorem.

REFERENCES

1. Bhatia, R. and L. Elsner – Symmetries and variation of spectra. Can. Math. J., to appear.
2. Choi, M.D. – Almost commuting matrices need not be nearly commuting. Proc. Amer. Math. Soc., **102**, 529-533 (1988).
3. Choi, M.D., G.A. Elliott and N. Yui – Gauss polynomials and the rotation algebra. Inventiones Math. **99**, 225-246 (1990).
4. Gasper G. and M. Rahman – Basic Hypergeometric Series. Encyclopedia of Mathematics and its Applications, Cambridge University Press (1990).
5. Horn, R. and C.R. Johnson – Matrix Analysis, Cambridge University Press (1985).