A MATRIX GENERALIZATION OF A THEOREM OF FINE

ERIC ROWLAND

ABSTRACT. In 1947 Nathan Fine gave a beautiful product for the number of binomial coefficients $\binom{n}{m}$, for m in the range $0 \leq m \leq n$, that are not divisible by p. We give a matrix product that generalizes Fine's formula, simultaneously counting binomial coefficients with p-adic valuation α for each $\alpha \geq 0$. For each n this information is naturally encoded in a polynomial generating function, and the sequence of these polynomials is p-regular in the sense of Allouche and Shallit. We also give a further generalization to multinomial coefficients.

1. Binomial coefficients

For a prime p and an integer $n \geq 0$, let $F_p(n)$ be the number of integers m in the range $0 \leq m \leq n$ such that $\binom{n}{m}$ is not divisible by p. Let the standard base-p representation of n be $n_\ell \cdots n_1 n_0$. Fine [6] showed that

$$F_p(n) = (n_0 + 1)\,(n_1 + 1)\,\cdots\,(n_\ell + 1).$$

Equivalently,

$$F_p(n) = \prod_{d=0}^{p-1} (d+1)^{|n|_d}, \tag{1}$$

where $|n|_w$ is the number of occurrences of the word w in the base-p representation of n. In the special case $p = 2$, Glaisher [7] was aware of this result nearly 50 years earlier.

Many authors have been interested in generalizing Fine's theorem to higher powers of p. Since Equation (1) involves $|n|_d$, a common approach is to express the number of binomial coefficients satisfying some congruence property modulo p^α in terms of $|n|_w$ for more general words w. Howard [8], Davis and Webb [4], Webb [16], and Huard, Spearman, and Williams [9, 10, 11] all produced results in this direction. Implicit in the work of Barat and Grabner [2, §3] is that the number of binomial coefficients $\binom{n}{m}$ with p-adic valuation α is equal to $F_p(n) \cdot G_{p^\alpha}(n)$, where $G_{p^\alpha}(n)$ is some polynomial in the subword-counting functions $|n|_w$. The present author [14] gave an algorithm for computing a suitable polynomial $G_{p^\alpha}(n)$. Spiegelhofer and Wallner [15] showed that $G_{p^\alpha}(n)$ is unique and greatly sped up its computation by showing that its coefficients can be read off from certain power series.

These general results all use the following theorem of Kummer [12, pages 115–116]. Let $\nu_p(n)$ denote the p-adic valuation of n, that is, the exponent of the highest power of p dividing n. Let $\sigma_p(m)$ be the sum of the standard base-p digits of m.

Kummer's theorem. *Let p be a prime, and let n and m be integers with $0 \leq m \leq n$. Then $\nu_p(\binom{n}{m})$ is the number of carries involved in adding m to $n - m$ in base p. Equivalently, $\nu_p(\binom{n}{m}) = \frac{\sigma_p(m)+\sigma_p(n-m)-\sigma_p(n)}{p-1}$.*

Date: April 21, 2017.

Kummer's theorem follows easily from Legendre's formula

$$\nu_p(m!) = \frac{m - \sigma_p(m)}{p-1}$$

for the p-adic valuation of $m!$.

Our first theorem is a new generalization of Fine's theorem. It provides a matrix product for the polynomial

$$T_p(n, x) = \sum_{m=0}^{n} x^{\nu_p(\binom{n}{m})}$$

whose coefficient of x^α is the number of binomial coefficients $\binom{n}{m}$ with p-adic valuation α. In particular, $T_p(n, 0) = F_p(n)$. For $p = 2$ the first few values of the sequence $(T_2(n, x))_{n \geq 0}$ are as follows.

n	$T_2(n,x)$	n	$T_2(n,x)$
0	1	8	$4x^3 + 2x^2 + x + 2$
1	2	9	$4x^2 + 2x + 4$
2	$x + 2$	10	$2x^3 + x^2 + 4x + 4$
3	4	11	$4x + 8$
4	$2x^2 + x + 2$	12	$2x^3 + 5x^2 + 2x + 4$
5	$2x + 4$	13	$2x^2 + 4x + 8$
6	$x^2 + 2x + 4$	14	$x^3 + 2x^2 + 4x + 8$
7	8	15	16

The polynomial $T_p(n, x)$ was identified by Spiegelhofer and Wallner [15] as an important component in the efficient computation of the polynomial $G_{p^\alpha}(n)$. Everett [5] was also essentially working with $T_p(n, x)$.

For each $d \in \{0, 1, \ldots, p-1\}$, let

$$M_p(d) = \begin{bmatrix} d+1 & p-d-1 \\ d\,x & (p-d)\,x \end{bmatrix}.$$

Theorem 1. *Let p be a prime, and let $n \geq 0$. Let $n_\ell \cdots n_1 n_0$ be the standard base-p representation of n. Then*

$$T_p(n, x) = \begin{bmatrix} 1 & 0 \end{bmatrix} M_p(n_0)\, M_p(n_1) \,\cdots\, M_p(n_\ell) \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

A sequence $s(n)_{n \geq 0}$, with entries in some field, is *p-regular* if the vector space generated by the set of subsequences $\{s(p^e n + i)_{n \geq 0} : e \geq 0 \text{ and } 0 \leq i \leq p^e - 1\}$ is finite-dimensional. Allouche and Shallit [1] introduced regular sequences and showed that they have several desirable properties, making them a natural class. The sequence $(F_p(n))_{n \geq 0}$ is included as an example of a p-regular sequence of integers in their original paper [1, Example 14]. It follows from Theorem 1 and [1, Theorem 2.2] that $(T_p(n, x))_{n \geq 0}$ is a p-regular sequence of polynomials.

Whereas Fine's product can be written as Equation (1), Theorem 1 cannot be written in an analogous way, since the matrices $M_p(i)$ and $M_p(j)$ do not commute if $i \neq j$.

The proof of Theorem 1 uses Lemma 4, which is stated and proved in general for multinomial coefficients in Section 2. The reason for including the following proof of Theorem 1 is that the outline is fairly simple. The details relegated to Lemma 4 are not essentially simpler in the case of binomial coefficients, so we do not include a separate proof.

Proof of Theorem 1. For $n \geq 0$ and $d \in \{0, 1, \dots, p-1\}$, let m be an integer with $0 \leq m \leq pn + d$. There are two cases. If $(m \bmod p) \in \{0, 1, \dots, d\}$, then there is no carry from the 0th position when adding m to $pn + d - m$ in base p; therefore $\nu_p(\binom{pn+d}{m}) = \nu_p(\binom{n}{\lfloor m/p \rfloor})$ by Kummer's theorem. Otherwise, there is a carry from the 0th position, and $\nu_p(\binom{pn+d}{m}) = \nu_p(n) + \nu_p(\binom{n-1}{\lfloor m/p \rfloor}) + 1$ by Lemma 4 with $i = 0$ and $j = 1$. (Note that $n - 1 \geq 0$ here, since if $n = 0$ then $0 \leq m \leq d$ and we are in the first case.) Since $\{0, 1, \dots, d\}$ has $d + 1$ elements and its complement has $p - d - 1$ elements, we have

$$\sum_{m=0}^{pn+d} x^{\nu_p(\binom{pn+d}{m})} = (d+1) \sum_{c=0}^{n} x^{\nu_p(\binom{n}{c})} + (p-d-1) \sum_{c=0}^{n-1} x^{\nu_p(n)+\nu_p(\binom{n-1}{c})+1}$$

by comparing the coefficient of x^α on each side for each $\alpha \geq 0$. Using the definition of $T_p(n, x)$, this equation can be written

$$T_p(pn+d, x) = (d+1)\, T_p(n, x) + \begin{cases} 0 & \text{if } n = 0 \\ (p-d-1)\, x^{\nu_p(n)+1}\, T_p(n-1, x) & \text{if } n \geq 1. \end{cases} \tag{2}$$

Similarly, let m be an integer with $0 \leq m \leq pn + d - 1$. If $(m \bmod p) \in \{0, 1, \dots, d-1\}$, then there is no carry from the 0th position when adding m to $pn + d - 1 - m$ in base p, and $\nu_p(pn+d) + \nu_p(\binom{pn+d-1}{m}) = \nu_p(\binom{n}{\lfloor m/p \rfloor})$ by Lemma 4 with $i = 1$ and $j = 0$. (Note that $pn + d - 1 \geq 0$ here, since if $n = d = 0$ there is no m in the range $0 \leq m \leq pn + d - 1$.) Otherwise there is a carry from the 0th position, and $\nu_p(pn+d) + \nu_p(\binom{pn+d-1}{m}) = \nu_p(n) + \nu_p(\binom{n-1}{\lfloor m/p \rfloor}) + 1$ by Lemma 4 with $i = 1$ and $j = 1$. Therefore

$$\sum_{m=0}^{pn+d-1} x^{\nu_p(pn+d)+\nu_p(\binom{pn+d-1}{m})} = d \sum_{c=0}^{n} x^{\nu_p(\binom{n}{c})} + (p-d) \sum_{c=0}^{n-1} x^{\nu_p(n)+\nu_p(\binom{n-1}{c})+1}.$$

Multiplying both sides by x and rewriting in terms of $T_p(n, x)$ gives

$$\begin{cases} 0 & \text{if } pn+d = 0 \\ x^{\nu_p(pn+d)+1}\, T_p(pn+d-1, x) & \text{if } pn+d \geq 1 \end{cases} = d\, x\, T_p(n, x) + \begin{cases} 0 & \text{if } n = 0 \\ (p-d)\, x \cdot x^{\nu_p(n)+1}\, T_p(n-1, x) & \text{if } n \geq 1. \end{cases} \tag{3}$$

We combine Equations (2) and (3) into a matrix equation by defining

$$T_p'(n, x) := \begin{cases} 0 & \text{if } n = 0 \\ x^{\nu_p(n)+1}\, T_p(n-1, x) & \text{if } n \geq 1. \end{cases}$$

For each $n \geq 0$, we therefore have the recurrence

$$\begin{bmatrix} T_p(pn+d, x) \\ T_p'(pn+d, x) \end{bmatrix} = \begin{bmatrix} d+1 & p-d-1 \\ d\,x & (p-d)\,x \end{bmatrix} \begin{bmatrix} T_p(n, x) \\ T_p'(n, x) \end{bmatrix}, \tag{4}$$

which expresses $T_p(pn+d, x)$ and $T_p'(pn+d, x)$ in terms of $T_p(n, x)$ and $T_p'(n, x)$. The 2×2 coefficient matrix is $M_p(d)$. We have

$$\begin{bmatrix} T_p(0, x) \\ T_p'(0, x) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

for the vector of initial conditions, so the product

$$T_p(n,x) = \begin{bmatrix}1 & 0\end{bmatrix} M_p(n_0)\, M_p(n_1) \,\cdots\, M_p(n_\ell) \begin{bmatrix}1\\0\end{bmatrix}$$

now follows by writing n in base p. □

We obtain Fine's theorem as a special case by setting $x = 0$. The definition of $T_p'(n,x)$ implies $T_p'(n,0) = 0$, so Equation (4) becomes

$$\begin{bmatrix}F_p(pn+d)\\0\end{bmatrix} = \begin{bmatrix}d+1 & p-d-1\\0 & 0\end{bmatrix}\begin{bmatrix}F_p(n)\\0\end{bmatrix},$$

or simply

$$F_p(pn+d) = (d+1)\,F_p(n).$$

Equation (2) was previously proved by Spiegelhofer and Wallner [15, Equation (2.2)] using an infinite product and can also be obtained from a equation discovered by Carlitz [3]. In fact Carlitz came close to discovering Theorem 1. He knew that the coefficients of $T_p(n,x)$ and $T_p'(n,x)$ can be written in terms of each other. In his notation, let $\theta_\alpha(n)$ be the coefficient of x^α in $T_p(n,x)$, and let $\psi_{\alpha-1}(n-1)$ be the coefficient of x^α in $T_p'(n,x)$. Carlitz gave the recurrence

$$\theta_\alpha(pn+d) = (d+1)\theta_\alpha(n) + (p-d-1)\psi_{\alpha-1}(n-1)$$

$$\psi_\alpha(pn+d) = \begin{cases}(d+1)\theta_\alpha(n) + (p-d-1)\psi_{\alpha-1}(n-1) & \text{if } 0 \le d \le p-2\\ p\psi_{\alpha-1}(n) & \text{if } d = p-1.\end{cases}$$

The first of these equations is equivalent to Equation (2). But to get a matrix product for $T_p(n,x)$, one needs an equation expressing $\psi_\alpha(pn+d-1)$, not $\psi_\alpha(pn+d)$, in terms of θ and ψ. That equation is

$$\psi_\alpha(pn+d-1) = d\theta_\alpha(n) + (p-d)\psi_{\alpha-1}(n-1),$$

which is equivalent to Equation (3). Therefore $\psi_\alpha(n-1)$ (or, more precisely, $\psi_{\alpha-1}(n-1)$) seems to be more natural than Carlitz's $\psi_\alpha(n)$.

In addition to making use of $T_p(n,x)$, Spiegelhofer and Wallner [15] also utilized the normalized polynomial

$$\overline{T}_p(n,x) = \frac{1}{F_p(n)} T_p(n,x).$$

It follows from Theorem 1 that the sequence $\left(\overline{T}_p(n,x)\right)_{n\ge 0}$ is also p-regular, since it is generated by the normalized matrices $\frac{1}{d+1}M_p(d)$.

We briefly investigate $T_p(n,x)$ evaluated at particular values of x. We have already mentioned $T_p(n,0) = F_p(n)$. It is clear that $T_p(n,1) = n+1$. When $p = 2$ and $x = -1$, we obtain a version of A106407 [13] with different signs. Let $t(n)_{n\ge 0}$ be the Thue–Morse sequence, and let $S(n,x)$ be the nth Stern polynomial, defined by

$$S(n,x) = \begin{bmatrix}1 & 0\end{bmatrix} A(n_0)\, A(n_1) \,\cdots\, A(n_\ell) \begin{bmatrix}0\\1\end{bmatrix},$$

where

$$A(0) = \begin{bmatrix}x & 0\\1 & 1\end{bmatrix}, \quad A(1) = \begin{bmatrix}1 & 1\\0 & x\end{bmatrix},$$

and as before $n_\ell \cdots n_1 n_0$ is the standard base-2 representation of n.

Theorem 2. *For each $n \geq 0$, we have $T_2(n, -1) = (-1)^{t(n)} S(n+1, -2)$.*

Proof. Define the rank of a regular sequence to be the dimension of the corresponding vector space. We bound the rank of $T_2(n, -1) - (-1)^{t(n)} S(n+1, -2)$ using closure properties of 2-regular sequences [1, Theorems 2.5 and 2.6]. Since the rank of $S(n, x)$ is 2, the rank of $S(n+1, -2)$ is at most 2. The rank of $(-1)^{t(n)}$ is 1. If two sequences have ranks r_1 and r_2, then their sum and product have ranks at most $r_1 + r_2$ and $r_1 r_2$. Therefore $T_2(n, -1) - (-1)^{t(n)} S(n+1, -2)$ has rank at most 4, so to show that it is the 0 sequence it suffices to check 4 values of n. ☐

It would be interesting to know if there is a combinatorial interpretation of this identity.

2. Multinomial coefficients

In this section we generalize Theorem 1 to multinomial coefficients. For a k-tuple $\mathbf{m} = (m_1, m_2, \ldots, m_k)$ of non-negative integers, define

$$\operatorname{total} \mathbf{m} := m_1 + m_2 + \cdots + m_k$$

and

$$\operatorname{mult} \mathbf{m} := \frac{(\operatorname{total} \mathbf{m})!}{m_1! \, m_2! \, \cdots \, m_k!}.$$

Specifically, we count k-tuples $\mathbf{m}$ with a fixed total, according to the p-adic valuation $\nu_p(\operatorname{mult} \mathbf{m})$. The result is a matrix product as in Theorem 1. The matrices are $k \times k$ matrices with coefficients from the following sequence.

Let $c_{p,k}(n)$ be the number of k-tuples $\mathbf{d} \in \{0, 1, \ldots, p-1\}^k$ with $\operatorname{total} \mathbf{d} = n$. Note that $c_{p,k}(n) = 0$ for $n < 0$. For example, let $p = 5$ and $k = 3$; the values of $c_{5,3}(n)$ for $-k+1 \leq n \leq pk - 1$ are

$$0, 0, 1, 3, 6, 10, 15, 18, 19, 18, 15, 10, 6, 3, 1, 0, 0.$$

For $k \geq 1$, every tuple counted by $c_{p,k}(n)$ has a last entry d; removing that entry gives a $(k-1)$-tuple with total $n - d$, so we have the recurrence

$$c_{p,k}(n) = \sum_{d=0}^{p-1} c_{p,k-1}(n-d).$$

Therefore $c_{p,k}(n)$ is an entry in the Pascal-like triangle generated by adding p entries on the previous row. For $p = 5$ this triangle begins as follows.

1																
1	1	1	1	1												
1	2	3	4	5	4	3	2	1								
1	3	6	10	15	18	19	18	15	10	6	3	1				
1	4	10	20	35	52	68	80	85	80	68	52	35	20	10	4	1

The entry $c_{p,k}(n)$ is also the coefficient of x^n in $(1 + x + x^2 + \cdots + x^{p-1})^k$.

For each $d \in \{0, 1, \dots, p-1\}$, let $M_{p,k}(d)$ be the $k \times k$ matrix whose (i, j) entry is $c_{p,k}(p\,(j-1) + d - (i-1))\, x^{i-1}$. The matrices $M_{5,3}(0), \dots, M_{5,3}(4)$ are

$$\begin{bmatrix} 1 & 18 & 6 \\ 0 & 15x & 10x \\ 0 & 10x^2 & 15x^2 \end{bmatrix}, \quad \begin{bmatrix} 3 & 19 & 3 \\ x & 18x & 6x \\ 0 & 15x^2 & 10x^2 \end{bmatrix}, \quad \begin{bmatrix} 6 & 18 & 1 \\ 3x & 19x & 3x \\ x^2 & 18x^2 & 6x^2 \end{bmatrix},$$
$$\begin{bmatrix} 10 & 15 & 0 \\ 6x & 18x & x \\ 3x^2 & 19x^2 & 3x^2 \end{bmatrix}, \quad \begin{bmatrix} 15 & 10 & 0 \\ 10x & 15x & 0 \\ 6x^2 & 18x^2 & x^2 \end{bmatrix}.$$

For $k = 2$, the matrix $M_{p,2}(d)$ is exactly the matrix $M_p(d)$ in Section 1.

We use $\mathbb{N}$ to denote the set of non-negative integers. Let

$$T_{p,k}(n, x) = \sum_{\substack{\mathbf{m} \in \mathbb{N}^k \\ \operatorname{total} \mathbf{m} = n}} x^{\nu_p(\operatorname{mult} \mathbf{m})}.$$

Theorem 3. *Let p be a prime, let $k \geq 1$, and let $n \geq 0$. Let $e = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix}$ be the first standard basis vector in $\mathbb{Z}^k$. Let $n_\ell \cdots n_1 n_0$ be the standard base-p representation of n. Then*

$$T_{p,k}(n, x) = e\, M_{p,k}(n_0)\, M_{p,k}(n_1) \,\cdots\, M_{p,k}(n_\ell)\, e^\top.$$

By setting $x = 0$ we recover a generalization of Fine's theorem for the number of multinomial coefficients not divisible by p; the top left entry of $M_{p,k}(d)$ is $c_{p,k}(d) = \binom{d+k-1}{k-1}$, so

$$T_{p,k}(n, 0) = \binom{n_0 + k - 1}{k - 1} \binom{n_1 + k - 1}{k - 1} \cdots \binom{n_\ell + k - 1}{k - 1}.$$

The proof of Theorem 3 uses the following generalization of Kummer's theorem. Recall that $\sigma_p(m)$ denotes the sum of the base-p digits of m. We write $\sigma_p(\mathbf{m})$, $\lfloor \mathbf{m}/p \rfloor$, and $\mathbf{m} \bmod p$ for the tuples obtained by applying these functions termwise to the entries of $\mathbf{m}$.

Kummer's theorem for multinomial coefficients. *Let p be a prime, and let $\mathbf{m} \in \mathbb{N}^k$ for some $k \geq 0$. Then*

$$\nu_p(\operatorname{mult} \mathbf{m}) = \frac{\operatorname{total} \sigma_p(\mathbf{m}) - \sigma_p(\operatorname{total} \mathbf{m})}{p - 1}.$$

This generalized version of Kummer's theorem also follows from Legendre's formula. The following lemma gives the relationship between $\nu_p(\operatorname{mult} \mathbf{m})$ and $\nu_p(\operatorname{mult} \lfloor \mathbf{m}/p \rfloor)$.

Lemma 4. *Let p be a prime, $k \geq 1$, $n \geq 0$, $d \in \{0, 1, \dots, p-1\}$, and $0 \leq i \leq k - 1$. Let $\mathbf{m} \in \mathbb{N}^k$ with $\operatorname{total} \mathbf{m} = pn + d - i$. Let $j = n - \operatorname{total} \lfloor \mathbf{m}/p \rfloor$. Then $\operatorname{total}(\mathbf{m} \bmod p) = pj + d - i$, $0 \leq j \leq k - 1$, and*

$$\nu_p\!\left(\frac{(pn + d)!}{(pn + d - i)!}\right) + \nu_p(\operatorname{mult} \mathbf{m}) = \nu_p\!\left(\frac{n!}{(n - j)!}\right) + \nu_p(\operatorname{mult} \lfloor \mathbf{m}/p \rfloor) + j.$$

Proof. Let $\mathbf{c} = \lfloor \mathbf{m}/p \rfloor$ and $\mathbf{d} = (\mathbf{m} \bmod p) \in \{0, 1, \dots, p-1\}^k$, so that $\mathbf{m} = p\,\mathbf{c} + \mathbf{d}$. We have

$$\begin{aligned}\operatorname{total} \sigma_p(\mathbf{m}) - \operatorname{total} \sigma_p(\mathbf{c}) &= \operatorname{total} \mathbf{d} \\ &= \operatorname{total} \mathbf{m} - p \operatorname{total} \mathbf{c} \\ &= pj + d - i \\ &= pj + \sigma_p(pn + d) - \sigma_p(n) - i.\end{aligned}$$

In particular, $\operatorname{total} \mathbf{d} = pj + d - i$, as claimed; solving this equation for j gives

$$j = \frac{-d + i + \operatorname{total} \mathbf{d}}{p},$$

which implies the bounds

$$-1 + \frac{1}{p} = \frac{-(p-1) + 0 + 0}{p} \leq j \leq \frac{0 + (k-1) + (p-1)k}{p} = k - \frac{1}{p}.$$

Since j is an integer, this implies $0 \leq j \leq k - 1$.

The generalized Kummer theorem gives

$$\begin{aligned}&(p-1)\left(\nu_p(\operatorname{mult} \mathbf{m}) - \nu_p(\operatorname{mult} \mathbf{c})\right) \\ &\quad = \left(\operatorname{total} \sigma_p(\mathbf{m}) - \sigma_p(\operatorname{total} \mathbf{m})\right) - \left(\operatorname{total} \sigma_p(\mathbf{c}) - \sigma_p(\operatorname{total} \mathbf{c})\right) \\ &\quad = \operatorname{total} \sigma_p(\mathbf{m}) - \sigma_p(pn + d - i) - \operatorname{total} \sigma_p(\mathbf{c}) + \sigma_p(n - j).\end{aligned}$$

Since we established

$$\operatorname{total} \sigma_p(\mathbf{m}) - \operatorname{total} \sigma_p(\mathbf{c}) = pj + \sigma_p(pn + d) - \sigma_p(n) - i$$

above, we can write

$$\begin{aligned}&(p-1)\left(\nu_p(\operatorname{mult} \mathbf{m}) - \nu_p(\operatorname{mult} \mathbf{c})\right) \\ &\quad = pj + \sigma_p(pn + d) - \sigma_p(n) - i - \sigma_p(pn + d - i) + \sigma_p(n - j) \\ &\quad = \left(\sigma_p(pn + d) - i - \sigma_p(pn + d - i)\right) + \left(-\sigma_p(n) + j + \sigma_p(n - j)\right) + (p-1)j \\ &\quad = (p-1)\left(-\nu_p\left(\frac{(pn+d)!}{(pn+d-i)!}\right) + \nu_p\left(\frac{n!}{(n-j)!}\right) + j\right),\end{aligned}$$

where the last equality uses Legendre's formula. Dividing by $p - 1$ and rearranging terms gives the desired equation. □

We are now ready to prove the main theorem of this section.

Proof of Theorem 3. Let $d \in \{0, 1, \dots, p-1\}$, $0 \leq i \leq k - 1$, and $\alpha \geq 0$. We claim that the map β defined by

$$\beta(\mathbf{m}) := (\lfloor \mathbf{m}/p \rfloor, \mathbf{m} \bmod p)$$

is a bijection from the set

$$A = \left\{\mathbf{m} \in \mathbb{N}^k : \operatorname{total} \mathbf{m} = pn + d - i \text{ and } \nu_p(\operatorname{mult} \mathbf{m}) = \alpha - \nu_p\left(\frac{(pn+d)!}{(pn+d-i)!}\right)\right\}$$

to the set

$$B = \bigcup_{j=0}^{k-1} \left(\left\{ \mathbf{c} \in \mathbb{N}^k : \operatorname{total} \mathbf{c} = n - j \text{ and } \nu_p(\operatorname{mult} \mathbf{c}) = \alpha - \nu_p\!\left(\frac{n!}{(n-j)!}\right) - j \right\} \right.$$
$$\left. \times \left\{ \mathbf{d} \in \{0, 1, \dots, p-1\}^k : \operatorname{total} \mathbf{d} = pj + d - i \right\} \right).$$

Note that the k sets in the union comprising B are disjoint, since each tuple $\mathbf{d}$ occurs for at most one index j. Lemma 4 implies that if $\mathbf{m} \in A$ then $\beta(\mathbf{m}) \in B$.

Clearly β is injective, since $\beta(\mathbf{m})$ preserves all the digits of the entries of $\mathbf{m}$. It is also clear that β is surjective, since a given pair $(\mathbf{c}, \mathbf{d}) \in B$ is the image of $p\,\mathbf{c} + \mathbf{d} \in A$. Therefore $\beta : A \to B$ is a bijection.

Consider the polynomial

$$\sum_{\substack{\mathbf{m} \in \mathbb{N}^k \\ \operatorname{total} \mathbf{m} = pn+d-i}} x^{\nu_p\left(\frac{(pn+d)!}{(pn+d-i)!}\right) + \nu_p(\operatorname{mult} \mathbf{m})}.$$

The coefficient of x^α in this polynomial is $|A|$. On the other hand, the coefficient of x^α in the polynomial

$$\sum_{j=0}^{k-1} c_{p,k}(pj + d - i) \sum_{\substack{\mathbf{c} \in \mathbb{N}^k \\ \operatorname{total} \mathbf{c} = n-j}} x^{\nu_p\left(\frac{n!}{(n-j)!}\right) + \nu_p(\operatorname{mult} \mathbf{c}) + j}$$

is $|B|$, since $c_{p,k}(pj + d - i)$ is the number of k-tuples $\mathbf{d} \in \{0, 1, \dots, p-1\}^k$ with $\operatorname{total} \mathbf{d} = pj + d - i$. Since A and B are in bijection for each $\alpha \geq 0$, these two polynomials are equal. Multiplying both polynomials by x^i and rewriting in terms of $T_{p,k}(n, x)$ gives

$$\begin{cases} 0 & \text{if } 0 \leq pn + d \leq i - 1 \\ x^{\nu_p\left(\frac{(pn+d)!}{(pn+d-i)!}\right) + i}\, T_{p,k}(pn + d - i, x) & \text{if } pn + d \geq i \end{cases}$$
$$= \sum_{j=0}^{k-1} \begin{cases} 0 & \text{if } 0 \leq n \leq j - 1 \\ c_{p,k}(pj + d - i)\, x^i \cdot x^{\nu_p\left(\frac{n!}{(n-j)!}\right) + j}\, T_{p,k}(n - j, x) & \text{if } n \geq j. \end{cases}$$

For each i in the range $0 \leq i \leq k - 1$, define

$$T_{p,k,i}(n, x) := \begin{cases} 0 & \text{if } 0 \leq n \leq i - 1 \\ x^{\nu_p\left(\frac{n!}{(n-i)!}\right) + i}\, T_{p,k}(n - i, x) & \text{if } n \geq i. \end{cases}$$

Note that $T_{p,k,0}(n, x) = T_{p,k}(n, x)$. For $n \geq 0$, we therefore have

$$T_{p,k,i}(pn + d, x) = \sum_{j=0}^{k-1} c_{p,k}(pj + d - i)\, x^i\, T_{p,k,j}(n, x).$$

For each i, this equation gives a recurrence for $T_{p,k,i}(pn+d, x)$ in terms of $T_{p,k,j}(n, x)$ for $0 \leq j \leq k - 1$. The coefficients of this recurrence are the entries of the matrix $M_{p,k}(d)$. It follows from the definition of $T_{p,k,i}(n, x)$ that $T_{p,k,0}(0, x) = 1$ and $T_{p,k,i}(0, x) = 0$ for $1 \leq i \leq k - 1$. Therefore the vector of initial conditions is $\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix}^\top$, and the matrix product follows. $\square$

A natural question suggested by this paper is whether various generalizations of binomial coefficients (Fibonomial coefficients, q-binomial coefficients, Carlitz binomial coefficients, etc.) and multinomial coefficients have results that are analogous to Theorems 1 and 3.

References

[1] Jean-Paul Allouche and Jeffrey Shallit, The ring of k-regular sequences, *Theoretical Computer Science* **98** (1992) 163–197.

[2] Guy Barat and Peter J. Grabner, Distribution of binomial coefficients and digital functions, *Journal of the London Mathematical Society* **64** (2001) 523–547.

[3] Leonard Carlitz, The number of binomial coefficients divisible by a fixed power of a prime, *Rendiconti del Circolo Matematico di Palermo* **16** (1967) 299–320.

[4] Kenneth Davis and William Webb, Pascal's triangle modulo 4, *The Fibonacci Quarterly* **29** (1989) 79–83.

[5] William Everett, Subprime factorization and the numbers of binomial coefficients exactly divided by powers of a prime, *Integers* **11** (2011) #A63.

[6] Nathan Fine, Binomial coefficients modulo a prime, *The American Mathematical Monthly* **54** (1947) 589–592.

[7] James W. L. Glaisher, On the residue of a binomial-theorem coefficient with respect to a prime modulus, *Quarterly Journal of Pure and Applied Mathematics* **30** (1899) 150–156.

[8] Fred T. Howard, The number of binomial coefficients divisible by a fixed power of 2, *Proceedings of the American Mathematical Society* **29** (1971) 236–242.

[9] James Huard, Blair Spearman, and Kenneth Williams, Pascal's triangle $\pmod 9$, *Acta Arithmetica* **78** (1997) 331–349.

[10] James Huard, Blair Spearman, and Kenneth Williams, On Pascal's triangle modulo p^2, *Colloquium Mathematicum* **74** (1997) 157–165.

[11] James Huard, Blair Spearman, and Kenneth Williams, Pascal's triangle $\pmod 8$, *European Journal of Combinatorics* **19** (1998) 45–62.

[12] Ernst Kummer, Über die Ergänzungssätze zu den allgemeinen Reciprocitätsgesetzen, *Journal für die reine und angewandte Mathematik* **44** (1852) 93–146.

[13] The OEIS Foundation, The On-Line Encyclopedia of Integer Sequences, `http://oeis.org`.

[14] Eric Rowland, The number of nonzero binomial coefficients modulo p^α, *Journal of Combinatorics and Number Theory* **3** (2011) 15–25.

[15] Lukas Spiegelhofer and Michael Wallner, Divisibility of binomial coefficients by powers of primes, `https://arxiv.org/abs/1604.07089`.

[16] William Webb, The number of binomial coefficients in residue classes modulo p and p^2, *Colloquium Mathematicum* **60/61** (1990) 275–280.

Department of Mathematics, Hofstra University, Hempstead, NY, USA