to appear in the
Journal of the London Mathematical Society

DISTRIBUTION OF BINOMIAL COEFFICIENTS AND DIGITAL FUNCTIONS

GUY BARAT AND PETER J. GRABNER

Abstract. The distribution of binomial coefficients in residue classes modulo prime powers and with respect to the p-adic valuation is studied. For this purpose general asymptotic results for arithmetic functions depending on blocks of digits with respect to q-ary expansions are established.

1. Introduction

The main purpose of this paper is the study of the asymptotic distribution of binomial coefficients modulo prime powers. D. Singmaster [**36**] has proved that "any integer divides almost all binomial coefficients" in the following sense: for any integer $m \geq 2$ we have

$$\lim_{N\to\infty} \frac{2}{N(N+1)} \# \left\{ (k,n) : 0 \leq k \leq n < N \text{ and } m \not| \binom{n}{k} \right\} = 0.$$

For prime m this result could be refined in [**3, 17**]. There a precise asymptotic formula for the quantity above was given, furthermore it was established that the distribution in the non-zero residue classes is asymptotically uniform.

Already in [**3**] it turned out that the key point to describe this behaviour was a result of [**19**] on so-called q-multiplicative functions, that is arithmetical functions defined by the digital expansion in base q. Digital functions, first of all the sum-of-digits function, have been studied from various points of view since the 1940's. Following [**4, 11, 18, 19, 34, 42**] the emphasis in this paper will be on the asymptotic behaviour of the summatory functions of such arithmetic functions, and especially on periodicity phenomena which occur in this context. We also note that distribution properties of digital functions have been studied intensively, cf. [**10, 13, 15**]. We will introduce and study various kinds of arithmetic functions which depend on blocks of digits.

In Section 2 we give asymptotic expansions for the summatory functions of "block-multiplicative" and mixed "block-multiplicative" and "block-additive" functions.

It has been recognized in the nineteenth century that the p-adic expansions of n and k yield information on the p-valuation of binomial coefficient $\binom{n}{k}$ and on its value modulo p.

In [**28**] (among many other things) E. E. Kummer expressed the p-valuation of $\binom{n}{k}$ in terms of the p-adic digits of n and k. This result has been rediscovered several times and also generalized to multinomial coefficients (cf. [**27, 37**]). It is the main combinatorial

Date: June 6, 2002.
1991 *Mathematics Subject Classification.* Primary: 11B65; Secondary: 11A63, 11N37.
Key words and phrases. binomial coefficients, digital functions.
The authors are supported by the START project Y96-MAT of the Austrian Science Fund.

tool for the study of the number of binomial coefficients in a given row of Pascal's triangle divisible by a given power of a prime. Partial results and exact formulæ in this direction have been obtained in [**6, 25, 26**].

The second historic ingredient that we use is É. Lucas' congruence for binomial coefficients modulo primes. This congruence has been generalized in various ways since then, for instance by H. Anton, L. Stickelberger, and K. Hensel (for a detailed history we refer to Dickson's book [**12**]). The last achievement, which gives a congruence similar to Lucas' modulo prime powers, is independently due to K. S. Davis and W. Webb [**8**] and to A. Granville [**21**].

In Section 3 we use the asymptotic techniques introduced in Section 2 to describe the asymptotic behaviour of the number of such binomial coefficients up to a given row. For this purpose we recall the notation $p^j \| M$ to express that p^j is the highest power of p that divides M (or equivalently that the p-adic valuation of M is j). We define

$$\vartheta_j(n) = \# \left\{ k : 0 \le k \le n \text{ and } p^j \| \binom{n}{k} \right\} \tag{1.1}$$

and prove an asymptotic formula which has

$$\sum_{n<N} \vartheta_j(n) \sim \frac{1}{j!} \left(\frac{p-1}{p+1} \right)^{2j} (\log_p N)^j \sum_{n<N} \vartheta_0(n)$$

as a consequence. The main step in the proof is getting an expression of ϑ_j as a polynomial in block digital functions of the type studied in Section 2.

In Section 4 we use Granville's congruence, Dirichlet characters and bivariate block multiplicative functions to derive a distribution result for binomial coefficients with given valuation modulo prime powers:

$$\# \left\{ (k,n) : 0 \le k \le n < N,\ p^j \| \binom{n}{k},\ \text{and}\ \ p^{-j} \binom{n}{k} \equiv a \bmod p^\ell \right\} \sim \frac{1}{\phi(p^\ell)} \sum_{n<N} \vartheta_j(n).$$

Furthermore, we show that the p-free parts of the binomial coefficients are uniformly distributed in $\mathbb{Z}_p^*$. (We recall that the p-free part of a non-zero integer m is $p^{-j}m$ where $p^j \| m$.) We also show that the distribution of their p-valuations in residue classes is uniform and independent of the distribution of the p-free parts.

In Section 5 we give indications how to generalize our ideas to multinomial coefficients. Special formulæ for the distribution of binomial coefficients in residue classes modulo primes and prime powers have been given in [**24, 43**]. Related studies can also be found in [**5**]. Furthermore, there exists a vast literature on special congruences involving binomial coefficients; for this we refer to Granville's survey [**21**].

Finally, we mention several papers that deal with properties of binomial coefficients modulo primes or prime powers from the point of view of the theory of cellular automata [**2, 20, 29, 44**]. A notion of complexity of Pascal's triangle modulo any integer is studied in [**1**].

2. Block-multiplicative functions

We consider digital expansions of the positive integers with respect to a base $q > 1$. For a positive integer ℓ, an arithmetic function $f : \mathbb{N} \to \mathbb{C}^*$ is called *ℓ-block-multiplicative*, if there exists a function $g : \{0, \dots, q-1\}^\ell \to \mathbb{C}^*$ with $g(0, 0, \dots, 0) = 1$ such that

$$f\left(\sum_{k=0}^{K} \varepsilon_k q^k\right) = \prod_{k=0}^{K} g(\varepsilon_k, \varepsilon_{k+1}, \dots, \varepsilon_{k+\ell-1}), \tag{2.1}$$

with $\varepsilon_k = 0$ for $k > K$.

Remark 1. Additive functions with respect to q-adic digital expansions can also be viewed in the context of probabilistic number theory (cf. for instance [**10, 13**]). For the introduction of the according Kubilius models we refer to [**33**]. These models make use of the fact that the q-adic digits of integers behave asymptotically like independent random variables. We remark here that probabilistic methods would only yield the main terms in the asymptotic expansions given later, but would not reveal the periodic oscillations.

Remark 2. The functions satisfying (2.1) are the multiplicative analogue of digital functions ("fonctions digitales") studied by E. Cateland [**7**], namely

$$f\left(\sum_{k=0}^{K} \varepsilon_k q^k\right) = \sum_{k=0}^{K} h(\varepsilon_k, \varepsilon_{k+1}, \dots, \varepsilon_{k+\ell-1}) \quad \text{with } h(0, 0, \dots, 0) = 0. \tag{2.2}$$

We will call these functions *ℓ-block-additive*. For instance, 1-block-multiplicative functions are completely q-multiplicative functions in the sense studied in [**19**].

Remark 3. We will allow ourselves to change freely between a number $n = \sum_{k=0}^{K} \varepsilon_k q^k$, its corresponding finite sequence of q-ary digits $(\varepsilon_0, \varepsilon_1, \dots, \varepsilon_K)$ and the infinite sequence $(\varepsilon_0, \varepsilon_1, \dots, \varepsilon_K, 0, 0, \dots)$. Therefore we write the digits from left to right.

The study of q-additive and q-multiplicative functions f is based on finding recurrence relations for $\sum_{n<q^k} f(n)$ and then $\sum_{n<N} f(n)$ for arbitrary N. In the case of block-additive and block-multiplicative functions these recurrences get more complicated and can be written as matrix relations. For a similar approach we refer to [**13**].

In the following $\mathcal{B}_s$ will denote the set of all digital blocks of length s. For $B \in \mathcal{B}_s$ an expression "Bn" will indicate the concatenation of the block B with the digits of n (or the corresponding positive integer). This convention will also give us the possibility to write $B < r$ for a block B and a positive integer r, if the number corresponding to B is smaller than r. This also gives a natural ordering on $\mathcal{B}_s$. Furthermore, we denote by $|B|$ the length of the block B.

For a given ℓ-block-multiplicative function f we define the summatory function

$$F(N) = \sum_{n<N} f(n).$$

For fixed s we introduce the functions

$$\begin{aligned}\mathbf{f}(n) &= (f(Bn))_{B\in\mathcal{B}_s},\\ F_B(N) &= \sum_{n<N} f(Bn)\\ \mathbf{F}(N) &= (F_B(N))_{B\in\mathcal{B}_s}\end{aligned}$$

Lemma 1. *Let f be an ℓ-block-multiplicative function. For fixed $s \geq \max(1, \ell-1)$ and every non-negative integer r the following equations hold*

$$f(B)f(ABC) = f(AB)f(BC) \quad \text{for } B \in \mathcal{B}_s, \text{ and } A, C \text{ blocks of arbitrary length,} \tag{2.3}$$

$$F(BN) = \sum_{C\in\mathcal{B}_s} F_C(N) + \sum_{C<B} f(CN) \tag{2.4}$$

$$\mathbf{F}(q^{r+s}) = U\mathbf{F}(q^r), \tag{2.5}$$

where the matrix U is given by

$$u_{B,C} = \begin{cases} \frac{f(BC)}{f(C)} & \text{for } f(C)\neq 0\\ 0 & \text{otherwise,}\end{cases}$$

$$\begin{aligned}\mathbf{F}(B_0B_1B_2\dots B_k) =& U^kU_{B_k}\mathbf{f}(0) + U^{k-1}U_{B_{k-1}}\mathbf{f}(B_k) + U^{k-2}U_{B_{k-2}}\mathbf{f}(B_{k-1}B_k) + \cdots\\ &+ UU_{B_1}\mathbf{f}(B_2\dots B_k) + U_{B_0}\mathbf{f}(B_1\dots B_k), \quad B_i\in\mathcal{B}_s \quad \text{for } i=0,1,\dots,k\end{aligned} \tag{2.6}$$

where $U_D = (u^D_{B,C})$ is a matrix given by

$$u^D_{B,C} = \begin{cases} \frac{f(BC)}{f(C)} & \text{for } C<D \text{ and } f(C)\neq 0\\ 0 & \text{otherwise.}\end{cases}$$

Proof. Equations (2.3) and (2.4) are clear. For (2.5) we observe that $f(BCn) = 0$ if $f(C) = 0$ and write

$$\begin{aligned}F_B(q^{r+s}) &= \sum_{C\in\mathcal{B}_s}\sum_{n<q^r} f(BCn)\\ &= \sum_{\substack{C\in\mathcal{B}_s\\ f(C)\neq 0}}\sum_{n<q^r}\frac{f(BC)}{f(C)}f(Cn) = \sum_{C\in\mathcal{B}_s}\frac{f(BC)}{f(C)}F_C(q^r).\end{aligned}$$

For equation (2.6) we first derive a recursion formula for $\mathbf{F}(BN)$ for $B\in\mathcal{B}_s$ and $N\in\mathbb{N}$. Indeed,

$$F_C(BN) = \sum_{D\in\mathcal{B}_s}\sum_{n<N} f(CDn) + \sum_{D<B} f(CDN) = {\sum_{D\in\mathcal{B}_s}}' \frac{f(CD)}{f(D)}F_D(N) + {\sum_{D<B}}' \frac{f(CD)}{f(D)}f(DN),$$

where $\sum'$ indicates that terms with $f(D) = 0$ are omitted. Writing this equation in matrix form we obtain

$$\mathbf{F}(BN) = U\mathbf{F}(N) + U_B\mathbf{f}(N).$$

In particular, inserting $N = 0$ gives $\mathbf{F}(B) = U_B\mathbf{f}(0)$. Then (2.6) follows by induction. □

Before deriving an asymptotic formula for $F(N)$ we fix some terminology and notations:

- For $g : \{0, \dots, q-1\}^\ell \mapsto \mathbb{C}$ we define $\|g\|_\infty = \max_{B \in \mathcal{B}_\ell} |g(B)|$.
- If $(a_1, \dots, a_t) \in \mathbb{C}^t$, $M = \operatorname{diag}(a_1, \dots, a_t)$ denotes the diagonal matrix with $m_{ii} = a_i$. More generally, if $A_1, \dots, A_t$ are square matrices, $\operatorname{diag}(A_1, \dots, A_t)$ is the block-diagonal matrix with diagonal blocks $A_1, \dots, A_t$.
- For $\lambda \in \mathbb{C}$ and an integer $k \geq 1$ we introduce the nilpotent $k \times k$-matrix

$$E_k = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \cdots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}$$

and the Jordan matrix $J_k(\lambda) = \lambda I_k + E_k$. Furthermore, real powers of $J_k(\lambda)$ can be defined as

$$[J_k(\lambda)]^x = \sum_{j=0}^{k-1} \binom{x}{j} \lambda^{x-j} E^j. \tag{2.7}$$

These matrices satisfy $[J_k(\lambda)]^x \cdot [J_k(\lambda)]^y = [J_k(\lambda)]^{x+y}$.

- For any square matrix U we denote by $\operatorname{spec}(U)$ the set of its eigenvalues.
- For any vector $\mathbf{x} \in \mathbb{C}^t$, $\|\mathbf{x}\|$ denotes its maximum norm. For t-dimensional square matrices U, $\|U\|$ denotes the corresponding operator norm. Thus we have $\|MN\| \leq \|M\| \cdot \|N\|$. Furthermore,

$$(\forall \lambda \in \operatorname{spec}(U) : |\lambda| < \nu) \Longrightarrow \|U^n\| = o(\nu^n) \text{ for } n \to \infty. \tag{2.8}$$

Moreover, for $B \in \mathcal{B}_t$ we have

$$\|\mathbf{f}(B)\| \leq \|g\|_\infty^{t+s}. \tag{2.9}$$

Theorem 1. *Let f be an ℓ-block-multiplicative function associated to a function g as in* (2.1) *and assume that there exists an $s \geq \max(1, \ell-1)$ such that the corresponding matrix U defined in* (2.5) *satisfies the following condition: there exists a unique dominating eigenvalue λ of U with $|\lambda| > \|g\|_\infty^s$.*
Then there exist continuous periodic functions ψ_k ($k = 0, \dots, K-1$) of period s such that

$$\sum_{n<N} f(n) = N^\delta e^{i\alpha \log_q N} \sum_{k=0}^{K-1} (\log_q N)^k \psi_k(\log_q N) + o(N^\mu) \tag{2.10}$$

for any $\mu > \max\left(\log_q \|g\|_\infty, \frac{1}{s} \log_q \max_{\nu \in \operatorname{spec}(U) \setminus \{\lambda\}} |\nu|\right)$, where $\delta = \frac{1}{s} \log_q |\lambda|$, $\alpha = \frac{1}{s} \arg \lambda$, and K is the dimension of the largest Jordan-block associated to λ.

Proof. First we transform the matrix U into its Jordan normal form

$$U = T \operatorname{diag}(J_{k_1}(\lambda), \dots, J_{k_p}(\lambda), A_1, \dots, A_d) T^{-1},$$

where the A_j's are Jordan blocks with eigenvalues strictly smaller in modulus than $|\lambda|$. Note that $\lambda \neq 0$ and define

$$V = T \operatorname{diag}(J_{k_1}^{-1}, \ldots, J_{k_p}^{-1}, 0, \ldots, 0)T^{-1}$$

For convenience we set the unusual convention $V^0 = T \operatorname{diag}(\underbrace{1, \ldots, 1}_{\kappa}, 0, \ldots, 0)T^{-1}$ with $\kappa = k_1 + \cdots + k_p$.

We now introduce the function

$$\varphi(0.\varepsilon_1\varepsilon_2\ldots) = \varphi(0.B_0B_1\ldots) = \sum_{k=0}^{\infty} V^k U_{\overline{B_k}} \mathbf{f}\left(\overline{B_{k-1}}\ldots\overline{B_0}\right), \tag{2.11}$$

where the B_k's are blocks of length s and $\overline{B} = (b_s, \ldots, b_1)$ denotes the reversion of $B = (b_1, \ldots, b_s)$. We choose a real number ξ such that $\max\left(\|g\|_\infty^s, \max_{\nu \in \operatorname{spec}(U)\setminus\{\lambda\}} |\nu|\right) < \xi < |\lambda|$. We first note that the series (2.11) converges normally, since (2.8) and (2.9) yield

$$\begin{aligned} \left\|V^k U_{\overline{B_k}} \mathbf{f}\left(\overline{B_{k-1}}\ldots\overline{B_0}\right)\right\| &\leq \left\|V^k\right\| \cdot \left\|U_{\overline{B_k}}\right\| \cdot \left\|\mathbf{f}\left(\overline{B_{k-1}}\ldots\overline{B_0}\right)\right\| \\ &\leq \left\|U_{\overline{B_k}}\right\| \cdot \|g\|_\infty^{(k+1)s} o(\xi^{-k}) = o\left(\frac{\|g\|_\infty^{ks}}{\xi^k}\right). \end{aligned} \tag{2.12}$$

Thus φ can be seen as a continuous function on the product space $\{0, \ldots, q-1\}^{\mathbb{N}}$. We now have to prove that φ descends to a well-defined function on $[0, 1]$. For this purpose it has to be verified that

$$\varphi(0.\varepsilon_1\ldots\varepsilon_k 0^\infty) = \varphi(0.\varepsilon_1\ldots\varepsilon_{k-1}(\varepsilon_k - 1)(q-1)^\infty) \quad \text{(for any } k \geq 1,\ \varepsilon_k \neq 0).$$

We rewrite the above strings in terms of blocks of length s: $\varepsilon_1\varepsilon_2\ldots\varepsilon_k 0^\infty = B_0B_1\ldots B_h(0^s)^\infty$ and $\varepsilon_1\ldots\varepsilon_{k-1}(\varepsilon_k - 1)(q-1)^\infty = B_0B_1\ldots B_{h-1}\tilde{B}_h(q^s-1)^\infty$, where $\overline{B_h}$ is the successor of $\overline{\tilde{B}_h}$ in the natural ordering on $\mathcal{B}_s$. Then we have by (2.11)

$$\varphi\left(0.B_0B_1\ldots B_h(0^s)^\infty\right) = \sum_{k=0}^{h-1} V^k U_{\overline{B_k}} \mathbf{f}(\overline{B_{k-1}}\ldots\overline{B_0}) + V^h U_{\overline{B_h}} \mathbf{f}(\overline{B_{h-1}}\ldots\overline{B_0}).$$

On the other hand we have

$$\begin{aligned} \varphi\left(0.B_0B_1\ldots B_{h-1}\tilde{B}_h(q^s-1)^\infty\right) = &\sum_{k=0}^{h-1} V^k U_{\overline{B_k}} \mathbf{f}(\overline{B_{k-1}}\ldots\overline{B_0}) + V^h U_{\overline{\tilde{B}_h}} \mathbf{f}(\overline{B_{h-1}}\ldots\overline{B_0}) + \\ &\sum_{k=h+1}^{\infty} V^k U_{(q^s-1)} \mathbf{f}((q^s-1)^{k-h-1}\overline{\tilde{B}_h}\,\overline{B_{h-1}}\ldots\overline{B_0}). \end{aligned}$$

Thus we have to prove

$$V^h U_{\overline{B_h}} \mathbf{f}(\overline{B_{h-1}}\ldots\overline{B_0}) = V^h U_{\overline{\tilde{B}_h}} \mathbf{f}(\overline{B_{h-1}}\ldots\overline{B_0}) + \sum_{k=h+1}^{\infty} V^k U_{(q^s-1)} \mathbf{f}((q^s-1)^{k-h-1}\overline{\tilde{B}_h}\,\overline{B_{h-1}}\ldots\overline{B_0}).$$

We introduce the matrices $X_B = U_{B+1} - U_B$ for $B \neq (q^s-1)$ and $X_{(q^s-1)} = U - U_{(q^s-1)}$, which by the definition of U_B only have one non-zero column. Then, using that $V^k U = V^{k-1}$ and

$X_B\mathbf{f}(\cdot) = \mathbf{f}(B\cdot)$ (which follows from the definition of X_B and (2.3)), we rewrite the right-hand side as

$$\begin{aligned}
&V^h U_{\overline{\tilde{B}_h}}\mathbf{f}(\overline{B_{h-1}}\dots\overline{B_0}) + \sum_{k=h+1}^{\infty} V^k (U - X_{(q^s-1)})\mathbf{f}((q^s-1)^{k-h-1}\overline{\tilde{B}_h}\overline{B_{h-1}}\dots\overline{B_0}) \\
&= V^h U_{\overline{\tilde{B}_h}}\mathbf{f}(\overline{B_{h-1}}\dots\overline{B_0}) + \sum_{k=h+1}^{\infty} V^{k-1}\mathbf{f}((q^s-1)^{k-h-1}\overline{\tilde{B}_h}\overline{B_{h-1}}\dots\overline{B_0}) - \\
&\qquad\qquad \sum_{k=h+1}^{\infty} V^k \mathbf{f}((q^s-1)^{k-h}\overline{\tilde{B}_h}\overline{B_{h-1}}\dots\overline{B_0}) \\
&= V^h U_{\overline{\tilde{B}_h}}\mathbf{f}(\overline{B_{h-1}}\dots\overline{B_0}) + V^h \mathbf{f}(\overline{\tilde{B}_h}\overline{B_{h-1}}\dots\overline{B_0}).
\end{aligned}$$

The last expression equals

$$V^h U_{\overline{\tilde{B}_h}}\mathbf{f}(\overline{B_{h-1}}\dots\overline{B_0}) + V^h X_{\overline{\tilde{B}_h}}\mathbf{f}(\overline{B_{h-1}}\dots\overline{B_0}) = V^h U_{\overline{B_h}}\mathbf{f}(\overline{B_{h-1}}\dots\overline{B_0})$$

and the assertion on φ is proved. Thus φ is a continuous function on $[0,1]$.

We now derive the asymptotic formula for $F(N)$. Set $N = B_0B_1\dots B_k$ (with $B_k \neq 0^{(s)}$ and $\lambda = \rho e^{i\vartheta}$. Denote as usual $\lfloor x\rfloor$ (respectively $\{x\}$) the integral (respectively the fractional) part of the real number x. Then $k = \left\lfloor \frac{1}{s}\log_q N\right\rfloor$ and $0.\overline{B_k}\dots\overline{B_0} = Nq^{-\left(s\left\lfloor\frac{1}{s}\log_q N\right\rfloor + s\right)}$. We now write $U^{k-r} = \tilde{U}^k V^r + W^{k-r}$ for $k \geq r$, with $\tilde{U} = T\,\mathrm{diag}(J_{k_1}(\lambda),\dots,J_{k_p}(\lambda),0,\dots,0)T^{-1}$, $W = U - \tilde{U}$ and, by convention again, $W^0 = T\,\mathrm{diag}(\underbrace{0,\dots,0}_{\kappa},1,\dots,1)T^{-1}$ and $\tilde{U}^0 = I - W^0$.

We now insert this into (2.6) to obtain

$$\mathbf{F}(B_0\dots B_k) = \tilde{U}^k\varphi(0.\overline{B_k}\dots\overline{B_0}) + R(B_0\dots B_k), \tag{2.13}$$

where

$$\begin{aligned}
R(B_0\dots B_k) &= W^k U_{B_k}\mathbf{f}(0) + W^{k-1}U_{B_{k-1}}\mathbf{f}(B_k) + W^{k-2}U_{B_{k-2}}\mathbf{f}(B_{k-1}B_k) + \cdots \\
&\quad + WU_{B_1}\mathbf{f}(B_2\dots B_k) + W^0 U_{B_0}\mathbf{f}(B_1\dots B_k).
\end{aligned} \tag{2.14}$$

By our choice of ξ and the assumptions on the eigenvalues of U there exists a constant C such that

$$\begin{aligned}
\|R(B_0\dots B_k)\| &\leq \max_{B\in\mathcal{B}_s}\|U_B\| \sum_{r=0}^{k} \left\|W^{k-r}\right\| \cdot \|\mathbf{f}(B_{k-r+1}\dots B_k)\| \\
&\leq \max_{B\in\mathcal{B}_s}\|U_B\| \sum_{r=0}^{k} C\xi^{k-r}\|g\|_\infty^{rs} = \mathcal{O}(\xi^k) = o(N^\mu)
\end{aligned} \tag{2.15}$$

where again we have used (2.8) and (2.9). Furthermore, (2.13) and $V\tilde{U}^k = U^{k-1}$ (for $k \geq 1$) gives

$$\begin{aligned}
&\mathbf{F}(B_1\dots B_k) - V\mathbf{F}(B_0\dots B_k) = \\
&\tilde{U}^{k-1}\left(\varphi(0.\overline{B_k}\dots\overline{B_1}) - \varphi(0.\overline{B_k}\dots\overline{B_0})\right) + R(B_1\dots B_k) - VR(B_0\dots B_k).
\end{aligned} \tag{2.16}$$

Then the definition of φ (2.11) yields

$$\varphi(0.\overline{B_k}\dots\overline{B_1})-\varphi(0.\overline{B_k}\dots\overline{B_0})=V^kU_{B_0}\mathbf{f}(B_1\dots B_k). \tag{2.17}$$

Putting (2.16) and (2.17) together and using (2.8), (2.12), and (2.15) yields

$$\begin{aligned}&\|\mathbf{F}(B_1\dots B_k)-V\mathbf{F}(B_0B_1\dots B_k)\|\le\\ &\|VU_{B_0}\mathbf{f}(B_1\dots B_k)\|+\|R(B_1\dots B_k)\|+\|VR(B_0\dots B_k)\|=\\ &\mathcal{O}(\|g\|_\infty^{ks})+\mathcal{O}(\xi^k)+\mathcal{O}(\xi^k)=\mathcal{O}(\xi^k)=o(N^\mu).\end{aligned} \tag{2.18}$$

We now define (here we use the definition of real powers of a matrix via the real powers of its Jordan decomposition as in (2.7))

$$\Psi(\log_q N)=\tilde V^{\{\frac1s\log_q N\}}\varphi\left(Nq^{-\left(s\left\lfloor\frac1s\log_q N\right\rfloor+s\right)}\right)$$

and note that Ψ can be extended to a continuous periodic function with period s on $\mathbb{R}\setminus s\mathbb{Z}$ by the above discussion on the continuity of φ. The continuity of this extension in the points $s\mathbb{Z}$ follows from the observation that $\varphi(q^{-s})=V\varphi(1)$.

By (2.13) and (2.15) we can write

$$\mathbf{F}(N)=\tilde U^{\frac1s\log_q N}\Psi(\log_q N)+o(N^\mu). \tag{2.19}$$

By (2.4) we have $F(B_0\dots B_k)=\sum_{C\in\mathcal{B}_s}F_C(B_1\dots B_k)+\sum_{C<B_0}f(CB_1\dots B_k)$ with

$$\begin{aligned}&\left|\sum_{C<B_0}f(CB_1\dots B_k)\right|<q^s\|g\|_\infty^{(k+1)s}=o(\xi^k)=o(N^\mu)\quad\text{and by (2.18)}\\ &\sum_{C\in\mathcal{B}_s}F_C(B_1\dots B_k)=(1,\dots,1)\mathbf{F}(B_1\dots B_k)=(1,\dots,1)V\mathbf{F}(N)+o(N^\mu).\end{aligned} \tag{2.20}$$

According to (2.19) and the definition of $\tilde U^{\frac1s\log_q N}$ (2.7) all the entries of $\mathbf{F}(N)$ are of the form

$$N^\delta e^{i\alpha\log_q N}\sum_{j=0}^{K-1}(\log_q N)^j\times\text{periodic function of }\log_q N.$$

Thus (2.10) is proved. □

Corollary 2. *Under the assumptions of Theorem 1 but allowing several dominating eigenvalues $\lambda_r=|\lambda|e^{is\alpha_r}$, $r=1,\dots,Q$, we obtain (with obvious notations)*

$$\sum_{n<N}f(n)=N^\delta\sum_{r=1}^{Q}\sum_{j=0}^{K_r-1}e^{i\alpha_r\log_q N}(\log_q N)^j\psi_{j,r}(\log_q N)+o(N^\mu).$$

Proof. This is an immediate but notationally inconvenient generalization of the proof of Theorem 1. □

Corollary 3. *Under the assumptions of Theorem 1 and for positive-valued function f we have*

$$\sum_{n<N}f(n)=N^\delta\psi(\log_q N)+o(N^\mu),$$

where ψ is a periodic function of period 1.

Proof. At first we notice that by for any $s \geq \max(1, \ell - 1)$ by Perron-Frobenius' theorem U has a unique dominating eigenvalue that is positive and of multiplicity 1. We now apply Theorem 1 to two coprime values of s to obtain 1-periodicity of ψ. □

We will now use Theorem 1 to describe the asymptotic behaviour of summatory functions of products of a block multiplicative function with a number of block additive functions. Such sums will occur in Section 3.

Proposition 4. *Let ϑ be a positive-valued block-multiplicative function satisfying the assumptions of Theorem 1 and $f_1, \dots, f_m$ arbitrary real-valued block-additive functions. Then the summatory function F of $\vartheta(n) f_1(n) \cdots f_m(n)$ satisfies*

$$F(N) = \sum_{n<N} \vartheta(n) f_1(n) \cdots f_m(n) = N^{\delta} \sum_{j=0}^{m} (\log_q N)^j \psi_j(\log_q N) + o(N^{\mu}), \tag{2.21}$$

where the functions ψ_j are continuous and periodic with period 1*; μ and δ are given by ϑ as in Theorem 1.*

Proof. Let $\mathbf{t} = (t_1, \dots, t_m) \in \mathbb{R}^m$. We define the function

$$h(n; \mathbf{t}) = \vartheta(n) \exp\left(\sum_{j=1}^{m} t_j f_j(n) \right).$$

We note that $n \mapsto h(n; \mathbf{t})$ is ℓ-block-multiplicative, where ℓ is the maximum of the block-lengths corresponding to ϑ and the f_j's. Then clearly h satisfies the hypotheses of Corollary 3 for $\mathbf{t}$ in some neighbourhood of $\mathbf{0}$. Thus we have

$$H(N; \mathbf{t}) = \sum_{n<N} h(n; \mathbf{t}) = N^{\delta(\mathbf{t})} \Psi(\log_q N; \mathbf{t}) + R(N; \mathbf{t}).$$

Defining δ and μ to be the exponents associated to ϑ by Theorem 1, we have $\delta(\mathbf{0}) = \delta$ and $R(N; \mathbf{t}) = o(N^{\mu})$ for $\mathbf{t}$ in some neighbourhood of $\mathbf{0}$. The dominating eigenvalue of the corresponding matrix has multiplicity 1 and is therefore an analytic function of $\mathbf{t}$ in a neighbourhood of $\mathbf{0}$.

Then differentiating at $\mathbf{t} = \mathbf{0}$ and disregarding the error term for a moment gives

$$\begin{aligned} \frac{\partial^m}{\partial t_1 \cdots \partial t_m} H(N; \mathbf{t}) \Big|_{\mathbf{t}=\mathbf{0}} &= \sum_{n<N} \vartheta(n) f_1(n) \cdots f_m(n) = \\ \sum_{(\varepsilon_1, \dots, \varepsilon_m) \in \{0,1\}^m} & \frac{\partial^{\varepsilon_1 + \cdots + \varepsilon_m}}{\partial t_1^{\varepsilon_1} \cdots \partial t_m^{\varepsilon_m}} N^{\delta(\mathbf{t})} \Big|_{\mathbf{t}=\mathbf{0}} \cdot \frac{\partial^{m - (\varepsilon_1 + \cdots + \varepsilon_m)}}{\partial t_1^{1-\varepsilon_1} \cdots \partial t_m^{1-\varepsilon_m}} \Psi(\log_q N; \mathbf{t}) \Big|_{\mathbf{t}=\mathbf{0}} + \text{remainder}. \end{aligned} \tag{2.22}$$

Writing $N^{\delta(\mathbf{t})} = N^{\delta}\exp[(\delta(\mathbf{t})-\delta)\log N]$ and expanding the exponential into its power series gives

$$\begin{aligned}\frac{\partial^{\varepsilon_1+\cdots+\varepsilon_m}}{\partial t_1^{\varepsilon_1}\cdots\partial t_m^{\varepsilon_m}}N^{\delta(\mathbf{t})}\Big|_{\mathbf{t}=\mathbf{0}} &= N^{\delta}\sum_{k=0}^{\infty}\frac{(\log N)^k}{k!}\,\frac{\partial^{\varepsilon_1+\cdots+\varepsilon_m}}{\partial t_1^{\varepsilon_1}\cdots\partial t_m^{\varepsilon_m}}\left(\delta(\mathbf{t})-\delta\right)^k\Big|_{\mathbf{t}=\mathbf{0}}\\ &= N^{\delta}\sum_{k=0}^{m}\frac{(\log N)^k}{k!}\,\frac{\partial^{\varepsilon_1+\cdots+\varepsilon_m}}{\partial t_1^{\varepsilon_1}\cdots\partial t_m^{\varepsilon_m}}\left(\delta(\mathbf{t})-\delta\right)^k\Big|_{\mathbf{t}=\mathbf{0}}. \qquad (2.23)\end{aligned}$$

Inserting (2.23) into (2.22) and collecting terms of like powers of $\log N$ gives (2.21) except for the error term.

It remains to give an argument that the error term can be pulled through differentiation. For this purpose we refer to the proof of Theorem 1, where the error term is made up of three parts: (2.14), (2.18), and (2.20), namely

$$\begin{aligned}R(N) = \sum_{C<B_0} h(CB_1\dots B_k;\mathbf{t}) + (1,\dots,1)V_{\mathbf{t}}R_{\mathbf{t}}(B_0\dots B_k)+\\ (1,\dots,1)\left(\mathbf{F}_{\mathbf{t}}(B_1\dots B_k)-V_{\mathbf{t}}\mathbf{F}_{\mathbf{t}}(B_0\dots B_k)\right).\end{aligned} \qquad (2.24)$$

We first study the first term: since this is a finite sum, it suffices to treat the single term

$$h(CB_1\dots B_k;\mathbf{t}) = \vartheta(CB_1\dots B_k)\exp\left(\sum_{j=1}^{m} f_j(CB_1\dots B_k)t_j\right).$$

Using (2.9) and since the block-additive functions satisfy $f_j(n)=\mathcal{O}(\log n)$ differentiation with respect to the variables $t_1,\dots,t_m$ yields only a contribution of

$$\vartheta(CB_1\dots B_k)\prod_{j=1}^{m} f_j(CB_1\dots B_k) = \mathcal{O}(k^m\|g\|_\infty^{sk}) = o\left(N^{\mu}\right).$$

For the second term in (2.24) the matrix W occurring in (2.14) is again analytic in a neighbourhood of $\mathbf{t}=\mathbf{0}$; the same is true for the matrices U_{B_k} and $\mathbf{f}$. Thus $R_{\mathbf{t}}(B_0\dots B_k)$ can be differentiated and the resulting sum can be estimated by

$$\mathcal{O}\left(k^m\sum_{r=0}^{k}\|W^{k-r}\|\cdot\|g\|_\infty^{s(r+1)}\right) = o(N^{\mu}).$$

For the last term in (2.24) observe that (2.6) gives

$$U\mathbf{F}(B_1\dots B_k)-\mathbf{F}(B_0\dots B_k) = -U_{B_0}\mathbf{f}(B_1\dots B_k).$$

Thus it can be treated similarly to the two other terms in (2.24). □

Remark 4. The leading term in the asymptotic formula (2.21) is

$$\prod_{j=1}^{m}\frac{\partial}{\partial t_j}\delta(\mathbf{t})\Big|_{\mathbf{t}=\mathbf{0}}\cdot N^{\delta}(\log N)^m\psi(\log_q N),$$

where $\psi(\cdot) = \Psi(\cdot, \mathbf{0})$ is the periodic function given by

$$\sum_{n<N} \vartheta(n) = N^{\delta}\psi(\log_q N) + o(N^{\mu}).$$

Furthermore, we have for $j = 1, \dots, m$

$$\sum_{n<N} \vartheta(n) f_j(n) = \frac{\partial}{\partial t_j}\delta(\mathbf{t})\Big|_{\mathbf{t}=\mathbf{0}} \cdot N^{\delta}(\log N) \cdot \psi(\log_q N) + \mathcal{O}(N^{\delta}).$$

Thus the constant in the leading term of $\sum_{n<N} \vartheta(n) f_1(n) \cdots f_m(n)$ is the product of the constants in the leading terms of $\sum_{n<N} \vartheta(n) f_j(n)$.

Remark 5. Proposition 4 can be used to compute the moments and correlations of block-additive functions by using $\vartheta(n) = 1$. Notice that in the case that $\vartheta(n) = 1$ the main term in the asymptotic expansion does not have a periodic factor.

3. p-Valuation of binomial coefficients

We consider the number of binomial coefficients $\binom{n}{k}$ whose p-adic valuation equals j: By a result of E. Kummer [**28**] the p-adic valuation of $\binom{n}{k}$ equals the number of carries in the addition of k and $n-k$ (or equivalently, in the subtraction of k from n) in base p. In particular we have

$$\vartheta_0(n) = \prod_{k=0}^{K}(1+\varepsilon_k) \quad \text{for } n = \sum_{k=0}^{K} \varepsilon_k p^k.$$

The quantities $\vartheta_j(n)$ and their summatory functions $S_j(N) = \sum_{n<N} \vartheta_j(n)$ have been studied by L. Carlitz [**6**], who gave a formula for $\vartheta_1(n)$, a recursion for $\vartheta_j(n)$, and computed $S_j(p^r)$. Further special values of $\vartheta_j(n)$ (for "lacunary" n) have been given by F. T. Howard [**26**].

Let us introduce the functions

$$a_\ell(\varepsilon_0, \dots, \varepsilon_\ell) = \frac{(p-\varepsilon_0-1)(p-\varepsilon_1)\cdots(p-\varepsilon_{\ell-1})\varepsilon_\ell}{\prod_{j=0}^{\ell}(1+\varepsilon_j)} \quad \text{for } \ell \geq 1.$$

According to Kummer's result the numerator of these functions counts the number of $k \leq m = \varepsilon_0 + p\varepsilon_1 + \cdots + p^{\ell}\varepsilon_\ell$, such that there are ℓ consecutive carries starting at the least significant digit in the subtraction of k from m. Splitting the j carries into s sets of consecutive ones gives

(3.1)

$$\vartheta_j(n) = \vartheta_0(n) \sum_{s=1}^{j} \sum_{\substack{\ell_1+\cdots+\ell_s=j \\ \forall i,\, k_i+\ell_i<k_{i+1}}} a_{\ell_1}(\varepsilon_{k_1}, \dots, \varepsilon_{k_1+\ell_1}) a_{\ell_2}(\varepsilon_{k_2}, \dots, \varepsilon_{k_2+\ell_2}) \cdots a_{\ell_s}(\varepsilon_{k_s}, \dots, \varepsilon_{k_s+\ell_s}).$$

For $j = 1$ we retrieve Carlitz' formula (cf. [**6**])

$$\vartheta_1(n) = \vartheta_0(n) \sum_{k=0}^{K} a_1(\varepsilon_k, \varepsilon_{k+1}) = \vartheta_0(n) \sum_{k=0}^{K} \frac{(p-\varepsilon_k-1)\varepsilon_{k+1}}{(1+\varepsilon_k)(1+\varepsilon_{k+1})}.$$

Theorem 5. *Let p be a prime. Then for $j \geq 0$ there exist continuous periodic functions of period 1, $\psi_r^{(j)}$, $r = 0, \dots, j$, such that*

$$\#\left\{(k,n) : 0 \leq k \leq n < N \text{ and } p^j \| \binom{n}{k}\right\} = N^\alpha \sum_{r=0}^{j} \psi_r^{(j)}(\log_p N)(\log_p N)^r + o(N^\varepsilon) \tag{3.2}$$

for $\alpha = \log_p \frac{p(p+1)}{2}$ and any $\varepsilon > 0$. Furthermore, $\psi_j^{(j)} = \frac{1}{j!}\left(\frac{p-1}{p+1}\right)^{2j} \psi_0^{(0)}$.

Proof. We first note that the case $j = 0$ has been established in [**3, 14**]; the error term vanishes in this case. We assume in the sequel that $j \geq 1$. The main idea of the proof is to make a recurrence on the number s of different "connected" digital functions involved in (3.1). In that context, the difficulty is to formalize this notion of connectedness in a suitable way, which respects in particular the symmetry of (3.1). This is done by introducing an action of the symmetric group on sub-sums of (3.1) and by associating a graph to the set of the different digital functions occurring in the sum.

In order to recognize $\vartheta_j(n)\vartheta_0(n)^{-1}$ as a linear combination of products of block-additive functions (as studied in Proposition 4), we introduce a further set of digital functions. For s-tuples $\mathbf{T} = (t_1, \dots, t_s)$ and $\mathbf{L} = (\ell_1, \dots, \ell_s)$ satisfying

$$t_1 = 0, \quad t_r \geq 0, \quad \ell_r > 0 \quad \text{for } r = 1, \dots, s$$

we define a graph on the pairs (ℓ_r, t_r) by connecting (ℓ_1, t_1) and (ℓ_2, t_2), if the intervals $\{t_1, \dots, t_1 + \ell_1\}$ and $\{t_2, \dots, t_2 + \ell_2\}$ intersect. We call pairs $(\mathbf{L}, \mathbf{T})$ *connected* if they correspond to connected graphs. For these we define the functions

$$b_{\mathbf{L}}^{\mathbf{T}}(\varepsilon_0, \dots, \varepsilon_{m(\mathbf{L},\mathbf{T})}) = \prod_{r=1}^{s} a_{\ell_r}(\varepsilon_{t_r}, \dots, \varepsilon_{t_r+\ell_r}) \quad \text{with } m(\mathbf{L},\mathbf{T}) = \max_r(t_r + \ell_r) \quad \text{and} \tag{3.3}$$

$$h_{\mathbf{L}}^{\mathbf{T}}(n) = \sum_{k=0}^{K} b_{\mathbf{L}}^{\mathbf{T}}(\varepsilon_k, \dots, \varepsilon_{k+m(\mathbf{L},\mathbf{T})}) \quad \text{for } n = \sum_{k=0}^{K} \varepsilon_k p^k. \tag{3.4}$$

Notice that for $s = 1$, $h_\ell^0(n)$ is the $(\ell+1)$-block-additive function defined by a_ℓ. Furthermore, the set of functions $b_{\mathbf{L}}^{\mathbf{T}}$ is stable under multiplication in the following sense: for given connected tuples $(\mathbf{L}_1, \mathbf{T}_1), \dots, (\mathbf{L}_s, \mathbf{T}_s)$ and $(\tau_1, \dots, \tau_s) \in \mathbb{N}^s$, $\tau_1 = 0$, there exists a connected pair $(\mathbf{L}', \mathbf{T}')$ such that

$$\prod_{r=1}^{s} b_{\mathbf{L}_r}^{\mathbf{T}_r}(\varepsilon_{\tau_r}, \dots, \varepsilon_{\tau_r+m(\mathbf{L}_r,\mathbf{T}_r)}) = b_{\mathbf{L}'}^{\mathbf{T}'}(\varepsilon_0, \dots, \varepsilon_{m(\mathbf{L}',\mathbf{T}')}) \tag{3.5}$$

provided that the pair $((m(\mathbf{L}_1, \mathbf{T}_1), \dots, m(\mathbf{L}_s, \mathbf{T}_s)), (\tau_1, \dots, \tau_s))$ is itself connected. If the pair $((m(\mathbf{L}_1, \mathbf{T}_1), \dots, m(\mathbf{L}_s, \mathbf{T}_s)), (\tau_1, \dots, \tau_s))$ is not connected, the product (3.5) can be written in a minimal way by collecting the factors a_ℓ according to the connected components of the corresponding graph.

We want to express the sum (3.1) as a linear combination of products of functions $h_{\mathbf{L}}^{\mathbf{T}}(n)$. It suffices to prove this for the inner sum for a fixed value of s. This we do by induction on

s. For $s = 1$ we get the block-additive function $h_j^0(n)$ and the assertion is clear. Assume that $s > 1$ and fix $\ell_1, \dots, \ell_s$. Then the symmetric group $\mathfrak{S}_s$ acts on the set of sums

$$S = \sum_{k_i+\ell_i<k_{i+1}} a_{\ell_1}(\varepsilon_{k_1}, \dots, \varepsilon_{k_1+\ell_1}) a_{\ell_2}(\varepsilon_{k_2}, \dots, \varepsilon_{k_2+\ell_2}) \cdots a_{\ell_s}(\varepsilon_{k_s}, \dots, \varepsilon_{k_s+\ell_s})$$

by

$$\sigma \cdot S = \sum_{k_i+\ell_{\sigma(i)}<k_{i+1}} a_{\ell_{\sigma(1)}}(\varepsilon_{k_1}, \dots, \varepsilon_{k_1+\ell_{\sigma(1)}}) a_{\ell_{\sigma(2)}}(\varepsilon_{k_2}, \dots, \varepsilon_{k_2+\ell_{\sigma(2)}}) \cdots a_{\ell_{\sigma(s)}}(\varepsilon_{k_s}, \dots, \varepsilon_{k_s+\ell_{\sigma(s)}}).$$

Observe that if a sum S occurs in (3.1) then also $\sigma \cdot S$ occurs and that the graph corresponding to $((\ell_{\sigma(1)}, \dots, \ell_{\sigma(s)}), (k_1, \dots, k_s))$ is totally disconnected by the condition on the summation indices. We split the inner sum in (3.1) into the orbits of s-tuples $(\ell_1, \dots, \ell_s)$ and define

$$S_{\ell_1,\dots,\ell_s} = \sum_{\sigma \in \mathfrak{S}_s} \sigma \cdot S.$$

Then we have

$$h_{\ell_1}^0(n) \cdots h_{\ell_s}^0(n) = S_{\ell_1,\dots,\ell_s} + \sum,$$

where $\sum$ is a sum over the same summands as $\sum_\sigma \sigma \cdot S$ but with at least one of the conditions $k_i + \ell_{\sigma(i)} < k_{i+1}$ violated. For given $(k_1, \dots, k_s)$ this means that the graph corresponding to $((\ell_1, \dots, \ell_s), (k_1, \dots, k_s))$ has at least one non-trivial component. We now split $\sum$ into subsums: For k's corresponding to one component we fix the according differences, for k's in different components we impose conditions such that the component structure is preserved. For every summand each component C of the graph yields a function $b_{\mathbf{L}_C}^{\mathbf{T}_C}$ in the following way:

$$a_{\ell_1}(\varepsilon_{k_1}, \dots, \varepsilon_{k_1+\ell_1}) \cdots a_{\ell_r}(\varepsilon_{k_r}, \dots, \varepsilon_{k_r+\ell_r}) = b_{\ell_1,\dots,\ell_r}^{0,k_2-\tilde{k}_C,\dots,k_r-\tilde{k}_C}(\varepsilon_{\tilde{k}_C}, \dots, \varepsilon_{\tilde{k}_C+m(\mathbf{L}_C,\mathbf{T}_C)}),$$

where we assume that $k_1 = \min(k_1, \dots, k_r)$ and set $\tilde{k}_C = k_1$ (otherwise reorder); furthermore, $t_1^C = 0, t_2^C = k_2 - \tilde{k}_C, \dots, t_r^C = k_r - \tilde{k}_C$. Thus for every component C we introduce a new variable $\tilde{k}_C$ and variables $t_1^C, \dots, t_{|C|}^C$ and split summation over $k_1, \dots, k_s$ into summation over $\tilde{k}_C$ (inner sum) and the corresponding t^C's. Therefore the inner sum involves less than s variables and by induction we have expressed $\vartheta_j \vartheta_0^{-1}$ as a sum of products of block-additive functions.

Finally, we have to consider the matrix U corresponding to the function ϑ_0; since this function is multiplicative, we have $\frac{\vartheta_0(BC)}{\vartheta_0(C)} = \vartheta_0(B)$, and therefore

$$U = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 2 & 2 & \dots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ p & p & \dots & p \end{pmatrix}.$$

Then the dominating eigenvalue of U is $\frac{p(p+1)}{2}$ and all the other eigenvalues are 0. Application of Proposition 4 and noting that $\|g\|_\infty = p$ finishes the proof of (3.2).

It remains to prove that $\psi_j^{(j)} = \frac{1}{j!}\left(\frac{p-1}{p+1}\right)^{2j}\psi_0^{(0)}$. It is clear that the two functions are proportional by Proposition 4. Thus it suffices to compare the two functions at one point. We choose $N = p^r$ and use a formula given in [**6**] ((7.14), p. 319):

$$\begin{aligned} S_j(p^r) &= \sum_{s=0}^{j}\binom{j-1}{s}\binom{r-j}{j-s}\binom{p}{2}^{2j-2s}\binom{p+1}{2}^{r-2j-2s} \\ &\sim \frac{1}{j!}\left(\frac{p-1}{p+1}\right)^{2j}\binom{p+1}{2}^{r} r^j \sim \binom{p+1}{2}^{r} r^j \psi_j^{(j)}(0). \end{aligned}$$

Using $\psi_0^{(0)}(0) = 1$ (cf. [**40**]) we get the proportionality factor. □

Remark 6. For $j = 1, 2, 3$ we have computed the formulæ for ϑ_j in terms of the functions $h_{\mathbf{L}}^{\mathbf{T}}$:

$$\begin{aligned} \vartheta_1(n) =& \vartheta_0(n) h_1^0(n) \\ \vartheta_2(n) =& \vartheta_0(n)\left(\frac{1}{2}(h_1^0(n))^2 + h_2^0(n) - h_{1,1}^{0,1}(n) - \frac{1}{2}h_{1,1}^{0,0}(n)\right) \\ \vartheta_3(n) =& \vartheta_0(n)\left(\frac{1}{6}(h_1^0(n))^3 + h_2^0(n)h_1^0(n) - h_1^0(n)h_{1,1}^{0,1}(n) - \frac{1}{2}h_1^0(n)h_{1,1}^{0,0}(n) + h_{1,1,1}^{0,1,1}(n) + \right. \\ & \left. h_{1,1,1}^{0,1,2}(n) + h_{1,1,1}^{0,0,1}(n) + \frac{1}{3}h_{1,1,1}^{0,0,0}(n) - h_{2,1}^{0,2}(n) - h_{2,1}^{0,1}(n) - h_{2,1}^{0,0}(n) - h_{1,2}^{0,1}(n) + h_3^0(n)\right) \end{aligned}$$

Remark 7. Several papers have been devoted to the detailed study of the function $\psi_0^{(0)}$. For instance it is the quotient of an increasing function and a differentiable function and therefore is itself differentiable almost everywhere; for $p = 2$ the points of non-differentiability have been characterized in [**30**]. The minimum has been studied for $p = 2$ in [**23, 30, 38, 39, 40, 41**] and for general p in [**16**] (for an extensive bibliography for the work before 1977 from the point of view of digital functions we refer to [**41**]). In particular, it is bounded away from 0. Therefore the first term in (3.2) is an asymptotic leading term.

4. Distribution of binomial coefficients modulo prime powers

In this section we will generalize results of R. Garfield and H. Wilf [**17**] and D. Barbolosi and the second author [**3**] concerning the distribution of binomial coefficients in the residue classes modulo primes to prime powers. It is clear from results of D. Singmaster [**36**] and also from section 3 that "almost all" (in the sense of density) binomial coefficients lie in the 0 residue class modulo p^ℓ.

In the sequel we will use the notation $m_{(p)}$ to denote the p-free part of m; this is $m_{(p)} = mp^{-v_p(m)}$. Furthermore, as in [**21**] we will use the notation $(n!)_p$ for the product of all integers less than or equal to n, which are not divisible by p. Since the subscript p will only occur with factorials this should not cause any confusion.

4.1. **Binomial coefficients with given p-valuation.** A famous formula due to É. Lucas [**32**] expresses the binomial coefficient $\binom{n}{k}$ modulo a prime p in terms of the p-adic digits of n and k. Recently this congruence was generalized by K. S. Davis and W. A. Webb [**8**] and A. Granville [**21**] to prime powers in the following sense (we give Granville's formulation):

Lemma 2. *Suppose that a prime power p^ℓ and positive integers $n = m + r$ are given. Write $n = n_0 + n_1 p + \cdots + n_d p^d$ in base p, and let N_j be the least positive residue of $[n/p^j] \mod p^\ell$ for each $j \geq 0$ (so that $N_j = n_j + n_{j+1}p + \cdots + n_{j+\ell-1}p^{\ell-1}$): also make the corresponding definitions for m_j, M_j, r_j, R_j. Let e_j be the number of 'carries', when adding m and r in base p, on or beyond the j-th digit. In particular, we have $p^{e_0} \| \binom{n}{m}$. Then*

$$\frac{1}{p^{e_0}}\binom{n}{m} \equiv (\pm 1)^{e_{\ell-1}} \left(\frac{(N_0!)_p}{(M_0!)_p (R_0!)_p}\right)\left(\frac{(N_1!)_p}{(M_1!)_p (R_1!)_p}\right)\cdots\left(\frac{(N_d!)_p}{(M_d!)_p (R_d!)_p}\right) \mod p^\ell, \tag{4.1}$$

where (± 1) is (-1) except if $p = 2$ and $\ell \geq 3$.

Remark 8. Special cases of this congruence were known earlier; for a more precise history we refer to [**21**] and Dickson's book [**12**].

In order to apply the ideas of section 2 to the distribution of binomial coefficients we have to define bivariate block-multiplicative functions: for given $g : \{0, \ldots, q-1\}^{2\ell} \to \mathbb{C}$ with $g(0, \ldots, 0; 0, \ldots, 0) = 1$ let

$$f(n; m) = f\left(\sum_{j=0}^{\infty} \varepsilon_j q^j; \sum_{j=0}^{\infty} \delta_j q^j\right) = \prod_{j=0}^{\infty} g\left(\varepsilon_j, \varepsilon_{j+1}, \ldots, \varepsilon_{j+\ell-1}; \delta_j, \delta_{j+1}, \ldots, \delta_{j+\ell-1}\right).$$

As for simple block-multiplicative functions we set

$$F(N) = \sum_{\substack{n<N \\ m<N}} f(n; m)$$
$$F_{B,B'}(N) = \sum_{\substack{n<N \\ m<N}} f(Bn; B'm)$$
$$\mathbf{F}(N) = (F_{B,B'}(N))_{B,B' \in \mathcal{B}_s}.$$

Similarly to Lemma 1 we have

Lemma 3. *For $s \geq \max(1, \ell - 1)$ and arbitrary $s' \geq 1$ the following equations hold*

$$f(B;B')f(ABC;A'B'C') = f(AB;A'B')f(BC;B'C') \tag{4.2}$$
$$\text{for } B, B' \in \mathcal{B}_s,\ |A| = |A'|,\ \text{and } |C| = |C'|$$

$$F(BN) = \sum_{C,C' \in \mathcal{B}_s} F_{C,C'}(N) + \sum_{n<N} \sum_{C \in \mathcal{B}_{s'}} \sum_{C'<B} f(Cn;C'N) + \tag{4.3}$$
$$\sum_{m<N} \sum_{C' \in \mathcal{B}_{s'}} \sum_{C<B} f(CN;C'm) + \sum_{C,C'<B} f(CN;C'N)$$

$$\mathbf{F}(q^{r+s}) = U\mathbf{F}(q^r), \quad \text{where } U = \left(u_{(B,B'),(C,C')}\right)_{(B,B'),(C,C') \in \mathcal{B}_s^2} \tag{4.4}$$

$$\text{with } u_{(B,B'),(C,C')} = \begin{cases} \frac{f(BC;B'C')}{f(C;C')} & \text{for } f(C;C') \neq 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.5}$$

$$F_{C,C'}(BN) = \sum_{D,D' \in \mathcal{B}_{s'}} \sum_{\substack{n<N \\ m<N}} f(CDn;C'D'm) + \sum_{n<N} \sum_{D \in \mathcal{B}_{s'}} \sum_{D'<B} f(CDn;C'D'N) + \tag{4.6}$$
$$\sum_{m<N} \sum_{D' \in \mathcal{B}_{s'}} \sum_{D<B} f(CDN;C'D'm) + \sum_{D,D'<B} f(CDN;C'D'N) \text{ for } |C| = |C'| = s \text{ and } |B| = s'.$$

The following three lemmas will be used in the proofs of Theorems 6 and 7.

Lemma 4. *Let A be an $m \times m$ matrix and B the matrix given by $b_{ij} = |a_{ij}|$. Suppose further that B is primitive (i.e. there is a power of B with all entries strictly positive, cf. [**35**]). Let $a_{ij}^{(n)}$ be the entries of the matrix A^n and $b_{ij}^{(n)}$ be the entries of the matrix B^n. If $|a_{ij}^{(n)}| < b_{ij}^{(n)}$ holds for one pair (i,j) and one n, then*

$$\max_{\mu \in \operatorname{spec} A} |\mu| < \max_{\lambda \in \operatorname{spec} B} |\lambda|.$$

Proof. It is clear that $|a_{ij}^{(n)}| \leq b_{ij}^{(n)}$. Assume now that $|a_{IJ}^{(N)}| < b_{IJ}^{(N)}$ and $b_{ij}^{(M)} \neq 0$ for all pairs (i,j). Then we have for all $\ell = 1, \ldots, m$

$$\left|a_{I\ell}^{(N+M)}\right| = \left|\sum_{j=1}^{m} a_{Ij}^{(N)} a_{j\ell}^{(M)}\right| \leq \sum_{j=1}^{m} \left|a_{Ij}^{(N)}\right| b_{j\ell}^{(M)} < b_{I\ell}^{(N+M)},$$

since $b_{J\ell}^{(M)} \neq 0$. Applying the same argument to the multiplication of $A^M \cdot A^{N+M}$ yields that

$$\left|a_{ij}^{(N+2M)}\right| < b_{ij}^{(N+2M)} \text{ for all pairs } (i,j).$$

Thus there is an $\alpha < 1$ and some K such that

$$\left|a_{ij}^{(K)}\right| \leq \alpha b_{ij}^{(K)} \text{ for all } i,j = 1, \ldots, m,$$

from which we conclude that

$$\left|a_{ij}^{(nK)}\right| \leq \alpha^n b_{ij}^{(nK)} \text{ for all } i,j = 1, \ldots, m, \text{ and all } n \geq 1.$$

This yields the following inequality for the norms of the matrices

$$\left\|A^{nK}\right\| \leq \alpha^n \left\|B^{nK}\right\|$$

and by taking nK-th roots and letting n tend to infinity we obtain that

$$\max_{\mu\in\operatorname{spec} A}|\mu|\leq\alpha^{\frac{1}{K}}\max_{\lambda\in\operatorname{spec} B}|\lambda|.$$

□

Lemma 5. *Let p be a prime and $n\in\mathbb{Z}$. Then the following congruence holds for all $\ell\geq 0$:*

$$\binom{p^{\ell+1}n}{p^{\ell+1}}\equiv\binom{p^{\ell}n}{p^{\ell}}\bmod p^{\ell+1}. \tag{4.7}$$

Proof. We use the following observation stated in [**8**]:

$$\binom{pN}{pK}=\binom{N}{K}\prod_{j=1}^{p-1}\prod_{k=1}^{K}\frac{p(k+N-K)-j}{pk-j}.$$

Inserting $N=p^{\ell}n$ and $K=p^{\ell}$ into this equation and observing that the denominators in the product are never divisible by p we obtain the congruence

$$\binom{p^{\ell+1}n}{p^{\ell+1}}=\binom{p^{\ell}n}{p^{\ell}}\prod_{j=1}^{p-1}\prod_{k=1}^{p^{\ell}}\frac{p(k+p^{\ell}n-p^{\ell})-j}{pk-j}\equiv\binom{p^{\ell}n}{p^{\ell}}\prod_{j=1}^{p-1}\prod_{k=1}^{p^{\ell}}\frac{pk-j}{pk-j}\equiv\binom{p^{\ell}n}{p^{\ell}}\bmod p^{\ell+1}.$$

□

Remark 9. This lemma would be a special case of Theorem 3 in [**9**], but is excluded in the statement there.

Lemma 6. *If n ranges over the numbers $0,1,\ldots,p^{\ell}-1$ then $\binom{p^{\ell}n}{p^{\ell}}$ ranges over all residue classes modulo p^{ℓ}.*

Proof. We proceed by induction on ℓ. For $\ell=1$ the assertion is clear by the fact that $\binom{pn}{p}\equiv n \bmod p$. Assume now that the assertion is true for ℓ and that

$$\binom{p^{\ell+1}n}{p^{\ell+1}}\equiv\binom{p^{\ell+1}m}{p^{\ell+1}}\bmod p^{\ell+1}. \tag{4.8}$$

Then we have

$$\binom{p^{\ell}n}{p^{\ell}}\equiv\binom{p^{\ell}m}{p^{\ell}}\bmod p^{\ell}.$$

as a consequence of (4.7) and by the induction hypothesis we get $m\equiv n \bmod p^{\ell}$. Inserting $n=n'+\varepsilon p^{\ell}$ and $m=n'+\delta p^{\ell}$ into (4.8) we obtain

$$(n'+\varepsilon p^{\ell})\prod_{r=1}^{p^{\ell+1}-1}\frac{p^{\ell+1-v_p(r)}(n'+\varepsilon p^{\ell})-r_{(p)}}{p^{\ell+1-v_p(r)}-r_{(p)}}\equiv(n'+\delta p^{\ell})\prod_{r=1}^{p^{\ell+1}-1}\frac{p^{\ell+1-v_p(r)}(n'+\delta p^{\ell})-r_{(p)}}{p^{\ell+1-v_p(r)}-r_{(p)}}\bmod p^{\ell+1}.$$

It is immediate that the factors in the product do not depend on ε and δ, from which we deduce that $\varepsilon\equiv\delta \bmod p$ and therefore $n\equiv m \bmod p^{\ell+1}$. □

Theorem 6. *Let p be prime, $j \geq 0$, $\ell \geq 1$, and $a \in \mathbb{Z}$, $(a,p) = 1$. Then there exists $\beta < \alpha = \log_p \frac{p(p+1)}{2}$ such that*

$$\begin{aligned} &\#\left\{(k,n) : 0 \leq k \leq n < N,\ p^j \| \binom{n}{k},\ \text{and}\ p^{-j}\binom{n}{k} \equiv a \bmod p^\ell\right\} = \\ &\frac{1}{\phi(p^\ell)} \#\left\{(k,n) : 0 \leq k \leq n < N \text{ and } p^j \| \binom{n}{k}\right\} + \mathcal{O}\left(N^\beta\right) = \\ &\frac{1}{\phi(p^\ell)} N^\alpha \sum_{r=0}^{j} \psi_r^{(j)}(\log_p N)(\log_p N)^r + \mathcal{O}\left(N^\beta\right), \end{aligned} \tag{4.9}$$

where $\psi_r^{(j)}$ are continuous periodic functions defined in Theorem 5 and ϕ is Euler's function.

Proof. We begin with the case $j = 0$ and fix $\ell \geq 2$ for technical reasons (the case $\ell = 1$ follows by summation from $\ell = 2$). Let χ be a Dirichlet character modulo p^ℓ. We introduce the notation $m \leq_p n$ to indicate that the k-th p-adic digit of m is less than or equal to the k-th digit of n for all k. For integers $0 \leq m, n < p^\ell$ we define the function

$$g_\chi(n;m) = \begin{cases} \chi\left(\dfrac{(n!)_p}{(m!)_p((n-m)!)_p}\right) & \text{if } m \leq_p n \\ 0 & \text{otherwise.} \end{cases} \tag{4.10}$$

Let f_χ be the corresponding bivariate block-multiplicative function. We note that $f_\chi(n;m) = \chi\left(\binom{n}{m}\right)$ by (4.1) and furthermore that

$$\#\left\{(k,n) : 0 \leq k \leq n < N \text{ and } \binom{n}{k} \equiv a \bmod p^\ell\right\} = \frac{1}{\phi(p^\ell)} \sum_\chi \overline{\chi(a)} \sum_{m,n<N} f_\chi(n;m) \tag{4.11}$$

for $(a,p) = 1$.

The term corresponding to the principal character ε in the right hand side of (4.11) is exactly the number of binomial coefficients not divisible by p. In the sequel we will use matrices U as in Lemma 3 for $s = \ell$. The matrix U_ε corresponding to f_ε has only entries 0 and 1. The matrices U_χ for non-principal χ have 0 entries exactly in the same places as U_ε. The other entries of U_χ are complex numbers of modulus 1. Since $u_{(B,B'),(C,C')} \neq 0$, if and only if $B' \leq_p B$ and $C' \leq_p C$, a suitable ordering of $\mathcal{B}_\ell \times \mathcal{B}_\ell$ gives a matrix U_ε of the form

$$\left(\tfrac{p(p+1)}{2}\right)^\ell \left\{ \begin{pmatrix} 1 & \dots & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & \dots & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}, \right.$$

which shows immediately that U_ε has rank 1 and that its only non-zero eigenvalue is $\left(\frac{p(p+1)}{2}\right)^\ell$. It is sufficient to deal with the sub-matrix of size $\left(\frac{p(p+1)}{2}\right)^\ell$ containing the non-zero entries. For convenience, we denote this matrix also by U_χ. By Lemma 4 it suffices to

find one entry in U_χ^2 which has smaller modulus than $\left(\frac{p(p+1)}{2}\right)^\ell$ (the corresponding entry in U_ε^2). For this purpose we observe that by (4.2) and (4.5)

$$u^\chi_{(0,0),(n,1)} u^\chi_{(n,1),(0,0)} = \chi\left(\binom{p^\ell n}{p^\ell}\right).$$

Thus as a consequence of Lemma 6 we have

$$\sum_n u^\chi_{(0,0),(n,1)} u^\chi_{(n,1),(0,0)} = 0,$$

where the summation is extended over all $n = 1, \dots, p^\ell - 1$, which are not divisible by p. Therefore, the entry $u^{\chi,(2)}_{(0,0),(0,0)}$ of U_χ^2 is strictly less than $\left(\frac{p(p+1)}{2}\right)^\ell$.

To conclude with the case $j = 0$, we first note that by (4.2) the first sum in (4.6) is

$$\sum_{D,D' \in \mathcal{B}_{s'}} u_{(C,C'),(D,D')} F_{D,D'}(N).$$

Thus, since $\|g_\chi\|_\infty = 1$, there exists a constant $c > 0$ (independent of B, N, and χ) such that

$$\|\mathbf{F}_\chi(BN) - U_\chi \mathbf{F}_\chi(N)\| \le cN. \tag{4.12}$$

For $\chi \neq \varepsilon$ iterating this inequality, using (2.8) and the estimate for λ_χ yields (for any $\eta > 0$)

$$\|\mathbf{F}_\chi(N)\| = \mathcal{O}\left(\max(|\lambda_\chi| + \eta, p^s)^{\frac{1}{s}\log_p N}\right) = \mathcal{O}(N^\beta) \tag{4.13}$$

for some $\beta < \alpha$.

We now consider arbitrary $j \ge 1$ for $N = B_0 \dots B_k$ with $B_0, \dots, B_k \in \mathcal{B}_\ell$. In order to handle the carries, which occur in this case, we denote by $\mathcal{C}(n, m)$ the set of indices, where a carry occurs in the subtraction of m from n. For a given set of carries A and a positive integer t we denote $A - t = \{k - t : k - t \ge 0, k \in A\}$. For fixed A, χ a fixed non-trivial character, and $N < p^{K+1}$ we introduce

$$F^{(A)}(N) = \sum_{\substack{n<N \\ m<N \\ \mathcal{C}(n,m)=A}} \chi\left(\frac{(N_0!)_p}{(M_0!)_p (R_0!)_p}\right) \cdots \chi\left(\frac{(N_K!)_p}{(M_K!)_p (R_K!)_p}\right),$$

and as before we use $F^{(A)}_{C,C'}(N)$ and $\mathbf{F}^{(A)}(N)$ (for readability we omit the dependence on χ in the notation).

We now proceed by induction on the size of A and claim that there exist $C = C(\#A) > 0$ and $\beta < \alpha$ such that for all N

$$\left\|\mathbf{F}^{(A)}(N)\right\| \le C(\#A) N^\beta.$$

This is true for $\#A = j = 0$ by (4.13). Until meeting the first carry we get similarly to (4.12)

$$\begin{aligned}\left\|\mathbf{F}^{(A-r\ell)}(B_rB_{r+1}\dots B_k) - U_\chi\cdot\mathbf{F}^{(A-(r+1)\ell)}(B_{r+1}\dots B_k)\right\| \le \\ \le c\times(B_{r+1}\dots B_k)\le cNp^{-(r+1)\ell} \\ \text{for } r\le r_0=\left\lfloor\frac{\min A}{\ell}\right\rfloor-1\end{aligned} \tag{4.14}$$

(note that $\#(A-(r+1)\ell) = \#A$). For $r = r_0+1$ we encounter the first carry. We form a block of length $t\ge\ell$ which contains at least this carry, ends with $\ell-1$ digits without carry, and such that the $\ell-1$ digits after this block have no carries. We choose t minimal with respect to these properties; it is clear that $t\le 2j(\ell-1)$. Then we define two matrices V and W of respective dimensions $p^{2\ell}\times p^{2t}$ and $p^{2t}\times p^{2\ell}$ (using the notation of (4.1), which explains the meaning of R and $\tilde R$):

$$v_{(B,B'),(C,C')} = \begin{cases}\dfrac{\displaystyle\prod_{i=0}^{t+\ell-1}\chi\left(\frac{((BC)_i!)_p}{((B'C')_i!)_p((R)_i!)_p}\right)}{\displaystyle\prod_{i=0}^{t}\chi\left(\frac{((C)_i!)_p}{((C')_i!)_p((\tilde R)_i!)_p}\right)} & \text{if subtraction of } C' \text{ from } C \text{ has carries exactly in the according positions}\\ 0 & \text{otherwise,}\end{cases}$$

and

$$w_{(C,C'),(B,B')} = \begin{cases}\dfrac{\displaystyle\prod_{i=0}^{t+\ell-1}\chi\left(\frac{((CB)_i!)_p}{((C'B')_i!)_p((R)_i!)_p}\right)}{\displaystyle\prod_{i=0}^{t}\chi\left(\frac{((B)_i!)_p}{((B')_i!)_p((\tilde R)_i!)_p}\right)} & \text{if subtraction of } C' \text{ from } C \text{ has carries exactly in the according positions}\\ 0 & \text{otherwise}\end{cases}$$

for $|B| = |B'| = \ell$ and $|C| = |C'| = t$.

We now rewrite $B_{r_0+1}\dots B_k = CB'_1\dots B'_{k'}$ (by adding leading 0's if necessary) with $|C| = t$ and $|B'_i| = \ell$. Similarly to (4.14) we have

$$\begin{aligned}\left\|\mathbf{F}^{(A-(r_0+1)\ell)}(B_{r_0+1}\dots B_k) - VW\mathbf{F}^{(A-(r_0+2)\ell-t)}(B'_2\dots B'_{k'})\right\| \le \\ \le c_V\times(B'_1\dots B'_{k'}) + \|V\|c_W\times(B'_2\dots B'_{k'})\le c'Np^{-(r_0+2)\ell-t},\end{aligned} \tag{4.15}$$

where the last inequality makes use of the fact that there are only finitely many matrices V and W, since $t\le 2j(\ell-1)$. By construction we have $\#(A-(r_0+2)\ell-t) < \#A$. By putting together (4.14) and (4.15) we obtain

$$\left\|\mathbf{F}^{(A)}(B_0\dots B_k) - U_\chi^{r_0}VW\mathbf{F}^{(A-(r_0+2)\ell-t)}(B'_2\dots B'_{k'})\right\| \le c'N\sum_{i=0}^{r_0}\left(\frac{\|U_\chi\|}{p^\ell}\right)^i.$$

Using the induction hypothesis and proceeding as in (4.13) we get

$$\left\|F^{(A)}(N)\right\| = C(\#A)N^\beta.$$

Summing up over the $\mathcal{O}((\log N)^j)$ possible choices of A and taking into account that the factor $(-1)^{e_{\ell-1}}$ in (4.1) depends only on A yield that the sum over the non-principal characters is $\mathcal{O}(N^{\beta+\varepsilon})$ (for any $\varepsilon > 0$). As for $j = 0$ the theorem follows. □

4.2. **p-free part of binomial coefficients.** A further natural question concerns the distribution of the p-free part of binomial coefficients in the p-adic integers. Since in Theorem 6 the valuation (or equivalently, the number of carries) was fixed, and the error depended on this, the following theorem is not a consequence of Theorem 6. Conversely, such a p-adic result could not imply Theorem 6 because the set of binomial coefficients with given valuation has zero density.

Theorem 7. *Let p be a prime. Then for any $a \in \mathbb{Z}_p^*$, any $\ell \geq 1$, and any $(m, r) \in \mathbb{N}^* \times \mathbb{N}$ we have*

$$\lim_{N\to\infty} \frac{2}{N^2} \# \left\{ (n,k) : 0 \leq k \leq n < N,\ v_p\left(\binom{n}{k}\right) \equiv r \bmod m, \right. \\ \left. \text{and } \binom{n}{k}_{(p)} \equiv a \bmod p^\ell \mathbb{Z}_p \right\} = \frac{1}{m\phi(p^\ell)}. \tag{4.16}$$

Before giving a proof we state two immediate corollaries.

Corollary 8. *For any $a \in \mathbb{Z}_p^*$ and any $\ell \geq 1$ we have*

$$\lim_{N\to\infty} \frac{2}{N^2} \# \left\{ (n,k) : 0 \leq k \leq n < N \text{ and } \binom{n}{k}_{(p)} \equiv a \bmod p^\ell \mathbb{Z}_p \right\} = \frac{1}{\phi(p^\ell)}; \tag{4.17}$$

in words: the p-free parts of the binomial coefficients are uniformly distributed in $\mathbb{Z}_p^*$.

By a classical result of Legendre [**31**] a positive integer n can be represented as a sum of three squares, if and only if it is not of the form $n = 2^{2k}(8m + 7)$. This fact was used in [**22**] to prove that the set of integers n, for which $\binom{2n}{n}$ can be represented as a sum of three squares, has asymptotic density $\frac{7}{8}$ in the set of all natural numbers. The following corollary will prove that this is true for all binomial coefficients. We remark here that the set of positive integers not representable as a sum of three squares has density $\frac{5}{6}$.

Corollary 9. *The asymptotic density of binomial coefficients representable as a sum of three squares is $\frac{7}{8}$:*

$$\lim_{N\to\infty} \frac{2}{N^2} \# \left\{ (n,k) : 0 \leq k \leq n < N \text{ and } \binom{n}{k} \textit{is a sum of three squares} \right\} = \frac{7}{8}. \tag{4.18}$$

Proof of Theorem 7. The proof will follow similar lines as in the proof of Theorem 6. The main difference will be that we will have to incorporate the effects of the carries into the definition of the matrices. We make use of the fact that $\mathbb{Q}_p^* \simeq \mathbb{Z} \times \mathbb{Z}_p^*$. Therefore the characters in $\widehat{\mathbb{Q}_p^*}$ can be written as

$$\chi_\zeta(x) = \zeta^{v_p(x)} \chi\left[(x)_{(p)} \bmod p^\ell\right], \tag{4.19}$$

where χ is a Dirichlet character modulo p^ℓ for some $\ell \geq 1$ and ζ is a complex number of modulus 1. We write G_ℓ for the Dirichlet characters modulo p^ℓ on $\mathbb{Z}_p^*$. For our purpose it suffices to use characters χ_ζ with ζ an m-th root of unity. Using the above notation we have

$$\#\left\{(n,k) : 0 \leq k \leq n < N,\ v_p\left(\binom{n}{k}\right) \equiv r \bmod m, \text{ and } \binom{n}{k}_{(p)} \equiv a \bmod p^\ell \mathbb{Z}_p\right\}$$
$$= \frac{1}{m\phi(p^\ell)} \sum_{\substack{\zeta \\ \zeta^m=1}} \zeta^{-r} \sum_{\chi \in G_\ell} \overline{\chi(a)} \sum_{0 \leq k \leq n < N} \chi_\zeta\left(\binom{n}{k}\right).$$

It now remains to show that

$$\lim_{N\to\infty} \frac{2}{N^2} \sum_{0 \leq k \leq n < N} \chi_\zeta\left(\binom{n}{k}\right) = 0$$

for all non-trivial characters of the type introduced above.

We fix $\ell \geq 1$ and a non-principal character χ_ζ. Using the notation of Lemma 2 we define the matrix $U = U_{\chi_\zeta}$ of dimension $p^{3\ell}$ with $A, B, C, A', B', C' \in \mathcal{B}_\ell$ by

$$u_{(A,B,C),(A',B',C')} = \tag{4.20}$$
$$(\zeta\chi(-1))^{e(A,B,A',B')} \prod_{i=0}^{2\ell-1} \chi\left(\frac{((AA')_i!)_p}{((BB')_i!)_p((CC')_i!)_p}\right) \prod_{i=0}^{\ell-1} \overline{\chi}\left(\frac{((A')_i!)_p}{((B')_i!)_p((C')_i!)_p}\right),$$

if either $BB' + CC' \equiv AA' \bmod p^{2\ell}$ or $BB' + CC' + 1 \equiv AA' \bmod p^{2\ell}$, where $e(A, B, A', B')$ denotes the number of carries in the subtraction $(AA')_{\ell-1} - (BB')_{\ell-1}$. Otherwise we set $u_{(A,B,C),(A',B',C')} = 0$. We define

$$f\left(A_0A_1 \dots A_t; B_0B_1 \dots B_t; C_0C_1 \dots C_t\right) = \tag{4.21}$$
$$u_{(A_0,B_0,C_0),(A_1,B_1,C_1)} u_{(A_1,B_1,C_1),(A_2,B_2,C_2)} \cdots u_{(A_{t-1},B_{t-1},C_{t-1}),(A_t,B_t,C_t)},$$

and

$$F_{A,B,C}(N) = \sum_{An,Bk,Cr<N} f(An; Bk; Cr).$$

Notice first that by definition of f only arguments $(n; k; r)$ with $k + r = n$ yield a non-zero value of f. Moreover, the two products in (4.20) yield the product in Granville's formula (4.1); similarly to Lemmas 1 and 3 the considerations of couples of blocks permits to take into account the effect of the carries between two consecutive blocks. In (4.21) the exponent $e(A, B, A', B')$ gives $\sum_{k=0}^{t-1} e(A_k, B_k, A_{k+1}, B_{k+1}) = e_{\ell-1}$ with the notation of (4.1). For the contribution of ζ we will have to take care of the carries within the first $\ell-1$ digits separately. For this purpose we define $h(B, C)$ to be the number of carries in the addition of B and C.

Then by Granville's congruence (4.1) and (4.21) we have for $N > p^\ell$

$$F(N) = \sum_{A,B,C \in \mathcal{B}_\ell} \delta(A, B, C)\zeta^{h(B,C)} F_{A,B,C}(N) = \sum_{0 \leq k \leq n < N} \chi_\zeta\left(\binom{n}{k}\right),$$

where $\delta(A,B,C)=1$, if $B+C\equiv A \bmod p^\ell$ and $\delta(A,B,C)=0$ otherwise. Since f as defined above is a "trivariate block-multiplicative function" we have similarly to Lemma 3 (we omit all sums, which are obviously zero)

$$\begin{aligned} F_{A,B,C}(DN) = &\sum_{A',B',C'\in\mathcal{B}_\ell}\sum_{n,k,r<N} f(AA'n;BB'k;CC'r) + \sum_{\substack{A'<D\\ B',C'\in\mathcal{B}_\ell}}\sum_{k,r<N} f(AA'N;BB'k;CC'r)\\ &+\sum_{\substack{A',B'<D\\ C'\in\mathcal{B}_\ell}}\sum_{r<N} f(AA'N;BB'N;CC'r) + \sum_{\substack{A',C'<D\\ B'\in\mathcal{B}_\ell}}\sum_{k<N} f(AA'N;BB'k;CC'N)\\ &+\sum_{A',B',C'<D} f(AA'N;BB'N;CC'N). \end{aligned} \tag{4.22}$$

Since only arguments $(n;k;r)$ give non-zero value, if $k+r=n$ the second sum is $\mathcal{O}(N)$; the third to fifth sums are trivially $\mathcal{O}(N)$. Thus we only have to treat the first sum. As in the proof of Theorem 6 we have to show that the dominating eigenvalue of U_{χ_ζ} is strictly smaller in modulus than the dominating eigenvalue of U_ε (ε being the trivial character). By reordering the entries of the matrix U_ε we can reach the following form, where $\mathbf{0}$ and $\mathbf{1}$ denote blocks of 0's and 1's of the size indicated

$$\begin{array}{rc} B+C=A & \{\\ B+C=A+p^\ell & \{\\ B+C+1=A & \{\\ B+C+1=C+p^\ell & \{ \end{array} \begin{pmatrix} \mathbf{1}_{(a\times a)} & \mathbf{1}_{(a\times b)} & \mathbf{0}_{(a\times b)} & \mathbf{0}_{(a\times a)} & \mathbf{0}\\ \mathbf{0}_{(b\times a)} & \mathbf{0}_{(b\times b)} & \mathbf{1}_{(b\times b)} & \mathbf{1}_{(b\times a)} & \mathbf{0}\\ \mathbf{1}_{(b\times a)} & \mathbf{1}_{(b\times b)} & \mathbf{0}_{(b\times b)} & \mathbf{0}_{(b\times a)} & \mathbf{0}\\ \mathbf{0}_{(a\times a)} & \mathbf{0}_{(a\times b)} & \mathbf{1}_{(a\times b)} & \mathbf{1}_{(a\times a)} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix},$$

where $a=\frac{p^\ell(p^\ell+1)}{2}$ and $b=\frac{p^\ell(p^\ell-1)}{2}$. We note that the sub-matrix of the first $2p^{2\ell}$ lines and columns is primitive, since its square has only non-zero entries.

As in the proof of Theorem 6 we use

$$u^{\chi_\zeta}_{(0,0,0),(n,n-1,1)}u^{\chi_\zeta}_{(n,n-1,1),(0,0,0)} = \chi_\zeta\left(\binom{p^\ell n}{p^\ell}\right).$$

In the case that $\chi\neq\varepsilon$ we take summation over all prime residue classes modulo p^ℓ and use Lemma 6 to prove that this entry in $U^2_{\chi_\zeta}$ is less than the corresponding entry in U_ε. In the case that $\chi=\varepsilon$ and $\zeta\neq 1$ we sum the entries for $n=1$ and $n=p$. Then an application of Lemma 4 finishes the proof. □

Remark 10. It follows from Corollary 3 and the above proof that we can bound the implicit error term in Theorem 7 by $o\left(N^{\frac{1}{\ell}\log_p\lambda-2}\right)$ for any λ larger than the moduli of all the eigenvalues of the matrices U_{χ_ζ}.

5. Extension to multinomial coefficients

The two main tools in Sections 3 and 4 are Kummer's formula for the valuation of binomial coefficients and Granville's extension of Lucas' congruence. For multinomial coefficients there exists a formula generalizing Kummer's result [**27, 37**]: the p-valuation of $\binom{n}{h_1,h_2,\ldots,h_t}$ is the number of carries (counted with multiplicity) in the addition $h_1+\cdots+h_t$ in base p.

Similarly, Granville's congruence (4.1) can be generalized to multinomial coefficients by replacing m and r by $h_1, \dots, h_t$.

The main ingredient for the proof of Theorem 5 was the stability of the set of functions $b_{\mathbf{L}}^{\mathbf{T}}$ generated by the functions a_ℓ. The combinatorics behind the carries in the addition $h_1 + \cdots + h_t$ is more complicated but again gives rise to a finite set of functions generalizing the a_ℓ's; after this the same procedure can be applied to derive a result similar to Theorem 5, where α has to be replaced by $\log_p \binom{p+t-1}{t}$. The case $j = 0$ has been treated in [**3**].

Generalizations of Theorem 7 and the case $j = 0$ Theorem 6 are quite immediate adaptations of the proofs above. For general j in Theorem 6 we expect that it could be done along the same lines with some more efforts.

Acknowledgment. We are indebted to an anonymous referee for having drawn our attention to the paper [**22**], which led us to a more general formulation of Theorem 7 and to Corollary 9. Furthermore, we are grateful for helpful remarks concerning the presentation of the paper.

References

[**1**] J.-P. Allouche and V. Berthé, Triangle de Pascal, complexité et automates, *Bull. Belg. Math. Soc. Simon Stevin* 4 (1997), 1–23, Journées Montoises (Mons, 1994).

[**2**] J.-P. Allouche, F. von Haeseler, H.-O. Peitgen, and G. Skordev, Linear cellular automata, finite automata and Pascal's triangle, *Discrete Appl. Math.* 66 (1996), 1–22.

[**3**] D. Barbolosi and P. J. Grabner, Distribution des coefficients multinomiaux et q-binomiaux modulo p, *Indag. Math.* 7 (1996), 129–135.

[**4**] R. Bellman and H. N. Shapiro, On a problem in additive number theory, *Ann. of Math. (2)* 49 (1948), 333–340.

[**5**] L. Carlitz, The distribution of binomial coefficients (mod p), *Arch. Math.* 14 (1963), 297–303.

[**6**] ———, The number of binomial coefficients divisible by a fixed power of a prime, *Rend. Circ. Matem. Palermo* 16 (1967), 299–320.

[**7**] E. Cateland, Suites digitales et suites k-régulières, Ph.D. thesis, Université Bordeaux I, 1992.

[**8**] K. S. Davis and W. Webb, Lucas congruence for prime powers, *European J. Combin.* 11 (1990), 229–233.

[**9**] K. S. Davis and W. Webb, A binomial coefficient congruence modulo prime powers, *J. Number Th.* 43 (1993), 20–23.

[**10**] H. Delange, Sur les fonctions q-additive ou q-multiplicatives, *Acta Arith.* 21 (1972), 285–298.

[**11**] ———, Sur la fonction sommatoire de la fonction "somme des chiffres", *l'Enseignement Math.(2)* 21 (1975), 31–47.

[**12**] L. E. Dickson, *History of the Theory of Numbers I*, Chelsea Publ., New York, 1971.

[**13**] M. Drmota, The distribution of patterns in digital expansions, *Algebraic Number Theory and Diophantine Analysis* (Graz 1998) (F. Halter-Koch and R. F. Tichy, eds., W. de Gruyter, Berlin, 2000), pp. 103–121.

[**14**] P. Flajolet, P. J. Grabner, P. Kirschenhofer, H. Prodinger, and R. F. Tichy, Mellin transforms and asymptotics: digital sums, *Theor. Comput. Sci.* 123 (1994), 291–314.

[**15**] E. Fouvry and C. Mauduit, Sommes des chiffres et nombres presque premiers, *Math. Ann.* 305 (1996), 571–599.

[**16**] Z. M. Franco, Distribution of binomial coefficients modulo p, *Number Theory* (Eger, 1996) (K. Győry, A. Pethő, and V. T. Sós, eds., de Gruyter, Berlin, 1998), pp. 199–209.

[**17**] R. Garfield and H. S. Wilf, The distribution of the binomial coefficients modulo p, *J. Number Th.* 41 (1992), 1–5.

[**18**] A. O. Gelfond, Sur les nombres qui ont des propriétés additives et multiplicatives données, *Acta Arith.* 13 (1968), 259–266.
[**19**] P. J. Grabner, Completely q-multiplicative functions: the Mellin-transform approach, *Acta Arith.* 65 (1993), 85–96.
[**20**] A. Granville, Zaphod Beeblebrox's brain and the fifty-ninth row of Pascal's triangle, *Amer. Math. Monthly* 99 (1992), 318–331, *Corrigendum: ibid.* 104 (1997), 848–851.
[**21**] ———, Arithmetic properties of binomial coefficients. I. Binomial coefficients modulo prime powers, *Organic Mathematics* (Burnaby, BC, 1995) (J. Borwein, P. Borwein, L. Jörgenson, and R. Corless, eds.), Amer. Math. Soc., Providence, RI, 1997, pp. 253–276,
see also: `http://www.math.uga.edu:80/~andrew/Binomial/index.html`.
[**22**] A. Granville and Y. Zhu, Representing binomial coefficients as sums of squares, *Amer. Math. Monthly* 97 (1990), 486–493.
[**23**] H. Harborth, Number of odd binomial coefficients, *Proc. Amer. Math. Soc.* 62 (1977), 19–22.
[**24**] E. Hexel and H. Sachs, Counting residues modulo a prime in Pascal's triangle, *Indian J. Math.* 20 (1978), 91–105.
[**25**] F. T. Howard, The number of binomial coefficients divisible by a fixed power of 2, *Proc. Amer. Math. Soc.* 29 (1971), 236–242.
[**26**] ———, Formulas for the number of binomial coefficients divisible by a fixed power of a prime, *Proc. Amer. Math. Soc.* 37 (1973), 358–362.
[**27**] ———, The number of multinomial coefficients divisible by a fixed power of a prime, *Pacific J. Math.* 50 (1974), 99–108.
[**28**] E. E. Kummer, Über die Ergänzungssätze zu den allgemeinen Reciprocitätsgesetzen, *J. reine angew. Math.* 44 (1852), 93–146.
[**29**] E. Lange, H.-O. Peitgen, and G. Skordev, Fractal patterns in Gaussian and Stirling number tables, *Ars Combin.* 48 (1998), 3–26.
[**30**] G. Larcher, On the number of odd binomial coefficients, *Acta Math. Hungar.* 71 (1996), 183–203.
[**31**] A.-M. Legendre, *Théorie des Nombres*, Paris, 1798.
[**32**] É. Lucas, Sur les congruences des nombres eulériens et des coefficients différentiels des fonctions trigonométriques, suivant un module premier, *Bull. Soc. Math. France* 6 (1878), 49–54.
[**33**] E. Manstavičius. Probabilistic theory of additive functions related to systems of numeration. In *New Trends in Probability and Statistics, Vol. 4 (Palanga, 1996)*, pages 413–429. VSP, Utrecht, 1997.
[**34**] J.-L. Mauclaire and L. Murata, On q-additive functions, II, *Proc. Japan Acad.* 59 (1983), 441–444.
[**35**] E. Seneta, *Non-Negative Matrices*, Halsted Press, New York, 1973.
[**36**] D. Singmaster, Notes on binomial coefficients, I – A generalization of Lucas' congruence, II – The least n such that p^e divides an r-nomial coefficient of rank n, III – Any integer divides almost all binomial coefficients, *J. London Math. Soc.* 8 (1974), 545–548, 549–554, 555–560.
[**37**] J.-P. Soublin, Une congruence pour les coefficients binomiaux, *C. R. Math. Acad. Sci. Soc. R. Can.* 21 (1999), 6–8.
[**38**] A. H. Stein, Exponential sums related to binomial coefficient parity, *Proc. Amer. Math. Soc.* 80 (1980), 526–530.
[**39**] ———, Exponential sums of sum-of-digits functions, *Ill. J. Math.* 30 (1986), 660–675.
[**40**] ———, Binomial coefficients not divisible by a prime, *Number Theory* (New York, 1985/1988) (D. V. Chudnovsky, G. V. Chudnovsky, H. Cohn, and M. B. Nathanson, eds.), Lecture Notes in Mathematics, vol. 1383, Springer, Berlin, 1989, pp. 170–177.
[**41**] K. B. Stolarsky, Power and exponential sums of digital sums related to binomial coefficient parity, *SIAM J. Appl. Math.* 32 (1977), 717–730.
[**42**] G. Tenenbaum, Sur la non-dérivabilité de fonctions périodiques associées à certaines formules sommatoires, *The Mathematics of Paul Erdős I* (R. L. Graham and J. Nešetřil, eds.), Algorithms Comb., vol. 13, Springer, Berlin, 1997.
[**43**] W. Webb, The number of binomial coefficients in residue classes modulo p and p^2, *Colloq. Math.* 60/61 (1990), 275–280.

[**44**] S. Wolfram, Geometry of binomial coefficients, *Amer. Math. Monthly* 91 (1984), 566–571.

(G. B.)
20, rue Fourcroy
75017 Paris
France
E-mail address: barat@finanz.math.tu-graz.ac.at

(P. G.)
Institut für Mathematik A
Technische Universität Graz
Steyrergasse 30
8010 Graz
Austria
E-mail address: peter.grabner@tugraz.at