

# ASYMPTOTIC PRIME-POWER DIVISIBILITY OF BINOMIAL, GENERALIZED BINOMIAL, AND MULTINOMIAL COEFFICIENTS

JOHN M. HOLTE

**ABSTRACT.** This paper presents asymptotic formulas for the abundance of binomial, generalized binomial, multinomial, and generalized multinomial coefficients having any given degree of prime-power divisibility. In the case of binomial coefficients, for a fixed prime  $p$ , we consider the number of  $(x, y)$  with  $0 \leq x, y < p^n$  for which  $\binom{x+y}{x}$  is divisible by  $p^{zn}$  (but not  $p^{zn+1}$ ) when  $zn$  is an integer and  $\alpha < z < \beta$ , say. By means of a classical theorem of Kummer and the probabilistic theory of large deviations, we show that this number is approximately  $p^{nD((\alpha, \beta))}$ , where  $D((\alpha, \beta)) := \sup\{D(z) : \alpha < z < \beta\}$  and  $D$  is given by an explicit formula. We also develop a “ $p$ -adic multifractal” theory and show how  $D$  may be interpreted as a multifractal spectrum of divisibility dimensions. We then prove that essentially the same results hold for a large class of the generalized binomial coefficients of Knuth and Wilf, including the  $q$ -binomial coefficients of Gauss and the Fibonomial coefficients of Lucas, and finally we extend our results to multinomial coefficients and generalized multinomial coefficients.

## INTRODUCTION

The purpose of this paper is to present asymptotic formulas regarding the prime-power divisibility of binomial coefficients and their generalizations that follow from a classical theorem of Kummer and its generalizations. Specifically, for a fixed prime  $p$  and large  $n$ , we consider initially how many binomial coefficients  $\binom{x+y}{x}$  with  $0 \leq x, y < p^n$  are divisible by  $p^{zn}$  (but not  $p^{zn+1}$ ) when  $zn$  is an integer and  $\alpha < z < \beta$ , say. We will find that this number is approximately  $p^{nD((\alpha, \beta))}$ , where  $D((\alpha, \beta)) := \sup\{D(z) : \alpha < z < \beta\}$  and  $D$  is a function we will determine explicitly.

First we will derive the exact answer by combinatorial means for ordinary binomial coefficients, and then, starting afresh, we will find the asymptotic result by means of the probabilistic theory of large deviations. In yet a third approach, we will show that the function  $D$  represents the “ $p$ -adic multifractal spectrum” of divisibility dimensions.

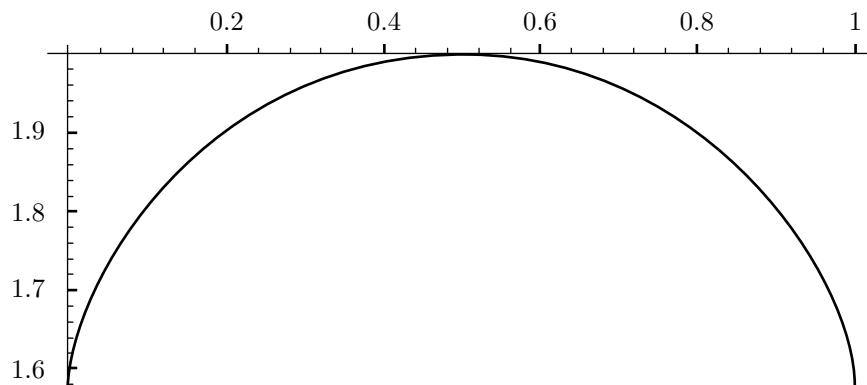
Then we will show that essentially the same results hold for a large class of generalized binomial coefficients, as defined by Knuth and Wilf [32], by means of their

---

Received by the editors December 8, 1995 and, in revised form, March 26, 1996.

1991 *Mathematics Subject Classification.* Primary 05A16; Secondary 11K16.

*Key words and phrases.* Binomial coefficients, generalized binomial coefficients, multinomial coefficients, asymptotic enumeration, prime-power divisibility, large deviation principle, carries, Markov chains, multifractals, fractals,  $p$ -adic numbers.

FIGURE 1. The function  $D(z)$  for  $m = 2$  and  $p = 2$ 

generalization of Kummer's theorem; special cases include the  $q$ -binomial coefficients of Gauss and the Fibonomial coefficients of Lucas. Passing from binomials to multinomials, we will apply Dickson's generalization of Kummer's theorem to obtain an asymptotic prime-power divisibility result for multinomial coefficients. Finally, we will subsume all these results under the rubric of generalized multinomial coefficients.

**Notation.** Throughout this paper,  $p$  will denote a fixed prime. For any positive integer  $m$ , let  $\text{ord}_p m$  denote the number of times  $p$  occurs as a factor of  $m$ . For real  $z$  and nonnegative integer  $n$  (and nonnegative integers  $x$  and  $y$ ), let  $N_n(z)$  denote the number of  $(x, y)$  with  $0 \leq x, y < p^n$  for which  $\text{ord}_p \binom{x+y}{x} = zn$ . We will soon see that  $N_n(z) = 0$  unless  $z \in \{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$ . If  $\Gamma$  is a set of real numbers, let

$$N_n(\Gamma) = \sum_{z \in \Gamma} N_n(z) = \# \left\{ (x, y) \mid 0 \leq x, y < p^n \text{ and } \frac{1}{n} \text{ord}_p \binom{x+y}{x} \in \Gamma \right\}.$$

Similarly, to deal with multinomial coefficients, let

$$N_n(\Gamma) = \left\{ (x_1, \dots, x_m) \mid 0 \leq x_1, \dots, x_m < p^n \text{ and } \frac{1}{n} \text{ord}_p \binom{x_1 + \dots + x_m}{x_1, \dots, x_m} \in \Gamma \right\}.$$

**Summary of Results.** In Theorem 1 we will present an exact formula for  $N_n(\Gamma)$ . In Theorems 2 and 5 we will show that for every nondegenerate interval  $\Gamma \subseteq [0, m-1]$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_p N_n(\Gamma) = \sup_{z \in \Gamma} D(z),$$

where  $D$  is a concave “dimension” function that is finite and continuous on  $[0, m-1]$  and symmetric about  $(m-1)/2$ . See Figure 1. Moreover, Theorem 2 will give an elementary formula for  $D$  in the binomial case. In Theorem 4 and the “Main Theorem” we will show that the conclusion of Theorem 2 also holds for a wide class of generalized binomial and multinomial coefficients. Meanwhile, a new, “ $p$ -adic” multifractal formalism will be developed (see Section 2, esp. Proposition 2), and Theorems 3 and 6 will establish that the function  $D$  gives the multifractal spectrum of divisibility dimensions.

## 1. BINOMIAL COEFFICIENTS

**1.1. Kummer's Theorem.** The basis for evaluating  $\text{ord}_p\binom{x+y}{x}$  is the following 1852 theorem of Kummer ([34], p. 116).

**Theorem A.** *If  $p$  is a prime, then  $\text{ord}_p\binom{x+y}{x}$  is equal to the number of carries (or, the total of the carries) that occur when  $x$  and  $y$  are added in base- $p$  arithmetic.*

For example, consider the highest power of 3 that divides  $\binom{88}{41} = \binom{41+47}{41}$ . In base 3 we have:

	Carries	1	1	0	1	0
41	Addend		1	1	1	2
+47	Augend	+	1	2	0	2
88	Sum		1	0	0	2

There are 3 carries, so  $\text{ord}_3\binom{41+47}{41} = 3$ . (In fact,  $\binom{88}{41} = 3^3 \cdot 2^5 \cdot 5 \cdot 7^2 \cdot 11 \cdot 17 \cdot 29 \cdot 43 \cdot 53 \cdot 59 \cdot 61 \cdot 67 \cdot 71 \cdot 73 \cdot 79 \cdot 83$ .)

More generally, if  $0 \leq x, y < p^n$ , let  $\{X_k\}$  and  $\{Y_k\}$  denote their base- $p$  digits:

$$x = X_{n-1}p^{n-1} + \cdots + X_1p + X_0$$

and

$$y = Y_{n-1}p^{n-1} + \cdots + Y_1p + Y_0$$

with  $0 \leq X_k, Y_k < p$ . Let  $\{C_k\}$  denote the carries in their base- $p$  addition:

Carries	$C_n$	$C_{n-1}$	$\cdots$	$C_1$	$C_0 = 0$
Addend		$X_{n-1}$	$\cdots$	$X_1$	$X_0$
Augend	+	$Y_{n-1}$	$\cdots$	$Y_1$	$Y_0$
Sum	$S_n$	$S_{n-1}$	$\cdots$	$S_1$	$S_0$

Let  $Z_n = (C_1 + \cdots + C_n)/n$ , the average carry. Then Kummer's theorem tells us that  $\text{ord}_p\binom{x+y}{x} = C_1 + \cdots + C_n = nZ_n$ . Notice that we always have  $Z_n \in \{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$ .

**1.2. A Probabilistic Perspective.** Assume that the digits  $\{X_k\}, \{Y_k\}$  are independent random variables uniformly distributed on  $\{0, 1, \dots, p-1\}$ . Then  $x = \sum_{k=0}^{n-1} X_k p^k$  and  $y = \sum_{k=0}^{n-1} Y_k p^k$  are (uniform) random numbers in  $\{0, 1, \dots, p^n - 1\}$ , and  $N_n(z)$  can be expressed in terms of probabilities: If  $z = h/n$ , then

$$\begin{aligned} N_n(z) &= \#\left\{(x, y) \mid 0 \leq x, y < p^n, \text{ord}_p\binom{x+y}{x} = h\right\} \\ &= \Pr(nZ_n = h) \cdot (p^n)^2 = \Pr(Z_n = z) \cdot p^{2n}, \end{aligned}$$

and, if  $\Gamma \subseteq \mathbb{R}$ ,

$$(1) \quad N_n(\Gamma) = p^{2n} \Pr(Z_n \in \Gamma).$$

**Key observation:** Because  $C_{k+1}$  depends only on  $C_k, X_k$ , and  $Y_k$ , the carries form a finite Markov chain:

$$\Pr(C_{k+1} = c_{k+1} \mid C_k = c_k, \dots, C_1 = c_1, C_0 = 0) = \Pr(C_{k+1} = c_{k+1} \mid C_k = c_k).$$

The state space of this Markov chain is  $\{0, 1\}$ , the set of possible carries, and the transition matrix is  $\mathbf{\Pi} = [\pi_{ij}]$  where

$$\pi_{ij} = \Pr(\text{carry-out} = j \mid \text{carry-in} = i).$$

(For background on Markov chains, see, e.g., [4].)

Consider the addition in the  $k^{\text{th}}$  place:

$$\begin{array}{r} C_{k+1} \quad C_k \\ \quad \quad X_k \\ + \quad \quad Y_k \\ \hline \end{array}$$

There will be a carry out ( $C_{k+1} = 1$ ) if and only if  $C_k + X_k + Y_k \geq p$ . By counting cases, we get

$$\begin{bmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{bmatrix} = p^{-2} \begin{bmatrix} p(p+1)/2 & p(p-1)/2 \\ p(p-1)/2 & p(p+1)/2 \end{bmatrix},$$

or,

$$(2) \quad \mathbf{\Pi} = \frac{1}{2p} \begin{bmatrix} p+1 & p-1 \\ p-1 & p+1 \end{bmatrix}.$$

Notice that this is an ergodic Markov chain with stationary vector  $\mathbf{v}' = (\frac{1}{2}, \frac{1}{2})$ :  $\mathbf{v}'\mathbf{\Pi} = \mathbf{v}'$ . Therefore, with probability 1,

$$(3) \quad Z_n \longrightarrow \frac{1}{2} \text{ as } n \longrightarrow \infty.$$

Thus, for large  $n$ ,  $N_n(\frac{1}{2})/p^{2n} \approx 1$ : predominantly, the highest power of  $p$  that divides the binomial coefficients in a  $p^n \times p^n$  block of Pascal's triangle is half the maximum possible.

**1.3. A Combinatorial Formula.** For completeness, we will derive a formula for  $N_n(h/n)$ . Although we could follow up on this with an asymptotic analysis via Stirling's formula, we will instead develop the asymptotics via large deviation theory in the following sections.

It should be noted that Carlitz ([2], p. 302) has also given a formula for the number of  $(x, y)$  with  $0 \leq x, y < p^n$  for which  $\text{ord}_p\binom{x+y}{x} = h$ , but assuming in addition that  $x+y < p^n$ , and that Knuth ([31], Section 4.3.1) has given a generating function for  $N_n(h/n)$ .

**Theorem 1.** *Let  $p$  be a prime. Let  $h$  and  $n$  be nonnegative integers. The number of entries in the first  $p^n \times p^n$  block of Pascal's triangle divisible by  $p^h$  but not  $p^{h+1}$  is*

$$N_n(h/n) = \sum_{k=0}^h \sum_{i=0}^{n-h} \binom{i+k}{k} \binom{h-1}{h-k} \binom{n-h}{i} \left(\frac{p-1}{2}\right)^{i+k} p^n.$$

*Proof.* By Kummer's theorem,  $N_n(h/n)$  is the number of pairs of  $n$ -digit base- $p$  numbers whose base- $p$  addition yields exactly  $h$  carries. Notice that each column in such an addition may be classified as follows:

I	(carry is impossible)	$X_j + Y_j < p - 1$	# cases = $p(p-1)/2$
M	(carry may or may not occur)	$X_j + Y_j = p - 1$	# cases = $p$
K	(carry is certain)	$X_j + Y_j > p - 1$	# cases = $p(p-1)/2$

For example, the addition of two numbers having  $n = 27$  digits each might lead to the following sequence having  $h = 10$  carries:

$$\overbrace{\text{MMMMMII}}^{\text{no carries}} \underbrace{\text{KK}}_2 \overbrace{\text{III}}^0 \underbrace{\text{MMMM}}_{5 \text{ ripple carries}} \underbrace{\text{KKK}}_3 \overbrace{\text{IIIIIIII}}^{\text{no carries}}$$

First we will count the number of sequences of  $i$   $I$ 's,  $m$   $M$ 's, and  $k$   $K$ 's, where  $i + m + k = n$ , for which the number of carries is exactly  $h$ . Let us start by arranging the  $i$   $I$ 's and  $k$   $K$ 's; this may be done in  $\binom{i+k}{k}$  ways. Next we put in the  $m$   $M$ 's so as to achieve exactly  $r = h - k$  ripple carries. Thus, we must put  $r$   $M$ 's to the left of various  $K$ 's (without any intervening  $I$ 's)—this may be done in  $\binom{r+k-1}{r} = \binom{h-k+k-1}{h-k} = \binom{h-1}{h-k}$  ways—and we must put the remaining  $m - r$   $M$ 's to the left of various  $I$ 's (without intervening  $K$ 's) or at the far right end of the string, which may be done in  $\binom{(m-r)+(i+1)-1}{m-r} = \binom{m-r+i}{i} = \binom{n-h}{i}$  ways. So the number of arrangements of  $i$   $I$ 's,  $k$   $K$ 's, and  $m$   $M$ 's yielding exactly  $h$  carries is

$$\binom{i+k}{k} \binom{h-1}{h-k} \binom{n-h}{i}.$$

Next we factor in the  $p(p-1)/2$  ways to choose digit pairs for each  $I$  and  $K$  case and the  $p$  choices for each  $M$ . Finally, summing over the possible  $k$  and  $i$  values, we get

$$N_n(h/n) = \sum_{k=0}^h \sum_{i=0}^{n-h} \binom{i+k}{k} \binom{h-1}{h-k} \binom{n-h}{i} \left( \frac{p(p-1)}{2} \right)^{i+k} p^{n-i-k}. \quad \square$$

As an immediate corollary of this theorem, we find that

$$\Pr(Z_n = h) = \sum_{k=0}^h \sum_{i=0}^{n-h} \binom{i+k}{k} \binom{h-1}{h-k} \binom{n-h}{i} \left( \frac{p-1}{2} \right)^{i+k} p^{-n}.$$

**1.4. Large Deviation Principle for Markov Chains.** The theory of large deviations provides the key technique for our asymptotic analysis. The following definitions therefrom will suffice for this paper.

**Definitions.** A *rate function*  $I$  is a lower semicontinuous mapping  $I : \mathbb{R} \rightarrow [0, \infty]$ . If  $I$  is a rate function and  $\Gamma \subseteq \mathbb{R}$ , then  $I(\Gamma)$  denotes  $\inf_{z \in \Gamma} I(z)$ . We say that a sequence of measures  $\langle \mu_n \rangle$  on the Borel sets of  $\mathbb{R}$  satisfies the *large deviation principle* with rate function  $I$  if, for every Borel set  $\Gamma$ ,

$$\begin{aligned} -I(\Gamma^\circ) &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(\Gamma) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(\Gamma) \\ &\leq -I(\bar{\Gamma}). \end{aligned}$$

The set  $\Gamma$  is an  *$I$ -continuity set* if  $I(\Gamma^\circ) = I(\bar{\Gamma})$ ; for such a set

$$\lim_{n \rightarrow \infty} \log \mu_n(\Gamma) = -I(\Gamma).$$

For Markov chains, an early version of the large deviation principle was given by Miller [40]; see also [8], [9], and [10], pp. 288–291. The Markov chain theorem we use is that given in Dembo and Zeitouni [7], p. 60.

We shall also make use of the Perron-Frobenius theory ([42], [19]) of nonnegative matrices (see [48] for a modern treatment). The “Perron-Frobenius” eigenvalue of such a matrix is the (necessarily nonnegative) eigenvalue of largest magnitude.

**Theorem B.** *Let  $\langle \xi_k \rangle$  be a finite Markov chain having irreducible transition matrix  $\mathbf{\Pi} = [\pi_{ij}]$ , and let  $f$  be a deterministic real-valued function on the state space. Define*

$$\mathbf{\Pi}_t = \left[ \pi_{ij} e^{tf(j)} \right],$$

$$\rho(\mathbf{\Pi}_t) = \text{Perron-Frobenius eigenvalue of } \mathbf{\Pi}_t,$$

$$I(z) = \sup_t \{tz - \log \rho(\mathbf{\Pi}_t)\},$$

and

$$Z_n = \frac{f(\xi_1) + \cdots + f(\xi_n)}{n}.$$

Then  $I$  is a convex function, and for any starting state  $\sigma$ , the measures defined by

$$\mu_n(\Gamma) = \Pr(Z_n \in \Gamma \mid \xi_0 = \sigma)$$

satisfy the large deviation principle.

The rate function  $I$  is the convex conjugate, or Fenchel-Legendre transform, of  $\log(\mathbf{\Pi}_t)$  (see [16]; [51], sect. 4.6; or [47], sect. 12). Thus, this theorem says that the time-averaged value of a function of a Markov chain has a probability distribution which satisfies the large deviation principle with rate function given by the Fenchel-Legendre transform of the logarithm of the spectral radius of  $\mathbf{\Pi}_t$ .

We will apply this theorem with  $\langle \xi_k \rangle$  the carries process  $\langle C_k \rangle$ ,  $\mathbf{\Pi}$  its transition matrix, and  $f$  the identity function, so that  $Z_n = (C_1 + \cdots + C_n)/n$ . In our case,  $C_0 = 0$  necessarily, and so we write  $\Pr(Z_n \in \Gamma)$  for  $\Pr(Z_n \in \Gamma \mid C_0 = 0)$ , but we note for later use that the following proposition would also be true if we had  $C_0 = 1$ , because the theorem allows an arbitrary starting state. The theorem implies that

$$(4) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log \Pr(Z_n \in \Gamma) = -I(\Gamma)$$

for each  $I$ -continuity set  $\Gamma$ , where  $I$  is defined as in the theorem. By equation (1),  $\Pr(Z_n \in \Gamma) = p^{2n} N_n(\Gamma)$ , and so (4) becomes

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_p N_n(\Gamma) = D(\Gamma)$$

where  $D(\Gamma) := 2 - I(\Gamma)/\log p$ . Working backwards from this definitional equation, we define  $D(z) = 2 - I(z)/\log p$  and note that

$$(5) \quad D(\Gamma) = \sup_{z \in \Gamma} D(z).$$

We have proved the following.

**Proposition 1.** *Let  $p$  be a prime. Let*

$$\mathbf{\Pi}_t = \frac{1}{2p} \begin{bmatrix} p+1 & (p-1)e^t \\ p-1 & (p+1)e^t \end{bmatrix},$$

$$\rho(\mathbf{\Pi}_t) = \text{Perron-Frobenius eigenvalue of } \mathbf{\Pi}_t,$$

$$I(z) = \sup_t \{tz - \log \rho(\Pi_t)\},$$

and

$$D(z) = 2 - I(z)/\log p.$$

Then, for every  $I$ -continuity set  $\Gamma$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_p N_n(\Gamma) = D(\Gamma).$$

In the next section we will show that  $I$  is continuous on  $[0, 1]$ . Therefore all nondegenerate subintervals of  $[0, 1]$  are  $I$ -continuity sets.

**1.5. A Prime-power Divisibility Spectrum.** In this section we present our main theorem for binomial coefficients, which includes an explicit, elementary formula for the rate function,  $I$ , and the related function  $D$  which characterizes the asymptotic prime-power divisibility.

**Theorem 2.** *Let  $p$  be a prime. Then, for every nondegenerate interval  $\Gamma \subseteq [0, 1]$ ,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_p N_n(\Gamma) = D(\Gamma)$$

where

$$D(z) = 2 - \frac{I(z)}{\log p}, \quad D(\Gamma) = \sup_{z \in \Gamma} D(z),$$

and  $I(z)$  is defined as follows. For  $0 < z < 1$ ,

$$I(z) = \log \frac{4p}{p+1} + z \log u - \log \{1 + u + \sqrt{1 + cu + u^2}\},$$

where

$$c = 2 - 16p/(p+1)^2,$$

and

$$u = \frac{b + \operatorname{sgn}(z - \frac{1}{2})\sqrt{b^2 - 4a^2}}{2a},$$

where

$$a = (2 - c)z(1 - z)$$

and

$$b = 1 - 2cz(1 - z) - c^2 \left(z - \frac{1}{2}\right)^2 = 1 - c^2/4 - ca.$$

Also,  $I(0) = I(1) = \log \frac{2p}{p+1}$ , and  $I(z) = \infty$  for  $z < 0$  and for  $z > 1$ . Furthermore,  $I(z)$  is symmetric about  $z = \frac{1}{2}$  and is continuous and convex for  $0 \leq z \leq 1$ .

*Proof.* In view of Proposition 1 and the above discussion, it remains to prove the assertions concerning the rate function,  $I$ . The characteristic polynomial for  $\Pi_t$  is

$$\begin{aligned} \begin{vmatrix} \lambda - \frac{p+1}{2p} & -\frac{p-1}{2p}e^t \\ -\frac{p-1}{2p} & \lambda - \frac{p+1}{2p}e^t \end{vmatrix} &= \left(\lambda - \frac{p+1}{2p}\right) \left(\lambda - \frac{p-1}{2p}e^t\right) - \left(\frac{p-1}{2p}\right)^2 e^t \\ &= \lambda^2 - \frac{p+1}{2p}(1 + e^t)\lambda + \frac{e^t}{p}, \end{aligned}$$

and so the eigenvalues are given by

$$\begin{aligned}\lambda &= \frac{1}{2} \left\{ \frac{p+1}{2p} (1+e^t) \pm \sqrt{\left(\frac{p+1}{2p}\right)^2 (1+e^t)^2 - \frac{4e^t}{p}} \right\} \\ &= \frac{p+1}{4p} \left\{ 1+e^t \pm \sqrt{1 + \left[2 - \frac{16p}{(p+1)^2}\right] e^t + e^{2t}} \right\}.\end{aligned}$$

Let

$$c = 2 - 16p/(p+1)^2.$$

Since  $p \geq 2$ , we have  $-14/9 \leq c < 2$ , and so  $1 + ce^t + e^{2t}$  is always positive. Therefore the eigenvalues are real and the larger eigenvalue is

$$\rho(\Pi_t) = \frac{p+1}{4p} \left\{ 1 + e^t + \sqrt{1 + ce^t + e^{2t}} \right\}.$$

Now we have

$$\begin{aligned}I(z) &= \sup_t \{zt - \log \rho(\Pi_t)\} \\ &= \sup_t \left[ zt - \log \frac{p+1}{4p} \{1 + e^t + \sqrt{1 + ce^t + e^{2t}}\} \right] \\ &= -\log \frac{p+1}{4p} + \sup_{u>0} [z \log u - \log \{1 + u + \sqrt{1 + cu + u^2}\}] \\ &= \log \frac{4p}{p+1} + \sup_{u>0} \log \frac{u^z}{1 + u + \sqrt{1 + cu + u^2}}.\end{aligned}$$

If  $z < 0$ , then  $u^z = u^{-|z|} \rightarrow \infty$  as  $u \rightarrow 0+$ , and so  $I(z) = \infty$ . If  $z = 0$ , then  $I(z)$  is

$$(6) \quad I(0) = \log \frac{4p}{p+1} + \log \frac{1}{2} = \log \frac{2p}{p+1}.$$

If  $z \geq 1$ , then, writing

$$\frac{u^z}{1 + u + \sqrt{1 + cu + u^2}} = \frac{u^{z-1}}{u^{-1} + 1 + \sqrt{u^{-2} + cu^{-1} + 1}},$$

we see that  $I(z) = \infty$  for  $z > 1$  and  $I(1) = \log \frac{2p}{p+1}$ .

Now assume  $0 < z < 1$ . We will evaluate the supremum (actually, maximum) in the above formula for  $I(z)$  by elementary calculus. To simplify the formulas, let  $Q = 1 + cu + u^2$ . Then

$$\begin{aligned}&\frac{\partial}{\partial u} [z \log u - \log \{1 + u + \sqrt{Q}\}] \\ &= \frac{z}{u} - \frac{1 + \frac{c+2u}{2\sqrt{Q}}}{1 + u + \sqrt{Q}} \\ &= \frac{z(1 + u + \sqrt{Q})\sqrt{Q} - u\sqrt{Q} - \frac{1}{2}cu - u^2}{u(1 + u + \sqrt{Q})\sqrt{Q}} \\ &= \frac{\{z - (1-z)u\}\sqrt{Q} + (z - \frac{1}{2})Q + \frac{1}{2}(1 - u^2)}{u(1 + u + \sqrt{Q})\sqrt{Q}}.\end{aligned}$$

If  $0 < z < \frac{1}{2}$ , then, as  $u \rightarrow 0+$ , the numerator  $\rightarrow z + (z - \frac{1}{2}) \cdot 1 + \frac{1}{2} > 0$ , while as  $u \rightarrow 1-$ , the numerator  $\rightarrow (2z - 1)\sqrt{2 + c} + (z - \frac{1}{2})\sqrt{2 + c} + 0 < 0$ ; so there is



a maximizing point  $u$  between 0 and 1. Similarly, if  $\frac{1}{2} < z < 1$ , then, as  $u \rightarrow 1+$ , the numerator  $\rightarrow 3\sqrt{2+c}(z - \frac{1}{2}) > 0$ , while as  $u \rightarrow \infty$ , the numerator  $\rightarrow -\infty$ ; so there is also a value of  $u \in (1, \infty)$  where the derivative is zero. Now, setting the (numerator of the) derivative equal to zero, we get

$$(7) \quad \{z - (1-z)u\}\sqrt{Q} + (z - \frac{1}{2})Q + \frac{1}{2}(1-u^2) = 0.$$

Rearranging, squaring, and simplifying, we get

$$(8) \quad u\{au^2 - bu + a\} = 0,$$

where

$$a = (2-c)z(1-z)$$

and

$$b = 1 - 2cz(1-z) - c^2 \left(z - \frac{1}{2}\right)^2 = 1 - c^2/4 - ca.$$

The root  $u = \lim_{t \rightarrow -\infty} e^t = 0$  is extraneous, because it does not satisfy (7) when  $z \neq 0$ . The other roots of (8) are

$$u = u_{\pm} = \frac{b \pm \sqrt{b^2 - 4a^2}}{2a}.$$

After ruling out extraneous roots here, we find that  $u = u_-$  if  $0 < z \leq \frac{1}{2}$  and  $u = u_+$  if  $\frac{1}{2} \leq z < 1$ , which is the formula stated in the theorem.

Finally, we check the claimed properties of  $I$ . The large deviation theorem for Markov chains tells us that  $I$  is convex. (Alternatively, a tedious calculation will show  $I''(z) \geq 0$ .) Our formula for  $I(z)$  implies that it is symmetric about  $z = \frac{1}{2}$  and is continuous for  $0 < z < 1$ . To check continuity from the right at 0 (and, by symmetry, from the left at 1), we calculate

$$\lim_{z \rightarrow 0+} I(z) = \log \frac{4p}{p+1} + \lim_{z \rightarrow 0+} z \log u - \lim_{z \rightarrow 0+} \log\{1 + u + \sqrt{Q}\}.$$

As  $z \rightarrow 0+$ , we have  $a = (2-c)z(1-z) \rightarrow 0$  and  $b \rightarrow 1 - c^2/4 > 0$ . Also, for  $0 < z < \frac{1}{2}$ ,

$$u = \frac{b - \sqrt{b^2 - 4a^2}}{2a} = \frac{2a}{b + \sqrt{b^2 - 4a^2}}.$$

Hence, as  $z \rightarrow 0+$ , we have  $u \rightarrow 0$  and

$$\begin{aligned} z \log u &= z \log\{2(2-c)\} + z \log z + z \log(1-z) - z \log\{b + \sqrt{b^2 - 4a^2}\} \\ &\rightarrow 0 + 0 + 0 - 0 \cdot \log\{1 - c^2/4 + \sqrt{(1 - c^2/4)^2 - 0}\} = 0. \end{aligned}$$

Therefore,

$$\lim_{z \rightarrow 0+} I(z) = \log \frac{4p}{p+1} + 0 - \log\{1 + 0 + \sqrt{1}\} = \log \frac{2p}{p+1} = I(0). \quad \square$$

**1.6. Special Dimensions.** The graph of  $D(z)$  for  $p = 2$  is shown in Figure 1. The appearance of the graph for other primes is very similar.

The value of  $D(z)$  at  $z = 0$  is particularly interesting:

$$D(0) = 2 - \frac{I(0)}{\log p} = \frac{\log p^2}{\log p} + \frac{\log(p+1)/(2p)}{\log p} = \frac{\log p(p+1)/2}{\log p}.$$

According to [56], this is the self-similarity dimension of the binomial coefficients *not* divisible by  $p$ . In fact, it is the Hausdorff dimension of the closure of the subset of the unit square representing the entire fractal pattern of the elements of Pascal's triangle not divisible by  $p$  "viewed from infinity"; this is described precisely and proved by Flath and Peele [17], who also prove that the Hausdorff dimension of the closure of the set corresponding to the elements divisible precisely by  $p^h$  for a fixed  $h$  (*not*  $p^{zn}$  as in our approach) is equal to  $D(0)$ . They also note that when  $h > 0$  the pattern is usually not self-similar ([17], p. 229). A paper of v. Haeseler, Peitgen, and Skordev "deciphers exactly the hierarchical self-similarity features of the Pascal triangle (mod  $p$ )<sup>s</sup>" ([23, p. 483]).

It is also of interest to note that there is a neat connection between results on the structure of Pascal's triangle modulo 2 and Chaitin's work on algorithmic information theory and Hilbert's tenth problem (see [3], especially section 2.2).

Using Theorem 1, we can establish related results for the (more elementary) box-counting dimension of appropriate sets (see, e.g., [14], sect. 3.1). First consider the case where  $h = 0$ . Map the initial  $p^n \times p^n$  block of Pascal's triangle into the unit square by making  $\binom{x+y}{x}$  correspond to  $[xp^{-n}, (x+1)p^{-n}) \times [yp^{-n}, (y+1)p^{-n})$ . Then define the fractal set  $F := \bigcap_{n \in \mathbb{N}} F_n$ , where

$$(9) \quad F_n = \bigcup \left[ \frac{x}{p^n}, \frac{x+1}{p^n} \right) \times \left[ \frac{y}{p^n}, \frac{y+1}{p^n} \right)$$

and where the union is taken over  $(x, y)$  satisfying  $0 \leq x, y < p^n$  and  $p \nmid \binom{x+y}{x}$ . The facts that  $\langle F_n \rangle$  are nested sets and that  $F$  is self-similar with scaling factor  $1/p$  may be deduced from Kummer's theorem. (Alternatively and more generally, the self-similarity of each set of residues modulo  $p$  is a consequence of a theorem of Lucas [36], Section 21; Long [35] and Sved [49] provide pictures as well as proofs. See also [46], Lemma 4; [1]; and [39], p. 329. Mandelbrot, who promoted the fractal perspective in [39], noted that when  $p = 2$  the resulting figure is one studied by Sierpiński [50], and dubbed it (p. 131) a *Sierpiński gasket*.)

The box-counting dimension of  $F$  is

$$D_F = \lim_{n \rightarrow \infty} \frac{\log N_n(0)}{\log(1/p^{-n})}.$$

Taking  $h = 0$  in Theorem 1, we get

$$N_n(0) = \sum_{i=0}^n \binom{n}{i} \left( \frac{p-1}{2} \right)^i p^n = \left( 1 + \frac{p-1}{2} \right)^n p^n = \left[ \frac{p(p+1)}{2} \right]^n,$$

so

$$\frac{\log N_n(0)}{\log(1/p^{-n})} = \frac{n \log(p(p+1)/2)}{n \log p} = D(0)$$

for every  $n$ . Thus, even though  $\Gamma = \{0\}$  is not an  $I$ -continuity set (because  $\Gamma^\circ$  is empty, making  $\inf_{z \in \Gamma^\circ} I(z) = \infty$ ), we still have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_p N_n(\{0\}) = D(0),$$

and  $D(0) = D_F$ .

For fixed  $h \geq 0$ , redefine  $F_n$  in equation (9) by taking the union over  $(x, y)$  satisfying  $0 \leq x, y < p^n$  and  $p^{h+1} \nmid \binom{x+y}{x}$ . Again the  $\langle F_n \rangle$  are nested sets, by Kummer's theorem. By Theorem 1,

$$N_n(h/n) = p^n \sum_{i=0}^{n-h} \binom{n-h}{i} \left(\frac{p-1}{2}\right)^i \sum_{k=0}^h \binom{h-1}{h-k} \binom{i+k}{k} \left(\frac{p-1}{2}\right)^k.$$

Here the inner sum is (crudely) bounded between 1 and  $(h+1)h! \binom{n}{h} [1 \vee ((p-1)/2)^h]$ , while

$$p^n \sum_{i=0}^{n-h} \binom{n-h}{i} \left(\frac{p-1}{2}\right)^i = p^n \left(1 + \frac{p-1}{2}\right)^{n-h} = \frac{[p(p+1)/2]^n}{[(p-1)/2]^h}.$$

Thus, for  $n$  large enough ( $n > 2h$ ),

$$\begin{aligned} \left[\frac{p(p+1)}{2}\right]^n &\leq \sum_{0 \leq h' \leq h} N_n\left(\frac{h'}{n}\right) \\ &\leq (h+1)(h+1)! \binom{n}{h} [1 \vee ((p-1)/2)^h] \frac{[p(p+1)/2]^n}{[(p-1)/2]^h}. \end{aligned}$$

Therefore,

$$\lim_{n \rightarrow \infty} \frac{\log \sum_{0 \leq h' \leq h} N_n(h'/n)}{\log(1/p^{-n})} = \frac{\log p(p+1)/2}{\log p}$$

is the box-counting dimension of the redefined fractal set  $F$  corresponding to the binomial coefficients not divisible by  $p^{h+1}$ .

Another interesting value occurs at  $z = \frac{1}{2}$ :  $D(\frac{1}{2}) = 2$ . In this case, in view of ergodicity, we see that if  $\frac{1}{2} \in \Gamma$ , then the set of points in the unit square for which, for some  $n$ , the corresponding  $Z_m \in \Gamma$  for every  $m > n$  is a set of probability (two-dimensional Lebesgue measure) 1, and thus it is a set of dimension 2.

Thus the minimum and maximum values of  $D$  may be interpreted as dimensions. Additionally,  $D$  satisfies  $D(A \cup B) = \max\{D(A), D(B)\}$ , a familiar property of dimension, by virtue of (5), and the expression

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_p N_n(\Gamma) = \lim_{n \rightarrow \infty} \frac{\log N_n(\Gamma)}{\log(1/p^{-n})}$$

superficially resembles a box-counting dimension, but it is not clear whether there is a natural fixed set for which this is the fractal dimension. Still, the function  $D$  can be shown to represent a multifractal spectrum of dimensions, as we shall see in the next section.

## 2. A $p$ -ADIC MULTIFRACTAL SPECTRUM

**2.1. Classical Multifractals.** Although the scientific literature tends to avoid a precise definition of the term “multifractal,” modern mathematical formulations

agree in identifying a multifractal with a *measure*. Our summary is based on Falconer ([14], ch. 17), Holley and Waymire [27], Feder ([15], ch. 6), and Evertsz and Mandelbrot [13].

Let  $\mu$  denote a probability measure supported by a bounded subset of  $\mathbb{R}^m$ . For each  $\delta \in (0, 1)$ , let  $\mathcal{M}_\delta$  denote the collection of  $\delta$ -mesh boxes that intersect the support of  $\mu$ . For each  $\alpha \in \mathbb{R}$ , let

$$\overline{N}_\delta(\alpha) = \#\{B \in \mathcal{M}_\delta : \mu(B) \geq \delta^\alpha\}.$$

Define the multifractal spectrum for  $\alpha \in \mathbb{R}$  by the double limit

$$f(\alpha) = \lim_{\varepsilon \rightarrow 0+} \lim_{\delta \rightarrow 0+} \frac{\log[\overline{N}_\delta(\alpha + \varepsilon) - \overline{N}_\delta(\alpha - \varepsilon)]}{-\log \delta},$$

so that

$$\overline{N}_\delta(\alpha + \varepsilon) - \overline{N}_\delta(\alpha - \varepsilon) \sim \delta^{-f(\alpha)} \quad \text{as } \delta \rightarrow 0+.$$

Also define the “partition function” for  $q \in \mathbb{R}$  by

$$S_\delta(q) = \sum_{B \in \mathcal{M}_\delta} \mu(B)^q.$$

Here  $S_\delta(0) = \#\mathcal{M}_\delta$ , the number of  $\delta$ -coordinate-mesh boxes covering  $\text{supp}(\mu)$ , and so

$$\tau(q) = \lim_{\delta \rightarrow 0+} \frac{\log S_\delta(q)}{-\log \delta}$$

for  $-\infty < q < \infty$  may be considered to provide generalizations of the box-counting dimension, and we have

$$S_\delta(q) \sim \delta^{-\tau(q)} \quad \text{as } \delta \rightarrow 0+.$$

Because  $\mu(\text{supp}(\mu)) < \infty$ , we have  $\tau(1) = 0$ , and the difference quotient  $D_q := \tau(q)/(q - 1)$  is known as a “generalized dimension” (it is normalized so that  $D_q = m = \dim(\text{supp}(\mu))$  when  $\mu$  has constant density) ([15], p. 87). In the scientific literature Legendre transformations connecting  $\tau(q)$  and  $f(\alpha)$  are often adduced. But it appears that only a few cases have been worked out rigorously ([27], p. 821).

It was Mandelbrot who introduced precursors of the multifractal notion to describe turbulence and other phenomena ([37], [38], [39], pp. 375–381). Frisch and Parisi [18] coined the term “multifractal” and, along with Jensen *et al.* [30], introduced the singularity spectrum  $f(\alpha)$ . The quantity  $D_q$  is related to *Rényi information* (see [45], ch. XI), and in the physics literature ([22], [26], [41]), as noted, it is called *generalized dimension*. Following [27], we will use the term *Rényi exponent* for  $\tau(q)$  and its counterpart below.

**2.2. A  $p$ -adic Approach to Multifractals.** Let  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ , let  $m$  denote the dimension, let  $p$  be a prime, and let  $\delta$  denote a real number in  $(0, 1)$ . Define  $\mathcal{M}_\delta$  anew as

$$\mathcal{M}_\delta = \{(x_1, \dots, x_m) \in \mathbb{N}_0^m : 0 \leq x_1, \dots, x_m < \delta^{-1}\}.$$

Now, instead of measuring the size of  $\delta$ -mesh boxes by a measure  $\mu$ , we shall measure the size of elements of  $\mathcal{M}_\delta$  by a fixed function  $\phi : \mathbb{N}_0^m \rightarrow \mathbb{N}_0$  and the  $p$ -norm ([33], p. 2) of  $p$ -adic analysis ([25]), which is defined on  $\mathbb{N}_0$  as follows:

$$|y|_p = \frac{1}{p^{\text{ord}_p y}}.$$

We take  $\text{ord}_p 0 = \infty$  so that  $|0|_p = 0$ .

For  $\alpha \in \mathbb{R}$ , let

$$\tilde{N}_\delta(\alpha) = \#\{\mathbf{x} \in \mathcal{M}_\delta : |\phi(\mathbf{x})|_p \geq \delta^\alpha\}.$$

Note that

$$\tilde{N}_\delta(\alpha) = \#\{\mathbf{x} \in \mathcal{M}_\delta : \text{ord}_p \phi(\mathbf{x}) \leq \alpha \log_p \delta^{-1}\}$$

and for  $\delta = p^{-n}$ ,  $n \in \mathbb{Z}^+$ ,

$$\begin{aligned} (10) \quad & \tilde{N}_{p^{-n}}(\alpha + \varepsilon) - \tilde{N}_{p^{-n}}(\alpha - \varepsilon) \\ &= \#\{(x_1, \dots, x_m) \in \mathbb{N}_0^m : 0 \leq x_1, \dots, x_m < p^n, \\ & \quad \alpha - \varepsilon < \frac{1}{n} \text{ord}_p \phi(x_1, \dots, x_m) \leq \alpha + \varepsilon\}. \end{aligned}$$

For  $\alpha \in \mathbb{R}$ , define the *multifractal spectrum* by

$$f(\alpha) = \lim_{\varepsilon \rightarrow 0+} \lim_{\delta \rightarrow 0+} \frac{\log[\tilde{N}_\delta(\alpha + \varepsilon) - \tilde{N}_\delta(\alpha - \varepsilon)]}{-\log \delta}$$

wherever the double limit exists in  $[-\infty, \infty)$ . Notice that  $0 \leq \tilde{N}_\delta(\alpha) \leq \#\mathcal{M}_\delta = \lceil \delta^{-1} \rceil^m$  and  $\tilde{N}_\delta(\alpha) = 0$  for  $\alpha < 0$ . Thus, if  $f$  is everywhere defined, then

$$(11) \quad -\infty \leq f(\alpha) \leq m \text{ for every } \alpha \quad \text{and} \quad f(\alpha) = -\infty \text{ for } \alpha < 0.$$

The new partition function is

$$S_\delta(q) = \sum_{\mathbf{x} \in \mathcal{M}_\delta} |\phi(\mathbf{x})|_p^q$$

for  $q \in \mathbb{R}$ . Note that  $|y|_p^q = \beta^{\text{ord}_p y}$  for  $\beta = p^{-q}$ . It is known ([33], pp. 3, 7) that all the norms  $\|y\| = \beta^{\text{ord}_p y}$  with  $0 < \beta < 1$ , i.e.,  $q > 0$ , induce equivalent metrics on  $\mathbb{Q}$ , so, for  $q > 0$ ,  $S_\delta(q)$  yields values based on norms equivalent to  $|\cdot|_p$ . Let

$$M_\delta(k) = \#\{\mathbf{x} \in \mathcal{M}_\delta : \text{ord}_p \phi(\mathbf{x}) = k\};$$

then

$$S_\delta(q) = \sum_{k=0}^{\infty} \frac{M_\delta(k)}{p^{qk}},$$

which is the generating function of  $\langle M_\delta(k) \rangle$  evaluated at  $p^{-q}$ . We shall find that when  $f(\alpha)$  exists for every  $\alpha$ , then

$$\tau(q) = \lim_{\delta \rightarrow 0+} \frac{\log S_\delta(q)}{-\log \delta}$$

exists at least for every  $q > 0$ , so that  $S_\delta(q) \sim \delta^{-\tau(q)}$ .

**2.3. Binomial and Multinomial Coefficients.** In order to apply the  $p$ -norm-based approach to binomial coefficient multifractals, we take

$$\phi(x, y) = \binom{x+y}{x};$$

in the case of multinomial coefficients, we take

$$\phi(x_1, \dots, x_m) = \binom{x_1 + \dots + x_m}{x_1, \dots, x_m}.$$

For  $\Gamma \subset \mathbb{R}$ , let

$$N_n(\Gamma) = \# \left\{ \mathbf{x} \in \mathcal{M}_{p^{-n}} : \frac{1}{n} \text{ord}_p \phi(\mathbf{x}) \in \Gamma \right\},$$

and note that this agrees with the earlier definitions of  $N_n(\Gamma)$  in the case of binomial and multinomial coefficients.

As a consequence of equation (10) and the definitions, we have the following.

**Lemma 1.** *Let  $\phi : \mathbb{N}_0^m \rightarrow \mathbb{N}_0$ . If  $\delta = p^{-n}$ , where  $n$  is a positive integer, then for  $\alpha \in \mathbb{R}$  and  $\varepsilon > 0$ , we have*

$$\tilde{N}_{p^{-n}}(\alpha + \varepsilon) - \tilde{N}_{p^{-n}}(\alpha - \varepsilon) = N_n((\alpha - \varepsilon, \alpha + \varepsilon]).$$

What happens if  $\delta \neq p^{-n}$ ?

**Lemma 2.** *Let  $\phi : \mathbb{N}_0^m \rightarrow \mathbb{N}_0$ . If  $1 < p^{n-1} < \delta^{-1} \leq p^n$ , then for  $\varepsilon > 0$*

$$\begin{aligned} N_{n-1} \left( \left( \frac{n}{n-1}(\alpha - \varepsilon), \alpha + \varepsilon \right] \right) &\leq \tilde{N}_\delta(\alpha + \varepsilon) - \tilde{N}_\delta(\alpha - \varepsilon) \\ &\leq N_n \left( \left( \frac{n-1}{n}(\alpha - \varepsilon), \alpha + \varepsilon \right] \right), \end{aligned}$$

where  $0 < \varepsilon < |\alpha|$  if  $\alpha \neq 0$ .

*Proof.* Assume  $1 < p^{n-1} < \delta^{-1} \leq p^n$  and  $0 < \varepsilon < \alpha$  or  $\alpha = 0 < \varepsilon$ . Then

$$\begin{aligned} \mathbf{x} \in \mathcal{M}_{p^{-(n-1)}} \quad &\text{and} \quad \frac{n}{n-1}(\alpha - \varepsilon) < \frac{1}{n-1} \text{ord}_p \phi(\mathbf{x}) \leq \alpha + \varepsilon \\ \implies \mathbf{x} \in \mathcal{M}_\delta \quad &\text{and} \quad \alpha - \varepsilon < \frac{\text{ord}_p \phi(\mathbf{x})}{\log_p \delta^{-1}} < \alpha + \varepsilon \\ \implies \mathbf{x} \in \mathcal{M}_\delta \quad &\text{and} \quad \delta^{\alpha+\varepsilon} \leq |\phi(\mathbf{x})|_p < \delta^{\alpha-\varepsilon} \\ \implies \mathbf{x} \in \mathcal{M}_{p^{-n}} \quad &\text{and} \quad \alpha - \varepsilon < \frac{\text{ord}_p \phi(\mathbf{x})}{\log_p \delta^{-1}} \leq \alpha + \varepsilon \\ \implies \mathbf{x} \in \mathcal{M}_{p^{-n}} \quad &\text{and} \quad \frac{n-1}{n}(\alpha - \varepsilon) < \frac{1}{n} \text{ord}_p \phi(\mathbf{x}) \leq \alpha + \varepsilon. \end{aligned}$$

Therefore, the inequalities asserted in the lemma follow when  $0 < \varepsilon < \alpha$  or  $\alpha = 0 < \varepsilon$ . If  $\alpha < 0$  and  $0 < \varepsilon < |\alpha|$ , then all the terms in the lemma inequalities become zero, and so the lemma holds in this case too.  $\square$

Now we are in a position to show that the function  $D$  of Theorem 2 coincides with the multifractal spectrum of the binomial coefficients.

**Theorem 3.** *Let  $\phi(x, y) = \binom{x+y}{x}$ . Then the multifractal spectrum  $f(\alpha)$  exists and equals  $D(\alpha)$  for every  $\alpha \in \mathbb{R}$ .*

*Remark.* Anticipating Theorem 5, we note that the same conclusion, and the proof given below, hold also for the multifractal spectra of multinomial coefficients.

*Proof.* Let  $\varepsilon, \eta > 0$ . First suppose  $\alpha \geq 0$ , and require  $0 < \eta < \varepsilon$  and, if  $\alpha > 0$ ,  $0 < \varepsilon < \alpha$ . Choose  $n$  so large that  $(\alpha - \varepsilon)/(n-1) < \eta$ . Then

$$(\alpha - \varepsilon + \eta, \alpha + \varepsilon] \subset A_{n-1} := \left( \frac{n}{n-1}(\alpha - \varepsilon), \alpha + \varepsilon \right]$$

and

$$B_n := \left( \left(1 - \frac{1}{n}\right)(\alpha - \varepsilon), \alpha + \varepsilon \right] \subset (\alpha - \varepsilon - \eta, \alpha + \varepsilon].$$

By Lemma 2)

$$N_{n-1}(A_{n-1}) \leq \tilde{N}_\delta(\alpha + \varepsilon) - \tilde{N}_\delta(\alpha - \varepsilon) \leq N_n(B_n),$$

so

$$\begin{aligned} \frac{n-1}{\log \delta^{-1}} \frac{\log N_{n-1}(\alpha - \varepsilon + \eta, \alpha + \varepsilon)]}{n-1} &\leq \frac{\log[\tilde{N}_\delta(\alpha + \varepsilon) - \tilde{N}_\delta(\alpha - \varepsilon)]}{-\log \delta} \\ &\leq \frac{n}{\log \delta^{-1}} \frac{\log N_n((\alpha - \varepsilon - \eta, \alpha + \varepsilon])}{n}. \end{aligned}$$

Let  $n \rightarrow \infty$ . By Theorem 2,

$$\frac{1}{n} \log N_n(\Gamma) \rightarrow \frac{D(\Gamma)}{\log p}.$$

Also  $n/\log \delta^{-1} \rightarrow \log p$ . So

$$\begin{aligned} (\log p) \frac{D((\alpha - \varepsilon + \eta, \alpha + \varepsilon])}{\log p} &\leq \liminf_{\delta \rightarrow 0+} \frac{\log[\tilde{N}_\delta(\alpha + \varepsilon) - \tilde{N}_\delta(\alpha - \varepsilon)]}{-\log \delta} \\ &\leq \limsup_{\delta \rightarrow 0+} \frac{\log[\tilde{N}_\delta(\alpha + \varepsilon) - \tilde{N}_\delta(\alpha - \varepsilon)]}{-\log \delta} \leq (\log p) \frac{D((\alpha - \varepsilon - \eta, \alpha + \varepsilon])}{\log p}. \end{aligned}$$

Let  $\varepsilon \rightarrow 0+$  (whence  $\eta \rightarrow 0+$ ). By the continuity of  $D$  from Theorem 2 for  $0 \leq \alpha \leq 1$  ( $0 \leq \alpha \leq m-1$  for multinomials),  $D(\alpha) = f(\alpha)$ . If  $\alpha < 0$  or  $\alpha > 1$  ( $\alpha > m-1$  for multinomials),  $\tilde{N}_\delta(\alpha + \varepsilon) - \tilde{N}_\delta(\alpha - \varepsilon) = 0$  for sufficiently small  $\varepsilon$ , so  $f(\alpha) = -\infty = D(\alpha)$ .  $\square$

**2.4. The Rényi Exponent  $\tau(q)$ .** Returning now to the general case, where  $\phi : \mathbb{N}_0^n \rightarrow \mathbb{N}_0$  is an arbitrary given function, we will find evidence that the  $p$ -adic multifractal formulation is a pleasant and appropriate one in a proof that the “Legendre transform” formalism works. Specifically, we will show that if

$$\tau(q) = \sup_{\alpha} \{f(\alpha) - q\alpha\} = -f^c(q)$$

where  $f^c(q) = \inf_{\alpha} \{q\alpha - f(\alpha)\}$  is the concave conjugate of  $f$ , then the partition function follows the power law,

$$S_\delta(q) \sim \delta^{-\tau(q)} \quad \text{as } \delta \rightarrow 0+.$$

**Proposition 2.** Assume that the double limit

$$f(\alpha) = \lim_{\varepsilon \rightarrow 0+} \lim_{\delta \rightarrow 0+} \frac{\log[\tilde{N}_\delta(\alpha + \varepsilon) - \tilde{N}_\delta(\alpha - \varepsilon)]}{-\log \delta}$$

exists in  $[-\infty, \infty)$  for every  $\alpha \in \mathbb{R}$ . Define

$$\tau(q) = \sup_{\alpha} \{f(\alpha) - q\alpha\}.$$

Then

$$\lim_{\delta \rightarrow 0+} \frac{\log S_\delta(q)}{-\log \delta} = \tau(q)$$

for  $q > 0$ . In addition, if there is a constant  $c > 0$  such that, for every  $\delta$ ,  $\text{ord}_p \phi(\mathbf{x}) \leq c \log_p \delta^{-1}$  for every  $\mathbf{x} \in \mathcal{M}_\delta$ , then this limit also holds for  $q \leq 0$ .

*Remark.* The additional hypothesis for  $q \leq 0$  rules out functions with very high  $p$ -power divisibility: If whenever  $\mathbf{x} \in \mathcal{M}_{p^{-n}}$  the highest power  $p^k$  dividing  $\phi(\mathbf{x})$  has  $k \leq cn$ , then this hypothesis is satisfied. The hypothesis is also equivalent to the existence of  $c > 0$  such that  $|\phi(\mathbf{x})|_p \geq \delta^c$  for every  $\mathbf{x} \in \mathcal{M}_\delta$ , for arbitrary  $\delta$ .

*Proof.* Our proof follows Falconer's ([14], p. 258) for  $q > 0$ . Fix  $q$ . Let  $\eta > 0$ . Because  $\tau(q) := \sup_\alpha (f(\alpha) - q\alpha)$ , we may choose  $\alpha$  so that  $f(\alpha) - q\alpha > \tau(q) - \eta$ . The assumption that  $f(\alpha)$  exists implies that there exists  $\varepsilon_0 > 0$  such that whenever  $0 < \varepsilon < \varepsilon_0$ ,

$$(12) \quad \delta^{-f(\alpha)+\eta} \leq \tilde{N}_\delta(\alpha + \varepsilon) - \tilde{N}_\delta(\alpha - \varepsilon) \leq \delta^{-f(\alpha)-\eta}$$

for every  $\delta$  smaller than some  $\delta_0 > 0$ . We require in addition that  $\varepsilon < \eta/q$  in the case that  $q > 0$ . Now,

$$\begin{aligned} S_\delta(q) &\geq [\tilde{N}_\delta(\alpha + \varepsilon) - \tilde{N}_\delta(\alpha - \varepsilon)](\delta^{\alpha \pm \varepsilon})^q \\ &\geq \delta^{-f(\alpha)+\eta+q(\alpha \pm \varepsilon)} \\ &\geq \delta^{-\tau(q)+2\eta \pm q\varepsilon} \geq \delta^{-\tau(q)+3\eta}, \end{aligned}$$

where we choose  $+$  if  $q > 0$  and  $-$  if  $q \leq 0$ . Therefore,

$$(13) \quad \liminf_{\delta \rightarrow 0^+} \frac{\log S_\delta(q)}{-\log \delta} \geq -\tau(q) - 3\eta.$$

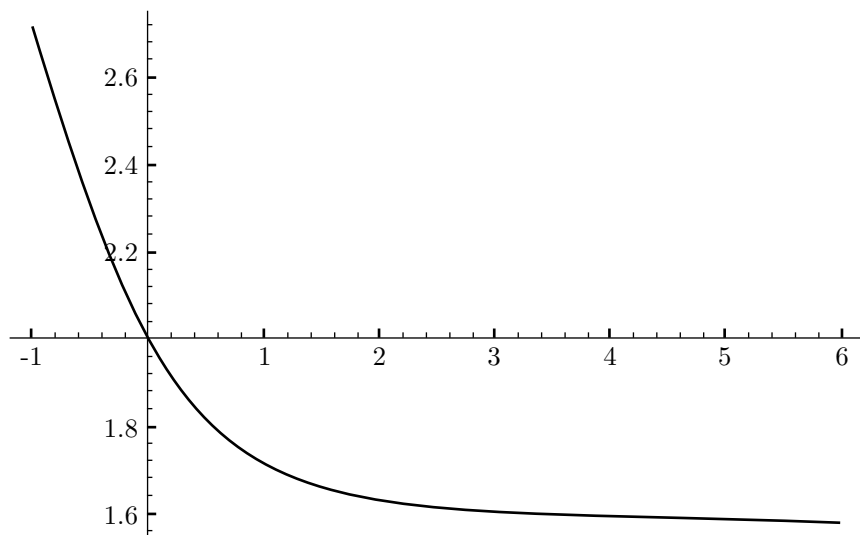
Next let us consider the limit supremum in the case that  $q > 0$ . As noted in (11),  $f(\alpha) = -\infty$  for  $\alpha < 0$ , so

$$\tau(q) = \sup_{\alpha \geq 0} \{f(\alpha) - q\alpha\}.$$

From this it is clear that  $\tau$  is a nonincreasing function. Since  $f(\alpha) \leq m$  for  $\alpha \geq 0$ ,  $\tau(q) \leq m$  for  $q > 0$ . Choose  $\beta \geq (m - \tau(q))/q \geq 0$ , so that  $m - q\beta \leq \tau(q)$ . Now for every  $\alpha$  in  $[0, \beta]$  there exists  $\varepsilon'_\alpha > 0$  such that whenever  $0 < \varepsilon < \varepsilon'_\alpha$ —say  $\varepsilon = \varepsilon_\alpha$  and  $\varepsilon_\alpha < \eta/q$ —we have that (12) holds for every  $\delta$  smaller than some  $\delta_\alpha$ . The intervals  $\{(\alpha - \varepsilon_\alpha, \alpha + \varepsilon_\alpha)\}$  cover  $[0, \beta]$ . By the Heine-Borel theorem, a finite subcollection of them, now denoted  $\{(\alpha_k - \varepsilon_k, \alpha_k + \varepsilon_k)\}_{k=1}^n$ , also covers  $[0, \beta]$ . Let  $\delta_0$  denote the minimum of the corresponding  $\delta$ 's. Now

$$\begin{aligned} S_\delta(q) &:= \sum_{\mathbf{x} \in \mathcal{M}_\delta} |\phi(\mathbf{x})|_p^q \\ &\leq \sum_{k=1}^n [\tilde{N}_\delta(\alpha_k + \varepsilon_k) - \tilde{N}_\delta(\alpha_k - \varepsilon_k)](\delta^{\alpha_k - \varepsilon_k})^q + (\#\mathcal{M}_\delta)(\delta^\beta)^q \\ &\leq \sum_{k=1}^n \delta^{-f(\alpha_k)-\eta+q(\alpha_k-\varepsilon_k)} + \lceil \delta^{-1} \rceil^m \delta^{q\beta} \\ &\leq n\delta^{-\tau(q)-\eta-\eta} + 2^m \delta^{-m+q\beta} \\ &\leq n\delta^{-\tau(q)-2\eta} + 2^m \delta^{-\tau(q)} \\ &\leq (n + 2^m) \delta^{-\tau(q)-2\eta} \end{aligned}$$



FIGURE 2. The Rényi exponent  $\tau(q)$  for  $m = 2$  and  $p = 2$ 

whenever  $0 < \delta < \delta_0$ . Therefore, for  $q > 0$ ,

$$(14) \quad \limsup_{\delta \rightarrow 0+} \frac{\log S_\delta(q)}{-\log \delta} \leq \tau(q) + 2\eta.$$

Because  $\eta$  was arbitrary in (13) and (14),

$$\lim_{\delta \rightarrow 0+} \frac{\log S_\delta(q)}{-\log \delta} = \tau(q).$$

Now assume  $q \leq 0$  and, for all  $\delta$ ,  $|\phi(\mathbf{x})|_p \geq \delta^c$  for every  $\mathbf{x} \in \mathcal{M}_\delta$ . In this case we use the interval  $[0, c]$  in place of  $[0, \beta]$  and get a finite subcover as before. Now, since  $q \leq 0$ ,

$$\begin{aligned} S_\delta(q) &\leq \sum_{k=1}^n \delta^{-f(\alpha_k) - \eta} (\delta^{\alpha_k + \varepsilon_k})^q \\ &\leq n \delta^{-f(\alpha_k) + q\alpha_k - \eta} \leq n \delta^{-\tau(q) - \eta} \end{aligned}$$

whenever  $0 < \delta < \delta_0$ . Therefore, for  $q \leq 0$ ,

$$(15) \quad \limsup_{\delta \rightarrow 0+} \frac{\log S_\delta(q)}{-\log \delta} \leq \tau(q) + \eta.$$

Because  $\eta$  was arbitrary in (13) and (15),

$$\lim_{\delta \rightarrow 0+} \frac{\log S_\delta(q)}{-\log \delta} = \tau(q). \quad \square$$

**2.5. Special Values of the Rényi Exponent.** Figure 2 depicts the graph of the Rényi exponent  $\tau(q)$  for binomial coefficients when  $p = 2$ . It was calculated numerically using the definition of  $\tau$  and the formula for  $D$  from Theorem 2. The graph shows a nonincreasing function. This is a general feature of  $\tau$ , because  $\tau(q) = \sup_{\alpha \geq 0} (f(\alpha) - q\alpha)$ .

The value at  $q = 0$  is of special interest. From the definition of  $\tau$  we have  $\tau(0) = \sup_{\alpha} f(\alpha)$ . Furthermore, if  $f$  is defined everywhere, then, by (11),  $\sup_{\alpha} f(\alpha) \leq m$ , and so  $\tau(0) \leq m$ . Under the additional hypothesis of Proposition 2 limiting the divisibility of  $\phi$ , we get

$$\tau(0) = \lim_{\delta \rightarrow 0+} \frac{\log S_{\delta}(0)}{-\log \delta} = \lim_{\delta \rightarrow 0+} \frac{\log [\delta^{-1}]^m}{\log \delta^{-1}} = m.$$

The value of  $\tau(1)$  for classical multifractals is always 0. But in Figure 2,  $\tau(1)$  is approximately 1.71. In the  $p$ -adic case  $\tau(1)$  is the value for which

$$\sum_{\mathbf{x} \in \mathcal{M}_{\delta}} |\phi(\mathbf{x})|_p \sim \delta^{-\tau(1)} \quad \text{as } \delta \rightarrow 0+,$$

and this varies with  $\phi$  and  $p$ .

What is  $\lim_{q \rightarrow \infty} \tau(q)$ ? Since  $\tau$  is nonincreasing, it has a limit in  $[-\infty, m]$ . In fact, because  $\tau(q) := \sup_{\alpha} (f(\alpha) - q\alpha) \geq f(0) - q \cdot 0 = f(0)$ , it satisfies  $\lim_{q \rightarrow \infty} \tau(q) \geq f(0)$ . If we make the additional assumption that  $f$  is continuous from the right at 0, we can show that  $\lim_{q \rightarrow \infty} \tau(q) = f(0)$ . For, letting  $\alpha_q = (m - f(0))/q$ , we have

$$\tau(q) = \sup_{0 \leq \alpha \leq \alpha_q} \{f(\alpha) - q\alpha\},$$

from which  $\lim_{q \rightarrow \infty} \tau(q) = f(0)$  follows.

In the case of binomial coefficients, then, we conclude that  $\tau(0) = 2$  and

$$\lim_{q \rightarrow \infty} \tau(q) = D(0) = \frac{\log(p(p+1)/2)}{\log p}.$$

For  $m$ -multinomial coefficients, we find  $\tau(0) = m$  and  $\lim_{q \rightarrow \infty} \tau(q) = D(0)$ . The value of  $D(0)$  for multinomials will be discussed further in Section 4.

### 3. GENERALIZED BINOMIAL COEFFICIENTS

**3.1. A Generalization of Kummer's Theorem.** Knuth and Wilf ([32]) have extended Kummer's theorem to a large class of generalized binomial coefficients. These include Gauss's  $q$ -binomial coefficients ([20], Section 5),

$$\binom{x+y}{x}_q := \frac{(1-q^{x+y})(1-q^{x+y-1}) \cdots (1-q^{x+1})}{(1-q^y)(1-q^{y-1}) \cdots (1-q)},$$

and the Fibonomial coefficients of Lucas ([36], Section 9),

$$\binom{x+y}{x}_{\mathcal{F}} := \frac{F_{x+y} F_{x+y-1} \cdots F_{x+1}}{F_y F_{y-1} \cdots F_1}$$

where  $\langle F_n \rangle_1^{\infty}$  is the Fibonacci sequence:  $F_1 = F_2 = 1$ ;  $F_n = F_{n-1} + F_{n-2}$  for  $n \geq 3$ . They identified "regular divisibility" as the key to a Kummer-like theorem.

**Definitions.** Let  $\mathcal{U} = \langle U_1, U_2, \dots \rangle$  be a sequence of positive integers. For every pair of nonnegative integers  $(x, y)$  define the  $\mathcal{U}$ -nomial coefficient

$$\binom{x+y}{x}_{\mathcal{U}} = \frac{U_{x+y} U_{x+y-1} \cdots U_{x+1}}{U_y U_{y-1} \cdots U_1}.$$

Also define, for  $m \in \mathbb{N}$ , the rank of apparition of  $m$  in  $\mathcal{U}$  by

$$r(m) = \min\{k : m | U_k\} \quad (r(m) = \infty \text{ if no } U_k \text{ is divisible by } m),$$

and for  $n \in \mathbb{N}$  define

$$d_m(n) = \#\{k : 1 \leq k \leq n, m|U_k\}.$$

The sequence  $\mathcal{U}$  is *regularly divisible* if, for each  $m \in \mathbb{N}$ ,  $r(m) = \infty$  or  $m|U_k \iff r(m)|k$ .

For a regularly divisible sequence one has

$$(16) \quad d_m(n) = \left\lfloor \frac{n}{r(m)} \right\rfloor.$$

Furthermore, the  $\mathcal{U}$ -nomial coefficients corresponding to a regularly divisible sequence  $\mathcal{U}$  must all be integers ([32], p. 214).

The following theorem is hinted at in Knuth and Wilf ([32], p. 215). It is stated explicitly in Wells ([54], p. 112) and proved (in a different fashion than below) in Webb and Wells [53].

**Theorem C.** *Let  $p$  be a prime. If  $\mathcal{U}$  is a regularly divisible sequence, then the order  $\text{ord}_p\left(\binom{x+y}{x}\right)_{\mathcal{U}}$  is equal to the number of carries that occur when  $x$  and  $y$  are added in the mixed-radix (mixed-base) system with radices  $b_1 := r(p) = \text{rank of apparition of } p \text{ in } \mathcal{U}$ ,  $b_2 := r(p^2)/r(p)$ ,  $b_3 := r(p^3)/r(p^2), \dots$ .*

*Note 1.* The mixed-radix representation of a nonnegative integer  $x$  is given by

$$x = \sum_{k \geq 0} X_k \prod_{j=1}^k b_j = \sum_{k \geq 0} X_k r(p^k)$$

where  $0 \leq X_k < b_{k+1}$  for each  $k$  and  $r(p^0) := 1$ .

*Note 2.* It is possible to have  $b_k = \infty$ , but this case will be ruled out by the hypotheses of Theorem 4.

*Note 3.* It is possible to have  $b_k = 1$ . Then if there is a carry into the  $k^{\text{th}}$  place, there is automatically a carry out.

*Proof.* By Proposition 1 of [32] and (16),

$$\begin{aligned} \text{ord}_p \left( \binom{x+y}{x} \right)_{\mathcal{U}} &= \sum_{k \geq 1} \{d_{p^k}(x+y) - d_{p^k}(x) - d_{p^k}(y)\} \\ &= \sum_{k \geq 1} \left\{ \left\lfloor \frac{x+y}{r(p^k)} \right\rfloor - \left\lfloor \frac{x}{r(p^k)} \right\rfloor - \left\lfloor \frac{y}{r(p^k)} \right\rfloor \right\} \\ &= \sum_{k \geq 1} C_k, \end{aligned}$$

the total number of carries. □

**Example 1.**  $\text{ord}_2 \left( \binom{8+11}{8} \right)_{\mathcal{F}}$

For the Fibonacci sequence and  $p = 2$  we have

$k$		4	3	2	1	0
$r(p^k)$		12	6	6	3	1
$b_k$		2	1	2	3	*
Carries		1	1	1	1	0
8	Augend		1	0	0	2
+11	Addend		1	0	1	2
19	Sum	1	1	0	0	1

There are 4 carries, so  $\text{ord}_2 \binom{8+11}{8}_{\mathcal{F}} = 4$ . (In fact,

$$\begin{aligned} \binom{8+11}{8}_{\mathcal{F}} &= \frac{(4181)(2584)(1597)(987)(610)(377)(233)(144)}{(21)(13)(8)(5)(3)(2)(1)(1)} \\ &= \frac{(4181)(2^3 \cdot 323)(1597)(7 \cdot 141)(2 \cdot 5 \cdot 61)(13 \cdot 29)(233)(2^4 \cdot 3^2)}{(3 \cdot 7)(13)(2^3)(5)(3)(2)}, \end{aligned}$$

which we see is an integer divisible by  $2^4$  but not  $2^5$ .)

**Example 2.**  $\text{ord}_p \binom{5+13}{5}_3$

The  $q$ -binomial coefficients may be generated by  $U_n = q^0 + q^1 + \cdots + q^{n-1}$  if  $q \in \mathbb{N}$ . Suppose  $q = 3$ . For  $p = 2$  and for  $p = 3$  we have

	$k$	5	4	3	2	1	0			1	0
$\boxed{p=2}$	$r(p^k)$	16	8	4	2	2	1	$\boxed{p=3}$	$\infty$	$\infty$	1
	$b_k$	2	2	2	1	2	*		$\infty$	$\infty$	*
	Carries	1	1	0	1	1	0		0	0	0
5	Augend		0	1	0	0	1				5
+13	Addend	+	1	1	0	1	1		+		13
18	Sum	1	0	0	1	0	0				18

When  $p = 2$  we get, by adding carries,  $\text{ord}_2 \binom{5+13}{5}_3 = 4$ . (In fact,

$$\begin{aligned} \text{ord}_2 \binom{5+13}{5}_3 &= \frac{(193710244)(64570081)(21523360)(7174453)(2391484)}{(121)(40)(13)(4)(1)} \\ &= \frac{(2^2 \cdot 48427561)(64570081)(2^5 \cdot 5 \cdot 134521)(11^2 \cdot 13 \cdot 4561)(2^2 \cdot 597871)}{(11^2)(2^3 \cdot 5)(13)(2^2)}, \end{aligned}$$

from which we also see that 4 is the highest power of 2 dividing this 3-binomial coefficient.) When  $p = 3$  there are no carries; in fact, there would be no carries for any  $x + y$ , because  $b_1 = \infty$ . So 3 does not divide any 3-binomial coefficient.

**3.2. Asymptotic Prime-power Divisibility.** Consider the carries process  $C_0 = 0, C_1, C_2, \dots, C_n$  arising from the addition of random  $n$ -digit mixed-radix numbers. (If  $r(p^g) = \infty$  for some  $g \in \mathbb{N}$ , we naturally have  $n < g$ .) As in the fixed-base case, we still have the Markov property,

$$\Pr(C_{k+1} = c_{k+1} \mid C_k = c_k, \dots, C_1 = c_1, C_0 = 0) = \Pr(C_{k+1} = c_{k+1} \mid C_k = c_k),$$

because the carry-out,  $C_{k+1}$ , is a function of only  $C_k$  and the random digits  $X_k$  and  $Y_k$ . There is a carry out ( $C_{k+1} = 1$ ) if and only if  $C_k + X_k + Y_k \geq b_{k+1}$ . So

the transition matrices are given by

$$(17) \quad \mathbf{\Pi}(k) = [\pi_{ij}(k)] = \frac{1}{2b_{k+1}} \begin{bmatrix} b_{k+1} + 1 & b_{k+1} - 1 \\ b_{k+1} - 1 & b_{k+1} + 1 \end{bmatrix}.$$

Thus, the Markov chain does not necessarily have stationary transition probabilities. But for the “ideal” primes of [32],  $\mathbf{\Pi}(k)$  is eventually constant.

**Definition** ([32], p. 215). The prime  $p$  is *ideal* for a regularly divisible sequence if there exists  $s = s(p) \in \mathbb{N}$  such that the radices of the Generalized Kummer Theorem satisfy  $b_k = 1$  for  $1 < k \leq s$  and  $b_k = p$  for  $k > s$ .

It is well known ([32], pp. 216-217) that every odd prime that does not divide  $q$  is ideal for the sequence that generates the  $q$ -binomial coefficients, and that every odd prime is ideal for the Fibonacci sequence.

For an ideal prime  $p$ , equation (17) becomes

$$(18) \quad \mathbf{\Pi}(k) = \frac{1}{2r(p)} \begin{bmatrix} r(p) + 1 & r(p) - 1 \\ r(p) - 1 & r(p) + 1 \end{bmatrix}, \text{ if } k = 0;$$

$$\mathbf{\Pi}(k) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \text{ if } 1 \leq k < s(p);$$

$$(19) \quad \mathbf{\Pi}(k) = \frac{1}{2p} \begin{bmatrix} p + 1 & p - 1 \\ p - 1 & p + 1 \end{bmatrix}, \text{ if } k \geq s(p).$$

Thus,  $\mathbf{\Pi}(k)$  is eventually the same as the transition matrix (2) for the carries process in the ordinary binomial case. The special values for an ideal prime specified by equation (18) turn out to be unimportant in the asymptotic analysis, and so we do not need to assume them. Property (19) alone is sufficient to guarantee similar asymptotic behavior for the number of generalized binomial coefficients divisible precisely by  $p^{zn}$  for  $z$  in an interval:

**Theorem 4.** Let  $\mathcal{U}$  be a regularly divisible sequence, let  $p$  be a prime, let  $r(m)$  denote the rank of apparition of  $m$  in  $\mathcal{U}$ , and assume that there exists  $s(p) \in \mathbb{N}$  such that  $b_k := r(p^k)/r(p^{k-1}) = p$  for every  $k > s(p)$ . Define  $D$  as in Theorem 2. For each interval  $\Gamma$ , let

$$N_n^{\mathcal{U}}(\Gamma) = \# \left\{ (x, y) \mid 0 \leq x, y < r(p^n) \text{ and } \frac{1}{n} \text{ord}_p \left( \begin{matrix} x+y \\ x \end{matrix} \right)_{\mathcal{U}} \in \Gamma \right\}.$$

Then, for every nondegenerate interval  $\Gamma \subseteq [0, 1]$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_p N_n^{\mathcal{U}}(\Gamma) = D(\Gamma).$$

**3.3. Proof of Theorem 4.** Let  $I$  be the rate function defined in Theorem 2. Let  $\Gamma = [\alpha, \beta] \subseteq [0, 1]$ , and assume  $\alpha < \beta$ . Since  $I$  is continuous and finite on  $[0, 1]$ ,  $\Gamma$  is an  $I$ -continuity set. We have

$$Z_n = \frac{1}{n} \text{ord}_p \left( \begin{matrix} x+y \\ x \end{matrix} \right)_{\mathcal{U}} = \frac{C_1 + \cdots + C_n}{n},$$

where  $\langle C_k \rangle$  is the carries process, by the Generalized Kummer Theorem. For  $n > s := s(p)$ , write

$$Z_n = \frac{C_1 + \cdots + C_s}{n} + \left(1 - \frac{s}{n}\right) \tilde{Z}_{n,s},$$

where  $\tilde{Z}_{n,s} := (C_{s+1} + \cdots + C_n)/(n-s)$ . Observe that  $\tilde{Z}_{n,s}$  conditioned on  $C_s = c_s$  has the same probability distribution as  $\tilde{Z}_{n-s}$  conditioned on  $C_0 = c_s$ , where  $\tilde{Z}_n$  is the average carry for the carries process with constant transition matrix given by (2). Write

$$\begin{aligned}
 (20) \quad & \Pr(\alpha \leq Z_n \leq \beta) \\
 &= \sum_{(c_1, \dots, c_s)} \Pr(C_1 = c_1, \dots, C_s = c_s) \\
 &\quad \times \Pr\left(\alpha \leq \frac{C_1 + \cdots + C_s}{n} + \left(1 - \frac{s}{n}\right)\tilde{Z}_{n,s} \leq \beta \mid C_1 = c_1, \dots, C_s = c_s\right) \\
 &= \sum_{(c_1, \dots, c_s)} \Pr(C_1 = c_1, \dots, C_s = c_s) \\
 &\quad \times \Pr\left(\alpha_n(c_1, \dots, c_s) \leq \tilde{Z}_{n,s} \leq \beta_n(c_1, \dots, c_s) \mid C_s = c_s\right)
 \end{aligned}$$

where

$$\alpha_n(c_1, \dots, c_s) = \frac{\alpha - (c_1 + \cdots + c_s)/n}{1 - s/n} \text{ and } \beta_n(c_1, \dots, c_s) = \frac{\beta - (c_1 + \cdots + c_s)/n}{1 - s/n}.$$

Thus  $\Pr(\alpha \leq Z_n \leq \beta)$  is a convex linear combination of probabilities of the form  $\Pr(\alpha' \leq \tilde{Z}_{n,s} \leq \beta' \mid C_s = c_s) = \Pr(\alpha' \leq \tilde{Z}_{n-s} \leq \beta' \mid C_0 = c_s)$ , which satisfy the large deviation principle with rate function  $I$ , by Proposition 1 and the note preceding it. Now, given arbitrary  $\varepsilon \in (0, (\beta - \alpha)/2)$ , we have, for all sufficiently large  $n$ ,

$$\begin{aligned}
 \alpha - \varepsilon &< \frac{\alpha - s/n}{1 - s/n} \leq \alpha_n(c_1, \dots, c_s) \leq \frac{\alpha}{1 - s/n} < \alpha + \varepsilon \\
 &< \beta - \varepsilon < \frac{\beta - s/n}{1 - s/n} \leq \beta_n(c_1, \dots, c_s) \leq \frac{\beta}{1 - s/n} < \beta + \varepsilon.
 \end{aligned}$$

Therefore, for  $\sigma \in \{0, 1\}$ ,

$$\begin{aligned}
 & \Pr\left(\alpha + \varepsilon \leq \tilde{Z}_{n-s} \leq \beta - \varepsilon \mid C_0 = \sigma\right) \\
 & \leq \Pr\left(\alpha_n(c_1, \dots, c_s) \leq \tilde{Z}_{n,s} \leq \beta_n(c_1, \dots, c_s) \mid C_s = \sigma\right) \\
 & \leq \Pr\left(\alpha - \varepsilon \leq \tilde{Z}_{n-s} \leq \beta + \varepsilon \mid C_0 = \sigma\right).
 \end{aligned}$$

Because the bounding terms here satisfy the large deviation principle with rate function  $I$ , the limit inferior and limit superior of the middle term lie between  $-I([\alpha + \varepsilon, \beta - \varepsilon])$  and  $-I([\alpha - \varepsilon, \beta + \varepsilon])$ . (Note that if  $\alpha - \varepsilon < 0$ , then  $\beta + \varepsilon > 0$ , and if  $\beta + \varepsilon > 1$ , then  $\alpha - \varepsilon < 1$ , so that  $[\alpha - \varepsilon, \beta + \varepsilon]$  is still an  $I$ -continuity set.) Because  $\varepsilon$  is arbitrary and  $I$  is continuous,

$$(21) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log \Pr\left(\alpha_n(c_1, \dots, c_s) \leq \tilde{Z}_{n,s} \leq \beta_n(c_1, \dots, c_s) \mid C_s = \sigma\right) = -I([\alpha, \beta]).$$

Finally, it is not hard to check that if a finite number of sequences of measures satisfy the large deviation principle with (the same) rate function  $I$ , then so does

any fixed convex linear combination of them. By (20) and (21), then,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \Pr(\alpha \leq Z_n \leq \beta) = -I([\alpha, \beta]),$$

and therefore

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_p N_n^{\mathcal{U}}([\alpha, \beta]) = D([\alpha, \beta]). \quad \square$$

**3.4. Special Cases.** We now apply Theorem 4 to the Gaussian  $q$ -binomial coefficients and the Fibonomial coefficients. The self-similarity of the Gaussian coefficients modulo  $p$  is discussed and depicted in Sved [49].

**Corollary 1** ( $q$ -binomial coefficients). *Let  $p$  be a prime, and let  $q > 1$  be an integer. If  $p \nmid q$ , then for every nondegenerate interval  $\Gamma \subseteq [0, 1]$ ,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_p \# \left\{ (x, y) \left| 0 \leq x, y < r(p^n), \frac{1}{n} \text{ord}_p \binom{x+y}{x}_q \in \Gamma \right. \right\} = D(\Gamma)$$

where  $D$  is as given in Theorem 2. If  $p \mid q$ , then  $p$  does not divide any  $q$ -binomial coefficient (or,  $D(\Gamma) = -\infty$  above).

*Proof.* Gauss's  $q$ -binomial coefficients may be generated by  $\mathcal{U} = \langle q^n - 1 \rangle$ . It is known that  $\mathcal{U}$  is regularly divisible ([32], p. 216) and that, as noted before, every odd prime that does not divide  $q$  is ideal for  $\mathcal{U}$ . If  $p = 2$ , then there exists  $f$  such that  $q \equiv 2^f \pm 1 \pmod{2^{f+1}}$ , and if  $q \equiv 2^f + 1$  then  $r(2) = r(2^2) = \cdots = r(2^f) = 1$  and  $r(2^k) = 2^{k-f}$  for  $k \geq f$ , while if  $q \equiv 2^f - 1$  then  $r(2) = 1, r(2^2) = \cdots = r(2^{f+1}) = 2$ , and  $r(2^k) = 2^{k-f}$  for  $k > f$  ([32], p. 216). Thus, whether  $p$  is 2 or an odd prime, there exists  $s(p)$  such that  $b_k := r(p^k)/r(p^{k-1}) = p$  for every  $k > s(p)$ , and so Theorem 4 applies.  $\square$

**Corollary 2** (Fibonomial coefficients). *Let  $p$  be a prime. Then the Fibonomial coefficients satisfy, for each nondegenerate interval  $\Gamma \subseteq [0, 1]$ ,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_p \# \left\{ (x, y) \left| 0 \leq x, y < r(p^n), \frac{1}{n} \text{ord}_p \binom{x+y}{x}_{\mathcal{F}} \in \Gamma \right. \right\} = D(\Gamma),$$

where  $D$  is as given in Theorem 2.

*Proof.* By [36], p. 206, or [32], p. 216, the Fibonacci sequence  $\mathcal{F}$  is regularly divisible. Also, by Theorem 1 of Halton [24], every odd prime is ideal for the Fibonacci sequence, and (by [24], p. 226) if  $p = 2$ , then  $r(p) = 3$  and  $r(p^2) = r(p^3) = 6 = p \cdot 3$  and  $r(p^k) = p^{k-2} \cdot 3$  for  $k \geq 3$ , so that  $b_k = p$  for  $k > s := 3$ . Thus, whether  $p$  is 2 or an odd prime, Theorem 4 applies.  $\square$

#### 4. MULTINOMIAL COEFFICIENTS

**4.1. Dickson's Theorem.** Now we consider the extension of our asymptotic results to multinomial coefficients,

$$\binom{x_1 + x_2 + \cdots + x_m}{x_1, x_2, \dots, x_m} = \frac{(x_1 + x_2 + \cdots + x_m)!}{x_1! x_2! \cdots x_m!},$$

where  $x_1, x_2, \dots, x_m$  are nonnegative integers and  $m$  is a fixed integer greater than 1.

The basis for our analysis is Dickson's extension of Kummer's theorem ([5], pp. 75–76; [6], p. 273). Now one expresses  $x_1, x_2, \dots, x_m$  in base  $p$  and adds:

	Carries	$C_n$	$C_{n-1}$	$C_{n-2}$	$\cdots$	$C_2$	$C_1$	$C_0 = 0$
$x_1$	Addends		$X_{1,n-1}$	$X_{1,n-2}$	$\cdots$	$X_{1,2}$	$X_{1,1}$	$X_{1,0}$
$\cdot$			$\cdot$	$\cdot$		$\cdot$	$\cdot$	$\cdot$
$\cdot$			$\cdot$	$\cdot$		$\cdot$	$\cdot$	$\cdot$
$\cdot$			$\cdot$	$\cdot$		$\cdot$	$\cdot$	$\cdot$
$x_m$		+	$X_{m,n-1}$	$X_{m,n-2}$	$\cdots$	$X_{m,2}$	$X_{m,1}$	$X_{m,0}$
$x_1 + \cdots + x_m$	Sum	$S_n$	$S_{n-1}$	$S_{n-2}$	$\cdots$	$S_2$	$S_1$	$S_0$

**Theorem D.** *Let  $p$  be a prime. Then*

$$\text{ord}_p \binom{x_1 + x_2 + \cdots + x_m}{x_1, x_2, \dots, x_m} = C_1 + C_2 + \cdots + C_n,$$

*the total of the carries one gets in adding  $x_1, x_2, \dots, x_m$  in base- $p$  arithmetic.*

**4.2. The Markov Chain of Carries.** Once again, the key observation is that if the  $\{X_{h,k}\}$  are independent (uniform) random digits, then the carries form a finite Markov chain:

$$\Pr(C_{k+1} = c_{k+1} \mid C_k = c_k, \dots, C_1 = c_1, C_0 = 0) = \Pr(C_{k+1} = c_{k+1} \mid C_k = c_k).$$

One may check by induction that the maximum possible value of the carry  $C_k$  is  $m - 1 - \lfloor (m - 1)/p^k \rfloor$ , and so, for  $k > \log_p(m - 1)$ , every carry value  $C_k$  from 0 through  $m - 1$  is possible. Consequently the state space of our Markov chain is  $\{0, 1, \dots, m - 1\}$ . Furthermore, it is possible to get from any state to any other state in  $\lfloor \log_p(m - 1) \rfloor + 1$  steps, so our Markov chain is aperiodic and irreducible. Let  $\mathbf{\Pi} = [\pi_{ij}]$  denote the transition matrix:

$$\pi_{ij} = \Pr(\text{carry-out} = j \mid \text{carry-in} = i).$$

The matrix  $\mathbf{\Pi}$  possesses remarkable combinatorial properties. Some of these are presented in [29], which includes more details regarding the results in this section.

**Lemma 1.** *The  $m$ -nomial carries process is an aperiodic, irreducible finite Markov chain with state space  $\{0, 1, \dots, m - 1\}$  and transition matrix  $\mathbf{\Pi} = [\pi_{ij}]$ , where*

$$(22) \quad \pi_{ij} = p^{-m} \sum_{r=0}^{j - \lfloor i/p \rfloor} (-1)^r \binom{m+1}{r} \binom{m-1-i+(j+1-r)p}{m}.$$

For example, when  $p = 3$  and  $m = 2, 3, 4$ , then  $\mathbf{\Pi}$  is

$$\begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}, \quad \begin{bmatrix} \frac{10}{27} & \frac{16}{27} & \frac{1}{27} \\ \frac{4}{27} & \frac{19}{27} & \frac{4}{27} \\ \frac{1}{27} & \frac{16}{27} & \frac{10}{27} \end{bmatrix}, \quad \begin{bmatrix} \frac{5}{27} & \frac{17}{27} & \frac{5}{27} & 0 \\ \frac{5}{81} & \frac{5}{9} & \frac{10}{27} & \frac{1}{81} \\ \frac{1}{81} & \frac{10}{27} & \frac{5}{9} & \frac{5}{81} \\ 0 & \frac{5}{27} & \frac{17}{27} & \frac{5}{27} \end{bmatrix}.$$



For an arbitrary prime  $p$  and  $m = 3$ ,

$$\Pi = \frac{1}{6p^2} \begin{bmatrix} p^2 + 3p + 2 & 4p^2 - 4 & p^2 - 3p + 2 \\ p^2 - 1 & 4p^2 + 2 & p^2 - 1 \\ p^2 - 3p + 2 & 4p^2 - 4 & p^2 + 3p + 2 \end{bmatrix}.$$

*Proof.* To calculate  $\pi_{ij} = \Pr(C_{k+1} = j | C_k = i)$ , consider the base- $p$  addition in the  $k^{\text{th}}$  place:

$$C_{k+1} = j \iff jp \leq i + X_{1,k} + \cdots + X_{m,k} \leq (j+1)p - 1$$

where  $0 \leq X_{1,k}, \dots, X_{m,k} \leq p - 1$ . Introducing the “slack” variable  $Y$ , this is equivalent to

$$X_{1,k} + \cdots + X_{m,k} + Y = (j+1)p - 1 - i,$$

where  $0 \leq X_{1,k}, \dots, X_{m,k}, Y \leq p - 1$ . The number of integer solutions of these inequalities is the same as the coefficient of  $x^{(j+1)p-1-i}$  in  $(1+x+x^2+\cdots+x^{p-1})^{m+1}$ . Because

$$(1+x+x^2+\cdots+x^{p-1})^{m+1} = (1-x^p)^{m+1}(1-x)^{-(m+1)}$$

and

$$(1-x^p)^{m+1} = \sum_r \binom{m+1}{r} (-x^p)^r$$

and

$$(1-x)^{-(m+1)} = \sum_{s \geq 0} \binom{m+s}{m} x^s,$$

the desired coefficient is

$$\sum_{r \leq j+1-(i+1)/p} (-1)^r \binom{m+1}{r} \binom{m+(j+1)p-1-i-rp}{m}.$$

Since  $r \leq j+1-(i+1)/p$  iff  $r \leq j - \lfloor i/p \rfloor$ , the lemma follows.  $\square$

The examples given above illustrate the cross-symmetry, or centrosymmetry [52], of  $\Pi$ . This is not obvious from the formula of Lemma 1, but it is easy to show from the definition of  $\pi_{ij}$  using the fact that if the  $\{X_{h,k}\}$  are independent random variables that are uniformly distributed on  $\{0, 1, \dots, p-1\}$ , then the equation  $\tilde{X}_{h,k} := p-1-X_{h,k}$  defines independent random variables which are also uniformly distributed on  $\{0, 1, \dots, p-1\}$ . We record this as a lemma.

**Lemma 2.** For  $i, j = 0, 1, \dots, m-1$ , we have  $\pi_{m-1-i, m-1-j} = \pi_{i,j}$ .

Another result that is easier to show from the probabilistic definition is that

$$(23) \quad \pi_{jj} > 0 \quad \text{for } j = 0, 1, \dots, m-1.$$

Here is the proof:

$$\begin{aligned} \pi_{jj} &= \Pr(C_{k+1} = j \mid C_k = j) \\ &= \Pr(j + X_{1,k} + \cdots + X_{m,k} = jb + r \text{ for some } r \in \{0, \dots, m-1\}) \\ &= \Pr((b-1)j \leq X_{1,k} + \cdots + X_{m,k} \leq (b-1)(j+1)), \end{aligned}$$

and this is positive, because  $X_{1,k} + \cdots + X_{m,k}$  has a positive probability of assuming each value between  $0 + \cdots + 0 = 0$  and  $(b-1) + \cdots + (b-1) = m(b-1)$ .

The stationary probability vector of the carries process is the row vector  $\mathbf{v}'$  of probabilities satisfying  $\mathbf{v}' = \mathbf{v}'\mathbf{\Pi}$ . Remarkably, its components turn out to be proportional to the Eulerian numbers ([11, pp. 485–487], [12, pp. 373–375], [21, sect. 6.2]),

$$\left\langle \begin{matrix} m \\ j \end{matrix} \right\rangle = \sum_{r=0}^j (-1)^r \binom{m+1}{r} (j+1-r)^m.$$

**Lemma 3.** *The stationary probability vector  $\mathbf{v}' = (p_0, p_1, \dots, p_{m-1})$  of  $\mathbf{\Pi}$  is given by*

$$p_j = \frac{1}{m!} \left\langle \begin{matrix} m \\ j \end{matrix} \right\rangle \quad \text{for } j = 0, 1, \dots, m-1.$$

*The mean of the stationary distribution is  $(m-1)/2$  and the variance is  $(m+1)/12$ .*

The proof (given, as noted, in [29]) is a calculation verifying that  $\mathbf{v}' = \mathbf{v}'\mathbf{\Pi}$  etc.

**4.3. The Large Deviation Principle.** Now let

$$Z_n = \frac{1}{n} \text{ord}_p \begin{pmatrix} x_1 + \dots + x_m \\ x_1, \dots, x_m \end{pmatrix}$$

and

$$N_n(\Gamma) = \#\{(x_1, \dots, x_m) | 0 \leq x_1, \dots, x_m < p^n \text{ and } Z_n \in \Gamma\}.$$

By Dickson's theorem,  $Z_n = (C_1 + \dots + C_n)/n$ . If  $(x_1, \dots, x_m)$  is chosen uniformly at random from  $\{0, \dots, p^n - 1\}^m$ , then the base- $p$  digits representing  $x_1, \dots, x_m$  are independent, uniform random digits, and so, by the “key observation” of the last section, the carries  $\langle C_k \rangle$  form a Markov chain. Thus (in just two steps!) we conclude that the large deviation principle applies to the sequence of measures  $\langle \Pr(Z_n \in \cdot) \rangle$ . Since

$$N_n(\Gamma) = p^{nm} \Pr(Z_n \in \Gamma),$$

we may re-express the large deviation principle in terms of  $N_n(\Gamma)$ :

**Proposition 3.** *Let  $p$  be a prime, and let  $m \geq 2$  be an integer. Let  $\pi_{ij}$  be as given in Lemma 1, and define*

$$\begin{aligned} \mathbf{\Pi}_t &= [\pi_{ij} e^{tj}], \\ \rho(\mathbf{\Pi}_t) &= \text{Perron-Frobenius eigenvalue of } \mathbf{\Pi}_t, \\ I(z) &= \sup_t \{tz - \log \rho(\mathbf{\Pi}_t)\}, \\ D(z) &= m - I(z)/\log p, \end{aligned}$$

*and, for  $\Gamma \subseteq \mathbb{R}$ ,  $D(\Gamma) = \sup_{z \in \Gamma} D(z)$ . Let*

$$N_n(\Gamma) = \left\{ (x_1, \dots, x_m) \mid 0 \leq x_1, \dots, x_m < p^n \text{ and } \frac{1}{n} \text{ord}_p \begin{pmatrix} x_1 + \dots + x_m \\ x_1, \dots, x_m \end{pmatrix} \in \Gamma \right\}.$$

*Then, for each  $I$ -continuity set  $\Gamma$ ,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_p N_n(\Gamma) = D(\Gamma).$$

In order to turn this proposition into a clear result, we need to get a better picture of the function  $D(z)$ . Although we will not find an elementary formula, as we did for the  $m = 2$  (binomial coefficient) divisibility spectrum in Theorem 2, we will be able to establish the corresponding properties and special values for the general  $m$ -nomial divisibility spectrum, and we turn now to that task.

#### 4.4. Asymptotic Prime-power Divisibility of Multinomial Coefficients.

Define  $\tilde{\Pi}(x) = [\pi_{ij}x^j]$ , and let  $\rho(x)$  be its spectral radius (the Perron-Frobenius eigenvalue for  $x \geq 0$ ). Notice that, if  $x > 0$  and  $t = \log x$ , then

$$\Pi_t = [\pi_{ij}e^{tj}] = [\pi_{ij}x^j] = \tilde{\Pi}(x) \quad \text{and} \quad \rho(\Pi_t) = \rho(x),$$

whence

$$I(z) = \sup_{t \in \mathbb{R}} \{zt - \log \rho(\Pi_t)\} = \sup_{x > 0} \{z \log x - \log \rho(x)\},$$

so

$$(24) \quad I(z) = \log \sup_{x > 0} \frac{x^z}{\rho(x)}.$$

**Lemma 4.** *The function  $\rho$  is continuous, convex, and strictly increasing on  $[0, \infty)$ ;  $\rho(x) = x^{m-1}\rho(1/x)$  for  $x > 0$ ;  $\rho(0) = \pi_{00} = p^{-m} \binom{p+m-1}{m}$ ; and  $\rho(x) \asymp x^{m-1}$  as  $x \rightarrow \infty$ , i.e., there are positive constants  $c_1, c_2$  such that  $c_1x^{m-1} \leq \rho(x) \leq c_2x^{m-1}$  for all sufficiently large  $x$ —in fact,  $\pi_{00}x^{m-1} \leq \rho(x) \leq x^{m-1}$  for  $x \geq 1$ .*

*Proof.* We use Wielandt's [55] variational characterization of the Perron-Frobenius eigenvalue: Assume  $x \geq 0$  (so that every  $\pi_{ij}x^j \geq 0$ ), and let

$$Q_{\mathbf{w}}(x) = \min_j \sum_{i=0}^{m-1} w_i \pi_{ij} x^j / w_j,$$

where the minimum is taken over  $j \in \{0, \dots, m-1\}$  for which  $w_j \neq 0$ ; then

$$\rho(x) = \sup_{\substack{\mathbf{w} \geq \mathbf{0} \\ \mathbf{w} \neq \mathbf{0}}} Q_{\mathbf{w}}(x) = \max_{\substack{\mathbf{w} \geq \mathbf{0} \\ \|\mathbf{w}\| \neq 0}} Q_{\mathbf{w}}(x),$$

where  $\|\cdot\|$  is any norm. Notice that, for each  $j$ ,

$$w_j Q_{\mathbf{w}}(x) \leq \sum_i w_i \pi_{ij} x^j,$$

so

$$\left(\sum_j w_j\right) Q_{\mathbf{w}}(x) \leq \sum_i w_i \sum_j \pi_{ij} x^j \leq \left(\sum_i w_i\right) \max_i \sum_j \pi_{ij} x^j.$$

Combining this with an easy lower bound, we have, for  $\mathbf{w} \geq \mathbf{0}$  and (say)  $w_k \neq 0$ ,

$$\pi_{kk} x^k \leq Q_{\mathbf{w}}(x) \leq \max_i \sum_j \pi_{ij} x^j.$$

Therefore,

$$(25) \quad 0 < \pi_{00} \leq \rho(x) \leq \max_i \sum_j \pi_{ij} x^j.$$

Since each  $\pi_{ij} \geq 0$  and, by (23), each  $\pi_{ii} > 0$ , we see that, for  $\mathbf{w} \geq \mathbf{0}$ ,  $Q_{\mathbf{w}}(x)$  is a strictly increasing function of  $x$  (for  $m \geq 2$ ), whence  $\rho(x)$  is also.

Invoking the cross-symmetry of  $\Pi$  (Lemma 2), we get

$$\begin{aligned} Q_{\mathbf{w}}(x) &= \min_j \sum_i w_i \pi_{m-1-i, m-1-j} x^j / w_j \\ &= \min_h \sum_k \tilde{w}_k \pi_{k, h} x^{m-1-h} / \tilde{w}_h \\ &= x^{m-1} \min_h \sum_k \tilde{w}_k \pi_{k, h} (1/x)^h / \tilde{w}_h \end{aligned}$$

for  $x > 0$ , where  $\tilde{w}_k = w_{m-1-k}$ . Therefore,

$$Q_{\mathbf{w}}(x) = x^{m-1} Q_{\tilde{\mathbf{w}}}(1/x) \quad \text{for } x > 0,$$

whence

$$\rho(x) = x^{m-1} \rho(1/x) \quad \text{for } x > 0.$$

Because  $\rho$  is increasing and satisfies the bounds given by (25) for  $0 < x \leq 1$  and  $\rho(1) = 1$ , the previous equation implies that

$$\pi_{00} x^{m-1} \leq \rho(x) \leq x^{m-1} \quad \text{for } x \geq 1,$$

whence  $\rho(x) \asymp x^{m-1}$ .

To establish the convexity of  $\rho$ , we start with the observation that, if  $p \geq 0, q \geq 0$ , and  $p + q = 1$ , then for  $x, y \geq 0$ ,

$$(px + qy)^j \leq px^j + qy^j \quad \text{for } j = 0, 1, \dots, m-1.$$

Then, for  $\mathbf{w} \geq \mathbf{0}$  and  $w_j > 0$ ,

$$\sum_i w_i \pi_{ij} (px + qy)^j / w_j \leq p \sum_i w_i \pi_{ij} x^j / w_j + q \sum_i w_i \pi_{ij} y^j / w_j,$$

whence  $Q_{\mathbf{w}}(px + qy) \leq pQ_{\mathbf{w}}(x) + qQ_{\mathbf{w}}(y)$ , so  $\rho(px + qy) \leq p\rho(x) + q\rho(y)$ . Therefore  $\rho(x)$  is convex for  $x \geq 0$ , and hence it is necessarily continuous for  $x > 0$ . It remains to prove the claimed value of  $\rho(0)$  and the continuity of  $\rho$  at 0.

Because

$$\tilde{\Pi}(0)\mathbf{w} = \begin{bmatrix} \pi_{00} & 0 & \dots & 0 \\ \pi_{10} & 0 & \dots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \pi_{m-1,0} & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \cdot \\ \cdot \\ \cdot \\ w_{m-1} \end{bmatrix} = w_0 \begin{bmatrix} \pi_{00} \\ \pi_{10} \\ \cdot \\ \cdot \\ \cdot \\ \pi_{m-1,0} \end{bmatrix},$$

we will have  $\tilde{\Pi}(0)\mathbf{w} = \lambda\mathbf{w}$  iff  $w_0 = 0$  or  $w_0 \neq 0$  and  $\lambda = \pi_{00}$ , whence  $\rho(0) = \pi_{00}$ . By Lemma 1,  $\pi_{i0} = p^{-m} \binom{m-1-i+p}{m}$ , so  $\max_i \pi_{i0} = \pi_{00}$ . Now we let  $x \rightarrow 0+$  in the inequalities (25) and get  $\lim_{x \rightarrow 0+} \rho(x) = \pi_{00} = \rho(0)$ .  $\square$

**Lemma 5.** *The rate function  $I$  is a nonnegative convex function that is finite and continuous precisely on  $[0, m-1]$ ; it is symmetric about  $(m-1)/2$ ;  $I((m-1)/2) = 0 = \min_z I(z)$ ; and  $I(0) = I(m-1) = m \log p - \log \binom{p+m-1}{m}$ .*

*Proof.* Recall equation (24):

$$I(z) = \log \sup_{x>0} \frac{x^z}{\rho(x)}.$$

By Lemma 4,  $\rho(x) \geq \pi_{00} > 0$  and  $\rho(x) \asymp x^{m-1}$  as  $x \rightarrow \infty$ , and so  $0 \leq I(z) < \infty$  for  $0 \leq z \leq m-1$  and  $I(z) = \infty$  otherwise. The rate function  $I$  is convex, by Theorem B, and so  $I$  is continuous on  $(0, m-1)$ .

By ergodicity of the Markov chain  $\langle C_k \rangle$ , with probability 1,  $(C_1 + \cdots + C_n)/n \rightarrow (m-1)/2$ , the mean of the stationary distribution (see Lemma 3). It is then a standard result that  $I((m-1)/2) = 0$ . For, if  $\Gamma$  is an arbitrary interval containing  $(m-1)/2$ , then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \Pr \left( \frac{C_1 + \cdots + C_n}{n} \in \Gamma \right) = 0,$$

but, by Proposition 3 and the continuity of  $I$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \Pr \left( \frac{C_1 + \cdots + C_n}{n} \in \Gamma \right) = - \inf_{z \in \Gamma} I(z),$$

and therefore  $\min I(z) = 0 = I((m-1)/2)$ .

By (24) and Lemma 4,

$$\begin{aligned} I(m-1-z) &= \log \sup_{x>0} \frac{x^{m-1-z}}{\rho(x)} \\ &= \log \sup_{x>0} \frac{x^{m-1-z}}{x^{m-1} \rho(1/x)} \\ &= \log \sup_{x>0} \frac{(1/x)^z}{\rho(1/x)} = I(z), \end{aligned}$$

so  $I$  is symmetric about  $(m-1)/2$ . Thus  $I(0) = I(m-1)$ , and, by (24) again and Lemma 4,

$$\begin{aligned} I(0) &= \log \sup_{x>0} \frac{x^0}{\rho(x)} = \log \frac{1}{\rho(0)} \\ &= \log \frac{1}{p^{-m} \binom{p+m-1}{m}} = m \log p - \log \binom{p+m-1}{m}. \end{aligned}$$

Finally, we check the continuity of  $I$  from the right at 0 (and, by symmetry, continuity from the left at  $m-1$ ). Assume  $0 \leq z \leq m-1$ . For  $0 < x \leq 1$ ,  $x^z/\rho(x) \leq 1/\rho(0) = 1/\pi_{00}$ , since both  $x^z$  and  $\rho(x)$  are nondecreasing functions of  $x$ ; for  $x \geq 1$ ,  $x^z/\rho(x) \leq x^z/(\pi_{00}x^{m-1}) = 1/(\pi_{00}x^{m-1-z})$ , by Lemma 4. Therefore,

$$\sup_{x>0} \frac{x^z}{\rho(x)} \leq \frac{1}{\pi_{00}}.$$

On the other hand, we claim this bound is the limit as  $z \rightarrow 0+$ . Let  $\varepsilon > 0$  (and  $\varepsilon < 1/\pi_{00}$ ). Let

$$\eta = \frac{\varepsilon \pi_{00}}{1 - \varepsilon \pi_{00}}, \text{ so that } \frac{1}{\pi_{00} + \eta} = \frac{1}{\pi_{00}} - \varepsilon.$$

Because  $\rho$  is increasing and continuous from the right at 0 and  $\rho(0) = \pi_{00}$ , there exists  $\delta > 0$  such that

$$0 < x < \delta \implies \pi_{00} < \rho(x) < \pi_{00} + \eta.$$

Then,

$$\sup_{x>0} \frac{x^z}{\rho(x)} \geq \sup_{0 < x < \delta} \frac{x^z}{\pi_{00} + \eta} = \frac{\delta^z}{\pi_{00} + \eta}.$$

But

$$\lim_{z \rightarrow 0+} \frac{\delta^z}{\pi_{00} + \eta} = \frac{1}{\pi_{00} + \eta} = \frac{1}{\pi_{00}} - \varepsilon.$$

Because  $\varepsilon$  was arbitrary,

$$\lim_{z \rightarrow 0+} \sup_{x > 0} \frac{x^z}{\rho(x)} = \frac{1}{\pi_{00}},$$

and so

$$\lim_{z \rightarrow 0+} I(z) = \log \frac{1}{\pi_{00}} = I(0). \quad \square$$

Let

$$(26) \quad D(z) = m - \frac{I(z)}{\log p}.$$

Combining Proposition 3 and the above lemmas, we have our main result for multinomial coefficients.

**Theorem 5.** *Let  $p$  be a prime, and let  $m \geq 2$  be an integer. Let*

$$N_n(\Gamma) = \left\{ (x_1, \dots, x_m) \mid 0 \leq x_1, \dots, x_m < p^n \text{ and } \frac{1}{n} \text{ord}_p \begin{pmatrix} x_1 + \dots + x_m \\ x_1, \dots, x_m \end{pmatrix} \in \Gamma \right\}.$$

*Then, for each nondegenerate interval  $\Gamma \subseteq [0, m-1]$ ,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_p N_n(\Gamma) = \sup_{z \in \Gamma} D(z)$$

*where  $D$  is given by (26);  $D$  is a concave function that is finite and continuous on  $[0, m-1]$ ;  $D$  is symmetric about  $(m-1)/2$ ;  $D((m-1)/2) = m$ ; and  $D(0) = D(m-1) = \log_p \binom{p+m-1}{m}$ .*

**4.5. Special Dimensions.** As in the binomial case, the extreme values of  $D(z)$  are special dimensions. When  $z = (m-1)/2$ , we have  $D(z) = m$ , the full dimension of the space of  $m$ -tuples. This corresponds, as noted in the proof of Lemma 5, to the almost sure convergence of  $(C_1 + \dots + C_n)/n$  to  $(m-1)/2$ .

The value of  $D(z)$  at  $z = 0$  is

$$D(0) = \frac{\log \binom{p+m-1}{m}}{\log p},$$

which is the self-similarity dimension of the multinomial coefficients *not* divisible by  $p$  (see [28], [43], [44]). (The self-similarity upon scaling by  $p$  may be deduced from Dickson's theorem.) Thus, as in the  $m = 2$  (binomial) case, Theorem 5 may be extended to include  $\Gamma = \{0\}$ , thereby incorporating this fractal dimension result at one endpoint of a full spectrum of divisibility dimensions for multinomial coefficients.

## 5. GENERALIZED MULTINOMIAL COEFFICIENTS

**5.1. Divisibility and Carries.** Both multinomial coefficients and Knuth and Wilf's generalized binomial coefficients are special cases of generalized multinomial coefficients.

**Definition.** Let  $\mathcal{U} = \langle U_1, U_2, \dots \rangle$  be a sequence of positive integers. For  $x \in \mathbb{N}$ , let  $[x]! = U_1 U_2 \cdots U_x$ ; for integer  $m \geq 2$ , define the *generalized  $m$ -nomial coefficient* corresponding to  $\mathcal{U}$  for nonnegative integers  $x_1, \dots, x_m$  by

$$\binom{x_1 + x_2 + \cdots + x_m}{x_1, x_2, \dots, x_m}_{\mathcal{U}} = \frac{[x_1 + x_2 + \cdots + x_m]!}{[x_1]![x_2]! \cdots [x_m]!}.$$

When  $m = 2$  this is the same as the  $\mathcal{U}$ -nomial coefficient of Knuth and Wilf.

The results of the previous sections may be subsumed under a more general asymptotic divisibility theory of generalized multinomial coefficients. In this section we shall give full statements of these comprehensive theorems. The hard work of the requisite proofs has already been done, for the most part, in the previous sections, so here we will streamline the exposition by merely outlining most of the arguments. First we state a generalization of Dickson's theorem, Theorem D, that relates prime-power divisibility of generalized multinomial coefficients to the total of the carries that arise when the components  $x_1, \dots, x_m$  are added in a suitable mixed-radix system.

**Theorem E.** Let  $p$  be a prime. If  $\mathcal{U}$  is a regularly divisible sequence, then we have that  $\text{ord}_p \binom{x_1 + x_2 + \cdots + x_m}{x_1, x_2, \dots, x_m}_{\mathcal{U}}$  is equal to the total of the carries that occur when  $x_1, \dots, x_m$  are added in the mixed-radix system with radices  $b_1 := r(p)$ , the rank of apparition of  $p$  in  $\mathcal{U}$ ,  $b_2 := r(p^2)/r(p)$ ,  $b_3 := r(p^3)/r(p^2)$ , etc.

The proof is a straightforward generalization of the (Knuth and Wilf-based) proof of Theorem C.

**5.2. The Markov Chain of Carries.** For the carries  $\langle C_k \rangle$  that come from the addition of  $m$  numbers  $x_i = \sum_j X_{ij} \prod_{k=1}^j b_k$  given by independent random digits  $X_{ij}$  in a mixed-radix system, the Markov property still holds:

$$\Pr(C_{k+1} = c_{k+1} \mid C_k = c_k, \dots, C_1 = c_1, C_0 = 0) = \Pr(C_{k+1} = c_{k+1} \mid C_k = c_k),$$

because the carry out  $C_{k+1}$  depends only on  $C_k$  (and the digits  $X_{1,k}, \dots, X_{m,k}$ ): if  $C_k = i$ ,

$$C_{k+1} = j \iff j b_{k+1} \leq i + X_{1,k} + \cdots + X_{m,k} < (j+1) b_{k+1}.$$

Now the proof of equation (22) works with  $p$  replaced by  $b_{k+1}$ , so the transition probabilities are given by

$$(27) \quad \pi_{ij}(k) = b_{k+1}^{-m} \sum_{t=0}^{j - \lfloor i/b_{k+1} \rfloor} (-1)^t \binom{m+1}{t} \binom{m-1-i-(j+1-t)b_{k+1}}{m}.$$

**5.3. The Large Deviation Principle.** Let

$$Z_n = \frac{1}{n} \text{ord}_p \binom{x_1 + \cdots + x_m}{x_1, \dots, x_m}_{\mathcal{U}}.$$

By Theorem E, we have  $Z_n = (C_1 + \cdots + C_n)/n$ , where  $\langle C_k \rangle$  is the Markov chain of carries. As in the case of generalized binomial coefficients, we suppose that the

transition probabilities for the carries process are eventually stationary:  $\pi_{ij}(k) = \pi_{ij}$  for  $k \geq s$  for some  $s$ . This happens when the radices  $b_k := r(p^k)/r(p^{k-1})$  become constant for  $k > s$ .

At this point we note a further generalization. It is not necessary that the stationary transition matrix  $\mathbf{\Pi}$  be based on the prime  $p$  as in equation (19) or equation (22), i.e., it is not necessary that the eventual value of  $r(p^k)/r(p^{k-1})$  be the prime  $p$ . We need only require that  $r(p^k)/r(p^{k-1})$  eventually have some integer value  $b > 1$ . Then  $\pi_{ij}$  is given by equation (27) with  $b_{k+1} = b$ . A class of regularly divisible sequences satisfying this criterion may be constructed as follows. Let  $\psi$  be any permutation of the primes, and let  $\mathcal{U} : \mathbb{N} \rightarrow \mathbb{N}$  be defined by the rule:

$$\text{if } n = p_1^{e_1} p_2^{e_2} \cdots, \text{ then } U_n = \psi(p_1)^{e_1} \psi(p_2)^{e_2} \cdots.$$

For  $\Gamma \subset \mathbb{R}$ , let

$$N_n(\Gamma) = \# \left\{ (x_1, \dots, x_m) \mid 0 \leq x_1, \dots, x_m < r(p^n) \text{ and } Z_n \in \Gamma \right\}.$$

Since  $r(p^n) = b_1 \cdots b_s \cdot b^{n-s}$ ,

$$(28) \quad N_n(\Gamma) = [r(p^n)]^m \Pr(Z_n \in \Gamma) = (b_1 \cdots b_s \cdot b^{n-s})^m b^{mn} \Pr(Z_n \in \Gamma).$$

Now we may modify the proof of the main theorem for generalized binomial coefficients, Theorem 4, to show that the “transient” transition probabilities  $\pi_{ij}(k)$  for  $k < s$  have no lingering effect on the asymptotic behavior of  $\log \Pr(Z_n \in \Gamma)$  when  $\Gamma$  is a nondegenerate subinterval of  $[0, m-1]$ . Then, taking (28) into account, we may generalize the large deviation principle for multinomial coefficients, Proposition 3, to conclude that

$$(29) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log_b N_n(\Gamma) = D(\Gamma),$$

where

$$(30) \quad D(\Gamma) = \sup_{z \in \Gamma} D(z) \quad \text{where} \quad D(z) = m - I(z)/\log b,$$

and  $I$  is the rate function for the finite Markov chain corresponding to the stationary transition probabilities given by (27), i.e.,

$$I(z) = \sup_t \{tz - \log \rho(\mathbf{\Pi}_t)\},$$

where

$$\mathbf{\Pi}_t = [\pi_{ij} e^{tj}]$$

and

$$\rho(\mathbf{\Pi}_t) = \text{Perron-Frobenius eigenvalue of } \mathbf{\Pi}_t.$$

**5.4. Main Theorem.** We have shown that our main conclusion, equation (29), follows from the large deviation principle for the chain given by (27). Let us also observe that the special properties of the dimension function  $D$  for the multinomial case were deduced using only the transition probabilities (22), and in the present case the transition matrix has the same form (just replace  $p$  by  $b$ ), so the dimension function given in (30) must enjoy the same properties. Thus we may summarize our discussion in the following statement.



**Main Theorem.** Let  $\mathcal{U}$  be a regularly divisible sequence, let  $p$  be a prime, let  $r(k) = \min\{j : k|U_j\}$ , and assume that for some  $b, s \in \mathbb{N}$  we have  $b_k := r(p^k)/r(p^{k-1}) = b > 1$  for every  $k > s$ .

Let  $m \geq 2$  be a fixed integer; for  $n \in \mathbb{N}$  and  $\Gamma \subset \mathbb{R}$ , define

$$N_n(\Gamma) = \# \left\{ (x_1, \dots, x_m) \mid \begin{array}{l} 0 \leq x_1, \dots, x_m < r(p^n) \\ \& \binom{x_1 + x_2 + \dots + x_m}{x_1, x_2, \dots, x_m}_{\mathcal{U}} \in \Gamma \end{array} \right\}.$$

Then, for each nondegenerate interval  $\Gamma \subseteq [0, m-1]$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_b N_n(\Gamma) = D(\Gamma),$$

where  $D(\Gamma)$  is given by (30). In the case that  $m = 2$ ,  $D$  is given explicitly by Theorem 2 with  $p$  replaced by  $b$ . In any case,  $D$  is a concave function that is finite and continuous on  $[0, m-1]$ ;  $D$  is symmetric about  $(m-1)/2$ ;  $D((m-1)/2) = m$ ; and  $D(0) = D(m-1) = \log_b \binom{b+m-1}{m}$ .

**5.5. Multifractal Spectrum of Generalized Multinomial Coefficients.** We conclude by applying our Main Theorem to the determination of the multifractal spectrum of the generalized multinomial coefficients, i.e., to the evaluation of

$$f(\alpha) = \lim_{\varepsilon \rightarrow 0+} \lim_{\delta \rightarrow 0+} \frac{\log[\tilde{N}_\delta(\alpha + \varepsilon) - \tilde{N}_\delta(\alpha - \varepsilon)]}{-\log \delta}$$

when

$$\tilde{N}_\delta(\alpha) = \#\{\mathbf{x} \in \mathcal{M}_\delta : |\phi(\mathbf{x})|_p \geq \delta^\alpha\},$$

where

$$\mathcal{M}_\delta = \{(x_1, \dots, x_m) \in \mathbb{N}_0^m : 0 \leq x_1, \dots, x_m < \delta^{-1}\}$$

and

$$\phi(\mathbf{x}) = \phi((x_1, \dots, x_m)) = \binom{x_1 + x_2 + \dots + x_m}{x_1, x_2, \dots, x_m}_{\mathcal{U}}.$$

The derivation follows the main lines of the proof that  $f(\alpha) = D(\alpha)$  for binomial coefficients, but the details are a little different, and the result differs in some cases too, so we will present them with some care. First we will establish a bound on  $\tilde{N}_\delta(\alpha + \varepsilon) - \tilde{N}_\delta(\alpha - \varepsilon)$ .

**Lemma.** Assume  $n > s$  and  $1 < r(p^{n-1}) < \delta^{-1} \leq r(p^n)$  and  $\varepsilon > 0$ . Write

$$\log_p r(p^n) = \log_p(b_1 \cdots b_s b^{-s}) + n \log_p b = \gamma + \theta n, \quad \text{where } \theta = \log_p b > 0.$$

Then

$$\begin{aligned} N_{n-1}((\theta + (\gamma + \theta)/(n-1))(\alpha - \varepsilon), (\theta + \gamma/(n-1))(\alpha + \varepsilon)) \\ \leq \tilde{N}_\delta(\alpha + \varepsilon) - \tilde{N}_\delta(\alpha - \varepsilon) \\ \leq N_n((\theta + (\gamma - \theta)/n)(\alpha - \varepsilon), (\theta + \gamma/n)(\alpha + \varepsilon)), \end{aligned}$$

where  $0 < \varepsilon < |\alpha|$  if  $\alpha \neq 0$ .

*Proof.* Notice that for  $n > s$ , we have  $r(p^{n-1}) > 1$  iff  $\theta(n-1) + \gamma > 0$ , i.e.,  $n > 1 - \gamma/\theta$ , so the coefficients of  $\alpha \pm \varepsilon$  above are positive. (Also the intervals are nondegenerate for sufficiently large  $n$  ( $n > \alpha/2\varepsilon + 1/2 - \gamma/\theta$ ).) Assume that  $0 < \varepsilon < \alpha$  or  $\alpha = 0 < \varepsilon$ . Now,

$$\begin{aligned}
 \mathbf{x} \in \mathcal{M}_{1/r(p^{n-1})} \quad & \text{and} \quad \left(\theta + \frac{\gamma + \theta}{n-1}\right)(\alpha - \varepsilon) < \frac{\text{ord}_p \phi(\mathbf{x})}{n-1} \\
 & \leq \left(\theta + \frac{\gamma}{n-1}\right)(\alpha + \varepsilon) \\
 \implies \mathbf{x} \in \mathcal{M}_\delta \quad & \text{and} \quad (\theta n + \gamma)(\alpha - \varepsilon) < \text{ord}_p \phi(\mathbf{x}) < (\theta(n-1) + \gamma)(\alpha + \varepsilon) \\
 \implies \mathbf{x} \in \mathcal{M}_\delta \quad & \text{and} \quad \alpha - \varepsilon < \frac{\text{ord}_p \phi(\mathbf{x})}{\log_p \delta^{-1}} < \alpha + \varepsilon \\
 \implies \mathbf{x} \in \mathcal{M}_\delta \quad & \text{and} \quad \delta^{\alpha+\varepsilon} \leq |\phi(\mathbf{x})|_p < \delta^{\alpha-\varepsilon} \\
 \implies \mathbf{x} \in \mathcal{M}_{1/r(p^n)} \quad & \text{and} \quad \alpha - \varepsilon < \frac{\text{ord}_p \phi(\mathbf{x})}{\log_p \delta^{-1}} \leq \alpha + \varepsilon \\
 \implies \mathbf{x} \in \mathcal{M}_{1/r(p^n)} \quad & \text{and} \quad (\theta(n-1) + \gamma)(\alpha - \varepsilon) < \text{ord}_p \phi(\mathbf{x}) \leq (\theta n + \gamma)(\alpha + \varepsilon) \\
 \implies \mathbf{x} \in \mathcal{M}_{1/r(p^n)} \quad & \text{and} \quad \left(\theta + \frac{\gamma - \theta}{n}\right)(\alpha - \varepsilon) < \frac{\text{ord}_p \phi(\mathbf{x})}{n} \leq \left(\theta + \frac{\gamma}{n}\right)(\alpha + \varepsilon).
 \end{aligned}$$

Therefore the conclusion of the lemma follows in the case that  $0 < \varepsilon < \alpha$  or  $\alpha = 0 < \varepsilon$ . If  $\alpha < 0$  and  $0 < \varepsilon < |\alpha|$ , then  $\alpha \pm \varepsilon < 0$  and all the terms in the lemma inequalities become zero.  $\square$

Now we are in a position to calculate  $f(\alpha)$ . Let  $\varepsilon > 0$  and  $\theta = \log_p b$ . First suppose  $\alpha \geq 0$ ; if  $\alpha > 0$ , require  $0 < \varepsilon < \alpha$ . Choose  $\eta$  such that  $0 < \eta < \theta$  and, if  $\alpha > 0$ , such that  $\eta < \theta\varepsilon/(\alpha + \varepsilon)$  (so that the intervals defined below will include  $\theta\alpha$ ). Require that  $n > s$ ,  $n > 1 - \gamma/\theta$ , and  $n$  be large enough that

$$\left| \frac{\gamma + \theta}{n-1} \right| < \eta, \quad \left| \frac{\gamma}{n-1} \right| < \eta, \quad \left| \frac{\gamma - \theta}{n} \right| < \eta, \quad \text{and} \quad \left| \frac{\gamma}{n} \right| < \eta.$$

Then

$$\begin{aligned}
 & ((\theta + \eta)(\alpha - \varepsilon), (\theta - \eta)(\alpha + \varepsilon)] \\
 & \subset \left( \left( \theta + \frac{\gamma + \theta}{n-1} \right)(\alpha - \varepsilon), \left( \theta + \frac{\gamma}{n-1} \right)(\alpha + \varepsilon) \right] =: A_{n-1}
 \end{aligned}$$

and

$$\begin{aligned}
 B_n & := \left( \left( \theta + \frac{\gamma - \theta}{n} \right)(\alpha - \varepsilon), \left( \theta + \frac{\gamma}{n} \right)(\alpha + \varepsilon) \right] \\
 & \subset ((\theta - \eta)(\alpha - \varepsilon), (\theta + \eta)(\alpha + \varepsilon)].
 \end{aligned}$$

Then, by the lemma,

$$N_{n-1}(A_{n-1}) \leq \tilde{N}_\delta(\alpha + \varepsilon) - \tilde{N}_\delta(\alpha - \varepsilon) \leq N_n(B_n),$$

so

$$\begin{aligned} \frac{n-1}{\log \delta^{-1}} \frac{\log N_{n-1}(((\theta+\eta)(\alpha-\varepsilon), (\theta-\eta)(\alpha+\varepsilon)))}{n-1} \\ \leq \frac{\log[\tilde{N}_\delta(\alpha+\varepsilon) - \tilde{N}_\delta(\alpha-\varepsilon)]}{-\log \delta} \\ \leq \frac{n}{\log \delta^{-1}} \frac{\log N_n(((\theta-\eta)(\alpha-\varepsilon), (\theta+\eta)(\alpha+\varepsilon)))}{n}. \end{aligned}$$

Let  $n \rightarrow \infty$  (whence  $\delta \rightarrow 0+$ ). By the Main Theorem  $(1/n) \log N_n(\Gamma) \rightarrow D(\Gamma) \log b$ , where  $D$  is as specified in the theorem. Also,  $n/\log \delta^{-1} \rightarrow 1/\log b$ . So the limit inferior and limit superior of the middle term lie between

$$D(((\theta+\eta)(\alpha-\varepsilon), (\theta-\eta)(\alpha+\varepsilon)))$$

and

$$D(((\theta-\eta)(\alpha-\varepsilon), (\theta+\eta)(\alpha+\varepsilon))).$$

Let  $\varepsilon \rightarrow 0+$  (whence also  $\eta \rightarrow 0+$ ). By the continuity of  $D$  on  $[0, m-1]$ ,

$$f(\alpha) = \lim_{\varepsilon \rightarrow 0+} \lim_{\delta \rightarrow 0+} \frac{\log[\tilde{N}_\delta(\alpha+\varepsilon) - \tilde{N}_\delta(\alpha-\varepsilon)]}{-\log \delta} = D(\theta\alpha)$$

for  $0 \leq \alpha \leq (m-1)/\theta$ .

If  $\alpha < 0$ , then  $\delta^{\alpha \pm \varepsilon} > 1$  for sufficiently small  $\varepsilon$ , while  $|\phi(\mathbf{x})|_p \leq 1$ , so  $\tilde{N}_\delta(\alpha \pm \varepsilon) = 0$ , and  $f(\alpha) = -\infty = D(\theta\alpha)$ .

If  $\theta\alpha > m-1$ , then for  $r(p^{n-1}) < \delta^{-1} \leq r(p^n)$  we have, since  $r(p^{n-1}) = p^{\gamma+\theta(n-1)}$ ,

$$\delta^{\alpha-\varepsilon} > \delta^{\alpha+\varepsilon} \geq \left(p^{-\gamma-\theta(n-1)}\right)^{\alpha+\varepsilon} = p^{(\theta-\gamma)(\alpha+\varepsilon)} p^{-\theta\alpha n} > p^{-(m-1)n}$$

for small  $\varepsilon$  and large  $n$ . But because  $\text{ord}_p \phi(\mathbf{x}) \leq n(m-1)$  (bound on carries),  $|\phi(\mathbf{x})|_p \geq p^{-(m-1)n}$ . So  $\tilde{N}(\alpha+\varepsilon) = \tilde{N}(\alpha-\varepsilon)$ , and again  $f(\alpha) = -\infty = D(\theta\alpha)$ . This proves our final result.

**Theorem 6.** *Let  $p$  be a prime, and let  $m \geq 2$  be an integer. Let  $\mathcal{U}$  be a regularly divisible sequence, and let*

$$\phi(x_1, \dots, x_m) = \begin{pmatrix} x_1 + x_2 + \dots + x_m \\ x_1, x_2, \dots, x_m \end{pmatrix}_{\mathcal{U}}.$$

*Suppose there exist  $b, s \in \mathbb{N}$  such that  $r(p^k)/r(p^{k-1}) = b > 1$  for every  $k > s$ . Then the multifractal spectrum  $f$  of  $\phi$  exists and is given by  $f(\alpha) = D(\alpha \log_p b)$ , where  $D$  is the dimension function given by equation (30).*

#### ACKNOWLEDGMENTS

I thank Benoit Mandelbrot, who encouraged me to consider a multifractal perspective in my research, and I thank Gerald Bergum for help with Fibonacci-sequence references.

## REFERENCES

1. W. A. Broomhead, *Pascal (mod p)*, Math. Gaz. **56** (1972), No. 398, 267–271. MR **58**:5250
2. L. Carlitz, *The number of binomial coefficients divisible by a fixed power of a prime*, Rend. Circ. Mat. Palermo **16** (1967), 299–320. MR **40**:2554
3. G. J. Chaitin, *Algorithmic Information Theory*, Cambridge U. Press, Cambridge, 1987. MR **89g**:68022
4. K. L. Chung, *Markov Chains with Stationary Transition Probabilities*, second ed., Springer, Berlin, 1967. MR **36**:961
5. L. E. Dickson, *Analytic representation of substitution*, I, Annals of Math. **11** (1897), nos. 3–4, 65–120.
6. L. E. Dickson, *History of the Theory of Numbers, Volume I: Divisibility and Primality*, Chelsea, New York, 1923; reprint, 1966. MR **39**:6807a
7. A. Dembo and O. Zeitouni, *Large Deviation Techniques and Applications*, Jones and Bartlett, Boston, 1993. MR **95a**:60034
8. M. D. Donsker and S. R. S. Varadhan, *Asymptotic evaluation of certain Markov process expectations for large time*, I, Commun. Pure Appl. Math. **28** (1975), 1–47. MR **52**:6883
9. R. S. Ellis, *Large deviations for a general class of random vectors*, Ann. Probab. **12** (1984), 1–12. MR **85e**:60032
10. R. S. Ellis, *Entropy, Large Deviations and Statistical Mechanics*, Springer-Verlag, New York, 1985. MR **87d**:82008
11. L. Euler, *Institutiones calculi differentialis*, Imperial Academy of Sciences, St. Petersburg, 1755.
12. L. Euler, *Opera Omnia* (1) **10**, Teubner, Leipzig and Berlin, 1913.
13. C. J. G. Evertsz and B. B. Mandelbrot, *Multifractal measures*, in H.-O. Peitgen, H. Jürgens, and D. Saupe, *Chaos and Fractals: New Frontiers of Science*, Springer-Verlag, New York, 1992, 921–953. MR **93k**:58157
14. K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, J. Wiley, Chichester, 1993. MR **92j**:28008
15. J. Feder, *Fractals*, Plenum Press, New York, 1988. MR **89i**:76001
16. W. Fenchel, *On conjugate convex functions*, Canad. J. Math., **1** (1949), 73–77. MR **10**:435b
17. D. Flath and R. Peele, *Hausdorff dimension in Pascal's triangle*, Applications of Fibonacci Numbers **5**, G. E. Bergum et al., eds., Kluwer, Netherlands, 1993, 229–244. MR **95a**:11068
18. U. Frisch and G. Parisi, *On the singularity structure of fully developed turbulence*, in Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics, M. Ghil, R. Benzi, and G. Parisi, eds., North-Holland, New York, 1985, 84–88.
19. G. Frobenius, *Über matrizen aus nicht-negativen elementen*, Sitzber. Kgl. Preuss. Akad. Wiss., 1912, 456–477.
20. C. F. Gauss, *Summatio quarundam serierum singularium*. Comment. Soc. R. Scient. Gotttingensis Rec. **1** (1811), reprinted in his Werke **2** (1863), 9–45.
21. R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*, second ed., Addison-Wesley, Reading, Mass., 1994. CMP 96:14; MR **91f**:00001 (1st ed.)
22. P. Grassberger, *Generalized dimensions of strange attractors*, Phys. Lett. A **97** (1983), 227–230. MR **84i**:58075
23. F. v. Haeseler, H.-O. Peitgen, and G. Skordev, *Pascal's triangle, dynamical systems and attractors*, Ergod. Th. & Dynam. Sys. **12** (1992), 479–486. MR **93k**:28009
24. J. H. Halton, *On the divisibility properties of Fibonacci numbers*, Fibonacci Quart. **4** (1966), 217–240. MR **34**:1257
25. K. Hensel, *Zahlentheorie*, Göschen, Berlin, 1913.
26. H. G. E. Hentschel and I. Procaccia, *The infinite number of generalized dimensions of fractals and strange attractors*, Physica D **8** (1983), 435–444. MR **85a**:58064
27. R. Holley and E. C. Waymire, *Multifractal dimensions and scaling exponents for strongly bounded random cascades*, Ann. Appl. Probab. **2** (1992), 819–845. MR **93k**:60122
28. J. Holte, *The dimension of the set of multinomial coefficients not divisible by n*, Abstracts of the AMS **12**, No. 1 (Jan. 1991), 21.
29. J. M. Holte, *Carries, combinatorics, and an amazing matrix*, Amer. Math. Monthly **104** (1997), 138–149.

30. M. H. Jensen, L. P. Kadanoff, A. Libchaber, I. Procaccia, and J. Stavans, *Global universality at the onset of chaos: Results of a forced Rayleigh-Bénard experiment*, Phys. Rev. Lett. **55** (1985), 2798–2801.
31. D. E. Knuth, *The Art of Computer Programming, Vol. 2, Seminumerical Algorithms*, Addison-Wesley, Reading, Mass., 1969, 1981. MR **44**:3531; MR **83i**:68003
32. D. E. Knuth and H. S. Wilf, *The power of a prime that divides a generalized binomial coefficient*, J. Reine Angew. Math. **396** (1989), 212–219. MR **90d**:11029
33. N. Koblitz, *p-adic Numbers, p-adic Analysis, and Zeta-Functions*, second ed., Springer-Verlag, New York, 1984. MR **86c**:11086
34. E. E. Kummer, *Über die Ergänzungssätze zu den allgemeinen Reciprocitätsgesetzen*, J. Reine Angew. Math. **44** (1852), 93–146.
35. C. T. Long, *Pascal's triangle modulo  $p$* , Fibonacci Quart. **19** (1981), 458–463. MR **83a**:05013b
36. E. Lucas, *Théorie des fonctions numériques simplement périodiques*, Amer. J. Math. **1** (1878), 184–240.
37. B. B. Mandelbrot, *Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence*, in Statistical Models and Turbulence, M. Rosenblatt and C. Van Atta, eds., Lecture Notes in Physics **12**, Springer-Verlag, New York, 1972, 333–351.
38. B. B. Mandelbrot, *Intermittent turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier*, J. Fluid Mech. **62** (1974), 331–358.
39. B. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman, New York, 1982, 1983. MR **84h**:00021
40. H. Miller, *A convexity property in the theory of random variables defined on a finite Markov chain*, Ann. Math. Statist. **32** (1961), 1260–1270. MR **23**:A4180
41. G. Paladin and A. Vulpiani, *Characterization of strange attractors as inhomogeneous fractals*, Lett. Nuovo Cimento **41** (1984), 82–86. MR **86b**:58075
42. O. Perron, *Zur Theorie der Matrizen*, Math. Ann. **64** (1907), 248–263.
43. A. M. Reiter, *Determining the dimension of fractals generated by Pascal's triangle*, Westinghouse Science Talent Search 1990–1991 winning paper.
44. A. M. Reiter, *Determining the dimension of fractals generated by Pascal's triangle*, Fibonacci Quart. **31** (1993), 112–120. MR **94b**:05011
45. A. Rényi, *Probability Theory*, North-Holland, Amsterdam, 1970. MR **47**:4296
46. J. B. Roberts, *On binomial coefficient residues*, Canadian J. of Math. **9** (1957), 363–370. MR **19**:250e
47. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970. MR **43**:445
48. E. Seneta, *Non-negative Matrices and Markov Chains*, second ed., Springer-Verlag, New York, 1981. MR **85i**:60058
49. M. Sved, *Divisibility—with visibility*, Math. Intelligencer **10** (1988), 56–64. MR **89b**:11019
50. W. Sierpiński, *Sur une courbe dont tout point est un point de ramification*, Comptes Rendus (Paris) **160** (1915), 302.
51. J. Stoer and C. Witzgall, *Convexity and Optimization in Finite Dimensions I*, Springer-Verlag, Berlin, 1970. MR **44**:3707
52. J. R. Weaver, *Centrosymmetric (cross-symmetric) matrices, their basic properties, eigenvalues, and eigenvectors*, Amer. Math. Monthly **92** (1985), 711–717. MR **87b**:15036
53. W. A. Webb and D. L. Wells, *Kummer's Theorem, Cantor Bases, and Generalized Binomial Coefficients*, preprint.
54. D. L. Wells, *The Fibonacci and Lucas triangles modulo 2*, Fibonacci Quart. **32** (1994), 111–123. MR **95d**:11022
55. H. Wielandt, *Unzerlegbare, nicht negative Matrizen*, Math. Z. **52** (1950), 642–648. MR **11**:710g
56. S. Wolfram, *Geometry of binomial coefficients*, Amer. Math. Monthly **91** (1984), 566–571. MR **86d**:05007

DEPARTMENT OF MATHEMATICS, GUSTAVUS ADOLPHUS COLLEGE, ST. PETER, MINNESOTA 56082  
*E-mail address*: holte@gac.edu