J. Fractal Geom. 4 (2017), 105–126
DOI 10.4171/JFG/46

Journal of Fractal Geometry

Spatial equidistribution of combinatorial number schemes

Florian Greinecker[1]

Abstract. In this paper we use a *generalized Lucas type congruence* for certain combinatorial number schemes to define IFS. We show some distribution properties of these fractals. As an application we prove that the binomial coefficients, the Stirling numbers of the first and second kind as well as the multinomial coefficients are spatially equidistributed in the nonzero residue classes modulo a prime p.

Mathematics Subject Classification (2010). Primary: 28A80; Secondary: 11A63, 11B65, 11B73.

Keywords. IFS, Stirling numbers, Lucas type congruences, equidistribution.

1. Introduction

Let us start with an investigation of the binomial coefficients, as well as the Stirling numbers of the first and second kind. If we color them according to their residue classes modulo a prime p, they have a self-similar, fractal structure, which we can observe in Figure 1. The most prominent member of this class of fractals is the Sierpinski triangle, which we obtain if we color the binomial coefficients according to their residue class modulo 2.

The binomial coefficients and Stirling numbers of the first and second kind share some properties; they have a combinatorial interpretation, they have a recursive structure and they can be easily expressed by generating functions.

We will focus on two further common properties.

[1] The author is supported by the Austrian Science Fund FWF projects W1230, Doctoral Program "Discrete Mathematics," and F5503 (part of the special research program (SFB) "Quasi-Monte Carlo Methods: Theory and Applications"). The author wants to express his gratitude to his advisor Peter Grabner and an anonymous referee whose comments and feedback greatly improved this article.

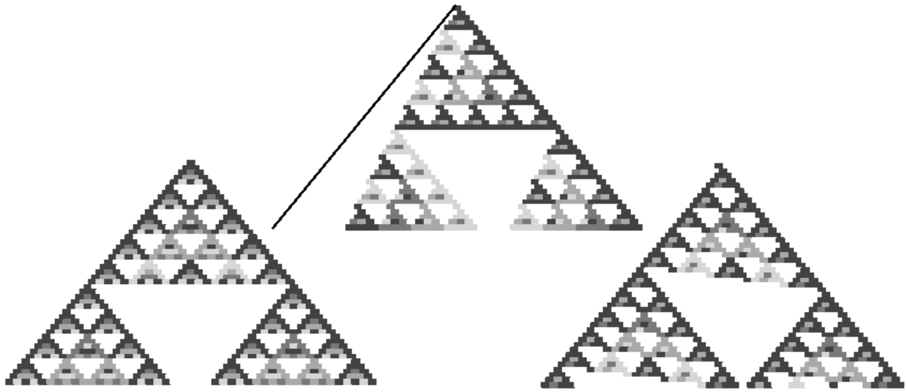

Figure 1. Binomial coefficients, Stirling numbers of the first and second kind modulo 5

The p-divisibility property. For a fixed prime p, almost all binomial coefficients, Stirling numbers of the first and the second kind are divisible by p.

The equidistribution property. The binomial coefficients, Stirling numbers of the first and second kind in the nonzero residue classes modulo a prime p approach an equidistribution in the first n lines, if n goes to infinity.

The first property is very well studied. Regarding the binomial coefficients, Fine [16] gave a formula for the number of zeros in a row modulo p. Carlitz proved the corresponding results for columns [10]. Singmaster showed in [26] that "any integer divides almost all binomial coefficients" with respect to four different definitions of "almost all." He also gives a nice survey of known results in [27]. An approach via cellular automata is described in [31]. This approach may be extended since in [2] and [3] it is shown that the binomial coefficients, unsigned Stirling numbers of the first kind modulo a prime p and the Gaussian q-binomials modulo m (if $\gcd(m, q) = 1$) are automatic sequences. The ideas of this paper can be expressed in terms of automatic sequences (see [1]).

The p-divisibility of the Stirling numbers of the first kind has been studied by Carlitz [8] and by Peele, Radcliffe and Wilf [25]. For the Stirling numbers of the second kind there are results by Carlitz [9] and Lundell [23]. A result for the q-binomial coefficients is given in Howard [19].

In this paper we will use a unified approach and show that a prime p divides almost all numbers in any number scheme which satisfies the generalized Lucas congruence. All previous examples turn out to share this property given below.

The equidistribution property is not so well studied. There are results on the equidistribution of the binomial coefficients from Garfield and Wilf [17] and by Barbolosi and Grabner [7] (with a generalization in [6]). The fractal structure of the binomial coefficients is analyzed in [30]. In our treatment of the fractal structures we will use matrices, an approach also followed in [29].

The original motivation of this paper was to extend these results to Stirling numbers of the first and second kind. We show that in all number schemes which satisfy the generalized Lucas congruence the nonzero residue classes modulo p are spatially equidistributed. This proves the equidistribution property for Stirling numbers and also generalizes the result for binomial coefficients.

The classic Lucas congruence for binomial coefficients is stated in the following theorem by Lucas (see [13], p.271).

Theorem 1.1 (Édouard Lucas 1878). *The binomial coefficients satisfy*

$$\binom{n}{k} \equiv \prod_{i=0}^{\ell} \binom{n_i}{k_i} \pmod{p}$$

with $n = \sum_{i=0}^{\ell} n_i p^i$ *and* $k = \sum_{i=0}^{\ell} k_i p^i$ *where* $0 \le n_i, k_i \le p-1$.

Lucas' theorem allows us to calculate the binomial coefficients modulo p digitwise if we write n and k in base p. We generalize this idea in the following way.

Instead of p-adic representatives, we use matrix digital systems. They allow us to define an iterated function system and a sequence U_N of sets which converge to a limit fractal. Then we assign "colors," their residue classes modulo p, to the points in U_N.

The primary idea of this paper is to define a *generalized Lucas congruence* (see Definition 2.2) for a number scheme. A number scheme is a matrix digital system together with coloring functions that assign residue classes to numbers and digits. A number scheme satisfies the generalized Lucas congruence if the residue class of a number n modulo p is the product of the residue classes of the digits of n.

The binomial coefficients and Stirling numbers of the first and second kind all have number schemes which satisfy the generalized Lucas congruence. We will show that for all number schemes which satisfy the generalized Lucas congruence the sequence U_N is p-divisible and has a generalized equidistribution property. We can even replace the equidistribution property with a stronger property and show that we have an equidistribution modulo p for all μ-continuity sets of a normalized Hausdorff measure μ. The paper is organized in the following way:

In Section 2 we define *matrix digital systems* and the *generalized Lucas congruence*. Then we state our two main results, which are stronger, formalized versions of the two properties stated in this section.

We use Section 3 to point out the fractal structure of number schemes which satisfy the generalized Lucas congruence.

Then we recall properties of Dirichlet characters in Section 4 and use them to prove the two main results.

Finally we show in Section 5 that our results are applicable to a range of well known number schemes before we finish with some concluding remarks in Section 6.

2. Results

First, we define *matrix digital systems* (based on *matrix number systems* in [22]). To avoid confusion between 2-dimensional vectors and binomial coefficients we write vectors with an arrow above them. So $\overrightarrow{\binom{5}{3}}$ denotes a vector while $\binom{5}{3}$ denotes a binomial coefficient.

For every dimension $d \geq 1$ we have the d-dimensional Euclidean space $\mathbb{R}^d$ and the embedded ring of integer vectors $\mathbb{Z}^d$. Let $A \in \mathbb{Z}^{d\times d}$ be a $d \times d$ matrix whose eigenvalues all have modulus greater then 1. Then $\mathcal{L} = A\mathbb{Z}^d$ is a subgroup of $\mathbb{Z}^d$ and the factor group $\mathbb{Z}^d/A\mathbb{Z}^d$ has order $\operatorname{ord}(\mathbb{Z}^d/A\mathbb{Z}^d) = |\det A| > 1$.

Definition 2.1. Let the *digit set* $\mathcal{D} \subseteq \mathbb{Z}^d$ be a complete residue system mod $\mathcal{L}$ with $\vec{0} \in \mathcal{D}$. We call the pair $(A, \mathcal{D})$ a *matrix digital system*.

We define the set T of all vectors $\vec{n} \in \mathbb{Z}^d$ with a representation of the form

$$\vec{n} = \vec{\varepsilon_0} + A\vec{\varepsilon_1} + A^2\vec{\varepsilon_2} + \cdots + A^\ell\vec{\varepsilon_\ell} \tag{1}$$

as

$$T := \left\{ \sum_{k=0}^{\ell} A^k \vec{\varepsilon_k} \,\middle|\, \ell \in \mathbb{N},\ \vec{\varepsilon_k} \in \mathcal{D} \right\} \subseteq \mathbb{Z}^d.$$

Since $\mathcal{D}$ is a complete residue system, the representation of the form (1) is unique, except for the number of leading zeros. The set of vectors having a representation of the form (1) with $\ell + 1$ digits is denoted by T_ℓ. We write $\vec{n}$ also as

$$\vec{n} = (\vec{\varepsilon_0}\vec{\varepsilon_1}\ldots\vec{\varepsilon_\ell}).$$

If $T = \mathbb{Z}^d$, we have a *matrix number system* in the sense of [22]. We are not interested in whether all $\vec{n} \in \mathbb{Z}^d$ have such a representation.

Definition 2.2. A function $f : T \to \mathbb{Z}/p\mathbb{Z}$ satisfies the *generalized Lucas congruence* modulo a prime p if there is a function f with $f : \mathcal{D} \to \mathbb{Z}/p\mathbb{Z}$ and a matrix digital system (with $\vec{n} = (\vec{\varepsilon_0}\vec{\varepsilon_1}\ldots\vec{\varepsilon_\ell})$) so that

$$f(\vec{n}) \equiv \prod_{i=0}^{\ell} f(\vec{\varepsilon_i}) \pmod{p}. \tag{2}$$

This implies that $f(\vec{0}) = 1$. We say that f is a *coloring* and the elements of $\mathbb{Z}/p\mathbb{Z}$ are *colors*. The generalized Lucas congruence extends the domain of f from digits to numbers.

Definition 2.3. Let f be a function that satisfies the generalized Lucas congruence. We say that f is p *multiplicatively complete* if there is a $k \in \mathbb{N}$ with

$$(\mathbb{Z}/p\mathbb{Z})^* \subseteq \{f(\vec{n}) \mid \vec{n} \in T_k\}. \tag{3}$$

Here $(\mathbb{Z}/p\mathbb{Z})^*$ denotes the unit group of $(\mathbb{Z}/p\mathbb{Z})$. If (3) holds for k, then it also holds for every $\ell \geq k$.

The following two theorems apply to all functions which satisfy the generalized Lucas congruence. Binomial coefficients and Stirling numbers of the first and second kind are the best known examples of functions which satisfy the generalized Lucas congruence. Numerous other examples can be found in Section 5 and in the comprehensive paper [24].

Theorem 2.4. *Let f be a function which satisfies the generalized Lucas congruence. If there is an $\vec{\varepsilon} \in \mathcal{D}$ so that $f(\vec{\varepsilon}) \equiv 0 \pmod{p}$, then*

$$\lim_{\ell\to\infty} \frac{\#\{\vec{n} \in T_\ell \mid f(\vec{n}) \not\equiv 0 \pmod{p}\}}{\#\{\vec{n} \in T_\ell\}} \longrightarrow 0.$$

If there is a matrix digital system with a coloring function which satisfies the generalized Lucas congruence, we speak of a *number scheme*. Let us recall that a *continuity set* of a measure μ is any Borel set B with a boundary set ∂B of measure zero $\mu(\partial B) = 0$.

The set $\mathcal{F}$ of all numbers $\vec{x} \in \mathbb{R}^d$ which can be written as

$$\vec{x} = \sum_{i=1}^{\infty} A^{-i}\vec{\varepsilon_i}$$

with $\vec{\varepsilon_i} \in \mathcal{D}$ and $f(\vec{\varepsilon_i}) \not\equiv 0 \pmod{p}$ is called the *fundamental region* with respect to $(A, \mathcal{D})$. Since

$$\mathcal{F} = \bigcup_{\substack{\vec{\varepsilon}\in\mathcal{D} \\ f(\varepsilon)\not\equiv 0 \pmod{p}}} (A^{-1}\vec{\varepsilon} + A^{-1}\mathcal{F}),$$

the set is self-similar.

Theorem 2.5. *Let $(A, \mathcal{D})$ be a matrix digital system in $\mathbb{Z}^d$ with a matrix A of the form $A = gQ$ where Q is an orthonormal matrix and $g \in \mathbb{N}$. Let f be a function which is p multiplicatively complete and let $\mu := \mathcal{H}^s|\mathcal{F}$ be the normalized Hausdorff measure of dimension*

$$s = \frac{\log \#\{\vec{\varepsilon} \in \mathcal{D} \mid f(\vec{\varepsilon}) \neq 0\}}{\log g}$$

restricted to the fundamental domain $\mathcal{F}$. If $\mathcal{B}$ is a μ-continuity set and $a \not\equiv 0 \pmod{p}$, then

$$\lim_{\ell\to\infty} \frac{\#\{\vec{n} \in T_\ell \mid f(\vec{n}) \equiv a \pmod{p},\ \vec{n} \in A^{\ell+1}\mathcal{B}\}}{\#\{\vec{n} \in T_\ell \mid f(\vec{n}) \not\equiv 0 \pmod{p}\}} \to \frac{\mu(\mathcal{B})}{p-1}.$$

Theorem 2.5 is a spatial distribution result. One way to visualize Theorem 2.5 is to take a μ-continuity set $\mathcal{B}$ in $\mathbb{R}^d$, that is $\mu(\partial\mathcal{B}) = 0$. We assign colors to numbers in T_ℓ and represent each number by a colored dot. If the assumptions of Theorem 2.5 are satisfied, the nonzero colors inside $\mathcal{B}$ converge to an equidistribution. The set $A^{-\ell-1}T_\ell \cap \mathcal{B}$ is either empty or the colors occur with the same asymptotic frequency if $\ell \to \infty$. Figure 2 is a graphic representation of this idea.

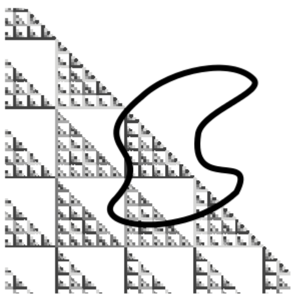

Figure 2. A μ-continuity set on the binomial coefficients.

3. Fractals

Matrix number systems are connected to fractals. We will use the term *fractal* for sets generated by an iterated function system (short IFS). An *iterated function system* is a finite collection $(w_1, w_2, \ldots, w_t)$ of contractions on a complete metric space $(\mathbf{X}, d)$. We will denote the IFS by

$$\mathcal{W} = \{\mathbf{X}; w_1, w_2, \ldots, w_t\}.$$

The main theorem about IFS is from Hutchinson [21].

Theorem 3.1 (Hutchinson 1981). *Let $\mathcal{W} = \{\mathbf{X}; w_1, w_2, \dots, w_t\}$ be an IFS on the complete metric space $(\mathbf{X}, d)$ and let $\mathcal{H}(\mathbf{X})$ denote all nonempty compact subsets of $\mathbf{X}$. The transformation $W : \mathcal{H}(\mathbf{X}) \to \mathcal{H}(\mathbf{X})$ defined by*

$$W(U) := \bigcup_{i=1}^{t} w_i(U) \qquad (\textit{Hutchinson operator})$$

has a unique fixed point $\mathcal{F} \in \mathcal{H}(\mathbf{X})$. It can be constructed as the limit of the iteratively defined sequence $U_N = W(U_{N-1})$ with $U_0 \in \mathcal{H}(\mathbf{X})$ chosen arbitrarily. This fixed point is called the ***attractor of the IFS****.*

We will follow the books of Egdar [14] and Falconer [15] to define measures on the IFS. If $(\mathbf{X}, d)$ is a compact metric space, let $\mathcal{P}(\mathbf{X})$ denote the space of normalized Borel measures on $\mathbf{X}$. To every contraction w_i in $\{\mathbf{X}; w_1, w_2, \dots, w_t\}$ we assign a weight $p_i > 0$ so that $\sum_{i=1}^{t} p_i = 1$.

In our case, for a self-similar IFS which satisfies the open set condition there is a privileged choice of weights. The maps are similarities since the matrix A is a multiple of an orthonormal matrix. The open set condition is fulfilled since the matrix A is an integer matrix and $\mathcal{D}$ is a complete residue system (see [5]).

Let the maps w_i be similarities with ratios r_i. The *similarity dimension* s is the solution of the equation

$$\sum_{i=1}^{t} r_i^s = 1.$$

The weights $p_i = r_i^s$ are called the *uniform weights*.

There is a unique probability measure $\mu \in \mathcal{P}(\mathbf{X})$, so that for all Borel sets $\mathcal{B}$ we have

$$\mu(\mathcal{B}) := \sum_{i=1}^{t} r_i^s \mu(w_i^{-1}(\mathcal{B})). \tag{4}$$

Since

$$\mu(\mathcal{F}) = \sum_{i=1}^{t} r_i^s \mu(w_i^{-1}(\mathcal{F})),$$

the measure μ is an invariant measure on the IFS with weights. It is known as the *uniform measure*. In our examples μ is always a Hausdorff measure. The support of μ is the fractal set $\mathcal{F}$.

Later we will use an IFS we use for the proof of Theorem 2.5. We take the IFS $\mathcal{W} = \{X; w_{1,2}, \dots, w_t\}$ where w_i is an affine map of the form

$$w_i : \vec{x} \longmapsto A^{-1}(\vec{x} + \vec{\varepsilon_i}),$$

with $\vec{\varepsilon_i} \in \mathcal{D} \setminus \{\vec{\varepsilon_i} \mid f(\vec{\varepsilon_i}) \equiv 0 \pmod{p}\}$. The invariant measure μ in this case is the normalized Hausdorff measure of dimension

$$s = \frac{\log \#\{\vec{\varepsilon} \in \mathcal{D} \mid f(\vec{\varepsilon}) \neq 0\}}{\log g}$$

restricted to $\mathcal{F}$. Recall that A is a matrix A of the form $A = gQ$ where Q is an orthonormal matrix and $g \in \mathbb{N}$.

Please note that the colorings are defined on the sequence U_n. It is not possible to define a coloring for the limit fractal.

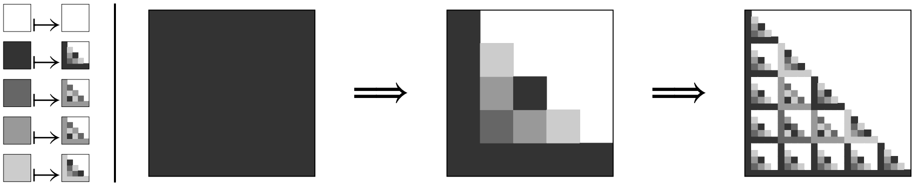

Figure 3. On the left we see the substitutions for $p = 5$, while on the right we see the first two steps of their action

4. Character sums and measures

The definition of p-colorings leads us to the question as to how many numbers have a certain color, i.e. belong to a certain residue class modulo p. To count the numbers in one residue class, we use *Dirichlet characters*.

Definition 4.1. A Dirichlet character modulo a prime p is a function $\chi\colon \mathbb{Z} \to \mathbb{C}$, so that

(i) $\chi(n) = \chi(n + p)$,

(ii) $\chi(n) = 0$ if and only if $\gcd(p, n) = 1$, and

(iii) $\chi(nm) = \chi(n)\chi(m)$.

Characters take roots of unity and 0 as values. The *principal character* $\chi_0(n)$ is given by $\chi_0(n) = 1$ if $\gcd(n, k) = 1$.

Nonprincipal characters satisfy an orthogonality relation

$$\sum_{a \bmod p} \chi(a) = 0. \tag{5}$$

Later we will use a well known identity for characters (see [4])

$$\frac{1}{p-1}\sum_{\chi}\overline{\chi(a)}\chi(b)=\begin{cases}1 & \text{if } a\equiv b \pmod{p},\\ 0 & \text{else,}\end{cases} \tag{6}$$

where the sum runs over all characters. In the proof of Theorem 2.4 we will use the principal character as the indicator function for the non-zero residue classes.

Proof of Theorem 2.4. Each vector $\vec{n} \in T_\ell$ has a unique representation of the form $\vec{n} = (\vec{\varepsilon_0}\vec{\varepsilon_1}\dots\vec{\varepsilon_\ell})$. There are $\#\mathcal{D}$ ways to choose each digit. Therefore, $\#T_\ell = (\#\mathcal{D})^{\ell+1}$.

Equation (2) allows us to write $f(\vec{n})$ as $f(\vec{n}) \equiv f(\vec{\varepsilon_0})f(\vec{\varepsilon_1})\dots f(\vec{\varepsilon_\ell}) \pmod{p}$. Since $f(\vec{\varepsilon})$ is in the finite field $\mathbb{Z}/p\mathbb{Z}$, we have $f(\vec{n}) \not\equiv 0 \pmod{p}$ if and only if $f(\vec{\varepsilon_i}) \not\equiv 0 \pmod{p}$ for $0 \le i \le \ell$. So we count

$$D := \sum_{\vec{\varepsilon}\in\mathcal{D}} \chi_0(f(\vec{\varepsilon})).$$

There is an $\vec{\varepsilon} \in \mathcal{D}$ so that $f(\vec{\varepsilon}) \equiv 0 \pmod{p}$ and therefore $D < \#\mathcal{D}$ and we get

$$\lim_{\ell\to\infty}\frac{\#\{\vec{n}\in T_\ell \mid f(\vec{n})\not\equiv 0 \pmod{p}\}}{\#\{\vec{n}\in T_\ell\}} = \lim_{\ell\to\infty}\frac{D^{\ell+1}}{(\#\mathcal{D})^{\ell+1}} = 0. \qquad \square$$

Remark 4.2. Of course we could state Theorem 2.4 in the same form as Theorem 2.5,

$$\lim_{\ell\to\infty}\frac{\#\{\vec{n}\in T_\ell \mid f(\vec{n})\not\equiv 0 \pmod{p},\ \vec{n}\in A^{\ell+1}\mathcal{B}\}}{\#\{\vec{n}\in T_\ell\}} \longrightarrow 0.$$

Since the set of all colored points is a measure-zero set on the whole fractal (with respect to the Lebesgue measure), it is a measure-zero set for every μ-continuity set $\mathcal{B}$.

Lemma 4.3. *The function f is p multiplicatively complete if and only if for all $\chi \neq \chi_0$ we have*

$$\left|\sum_{\vec{\varepsilon}\in\mathcal{D}}\chi(f(\vec{\varepsilon}))\right| < \sum_{\substack{\vec{\varepsilon}\in\mathcal{D}\\ f(\vec{\varepsilon})\neq 0}} 1.$$

Proof. Since $(\mathbb{Z}/p\mathbb{Z})^* \subseteq \{f(\vec{n}) \mid \vec{n} \in T_k\}$ we can use the orthogonality relation (5) and get for every $\chi \neq \chi_0$

$$\left|\sum_{\vec{n}\in T_k}\chi(f(\vec{n}))\right| \le \#\{\vec{n}\in T_k \mid f(\vec{n})\neq 0\} - (p-1) < \sum_{\substack{\vec{n}\in T_k\\ f(\vec{n})\neq 0}} 1. \tag{7}$$

Now we use the generalized Lucas property and the multiplicativity of characters to write inequality (7) as

$$\left|\prod_{i=0}^{k}\Big(\sum_{\vec{\varepsilon}\in\mathcal{D}}\chi(f(\vec{\varepsilon}))\Big)\right| < \prod_{i=0}^{k}\Big(\sum_{\substack{\vec{n}\in\mathcal{D}\\ f(\vec{\varepsilon})\not\equiv 0}} 1\Big),$$

which proves one direction of Lemma 4.3.

For the other direction we choose k so large that

$$\left|\sum_{\vec{\varepsilon}\in\mathcal{D}}\chi(f(\vec{\varepsilon}))\right|^{k} \le \frac{1}{2p}\Big(\sum_{\vec{\varepsilon}\in\mathcal{D}}\chi_0(f(\vec{\varepsilon}))\Big)^{k}.$$

Now we use identity(6) and the reverse triangle inequality to get

$$\begin{aligned}
&\#\{\vec{n}\in T_k \mid f(\vec{n})\equiv a \pmod p\}\\
&= \frac{1}{p-1}\sum_{\chi}\overline{\chi(a)}\Big(\sum_{\vec{\varepsilon}\in\mathcal{D}}\chi(f(\vec{n}))\Big)^{k}\\
&\ge \frac{1}{p-1}\bigg(\Big(\sum_{\vec{\varepsilon}\in\mathcal{D}}\chi_0(f(\vec{\varepsilon}))\Big)^{k} - \frac{p-2}{2p}\Big(\sum_{\vec{\varepsilon}\in\mathcal{D}}\chi_0(f(\vec{\varepsilon}))^{k}\Big)\bigg)\\
&> 0.
\end{aligned}$$

Therefore $a\in\{f(\vec{n})\mid \vec{n}\in T_k\}$. □

Proof of Theorem 2.5. If we define a function

$$F_\chi(\vec{x}) := \frac{\sum_{\vec{\varepsilon}\in\mathcal{D}}\chi(f(\vec{\varepsilon}))e(\langle\vec{\varepsilon},\vec{x}\rangle)}{\sum_{\vec{\varepsilon}\in\mathcal{D}}\chi_0(f(\vec{\varepsilon}))},$$

Lemma 4.3 tells us that

$$|F_\chi(\vec{0})| < 1 \tag{8}$$

for $\chi\ne\chi_0$. As usual, we have $e(x) := e^{2\pi i x}$.

Now we define a sequence of measures on T_ℓ which, as we will show, converges weakly to a measure μ,

$$\frac{\sum_{\substack{\vec{n}\in T_\ell\\ f(\vec{n})\not\equiv 0}}\delta_{A^{-\ell-1}\vec{n}}}{\sum_{\substack{\vec{n}\in T_\ell\\ f(\vec{n})\not\equiv 0}} 1} \longrightarrow \mu. \tag{9}$$

The measure μ satisfies equation (4). Then we define another sequence of measures on T_ℓ which converges weakly to a measure μ_a, as we will show later,

$$\frac{\sum_{\substack{\vec{n}\in T_\ell\\ f(\vec{n})\equiv a}} \delta_{A^{-\ell-1}\vec{n}}}{\sum_{\substack{\vec{n}\in T_\ell\\ f(\vec{n})\not\equiv 0}} 1} \longrightarrow \mu_a = \frac{1}{p-1}\mu \tag{10}$$

The measure μ_a also satisfies equation (4).

Then we look at the characteristic function $\hat{\mu}_a(\vec{x})$ of μ_a

$$\hat{\mu}_a(\vec{x}) = \lim_{\ell\to\infty} \frac{\sum_{\substack{\vec{n}\in T_\ell\\ f(\vec{n})\equiv a}} e(\langle A^{-\ell-1}\vec{n}, \vec{x}\rangle)}{\sum_{\substack{\vec{n}\in T_\ell\\ f(\vec{n})\not\equiv 0}} 1}.$$

We use Dirichlet characters to rewrite $\hat{\mu}_a(\vec{x})$. In the denominator we use the principal character as indicator function and get

$$\#\{\vec{n}\in T_\ell \mid f(\vec{\varepsilon})\not\equiv 0 \pmod p\} = \sum_{\vec{n}\in T_\ell} \chi_0(f(\vec{n})).$$

In the numerator we can use equation (6) to get

$$\#\{\vec{n}\in T_\ell \mid f(\vec{\varepsilon})\equiv a \pmod p\} = \frac{1}{p-1}\sum_{\chi}\overline{\chi(a)}\sum_{\vec{n}\in T_\ell}\chi(f(\vec{n})).$$

Hence, we have

$$\begin{aligned}
\hat{\mu}_a(\vec{x}) &= \lim_{\ell\to\infty}\frac{1}{p-1}\sum_{\chi}\overline{\chi(a)}\frac{\sum_{\vec{n}\in T_\ell}\chi(f(\vec{n}))e(\langle A^{-\ell-1}\vec{n}, \vec{x}\rangle)}{\sum_{\vec{n}\in T_\ell}\chi_0(f(\vec{n}))}\\
&= \lim_{\ell\to\infty}\frac{1}{p-1}\frac{\sum_{\vec{n}\in T_\ell}\chi_0(f(\vec{n}))e(\langle A^{-\ell-1}\vec{n}, \vec{x}\rangle)}{\sum_{\vec{n}\in T_\ell}\chi_0(f(\vec{n}))} \qquad (11)\\
&\quad + \lim_{\ell\to\infty}\frac{1}{p-1}\frac{\sum_{\chi\neq\chi_0}\overline{\chi(a)}\sum_{\vec{n}\in T_\ell}\chi(f(\vec{n}))e(\langle A^{-\ell-1}\vec{n}, \vec{x}\rangle)}{\sum_{\vec{n}\in T_\ell}\chi_0(f(\vec{n}))}. \qquad (12)
\end{aligned}$$

The first part of the sum (11) comes from the principal character, while the second part (12) contains all other characters. The summand in equation (11) is just the characteristic function $\hat{\mu}(\vec{x})$ of μ,

$$\lim_{\ell\to\infty} \frac{1}{p-1} \frac{\sum_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))e(\langle A^{-\ell-1}\vec{n}, \vec{x}\rangle)}{\sum_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))} = \frac{1}{p-1}\hat{\mu}(\vec{x}).$$

We use the conjugate transpose and write it as

$$\frac{\sum_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))e(\langle A^{-\ell-1}\vec{n}, \vec{x}\rangle)}{\sum_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))} = \prod_{k=1}^{\ell+1} \left(\frac{\sum_{\vec{\varepsilon}\in\mathcal{D}} \chi_0(f(\vec{\varepsilon}))e(\langle \vec{\varepsilon}, (A^T)^{-k}\vec{x}\rangle)}{\sum_{\vec{\varepsilon}\in\mathcal{D}} \chi_0(f(\vec{\varepsilon}))} \right)$$
$$= \prod_{k=1}^{\ell+1} F_{\chi_0}(\vec{\varepsilon}, (A^T)^{-k}\vec{x}).$$

$F_{\chi_0}(\vec{0}) = 1$ and F_{χ_0} is continuously differentiable since it is a composition of continuously differentiable functions. Therefore, it is also Lipschitz continuous and since $A = gQ$ where Q is a orthonormal matrix, we get

$$|F_{\chi_0}((A^T)^{-k}\vec{x}) - 1| \le L\|(A^T)^{-k}\vec{x}\| \le Lg^{-k}\|\vec{x}\|.$$

Hence,

$$\prod_{k=1}^{\ell+1} F_{\chi_0}(\vec{\varepsilon}, (A^T)^{-k}\vec{x})$$

converges uniformly on every compact subset of $\mathbb{R}^d$ since $\|\vec{x}\|$ is bounded on a compact subset. Therefore, we can use the uniform convergence theorem and know that $\hat{\mu}(\vec{x})$ is continuous on every compact subset and especially continuous in 0. We use Lèvy's continuity theorem for characteristic functions (see [11, Theorem 2.6.9]) and prove (9).

Now we focus on the nonprincipal characters (12). We use the same idea as before and write

$$\frac{\sum_{\vec{n}\in T_\ell} \chi(f(\vec{n}))e(\langle A^{-\ell-1}\vec{n}, \vec{x}\rangle)}{\sum_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))} = \prod_{k=1}^{\ell+1} F_{\chi}(\vec{\varepsilon}, (A^T)^{-k}\vec{x}).$$

Because of inequality (8) we know that there exist $\epsilon, \delta > 0$ so that

$$|F_\chi(\vec{x})| \leq 1 - \epsilon \quad \text{for } \|\vec{x}\| \leq \delta. \tag{13}$$

For $R > 0$ there is a $k_0 \in \mathbb{N}_0$, so that for all $k \geq k_0$

$$\|(A^T)^{-k}\vec{x}\| \leq \delta \qquad \text{for } \|\vec{x}\| \leq R. \tag{14}$$

Thus,

$$\left|\prod_{k=1}^{\ell+1} F_\chi((A^T)^{-k}\vec{x})\right| \leq \prod_{k=k_0}^{\ell+1}(1-\epsilon) = (1-\epsilon)^{\ell-k_0+2}$$

and

$$\lim_{\ell\to\infty}(1-\epsilon)^{\ell-k_0+2} \longrightarrow 0.$$

Therefore, the limit in (12) is zero and the weak converge in (10) is proved. □

5. Applications

In this section we give several examples of combinatorially defined number schemes which satisfy the generalized Lucas congruence. Many more such examples (including some infinite classes) can be found in [24].

For each of these examples an analogon to Lucas theorem can be found in the literature. Each time we will give an extended matrix digital system and a coloring function so that the generalized Lucas congruence is satisfied. Therefore, Theorem 2.4 and Theorem 2.5 apply to the binomial coefficients, the Stirling numbers of the first and second kind, the Gaussian q-nomial coefficients as well as to the multinomial coefficients.

In Figure 1 we see the binomial coefficients and Stirling numbers of the first and second kind modulo $p = 5$. All triangles have the same size but only the binomial coefficients have a bilateral symmetry. For the Stirling numbers of the first kind the leftmost numbers are all zero. The Stirling numbers of the second kind have a slanted structure.

With exception of the Apéry numbers it is very easy to check the conditions of Theorem 2.4 and Theorem 2.5 because all examples involve binomial coefficients.

5.1. Apéry numbers. The Apéry numbers $A_1(n)$ and $A_2(n)$ were introduced by Apéry in his 1979 proof that $\zeta(3)$ is irrational. They are defined as

$$A_1(n) := \sum_{k=0}^{n}\binom{n}{k}^2\binom{n+k}{k}^2, \quad A_2(n) := \sum_{k=0}^{n}\binom{n}{k}^2\binom{n+k}{k}.$$

Both of them satisfy the Lucas congruence and for $j \in \{1,2\}$ we have

$$A_j(n) \equiv \prod_{i=0}^{\ell} A_j(n_i) \pmod{p}.$$

The proof that the Apéry numbers $A_1(n)$ satisfy the Lucas congruence has been published in [18], while the proof for the numbers $A_2(n)$ can be found in [12]. Here we have the usual p-adic representation which can be treated as a matrix digital system with the degenerate matrix $A = (p)$ and digit set $\mathcal{D} = \{v \mid 0 \leq v \leq p-1\}$.

The Apéry numbers differ from the other examples as there is no obvious way to check the conditions of Theorem 2.4 and Theorem 2.5.

In order to apply Theorem 2.4, there has to be an $\varepsilon \in \mathcal{D}$ with $f(\varepsilon) = 0$. Often this condition is not satisfied. For instance, there is no $\varepsilon \in \mathcal{D}$ with $f(\varepsilon) = 0$ for both Apéry numbers and $p = 13$. If there is no $\varepsilon \in \mathcal{D}$ with $f(\varepsilon) = 0$, the measure in Theorem 2.5 is the Lebesgue measure.

Based on calculations for the primes $p < 500$, I conjecture that the conditions of Theorem 2.5 are always satisfied for $p > 3$.

5.2. Binomial coefficients. As stated in the introduction, the binomial coefficients satisfy Lucas' original Theorem 1.1. We can treat the linear equation

$$\overrightarrow{\binom{np+r}{kp+s}} = \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix} \overrightarrow{\binom{n}{k}} + \overrightarrow{\binom{r}{s}} \tag{15}$$

as a matrix digital system $(A, \mathcal{D})$ with

$$A = \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix} \quad \text{and} \quad \mathcal{D} = \left\{ \overrightarrow{\binom{v}{w}} \;\middle|\; 0 \leq v, w \leq p-1 \right\}.$$

If we interpret the vectors as binomial coefficients and use induction, Lucas' theorem tells us that they satisfy the generalized Lucas congruence with the coloring function

$$f\left(\overrightarrow{\binom{n}{k}}\right) = \binom{n}{k} \operatorname{Mod} p.$$

The next lemma shows that Theorems 2.4 and 2.5 are satisfied. As usual, we define $a \operatorname{Mod} p := a - \lfloor \frac{a}{p} \rfloor p$.

Lemma 5.1. *The binomial coefficients satisfy the conditions of Theorem 2.4 and Theorem 2.5.*

Proof. The binomial coefficients fulfill Theorem 1.1. There is a corresponding matrix digital system $(A, \mathcal{D})$ with

$$A = \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix}, \quad \mathcal{D} = \left\{ \overrightarrow{\begin{pmatrix} v \\ w \end{pmatrix}} \;\middle|\; 0 \le v, w \le p-1 \right\},$$

and coloring function $f\left(\overrightarrow{\binom{v}{w}}\right) := \binom{v}{w} \operatorname{Mod} p$ where p is a prime.

We have $f\left(\overrightarrow{\binom{v}{w}}\right) = 0$ for the $\frac{p(p-1)}{2}$ digits with $v < w$. Since there are digits $\vec{\varepsilon} \in \mathcal{D}$ so that $f(\vec{\varepsilon}) \equiv 0 \pmod{p}$, according to Theorem 2.4 almost all numbers are 0 modulo p.

For every $0 \le v \le p-1$ we have $f\left(\overrightarrow{\binom{v}{1}}\right) = f\left(\overrightarrow{\binom{v}{v-1}}\right) = v$. Thus, the binomial coefficients modulo p generate the whole group $(\mathbb{Z}/p\mathbb{Z})^*$ and we have equidistribution in the non-zero residue classes modulo p according to Theorem 2.5. $\square$

The binomial coefficients are an archetype for all other examples. We always look for a generalized Lucas congruence, the corresponding matrix digital system and coloring functions that satisfy the congruence. Then we show that the digital function fulfills the hypothesis of Theorem 2.4 and Theorem 2.5.

5.3. Stirling numbers of the first kind. Theorem 5.2 is an analogon to Lucas' theorem for Stirling numbers of the first kind (see [25]). This is one of the more interesting examples and the first one with an extended matrix digital system, so we will look at it in more detail.

Theorem 5.2. *The Stirling numbers of the first kind satisfy the following congruence*

$$\begin{bmatrix} n \\ k \end{bmatrix} \equiv \begin{bmatrix} r \\ t \end{bmatrix} \binom{m}{s} (-1)^{m-s} \pmod{p} \tag{16}$$

with $n = mp + r$ *and* $0 \le r \le p-1$. *The numbers* s *and* t *are defined as* $k - m = s(p-1) + t$ *with* $0 \le t < p-1$ *if* $r = 0$ *and* $0 < t \le p-1$ *if* $r > 0$.

Again we use a linear equation

$$\overrightarrow{\begin{pmatrix} n \\ k \end{pmatrix}} = \begin{pmatrix} p & 0 \\ 1 & p-1 \end{pmatrix} \overrightarrow{\begin{pmatrix} m \\ s \end{pmatrix}} + \overrightarrow{\begin{pmatrix} r \\ t \end{pmatrix}}$$

to define a matrix digital system $(B, \mathcal{D}_1)$ with the matrix $B = \begin{pmatrix} p & 0 \\ 1 & p-1 \end{pmatrix}$ and the digit set

$$\mathcal{D}_1 = \left\{ \overrightarrow{\begin{pmatrix} v \\ w \end{pmatrix}} \,\middle|\, 1 \le v, w \le p-1 \right\} \cup \left\{ \overrightarrow{\begin{pmatrix} 0 \\ w \end{pmatrix}} \,\middle|\, 0 \le w \le p-2 \right\}.$$

After we split off the least significant digit, let us call it $\overrightarrow{\varepsilon_{-1}}$, we are left with an expression of the form $\binom{m}{s}(-1)^{m-s}$. We already know that the binomial coefficients satisfy the generalized Lucas congruence if we write m and s as p-adic numbers. Since $(-1)^p = -1$ for odd p and $-1 \equiv 1 \pmod 2$, we have

$$(-1)^{(m_\ell p^\ell + \cdots + m_1 p + m_0) - (s_\ell p^\ell + \cdots + s_1 p + s_0)} = \prod_{i=0}^{\ell} (-1)^{m_i - s_i}$$

so $(-1)^{m-s}$ also satisfies the generalized Lucas congruence if we write m and s as p-adic numbers. Therefore, we use the matrix digital system $A = \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix}$ and $\mathcal{D} = \{\overrightarrow{\binom{v}{w}} \mid 0 \le v, w \le p-1\}$ from the binomial coefficients again to write $\overrightarrow{\binom{m}{s}}$ as

$$f\left(\overrightarrow{\begin{pmatrix} m \\ s \end{pmatrix}}\right) \equiv \prod_{i=0}^{\ell} f(\overrightarrow{\varepsilon_i}) \pmod p$$

with coloring function

$$f\left(\overrightarrow{\begin{pmatrix} m \\ s \end{pmatrix}}\right) = \begin{pmatrix} m \\ s \end{pmatrix} (-1)^{m-s} \operatorname{Mod} p.$$

What we finally get is an expression of the form

$$\overrightarrow{\varepsilon_{-1}} + B(\overrightarrow{\varepsilon_0} + A\overrightarrow{\varepsilon_1} + A^2\overrightarrow{\varepsilon_2} + \cdots + A^\ell \overrightarrow{\varepsilon_\ell}). \tag{17}$$

Let us now look at the set V of all numbers of the form (17). We call this construction *nested matrix digital system*. Inside the brackets we have a matrix digital system with a similarity A. This matrix digital system satisfies a generalized Lucas congruence and its fundamental domain $\mathcal{F}$ is the unit square. Outside the brackets we have $\#\mathcal{D}_1 = (p-1)p$ affine maps. It is easy to see that the interior of the unit square is mapped to nonoverlapping images.

If we define coloring functions

$$f_1\left(\overrightarrow{\binom{n}{k}}\right) = \left[{n \atop k}\right] \operatorname{Mod} p, \quad f\left(\overrightarrow{\binom{m}{s}}\right) = \binom{m}{s}(-1)^{m-s} \operatorname{Mod} p$$

equation (16) allows us to calculate $f_1\left(\overrightarrow{\binom{n}{k}}\right)$ as

$$f_1\left(\overrightarrow{\binom{n}{k}}\right) \equiv f_1(\overrightarrow{\varepsilon_{-1}}) \prod_{i=0}^{\ell} f(\overrightarrow{\varepsilon_i}) \pmod{p}. \tag{18}$$

Equation (18) and the fact that the images of the affine maps are nonoverlapping allows us to extend the generalized Lucas congruence and the measures μ and μ_a to all numbers in V. The affine maps will deform the measures, but they do not change their equidistribution properties. Therefore, Theorem 2.4 and Theorem 2.5 apply to the Stirling numbers of the first kind.

5.4. Stirling numbers of the second kind. The analogon to Lucas' theorem for Stirling numbers of the second kind can be found in Howard [20].

Theorem 5.3. *If p is prime and $n - (r+1)(p-1) = h$, then*

$$\left\{{n \atop hp}\right\} \equiv \binom{r}{h-1} \bmod p. \tag{19}$$

If $n - (p-1)r - i = h$ and $1 \le m \le i \le p-1$, then

$$\left\{{n \atop hp+m}\right\} \equiv \binom{r}{h}\left\{{i \atop m}\right\} \bmod p. \tag{20}$$

If we take a look at the actual proof by Howard, we see that in all cases not covered by these equations we have $\left\{{n \atop k}\right\} \equiv 0 \bmod p$.

To find a more convenient form for equation (19), we use the recursion formula for Stirling numbers of the second kind

$$\left\{{n \atop hp}\right\} = \left\{{n-1 \atop hp-1}\right\} + \left\{{n-1 \atop hp}\right\} \cdot hp \equiv \left\{{n-1 \atop (h-1)p+(p-1)}\right\} \bmod p$$

and apply equation (20). We can use the linear equation

$$\overrightarrow{\binom{n}{k}} = \begin{pmatrix} p-1 & 1 \\ 0 & p \end{pmatrix} \overrightarrow{\binom{r}{h}} + \overrightarrow{\binom{i}{m}}$$

with matrix $B = \left(\begin{smallmatrix} p-1 & 1 \\ 0 & p \end{smallmatrix}\right)$ and digit set

$$\mathcal{D}_1 = \left\{ \overrightarrow{\begin{pmatrix} v \\ w \end{pmatrix}} \,\middle|\, 1 \le v, w \le p-1 \right\} \cup \left\{ \overrightarrow{\begin{pmatrix} v \\ p \end{pmatrix}} \,\middle|\, 2 \le w \le p \right\}$$

to define a nested matrix digital system. The matrix A and the digit set $\mathcal{D}$ are the same as in the previous example and we use the same argumentation to show that Theorem 2.4 and Theorem 2.5 apply to the Stirling numbers of the second kind.

5.5. Gaussian q-nomial coefficients. Here the analogon to Lucas‘ theorem is given by M. Sved in [28]. The Gaussian q-nomial coefficients are defined as

$$\begin{bmatrix} a \\ b \end{bmatrix}_q = \begin{cases} \dfrac{(1-q^a)(1-q^{a-1})\cdots(1-q^{a-b+1})}{(1-q)(1-q^2)\cdots(1-q^b)} & \text{if } b \le a, \\ 0 & \text{if } b > a, \end{cases}$$

for nonnegative integers a and b.

The Gaussian q-nomial coefficients have a regular, bilateral symmetric structure, which resembles that of the binomial coefficients. This is to be expected since the matrix digital systems are also very similar.

Theorem 5.4. *Let p be a prime, $q > 1$ a positive integer not divisible by p and let $a \ne 1$ be the minimal exponent for which $q^a \equiv 1 \pmod p$; then by Fermat's little theorem it follows that $a|(p-1)$. Furthermore, if $0 \le r, s < a$, then*

$$\begin{bmatrix} na+r \\ ka+s \end{bmatrix}_q \equiv \binom{n}{k} \begin{bmatrix} r \\ s \end{bmatrix}_q \pmod p.$$

The linear equation

$$\overrightarrow{\begin{pmatrix} na+r \\ ka+s \end{pmatrix}} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \overrightarrow{\begin{pmatrix} n \\ k \end{pmatrix}} + \overrightarrow{\begin{pmatrix} r \\ s \end{pmatrix}} \tag{21}$$

gives us the nested matrix digital system with matrix $B = \left(\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right)$ and digit set $\mathcal{D}_1 = \left\{ \overrightarrow{\left(\begin{smallmatrix} v \\ w \end{smallmatrix}\right)} \,\middle|\, 0 \le v, w \le a-1 \right\}$. Now we are left with the binomial coefficients for the infinite part and can use same arguments as for the Stirling numbers.

5.6. Multinomial coefficients. There is a d-dimensional analogon of Lucas' theorem for the multinomial coefficients (see [13], p.273):

Theorem 5.5. *The multinomial coefficients satisfy*

$$\binom{n}{k^{(1)},\dots,k^{(d)}} \equiv \prod_{d=0}^{L} \binom{n_\ell}{k_\ell^{(1)},\dots,k_\ell^{(d)}} \pmod{p}$$

with $n = \sum_{\ell=0}^{L} n_\ell p^\ell$ *and* $k^{(i)} = \sum_{\ell=0}^{L} k_\ell^{(i)} p^\ell$ *where* $0 \le n_\ell, k_\ell^{(i)} \le p-1$.

$$\overrightarrow{\begin{pmatrix} np+r \\ k^{(1)}p+s^{(1)} \\ \vdots \\ k^{(d)}p+s^{(d)} \end{pmatrix}} = \begin{pmatrix} p & 0 & \cdots & 0 \\ 0 & p & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p \end{pmatrix} \overrightarrow{\begin{pmatrix} n \\ k^{(1)} \\ \vdots \\ k^{(d)} \end{pmatrix}} + \overrightarrow{\begin{pmatrix} r \\ s^{(1)} \\ \vdots \\ s^{(d)} \end{pmatrix}}. \tag{22}$$

Again, we have a matrix digital system $(A, \mathcal{D})$ with $A = pI_{d+1}$ where I_d is the d-dimensional identity matrix and $\mathcal{D} = \{(v_1,\dots,v_{d+1})^T \mid 0 \le v_i \le p-1\}$. We can use Lemma 5.1 since the multinomial coefficients contain the binomial coefficients.

6. Concluding Remarks

In this section we studied several examples of combinatorial defined number schemes. For these examples we showed that almost all entries in the number scheme are divisible by a given prime p and the nonzero residue classes are equidistributed modulo p. It turned out that all multidimensional examples of combinatoric functions with the Lucas property are based on the binomial coefficients, for which the matrix A is just the p-fold identity matrix.

Theorems 2.4 and 2.5 allow us to treat more complicated examples. For Theorem 2.4 we only need an extended matrix digital system with a coloring function which satisfies the generalized Lucas congruence. The proof will work with any two integer matrices A and B which are affine expansions.

For the proof of Theorem 2.5 we required the matrix B to be a similarity. If B is a similarity, the measure μ is a uniform measure with identical weights $p_i = \frac{1}{t}$ and we can determine the dimension s and describe the measure μ, which is the Hausdorff measure.

If B is not a similarity map but any affine map given by an expanding integer matrix, the whole proof of Theorem 2.5 and the arguments with Dirichlet characters and Fourier transforms of measures work in exactly the same way. We still get an equidistribution result with respect to a probability measure μ, though μ will not be the Hausdorff measure but a less explicit measure.

References

[1] J.-P. Allouche and J. Shallit, *Automatic sequences.* Theory, applications, generalizations. Cambridge University Press, Cambridge, 2003. MR 1997038 Zbl 1086.11015

[2] J.-P. Allouche, F. von Haeseler, H.-O. Peitgen, A. Petersen, and G. Skordev, Automaticity of double sequences generated by one-dimensional linear cellular automata. *Theoret. Comput. Sci.* **188** (1997), no. 1-2, 195–209. MR 1479629 Zbl 0943.11020

[3] J.-P. Allouche, F. von Haeseler, H.-O. Peitgen, and G. Skordev, Linear cellular automata, finite automata and Pascal's triangle. *Discrete Appl. Math.* **66** (1996), no. 1, 1–22. MR 1387674 Zbl 0854.68065

[4] T. M. Apostol, *Introduction to analytic number theory.* Undergraduate Texts in Mathematics. Springer-Verlag, New York and Heidelberg, 1976. MR 0434929 Zbl 0335.10001

[5] C. Bandt, Self-similar sets V. Integer matrices and fractal tilings of $\mathbb{R}^n$. *Proc. Amer. Math. Soc.* **112** (1991), no. 2, 549–562. MR 1036982 Zbl 0743.58027

[6] G. Barat and P. Grabner, Digital functions and distribution of binomial coefficients. *J. London Math. Soc.* (2) **64** (2001), no. 3, 523–547. MR 1865548 Zbl 1049.11016

[7] D. Barbolosi and P. Grabner, Distribution des coefficients multinomiaux et q-binomiaux modulo p. *Indag. Math.* (*N.S.*) **7** (1996), no. 2, 129–135. MR 1621308 Zbl 0863.11009

[8] L. Carlitz, A note on Stirling numbers of the first kind. *Math. Mag.* **37** (1964), no. 5, 318–321. MR 1571453 Zbl 0124.02504

[9] L. Carlitz, Some partition problems related to the Stirling numbers of the second kind. *Acta Arith.* **10** (1965) 409–422. MR 0172813 Zbl 0151.02601

[10] L. Carlitz, The number of binomial coefficients divisible by a fixed power of a prime. *Rend. Circ. Mat. Palermo* (2) **16** (1967), 299–320. MR 0249308 Zbl 0177.06701

[11] R. Cuppens, *Decomposition of multivariate probabilities.* Probability and Mathematical Statistics, 29. Academic Press, New York and London, 1975. MR 0517412 Zbl 0363.60012

[12] É. Delaygue, Arithmetic properties of Apéry-like numbers. Preprint 2013. arXiv:1310.4131v1 [math.NT]

[13] L. E. Dickson, *History of the theory of numbers.* Vol. I. Divisibility and primality. Carnegie Institute of Washington, Washington, 1919. JFM 47.0100.04

[14] G. Edgar, *Integral, probability, and fractal measures.* Springer-Verlag, New York, 1998. MR 1484412 Zbl 0893.28001

[15] K. Falconer, *The geometry of fractal sets.* Cambridge Tracts in Mathematics, 85. Cambridge University Press, Cambridge, 1986. MR 0867284 Zbl 0587.28004

[16] N. J. Fine, Binomial coefficients modulo a prime. *Amer. Math. Monthly* **54** (1947), 589–592. MR 0023257 Zbl 0030.11102

[17] R. Garfield and H. S. Wilf, The distribution of the binomial coefficients modulo p. *J. Number Theory* **41** (1992), no. 1, 1–5. MR 1161140 Zbl 0765.11008

[18] I. Gessel, Some congruences for Apéry numbers. *J. Number Theory* **14** (1982), no. 3, 362–368. MR 0660381 Zbl 0482.10003

[19] F. T. Howard, Prime divisors of q-binomial coefficients. *Rend. Sem. Mat. Univ. Padova* **48** (1972), 181–188 (1973). MR 0337743 Zbl 0274.10005

[20] F. T. Howard, Congruences for the Stirling numbers and associated Stirling numbers. *Acta Arith.* **55** (1990), no. 1, 29–41. MR 1056113 Zbl 0648.10008

[21] J. Hutchinson, Fractals and self-similarity. *Indiana Univ. Math. J.* **30** (1981), no. 5, 713–747. MR 0625600 Zbl 0598.28011

[22] I. Kátai, Generalized number systems in Euclidean spaces. *Math. Comput. Modelling* **38** (2003), no. 7-9, 883–892. MR 2025175 Zbl 1083.11011

[23] A. T. Lundell, A divisibility property for Stirling numbers. *J. Number Theory* **10** (1978), no. 1, 35–54. MR 0460135 Zbl 0368.10012

[24] R. Meštrović, Lucas' theorem: its generalizations, extensions and applications (1878–2014). Preprint 2014. arXiv:1409.3820v1 [math.NT]

[25] R. Peele, A. J. Radcliffe, and H. Wilf, Congruence problems involving Stirling numbers of the first kind. *Fibonacci Quart.* **31** (1993), no. 1, 73–80. MR 1202346 Zbl 0776.11010

[26] D. Singmaster, Notes on binomial coefficients. III. Any integer divides almost all binomial coefficients. *J. London Math. Soc.* (2) **8** (1974), 555–560. MR 0396285 Zbl 0293.05007

[27] D. Singmaster, Divisibility of binomial and multinomial coefficients by primes and prime powers. In V. E. Hoggatt, Jr, and M. Bicknell-Johnson (eds.), *A collection of manuscripts related to the Fibonacci sequence.* 18$^{\text{th}}$ anniversary volume of the Fibonacci Association. Fibonacci Association, Santa Clara, CA., 1980, 98–113. MR 0624104 Zbl 0517.10013

[28] M. Sved, Divisibility – with visibility. *Math. Intelligencer* **10** (1988), no. 2, 56–64. MR 0932163 Zbl 0644.05002

[29] F. von Haeseler, H.-O. Peitgen, and Skorde, Cellular automata, matrix substitutions and fractals. *Ann. Math. Artificial Intelligence* **8** (1993), no. 3-4, 345–362. Theorem proving and logic programming (1992). MR 1280046 Zbl 0866.68069

[30] F. von Haeseler, H.-O. Peitgen, and G. Skordev, Pascal's triangle, dynamical systems and attractors. *Ergodic Theory Dynam. Systems* **12** (1992), no. 3, 479–486. MR 1182659 Zbl 0784.58032

[31] S. Wolfram, Geometry of binomial coefficients. *Amer. Math. Monthly* **91** (1984), no. 9, 566–571. MR 0764797 Zbl 0553.05004

Received August 11, 2015; revised February 25, 2016

Florian Greinecker, Department of Analysis and Computational Number Theory,
TU Graz, Steyrergasse 30, A-8010 Graz, Austria

e-mail: greinecker@math.tugraz.at