ACTA ARITHMETICA
186.4 (2018)

Calculating the density of solutions of equations related to the Pólya–Ostrowski group through Markov chains

by

Dario Spirito (Roma)

1. Introduction. Let $q \geq 2$ be a positive integer. Following [1] and [2], we define the following three functions on $\mathbb{N}$:

$$v_q(n) := \begin{cases} 0 & \text{if } n = 0, \\ \max\{k \in \mathbb{N} : q^k \text{ divides } n\} & \text{if } n > 0, \end{cases}$$

$$w_q(n) := \sum_{i=0}^{n} v_q(i) = \sum_{k\geq 1} \left\lfloor \frac{n}{q^k} \right\rfloor,$$

$$u_q(n) := \sum_{i=0}^{n} w_q(i).$$

The functions $w_q(n)$ and $u_q(n)$ arise naturally during the study of the *Pólya–Ostrowski group* $\mathrm{Po}(D)$ of a Dedekind domain D, a subgroup of the class group of D closely related to the ring $\mathrm{Int}(D)$ of integer-valued polynomials over D. In particular, $\mathrm{Po}(D)$ is linked with the module structure of two sequences of D-modules, namely the sequence of *characteristic ideals* $\mathfrak{J}_n(D)$ of D and the sequence of the subsets $\mathrm{Int}_n(D)$ of $\mathrm{Int}(D)$ formed by the polynomials $f \in \mathrm{Int}(D)$ of degree at most n. More precisely, $\mathfrak{J}_n(D)$ and $\mathrm{Int}_n(D)$ are free if and only if, respectively, $w_q(n) \equiv 0 \bmod d$ and $u_q(n) \equiv 0 \bmod d$ for every q, where d is the order of the ideal Π_q (defined as the product of the maximal ideals $\mathfrak{m}$ of D such that $|D/\mathfrak{m}| = q$) in the class group of D. See Section 5 for a more detailed explanation.

These ideas led Elliott [2] to study the equations $w_q(n) \equiv x \bmod d$ and $u_q(n) \equiv x \bmod d$, where d and x are integers; in particular, he was interested

2010 *Mathematics Subject Classification*: Primary 11B05; Secondary 11A07, 11A63, 13F05, 13F20.

Key words and phrases: distribution modulo d, natural density, integer-valued polynomials, Pólya–Ostrowski group, Markov chain, q-additive functions.

Received 5 June 2017; revised 14 December 2017.

Published online 9 November 2018.

DOI: 10.4064/aa170605-6-3

in the *density* of the set of solutions, where the (*natural*) *density* of a subset $T \subseteq \mathbb{N}$ is the limit

$$\delta(T) := \lim_{N\to\infty} \frac{|\{n \in T : 0 \le n < N\}|}{N}$$

(provided that it exists). Through a mixture of special cases and experimental evidence, he conjectured [2, Conjecture 2.2] that the density of the solutions always exists and is rational, and that if $x = 0$ then it is at least $1/d$.

In this paper, we consider the more general equation

$$\theta_u u_q(n) + \theta_w w_q(n) + \theta_2 \frac{n(n+1)}{2} + \theta_1 n + \theta_0 \equiv 0 \bmod d, \tag{1}$$

where $\theta_u, \theta_w, \theta_2, \theta_1, \theta_0$ are integer coefficients; this form appears when trying to express the function $u_q(aq + \lambda)$ in terms of $u_q(a)$.

Our starting point is the possibility of expressing the number of solutions of (1) in $[0, qN)$ in terms of the number of solutions in $[0, N)$ of equations of the same form, but with different coefficients (Proposition 2.3). While we are not able to prove that the density $\delta(q, d; \theta_u, \theta_w, \theta_2, \theta_1, \theta_0)$ of the solutions of (1) exists for every choice of q, d and the coefficients, we can use this recurrence relation to associate to the equation (1) (for any fixed q and d) a stochastic matrix $P := P(q, d)$ and a Markov chain, studying which we can calculate these densities in several cases.

More precisely, fix q and d. Suppose that $\delta(q, d; \theta_u, \theta_w, \theta_2, \theta_1, \theta_0)$ exists for every choice of $\theta_u, \theta_w, \theta_2, \theta_1, \theta_0$. We prove that:

- (Theorem 2.6) every $\delta(q, d; \theta_u, \theta_w, \theta_2, \theta_1, \theta_0)$ is rational;
- (Theorem 3.2)

$$\delta(q, d; 0, \psi, 0, \theta, x) = \begin{cases} \frac{1}{d} \gcd(\psi, \theta, d) & \text{if } \gcd(\psi, \theta, d) \mid \gcd(x, d), \\ 0 & \text{otherwise;} \end{cases}$$

- (Theorem 4.2) if $d \mid q$, then (φ is the Euler function)

$$\delta(q, d; 1, 0, 0, 0, x) = \frac{1}{d^2} \sum_{f \mid \gcd(x,d)} f \cdot \varphi\left(\frac{d}{f}\right);$$

- (Theorem 4.4) if d and q are coprime and θ_u is coprime to d, then

$$\delta(q, d; \theta_u, \theta_w, 0, 0, x) = 1/d.$$

Section 5 translates the result obtained back to the setting of integer-valued polynomials.

2. The general transformation. Our first step is expressing $w_q(aq+\lambda)$ and $u_q(aq + \lambda)$ as functions of a and λ.

Lemma 2.1. *Let q be a positive integer, and let $\lambda \in \{0, \ldots, q-1\}$. Then*

$$w_q(aq+\lambda) = w_q(a) + a \quad \textit{for every } a \in \mathbb{N}.$$

Proof. Since $v_q(i) = 0$ if i is not a multiple of q, we have

$$w_q(aq+\lambda) = \sum_{i=0}^{aq+\lambda} v_q(i) = \sum_{j=0}^{a} v_q(jq).$$

Moreover, $v_q(jq) = 1 + v_q(j)$, and thus

$$w_q(aq+\lambda) = \sum_{j=0}^{a}(1+v_q(j)) = a + \sum_{j=0}^{a} v_q(j) = a + w_q(a),$$

as claimed. ■

Lemma 2.2. *Let q be a positive integer, and let $\lambda \in \{0, \ldots, q-1\}$. Then*

$$\begin{aligned} u_q(aq+\lambda) &= qu_q(a) + (\lambda+1-q)w_q(a) + \frac{q}{2}a^2 + \left(\lambda + 1 - \frac{q}{2}\right)a \\ &= qu_q(a) + (\lambda+1-q)w_q(a) + q\frac{a(a+1)}{2} + (\lambda+1-q)a \end{aligned}$$

for every $a \in \mathbb{N}$.

Proof. We start by calculating $u_q(aq-1)$. We have

$$u_q(aq-1) = \sum_{i=0}^{aq-1} w_q(i) = \sum_{b=0}^{a-1}\sum_{t=0}^{q-1} w_q(bq+t).$$

Since $w_q(bq+t) = w_q(bq)$ for every $t \in \{0, \ldots, q-1\}$, this implies that

$$u_q(aq-1) = \sum_{b=0}^{a-1} qw_q(bq) = q\sum_{b=0}^{a-1}(w_q(b)+b) = qu_q(a-1) + q\frac{(a-1)a}{2}.$$

Hence, using again $w_q(bq+t) = w_q(bq)$, we have

$$\begin{aligned} u_q(aq+\lambda) &= u_q(aq-1) + (\lambda+1)w_q(aq) \\ &= qu_q(a-1) + q\frac{(a-1)a}{2} + (\lambda+1)(w_q(a)+a) \\ &= q(u_q(a) - w_q(a)) + \frac{(a-1)aq}{2} + (\lambda+1)w_q(a) + (\lambda+1)a, \end{aligned}$$

rearranging which we obtain our claim. ■

The previous lemmas suggest considering the more general equation

$$\theta_u u_q(n) + \theta_w w_q(n) + \theta_2\frac{n(n+1)}{2} + \theta_1 n + \theta_0 \equiv 0 \bmod d, \tag{2}$$

where $\theta_u, \theta_w, \theta_2, \theta_1, \theta_0$ vary in $\mathbb{Z}$. Clearly, if $\theta'_u \equiv \theta_u \bmod d$, and analogously for $\theta'_w, \theta'_2, \theta'_1$ and θ'_0, then n is a solution of (2) if and only if it is a solution of

$$\theta'_u u_q(n) + \theta'_w w_q(n) + \theta'_2 \frac{n(n+1)}{2} + \theta'_1 n + \theta'_0 \equiv 0 \bmod d;$$

for this reason, we will sometimes consider equation (2) as having the coefficients $\theta_u, \theta_w, \theta_2, \theta_1, \theta_0$ in $\mathbb{Z}/d\mathbb{Z}$; this should not cause confusion.

Note that if we were using n^2 instead of $n(n+1)/2$, we may need to consider also half-integer values of θ_2 and θ_1, and the situation may become troublesome when d is even.

Let now $\mathbf{s} := (\theta_u, \theta_w, \theta_2, \theta_1, \theta_0) \in \mathbb{Z}^5$. For any $A \in \mathbb{N}$, we denote by $\gamma(A, q, d; \theta_u, \theta_w, \theta_2, \theta_1, \theta_0)$, or by $\gamma(A, q, d; \mathbf{s})$, the number of natural numbers $n < A$ that satisfy (2).

Proposition 2.3. *Let $q, d \geq 2$ be integers. For every $A \in \mathbb{N}$ and every $\mathbf{s} \in \mathbb{Z}^5$, we have*

$$\gamma(qA, q, d; \mathbf{s}) = \sum_{\lambda=0}^{q-1} \gamma(A, q, d; \mathbf{s}M_\lambda) \tag{3}$$

where

$$M_\lambda := \begin{pmatrix} q & \lambda - q + 1 & q & \lambda - q + 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & q^2 & \lambda q - q(q-1)/2 & \lambda(\lambda+1)/2 \\ 0 & 0 & 0 & q & \lambda \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

for every λ.

Proof. Let $\mathbf{s} := (\theta_u, \theta_w, \theta_2, \theta_1, \theta_0)$. Moreover, for any $\lambda \in \{0, \ldots, q-1\}$, let $\gamma_{(\lambda)}(A, q, d; \mathbf{s})$ be the number of solutions to (2) that are smaller than A and congruent to λ modulo q.

Each row of M_λ is the expansion of $u_q(n), w_q(n), n(n+1)/2, n, 1$ in terms of $u_q(a), w_q(a), a(a+1)/2, a, 1$ when $n = aq + \lambda$: indeed, for u_q and w_q this follows from Lemmas 2.1 and 2.2, and it is obvious for n and 1. Moreover,

$$\frac{(aq+\lambda)(aq+\lambda+1)}{2} = \frac{q^2}{2}a^2 + \frac{q^2}{2}a - \frac{q^2}{2}a + \left(q\lambda + \frac{q}{2}\right)a + \frac{\lambda(\lambda+1)}{2},$$

which gives the third row of M_λ after rearrangement.

Now $aq + \lambda < Aq$ if and only if $a < A$; this means exactly that

$$\gamma_{(\lambda)}(qA, q, d; \mathbf{s}) = \gamma(A, q, d; \mathbf{s}M_\lambda).$$

Summing over λ we get the claim. ∎

As remarked before the statement of the proposition, we can consider $\mathbf{s}$ to be an element of $(\mathbb{Z}/d\mathbb{Z})^5$ instead of $\mathbb{Z}^5$; in particular, we can define $\gamma(A,q,d;\mathbf{s})$ even with $\mathbf{s}\in(\mathbb{Z}/d\mathbb{Z})^5$, and Proposition 2.3 carries over without problems, with the only difference that each M_λ must be considered as a matrix over $\mathbb{Z}/d\mathbb{Z}$.

This convention is useful because it makes the space of possible $\mathbf{s}$ finite, and in particular it allows us to rearrange equation (3) in matrix form. Indeed, for every $\mathbf{s},\mathbf{t}\in(\mathbb{Z}/d\mathbb{Z})^5$, let $\mu(\mathbf{s},\mathbf{t})$ be the number of $\lambda\in\{0,\dots,q-1\}$ such that $\mathbf{t}=\mathbf{s}M_\lambda$. Then, for every $\mathbf{s}\in(\mathbb{Z}/d\mathbb{Z})^5$, we have the finite sum

$$\gamma(qA,q,d;\mathbf{s})=\sum_{\mathbf{t}\in(\mathbb{Z}/d\mathbb{Z})^5}\mu(\mathbf{s},\mathbf{t})\gamma(A,q,d;\mathbf{t}).$$

Let now

$$\widetilde{\gamma}(A,q,d;\mathbf{s}):=\frac{\gamma(A,q,d;\mathbf{s})}{A},$$

and let $\widetilde{\boldsymbol{\gamma}}(A,q,d)$ be the column vector composed of the $\widetilde{\gamma}(A,q,d;\mathbf{s})$ as $\mathbf{s}$ ranges in $(\mathbb{Z}/d\mathbb{Z})^5$. Then the previous equality can be written as

$$\widetilde{\boldsymbol{\gamma}}(qA,q,d)=P(q,d)\widetilde{\boldsymbol{\gamma}}(A,q,d) \tag{4}$$

where $P(q,d):=(\mu(\mathbf{s},\mathbf{t})/q)_{\mathbf{s},\mathbf{t}}$ is a (rational) matrix of order d^5. It is a *stochastic matrix*, i.e., each entry is nonnegative and the sum of each row is 1: indeed, the sum of $\mu(\mathbf{s},\mathbf{t})$, as $\mathbf{s}$ is fixed and $\mathbf{t}$ varies, must be q, since for each λ there is a $\mathbf{t}$ such that $\mathbf{t}=\mathbf{s}M_\lambda$.

We introduce the following definition.

Definition 2.4. Let $q,d\geq 2$ be integers. The *density of solutions* for $\mathbf{s}\in\mathbb{Z}^5$ (or $\mathbf{s}\in(\mathbb{Z}/d\mathbb{Z})^5$) with respect to q and d is

$$\delta(q,d;\mathbf{s}):=\lim_{N\to\infty}\widetilde{\gamma}(N,q,d;\mathbf{s})=\lim_{N\to\infty}\frac{\gamma(N,q,d;\mathbf{s})}{N},$$

provided that the limit exists; if q and d are clear from the context, we also write $\delta(\mathbf{s})$ for $\delta(q,d;\mathbf{s})$. The column vector $(\delta(\mathbf{s}))_{\mathbf{s}}$ is called the *vector of densities* of the solutions of (2) and is denoted by $\boldsymbol{\delta}(q,d)$ (or simply $\boldsymbol{\delta}$).

Fix now $\theta_u,\theta_w,\theta_2,\theta_1$. If the density $\delta(q,d;\theta_u,\theta_w,\theta_2,\theta_1,x)$ exists for every $x\in\mathbb{Z}$, we say that the function

$$f:n\mapsto\theta_u u_q(n)+\theta_w w_q(n)+\theta_2\frac{n(n+1)}{2}+\theta_1 n$$

has a limit distribution modulo d (and we call the assignment

$$x\mapsto\delta(q,d;\theta_u,\theta_w,\theta_2,\theta_1,x)$$

the *limit distribution*). If the densities are all equal (and so are equal to $1/d$), we say that f is *uniformly distributed modulo* d (see e.g. [9]).

Lemma 2.5. *Let M be a square matrix of order n over $\mathbb{C}$; suppose that each eigenvalue λ of M satisfies $|\lambda| = 1$, $\lambda \neq 1$. If $\mathbf{v}$ is a vector such that $M^k\mathbf{v}$ has a limit when $k \to \infty$, then $\mathbf{v} = 0$.*

Proof. By conjugation, we can suppose that $M = (m_{ij})_{i,j}$ is an upper triangular matrix.

Suppose $\mathbf{v} = (v_1, \ldots, v_n) \neq 0$, and let t be the largest i such that $v_i \neq 0$. Then the tth component of $M^k\mathbf{v}$ is equal to $m_{tt}^k v_t$; in particular, $m_{tt}^k v_t$ has a limit as $k \to \infty$. However, m_{tt} is an eigenvalue of M, and thus it is a complex number of norm 1 different from 1; hence, m_{tt}^k does not have a limit. This would imply $v_t = 0$, contrary to our choice of t. The claim is proved. ■

Theorem 2.6. *Let $q, d \geq 2$ be integers. If the vector of densities $\boldsymbol{\delta}(q,d)$ exists, then it is a right eigenvector of $P(q,d)$ with eigenvalue* 1*, and all its entries are rational.*

Proof. For every $k \in \mathbb{N}$, let $\widetilde{\boldsymbol{\gamma}}_k$ be the column vector whose entries are $\widetilde{\gamma}(q^k, q, d; \mathbf{s})$ as $\mathbf{s}$ ranges in $(\mathbb{Z}/d\mathbb{Z})^5$. Then (4) becomes

$$\widetilde{\boldsymbol{\gamma}}_k = P\widetilde{\boldsymbol{\gamma}}_{k-1} = P^k\widetilde{\boldsymbol{\gamma}}_0.$$

Clearly, if $\widetilde{\boldsymbol{\gamma}}_k$ has a limit $k \to \infty$ then it must be $\boldsymbol{\delta} := \boldsymbol{\delta}(q,d)$; in particular, the first equality of the previous equation becomes

$$\boldsymbol{\delta} = P\boldsymbol{\delta},$$

and thus $\boldsymbol{\delta}$ is an eigenvector of P with eigenvalue 1.

Moreover, the existence of $\boldsymbol{\delta}$ implies that $P^k\widetilde{\boldsymbol{\gamma}}_0$ has a limit as $k \to \infty$. Since $P := P(q,d)$ is a stochastic matrix with rational entries, the algebraic and geometric multiplicities of its eigenvalue 1 coincide (see e.g. [4, Section 9.4, Fact 1(b)] or [8, p. 696]), and we can find a rational matrix A such that $A^{-1}PA$ is a block matrix

$$N := A^{-1}PA = \begin{pmatrix} I & 0 & 0 \\ 0 & R & 0 \\ 0 & 0 & Q \end{pmatrix},$$

where I is the identity matrix, the eigenvalues of R have norm 1 but are different from 1, and the norm of each eigenvalue of Q is smaller than 1. The limit $P^k\widetilde{\boldsymbol{\gamma}}_0 \to \boldsymbol{\delta}$ can we rewritten as $N^k(A^{-1}\widetilde{\boldsymbol{\gamma}}_0) \to A^{-1}\boldsymbol{\delta}$. Let $\mathbf{v} := A^{-1}\widetilde{\boldsymbol{\gamma}}_0$; then

$$N^k\mathbf{v} = \begin{pmatrix} I & 0 & 0 \\ 0 & R^k & 0 \\ 0 & 0 & Q^k \end{pmatrix}\begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1 \\ R^k\mathbf{v}_2 \\ Q^k\mathbf{v}_3 \end{pmatrix},$$

where $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are subvectors of $\mathbf{v}$ of appropriate length. The existence of the limit implies that both $R^k\mathbf{v}_2$ and $Q^k\mathbf{v}_3$ have limit as $k \to \infty$. Since R

satisfies the hypothesis of Lemma 2.5, we have $\mathbf{v}_2 = 0$; on the other hand, by construction, $Q^k \to 0$ [8, p. 617], and thus $Q^k\mathbf{v}_3 \to 0$.

Both A and $\widetilde{\boldsymbol{\gamma}}_0$ have rational entries (since $\gamma(1, q, d; \mathbf{s})$ is either 1 or 0). Hence, $\mathbf{v}$ has rational entries, and by the previous reasoning so does the limit of $N^k\mathbf{v}$, i.e., $A^{-1}\boldsymbol{\delta}$; therefore, $\boldsymbol{\delta}$ has rational entries, as claimed. ∎

The fact that the entries are rational supports part of [2, Conjecture 2.2(1)].

As observed before Definition 2.4, $P := P(q, d)$ is a stochastic matrix; hence, we can interpret it in a probabilistic way. A (discrete) *Markov chain* $\mathcal{M}$ is a family of random variables $\{X_n\}_{n\in\mathbb{N}}$, whose range is a finite set $S := \{s_1, \ldots, s_k\}$ (called the *state space* of $\mathcal{M}$), such that

$$P(X_{n+1} = t_{n+1} \mid X_n = t_n, X_{n-1} = t_{n-1}, \ldots, X_0 = t_0) \\ = P(X_{n+1} = t_{n+1} \mid X_n = t_n)$$

for all n and all $t_1, \ldots, t_{n+1} \in S$.

If, furthermore, the probability $P(X_{n+1} = t_{n+1} \mid X_n = t_n)$ of going from t_n to t_{n+1} does not depend on n, then $\mathcal{M}$ is said to be *time-homogeneous*, and the matrix $M := (m_{ij})_{i,j}$, where $m_{ij} := P(X_{n+1} = s_j \mid X_n = s_i)$, is a stochastic matrix, called the *transition matrix* of $\mathcal{M}$.

Conversely, to every stochastic matrix $M = (m_{ij})_{i,j}$ of order k is associated a discrete, time-homogeneous Markov chain $\mathcal{M}$ on a state space $S = \{s_1, \ldots, s_k\}$ of cardinality k, such that $P(X_{n+1} = s_j \mid X_n = s_i) = m_{ij}$ for all n, that is, M is the transition matrix of $\mathcal{M}$. When M and $\mathcal{M}$ are linked in this way, we call $\mathcal{M}$ the Markov chain *represented* by M. See e.g. [8, Section 8.4] for further details.

When $M = P(q, d)$, we denote the Markov chain arising in this way by $\mathcal{C}(q, d)$; more explicitly, $\mathcal{C}(q, d)$ is the Markov chain such that the probability of going from state $\mathbf{s}$ to state $\mathbf{t}$ is $\mu(\mathbf{s}, \mathbf{t})/q$.

Let $\mathcal{M}$ be a Markov chain with state space X, and let $i, j \in X$. We say that j is *reachable* from i (or that i *leads to* j) if there is a $k \in \mathbb{N}$ such that $P(X_{n+k} = j \mid X_n = i) > 0$, that is, if the probability of going from i to j in k steps is positive. We say that i and j are *communicating* (and we write $i \leftrightarrow j$) if i is reachable from j and j is reachable from i. The relation "$\leftrightarrow$" is an equivalence relation; if C is an equivalence class, we say that C is *ergodic* if, when $i \in C$ and j is reachable from i, then also $j \in C$, that is, if once the chain arrives in C then it cannot leave C; equivalently, $P(X_{n+k} = l \mid X_n = i) = 0$ for all $l \notin C$ and all $k \in \mathbb{N}$. A state is *ergodic* if the equivalence class it belongs to is an ergodic class.

PROPOSITION 2.7. *Let P be a stochastic matrix, and let $\mathbf{v} := (v_1, \ldots, v_n)^\mathsf{T}$ be such that $\mathbf{v} = P\mathbf{v}$. If i and j belong to the same ergodic class, then $v_i = v_j$.*

Proof. This is essentially a consequence of the Perron–Frobenius theorem (see e.g. [4, Section 9.2, Fact 5] or [8, Chapter 8]). Let $C_1, \ldots, C_t$ be the ergodic classes of the Markov chain associated to P. By [4, Section 9.4, Fact 1(g)], the space V of right eigenvectors of P with eigenvalue 1 has dimension t, and there is a basis $\underline{\mathbf{u}} := \{\mathbf{u}^{C_1}, \dots, \mathbf{u}^{C_t}\}$ of V such that if a is an ergodic state and C_k is its equivalence class, then the ath component of $\mathbf{u}^{C_l}$ is 1 if $l = k$ and 0 otherwise. In particular, if both $i, j \in C_k$ and $\mathbf{v} \in V$, then v_i and v_j are equal to the coefficients of $\mathbf{u}^{C_k}$ along the basis $\underline{\mathbf{u}}$ of V. In particular, $v_i = v_j$. ∎

3. The case $\theta_u = \theta_2 = 0$. A function $f : \mathbb{N} \to \mathbb{R}$ is said to be *q-additive* if $f(0) = 0$ and

$$f(n) = \sum_{j\geq 0} f(a_{q,j}(n)q^j) \quad \text{for} \quad n = \sum_{j\geq 0} a_{q,j}(n)q^j,$$

where $a_{q,j}(n) \in \{0, \dots, q-1\}$ are the digits of n in base q. The prototype of q-additive functions is the sum of digits of n in base q, which we denote by $s_q(n)$.

By [1, Exercise II.8 and Lemma II.4], we can write

$$w_q(n) := \frac{n - s_q(n)}{q-1};$$

thus, for every $\psi, \theta \in \mathbb{Z}$, the function

$$\psi w_q(n) + \theta n = \frac{\theta(q-1) + \psi}{q-1} \cdot n - \frac{\psi}{q-1} \cdot s_q(n)$$

is q-additive. On the other hand, $u_q(n)$ is not q-additive, since (for example) $u_q(q+1) = 2 \neq 1 = u_q(q) + u_q(1)$. This point of view yields the following, the second part of which is another partial answer to [2, Corollary 2.2(1)].

Theorem 3.1. *Let $d, q \geq 2$ be positive integers. Then:*

(a) *for every $\theta_w, \theta_1, \theta_0 \in \mathbb{Z}$, the density $\delta(q, d; 0, \theta_w, 0, \theta_1, \theta_0)$ exists;*
(b) *if $d \mid q^n$ for some n, then the vector of densities $\boldsymbol{\delta}(q, d)$ exists.*

Proof. (a) By [11, Theorem 1.1], every q-additive function with integer values has a limit distribution modulo d, for every d; in particular, so does $\theta_w w_q(n) + \theta_1 n$, and the density $\delta(q, d; 0, \theta_w, 0, \theta_1, \theta_0)$ exists.

(b) Let $\lambda_1, \dots, \lambda_n \in \{0, \dots, q-1\}$, and let $\mathbf{s} \in (\mathbb{Z}/d\mathbb{Z})^5$. Applying Proposition 2.3 repeatedly, we obtain

$$\frac{\gamma(q^n N, q, d; \mathbf{s})}{q^n N} = \sum_{\lambda_1,\dots,\lambda_n=0}^{q-1} \frac{1}{q^n} \frac{\gamma(N, q, d; \mathbf{s}M_{\lambda_1} \cdots M_{\lambda_n})}{N}.$$

The first and third columns of $M_{\lambda_1} \cdots M_{\lambda_n}$ (as integer matrices) are divisible by q^n; hence, they are 0 when reduced modulo d. It follows that the first and third components of $\mathbf{s}M_{\lambda_1} \cdots M_{\lambda_n}$ are always 0.

Therefore, by the previous point, for each summand of the right hand side the limit (as $N \to \infty$) exists; it follows that so does the limit of the left hand side, i.e., the density $\delta(q,d;\mathbf{s})$ exists. ■

The first part of the previous theorem is especially useful since, if we are interested in the distribution of the function w_q, we do not need to consider θ_u or θ_2; that is, we can study the Markov chain limited to the subset of $(\mathbb{Z}/d\mathbb{Z})^5$ where $\theta_u = \theta_2 = 0$. The next result calculates these densities.

THEOREM 3.2. *Let $q, d \geq 2$ be integers, and let $\psi \in \mathbb{Z}$, $\psi \neq 0$. Then*

$$\delta(q,d;0,\psi,0,\theta,x) = \begin{cases} \frac{1}{d}\gcd(\psi,\theta,d) & \textit{if } \gcd(\psi,\theta,d) \mid \gcd(x,d), \\ 0 & \textit{otherwise}, \end{cases}$$

for every $x \in \mathbb{Z}$.

Proof. By Theorem 3.1(a) the density exists.

We first note that if $\gcd(\psi,\theta,d)$ does not divide $\gcd(x,d)$, then the equation $\psi w_q(n) + \theta n + x \equiv 0 \bmod d$ cannot have solutions, so the density is 0.

Suppose that $g := \gcd(\psi,\theta,d)$ divides $\gcd(x,d)$. In this case, $\psi w_q(n) + \theta n + x \equiv 0 \bmod d$ is equivalent to $\frac{\psi}{g} w_q(n) + \frac{\theta}{g} n + \frac{x}{g} \equiv 0 \bmod \frac{d}{g}$; therefore,

$$\gamma(A,q,d;0,\psi,0,\theta,x) = \gamma(A,q,d/g;0,\psi/g,0,\theta/g,x/g).$$

Moreover, $\gcd(\psi/g,\theta/g,d/g) = 1$; hence, it is enough to prove the claim when $\gcd(\psi,\theta,d) = 1$.

The set $X := \{(0,\psi,0,\theta,x) \mid \psi,\theta,x \in \mathbb{Z}/d\mathbb{Z}\}$ is invariant under right multiplication by M_λ; hence, the Markov chain $\mathcal{C}(q,d)$ restricts to a chain $\mathcal{C}'(q,d)$ on X, which can also be defined as the Markov chain with transition matrix $Q := (\mu'(\mathbf{s},\mathbf{t})/q)$, where $\mu'(\mathbf{s},\mathbf{t})$ is the number of λ's such that $\mathbf{s}N_\lambda = \mathbf{t}$ and

$$N_\lambda := \begin{pmatrix} 1 & 1 & 0 \\ 0 & q & \lambda \\ 0 & 0 & 1 \end{pmatrix}$$

is just the submatrix of M_λ relative to θ_w, θ_1 and θ_0.

Let $k \geq 1$. Consider the stochastic matrix Q^k; its $\mathbf{s},\mathbf{t}$ entry is $\mu_k(\mathbf{s},\mathbf{t})/q^k$, where $\mu_k(\mathbf{s},\mathbf{t})$ is the number of k-tuples $(\lambda_1,\ldots,\lambda_k)$ such that $\mathbf{s}M_{\lambda_1}\cdots M_{\lambda_k} = \mathbf{t}$. Consider the Markov chain $\mathcal{C}'_k(q,d)$ represented by Q^k; the probability of going from $\mathbf{s}$ to $\mathbf{t}$ is $\mu_k(\mathbf{s},\mathbf{t})/q^k$, which is equal to the probability of going from $\mathbf{s}$ to $\mathbf{t}$ in k steps in $\mathcal{C}'(q,d)$.

We claim that, for $k \geq 1$,

$$(\psi\ \theta\ x)\, N_{\lambda_1}\cdots N_{\lambda_k} = \begin{pmatrix} \psi \\ q^k\theta + q^{k-1}\psi + q^{k-2}\psi + \cdots + q\psi + \psi \\ x + P_k(\lambda_1,\ldots,\lambda_k,\psi,\theta,q) \end{pmatrix}^{\mathsf{T}},$$

where

$$\begin{aligned}&P_k(\lambda_1,\dots,\lambda_k,\psi,\theta,q)\\&\qquad:=\lambda_1\theta+\lambda_2(q\theta+\psi)+\dots+\lambda_k(q^{k-1}\theta+q^{k-2}\psi+\dots+\psi).\end{aligned}$$

Indeed, this is clear for $k=1$ and follows easily by induction for arbitrary k.

Set $a_k(\theta):=q^k\theta+q^{k-1}\psi+q^{k-2}\psi+\dots+q\psi+\psi$, and suppose there is a $k>1$ such that $a_k(\theta)\equiv\theta \bmod d$. Then, for any ψ and θ, the subset $X(\psi,\theta):=\{(\psi,\theta,x):x\in\mathbb{Z}/d\mathbb{Z}\}$ of X is invariant by $\mathcal{C}'_k(q,d)$, and thus we can restrict the Markov chain to $X(\psi,\theta)$; we claim that this chain is irreducible, i.e., that any state can be reached from any other one.

The probability of going from (ψ,θ,x) to (ψ,θ,y) is nonzero if and only if there are $\lambda_1,\dots,\lambda_k$ such that $y-x=P_k(\lambda_1,\dots,\lambda_k,q,\psi,\theta,q)$; hence, x and y are communicating if and only if they belong to the same coset of the subgroup G of $\mathbb{Z}/d\mathbb{Z}$ generated by $P_k(\lambda_1,\dots,\lambda_k,\psi,\theta,q)$, as $\lambda_1,\dots,\lambda_k$ vary in $\{0,\dots,q-1\}$.

The group G contains both θ and ψ, since $P_k(1,0,\dots,0)=\theta$ and $P_k(0,1,0,\dots,0)=q\theta+\psi$. Since $\gcd(\psi,\theta,d)=1$, this implies that $G=\mathbb{Z}/d\mathbb{Z}$, and thus any state can be reached from any x. By Proposition 2.7, $\delta(0,\psi,0,\theta,x)=\delta(0,\psi,0,\theta,y)$ for all x,y. However,

$$\sum_{x=0}^{d-1}\delta(0,\psi,0,\theta,x)=1;$$

hence, $\delta(0,\psi,0,\theta,x)=1/d$ for all x, as claimed.

Suppose now that $a_k(\theta)\not\equiv\theta \bmod d$ for every k. Even in this case, there must be $m<m'$ such that $a_m(\theta)\equiv a_{m'}(\theta) \bmod d$; since $a_{i+j}(\theta)=a_i(a_j(\theta))$ for every i,j,θ, this means that $a_k(a_m(\theta))\equiv a_m(\theta) \bmod d$, where $k:=m'-m$. By the previous part of the proof, $\delta(0,\psi,0,a_m(\theta),x)=1/d$ for every $x\in\mathbb{Z}$. Since the set of the densities $\delta(0,\psi,0,\theta,x)$ is an eigenvector of Q^m, we have

$$\delta(\mathbf{s})=\sum_{\mathbf{t}\in X}\frac{\mu_m(\mathbf{s},\mathbf{t})}{q^m}\delta(\mathbf{t}).$$

Now $\mu_m((0,\psi,0,\theta,x),(0,\psi',0,\theta',x))=0$ unless $\psi'=\psi$ and $\theta'=a_m(\theta)$; hence, if $\mathbf{s}:=(0,\psi,0,\theta,x)$ we have

$$\delta(\mathbf{s})=\sum_{y=0}^{d-1}\frac{\mu_m(\mathbf{s},(0,\psi,0,a_m(\theta),y))}{q^m}\delta(0,\psi,0,a_m(\theta),y).$$

By the previous part of the proof, each $\delta(0,\psi,0,a_m(\theta),y)$ is equal to $1/d$; since the sum of all $\mu_m(\mathbf{s},\mathbf{t})$ is q^m, this means that $\delta(0,\psi,0,\theta,x)=1/d$, as claimed. ∎

Corollary 3.3. *Let $q,d\geq 2$ be integers. For every θ, the map $n\mapsto w_q(n)+\theta n$ is uniformly distributed.*

Proof. Let $\psi = 1$ in the previous theorem. ∎

In some cases we can also prove the existence of the limit without using the theory of q-additive functions.

Proposition 3.4. *Let $q \geq 2$ be an integer and $x, \theta \in \mathbb{Z}$. If θ is coprime to q, or if $q \mid \theta$, then the map $n \mapsto w_q(n) + \theta n$ is uniformly distributed modulo q.*

Proof. Suppose first that θ is coprime to q. In each block $\{qa, \ldots, qa + q - 1\}$, $w_q(n)$ is constant; hence, the equation $w_q(n) + \theta n \equiv x \bmod q$ has a unique solution, namely $n = qa + r$ with $r \equiv \theta^{-1}(x - w_q(qa)) \bmod q$. Hence, the number of solutions of $w_q(n) + \theta n \equiv x \bmod q$ in $[0, \ldots, N)$ is $N/d + O(1)$; i.e., $\gamma(N, q, q; 0, 1, 0, \theta, x) = N/d + O(1)$. If we divide by N, the limit on the right hand side exists and is equal to $1/d$; hence, the same applies to the left hand side.

For $q \mid \theta$ (and it is enough to consider $\theta = 0$), we note that

$$\frac{\gamma(N, q, q; 0, 1, 0, 0, x)}{N} = q\frac{\gamma(N/q, q, q; 0, 1, 0, 1, x)}{N} + O(1),$$

and the right hand side goes to $1/d$ by the previous reasoning. The claim is proved. ∎

4. The case $\mathbf{s} = (1, 0, 0, 0, x)$.

The methods used in the proof of Theorem 3.2 can also be applied to study the full equation (2); in particular, we shall be interested in the equation with $(\theta_u, \theta_w, \theta_2, \theta_1, \theta_0) = (1, 0, 0, 0, x) =: \mathbf{s}_x$. We are not able to obtain a full picture of the situation, so we will concentrate on two special cases. Before analyzing them, we note that we can obtain a lower limit for the density of the solutions.

Proposition 4.1. *Let d, q be positive integers and let $x \in \mathbb{Z}$. Then*

$$\liminf_{N\to\infty} \frac{\gamma(N, q, d; \mathbf{s}_x)}{N} \geq \frac{\varphi(d)}{dq}\left\lfloor \frac{q}{d} \right\rfloor,$$

where φ is the Euler function.

Proof. Consider the blocks $\{aq, aq + 1, \ldots, aq + q - 1\}$ of q consecutive natural numbers, starting from a multiple of q. In any block, the map $n \mapsto w_q(n)$ is constant, and thus the map $n \mapsto u_q(n)$ is of constant difference; in particular,

$$u_q(aq + r) = u_q(aq) + rw_q(aq).$$

If $w_q(aq)$ is coprime to d, then $u_q(aq) + rw_q(aq)$ passes through every residue class modulo d, as r goes from 0 to $d-1$; hence, the equation $u_q(n) \equiv x \bmod d$ has at least $\lfloor q/d \rfloor$ solutions in each block $\{aq, aq + 1, \ldots, aq + q - 1\}$.

Let now N be an integer, and divide $\{1, \ldots, N\}$ into blocks of length q. By Corollary 3.3, in approximately $\varphi(d)/d$ of these blocks $w_q(n)$ is coprime to d; hence,

$$\gamma(N,q,d;1,0,0,0,x) \geq \frac{N}{q}\frac{\varphi(d)}{d}\left\lfloor\frac{q}{d}\right\rfloor + O(1).$$

Dividing by N and taking the limit inferior we get our claim. ■

This result is far from being completely satisfactory (for example, it says nothing when $q < d$). However, it sometimes hits what is actually the real density, as the following theorem shows.

THEOREM 4.2. *Let $d, q \geq 2$ be integers, and suppose that $d \mid q$. Then, for every $x \in \mathbb{Z}$,*

$$\delta(q,d;1,0,0,0,x) = \frac{1}{d^2}\sum_{f \mid \gcd(x,d)} f \cdot \varphi\left(\frac{d}{f}\right).$$

This result can be seen as a generalization of [2, Proposition 3.4].

Proof. Let $\mathbf{s}_x := (1,0,0,0,x)$ with $x \in \mathbb{Z}/d\mathbb{Z}$; by Theorem 3.1(b), the density $\delta(q,d;\mathbf{s}_x)$ exists. We have $\mathbf{s}_x M_\lambda = (0,\lambda+1,0,\lambda+1,x)$; hence, applying Theorem 3.2, we obtain

$$\delta(q,d;\mathbf{s}_x) = \frac{1}{q}\sum_{\substack{\lambda=1,\ldots,q\\ \lambda \mid \gcd(d,x)}} \frac{\gcd(\lambda,d)}{d}.$$

However, since $d \mid q$, the summand for λ is equal to the summand for $\lambda + d$. Hence, the previous formula reads

$$\frac{1}{q}\frac{q}{d}\sum_{f \mid \gcd(x,d)} \frac{f}{d}|\{t \in \{1,\ldots,d\} : \gcd(t,d) = f\}| = \frac{1}{d^2}\sum_{f \mid \gcd(x,d)} f \cdot \varphi\left(\frac{d}{f}\right),$$

as claimed. ■

REMARK 4.3. If $\gcd(d,x) = 1$, then the theorem gives $\delta(q,d;\mathbf{s}) = \varphi(d)/d^2$, exactly the limit inferior obtained in Proposition 4.1.

The second case we consider is when d and q are coprime; it can be seen as a generalization of [2, Proposition 3.6], albeit with a stronger hypothesis (since we need all the densities to exist). We denote by $\mathcal{U}(\mathbb{Z}/d\mathbb{Z})$ the set of units of $\mathbb{Z}/d\mathbb{Z}$.

THEOREM 4.4. *Let $q, d \geq 2$ be coprime integers. Suppose that, for every $\mathbf{s} \in \mathbb{Z}^5$, the density $\delta(q,d;\mathbf{s}) = \delta(\mathbf{s})$ exists. Then, for every $\theta_w, x \in \mathbb{Z}/d\mathbb{Z}$ and each $\theta_u \in \mathcal{U}(\mathbb{Z}/d\mathbb{Z})$, we have*

$$\delta(q,d;\theta_u,\theta_w,0,0,x) = 1/d.$$

In particular, the map $n \mapsto u_q(n)$ is uniformly distributed modulo d.

Proof. The proof is similar to the proof of Theorem 3.2. Note first that we can suppose $\theta_u = 1$, since $\gamma(N,q,d;\mathbf{s}) = \gamma(N,q,d;u\mathbf{s})$ for every $u \in \mathcal{U}(\mathbb{Z}/d\mathbb{Z})$.

Since all the limits exist, by Theorem 2.6 the density vector $\boldsymbol{\delta}$ is a right eigenvector of the transition matrix of the Markov chain $\mathcal{C}(q,d)$.

We explicitly calculate the inverse M_0^{-1} as a matrix with rational entries:

$$M_0^{-1} = \frac{1}{q^2}\begin{pmatrix} q & q(q-1) & -1 & -(q-1)/2 & 0 \\ 0 & q^2 & 0 & -q & 0 \\ 0 & 0 & 1 & (q-1)/2 & 0 \\ 0 & 0 & 0 & q & 0 \\ 0 & 0 & 0 & 0 & q^2 \end{pmatrix}.$$

Hence,

$$M_1M_0^{-1} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$M_1^2M_0^{-2} = \begin{pmatrix} 1 & q+1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & q+1 & (q+1)(q+2)/2 \\ 0 & 0 & 0 & 1 & q+1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$M_2^2M_0^{-2} = \begin{pmatrix} 1 & 2(q+1) & 0 & 0 & q+6 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 2(q+1) & (q+1)(2q+3) \\ 0 & 0 & 0 & 1 & 2(q+1) \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

In particular, $M_1M_0^{-1}$, $M_1^2M_0^{-2}$ and $M_2^2M_0^{-2}$ all have integer coefficients, and so they can always be reduced modulo d.

Consider now the matrices M_λ modulo d. Each determinant is equal to q^4; since q and d are coprime, these matrices are all invertible modulo d, and their inverses are the reduction modulo d of their rational inverses. Moreover, since $\mathrm{GL}_5(\mathbb{Z}/d\mathbb{Z})$ is a finite group, each M_λ has a finite order h_λ; hence, each state of the Markov chain $\mathcal{C}(q,d)$ is ergodic. Indeed, if $\mathbf{t}$ is reachable from $\mathbf{s}$, then $\mathbf{t} = \mathbf{s}M_{\lambda_1}\cdots M_{\lambda_k}$ for some $\lambda_1,\ldots,\lambda_k$; but then

$$\mathbf{s} = \mathbf{t}M_{\lambda_k}^{h_{\lambda_k}-1}\cdots M_{\lambda_1}^{h_{\lambda_1}-1},$$

so $\mathbf{s}$ is reachable from $\mathbf{t}$.

We claim that, for every a and x, the tuples $(1,a,0,0,x)$ and $(1,b,0,0,x)$ are communicating.

For every θ_w and every x, we have the three equalities

$$(1\ \theta_w\ 0\ 0\ x)\,M_1M_0^{-1} = (1\ \theta_w+1\ 0\ 0\ x),$$
$$(1\ 0\ 0\ 0\ x)\,M_1^2M_0^{-2} = (1\ q+1\ 0\ 0\ x+2)$$

and

$$(1\ 0\ 0\ 0\ x)\,M_2^2M_0^{-2} = (1\ 2(q+1)\ 0\ 0\ x+q+6).$$

Since $q \geq 2$, we can always use M_0 and M_1; hence, the first equality implies that, for every x, the 5-tuples $(1,a,0,0,x)$ and $(1,b,0,0,x)$ are communicating, while the first and second imply that $(1,0,0,0,x)$ and $(1,a,0,0,x+2)$ are communicating for all a and x (in particular, $a=0$).

If d is odd, this implies that $(1,0,0,0,x) \leftrightarrow (1,0,0,0,y)$ for every $x,y \in \mathbb{Z}/d\mathbb{Z}$, and thus $(1,a,0,0,x)$ and $(1,b,0,0,y)$ are communicating whatever a and b are.

If d is even, then $q \geq 3$ (since $\gcd(d,q)=1$); hence, we can also use M_2, obtaining $(1,0,0,0,x) \leftrightarrow (1,2(q+1),0,0,x+q+6)$, and so $(1,0,0,0,x) \leftrightarrow (1,a,0,0,x+q+6)$ for every a. Hence, $(1,0,0,0,x)$ is communicating with

$$(1,a,0,0,x+2z_1+(q+6)z_2)$$

for every choice of $z_1,z_2 \in \mathbb{N}$. However, q is odd, and thus $\gcd(2,q+6,d)=1$: hence, $2z_1+(q+6)z_2$ can be equal to any $s \in \mathbb{Z}/d\mathbb{Z}$. Therefore, $(1,0,0,0,x) \leftrightarrow (1,a,0,0,y)$ for all a and y, as claimed.

Consequently, for every d, we have $\delta(1,a,0,0,x) = \delta(1,b,0,0,y)$ for any a,b,x,y. However,

$$\sum_{x=0}^{d-1} \delta(1,\theta_w,0,0,x) = 1;$$

hence, for all a and x we have $\delta(1,a,0,0,x) = 1/d$, as claimed. ∎

5. Algebraic interpretation.

Let D be an integral domain with quotient field K. The set of *integer-valued polynomials* on D is

$$\mathrm{Int}(D) := \{f \in K[X] : f(D) \subseteq D\}.$$

The set $\mathrm{Int}(D)$ is always an integral domain contained between $D[X]$ and $K[X]$. There are two sequences of D-modules associated to $\mathrm{Int}(D)$: the first is formed by the *characteristic ideals* $\mathfrak{I}_n := \mathfrak{I}_n(D)$, defined as the union of (0) with the leading coefficients of the polynomials of $\mathrm{Int}(D)$ of degree n; the second contains the modules of the form

$$\mathrm{Int}_n(D) := \{f \in \mathrm{Int}(D) : \deg f \leq n\}.$$

If D is a Dedekind domain, these two sequences are linked by the relation [1, Corollary II.3.6]

$$\mathrm{Int}_n(D) \simeq \bigoplus_{k=0}^{n} \mathfrak{J}_k \simeq D^n \oplus \prod_{k=0}^{n} \mathfrak{J}_k. \tag{5}$$

For any maximal ideal $\mathfrak{m}$ of D, let $N(\mathfrak{m})$ be the *norm* of $\mathfrak{m}$, that is, the cardinality of $D/\mathfrak{m}$; for any $q \in \mathbb{N}$, let $\mathit{\Pi}_q$ be the product of the maximal ideals of norm q. The subgroup of the class group generated by the $\mathit{\Pi}_q$ is called the *Pólya–Ostrowski group* of D, and is denoted by $\mathrm{Po}(D)$. Several papers studied $\mathrm{Po}(D)$ when D is an integral extension of $\mathbb{Z}$, with special focus on D for which the group $\mathrm{Po}(D)$ is trivial (in this case, the quotient field K of D is said to be a *Pólya field*) [12, 7, 3]: this happens if and only if $\mathrm{Int}(D)$ has a *regular basis*, i.e., a basis $\{f_0, f_1, \ldots,\}$ over D such that $\deg f_i = i$ for every i. For example, every cyclotomic extension of $\mathbb{Q}$ is a Pólya field [1, Proposition II.4.3].

The (classes of the) characteristic ideals of D naturally belong to $\mathrm{Po}(D)$, and by [1, Proposition II.3.9] we have

$$\mathfrak{J}_n = \prod_{q=2}^{n} \mathit{\Pi}_q^{-w_q(n)}.$$

On the other hand, the modules $\mathrm{Int}_n(D)$ do not belong, by themselves, to $\mathrm{Po}(D)$, for the trivial reason that they are not fractional ideals of D. However, by (5), we can write $\mathrm{Int}_n(D)$ as the direct sum $D^n \oplus \widehat{\mathrm{Int}_n}(D)$, where

$$\widehat{\mathrm{Int}_n}(D) := \prod_{k=1}^{n} \mathfrak{J}_n(D)$$

is a fractional ideal of D; by construction, the isomorphism class of $\widehat{\mathrm{Int}_n}(D)$ belongs to $\mathrm{Po}(D)$. Moreover, since D is a Dedekind domain, $\widehat{\mathrm{Int}_n}(D)$ is a projective module of rank 1, and thus $\mathrm{Int}_n(D)$ is free if and only if $\widehat{\mathrm{Int}_n}(D)$ is free [6, Theorem 4.11]. Applying again (5), we see that

$$\widehat{\mathrm{Int}_n}(D) \simeq \prod_{q=2}^{n} \mathit{\Pi}_q^{-u_q(n)}$$

for every $n \in \mathbb{N}$.

Therefore, we have two maps $\mathbb{N} \to \mathrm{Po}(D)$ given by

$$n \mapsto [\mathfrak{J}_n(D)] \quad \text{and} \quad n \mapsto [\widehat{\mathrm{Int}_n}(D)],$$

and studying how many times $\mathfrak{J}_n(D)$ and $\widehat{\mathrm{Int}_n}(D)$ are isomorphic to a specific module is essentially equivalent to studying the limit distribution of these maps in $\mathrm{Po}(D)$. Elliott [2] conjectured that the density of the natural numbers such that $\mathrm{Int}_n(D)$ is free exists and is rational, and it is always at least $1/|\mathrm{Po}(D)|$. He also makes several conjectures for more specific cases, mostly expressed in terms of multisets.

Suppose now that there is a unique q such that Π_q is not a principal ideal of D, and let $d := |\mathrm{Po}(D)|$. Then $\mathrm{Po}(D) \simeq \mathbb{Z}/d\mathbb{Z}$ is a cyclic group and $[\Pi_q]$ is a generator; moreover,

$$\mathfrak{J}_n \simeq \Pi_q^{-w_q(n) \bmod d} \quad \text{and} \quad \widehat{\mathrm{Int}_n}(D) \simeq \Pi_q^{-u_q(n) \bmod d}.$$

Therefore, the distribution of the maps $n \mapsto [\mathfrak{J}_n(D)]$ and $n \mapsto [\widehat{\mathrm{Int}_n}(D)]$ in $\mathrm{Po}(D)$ is determined by the distribution of $n \mapsto w_q(n)$ and $n \mapsto u_q(n)$ modulo d, and we can simply translate the results in Sections 3 and 4 to this context; the definitions of limit distribution and of being uniformly distributed in $\mathrm{Po}(D)$ is analogous to the ones after Definition 2.4.

PROPOSITION 5.1. *Let D be a Dedekind domain, and suppose that Π_n is nonprincipal only for $n = q$; let $d := |\mathrm{Po}(D)|$.*

(a) *The map $n \mapsto [\mathfrak{J}_n(D)]$ is uniformly distributed in* $\mathrm{Po}(D)$.
(b) *If, for every $\mathbf{s}$, the density $\delta(q, d; \mathbf{s})$ exists, then for every $g \in \mathrm{Po}(D)$ the density of n such that $[\widehat{\mathrm{Int}_n}(D)] = g$ is rational.*
(c) *If $d \mid q$, the density of n such that $\mathrm{Int}_n(D) \simeq \Pi_q^x \oplus D^n$ is equal to*

$$\frac{1}{d^2} \sum_{f \mid \gcd(x,d)} f \cdot \varphi\left(\frac{d}{f}\right).$$

(d) *If q and d are coprime and, for every $\mathbf{s} \in \mathbb{Z}^5$, the density $\delta(q, d; \mathbf{s})$ exists, then the map $n \mapsto [\widehat{\mathrm{Int}_n}(D)]$ is uniformly distributed in* $\mathrm{Po}(D)$.

Proof. The four statements are the translation, respectively, of Corollary 3.3, Theorem 2.6, Theorem 4.2 and Theorem 4.4. ∎

Suppose D is the ring of integers of a number field K. Some examples in which D satisfies the hypotheses of the previous proposition are given in [2, Examples 7.3 and 7.4]. Other examples can be constructed using [1, proof of Proposition II.4.2]: if K is Galois over $\mathbb{Q}$, then Π_q is principal for every $q = p^r$ such that p is not ramified in K. Since p is ramified if and only if it divides the discriminant [10, Chapter II, Corollary 2.12], Proposition 5.1 can be applied if the discriminant of K is a prime power. Unfortunately, the simplest cases in which this happens (the quadratic fields $K = \mathbb{Q}(\sqrt{p})$, where $|p|$ is a prime and $p \equiv 1 \bmod 4$, and the cyclotomic extensions $K = \mathbb{Q}(\zeta_p)$ for p prime) are also Pólya fields [1, Proposition II.4.3 and Corollary II.4.5], and thus $\mathrm{Po}(D)$ is actually trivial; other examples where the discriminant is a prime power (with unknown Pólya–Ostrowski group) are collected in [5].

When there is more than one Π_q which is nonprincipal, it is necessary to study the density of the n that are simultaneous solutions to the equations

$$\theta_u^{(i)} u_{q_i}(n) + \theta_w^{(i)} w_{q_i}(n) + \theta_2^{(i)} \frac{n(n+1)}{2} + \theta_1^{(i)} n + \theta_0^{(i)} \equiv 0 \bmod d_i$$

for $i \in \{1, \dots, k\}$, where $q_1, \dots, q_k$ and $d_1, \dots, d_k$ are arbitrary integers (and $q_1, \dots, q_k$ are pairwise different). The problem is essentially in determining how much these equations are correlated; it is reasonable to think that (under the obvious hypothesis that at least one of $\theta_u^{(i)}$ and $\theta_w^{(i)}$ is nonzero modulo d_i, for each i) such equations are actually independent, so that the density of the solutions of all the equations is determined by the densities of the solutions of the single equations. The major obstruction seems to be the problem of understanding the behaviour of $u_{q_1}(aq_2 + \lambda)$ and $w_{q_1}(aq_2 + \lambda)$ when $q_1 \neq q_2$.

Acknowledgements. The author wishes to thank Francesco Pappalardi and the referee for their suggestions.

References

[1] P.-J. Cahen and J.-L. Chabert, *Integer-Valued Polynomials*, Math. Surveys Monogr. 48, Amer. Math. Soc., Providence, RI, 1997.

[2] J. Elliott, *The probability that* $\mathrm{Int}_n(D)$ *is free*, in: M. Fontana et al. (eds.), Commutative Algebra, Springer, New York, 2014, 133–151.

[3] B. Heidaryan and A. Rajaei, *Biquadratic Pólya fields with only one quadratic Pólya subfield*, J. Number Theory 143 (2014), 279–285.

[4] L. Hogben (ed.), *Handbook of Linear Algebra*, Chapman & Hall/CRC, Boca Raton, FL, 2007.

[5] J. W. Jones and D. P. Roberts, *Number fields ramified at one prime*, in: Algorithmic Number Theory, Lecture Notes in Comput. Sci. 5011, Springer, Berlin, 2008, 226–239.

[6] T. Y. Lam, *Serre's Problem on Projective Modules*, Springer Monogr. Math., Springer, Berlin, 2006.

[7] A. Leriche, *Cubic, quartic and sextic Pólya fields*, J. Number Theory 133 (2013), 59–71.

[8] C. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, PA, 2000.

[9] W. Narkiewicz, *Uniform Distribution of Sequences of Integers in Residue Classes*, Lecture Notes in Math. 1087, Springer, Berlin, 1984.

[10] J. Neukirch, *Algebraic Number Theory*, Grundlehren Math. Wiss. 322, Springer, Berlin, 1999.

[11] M. Peter and J. Spilker, *Value distribution of g-additive functions*, Manuscripta Math. 105 (2001), 519–536.

[12] H. Zantema, *Integer valued polynomials over a number field*, Manuscripta Math. 40 (1982), 155–203.

Dario Spirito
Dipartimento di Matematica e Fisica
Università degli Studi “Roma Tre”
Roma, Italy
E-mail: spirito@mat.uniroma3.it