Combinatorics, Automata and Number Theory

CANT

Valérie Berthé, LIRMM - Univ. Montpelier II - CNRS UMR 5506
161 rue Ada, 34392 Montpellier Cedex 5, France

Michel Rigo, Université de Liège, Institut de Mathématiques
Grande Traverse 12 (B 37), B-4000 Liège, Belgium

Contents

8

Analysis of digital functions and applications

8.1 Introduction: digital functions

Digital functions in a rather informal and general sense are functions defined in a way depending on the digits in some digital representation of the integers. In the simplest case the digital representation is the q-adic representation and the dependence of the function on the digits is additive as for the sum-of-digits function given by

$$s_q\left(\sum_{k=0}^{K}\varepsilon_k q^k\right)=\sum_{k=0}^{K}\varepsilon_k,$$

which also serves as the most prominent example for such functions. As a very general reference for results on digital functions, we refer to (Allouche and Shallit 2003). We remark that depending on the point of view such maps $f:\mathbb{N}\to A$ can be seen as (arithmetic) functions or sequences. The aim of this chapter is to study various asymptotic and limiting properties of such functions.

For the convenience of the reader we collect the basic definitions as given in (Allouche and Shallit 2003).

Automatic sequences

Definition 8.1.1 A sequence $(v(n))_{n\in\mathbb{N}}$ is called *q-automatic*, if the collection of sequences

$$K_q(v)=\left\{\left(v\left(q^k n+\ell\right)\right)_{n\in\mathbb{N}} \mid k\in\mathbb{N}, 0\le\ell<q^k\right\},\quad \text{"the } q\text{-kernel"},\quad (8.1)$$

is finite.

This definition is equivalent to saying that the value $v(n)$ can be determined by a deterministic finite automaton on the q-adic digits of n. Furthermore, this definition is equivalent to saying that the sequence $(v(n))_{n\in\mathbb{N}}$ is the

image of a fixed point of a morphism of constant length q on a finite alphabet (see the discussion in Chapters˜5 and˜6 of (Allouche and Shallit 2003) and Sections˜**??** and˜**??**).

We remark that for this definition the set of values A of v is simply a finite set without any further structure.

Regular sequences

If the set A is a ring (in most of the examples this is $\mathbb{Z}$, $\mathbb{F}_q$, $\mathbb{R}$, or $\mathbb{C}$), then the structure of the ring A can be used for the following definition (cf.˜Chapter˜16 of (Allouche and Shallit 2003)).

Definition 8.1.2 Let A be a ring. Then an A-valued sequence $(f(n))_{n\in\mathbb{N}}$ is called *q-regular*, if the A-module generated by the q-kernel (8.1) is finitely generated.

This is equivalent to saying that there is a positive integer r, sequences $(f_1(n))_{n\in\mathbb{N}} = (f(n))_{n\in\mathbb{N}}, \dots, (f_r(n))_{n\in\mathbb{N}}$ and map $\mathbf{M} : \{0, \dots, q-1\} \to A^{r\times r}$, such that for all $n \in \mathbb{N}$ and all $b \in \{0, \dots, q-1\}$

$$\begin{pmatrix} f_1(qn+b) \\ f_2(qn+b) \\ \vdots \\ f_r(qn+b) \end{pmatrix} = \mathbf{M}(b) \begin{pmatrix} f_1(n) \\ f_2(n) \\ \vdots \\ f_r(n) \end{pmatrix}. \tag{8.2}$$

q-Additive functions

For the following definition it is more convenient to view the sequence of values $(v(n))_{n\in\mathbb{N}}$ as a function on the positive integers taking its values in an abelian group. The most important examples use the groups $\mathbb{R}$, $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, or $\mathbb{Z}/m\mathbb{Z}$. The group law will be written additively.

Definition 8.1.3 Let A be an abelian group. A function $f : \mathbb{N} \to A$ is called q-additive, if for all $n \in \mathbb{N}$, all $k \in \mathbb{N}$, and all $0 \le \ell < q^k$

$$f\left(q^k n + \ell\right) = f(q^k n) + f(\ell) \tag{8.3}$$

holds. If there is no dependence on k on the right-hand-side, i.e.

$$f\left(q^k n + \ell\right) = f(n) + f(\ell), \tag{8.4}$$

then f is called *completely q-additive*.

From (8.3) it follows by induction that

$$f\left(\sum_{k=0}^{K}\varepsilon_k q^k\right)=\sum_{k=0}^{K}f(\varepsilon_k q^k),$$

which shows that a q-additive function f is determined by the values $f(\varepsilon q^k)$, $\varepsilon\in\{1,\dots,q-1\}$, $k\in\mathbb{N}$, and $f(0)=0$. A completely q-additive function is determined by the values $f(1),\dots,f(q-1)$.

The values εq^k can be viewed as the (additive) "building blocks" of the positive integers with respect to q-adic numeration. This is a direct analogy to classical additive functions as studied in analytic number theory (cf.˜(Tenenbaum 1995, Elliott 1979, Elliott 1980, Elliott 1985)). In the case the (multiplicative) "building blocks" of the positive integers, namely the prime powers are used to define

$$f(p_1^{e_1}p_2^{e_2}\cdots p_k^{e_k})=f(p_1^{e_1})+f(p_2^{e_2})+\cdots+f(p_k^{e_k}).$$

These functions and their properties of their value distribution are usually studied by probabilistic methods ("Kubilius models"). In Section˜8.3 analogous models will be presented for q-additive functions.

Remark 8.1.4 A completely q-additive function f taking its values in a ring A is q-regular, since the A-module generated by f and the constant function 1 equals the A-module generated by the q-kernel.

q-Multiplicative functions

Multiplicative functions are defined analogous to additive functions using a multiplicative structure (usually the multiplicative group of the field $\mathbb{R}$ or $\mathbb{C}$ or the multiplicative semigroup of the ring $\mathbb{Z}$).

Definition 8.1.5 Let A be a monoid (written multiplicatively). A function $f:\mathbb{N}\to A$ is called *q-multiplicative*, if for all $n\in\mathbb{N}$, all $k\in\mathbb{N}$, and all $0\le\ell<q^k$

$$f\left(q^k n+\ell\right)=f(q^k n)f(\ell) \tag{8.5}$$

holds. If there is no dependence on k on the right-hand-side, i.e.

$$f\left(q^k n+\ell\right)=f(n)f(\ell), \tag{8.6}$$

then f is called *completely q-multiplicative*.

From (8.5) it follows by induction that

$$f\left(\sum_{k=0}^{K}\varepsilon_k q^k\right)=\prod_{k=0}^{K}f(\varepsilon_k q^k), \tag{8.7}$$

which shows that a q-additive function f is determined by the values $f(\varepsilon q^k)$, $\varepsilon\in\{1,\dots,q-1\}$, $k\in\mathbb{N}$, and $f(0)=1$. A completely q-multiplicative function is determined by the values $f(1),\dots,f(q-1)$.

Remark 8.1.6 A completely q-multiplicative function f taking its values in a ring A is q-regular, since the A-module generated by f equals the A-module generated by the q-kernel.

Remark 8.1.7 If $f(n)$ is a q-automatic sequence taking its values in the finite set A, the indicator function $\mathbb{1}_{\{a\}}(f(n))$ (for $a\in A$) can be expressed in terms a matrix product involving the transition matrices of the underlying finite automaton: let $\mathcal{A}=(Q,\{0,\dots,q-1\},\delta,\{i_0\},A,\tau)$ be the DFAO defining f. For $a\in\{0,\dots,q-1\}$ define the $Q\times Q$-matrix $\mathbf{M}_\delta(a)$ by

$$(m_\delta(a))_{ij}=\begin{cases}1 & \text{if } \delta(i,a)=j\\ 0 & \text{otherwise}\end{cases}$$

and the vectors $\mathbf{v}=(1,0,\dots,0)$ with the entry 1 in position i_0 and $\mathbf{w}$

$$w_\ell=\begin{cases}1 & \text{if } \tau(\ell)=a\\ 0 & \text{otherwise.}\end{cases}$$

Then

$$\mathbb{1}_{\{a\}}(f(n))=\mathbf{v}\prod_{j=0}^{J}\mathbf{M}_\delta(\varepsilon_j(n))\mathbf{w} \tag{8.8}$$

gives this representation.

Since the underlying DFA $(Q,\{0,\dots,q-1\},\delta,\{i_0\})$ recognises all q-adic representations, the matrix

$$\mathbf{M}_\delta=\mathbf{M}_\delta(0)+\mathbf{M}_\delta(1)+\cdots+\mathbf{M}_\delta(q-1) \tag{8.9}$$

has the eigenvalue q and all eigenvalues are of modulus $\le q$. Furthermore, the eigenvalue q has the same geometric and algebraic multiplicity (cf.˜Section˜**??**).

8.2 Asymptotic analysis of digital functions

In this section we study the asymptotic behaviour of summatory functions of various digital functions. We develop several techniques and discuss their strengths and weaknesses. As a general theme, we can say that the values $(f(n))_{n\in\mathbb{N}}$ for digital functions usually vary very irregularly. Nevertheless, the summatory function

$$F(N) = \sum_{n<N} f(n) \tag{8.10}$$

shows a rather "smooth" asymptotic behaviour, in many cases there even exists an exact formula for $F(N)$. Usually, this formula involves a periodic continuous function of $\log_q N$.

8.2.1 Completely additive functions

Completely additive functions, such as the sum-of-digits function have been the first and simplest type of digital functions that has been studied. The study of the sum (8.10) is especially easy in this case and will be used to develop several techniques. Furthermore, this will be a preparation for the development in Section˜8.3, where we consider f as a random variables on the finite probability space

$$\Omega_N = \{0, \dots, N-1\}, \quad \mathbb{P}_N(A) = \frac{\#A}{N}.$$

In this interpretation, $F(N)/N$ is simply the mean of the function f on the space Ω_N.

Theorem 8.2.1 *Let $f : \mathbb{N} \to \mathbb{R}$ be a completely q-additive function given by the values $f(0) = 0, f(1), \dots, f(q-1)$. Then there exists a continuous periodic function Φ_f of period 1, such that*

$$\sum_{n<N} f(n) = C_f N \log_q N + N\Phi_f(\log_q N) \tag{8.11}$$

with $C_f = \frac{1}{q}\left(f(1) + \cdots + f(q-1)\right)$

Remark 8.2.2 Notice that the fractional parts of $\log_q N$ are dense in the interval $[0, 1]$. This allows to interpret continuity of Φ in the sense that the discrete set of points

$$\left\{(\{\log_q N\}, F(N)/N - C_f \log_q N) \mid N \in \mathbb{N}\right\}$$

can be fit by one (and only one!) graph of a continuous function.

Remark 8.2.3 The function Φ_f is continuous and nowhere differentiable (except for trivial cases). This fact has been observed in (Tenenbaum 1997) for rather general periodic functions occurring in the context of digital functions.

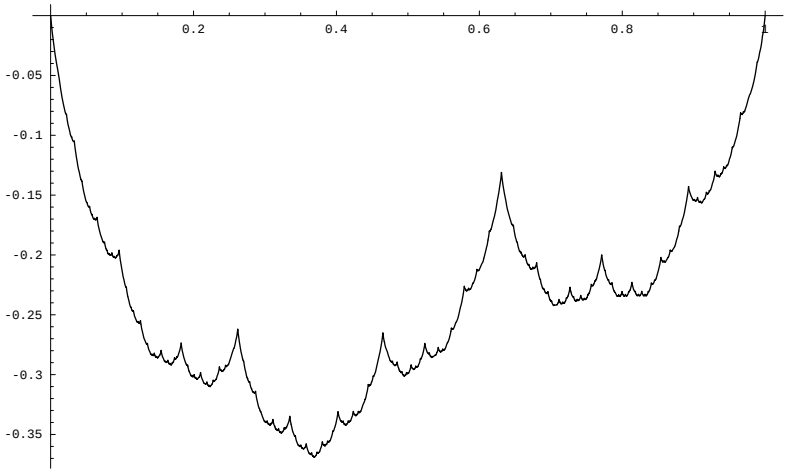

Fig. 8.1. Plot of the function Φ_f for the sum-of-digits function in base 3

Proof [(Delange 1975)] We will present Delange's method for general completely q-additive functions f. As a first step we rewrite $F(N)$ as

$$\sum_{n<N} f(n) = \sum_{k=0}^{K} \sum_{n<N} f(\varepsilon_k(n)).$$

The k-th digit $\varepsilon_k(x)$ of a real number x is a periodic function of period q^{k+1}, which is constant on intervals $[nq^k, (n+1)q^k)$. Using $\varepsilon_k(x) = \varepsilon_0(xq^{-k})$, the inner sum can be rewritten as an integral

$$\sum_{n<N} f(\varepsilon_k(x)) = \int_0^N f(\varepsilon_k(x))\,dx = NC_f + q^k \int_0^{Nq^{-k}} (f(\varepsilon_0(x)) - C_f)\,dx.$$

Thus we have by inverting the order of summation

$$\sum_{n<N} f(n) = C_f KN + q^K \sum_{k=0}^{K} q^{-k} \int_0^{q^k N q^{-K}} (f(\varepsilon_0(x)) - C_f)\,dx.$$

Since $C_f = \frac{1}{q}\int_0^q f(\varepsilon_0(x))\,dx$, the last integral vanishes for $k > K$. Thus the sum can be extended to an infinite sum, and we get

$$\sum_{n<N} f(n) = C_f N \log_q N + N\left(q^{-\{\log_q N\}} \sum_{k=0}^{\infty} q^{-k} g(q^{\{\log_q N\}+k}) - C_f\{\log_q N\}\right),$$

where $g(t) = \int_0^t f(\varepsilon_0(x) - C_f)\,dx$ is a piecewise linear periodic function of period q. Taking

$$\Phi_f(x) = q^{-x}\sum_{k=0}^{\infty} q^{-k} g(q^{k+x}) - C_f x \quad x \in [0,1]$$

gives the desired result. □

Remark 8.2.4 Delange's method is applicable in cases of systems of numeration, which allow to give a closed expression for the single digits. This is the case for instance for canonical number systems (see Section˜1.3), where this method has been applied by (Grabner, Kirschenhofer, and Prodinger 1998) to prove an asymptotic formula for the sum of digits on the Gaussian integers.

We compute the Fourier-coefficients of Φ_f

$$\widehat{\Phi_f}(m) = \int_0^1 \Phi_f(x) e^{-2\pi i m x}\,dx = \sum_{k=0}^{\infty}\int_0^1 q^{-k-x} g(q^{k+x}) e^{-2\pi i m x}\,dx - C_f\int_0^1 x e^{-2\pi i m x}\,dx.$$

We substitute $u = q^{k+x}$ in the integral and observe that the ranges of integration for u add then up to $[1,\infty)$

$$\widehat{\Phi_f}(m) = \frac{1}{\log q}\int_1^{\infty} g(u) u^{-\frac{2\pi i m}{\log q}-2}\,du + \frac{C_f}{2\pi i m}.$$

This integral is easily recognised as a Mellin transform and can be computed by partial integration using the fact that the integrand is periodic with period q.

We write

$$\mathcal{M}g(s) = \int_1^{\infty} g(u) u^{s-1}\,du, \tag{8.12}$$

and using the notation $\chi_m = \frac{2\pi m}{\log q}$

$$\widehat{\Phi_f}(m) = \frac{1}{\log q}\mathcal{M}g(-1-\chi_m) + \frac{C_f}{2\pi i m}.$$

Since g is bounded, the integral (8.12) converges for $\Re s < 0$. By partial integration we get

$$\begin{aligned}\mathcal{M}g(s) = g(u)\frac{u^s}{s}\Big|_{u=1}^{\infty} - \frac{1}{s}\int_1^\infty (f(\varepsilon_0(u)) - C_f)u^s\,du \\ = \frac{C_f}{s} - \frac{1}{s}\int_1^\infty (f(\varepsilon_0(u)) - C_f)u^s\,du.\end{aligned}$$

The function $f(\varepsilon_0(u))$ is periodic with period q and constant on intervals between consecutive integers, which allows to compute the last integral

$$\begin{aligned}&\frac{C_f}{s} - \frac{1}{s}\int_1^\infty (f(\varepsilon_0(u)) - C_f)u^s\,du \\ &= \frac{C_f}{s} - \frac{1}{s(s+1)}\sum_{k=1}^{q-1} f(k)\sum_{n=0}^{\infty}\left((qn+k+1)^{s+1} - (qn+k)^{s+1}\right) - \frac{C_f}{s(s+1)} \\ &= \frac{C_f}{s+1} - \frac{q^{s+1}}{s(s+1)}\sum_{k=1}^{q-1} f(k)\left(\zeta\left(-s-1, \frac{k+1}{q}\right) - \zeta\left(-s-1, \frac{k}{q}\right)\right),\end{aligned} \tag{8.13}$$

where

$$\zeta(s,\alpha) = \sum_{n=0}^{\infty}\frac{1}{(n+\alpha)^s}$$

denotes the Hurwitz zeta function; the Riemann zeta function is given by $\zeta(s) = \zeta(s,1)$. The poles of the Hurwitz zeta functions in (8.13) at $s=-2$ cancel; furthermore, the poles at $s=-1$ in (8.13) cancel by the fact that $\zeta(0,\alpha) = \frac{1}{2} - \alpha$.

Putting everything together gives (for $m \neq 0$)

$$\widehat{\Phi_f}(m) = -\frac{1}{\chi_m(\chi_m+1)}\sum_{k=1}^{q-1} f(k)\left(\zeta\left(\chi_m, \frac{k+1}{q}\right) - \zeta\left(\chi_m, \frac{k}{q}\right)\right)$$

and

$$\begin{aligned}\widehat{\Phi_f}(0) = \frac{1}{\log q}\int_1^\infty g(u)u^{-2}\,du - \frac{C_f}{2} \\ = C_f\left(\frac{1}{\log q} - \frac{1}{2}\right) - \frac{1}{\log q}\sum_{k=1}^{q-1} f(k)\log\frac{\Gamma((k+1)/q)}{\Gamma(k/q)}.\end{aligned}$$

Here we have used

$$\zeta'(0,\alpha) = -\frac{1}{2}\log(2\pi) + \log\Gamma(\alpha).$$

If $f = s_q$ is the sum of digits function, these formulæ for the Fourier-coefficients can be simplified further. Since $s_q(k) = k$ for $k = 0, \ldots, q-1$, Abelian summation yields

$$\widehat{\Phi_{s_q}}(m) = -\frac{(q-1)}{\chi_m(\chi_m+1)}\zeta(\chi_m) \quad \text{for } m \neq 0$$

and

$$\widehat{\Phi_{s_q}}(0) = \frac{q-1}{2}\left(\frac{1}{\log q} - \frac{\log 2\pi}{\log q} - \frac{1}{2}\right) + \frac{1}{2},$$

which is the result of (Delange 1975).

Remark 8.2.5 Finding the maxima and minima of the periodic function Φ_f is a tricky task. For the sum-of-digits function that has been done in (Foster 1987), (Foster 1991), and (Foster 1992). For the counting functions of q-adic digits $\geq d$, the minimum of the corresponding periodic function has been determined in (Grabner 2004).

Remark 8.2.6 There is an elementary method for proving (8.11) which runs along the same lines as the proof of Theorem˜8.2.7 below. This method can be generalised to recurrence based systems of numeration and other generalisations of base q numeration systems (cf. (Grabner and Tichy 1990), (Grabner and Tichy 1991), (Kirschenhofer and Tichy 1984), (Grabner and Rigo 2003)). Working out this proof of Theorem˜8.2.1 is left as an exercise.

8.2.2 Multiplicative functions

Multiplicative functions with respect to numeration have been introduced to study statistical properties of additive functions via exponential sums (cf.˜(Delange 1972)). This is used in Section˜8.3 to derive limit theorems of various kinds for additive functions.

In this section we study the summatory functions of completely q-multiplicative functions by elementary means. The following theorem was proved in (Grabner 1993).

Theorem 8.2.7 *Let f be a complex valued completely q-multiplicative function satisfying*

$$|1 + f(1) + \ldots + f(q-1)| > \max_{0 \leq a < q} |f(a)|. \tag{8.14}$$

Then there exists a continuous periodic function ψ of period 1, such that

$$F(N) = \sum_{n<N} f(n) = N^{\rho} e^{i\alpha \log_q N} \psi(\log_q N), \tag{8.15}$$

where $\rho = \log_q |F(q)|$ and $\alpha = \arg(F(q))$.

Proof We write

$$N = \sum_{k=0}^{K} \varepsilon_k(N) q^k$$

and set

$$N_\ell = \sum_{k=\ell}^{K} \varepsilon_k(N) q^k.$$

Then we split the sum for $F(N)$

$$F(N) = \sum_{\ell=0}^{K} \sum_{n<\varepsilon_\ell(N)q^\ell} f(N_{\ell+1}+n) = \sum_{\ell=0}^{K} f(N_{\ell+1}) F(\varepsilon_\ell(N) q^\ell). \tag{8.16}$$

Thus we have reduced the problem to the computation of $F(aq^\ell)$ for $a \in \{0, \dots, q-1\}$, which can be done using the independence of the digits

$$F(aq^\ell) = \sum_{\varepsilon_0=0}^{q-1} \cdots \sum_{\varepsilon_{\ell-1}=0}^{q-1} \sum_{\varepsilon_\ell=0}^{a-1} f(\varepsilon_0) \cdots f(\varepsilon_\ell) = F(q)^\ell F(a).$$

Inserting this into (8.16) gives

$$F(N) = \sum_{\ell=0}^{K} F(q)^\ell F(\varepsilon_\ell(N)) f(\varepsilon_{\ell+1}(N)) \cdots f(\varepsilon_K(N)). \tag{8.17}$$

We define a function $\varphi : [1, q] \to \mathbb{C}$ by

$$\varphi\left(\sum_{\ell=0}^{\infty} \frac{\varepsilon_\ell}{q^\ell} \right) = \sum_{\ell=0}^{\infty} F(\varepsilon_\ell) F(q)^{-\ell} \prod_{k=0}^{\ell-1} f(\varepsilon_k). \tag{8.18}$$

Notice that this series is geometrically convergent by the assumption (8.14).

The function is well defined and continuous, because

$$\begin{aligned}\varphi\left(\sum_{\ell=0}^{L}\frac{\varepsilon_\ell}{q^\ell}\right) &= \sum_{\ell=0}^{L} F(\varepsilon_\ell)F(q)^{-\ell}\prod_{k=0}^{\ell-1} f(\varepsilon_k)\\ &= \sum_{\ell=0}^{L-1} F(\varepsilon_\ell)F(q)^{-\ell}\prod_{k=0}^{\ell-1} f(\varepsilon_k) + F(\varepsilon_L-1)F(q)^{-L}\prod_{k=0}^{L-1} f(\varepsilon_k)\\ &\quad + \prod_{k=0}^{L} f(\varepsilon_k)\sum_{\ell=L+1}^{\infty} F(q-1)F(q)^{-\ell}f(q-1)^{\ell-L-1}\\ &= \varphi\left(\sum_{\ell=0}^{L-1}\frac{\varepsilon_\ell}{q^\ell} + \frac{\varepsilon_L-1}{q^L} + \sum_{\ell=L+1}^{\infty}\frac{q-1}{q^\ell}\right).\end{aligned}$$

Furthermore, we have

$$\varphi(1) = 1 \quad \text{and} \quad \varphi(q) = F(q).$$

Using the function φ we can rewrite˜(8.17)

$$F(N) = F(q)^K\varphi\left(Nq^{-K}\right) = N^\rho e^{i\alpha\log_q N}F(q)^{-\{\log_q N\}}\varphi\left(q^{\{\log_q N\}}\right).$$

Setting

$$\psi(\log_q N) = F(q)^{-\{\log_q N\}}\varphi\left(q^{\{\log_q N\}}\right)$$

yields the desired result (notice that $\psi(0) = \psi(1)$). □

Remark 8.2.8 Notice that condition (8.14) is trivially satisfied for f taking only positive values. In this case the function ψ is the quotient of a monotonically increasing function and a differentiable function. It is therefore differentiable almost everywhere. Nevertheless, it is not the integral of its derivative (except for trivial cases like $f \equiv 1$). An explanation for this phenomenon is given by (Tenenbaum 1997).

The following corollary describes the behaviour of $F(N)$, if (8.14) is not satisfied.

Corollary 8.2.9 *Let f be a completely q-multiplicative function satisfying*

$$|F(q)| = |1 + f(1) + \ldots + f(q-1)| = \max_{0\le a<q}|f(a)| = M, \tag{8.19}$$

and $Q = \max_{1\le a<q}|F(a)|$. Then

$$|F(N)| \le Q(K+1)M^K \le QN^{\log_q M}(\log_q N + 1).$$

If

$$|F(q)| = |1 + f(1) + \ldots + f(q-1)| < \max_{0 \le a < q} |f(a)| = M, \tag{8.20}$$

then

$$|F(N)| \le \frac{QM^{K+1}}{M - |F(q)|} \le \frac{QM}{M - |F(q)|} N^{\log_q M}.$$

Proof Every summand on the right hand side of (8.17) can be estimated by $Q|F(q)|^\ell M^{K-\ell}$. Considering the two cases $M = |F(q)|$ and $M > |F(q)|$ gives the two assertions. □

Example 8.2.10 (Barbolosi and Grabner 1996) Let p be a prime and write n and k in base p. Then Lucas' congruence asserts that

$$\binom{n}{k} \equiv \binom{\varepsilon_0(n)}{\varepsilon_0(k)} \binom{\varepsilon_1(n)}{\varepsilon_1(k)} \cdots \binom{\varepsilon_L(n)}{\varepsilon_L(k)} \pmod{p}. \tag{8.21}$$

From this congruence it follows immediately that (see also˜(Stein 1989))

$$\# \left\{ 0 \le k \le n \mid \binom{n}{k} \not\equiv 0 \pmod{p} \right\} = \prod_{\ell=0}^{L} (1 + \varepsilon_\ell(n)),$$

which is a completely p-multiplicative function. Furthermore, for any multiplicative character χ modulo p the function

$$f_\chi(n) = \sum_{k=0}^{n} \chi\left(\binom{n}{k} \right) = \prod_{\ell=0}^{L} f_\chi(\varepsilon_\ell(n)),$$

is completely p-multiplicative.

Using Fourier analysis on the finite group $(\mathbb{Z}/p\mathbb{Z})^*$ we obtain for $a \not\equiv 0 \pmod{p}$

$$\# \left\{ 0 \le k \le n < N \mid \binom{n}{k} \equiv a \pmod{p} \right\} = \frac{1}{p-1} \sum_{\chi} \sum_{n<N} \overline{\chi}(a) f_\chi(n).$$

Since for $\chi \ne \chi_0$ (χ_0 denotes the principal character) $|F_\chi(p)| < F_{\chi_0}(p) = \frac{p(p-1)}{2}$ and $\max_{0 \le a < p} |f_\chi(a)| \le p-1$, the term for $\chi = \chi_0$ is the asymptotic main term. This implies that the binomial coefficients not divisible by p are uniformly distributed in the prime residue classes modulo p.

For $p = 3$ we have the simple formulæ (cf. also˜(Wolfram 1984))

$$\# \left\{ 0 \le k \le n \mid \binom{n}{k} \equiv 1 \pmod{3} \right\} = \frac{1}{2} 2^{c_1(n)} \left(3^{c_2(n)} + 1 \right)$$

and

$$\#\left\{0 \le k \le n \mid \binom{n}{k} \equiv 2 \pmod 3\right\} = \frac{1}{2}2^{c_1(n)}\left(3^{c_2(n)} - 1\right),$$

where $c_\varepsilon(n)$ $(\varepsilon = 1, 2)$ denotes the number of digits ε in the base 3 expansion of n.

Example 8.2.11 (Larcher 1996) The number of odd binomial coefficients in the n-th row of Pascal's triangle is given by $2^{s_2(n)}$, where $s_2(n)$ denotes the binary sum-of-digits function. This is just the special case $p = 2$ of the last example. In this case the corresponding periodic function ψ has been investigated further by (Harborth 1977). Since the function is singular, the minimum cannot be found by differential calculus (the maximum is easily found to be 1 and to be attained at integer points). (Larcher 1996) gives an algorithm, which finds arbitrarily good approximations to the minimum.

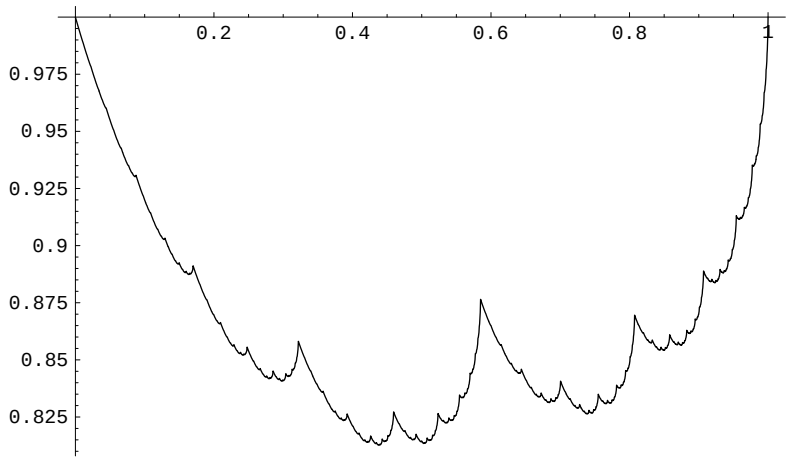

Fig. 8.2. Plot of the function ψ for the multiplicative function $2^{s_2(n)}$

Remark 8.2.12 The two examples show that the number of binomial coefficients $\binom{n}{k}$ not divisible by p in the region $0 \le k \le n < N$ is of order of magnitude N^{ρ_p} with $\rho_p = \log_p \frac{p(p+1)}{2} < 2$, which implies that "almost all" (in the sense of density) binomial coefficients are divisible by p. This fact and further results on the divisibility of binomial coefficients have been observed in (Singmaster 1974a), (Singmaster 1974b), and (Singmaster 1974c).

8.2.3 Divide and conquer recursions and Mellin-Perron techniques

A divide and conquer recursions is a relation of the form

$$a_n = \alpha\, a_{\lfloor n/2 \rfloor} + \beta\, a_{\lceil n/2 \rceil} + g_n \tag{8.22}$$

(with suitable initial conditions). Such kinds of recurrences appear in several applications in the analysis of algorithms, for example in the analysis of the number of comparisons in sorting networks (Bose and Nelson 1962), (Hwang 1998), the Karatsuba multiplication (Knuth 1981), or the Euclidean matching heuristic (Reingold and Tarjan 1981). The general scheme of all these algorithms is that to perform an operation (for instance, sorting) on a set of data of size n, at first the set is divided into parts of respective sizes $\lceil \frac{n}{2} \rceil$ and $\lfloor \frac{n}{2} \rfloor$. Then the algorithm is applied recursively to these smaller sets of data. The costs for merging the results for the original set of data are measured by the term g_n. A first approach to a unified study of such recurrences was done in (Flajolet and Golin 1993). This was extended further in (Hwang 1998).

Furthermore, several digital functions satisfy a relation of the form (8.22). For example, if we consider the summatory function $S(N) = \sum_{n<N} s(n)$, where $s(n)$ denotes the binary sum-of-digits function, then by using the relations $s(2k) = s(k)$ and $s(2k+1) = s(k) + 1$ we directly get

$$\begin{aligned} S(N) &= \sum_{2k<N} s(2k) + \sum_{2k+1<N} s(2k+1) \\ &= \sum_{k<\lceil N/2 \rceil} s(k) + \sum_{k<\lfloor N/2 \rfloor} (s(k)+1) \\ &= S(\lceil N/2 \rceil) + S(\lfloor N/2 \rfloor) + \lfloor N/2 \rfloor. \end{aligned}$$

It is not difficult to guess the (correct) growth rate of a_n if one knows the growth rate of g_n. However, it is usually a non-trivial problem to get precise asymptotic information on a_n. There is, however, a general method based on Dirichlet series. This method has its origin in classical analytic number theory (cf.˜(Tenenbaum 1995)). Set

$$A(s) = \sum_{n\geq 1} \frac{a_{n+1} - a_n}{n^s} \quad \text{and} \quad G(s) = \sum_{n\geq 1} \frac{g_{n+1} - g_n}{n^s} \tag{8.23}$$

Then, by distinguishing between odd and even numbers we get

$$\begin{aligned} A(s) &= \alpha \sum_{n\geq 1} \frac{a_{\lfloor (n+1)/2 \rfloor} - a_{\lfloor n/2 \rfloor}}{n^s} + \beta \sum_{n\geq 1} \frac{a_{\lceil (n+1)/2 \rceil} - a_{\lceil n/2 \rceil}}{n^s} + G(s) \\ &= \alpha \sum_{k\geq 0} \frac{a_{k+1} - a_k}{(2k+1)^s} + \beta \sum_{k\geq 1} \frac{a_{k+1} - a_k}{(2k)^s} + G(s) \\ = \frac{\alpha+\beta}{2^s} A(s) &+ \alpha(a_1 - a_0) + \alpha \sum_{k\geq 1} (a_{k+1} - a_k) \left(\frac{1}{(2k+1)^s} - \frac{1}{(2k)^s} \right) + G(s) \end{aligned}$$

which leads to the representation

$$A(s) = \frac{\alpha(a_1 - a_0) + F(s) + G(s)}{1 - (\alpha+\beta)2^{-s}}, \tag{8.24}$$

where

$$F(s) = \alpha \sum_{k\geq 1} (a_{k+1} - a_k) \left(\frac{1}{(2k+1)^s} - \frac{1}{(2k)^s} \right)$$

usually has a smaller abscissa of convergence than $A(s)$. Thus the expression (8.24) provides us with the analytic continuation of $A(s)$ to a larger domain together with information about the poles of $A(s)$.

Applying the Mellin-Perron summation formula (cf.˜(Tenenbaum 1995)) we obtain

$$a_n = a_1 + \sum_{k=1}^{n-1} (a_{k+1} - a_k) = a_1 + \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} A(s) \frac{n^s}{s}\, ds. \tag{8.25}$$

Here c has to be chosen large enough to make the series $A(s)$ absolutely (and therefore uniformly) convergent on the line $\Re(s) = c$. Usually, the integral in (8.25) is only convergent in the sense of a Cauchy principal value, which makes this analysis a bit delicate.

Using the analytic continuation of $A(s)$ makes it possible to deform the contour of integration to the left (in order to make the exponent of n occurring in the integral smaller). This is again a standard technique in analytic number theory. The problem when shifting the line of integration to the left comes from the convergence of the integral in (8.25). As a general fact about Dirichlet series, the growth along vertical lines becomes stronger for smaller values of $\Re(s)$ (cf.˜(Hardy and Riesz 1964)). Thus additional information on the growth of $F(s)$ and $G(s)$ along vertical lines is needed. This is usually the technically most elaborate step of this method.

If the convergence of the integral is proved, then (8.25) can be rewritten

as (for $c' < c$)

$$a_n = a_1 + \sum_{\substack{\text{poles of } A(s) \\ \text{with } c'<\Re(s)<c}} \operatorname{Res} \frac{n^s A(s)}{s} + \frac{1}{2\pi i} \int_{c'-i\infty}^{c'+i\infty} A(s) \frac{n^s}{s}\, ds. \qquad (8.26)$$

Usually, the poles mainly come from the zeroes of the denominator in (8.24), which are given by

$$s_k = \frac{\log(\alpha+\beta)}{\log 2} + \frac{2k\pi i}{\log 2}, \qquad k \in \mathbb{Z}.$$

The residues at these points take the form

$$\operatorname{Res}_{s=s_k} \frac{n^s A(s)}{s} = n^\rho \frac{e^{2k\pi i \log_2 n}}{s_k \log 2} \left(\alpha(a_1 - a_0) + F(s_k) + G(s_k)\right) \qquad (8.27)$$

with $\rho = \log_2(\alpha+\beta)$. The sum of these residues is then the main term in the asymptotic expansion of a_n and can be written as

$$a_n = a_1 + n^\rho \sum_{k\in\mathbb{Z}} \frac{e^{2k\pi i \log_2 n}}{s_k \log 2} \left(\alpha(a_1 - a_0) + F(s_k) + G(s_k)\right) + \mathcal{O}\left(n^{c'}\right)$$
$$= n^\rho H(\log_2 n) + \mathcal{O}\left(n^{c'}\right).$$

The series can be interpreted as the Fourier series of a periodic function H in the dyadic logarithm of n. This is the same periodicity phenomenon that we encountered in the study of additive and multiplicative functions before. The Fourier series usually converges only very slowly, which reflects the lack of smoothness of the function H. A collection of arguments, which can be used to prove growth estimates for Dirichlet series in the context of digital functions and divide-and-conquer recurrences as well as arguments for the convergence of the Fourier series of the occurring periodic functions is given in (Grabner and Hwang 2005). The fact that Dirichlet generating functions of q-regular functions have an analytic continuation to the whole complex plane, as well as information on their poles, was derived in (Allouche, Mendès France, and Peyrière 2000).

Example 8.2.13 For the binary sum-of-digits function $s_2(n)$, we have (cf.˜(Flajolet, Grabner, Kirschenhofer, et˜al. 1994))

$$\sum_{n=1}^{\infty} \frac{s_2(n) - s_2(n-1)}{n^s} = \frac{2^s - 2}{2^s - 1}\zeta(s),$$

where $\zeta(s)$ is the Riemann ζ-function. Using the according version of the

Mellin-Perron formula, we get

$$\sum_{n<N} s_2(n) = \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \frac{2^s-2}{2^s-1}\zeta(s)\frac{N^{s+1}}{s(s+1)}\,ds. \tag{8.28}$$

Since the growth of $\zeta(\sigma+it)$ for fixed σ and $|t| \to \infty$ is very well understood (cf.˜(Titchmarsh 1986)), the line of integration can be shifted to $\Re(s) = -\frac{1}{4}$ (it is known that $\zeta(-\frac{1}{4}+it) = \mathcal{O}(|t|^{3/4})$). Collecting residues (there is a double pole at $s=0$, which corresponds to the $N\log N$-term) gives

$$\begin{aligned}&\sum_{n<N} s_2(n)\\ &= \frac{1}{2}N\log_2 N - \frac{\log_2\pi}{2} - \frac{1}{2\log 2} - \frac{1}{4} - N\sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{\zeta(s_k)}{s_k(s_k+1)\log 2} e^{2k\pi i\log_2 N}.\end{aligned} \tag{8.29}$$

In this case it can be shown that the remainder term vanishes. This is exactly the Fourier series that we got by Delange's method before.

Remark 8.2.14 The vanishing of the remainder term in (8.29) comes from the fact that the integral

$$\frac{1}{2\pi i} \int_{-\frac{1}{4}-i\infty}^{-\frac{1}{4}+i\infty} \frac{2^s-2}{2^s-1}\zeta(s)\frac{N^{s+1}}{s(s+1)}\,ds$$

vanishes for $N \in \mathbb{N}$. There is a rather general theorem, which ensures the vanishing of of integrals occurring as remainder terms in this context (cf.˜(Hwang 1998) and for a slightly more general formulation (Grabner and Hwang 2005)).

In order to overcome the difficulties originating from the slow convergence of the integral in (8.25), in (Grabner and Hwang 2005) double differences and higher order Mellin-Perron formulæ were studied. Instead of the function $A(s)$ in (8.23) the function

$$\tilde{A}(s) = \sum_{n=1}^{\infty} \frac{a_{n+1}-2a_n+a_{n-1}}{n^s}$$

was used, which allows to compute a_n by the formula

$$a_n = na_1 + \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \tilde{A}(s)\frac{n^{s+1}}{s(s+1)}\,ds.$$

This formula gives a gain in convergence. On the other hand this gain has to be payed by more complicated expressions and the fact that the poles of $\tilde{A}(s)$ are one unit further to the left of the poles of $A(s)$, which makes the growth of $\tilde{A}(s)$ worse on vertical lines. Nevertheless, in many cases the double differencing technique gives easier proves for the convergence of the Mellin-Perron integrals.

Even if the integral in (8.25) is only conditionally convergent, in (Drmota, Grabner, and Liardet 2008) rather elaborate estimates could be used to obtain an analogue to Theorem˜8.3.17 for the summatory function of a block-multiplicative function on the Gaussian integers. In this case the method was applied to the Gaussian integers, which made alternative techniques still more difficult or even impossible to apply.

8.2.4 Generalisations

8.2.4.1 q-regular functions

The study of the summatory functions of q-regular functions follows the same line of ideas as the study of completely multiplicative functions. The only difference is that the products of scalars used there have to be replaced by matrix products. By the lack of commutativity, this makes the order of factors in all occurring products significant.

More precisely, let $f : \mathbb{N} \to \mathbb{R}$ be a real-valued q-regular function. Then by Definition˜8.1.2 and the discussion after the definition there exist functions $f = f_1, f_2, \dots, f_r$ and a map $\mathbf{M} : \{0, \dots, q-1\} \to \mathbb{R}^{r\times r}$ such that (8.2) holds. We write $\mathbf{f}(n) = (f_1(n), \dots, f_r(n))^T$ and use (8.2) to obtain

$$\mathbf{f}\left(\sum_{\ell=0}^{L} \varepsilon_\ell q^\ell\right) = \prod_{\ell=0}^{L} \mathbf{M}(\varepsilon_\ell)\mathbf{f}(0), \tag{8.30}$$

which allows to write

$$f(n) = \mathbf{v}_1 \prod_{\ell=0}^{L} \mathbf{M}(\varepsilon_\ell(n))\mathbf{v}_2$$

with $\mathbf{v}_1 = (1, 0, \dots, 0)$ and $\mathbf{v}_2 = \mathbf{f}(0)$.

Then by arguing along the same lines as in the proof of Theorem˜8.2.7 we get the following theorem.

Theorem 8.2.15 *Let $f : \mathbb{N} \to \mathbb{R}$ be a q-regular function and let $\mathbb{M}$ be the matrix function related to f by (8.2). Let $\mathbf{F} = \mathbf{M}(0) + \cdots + \mathbf{M}(q-1)$ and assume that $\mathbf{F}$ has a unique eigenvalue $\lambda > 0$ of maximal modulus and that*

this eigenvalue has algebraic multiplicity 1. *Assume further that*

$$\lambda > \max_{\varepsilon} \|\mathbf{M}(\varepsilon)\|$$

for some matrix norm $\|\cdot\|$. *Denote by* λ_2 *the modulus of the second largest eigenvalue. Then there exists a periodic continuous function* Φ *such that*

$$\sum_{n<N} f(n) = N^{\log_q \lambda}\Phi(\log_q N) + \mathcal{O}\left(N^{\log_q \lambda_2}\right) + \mathcal{O}(\log N).$$

Remark 8.2.16 The asymptotic behaviour of $\sum_{n<N} f(n)$ like a pure power of N corresponds to the fact that the part of $\mathbf{F}$ corresponding to λ is diagonalisable. If there are different Jordan-blocks occurring for λ, then powers of the logarithm occur in the asymptotic main terms. This happens, as can be seen from the result for q-additive functions.

Remark 8.2.17 The value $f(n)$ of a q-regular function in terms of the q-adic digits of n is given the matrix product (8.30). Since all possible finite sequences of digits occur as digital expansions of the positive integers, the question of finding the maximal growth order of $f(n)$ is related to extremal matrix products as studied in Chapter˜10.

8.2.4.2 q-automatic functions

Let $f(n)$ be an A-valued q-automatic function and $a \in A$. By Remark˜8.1.7 the indicator function $\mathbb{1}_{\{a\}}(f(n))$ can be expressed in terms of a matrix valued completely q-multiplicative function by (8.8). This makes the ideas developed before applicable for the computation of

$$F(N) = \sum_{n<N} \mathbb{1}_{\{a\}}(f(n)). \tag{8.31}$$

The question of existence of the limit $\lim_{N\to\infty} F(N)/N$, the density of the set $\{n \in \mathbb{N} \mid f(n) = a\}$ is of special interest in this context (cf.˜Remark˜8.2.20 below).

Applying the same reasoning as above we can prove

Theorem 8.2.18 *Let* $f(n)$ *be an* A*-valued* q*-automatic function and* $a \in A$. *Assume that* q *is the dominating eigenvalue of the matrix* M_δ *defined by* (8.9) *(i.e. all other eigenvalues have modulus* $< q$*). Then there is a continuous periodic function* Ψ *of period* 1 *such that*

$$F(N) = N\Psi(\log_q N) + o(N).$$

Corollary 8.2.19 *Under the assumptions of Theorem˜8.2.18 let* $\mathbf{A} = \lim_{k\to\infty} q^{-k}\mathbf{M}_\delta^k$. *Then the function* Ψ *is constant, if*

$$\mathbf{A}\mathbf{M}_\delta(0) = \mathbf{A}\mathbf{M}_\delta(1) = \ldots = \mathbf{A}\mathbf{M}_\delta(q-1) = \mathbf{Q} \tag{8.32}$$

and

$$\mathbf{Q}\mathbf{M}_\delta(a) = \mathbf{Q} \quad \text{for } a = 0, \ldots, q-1. \tag{8.33}$$

Proof Using the notation of Remark˜8.1.7 we can write for $K = \lfloor \log_q N \rfloor$

$$F(N) = q^K \mathbf{v}\mathbf{G}(Nq^{-K})\mathbf{w} + o(N),$$

where

$$\begin{aligned}\mathbf{G}\left(\sum_{j=0}^{\infty}\frac{\varepsilon_j}{q^j}\right) &= \mathbf{A}\left(\mathbf{M}(1) + \cdots + \mathbf{M}(q-1)\right)\\ &\quad + \sum_{j=1}^{\infty} q^{-j}\mathbf{A}\left(\mathbf{M}(0) + \cdots + \mathbf{M}(\varepsilon_j - 1)\right)\mathbf{M}(\varepsilon_{j-1})\cdots\mathbf{M}(\varepsilon_0)\\ &\quad + \frac{1}{q-1}\mathbf{A}\left(\mathbf{M} - \mathbf{M}(0)\right). \end{aligned} \tag{8.34}$$

Here and in the sequel we omit the subscript δ. This function is continuous on the interval $[1, q]$ and

$$\mathbf{G}(1) = \frac{1}{q-1}\mathbf{A}(\mathbf{M} - \mathbf{M}(0)) \quad \text{and } \mathbf{G}(q) = \frac{q}{q-1}\mathbf{A}(\mathbf{M} - \mathbf{M}(0)).$$

This proves the theorem.

The function Ψ in the theorem is constant, if $\mathbf{G}(x)$ is proportional to x for $x \in [1, q]$. Inserting the integer values $\{1, \ldots, q\}$ for x gives the conditions (8.32). Inserting $1 + \frac{a}{q}$ gives (8.33). Inserting these two conditions into (8.34) gives that $\mathbf{G}(x) = x\mathbf{Q}$ for $x \in [1, q]$. This proves the corollary. □

Remark 8.2.20 Corollary˜8.2.19 gives a condition for the existence of the density of the set

$$S = \{n \in \mathbb{N} \mid f(n) = a\}$$

for a given q-automatic function f. It is known (cf.˜(Allouche and Shallit 2003, Chapter˜8)) that the density does not always exist. It is also known that the logarithmic density of S

$$\lim_{N\to\infty} \frac{1}{\log N} \sum_{n<N} \frac{\mathbb{1}_S(n)}{n}$$

always exists (cf.˜(Allouche and Shallit 2003, Theorem˜8.4.8)) and that the two densities are equal, if the density exists.

Remark 8.2.21 In the proof of (Allouche and Shallit 2003, Theorem˜8.4.8) the logarithmic density is related to the residue of the Dirichlet series

$$\varphi_S(s) = \sum_{n=1}^{\infty} \frac{\mathbb{1}_S(n)}{n^s}$$

at $s = 1$. The logarithmic averaging process has the effect that the (possible) other poles of $\varphi_S(s)$ with $\Re(s) = 1$ (that correspond to the Fourier series of the periodic oscillation occurring in Theorem˜8.2.18 as was shown in Section˜8.2.3) can be disregarded. This corresponds to the fact that

$$\sum_{n<N} \frac{1}{n^{1+it}} = \begin{cases} \log N + \mathcal{O}(1) & \text{if } t = 0 \\ \mathcal{O}(1) & \text{if } t \neq 0, \end{cases}$$

which means that the logarithmic averaging singles out the Fourier coefficient of index zero (the mean of the periodic function).

8.2.4.3 Block-additive and block-multiplicative functions

Block-additive functions have been introduced in (Cateland 1992) as a more flexible generalisation of q-additive functions. Given a map $f : \{0, \dots, q-1\}^L \to \mathbb{R}$ with $f(0, \dots, 0) = 0$ the corresponding block-additive function is given by

$$s_f(n) = \sum_{k=0}^{\infty} f(\varepsilon_k(n), \dots, \varepsilon_{k+L-1}(n)). \tag{8.35}$$

Examples for such functions are block-counting functions. Block-additive functions are q-regular by the observation that the $\mathbb{R}$-module generated by the kernel is generated by the functions f and the functions $n \to f(b_1, b_2, b_r, \varepsilon_0(n), \dots, \varepsilon_{L-r-1}(n))$, $1 \leq r \leq L-2$, $b_1, \dots, b_r \in \{0, \dots, q-1\}$. As for q-additive functions, we can expect that the dominating eigenvalue occurring in the matrix $\mathbf{F}$ in Theorem˜8.2.15 has different algebraic and geometric multiplicity, which makes this theorem inapplicable.

A means to study block-additive functions, and also an object of study in their own right, are block-multiplicative functions given by a map $g : \{0, \dots, q-1\}^L \to \mathbb{R}$ with $g(0, \dots, 0) = 1$ and

$$m_g(n) = \prod_{k=0}^{\infty} g(\varepsilon_k(n), \dots, \varepsilon_{k+L-1}(n)). \tag{8.36}$$

Such functions are again q-regular.

A block-multiplicative function defines a $q^L \times q^L$-matrix $\mathbf{U}$ given by

$$u_{B,C} = \begin{cases} \frac{m_g(BC)}{m_g(C)} & \text{for } m_g(C) \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{for } B, C \in \{0, \ldots, q-1\}^L.$$

Here we have used the convention that m_g evaluated at a block of digits is the same as m_g evaluated at the number represented by that block. As usual BC denotes the concatenation of the blocks B and C. As for the (less explicit) matrix $\mathbf{F}$ in Theorem˜8.2.15, this matrix U allows to express $\sum_{n<N} f(n)$ in terms of sums of matrix products involving $\mathbf{U}$. The asymptotic behaviour of $\sum_{n<N} f(n)$ then depends on the dominating eigenvalue λ of $\mathbf{U}$ and its algebraic and geometric multiplicities. For a more detailed discussion we refer to (Barat and Grabner 2001).

Given a block additive function s_f, $\exp(ts_f(n))$ clearly defines a block-multiplicative function. A simple idea to study the moments of a block-additive function is to use

$$\sum_{n<N} s_f(n)^k = \left(\frac{d}{dt}\right)^k \left(\sum_{n<N} \exp(ts_f(n))\right)\Bigg|_{t=0}.$$

The most general theorem that was obtained in (Barat and Grabner 2001) gives an asymptotic formula for the summatory function of a product of several block-additive functions with a multiplicative function.

Proposition 8.2.22 *Let θ be a positive-valued block-multiplicative function and $f_1, \ldots, f_m$ arbitrary real-valued block-additive functions. Then the summatory function F of $\theta(n) f_1(n) \cdots f_m(n)$ satisfies*

$$\begin{aligned} F(N) &= \sum_{n<N} \theta(n) f_1(n) \cdots f_m(n) \\ &= N^{\log_q \lambda} \sum_{j=0}^{m} (\log_q N)^j \psi_j(\log_q N) + o(N^{\log_q \lambda_2}), \end{aligned}$$

where the functions ψ_j are continuous and periodic with period 1. λ and λ_2 are the eigenvalues of the matrix $\mathbf{U}$ corresponding to θ of largest and second largest modulus.

Remark 8.2.23 This result includes, for instance, moments of q-additive functions such as the sum-of-digits function (cf.˜(Coquet 1986)), digital functions occurring in the study of binomial coefficients with given divisibility by a prime power (cf.˜(Carlitz 1967) and Example˜8.2.25 below).

Remark 8.2.24 Since θ only attains positive values, the matrix U is fully

populated with positive entries, which by the Perron-Frobenius theorem ensures the existence of a dominating eigenvalue λ of multiplicity 1.

Example 8.2.25 (Barat and Grabner 2001) As an application of block-multiplicative functions, further distribution properties of binomial coefficients can be obtained. These use an extension of Lucas' congruence˜(8.21) to prime powers by (Granville 1997). This congruence involves digital blocks of length L for a congruence $\pmod{p^L}$. Results of the following kind could then be proved by applying summation formulæ for block-additive and block-multiplicative functions. For $(a,p)=1$ and $v_p(n)$ denoting the p-adic valuation (*i.e.* the highest power of p dividing n)

$$\begin{aligned}\#\left\{(k,n)\mid 0\le k\le n<N,\ v_p\left(\binom{n}{k}\right)=j,\ \text{and}\ \ p^{-j}\binom{n}{k}\equiv a \bmod p^{\ell}\right\}=\\ \frac{1}{\varphi(p^{\ell})}\#\left\{(k,n): 0\le k\le n<N \text{ and } \ v_p\left(\binom{n}{k}\right)=j\right\}+\mathcal{O}\left(N^{\beta}\right)=\\ \frac{1}{\varphi(p^{\ell})}N^{\alpha}\sum_{r=0}^{j}\psi_r^{(j)}(\log_p N)(\log_p N)^r+\mathcal{O}\left(N^{\beta}\right),\end{aligned} \tag{8.37}$$

where $\psi_r^{(j)}$ are continuous periodic functions of period 1 and $\beta<\alpha=\log_p\frac{p(p+1)}{2}$.

8.2.4.4 A measure-theoretic method for the analysis of digital functions

Looking back at the derivation of asymptotic formulas for the summatory functions of regular, multiplicative, or block-multiplicative functions, we observe that the expression for

$$F(q^k)=\sum_{n<q^k} f(n)$$

is usually much simpler and much simpler to obtain than the formula for general N. In (Grabner and Heuberger 2006) and (Grabner, Heuberger, and Prodinger 2005) a rather general technique has been developed, which can even be applied in multidimensional settings. This technique is based on the simple observation that the sequence of measures

$$\mu_k=\frac{1}{F(q^k)}\sum_{n<q^k} f(n)\delta_{nq^{-k}} \tag{8.38}$$

converges weakly to a limiting measure μ. (Here δ_x as usual denotes a unit point mass at x) Then the sum of $f(n)$ can be rewritten in terms of μ_k

$$\sum_{n<N} f(n) = F(q^k)\mu_k\left([0, Nq^{-k})\right),$$

where k has to be chosen such that $q^k > N$. If estimates for the error $|\mu([0,x)) - \mu_k([0,x))|$ are known, the right hand side can be rewritten as

$$\sum_{n<N} f(n) = F(q^k)\mu\left([0, Nq^{-k})\right) + o(F(q^k)).$$

Estimates for the difference of measures of intervals can be obtained by (versions of) the Berry-Esseen inequality, which estimates $|\mu([0,x)) - \mu_k([0,x))|$ in terms of the difference of the Fourier transforms of the measures. By the product structure of the functions f and $F(q^k)$, the Fourier transform of μ_k can be expressed as a product (of scalar or matricial functions).

We explain the technique by an example that is motivated by applications of digital expansions in cryptography. For a more detailed explanation of the background and for all the details left out in the exposition we refer to (Grabner and Heuberger 2006).

Every positive integer n can be represented in the form

$$n = \sum_{k=0}^{K} \varepsilon_k 2^k \text{ with } \varepsilon_k \in \{-1, 0, 1\}.$$

Adding the extra digit -1 introduces some freedom, which is used to minimise the number of non-zero digits (the "weight" of the representation). The weight corresponds to the number of additions needed for multiplication by n in an abelian group using Horner's scheme. In cryptographic applications, especially in elliptic curve cryptography, the number of operations needed for the computation of multiples is an important parameter (see for instance the discussion of optimal multiplication algorithms in (Cohen, Frey, Avanzi, et˜al. 2006)).

In (Heuberger and Prodinger 2006) it was shown that the automaton in Figure˜8.3 recognises all representations of minimal weight. We define $f(n)$ as the number of representations of n recognised by this automaton. By a careful investigation of the transitions in the automaton it can be proved that

$$f(n) = \mathcal{O}(n^\rho) \text{ with } \rho = \log_4\left(\frac{1+\sqrt{5}}{2}\right). \tag{8.39}$$

Remark 8.2.26 As was pointed out earlier, the question of determining

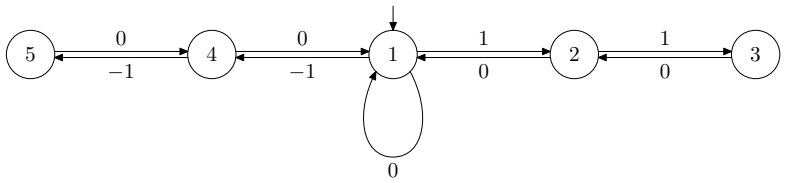

Fig. 8.3. Automaton recognising signed binary expansions of minimal weight from right to left. All states are terminal.

the maximal growth order of such functions is related to extremal matrix products as studied in Chapter˜10 and therefore is rather hard in general.

Let $f_n(k)$ denote the number of representations of an integer k of minimal weight and length at most n. Since any representation of minimal weight is at most 1 digit longer than the usual binary expansion, $f_n(k) = f_{\lfloor \log_2 |k| \rfloor + 2}(k) = f(k)$ for $n \geq \lfloor \log_2 |k| \rfloor + 2$. We define a sequence of measures by

$$\mu_n = \frac{1}{M_n} \sum_{k \in \mathbb{Z}} f_n(k) \delta_{k2^{-n}}, \tag{8.40}$$

where δ_x denotes the unit point mass concentrated in x and

$$M_n = \sum_{k \in \mathbb{Z}} f_n(k).$$

We notice that all points $k2^{-n}$ with $f_n(k) > 0$ lie in the interval $[-1, 1]$.

In order to compute the characteristic function of μ_n we consider the weighted adjacency matrix of the automaton in Figure˜8.3 (using the notation $e(t) = e^{2\pi i t}$):

$$A(t) = \begin{pmatrix} 1 & e(t) & 0 & e(-t) & 0 \\ 1 & 0 & e(t) & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & e(-t) \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

In the matrix $A(t)$ a transition with label ℓ is represented by an entry $e(\ell t)$. Then we have

$$\begin{aligned} \widehat{\mu}_n(t) &= \frac{1}{M_n} \sum_{k \in \mathbb{Z}} f_n(k) e\left(k2^{-n}t\right) \\ &= \frac{1}{M_n} v_1 A\left(t2^{-n}\right) A\left(t2^{-n+1}\right) \cdots A\left(t/2\right) v_2 \end{aligned} \tag{8.41}$$

with $v_1 = (1, 0, 0, 0, 0)$ and $v_2 = (1, 1, 1, 1, 1)^T$.

We notice that

$$M_n = (1,0,0,0,0)A(0)^n(1,1,1,1,1)^T = C\alpha^n + \mathcal{O}(|\alpha_2|^n), \tag{8.42}$$

where α and α_2 are the largest and second largest roots of the characteristic polynomial of $A(0)$ given by

$$(x-1)(x+1)(x^3-x^2-3x+1),$$

and $C = \frac{1}{37}(14\alpha^2+5\alpha-22)$, numerically

$$\alpha = 2.17009\ldots, \qquad \alpha_2 = -1.48119\ldots, \qquad C = 1.48055\ldots.$$

We will prove that (μ_n) weakly tends to a limit measure by showing that $\widehat{\mu}_n(t)$ tends to a limit $\widehat{\mu}(t)$.

Lemma 8.2.27 *The sequence of measures μ_n defined by* (8.40) *converges weakly to a probability measure μ. The characteristic functions satisfy the inequality*

$$|\widehat{\mu}_n(t) - \widehat{\mu}(t)| = \begin{cases} \mathcal{O}\left(|t|2^{-\eta n}\right) & \text{for } |t| \le 1 \\ \mathcal{O}\left(|t|^{\eta}2^{-\eta n}\right) & \text{for } |t| \ge 1 \end{cases} \tag{8.43}$$

with

$$\eta = \frac{\log\alpha - \log|\alpha_2|}{\log 2 + \log\alpha - \log|\alpha_2|} = 0.355251\ldots.$$

The constants implied by the $\mathcal{O}$-symbol are absolute.

Proof Equation (8.41) allows to express $\widehat{\mu}_n(t)$ in terms of matrix products. Standard analysis of such products allows to give estimates for $|\widehat{\mu}_m(t) - \widehat{\mu}_n(t)|$ for $n > m$, which give the desired estimates by letting n tend to infinity. The exponent η comes from a balancing argument used in an intermediate estimate. □

In the next lemma we prove continuity of the measure μ. Our study of the Fourier expansion of the periodic main term as well as the remainder term estimate in (8.10) will depend on the modulus of continuity given here.

Lemma 8.2.28 *The measure μ satisfies*

$$\mu([x,y]) = \mathcal{O}\left((y-x)^{\beta}\right) \tag{8.44}$$

for $\beta = \log_2\alpha - \log_4\varphi = 0.770632\ldots > \frac{1}{2}$.

Proof Every interval $[x, y]$ can be approximated by dyadic intervals $[2^{-n}\lfloor x2^n \rfloor, 2^{-n}\lceil y2^n \rceil]$. The measure μ of such intervals can be computed as the limit of the measures μ_k of these intervals. The estimate (8.39) is used as a trivial bound for the measure. □

In order to give an error bound for the rate of convergence of the measures μ_n to the measure μ, we will use the following version of the Berry-Esseen inequality, which was proved in (Grabner 1997).

Proposition 8.2.29 *Let μ_1 and μ_2 be two probability measures with their Fourier transforms defined by*

$$\widehat{\mu}_k(t) = \int_{-\infty}^{\infty} e^{2\pi i t x}\, d\mu_k(x), \quad k = 1, 2.$$

Suppose that $(\widehat{\mu}_1(t) - \widehat{\mu}_2(t))t^{-1}$ is integrable on a neighbourhood of zero and μ_2 satisfies

$$\mu((x, y)) \le c|x - y|^{\beta}$$

for some $0 < \beta < 1$. Then the following inequality holds for all real x and all $T > 0$

$$\begin{aligned} &|\mu_1((-\infty, x)) - \mu_2((-\infty, x))| \\ &\qquad \le \left| \int_{-T}^{T} \widehat{J}(T^{-1}t)(2\pi i t)^{-1} \left(\widehat{\mu}_1(t) - \widehat{\mu}_2(t)\right) e^{-2\pi i x t}\, dt \right| \\ &\quad + \left(c + \frac{1}{\pi^2}\right) T^{-\frac{2\beta}{2+\beta}} + \left| \frac{1}{2T} \int_{-T}^{T} \left(1 - \frac{|t|}{T}\right) \left(\widehat{\mu}_1(t) - \widehat{\mu}_2(t)\right) e^{-2\pi i x t}\, dt \right|, \end{aligned}$$

where

$$\widehat{J}(t) = \pi t(1 - |t|) \cot \pi t + |t|.$$

Lemma 8.2.30 *The measures μ_n satisfy*

$$|\mu_n((x, y)) - \mu((x, y))| = \mathcal{O}\left(2^{-\theta n}\right) \tag{8.45}$$

uniformly for all $x, y \in \mathbb{R}$ with $\theta = \frac{2\beta\eta}{\eta(\beta+2)+2\beta} = 0.2168\ldots$.

Proof We apply Proposition˜8.2.29 to the measures μ_n and μ. For this

purpose we use the inequalities (8.43) to obtain

$$|\mu_n((-\infty,x)) - \mu((-\infty,x))| \ll 2^{-\eta n}\int_{-1}^{1} dt + 2^{-\eta n}\int_{1\le|t|\le T} |t|^{\eta-1}\,dt$$
$$+ T^{-\frac{2\beta}{2+\beta}} + 2^{-\eta n}\frac{1}{T}\int_{-1}^{1}|t|\,dt + 2^{-\eta n}\frac{1}{T}\int_{1\le|t|\le T}|t|^{\eta}\,dt \ll 2^{-\theta n}$$

by choosing $T = 2^{\theta\frac{2+\beta}{2\beta}n}$. □

Putting everything together, we have found

$$\sum_{n<N} f(n) = M_k\mu_k([0,n2^{-k})) = C\alpha^k\mu([0,n2^{-k})) + \mathcal{O}(|\alpha_2|^k) + \mathcal{O}(\alpha^k 2^{-\theta k}) \tag{8.46}$$

with $k = \lfloor \log_2 N\rfloor + 2$. We set $\Phi(x) = C\alpha^{-x+2}\mu([0,2^{x-2}])$ for $0 \le x \le 1$ and observe that $\Phi(0) = \Phi(1)$ by the fact that we can also choose $= \lfloor \log_2 N\rfloor + 3$ by the discussion in the beginning. Thus we can write

$$\sum_{n<N} f(n) = N^{\log_2\alpha}\Phi(\{\log_2 N\}) + \mathcal{O}(N^{\log_2\alpha-\theta}). \tag{8.47}$$

Remark 8.2.31 The main ingredients for the method to work are the following

(i) a good understanding of the Fourier-transforms of $\widehat{\mu}_k$ and $\widehat{\mu}$, for instance in terms of products of scalar or matricial functions, which come from the underlying q-regular or block-multiplicative structure of the function f
(ii) an estimate for the difference $|\widehat{\mu}_k(t) - \widehat{\mu}(t)|$ (Lemma˜8.2.27)
(iii) an estimate for the measure-dimension of μ (Lemma˜8.2.28) is needed in the Berry-Esseen-type inequality (Proposition˜8.2.29). This estimate can be obtained by a priori estimates for $f(n)$ (like (8.39)), which can be tricky (cf.˜Chapter˜10).
(iv) in higher dimensional applications different versions of the Berry-Esseen inequality, for instance for balls in Euclidian space, are needed (cf.˜Proposition˜1 in (Grabner, Heuberger, and Prodinger 2005))

The method usually produces rather weak error terms, since the estimate for $|\widehat{\mu}_k(t) - \widehat{\mu}(t)|$ may not be best possible, and this estimate is pulled through the Berry-Esseen inequality, which needs one further balancing. Nevertheless, the method avoids the somehow intricate computations with the complicated explicit expressions for f needed for other approaches.

Remark 8.2.32 In (Okada, Sekiguchi, and Shiota 1995) suitably defined measures on $[0,1]$ were used to give exact formulæ for the moments of the binary sum-of-digits function. Their construction uses the measure μ related to the function $e^{ts_2(n)}$. Since in this case $\mu_k([0,a2^{-k}]) = \mu([0,a2^{-k}])$ the approximating measures μ_k are not needed, as well as the application of the Berry-Esseen argument.

Remark 8.2.33 The measures μ occurring in this context in many cases can be interpreted as distributions of infinite series of dependent random variables. The dependence is coded by the underlying matrix structure and is therefore of a Markov type. This relates the measures to Bernoulli convolutions (i.e. infinite series if independent random variables) such as studied by (Erdős 1939) and (Erdős 1940). For a survey on this subject we refer to (Peres, Schlag, and Solomyak 2000). Furthermore, we remark that all these measures are of pure type (either purely absolutely continuous, purely singular continuous, or consists only of point masses) by the Jessen-Wintner theorem (cf.˜(Elliott 1979, Lemma˜1.22)).

8.3 Statistics on digital functions

Let f be a q-additive function and define the shorthand-notation $f_j(n) = f(\varepsilon_j(n)q^j)$. Then $f(n) = \sum_{k=0}^{K} f_j(n)$ with $K = \lfloor \log_q n \rfloor$.

We will make now extensive use of the probabilistic interpretation of f as a random variable. As above the underlying probability space is $\Omega_N = \{0,1,\ldots,N-1\}$ with the uniform distribution, that is, every $n \in \Omega_N$ is equally likely.

The digits $\varepsilon_j(n)$ and also $f_j(n)$ are then random variables, too. However, the essential observation is that the digits $\varepsilon_j(n)$, $j \le K$, are almost independent (in what follows, we will make this more precise). Thus, we can consider $f(n)$ as a sum of K almost independent random variables. It is therefore not unexpected that several results from sums of independent random variables transfer to asymptotic and distributional properties of q-additive functions. Note that for $N = q^L$ for some positive integer L the digits $(\varepsilon_0(n),\ldots,\varepsilon_{L-1}(n))$ are actually independent.

We will first survey on general distributional results on (general) q-additive functions. In Section˜8.3.3 we focus on completely q-additive functions where we can get much more precise results by using a generating function approach. Note that completely q-additive functions correspond to sums of almost independent and identically distributed random variables. More precise results are thus not unexpected.

There are several different types of distribution results known for q-additive functions that can be unified to some extent:

(i) existence of an asymptotic distribution of the values of f on $\mathbb{R}$ (Erdős-Wintner-type theorems): Section˜8.3.1.3,
(ii) existence of a normal limit distribution for suitably renormalised values of f on $\mathbb{R}$ (central limit theorems): Section˜8.3.1.4,
(iii) some results of these kinds are also known for n ranging through subsequences of the integers, such as the values of a polynomial or the primes: Section˜8.3.2,
(iv) precise estimates for the number of n, where $f(n)$ attains a fixed value, for integer valued f (local limit theorems): Section˜8.3.3,
(v) uniform distribution of the values of f in a compact abelian group (usually $\mathbb{Z}/m\mathbb{Z}$ and $\mathbb{T}$): Section˜8.3.4.

8.3.1 General distributional results for additive functions

8.3.1.1 Approximation of digits by independent random variables

Our first goal is to make the statement that a q-additive function is a sum of almost independent random variables more precise. For this purpose we introduce an analogue to the number theoretic Kubilius model (see (Elliott 1979, Elliott 1980)) to the digital situation which was formulated by (Manstavičius 1997).

We start by considering infinite subsets of the non-negative integers $\mathbb{N} = \{0, 1, 2, \ldots\}$ that are defined by *digital restrictions*. For $0 \le d < q$ and $j \ge 0$ set

$$E_j(d) = \{n \in \mathbb{N} \mid \varepsilon_j(n) = d\}.$$

Furthermore, for every non-negative integer $k < q^{r+1}$ we consider the sets

$$K_r(k) = \bigcap_{j \le r} E_j(\varepsilon_j(k)) = \{n \in \mathbb{N} \mid n \equiv k \bmod q^{r+1}\}$$

that consist exactly those $n \in \mathbb{N}$ with $\varepsilon_j(n) = \varepsilon_j(k)$ for all $j \le r$. Note that the sets $K_r(k)$, $0 \le k < q^{r+1}$, are disjoint arithmetic progressions. It is clear that the algebra $\mathcal{F}_r$ of subsets of $\mathbb{N}$ generated by the sets $E_j(d)$ for $0 \le d < q$ and $j \le r$ are precisely sets of the form

$$A = \bigcup_{k \in I} K_r(k), \tag{8.48}$$

where I is any subset of $\{k \in \mathbb{N} \mid k < q^{r+1}\}$. Furthermore the asymptotic

density of the sets $K_r(k)$ in $\mathbb{N}$ equals q^{-r-1}. It is therefore natural to define the probability of $A \subset \mathcal{F}_r$ by

$$\mathbb{P}(A) = \mathbb{P}_r(A) = \frac{\#I}{q^{r+1}}.$$

By this definition $(\mathbb{N}, \mathcal{F}_r, \mathbb{P}_r)$ is a finite probability space, where the events $E_0(d_0), E_1(d_1), \ldots, E_r(d_r)$ are independent for any choice of numbers $0 \leq d_j < q$, $j \leq r$. Namely if we set $k_0 = d_0 + d_1 q + \cdots + d_r q^r$ then

$$\mathbb{P}_r\left(\bigcap_{0\leq j\leq r} E_j(d_j)\right) = \mathbb{P}_r\left(K_k(k_0)\right) = q^{-r-1} = \prod_{0\leq j\leq r} \mathbb{P}_r\left(E_j(d_j)\right).$$

Next observe that the function $f_j(n) = f(\varepsilon_j(n)q^j)$ just depends on the j-digit $\varepsilon_j(n)$ and is thus a $\mathcal{F}_r$-measurable function $f_j : \mathbb{N} \to \mathbb{R}$ (for $j \leq r$). Hence, it can be considered as a random variable Y_j. Due to the independence property the sets $E_j(d_j)$ the random variables $Y_0, Y_1, \ldots, Y_r$ are independent, too.

The following *Fundamental Lemma* that is due to (Manstavičius 1997) quantifies the difference between $\mathbb{P}_r$ and the counting measure.

Lemma 8.3.1 *Let $N \geq 1$ be given and set $K = \lfloor \log_q N \rfloor$. Then we have uniformly for all $r < K$ and all sets $A \in \mathcal{F}_r$*

$$\frac{1}{N}\#\left\{n < N \mid n \in A\right\} = \mathbb{P}_r(A) + O\left(\frac{q^r}{N}\right), \tag{8.49}$$

where the constant implied by the error term is universal.

Proof Let A be a set of the form (8.48) for some $r < K$. Since

$$\frac{1}{N}\#\{n < N \mid n \in K_r(k)\} = \left\lfloor \frac{N-k}{q^{r+1}} \right\rfloor + \theta_{r,k,N}$$

with $\theta_{r,k,N} \in \{0, 1\}$ we hence obtain

$$\begin{aligned}\frac{1}{N}\#\{n < N \mid n \in A\} &= \frac{1}{N}\sum_{k\in I}\left(\left\lfloor \frac{N-k}{q^{r+1}} \right\rfloor + \theta_{r,k,N}\right) \\ &= \frac{\#I}{q^{r+1}} + O\left(\frac{q^{r+1}}{N}\right) \\ &= \mathbb{P}_r(A) + O\left(\frac{q^r}{N}\right).\end{aligned}$$

□

In particular, if $A = \left\{ n \in \mathbb{N} \mid \sum_{j \le r} f_j(n) \in B \right\}$ (for some Borel set B) then we obtain

$$\frac{1}{N} \# \left\{ n < N \mid \sum_{j \le r} f_j(n) \in B \right\} = \mathbb{P}_r \left\{ n \in \mathbb{N} \mid \sum_{j \le r} f_j(n) \in B \right\} + O\left(\frac{q^r}{N} \right)$$
$$= \mathbb{P}_r \left\{ \sum_{j \le r} Y_j \in B \right\} + O\left(\frac{q^r}{N} \right) \qquad (8.50)$$

Note that (8.50) gives very precise bounds for partial sums $\sum_{j \le r} f_j(n)$ but not for $f(n)$. However, it is easy to extend the above model.

Let $\varepsilon_K(N) \ge 1$ denote the leading digit of N and let $\mathcal{F}$ be the algebra generated by $E_j(d)$ $(0 \le d < q,\, j < K)$ and $E_K(d)$ $(0 \le d \le \varepsilon_K(N))$, where we also set

$$\mathbb{P}(E_K(d)) = \frac{1}{\varepsilon_K(N)}, \qquad 0 \le d \le \varepsilon_K(N).$$

In this new probability space the K-th term $f_K(n) = f(\varepsilon_K(n) q^K)$ is also a random variable and $f(n)$ can be considered, too. Note also that $\mathbb{P}_r(A) = \mathbb{P}(A)$ for all $A \in \mathcal{F}_r$ and $r < K$. In particular if follows easily that for all $A \in \mathcal{F}$ (see (Manstavičius 1997))

$$\frac{1}{N} \# \{ n < N \mid n \in A \} \le 2\, \mathbb{P}(A).$$

Consequently

$$\frac{1}{N} \# \{ n < N \mid f(n) \in B \} \le 2\, \mathbb{P} \{ n \in \mathbb{N} \mid f(n) \in B \}.$$

8.3.1.2 A Turán-Kubilius inequality for additive functions

Let $f(n)$ denote a q-additive function and set

$$m_{j,q} := \mathbb{E}\, Y_j = \frac{1}{q} \sum_{d=1}^{q-1} f(dq^j),$$
$$m^2_{2;j,q} := \mathbb{E}\, Y_j^2 = \frac{1}{q} \sum_{d=1}^{q-1} f(dq^j)^2,$$

where $Y_j = f_j$ is the random variable related to the probability space $(\mathbb{N}, \mathcal{F}_r, \mathbb{P}_r)$ for some $j \le r$ and

$$M_q(N) := \sum_{j=0}^{\lfloor \log_q N \rfloor} m_{j,q}, \qquad D_q^2(N) = \sum_{j=0}^{\lfloor \log_q N \rfloor} \left(m^2_{2;j,q} - m^2_{j,q} \right).$$

Then the following general property holds which can be seen as an analogue of the celebrated Turán-Kubilius inequality (see (Kubilius 1964)) which has many applications in number theory.

Theorem 8.3.2 *Let f be a q-additive function. Then we have*

$$\frac{1}{N}\sum_{n<N}\left(f(n)-M_q(N)\right)^2 \le 2D_q^2(N). \tag{8.51}$$

Proof We use the relation

$$\mathbb{E}\,Y^2 = \int_0^\infty \mathbb{P}\{|Y|\ge u\}\,2u\,du.$$

Hence, if we the inequality

$$\begin{aligned}\frac{1}{N}\,&|\{n<N \mid |f(n)-M_q(N)|\ge u\}|\\ &\le 2\,\mathbb{P}\,\{n\in\mathbb{N} \mid |f(n)-M_q(N)|\ge u\}\end{aligned}$$

with respect to $2u\,du$ and apply the Burkholder inequality we obtain the proposed result (compare also with (Ruzsa 1984)). □

A direct application of Theorem˜8.3.2 is a very general property for the mean value of q-additive functions.

Corollary 8.3.3 *Let f be a q-additive function. Then we have*

$$\frac{1}{N}\sum_{n<N} f(n) = M_q(N) + O(D_q(N)).$$

Note that this corollary is in accordance with Theorem˜8.2.1. If f is completely q-additive then

$$M_q(N) = \big(\lfloor \log_q N\rfloor + 1\big)\,C_f \sim C_f \log_q N$$

and

$$D_q(N)^2 = \left(\frac{1}{q}\sum_{d=1}^{q-1} f(d)^2 - C_f^2\right)\log_q N.$$

8.3.1.3 An Erdős-Wintner theorem for additive functions

The above inequalities provide only a very rough idea of the overall behaviour of q-additive functions. We are now interested in conditions which ensure that the values of f have an asymptotic limiting distribution. In the context of classical additive functions Erdős and Wintner proved a necessary and sufficient condition for the existence of a limiting distribution (cf.˜(Elliott 1979)).

For additive functions the situation is very similar. (Delange 1972) could prove the following theorem, which is the analogue of Erdős' and Wintner's theorem for q-additive functions.

Theorem 8.3.4 *Let $f(n)$ be a q-additive function. Then $f(n)$ has a distribution function $G(y)$, that is*

$$\lim_{N\to\infty} \frac{1}{N} \#\left\{n < N \mid f(n) < y\right\} = G(y), \tag{8.52}$$

if and only if the two series

$$\sum_{j\geq 0}\sum_{d=1}^{q-1} f(dq^j) \quad \textit{and} \quad \sum_{j\geq 0}\sum_{d=1}^{q-1} f(dq^j)^2 \tag{8.53}$$

converge.

Proof The idea of the original proof of (Delange 1972) is to discuss convergence properties of q-multiplicative functions $F(n) = e(tf(n))$, in particular by using the identity

$$\sum_{n<q^L} F(n) = \prod_{j<L} \left(1 + \sum_{d=1}^{q-1} F(dq^j)\right).$$

Since $e(u) = 1 + 2\pi i u + O(u^2)$ for real u we have

$$\begin{aligned}\log\left(\frac{1+\sum_{d=1}^{q-1} F(dq^j)}{q}\right) &= \log\left(1 + 2\pi i t m_{j,q} + O\left(t^2 m_{2;j,q}^2\right)\right)\\ &= 2\pi i t m_{j,q} + O\left(t^2 m_{2;j,q}^2\right).\end{aligned}$$

Hence the limit

$$\lim_{L\to\infty} \frac{1}{q^L} \prod_{j<L} \left(1 + \sum_{d=1}^{q-1} F(dq^j)\right) = \prod_{j=0}^{\infty} \frac{1}{q} \left(1 + \sum_{d=1}^{q-1} e(tf(dq^j))\right) \tag{8.54}$$

exists if the two series (8.53) converge. The converse statement is also true.

Finally this easily extends to the convergence of

$$\frac{1}{N} \sum_{n\to N} F(n) = \frac{1}{N} \sum_{n\to N} e(tf(n)), \tag{8.55}$$

by comparing the sums with partial products of (8.54). By Lévy's criterion this is equivalent to the existence of a distribution function. □

Remark 8.3.5 The expression (8.54) for the characteristic function of the limiting distribution $G(y)$ shows that this distribution can be interpreted

as an infinite Bernoulli convolution. The Theorem of Jessen and Wintner asserts that such distribution measure given by $G(y)$ is either purely absolutely continuous, purely singular continuous, or consists only of point masses. A theorem of Lévy applied to the present setting asserts that the last alternative can only occur, if there exists a J such that $f(dq^j) = 0$ for $j > J$. In this case the distribution consists only of finitely many point masses. The two theorems cited are the contents of Lemma˜1.22 in (Elliott 1979).

Remark 8.3.6 A totally different proof of Theorem˜8.3.4 was given in (Barat and Grabner 2008). There the addition-by-one map τ is studied on the compact space $\mathbb{Z}_q = \operatorname{proj\,lim}_j \mathbb{Z}/q^j\mathbb{Z}$, which can be viewed as the compactification of $\mathbb{N}$ associated to the q-adic expansion. By Kolmogorov's three series theorem the conditions (8.53) are necessary and sufficient that the random series

$$\sum_{j=0}^{\infty} f(X_j q^j)$$

converges almost surely for $X_j \in \{0, \dots, q-1\}$ independent and identically uniformly distributed random variables (the convergence of the third series in the three series theorem is trivial in this case). In the setting of the dynamical system $(\mathbb{Z}_q, \tau)$ this simply means that f can be extended to an almost everywhere defined measurable function on $\mathbb{Z}_q$. Since $(\mathbb{Z}_q, \tau)$ is ergodic with respect to the Haar measure μ on $\mathbb{Z}_q$ Birkhoff's ergodic theorem (Theorem˜REF-BIRKHOFF) asserts that

$$\lim_{N\to\infty} \frac{1}{N} \#\{n < N \mid f(\tau^n(x)) < t\} = \mu\left(\{y \in \mathbb{Z}_q \mid f(y) < t\}\right) = G(t).$$

for μ-almost all $x \in \mathbb{Z}_q$. It remains to prove that 0 is one of the points for which this is valid (i.e. 0 is a *generic point*).

This point of view allows to generalise Delange's theorem to other types of digital expansions, such as expansions with linear recurrent base sequences, which involve dependent digits.

Finally we want to remark that there is also an alternative proof by (Manstavičius 1997) that uses the approximation properties of the form stated in Lemma˜8.3.1 and applies for Cantor expansions.

Remark 8.3.7 Theorem˜8.3.4 was generalised by (Kátai 1992) who proved that there exists a a distribution function $G(y)$ such that,

$$\lim_{N\to\infty} \frac{1}{N} \#\left\{n < N \mid f(n) - M_q(N) < y\right\} = G(y)$$

if and only if the series $\sum_{j\geq 0}\sum_{d=1}^{q-1} f(dq^j)^2$ converges.

Example 8.3.8 The q-additive function

$$v(n) = \sum_{j=0}^{\infty} \frac{\varepsilon_j(n)}{q^{j+1}}$$

defines the van˜der˜Corput sequence (cf.˜(Kuipers and Niederreiter 1974) and (Drmota and Tichy 1997)). It is easy to see that the distribution of this sequence is the uniform distribution on $[0,1]$. This sequence and related constructions are used in numerical integration to define sequences of low discrepancy, which give a small error in integration (cf.˜(Niederreiter 1992)).

8.3.1.4 A general central limit theorem for additive functions

Theorem˜8.3.4 and its variant by Kátai do not apply for completely q-additive function $f(n)$ or for functions where f_j does not converge to 0. In these cases we expect a central limit theorem which is ubiquitous in the context of sums of independent random variables.

The most general central limit theorem for q-additive functions is due to (Manstavičius 1997).

Theorem 8.3.9 *Suppose that, as $N \to \infty$,*

$$\max_{j\leq \log_q N} \max_{0\leq d<q} |f(dq^j)| = o(D_q(N)) \tag{8.56}$$

and $D_q(N) \to \infty$. Then,

$$\lim_{N\to\infty} \frac{1}{N} \# \left\{ n < N \mid \frac{f(n) - M_q(N)}{D_q(N)} < y \right\} = \Phi(y),$$

where Φ is the normal distribution function.

Proof For $N \geq 1$ let $(\mathbb{N}, \mathcal{F}, \mathcal{P})$ be the probability space constructed after the proof of Lemma˜8.3.1 for which the random variables $Y_j = f_j$, $0 \leq j \leq K = \lfloor \log_q N \rfloor$, are independent. For $r \leq K$ let

$$F_{N,r}(y) = \frac{1}{N} \left\{ n < N \mid \frac{\sum_{j\leq r} f_j(n) - M_q(N)}{D_q(N)} \leq y \right\},$$
$$V_{N,r}(y) = \mathbb{P} \left\{ n \in \mathbb{N} \mid \frac{\sum_{j\leq r} f_j(n) - M_q(N)}{D_q(N)} \leq y \right\}$$

denote the distribution functions of the normalised and truncated functions and

$$F_N(y) = F_{N,K}(y) \quad \text{and} \quad V_N(y) = V_{N,K}(y)$$

the distribution function of (normalised) $f(n)$ according to the counting measure and to the measure $\mathbb{P}$, respectively.

By the central limit theorem for sums of independent random variables (Billingsley 1968, Theorem 7.2) it is obvious that $V_N(y) \to \Phi(y)$, where $\Phi(y)$ denotes the distribution of the standard normal distribution, since the assumption (8.56) implies the Lindeberg condition.

Thus, is remains to show that $V_N(y)$ and $F_N(y)$ are close. For this purpose one can use the Lévy metric

$$L(F,G) = \inf\{\varepsilon > 0 \mid \forall y \in \mathbb{R} : \ F(y-\varepsilon) - \varepsilon \le G(y) \le F(y+\varepsilon) + \varepsilon\}.$$

between two distribution function F and G which quantifies and characterised weak convergence.

By the triangle inequality we have

$$L(F_N, V_N) \le L(F_N, F_{N,K-r}) + L(F_{N,K-r}, V_{N,K-r}) + L(V_{n,K-r}, V_N).$$

First by Lemma˜8.3.1 it follows for all $r > 0$

$$L(F_{N,K-r}, V_{N,K-r}) = O(q^{-r}).$$

Furthermore, for every $r > 0$ we obtain for every $\varepsilon > 0$ by another application of (8.56), as $N \to \infty$,

$$\begin{aligned} L(V_{n,K-r}, V_N) &\le \varepsilon + \mathbb{P}\left\{ n \in \mathbb{N} \mid \left| \sum_{K-r<j\le K} f_j(n) \right| \ge \varepsilon\, D_q(N) \right\} \\ &= \varepsilon + o_\varepsilon(1). \end{aligned}$$

A similar estimate holds for the distance $L(F_N, F_{N,K-r})$. Hence, we obtain $L(F_N, V_N) \to 0$ as $N \to \infty$ and consequently $L(F_N, \Phi) \to 0$. □

8.3.2 A central limit theorem for subsequences

The advantage of Theorem˜8.3.9 is its generality. However, it cannot be applied if we are dealing with certain subsequences of the integers, that is, the underlying probability space $\Omega_N = \{0, 1, \dots, N-1\}$ is replaced by a certain subset of integers, for example by $\Omega_N = \{2, 3, 5, \dots, p_N\}$, the first N primes, or by $\Omega_N = \{0^2, 1^2, 2^2, \dots, (N-1)^2\}$, the first N squares.

In this section we describe a general method that is due to (Bassily and Kátai 1995). In particular they could cover polynomial sequences and polynomial sequences of primes.

Theorem 8.3.10 *Let f be a q-additive function such that*

$$\sup_{j\geq 0}\max_{0\leq d<q} f(dq^j) = O(1)$$

Assume that $\frac{D_q(N)}{(\log N)^\eta} \to \infty$ as $N \to \infty$ for some $\eta > 0$ and let $P(x)$ be a polynomial with integer coefficients, degree r, and positive leading term. Then,

$$\lim_{N\to\infty}\frac{1}{N}\#\left\{n < N \mid \frac{f(P(n)) - M_q(N^r)}{D_q(N^r)} < y\right\} = \Phi(y)$$

and

$$\lim_{N\to\infty}\frac{1}{N}\#\left\{p < N \mid p \ \ prime, \quad \frac{f(P(p)) - M_q(N^r)}{D_q(N^r)} < y\right\} = \Phi(y).$$

In what follows we present the framework of their method in a slightly more general setting. We consider general subsets Ω_N of the non-negative integers of size N that satisfy the following property. It has a similar flavour as Lemma˜8.3.1, it quantifies the difference between the counting measure and and a measure with independent digits. However, it only needs properties of finitely many different digits.

Property 8.3.11 (BK-Property) *Let Ω_N be subsets of the non-negative integers of size N. We assume that $M_N = \max\Omega_N = O(N^k)$ for some $k \geq 1$ and that for every $\eta > 0$, $\lambda > 0$ and for every integer $L \geq 1$ we have*

$$\frac{1}{N}\#\{n \in \Omega_N \mid \varepsilon_{j_1}(n) = \ell_1, \ldots, \varepsilon_{j_L}(n) = \ell_L\} = q^{-L} + O\left((\log N)^{-\lambda}\right),$$

uniformly for all $j_1, \ldots, j_L$ with

$$(\log N)^\eta \leq j_1 < j_2 < \cdots < j_L \leq \log_q M_N - (\log N)^\eta$$

and for all $\ell_1, \ldots, \ell_L \in \{0, 1, \ldots, q-1\}$.

Note that the constant implied by the error term might depend on η, λ, and L.

Example 8.3.12 Let $\Omega_N = \{0, 1, \ldots, N-1\}$. Then the BK-Property is trivially satisfied. We actually have an error bound of the form

$$O\left(\frac{q^{j_L}}{N}\right) = O\left(q^{-(\log N)^\eta}\right) = O\left((\log N)^{-\lambda}\right),$$

since $j_L \leq \log_q N - (\log N)^\eta$ and $\eta > 0$.

The essential observation is that the BK-Property implies a central limit theorem.

Theorem 8.3.13 *Suppose that f satisfies the same assumptions as in Theorem˜8.3.10 and for every $N \geq 1$ let Ω_N be subset of the non-negative integers of size N.*

If the BK-Property, then we have

$$\lim_{N\to\infty} \frac{1}{N} \# \left\{ n \in \Omega_N \mid \frac{f(n) - M_q(M_N)}{D_q(M_N)} < y \right\} = \Phi(y).$$

Proof The idea of the proof is to compare moments, however, one has to be careful. We choose $0 < \eta' < \eta/(2k)$ (where $\eta > 0$ satisfies $D_(N)/(\log N)^\eta \to \infty$) and replace f by $\tilde{f}$ that is defined by

$$\tilde{f}(n) = \sum_{(\log N)^{\eta'} \leq j \leq \log M_N - (\log N)^{\eta'}} f(\varepsilon_j(n) q^j),$$

that is, we cut off some of the first and of the last digits. Let $\tilde{T}_N$ denote the random variable associated to $\tilde{f}$ and the counting measure on Ω_N, that is, the distribution function $F_{\tilde{T}_N}$ is given by

$$F_{\tilde{T}_N}(u) = \frac{1}{N} \# \left\{ n \in \Omega_n \mid \tilde{f}(n) \leq u \right\}.$$

Furthermore set

$$\tilde{S}_{M_N} = \sum_{(\log N)^{\eta'} \leq j \leq \log M_N - (\log N)^{\eta'}} Y_j,$$

where $Y_j = f_j$ are the independent random variables from above. We also set

$$\tilde{M}_q(M_N) = \mathbb{E}\, \tilde{S}_{M_N} = \sum_{(\log N)^{\eta} \leq j \leq \log M_N - (\log N)^{\eta}} m_{j;q}$$

$$\tilde{D}_q(M_N)^2 = \mathbb{V}\, \tilde{S}_{M_N} = \sum_{(\log N)^{\eta} \leq j \leq \log M_N - (\log N)^{\eta}} \left(m_{2,j;q}^2 - m_{j;q}^2 \right).$$

Note that by assumption $\tilde{D}_q(M_N) \sim D_q(M_N)$ as $N \to \infty$.

Next we expand the difference

$$\delta_L = \mathbb{E} \left(\tilde{T}_N - \tilde{M}_q(M_N) \right)^L - \mathbb{E} \left(\tilde{S}_N - \tilde{M}_q(M_N) \right)^L$$

in terms of the probabilities

$$\frac{1}{N} \# \{ n \in \Omega_N \mid \varepsilon_{j_1}(n) = \ell_1, \ldots, \varepsilon_{j_L}(n) = \ell_L \}$$

and compare them with help of the BK-Property. In fact, we have to take into account $\leq (q \log_q M_N)^L$ terms and, thus, we get

$$|\delta_L| = O\left((q \log_q M_N)^L (\log N)^{-\lambda} \right) = O\left((\log N)^{kL - \lambda} \right).$$

By assumption it follows that

$$\mathbb{E}\left(\frac{\tilde{T}_N - \tilde{M}_q(M_N)}{\tilde{D}_q(M_N)}\right)^L - \mathbb{E}\left(\frac{\tilde{S}_{M_N} - \tilde{M}_q(M_N)}{\tilde{D}_q(M_N)}\right)^L \to 0.$$

By standard tools in probability (see (Billingsley 1968)) we know that $(\tilde{S}_{M_N} - \tilde{M}_q(M_N))/\tilde{D}_q(M_N)$ converges to the Gaussian distribution $N(0,1)$ and we have convergence of all moments. Hence, the same is true for $(\tilde{T}_N - \tilde{M}_q(M_N))/\tilde{D}_q(M_N)$. Finally, since $f(n) - \tilde{f}(n) = O((\log N)^{\eta'})$, $M_q(N) - \tilde{M}_q(N) = O((\log N)^{\eta'})$ and $\tilde{D}_q(M_N) \sim D_q(M_N) \geq (\log N)^{2\eta'}$ we also deduce a central limit theorem for f. □

It remains to verify the BK-Property in several examples. Interestingly enough, proper exponential sum estimates are sufficient to derive this property.

Lemma 8.3.14 *Suppose that for all $\eta > 0$ and $\lambda > 0$ the exponential sum estimate*

$$\frac{1}{N}\sum_{n\in\Omega_N} e\left(\frac{a}{q^r}n\right) = O\left((\log N)^{-\lambda}\right) \tag{8.57}$$

holds uniformly for $(\log N)^\eta \leq r \leq \log_q M_N - (\log N)^\eta$, for all integers a with $1 \leq a < (\log N)^\lambda$ and for all integers a with $1 \leq a < q^r$ which are not divisible by q.

Then the BK-Property holds.

Proof We just sketch the idea of the proof. For a detailed analysis we refer to (Bassily and Kátai 1995).

Since $\varepsilon_j(n) = \ell$ if and only if $\{nq^{-j-1}\} = \ell/q$ we have

$$\frac{1}{N}\#\{n \in \Omega_N \mid \varepsilon_j(n) = \ell\} = \frac{1}{N}\sum_{n\in\Omega_N} \mathbb{1}_{[\ell/q,(\ell+1)/q)}(\{nq^{-j-1}\}),$$

where $\mathbb{1}_S$ denotes the indicator function of the set S. Let

$$\mathbb{1}_{[\ell/q,(\ell+1)/q)}(x) = \sum_{h\in\mathbb{Z}} a_h e(hx)$$

denote the Fourier series; note that $a_0 = 1/q$. Then it also follows that

$$\begin{aligned}\frac{1}{N}\#\{n \in \Omega_N \mid \varepsilon_j(n) = \ell\} &= \sum_{h\in\mathbb{Z}} a_h \frac{1}{N}\sum_{n\in\Omega_N} e\left(\frac{hn}{q^{j+1}}\right)\\ &= \frac{1}{q} + \sum_{h\neq 0} a_h \frac{1}{N}\sum_{n\in\Omega_N} e\left(\frac{hn}{q^{j+1}}\right).\end{aligned}$$

Hence, exponential sum estimates provide asymptotic information for the probabilities $\frac{1}{N}\#\{n \in \Omega_N \mid \varepsilon_j(n) = \ell\}$.

However, one has to be more precise since the Fourier series of the characteristic function is not absolutely convergent. In fact one can use smoothing arguments so that only few exponential sums are sufficient. Furthermore, this method extends to several digits (see (Bassily and Kátai 1995)). □

Example 8.3.15 By using standard estimates for exponential sums for polynomials and polynomials of primes (see (Iwaniec and Kowalski 2004)) it is clear that (8.57) is satisfied for the sets

$$\Omega_N = \{P(n) \mid n < N\}$$

and

$$\Omega_N = \{P(p_1), \dots, P(p_N)\},$$

where $p_1, \dots, p_N$ are the first N primes.

Example 8.3.16 Let $c > 1$ be a real non-integral number. and set

$$\Omega_N = \{\lfloor n^c \rfloor \mid n < N\}.$$

In this case we first observe that the digits $\varepsilon_0, \varepsilon_1, \dots$ coincide for $\lfloor n^c \rfloor$ and n^c in the q-ary digital expansion. Furthermore we have $\varepsilon_j(n^c) = \ell$ if and only if $\{n^c q^{-j-1}\} = \ell/q$. Thus, we can replace the exponential sums from (8.57) by the sums

$$\frac{1}{N} \sum_{n<N} e\left(\frac{a}{q^r} n^c\right).$$

For non-integral c these kinds of exponential sums can be easily estimated by Van der Corputs theorem (Iwaniec and Kowalski 2004, Theorem 8.20) and provide upper bounds which are of the same kind as those from (8.57). Hence, the BK-Property holds, too.

An important feature of the method of Bassily and Kátai is its flexibility. It also applies for so-called block additive functions as well as for other digital expansions. We will comment on this in Section˜8.4.

8.3.3 A generating function approach to completely q-additive functions

We have observed that due to the (almost) independence properties of the digits, a central limit theorem appears in very general situations.

We now concentrate on a very special situation, where we can obtain much more precise results. We discuss properties of the *generating function*

$$S(N,x) = \sum_{n<N} x^{f(n)},$$

where $x \neq 0$ is a complex variable. If we assume that f is integer valued then $S(N,x)$ can be rewritten as

$$S(N,x) = \sum_{k\in\mathbb{Z}} \#\{n < N \mid f(n) = k\}\, x^k$$

which explains the notion generating function. In fact, we will use this interpretation in the proof of Corollary˜8.3.21.

Obviously, the function $n \mapsto x^{f(n)}$ is a completely q-multiplicative function. Thus, we can apply the method of Theorem˜8.2.7 and obtain the following representation.

Theorem 8.3.17 *Suppose that f is a completely q-additive function and let $G \subseteq \mathbb{C}$ be defined by*

$$G = \left\{ x \in \mathbb{C} \mid \left|1 + x^{f(1)} + \cdots + x^{f(q-1)}\right| > \max_{0\le d<q} |x^{f(d)}| \right\}.$$

Then there exists a function $\Psi(x,t)$ ($x \in G$, $t \in \mathbb{R}$) that is analytic for $x \in G$ and Hölder continuous and periodic in t with period 1 such that

$$S(N,x) = \Psi(x, \log_q N)\,(1 + x^{f(1)} + \cdots + x^{f(q-1)})^{\log_q N}. \tag{8.58}$$

Furthermore there exists a continuous function $C(x)$ ($x \neq 0$) such that

$$|S(N,x)| \le C(x) \sum_{k\le \log_q N} \left|1 + x^{f(1)} + \cdots + x^{f(q-1)}\right|^k. \tag{8.59}$$

Proof The proof of (8.58) is just a refinement of the proof of Theorem˜8.2.7.

The proof of (8.59) is similar but even easier, since we are only interested in upper bounds, compare with Corollary˜8.2.9. □

It is an important feature of this lemma that the function $\Psi(x,t)$ represents an analytic function in x if x is sufficiently close to the real axis. It is interesting that Theorem˜8.3.17 has several corollaries (compare with (Drmota and Rivat 2005, Drmota, Grabner, and Liardet 2008)): We start with a representation for moments.

Corollary 8.3.18 *Suppose that f is a completely q-additive function. Then*

for every integer $r \geq 1$ we have

$$\frac{1}{N}\sum_{n<N} f(n)^r = C_f(\log_q N)^r + \sum_{\ell=0}^{r-1} \Psi_{r,\ell}(\log_q N)\cdot(\log_q N)^\ell, \tag{8.60}$$

where the functions $\Psi_{r,\ell}(t)$ $(0 \leq \ell < r)$ are continuous and periodic (with period 1).

Proof We just set $x = e^t$ in (8.58) and evaluate the r-th derivative (compare also with Proposition˜8.2.22). Furthermore, note that $\Psi(1,t) = C_f$. Hence, the asymptotic leading term is given by $C_f(\log_q N)^r$ and has no periodic fluctuations. □

Remark 8.3.19 The idea of taking the derivative also applies if a formula of the kind (8.58) is not exact but has an error term that is is uniform in a neighbourhood of $x = 1$. Due to analyticity in x one can take derivatives at $x = 1$ at arbitrary order by using the formula

$$G^{(r)}(1) = \frac{r!}{2\pi i}\int_{|x-1|=\delta} \frac{G(x)}{(x-1)^{r+1}}dx.$$

Next we derive a global and a local central limit theorem.

Corollary 8.3.20 *Suppose that f is a completely q-additive function and suppose that*

$$D_f^2 = \frac{1}{q}\sum_{d=1}^{q-1} f(d)^2 - C_f^2 > 0.$$

Then,

$$\lim_{N\to\infty}\frac{1}{N}\#\left\{n<N \mid \frac{f(n)-C_f\log_q N}{\sqrt{D_f^2\log_q N}} < y\right\} = \Phi(y). \tag{8.61}$$

and for all $r \geq 1$

$$\frac{1}{N}\sum_{n<N}\left(\frac{f(n)-C_f\log_q N}{\sqrt{D_f^2\log_q N}}\right)^r = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} u^r e^{-\frac{1}{2}u^2}\,du + o(1). \tag{8.62}$$

Furthermore, we have exponential tail estimates of the form

$$\begin{aligned}\frac{1}{N}\#\left\{n<N \mid \left|f(n)-C_f\log_q N\right| \geq y\sqrt{\log_q N}\right\} \\ \ll \min\left(e^{-cy}, e^{-cy^2+O(y^3/\sqrt{\log N})}\right)\end{aligned} \tag{8.63}$$

for some constant $c > 0$.

Proof Let T_N denote the random variable that is induced by the distribution of $f(n)$ on $\Omega_N = \{0, 1, \dots, N-1\}$. Then the moment generating function $\mathbb{E}\, e^{tT_N}$ is given by

$$\begin{aligned}\mathbb{E}\, e^{tT_N} &= \frac{1}{N}\sum_{n<N} e^{tf(n)} = \frac{1}{N} S(N, e^t) \\ &= \Phi(e^t, \log_q N)\left(1 + e^{tf(1)} + \cdots e^{tf(q-1)}\right)^{\log_q N}.\end{aligned}$$

Hence, by using the local expansion

$$\log\left(1 + e^{tf(1)} + \cdots e^{tf(q-1)}\right) = \log q + C_f t + \frac{D_f^2}{2} t^2 + O(t^3)$$

we obtain that the moment generating function of the normalised random variable

$$Z_N = \frac{T_N - C_f \log_q N}{\sqrt{D_f^2 \log_q N}}$$

is given by

$$\begin{aligned}\mathbb{E}\, e^{tZ_N} &= e^{t(C_f/D_f)\sqrt{\log_q N}}\, \mathbb{E}\, e^{(t/\sqrt{D_f^2 \log_q N})Y_N} \\ &= e^{\frac{1}{2}t^2 + O(t^3/\sqrt{\log N})}.\end{aligned}$$

Of course, this translates to (8.61).

Further, convergence of the moment generating function also implies convergence of all moments, that is, we get (8.62). Finally, the tail estimates (8.63) are a direct consequence of Chernov type inequalities. □

Corollary 8.3.21 *Suppose that f is an integer valued completely q-additive function such that*

$$d = \gcd\{f(c) \mid 0 \le c < q\} = 1. \tag{8.64}$$

Set

$$\mu(x) = \frac{x\lambda'(x)}{\lambda(x)} \quad \text{and} \quad \sigma^2(x) = \frac{x^2\lambda''(x)}{\lambda(x)} + \mu(x) - \mu(x)^2,$$

where $\lambda(x)$ abbreviates $\lambda(x) = 1 + x^{f(1)} + \cdots + x^{f(q-1)}$. Furthermore, for $k \in K(N) = \mathbb{Z} \cap \left[\delta \log_q N, (1-\delta)\log_q N\right]$ we define $x_{k,N}$ by $\mu(x_{k,N}) = k/\log_q N$, where $\delta > 0$ is arbitrary. Then we have uniformly for $k \in K(N)$

$$\begin{aligned}&\#\{n < N,\ f(n) = k\} \\ &\quad = \frac{\Phi(x_{k,N}, \log_q N)}{\sqrt{2\pi\sigma^2(x_{k,N})\log_q N}}\, N^{\log_q \lambda(x_{k,N})}\, x_{k,N}^{-k}\left(1 + O\left(\frac{1}{\log N}\right)\right)\end{aligned} \tag{8.65}$$

Furthermore, if $|k - C_f \log_q N| \le C\sqrt{\log_q N}$ *(for some* $C > 0$*) we also have*

$$\begin{aligned}&\#\{n < N,\ f(n) = k\}\\ &\quad = \frac{\pi N}{\sqrt{2\pi\sigma^2 \log_q N}} \exp\left(-\frac{(k - C_f \log_q N)^2}{2D_f^2 \log_q N}\right)\left(1 + O\left(\frac{1}{\sqrt{\log N}}\right)\right).\end{aligned} \tag{8.66}$$

Note that $C_f = \mu(1)$ and $D_f^2 = \sigma^2(1)$.

Proof We apply (8.58) and (8.59) and use Cauchy's formula:

$$\#\{n < N,\ f(n) = k\} = \frac{1}{2\pi i}\int_{|x|=x_{k,N}} \left(\sum_{n<N} x^{f(n)}\right) x^{-k-1}\,dx,$$

where $x_{k,N}$ is the saddle point of the asymptotic leading term of the integrand:

$$\lambda(x)^{\log_q N} x^{-k} = e^{\log\lambda(x)\cdot\log_q N - k\log x}.$$

We do not work out the details of standard saddle point techniques. We just refer to (Mauduit and Sárközy 1997) and (Drmota and Rivat 2005), where problems of almost the same kind have been discussed. □

8.3.4 Uniform distribution of q-additive functions

A last type of distribution results for additive functions is the distribution of values in a compact abelian group. Usually, the group under consideration is $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ or $\mathbb{Z}/m\mathbb{Z}$.

Definition 8.3.22 Let A be a compact abelian group equipped with its Haar measure λ and $f : \mathbb{N} \to A$ an A-valued arithmetic function. The sequence $(f(n))_{n\in\mathbb{N}}$ is called uniformly distributed, if for all measurable subsets $B \subseteq A$ with $\lambda(\partial B) = 0$

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n<N} \mathbb{1}_B(f(n)) = \lambda(B).$$

By harmonic analysis on the group A this is equivalent to saying that for all characters $\chi \in \widehat{A} \setminus \{\chi_0\}$ ($\chi_0 \equiv 1$ denotes the trivial character) one has

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n<N} \chi(f(n)) = 0 \tag{8.67}$$

(Weyl's criterion, cf.˜(Kuipers and Niederreiter 1974)).

For a q-additive function f the function $\chi \circ f$ is a $\mathbb{C}$-valued q-multiplicative function, for which we can apply Theorems˜8.2.7 and˜8.3.17 or Corollary˜8.2.9 to obtain the following corollaries.

Corollary 8.3.23 *Suppose that f is an integer valued completely q-additive function and that* (8.64) *holds. Then for every integer $M \geq 1$ and all $m \in \{0, 1, \ldots, M-1\}$ we have*

$$\frac{1}{N} \# \{n < N \mid f(n) \equiv m \bmod M\} = \frac{1}{M} + O(N^{-\eta})$$

for some $\eta > 0$.

Remark 8.3.24 Alternatively to condition (8.64) we can assume that f attains a value that is relatively prime to M. Then the same assertion holds.

Proof We use (8.59) for all M-th roots of unity $x = e^{2\pi i m/M}$ and apply simple discrete Fourier techniques. The exponent η comes from Corollary˜8.2.9 or Theorem˜8.2.7 as

$$\eta = 1 - \log_q \left(\max_{1 \leq h < M} \left| \sum_{\ell=0}^{q-1} e\left(\frac{h}{M} f(\ell) \right) \right| \right) > 0.$$

□

Corollary 8.3.25 *Let f be a real-valued completely q-additive function which attains one irrational value. Then the sequence $(f(n))_{n \geq 0}$ is uniformly distributed modulo* 1.

Remark 8.3.26 Note that Corollary˜8.3.25 in particularly applies to sequences of the kind $(\alpha f(n))_{n \geq 0}$ if f is integer valued and if α is irrational.

Proof We just set $x = e^{2\pi i h}$ for a non-zero integer h and use (8.59) to show that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n < N} e(hf(n)) = 0.$$

This is just Weyl's criterion (8.67) for the group $\mathbb{T}$. □

Corollary 8.3.27 *Let f be a real-valued completely additive function, which attains one irrational value and $\beta \in \mathbb{R} \setminus \mathbb{Q}$. Then the sequence* $(n\beta \pmod 1, f(n) \pmod 1)$ *is uniformly distributed in $\mathbb{T}^2$.*

Proof Simply realise that $n \mapsto n$ is q-additive and study the exponential sum

$$\sum_{n<N} e(h_1\beta n + h_2 f(n))$$

for $(h_1, h_2) \in \mathbb{Z}^2 \setminus \{(0,0)\}$ using the same ideas as in the proof of Theorem˜8.3.4. □

Remark 8.3.28 After the explanation of the probabilistic point of view in Section˜8.3.1.1 such results can also be seen as limiting distribution results for sums of independent random variables taking values in a compact group. For this point of view we refer to (Bártfai 1966).

Remark 8.3.29 As it was explained in Remark˜8.3.6 distribution results can also be seen from an ergodic point of view. Contrary to the situation explained there, the function f cannot be extended in a consistent way to $\mathbb{Z}_q$. The following idea originating from (Kamae 1977, Kamae 1978, Kamae 1987) defines a *cocycle* by

$$a_f(x,n) = \lim_{\substack{m\to x\\ m\in\mathbb{N}}} f(m+n) - f(n),$$

which is easily seen to exist almost everywhere in $\mathbb{Z}_q$. Then the dynamical system $(\mathbb{Z}_q \times A, T_a, \mu_q \otimes \lambda)$ (cf.˜Section˜**??**) with

$$T_a(x,\alpha) = (x+1, \alpha + a_f(x,1))$$

is used to study the distribution properties of f. Such systems are called *skew-products*. Then we have $T_a^n(0,0) = (n, f(n))$, which motivates the definition of T_a. By arguments explained in (Grabner and Liardet 1999) and (Drmota, Grabner, and Liardet 2008) a special property of the null-set, where a_f is not defined ("uniform negligibility", cf.˜(Liardet 1978)), is used to prove that all points are generic for the dynamical system, i.e. the ergodic theorem holds for all points, especially $(0,0)$.

This point of view has the advantage that it generalises to other situations, such as multidimensional settings. Furthermore, (Kamae 1987) used this idea to prove spectral disjointness for such dynamical systems with respect to multiplicatively independent bases. This was generalised to the Gaussian integers in (Grabner, Liardet, and Tichy 2005).

8.4 Further results

8.4.1 Gelfond Problems

In 1968 (Gelfond 1968) proved for the q-ary sum-of-digits function $s_q(n)$

$$\#\{n < N \mid n \equiv \ell \bmod r,\ s_q(n) \equiv a \bmod m\} = \frac{N}{mr} + O\left(N^{\lambda}\right)$$

provided that $\gcd(m, q-1) = 1$, where

$$\lambda = \frac{1}{2\log q} \log \frac{q \sin(\pi/2m)}{\sin(\pi/2mq)}.$$

This means that the sum-of-digits $s_q(n)$ is asymptotically uniformly distributed modulo m if we restrict n to arithmetic subsequences $n = \ell + kr, k \geq 0$. The proof is based on a subtle but elementary analysis of the expression

$$\sum_{n<q^K} e\left(\alpha n + \beta s_q(n)\right) = \prod_{r<K} \left(1 + e(\alpha 2^k + \beta) + \cdots + e((q-1)(\alpha 2^k + \beta))\right). \tag{8.68}$$

In this paper Gelfond formulated three problems. He first conjectured that a corresponding property is true if one uses two coprime bases q_1, q_2 at once, namely

$$\#\{n < N \mid s_{q_1}(n) \equiv a_1 \bmod m_1,\ s_{q_2}(n) \equiv a_2 \bmod m_2\} = \frac{N}{m_1 m_2} + O\left(N^{\lambda}\right) \tag{8.69}$$

provided that $\gcd(m_1, q_1 - 1) = 1$ and $\gcd(m_2, q_2 - 1) = 1$. Few years later (Bésineau 1972) proved this property, however, without an error term. Eventually (Kim 1999) provided also the proposed error term (even for a system of completely q-additive functions). It is interesting that these methods can be extended to non-trivial exponential sum estimates for

$$\sum_{n<N} e\left(\alpha n^k + \beta s_q(n)\right)$$

which were used in (Thuswaldner and Tichy 2005) to discuss Waring's problem under digital congruence conditions.

Gelfond also asked on the number of primes $p < N$ for which $s_q(p) \equiv a \bmod m$ and on the number $n < N$ for which $s_q(P(n)) \equiv a \bmod m$, where $P(x)$ is an integer polynomial. These challenging problems were unsolved for almost 40 years and only partial results have been obtained (see (Fouvry and Mauduit 1996a), (Fouvry and Mauduit 1996b), (Drmota and Rivat 2005), (Dartyge and Tenenbaum 2005), (Dartyge and Tenenbaum 2006)). Finally the problem for the subsequence of primes was completely solved by (Mauduit and Rivat 2009b). The problem for the subsequence of squares (more precisely on $s_q(n^2)$)

was solved by (Mauduit and Rivat 2009a). Interestingly the approach of Mauduit and Rivat again uses a subtle analysis of the product representation (8.68) combined with Fourier analytic tools and tricky but classical exponential sum techniques. Basically they showed that

$$\sum_{p<N} e(\alpha s_q(p)) \ll N^{1-\eta} \tag{8.70}$$

and

$$\sum_{n<N} e(\alpha s_q(n^2)) \ll N^{1-\eta}, \tag{8.71}$$

where $\eta = \eta(\alpha) > 0$ if $\alpha(q-1) \notin \mathbb{Z}$. With the help of these estimates one obtains asymptotic distributions in residue classes (as asked by Gelfond) and also uniform distribution modulo 1 for the sequences $(\alpha s_q(p))$ and $(\alpha s_q(n^2))$ for irrational α.

There are already some extensions of these results (see (Drmota, Rivat, and Stoll 2008), (Drmota, Mauduit, and Rivat 2009)). The main open problem in this context is to generalise (8.71) to polynomials $P(x)$ of degree ≥ 3.

8.4.2 Odometers and systems of numeration

The q-adic digital representation presented here is one special case of a rather general definition of numeration system. Every strictly increasing sequence of positive integers $(G_k)_{k\in\mathbb{N}}$ with $G_0 = 0$ gives rise to a representation all $n \in \mathbb{N}$. Every $n \in \mathbb{N}$ can be written in the form

$$n = \sum_{k=0}^{K} \varepsilon_k(n) G_k \quad \text{with } \forall k : 0 \leq \varepsilon_k(n) < \frac{G_{k+1}}{G_k}.$$

The additional requirement

$$\forall k : \varepsilon_0(n)G_0 + \cdots \varepsilon_k(n)G_k < G_{k+1} \tag{8.72}$$

makes this representation unique. The representation satisfying (8.72) can be determined by the *greedy algorithm*. The main difference between this general notion of digital expansion and the q-adic case presented here is the dependence between the digits given by (8.72).

One possible approach to extend distribution results of various kinds to

this more general setting is to define an according compactification of $\mathbb{N}$ by

$$\mathcal{K}_G = \{(\varepsilon_0, \varepsilon_1, \ldots) \mid \forall k : \varepsilon_0 G_0 + \cdots \varepsilon_k G_k < G_{k+1}\} \subseteq \prod_{k=0}^{\infty} \left\{0, 1, \ldots, \left\lceil \frac{G_{k+1}}{G_k} \right\rceil - 1\right\}.$$

This space is equipped with the product topology of the discrete spaces and therefore compact. All representation of positive integers are then in $\mathcal{K}_G$ by (8.72). The addition-by-one map τ on $\mathbb{N}$ can then be extended to $\mathcal{K}_G$ by

$$\tau(x) = \lim_{\substack{n \to x \\ n \in \mathbb{N}}} \tau(n).$$

Under additional conditions on $(G_k)_{k\in\mathbb{N}}$ there exists a unique τ-invariant measure μ_G on $\mathcal{K}_G$ (cf.˜(Barat, Downarowicz, and Liardet 2002)). The properties of the dynamical system $(\mathcal{K}_G, \tau, \mu_G)$ (the *odometer*) have been studied from combinatorial, topological, and dynamical point of view by (Grabner, Liardet, and Tichy 1995), (Barat, Downarowicz, Iwanik, et˜al. 2000), and (Barat, Downarowicz, and Liardet 2002).

A different point of view was taken in (Lecomte and Rigo 2001), where a regular language $\mathcal{L}$ on an ordered alphabet was used to define numeration: the ordering on the alphabet induces the genealogical ordering (see Definition˜**??**) on the language $\mathcal{L}$, and the positive integer n is then represented by the n-th word in the language $\mathcal{L}$. For a detailed description of this numeration we refer to Chapter˜2

This generalises the numeration systems with linear recurrent base sequence. Again additive functions on such numeration systems can be defined. In (Grabner and Rigo 2003) it was shown that theorems analogous to Theorem˜8.2.1 do not hold in this very general setting, but only under additional combinatorial assumptions on the language $\mathcal{L}$. Furthermore, in (Grabner and Rigo 2007) limiting distributions of additive functions on regular languages were studied. Again these distributions exist only under additional assumptions on the language $\mathcal{L}$; there are cases, where the limiting distribution is not Gaussian. In (Berthé and Rigo 2007) a compactification of $\mathbb{N}$ is constructed from this type of number representations, and the according odometer is studied.

Another different approach to numeration with respect to (certain) linear recurring sequences uses substitutions. Let σ be primitive substitution on the alphabet A such that for some $a \in A$, a is a prefix of $\sigma(a)$. A sequence of words $m_1, \ldots, m_k$ is called a-admissible, if there exist letters $a = a_0, a_1, \ldots, a_k$ such that $m_i a_i$ is a prefix of $\sigma(a_{i-1})$ for $i = 1, \ldots, k$. Then

every positive integer n can be represented by an a-admissible sequence of words $m_1, \dots, m_k$ satisfying

$$n = |\sigma^{k-1}(m_k)| + \cdots + |\sigma^0(m_1)|.$$

Notice that by definition the length of the occurring words m_i is bounded. The words $m_1, \dots, m_k$ are considered as the digits in this representation. Additive functions are then defined as

$$s_f(n) = \sum_{\ell=1}^{k} f(m_\ell).$$

In a series of papers J.-M.˜Dumont and A.˜Thomas (Dumont 1990), (Dumont and Thomas 1991), (Dumont and Thomas 1993), and (Dumont and Thomas 1997) derived analogues of the theorems in Sections˜8.2 and 8.3.1 for this notion of additive functions.

The study of digital functions in the context of harmonic analysis dates back to (Mahler 1927) and (Wiener 1927). They computed, what one would call today the Fourier coefficients of the spectral measure associated to a dynamical system given by the sum-of-digits function. Their work was then continued by M.˜Mendès France, J.˜Coquet, and P.˜Liardet who worked out the aspect of dynamical systems and uniform distribution in a series of papers (Mendès France 1967), (Mendès France 1971), (Mendès France 1973), (Mendès France 1974), (Coquet and Mendès France 1977), (Coquet, Kamae, and Mendès France 1977), (Liardet 1978), and (Coquet 1979). As overviews over this aspect we refer to (Queffélec 1987) and (Barat, Berthé, and Liardet 2006).

8.4.3 Distributional results for general numeration systems

Following A.˜O.˜Gelfond's question (8.69) on the joint distribution of the sum-of-digits functions $s_{q_1}(n)$ and $s_{q_2}(n)$ (for coprime bases q_1, q_2) it is natural to aks on the joint distribution of a q_1-additive function $f(n)$ and a q_2-additive function $g(n)$.

It turns out that Theorem˜8.3.4 directly extends to several pairwise coprime bases. For example one has

$$\lim_{N\to\infty} \frac{1}{N} \#\{n < N \mid f(n) < y_1,\, g(n) < y_2\} = G(y_1)G(y_2)$$

for certain distribution function $G_1(y)$, $G_2(y)$ if and only if the corresponding series (8.53) for $f(n)$ and $g(n)$ converge. This was observed by Hildebrandt (personal communication), the only additional ingredient for the proof is the Chinese remainder theorem.

Interestingly the general central limit theorem (Theorem˜8.3.9) has no direct analogue. The reason is that the Fundamental Lemma˜8.3.1 only generalises (with the help of the Chinese remainder theorem) to a property of the kind

$$\frac{1}{N}\#\{n < N \mid n \in A\} = \mathbb{P}_{r_1,r_2}(A) + O\left(\frac{q_1^{r_1}q_2^{r_2}}{N}\right), \tag{8.73}$$

where A is a set depending on the first r_1 q_1-ary digits and on the first r_2 q_2-ary digits of n and $\mathbb{P}_{r_1,r_2}$ is the *natural measure* for these sets. Thus (8.73) just provides a proper approximation for half the range. Nevertheless, the BK-Property˜8.3.11 generalises to two (coprime) expansions, see (Drmota 2001), (Drmota, Fuchs, and Manstavičius 2003). In particular, Theorem˜8.3.10 has a bivariate extension. In contrast to the Fundamental Lemma this method is based on exponential sum estimates of the form of Lemma˜8.3.14, where one can apply Bakers's theory on linear forms of logarithms of algebraic numbers.

Another problem is to generalise distributional results (Theorems˜8.3.4, 8.3.9, 8.3.10) to numeration systems $(G_k)_{k\in\mathbb{N}}$ that have been described in Section˜8.4.2. One of the easiest extensions of the q-ary system is the Cantor system, where $G_k = q_1q_2\cdots q_k$ with integers $q_j \geq 2$. Here the digits $\varepsilon_j(n)$ can be independently chosen from the sets $\varepsilon_j(n) \in \{0, 1, \ldots, q_j - 1\}$. This independence property gives also rise to a corresponding Kubilius model (compare with Section˜8.3.1.1) so that Theorems˜8.3.4 and 8.3.9 directly extend to the Cantor case provided that the q_j are uniformly bounded.

Numeration systems $(G_k)_{k\in\mathbb{N}}$, where the sequence G_k satisfies a linear recurrence

$$G_k = a_1G_{k-1} + a_2G_{k-2} + \cdots + a_dG_{k-d}, \quad k \geq d, \tag{8.74}$$

with constant coefficients $a_1, \ldots, a_d$ are also very well studied. The most prominent one is the Zeckendorf system ($d = 2$, $a_1 = a_2 = 1$) that is based on the Fibonacci numbers. By assuming that the coefficients satisfy the relations

$$(a_j, a_{j+1}, \ldots, a_d) \leq (a_1, a_2, \ldots, a_{d-j+1}), \quad 2 \leq j \leq d,$$

where $\leq$ denotes the lexicographic order, then every non-negative integer n has the unique (greedy) expansion $n = \sum_{j\geq 0}\varepsilon_j(n)G_j$ with digits $\varepsilon_j(n)$ if and only if

$$(\varepsilon_k(n), \varepsilon_{k-1}(n), \ldots) < (a_1, a_2, \ldots, a_k), \quad k \geq 0.$$

This already shows that the digits are not independent. However, this system is closely related to a digital representation associated to substitutions

and is thus related to a Markov process (see (Dumont and Thomas 1997)). In particular, this implies that a completely additive function (related to such a numeration system) satisfies a central limit theorem. Theorems similar to Theorem˜8.3.4 were proved in (Barat and Grabner 1996) and (Barat and Grabner 2008).

However, if one additionally assumes that the characteristic polynomial of the recurrence (8.74) is irreducible and that its dominant root is a Pisot unit then there is an analogue of Theorem˜8.3.10 (see (Drmota and Steiner 2002) (Steiner 2002)). For example, if $a_1 \geq a_2 \geq \cdots \geq a_d = 1$ then these assumptions are satisfied. The essential point is again a proper variant of the BK-Property˜8.3.11 which can be proved with the help of exponential sum estimates. However, at this stage one has to use an interesting relation to Rauzy fractals. In the q-ary case the q-ary digit $\varepsilon_j(n)$ can be determined by considering the fractional part of n/q^{j+1}, that is, $\varepsilon_j(n) = d$ if and only if $\{n/q^{j+1}\} \in [d/q, (d+1)/q)$. In the Pisot case there is a tiling $(T_d)_{0\leq d\leq a_1}$ of $\mathbb{R}^d$ that is deduced from the Rauzy fractal related to the α-shift with the property that

$$\operatorname{dist}\left(\mathbf{v}(n,k), T_{\varepsilon_k(n)}\right) = O(\alpha^{-k}),$$

where

$$\mathbf{v}(n,k) = \frac{n}{\alpha^k}\frac{\alpha-1}{\alpha^d-1}\left(\alpha^{d-1},\ldots,\alpha,1\right).$$

This means that the digits $\varepsilon_k(n)$ can be almost determined by looking at n/α^k modulo the tiling. Thus a Fourier series approach similarly to that of Lemma˜8.3.14 applies.

References

[Allouche and Shallit 2003] Allouche, J.-P. and Shallit, J.˜O. *Automatic Sequences, Theory, Applications, Generalizations*. Cambridge University Press, 2003.

[Allouche, Mendès France, and Peyrière 2000] Allouche, J.-P., Mendès France, M., and Peyrière, J. Automatic Dirichlet series. *J. Number Theory* **81**, (2000) 359–373.

[Barat and Grabner 1996] Barat, G. and Grabner, P.˜J. Distribution properties of G-additive functions. *J. Number Theory* **60**, (1996) 103–123.

[Barat and Grabner 2001] Barat, G. and Grabner, P.˜J. Distribution of binomial coefficients and digital functions. *J. London Math. Soc. (2)* **64**, (2001) 523–547.

[Barat and Grabner 2008] Barat, G. and Grabner, P.˜J. Limit distribution of Q-additive functions from an ergodic point of view. *Ann. Univ. Sci. Budapest. Sect. Comput.* **28**, (2008) 55–78.

[Barat, Berthé, and Liardet 2006] Barat, G., Berthé, V., and Liardet, P. Dynamical directions in numeration. *Ann. Inst. Fourier (Grenoble)* **56**, (2006) 1987–2092.

[Barat, Downarowicz, and Liardet 2002] Barat, G., Downarowicz, T., and Liardet, P. Dynamiques associées à une échelle de numération. *Acta Arith.* **103**, (2002) 41–78.

[Barat, Downarowicz, Iwanik, et˜al. 2000] Barat, G., Downarowicz, T., Iwanik, A., and Liardet, P. Propriétés topologiques et combinatoires des échelles de numération. *Colloq. Math.* **84/85**, (2000) 285–306. Dedicated to the memory of Anzelm Iwanik.

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

[Bártfai 1966] Bártfai, P. Grenzverteilungssätze auf der Kreisperipherie und auf kompakten Abelschen Gruppen. *Studia Sci. Math. Hungar.* **1**, (1966) 71–85.

[Bassily and Kátai 1995] Bassily, N.˜L. and Kátai, I. Distribution of the values of q-additive functions on polynomial sequences. *Acta Math. Hung.* **68**, (1995) 353–361.

[Berthé and Rigo 2007] Berthé, V. and Rigo, M. Odometers on regular languages. *Theory Comput. Syst.* **40**, (2007) 1–31.

[Bésineau 1972] Bésineau, J. Indépendance statistique d'ensembles liés à la fonction "somme des chiffres". *Acta Arith.* **20**, (1972) 401–416.

[Billingsley 1968] Billingsley, P. *Convergence of probability measures*. John Wiley

& Sons Inc., New York, 1968.

[Bose and Nelson 1962] Bose, R.˜C. and Nelson, R.˜J. A sorting problem. *J. Assoc. Comput. Mach.* **9**, (1962) 282–296.

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

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

[Cohen, Frey, Avanzi, et˜al. 2006] Cohen, H., Frey, G., Avanzi, R., Doche, C., Lange, T., Nguyen, K., and Vercauteren, F. *Handbook of elliptic and hyperelliptic curve cryptography*. Discrete Mathematics and its Applications (Boca Raton). Chapman & Hall/CRC, Boca Raton, FL, 2006.

[Coquet and Mendès France 1977] Coquet, J. and Mendès France, M. Suites à spectre vide et suites pseudo-aléatoires. *Acta Arith.* **32**, (1977) 99–106.

[Coquet, Kamae, and Mendès France 1977] Coquet, J., Kamae, T., and Mendès France, M. Sur la mesure spectrale de certaines suites arithmétiques. *Bull. Soc. Math. France* **105**, (1977) 369–384.

[Coquet 1979] Coquet, J. Sur la mesure spectrale des suites q-multiplicatives. *Ann. Inst. Fourier (Grenoble)* **29**, (1979) 163–170.

[Coquet 1986] Coquet, J. Power sums of digital sums. *J. Number Theory* **22**, (1986) 161–176.

[Dartyge and Tenenbaum 2005] Dartyge, C. and Tenenbaum, G. Sommes des chiffres de multiples d'entiers. *Ann. Inst. Fourier (Grenoble)* **55**(7), (2005) 2423–2474.

[Dartyge and Tenenbaum 2006] Dartyge, C. and Tenenbaum, G. Congruences de sommes de chiffres de valeurs polynomiales. *Bull. London Math. Soc.* **38**(1), (2006) 61–69.

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

[Delange 1975] Delange, H. Sur la fonction sommatoire de la fonction "somme des chiffres". *Enseign. Math.* **21**, (1975) 31–47.

[Drmota and Rivat 2005] Drmota, M. and Rivat, J. The sum-of-digits function of squares. *J. London Math. Soc.* **72**(2), (2005) 273–292.

[Drmota and Steiner 2002] Drmota, M. and Steiner, W. The Zeckendorf expansion of polynomial sequences. *J. Théor. Nombres Bordeaux* **14**, (2002) 439–475.

[Drmota and Tichy 1997] Drmota, M. and Tichy, R.˜F. *Sequences, Discrepancies, and Applications*, vol. 1651 of *Lecture Notes in Mathematics*. Springer-Verlag, 1997.

[Drmota, Fuchs, and Manstavičius 2003] Drmota, M., Fuchs, M., and Manstavičius, E. Functional limit theorems for digital expansions. *Acta Math. Hungar.* **98**(3), (2003) 175–201.

[Drmota, Grabner, and Liardet 2008] Drmota, M., Grabner, P.˜J., and Liardet, P. Block Additive Functions on the Gaussian Integers. *Acta Arith.* **135**, (2008) 299–332.

[Drmota, Mauduit, and Rivat 2009] Drmota, M., Mauduit, C., and Rivat, J. Primes with an Average Sum of Digits. *Compositio Math.* To appear.

[Drmota, Rivat, and Stoll 2008] Drmota, M., Rivat, J., and Stoll, T. The sum of digits of primes in $\mathbb{Z}[i]$. *Monatshefte Math.* **155**(3–4), (2008) 317–347. Special volume dedicated to the conference *Journées de Numération*, Graz, April 2007.

[Drmota 2001] Drmota, M. The joint distribution of q-additive functions. *Acta Arith.* **100**(1), (2001) 17–39.

[Dumont and Thomas 1991] Dumont, J.-M. and Thomas, A. Digital sum problems and substitutions on a finite alphabet. *J. Number Theory* **39**, (1991) 351–366.

[Dumont and Thomas 1993] Dumont, J.-M. and Thomas, A. Digital sum moments and substitutions. *Acta Arith.* **64**, (1993) 205–225.

[Dumont and Thomas 1997] Dumont, J.˜M. and Thomas, A. Gaussian asymptotic properties of the sum-of-digits function. *J. Number Theory* **62**, (1997) 19–38.

[Dumont 1990] Dumont, J.˜M. Summation formulae for substitutions on a finite alphabet. In J.-M. Luck, P.˜Moussa, and M.˜Waldschmidt, eds., *Number Theory and Physics*, vol.˜47 of *Springer Proceedings in Physics*, pp. 185–194. Springer-Verlag, 1990.

[Elliott 1979] Elliott, P. D. T.˜A. *Probabilistic number theory. I*, vol. 239 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science]*. Springer-Verlag, New York, 1979. Mean-value theorems.

[Elliott 1980] Elliott, P. D. T.˜A. *Probabilistic number theory. II*, vol. 240 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1980. Central limit theorems.

[Elliott 1985] Elliott, P. D. T.˜A. *Arithmetic functions and integer products*, vol. 272 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1985.

[Erdős 1939] Erdős, P. On a family of symmetric Bernoulli convolutions. *Amer. J. Math.* **61**, (1939) 974–976.

[Erdős 1940] Erdős, P. On the smoothness properties of a family of Bernoulli convolutions. *Amer. J. Math.* **62**, (1940) 180–186.

[Flajolet and Golin 1993] Flajolet, P. and Golin, M. Exact asymptotics of divide-and-conquer recurrences. In *Automata, languages and programming (Lund, 1993)*, vol. 700 of *Lecture Notes in Comput. Sci.*, pp. 137–149. Springer, Berlin, 1993.

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

[Foster 1987] Foster, D. M.˜E. Estimates for a remainder term associated with the sum of digits function. *Glasgow Math. J.* **29**, (1987) 109–129.

[Foster 1991] Foster, D. M.˜E. A lower bound for a remainder term associated with the sum of digits function. *Proc. Edinburgh Math. Soc.* **34**, (1991) 121–142.

[Foster 1992] Foster, D. M.˜E. Averaging the sum of digits function to an even base. *Proc. Edinburgh Math. Soc.* **35**, (1992) 449–455.

[Fouvry and Mauduit 1996a] Fouvry, E. and Mauduit, C. Méthodes de crible et fonctions sommes des chiffres. *Acta Arith.* **77**, (1996) 339–351.

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

[Gelfond 1968] Gelfond, A.˜O. Sur les nombres qui ont des propriétés additives et multiplicatives données. *Acta Arith.* **13**, (1968) 259–265.

[Grabner and Heuberger 2006] Grabner, P.˜J. and Heuberger, C. On the number of optimal base 2 representations of integers. *Des. Codes Cryptogr.* **40**, (2006) 25–39.

[Grabner and Hwang 2005] Grabner, P.˜J. and Hwang, H.-K. Digital sums and divide-and-conquer recurrences: Fourier expansions and absolute convergence. *Constructive Approximation* **21**, (2005) 149–179.

[Grabner and Liardet 1999] Grabner, P.˜J. and Liardet, P. Harmonic properties of the sum-of-digits function for complex bases. *Acta Arith.* **91**, (1999) 329–349.

[Grabner and Rigo 2003] Grabner, P.˜J. and Rigo, M. Additive functions with respect to numeration systems on regular languages. *Monatsh. Math.* **139**, (2003) 205–219.

[Grabner and Rigo 2007] Grabner, P.˜J. and Rigo, M. Distribution of additive functions with respect to numeration systems on regular languages. *Theory Comput. Syst.* **40**, (2007) 205–223.

[Grabner and Tichy 1990] Grabner, P.˜J. and Tichy, R.˜F. Contributions to digit expansions with respect to linear recurrences. *J. Number Theory* **36**, (1990) 160–169.

[Grabner and Tichy 1991] Grabner, P.˜J. and Tichy, R.˜F. α-expansions, linear recurrences, and the sum-of-digits function. *Manuscripta Math.* **70**, (1991) 311–324.

[Grabner, Heuberger, and Prodinger 2005] Grabner, P.˜J., Heuberger, C., and Prodinger, H. Counting optimal joint digit expansions. *Integers* **5**, (2005) A9, 19 pp. (electronic).

[Grabner, Kirschenhofer, and Prodinger 1998] Grabner, P.˜J., Kirschenhofer, P., and Prodinger, H. The sum-of-digits function for complex bases. *J. London Math. Soc.* **57**, (1998) 20–40.

[Grabner, Liardet, and Tichy 1995] Grabner, P.˜J., Liardet, P., and Tichy, R.˜F. Odometers and systems of numeration. *Acta Arith.* **70**, (1995) 103–123.

[Grabner, Liardet, and Tichy 2005] Grabner, P.˜J., Liardet, P., and Tichy, R.˜F. Spectral disjointness of dynamical systems related to some arithmetic functions. *Publ. Math. Debrecen* **66**, (2005) 213–243.

[Grabner 1993] Grabner, P.˜J. Completely q-multiplicative functions: the Mellin transform approach. *Acta Arith.* **65**, (1993) 85–96.

[Grabner 1997] Grabner, P.˜J. Functional iterations and stopping times for Brownian motion on the Sierpiński gasket. *Mathematika* **44**, (1997) 374–400.

[Grabner 2004] Grabner, P.˜J. Minima of digital functions related to large digits in q-adic expansions. *Quaest. Math.* **27**, (2004) 75–87.

[Granville 1997] Granville, A. Arithmetic properties of binomial coefficients. I. Binomial coefficients modulo prime powers. In *Organic mathematics (Burnaby, BC, 1995)*, vol.˜20 of *CMS Conf. Proc.*, pp. 253–276. Amer. Math. Soc., Providence, RI, 1997.

[Harborth 1977] Harborth, H. Number of odd binomial coefficients. *Proc. Amer. Math. Soc.* **62**, (1977) 19–22.

[Hardy and Riesz 1964] Hardy, G.˜H. and Riesz, M. *The general theory of Dirichlet's series*. Cambridge Tracts in Mathematics and Mathematical Physics, No. 18. Stechert-Hafner, Inc., New York, 1964.

[Heuberger and Prodinger 2006] Heuberger, C. and Prodinger, H. Analysis of alternative digit sets for nonadjacent representations. *Monatsh. Math.* **147**, (2006) 219–248.

[Hwang 1998] Hwang, H.-K. Asymptotics of divide-and-conquer recurrences: Batcher's sorting algorithm and a minimum Euclidean matching heuristic. *Algorithmica* **22**(4), (1998) 529–546. Average-case analysis of algorithms.

[Iwaniec and Kowalski 2004] Iwaniec, H. and Kowalski, E. *Analytic number theory*, vol.˜53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[Kamae 1977] Kamae, T. Mutual singularity of spectra of dynamical systems given by "sums of digits" to different bases. In *Dynamical Systems I – Warsaw*, vol.˜49 of *Astérisque*, pp. 109–114. Soc. Math. France, 1977.

[Kamae 1978] Kamae, T. Sum of digits to different bases and mutual singularity of their spectral measures. *Osaka J. Math.* **15**, (1978) 569–574.

[Kamae 1987] Kamae, T. Cyclic extensions of odometer transformations and spectral disjointness. *Israel J. Math.* **59**, (1987) 41–63.

[Kátai 1992] Kátai, I. Distribution of q-additive function. In *Probability theory*

and applications, vol.˜80 of *Math. Appl.*, pp. 309–318. Kluwer Acad. Publ., Dordrecht, 1992.

[Kim 1999] Kim, D.-H. On the joint distribution of q-additive functions in residue classes. *J. Number Theory* **74**(2), (1999) 307–336.

[Kirschenhofer and Tichy 1984] Kirschenhofer, P. and Tichy, R.˜F. On the distribution of digits in Cantor representations of integers. *J. Number Theory* **18**, (1984) 121–134.

[Knuth 1981] Knuth, D.˜E. *The Art of Computer Programming. Volume 2: Seminumerical Algorithms*. Addison-Wesley, 1981. 2nd edition.

[Kubilius 1964] Kubilius, J. *Probabilistic methods in the theory of numbers*. Translations of Mathematical Monographs, Vol. 11. American Mathematical Society, Providence, R.I., 1964.

[Kuipers and Niederreiter 1974] Kuipers, L. and Niederreiter, H. *Uniform Distribution of Sequences*. Wiley, 1974.

[Larcher 1996] Larcher, G. On the number of odd binomial coefficients. *Acta Math. Hung.* **71**, (1996) 183–203.

[Lecomte and Rigo 2001] Lecomte, P. B.˜A. and Rigo, M. Numeration systems on a regular language. *Theory Comput. Systems* **34**, (2001) 27–44.

[Liardet 1978] Liardet, P. Répartition et ergodicité. In *Séminaire Delange-Pisot-Poitou, 19e année: 1977/78, Théorie des nombres, Fasc. 1*, pp. Exp. No. 10, 12. Secrétariat Math., Paris, 1978.

[Mahler 1927] Mahler, K. On the translation properties of a simple class of arithmetical functions. *J. Math. and Phys.* **6**, (1927) 158–163.

[Manstavičius 1997] Manstavičius, E. Probabilistic theory of additive functions related to systems of numeration. In *New trends in probability and statistics, Vol. 4 (Palanga, 1996)*, pp. 413–429. VSP, Utrecht, 1997.

[Mauduit and Rivat 2009a] Mauduit, C. and Rivat, J. La somme des chiffres des carrés. *Acta Math.* To appear.

[Mauduit and Rivat 2009b] Mauduit, C. and Rivat, J. Sur un problème de Gelfond: la somme des chiffres des nombres premiers. *Ann. of Math.* To appear.

[Mauduit and Sárközy 1997] Mauduit, C. and Sárközy, A. On the arithmetic structure of the integers whose sum of digits is fixed. *Acta Arith.* **81**, (1997) 145–173.

[Mendès France 1967] Mendès France, M. Nombres normaux, applications aux fonctions pseudoaléatoires. *J. d'Analyse Math.* **20**, (1967) 1–56.

[Mendès France 1971] Mendès France, M. Fonctions g-additives et les suites à spectre vide. In *Séminaire Delange-Pisot-Poitou*, pp. 10.01–10.06. 1970/1971.

[Mendès France 1973] Mendès France, M. Les suites à spectre vide et la répartition modulo 1. *J. Number Theory* **5**, (1973) 1–15.

[Mendès France 1974] Mendès France, M. Les suites additives et leur répartition (mod. 1). In *Séminaire de Théorie des Nombres de Bordeaux*, pp. 8.01–8.06. 1973/1974.

[Niederreiter 1992] Niederreiter, H. *Random number generation and quasi-Monte Carlo methods*, vol.˜63 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[Okada, Sekiguchi, and Shiota 1995] Okada, T., Sekiguchi, T., and Shiota, Y. Applications of binomial measures to power sums of digital sums. *J. Number Theory* **52**, (1995) 256–266.

[Peres, Schlag, and Solomyak 2000] Peres, Y., Schlag, W., and Solomyak, B. Sixty years of Bernoulli convolutions. In *Fractal geometry and stochastics, II (Greifswald/Koserow, 1998)*, vol.˜46 of *Progr. Probab.*, pp. 39–65. Birkhäuser, Basel, 2000.

[Queffélec 1987] Queffélec, M. *Substitution Dynamical Systems — Spectral Analysis*, vol. 1294 of *Lecture Notes in Mathematics*. Springer-Verlag, 1987.

[Reingold and Tarjan 1981] Reingold, E.˜M. and Tarjan, R.˜E. On a greedy heuristic for complete matching. *SIAM J. Comput.* **10**(4), (1981) 676–681.

[Ruzsa 1984] Ruzsa, I.˜Z. Generalized moments of additive functions. *J. Number Theory* **18**(1), (1984) 27–33.

[Singmaster 1974a] Singmaster, D. Notes on binomial coefficients I — a generalization of Lucas' congruence. *J. London Math. Soc.* **8**, (1974) 545–548.

[Singmaster 1974b] Singmaster, D. Notes on binomial coefficients II — the least n such that p^e divides an r-nomial coefficient of rank n. *J. London Math. Soc.* **8**, (1974) 549–554.

[Singmaster 1974c] Singmaster, D. Notes on binomial coefficients II — any integer divides almost all binomial coefficients. *J. London Math. Soc.* **8**, (1974) 555–560.

[Stein 1989] Stein, A.˜H. Binomial coefficients not divisible by a prime. In D.˜V. Chudnovsky, G.˜V. Chudnovsky, H.˜Cohn, and M.˜B. Nathanson, eds., *Number Theory (New York, 1985/1988)*, vol. 1383 of *Lecture Notes in Mathematics*, pp. 170–177. Springer-Verlag, 1989.

[Steiner 2002] Steiner, W. Parry expansions of polynomial sequences. *Integers* **2** (2002), A14 (electronic), `http://www.integers-ejcnt.org/vol2.html`

[Tenenbaum 1995] Tenenbaum, G. *Introduction to Analytic and Probabilistic Number Theory*. Cambridge University Press, 1995.

[Tenenbaum 1997] Tenenbaum, G. Sur la non-dérivabilité de fonctions périodiques associées à certaines formules sommatoires. In R.˜L. Graham and J.˜Nešetřil, eds., *The Mathematics of Paul Erdős*, pp. 117–128. Springer-Verlag, 1997.

[Thuswaldner and Tichy 2005] Thuswaldner, J.˜M. and Tichy, R.˜F. Waring's problem with digital restrictions. *Israel J. Math.* **149**, (2005) 317–344. Probability in mathematics.

[Titchmarsh 1986] Titchmarsh, E.˜C. *The theory of the Riemann zeta-function*. The Clarendon Press Oxford University Press, New York, second edn., 1986. Edited and with a preface by D. R. Heath-Brown.

[Wiener 1927] Wiener, N. The spectrum of an array and its applications to the study of the translation properties of a simple class of arithmetical functions. *J. Math. and Phys.* **6**, (1927) 145–157.

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

Notation Index

General Index