Distribution of arithmetic functions on particular subsets of integers

Jean-Marie De Koninck[1] and Imre Kátai

Dedicated to Professors Antanas Laurincikas and Eugenijus Manstavicius on the occasion of their 70-th anniversary

Abstract

Let $Q_1, \dots, Q_t \in \mathbb{R}[x]$ be polynomials with no constant term for which each linear combination $m_1 Q_1(x) + \cdots + m_t Q_t(x)$, with $m_1, \dots, m_t \in \mathbb{Z}$ and not all 0, always has an irrational coefficient. Let $I_1, \dots, I_t$ be sets included in the interval $[0,1)$, each of which being a union of finitely many subintervals of $[0,1)$. Furthermore, let $\mathcal{T}$ be the set of those positive integers n for which $\{Q_1(n)\} \in I_1, \dots, \{Q_t(n)\} \in I_t$ holds simultaneously, where $\{y\}$ stands for the fractional part of y. Let $t_1, t_2, \dots$ be a sequence of complex numbers uniformly summable and set $T(x) = \sum_{n \le x} t_n$ and $T(x|\mathcal{T}) = \sum_{\substack{n \le x \\ n \in \mathcal{T}}} t_n$. We prove that, as $x \to \infty$, $T(x)/x \sim T(x|\mathcal{T})/(\lambda(I_1) \cdots \lambda(I_t) x)$, where $\lambda(I)$ stands for the Lebesgue measure of the set I.

AMS Subject Classification numbers: 11K16, 11N37, 11A41
Key words: polynomials, primes, arithmetic function

1 Introduction

We say that sequence of real numbers t_n is *uniformly summable* if

$$\limsup_{x \to \infty} \frac{1}{x} \sum_{\substack{n \le x \\ |t_n| \ge K}} |t_n| \le \delta(K)$$

for some sequence $\delta(K)$ tending to 0 as $K \to \infty$.

Let $I_1, \dots, I_t$ be sets included in the interval $[0,1)$, each of which being a union of finitely many subintervals of $[0,1)$. For each $j \in \{1, \dots, t\}$, let $\ell_j(x)$ be a mod 1 periodic function, defined by

$$\ell_j(x) = \begin{cases} 1 & \text{if } x \in I_j, \\ 0 & \text{if } x \in [0,1) \setminus I_j. \end{cases}$$

It is easy to see that if $\sum_{n=-\infty}^{\infty} a_n^{(j)} e(nx)$ stands for the Fourier series associated with $\ell_j(x)$ (here, $e(y)$ stands for $\exp\{2\pi i y\}$), then

$$\left|a_n^{(j)}\right| \le \frac{c_j}{|n|} \qquad \text{and} \qquad \left|a_n^{(j)}\right| \le 1 \quad \text{for all } n \in \mathbb{Z},\ j = 1, \dots, t,$$

[1]Research supported in part by a grant from NSERC.

where c_j is the number of end points of the intervals occurring in the set I_j.

Given a small constant $\Delta > 0$, we set

$$\ell_j^{(\Delta)}(x) := \frac{1}{(2\Delta)^2} \int_{-\Delta}^{\Delta} \int_{-\Delta}^{\Delta} \ell_j(x + u_1 + u_2)\, du_1\, du_2.$$

Further let

$$\kappa(n) = \frac{\sin 2\pi\Delta n}{4\pi\Delta n}.$$

Then,

$$\begin{aligned} \ell_j^{(\Delta)}(x) &= \sum_{n=-\infty}^{\infty} b_n^{(j)} e(nx), \\ b_n^{(j)} &= \kappa(n)\, a_n^{(j)}, \\ \left|b_n^{(j)}\right| &\leq \min\left(1, \frac{1}{\Delta|n|}\right)^2. \end{aligned}$$

We further define, for each $j \in \{1, \ldots, t\}$,

$$\begin{aligned} I_j^{(-\Delta)} &= \{x : (x - 2\Delta, x + 2\Delta) \subseteq I_j\} \\ I_j^{(\Delta)} &= \{x : (x - 2\Delta, x + 2\Delta) \cap I_j = \emptyset\} \end{aligned}$$

Observe that $\lambda(I_j^{(\Delta)} \setminus I_j^{(-\Delta)}) \leq c_j\Delta$, where $\lambda(I)$ stands for the Lebesgue measure of the set I. Moreover, observe that

$$\ell_j^{(\Delta)}(x) = \begin{cases} 1 & \text{if } x \in I_j^{(-\Delta)}, \\ 0 & \text{if } x \in [0,1) \setminus I_j^{(\Delta)} \end{cases}$$

and

$$0 \leq \ell_j^{(\Delta)}(x) \leq 1 \quad \text{for all } x. \tag{1.1}$$

We now introduce the truncated sum

$$\ell_j^{(\Delta,K)}(x) = \sum_{|n|<K} b_n^{(j)} e(nx).$$

Choosing $K \geq (1/\Delta)^4$, we get that

$$\sum_{|n|\geq K} |b_n^{(j)}| \leq 2 \sum_{n \geq K} \frac{1}{(\Delta n)^2} \leq 2\Delta^2. \tag{1.2}$$

From this estimate, we can prove that, given t points $x_1, \ldots, x_t$,

$$\left|\ell_1^{(\Delta)}(x_1) \cdots \ell_t^{(\Delta)}(x_t) - \ell_1^{(\Delta,K)}(x_1) \cdots \ell_t^{(\Delta,K)}(x_t)\right| \leq 3t\Delta^2. \tag{1.3}$$

To see this, we proceed as follows. First, for each j, we write

$$\ell_j^{(\Delta)}(x) = \ell_j^{(\Delta,K)}(x) + T_j(x), \text{ so that } T_j(x) = \sum_{|n|\geq K} b_n^{(j)} e(nx).$$

Using (1.2), one can easily see that

$$|T_j(x)| \leq 2\Delta^2. \tag{1.4}$$

It then follows from (1.1) and (1.4) that

$$\left|\ell_j^{(\Delta,K)}(x)\right| \leq |T_j(x)| + \left|\ell_j^{(\Delta)}(x)\right| \leq 1 + 2\Delta^2. \tag{1.5}$$

We shall now estimate the size of

$$R_h(x) := \ell_1^{(\Delta)}(x)\cdots\ell_h^{(\Delta)}(x) - \ell_1^{(\Delta,K)}(x)\cdots\ell_h^{(\Delta,K)}(x).$$

We have

$$\begin{aligned} R_h(x) &= \ell_1^{(\Delta)}(x)\cdots\ell_{h-1}^{(\Delta)}(x)\left(\ell_h^{(\Delta,K)}(x) + T_h(x)\right) - \ell_1^{(\Delta,K)}(x)\cdots\ell_h^{(\Delta,K)}(x) \\ &= T_h(x)\ell_1^{(\Delta)}(x)\cdots\ell_{h-1}^{(\Delta)}(x) + \ell_h^{(\Delta,K)}(x)R_{h-1}(x). \end{aligned} \tag{1.6}$$

In light of (1.1), (1.4) and (1.5), it follows from (1.6) that

$$|R_h(x)| \leq 2\Delta^2 + (1+2\Delta^2)\,|R_{h-1}(x)|\,. \tag{1.7}$$

Setting $C(1) = 2$ and thereafter $C(h) = (1+2\Delta^2)C(h-1)+2$, it follows from (1.7) that

$$|R_h(x)| \leq C(h)\Delta^2.$$

Since one can easily obtain from the above definition of $C(h)$ that

$$C(h) \leq 3h \qquad (h = 1, 2, \ldots, t),$$

provided Δ is sufficiently small, (1.3) follows immediately.

So, if we introduce the notations

$$\begin{aligned} s(x_1,\ldots,x_t) &= \ell_1(x_1)\cdots\ell_t(x_t), \\ s^{(\Delta)}(x_1,\ldots,x_t) &= \ell_1^{(\Delta)}(x_1)\cdots\ell_t^{(\Delta)}(x_t), \\ s^{(\Delta,K)}(x_1,\ldots,x_t) &= \ell_1^{(\Delta,K)}(x_1)\cdots\ell_t^{(\Delta,K)}(x_t), \end{aligned}$$

it follows from (1.3) that

$$\left|s^{(\Delta)}(x_1,\ldots,x_t) - s^{(\Delta,K)}(x_1,\ldots,x_t)\right| \leq 2t\Delta^2. \tag{1.8}$$

Now, let the *discrepancy* of a sequence $y_1, \ldots, y_n$ be defined as usual as

$$D_N(y_1, \ldots, y_N) = \sup_{[\alpha,\beta)\subseteq[0,1)} \left| \frac{1}{N} \sum_{\substack{j=1 \\ \{y_j\}\in[\alpha,\beta)}}^{N} 1 - (\beta - \alpha) \right|,$$

where $\{y\}$ stands for the fractional part of y. Then by the Erdős-Turán Theorem [4], it is known that there exists an absolute constant $c > 0$ such that, given an arbitrary positive integer T,

$$D_N(y_1, \ldots, y_N) \leq c \left(\sum_{k=1}^{T} \frac{|\Psi_k|}{k} + \frac{1}{T} \right),$$

where $\Psi_k = \frac{1}{N} \sum_{j=1}^{N} e(ky_j)$.

2 Main results and their proofs

Let $Q_1, \ldots, Q_t \in \mathbb{R}[x]$ be polynomials satisfying $Q_j(0) = 0$ for each $j \in \{1, \ldots, t\}$ and for which each linear combination $U_{m_1,\ldots,m_t}(x) := m_1 Q_1(x) + \cdots + m_t Q_t(x)$ (with $m_1, \ldots, m_t \in \mathbb{Z}$ with the exception of $m_1 = \cdots = m_t = 0$) always has an irrational coefficient.

Let $\mathcal{T}$ be the set of those positive integers n for which

$$\{Q_1(n)\} \in I_1, \ldots, \{Q_t(n)\} \in I_t \qquad \text{hold simultaneously.}$$

Then, it is clear that

$$n \in \mathcal{T} \text{ if and only if } s(Q_1(n), \ldots, Q_t(n)) = 1.$$

Let $t_1, t_2, \ldots$ be a sequence of complex numbers such that $|t_n| \leq 1$ and set

$$T(x) = \sum_{n \leq x} t_n \qquad \text{and} \qquad T(x|\mathcal{T}) = \sum_{\substack{n \leq x \\ n \in \mathcal{T}}} t_n.$$

Let $\delta(M)$ be a sequence which is such that $\frac{1}{x} \sum_{\substack{n \leq x \\ |t_n| \geq M}} |t_n| \leq \delta(M)$ if $x > x_0(M)$. Assuming that $x > x_0(M)$, then, one can see that

$$\begin{aligned} T(x|\mathcal{T}) &= \sum_{\substack{n \leq x \\ |t_n| \leq M}} t_n s(Q_1(n), \ldots, Q_t(n)) \\ &= \sum_{n \leq x} t_n s^{(\Delta)}(Q_1(n), \ldots, Q_t(n)) + O(\delta(M)x) \end{aligned}$$

$$+O\left(M\sum_{j=1}^{t}\sum_{\{Q_j(n)\}\in I_j^{(\Delta)}\setminus I_j^{(-\Delta)}}1\right). \tag{2.1}$$

Setting

$$\Sigma^{(j)} := \sum_{\{Q_j(n)\}\in I_j^{(\Delta)}\setminus I_j^{(-\Delta)}}1$$

and using the Erdős-Turán Theorem mentioned above, we get that, for some $C>0$,

$$\Sigma^{(j)} \le c_j C\left(\frac{x}{T}+\sum_{k=1}^{T}\frac{1}{k}\left|\sum_{n\le x}e(kQ_j(n))\right|\right)+c_j\Delta x. \tag{2.2}$$

Then, an old theorem of Weyl [12] tells us that, if $k\neq 0$, then

$$\lim_{x\to\infty}\frac{1}{x}\sum_{n\le x}e(kQ_j(n))=0. \tag{2.3}$$

Substituting (2.3) in (2.2), it follows that

$$\limsup_{x\to\infty}\frac{1}{x}\Sigma^{(j)} \le c_j\Delta+\frac{c_jC}{T}. \tag{2.4}$$

Since T can be taken arbitrarily large, it follows, in light of (2.4), that the O(...) term in (2.1) is $\ll M\Delta x$.

Hence, using (1.8), estimate (2.1) becomes

$$\begin{aligned}
T(x|\mathcal{T}) &= \sum_{\substack{n\le x\\ |t_n|\le M}} t_n s^{(\Delta,K)}(Q_1(n),\ldots,Q_t(n))+O(Mt\Delta^2x)+O(M\Delta x)+O(\delta(M)x)\\
&= \sum_{\substack{n\le x\\ |t_n|\le M}} t_n \sum_{\substack{m_1,\ldots,m_t\in\mathbb{Z}\\ |m_\nu|\le K}} b_{m_1}^{(1)}\cdots b_{m_t}^{(t)} e(m_1Q_1(n)+\cdots+m_tQ_t(n))\\
&\qquad +O(Mt\Delta^2x)+O(M\Delta x)+O(\delta(M)x).
\end{aligned} \tag{2.5}$$

Since one can easily see that

$$b_0^{(j)}=\lambda(I_j) \qquad (j=1,\ldots,t),$$

it follows from (2.5) that

$$\begin{aligned}
T(x|\mathcal{T}) &= \lambda(I_1)\cdots\lambda(I_t)T(x)\\
&\quad + \sum_{\substack{m_1,\ldots,m_t\\ (m_1,\ldots,m_t)\neq(0,\ldots,0)\\ |m_\nu|\le K}} b_{m_1}^{(1)}\cdots b_{m_t}^{(t)} \sum_{\substack{n\le x\\ |t_n|\le M}} t_n e(m_1Q_1(n)+\cdots+m_tQ_t(n))\\
&\qquad +O(Mt\Delta^2x)+O(M\Delta x)+O(\delta(M)x).
\end{aligned}$$

For convenience, from here on, let

$$D := \lambda(I_1) \cdots \lambda(I_t). \tag{2.6}$$

Since Δ can be taken arbitrarily small, and since $\delta(1/\sqrt{\Delta}) \to 0$ as $\Delta \to 0$, we have thus proven the following result.

Theorem 1. *Let t_n be a uniformly summable sequence. Assume that*

$$\frac{1}{x}\sum_{n\le x} t_n e(m_1Q_1(n) + \cdots + m_tQ_t(n)) \to 0 \qquad \text{as } x \to \infty$$

for every t-tuple $(m_1, \ldots, m_t)$ with $(m_1, \ldots, m_t) \neq (0, \ldots, 0)$. Then,

$$\lim_{x\to\infty} \left(\frac{T(x)}{x} - \frac{T(x|\mathcal{T})}{Dx} \right) = 0.$$

Let $\wp$ stand for the set of all primes and then set

$$\mathcal{T}_\wp = \{p : p \in \wp \cap \mathcal{T}\}.$$

Further set

$$S(x) = \sum_{p\le x} t_p \qquad \text{and} \qquad S(x|\mathcal{T}_\wp) = \sum_{\substack{p\le x\\ p\in\mathcal{T}_\wp}} t_p$$

and assume that $Q_1, \ldots, Q_t$ are polynomials satisfying the conditions stated above.

Theorem 2. *Let t_p be a uniformly summable sequence and let D be as in (2.6), then,*

$$\lim_{x\to\infty} \left(\frac{S(x)}{\pi(x)} - \frac{S(x|\mathcal{T}_\wp)}{D\pi(x)} \right) = 0.$$

The proof of Theorem 2 is very similar to that of Theorem 1 with the exception that, instead of Weyl's Theorem, one should use the theorem of I.M. Vinogradov [11], that is the one that states that

$$\frac{1}{\pi(x)}\sum_{p\le x} e(kQ_j(p)) \to 0 \qquad \text{as } x \to \infty.$$

Remark 1. *According to a classical theorem of Daboussi (see Daboussi and Delange [1], as well as Daboussi and Delange [2]),*

$$\sup_{f\in\mathcal{M}_1} \frac{1}{x}\left|\sum_{n\le x} f(n)e(n\alpha)\right| \to 0 \ \text{as } x \to \infty$$

for every irrational number α. Here $\mathcal{M}_1$ stands for the set of multiplicative functions $f : \mathbb{N} \to \mathbb{C}$ such that $|f(n)| \le 1$.

This result has been generalized by Kátai [6] who proved that, given any polynomial $F(x) = \alpha_k x^k + \cdots + \alpha_1 x \in \mathbb{R}[x]$ with at least one irrational coefficient,

$$\sup_{f \in \mathcal{M}_1} \frac{1}{x} \left| \sum_{n \le x} f(n) e(F(n)) \right| \to 0 \qquad \text{as } x \to \infty.$$

Later, this assertion was generalized by Indlekofer and Kátai [9] for uniformly summable multiplicative sequences $f(n)$.

Recently, Kátai [5] proved that

$$\sup_{g \in \mathcal{M}_1} \frac{1}{x} \left| \sum_{\substack{n \le x \\ n \in \mathcal{T}}} g(n) - \frac{1}{D} \sum_{n \le x} g(n) \right| \to 0 \text{ as } x \to \infty. \tag{2.7}$$

Thus, Theorem 1 in the case $t_n = g(n) \in \mathcal{M}_1$ has been proved earlier.

3 Applications

3.1 First set of applications

Theorem 3. *Let f be an additive function for which the necessary conditions of the Erdős-Wintner Theorem hold, namely the three conditions*

$$\sum_{|f(p)|>1} \frac{1}{p} < \infty, \quad \sum_{|f(p)| \le 1} \frac{f(p)}{p} \text{ is convergent}, \quad \sum_{|f(p)| \le 1} \frac{f^2(p)}{p} < \infty.$$

Let $F(y)$ be the limit distribution function of f. Then,

$$\lim_{x \to \infty} \frac{1}{Dx} \#\{n \le x : n \in \mathcal{T},\ f(n) < y\} = F(y).$$

Theorem 4. *Let f be an additive function satisfying the two conditions*

$$\sum_{|f(p)|>1} \frac{1}{p} < \infty \qquad \text{and} \qquad \sum_{|f(p)| \le 1} \frac{f^2(p)}{p} < \infty.$$

Let $A(x) = \sum_{p \le x} \frac{f(p)}{p}$ and

$$F^*(y) = \lim_{x \to \infty} \frac{1}{x} \#\{n \le x : f(n) - A(x) < y\},$$

which exists for almost all y. Then,

$$\lim_{x \to \infty} \frac{1}{Dx} \#\{n \le x : n \in \mathcal{T},\ f(n) - A(x) < y\} = F^*(y).$$

In what follows, we shall let $f(n)$ be a strongly additive function and set

$$A(x) := \sum_{p\le x} \frac{f(p)}{p} \qquad \text{and} \qquad B(x) := \left(\sum_{p\le x} \frac{f^2(p)}{p}\right)^{1/2}.$$

Following Kubilius, we shall say that f belongs to the class H if there exists a function $r = r(x)$ such that, as $x \to \infty$,

$$\frac{\log r}{\log x} \to 0, \qquad \frac{B(r)}{B(x)} \to 1, \qquad B(x) \to \infty.$$

And, as usual, let $\Phi(z)$ be the normal distribution function, that is

$$\Phi(z) = \frac{1}{2\pi} \int_{-\infty}^{z} e^{-u^2/2}\, du \qquad (z \in \mathbb{R}).$$

Then, the following result can be proved to be a consequence of (2.7).

Theorem 5. *(Kubilius, Shapiro) Let $f(n)$ be a strongly additive function. In order to have*

$$\lim_{x\to\infty} \frac{1}{xD} \#\{n \le x : n \in \mathcal{T},\ f(n) - A(x) < zB(x)\} = \Phi(z),$$

it is sufficient that for each fixed $\varepsilon > 0$,

$$\frac{1}{B^2(x)} \sum_{\substack{p\le x \\ |f(p)|>\varepsilon B(x)}} \frac{f^2(p)}{p} \to 0 \qquad \text{as } x \to \infty.$$

Moreover, if $f(n)$ belongs to the class H, then this condition is also necessary.

The above result is Theorem 12.2, with $\mathcal{T} = \mathbb{N}$, in the book of Elliott [3].

As a special case, we obtain the following analogue of the Erdős-Kac Theorem, which can also be found in the book of Elliott [3]:

Theorem 6. *Let $f(n)$ be a strongly additive function which satisfies $|f(p)| \le 1$ for all primes p. Assume that $B(x) \to \infty$ as $x \to \infty$. Then,*

$$\lim_{x\to\infty} \frac{1}{Dx} \#\{n \le x : n \in \mathcal{T},\ f(n) - A(x) < zB(x)\} = \Phi(z).$$

Observe that Theorems 3, 4 and 5 can be deduced directly from (2.7), namely choosing $g(n) = e^{itf(n)}$ and then using it for $t_n = g(n)$ in Theorem 3, $t_n = t_n^{(t)} = g(n)e^{-itf(n)}$ in Theorem 4 and finally $t_n = t_n^{(t)} = g(n)e^{-it(f(n)-A(x))/B(x)}$ in Theorem 5.

3.2 Second set of applications

Let g be a multiplicative function satisfying $|g(n)| = 1$ for all $n \in \mathbb{N}$. Given a real number $Y \geq 2$, consider the multiplicative function g_Y defined on the prime powers p^a by

$$g_Y(p^a) = \begin{cases} g(p^a) & \text{if } p \leq Y, \\ 1 & \text{if } p > Y. \end{cases}$$

Let $h(n)$ be the Moebius inverse of g, that is $\sum_{d|n} h(d) = g(n)$. Similarly, let $g_Y(n)$ be the Moebius inverse of $h_Y(n)$. Finally, let $f(n)$ be the additive function defined on prime powers p^a by $f(p^a) = \arg g(p^a)$, so that $g(n) = e^{if(n)}$.

Assume that

$$\sum_p \frac{1-g(p)}{p} \qquad \text{is convergent.} \tag{3.1}$$

From the Turán-Kubilius inequality applied to the additive function f, we obtain that

$$\limsup_{x\to\infty} \frac{1}{x} \sum_{n\leq x} |g(n) - g_Y(n)| \leq \delta(Y),$$

where $\delta(Y) \to 0$ as $Y \to \infty$.

Moreover, $h_Y(p^a) = 0$ if $p > Y$, $h_Y(p^a) = h(p^a)$ if $p \leq Y$.

Recalling that $P(n)$ stands for the largest prime factor of n, observe that

$$\frac{1}{x}\#\{n \leq x : \exists d|n,\ d > Y^{K_Y},\ P(d) < Y\} \to 0 \qquad \text{as } K_Y \to \infty.$$

Consequently,

$$g_Y(n) = \sum_{\substack{d|n \\ d<Y^{K_Y}}} h_Y(d)$$

for all but at most $\delta(Y)x$ integers $n \leq x$.

Now, consider the k linear functions $L_\ell(n) = a_\ell n + b_\ell$ ($\ell = 1, \ldots, k$), where each a_ℓ is a positive integer and each $b_\ell \in \mathbb{Z}$, and let $g^{(\ell)}(n)$, $\ell = 1, \ldots, k$, be multiplicative functions such that $|g^{(\ell)}(n)| = 1$ and satisfying condition (3.1). Then,

$$\frac{1}{x}\sum_{n\leq x} \left| \prod_{\ell=1}^k g^{(\ell)}(a_\ell n + b_\ell) - \prod_{\ell=1}^k g_Y^{(\ell)}(a_\ell n + b_\ell) \right| \leq k\delta(Y) \tag{3.2}$$

and

$$\prod_{\ell=1}^k g^{(\ell)}(a_\ell n + b_\ell) = \sum_{\substack{d_1,\ldots,d_k \\ d_\ell \leq Y^{K_Y} \\ d_\ell | a_\ell n + b_\ell,\ P(d_\ell)\leq Y}} \prod_{\ell=1}^k h_Y^{(\ell)}(d_\ell) \tag{3.3}$$

for all but no more than $\varepsilon(Y, K_Y)x$ integers $n \le x$. Here $\varepsilon(Y, K_Y) \to 0$ as $Y \to \infty$ and $K_Y \to \infty$.

Further set

$$t_n = \prod_{\ell=1}^{k} g^{(\ell)}(a_\ell n + b_\ell).$$

Then, $a_\ell n + b_\ell \equiv 0 \pmod{d_\ell}$ holds for some residue classes modulo $LCM[d_1, \ldots, d_k]$. Let these residue classes be $n \equiv u_j \pmod{LCM[d_1, \ldots, d_k]}$ for $j = 1, \ldots, s$ (here, the u_j's may depend on $d_1, \ldots, d_k$).

Hence in light of (3.2) and (3.3), we get that

$$\begin{aligned} &\sum_{n \le x} t_n e(m_1 Q_1(n) + \cdots + m_t Q_t(n)) \\ &\quad = \sum_{\substack{d_1, \ldots, d_k \\ d_\ell \le Y^{K_Y} \\ P(d_\ell) \le Y}} \prod_{\ell=1}^{k} h_Y^{(\ell)}(d_\ell) \sum_{j=1}^{s} \sum_{n \equiv u_j \pmod{\frac{LCM[d_1, \ldots, d_k]}{D}}} e(m_1 Q_1(n) + \cdots + m_t Q_t(n)) \\ &\qquad + O(\delta(Y)x) + O(\varepsilon(Y, K_Y)x). \end{aligned} \tag{3.4}$$

The inner sum on the right hand side of (3.4) can be written as

$$\frac{1}{D} \sum_{a=1}^{D} \sum_{n \le x} e\left(\frac{(n - u_j)a}{D}\right) e(m_1 Q_1(n) + \cdots + m_t Q_t(n)). \tag{3.5}$$

Since the polynomial

$$\frac{(y - u_j)a}{D} + m_1 Q_1(y) + \cdots + m_t Q_t(y)$$

has an irrational coefficient, by a classical theorem of Weyl, the sum (3.5) must be $o(x)$ as $x \to \infty$. We are thus in the range of the conditions of Theorem 1. Therefore, the following result is an application of Theorem 1.

Theorem 7. *Assume that condition (3.1) holds for $g = g_\ell$ ($\ell = 1, \ldots, k$) and that $|g_\ell(n)| = 1$ for all $n \in \mathbb{N}$. Let*

$$H(n) := \prod_{\ell=1}^{k} g_\ell(a_\ell n + b_\ell) \quad \textit{with } a_\ell \in \mathbb{N},\ b_\ell \in \mathbb{Z}.$$

Then,

$$\mathcal{L} = \lim_{x \to \infty} \frac{1}{Dx} \sum_{\substack{n \le x \\ n \in \mathcal{T}}} H(n) = \lim_{x \to \infty} \frac{1}{x} \sum_{n \le x} H(n) \textit{ exists.}$$

Moreover, $\mathcal{L} = \mathcal{L}_1 \mathcal{L}_2$, where

$$\mathcal{L}_1 = \sum_{\{d_1, \ldots, d_k\} \in D_Y} \frac{h^{(1)}(d_1) \cdots h^{(k)}(d_k)}{LCM[d_1, \ldots, d_k]} \rho(d_1, \ldots, d_k),$$

$$\mathcal{L}_2 = \prod_{p>Y} m(p), \quad \textit{where} \quad m(p) = 1 + \sum_{a=1}^{\infty} \frac{1}{p^a} \sum_{\ell=1}^{k} (g_\ell(p^a) - 1).$$

Here, D_Y stands for the set of those $\{d_1, \ldots, d_k\}$ for which $P(d_j) \le Y$ and Y is so large that $\rho(d_i, d_j) = 0$ if $i \neq j$ and $d_i d_j$ has a prime factor larger than Y.

As a corollary, we have the following result.

Theorem 8. *Let $f_\ell(n)$, for $\ell = 1, \ldots, k$, be additive functions each satisfying the three conditions*

$$\sum_{|f_\ell(p)|>1} \frac{1}{p} < \infty, \qquad \sum_{|f_\ell(p)|\le 1} \frac{f_\ell(p)}{p} \ \textit{is convergent}, \qquad \sum_{|f_\ell(p)|\le 1} \frac{f_\ell^2(p)}{p} < \infty.$$

Then, the distribution function

$$F(y_1, \ldots, y_k) := \lim_{x\to\infty} \frac{1}{Dx} \#\{n \le x : n \in \mathcal{T},\ f_\ell(a_\ell n + b_\ell) < y_\ell,\ \ell = 1, \ldots, k\}$$

exists for almost all $y_1, \ldots, y_k$ and, moreover,

$$F(y_1, \ldots, y_k) = \lim_{x\to\infty} \frac{1}{x} \#\{n \le x :\ f_\ell(a_\ell n + b_\ell) < y_\ell,\ \ell = 1, \ldots, k\}$$

for almost all $y_1, \ldots, y_k$.

We can also prove the following.

Theorem 9. *Let $f_\ell(n)$, for $\ell = 1, \ldots, k$, be additive functions each satisfying the two conditions*

$$\sum_{|f_\ell(p)|>1} \frac{1}{p} < \infty \qquad \textit{and} \qquad \sum_{|f_\ell(p)|\le 1} \frac{f_\ell^2(p)}{p} < \infty,$$

and let $A_\ell(x) = \sum_{\substack{p\le x \\ |f_\ell(p)|\le 1}} \frac{f(p)}{p}$. Then, the distribution function

$$F(y_1, \ldots, y_k) := \lim_{x\to\infty} \frac{1}{Dx} \#\{n \le x : n \in \mathcal{T},\ f_\ell(a_\ell n + b_\ell) - A_\ell(x) < y_\ell,\ \ell = 1, \ldots, k\}$$

exists for almost all $y_1, \ldots, y_k$ and, moreover,

$$F(y_1, \ldots, y_k) = \lim_{x\to\infty} \frac{1}{x} \#\{n \le x :\ f_\ell(a_\ell n + b_\ell) - A_\ell(x) < y_\ell,\ \ell = 1, \ldots, k\}$$

for almost all $y_1, \ldots, y_k$.

Following the method used in Kátai [8], we can also prove the following results.

Theorem 10. *Assume that the conditions of Theorem 9 hold and that* $b_\ell \neq 0$ *for* $\ell = 1, \ldots, k$*. Then, the distribution function*

$$G(y_1, \ldots, y_k) = \lim_{x \to \infty} \frac{1}{\pi(x)} \#\{p \leq x : \ f_\ell(a_\ell p + b_\ell) - A_\ell(x) < y_\ell, \ \ell = 1, \ldots, k\}$$

exists for almost all $y_1, \ldots, y_k$ *and*

$$G(y_1, \ldots, y_k) = \lim_{x \to \infty} \frac{1}{D\pi(x)} \#\{p \leq x : p \in \mathcal{T}, \ f_\ell(a_\ell p + b_\ell) - A_\ell(x) < y_\ell, \ \ell = 1, \ldots, k\}$$

for almost all $y_1, \ldots, y_k$.

Theorem 11. *Assume that the conditions of Theorem 8 hold and that* $b_\ell \neq 0$ *for* $\ell = 1, \ldots, k$*. Then, the distribution function*

$$\widetilde{G}(y_1, \ldots, y_k) = \lim_{x \to \infty} \frac{1}{\pi(x)} \#\{p \leq x : \ f_\ell(a_\ell p + b_\ell) < y_\ell, \ \ell = 1, \ldots, k\}$$

exists for almost all $y_1, \ldots, y_k$ *and*

$$\widetilde{G}(y_1, \ldots, y_k) = \lim_{x \to \infty} \frac{1}{D\pi(x)} \#\{p \leq x : p \in \mathcal{T}, \ f_\ell(a_\ell p + b_\ell) < y_\ell, \ \ell = 1, \ldots, k\}$$

for almost all $y_1, \ldots, y_k$.

3.3 Further applications

We can obtain the analogues of all the theorems proved in Kátai [7]. To illustrate this, we will only explicitly formulate the analogue of Theorem 5.

Theorem 12. *Let* $F_j(x) \in \mathbb{Z}[x]$, $j = 1, \ldots, k$, *be* k *polynomials each of which has a positive leading coefficient. Let* x_0 *be chosen in such a way that for all* $j \in \{1, \ldots, k\}$, $F_j(n) > 0$ *if* $n \geq x_0$. *Let also* $\gamma \in \mathbb{N}$ *be such that* $F_i(n) \equiv 0 \pmod{p}$ *and* $F_j(n) \equiv 0 \pmod{p}$, *with* $i \neq j$, *do not hold simultaneously if* $p > \gamma$. *(Such an integer* γ *exists (see Tamaka [10]).) Further let* D_γ *be the set of those* k*-tuples of natural numbers* $\{d_1, \ldots, d_k\}$ *such that* $P(d_j) \leq \gamma$ *for all* $j \in \{1, \ldots, k\}$. *Let* $\rho(d_1, \ldots, d_k)$ *stand for the number of those* $n \pmod{LCM[d_1, \ldots, d_k]}$ *for which* $F_j(n) \equiv 0 \pmod{d_j}$, $j = 1, \ldots, k$, *simultaneously hold. Furthermore, let* $\lambda(d_1, \ldots, d_k)$ *be the number of solutions* n *for which the additional condition* $GCD(n, \prod_{j=1}^k d_j) = 1$ *holds. Also, for each* $j \in \{1, \ldots, j\}$, *let* $\rho_j(d)$ *be the number of solutions* n *of* $F_j(n) \equiv 0 \pmod{d}$ *and let* $\lambda_j(d)$ *be the number of solutions* n *for which the additional condition* $gcd(n, d) = 1$ *holds. Finally, assume that the polynomials* $F_i(x)$ *and* $F_j(x)$ *are coprime when* $i \neq j$. *For each* $j \in \{1, \ldots, k\}$, *let* ν_j *stand for the degree of* $F_j(x)$ *and* $g_j(n)$ *be a multiplicative function satisfying* $|g_j(n)| = 1$ *for all* $n \in \mathbb{N}$. *Furthermore, assume that*

$$\sum_p \frac{(1 - g_j(p))\rho_j(p)}{p} \qquad \text{converges for } j = 1, \ldots, k$$

and that

$$(1-g_j(p^a))\rho_j(p^a) \to 0 \qquad \text{as } p\to\infty$$

for $a=1$ *when* $\nu_j=2$ *and for* $a=1,2,\ldots,\nu_j-2$ *if* $\nu_j\geq 3$.

Let

$$H(n) := \prod_{j=1}^{k} g_j(F_j(n)).$$

Then,

$$\mathcal{M} := \lim_{x\to\infty}\frac{1}{Dx}\sum_{\substack{n\leq x\\ n\in\mathcal{T}}} H(n) = \lim_{x\to\infty}\frac{1}{x}\sum_{n\leq x} H(n)$$

exists, and moreover

$$\mathcal{M}=\mathcal{M}_1\mathcal{M}_2,$$

where

$$\mathcal{M}_1 = \sum_{\{d_1,\ldots,d_k\}\in D_\gamma} \frac{h_1(d_1)\cdots h_k(d_k)}{LCM[d_1,\ldots,d_k]}\rho(d_1,\ldots,d_k),$$

$$\text{where } h_j(n)=\sum_{d|n}\mu(d)g_j(n/d), \quad (j=1,\ldots,k),$$

$$\mathcal{M}_2 = \prod_{p>\gamma}\widetilde{m}(p), \qquad \text{where} \quad \widetilde{m}(p)=1+\sum_{a=1}^{\infty}\frac{1}{p^a}\sum_{j=1}^{k}(g_j(p^a)-1)\rho_j(p^a).$$

If we also assume that, for $j=1,\ldots,k$,

$$(1-g_j(p^a))\rho_j(p^a) \to 0 \qquad \text{as } p\to\infty$$

for $a=\nu_j-1$ *when* $\nu_j\geq 2$, *then*

$$\mathcal{N} := \lim_{x\to\infty}\frac{1}{D\pi(x)}\sum_{\substack{p\leq x\\ p\in\mathcal{T}}} H(p) = \lim_{x\to\infty}\frac{1}{\pi(x)}\sum_{p\leq x} H(p)$$

exists, and moreover

$$\mathcal{N}=\mathcal{N}_1\mathcal{N}_2,$$

where

$$\mathcal{N}_1 = \sum_{\{d_1,\ldots,d_k\}\in D_\gamma} \frac{h_1(d_1)\cdots h_k(d_k)}{\varphi(LCM[d_1,\ldots,d_k])}\lambda(d_1,\ldots,d_k),$$

$$\text{where } h_j(n)=\sum_{d|n}\mu(d)g_j(n/d), \quad (j=1,\ldots,k),$$

$$\mathcal{N}_2 = \prod_{p>\gamma}\widetilde{m}(p), \qquad \text{where} \quad \widetilde{m}(p)=1+\sum_{a=1}^{\infty}\frac{1}{p^{a-1}(p-1)}\sum_{j=1}^{k}(g_j(p^a)-1)\rho_j(p^a).$$

3.4 Further results

Let $R(x) \in \mathbb{R}[x]$ be a polynomial of degree k taking only positive values. Set $\varphi_0(n) = \varphi(n)/n$, so that

$$\varphi_0(n) = \prod_{p|n} \left(1 - \frac{1}{p}\right) = \sum_{d|n} \frac{\mu(d)}{d}.$$

Set $t_n = \varphi_0(R(n))$ and for $Y > 0$, set

$$t_n^{(Y)} = \prod_{\substack{p<Y \\ p|R(n)}} \left(1 - \frac{1}{p}\right). \tag{3.6}$$

Then,

$$\begin{aligned} 0 \leq t_n^{(Y)} - t_n &\leq \prod_{\substack{p<Y \\ p|R(n)}} \left(1 - \frac{1}{p}\right) \left\{1 - \prod_{\substack{Y \leq p < x^{1/k} \\ p|R(n)}} \left(1 - \frac{1}{p}\right)\right\} + O(x^{-1/k}) \\ &= \prod_{\substack{p<Y \\ p|R(n)}} \left(1 - \frac{1}{p}\right) \left\{1 - \exp\left(\sum_{\substack{p|R(n) \\ Y<p\leq x^{1/k}}} \log\left(1 - \frac{1}{p}\right)\right)\right\} + O(x^{-1/k}) \\ &\leq c \sum_{\substack{p|R(n) \\ Y<p<x}} \frac{1}{p} + O\left(\frac{1}{x^{1/k}}\right). \end{aligned}$$

Thus,

$$\sum_{n\leq x} \left(t_n^{(Y)} - t_n\right) \leq O\left(x^{1-1/k}\right) + cx \sum_{p>Y} \frac{\rho(p)}{p^2} \leq O\left(x^{1-1/k}\right) + \frac{cx}{\log Y}.$$

Now, in light of (3.6), we have

$$t_n^{(Y)} = \sum_{\substack{d|R(n) \\ P(d)\leq Y}} \frac{\mu(d)}{d},$$

so that we may write that

$$\sum_{n\leq x} t_n^{(Y)} e(m_1Q_1(n)+\cdots+m_tQ_t(n)) = \sum_{\substack{d\leq x \\ P(d)\leq Y}} \frac{\mu(d)}{d} \sum_{\substack{n\leq x \\ R(n)\equiv 0 \pmod d}} e(m_1Q_1(n)+\cdots+m_tQ_t(n)). \tag{3.7}$$

Since the inner sum on the right hand side of (3.7) runs over some arithmetical progression mod d and since the number of d's is limited by $2^{\pi(Y)}$, then one may conclude that

$$\sum_{n\leq x} t_n^{(Y)} e(m_1Q_1(n) + \cdots + m_tQ_t(n)) = o(x) \qquad \text{as } x \to \infty.$$

This allows us to state the following theorem.

Theorem 13. *The following limits all exist:*

$$\lim_{x\to\infty} \frac{1}{Dx} \sum_{\substack{n\le x\\ n\in\mathcal{T}}} \varphi_0(R(n)) = \lim_{x\to\infty} \frac{1}{x} \sum_{n\le x} \varphi_0(R(n)),$$
$$\lim_{x\to\infty} \frac{1}{D\pi(x)} \sum_{\substack{p\le x\\ p\in\mathcal{T}}} \varphi_0(R(p)) = \lim_{x\to\infty} \frac{1}{\pi(x)} \sum_{p\le x} \varphi_0(R(p)).$$

Note that a similar theorem could be proved for $\sigma_0(n) := \sigma(n)/n$ instead of $\varphi_0(n)$.

4 Open problems

4.1 A question related to the divisor function

Let $\tau(n)$ stand for the number of divisors of n. Is it true that

$$\frac{1}{x\log x} \sum_{n\le x} \tau(n) - \frac{1}{Dx\log x} \sum_{\substack{n\le x\\ n\in\mathcal{T}}} \tau(n) \to 0 \qquad \text{as } x\to\infty$$

or not ?

4.2 A question related to shifted primes

Assume that the necessary conditions of the Erdős-Kac Theorem hold, that is that f is a strongly additive function satisfying $f(p) = O(1)$. Letting

$$A(x) := \sum_{p\le x} \frac{f(p)}{p} \qquad \text{and} \qquad B^2(x) = \sum_{p\le x} \frac{f^2(p)}{p}$$

and setting

$$G_x(z) = \frac{1}{D\pi(x)} \#\{p \le x : p\in\mathcal{T},\ f(p+1) - A(x) < zB(x)\}.$$

Then, is it true that

$$\lim_{x\to\infty} G_x(z) = \Phi(z)$$

for all real numbers z or not ?

References

[1] H. Daboussi and H. Delange, *Quelques propriétés des fonctions multiplicatives de module au plus égal à 1*, C.R. Acad. Sci. Paris, Ser. A **278** (1974), 245-264.

[2] H. Daboussi and H. Delange, *On multiplicative arithmetical functions whose modulus does not exceed one*, J. London Math. Soc. **26** (1982), 245-264.

[3] P.D.T.A. Elliott, *Probabilistic Number Theory II*, Springer, New York, 1979/80.

[4] P. Erdős and P. Turán, *On a problem in the theory of uniform distribution I, II*, Indag. Math. **10** (1948), 370–378, 406–413.

[5] I. Kátai, *On the sum of bounded multiplicative functions over some special subsets of integers*, Uniform Distribution Theory **3** (2008), no.2, 37-43.

[6] I. Kátai, *A remark on a theorem of H. Daboussi*, Acta Math. Hung. **47** (1986), 223-225.

[7] I. Kátai, *On the distribution of arithmetical functions*, Acta Math. Hungarica **20** (1969), 69-87.

[8] I. Kátai, *On the distribution of arithmetical functions on the set of primes plus one*, Compositio Math. **19** (1969), 278-288.

[9] K.H. Indlekofer and I. Kátai, *Exponential sums with multiplicative coefficients*, Acta Math. Hung. **54** (1989), 263-268.

[10] M. Tamaka, *On the number of prime factors of integers I*, Japan J. Math. **25** (1955), 1-20.

[11] I.M. Vinogradov, *The method of trigonometric sums in the theory of numbers*, Translated from the Russian, revised and annotated by K.F. Roth and A. Davenport. Reprint of the 1954 translation. Dover, Mineola, NY, 2004.

[12] H. Weyl, *Ueber die Gleichverteilung von Zahlen mod. Eins*, Math. Ann. 77 (3) (1916), 313–352.

Jean-Marie De Koninck
Dép. de mathématiques et de statistique
Université Laval
Québec
Québec G1V 0A6
Canada
jmdk@mat.ulaval.ca

Imre Kátai
Computer Algebra Department
Eötvös Loránd University
1117 Budapest
Pázmány Péter Sétány I/C
Hungary
katai@inf.elte.hu

JMDK, le 20 octobre 2018; fichier: distribution-arithmetic-functions-subsets.tex