A Turán–Kubilius type inequality on sum sets

by

JOËL RIVAT (Marseille) and ANDRÁS SÁRKÖZY (Budapest)

1. Introduction. \( \mathbb{N}, \mathbb{R} \) and \( \mathbb{C} \) denote the set of positive integers, real numbers, resp. complex numbers. We write \( e(\alpha) = \exp(2i\pi \alpha) \). The letter \( p \) denotes a prime number. \( \omega(n) \) denotes the number of distinct prime factors of \( n \), while \( \Omega(n) \) denotes the number of prime factors of \( n \) counted with multiplicity.

In 1917 Hardy and Ramanujan \cite{8} proved that for almost all positive integers \( m \leq n \) the value of \( \omega(m) \) is “near” \( \log \log n \). Their proof was based on the estimate of the number of positive integers \( m \) with \( m \leq n \) and \( \omega(m) = k \) for any fixed \( k \). In 1934 Turán \cite{13} proved in a simpler way that

\[
\sum_{m \leq n} (\omega(m) - \log \log n)^2 = O(n \log \log n),
\]

from which the result of Hardy and Ramanujan follows immediately. Later Turán extended \cite{11} to general additive arithmetic functions \( f(n) \), and he showed that if \( f(n) \) is a real-valued additive arithmetic function with

\[
f(p) = f(p^2) = \cdots = f(p^k) = \cdots
\]

for every prime number \( p \) (in which case it is said to be strongly additive) and it is bounded:

\[
|f(p)| = O(1),
\]

then, writing

\[
A_f(n) = \sum_{p \leq n} \frac{f(p)}{p},
\]

we have

\[
\sum_{m \leq n} (f(m) - A_f(n))^2 = O(nA_f(n)).
\]

2010 Mathematics Subject Classification: Primary 11K65; Secondary 11N64.
Key words and phrases: additive arithmetic function, probabilistic number theory, sum set, large sieve.
In [9] Kubilius showed that Turán’s conditions \( f(n) \in \mathbb{R}, \) \((2)\) and \((3)\) can be dropped, and still there is an inequality of type \((5)\): if \( f(n) \) is a complex-valued additive arithmetic function, \( A_f(n) \) is defined by \((4)\), and we also write

\[
D_f(n) = \left( \sum_{p^\alpha \leq n} \frac{|f(p^\alpha)|^2}{p^\alpha} \right)^{1/2},
\]

then

\[
\sum_{m \leq n} |f(m) - A_f(n)|^2 = O(nD_f^2(n)).
\]

This is called the Turán–Kubilius inequality.

In the last 25 years numerous papers have been written on the arithmetic properties of sum sets (see, e.g., [1–3], [4], [5], [6], [7], and [11, 12]). Typically, these results say that if \( A, B \) are “large” subsets of \( \{1, \ldots, n\} \) then a certain property of the sums simulates the behaviour of the consecutive integers \( 1, \ldots, n \). In particular, Elliott and Sárközy [5] showed that if \( A, B \) are large subsets of \( \{1, \ldots, n\} \), then the sums \( a + b \) satisfy an Erdős–Kac type theorem.

G. Halász (oral communication) asked whether the Turán–Kubilius inequality has a similar sum set analogue. We will show that, indeed, there is such an inequality which is, however, not quite as strong as \((7)\): we will prove a similar result midway between Turán’s and Kubilius’s inequality.

2. Theorem and comments. We will prove the following theorem:

**Theorem 1.** Let \( f \) be a complex-valued additive arithmetic function, define

\[
K_f(m) = \max\{|f(p^\alpha)| : p \text{ prime}, \alpha \in \mathbb{N}, p^\alpha \leq m\},
\]

let \( A_f(n) \) be defined by \((4)\), \( C \) any fixed positive number, \( n \in \mathbb{N} \) (with \( n \to +\infty \)) and \( A, B \subseteq \{1, \ldots, n\} \) with

\[
\frac{n}{\sqrt{|A||B|}} < \exp(C \sqrt{\log \log n \log \log \log n}).
\]

Then

\[
\frac{1}{|A||B|} \sum_{a \in A} \sum_{b \in B} |f(a + b) - A_f(2n)|^2 = O(C^2 K_f^2(2n) \log \log(2n)).
\]

**Remarks.** (i) If \( f(n) = \omega(n) \) and \( A = B = \mathbb{N} \), then the order of magnitude of the left hand side is the same as that of the function in the ordo term on the right hand side. This shows that in general \((10)\) is sharp apart from a constant factor. Indeed, one could construct much thinner sets \( A, B \) (but not thinner than the lower bound implied by \((9)\)) so that \((10)\) is sharp for these sets \( A, B \) and \( f(n) = \omega(n) \).
(ii) Condition \((9)\) is also sharp, i.e., to ensure that the left hand side of \((10)\) is \(O(K_f^2(2n) \log \log(2n))\) one needs assumption \((9)\). Indeed, fix a positive number \(C\), let \(Q\) denote the products of the primes not exceeding \((C/2)\sqrt{\log \log n} \log \log \log n\), and let \(A = B = \{iQ : i \leq n/Q\}\). Then it follows easily from the prime number theorem (or from a more elementary theorem) that \((9)\) holds. Moreover, let \(f(m) = \omega(m)\) (which implies \(K_f(2n) = 1\)). Then it can be shown that the left hand side of \((10)\) is

\[
\gg C'K_f^2(2n) \log \log(2n) = C' \log \log(2n),
\]

with some \(C' = C'(C)\) such that \(C' \to +\infty\) as \(C \to +\infty\). (The reason is that for a typical pair \(a = iQ \in A, b = jQ \in B\) we have \(Q | a + b = (i + j)Q\), thus every prime \(p \leq (C/2)\sqrt{\log \log n} \log \log \log n\) divides \(a + b\). The number of these primes is \((1 + o(1))C\sqrt{\log \log n}\) so that their contribution makes a typical difference \(f(a + b) - A_f(2n) = \omega(a + b) - A_\omega(2n)\) greater by \((1 + o(1))C\sqrt{\log \log n}\), which adds \(C' \log \log n\) to the left hand side of \((10)\) with some \(C' = C'(C)\) such that \(C' \to +\infty\) as \(C \to +\infty\).)

(iii) While Theorem 1 is sharp for \(f(m) = \omega(m)\), it gives only a very weak upper bound for the left hand side of \((10)\) if \(f(m) = \Omega(m)\). The reason is that the prime powers \(p^\alpha\) with small \(p\) and large \(\alpha\) may influence the distribution of the values \(\Omega(a + b)\) (with \(a \in A, b \in B\)) significantly. E.g. let \(\alpha = \lfloor C' \sqrt{\log \log n}\rfloor\), and set first \(A = B = \{m : m \leq n, 2^\alpha | m\}\) and then \(A = \{m : m \leq n, 2^\alpha | m + 1\}\). In the first case the powers \(2^\beta\) with \(\beta \geq \alpha\) make a contribution \(\gg \sqrt{\log \log n}\) to every value \(f(a + b) = \Omega(a + b)\) \((a \in A, b \in B)\), while in the second case they make no contribution at all. In both cases the left hand side of \((9)\) is \(< \exp(C'' \sqrt{\log \log n})\) so that a stronger inequality holds than \((9)\); however, the left hand side of \((10)\) is \(O(\log \log(2n))\) in the first case, and in the second case it is \(\neq O(\log \log(2n))\) if \(C' \to +\infty\) slowly.

### 3. Structure of the proof

Let \(\mathcal{P}\) denote the set of prime powers \(p^\alpha \leq 2n\), and write

\[
V = \frac{n}{\sqrt{|A||B|}}.
\]

We split \(\mathcal{P}\) into three parts: let

\[
\mathcal{P}_1 = \{p^\alpha : p \leq V, p^\alpha \leq 2n\},
\]

\[
\mathcal{P}_2 = \{p^\alpha : V < p, p^\alpha \leq \sqrt{2n}\},
\]

\[
\mathcal{P}_3 = \{p^\alpha : V < p, \sqrt{2n} < p^\alpha \leq 2n\},
\]

so that \(\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2 \cup \mathcal{P}_3\) and \(\mathcal{P}_i \cap \mathcal{P}_j = \emptyset\) for \(1 \leq i < j \leq 3\). Define the additive arithmetic functions \(f_1(m), f_2(m), f_3(m)\) by
\[ f_i(p^\alpha) = \begin{cases} f(p^\alpha) & \text{if } p^\alpha \in \mathcal{P}_i \\ 0 & \text{if } p^\alpha \notin \mathcal{P}_i \end{cases} \quad \text{(for } i = 1, 2, 3). \]

Then clearly \( f(m) = f_1(m) + f_2(m) + f_3(m) \). Thus by using the elementary inequality \( |z_1 + z_2 + z_3|^2 \leq 3(|z_1|^2 + |z_2|^2 + |z_3|^2) \) (where \( z_1, z_2, z_3 \) are any complex numbers) we can estimate the sum on the left hand side of (10) in the following way:

\[
1 \frac{|A||B|}{|A|} \sum_{a \in A} \sum_{b \in B} |f(a + b) - A_f(2n)|^2 \leq 3(T_1 + T_2 + T_3)
\]

where

\[
T_i = \frac{1}{|A|} \sum_{a \in A} \sum_{b \in B} \left| f_i(a + b) - \sum_{p \leq 2n} \frac{f_i(p)}{p} \right|^2 \quad \text{(for } i = 1, 2, 3). \]

The crucial part of the proof is the estimate of \( T_2 \), which is based on the large sieve; this estimate will be carried out in Sections 4 and 5 \( T_1 \) will be estimated in Section 6, while the (nearly trivial) estimate of \( T_3 \) and the completion of the proof of Theorem 1 will be presented in Section 7.

### 4. The estimate of \( T_2 \). Preliminary lemmas.

For \( A, B \subseteq \{1, \ldots, n\} \) and \( m \in \mathbb{N} \) we define

\[
R(m) = \sum_{(a, b) \in A \times B, \, a + b \equiv 0 \mod m} 1 - \frac{|A| |B|}{m}.
\]

**Lemma 1.** For \( n \in \mathbb{N} \) and \( w \in \{1, 2\} \) let

\[ \mathcal{M}_{n, w} = \{m : 1 \leq m \leq 2n, 1 \leq \omega(m) \leq w, p^\alpha \parallel m \Rightarrow p^\alpha \in \mathcal{P}_2\}. \]

Then for \( n \geq 8 \),

\[
\sum_{m \in \mathcal{M}_{n, w}} |R(m)| \leq 3 \kappa_w(2n) |A| |B|
\]

where

\[
\kappa_w(u) = \max_{k \leq u} \sum_{d \leq u/k, \, 1 \leq \omega(kd) \leq w} \frac{1}{d}.
\]

Moreover,

\[
\kappa_1(u) \leq 2, \quad \kappa_2(u) \leq 3 + \sum_{p \leq u} \frac{1}{p} = O(\log \log u).
\]

**Proof.** We first observe that we may assume

\[
V < \sqrt[n]{2n},
\]

for otherwise \( \mathcal{P}_2 = \emptyset \), thus \( \mathcal{M}_{n, w} = \emptyset \) and (15) is trivially true. We can write
\[ \sum_{m \in M_{n,w}} |R(m)| = \sum_{m \in M_{n,w}} \frac{1}{m} \left| \sum_{1 \leq h < m} \left( \sum_{a \in A} e\left(\frac{ha}{m}\right) \right) \left( \sum_{b \in B} e\left(\frac{hb}{m}\right) \right) \right| \]

and by Cauchy’s inequality we get

(19) \[ \sum_{m \in M_{n,w}} |R(m)| \leq \left( \sum_{m \in M_{n,w}} \frac{1}{m} \sum_{1 \leq h < m} \left| \sum_{a \in A} e\left(\frac{ha}{m}\right) \right|^2 \right)^{1/2} \times \left( \sum_{m \in M_{n,w}} \frac{1}{m} \sum_{1 \leq h < m} \left| \sum_{b \in B} e\left(\frac{hb}{m}\right) \right|^2 \right)^{1/2} . \]

Let us consider the first term on the right hand side of (19). We arrange the summation according to the greatest common factor of \( h \) and \( m \):

\[ \sum_{m \in M_{n,w}} \frac{1}{m} \sum_{1 \leq h < m} \left| \sum_{a \in A} e\left(\frac{ha}{m}\right) \right|^2 = \sum_{m \in M_{n,w}} \frac{1}{m} \sum_{d|l} \sum_{1 \leq h < m} \left( \frac{h}{d} \right) \left| \sum_{a \in A} e\left(\frac{la}{d}\right) \right|^2 . \]

Notice that this last sum is empty for \( d = m \) so we may assume that \( 1 < d < m \). We put \( k = m/d \); observe that \( k > 1 \). Since \( k \mid m \) we have \( 1 \leq \omega(k) \leq w \) and \( p^\alpha \mid k \Rightarrow p^\alpha \in P_2 \), i.e. \( k \in M_{n,w} \). Hence changing the order of summation we get

\[ \sum_{m \in M_{n,w}} \frac{1}{m} \sum_{1 \leq h < m} \left| \sum_{a \in A} e\left(\frac{ha}{m}\right) \right|^2 = \sum_{k \in M_{n,w}} \frac{1}{k} \sum_{d|l} \sum_{1 \leq k < \omega(kd)} \left| \sum_{a \in A} e\left(\frac{la}{k}\right) \right|^2 . \]

Using (16) we get

\[ \sum_{m \in M_{n,w}} \frac{1}{m} \sum_{1 \leq h < m} \left| \sum_{a \in A} e\left(\frac{ha}{m}\right) \right|^2 \leq \kappa_w(2n) \sum_{k \in M_{n,w}} \frac{1}{k} \sum_{1 \leq l < k} \left| \sum_{a \in A} e\left(\frac{la}{k}\right) \right|^2 . \]

Let \( z_j = 2n/2^j \). For \( k \in M_{n,w} \) with \( w \in \{1, 2\} \), by the definition of \( P_2 \) we have \( V < k \) and \( k \leq \left( \frac{\sqrt{2n}}{2} \right)^2 = \sqrt{2n} \). It follows that if there is a \( k \in M_{n,w} \) with \( z_{j+1} < k \leq z_j \), then \( V < z_j = 2n/2^j \) and \( z_{j+1} = 2n/2^{j+1} < \sqrt{2n} \), whence \( \sqrt{n/2} < 2^j < 2n/V \). Thus

\[ \sum_{k \in M_{n,w}} \frac{1}{k} \sum_{1 \leq l < k} \left| \sum_{a \in A} \right| e\left(\frac{la}{k}\right) \right|^2 \leq \sum_{\sqrt{n/2} < 2^j < 2n/V} \frac{1}{z_j} \sum_{z_{j+1} < k \leq z_j} \sum_{1 \leq l < k} \left| \sum_{a \in A} e\left(\frac{la}{k}\right) \right|^2 . \]
For fixed $j$ the points $l/k$ with $(l, k) = 1$ are at least $(z_{j+1}^{-1})^{-2}$ spaced modulo 1 and $a \in \mathcal{A}$ satisfies $1 \leq a \leq n$, hence by the large sieve inequality (see for example [10]),

$$\sum_{z_{j+1} \leq k \leq z_j} \sum_{1 \leq l < k \atop (l,k)=1} \left| \sum_{a \in \mathcal{A}} e\left(\frac{la}{k}\right) \right|^2 \leq (n - 1 + z_{j+1}^2)|\mathcal{A}|.$$  

But

$$\sum_{\sqrt{n/2} < 2^j < n/V} \frac{1}{z_j} (n - 1 + z_{j+1}^2)$$

$$= \frac{n - 1}{2n} \sum_{\sqrt{n/2} < 2^j < n/V} 2^j + \frac{2n}{\sqrt{n/2} < 2^j < n/V} \sum_{\sqrt{n/2} < 2^j < n/V} \frac{1}{2j + 2}$$

$$\leq \frac{2n}{V} + \sqrt{2n},$$

and by (18),

$$\frac{2n}{V} + \sqrt{2n} = \frac{2n}{V} \left(1 + \frac{V}{\sqrt{2n}}\right) \leq \frac{2n}{V} \left(1 + \frac{\sqrt{2n}}{\sqrt{2n}}\right) \leq 3n/V$$

for $n \geq 8$, so that

$$\sum_{m \in \mathcal{M}_{n,w}} \frac{1}{m} \sum_{1 \leq h < m} \left| \sum_{a \in \mathcal{A}} e\left(\frac{ha}{m}\right) \right|^2 \leq \frac{3n}{V} \kappa_w(2n)|\mathcal{A}|.$$  

We may replace $\mathcal{A}$ by $\mathcal{B}$ in the argument above and get

$$\sum_{m \in \mathcal{M}_{n,w}} \frac{1}{m} \sum_{1 \leq h < m} \left| \sum_{b \in \mathcal{B}} e\left(\frac{hb}{m}\right) \right|^2 \leq \frac{3n}{V} \kappa_w(2n)|\mathcal{B}|.$$  

Finally, if we apply these two estimates in (19) and use the definition of $V$ given by (11), we get (15):

$$\sum_{m \in \mathcal{M}_{n,w}} |R(m)| \leq \frac{3n}{V} \kappa_w(2n) \sqrt{|\mathcal{A}| |\mathcal{B}|} = 3\kappa_w(2n)|\mathcal{A}| |\mathcal{B}|.$$  

It remains to prove (17). When $w = 1$, we observe that $k$ in (16) is a prime power $p^\gamma$, and $d$ must be a power of the same prime number $p$ (say $d = p^\beta$). More precisely, by the definition (16) we have

$$\kappa_1(u) = \max_{k \leq u} \sum_{d \leq u/k} \frac{1}{d} = \max_{p^\gamma \leq u} \sum_{\beta \geq 0} \frac{1}{p^\beta} \leq \max_p \sum_{\beta \geq 0} \frac{1}{p^\beta} = \max_p \frac{1}{1 - 1/p} = 2,$$

which establishes the first statement of (17). The second is slightly more
complicated. Observe that \( \omega(kd) \geq \omega(d) \), hence by (16) we have

\[
\kappa_2(u) = \max \left( \max_{k \leq u} \frac{1}{d}, \max_{k \leq u} \frac{1}{d} \right).
\]

Now if \( \omega(k) = 2 \) then \( k = p^\gamma p'^\gamma \) and \( d \) must be of the form \( d = p^\beta p'^\beta \) with the same pair of prime numbers \( (p, p') \) and \( \beta, \beta' \geq 0 \). Therefore

\[
\kappa_2(u) = \max_{k \leq u} \sum_{\omega(k) = 2} \sum_{d \leq u/k} \frac{1}{d} \leq \max_{k \leq u} \sum_{d \leq u/k} \frac{1}{p^{\beta}} = \frac{1}{1 - 1/2} \frac{1}{1 - 1/3} = 3.
\]

Moreover

\[
\kappa_1(u) + \max_{k \leq u} \sum_{\omega(k) = 1} \sum_{d \leq u/k} \frac{1}{d}.
\]

If \( \omega(k) = 1 \) and \( \omega(kd) = 2 \) this implies that \( d \) is a prime power (say \( d = p^\beta \)) coprime to \( k \). Hence if \( \omega(k) = 1 \) we have

\[
\sum_{d \leq u/k} \frac{1}{d} \leq \sum_{p \leq u/k} \frac{1}{p^{\beta}} = \sum_{p \leq u/k} \frac{1}{p} + \sum_{p \leq u/k} \frac{1}{p(p-1)} \leq \sum_{p \leq u/k} \frac{1}{p} + 1.
\]

Finally, we obtain

\[
\kappa_2(u) \leq \max \left( \kappa_1(u) + 1 + \sum_{p \leq u} \frac{1}{p}, 3 \right) \leq 3 + \sum_{p \leq u} \frac{1}{p},
\]

which is the second statement of (17). ■

**Lemma 2.** For any complex-valued additive arithmetic function \( f_2 \) such that \( f_2(p^\alpha) = 0 \) whenever \( p^\alpha \not\in \mathcal{P}_2 \) and \( n \geq 8 \) we have

\[
(20) \quad \left| \sum_{a \in A} \sum_{b \in B} f_2(a+b) - |A||B| \sum_{p^\alpha \leq 2n} \frac{f_2(p^\alpha)}{p^\alpha} \left(1 - \frac{1}{p}\right) \right| \leq 14Kf_2(2n)|A||B|
\]

where \( K_{f_2} \) is defined by (8).

Proof. Let

\[
(21) \quad S_1 = \sum_{a \in A} \sum_{b \in B} f_2(a + b).
\]
Since $f_2$ is an additive arithmetic function we have

$$S_1 = \sum_{a \in A} \sum_{b \in B} \sum_{p^\alpha \mid a+b} f_2(p^\alpha) = \sum_{p^\alpha \leq 2n} f_2(p^\alpha) \sum_{(a,b) \in A \times B} 1,$$

and

$$\sum_{(a,b) \in A \times B} 1 = \sum_{p^\alpha \mid a+b} 1 - \sum_{p^\alpha+1 \mid a+b} 1.$$

Using (14) we have

$$\left| S_1 - |A||B| \sum_{p^\alpha \leq 2n} f_2(p^\alpha) \left(1 - \frac{1}{p}\right) + \sum_{p^\alpha \leq 2n} f_2(p^\alpha) R(p^\alpha) - \sum_{p^\alpha \leq 2n} f_2(p^\alpha) R(p^\alpha+1) \right|$$

so that using (8) and $f_2(p^\alpha) = 0$ whenever $p^\alpha \not\in \mathcal{P}_2$ we obtain

$$\left| S_1 - |A||B| \sum_{p^\alpha \leq 2n} f_2(p^\alpha) \left(1 - \frac{1}{p}\right) \right| \leq K f_2(2n) \left( \sum_{p^\alpha \in \mathcal{P}_2} |R(p^\alpha)| + \sum_{p^\alpha \in \mathcal{P}_2} |R(p^\alpha+1)| \right).$$

By (15) and (17) we have

$$\sum_{p^\alpha \in \mathcal{P}_2} |R(p^\alpha)| \leq 6 |A||B|$$

and

$$\sum_{p^\alpha \in \mathcal{P}_2} |R(p^\alpha+1)| = \sum_{p^\alpha \in \mathcal{P}_2, p^\alpha+1 \not\in \mathcal{P}_2} |R(p^\alpha+1)| + \sum_{p^\alpha+1 \in \mathcal{P}_2} |R(p^\alpha+1)|$$

$$\leq \sum_{p^\alpha \in \mathcal{P}_2, p^\alpha+1 \not\in \mathcal{P}_2} |R(p^\alpha+1)| + 6 |A||B|.$$

Using (14) we have

$$|R(p^\alpha+1)| \leq \sum_{(a,b) \in A \times B, a+b \equiv 0 \mod p^\alpha+1} 1 + \frac{|A||B|}{p^\alpha+1}$$
and counting trivially we can write
\[ \sum_{(a,b)\in A\times B, a+b\equiv 0 \mod p^{\alpha+1}} 1 \leq \min \left( \frac{n|A|}{p^{\alpha+1}}, \frac{n|B|}{p^{\alpha+1}} \right) \leq \frac{n}{p^{\alpha+1}} \sqrt{|A||B|}, \]
so that
\[ \sum_{p^\alpha \in \mathcal{P}_2, p^{\alpha+1} \notin \mathcal{P}_2} |R(p^{\alpha+1})| \leq (n\sqrt{|A||B|} + |A||B|) \sum_{p^\alpha \in \mathcal{P}_2, p^{\alpha+1} \notin \mathcal{P}_2} \frac{1}{p^{\alpha+1}}. \]

We will show that
\[ (24) \sum_{p^\alpha \in \mathcal{P}_2, p^{\alpha+1} \notin \mathcal{P}_2} \frac{1}{p^{\alpha+1}} \leq \frac{1}{V}. \]
Indeed, if $p^\alpha \in \mathcal{P}_2$ and $p^{\alpha+1} \notin \mathcal{P}_2$ then by the definition of $\mathcal{P}_2$ we have $V < p$ and $p^\alpha \leq \sqrt{2n} < p^{\alpha+1}$ so that $\alpha$ is uniquely defined. Hence
\[ \sum_{p^\alpha \in \mathcal{P}_2, p^{\alpha+1} \notin \mathcal{P}_2} \frac{1}{p^{\alpha+1}} \leq \sum_{p > V} \frac{1}{p^2}. \]
But for $n_0 \in \mathbb{N}$,
\[ \sum_{n \geq n_0} \frac{1}{(2n+1)^2} \leq \sum_{n \geq n_0} \frac{1}{4} \left( \frac{1}{n} - \frac{1}{n+1} \right) \leq \frac{1}{4n_0}, \]
so that if $V \geq 2$ then
\[ \sum_{p > V} \frac{1}{p^2} = \sum_{p \geq [V]+1} \frac{1}{p^2} \leq \sum_{n \geq [V]/2} \frac{1}{(2n+1)^2} \leq \frac{1}{2[V]} \leq \frac{1}{2V - 2} \leq \frac{1}{V}, \]
and if $1 \leq V < 2$ then
\[ \sum_{p > V} \frac{1}{p^2} = \sum_{p \geq 2} \frac{1}{p^2} \leq \frac{1}{4} + \sum_{n \geq 1} \frac{1}{(2n+1)^2} \leq \frac{1}{4} + \frac{1}{4} \leq \frac{1}{V}. \]

By these estimates and the definition of $V$ given by (11) we get
\[ \sum_{p^\alpha \in \mathcal{P}_2, p^{\alpha+1} \notin \mathcal{P}_2} |R(p^{\alpha+1})| \leq \frac{n\sqrt{|A||B|} + |A||B|}{V} \leq 2|A||B|. \]

The sum of the last three upper bounds above gives
\[ (25) \sum_{p^\alpha \in \mathcal{P}_2} |R(p^\alpha)| + \sum_{p^{\alpha+1} \notin \mathcal{P}_2} |R(p^{\alpha+1})| \leq 14|A||B|. \]
Now (20) follows from (23) and (25). ■
Lemma 3. For any complex-valued additive arithmetic function \( f_2 \) such that \( f_2(p^\alpha) = 0 \) whenever \( p^\alpha \not\in \mathcal{P}_2 \), and \( n \geq 8 \), we have

\[
\left| \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} |f_2(a + b)|^2 - |\mathcal{A}| \cdot |\mathcal{B}| \right| \sum_{p^\alpha \leq 2n} \frac{f_2(p^\alpha)}{p^\alpha} \left( 1 - \frac{1}{p} \right)^2 \ll K_{f_2}^2 (2n)|\mathcal{A}| \cdot |\mathcal{B}| \log \log (2n)
\]

where \( K_{f_2} \) is defined by (8).

Proof. Let

\[
S_2 = \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} |f_2(a + b)|^2.
\]

Since \( f_2 \) is an additive arithmetic function we can write

\[
S_2 = \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} \left| \sum_{p^\alpha \parallel a + b} f_2(p^\alpha) \right|^2
\]

and expanding the square we get

\[
S_2 = \sum_{p^\alpha \leq 2n} \left| f_2(p^\alpha) \right|^2 \sum_{(a, b) \in \mathcal{A} \times \mathcal{B}} 1 + \sum_{p^\alpha q^\beta \leq 2n} \sum_{p \neq q} f_2(p^\alpha) \overline{f_2(q^\beta)} \sum_{(a, b) \in \mathcal{A} \times \mathcal{B}} 1.
\]

First we will give an upper bound for the first term. Using (22) we can write

\[
\sum_{p^\alpha \leq 2n} \left| f_2(p^\alpha) \right|^2 \sum_{(a, b) \in \mathcal{A} \times \mathcal{B}} 1 \leq \sum_{p^\alpha \leq 2n} \left| f_2(p^\alpha) \right|^2 \left( \frac{|\mathcal{A}| \cdot |\mathcal{B}|}{p^\alpha} \left( 1 - \frac{1}{p} \right) + R(p^\alpha) - R(p^{\alpha + 1}) \right)
\]

\[
\leq K_{f_2}^2 (2n) \sum_{p^\alpha \in \mathcal{P}_2} \left( \frac{|\mathcal{A}| \cdot |\mathcal{B}|}{p^\alpha} \left( 1 - \frac{1}{p} \right) + |R(p^\alpha)| + |R(p^{\alpha + 1})| \right).
\]

Now

\[
\sum_{p^\alpha \in \mathcal{P}_2} \frac{|\mathcal{A}| \cdot |\mathcal{B}|}{p^\alpha} \left( 1 - \frac{1}{p} \right) \leq |\mathcal{A}| \cdot |\mathcal{B}| \sum_{p \leq 2n} \sum_{\alpha \geq 1} \frac{1}{p^\alpha} \left( 1 - \frac{1}{p} \right)
\]

\[
= |\mathcal{A}| \cdot |\mathcal{B}| \sum_{p \leq 2n} \frac{1}{p} = O(|\mathcal{A}| \cdot |\mathcal{B}| \log \log (2n)),
\]

while by (25) we have

\[
\sum_{p^\alpha \in \mathcal{P}_2} |R(p^\alpha)| + \sum_{p^\alpha \in \mathcal{P}_2} |R(p^{\alpha + 1})| \leq 14 |\mathcal{A}| \cdot |\mathcal{B}|.
\]
Using these two estimates we obtain, from \(28\),

\[
S_2 = \sum_{p^\alpha q^\beta \leq 2n \atop p \neq q} f_2(p^\alpha) f_2(q^\beta) \sum_{(a,b) \in A \times B \atop p^\alpha \| a+b \atop q^\beta \| a+b} 1 + O(K^2_{f_2}(2n)|A||B| \log \log(2n)).
\]

For \(q \neq p\),

\[
\sum_{(a,b) \in A \times B \atop p^\alpha \| a+b \atop q^\beta \| a+b} 1 = \sum_{(a,b) \in A \times B \atop p^\alpha q^\beta \| a+b \atop q^\beta+1 \| a+b} 1 - \sum_{(a,b) \in A \times B \atop p^\alpha q^\beta+1 \| a+b \atop q^\beta \| a+b} 1 + \sum_{(a,b) \in A \times B \atop p^\alpha q^\beta+1 \| a+b \atop q^\beta+1 \| a+b} 1,
\]

thus using \(14\) we get

\[
\sum_{(a,b) \in A \times B \atop p^\alpha q^\beta \| a+b \atop q^\beta \| a+b} 1 = \frac{|A||B|}{p^\alpha q^\beta} \left(1 - \frac{1}{p}\right) \left(1 - \frac{1}{q}\right) + R(p^\alpha q^\beta) - R(p^{\alpha+1} q^\beta) - R(p^\alpha q^{\beta+1}) + R(p^{\alpha+1} q^{\beta+1}).
\]

Writing

\[
S'_2 = \sum_{p^\alpha q^\beta \leq 2n \atop p \neq q} \frac{f_2(p^\alpha) f_2(q^\beta)}{p^\alpha q^\beta} \left(1 - \frac{1}{p}\right) \left(1 - \frac{1}{q}\right)
\]

and

\[
R_2 = \sum_{p^\alpha \in P_2, q^\beta \in P_2 \atop p \neq q} (|R(p^\alpha q^\beta)| + |R(p^{\alpha+1} q^\beta)| + |R(p^\alpha q^{\beta+1})| + |R(p^{\alpha+1} q^{\beta+1})|),
\]

we deduce from \(29\) and \(30\) that

\[
|S_2 - |A||B|S'_2| \ll K^2_{f_2}(2n)(|A||B| \log \log(2n) + R_2).
\]

By \(15\) and \(17\) we have

\[
\sum_{p^\alpha \in P_2, q^\beta \in P_2 \atop p \neq q} |R(p^\alpha q^\beta)| \ll |A||B| \log \log(2n).
\]

Using \(14\) we have

\[
|R(p^{\alpha+1} q^\beta)| \leq \sum_{(a,b) \in A \times B \atop a+b \equiv 0 \mod p^{\alpha+1} q^\beta} 1 + \frac{|A||B|}{p^{\alpha+1} q^\beta}.
\]
and counting trivially we can write
\[ \sum_{(a,b) \in A \times B} 1 \leq \min \left( \frac{n|A|}{p^{\alpha+1}q^\beta}, \frac{n|B|}{p^{\alpha+1}q^\beta} \right) \leq \frac{n}{p^{\alpha+1}q^\beta} \sqrt{|A||B|}, \]
thus using (24)
\[ \sum_{p^\alpha \in P_2} \sum_{q^\beta \in P_2 \atop q^\beta \neq p} |R(p^{\alpha+1}q^\beta)| \ll n\sqrt{|A||B|} \sum_{p^\alpha \in P_2} \frac{1}{p^{\alpha+1}} \sum_{q^\beta \in P_2 \atop q^\beta \neq p} \frac{1}{q^\beta} \ll \frac{n\sqrt{|A||B|}}{V} \log \log n, \]
and by the definition of V given by (11) we get
\[ \sum_{p^\alpha \in P_2} \sum_{q^\beta \in P_2 \atop q^\beta \neq p} |R(p^{\alpha+1}q^\beta)| \ll |A||B| \log \log n. \]
Similarly
\[ \sum_{q^\beta \in P_2} \sum_{p^\alpha \in P_2 \atop p^\alpha \neq q} |R(p^\alpha q^{\beta+1})| \ll |A||B| \log \log n \]
and
\[ \sum_{p^\alpha \in P_2} \sum_{q^\beta \in P_2 \atop q^\beta \neq p} |R(p^{\alpha+1}q^{\beta+1})| \ll |A||B|. \]
Thus it follows from (31) that
\[ |S_2 - |A||B|S'_2| \ll K^2_{f_2}(2n)|A||B| \log \log(2n). \]
In order to prove (26) it is sufficient to show that
\[ \left| \sum_{p^\alpha \leq 2n} \frac{f_2(p^\alpha)}{p^\alpha} \left( 1 - \frac{1}{p} \right) \right|^2 = S'_2 + O(K^2_{f_2}(2n)). \]
We write
\[ \left| \sum_{p^\alpha \leq 2n} \frac{f_2(p^\alpha)}{p^\alpha} \left( 1 - \frac{1}{p} \right) \right|^2 = \sum_{p^\alpha \leq 2n} \frac{f_2(p^\alpha)}{p^\alpha q^\beta} \left( 1 - \frac{1}{p} \right) \left( 1 - \frac{1}{q} \right) \]
\[ = \sum_{p^\alpha \leq 2n} \cdots + \sum_{p^\alpha q^\beta \leq 2n} \cdots + \sum_{p^\alpha q^\beta q^\gamma \leq 2n} \cdots. \]
By the definition of $f_2$ we have $f_2(p^\alpha)\overline{f_2(q^{\beta})} \neq 0$ only if $p^\alpha \leq \sqrt{2n}$ and $q^\beta \leq \sqrt{2n}$. This implies that the third sum above is empty ($p^\alpha q^\beta > 2n$ is not possible). $S'_2$ is the second sum above. The first sum can be majorized easily:

$$
\left| \sum_{\substack{p^\alpha \leq 2n \\ q^{\beta} \leq 2n \\ p=q}} \frac{f_2(p^\alpha)\overline{f_2(q^{\beta})}}{p^\alpha q^{\beta}} \left(1 - \frac{1}{p}\right)\left(1 - \frac{1}{q}\right) \right|
\leq K^2 f_2(2n) \sum_p \left(1 - \frac{1}{p}\right)^2 \sum_{\alpha \geq 1} \frac{1}{p^\alpha} \sum_{\beta \geq 1} \frac{1}{p^\beta} = K^2 f_2(2n) \sum_p \frac{1}{p^2} = O(K^2 f_2(2n)).
$$

This completes the proof of (26). $\blacksquare$

5. Completion of the estimate of $T_2$. Our first step is to replace the function $A f_2(n) = \sum_{p \leq n} f_2(p) / p$ in the definition of $T_2$ by

$$
E f_2(n) = \sum_{p^\alpha \leq n} \frac{f_2(p^\alpha)}{p^\alpha} \left(1 - \frac{1}{p}\right).
$$

We have

$$
E f_2(n) - A f_2(n) = \sum_{p^\alpha \leq n \\ \alpha \geq 2} \frac{f_2(p^\alpha)}{p^\alpha} - \sum_{p^\alpha \leq n} \frac{f_2(p^\alpha)}{p^\alpha+1},
$$

so that

$$
|E f_2(n) - A f_2(n)| \leq 2K f_2(n) \sum_{p^\alpha \leq n \\ \alpha \geq 2} \frac{1}{p^\alpha}.
$$

Observing that

$$
\sum_{p^\alpha \leq n \\ \alpha \geq 2} \frac{1}{p^\alpha} \leq \sum_p \sum_{\alpha \geq 2} \frac{1}{p^\alpha} = \sum_p \frac{1}{p(p-1)} \leq \sum_{n \geq 2} \frac{1}{n(n-1)} = 1,
$$

we obtain

$$
|E f_2(n) - A f_2(n)| \leq 2K f_2(n).
$$

Using the inequality $|u + v|^2 \leq 2|u|^2 + 2|v|^2$ with $u = f_2(a+b) - E f_2(2n)$ and $v = E f_2(2n) - A f_2(2n)$ (so that $|v| \leq 2K f_2(2n)$), we get

$$
(32) \sum_{a \in A} \sum_{b \in B} |f_2(a+b) - A f_2(2n)|^2 \leq 2 \sum_{a \in A} \sum_{b \in B} |f_2(a+b) - E f_2(2n)|^2 + 8|A||B|K^2 f_2(2n).
$$
Now we will prove
\[ (33) \sum_{a \in A} \sum_{b \in B} |f_2(a + b) - E_{f_2}(2n)|^2 = O(K_{f_2}(2n)|A||B| \log \log(2n)). \]

We have
\[
\sum_{a \in A} \sum_{b \in B} |f_2(a + b) - E_{f_2}(2n)|^2 \\
= S_2 - S_1 \overline{E_{f_2}(2n)} - S_1 E_{f_2}(2n) + |A||B| |E_{f_2}(2n)|^2,
\]
where \( S_1 \) and \( S_2 \) are defined by (21) and (27) respectively. We can rewrite this as
\[
\sum_{a \in A} \sum_{b \in B} |f_2(a + b) - E_{f_2}(2n)|^2 \\
= (S_2 - |A||B| |E_{f_2}(2n)|^2) - (S_1 - |A||B| E_{f_2}(2n)) \overline{E_{f_2}(2n)} \\
- (S_1 - |A||B| E_{f_2}(2n)) E_{f_2}(2n).
\]

By Lemma 2 we have
\[
|S_1 - |A||B| E_{f_2}(2n)| \ll K_{f_2}(2n)|A||B|,
\]
and Lemma 3 yields
\[
|S_2 - |A||B| E_{f_2}(2n)|^2 \ll K_{f_2}^2(2n)|A||B| \log \log(2n),
\]
thus we obtain
\[
\sum_{a \in A} \sum_{b \in B} |f_2(a + b) - E_{f_2}(2n)|^2 \\
\ll K_{f_2}(2n)(K_{f_2}(2n) \log \log(2n) + |E_{f_2}(2n)|)|A||B|.
\]

Now observing that
\[
|E_{f_2}(2n)| \leq |A_{f_2}(2n)| + 2K_{f_2}(2n) \\
\leq K_{f_2}(2n) \sum_{p \leq 2n} \frac{1}{p} + 2K_{f_2}(2n) \ll K_{f_2}(2n) \log \log(2n),
\]
we get (33). It follows from (32) and (33) that
\[ (34) \quad T_2 = O(K_{f_2}(2n) \log \log(2n)). \]

6. The estimate of \( T_1 \). Let \( \omega_V(m) \) be the number of distinct prime factors of \( m \) not exceeding \( V \):
\[
\omega_V(m) = \sum_{\substack{p|m \leq V \atop p|m}} 1.
\]
Then for all $m \leq 2n$ we have
\[
|f_1(m)| = \left| \sum_{p \mid m, p \in \mathcal{P}_1} f_1(p^\alpha) \right| \leq \sum_{p \mid m, p \in \mathcal{P}_1} |f_1(p^\alpha)|,
\]
so that
\[
|f_1(m)| \leq K_{f_1}(2n) \sum_{p \mid m, p \in \mathcal{P}_1} 1 = K_{f_1}(2n) \sum_{p \mid m} 1 = K_{f_1}(2n) \omega_V(m).
\]
Moreover by (8), (11) and (9) we have
\[
\left| \sum_{p \leq 2n} f_1(p) \right| \leq K_{f_1}(2n) \sum_{p \leq V} \frac{1}{p} \ll K_{f_1}(2n) \log \log V \ll K_{f_1}(2n) \log \log \log n.
\]
Using the inequality $|z_1 + z_2|^2 \leq 2(|z_1|^2 + |z_2|^2)$ it follows that for $a \in \mathcal{A}$ and $b \in \mathcal{B},$
\[
\left| f_1(a + b) - \sum_{p \leq 2n} \frac{f_1(p)}{p} \right|^2 \leq 2 \left| f_1(a + b) \right|^2 + 2 \left| \sum_{p \leq 2n} \frac{f_1(p)}{p} \right|^2 \leq 2 K_{f_1}(2n) \omega_V^2(a + b) + O(K_{f_1}(2n)(\log \log \log n)^2),
\]
so that
\[
T_1 = \frac{1}{|\mathcal{A}| |\mathcal{B}|} \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} \left| f_1(a + b) - \sum_{p \leq 2n} \frac{f_1(p)}{p} \right|^2 \leq 2 K_{f_1}(2n) \omega_V^2(a + b) + O(K_{f_1}(2n)(\log \log \log n)^2).
\]
We split this double sum into two parts:
\[ (35) \quad T_1 \leq 2 K_{f_1}(2n)(X_1 + X_2) \]
where
\[
X_1 = \frac{1}{|\mathcal{A}| |\mathcal{B}|} \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} \omega_V^2(a + b), \quad \omega_V(a+b) \leq 5C(\log \log(2n))^{1/2}
\]
and
\[
X_2 = \frac{1}{|\mathcal{A}| |\mathcal{B}|} \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} \omega_V^2(a + b), \quad \omega_V(a+b) > 5C(\log \log(2n))^{1/2}
\]
Then clearly we have
\[ (36) \quad X_1 \leq \frac{1}{|\mathcal{A}| |\mathcal{B}|} \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} (5C(\log \log(2n))^{1/2})^2 = 25C^2 \log \log(2n). \]
In order to estimate $X_2$, we may assume that

$$|A| \leq |B|.$$  \hfill (37)

Then we have

$$X_2 = \frac{1}{|A||B|} \sum_{a \in A} \sum_{b \in B} \omega_V^2(a + b) \quad \omega_V(a + b) > 5C(\log \log(2n))^{1/2}$$

$$\leq \frac{1}{|A||B|} \sum_{a \in A} \sum_{m \leq 2n} \omega_V^2(m) \quad \omega_V(m) > 5C(\log \log(2n))^{1/2}$$

$$\leq \frac{1}{|B|} \sum_{m \leq 2n} \omega_V^2(m).$$  \hfill (38)

The last sum can be rewritten as

$$\sum_{\omega_V(m) > 5C(\log \log(2n))^{1/2}} \omega_V^2(m) = \sum_{t > 5C(\log \log(2n))^{1/2}} \sum_{\omega_V(m) = t} t^2 \sum_{m \leq 2n} \frac{1}{\omega_V(m)}.$$  \hfill (39)

Denote the smallest prime factor of a positive integer $i$ by $p(i)$ (and $p(1) = 1$). If an integer $m$ is counted in the inner sum, then there are prime powers $q_1^{a_1}, \ldots, q_t^{a_t} \leq 2n$ and an integer $r$ such that $q_1 < \cdots < q_t \leq V$, $p(r) > V$ and $q_1^{a_1} \cdots q_t^{a_t} r = m \leq 2n$. Thus the last sum in (39) is

$$\leq \sum_{m \leq 2n} 1 \leq \sum_{q_1^{a_1} \cdots q_t^{a_t} \leq 2n / q_1^{a_1} \cdots q_t^{a_t}} 1 \leq \sum_{q_1^{a_1} \cdots q_t^{a_t} \leq 2n} 1 \leq \sum_{q_1^{a_1} \cdots q_t^{a_t} \leq 2n} \frac{2n}{q_1^{a_1} \cdots q_t^{a_t}} \leq 2n \left( \sum_{q \leq V, q^a \leq 2n} \frac{1}{q^a} \right)^t \cdot \frac{1}{t!} \leq 2n \frac{(1 + o(1)) \log \log V^t}{t!}.$$  \hfill (39)

Inserting this estimate in (39) we get

$$\sum_{m \leq 2n} \omega_V^2(m) \ll n \sum_{t > 5C(\log \log(2n))^{1/2}} \frac{(1 + o(1)) \log \log V^t}{(t - 2)!}.$$
Hence, by the definition of $V$ given by (11), (9) and Stirling’s formula, for large $n$ we have

$$\sum_{m \leq 2n} \frac{\omega^2_V(m)}{\omega_V(m) > 5C(\log \log(2n))^{1/2}} \ll n \sum_{t > 5C(\log \log(2n))^{1/2}} ((1/2 + o(1)) \log \log(2n))^{t} (3/t)^{t-2} \ll n \exp(-(1 + o(1))5C(\log \log(2n))^{1/2} \log(5C(\log \log(2n))^{1/2})) \ll n \exp(-(1 + o(1))5C(\log \log(2n))^{1/2} \log \log(2n)),$$

so that

$$\sum_{m \leq 2n} \frac{\omega^2_V(m)}{\omega_V(m) > 5C(\log \log(2n))^{1/2}} \ll n \exp(-2C(\log \log(2n))^{1/2} \log \log(2n)).$$

It follows from (9), (37), (38) and (40) that

$$X_2 \ll \frac{n}{\sqrt{|A||B|}} \exp(-2C(\log \log(2n))^{1/2} \log \log(2n)) \ll \exp(-C(\log \log(2n))^{1/2} \log \log(2n)) = o(1).$$

Combining (35), (36) and (41) we obtain

$$T_1 = O(C^2K^2f_3(2n) \log \log(2n)).$$

7. The estimate of $T_3$ and the completion of the proof of Theorem 1

If $m \leq 2n$ then

$$|f_3(m)| = \left| \sum_{p^\alpha \parallel m} f_3(p^\alpha) \right| \leq \sum_{p^\alpha \parallel m} |f_3(p^\alpha)|,$$

so that, since $f_3(p^\alpha) = 0$ whenever $p^\alpha \not\in P_3$ and using (8),

$$|f_3(m)| \leq \sum_{p^\alpha \in P_3} Kf_3(2n) = Kf_3(2n) \sum_{p^\alpha \in P_3} 1.$$

Here the last sum is $\leq 3$ since otherwise we would have

$$m \geq \prod_{p^\alpha \in P_3} p^\alpha > \prod_{p^\alpha \in P_3} \sqrt[\sqrt{2n}]{} \geq 4^{\sqrt{2n}} = 2n,$$

which contradicts our assumption $m \leq 2n$. Thus it follows from (43) that

$$|f_3(a + b)| \leq 3Kf_3(2n) \quad \text{for all} \ a \in A, \ b \in B.$$
Moreover
\[ \left| \sum_{p \leq 2n} \frac{f_3(p)}{p} \right| \leq \sum_{p \leq 2n} \left| \frac{f_3(p)}{p} \right| \leq K_{f_3}(2n) \sum_{p \in P_3} \frac{1}{p}, \]
so that
\[ (45) \quad \left| \sum_{p \leq 2n} \frac{f_3(p)}{p} \right| \leq K_{f_3}(2n) \sum_{\mathcal{P} < p \leq 2n} \frac{1}{p} = O(K_{f_3}(2n)). \]

It follows from (44) and (45) that
\[ (46) \quad T_3 = \frac{1}{|A||B|} \sum_{a \in A} \sum_{b \in B} \left| O(K_{f_3}(2n)) + O(K_{f_3}(2n)) \right|^2 = O(K_{f_3}^2(2n)). \]

(10) follows from (12), (34), (42) and (46), observing that \( K_f(2n) = \max(K_{f_1}(2n), K_{f_2}(2n), K_{f_3}(2n)) \), and this completes the proof of Theorem 1.

Acknowledgements. The authors would like to thank the referee for the careful reading of the manuscript and his/her very helpful and valuable comments.

This research was partially supported by the Hungarian National Foundation for Scientific Research, Grants No K67676 and K72731, and by “Balaton” French-Hungarian exchange program F-48/06.

References

A Turán–Kubilius type inequality


Joël Rivat
Institut de Mathématiques de Luminy
CNRS-UMR 6206
Université de la Méditerranée
163 avenue de Luminy, Case 907
13288 Marseille Cedex 9, France
E-mail: rivat@iml.univ-mrs.fr

András Sárközy
Department of Algebra and Number Theory
Eötvös Loránd University
Pázmány Péter sétány 1/c
H-1117 Budapest, Hungary
E-mail: sarkozy@cs.elte.hu

Received on 7.11.2008
and in revised form on 12.10.2009 (5853)