ACTA ARITHMETICA
XCI.2 (1999)

On the quotient sequence of sequences of integers

by

RUDOLF AHLSWEDE (Bielefeld), LEVON H. KHACHATRIAN (Bielefeld),
and ANDRÁS SÁRKÖZY (Budapest)

1. Introduction. The set of positive integers is denoted by $\mathbb{N}$. If $m, n \in \mathbb{N}$ then $\omega_m(n)$ denotes the number of distinct prime factors of n not exceeding m, while $\Omega_m(n)$ denotes the number of prime factors of n not exceeding m counted with multiplicity:

$$\omega_m(n) = \sum_{\substack{p \le m \\ p \mid n}} 1, \qquad \Omega_m(n) = \sum_{\substack{p \le m \\ p^\alpha \| n}} \alpha,$$

and we write

$$\omega_n(n) = \omega(n), \qquad \Omega_n(n) = \Omega(n).$$

The smallest and greatest prime factors of the positive integer n are denoted by $p(n)$, and $P(n)$, respectively. The *counting function* of a set $\mathcal{A} \subset \mathbb{N}$, denoted by A, is defined by

$$A(x) = |\mathcal{A} \cap [1, x]|, \qquad x \in \mathbb{N}.$$

The *upper density* $\bar{d}(\mathcal{A})$ and the *lower density* $\underline{d}(\mathcal{A})$ are defined by

$$\bar{d}(\mathcal{A}) = \limsup_{x \to \infty} \frac{A(x)}{x} \quad \text{and} \quad \underline{d}(\mathcal{A}) = \liminf_{x \to \infty} \frac{A(x)}{x},$$

respectively, and if $\bar{d}(\mathcal{A}) = \underline{d}(\mathcal{A})$, then the *density* $d(\mathcal{A})$ of $\mathcal{A}$ is defined as

$$d(\mathcal{A}) = \bar{d}(\mathcal{A}) = \underline{d}(\mathcal{A}).$$

The *upper logarithmic density* $\overline{\delta}(\mathcal{A})$ is defined by

$$\overline{\delta}(\mathcal{A}) = \limsup_{x \to \infty} \frac{1}{\log x} \sum_{\substack{a \in \mathcal{A} \\ a \le x}} \frac{1}{a},$$

1991 *Mathematics Subject Classification*: Primary 11N25.

Research of the third author partially supported by the Hungarian National Foundation for Scientific Research, Grant no. T017433. This paper was written while he was visiting the Universität Bielefeld.

and the definitions of the lower logarithmic density $\underline{\delta}(\mathcal{A})$ and logarithmic density $\delta(\mathcal{A})$ are similar.

A set $\mathcal{A} \subset \mathbb{N}$ is said to be *primitive* if there are no a, a' with $a \in \mathcal{A}$, $a' \in \mathcal{A}$, $a \neq a'$ and $a \mid a'$. There are two classical results on primitive sequences: Behrend [2] proved that if $\mathcal{A} \subset \{1, \ldots, N\}$ and $\mathcal{A}$ is primitive, then

$$\sum_{a\in\mathcal{A}} \frac{1}{a} < c_1 \frac{\log N}{\sqrt{\log\log N}} \tag{1.1}$$

(so that an infinite primitive sequence must be of 0 logarithmic density), and Erdős [4] proved that if $\mathcal{A} \subset \mathbb{N}$ is a (finite or infinite) primitive sequence then

$$\sum_{a\in\mathcal{A}} \frac{1}{a\log a} < c_2. \tag{1.2}$$

These results have been extended in various directions; surveys of this field are given in [1], [8], [9], [13].

For $\mathcal{A} \subset \mathbb{N}$ and $a \in \mathcal{A}$ let $Q^a_{\mathcal{A}}$ denote the set of integers q such that $q > 1$ and $aq \in \mathcal{A}$, and write

$$Q_{\mathcal{A}} = \bigcup_{a\in\mathcal{A}} Q^a_{\mathcal{A}}. \tag{1.3}$$

Then $Q_{\mathcal{A}}$ consists of the integers $q > 1$ that can be represented in the form $q = a'/a$ with $a, a' \in \mathcal{A}$. We call $Q_{\mathcal{A}}$ the *quotient set* of the set $\mathcal{A}$. By Behrend's and Erdős's theorems, the quotient set of a "dense" set $\mathcal{A}$ is non-empty. We will also study the set $Q^\infty_{\mathcal{A}}$ defined by

$$Q^\infty_{\mathcal{A}} = \bigcap_{n=1}^{\infty} \bigcup_{\substack{a\geq n\\ a\in\mathcal{A}}}^{\infty} Q^a_{\mathcal{A}}.$$

This set consists of the integers $q > 1$ which have infinitely many representations in the form $q = a'/a$ with $a, a' \in \mathcal{A}$. We will call $Q^\infty_{\mathcal{A}}$ the *infinite quotient set* of $\mathcal{A}$.

Pomerance and Sárközy [12] initiated the study of quotient sets of "dense" sets. They investigated the arithmetic properties of $Q_{\mathcal{A}}$ and, in particular, they proved the following theorem:

Theorem A. *There exist constants c_3 and N_0 such that if $N \in \mathbb{N}$, $N > N_0$, $\mathcal{P}$ is a set of primes not exceeding N with*

$$\sum_{p\in\mathcal{P}} \frac{1}{p} > c_3, \tag{1.4}$$

$\mathcal{A} \subset \{1, \ldots, N\}$ *and*

$$\sum_{a\in\mathcal{A}} \frac{1}{a} > 10 \log N \biggl(\sum_{p\in\mathcal{P}} \frac{1}{p}\biggr)^{-1/2}, \tag{1.5}$$

then there is a $q \in Q_{\mathcal{A}}$ *such that* $q \mid \prod_{p\in\mathcal{P}} p$.

They discussed various consequences of this theorem, and they also studied the occurrence of numbers of the form $p-1$ (p prime) in $Q_{\mathcal{A}}$.

In this paper our goal is to continue the study of the quotient set by studying the density related properties of it.

2. The problems and results. Our first goal is to study the connection between $\overline{\delta}(\mathcal{A})$ and $\overline{\delta}(Q_{\mathcal{A}})$. First we thought that for all $\mathcal{A} \subset \mathbb{N}$ we have

$$\overline{\delta}(Q_{\mathcal{A}}) \ge \overline{\delta}(\mathcal{A}). \tag{2.1}$$

However, it is not so, as the following example shows: Let $\mathcal{A}$ be the set of integers that can be represented in the form $2m$, $3m$ or $5m$ with $m \in \mathbb{N}$, $(m, 30) = 1$. Then a simple computation shows that

$$\overline{\delta}(\mathcal{A}) = \delta(\mathcal{A}) = d(\mathcal{A}) = \frac{62}{225}$$

and

$$\overline{\delta}(Q_{\mathcal{A}}) = \delta(Q_{\mathcal{A}}) = d(Q_{\mathcal{A}}) = \frac{4}{15} = \frac{30}{31}\overline{\delta}(\mathcal{A}),$$

so that (2.1) does not hold. Later we prove that there is a connection between the densities in (2.1), however, they can be far apart:

Theorem 1. (i) *If a set* $\mathcal{A} \subset \mathbb{N}$ *has positive upper logarithmic density then* $Q_{\mathcal{A}}$ *also has positive upper logarithmic density.*

(ii) *For all* $\varepsilon, \delta > 0$ *there is a set* $\mathcal{A} \subset \mathbb{N}$ *such that*

$$\underline{d}(\mathcal{A}) > 1 - \varepsilon, \tag{2.2}$$

but

$$\overline{d}(Q_{\mathcal{A}}) < \delta. \tag{2.3}$$

Next we will study the following problem: what density assumptions are needed to ensure that $Q_{\mathcal{A}}^{\infty}$ is non-empty, resp. infinite? We will prove

Theorem 2. (i) *If a set* $\mathcal{A} \subset \mathbb{N}$ *has positive upper logarithmic density then* $Q_{\mathcal{A}}^{\infty}$ *is infinite.*

(ii) *For all* $\varepsilon(x) \searrow 0$ *there is a set* $\mathcal{A} \subset \mathbb{N}$ *such that*

$$A(x) > \varepsilon(x)x \quad \textit{for } x > x_0, \tag{2.4}$$

but $Q_{\mathcal{A}}^{\infty}$ *is empty.*

By Theorem 2(i), if $\mathcal{A}$ has positive upper logarithmic density, then $Q_{\mathcal{A}}^{\infty}$ is non-empty, so that there are integers $q > 1$ which have infinitely many representations in the form

$$q = a'/a \qquad \text{with } a, a' \in \mathcal{A}. \tag{2.5}$$

This result can be sharpened by showing that under the same assumption, there is a $q > 1$ which for infinitely many x has "many" representations of the form (2.5) with a not exceeding x:

Theorem 3. *If $\mathcal{A}$ has positive upper logarithmic density, then there is a $q \in Q_{\mathcal{A}}^{\infty}$ such that*

$$\limsup_{x\to\infty} \frac{\sum_{t\in\mathcal{A}, qt\in\mathcal{A}, t\le x} 1/t}{\log x} > 0. \tag{2.6}$$

By Theorem 2(i),

$$\overline{\delta}(\mathcal{A}) > 0 \tag{2.7}$$

implies that $Q_{\mathcal{A}}^{\infty}$ is infinite. Next we will sharpen this result by estimating the counting function $Q_{\mathcal{A}}^{\infty}(x)$ under assumption (2.7):

Theorem 4. (i) *If $\mathcal{A} \subset \mathbb{N}$ is a set of positive upper logarithmic density*:

$$\overline{\delta}(\mathcal{A}) = \eta > 0, \tag{2.8}$$

then for $x > x_0$ we have

$$\sum_{\substack{q\in Q_{\mathcal{A}}^{\infty}\\ q\le x}} \frac{1}{q} > \exp\{c(\log\log x)^{1/2}\log\log\log x\} \tag{2.9}$$

with a positive constant $c = c(\eta)$.

(ii) *For all $\varepsilon, \delta > 0$ there is a set $\mathcal{A} \subset \mathbb{N}$ such that*

$$\underline{d}(\mathcal{A}) > 1 - \varepsilon \tag{2.10}$$

and

$$Q_{\mathcal{A}}^{\infty}(y) < \frac{y}{\log y}\exp\{(\log\log y)^{1/2+\delta}\} \qquad \textit{for } y > y_0. \tag{2.11}$$

Note that, clearly, (i) implies that

$$Q_{\mathcal{A}}^{\infty}(y) > \frac{y}{\log y}\exp\{c'(\log\log y)^{1/2}\log\log\log y\}$$

for infinitely many positive integers y.

Moreover, we remark that by using a result of Erdős [5], for all $\varepsilon(x) \searrow 0$ one can construct a set $\mathcal{A}$ such that (2.10) holds and $Q_{\mathcal{A}}^{\infty}(x) < x^{1-\varepsilon(x)}$ for infinitely many positive integers x.

3. Proof of Theorem 1. (i) By a theorem of Davenport and Erdős [3], $\overline{\delta}(\mathcal{A}) > 0$ implies that there is an $a \in \mathcal{A}$ with

$$\overline{\delta}(Q_{\mathcal{A}}^a) > 0. \tag{3.1}$$

By definition (1.3) we have $Q_{\mathcal{A}}^a \subset Q_{\mathcal{A}}$ and thus (3.1) implies $\overline{\delta}(Q_{\mathcal{A}}) > 0$.

(ii) For some $b \in \mathbb{N}$ and $K > 0$ write

$$\mathcal{A} = \{n : n \in \mathbb{N},\ |\Omega_b(n) - \log\log b| < K\sqrt{\log\log b}\}.$$

We will show that if b and K are large enough in terms of ε and δ, then this set $\mathcal{A}$ satisfies (2.2) and (2.3).

If K is large enough in terms of ε, and then b is large enough in terms of ε and K, then (2.2) holds by the Turán–Kubilius inequality [9] (see also [5] and [10]). Moreover, if $q \in Q_{\mathcal{A}}$, then q can be represented in the form $q = a'/a$ with $a, a' \in \mathcal{A}$, $a < a'$. It follows from the definition of $\mathcal{A}$ that

$$\begin{aligned}\Omega_b(q) &= \Omega_b(a'/a) = \Omega_b(a') - \Omega_b(a)\\ &< (\log\log b + K\sqrt{\log\log b}) - (\log\log b - K\sqrt{\log\log b})\\ &= 2K\sqrt{\log\log b}\end{aligned}$$

so that we have

$$Q_{\mathcal{A}} \subset \{q : q \in \mathbb{N},\ \Omega_b(q) < 2K\sqrt{\log\log b}\}.$$

Again by the Turán–Kubilius inequality, if K is large enough in terms of δ and then b is large enough in terms of K, then the upper density of this set is $< \delta$ so that (2.3) also holds.

4. Proof of Theorem 2. (i) We argue by contradiction: assume that

$$\overline{\delta}(\mathcal{A}) = \eta > 0, \tag{4.1}$$

but $Q_{\mathcal{A}}^{\infty}$ is finite so that there is a $K > 0$ with

$$Q_{\mathcal{A}}^{\infty} \cap [K, \infty) = \emptyset. \tag{4.2}$$

It follows trivially from (4.1) that there is an infinite set $\mathcal{K}$ of positive integers k such that, writing

$$\mathcal{A}_k = \mathcal{A} \cap (2^{2^{k-1}}, 2^{2^k}], \tag{4.3}$$

we have

$$\frac{1}{\log 2^{2^k}} \sum_{a \in \mathcal{A}_k} \frac{1}{a} > \frac{\eta}{4} \qquad \text{for all } k \in \mathcal{K}. \tag{4.4}$$

Since the sum $\sum 1/p$ is divergent, there is a positive integer L such that

$$\sum_{K < p \le L} \frac{1}{p} > \min\left\{c_3, \left(\frac{40}{\eta}\right)^2\right\} \tag{4.5}$$

(where c_3 is the constant defined in Theorem A). Then if we write $\mathcal{P} = \{p : p$ prime, $K < p \le L\}$, then (1.4) holds and, writing also $N = 2^{2^k}$, by (4.4) and (4.5) we have

$$\sum_{a\in\mathcal{A}_k} \frac{1}{a} > \frac{\eta}{4}\log N > 10\log N\Big(\sum_{p\in\mathcal{P}}\frac{1}{p}\Big)^{-1}$$

so that Theorem A can be applied with 2^{2^k} and $\mathcal{A}_k$ in place of N and $\mathcal{A}$, respectively. It follows that if $k \in \mathcal{K}$ is large enough, then there is a number $q(k)$ which can be represented in the form

$$q(k) = a'/a \qquad \text{with } a, a' \in \mathcal{A}_k, a \neq a',\ a \mid a'$$

and which also satisfies

$$q(k) \mid \prod_{p\in\mathcal{P}} p = \prod_{K<p\le L} p.$$

Since this product has only finitely many divisors, $q(k)$ divides it, and since k can assume infinitely many values ($\mathcal{K}$ being infinite), by the pigeon hole principle there is a q_0 such that

$$q_0 \mid \prod_{K<p\le L} p \tag{4.6}$$

and $q_0 = q(k)$ for infinitely many values of k; denote the set of those k's by $\mathcal{K}_0$. Then q_0 can be represented in the form

$$q_0 = a'/a \qquad \text{with } a, a' \in \mathcal{A}_k,\ a \neq a' \text{ (for all } k \in \mathcal{K}_0). \tag{4.7}$$

Since $\mathcal{K}_0$ is infinite and the sets $\mathcal{A}_k$ are disjoint, (4.7) implies $q_0 \in Q^{\infty}_{\mathcal{A}}$, and by (4.6) and (4.7) we have $q_0 > K$, which contradicts (4.2) and completes the proof of (i).

(ii) It is well known that if $x > x_0$, then uniformly for $2 \le K \le \sqrt{x}$ we have

$$|\{n : n \le x,\ p(n) > K\}| > c_4 x \prod_{p\le K}\Big(1 - \frac{1}{p}\Big),$$

and by Mertens's formula, this is $> c_5 x/\log K$, which is $> \varepsilon(x)x$ if $K < e^{c_5/\varepsilon(x)}$. It follows that if we define $\mathcal{A} = \{n : p(n) > K(n)\}$ with $K(n) = \min\{\sqrt{n}, e^{c_6/\varepsilon(n)}\}$, where c_6 is a small positive constant, then $\mathcal{A}$ satisfies (2.4).

Moreover, for this $\mathcal{A}$ we clearly have

$$p(a) \to \infty \qquad \text{as } a \in \mathcal{A},\ a \to \infty. \tag{4.8}$$

If $q > 1$ and $q \in \mathbb{N}$, then if we represent q in the form $q = a'/a$ with $a, a' \in \mathcal{A}$, then a' must have a prime factor $\le q$, and thus by (4.8), a' must be bounded. This implies $q \notin Q_{\mathcal{A}}^{\infty}$ so that $Q_{\mathcal{A}}^{\infty}$ is empty, which completes the proof of the theorem.

5. Proof of Theorem 3. Write $\overline{\delta}(\mathcal{A}) = \eta$ (> 0). For $k \in \mathbb{N}$, let

$$\mathcal{A}_k = \{a : a \in \mathcal{A},\ 2^{2^{k-1}} < a \le 2^{2^k}\}.$$

Let $\mathcal{K}$ denote the set of positive integers k such that

$$\sum_{a \in \mathcal{A}_k} \frac{1}{a} > \frac{\eta}{4} \log 2^{2^k}. \tag{5.1}$$

Clearly, $\mathcal{K}$ is infinite. Let L denote the smallest positive integer such that

$$\sum_{p \le L} \frac{1}{p} > \min\left\{c_3, \left(\frac{80}{\eta}\right)^2\right\}, \tag{5.2}$$

and write $\prod_{p \le L} p = V$. For $q, k \in \mathbb{N}$ write

$$\mathcal{B}_{(q,k)} = \{a : 2^{2^{k-1}} < a \le 2^{2^k},\ a \in \mathcal{A},\ aq \in \mathcal{A}\}.$$

We will show that for $k \in \mathcal{K}$, $k > k_0$ there is a q such that $q \mid V$ and

$$\sum_{a \in \mathcal{B}_{(q,k)}} \frac{1}{a} > \frac{\eta}{8V} \log 2^{2^k}. \tag{5.3}$$

We argue by contradiction: assume that for all $q \mid V$ we have

$$\sum_{a \in \mathcal{B}_{(q,k)}} \frac{1}{a} \le \frac{\eta}{8V} \log 2^{2^k}. \tag{5.4}$$

Write

$$\mathcal{A}_k^c = \mathcal{A}_k \setminus \bigcup_{q \mid V} \mathcal{B}_{(q,k)}. \tag{5.5}$$

Then since $k \in \mathcal{K}$, (5.1), (5.4) and (5.5) yield

$$\begin{aligned}\sum_{a \in \mathcal{A}_k^c} \frac{1}{a} &\ge \sum_{a \in \mathcal{A}_k} \frac{1}{a} - \sum_{q \mid V} \sum_{a \in \mathcal{B}_{(q,k)}} \frac{1}{a} \\ &> \left(\frac{\eta}{4} - \sum_{q \mid V} \frac{\eta}{8V}\right) \log 2^{2^k} \ge \left(\frac{\eta}{4} - \frac{\eta}{8}\right) \log 2^{2^k} = \frac{\eta}{8} \log 2^{2^k}.\end{aligned}$$

By (5.2), it follows that

$$\sum_{a \in \mathcal{A}_k^c} \frac{1}{a} > 10 \frac{\log 2^{2^k}}{\sqrt{\sum_{p \le L} 1/p}}. \tag{5.6}$$

By (5.2) and (5.6), we may apply Theorem A with 2^{2^k}, $\mathcal{A}_k^c$ and $\{p : p$ prime, $p \le L\}$ in place of N, $\mathcal{A}$ and $\mathcal{P}$, respectively. It follows that if $k \in \mathcal{K}$

and k is large enough, then there is a q' which can be represented in the form

$$q' = a'/a \quad \text{with } a, a' \in \mathcal{A}_k^c,\ a \neq a',\ a \mid a' \tag{5.7}$$

and which also satisfies

$$q' \mid \prod_{p \le L} p = V. \tag{5.8}$$

For these a and q' we have $a \in \mathcal{A}_k$ and $aq' \in \mathcal{A}_k$, and thus

$$a \in \mathcal{B}_{(q',k)}. \tag{5.9}$$

It follows from (5.5), (5.8) and (5.9) that $a \notin \mathcal{A}_k^c$. This contradicts (5.7), which proves that, indeed, for all $k \in \mathcal{K}$, $k < k_0$ there is a q such that $q \mid V$ and (5.3) holds. To each $k \in \mathcal{K}$, $k > k_0$ assign a $q = q(k)$ with these properties. Since $\mathcal{K}$ is infinite and, as $q(k) \mid V$, $q(k)$ may assume only finitely many distinct values, there is a q_0 (with $q_0 \mid V$) which has infinitely many representations in the form $q_0 = q(k)$. For this q_0 we have

$$\frac{1}{\log 2^{2^k}} \sum_{\substack{a \in \mathcal{A},\, aq_0 \in \mathcal{A} \\ a \le 2^{2^k}}} \frac{1}{a} > \frac{\eta}{8V}$$

for infinitely many $k \in \mathbb{N}$, which proves (2.6) and completes the proof of Theorem 3.

6. Proof of Theorem 4(i). Combinatorial lemmas

LEMMA 1. *For all $\mu > 0$ there are numbers r_0 and $c = c(\mu) > 0$ such that if $r \in \mathbb{N}$, $r > r_0$, $\mathcal{U}$ is a finite set with $|\mathcal{U}| = r$, and $\mathcal{U}_1, \ldots, \mathcal{U}_k$ are subsets of $\mathcal{U}$ with*

$$k > \mu 2^r, \tag{6.1}$$

then there is a j $(1 \le j \le k)$ such that

$$|\{i : 1 \le i \le k,\ \mathcal{U}_i \subset \mathcal{U}_j\}| > \exp\{c\sqrt{r}\log r\}. \tag{6.2}$$

Proof. This is Theorem 2 of [7].

LEMMA 2. *For all $\mu > 0$ there are numbers r_0 and $c = c(\mu) > 0$ such that if $r \in \mathbb{N}$, $r > r_0$, $\mathcal{T}$ is a finite set with $|\mathcal{T}| = t$,*

$$\mathcal{T} = \mathcal{U} \cup \mathcal{V}, \qquad \mathcal{U} \cap \mathcal{V} = \emptyset, \qquad |\mathcal{U}| = r,$$

and $\mathcal{T}_1, \ldots, \mathcal{T}_l$ are subsets of $\mathcal{T}$ with

$$l > \mu 2^t, \tag{6.3}$$

then there is an h $(1 \le h \le l)$ such that

$$|\{i : 1 \le i \le l,\ \mathcal{T}_i \cap \mathcal{U} \subset \mathcal{T}_h \cap \mathcal{U},\ \mathcal{T}_i \cap \mathcal{V} = \mathcal{T}_h \cap \mathcal{V}\}| > \exp\{c\sqrt{r}\log r\}. \tag{6.4}$$

Proof. By the pigeon hole principle, it follows from (6.3) that $\mathcal{V}$ has a subset $\mathcal{V}_0$ such that

$$|\{h : 1 \le h \le l,\ \mathcal{T}_h \cap \mathcal{V} = \mathcal{V}_0\}| \ge \frac{l}{2^{|\mathcal{V}|}} > \frac{\mu 2^t}{2^{|\mathcal{V}|}} = \mu 2^{|\mathcal{U}|} = \mu 2^r. \tag{6.5}$$

Let $\mathcal{T}_{h_1}, \ldots, \mathcal{T}_{h_k}$ $(h_1 < \ldots < h_k)$ be the subsets of $\mathcal{T}$ with $\mathcal{T}_{h_i} \cap \mathcal{V} = \mathcal{V}_0$, $i = 1, \ldots, k$, so that (6.1) holds by (6.5). Write $\mathcal{U}_i = \mathcal{T}_{h_i} \cap \mathcal{U}$ for $1 \le i \le k$. By Lemma 1, there is a j $(1 \le j \le k)$ such that (6.2) holds. Then clearly $\mathcal{T}_{h_j}$ satisfies (6.4) with h_j in place of h, which completes the proof of Lemma 2.

7. Proof of Theorem 4(i). Arithmetic lemmas

LEMMA 3. *For all $\gamma > 0$ there are constants $c = c(\gamma) > 0$, N_0 and R_0 such that if $N > N_0$, $\mathcal{A} \subset \{1, \ldots, N\}$,*

$$\sum_{a \in \mathcal{A}} \frac{1}{a} > \gamma \log N \tag{7.1}$$

and $R_0 \le R \le N$, then, writing

$$f(\mathcal{A}, R, n) = |\{a : a \in \mathcal{A},\ a \mid n,\ P(n/a) \le R\}| \tag{7.2}$$

and

$$\mathcal{A}^*(R, c) = |\{a : a \in \mathcal{A},\ f(\mathcal{A}, R, a) > \exp(c(\log\log R)^{1/2} \log\log\log R)\}|,$$

we have

$$\sum_{a \in \mathcal{A}^*(R,c)} \frac{1}{a} > \frac{1}{2} \sum_{a \in \mathcal{A}} \frac{1}{a}. \tag{7.3}$$

Proof. We argue by contradiction: assume that contrary to (7.3) we have

$$\sum_{a \in \mathcal{A}^*(R,c)} \frac{1}{a} \le \frac{1}{2} \sum_{a \in \mathcal{A}} \frac{1}{a}. \tag{7.4}$$

We will show that if $c = c(\gamma)$ (> 0) is small enough (in terms of γ) then (7.4) leads to a contradiction.

Write $\mathcal{A}^c = \mathcal{A} \setminus \mathcal{A}^*(R, c)$ so that

$$\mathcal{A}^c = \{a : a \in \mathcal{A},\ f(\mathcal{A}, R, a) \le \exp(c(\log\log R)^{1/2} \log\log\log R)\} \tag{7.5}$$

and, by (7.1) and (7.4),

$$\sum_{a \in \mathcal{A}^c} \frac{1}{a} \ge \frac{1}{2} \sum_{a \in \mathcal{A}} \frac{1}{a} > \frac{\gamma}{2} \log N. \tag{7.6}$$

Write every $a \in \mathcal{A}^c$ as the product of a square $(r(a))^2$ and a squarefree integer $s(a)$:

$$a = (r(a))^2 s(a), \qquad |\mu(s(a))| = 1$$

(where $\mu(n)$ denotes the Möbius function). Then (7.6) can be rewritten as

$$\frac{\gamma}{2}\log N < \sum_{a\in\mathcal{A}} \frac{1}{(r(a))^2 s(a)} = \sum_{r=1}^{\infty}\frac{1}{r^2}\sum_{\substack{a\in\mathcal{A}^c\\ r(a)=r}}\frac{1}{s(a)}.$$

Since $\sum_{r=1}^{\infty} 1/r^2 = \pi^2/6 < 2$, it follows that there is an integer r_0 such that

$$\sum_{\substack{a\in\mathcal{A}^c\\ r(a)=r_0}}\frac{1}{s(a)} > \frac{\gamma}{4}\log N. \tag{7.7}$$

Write $S = \{s : \text{there is an } a\in\mathcal{A}^c \text{ with } r(a)=r_0,\ s(a)=s\}$. Then, by (7.7),

$$\sum_{s\in S}\frac{1}{s} > \frac{\gamma}{4}\log N, \tag{7.8}$$

and clearly

$$S\subset\{1,\dots,N\}, \tag{7.9}$$

$$\text{every } s\in S \text{ is squarefree.} \tag{7.10}$$

Set $d_S(n) = |\{s : s\in S,\ s\,|\,n\}|$ and let $d(n) = |\{d : d\in\mathbb{N},\ d\,|\,n\}|$ denote the divisor function. Then it is well known that for large N we have

$$\sum_{n=1}^{N} d(n) < 2N\log N. \tag{7.11}$$

Write

$$\mathcal{H}(N,R) = \left\{n : n\le N,\ \omega_R(n) > \tfrac{1}{2}\log\log R\right\}.$$

Now we will show that there is an integer n with

$$n\in\mathcal{H}(N,R), \qquad d_S(n) > \frac{\gamma}{32}d(n). \tag{7.12}$$

Clearly we have

$$\begin{aligned}\sum_{n\in\mathcal{H}(N,R)} d_S(n) &= \sum_{n\in\mathcal{H}(N,R)}\sum_{\substack{s\in S\\ s|n}} 1 = \sum_{s\in S}\sum_{\substack{n\le N,\, s|n\\ \omega_R(n)>\frac{1}{2}\log\log R}} 1\\ &= \sum_{s\in S}\sum_{\substack{st\le n\\ \omega_R(st)>\frac{1}{2}\log\log R}} 1 \ge \sum_{\substack{s\in S\\ S<N^{1-\gamma/10}}}\sum_{\substack{t\le N/S\\ \omega_R(t)>\frac{1}{2}\log\log R}} 1.\end{aligned}$$

By the Turán–Kubilius inequality [11], for $R_0\le R\le N$ the inner sum is $> \frac{1}{2}\frac{N}{S}$ so that, by (7.8), for large N we have

$$\begin{aligned}\sum_{n\in\mathcal{H}(N,R)} d_S(n) &\ge \frac{N}{2}\sum_{\substack{s\in S\\ s<N^{1-\gamma/10}}}\frac{1}{s} \ge \frac{N}{2}\Bigg(\sum_{s\in S}\frac{1}{s} - \sum_{N^{1-\gamma/10}\le S\le N}\frac{1}{s}\Bigg)\\ &> \frac{N}{2}\left(\frac{\gamma}{4}\log N - \frac{\gamma}{8}\log N\right) = \frac{\gamma}{16}N\log N.\end{aligned} \tag{7.13}$$

Now assume that contrary to our statement there is no n satisfying (7.12). Then it follows from (7.11) that

$$\sum_{n\in\mathcal{H}(N,R)} d_S(n) \le \sum_{n\in\mathcal{H}(N,R)} \frac{\gamma}{32} d(n) \le \frac{\gamma}{32}\sum_{n=1}^{N} d(n) < \frac{\gamma}{16} N\log N,$$

which contradicts (7.13), and this completes the proof of the existence of an n satisfying (7.12). Consider such an n, and write

$$n_1 = \prod_{p\mid n} p.$$

Then as $n \in \mathcal{H}(N,R)$ we clearly have

$$\omega_R(n_1) = \omega_R(n) > \tfrac{1}{2}\log\log R, \tag{7.14}$$

and, by (7.10), it follows from (7.12) that

$$d_S(n_1) = d_S(n) > \frac{\gamma}{32} d(n) \ge \frac{\gamma}{32} d(n_1). \tag{7.15}$$

Let $s_{i_1} < \ldots < s_{i_l}$ (with $l = d_S(n_1)$) be the elements of S dividing n_1. Write

$$\mathcal{T} = \{p : p \text{ prime},\ p \mid n_1\}, \qquad t = |\mathcal{T}| = \omega(n_1),$$
$$\mathcal{U} = \{p : p \text{ prime},\ p \le R,\ p \mid n_1\},$$
$$r = |\mathcal{U}| = \omega_R(n_1) \quad\text{and}\quad \mathcal{T}_j = \{p : p \text{ prime},\ p \mid s_{i_j}\} \qquad \text{for } j = 1,\ldots,l.$$

Then $\mathcal{T}_1,\ldots,\mathcal{T}_l$ are subsets of $\mathcal{T}$ and, by (7.15), their number is

$$l = d_S(n_1) > \frac{\gamma}{32} d(n_1) = \frac{\gamma}{32} 2^t. \tag{7.16}$$

Moreover, by (7.14) we have

$$|\mathcal{U}| = r = \omega_R(n_1) = \omega_R(n) > \tfrac{1}{2}\log\log R. \tag{7.17}$$

If R_0 is large enough in terms of γ then, since $R \ge R_0$, by (7.16) and (7.17) all the conditions in Lemma 2 hold with $\gamma/32$ in place of μ. Thus by Lemma 2 and (7.17), there is an h $(1 \le h \le l)$ such that

$$\begin{aligned} &|\{j : 1 \le j \le l,\ \mathcal{T}_j \cap \mathcal{U} \subset \mathcal{T}_h \cap \mathcal{U},\ \mathcal{T}_j \cap \mathcal{V} = \mathcal{T}_h \cap \mathcal{V}\}| \\ &\qquad > \exp\{c\sqrt{r}\log r\} > \exp\{c'(\log\log R)^{1/2}\log\log\log R\} \end{aligned} \tag{7.18}$$

with positive constants $c = c(\gamma)$, $c' = c'(\gamma)$. If $\mathcal{T}_j \cap \mathcal{U} \subset \mathcal{T}_h \cap \mathcal{U}$ and $T_j \cap \mathcal{V} = \mathcal{T}_h \cap \mathcal{V}$ then

$$r_0^2 s_{i_j} \mid r_0^2 s_{i_h} \quad\text{and}\quad P\left(\frac{r_0^2 s_{i_h}}{r_0^2 s_{i_j}}\right) \le R. \tag{7.19}$$

Here $r_0^2 s_{i_j} \in \mathcal{A}^c \subset \mathcal{A}$ (for all j) and $\overline{a} = r_0^2 s_{i_h} \in \mathcal{A}^c$, so that by (7.18) and (7.19) we have

$$\begin{aligned} f(\mathcal{A}, R, \overline{a}) &= |\{a : a \in \mathcal{A},\ a \mid \overline{a},\ P(\overline{a}/a) \le R\}| \\ &> \exp\{c'(\log\log R)^{1/2}\log\log\log R\}. \end{aligned}$$

This contradicts the definition (7.5) of $\mathcal{A}^c$ if we choose $c = c'$ there, and this completes the proof of Lemma 3.

LEMMA 4. *For all* $\gamma > 0$, *if* $N > N_0$, $\mathcal{A} \subset \{1, \dots, N\}$, $\sum_{a \in \mathcal{A}} 1/a > \gamma \log N$ *and* $R_1 \le R \le N$, *then, writing*

$$Q'(R) = \{q : P(q) \le R \text{ and there is an } a \in \mathcal{A} \text{ with } aq \in \mathcal{A}\},$$

we have

$$\sum_{q \in Q'(R)} \frac{1}{q} > \exp(c'(\log\log R)^{1/2} \log\log\log R) \tag{7.20}$$

where $c' = c/2$ *with the constant* $c = c(\gamma) > 0$ *defined in Lemma* 3.

Proof. Write

$$S = \sum_{a \in \mathcal{A}} \frac{f(\mathcal{A}, R, a)}{a}$$

where $f(\mathcal{A}, R, a)$ is defined by (7.2). Assume that contrary to (7.20) we have

$$\sum_{q \in Q'(R)} \frac{1}{q} \le \exp(c'(\log\log R)^{1/2} \log\log\log R).$$

Then

$$\begin{aligned} S &= \sum_{a \in \mathcal{A}} \frac{f(\mathcal{A}, R, a)}{a} = \sum_{a \in \mathcal{A}} \frac{1}{a} \sum_{\substack{a' \in \mathcal{A},\, a'q = a \\ P(q) \le R}} 1 = \sum_{a' \in \mathcal{A}} \frac{1}{a'} \sum_{\substack{a'q \in \mathcal{A} \\ P(q) \le R}} \frac{1}{q} \\ &\le \sum_{a' \in \mathcal{A}} \frac{1}{a'} \sum_{q \in Q'(R)} \frac{1}{q} \\ &\le \exp(c'(\log\log R)^{1/2} \log\log\log R) \sum_{a' \in \mathcal{A}} \frac{1}{a'}. \end{aligned} \tag{7.21}$$

On the other hand, by Lemma 3 we have

$$\begin{aligned} S &= \sum_{a \in \mathcal{A}} \frac{f(\mathcal{A}, R, a)}{a} > \sum_{a \in \mathcal{A}^*(R,c)} \frac{\exp(c(\log\log R)^{1/2} \log\log\log R)}{a} \\ &= \exp(c(\log\log R)^{1/2} \log\log\log R) \sum_{a \in \mathcal{A}^*(R,c)} \frac{1}{a} \\ &> \frac{1}{2} \exp(c(\log\log R)^{1/2} \log\log\log R) \sum_{a \in \mathcal{A}} \frac{1}{a}. \end{aligned}$$

If $c' = c/2$ and R is large enough then this lower bound contradicts the upper bound in (7.21), which completes the proof of Lemma 4.

LEMMA 5. *For all* $\gamma > 0$ *there are constants* N_0, U_0 *such that if* $N > N_0$, $\mathcal{A} \subset \{1, \dots, N\}$,

$$\sum_{a \in \mathcal{A}} \frac{1}{a} > \gamma \log N \tag{7.22}$$

and $U_0 \le U \le \exp((\log N)^2)$, *then, writing*

$$Q^*(U) = \{q : q \le U \text{ and there is an } a \in \mathcal{A} \text{ with } aq \in \mathcal{A}\},$$

we have

$$\sum_{q \in Q^*(U)} \frac{1}{q} > \exp(c''(\log\log U)^{1/2} \log\log\log U) \tag{7.23}$$

where $c'' = c'/2$ *with the constant* $c' = c'(\gamma)$ *defined in Lemma* 4.

P r o o f. Define R by

$$U = \exp((\log R)^2)$$

so that $\frac{1}{2}\log\log U = \log\log R$. If U is large enough then, by Lemma 4, (7.22) implies that

$$\begin{aligned} \sum_{q \in Q'(R)} \frac{1}{q} &> \exp(c'(\log\log R)^{1/2} \log\log\log R) \\ &= \exp\biggl((1 + O(1))\frac{c'}{\sqrt{2}}(\log\log U)^{1/2} \log\log\log U\biggr). \end{aligned} \tag{7.24}$$

Moreover, we clearly have

$$Q'(R) \setminus Q^*(U) \subset \{q : U < q,\, P(q) \le R\},$$

so that

$$\sum_{q \in Q^*(U)} \frac{1}{q} \ge \sum_{q \in Q'(R)} \frac{1}{q} - \sum_{\substack{q \in Q'(R) \\ q \notin Q^*(U)}} \frac{1}{q} \ge \sum_{q \in Q'(R)} \frac{1}{q} - \sum_{\substack{U<q \\ P(q) \le R}} \frac{1}{q}. \tag{7.25}$$

It remains to estimate the last sum.

Write $\sigma = 1/\log R$ so that $U^\sigma = R$. Then, since

$$\sum_{p \le x} \frac{1}{p} = \log\log x + O(1),$$

we have

$$\begin{aligned} \sum_{\substack{U<q \\ P(q) \le R}} \frac{1}{q} &< \sum_{\substack{U<q \\ P(q) \le R}} \frac{1}{q}\left(\frac{q}{U}\right)^{\sigma} < U^{-\sigma} \sum_{P(q) \le R} q^{-1+\sigma} \\ &= \frac{1}{R}\prod_{p \le R}(1 - p^{-1+\sigma})^{-1} = \frac{1}{R}\exp\Bigl\{-\sum_{p \le R}\log(1 - p^{-1+\sigma})\Bigr\} \\ &= \frac{1}{R}\exp\Bigl\{O\Bigl(\sum_{p \le R} p^{-1+\sigma}\Bigr)\Bigr\} \le \frac{1}{R}\exp\Bigl\{O\Bigl(R^\sigma \sum_{p \le R} p^{-1}\Bigr)\Bigr\} \\ &= \frac{1}{R}\exp\{O(\log\log R)\} = \frac{(\log R)^{O(1)}}{R} = o(1) \qquad (\text{as } R \to \infty). \end{aligned} \tag{7.26}$$

For large U, (7.23) follows from (7.24), (7.25) and (7.26), and this completes the proof of Lemma 5.

8. Completion of the proof of Theorem 4(i). By (2.8), there is an infinite set $N_1 < N_2 < \ldots$ of positive integers such that $N_{k+1} > N_k^2$ for $k = 1, 2, \ldots$, and, writing

$$\mathcal{A} \cap (N_{k-1}, N_k] = \mathcal{A}_k \qquad \text{for } k = 2, 3, \ldots,$$

we have

$$\sum_{a \in \mathcal{A}_k} \frac{1}{a} > \frac{\eta}{4} \log N_k.$$

Then for large k, by using Lemma 5 with $\eta/4$, N_k, $\mathcal{A}_k$ and x in place of $\gamma, N, \mathcal{A}$ and U, respectively, we find that, writing

$$Q_k^*(x) = \{q : q \le x \text{ and there is an } a \in \mathcal{A}_k \text{ with } aq \in \mathcal{A}_k\},$$

for $x > x_0$ and large enough k we have

$$\sum_{q \in Q_k^*(x)} \frac{1}{q} > \exp\{c''(\log\log x)^{1/2} \log\log\log x\}. \tag{8.1}$$

Since for every large k there is such a set $Q_k^*(x)$ and we have $Q_k^*(x) \subset \{1, \ldots, [x]\}$, by the pigeon hole principle there is a set

$$Q_0(x) \subset \{1, \ldots, [x]\} \tag{8.2}$$

which can be represented in the form

$$Q_0(x) = Q_k^*(x) \tag{8.3}$$

for an infinite set $\mathcal{K}$ of positive integers k. If $q \in Q_0(x)$ and $k \in \mathcal{K}$, then q can be represented in the form $q = a'/a$, $a \in \mathcal{A}_k$, $a' = aq \in \mathcal{A}_k$. Since $\mathcal{A}_k \subset \mathcal{A}$, the sets $\mathcal{A}_k$ are disjoint, and $\mathcal{K}$ is infinite, (8.2) implies

$$Q_0(x) \subset Q_{\mathcal{A}}^{\infty} \cap [1, x]. \tag{8.4}$$

(2.9) follows from (8.1), (8.3) and (8.4), and this completes the proof of Theorem 4(i).

9. Proof of Theorem 4(ii). Let K be a large but fixed number, and let $\mathcal{A}$ denote the set of integers a such that

$$|\Omega_b(a) - \log\log b| < (\log\log b)^{1/2+\delta/2}$$

for all $K < b \le a$. We will show that if K is large enough then $\mathcal{A}$ satisfies (2.10) and (2.11).

Indeed, it follows from Erdős's result [6, p. 4] that if K is large enough in terms of δ and ε then (2.10) holds.

Moreover, if $q \in Q_{\mathcal{A}}^{\infty}$ and $q > K$, then q can be represented infinitely often as $q = a'/a$ with $a, a' \in \mathcal{A}$, $a \mid a'$, $q < a < a'$. Then by the construction of $\mathcal{A}$,

$$\begin{aligned}\Omega(q) = \Omega_q(q) = \Omega_q\left(\frac{a'}{a}\right) &= \Omega_q(a') - \Omega_q(a) \\ &< (\log\log q + (\log\log q)^{1/2+\delta/2}) - (\log\log q - (\log\log q)^{1/2+\delta/2}) \\ &= 2(\log\log q)^{1/2+\delta/2}.\end{aligned}$$

Thus by a theorem of Sathe [14] and Selberg [15] we have

$$\begin{aligned}Q_{\mathcal{A}}^{\infty}(y) &\le K + |\{q : K < q \le y,\ q \in Q_{\mathcal{A}}^{\infty}\}| \\ &\le K + \sum_{i \le 2(\log\log y)^{1/2+\delta/2}} |\{q : q \le y,\ \Omega(q) = i\}| \\ &= O\Bigg(1 + \sum_{i \le 2(\log\log y)^{1/2+\delta/2}} \frac{y}{\log y} \cdot \frac{(\log\log y)^{i-1}}{(i-1)!}\Bigg) \\ &= O\bigg(\frac{y}{\log y}(\log\log y)^{2(\log\log y)^{1/2+\delta/2}}\bigg) \\ &= o\bigg(\frac{y}{\log y}\exp((\log\log y)^{1/2+\delta})\bigg),\end{aligned}$$

which proves (2.11).

References

[1] R. Ahlswede and L. H. Khachatrian, *Classical results on primitive and recent results on cross-primitive sequences*, in: The Mathematics of Paul Erdős, Vol. I, R. L. Graham and J. Nešetřil (eds.), Algorithms Combin. 13, Springer, 1997, 104–116.

[2] F. Behrend, *On sequences of numbers not divisible by one another*, J. London Math. Soc. 10 (1935), 42–44.

[3] H. Davenport and P. Erdős, *On sequences of positive integers*, Acta Arith. 2 (1936), 147–151.

[4] P. Erdős, *Note on sequences of integers no one of which is divisible by any other*, J. London Math. Soc. 10 (1935), 126–128.

[5] —, *A generalization of a theorem of Besicovitch*, ibid. 11 (1936), 92–98.

[6] —, *On the distribution of additive functions*, Ann. of Math. 97 (1946), 1–20.

[7] P. Erdős, A. Sárközy and E. Szemerédi, *On the divisibility properties of sequences of integers I*, Acta Arith. 11 (1966), 411–418.

[8] —, —, —, *On divisibility properties of sequences of integers*, in: Colloq. Math. Soc. János Bolyai 2, 1970, 35–49.

[9] H. Halberstam and K. F. Roth, *Sequences*, Springer, Berlin, 1983.

[10] G. H. Hardy and S. Ramanujan, *The normal number of prime factors of a number n*, Quart. J. Math. 48 (1920), 76–92.

[11] J. Kubilius, *Probabilistic Methods in the Theory of Numbers*, Amer. Math. Soc. Transl. Math. Monographs 11, Providence, 1964.

[12] C. Pomerance and A. Sárközy, *On homogeneous multiplicative hybrid problems in number theory*, Acta Arith. 49 (1988), 291–302.

[13] A. Sárközy, *On divisibility properties of sequences of integers*, in: The Mathematics of Paul Erdős, Vol. I, R. L. Graham and J. Nešetřil (eds.), Algorithms Combin. 13, Springer, 1997, 241–250.

[14] L. G. Sathe, *On a problem of Hardy on the distribution of integers having given number of prime factors*, J. Indian Math. Soc. 17 (1953), 63–141; and 18 (1954), 27–81.

[15] A. Selberg, *Note on a paper by L. G. Sathe*, ibid. 18 (1954), 83–87.

Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
D-33501 Bielefeld, Germany
E-mail: hollmann@mathematik.uni-bielefeld.de

Department of Algebra and Number Theory
Eötvös University
Múzeum krt. 6–8
H-1088 Budapest, Hungary
E-mail: sarkozy@cs.elte.hu

Received on 18.9.1998
and in revised form on 24.5.1999 (3466)