MULTIPLICATIVE PROPERTIES OF SETS OF POSITIVE INTEGERS

L. HAJDU, A. SCHINZEL, M. SKAŁBA

Introduction. Let $\mathbb{N}$ be the set of positive integers. It is easy to see that a set $\mathcal{A} \subseteq \mathbb{N}$ of upper asymptotic density

$$\overline{\delta^*}(\mathcal{A}) = \limsup_{x\to\infty} \frac{1}{x} \sum_{n\in\mathcal{A}, n\le x} 1 > \frac{1}{2}$$

contains two numbers and their sum. An analogous statement for the product fails, even if $\overline{\delta^*}(\mathcal{A})$ is arbitrarily close to 1, as the following example shows

$$\mathcal{A}_l = \bigcup_{k=0}^{\infty} \{n \in \mathbb{N} : l^{3k+1} \le n < l^{3k+2}\}, \quad \overline{\delta^*}(\mathcal{A}_l) = 1 - \frac{1}{l}.$$

However we shall prove

Theorem 1. *If a set $\mathcal{A} \subseteq \mathbb{N}$ has upper Dirichlet density*

$$\overline{D}(\mathcal{A}) = \limsup_{s\to 1+} (s-1) \sum_{n\in\mathcal{A}} \frac{1}{n^s} > \frac{1}{2},$$

then for every x there exist three distinct numbers h_1, h_2, h_3 in $\mathcal{A}$ all greater than x such that

$$h_1 h_2 h_3 = \square. \tag{1}$$

Corollary 1. *If a set $\mathcal{A} \subseteq \mathbb{N}$ has lower asymptotic density*

$$\delta^*(\mathcal{A}) = \liminf_{x\to\infty} \frac{1}{x} \sum_{n\in\mathcal{A}, n\le x} 1 > \frac{1}{2}$$

or lower logarithmic density

$$\underline{D_l}(\mathcal{A}) = \liminf_{x\to\infty} \frac{1}{\log x} \sum_{n\in\mathcal{A}, n\le x} \frac{1}{n} > \frac{1}{2},$$

then the assertion of Theorem 1 holds.

Remark 1. The example of the set of integers with odd total number of prime factors, which has density 1/2 (see [1],§167), but no three elements h_i satisfying (1) shows that the constant 1/2 in Theorem 1 ist best possible.

In the sequel we shall use the following notation: $\omega(m)$, $\Omega(m)$ and $\tau(m)$ are the number of distinct prime factors, the total number of prime factors and the number of divisors of m, respectively.

For $S \subseteq \mathbb{N}$ we put

$$S(x) = \sum_{n\in S, n\le x} 1 \text{ and } \tau(n,S) = \sum_{d|n, d\in S} 1 \text{ for } x > 0, n \in \mathbb{N}$$

Theorem 1 is a consequence of the following two theorems

Theorem 2. *Let m be a positive integer, and write*

$$\mathcal{D} := \{d : d \textit{ divides } m\}.$$

Let $\mathcal{H}$ be an arbitrary subset of $\mathcal{D}$ with $|\mathcal{H}| > |\mathcal{D}|/2$. Then there exist $h_1, h_2, h_3 \in \mathcal{H}$ such that $h_1h_2h_3 = \square$. Further, if $m > 1$ is not squarefree, and is neither of the form $p_1^2p_2$ or $p_1^{n_1}$ $(2 \le n_1 \le 4)$, then the above $h_1, h_2, h_3 \in \mathcal{H}$ can be chosen to be distinct.

Remark 2. For every $\epsilon > 0$ there exist $m \in \mathbb{N}$ and a set $\mathcal{H}$ of divisors of m such that $|\mathcal{H}| > (1-\epsilon)\tau(m)$ and $\mathcal{H}$ does not contain two numbers together with their product.

Theorem 3. *Assume that $\mathcal{A}, \mathcal{B} \subseteq \mathbb{N}$ satisfy the condition:*

$$\overline{D}(\mathcal{A},\mathcal{B}) = \limsup_{s\to 1+} \frac{\sum_{n\in\mathcal{A}} \frac{1}{n^s}}{\sum_{n\in\mathcal{B}} \frac{1}{n^s}} > \alpha. \tag{2}$$

Then there exists $m \in \mathbb{N}$ such that

$$\tau(m,\mathcal{A}) > \alpha\tau(m,\mathcal{B}). \tag{3}$$

Corollary 2. *Assume that $\mathcal{A}, \mathcal{B} \subseteq \mathbb{N}$ satisfy the condition: either*

$$\liminf_{n\to\infty} \frac{\mathcal{A}(n)}{\mathcal{B}(n)} > \alpha > 0 \textit{ and } \sum_{n\in\mathcal{B}} \frac{1}{n} = \infty, \tag{4}$$

or

$$\overline{D}(\mathcal{A}) > \alpha\overline{D}(\mathcal{B}), \textit{ or } \underline{D}(\mathcal{A}) > \alpha\underline{D}(\mathcal{B}). \tag{5}$$

Then there exists $m \in \mathbb{N}$ satisfying (3).

Theorems 2 and 3 imply also the following

Theorem 4. *If a set $\mathcal{A}$ consists entirely of squarefree numbers and*

$$\overline{D}(\mathcal{A}) > \frac{3}{\pi^2}$$

then for every $x > 0$ there exist three distinct numbers h_1, h_2, h_3 in $\mathcal{A}$ all greater than x satisfying (1). Also there exist two numbers a, b in $\mathcal{A}$ greater than x such that a/b is a prime.

2. Proof of Theorem 2 We start with proving the existence of some (not necessarily distinct) $h_1, h_2, h_3 \in \mathcal{H}$ with $h_1h_2h_3 = \square$.

For $m = 1$ the statement is trivial. So assume that $m > 1$ and write $m = p_1^{n_1} \dots p_k^{n_k}$, where $p_1, \dots, p_k$ are distinct primes and the exponents $n_1, \dots, n_k$ are positive integers. Introduce the following equivalence relation on $\mathcal{D}$: for $d_1, d_2 \in \mathcal{D}$ put

$$d_1 \sim d_2 \text{ if and only if } d_1 d_2 = \square.$$

We label the equivalence classes of $(\mathcal{D}, \sim)$ by binary k-tuples $\underline{a}$, i.e. by tuples of the form

$$\underline{a} = (a_1, \dots, a_k) \text{ with } a_i \in \{0, 1\} \text{ for } i = 1, \dots, k.$$

A divisor d of m of the form $d = p_1^{t_1} \dots p_k^{t_k}$ with $0 \leq t_i \leq n_i$ ($i = 1, \dots, k$) belongs to the class $\underline{a}$ if and only if $t_i \equiv a_i \pmod 2$ for all $i = 1, \dots, k$. First observe that if there exists an $h \in \mathcal{H}$ such that h belongs to the class $\underline{0} = (0, 0, \dots, 0)$, then by the choice $h_1 = h_2 = h_3 = h$ we have $h_1h_2h_3 = \square$, and the statement follows. So from this point on we assume that $\mathcal{H}$ does not contain such an h.

Denote by $c_{\underline{a}}$ the number of elements in the class $\underline{a}$ of $(\mathcal{D}, \sim)$. A simple calculation yields that

$$c_{\underline{a}} = \prod_{i=1}^{k} \left(\frac{n_i}{2} + \varepsilon_{\underline{a},i} \right),$$

where

$$\varepsilon_{\underline{a},i} = \begin{cases} 1/2, & \text{if } n_i \text{ is odd,} \\ 0, & \text{if } n_i \text{ is even and } a_i = 1, \\ 1, & \text{if } n_i \text{ is even and } a_i = 0. \end{cases}$$

Let

$$A = \{\underline{a} : \text{there exists an } h \in \mathcal{H} \text{ belonging to the class } \underline{a} \text{ of } (\mathcal{D}, \sim)\}.$$

Then our assumption $\mathcal{H} > \mathcal{D}/2$ implies that

$$\sum_{\underline{a} \in A} c_{\underline{a}} > \frac{|\mathcal{D}|}{2},$$

which yields

$$\sum_{\underline{a} \in A} \prod_{i=1}^{k} \left(\frac{n_i}{2} + \varepsilon_{\underline{a},i} \right) > \frac{1}{2} \prod_{i=1}^{k} (n_i + 1).$$

By the definition of $\varepsilon_{\underline{a},i}$, after multiplying both sides by 2^k and cancelling the factors corresponding to the odd exponents n_i, we obtain

$$\sum_{\underline{a}\in A}\prod_{i\in I}(n_i+\delta_{\underline{a},i}) > 2^{k-1}\prod_{i\in I}(n_i+1), \tag{6}$$

where $I=\{i\in\{1,\dots,k\}: n_i \text{ is even}\}$ and for all $i\in I$

$$\delta_{\underline{a},i}=\begin{cases}0, & \text{if } a_i=1,\\ 2, & \text{if } a_i=0.\end{cases}$$

After expanding both sides of inequality (6), we get linear combinations of terms of the shape $n_{i_1}\dots n_{i_l}$ with distinct indices $i_1,\dots,i_l\in I$. Obviously, the coefficients of all terms $n_{i_1}\dots n_{i_l}$ at the right hand side of (6) are 2^{k-1}. If $I\neq\emptyset$, since $\underline{0}\notin A$, the constant term at the left hand side of (6) is zero. Let $s_{i_1,\dots,i_l}$ denote the coefficient of the corresponding non-constant term at the left hand side. Observe that in the summand corresponding to an $\underline{a}\in A$ the term $n_{i_1}\dots n_{i_l}$ occurs if and only if $\delta_{\underline{a},i}\neq 0$ for all $i\in T$, where $T=I\setminus\{i_1,\dots,i_l\}$. Note that by $l>0$ we have $|T|<k$. By the definition of the $\delta_{\underline{a},i}$ we have

$$s_{i_1,\dots,i_l}=2^{|T|}\cdot|B|, \tag{7}$$

where

$$B=\{\underline{a}\in A: a_i=0 \text{ for all } i\in T\}.$$

Then to have inequality (6), for at least one of these coefficients

$$s_{i_1,\dots,i_l}>2^{k-1} \tag{8}$$

must be valid. Combining (7) and (8) we obtain that

$$|B|>2^{k-|T|-1}. \tag{9}$$

The same inequality is true for $I=\emptyset=T$, $B=A$. Observe that if $|T|=k-1$, then by $\underline{0}\notin B$ we have $|B|\le 1$, contradicting (9). Hence we may suppose that $|T|\le k-2$. In this case we define a graph V in the following way. The vertices of V are those binary k-tuples $(r_1,\dots,r_k)$ for which $r_i=0$ holds for all $i\in T$. Two such tuples x and y are connected with an edge if and only if their sum modulo 2 belongs to B. Obviously, the number of vertices of V is $2^{k-|T|}$, and each vertex x of V is connected exactly to $|B|$ other vertices y of V. (Note that as $\underline{0}\notin B$, for all such y we have $x\neq y$.) Thus using (9), for the number of edges $|E|$ of V we get

$$|E|=2^{k-|T|-1}\cdot|B|>2^{k-|T|-1}\cdot 2^{k-|T|-1}.$$

Now Turán's theorem (see [3]) yields that V contains a triangle. If the vertices of this triangle are x,y,z, then we have that $b_1=x+y$,

$b_2 = x + z$ and $b_3 = y + z$ (all taken modulo 2) are distinct elements of B. This yields that

$$b_1 + b_2 = 2x + y + z = y + z = b_3 \tag{10}$$

modulo 2. Hence taking arbitrary $h_1, h_2, h_3 \in \mathcal{H}$ from the classes b_1, b_2, b_3 of $(\mathcal{D}, \sim)$, respectively, we have $h_1 h_2 h_3 = \square$, and the statement follows.

Now we prove that under the further assumptions, the elements $h_1, h_2, h_3 \in \mathcal{H}$ with $h_1 h_2 h_3 = \square$ can be chosen to be distinct. First observe that if $h_1 h_2 h_3 = \square$ such that none of h_1, h_2, h_3 belongs to the class $\underline{0}$ of $(\mathcal{D}, \sim)$, then they are necessarily distinct and we are done. So we can restrict our attention to the case when there exists a $h_1 \in \mathcal{H}$ such that h_1 belongs to the class $\underline{0}$, that is, h_1 is a square. Observe that then if for any distinct $h_2, h_3 \in \mathcal{H}$ with $h_1 \notin \{h_2, h_3\}$ we have $h_2 \sim h_3$, then $h_1 h_2 h_3 = \square$, and the statement follows. Thus in this case $\mathcal{H}$ may contain only at most one element from each class of $(\mathcal{D}, \sim)$ different from $\underline{0}$. Further, obviously $\mathcal{H}$ may contain at most two elements from $\underline{0}$. Moreover, from the proof of the first part of the theorem it follows that if $\mathcal{H}$ contains elements from more than 2^{k-1} classes of $(\mathcal{D}, \sim)$ outside the class $\underline{0}$, then we are done. (In fact this follows from the fact that b_1, b_2, b_3 are distinct in (10).) This altogether yields that

$$2^{k-1} + 2 \geq |\mathcal{H}| > \frac{|\mathcal{D}|}{2} = \frac{1}{2} \prod_{i=1}^{k} (n_i + 1)$$

must be valid, which gives

$$2^k + 4 > \prod_{i=1}^{k} (n_i + 1). \tag{11}$$

Consider first the case when $k \geq 3$. Then since m is not square-free by assumption, we have

$$\prod_{i=1}^{k} (n_i + 1) \geq 3 \cdot 2^{k-1} = 2^k + 2^{k-1},$$

which by (11) yields a contradiction. Assume next that $k = 2$. Then (11) gives

$$8 > (n_1 + 1)(n_2 + 1).$$

However, since m is assumed to be neither of the form $p_1 p_2$ or $p_1^2 p_2$, we get a contradiction again. Finally, let $k = 1$ and m is of the form $m = p_1^{n_1}$ with $n_1 \geq 5$. Then (11) provides a trivial contradiction, and the statement follows.

Remark 3. The prescribed assumptions to have three distinct divisors $h_1, h_2, h_3 \in \mathcal{H}$ such that $h_1h_2h_3 = \square$ are necessary. One can easily check that in each case below we have $|\mathcal{H}| > \mathcal{D}/2$, however, we do not have three distinct elements in $\mathcal{H}$ with the required property.

• If $m = 1$, then take $\mathcal{H} = \{1\}$.

• Let m be of the shape $m = p_1^2 p_2$, where p_1, p_2 are distinct primes. Take $\mathcal{H} = \{1, p_1, p_2, p_1^2 p_2\}$.

• Let m be of the form $m = p_1^{n_1}$, where p_1 is a prime and $2 \le n_1 \le 4$. Choose $\mathcal{H} = \{1, p_1, p_1^2\}$.

• Finally, if $m > 1$ is an arbitrary square-free integer, then let

$$\mathcal{H} = \{d : d \mid m \text{ and has an odd number of prime divisors}\} \cup \{1\}.$$

Proof of Remark 2. Take y so large that

$$\frac{1}{\sqrt{2\pi}} \int_{-y}^{y} e^{-x^2/2} dx > 1 - \frac{\epsilon}{2},$$

and $n > 3y^2$, m squarefree with $\omega(m) = n$ and

$$\mathcal{H} = \{d|m : \frac{1}{3}\Omega(m) \le \Omega(d) < \frac{2}{3}\Omega(m)\}.$$

Clearly $a, b \in \mathcal{H}$ implies $ab \notin \mathcal{H}$. On the other hand

$$\frac{1}{\tau(m)}|\mathcal{H}| = 2^{-n} \sum_{\frac{1}{3}n \le k < \frac{2}{3}n} \binom{n}{k} > 2^{-n} \sum_{\frac{n}{2} - \frac{y}{2}\sqrt{n} \le k < \frac{n}{2} + \frac{y}{2}\sqrt{n}} \binom{n}{k}.$$

By de Moivre-Laplace theorem the right hand side tends to

$$\frac{1}{\sqrt{2\pi}} \int_{-y}^{y} e^{-x^2/2} dx > 1 - \frac{\epsilon}{2},$$

hence for n large enough it is greater than $1 - \epsilon$.

3. Proof of Theorem 3. By virtue of (2) there exists $s > 1$ such that

$$\sum_{n \in \mathcal{A}} \frac{1}{n^s} > \alpha \sum_{n \in \mathcal{B}} \frac{1}{n^s}.$$

Multiplying this inequality by $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ we obtain

$$\sum_{n=1}^{\infty} \frac{\tau(n, \mathcal{A})}{n^s} > \alpha \sum_{n=1}^{\infty} \frac{\tau(n, \mathcal{B})}{n^s},$$

thus there exists $m \in \mathbb{N}$ satisfying (3).

Proof of Corollary 2. If (4) holds, then by a known theorem (see [2],p. 93)

$$\overline{D}(\mathcal{A},\mathcal{B}) \geq \liminf_{n\to\infty} \frac{\mathcal{A}(n)}{\mathcal{B}(n)} > \alpha$$

and Theorem 3 applies.

If (5) holds, then

$$\overline{D}(\mathcal{A},\mathcal{B}) = \limsup_{s\to 1+} \left((s-1) \sum_{n\in\mathcal{A}} \frac{1}{n^s} / (s-1) \sum_{n\in\mathcal{B}} \frac{1}{n^s} \right)$$

$$\geq \max \left(\frac{\overline{D}(\mathcal{A})}{\overline{D}(\mathcal{B})}, \frac{\underline{D}(\mathcal{A})}{\underline{D}(\mathcal{B})} \right) > \alpha$$

where $c/0 = \infty$ for $c > 0$. Theorem 3 applies again and gives (3).

Proof of Theorem 1. Apply Theorem 3 to the sets $\mathcal{B} = \mathbb{N}$ and

$$\mathcal{A}' = \mathcal{A} \setminus \{1, 2, \ldots, [x]\} \setminus \{1^2, 2^2, \ldots\}. \tag{12}$$

Since $\overline{D}(\mathcal{A}') = \overline{D}(\mathcal{A})$ we have $\overline{D}(\mathcal{A}') > 1/2$ and by Theorem 3 there exists an m such that

$$\tau(m, \mathcal{A}') > \frac{1}{2}\tau(m). \tag{13}$$

Now by Theorem 2 m has divisors $h_i \in \mathcal{A}' (i = 1, 2, 3)$ satisfying (1). However by the definition of $\mathcal{A}'$: $h_i > x$ and $h_i \neq \square$, thus h_i are distinct.

Proof of Corollary 1. We have (see [2], p. 87 and 97)

$$\overline{D}(\mathcal{A}) \geq \underline{D}(\mathcal{A}) \geq D_l(\mathcal{A}) \geq \delta^*(\mathcal{A})$$

Proof of Theorem 4. Apply Theorem 3 to the set $\mathcal{B}$ of squarefree numbers and the set $\mathcal{A}'$ given by (12). Since $D(\mathcal{B}) = 6/\pi^2$ (see [1], §152) we infer from Theorem 3 the existence of a number n such that

$$\tau(n, \mathcal{A}') > \frac{1}{2}\tau(n, \mathcal{B}).$$

Let m be the greatest squarefree divisor of n. Then every squarefree divisor of n is a divisor of m and we obtain (13). Further proof is the same as for Theorem 1. In order to prove the second part of Theorem 4 take a prime factor p of m. All divisors of m split into $\frac{1}{2}\tau(m)$ pairs $\{d, pd\}$, where $d | \frac{m}{p}$. By (13) there exists d such that $d \in \mathcal{A}'$ and $pd \in \mathcal{A}'$. It suffices to take $a = pd, b = d$.

References

[1] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Reprint Chelsea 1953.
[2] H. Ostmann, *Additive Zahlentheorie*, Erster Teil, Berlin 1956.
[3] P. Turan, *An extremal problem in graph theory*, Collected papers, 231–250.