On the density of primitive sets

R. Ahlswede and L. Khachatrian
Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
33501 Bielefeld

and

A. Sárközy *
Eötvös Loránd University
Department of Algebra and Number Theory
H–1117 Budapest, Pázmány Péter sétány 1/c
Hungary

* Research partially supported by Hungarian National Foundation for Scientific Research, Grant No. T029759, and DFG–Schwerpunkt Nr. 1078 "Quanteninformationsverarbeitung". This paper was written while A. Sárközy was visiting the Universität Bielefeld.

1 Introduction

c_1, c_2 denote positive absolute constants. If $f(x) = O\big(g(x)\big)$ as $x \to \infty$, then we write $f(x) \ll g(x)$; $f(x) \gg g(x)$ is defined analogously. If both $f(x) \ll g(x)$ and $f(x) \gg g(x)$ hold, then we write $f(x) \asymp g(x)$. The set of the positive integers is denoted by $\mathbb{N}$. A set $\mathcal{A} \subset \mathbb{N}$ is said to be primitive if there are no $a \in \mathcal{A}$, $a' \in \mathcal{A}$ with $a \neq a'$, $a|a'$. The family of the primitive sets $\mathcal{A} \subset \mathbb{N}$ is denoted by $\mathbb{P}$. A subscript N indicates if we restrict ourselves to integers not exceeding N, so that $\mathbb{N}_N = \{1, 2, \dots, N\}$, and $\mathbb{P}_N$ denotes the family of the primitive subsets of $\{1, 2, \dots, N\}$. The number of distinct prime factors of n is denoted by $\omega(n)$, while $\Omega(n)$ denotes the total number (counted with multiplicity) of prime factors of n:

$$\Omega(n) = \sum_{p^\alpha \| n} \alpha.$$

(Here $p^\alpha \| n$ denotes that $p^\alpha | n$ but $p^{\alpha+1} \nmid n$.)

It is well–known and easy to prove (see, e.g., [11] p. 244) that

$$\max_{\mathcal{A} \in \mathbb{P}_N} |\mathcal{A}| = N - [N/2] \quad \left(= \left(\frac{1}{2} + o(1) \right) N \right). \tag{1.1}$$

Behrend [3] proved that

$$\max_{\mathcal{A} \in \mathbb{P}_N} \sum_{a \in \mathcal{A}} \frac{1}{a} < c_1 \frac{\log N}{(\log \log N)^{1/2}} \tag{1.2}$$

for some absolute constant c_1 and all $N \geq 3$, and Pillai [13] showed that

$$\max_{\mathcal{A} \in \mathbb{P}_N} \sum_{a \in \mathcal{A}} \frac{1}{a} > c_2 \frac{\log N}{(\log \log N)^{1/2}} \tag{1.3}$$

for $N \geq 3$. Erdős [5] conjectured and Erdős, Sárközy and Szemerédi ([7], [8]) proved that

$$\max_{\mathcal{A} \in \mathbb{P}_N} \sum_{a \in \mathcal{A}} \frac{1}{a} = \big(1 + o(1)\big) \frac{\log N}{(2\pi \log \log N)^{1/2}} \quad \text{as} \quad N \to \infty. \tag{1.4}$$

Erdős [4] proved that

$$\sum_{a \in \mathcal{A}} \frac{1}{a \log a} < c_3 \quad \text{for all} \quad \mathcal{A} \in \mathbb{P} \quad \text{with} \quad 1 \notin \mathcal{A}. \tag{1.5}$$

These results have been extended in various directions. Surveys of the results on primitive sets are given in [11], [8], [1] and [15].

Each of (1.1), (1.2) and (1.5) provides an upper bound for a certain type of density of a primitive set $\mathcal{A} \in \mathbb{P}_N$. To make this assertion more precise, we use the following notation and definitions:

If f is a non–negative arithmetic function, then we write

$$S(f,\mathcal{A})=\sum_{a\in\mathcal{A}}f(a) \tag{1.6}$$

and, for $N\in\mathbb{N}$ and $\mathcal{A}\subset\{1,2,\dots,N\}$,

$$\delta(f,\mathcal{A},N)=\frac{S(f,\mathcal{A})}{S\big(f,\{1,2,\dots,N\}\big)}. \tag{1.7}$$

We call the function f a *weighting*, and for a given weighting f,

Definition 1. $\delta(f,\mathcal{A},N)$ is called the *f–density of $\mathcal{A}$ in* $\mathbb{N}_N$.

The estimates (1,1), (1,2) and (1,5) correspond to the weightings f_1, f_2 and f_3, where for $n\in\mathbb{N}$

$$f_1(n)=1, \tag{1.8}$$

$$f_2(n)=\frac{1}{n}, \tag{1.9}$$

resp.

$$f_3(n)=\frac{1}{n\log n}\left(\text{more precisely,}\;\; f_3(n)=\begin{cases}0 & \text{for } n=1\\ \frac{1}{n\log n} & \text{for } n>1\end{cases}\right). \tag{1.10}$$

Using this terminology, the results quoted above say that the maximal f–density of a primitive set in $\mathbb{N}_N$ is $\frac{1}{2}$, $O\left(\frac{1}{(\log\log N)^{1/2}}\right)$, resp. $O\left(\frac{1}{\log\log N}\right)$ for the weightings f_1, f_2 and f_3.

It is a natural question to ask what happens for other weightings. Can one make the maximal f–density of a primitive set even smaller under a suitable weighting f? In this form, of course, the question is too general; one needs certain restrictions on the weight function f. There are two natural directions of posing restrictions: first, one might want to study (analytically) "smooth" weightings and, secondly, in some applications it can be useful to have results on multiplicative weightings. Correspondingly, we introduce

Definition 2. The weighting f is said to be *smooth* if

(i) $0\le f(n)\le 1$ for all n

and there is a number $n_o\in\mathbb{N}$ such that

(ii) $f(n_o)>0$,

(iii) $f(n)\le f(n-1)$ for $n>n_o$.

Definition 3. The weighting f is said to be a *multiplicative* weighting if

(i) f is a multiplicative function;

(ii) $f(n) \geq 0$ for all n;

(iii) $f(1) = 1$ (so that $f(n) \not\equiv 0$).

Each of the weightings f_1, f_2 and f_3 is smooth, and f_1 and f_2 are also multiplicative, but f_3 is not multiplicative.

If f is a smooth or multiplicative weighting and $N \in \mathbb{N}$, then let $F(f, N)$ denote the maximal f-density of a primitive set in $\mathbb{N}_N$:

$$F(f, N) = \max_{\mathcal{A} \in \mathbb{P}_N} \delta(f, \mathcal{A}, N).$$

Then by (1.1), (1.4) and (1.5) we have

$$F(f_1, N) = \frac{1}{2} + o(1), \tag{1.11}$$

$$F(f_2, N) = \big(1 + o(1)\big) \frac{1}{(2\pi \log\log N)^{1/2}} \tag{1.12}$$

and

$$F(f_3, N) < \frac{c_4}{\log\log N} \tag{1.13}$$

so that out of these three weightings the Erdős weighting f_3 provides the best upper bound for the density of a primitive set.

In this paper our goal is to study the $F(f, N)$ function for both smooth and multiplicative weightings f.

2 The results

First we will study the following problem: how small can one make $F(f, N)$ for a suitable smooth weighting. It will turn out that the Erdős weighting f_3 is superior not only to the weightings f_1 and f_2 but, apart from at most a constant factor, it is optimal amongst all smooth weightings:

Theorem 1. *If $\varepsilon > 0$, f is a smooth weighting and $N > N_o(\varepsilon, f)$, then there is a set $\mathcal{A} \in \mathbb{P}_N$ with*

$$\delta(f, \mathcal{A}, N) > (1 - \varepsilon) \frac{1}{\log\log N}. \tag{2.1}$$

Note that Erdős, Sárközy and Szemerédi proposed [8]:

Conjecture. For all $\varepsilon > 0$ there is a number $K = K(\varepsilon)$ such that if $\mathcal{A} \in \mathbb{P}$ and

$$\mathcal{A} \cap [1, K] = \varnothing, \tag{2.2}$$

then we have

$$\sum_{a \in \mathcal{A}} \frac{1}{a \log a} < 1 + \varepsilon. \tag{2.3}$$

If this conjecture is true and $0 < \varepsilon < 1$, then for N large enough (in terms of ε) and all $\mathcal{A} \in \mathbb{P}_N$, for the smooth weighting

$$f(n) = \begin{cases} 0 & \text{for } n \le K \\ \frac{1}{nlogn} & \text{for } n > K \end{cases}$$

we have

$$\delta(f, \mathcal{A}, N) = \frac{\sum_{a \in \mathcal{A}} f(a)}{\sum_{n \le N} f(n)} = \frac{\sum_{\substack{a \in \mathcal{A} \\ K < a}} \frac{1}{a \log a}}{\sum_{K < n \le N} \frac{1}{nlogn}} < \frac{1 + \varepsilon}{(1 - \frac{\varepsilon}{2}) \log \log N} < \frac{1 + 2\varepsilon}{\log \log N}.$$

Thus for all $\varepsilon > 0$ there is a smooth weighting f with

$$F(f, N) < \frac{1 + \varepsilon}{\log \log N},$$

and by Theorem 1, the constant factor $1 + \varepsilon$ here cannot be replaced by $1 - \varepsilon$. In other words, assuming that the conjecture is true, even the best constant factor is known.

Next we will determine the order of magnitude of $F(f, N)$ for the most important special family of weightings, namely, for f_σ, where

$$f_\sigma(n) = \frac{1}{n^\sigma} \text{ and } 0 \le \sigma < \infty.$$

Indeed, by a theorem of Erdős [6] these are the only weightings which are simultaneously both, smooth and multiplicative. Besides, this family includes the important special cases $\sigma = 0$ and 1 when we obtain the weightings (1.8), resp. (1.9).

Theorem 2. *We have*

$$F(f_\sigma, N) \asymp 1 \quad \textit{for} \quad \sigma \ge 11/10, \tag{2.4}$$

$$F(f_\sigma, N) \asymp \big(\log\big(1/(\sigma - 1)\big)\big)^{-1/2} \quad \textit{for} \quad 11/10 > \sigma > 1 + 3/\log N, \tag{2.5}$$

$$F(f_\sigma, N) \asymp (\log \log N)^{-1/2} \quad \textit{for} \quad |\sigma - 1| \le 3/\log N, \tag{2.6}$$

$$F(f_\sigma, N) \asymp \big(\log\big(1/(1 - \sigma)\big)\big)^{-1/2} \quad \textit{for} \quad 1 - 3/\log N > \sigma > 9/10 \tag{2.7}$$

and

$$F(f_\sigma, N) \asymp 1 \quad \textit{for} \quad 9/10 \geq \sigma \geq 0. \tag{2.8}$$

While for most σ values these estimates are connected with known results, the proof will also contain two important new elements. First, there will be a new large family of primitive sets constructed for the case $\sigma \to 1-$ (formula (2.7)) which leads to a new problem of independent interest that we shall settle in the form of Theorem 3 below. Secondly, in the proof of Theorem 3, and also implicity in the proofs of the other cases in Theorem 2, there will be a new large family of primitive sets constructed (see formula (5.21)). This construction seems to be canonical in a certain sense, we will return to this problem in a subsequent paper.

Theorem 3. *If $N \in \mathbb{N}$, $3 \leq Q \leq N$ and $\mathcal{A}$ is a primitive set all whose elements a satisfy $N/Q < a \leq N$, then we have*

$$\sum_{a\in\mathcal{A}} \frac{1}{a} < c_5 \frac{\log Q}{(\log\log Q)^{1/2}}, \tag{2.9}$$

and this estimate is the best possible, i.e., there is a set $\mathcal{A}$ with

$$\mathcal{A} \subset \mathbb{P},\ \mathcal{A} \subset \mathbb{N}_N - \mathbb{N}_{N/Q} \tag{2.10}$$

and

$$\sum_{a\in\mathcal{A}} \frac{1}{a} > c_6 \frac{\log Q}{(\log\log Q)^{1/2}}. \tag{2.11}$$

Note that (2.9) was stated by Erdős, Sárközy and Szemerédi, however, no proof has ever been given. Since there are certain difficulties in adopting Behrend's method to prove this, for the sake of completeness we will present the proof here.

In the remaining part of the paper we will study multiplicative weightings.

The multiplicative analog of Theorem 1 is nearly trivial:

Proposition.

(1) If f is any weighting and $N \in \mathbb{N}$, with $\sum\limits_{n=1}^{N} f(n) > 0$, then

$$F(f, N) \geq \left(\left\lceil \frac{\log N}{\log 2} \right\rceil + 1 \right)^{-1}. \tag{2.12}$$

(ii) There is a multiplicative weighting f such that

$$F(f, N) = \left(\left\lfloor \frac{\log N}{\log 2} \right\rfloor + 1 \right)^{-1}. \tag{2.13}$$

(So that in the special case $N = 2^k$ we have

$$\min_f F(f, 2^k) = \frac{1}{k+1} \equiv \left(\frac{\log N}{\log 2} + 1 \right)^{-1}.)$$

Proof:

(i) Write $K = \left\lceil \frac{\log N}{\log 2} \right\rceil$. Then we have

$$S\big(f, \{1, 2, \dots, N\}\big) = \sum_{n=1}^{N} f(n) = \sum_{k=0}^{K} \sum_{2^{k-1} < n \le \min(N, 2^k)} f(n).$$

Here the greatest of the inner sums satisfies

$$\sum_{2^{k-1} < n \le \min(N, 2^k)} f(n) \ge \frac{1}{K+1} S\big(f, \{1, 2, \dots, N\}\big). \tag{2.14}$$

Let $\mathcal{A} = \big\{n : 2^{k-1} < n \le \min(N, 2^k)\big\}$ with a k satisfying (2.14). Then clearly $\mathcal{A} \in \mathbb{P}_N$, and by (2.14), we have

$$S(f, \mathcal{A}) \ge \frac{1}{K+1} S\big(f, \{1, 2, \dots, N\}\big)$$

and whence

$$\delta(f, \mathcal{A}, N) = \frac{S(f, \mathcal{A})}{S\big(f, \{1, 2, \dots, N\}\big)} \ge \frac{1}{K+1},$$

which proves (2.12).

(ii) Define the multiplicative function f by

$$f(n) = \begin{cases} 1 & \text{for } n = 2^\alpha \\ 0 & \text{for } n \ne 2^\alpha, \end{cases}$$

and write $L = \left\lfloor \frac{\log N}{\log 2} \right\rfloor$. Then we have

$$S\big(f, \{1, 2, \dots, N\}\big) = \sum_{n=1}^{N} f(n) = \sum_{2^\alpha \le N} 1 = L + 1, \tag{2.15}$$

and clearly, any set $\mathcal{A} \in \mathbb{P}_N$ may contain only one of the numbers $1, 2, \dots, 2^L$ and thus we have

$$S(f, \mathcal{A}) = \sum_{n \in \mathcal{A}} f(n) \le 1 \quad \text{for all} \quad \mathcal{A} \subset \mathbb{P}_N. \tag{2.16}$$

(2.13) follows from (2.15) and (2.16).

The proof of the Proposition above warns that if we want a reasonable lower bound for $F(f, n)$, then we must be able to control the values of $f(p^\alpha)$ for prime powers p^α with $\alpha > 1$. If f is completely multiplicative, then the primes p with $f(p) > 1$ also may cause a problem:

Theorem 4. *If f is a completely multiplicative function such that there is a prime p with $f(p) > 1$, then there are numbers $C = C(p, f(p)) > 0$, $N_o = N_o(p)$ such that for $N \in \mathbb{N}$, $N > N_o$, we have*

$$F(f, N) \geq C. \tag{2.17}$$

If we want a good upper bound for $F(f, N)$ (for multiplicative weightings f) then, by the Proposition and Theorem 4 above, it is reasonable to assume that

$$0 \leq f(p) \leq 1 \tag{2.18}$$

and

$$f(p^\alpha) = 0 \text{ for } \alpha \geq 2.$$

However, (2.18) is still too general to handle it, thus we will restrict ourselves to the most important special case when $f(p) = 0$ or 1:

Definition 4. A multiplicative weighting f is said to be a *combinatorial* weighting if

(i) $f(p) = 0$ or 1 for every prime p;

(ii) $f(p^\alpha) = 0$ for every prime p and $\alpha = 2, 3, \ldots$.

For fixed N and suitable combinatorial weighting f, $F(f, N)$ can be made as small as $c\left(\frac{\log\log N}{\log N}\right)^{1/2}$ (compare this with the bounds $\frac{c}{\log\log N}$ and $\frac{c}{\log N}$ obtained in the case of smooth weighting, resp. general multiplicative weighting):

Theorem 5. *If $N \in \mathbb{N}$, then there is a combinatorial weighting f with*

$$F(f, N) < c_7 \left(\frac{\log\log N}{\log N}\right)^{1/2}. \tag{2.19}$$

We conjecture that Theorem 5 is best possible apart from the value of the constant in (2.19):

Conjecture 1. For any fixed $N \in \mathbb{N}$ and every combinatorial weighting f we have

$$F(f, N) > c_8 \left(\frac{\log\log N}{\log N}\right)^{1/2}. \tag{2.20}$$

Unfortunately, we have not been able to prove this and, indeed, this seems to be difficult. However, we have been able to show that (2.20) holds at least in the two extreme cases when $f(p) = +1$ holds for "very few", resp. "almost all" primes. Moreover, in the first case we can prove a stronger result under the assumption that a well–known conjecture of Frankl holds.

A family of sets is said to be an *antichain* if none of the given sets contains another one. We say that a family $\mathcal{F}$ of sets is *convex* if for all $\mathcal{A}, \mathcal{B} \in \mathcal{F}$, $\mathcal{A} \subset \mathcal{B}$ and $\mathcal{A} \subset \mathcal{C} \subset \mathcal{B}$ we also have

$\mathcal{C} \in \mathcal{F}$, and the family $\mathcal{F}$ is said to be a *downset* if whenever $\mathcal{A} \in \mathcal{F}$ and $\mathcal{B} \subset \mathcal{A}$, then we also have $\mathcal{B} \in \mathcal{F}$.

Conjecture 2 (Frankl [9]). If $M \in \mathbb{N}$ and $\mathcal{F}$ is a non–empty convex family of subsets of a set $\mathcal{S}$ of cardinality M, then there is an antichain $\mathcal{B} \subset \mathcal{F}$ satisfying

$$\frac{|\mathcal{B}|}{|\mathcal{F}|} \geq \binom{M}{[M/2]} 2^{-M}.$$

For our purposes it suffices to use the following slightly weaker form of Frankl's conjecture:

Conjecture 2'. The statement of Conjecture 2 holds if we specialize it to downsets.

We will prove:

Theorem 6.

(i) If $N \in \mathbb{N}$ and f is a combinatorial weighting with

$$\prod_{\substack{p \leq N \\ f(p)=+1}} p \leq N, \tag{2.21}$$

then (2.20) holds (unconditionally).

(ii) If $N \in \mathbb{N}$ and f is a combinatorial weighting with

$$\prod_{\substack{p \leq N \\ f(p)=+1}} p \leq N^C, \tag{2.22}$$

then, assuming that Conjecture 2' is true, (2.20) must hold.

(We remark that in the last section we will return to Conjecture 2' and Theorem 6, (ii).)

Theorem 7. *If $N \in \mathbb{N}$ and f is a combinatorial weighting with*

$$\sum_{\substack{p \leq N \\ f(p)=0}} \frac{1}{p} < \frac{1}{4}, \tag{2.23}$$

then we have

$$F(f, N) > \frac{1}{5}.$$

(Which is much stronger than (2.20).)

We can also handle the case when all the primes p with $f(p) = 1$ lie in a short interval (this is certainly the most interesting case that we can handle):

Theorem 8. *If $N \in \mathbb{N}$, $N \geq 3$, x and y are positive real numbers with*

$$\exp\left(\left((\log x)^{-1} + (\log N)^{-1/2}\right)^{-1}\right) \leq y < x \leq N, \tag{2.24}$$

$\mathcal{P}$ is a set of primes so that

$$y < p \le x \text{ for all } p \in \mathcal{P}, \tag{2.25}$$

and f denotes the combinatorial weighting defined by

$$f(p) = \begin{cases} 1 & \text{for } p \in \mathcal{P} \\ 0 & \text{for } p \notin \mathcal{P}, \end{cases}$$

then we have

$$F(f,N) > c_9 \frac{1}{(\log N)^{1/2}}$$

(Note that, e.g., we may take $x = N$, $y = \exp\big((\log N)^{1/2}\big)$, or $x = \exp\big((\log N)^{1/2}\big)$, $y = \exp\left(\frac{1}{2}(\log N)^{1/2}\right)$).

With further assumptions (just a weak lower bound for $|\mathcal{P}|$ and making interval (2.24) slightly shorter) the lower bound for $F(f,N)$ could be improved considerably; see the remark at the end of the proof of the theorem.

Probably for a combinatorial weighting f, one cannot make $F(f,N)$ as small as in (2.19) *uniformly* in N. In this direction we will prove:

Theorem 9. *There is a combinatorial weighting f satisfying*

$$F(f,N) < c_{10} \frac{1}{(\log\log N)^{1/2}} \text{ for all } N \in \mathbb{N}, N \ge 3. \tag{2.26}$$

Again we conjecture that this is best possible apart from the value of the constant factor:

Conjecture 3. For every combinatorial weighting f we have

$$F(f,N) > c_{11} \frac{1}{(\log\log N)^{1/2}} \text{ for infinitely many } N \in \mathbb{N}. \tag{2.27}$$

Again, we can prove this only in the two extreme cases when $f(p) = +1$ holds for "very few", resp. "many" primes:

Theorem 10. *If f is a combinatorial weighting satisfying*

$$|\{p : p \text{ prime }, p \le N, f(p) = 1\}| < \log\log N \text{ for } n > N_o, \tag{2.28}$$

then (2.27) holds for infinitely many $N \in \mathbb{N}$.

Theorem 11. *If there are $C > O$ and N_o so that we have*

$$|\{p : p \text{ prime }, p \le N, f(p) = 1\}| > N^{C(\log\log N)^{-1/2}} \text{ for } N > N_o, \tag{2.29}$$

then there is a $c_{11} = c_{11}(C)$ so that (2.27) holds for infinitely many N.

3 Proof of Theorem 1

Assume first that

$$R \triangleq \sum_{n=1}^{+\infty} f(n) < +\infty. \tag{3.1}$$

Clearly $\{n_o\} \in \mathbb{P}_N$ (where n_o is the number defined in (ii) in Definition 2) for all $N \geq n_o$ whence

$$F(f,N) = \max_{\mathcal{A}\in\mathbb{P}_N} \delta(f,\mathcal{A},N) \geq \delta\big(f,\{n_o\},N\big\} = \frac{S\big(f,\{n_o\}\big)}{S\big(f,\{1,2,\dots,N\}\big)} = \frac{f(n_o)}{\sum\limits_{n=1}^{N} f(n)} \geq \frac{f(n_o)}{R}$$

and, by (3.1) and by (ii) in Definition 2 this is $> 1/(\log\log N)$ if N is large enough.

Assume now that

$$\lim_{N\to\infty} \sum_{n=1}^{N} f(n) = +\infty. \tag{3.2}$$

In this case, the proof will be based on

Lemma 1. *If $\varepsilon > 0$, $\eta > 0$ and $N > N_o(\varepsilon,\eta)$, then for all but ηN integers n not exceeding N we have*

$$|\Omega(n) - \log\log N| < \varepsilon \log\log N.$$

Proof of Lemma 1: This is a well–known result of Hardy and Ramanujan [10]. (See Lemma 6 for a sharper version of this result.)

Now write

$$\mathcal{B}(N,t) = \big\{n : n \leq N, \Omega(n) = t\big\},$$

$$\mathcal{M}(N,\varepsilon) = \bigcup_{t<(1+\varepsilon)\log\log N} \mathcal{B}(N,t)$$

and

$$\overline{\mathcal{M}}(N,\varepsilon) = \{1,2,\dots,N\} \setminus \mathcal{M}(N,\varepsilon) = \big\{n : n \leq N, \Omega(n) \geq (1+\varepsilon)\log\log N\big\}.$$

By Lemma 1, for $\varepsilon > 0$, $\eta > 0$ and $N > N_1(\varepsilon,\eta)$ we have

$$|\overline{\mathcal{M}}(N,\varepsilon)| < \eta N. \tag{3.3}$$

By (iii) in Definition 2 and (3.3), there is a number $k_o = k_o(\varepsilon) \in \mathbb{N}$ such that for $k \geq k_o$, $k \in \mathbb{N}$ we have

$$\sum_{\substack{2^k < n \leq 2^{k+1} \\ n \in \overline{\mathcal{M}}(2^{k+1}, \varepsilon/4)}} f(n) \leq f(2^k) \sum_{\substack{2^k < n \leq 2^{k+1} \\ n \in \overline{\mathcal{M}}(2^{k+1}, \varepsilon/4)}} 1 \leq f(2^k) |\overline{\mathcal{M}}(2^{k+1}, \varepsilon/4)| < \frac{\varepsilon}{8} 2^{k+1} f(2^k) \leq \frac{\varepsilon}{4} \sum_{n=2^{k-1}+1}^{2^k} f(n)$$

whence, defining K by $2^{K-1} < N \leq 2^K$, by (3.2) we have for some N sufficiently large $\sum\limits_{n \leq 2^{k_o}} f(n) \leq \frac{\varepsilon}{4} \sum\limits_{n=1}^{n} f(n)$ and thus

$$\sum_{n \in \overline{\mathcal{M}}(N, \varepsilon/2)} f(n) \leq \sum_{n \leq 2^{k_o}} f(n) + \sum_{k=k_o+1}^{K-1} \sum_{\substack{2^k < n \leq 2^{k+1} \\ n \in \overline{\mathcal{M}}(N, \varepsilon/2)}} f(n) \leq \frac{\varepsilon}{4} \sum_{n=1}^{N} f(n) + \sum_{k=k_o+1}^{K-1} \sum_{\substack{2^k < n \leq 2^{k+1} \\ n \in \overline{\mathcal{M}}(2^k, \varepsilon/4)}} f(n)$$

$$\leq \frac{\varepsilon}{4} \sum_{n=1}^{N} f(n) + \sum_{k=k_o+1}^{K-1} \frac{\varepsilon}{4} \sum_{n=2^{k-1}+1}^{2^k} f(n) \leq \left(\frac{\varepsilon}{4} + \frac{\varepsilon}{4} \right) \sum_{n=1}^{N} f(n) = \frac{\varepsilon}{2} \sum_{n=1}^{N} f(n).$$

It follows that

$$\sum_{t < (1+\varepsilon/2) \log\log N} \sum_{n \in \mathcal{B}(N,t)} f(n) = \sum_{n \in \mathcal{M}(N, \varepsilon/2)} f(n) = \sum_{n=1}^{N} f(n) - \sum_{n \in \overline{\mathcal{M}}(N, \varepsilon/2)} f(n) > \left(1 - \frac{\varepsilon}{2}\right) \sum_{n=1}^{N} f(n).$$

Let T denote the t value for which here the inner sum is maximal. Then by the pigeon hole principle we have

$$S\big(f, \mathcal{B}(N,T)\big) = \sum_{n \in \mathcal{B}(N,T)} f(n) > \frac{1 - \varepsilon/2}{(1+\varepsilon/2) \log\log N} \sum_{n=1}^{N} f(n) > \frac{1-\varepsilon}{\log\log N} \sum_{n=1}^{N} f(n)$$

$$= \frac{1-\varepsilon}{\log\log N} S\big(f, \{1, \dots, N\}\big)$$

whence

$$\delta\big(f, \mathcal{B}(N,T), N\big) = \frac{S\big(f, \mathcal{B}(N,T)\big)}{S\big(f, \{1, \dots, N\}\big)} > \frac{1-\varepsilon}{\log\log N}.$$

Since $\mathcal{B}(N,T) \in \mathbb{P}_N$ also holds trivially, this completes the proof of Theorem 1.

4 Proof of Theorem 2

We have to distinguish several cases. It will turn out that the cases when σ is outside the interval $1-\varepsilon<\sigma<1+\varepsilon$ are trivial, i.e., the problem can be reduced to $\sigma \to 1$. Moreover, if σ is "very close" to 1 (in terms of N) or $1-\varepsilon<\sigma<1$ but σ is "not very close" to 1, then we will reduce the problem to the theorems of Behrend, resp. Pillai. Thus the only really interesting case is when $1<\sigma<1+\varepsilon$ but σ is "not very close" to 1.

Case 1. Assume that $\sigma \geq \frac{11}{10}$. Then by $\{1\} \in \mathbb{P}_N$ we have

$$\max_{\mathcal{A}\in\mathbb{P}_N} S(f_\sigma,\mathcal{A}) \geq S\big(f_\sigma,\{1\}\big) = 1 \tag{4.1}$$

and, on the other hand,

$$S\big(f_\sigma,\{1,2,\ldots,N\}\big) < \sum_{n=1}^{+\infty}\frac{1}{n^\sigma} \leq \sum_{n=1}^{+\infty}\frac{1}{n^{11/10}} = c_{12}. \tag{4.2}$$

(2.4) follows from (4.1) and (4.2).

Case 2. Assume that $0 \leq \sigma < \frac{9}{10}$. Then with $\big\{n : [N/2] < n \leq N\big\} \in \mathbb{P}_N$ we have

$$\max_{\mathcal{A}\in\mathbb{P}_N} S(f_\sigma,\mathcal{A}) \geq \sum_{[N/2]<n\leq N}\frac{1}{n^\sigma} \geq \sum_{[N/2]<n\leq N}\frac{1}{N^\sigma} \geq \frac{N}{2}\cdot\frac{1}{N^\sigma} = \frac{1}{2}N^{1-\sigma}. \tag{4.3}$$

On the other hand,

$$S\big(f_\sigma,\{1,2,\ldots,N\}\big) = \sum_{n=1}^{N}\frac{1}{n^\sigma} < 1 + \int_1^N x^{-\sigma}dx < 1 + \frac{1}{1-\sigma}N^{1-\sigma} < c_{13}N^{1-\sigma}. \tag{4.4}$$

(2.8) follows from (4.3) and (4.4).

Case 3. Assume that $|\sigma-1| \leq \frac{3}{\log N}$. Then by Pillai's theorem, there is a primitive set $\mathcal{A} \in \mathbb{P}_N$ satisfying (1.3); let $\mathcal{A}_o$ denote such a set. Then we have

$$\begin{aligned}\max_{\mathcal{A}\in\mathbb{P}_N} S(f_\sigma,\mathcal{A}) \geq S(f_\sigma,\mathcal{A}_o) &= \sum_{a\in\mathcal{A}_o}\frac{1}{a^\sigma} \geq \sum_{a\in\mathcal{A}_o}\frac{1}{a^{1+(3/\log N)}} \\ &\geq \sum_{a\in\mathcal{A}_o}\frac{1}{a\cdot N^{3/\log N}} = e^{-3}\sum_{a\in\mathcal{A}_o}\frac{1}{a} \gg \frac{\log N}{(\log\log N)^{1/2}}.\end{aligned} \tag{4.5}$$

On the other hand, by Behrend's theorem (1.2), for all $\mathcal{A} \in \mathbb{P}_N$ we have

$$S(f_\sigma, \mathcal{A}) = \sum_{a\in\mathcal{A}} \frac{1}{a^\sigma} < \sum_{a\in\mathcal{A}} \frac{1}{a^{1-(3/\log N)}} \le \sum_{a\in\mathcal{A}} \frac{1}{a\cdot N^{-3/\log N}}$$
$$= e^3 \sum_{a\in\mathcal{A}} \frac{1}{a} \ll \frac{\log N}{(\log\log N)^{1/2}} \quad \text{(for all } \mathcal{A}\in\mathbb{P}_N). \tag{4.6}$$

Moreover, clearly we have

$$S\big(f_\sigma, \{1,2,\dots,N\}\big) = \sum_{n=1}^{N} \frac{1}{n^\sigma} \le \sum_{n=1}^{N} \frac{1}{n^{1-3/\log N}} \le \sum_{n=1}^{N} \frac{1}{nN^{-3/\log N}} = e^3 \sum_{n=1}^{N} \frac{1}{n} \ll \log N, \tag{4.7}$$

and a similar computation shows that

$$S\big(f_\sigma, \{1,2,\dots,N\}\big) \gg \log N. \tag{4.8}$$

(2.6) follows from (4.5), (4.6), (4.7) and (4.8).

Case 4. Assume that

$$1 + \frac{3}{\log N} < \sigma < \frac{11}{10}. \tag{4.9}$$

Write $M = \big[\exp\big\{1/(\sigma-1)\big\}\big]$. By (4.9) we have

$$\exp\big\{1/(\sigma-1)\big\} > \exp\{10\}$$

and

$$\exp\big\{1/(\sigma-1)\big\} < \exp\big\{(\log N)/3\big\} < N$$

whence

$$e^9 < M < N. \tag{4.10}$$

By (4.10) and Pillai's theorem (1.3), there is a set $\mathcal{A}_o \in \mathbb{P}_M \subset \mathbb{P}_N$ with

$$\sum_{a\in\mathcal{A}_o} \frac{1}{a} \gg \frac{\log M}{(\log\log M)^{1/2}} \gg \frac{1}{(\sigma-1)\big(\log(1/(\sigma-1))\big)^{1/2}}$$

whence

$$\max_{\mathcal{A}\in\mathbb{P}_N} S(f_\sigma, \mathcal{A}) > S(f_\sigma, \mathcal{A}_o) \gg \frac{1}{(\sigma-1)\big(\log(1/(\sigma-1))\big)^{1/2}}. \tag{4.11}$$

Moreover, a simple computation shows that

$$S\big(f_\sigma, \{1,2,\dots,N\}\big) = \sum_{n=1}^{N} \frac{1}{n^\sigma} \asymp \int_1^\infty \frac{dx}{x^\sigma} = \frac{1}{\sigma-1}. \tag{4.12}$$

It remains to give an upper bound for the maximum in (4.11). It follows from Behrend's theorem by partial summation that, writing

$$Z(\mathcal{A}, n) = \sum_{a \in \mathcal{A}, a \le n} \frac{1}{a},$$

for $\mathcal{A} \in \mathbb{P}_N$ we have

$$\sum_{a \in \mathcal{A}} \frac{1}{a^\sigma} = \sum_{\substack{a \in \mathcal{A} \\ a \le M}} \frac{1}{a^\sigma} + \sum_{\substack{a \in \mathcal{A} \\ a > M}} \frac{1}{a^\sigma} \le \sum_{\substack{a \in \mathcal{A} \\ a \le M}} \frac{1}{a} + \sum_{n=M+1}^{+\infty} \frac{Z(n) - Z(n-1)}{n^{\sigma-1}}$$

$$\le \frac{\log M}{(\log \log M)^{1/2}} + \sum_{n=M+1}^{+\infty} Z(n) \left(\frac{1}{n^{\sigma-1}} - \frac{1}{(n+1)^{\sigma-1}} \right)$$

$$\ll \frac{\log M}{(\log \log M)^{1/2}} + \sum_{n=M+1}^{+\infty} \frac{\log n}{(\log \log n)^{1/2}} \left(\frac{1}{n^{\sigma-1}} - \frac{1}{(n+1)^{\sigma-1}} \right)$$

$$\le \frac{1}{(\log \log M)^{1/2}} \left(\log M + \sum_{n=M+1}^{+\infty} (\log n) \left(\frac{1}{n^{\sigma-1}} - \frac{1}{(n+1)^{\sigma-1}} \right) \right)$$

$$= \frac{1}{(\log \log M)^{1/2}} \left(\log M + \sum_{n=M+1}^{+\infty} \frac{1}{n^{\sigma-1}} \big(\log n - \log(n-1) \big) + \frac{\log M}{(M+1)^{\sigma-1}} \right)$$

$$\ll \frac{1}{(\log \log M)^{1/2}} \left(\log M + \sum_{n=M+1}^{+\infty} \frac{1}{n^\sigma} \right) \ll \frac{1}{(\log \log M)^{1/2}} \left(\log M + \frac{M^{1-\sigma}}{\sigma - 1} \right)$$

$$\ll \frac{\log M}{(\log \log M)^{1/2}} \ll \frac{1}{(\sigma - 1)\big(\log(1/(\sigma-1))\big)^{1/2}} \qquad \text{(for all } \mathcal{A} \in \mathbb{P}_N). \tag{4.13}$$

(2.5) follows from (4.11), (4.12) and (4.13).

Case 5. Assume finally that

$$\frac{9}{10} < \sigma < 1 - \frac{3}{\log N}. \tag{4.14}$$

This is the most interesting case, and to handle it we need Theorem 3 which will be proved in the next section. Here we will show that, indeed, (2.7) in Theorem 2 follows from Theorem 3.

To give a lower bound for $F(f_\sigma, N)$, write

$$Q = \exp\big\{ 1/(1-\sigma) \big\}. \tag{4.15}$$

Then

$$3 \le Q \le N \tag{4.16}$$

follows from (4.14) so that Theorem 3 can be applied. By Theorem 3 there is a set $\mathcal{A}$ satisfying (2.10) and (2.11). For this set $\mathcal{A}$ we have

$$\begin{aligned} S(f_\sigma, \mathcal{A}) &= \sum_{a\in\mathcal{A}} \frac{1}{a^\sigma} = N^{1-\sigma} \sum_{a\in\mathcal{A}} \frac{1}{a}\left(\frac{a}{N}\right)^{1-\sigma} > N^{1-\sigma} \sum_{a\in\mathcal{A}} \frac{1}{a}\left(\frac{N/Q}{N}\right)^{1-\sigma} \\ &= \frac{1}{e} N^{1-\sigma} \sum_{a\in\mathcal{A}} \frac{1}{a} \gg N^{1-\sigma} \frac{\log Q}{(\log\log Q)^{1/2}} \\ &= \frac{N^{1-\sigma}}{(1-\sigma)\big(\log(1/(1-\sigma))\big)^{1/2}} \qquad \text{(for some } \mathcal{A} \in \mathbb{P}_N\text{).} \end{aligned} \tag{4.17}$$

Moreover, we have

$$S\big(f_\sigma, \{1, 2, \dots, N\}\big) = \sum_{n=1}^{N} \frac{1}{n^\sigma} \asymp \int_1^N \frac{dx}{x^\sigma} \asymp \frac{N^{1-\sigma}}{1-\sigma}. \tag{4.18}$$

Now consider a set $\mathcal{A} \in \mathbb{P}_N$, define Q again by (4.15), and define the positive integer K by

$$\frac{N}{Q^K} < Q \le \frac{N}{Q^{K-1}}.$$

Then clearly we have

$$S(f_\sigma, \mathcal{A}) = \sum_{a\in\mathcal{A}} \frac{1}{a^\sigma} \le \sum_{\substack{a\in\mathcal{A}\\ a\le Q}} \frac{1}{a^\sigma} + \sum_{k=1}^{K} \sum_{\substack{a\in\mathcal{A}\\ N/Q^k < a \le N/Q^{k-1}}} \frac{1}{a^\sigma}. \tag{4.19}$$

Here both the first sum and the inner sum in the second term are of the form

$$\sum_{\substack{a\in\mathcal{A}\\ M/Q < a \le M}} \frac{1}{a^\sigma}$$

where $\mathcal{A} \in \mathbb{P}$ and, by (4.16), $3 \le Q \le M$. Thus by Theorem 3 this sum is

$$= \sum_{\substack{a\in\mathcal{A}\\ M/Q < a \le M}} \frac{a^{1-\sigma}}{a} \le M^{1-\sigma} \sum_{\substack{a\in\mathcal{A}\\ M/Q < a \le M}} \frac{1}{a} \ll M^{1-\sigma} \frac{\log Q}{(\log\log Q)^{1/2}}$$

(note that, clearly, in Theorem 3 the assumption $N \in \mathbb{N}$ can be dropped). Thus it follows from (4.15), (4.16) and (4.19) that

$$S(f_\sigma, \mathcal{A}) \ll \left(Q^{1-\sigma} + \sum_{k=1}^{K} \left(\frac{N}{Q^{k-1}} \right)^{1-\sigma} \right) \frac{\log Q}{(\log\log Q)^{1/2}}$$

$$\leq \left(N^{1-\sigma} + N^{1-\sigma} \sum_{k=1}^{N} \frac{1}{e^{k-1}} \right) \frac{1}{(1-\sigma)\bigl(\log(1/(1-\sigma))\bigr)^{1/2}}$$

$$\ll \frac{N^{1-\sigma}}{(1-\sigma)\bigl(\log(1/(1-\sigma))\bigr)^{1/2}} \qquad \text{(for all } \mathcal{A} \in \mathbb{P}_N). \tag{4.20}$$

(2.7) follows from (4.17), (4.18) and (4.20), and this completes the proof of Theorem 2.

5 Proof of Theorem 3

In order to prove (2.9), we have to show that if $N \in \mathbb{N}$, $3 \leq Q \leq N$, $\mathcal{A} \subset \mathbb{N}$

$$N/Q < a \leq N \quad \text{for all} \quad a \in \mathcal{A} \tag{5.1}$$

and

$$\sum_{a \in \mathcal{A}} \frac{1}{a} \geq C \frac{\log Q}{(\log\log Q)^{1/2}} \tag{5.2}$$

where C is large enough, then there are a, a' with

$$a \in \mathcal{A},\ a' \in \mathcal{A},\ a \neq a',\ a | a'. \tag{5.3}$$

Write all $a \in \mathcal{A}$ as the product of a square and a squarefree integer:

$$a = \bigl(u(a)\bigr)^2 v(a),\ u(a) \in \mathbb{N},\ v(a) \in \mathbb{N},\ |\mu\bigl(v(a)\bigr)| = 1$$

(where μ is the Möbius function). Let

$$\mathcal{A}_u = \bigl\{ a : a \in \mathcal{A},\ u(a) = u \bigr\}$$

so that

$$\sum_{a \in \mathcal{A}} \frac{1}{a} = \sum_{u=1}^{+\infty} \frac{1}{u^2} \sum_{\substack{a \in \mathcal{A} \\ u(a) = u}} \frac{1}{v(a)}. \tag{5.4}$$

Since

$$\sum_{u=1}^{+\infty} \frac{1}{u^2} = \frac{\pi^2}{6} < 2,$$

it follows from (5.2) and (5.4) that there is a number $u = u_o$ for which the innermost sum in (5.4) is at least half of the lower bound in (5.2):

$$\sum_{\substack{a \in \mathcal{A} \\ u(a)=u_o}} \frac{1}{v(a)} \geq \frac{C}{2} \frac{\log Q}{(\log\log Q)^{1/2}}. \tag{5.5}$$

Write $M = N/u_o^2$ and $\mathcal{B} = \{b : \text{there is an } a \in \mathcal{A} \text{ with } u(a) = u_o,\ v(a) = b\}$. It follows from (5.1), (5.5) and the definition of $\mathcal{B}$ that

$$M/Q < b \leq M \quad \text{for all} \quad b \in \mathcal{B}, \tag{5.6}$$

$$\text{all} \quad b \in \mathcal{B} \quad \text{are squarefree} \tag{5.7}$$

and

$$\sum_{b \in \mathcal{B}} \frac{1}{b} \geq \frac{C}{2} \frac{\log Q}{(\log\log Q)^{1/2}}. \tag{5.8}$$

Moreover, if

$$b \in \mathcal{B},\ b' \in \mathcal{B},\ b \neq b',\ b|b', \tag{5.9}$$

then defining a, a' by $a = u_o^2 b$, $a' = u_o^2 b'$, clearly (5.3) holds so that it suffices to show that assuming (5.6), (5.7) and (5.8), there are b, b' satisfying (5.9). (In other words, we have reduced the problem to the case when the given set consists of squarefree integers.)

Write $d(n, \mathcal{B}) = |\{b : b \in \mathcal{B}, b|n\}|$, $w_Q(n) = |\{p : p \text{ prime }, p \leq Q, p|n\}|$ and $\overline{n} = \prod_{p|n} p$. We need several lemmas.

Lemma 2.

$$\prod_{p \leq Q} \left(1 + \frac{1}{p}\right) < c_{14} \log Q.$$

Proof of Lemma 2: By the well–known formula

$$\sum_{p \leq x} \frac{1}{p} = \log\log x + c_{15} + O\left(\frac{1}{\log x}\right)$$

(see, e.g., [12, p. 20]) and since $1 + x < e^x$, we have

$$\prod_{p \leq Q} \left(1 + \frac{1}{p}\right) < \prod_{p \leq Q} \exp(1/p) = \exp\left(\sum_{p \leq Q} \frac{1}{p}\right) \ll \exp(\log\log Q) = \log Q.$$

Lemma 3.

$$\sum_{n \leq x} 2^{w_Q(n)} < c_{16} x \log Q.$$

Proof of Lemma 3: By Lemma 2 we have

$$\sum_{n\le x} 2^{w_Q(n)} = \sum_{n\le x}\sum_{k=0}^{w_Q(n)} \binom{w_Q(n)}{k} = \sum_{n\le x}\left(1+\sum_{k=1}^{w_Q(n)} \sum_{\substack{p_{i_1}<\dots<p_{i_k}\le Q\\ p_{i_1}\dots p_{i_k}\mid n}} 1\right)$$

$$= x + \sum_{p_{i_1}<\dots<p_{i_k}\le Q} \sum_{\substack{n\le x\\ p_{i_1}\dots p_{i_k}\mid n}} 1 = x + \sum_{p_{i_1}<\dots<p_{i_k}\le Q} \left[\frac{x}{p_{i_1}\dots p_{i_k}}\right]$$

$$\le x + x\sum_{p_{i_1}<\dots<p_{i_k}\le Q} \frac{1}{p_{i_1}\dots p_{i_k}} = x\prod_{p\le Q}\left(1+\frac{1}{p}\right) \ll x\log Q.$$

Lemma 4. *If $\mathcal{B}$ satisfies (5.6), (5.7) and (5.8), then*

$$\sum_{M/Q<n\le M} d(\overline{n},\mathcal{B}) > \frac{C}{4}M\frac{\log Q}{(\log\log Q)^{1/2}}.$$

Proof of Lemma 4: By (5.6), (5.7) and (5.8) we have

$$\sum_{M/Q<n\le M} d(\overline{n},\mathcal{B}) = \sum_{M/Q<n\le M} d(n,\mathcal{B}) = \sum_{M/Q<n\le M}\sum_{\substack{b\mid n\\ b\in\mathcal{B}}} 1 = \sum_{b\in\mathcal{B}}\sum_{\substack{M/Q<n\le M\\ b\mid n}} 1$$

$$= \sum_{b\in\mathcal{B}}\sum_{\substack{n\le M\\ b\mid n}} 1 = \sum_{b\in\mathcal{B}}\left[\frac{M}{b}\right] > \sum_{b\in\mathcal{B}}\frac{M}{2b} = \frac{M}{2}\sum_{b\in\mathcal{B}}\frac{1}{b} > \frac{C}{4}M\frac{\log Q}{(\log\log Q)^{1/2}}.$$

Lemma 5. (Sperner [16]) *If S is a finite set and $S_1, S_2, \dots, S_t$ are distinct subsets of S with*

$$t > \binom{|S|}{[|S|/2]}, \tag{5.10}$$

then there are i, j such that $i \ne j$ and $S_i \subset S_j$.

Now we are ready to complete the proof of Theorem 3. First we will show that there is a positive integer n satisfying

$$M/Q < n \le M, \tag{5.11}$$

$$d(\overline{n},\mathcal{B}) > \frac{C}{10c_{13}}\,\frac{2^{w_Q(n)}}{(\log\log Q)^{1/2}} \tag{5.12}$$

and

$$2^{w_Q(n)} > \frac{C}{10} \frac{\log Q}{(\log \log Q)^{1/2}}. \tag{5.13}$$

Let $\mathcal{N}$ denote the set of the integers n satisfying (5.11), (5.12) and (5.13), write $\mathcal{N}_0 = \{n : M/Q < n \leq M\}$, let $\mathcal{N}_1$ denote the set of the integers n satisfying (5.11) and

$$d(\overline{n}, \mathcal{B}) \leq \frac{C}{10c_{15}} \frac{2^{w_Q(n)}}{(\log \log Q)^{1/2}},$$

finally, let $\mathcal{N}_2$ denote the set of the integers n satisfying (5.11), (5.12) and

$$2^{w_Q(n)} \leq \frac{C}{10} \frac{\log Q}{(\log \log Q)^{1/2}}.$$

Then clearly we have

$$\mathcal{N} = (\mathcal{N}_0 \setminus \mathcal{N}_1) \setminus \mathcal{N}_2$$

so that, by Lemmas 3 and 4,

$$\begin{aligned}
&\sum_{n \in \mathcal{N}} d(\overline{n}, \mathcal{B}) \geq \sum_{n \in \mathcal{N}_0} d(\overline{n}, \mathcal{B}) - \sum_{n \in \mathcal{N}_1} d(\overline{n}, \mathcal{B}) - \sum_{n \in \mathcal{N}_2} d(\overline{n}, \mathcal{B}) \\
&> \frac{C}{4} M \frac{\log Q}{(\log \log Q)^{1/2}} - \sum_{n \in \mathcal{N}_1} \frac{C}{10c_{16}} \frac{2^{w_Q(n)}}{(\log \log Q)^{1/2}} - \sum_{n \in \mathcal{N}_2} \frac{C}{10} \frac{\log Q}{(\log \log Q)^{1/2}} \\
&\geq \frac{C}{4} m \frac{\log Q}{(\log \log Q)^{1/2}} - \frac{C}{10c_{16}(\log \log Q)^{1/2}} \sum_{n \leq M} 2^{w_Q(n)} - \frac{C}{10} M \frac{\log Q}{(\log \log Q)^{1/2}} \\
&> \left(\frac{C}{4} - \frac{C}{10} - \frac{C}{10} \right) M \frac{\log Q}{(\log \log Q)^{1/2}} = \frac{C}{20} M \frac{\log Q}{(\log \log Q)^{1/2}} > 0.
\end{aligned}$$

So that, indeed, $\mathcal{N}$ is non–empty, i.e., there is an integer n satisfying (5.11), (5.12) and (5.13).

For such an integer n, let $b_1 < b_2 < \cdots < b_t$ denote all the integers b with $b \in \mathcal{B}$, $b|n$, so that, by (5.12),

$$t = d(\overline{n}, \mathcal{B}) > \frac{C}{10c_{16}} \frac{2^{w_Q(n)}}{(\log \log Q)^{1/2}}. \tag{5.14}$$

For $i = 1, 2, \ldots, t$, define the integer q_i by

$$q_i = \frac{\overline{n}}{b_i}. \tag{5.15}$$

Then by (5.6) and (5.11) we have

$$q_i = \frac{\overline{n}}{b_i} \leq \frac{n}{b_i} < \frac{M}{M/Q} = Q. \tag{5.16}$$

Now write

$$S = \{p : p \text{ prime}, \ p \le Q, \ p|n\}$$

and

$$S_i = \{p : p \text{ prime}, \ p|q_i\} \ \text{ for } \ i = 1, 2, \dots, t$$

so that

$$|S| = w_Q(n), \tag{5.17}$$

and, by (5.16), $S_1, \dots, S_t$ are subsets of S. In order to use Lemma 5, first we have to show that (5.10) in the Lemma holds. By Stirling's formula and (5.17), the right hand side of (5.10) is

$$\binom{|S|}{[|S|/2]} = \binom{w_Q(n)}{[w_Q(n)/2]} < c_{17} \frac{2^{w_Q(n)}}{(w_Q(n))^{1/2}}. \tag{5.18}$$

If C is large enough, then it follows from (5.13) for all $Q \ge 3$ that

$$w_Q(n) > \log\log Q. \tag{5.19}$$

It follows from (5.18) and (5.19) that

$$\binom{|S|}{[|S|/2]} < c_{17} \frac{2^{w_Q(n)}}{(\log\log Q)^{1/2}}. \tag{5.20}$$

By (5.14) and (5.20), (5.10) would follow from

$$c_{14} \frac{2^{w_Q(n)}}{(\log\log Q)^{1/2}} < \frac{C}{10c_{16}} \ \frac{2^{w_Q(n)}}{(\log\log Q)^{1/2}}.$$

If C is large enough ($C > 10c_{13}c_{14}$), then this holds so that, indeed, Lemma 5 can be applied. We obtain that there are i, j with $i \ne j$, $S_i \subset S_j$. Then, clearly, $q_i|q_j$. By (5.15), it follows that $b_j|b_i$ with $j \ne i$, Then (5.9) holds with $b = b_j$, $b' = b_i$, and this completes the proof of (2.9).

In order to prove (2.10) and (2.11), we need the following lemma:

Lemma 6. *For all $\varepsilon > 0$ there is a number $K = K(\varepsilon)$ such that if $3 \le Q \le N$ then*

$$|\Omega_Q(n) - \log\log Q| < K(\log\log Q)^{1/2}$$

holds for all but εN positive integers n not exceeding N.

Proof of Lemma 6: This follows from the Turán–Kubilius inequality [12].

Write

$$\mathcal{A}(N, Q, t) = \big\{n : N/Q < n \le N, \Omega_Q(n) = t\big\}. \tag{5.21}$$

Clearly, each of these sets satisfies (2.10). It follows from Lemma 6 by partial summation that for all $\varepsilon > 0$ there is a number $L = L(\varepsilon)$ such that for all $3 \le Q \le N$ we have

$$\sum_t \sum_{a \in \mathcal{A}(N,Q,t)} \frac{1}{a} > \left(1 - \frac{\varepsilon}{2}\right) \sum_{N/Q < n \le N} \frac{1}{n} > (1-\varepsilon) \log Q \tag{5.22}$$

where in $\sum_t$ we sum over all $t \in \mathbb{N}$ such that

$$|t - \log\log Q| < L (\log\log Q)^{1/2}. \tag{5.23}$$

Now we fix an ε value, say let $\varepsilon = 1/2$, and let T denote a t value (satisfying (5.23)) for which the innermost sum in (5.22) is maximal. Then by the pigeon hole principle we have

$$\sum_{a \in \mathcal{A}(N,Q,T)} \frac{1}{a} > \frac{(1/2)\log Q}{3L(1/2)(\log\log Q)^{1/2}}$$

so that (2.11) also holds and this completes the proof of Theorem 3.

We remark that the construction at the end of the proof could be made more explicit by using deeper information on the distribution of the number of prime factors and, indeed, it could be shown with a little work that $\mathcal{A} = \mathcal{A}\big(N, Q, [\log\log Q]\big)$ satisfies (2.10) and (2.11).

6 Proof of Theorem 4

Write $f(p) = D$ (> 1). Let $N > p^2$ and define the positive integer k by

$$p^k < N \le p^{k+1}. \tag{6.1}$$

Since f is completely multiplicative, by (6.1) we have

$$\sum_{n=1}^{N} f(n) \ge \sum_{n=1}^{p^k} f(n) \ge \sum_{i=1}^{p^{k-1}} f(ip) = f(p) \sum_{i=1}^{p^{k-1}} f(i) = D \sum_{n=1}^{p^{k-1}} f(n).$$

It follows that

$$\sum_{n=p^{k-1}+1}^{N} f(n) = \sum_{n=1}^{N} f(n) - \sum_{n=1}^{p^{k-1}} f(n) \ge \sum_{n=1}^{N} f(n) - \frac{1}{D} \sum_{n=1}^{N} f(n) = \frac{D-1}{D} \sum_{n=1}^{N} f(n). \tag{6.2}$$

Define L by

$$2^{L-1} p^{k-1} < N \le 2^L p^{k-1}. \tag{6.3}$$

Then by (6.1) and (6.3) we have

$$2^L < \frac{2N}{p^{k-1}} = \frac{2p^2 N}{p^{k+1}} \le 2p^2 \le p^3$$

whence

$$L < \frac{\log p^3}{\log 2} < 6 \log p.$$

Write

$$\mathcal{A}_t = \{2^{t-1}p^{k-1} + 1, 2^{t-1}p^{k-1} + 2, \ldots, 2^t p^{k-1}\} \text{ for } 1 \le t \le L - 1$$

and

$$\mathcal{A}_L = \{2^{L-1}p^{k-1} + 1, \ldots, N\}.$$

Then clearly

$$\mathcal{A}_t \in \mathbb{P}_N \text{ for } 1 \le t \le L, \tag{6.4}$$

and we have

$$\sum_{n=p^{k-1}+1}^{N} f(n) = \sum_{t=1}^{L} \sum_{n \in \mathcal{A}_t} f(n). \tag{6.5}$$

Let T denote the t value for which the inner sum on the right hand side is maximal. Then by (6.2) and (6.5) we have

$$S(f, \mathcal{A}_T) = \sum_{n \in \mathcal{A}_T} f(n) \ge \frac{1}{L} \sum_{n=p^{k-1}+1}^{N} f(n) \ge \frac{D-1}{LD} \sum_{n=1}^{N} f(n) = \frac{D-1}{LD} S(f, \{1, 2, \ldots, N\})$$

whence

$$\delta(f, \mathcal{A}_T, N) = \frac{S(f, \mathcal{A}_T)}{S(f, \{1, 2, \ldots, N\})} \ge \frac{D-1}{LD}. \tag{6.6}$$

It follows from (6.4) and (6.6) that

$$F(f, N) \ge \frac{D-1}{LD}$$

so that (2.17) holds with $\frac{D-1}{LD}$ in place of C and this completes the proof of Theorem 4.

7 Proof of Theorem 5

Define the set $\mathcal{P}$ by

$$\mathcal{P} = \{p : p \text{ prime}, p \le \frac{1}{2} \log N\}$$

so that, by the prime number theorem, for $N > N_o$ we have

$$|\mathcal{P}| > \frac{1}{3} \frac{\log N}{\log \log N} \tag{7.1}$$

and

$$\prod_{p \in \mathcal{P}} p < \exp(\log N) = N, \tag{7.2}$$

and define the arithmetic function f by

(a) f is multiplicative;

(b) if p is a prime, then

$$f(p) = \begin{cases} 1 & \text{for } p \in \mathcal{P}, \\ 0 & \text{for } p \notin \mathcal{P}; \end{cases}$$

(c) if p is a prime, then we have

$$f(p^\alpha) = 0 \text{ for } \alpha = 2, 3, \dots .$$

Then clearly, f is a combinatorial weighting. Moreover, $f(n) = 1$ if and only if n is of the form

$$n = \prod_{p \in \mathcal{P}(n)} p \text{ with some } \mathcal{P}(n) \subset \mathcal{P}, \tag{7.3}$$

and for every other $n \in \mathbb{N}$ we have $f(n) = 0$.

By (7.2), for each n of this form we have

$$n = \prod_{p \in \mathcal{P}(n)} p \leq \prod_{p \in \mathcal{P}} p < N.$$

The number of the integers of form (7.3) is equal to the number of subsets $\mathcal{P}(n)$ of $\mathcal{P}$, so that there are $2^{|\mathcal{P}|}$ integers n of this form. It follows that

$$S(f, \{1, 2, \dots, N\}) = \sum_{n=1}^{N} f(n) = \sum_{n: f(n)=1} 1 = 2^{|\mathcal{P}|}. \tag{7.4}$$

Now consider a set $\mathcal{A} \in \mathbb{P}_N$, and let

$$\mathcal{A}^* = \{a : a \in \mathcal{A}, f(a) = 1\} = \{a : a \in \mathcal{A}, a | \prod_{p \in \mathcal{P}} p\}.$$

Then every $a \in \mathcal{A}^*$ is of the form (7.3), and $\mathcal{A}^*$ is primitive set, thus for $a_1 \in \mathcal{A}^*$, $a_2 \in \mathcal{A}^*$, $a_1 \neq a_2$ we cannot have

$$\mathcal{P}(a_1) \subset \mathcal{P}(a_2).$$

Thus by Lemma 5 (Sperner's theorem) and (7.1) we have

$$|\mathcal{A}^*| \leq \binom{|\mathcal{P}|}{[|\mathcal{P}|/2]} < c_{18} \frac{2^{|\mathcal{P}|}}{|\mathcal{P}|^{1/2}} < c_{19} \frac{2^{|\mathcal{P}|}}{\left(\frac{\log N}{\log \log N}\right)^{1/2}}. \tag{7.5}$$

It follows from (7.4) and (7.5) that for all $\mathcal{A} \in \mathbb{P}_N$ we have

$$\delta(f, \mathcal{A}, N) = \frac{S(f, \mathcal{A})}{S(f, \{1, 2, \dots, N\})} = \frac{S(f, \mathcal{A}^*)}{S(f, \{1, 2, \dots, N\})} = \frac{|\mathcal{A}^*|}{2^{|\mathcal{P}|}} < c_{19} \left(\frac{\log \log N}{\log N}\right)^{1/2}$$

which proves (2.19).

8 Proof of Theorem 6

(i) Let p_i denote the i^{th} prime: $p_1 = 2$, $p_2 = 3$, $p_3 = 5, \ldots$, and write

$$\mathcal{P} = \{p_{i_1}, p_{i_2}, \ldots, p_{i_t}\} = \{p : \text{ prime }, p \le N, f(p) = +1\}. \tag{8.1}$$

Then by (2.21) and the prime number theorem we have

$$N \ge \prod_{p\in\mathcal{P}} p = \prod_{j=1}^{t} p_{i_j} \ge \prod_{j=1}^{t} p_j = \exp\big((1+o(1))t\log t\big)$$

whence

$$(1+o(1))t\log t \le \log N$$

so that

$$t \le (1+o(1))\frac{\log N}{\log\log N}. \tag{8.2}$$

By (2.21) clearly we have $n \le N$, $f(n) = 1$ if and only if $n \mid \prod_{p\in\mathcal{P}} p$ so that

$$S\big(f, \{1,2,\ldots,N\}\big) = \sum_{n=1}^{N} f(n) = \left|\left\{ n : n \mid \prod_{p\in\mathcal{P}} p \right\}\right| = 2^t. \tag{8.3}$$

Moreover, clearly the set

$$\mathcal{A} = \left\{ a : a \mid \prod_{p\in\mathcal{P}} p, w(a) = [t/2] \right\}$$

satisfies $\mathcal{A} \in \mathbb{P}_N$, and we have

$$S(f,\mathcal{A}) = \sum_{a\in\mathcal{A}} f(a) = \sum_{a\in\mathcal{A}} 1 = |\mathcal{A}| = \binom{t}{\left[\frac{t}{2}\right]}. \tag{8.3}$$

By (8.2) and (8.3) we have

$$F(f,N) \ge \delta(f,\mathcal{A},N) = \frac{S(f,\mathcal{A})}{S(f,\{1,2,\ldots,N\})} = \binom{t}{\left[\frac{t}{2}\right]} 2^{-t} > c_{20} t^{-1/2}$$

whence, by (8.1), (2.20) follows.

(ii) Define $\mathcal{P}$ and t again by (8.1). Replacing (2.21) by (2.22) in the proof of (8.2), in the same way we obtain

$$t \le (c+o(1))\frac{\log N}{\log\log N}, \tag{8.4}$$

and again (8.3) holds. Now to define a "large" set $\mathcal{A} \subset \mathbb{P}_N$ ("large" in terms of the weighting f), we will use the statement of Conjecture 2' (which is assumed to be true now). We use Conjecture 2' with $\mathcal{P}$ in place of S, and we define the family $\mathcal{F}$ so that for $\mathcal{R} \subset \mathcal{P}$ we have $\mathcal{R} \in \mathcal{F}$ if and only if $\prod_{p\in\mathcal{R}} p \le N$. Then clearly $\mathcal{F}$ is a downset, so that we may apply Conjecture 2'. We obtain that there is an antichain $\mathcal{B}$ with

$$|\mathcal{B}| \ge |\mathcal{F}| \binom{M}{[M/2]} 2^{-M}. \tag{8.5}$$

Here we have

$$|\mathcal{F}| = \left| \left\{ \mathcal{R} : \mathcal{R} \subset \mathcal{P}, \prod_{p\in\mathcal{R}} p \le N \right\} \right| = \left| \left\{ n : n \mid \prod_{p\in\mathcal{P}} p, n \le N \right\} \right| = |\{n : f(n) = +1, n \le N\}|$$

$$= \sum_{n=1}^{N} f(n) = S\big(f, \{1, 2, \dots, N\}\big) \tag{8.6}$$

and

$$M = |\mathcal{S}| = |\mathcal{P}| = t. \tag{8.7}$$

Now define the set $\mathcal{A}$ of positive integers so that $a \in \mathcal{A}$ if and only if there is an $\mathcal{R} \in \mathcal{B}$ with $\prod_{p\in\mathcal{R}} p = a$. It follows from the definition of $\mathcal{F}$ that for all $a \in \mathcal{A}$ we have $a \le N$, and $\mathcal{A}$ is primitive since $\mathcal{B}$ is an antichain, so that we have $\mathcal{A} \subset \mathbb{P}_N$.
By (8.5), (8.6) and (8.7) we have

$$S(f, \mathcal{A}) = \sum_{a\in\mathcal{A}} f(a) = \sum_{a\in\mathcal{A}} 1 = |\mathcal{A}|$$
$$= |\mathcal{B}| \ge S\big(f, \{1, 2, \dots, N\}\big) \binom{t}{[t/2]} 2^{-t} \gg S\big(f, \{1, 2, \dots, N\}\big) t^{-1/2}$$

whence

$$\delta(f, \mathcal{A}, N) = \frac{S(f, \mathcal{A})}{S(f, \{1, 2, \dots, N\})} \gg t^{-1/2}.$$

By (8.2), the result follows.

9 Proof of Theorem 7

Write

$$\mathcal{A} = \{a : N/2 < a \le N, f(a) = 1\}.$$

Then clearly we have $\mathcal{A} \in \mathbb{P}_N$. Moreover, by (2.23) clearly we have

$$\begin{aligned}
&S(f,\mathcal{A}) = \sum_{a\in\mathcal{A}} f(a) = |\mathcal{A}| = |\{a : N/2 < a \le N, f(a) = 1\}| \\
&= |\{a : N/2 < a \le N\}| - |\{a : N/2 < a \le N, f(a) = 0\}| \\
&= \big(N - [N/2]\big) - |\{a : N/2 < a \le N, \exists\, p \text{ prime with } p \le N, f(p) = 0, p \mid a\}| \\
&\ge \frac{N}{2} - \sum_{\substack{p\le N\\ f(p)=0}} |\{a : N/2 < a \le N, p \mid a\}| \\
&\ge \frac{N}{2} - \sum_{\substack{p\le N\\ f(p)=0}} \left(\frac{N}{2p} + 1\right) \ge \frac{N}{2} - \frac{N}{2} \sum_{\substack{p\le N\\ f(p)=0}} \frac{1}{p} - \sum_{p\le N} 1 \ge \frac{N}{2} - \frac{N}{4} - o(N) > \frac{N}{5}
\end{aligned}$$

if N is large enough. Thus we have

$$\delta(f,\mathcal{A},N) = \frac{S(f,\mathcal{A})}{S(f,\{1,2,\dots,N\})} = \frac{S(f,\mathcal{A})}{\sum\limits_{n=1}^{N} f(n)} > \frac{N/5}{\sum\limits_{n=1}^{N} 1} = \frac{1}{5}$$

whence the result follows.

10 Proof of Theorem 8

The crucial tool in the proof will be a variant of the statement of Conjecture 2'. Indeed, we will be able to prove a lemma which is weaker than Conjecture 2' in the sense that we need an additional assumption and we also lose a constant factor but, on the other hand, it controls the situation better when $|\mathcal{E}|$ is small.

If $\mathcal{S}$ is a finite set then we say that the subsets $\mathcal{R} \subset \mathcal{S}$ with $|\mathcal{R}| = \ell$ are at level ℓ. If $\mathcal{E}$ is a family of subsets of $\mathcal{S}$ which contains all the subsets of $\mathcal{S}$ at level ℓ, then we say that $\mathcal{E}$ is full at level ℓ. If $\ell < k$ then we say that the level k is higher than level ℓ.

Definition 4. If $\mathcal{E}$ is a non–empty family of subsets of a set $\mathcal{S}$, its highest full level is level ℓ (if there is no full level we put $\ell = 0$), and level k is the highest level which contains at least one subset belonging to $\mathcal{E}$, then $k - \ell$ is said to be the height of the family $\mathcal{E}$.

Lemma 7. *If $\mathcal{S}$ is a finite set with $|\mathcal{S}| = s$, $\mathcal{E}$ is a non–empty downset of subsets of $\mathcal{S}$, the highest full level of $\mathcal{E}$ is level ℓ, and the hight of $\mathcal{E}$ is H, then $\mathcal{E}$ contains an antichain $\mathcal{A}$ of length*

$$|\mathcal{A}| \ge \frac{1}{2}\frac{1}{\max\big(H, \varphi(s,\ell)\big)}|\mathcal{E}| \tag{10.1}$$

where $\varphi(s,\ell)$ is defined by

$$\varphi(s,\ell) = \frac{2^{s-1}}{\binom{s}{[s/2]}} \quad \text{for } \ell \ge \left[\frac{s}{2}\right] \tag{10.2}$$

and

$$\varphi(s,\ell)=\frac{\binom{s}{0}+\binom{s}{1}+\dots+\binom{s}{\ell}}{\binom{s}{\ell}}\quad \text{for } \ell<\left[\frac{s}{2}\right]. \tag{10.3}$$

Moreover, here we have

$$\varphi(s,\ell)<c_{21}s^{1/2} \text{ for } [s/2]-s^{1/2}<\ell\le s \tag{10.4}$$

and, writing $\Delta=[s/2]-\ell$,

$$\varphi(s,\ell)<c_{22}\frac{s}{\Delta} \text{ for } \ell\le[s/2]-s^{1/2}. \tag{10.5}$$

Finally, independently of ℓ, *(10.1) can be replaced by*

$$|\mathcal{A}|>c_{23}\frac{1}{\max\big(H,(\log|\mathcal{E}|)^{1/2},1\big)}|\mathcal{E}|. \tag{10.6}$$

(Note that we will need only (10.6), however, the sharper (10.1) also can be useful in some applications.)

Proof of Lemma 7. In order to prove (10.1), we have to distinguish two cases.

Case 1. Assume first that the levels $\ell+1,\ell+2,\dots,\ell+H$ in total contain at least $|\mathcal{E}|/2$ of the subsets in $\mathcal{E}$. Then by the pigeon hole principle, one of these levels contains at least $\frac{|\mathcal{E}|}{2H}$ subsets in $\mathcal{E}$; denote the family of these subsets by $\mathcal{A}$. Then clearly $\mathcal{A}\subset\mathcal{E}$, $\mathcal{A}$ is an antichain and

$$|\mathcal{A}|\ge\frac{|\mathcal{E}|}{2H} \tag{10.7}$$

which proves (10.1).

Case 2. Assume now that the levels $0,\dots,\ell$ in total contain at least $|\mathcal{E}|/2$ of the subsets in $\mathcal{E}$. It follows that

$$\binom{s}{0}+\binom{s}{1}+\dots+\binom{s}{\ell}>\frac{|\mathcal{E}|}{2}. \tag{10.8}$$

Now we choose $\mathcal{A}$ as the family of all the subsets of $\mathcal{S}$ at level $[s/2]$ if $\ell\ge[s/2]$, resp. at level ℓ if $\ell<[s/2]$. Then clearly $\mathcal{A}\subset\mathcal{E}$ and $\mathcal{A}$ is an antichain. It remains to estimate $|\mathcal{A}|$.

If $\ell\ge[s/2]$, then

$$|\mathcal{A}|=\binom{s}{[s/2]}\ge\binom{s}{[s/2]}\frac{|\mathcal{E}|}{2^s}=\frac{1}{2}\frac{\binom{s}{[s/2]}}{2^{s-1}}|\mathcal{E}|=\frac{1}{2}\frac{1}{\varphi(s,\ell)}|\mathcal{E}| \text{ (for } \ell\ge[s/2])$$

while for $\ell<[s/2]$, by (10.8), we have

$$|\mathcal{A}|=\binom{s}{\ell}>\binom{s}{\ell}\frac{|\mathcal{E}|}{2\left(\binom{s}{0}+\binom{s}{1}+\dots+\binom{s}{\ell}\right)}=\frac{1}{2}\frac{1}{\varphi(s,\ell)}|\mathcal{E}| \text{ (for } \ell<[s/2])$$

so that in both cases we have

$$|\mathcal{A}| > \frac{1}{2}\frac{1}{\varphi(s,\ell)}|\mathcal{E}| \tag{10.9}$$

which proves (10.1).

Next we will estimate $\varphi(s,\ell)$. If $[s/2]-s^{1/2} < \ell \le s$, then (10.4) follows trivially from (10.2), (10.3),

$$\binom{s}{[s/2]} \gg \frac{2^s}{s^{1/2}},$$

$$\binom{s}{0}+\binom{s}{1}+\dots+\binom{s}{\ell} \le 2^s$$

and

$$\binom{s}{\ell} \ge \binom{s}{[s/2]-[s^{1/2}]} \text{ (for } [s/2]-s^{1/2} < \ell \le [s/2]).$$

Assume now that

$$\Delta = [s/2]-\ell \ge s^{1/2}. \tag{10.10}$$

Write $\delta = s-2\ell-1$ so that, by (10.10),

$$\delta = s-2\big([s/2]-\Delta\big)-1 \ge 2\Delta-1 \ge \Delta \ge s^{1/2}. \tag{10.11}$$

By the inequality

$$1-x \le e^{-x},$$

for $i=1,2,\dots,\ell$ we have

$$\begin{aligned}
\binom{s}{\ell-i} &= \binom{s}{\ell-i+1}\frac{\ell-i+1}{s-\ell+i}\\
&= \binom{s}{\ell-i+1}\left(1-\frac{s-2\ell+2i-1}{s-\ell+i}\right) = \binom{s}{\ell-i+1}\left(1-\frac{\delta+2i}{s-(\ell-i)}\right)\\
&\le \binom{s}{\ell-i+1}\left(1-\frac{\delta+2i}{s}\right) \le \binom{s}{\ell-i+1}\exp\left(-\frac{\delta+2i}{s}\right).
\end{aligned}$$

It follows that, for $j=1,2,\dots,\ell$,

$$\begin{aligned}
\binom{s}{\ell-j} &= \binom{s}{\ell}\prod_{i=1}^{j}\binom{s}{\ell-i}\Big/\binom{s}{\ell-i+1}\\
&\le \binom{s}{\ell}\prod_{i=1}^{j}\exp\left(-\frac{\delta+2i}{s}\right) = \binom{s}{\ell}\exp\left(-\frac{\delta j}{s}-\frac{2}{s}\sum_{i=1}^{j} i\right)\\
&\le \binom{s}{\ell}\exp\left(-\frac{\delta j}{s}\right)
\end{aligned}$$

so that by (10.3) and (10.11) we have

$$\begin{aligned}
\varphi(s,\ell) &= \binom{s}{\ell}^{-1} \sum_{j=0}^{\ell} \binom{s}{\ell-j} \\
&\le \binom{s}{\ell}^{-1} \left(\binom{s}{\ell} + \sum_{j=1}^{\ell} \binom{s}{\ell} \exp\left(-\frac{\delta j}{s}\right)\right) \\
&= \sum_{j=0}^{\ell} \exp\left(-\frac{\delta j}{s}\right) < \sum_{j=0}^{\infty} \exp\left(-\frac{\delta j}{s}\right) \\
&= \frac{1}{1-\exp(-\delta/s)} \ll \frac{s}{\delta} \le \frac{s}{\Delta}
\end{aligned}$$

since uniformly for $0 < x < 1$ we have $1 - \exp(-x) \gg x$, and this proves (10.5).

Finally, in Case 1 (10.6) follows from (10.7). Since in Case 2 (10.9) holds, thus in this case to prove (10.6) it suffices to show that in this case $\varphi(s,\ell)$ in the denominator of (10.9) satisfies

$$\varphi(s,\ell) \ll \max\big((\log|\mathcal{E}|)^{1/2}, 1\big). \tag{10.12}$$

If $\ell \ge s/4$, then we have

$$\begin{aligned}
|\mathcal{E}| &\ge \binom{s}{0} + \binom{s}{1} + \cdots + \binom{s}{\ell} \ge \binom{s}{[s/4]} \\
&= \prod_{i=1}^{[s/4]} \frac{s-i+1}{i} > \prod_{i=1}^{[s/4]} \frac{s-s/4}{s/4} = 3^{[s/4]},
\end{aligned} \tag{10.13}$$

and, on the other hand, by (10.4) and (10.5) for all ℓ we have

$$\varphi(s,\ell) \ll s^{1/2}, \tag{10.14}$$

(10.12) follows from (10.13) and (10.14).

If $\ell < s/4$, then we have

$$\Delta = [s/2] - \ell \gg s$$

so that by (10.5) we have

$$\varphi(s,\ell) = O(1) \text{ (for } \ell < s/4) \tag{10.15}$$

and thus (10.12) holds trivially, which completes the proof of Lemma 7.

Now we may complete the proof of the theorem. We will use Lemma 7 with $\mathcal{S} = \mathcal{P}$ and with

$$\mathcal{E} = \left\{ \mathcal{R} : \mathcal{R} \subset \mathcal{P}, \prod_{p \in \mathcal{R}} p \le N \right\}.$$

Then clearly $\mathcal{E}$ is a downset, and it follows from the condition $\prod_{p\in\mathcal{R}} p \le N$ that

$$|\mathcal{E}| \le N. \tag{10.16}$$

Now we will estimate ℓ and H (both defined as in Lemma 7).

Assume that $\ell' \in \mathbb{N}$ and

$$\ell' \leq \left[\frac{\log N}{\log x}\right], \tag{10.17}$$

and consider a set $\mathcal{R} = \{p_1, p_2, \dots, p_{\ell'}\} \subset \mathcal{P}$ with $|\mathcal{R}| = \ell'$. Then by (2.25) and (10.17) we have

$$\prod_{p \in \mathcal{R}} p = p_1 p_2 \dots p_{\ell'} \leq x^{\ell'} \leq x^{\log N/\log x} = N$$

and thus $\mathcal{R} \subset \mathcal{E}$, so that $\mathcal{E}$ is full at level ℓ'. It follows that

$$\ell \geq \left[\frac{\log N}{\log x}\right]. \tag{10.18}$$

By the definition of ℓ and H, the family $\mathcal{E}$ is not empty at level $\ell + H$, i.e., there is a set $\mathcal{R}$ with

$$\mathcal{R} \in \mathcal{E} \tag{10.19}$$

and $|\mathcal{R}| = \ell + H$. Write $\mathcal{R} = \{p_1, p_2, \dots, p_{\ell+H}\}$. Then by (2.25) and (10.19) we have

$$N \geq p_1 p_2 \dots p_{\ell+H} > y^{\ell+H}$$

whence, by (2.24),

$$\ell + H < \frac{\log N}{\log y} \leq \log N\big((\log x)^{-1} + (\log N)^{-1/2}\big) = \frac{\log N}{\log x} + (\log N)^{1/2}. \tag{10.20}$$

It follows from (10.18) and (10.20) that

$$H = (\ell+H) - \ell < \left(\frac{\log N}{\log x} + (\log N)^{1/2}\right) - \left[\frac{\log N}{\log x}\right] \leq 1 + (\log N)^{1/2} < 2(\log N)^{1/2}. \tag{10.21}$$

By (10.6) in Lemma 7, (10.16) and (10.21), there is an antichain $\mathcal{A} \subset \mathcal{E}$ of length

$$|\mathcal{A}| \gg \frac{1}{\max(H, (\log|\mathcal{E}|)^{1/2}, 1)}|\mathcal{E}| \gg \frac{1}{\max((\log N)^{1/2}, (\log N)^{1/2}, 1)}|\mathcal{E}| = \frac{|\mathcal{E}|}{(\log N)^{1/2}}. \tag{10.22}$$

Now let $\mathcal{B}$ denote the set of the squarefree integers b with

$$\{p : p \text{ prime }, p \mid b\} \in \mathcal{A}.$$

Then we have

$$|\mathcal{B}| = |\mathcal{A}|, \tag{10.23}$$

$\mathcal{B} \subset \{1, 2, \dots, N\}$ by $\mathcal{A} \in \mathcal{E}$, and $\mathcal{B}$ is primitive since $\mathcal{A}$ is an antichain, so that we have $\mathcal{B} \subset \mathbb{P}_N$. It follows from (10.22) and (10.23) that

$$F(f, N) \geq \delta(f, \mathcal{B}, N) = \frac{S(f, \mathcal{B})}{S(f, \{1, 2, \dots, N\})} = \frac{|\mathcal{B}|}{|\mathcal{E}|} = \frac{|\mathcal{A}|}{|\mathcal{E}|} \gg \frac{1}{(\log N)^{1/2}}$$

which completes the proof of the theorem.

Note that assuming that there is just a weak lower bound for $|\mathcal{P}|$ we could make (10.15) effective, and then replacing the interval in (2.24) by a slightly shorter one, the lower bound for $F(f, N)$ could be improved considerably.

11 Proof of Theorem 9

For $k = 1, 2, \dots$, let p_k denote the smallest prime with $3^{3^k} < p_k$. Then by the prime number theorem we have

$$p_k = \big(1 + o(1)\big)3^{3^k}. \tag{11.1}$$

It follows that for $k > k_o$ we have

$$p_1 p_2 \dots p_{k-1} < p_k. \tag{11.2}$$

Write $\mathcal{P} = \{p_1, p_2, \dots\}$, and define the combinatorial weighting f by

$$f(p) = \begin{cases} 1 & \text{if } p \in \mathcal{P} \\ 0 & \text{if } p \notin \mathcal{P}. \end{cases}$$

Then $f(n) = 1$ if and only if n is of the form (7.3), and for every other $n \in \mathbb{N}$ we have $f(n) = 0$.

Now fix some $N \in \mathbb{N}$ with $N \geq p_1$ (the case $3 \leq N < p_1$ is trivial), and define the positive integer K by

$$p_K \leq N < p_{K+1} \tag{11.3}$$

so that, by (11.1), we have

$$K = \left(\frac{1}{\log 3} + o(1)\right) \log\log N. \tag{11.4}$$

It follows from (7.3), (11.2) and (11.3) that

$$S\big(f, \{1, 2, \dots, N\}\big) = \sum_{n=1}^{N} f(n) = |\{n : n \leq N, f(n) = 1\}| \geq |\{n : n \mid p_1 \dots p_{K-1}\}| = 2^{K-1}. \tag{11.5}$$

Now consider a set $\mathcal{A} \in \mathbb{P}_N$, and let

$$\mathcal{A}^* = \{a : a \in \mathcal{A}, f(a) = 1\}.$$

Then every $a \in \mathcal{A}^*$ is of the form (7.3) and, indeed, writing $\mathcal{P}_K = \{p_1, p_2, \dots, p_K\}$, by (11.3) for each of these a's we have

$$\mathcal{P}(a) \subset \mathcal{P}_K.$$

$\mathcal{A}^*$ is a primitive set, thus for $a_1 \in \mathcal{A}^*$, $a_2 \in \mathcal{A}^*$, $a_1 \neq a_2$ we cannot have

$$\mathcal{P}(a_1) \subset \mathcal{P}(a_2).$$

Thus by Lemma 5 (Sperner's theorem) we have

$$|\mathcal{A}^*| \leq \binom{|\mathcal{P}_K|}{[|\mathcal{P}_K|/2]} = \binom{K}{[K/2]} < c_{24}\frac{2^K}{K^{1/2}}. \tag{11.6}$$

It follows from (11.4), (11.5) and (11.6) that for all $\mathcal{A} \in \mathbb{P}_N$ we have

$$\delta(f, \mathcal{A}, N) = \frac{S(f, \mathcal{A})}{S(f, \{1, 2, \dots, N\})} = \frac{S(f, \mathcal{A}^*)}{S(f, \{1, 2, \dots, N\})} = \frac{|\mathcal{A}^*|}{S(f, \{1, 2, \dots, N\})} < 2c_{24}K^{-1/2}$$

which, by (11.4), completes the proof of (2.26).

12 Proof of Theorem 10

Write $\mathcal{P} = \{p_1, p_2, \dots\} = \{p : p \text{ prime }, f(p) = 1\}$ with $p_1 < p_2 < \dots$.
Clearly, we may assume that $\mathcal{P}$ is infinite. First we will show that there are infinitely many $k \in \mathbb{N}$ with

$$p_1 p_2 \dots p_k < p_{k+1}. \tag{12.1}$$

We will prove this by contradiction: assume that there is a $k_o (\geq 1)$ so that

$$p_1 p_2 \dots p_k > p_{k+1} \text{ for } k \geq k_o. \tag{12.2}$$

Write $p_1 p_2 \dots p_{k_o} = U$. It follows from (12.2) by induction that

$$p_1 p_2 \dots p_{k_o+i} \geq U^{2^i} \text{ for } i = 0, 1, \dots \ . \tag{12.3}$$

Consider a large $i \in \mathbb{N}$, and write $N = N_0^{2^i}$ so that $i = \left(\frac{1}{\log 2} + o(1)\right) \log\log N$ (as $i \to \infty$). Then by (12.3) for large i we have

$$|\mathcal{P} \cap (0, N)| \geq k_o + i \geq \left(\frac{1}{\log 2} + o(1)\right) \log\log N > \log\log N$$

which contradicts (2.28), and this proves that there are infinitely many k satisfying (12.1).
Now consider a large k satisfying (12.1), and write $N = p_{k+1} - 1$ so that by (2.28) we have

$$k = |\mathcal{P} \cap (0, N)]| < \log\log N. \tag{12.4}$$

Write $\mathcal{A} = \{a : a | p_1 \dots p_k, \omega(a) = [k/2]\}$. Then by (12.1) we have $\mathcal{A} \subset \{1, 2, \dots, N\}$ and clearly $\mathcal{A} \in \mathbb{P}$ so that $\mathcal{A} \in \mathbb{P}_N$. Moreover, we have

$$\delta(f, \mathcal{A}, N) = \frac{S(f, \mathcal{A})}{S(f, \{1, 2, \dots, N\})} = \frac{\binom{k}{[k/2]}}{2^k} \gg k^{-1/2}$$

whence, by (12.4), the result follows.

13 Proof of Theorem 11

Write $H(N) = S\big(f, \{1, 2, \dots, N\}\big) = |\{n : n \le N, f(n) = 1\}|$. First we will show that there are infinitely many $M \in \mathbb{N}$ with

$$H(2M) > \left(1 + \frac{C}{3}(\log\log M)^{-1/2}\right) H(M). \tag{13.1}$$

We will prove this by contradiction: assume that there is an M_o so that

$$H(2M) \le \left(1 + \frac{C}{3}(\log\log M)^{-1/2}\right) H(M) \text{ for all } M \ge M_o;$$

we may assume that $H(M_o) \ge 1$. Then for large enough $k \in \mathbb{N}$ we have

$$\begin{aligned} H(2^k M_o) &= H(M_o) \prod_{i=1}^{k} \frac{H(2^i M_o)}{H(2^{i-1} M_o)} \\ &\le H(M_o) \prod_{i=1}^{k} \left(1 + \frac{C}{3}(\log\log 2^{i-1} M_o)^{-1/2}\right) \\ &< H(M_o) \exp\left(\sum_{i=1}^{k} \frac{C}{3}(\log\log 2^{i-1} M_o)^{-1/2}\right) \\ &< H(M_o) \exp\left(\frac{C}{2} k (\log k)^{-1/2}\right). \end{aligned} \tag{13.2}$$

On the other hand, by (2.29) for large k we have

$$\begin{aligned} &H(2^k M_o) = |\{n : n \le 2^k M_o, f(n) = 1\}| \ge |\{p : p \text{ prime }, p \le 2^k M_o, f(p) = 1\}| \\ &> (2^k M_o)^{C(\log\log 2^k M_o)^{-1/2}} = \exp\big(C(\log 2^k M_o)(\log\log 2^k M_o)^{-1/2}\big) \\ &> \exp\left(\frac{C}{2} k (\log k)^{-1/2}\right) \end{aligned}$$

which contradicts (13.2), and this shows that there are infinitely many M satisfying (13.1).

Now consider a large M satisfying (13.1), write $N = 2M$, and let $\mathcal{A} = \{n : M < n \le 2M, f(n) = 1\}$. Then clearly we have $\mathcal{A} \subset \mathbb{P}_N$, and by (13.1), for large M we have

$$\begin{aligned} \delta(f, \mathcal{A}, N) &= \frac{S(f, \mathcal{A})}{S(f, \{1, 2, \dots, N\})} = \frac{H(2M) - H(M)}{H(2M)} \\ &= 1 - \frac{H(M)}{H(2M)} > 1 - \left(1 + \frac{C}{3}(\log\log M^{-1/2}\right)^{-1} \\ &> 1 - \left(1 - \frac{C}{6}(\log\log N)^{-1/2}\right) = \frac{C}{6}(\log\log N)^{-1/2} \end{aligned}$$

so that (2.27) holds infinitely often with $c = \frac{C}{6}$ which completes the proof of Theoem 11.

14 Remarks

1. We remark first that Lemma 7 can be extended to the case when the elements of the sets $\mathcal{S}$ are weighted (here we did not need this generality). Indeed, let $\mathcal{S}$ be a finite set, and to each $n \in \mathcal{S}$ assign a positive number $\gamma(n)$. For $\mathcal{R} \subset \mathcal{S}$ write $\gamma(\mathcal{R}) = \sum_{r\in\mathcal{R}} \gamma(r)$, and if $\mathcal{E}$ is a family of subsets of $\mathcal{S}$, then write $\gamma(\mathcal{E}) = \sum_{\mathcal{R}\in\mathcal{E}} \gamma(\mathcal{R})$. By a slight modification of the proof of Lemma 7 one can prove:

   Lemma 7'. *If $\mathcal{S}$ is a finite set with a weight function γ as described above, $|\mathcal{S}| = s$, $\mathcal{E}$ is a non–empty downset of subsets of $\mathcal{S}$, the highest full level of $\mathcal{E}$ is level ℓ, and the height of $\mathcal{E}$ is H, then $\mathcal{E}$ contains an antichain $\mathcal{A}$ of weight*

$$\gamma(\mathcal{A}) \geq \frac{1}{2}\frac{1}{\max(H, \varphi(s,\ell))}\gamma(\mathcal{E})$$

   where $\varphi(s,\ell)$ is the same function as in Lemma 7.

2. In an earlier paper [2] we extended the study of divisibility properties to prefix free sets. Most of the problems and methods studied above could be adopted in the prefix free case; we leave the details to the reader. Here we will discuss only one related question. Namely, the proof of Theorem 6, (ii) was based on the assumption that the combinatorial Conjecture 2' is true. This conjecture has only recently been given more attention by combinatorialists, but may well be hard to prove. On the other hand, we can settle that analogue of this problem which is needed in the prefix free situation.

   Let $p(n)$ and $P(n)$ denote the smallest and greatest prime factor of n, respectively, and let $P^+(n)$ denote the smallest prime greater than $P(n)$.

   Recall that for $a, b \in \mathbb{N}^*$ (square free integers) with the properties $a|b$ and $p\left(\frac{b}{a}\right) > P(a)$, i.e. they are of the form $a = p_1 \dots p_r$, $b = p_1 \dots p_r p_{r+1} \dots p_t$ where $p_1 < p_2 < \dots < p_r < p_{r+1} < \dots < p_t$ are distinct primes (with $t > r$), we said in [2] that a is prefix of b and we wrote $a|_p b$.

   If $\mathcal{A} \subset \mathbb{N}^*$ is set such that there are no $a \in \mathcal{A}$, $b \in \mathcal{A}$ with $a|_p b$, then $\mathcal{A}$ is said to be prefix–free. Theorem 1 of [2] states that $\mathcal{B}_N = \left\{b : b \in \mathbb{N}^*, b\, P^+(b) > N\right\}$ is the largest prefix–free subset of $\mathbb{N}^*$.

   (It is also shown in [2] that $\lim_{N\to\infty} \frac{|B_N|}{|\mathbb{N}^*_N|} = 1$.) This corresponds to the combinatorial weighting with value 1 on all primes. Now for any combinatorial weighting f define

   $\mathcal{B}_N(f) = \{b : b \in \mathbb{N}^*_N$ with weight 1 such that $P(b)$ is the largest prime of weight 1 or $b\, P^+(f,b) > N\}$, where $P^+(f,b)$ is the smallest prime of weight 1 bigger than $P(b)$.

   Inspection of the proof of Theorem 1 shows that also in this generality $\mathcal{B}_N(f)$ is the largest prefix–free subset of $\mathbb{N}^*_N(f) = \{n : n \leq N, f(n) = 1\}$. Actually $\mathcal{B}_N(f)$ is the set of its maximal elements.

   Now the prefix relation makes $\mathbb{N}^*_N(f)$ to a partially ordered set of a simple tree structure. In it for any upset $\mathcal{U} \subset \mathbb{N}^*_N(f)$ a largest antichain $\mathcal{A} \subset \mathcal{U}$ satisfies $|\mathcal{A}| \geq \frac{1}{2}|\mathcal{U}|$. Indeed, let p be the smallest prime with $\mathcal{U}_p = \{u \in \mathcal{U} : p|u\} \neq \emptyset$ and set $\mathcal{U}_{\bar{p}} = \mathcal{U} \setminus \mathcal{U}_p$, then by induction hypothesis there are antichains $\mathcal{A}_p \subset \mathcal{U}_p$, $\mathcal{A}_{\bar{p}} \subset \mathcal{U}_{\bar{p}}$ with $|\mathcal{A}_p| \geq \frac{1}{2}|\mathcal{U}_p|$, $|\mathcal{A}_{\bar{p}}| \geq \frac{1}{2}|\mathcal{U}_{\bar{p}}|$ and therefore $\mathcal{A} = \mathcal{A}_p \cup \mathcal{A}_{\bar{p}} \subset \mathcal{U}$ satisfies $|\mathcal{A}| \geq \frac{1}{2}|\mathcal{U}|$.

It is readily shown that in general $\frac{1}{2}$ is the largest lower bound for $\frac{|\mathcal{B}_N(f)|}{|\mathbb{N}_N^*(f)|}$. For instance if f takes the value 1 on the primes $p_1, p_2, \dots, p_t$ and $N \geq \prod_{i=1}^{t} p_i$ then

$$\mathcal{B}_N(f) = \left\{p_1^{\varepsilon_1} p_2^{\varepsilon_2} \dots p_{t-1}^{\varepsilon_{t-1}} p_t : \varepsilon_i \in \{0,1\}\right\} \text{ and } |\mathcal{B}_N(f)| = 2^{t-1}, |\mathbb{N}_N^*(f)| = 2^t.$$

Similarily optimal suffix–free subsets are constructed by choosing all numbers of $\mathbb{N}_N^*(f)$ divisible by p_1.

References

[1] R. Ahlswede and L. Khachatrian, Classical results on primitive and recent results on cross–primitive sequences, in: The Mathematics of Paul Erdős, Vol. I, eds. R.L. Graham et al., Algorithms and Combinatorics 13, Springer–Verlag, Berlin, 104–116, 1997.

[2] R. Ahlswede, L. Khachatrian and A. Sárközy, On prefix–free and suffix–free sequences of integers, in: Numbers, Information and Complexity, eds. I. Althöfer et al., Kluwer Academic Publishers, Boston, 1–16, 2000.

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

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

[5] P. Erdős, On the integers having exactly k prime factors, Annals of Math. 49, 53–66, 1948.

[6] P. Erdős, On the distribution function of additive functions, Ann. of Math. 47, 1–20, 1946.

[7] P. Erdős, A. Sárközy and E. Szemerédi, On an extremal problem concerning primitive sequences, J. London Math. Soc. 42, 484–488, 1967.

[8] P. Erdős, A. Sárközy and E. Szemerédi, On divisibility properties of sequences of integers, Coll. Math. Soc. J. Bolyai 2, 35–49, 1970.

[9] P. Frankl, Old and new problems on finite sets, Congressus Numerantium 67, 243–256, 1988.

[10] G.H. Hardy and S. Ramanuyan, The normal number of prime factors of a number n, Quart. J. Math. 48, 76–92, 1917.

[11] H. Halberstam and K.F. Roth, Sequences, Springer–Verlag, Berlin–Heidelberg–New York, 1983.

[12] J. Kubilius, Probabilistic methods in the theory of numbers, Transl. Math. Monographs, Vol. 11, Amer. Math. Soc., Providence, Rhode Island, 1964.

[13] S. Pillai, On numbers which are not multiples of any other in the set, Proc. Indian Acad. Sci. A 10, 392–394, 1939.

[14] K. Prachar, Primzahlverteilung, Springer–Verlag, Berlin–Göttingen–Heidelberg, 1957.

[15] A. Sárközy, On divisibility properties of sequences of integers, in: The Mathematics of Paul Erdős, Vol. I, eds. R.L. Graham et al., Algorithms and Combinatorics 13, Springer–Verlag, Berlin, 241–250, 1997.

[16] E. Sperner, Ein Satz über Untermengen einer endlichen Menge, Math. Z. 27, 544–548, 1928.

[17] P. Turán, On a theorem of Hardy and Ramanujan, J. London Math. Soc. 9, 274–276, 1934.