JOURNAL OF NUMBER THEORY 12, 385–394 (1980)

On the Small Sieve. I. Sifting by Primes

P. ERDŐS AND I. Z. RUZSA

Mathematical Institute of the Hungarian Academy of Sciences, Budapest, Reáltanoda u. 13-15, Hungary H-1053

Communicated by H. Zassenhaus

Received September 29, 1978

The main object of the paper is to prove that if P is a set of primes with sum of reciprocals $\leqslant K$, then the number of natural numbers up to x, divisible by no element of P, is $\geqslant cx$, where c is a positive constant depending only on K. A lower estimate is given for c and a similar result is achieved in the case when the condition of primality is substituted by the weaker condition that any m elements of the sifting set are coprime.

1. INTRODUCTION

For a set A of natural numbers let $F(x, A)$ denote the number of natural numbers $n \leqslant x$ divisible by no element of A. Let

$$G(x, K) = \min F(x, P), \tag{1.1}$$

where P runs over all sets of primes satisfying

$$\sum_{p \in P} 1/p \leqslant K. \tag{1.2}$$

Our main aim is to prove that

$$G(x, K) > cx \tag{1.3}$$

with a positive constant c depending only on K.

At first sight this may seem obvious ("easy to see," the first-named author wrote [3]), but it is not. The sieves of Brun and Selberg give this result only if the sifting primes all lie below x^a, $a < 1$. The reason is that these sieves give a main term, which is the expected number of unsifted elements, and a remainder term. In our case the expectation is

$$x \prod_{p \in P} (1 - 1/p) \succ xe^{-K},$$

0022-314X/80/030385–10\$02.00/0

but the real order is much smaller. If we choose the largest primes up to x whose sum of reciprocals does not exceed K (roughly speaking, the interval $(x^{e^{-K}}, x)$), then (see de Bruijn [2]), the number of unsifted elements is

$$\approx xe^{-Ke^K};$$

this fact makes our problem nonstandard.

PROBLEM 1 (cf. Erdös [3]). Is $G(x, K)$ asymptotically given by the primes in $(x^{e^{-K}}, x)$?

The most we can achieve in this direction is

THEOREM 1. *We have*

$$G(x, K) \geqslant e^{-e^{cK}} \tag{1.4}$$

with a positive absolute constant c.

PROBLEM 2. What happens if we sift by other residue classes? Suppose $p_1, ..., p_k \leqslant x$ are primes with sum of reciprocals $\leqslant K$ and to each p_i corresponds a residue class $a_i \pmod{p_i}$. Is it true that the number of natural numbers $n \leqslant x$ satisfying $n \not\equiv a_i \pmod{p_i}$ for all i is at least cx, $c = c(K) > 0$?

Another surprising feature is that we cannot omit the condition that the elements of P be primes. Put

$$H(x, K) = \min F(x, A),$$

where A is subject to the conditions

$$\sum_{a \in A} 1/a \leqslant K, \qquad 1 \notin A. \tag{1.5}$$

In the second part of the paper we shall show that

$$H(x, K) < x^{\varepsilon}, \qquad K > K_0(\varepsilon);$$

more exactly, that

$$\lim_{x \to \infty} \frac{\log H(x, K)}{\log x} = e^{1-K} \quad (K \geqslant 1).$$

$H(x, 1) = \sigma(x)$ has been shown by Schinzel and Szekeres [8] (not stated explicitly).

The case when A is fixed and x tends to infinity is considerably different; we have

$$\Delta(A) = \lim_{x \to \infty} \frac{F(x, A)}{x} \geqslant \prod_{a \in A} (1 - 1/a).$$

This inequality is due to Heilbronn [5] and Rohrbach [6]; cf. also Behrend [1], Halberstam and Roth [4, Chap. V, Sect. 6] and Ruzsa [7].

A similar estimate holds under the weaker condition that $a < x^{1-\delta}$ for $a \in A$.

THEOREM 2. *If*

$$A \subset [2, x^{1-\delta}], \qquad \sum_{a \in A} 1/a \leqslant K, \tag{1.7}$$

then

$$F(x, A) \geqslant c_1 \delta e^{-K} x \tag{1.8}$$

with an absolute constant c_1.

Though the condition of primality cannot be dropped in Theorem 1, it can be weakened to some extent. Let

$$H_m(x, K) = \min F(x, A), \tag{1.9}$$

where A is subject to (1.5) and any m of its elements are coprime.

THEOREM 3. *We have*

$$H_m(x, K) \leqslant cx, \qquad c = c(m, K) > 0. \tag{1.10}$$

The proof actually gives

$$H_m(x, K) \geqslant c_2 e^{-K} G(x, K) \tag{1.11}$$

for $x > x_0(m, K)$; with a slight modification we can even prove

$$H_m(x, K) \geqslant G(x, K) - \varepsilon x, \qquad x > x_0(\varepsilon, m, K). \tag{1.12}$$

COROLLARY. *If P is a set of primes satisfying* (1.2), *then the number of squarefree integers up to x which are divisible by no element of P is $\geqslant cx$, $c = c(K) > 0$.*

This is obtained by applying Theorem 3 to the set

$$A = P \cup \{q^2 : q \text{ is prime}, q \notin P\}.$$

2. Proof of Theorem 2

Let B denote the set of natural numbers divisible by no element of A.

Lemma 2.1. *For all y we have*

$$\sum_{\substack{b \leqslant y \\ b \in B}} 1/b \geqslant \prod_{a \in A} (1 - 1/a) \log(y+1). \tag{2.2}$$

Proof. Every number has (one or more) decompositions of the form

$$a_1^{\alpha_1} \cdots a_k^{\alpha_k} b, \qquad b \in B, \quad a_i \in A.$$

Hence

$$\sum_{n \leqslant y}' 1/n \leqslant \sum_{\substack{b \leqslant y \\ b \in B}} 1/b \prod_{a \in A} (1 + a^{-1} + a^{-2} + \cdots),$$

which immediately yields (2.2).

Note. As a by-product, this gives a proof for the Heilbron–Rohbach inequality (1.6).

Proof of Theorem 2. Consider the numbers

$$bp \leqslant x, \quad p > x^{1-\delta}, \qquad b \in B, \quad p \text{ prime}. \tag{2.3}$$

We may assume $\delta < \frac{1}{2}$ and then these numbers are different. They all belong to B: if $a | bp$, then either $a | b$, or $p | a$; the first contradicts the definition of B, the second contradicts $a \leqslant x^{1-\delta} < p$. Therefore

$$F(x, A) \geqslant \sum_{\substack{bp \leqslant x \\ p > x^{1-\delta} \\ b \in B}} 1 = \sum_{\substack{b \in B \\ b \leqslant x^{\delta}}} (\pi(x/b) - \pi(x^{1-\delta})). \tag{2.4}$$

By the prime number theorem we have

$$\pi(x/b) - \pi(x^{1-\delta}) \geqslant c_3 x/(b \log x)$$

if $b \leqslant y = x^{\delta}/2$, so (2.4) yields

$$\begin{aligned} F(x, A) &\geqslant \frac{c_3 x}{\log x} \sum_{b \leqslant y} 1/b \\ &\geqslant \frac{c_3 x \log(y+1)}{\log x} \prod_{a \in A} (1 - 1/a) \end{aligned} \tag{2.5}$$

according to Lemma 2.1. Obviously $\log(y+1) \geqslant (\delta/2)\log x$ and

$$\prod_{a\in A}(1-1/a) \geqslant c_4 \exp\left(-\sum_{a\in A} 1/a\right),$$

so (2.5) gives (1.8) with $c_1 = c_3 c_4/2$.

3. Proof of Theorem 1

Let

$$\gamma(K) = \inf_x \frac{G(x,K)}{x};$$

our aim is to show

$$\gamma(K) > e^{-e^{cK}} \tag{3.1}$$

with a suitable constant c. We shall use a real-type induction, that is, we shall deduce (3.1) supposing it to hold for $K-h$, where h will be a positive number, depending on K explicitly and monotonically decreasing.

Evidently

$$F(x,P) \geqslant x - \sum_{p\in P}[x/p] \geqslant x(1-K);$$

hence

$$\gamma(K) \geqslant 1-K,$$

which proves (3.1) for $K \leqslant \frac{1}{2}$.

We are going to estimate $F(x,P)$ for a set P satisfying (1.2). As $F(x,P) \geqslant 1$,

$$G(x,K) > e^{-e^{cK}} \qquad (x < e^{e^{cK}})$$

is obvious, thus we may assume

$$x \geqslant e^{e^{cK}}. \tag{3.2}$$

Put $k = e^{K+2}$ and let Q be the set of primes lying in

$$[x^{1/k}, x]\backslash P.$$

Let B denote the set of numbers divisible by no prime from P. If $q \in Q$ and

$b \in B$, then $n = qb \in B$; as $q \geqslant x^{1/k}$, a number $n \leqslant x$ may have at most k divisors from Q, so it has at most k representations of this form. Hence we have

$$F(x, P) \geqslant \frac{1}{k} \sum_{q \in Q} F(x/q, P). \tag{3.3}$$

Let

$$\alpha = \sum_{\substack{p \in P \\ p > x^{1-1/k}}} 1/p.$$

Since $x/q \leqslant x^{1-1/k}$ for $q \in Q$, we have

$$F(x/q, P) \geqslant (x/q)\, \gamma(K - \alpha),$$

so that (3.3) yields

$$F(x, P) \geqslant e^{-K-2} \gamma(K - \alpha) x \sum_{q \in Q} 1/q. \tag{3.4}$$

By (3.2) we have

$$\sum_{q \in Q} 1/q \geqslant \sum_{p \in [x^{1/k}, x]} 1/p - K \geqslant 1$$

for c large enough, whence (3.4) gives

$$F(x, P) \geqslant e^{-K-2} \gamma(K - \alpha) x. \tag{3.5}$$

This inequality will be sufficient if α is not too small, and otherwise we may apply Theorem 2. To see this, set

$$P^* = P \cap [2, x^{1-1/k}];$$

we have evidently

$$F(x, P) \geqslant F(x, P^*) - \alpha x$$

and

$$F(x, P^*) \geqslant c_5 e^{-2K} x \qquad (c_5 = c_1 e^{-2})$$

by Theorem 2. Therefore, with $c_6 = c_5/2$ we have

$$F(x, P) \geqslant c_6 e^{-2K} x \quad \text{ if } \quad \alpha \leqslant c_6 e^{-2K}. \tag{3.6}$$

If this is not the case, (3.5) yields

$$F(x, P) \geqslant e^{-K-2}\gamma(K - c_6 e^{-2K})x. \tag{3.7}$$

Taking the minimum over the sets P we get

$$G(x, K) \geqslant \min(c_6 e^{-2K}, e^{-K-2}\gamma(K - c_6 e^{-2K}))x \tag{3.8}$$

if x satisfies (3.2).

An easy calculation yields

$$c_6 e^{-2K} > e^{-e^{cK}}$$

and

$$e^{-K-2}\exp(-\exp c(K - c_6 e^{-2K})) > e^{-e^{cK}}$$

if $K > \frac{1}{2}$ and c is large enough; this completes the proof.

4. Proof of Theorem 3

We do not actually need the condition that any m elements of A be relatively prime; what we shall use is the fact that the composite elements of A grow rapidly. Theorem 3 follows from the next two lemmas.

Lemma 4.1. *Let (w_j), $w_j > 0$, be a fixed sequence satisfying*

$$\sum 1/w_j < \infty.$$

Suppose A is a set of natural numbers, not containing 1, such that $A = P \cup A_1$, where $A_1 = \{a_1, a_2, \ldots\}$, $a_j > w_j$, P consists of primes and

$$\sum_{a \in A} 1/a \leqslant K.$$

Then we have

$$F(x, A) > cx,$$

where c depends on K and the sequence (w_j).

Lemma 4.2. *If $a_1 < a_2 < \cdots$ are composite numbers, any m of which are relatively prime, then we have*

$$a_j > j^2/(m-1)^2.$$

Proof. Let r_j be the smallest prime divisor of a_j. Since a prime can occur at most $(m-1)$ times among the r_j's, we have $r_i > j/(m-1)$ for some $i \leqslant j$. Hence

$$a_j \geqslant a_i \geqslant r_i^2 > j^2/(m-1)^2.$$

To prove Lemma 4.1 we need some preparation.

LEMMA 4.3. *Let P be a set of primes satisfying* (1.2) *and $F(x) = F(x, P)$. Uniformly for $c \in [0, 1]$ we have*

$$F(cx) = cF(x) + O(e^K x/\log x).$$

Proof. Let D be the set of numbers composed exclusively of the primes of P. We have

$$F(x) = \sum_{d \in D} \mu(d)[x/d].$$

Hence

$$|F(cx) - cF(x)| = \left| \sum_{d \in D} \mu(d) \left(\left[\frac{cx}{d}\right] - \left[\frac{x}{d}\right] \right) \right|$$

$$\leqslant \sum_{d \in D, d \leqslant x} 1 = O(e^K x/\log x).$$

Here the last inequality follows easily by Selberg's sieve.

LEMMA 4.4. *Let A be a set of k natural numbers and P a set of primes satisfying* (1.2). *Suppose that no element of A is divisible by any prime of P. Then we have, with $\Delta(A)$ as defined in* (1.6),

$$F(x, P \cup A) = \Delta(A)\, F(x, P) + O(2^k e^K x/\log x).$$

Proof. Again write $F(x; P) = F(x)$. By the sieve formula

$$F(x, P \cup A) = F(x) - \sum_{a \in A} F(x/a) + \sum_{\substack{a_1 < a_2 \\ a_1, a_2 \in A}} F(x/[a_1, a_2]) - \cdots.$$

Lemma 4.3 yields

$$F(x, P \cup A) = F(x) \left(1 - \sum \frac{1}{a} + \sum \frac{1}{[a_1, a_2]} - \cdots \right)$$
$$+ O(2^k e^K x/\log x).$$

The coefficient of $F(x)$ is just $\Delta(x)$, again by the sieve formula.

Proof of Lemma 4.1. Let $A_1 = A_2 \cup A_3$, $A_2 = \{a_1, \ldots, a_k\}$, $A_3 = \{a_{k+1}, \ldots\}$, $k = [\log\log x]$. Evidently

$$\sum_{a \in A_3} 1/a < \sum_{j > \log\log x} 1/w_j \to 0,$$

hence

$$F(x, A) \geqslant F(x, P \cup A_2) - \sum_{a \in A_3} [x/a] = F(x, P \cup A_2) + O(x). \tag{4.5}$$

We may assume that the elements of A_1 are not divisble by any prime from P, since any that are divisible may be dropped without influencing $F(x, A)$, and then Lemma 4.4 yields

$$\begin{aligned} F(x, P \cup A_2) &= \Delta(A_2)\, F(x, P) + O(2^k e^K x/\log x) \\ &= \Delta(A_2)\, F(x, P) + O(x). \end{aligned} \tag{4.6}$$

Now we have

$$\Delta(A_2) \geqslant \prod_{a \in A_2} (1 - 1/a) > c_1 e^{-K} \tag{4.7}$$

by the Heilbronn–Rohrbach inequality (1.6) and

$$F(x, P) > c_2 x, \qquad c_2 = c_2(K), \tag{4.8}$$

by Theorem 1. Formulas (4.5)–(4.8) give Lemma 4.1 for $x > x_0(K)$; for small x we may use the trivial estimate $F(x, A) \geqslant 1$.

To deduce Theorem 3 let A be a set, any m of whose elements are coprime and let $a_1 < a_2 < \cdots$ be its composite elements. Lemma 4.2 implies

$$a_j > w_j = j^2/(m-1)$$

and now Lemma 4.1 yields (1.10) since

$$\sum 1/w_j = (m-1) \sum j^{-2} < \infty$$

obviously holds.

References

1. F. A. Behrend, Generalization of an inequality of Heilbronn and Rohrbach, *Bull. Amer. Math. Soc.* **54** (1948), 681–684.
2. N. G. de Bruijn, On the number of positive integers $\leqslant x$ and free of prime factors $\geqslant y$, *Indag. Math.* **13** (1951), 50–60.

3. P. Erdös, Problem 2, *in* "Number Theory," Coll. Math. Soc. J. Bolyai, Vol. 2, p. 232, 1968.
4. H. Halberstam and K. F. Roth, "Sequences," Vol. I, Clarendon, Oxford, 1966.
5. H. A. Heilbronn, On an inequality in the elementary theory of numbers, *Proc. Cambridge Philos. Soc.* **33** (1937), 207–209.
6. H. Rohrbach, Beweis einer zahlentheoretischen Ungleichung, *J. Reine Angew. Math.* **177** (1937), 153–156.
7. I. Z. Ruzsa, Probabilistic generalization of a number-theoretical inequality, *Amer. Math. Monthly* **83** (1976), 723–725.
8. A. Schinzel and G. Szekeres, Sur un problème de M. Paul Erdös, *Acta Sci. Math. (Szeged)* **20** (1959), 221–229.