Some problems and results in elementary number theory.

By P. ERDŐS in Aberdeen (Scotland).

Throughout this paper $c_1, c_2, \ldots$ denote absolute constants, $p_1, p_2, \ldots$ are primes, $P_1, P_2, \ldots$ is the sequence of all primes, letters u, x, y, n, a, b, etc. denote positive integers.

Denote by $u_1 = 1 < u_2 < \ldots$ the sequence of integers of the form $x^2 + y^2$. CHOWLA and BAMBAH remarked that $u_{i+1} - u_i < c_1 u_i^{1/4}$. It has been conjectured that $u_{i+1} - u_i = o(u_i^{1/4})$. In other words for any $\varepsilon > 0$ an sufficiently large n there always is an integer of the form $x^2 + y^2$ in the interval $(n, n + \varepsilon n^{1/4})$. This conjencture is still unproved. This is rather surprising since the proof of BAMBAH and CHOWLA is immediate. It suffices to put $u_i = x^2 + y^2$ where

$$x^2 < u_{i+1} \leqq (x+1)^2, \quad y^2 < u_{i+1} - x^2 \leqq (y+1)^2.$$

TURÁN (in a letter) observed that it is easy to prove that

$$u_{i+1} - u_i > c_2 \frac{\log u_i}{\log \log u_i} \tag{1}$$

for inifinitely many i. He asked whether I can improve (1).

In fact I can show that for infinitely many i

$$u_{i+1} - u_i > c_2 \frac{\log u_i}{(\log \log u_i)^{1/2}}. \tag{2}$$

In fact I prove a somewhat more general theorem. Let $p_1 < p_2 < \ldots$ be a sequence of primes satisfying

$$\sum_{p_i < x} \frac{1}{p_i} = f(x) \to \infty \quad \text{as} \quad x \to \infty,$$

in other words $\sum \frac{1}{p_i} = \infty$. Denote by $v_1 < v_2 < \ldots$ the integers which either are not divisible by p_i or are divisible by p_i^2. Then we prove

Theorem 1. *For infinitely many i we have*

$$v_{i+1} - v_i > c_4 e^{f(\log v_i)} \frac{\log v_i}{\log \log v_i}. \tag{3}$$

Before proving Theorem 1 we show that Theorem 1 implies (2). If the p_i are the primes $\equiv 3 \pmod 4$, then as is well known the u's are included among the v's, also it is well known that

$$\sum_{\substack{p \equiv 3 \pmod 4 \\ p \leqq x}} \frac{1}{p} = \frac{1}{2} \log \log x + O(1).$$

Thus from $e^z > 1 + z$,

$$e^{f(\log v_i)} > c_5 (\log \log v_i)^{1/2},$$

or (3) implies (2).

It seems rather difficult to prove a stronger result than Theorem 1. In fact if $w_1, w_2, \ldots$ denote the integers not divisible by any p_i, the w's are clearly included among the v's; but nevertheless I cannot prove any stronger result than (3) for the w's.

Before proving Theorem 1 we state a few well known results on primes which we will need. Let denote by $D_l(y)$ the number of integers less than y which are not divisible by $p_1, p_2, \ldots, p_l$. It is well known [and follows from BRUN's method [1])] that there exist two absolute constants c_6 and c_7 so that

$$D_l(y) < c_6 y \prod_{p_i < y^{c_7}} \left(1 - \frac{1}{p_i}\right). \tag{4}$$

Now [with $\exp z = e^z$]

$$\prod_{y^{c_7} \leqq p_i \leqq y^4} \left(1 - \frac{1}{p_i}\right) \geqq \prod_{y^{c_7} \leqq P \leqq y^4} \left(1 - \frac{1}{P}\right) \geqq \exp\left(\sum_{y^{c_7} \leqq P \leqq y^4} \frac{1}{P}\right) > c_8, \tag{5}$$

since

$$\sum_{P < x} \frac{1}{P} = \log \log x + O(1). \tag{6}$$

Thus from (4) and (5)

$$D_l(y) < c_9 y \prod_{p_i < y^4} \left(1 - \frac{1}{p_i}\right) \qquad \left(c_9 = \frac{c_6}{c_8}\right). \tag{7}$$

Now we prove Theorem 1. Since $\sum \frac{1}{p_i} = \infty$, there exist infinitely many k satisfying $p_k < k^2$. Put

$$l = \left[\frac{k}{2}\right], \quad t = c_{10} k e^{f(p_k)},$$

where c_{10} is sufficiently small and will be determined later. Denote by $a_1, a_2, \ldots, a_z$ the integers not excending t which are either not divisible by p_i, $(1 \leqq i \leqq l)$ or are divisible by p_i^2. First we prove that for sufficiently large k

$$z < l. \tag{8}$$

[1]) See for example P. ERDŐS, On the easier Waring problem for powers of primes. I. *Proc. Cambridge Phil. Soc.* **33** (1937), 6—12.

First we show that for sufficiently large k

$$k < t < k(\log k)^2. \tag{9}$$

Obviously $t > k$ since $f(p_k) \to \infty$, $t < k(\log k)^2$ follows immediately from (6). To prove (8) write $a_j = b_j d_j$ where d_j is not divisible by any of $p_1, p_2, \ldots, p_l$ and b_j is entirely composed of the p_i, $1 \leqq i \leqq l$ and all the exponents of its prime factors are greater than 1. We split the a's into two classes. In the first class are the a's with $b_j \geqq t^{1/2}$ and in the second class the a's with $b_j < t^{1/2}$. Clearly b_j is divisible by a square not less than $b_j^{2/3}$ ($b_j = p_1^{2\alpha_1} \ldots p_h^{2\beta_1+1} \ldots$) and the largest square dividing b_j is $p_1^{2\alpha_1} \ldots p_h^{2\beta_1} \ldots$ which is clearly $\geqq b_j^{2/3}$. Thus the integers of the first class are divisible by a square $\geqq t^{1/3}$. Therefore by (9) the number of the a's of the first class is not greater than

$$\sum_{m \geqq t^{1/2}} \frac{t}{m^2} < t^{5/6} < k^{5/6}(\log k)^2 < \frac{k}{4}. \tag{10}$$

The number of integers ot the second class clearly equals

$$\sum_{b_j \leqq t^{1/6}} D_l\left(\frac{t}{b_j}\right). \tag{11}$$

We have by (9) and $p_l < p_k < k^2$

$$\left(\frac{t}{b_j}\right)^4 \geqq t^2 > p_k.$$

Thus by (7)

$$D_l\left(\frac{t}{b_j}\right) < c_9 \frac{t}{b_j} \prod_{i=1}^{k}\left(1 - \frac{1}{p_i}\right).$$

Thus

$$D_l\left(\frac{t}{b_j}\right) < c_9 \frac{t}{b_j} \prod_{i=1}^{k}\left(1 - \frac{1}{p_i}\right) < c_{11} \frac{t}{b_j} e^{-f(p_k)} = c_{10}c_{11}\frac{k}{b_j}. \tag{12}$$

Now

$$\sum \frac{1}{b_j} < \sum_{r=1}^{\infty} \frac{1}{r^2} \sum_{s=1}^{\infty} \frac{1}{s^3} < c_{12}. \tag{13}$$

Thus from (10), (12) and (13) the number of integers of the second class is less than

$$c_{10}c_{11}c_{12}k < \frac{k}{4} \tag{14}$$

for sufficiently small c_{10}. Thus (10) and (14) imply (8).

Let now $0 < x < (p_1 p_2 \cdots p_k)^2$ satisfy the following congruences:

$$x \equiv 0 \pmod{(p_1 p_2 \cdots p_l)^2}$$

$$x + a_i \equiv p_{l+i} \pmod{p_{l+i}^2}, \qquad i = 1, 2, \cdots, z.$$

Since by (8) $z < l$, all the p's are less than p_k.

Now we show that none of the integers

$$x+u,\ 0<u\leqq t$$

are v's. If u is not one of the a's then there exists a p_i, $1\leqq i\leqq l$ so that $p_i|u, p_i^2\nmid u$. But then clearly $p_i|(x+u)$, $p_i^2\nmid(x+u)$, i. e, $x+u$ is not a v. If $u=a_i$ then $p_{i+i}|(x+u)$, $p_{i+i}^2\nmid(x+u)$ i. e. $x+u$ is not a v. Define now

$$v_j\leqq x<x+t<v_{j+1}.$$

Thus

$$v_{j+1}-v_j>t=c_{10}ke^{f(p_k)}. \tag{15}$$

Further from $p_k<k^2$

$$v_j\leqq x<(p_1p_2\cdots p_k)^2<p_k^{2k}<k^{4k},$$

hence

$$k>c_{13}\frac{\log v_j}{\log\log v_j}. \tag{16}$$

Thus from (15) and (16)

$$v_{j+1}-v_j>c_{14}e^{f(p_k)}\cdot\frac{\log v_j}{\log\log v_j}. \tag{17}$$

From (16) $k^2>\log v_j$. Thus from (6)

$$f(\log v_j)-f(p_k)\leqq f(k^2)-f(p_k)\leqq f(k^2)-f(k)<c_{15}. \tag{18}$$

Hence finally from (17) and (18)

$$v_{j+1}-v_j>c_4e^{f(\log v_j)}\frac{\log v_j}{\log\log v_j}$$

which proves Theorem 1.

It is not difficult to show that Theorem 1 remains true if we do not assume that the p's are primes but only assume that $(p_i, p_j)=1$.

As stated before it seems hard to improve Theorem 1. In only succeeded in doing this if we assume that

$$\sum\frac{1}{q}<\infty \tag{19}$$

where q runs through the primes which are not p's. In this case Theorem 1 gives

$$v_{i+1}-v_i>c_{16}\log v_i,$$

and I can prove $\lim\frac{v_{i+1}-v_i}{\log v_i}=\infty$.

If we only want to prove (2) we do not need BRUN's method. We can use instead LANDAU's well known result that the number of integers $\leqq t$ of the form u^2+v^2 is $O\left(\frac{t}{(\log t)^{1/2}}\right)$. In fact LANDAU showed that it is

$$A\frac{t}{(\log t)^{1/2}}+o\frac{t}{(\log t)^{1/2}}.$$

Let us now assume that $\sum \frac{1}{p_i} < \infty$. Perhaps the most interesting case is if the v's are the squarefree numbers. Here the p's are not primes but the squares of primes. Denote by $s_1 < s_2 < \cdots$ the sequence of squarefree numbres. Our above method gives that for infinitely many i

$$s_{i+1} - s_i > (1 + o(1)) \frac{\pi^3}{6} \frac{\log s_i}{\log \log s_i}. \tag{20}$$

I do not know if (20) has been published before, but it was certainly known to several mathematicians e. g. BATEMAN, CHOWLA and MIRSKY. The curious thing is that it seems to be extremely hard to replace $\frac{\pi^2}{6}$ by any larger constant, in fact it seems possible that for $i > i_0$

$$s_{i+1} - s_i < (1 + \varepsilon) \frac{\pi^2}{6} \frac{\log s_i}{\log \log s_i}. \tag{21}$$

The strongest result in the direction of (21) is due to ROTH[2]) and states that

$$s_{i+1} - s_i < s_i^{3/13} (\log s_i)^{\frac{4}{13} + \varepsilon}. \tag{22}$$

When I heard of ROTH's result I thought of trying to prove that for every α

$$\sum_{s_{k+1} < x} (s_{k+1} - s_k)^{\alpha} = C_{\alpha} x + o(x). \tag{23}$$

The proof of (23) seems very difficult; it would of course imply $s_{i+1} - s_i = o(s_i^{\varepsilon})$. In can prove (23) only if $\alpha < A$ where A is a certain constant between 2 ane 3. Here I only sketch the proof for $\alpha = 2$.

Denote by $g_t(x)$ the number of $s_i < x$ satisfying $s_{i+1} - s_i = t$.

Lemma 1. *For fixed t as $x \to \infty$*

$$g_t(x) = \beta_t x + o(x).$$

In other words the density of the s_i with $s_{i+1} - s_i = t$ exists.

Lemma 1 is known.[3])

Lemma 2. *There exists an absolute constant c_{17} so that*

$$\sum_{l > t} g_l(x) < c_{17} \frac{x}{t^2 (\log t)^2}.$$

Put $s_{i+1} - s_i = r > t$. We show that there are at least $\frac{r}{16}$ integers z satisfying $s_i < z < s_{i+1}$ which are divisible by the square of a prime $P > \frac{t \log t}{100}$. To

[2]) K. F. ROTH, On the gaps between squarefree numbers, *Journal London Math. Soc.* Vol. **26** (1951), 263—268.

[3]) L. MIRSKY, Arithmetical pattern problems relating to divisibility by r-th powers. *Proc. London Math. Soc.* **50** (1949), 497—508. — See Theorem 4, p. 507.

show this we observe that the number of integers $s_i < y < s_{i+1}$ which are divisible by the square P^2 of a prime $P \leqq \frac{t \log t}{100}$ is at most

$$(24) \qquad \sum_{P \leqq \frac{t \log t}{100}} \left(\left[\frac{r}{P^2} \right] + 1 \right) < r \sum \frac{1}{P^2} + \pi\left(\frac{t \log t}{100} \right) < \frac{3}{4} r + \frac{r}{8} = \frac{7}{8} r,$$

since $\sum \frac{1}{p^2} < \frac{1}{4} + \frac{1}{2.3} + \frac{1}{3.4} + \cdots < \frac{3}{4}$ and by the results of TCHEBICHEFF for $t > 1$

$$\pi\left(\frac{t \log t}{100} \right) < \frac{t}{8} < \frac{r}{8}.$$

Since by assumption all the $r-1$ integers z satisfying $s_i < z < s_{i+1}$ are divisable by squares, we obtain from (24) that for $r > 16$ there are at least $r - 1 - \frac{7}{8} r > \frac{r}{16}$ integers z divisible by a P^2 with $P > \frac{t \log t}{100}$. [For $r \leqq 16$, $t < 16$, $\frac{t \log t}{100} < 1$ thus all the $r-1$ integers are divisible by P^2 with $P > \frac{t \log t}{100}$.] Hence our assertion is proved. Thus there are at least

$$(25) \qquad \frac{1}{16} \sum_{\substack{s_{i+1} \leqq x \\ s_{i+1} - s_i > t}} (s_{i+1} - s_i)$$

integers not exceeding x which are divisible by the square of a prime $P > \frac{t \log t}{100}$. But the number of these integers is less than

$$(26) \qquad \sum_{P > \frac{t \log t}{100}} \frac{x}{P^2} < c_{18} \frac{x}{t (\log t)^2}$$

[since $P_k > c_{11} k \log k$]. Thus from (25) and (26)

$$(27) \qquad \sum_{\substack{s_{i+1} \leqq x \\ s_{i+1} - s_i > t}} (s_{i+1} - s_i) < 16 c_{18} \frac{x}{t (\log t)^2}$$

or the number of i's with $s_{i+1} \leqq x$, $s_{i+1} - s_i > t$ is less than

$$16 c_{18} \frac{x}{t^2 (\log t)^2}$$

which proves Lemma 2.

From Lemma 2 we have by a simple computation that for every $\varepsilon > 0$ there exist an r so that

$$(28) \qquad \sum_{\substack{s_{i+1} \leqq x \\ s_{i+1} - s_i > 2^r}} (s_{i+1} - s_i)^2 = \sum_{j=0}^{\infty} \sum_{2^{r+j} < s_{i+1} - s_i \leqq 2^{r+j+1}} (s_{i+1} - s_i)^2 < \sum_{j=0}^{\infty} 2^{2r+2j+2} \sum_{l < 2^{r+j}} g_l(x) <$$

$$< c_{17} x \sum_{j=0}^{\infty} \frac{2^{2r+2j+2}}{2^{2r+2j} (r+j)^2} < \varepsilon x.$$

Further from Lemma 1

$$\sum_{\substack{s_{i+1}\leqq x \\ s_{i+1}-s_i\leqq 2^r}} (s_{i+1}-s_i)^2 = x\sum_{t=1}^{2^r} t^2\beta_t + o(x). \tag{29}$$

From (28) and (29) we obtain by a simple argument that

$$\sum_{s_{i+1}\leqq x} (s_{i+1}-s_i)^2 = O(x),\quad \sum_{t=1}^{\infty} t^2\beta_t < \infty$$

and finally that

$$\sum_{s_{i+1}\leqq x} (s_{i+1}-s_i)^2 = x\sum_{t=1}^{\infty} t^2\beta_t + o(x)$$

which proves (23) for $\alpha = 2$.

Lemma 1 can be generalized as follows: Let $a_1 < a_2 < \cdots$ be any sequence of integers. Denote by $b_1 < b_2 < \cdots$ the set of integers not divisible by any of the a's. Assume that the density of the b's exists. [This is certainly the case if $\sum \frac{1}{a_i} < \infty$.] Then the density of integers satisfying $b_{i+1} - b_i = t$ also exists. We do not discuss the proof, but only remark that it follows almost immediately from the following theorem of DAVENPORT and myself[4]): Denote by c_k the density of the integers not divisible by $a_1, a_2, \ldots, a_k$, and by c the density of the integers not divisible by $a_1, a_2, \ldots$. Then $c = \lim_{k\to\infty} c_k$.

By the theorem of DAVENPORT and myself it is easy to see that if the density of the b's exists and is positive, then to every ε there exist a c_ε so that

$$\sum_{\substack{b_{i+1}\leqq x \\ b_{i+1}-b_i > c_\varepsilon}} (b_{i+1}-b_i) < \varepsilon x.$$

On the other hand no stronger result can hold in general even if $\sum \frac{1}{a_i} < \infty$. Define the a's for example as the integers in the intervals $\left[2^k, 2^k\left(1+\frac{1}{k^2}\right)\right]$, then $\sum \frac{1}{a_i} < \infty$ but

$$\lim_{x\to\infty} \frac{1}{x}\sum_{b_i<x} (b_{i+1}-b_i)^{1+\varepsilon} = \infty.$$

I do not know whether this can happen if $\sum \frac{1}{a_i} < \infty$ and $(a_i, a_j) = 1$.

(Recieved May 4, 1951.)

---

[4]) H. DAVENPORT and P. ERDŐS, On sequences of integers. *Acta arithmetica* **2** (1936), 147—151. — See also: H. DAVENPORT and P. ERDŐS, On sequences of positive integers. *J. Indian Math. Soc.* **15**, Part A (1951), 19—24.