On a Problem in Elementary Number Theory and a Combinatorial Problem

By P. Erdös

In a recent paper [1] I considered among others the following little problem: Denote by $f_t(n)$ the smallest integer l so that if

$$1 \leqq a_1 < a_2 < \cdots < a_l \leqq n, \qquad l = f_t(n),$$

is an arbitrary sequence of integers one can always find t a's $a_{i_1}, \cdots, a_{i_t}$ which have pairwise the same greatest common divisor. I proved in [1] that for fixed t

$$f_t(n) < \frac{n}{\exp [(\log n)^{1/2}]^{-\epsilon}}. \tag{1}$$

Recently, I observed that using a combinatorial theorem due to Rado and myself (1) can be considerably improved and it might, in fact, be possible to obtain the correct order of magnitude for $f_t(n)$. The combinatorial theorem in question states as follows [2]: Let $g(k, t)$ be the smallest integer so that if $A_1, \cdots, A_s$, $s = g(k, t)$, are sets each having k or fewer elements then there are always t A's $A_{i_1}, \cdots, A_{i_t}$ which have pairwise the same intersection. We have

$$g(k, t) < k!(t - 1)^{k+1}. \tag{2}$$

We conjectured that (2) can be improved to ($c_1, c_2, \cdots$ are absolute constants)

$$g(k, t) < c_1^k(t - 1)^{k+1}. \tag{3}$$

The conjectured (3) would have applications to several questions in number theory. It is not difficult to show that

$$\lim_{k=\infty} g(k, t)^{1/k}$$

exists, but I cannot show that it is finite.

Now we prove the following:

Theorem. *For every t and $\epsilon > 0$ there is an n_0 so that for all $n > n_0(t, \epsilon)$,*

$$2^{c_t \log n/\log \log n} < f_t(n) < n^{3/4+\epsilon}. \tag{4}$$

First we prove the upper bound in (4).

Let $1 \leqq a_1 < a_2 < \cdots < a_l \leqq n$, $l = [n^{3/4+\epsilon}]$ be an arbitrary sequence of integers. We split the a's into two classes. In the first class are the a's which have at least

$$\left[\frac{\log n}{4 \log \log n}\right] = u$$

Received March 20, 1964.

distinct prime factors. Denote by $w_1, w_2, \cdots$ the squarefree integers not exceeding n which have exactly u prime factors. Clearly every number of the first class is a multiple of some w_i, hence the number of integers of the first class is by a simple calculation at most

$$\sum_i \frac{n}{w_i} < n \cdot \sum_{p_i \leqq n} \left(\frac{1}{p_i}\right)^u \Big/ u! < n(\log \log n + c_2)^u/u!$$
$$< n(e(\log \log n + c_2))^u/u! < \tfrac{1}{2} \cdot n^{3/4+\epsilon}$$

for every ϵ if n is sufficiently large.

Hence the number of integers of the second class is greater than $\frac{1}{2} \cdot n^{3/4+\epsilon}$. Consider the (unique) factorization

$$(5) \qquad a_i = A_i B_i, \qquad (A_i, B_i) = 1,$$

where each prime factor of A_i occurs with an exponent greater than one and B_i is squarefree. It is well known [3] and easy to prove that the number of integers $m \leqq n$ all of whose prime factors occur with an exponent >1 is less than $c_3 n^{1/2}$. Hence there are at least $(1/2c_3)n^{1/4+\epsilon}$ integers a_{i_j} with the same A_i:

$$(6) \qquad a_{i_j} = A_{i_j} B_{i_j}, \qquad 1 \leqq j \leqq r, \quad r > \frac{1}{2c_3} n^{1/4+\epsilon}, \quad A_{i_j} = A.$$

Clearly the number of prime factors of the squarefree number B_i is less than u. A simple computation gives, for $n > n_0(\epsilon, t)$,

$$\frac{1}{2c_3} n^{1/4+\epsilon} > u!(t-1)^{u+1} \qquad \left(u = \left[\frac{1}{4} \frac{\log n}{\log \log n}\right]\right).$$

Hence from (2) there are at least t B's and hence by (6) at least t a's which have pairwise the same common factor, which proves the upper bound in (4).

To prove the lower bound in (4) put

$$k = \left[\frac{\log n}{3 \log \log n}\right]$$

and denote by $p_i^{(j)}$, $1 \leqq i \leqq 3$, $1 \leqq j \leqq k$, the first $3k$ primes. Put

$$b_1^{(j)} = p_1^{(j)} p_2^{(j)}, \quad b_2^{(j)} = p_1^{(j)} p_3^{(j)}, \quad b_3^{(j)} = p_2^{(j)} p_3^{(j)}.$$

The a's are the 3^k integers of the form

$$\prod_{j=1}^{k} b_i^{(j)}, \qquad i = 1, 2, \text{ or } 3.$$

A simple computation using the prime number theorem (or a more elementary result) shows that all the a's are less than n. Further, obviously no three of them have pairwise the same greatest common divisor, also $f_t(n) \geqq f_3(n)$, thus the lower bound in (4) is proved and the proof of our theorem is complete.

The inequality (3) would easily imply

$$(7) \qquad f_t(n) < (c_t')^{\log n/\log \log n}.$$

The proof of (7) (using the unproved conjecture (3)) would be similar to the proof of our theorem. Instead of the decomposition (5) we would have to put $a_i = C_iD_i$ where all prime factors of C_i are less than $\log n$ and all prime factors of D_i are $\geqq \log n$. We suppress the details.

Very likely

$$\lim_{n=\infty} \log\left(f_t(n)\right)\cdot\frac{\log\log n}{\log n} \tag{8}$$

exists and perhaps it might be possible to determine its value, but it will probably not be possible to express $f_t(n)$ by a simple function of n and t (even for $t = 3$).

If t is large compared to n our method used in the proof of our theorem no longer gives a good estimation, but it is not difficult to prove by a different method the following result. Let $1 \leqq a_1 < a_2 < \cdots < a_l \leqq n$, $l = Cn$ be given, then there are always n^{ϵ_C} integers $a_{i_1}, \cdots, a_{i_r}$ which have pairwise the same common factor (ϵ_C depends only on C), but we do not investigate this question here any further.

I have not been able to decide if to every $\alpha > 0$ there is an $n_0(\alpha)$ so that if $n > n_0(\alpha)$ and

$$1 \leqq a_1 < a_2 < \cdots < a_l \leqq n, \qquad l \geqq \alpha n,$$

is any sequence of integers, then there always are three a's which have pairwise the same least common multiple. This is certainly true (and trivial) if α is close enough to 1; perhaps the whole question is trivial and I overlooked an obvious approach.

McMaster University
Hamilton, Ontario

1. P. Erdös, "Extremal problems in number theory," *Mat. Lapok*, v. 13, 1962, p. 228–255. (Hungarian)

2. P. Erdös & R. Rado, "Intersection theorems for systems of sets," *J. London Math. Soc.*, v. 35, 1960, p. 85–90.

3. P. Erdös & G. Szekeres, "Über die Anzahl der Abelschen Gruppen gegebener Ordnung und über ein verwandtes zahlentheoretisches Problem," *Acta Sci. Math. (Szeged)*, v. 7, 1934, p. 94–102.

On Maximal Gaps between Successive Primes

By Daniel Shanks

In personal correspondence Paul A. Carlson asked the author if he could give a rough "ball-park" estimate of where one would first find a run of a million or more consecutive composite integers. For notation let us define $p(g)$ to be the first prime that follows a gap of g or more consecutive composites. Thus $p(1) = 5$, $p(2) = p(3) = 11$, $p(4) = p(5) = 29$, $p(6) = p(7) = 97$, etc. We seek to estimate $p(10^6)$. Conversely, by $g(n)$ we mean the largest gap that occurs below any prime $p \leqq n$. We may call these values of g *maximal gaps*.

That $p(g)$ is finite for every g is well known. The famous proof by Lucas [1] merely notes that the g consecutive integers:

$$(g+1)!+2, (g+1)!+3, (g+1)!+4, \cdots, (g+1)!+g+1$$

Received April 27, 1964.